(12) STANDARD PATENT (11) Application No. **AU 2004311384 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Cell cycle genes and related methods of using

(51) International Patent Classification(s)
C12N 15/29 (2006.01) ***C07K 14/415*** (2006.01)
A01H 5/00 (2006.01) ***C12N 15/82*** (2006.01)

(21) Application No: **2004311384** (22) Date of Filing: **2004.12.30**

(87) WIPO No: **WO05/065339**

(30) Priority Data

(31) Number	(32) Date	(33) Country
60/533,036	**2003.12.30**	**US**

(43) Publication Date: **2005.07.21**
(44) Accepted Journal Date: **2012.01.12**

(71) Applicant(s)
ArborGen Inc.

(72) Inventor(s)
Grigor, Murray Robert;Magusin, Andreas;Connett, Marie B.;Emerson, Sarah Jane;Higgins, Colleen M.;Lund, Steven Troy;Kodrzycki, Robert John;Forster, Richard L.S.

(74) Agent / Attorney
Knightsbridge Patent Attorneys, Level 14 200 Queen Street, Melbourne, VIC, 3000

(56) Related Art
Database EMBL, Accession No. CF477422, 08 September 2003

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 July 2005 (21.07.2005)

PCT

(10) International Publication Number
WO 2005/065339 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/US2004/043804

(22) International Filing Date:
30 December 2004 (30.12.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/533,036 30 December 2003 (30.12.2003) US

(71) Applicants *(for all designated States except US)*: **ARBOGEN, LLC** [US/US]; 180 Westvaco Road, P.O. Box 840001, Summerville, SC 29484 (US). **GENESIS RESEARCH AND DEVELOPMENT CORPORATION LIMITED** [NZ/NZ]; One Fox Street, Parnell, Auckland (NZ).

(72) Inventors; and

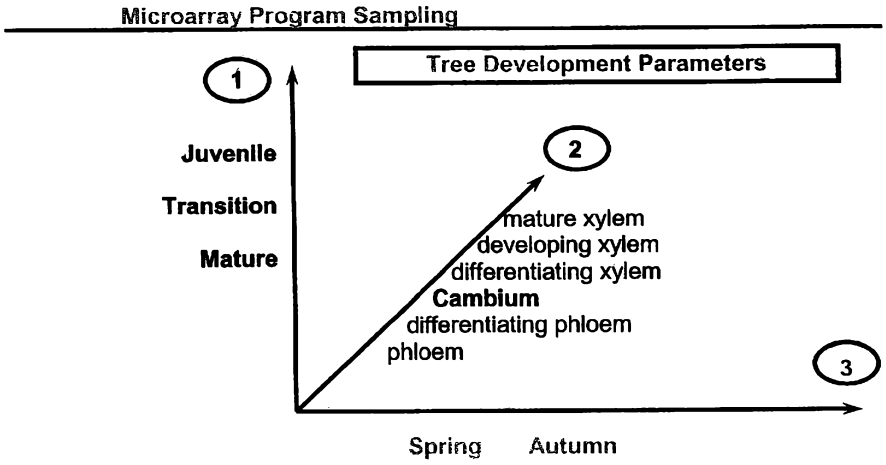
(75) Inventors/Applicants *(for US only)*: **FORSTER, Richard, L.** [NZ/NZ]; c/o Genesis Research and Development Corporation, Limited, One Fox Street, Parnell, P.O. Box 50, Auckland (NZ). **CONNETT, Marie, B.** [US/US]; c/o Cambia, GPO Box 3200, Canberra, ACT 2601 (AU). **EMERSON, Sarah, Jane** [NZ/NZ]; c/o Genesis Research and Development Corporation, Limited, One Fox Street, Parnell, P.O. Box 50, Auckland (NZ). **GRIGOR, Murray, Robert** [NZ/NZ]; c/o Genesis Research and Development Corporation, Limited, One Fox Street, Parnell, P.O. Box 50, Auckland (NZ). **HIGGINS, Collen, M.** [NZ/NZ]; c/o Genesis Research and Development Corporation, Limited, One Fox Street, Parnell, P.O. Box 50, Auckland (NZ). **LUND, Steven, Troy** [CA/CA]; c/o Genesis Research and Development Corporation, Limited, One Fox Street, Parnell, P.O. Box 50, Auckland (NZ). **MAGUSIN, Andreas** [SE/NZ]; c/o Genesis Research and Development Corporation, Limited, One Fox Street, Parnell, P.O. Box 50, Auckland (NZ). **KODRZYCKI, Robert, J.** [US/US]; c/o ArborGen LLC, 180 Westvaco Road, P.O. Box 840001, Summerville, SC 29484 (US).

(74) Agents: **PEET, Richard, C.** et al.; Foley & Lardner LLP, Washington Harbour, 3000 K Street, N.W., Suite 500, Washington, D.C.20007-5143 (US).

[Continued on next page]

(54) Title: CELL CYCLE GENES AND RELATED METHODS OF USING

Exemplary Microarray Sampling Parameters.



(57) Abstract: Novel plant polysaccharide synthesis genes and polypeptides encoded by such genes are provided. These genes and polynucleotide sequences are useful regulating polysaccharide synthesis and plant phenotype. Moreover, these genes are useful for expression profiling of plant polysaccharide synthesis genes. The invention specifically provides cell cycle polynucleotide and polypeptide sequences isolated from *Eucalyptus* and *Pinus*.

WO 2005/065339 A2

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

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

Published:

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

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

CELL CYCLE GENES AND RELATED METHODS OF USING

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to U.S. provisional application serial number 60/533,036, filed on December 30, 2003, which is specifically incorporated in its entirety herein by reference.

FIELD OF THE INVENTION

[0002] The present invention relates generally to the field of plant cell cycle genes and polypeptides encoded by such genes, and the use of such polynucleotide and polypeptide sequences for regulating a plant cell cycle. The invention specifically provides cell cycle polynucleotide and polypeptide sequences isolated from *Eucalyptus* and *Pinus* and sequences related thereto.

BACKGROUND OF THE INVENTION

[0003] Cell growth and division are controlled by the temporal expression of different sets of genes, allowing the dividing cell to progress through the different phases of the cell cycle. Continued growth and organogesis in plants requires precise function of the cell cycle machinery. Plant development, which is directly affected by cell division rates and patterns, also is influenced by environmental factors, such as temperature, nutrient availability, light, etc. *See* Gastal and Nelon, *Plant Physiol.* 105:191-7 (1994), Ben-Haj-Sahal and Tardieu, *Plant Physiol.* 109:861-7 (1995), and Sacks *et al., Plant Physiol.* 114:519-27 (1997). Plant development and phenotype are connected with the cell cycle, and altering expression of the genes involved in the cell cycle can be a useful method of modifying plant development and altering plant phenotype.

[0004] The ability to alter expression of cell cycle genes is extremely powerful because the cell cycle drives plant development, including growth rates, responses to environmental cues, and resulting plant phenotype. Control of the plant cell cycle and phenotypes associated with alteration of cell cycle gene expression, in

2004311384 31 May 2011

the vascular cambium, in particular, has applications for, *inter alia*, alteration of wood properties and, in particular, lumber and wood pulp properties. For example, improvements to wood pulp that can be effected by altering cell cycle gene expression include increased or decreased lignin and cellulose content, and altered length, diameter, and lumen diameter of cells. Manipulating the plant cell cycle, and in particular the cambium cell cycle (*i.e.* the rate and angle of cell division), can also engineer better lumber having increased dimensional stability, increased tensile strength, increased shear strength, increased compression strength, increased shock resistance, increased stiffness, increased or decreased hardness, decreased spirality, decreased shrinkage, and desirable characteristics with respect to weight, density, and specific gravity.

A. Cell Cycle Genes and Proteins

1. Cyclin dependent protein kinase

[0005] Progression through the cell cycle is regulated primarily by cyclin-dependent kinases (CDKs). CDKs are a conserved family of eukaryotic serine/threonine protein kinases, which require heterodimer formation with a cyclin subunit for activity. For review see, *e.g.* Joubes *et al.*, *Plant Mol. Biol.* 43: 607-20 (2000), Stals and Inze, Trends Plant Sci. 6:359-64 (2001), and John *et al.*, *Protoplasma* 216: 119-42 (2001).

[0006] The are five subclasses of CDK's, each having a different cyclin binding consensus sequence. In CDK type A the cyclin binding consensus sequence is PSTAIRE (SEQ ID NO: 778). *Id.* The cyclin binding consensus sequence in CDK types B-1, B-2, and C are PPTTLRE (SEQ ID NO: 779), PPTALRE (SEQ ID NO: 780), and PITAIRE (SEQ ID NO: 781), respectively. Joubes *et al, Plant Physiol,* 126: 1403-15 (2001).

[0007] Cell cycle progression is directed, in part, by changes in CDK activity. CDK activity is modulated by a number of different cell cycle protein components, such as changes in the abundance of individual cyclins due to changing rates of biosynthesis and proteolysis. Fluctuations in cyclin concentrations result in commensurate fluctuations in CDK activity. Cyclin accumulation is especially

important in terminating the G1 phase of the cell cycle because DNA replication is initiated by an increase in CDK activity.

[0008] Activation of CDK also requires phosphorylation of a threonine residue within the T-loop of CDK by a CDK-activating kinase (CAK). Umeda *et al., Proc. Nat'l Acad. Sci. U. S. A.* 97: 13396-400 (2000). It was suggested by Yamaguchi *et al., Plant J.* 24: 11-20 (2000), that cyclin H is a regulatory subunit of CAK. CDK activity is further regulated by interaction with a CDK regulatory subunit, a small (70-100 AA) protein involved in cell cycle regulation.

[0009] A cell must exit the cell cycle in order to commit to differentiation, senescence or apoptosis. This process involves the down-regulation of CDK activities. CDK inhibitors (CKI) are low molecular weight proteins, which are important for cell cycle regulation and development. CKIs bind stoichiometrically to CDK and down-regulate the activity of CDKs.

[0010] Many biochemical properties of ICK1, the first plant CKI to be identified from *Arabidopsis thaliana*, are known. Wang *et al., Nature* 386:451-2 (1997) Wang *et al., Plant J.* 24: 613-23 (2000). ICK1 is expressed at low levels in many tissue types, and there can be a threshold level of ICK1 that must be overcome before a cell can enter the cell cycle. Wang *et al., Plant J.* 24: 613-23 (2000). ICK1 is induced by the plant growth regulator abscisic acid (ABA), which inhibits cell division by blocking DNA replication. When the expression of ICK1 increases, there is a corresponding decrease in Cdc2-like H1 histone activity. ICK1 has been shown to bind *in vitro* with the cyclins C2c2a and CycD3, and deletion experiments have identified different domain regions for these two interactions.

[0011] Altering the expression of CDK regulatory protein or a subunit thereof is known to cause changes in plant phenotype. Overexpression of the *Arabidopsis* CDK regulatory subunit, CKS1At, resulted in a reduction of leaf size, root growth rates and meristem size. Additionally, overexpression of CKS1At resulted in inhibition of cell-cycle progression, with an extension in the duration of the G1 and G2 phases of the cell cycle.

2004311384 31 May 2011

2. Cyclins

[0012] Cyclins are positive regulatory subunits of cyclin-dependent kinase (CDK) enzymes and are required for CDK activity. Fowler *et al.*, *Mol. Biotech.* 10, 123, 126. Cyclins and CDK complexes provide temporal regulation of transition through the cell cycle. Evidence also suggests that cyclins provide spatial regulation of specific CDK activity, differentially targeting the cytoskeleton, spindle, phragmoplast, nuclear envelope, and chromosomes.

[0013] Plant cyclins are classified into five major groups: A, B, C, D, and H. Renaudin *et al.*, *Plant Mol. Biol.* 32: 1003-18 (1996) and Yamaguchi *et al.*, (*supra* 2000). Cyclins can be divided into mitotic cyclins (A and B) and G_1 cyclins.

[0014] The mitotic cyclins possess a consensus sequence (R-x-x-L-x-x-I-x-N, SEQ ID NO: 782) located at the N-terminal region, termed a destruction box, adjacent to a lysine-rich region. The destruction box and lysine-rich region target the mitotic cyclins for ubiquitin-dependent proteolysis during mitosis. Stals, *supra* at 361, and Fowler, *supra* at 126. The destruction box in A versus B cyclins differs slightly and this difference is thought to result in slightly different timing of degradation of A versus B cyclins. Fowler, *supra* at 126. A-type cyclins accumulate during the S, G2, and early M phase of the cell cycle, whereas B-type cyclins accumulate during the late G2 and early M phase. Mironov *et al.*, *Plant Cell* 11: 509-22 (1999). Three subgroups of A-type cyclins are known in plants, but only one is known in animals. Cyclin A1 (cycA1;zm;1 from *Zea cans*) is most concentrated during cytokinesis at the microtubule-containing phragmoplast. Expression of cyclin A2 is upregulated by auxins in roots, and by cytokinins in the shoot apex. Abrahams *et al.*, *Biochim. Biophys. Acta* 28: 1-2 (2001).

[0015] D-type cyclins, of which five subgroups are known, are thought to control the progression through the G1 phase in response to growth factors and nutrients. Riou-Khamlichi *et al.*, *Mol. Cell Biol.* 20: 4513-21 (2000). For example, the expression of D-type cyclins is upregulated by sucrose as shown by an increase in cycD2 mRNA 30 minutes after sucrose exposure, and an increase in cycD3 four hours

after sucrose exposure. This timing corresponds to early G1-phase and late G1-phase, respectively. Cockcroft *et al., Nature* 405: 575-9 (2000). Furthermore, in *Arabidopsis*, a D3 cyclin was shown to be upregulated by the brassinosteroid, epi-brassinolide.

[0016] Cyclin D2 proteins bind with CDKA to produce an active complex, which binds to and phosphorylates retinoblastoma-related protein (Rb). This process is found in actively proliferating tissue, suggesting it plays an important function during late G1- and early S-phase. Three different D3-type cyclins are active during tomato fruit development. These proteins all contain a retinoblastoma binding motif and a PEST-destruction motif. There are differences in the spatial and temporal expression of these D3 cyclins, inferring different roles during fruit development.

[0017] Overexpression of cyclin D was shown to increase overall growth rate. Over-expression of cyclin D2 in tobacco increases causes shortening the G1-phase which producing a faster rate of cell cycling.

[0018] C- and H-type cyclins were characterized in poplar (*Populus tremula* × *tremuloides*) and rice (*Oryza sativa*) but their exact function is still unclear. Putative cyclins with a lesser degree of peptide sequence conservation have also been identified. For example, *Arabidopsis* CycJ18 has only 20% identity with homologues over the cyclin box domain. CycJ18 is expressed predominantly in young seedlings. *Arabidopsis* F3O9.13 protein also has similarity to the cyclin family.

3. Histone acetyltransferase/deacetyltransferase

[0019] Histone acetyltransferase (HA) and histone deacetyltransferase (HAD) control the net level of acetylation of histones. Histone acetylation and deacetylation are thought to exert their regulatory effects on gene expression by altering the accessibility of nucleosomal DNA to DNA-binding transcriptional activators, other chromatin-modifying enzymes or multi-subunit chromatin remodeling complexes capable of displacing nucleosomes. Lusser *et al., Nucleic Acids Res.* 27: 4427-35 (1999). Therefore, in general, the HDAs are involved in the repression of gene expression, while HAs are correlated with gene activation.

[0020] HA effects acetylation at the ε-amino group of conserved lysine residues clustered near the amino terminus of core histones which up-regulates gene expression.

[0021] HDAs remove acetyl groups from the core histones of the nucleosome. There are numerous family members in the HDA group, many of which are conserved throughout evolution. Lechner *et al.*, *Biochim Biophys Acta* 5:181-8 (1996). HDAs function as part of multi-protein complexes facilitating chromatin condensation.

[0022] HDAs and HAs recognize highly distinct acetylation patterns on the nucleosome. It is thought that different types of HDAs interact with specific regions of the genome, to influence gene silencing.

[0023] Schultz *et al.*, *Genes Dev.* 15: 428-43 (2001), demonstrated that the superfamily of Kruppel-associated-box zinc finger proteins (KRAB-ZFPs) are linked to the nucleosome remodelling and histone deacetylation complex via the PHD (plant homeodomain) and bromodomains of co-repressor KAP-1, to form a cooperative unit that is required for transcriptional repression. A maize HDAC (HD2) has been identified that has no sequence homology to other eukaryotic HDACs, but instead contains sequence similarity to peptidyl-prolyl cis-trans isomerases (PPIases).

[0024] The effects of interfering with histone deacetylation are discussed in *e.g.* Tian and Chen, *Proc. Nat'l Acad. Sci. USA* 98: 200-5 (2001).

4. Peptidyl prolyl cis-trans isomerase

[0025] Peptidylprolyl isomerases (*e.g.*, peptidylprolyl cis-trans isomerase, peptidyl-prolyl cis-trans isomerase, PPIase, rotamase, cyclophilin) catalyze the interconversion of peptide bonds between the cis and trans conformations at proline residues. Sheldon and Venis, *Biochem J.* 315: 965-70 (1996). This interconversion is thought to be the rate limiting step of protein folding. PPIases belong to a conserved family of proteins that are present in animals, fungi, bacteria and plants. PPIases are implicated in a number of responses including the response to environmental stress,

calcium signals, transcriptional repression, cell cycle control, *etc.* Viaud, *et al.*, *Plant Cell* 14: 917-30 (2002).

5. Retinoblastoma-related protein

[0026] Retinoblastoma (Rb)-related protein putatively regulates progression of the cell cycle through the G1 phase and into S phase. Xie *et al., EMBO J.* 15: 4900-8 (1996) and Ach et al., *Mol. Cell Biol.* 17: 5077-86 (1997).

[0027] Although Rb is well-characterized in mammalian systems, the role of Rb-related proteins in regulation of G1 phase progression and S phase entry is not well characterized in plants. It is known, however, that RB-related protein functions through its association with various other cellular proteins involved in cell cycle regulation, such as the cyclins, WD40 proteins, Soni *et al.*, *Plant. Cell.* 7:85-103 (1995); Grafi *et al., Proc. Natl. Acad. Sci. U.S.A.* 93:8962 (1996); Ach *et al., Plant Cell* 9:1595-606 (1997); Umen and Goodenough, *Genes Dev.* 15:1652-61 (2001); Mariconti *et al., J. Biol. Chem.* 277:9911-9 (2002).

6. WD40 repeat protein

[0028] WD40 is a common repeating motif involved in many different protein-protein interactions. The WD40 domain is found in proteins having a wide variety of functions including adaptor/regulatory modules in signal transduction, pre-mRNA processing and cytoskeleton assembly. Goh *et al., Eur. J. Biochem.* 267: 434-49 (2000).

[0029] The WD40 domain, which is 40 residues long, typically contains a GH dipeptide 11-24 residues from the N-terminus and the WD dipeptide at the C-terminus. *Id.* Between the GH dipeptide and the WD dipeptide lies a conserved core which serves as a stable platform where proteins can bind either stably or reversibly. The core forms a propeller-like structure with several blades. Each blade is composed of a four-stranded anti-parallel β-sheet. Each WD40 sequence repeat forms the first three strands of one blade and the last strand in the next blade. The last C-terminal WD40 repeat completes the blade structure of the first WD40 repeat to create the

closed ring propeller-structure. The residues on the top and bottom surface of the propeller are proposed to coordinate interactions with other proteins and/or small ligands.

[0030] Studies in yeast demonstrated that Cdc20, which contains the WD40 motif, is required for the proteolysis of mitotic cyclins. This process is mediated by an ubiquitin-protein ligase called anaphase-promoting complex (APC) or cyclosome. Following ubiquitination and proteolysis by the 26S proteasome, the cell can segregate chromosomes, and exit from mitosis. Cdc20 also contains a destruction-box domain.

7. WEE1-like protein

[0031] WEE1 controls the activity of cyclin-dependent kinases. WEE1 itself is a serine/threonine kinase. Sorrell *et al.*, *Planta* 215: 518-22 (2002). The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain, sometimes combined with reversible conformational changes in the C-terminal autoregulatory tail. This process is conserved among eukaryotes, from fungi to animals and plants. Similarly, there is a high degree of homology between WEE1 proteins from various organisms. For example, there is 50% identity between the protein kinase domains of the human and maize WEE1 proteins.

[0032] Expression of WEE1 is shown to occur only in actively dividing tissues and is believed to inhibit cell division by acting as a negative regulator of mitosis. WEE1 is believed to prevent entry from G2 to M by protecting the nucleus from cytoplasmically-activated cyclin B1-complexed CDC2 before the onset of mitosis. For example, over-expression of AtWEE1 (from *Arabidopsis*) and ZmWEE1 (from *Zea cans*) in fission yeast inhibits cell division which results in elongated cells. Sun *et al.*, *Proc. Nat'l Acad. Sci.* U S A 96: 4180-5 (1999).

B. Expression Profiling and Microarray Analysis in Plant Development

[0033] The multigenic control of plant phenotype presents difficulties in determining the genes responsible for phenotypic determination. One major obstacle to identifying genes and gene expression differences that contribute to phenotype in plants is the difficulty with which the expression of more than a handful of genes can be studied concurrently. Another difficulty in identifying and understanding gene expression and the interrelationship of the genes that contribute to plant phenotype is the high degree of sensitivity to environmental factors that plants demonstrate.

[0034] There have been recent advances using genome-wide expression profiling. In particular, the use of DNA microarrays has been useful to examine the expression of a large number of genes in a single experiment. Several studies of plant gene responses to developmental and environmental stimuli have been conducted using expression profiling. For example, microarray analysis was employed to study gene expression during fruit ripening in strawberry, Aharoni *et al., Plant Physiol.* 129:1019-1031 (2002), wound response in *Arabodopsis*, Cheong *et al., Plant Physiol.* 129:661-7 (2002), pathogen response in *Arabodopsis*, Schenk *et al., Proc. Nat'l Acad. Sci.* 97:11655-60 (2000), and auxin response in soybean, Thibaud-Nissen *et al., Plant Physiol.* 132:118. Whetten *et al., Plant Mol. Biol.* 47:275-91 (2001) discloses expression profiling of cell wall biosynthetic genes in *Pinus taeda L.* using cDNA probes. Whetten *et al.* examined genes which were differentially expressed between differentiating juvenile and mature secondary xylem. Additionally, to determine the effect of certain environmental stimuli on gene expression, gene expression in compression wood was compared to normal wood. 156 of the 2300 elements examined showed differential expression. Whetten, *supra* at 285. Comparison of juvenile wood to mature wood showed 188 elements as differentially expressed. *Id.* at 286.

[0035] Although expression profiling and, in particular, DNA microarrays provide a convenient tool for genome-wide expression analysis, their use has been limited to organisms for which the complete genome sequence or a large cDNA

collection is available. *See* Hertzberg *et al., Proc. Nat'l Acad. Sci.* 98:14732-7 (2001a), Hertzberg *et al., Plant J.*, 25:585 (2001b). For example, Whetten, *supra,* states, "A more complete analysis of this interesting question awaits the completion of a larger set of both pine and poplar ESTs." Whetten *et al.* at 286. Furthermore, microarrays comprising cDNA or EST probes may not be able to distinguish genes of the same family because of sequence similarities among the genes. That is, cDNAs or ESTs, when used as microarray probes, may bind to more than one gene of the same family.

[0036] Methods of manipulating gene expression to yield a plant with a more desirable phenotype would be facilitated by a better understanding of cell cycle gene expression in various types of plant tissue, at different stages of plant development, and upon stimulation by different environmental cues. The ability to control plant architecture and agronomically important traits would be improved by a better understanding of how cell cycle gene expression effects formation of plant tissues, how cell cycle gene expression causes plant cells to enter or exit cell division, and how plant growth and the cell cycle are connected. Among the large number of genes, the expression of which can change during development of a plant, only a fraction are likely to effect phenotypic changes during any given stage of the plant development.

SUMMARY

[0037] Accordingly, there is a need for tools and methods useful in determining the changes in the expression of cell cycle genes that occur during the plant cell cycle. There is also a need for polynucleotides useful in such methods. There is a further need for methods which can correlate changes in cell cycle gene expression to phenotype or stage of plant development. There is a further need for methods of identifying cell cycle genes and gene products that impact plant phenotype, and that can be manipulated to obtain a desired phenotype.

[0038] In one aspect, the present invention provides an isolated polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260 and conservative variants thereof.

[0039] In other aspect, the present invention provides a DNA construct comprising at least one polynucleotide having the sequence of any one of SEQ ID NOs: 1-260 and conservative variants thereof.

[0039A] In a broad aspect, the invention specifically provides a DNA construct comprising a polynucleotide which comprises:

(i) the sequence of SEQ ID NO. 136 operably linked in a sense or an antisense orientation to a promoter; or

(ii) a sequence that encodes the protein sequence of SEQ ID NO. 396, operably linked in sense or antisense orientation to a promoter,

wherein the promoter in (i) or (ii) is selected from the group consisting of:

(a) a constitutive promoter;

(b) an inducible promoter;

(c) a regulatable promoter;

(d) a temporally regulated promoter, and

(e) a tissue-preferred promoter.

[0039B] In one preferred embodiment of the invention, the DNA construct is one in which the sequence that encodes the protein sequence of SEQ ID NO. 396 is operably linked in a sense orientation to the promoter. In an alternative preferred embodiment, the sequence that encodes the protein sequence of SEQ ID NO. 396 is operably linked in an antisense orientation to the promoter.

[0039C] In this aspect of the invention, it is also preferred that an RNA transcript of the polynucleotide is complementary to a nucleic acid sequence of SEQ ID NO. 136.

[0040] Another aspect of the invention is a plant cell transformed with a DNA construct comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260 and conservative variants thereof.

[0040A] In the present aspect, the invention further provides a plant cell that contains the DNA construct according to any of the embodiments mentioned in any of paragraphs [0039A] to [0039C].

[0040B] Preferably, at least one plant cell according to this aspect of the invention is in a plant, where the phenotype of the plant is different from a plant of the same species which does not comprise the plant cell, wherein the difference in phenotype is in any one or more of:

(a) lignin quality;

(b) lignin structure;

(c) wood composition;

(d) wood appearance;

(e) wood density;

(f) wood strength;

(g) wood stiffness;

(h) cellulose polymerization;

(i) fibre dimensions;

(j) lumen size;

(k) other plant components;

(l) plant cell division;

(m) plant cell development;

(n) number of cells per unit area;

(o) cell size;

(p) cell shape;

(q) cell wall composition;

(r) rate of wood formation;

(s) formation of stem defects;

(t) average microfibril angle;

(u) width of the S2 cell wall layer;

(v) rate of growth;

(w) rate of root formation ratio of root to branch vegetative development;

(x) leaf area index; and

(y) leaf shape.

[0041] A further aspect of the invention is a transgenic plant comprising a plant cell transformed with a DNA construct comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260 and conservative variants thereof.

[0041A] Preferably, the at least one plant cell that contains the DNA construct forms part of a transgenic plant, where the plant is of a species of *Eucalyptus* or *Pinus*.

[0041B] Preferably further, the transgenic plant exhibits one or more traits selected from the group consisting of:

(a) increased drought tolerance;

(b) herbicide resistance;

(c) reduced or increased height;

(d) reduced or increased branching;

(e) enhanced cold and frost tolerance;

(f) improved vigour;

(g) enhanced colour;

(h) enhanced health and nutritional characteristics;

(i) improved storage;

(j) enhanced yield;

(k) enhanced salt tolerance;

(l) enhanced resistance of the wood to decay;

(m) enhanced resistance to fungal diseases;

(n) altered attractiveness to insect pests;

(o) enhanced heavy metal tolerance;

(p) increased disease tolerance;

(q) increased insect tolerance;

(r) increased water-stress tolerance;

(s) enhanced sweetness, improved texture;

(t) decreased phosphate content;

(u) increased germination;

(v) increased micronutrient uptake;

(w) improved starch composition;

(y) improved flower longevity;

(z) production of novel resins;

(za) production of novel proteins or peptides;

(zb) reduced period of juvenility;

(zc) an increased period of juvenility;

(zd) propensity to form reaction wood;

(ze) self-abscising branches; and

(zf) accelerated reproductive development or delayed reproductive development as compared to a plant of the same species that has not been transformed with the DNA construct.

[0042] Another aspect of the invention is an isolated polynucleotide comprising a sequence encoding the catalytic or substrate-binding domain of a polypeptide selected from of any one of SEQ ID NOs: 261-520, wherein the polynucleotide encodes a polypeptide having the activity of said polypeptide selected from any one of SEQ ID NOs: 261-520.

[0043] A further aspect of the invention is a method of making a transformed plant comprising transforming a plant cell with a DNA construct comprising at least one polynucleotide having the sequence of any of SEQ ID NOs: 1-260; and culturing the transformed plant cell under conditions that promote growth of a plant.

[0044] In another aspect, the invention provides a wood obtained from a transgenic tree.

[0045] In a further aspect, the invention provides a wood pulp obtained from a transgenic tree which has been transformed with the DNA construct of the invention.

[0046] Another aspect of the invention is a method of making wood, comprising transforming a plant with a DNA construct comprising a polynucleotide having a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260 and conservative variants thereof; culturing the transformed plant under conditions that promote growth of a plant; and obtaining wood from the plant.

[0047] The invention further provides a method of making wood pulp, comprising transforming a plant with a DNA construct comprising a polynucleotide having a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260 and conservative variants thereof; culturing the transformed plant under conditions that promote growth of a plant; and obtaining wood pulp from the plant.

[0048] In another aspect, the invention provides an isolated polypeptide comprising an amino acid sequence encoded by the isolated polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260 and conservative variants thereof.

[0049] The invention also provides, an isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 261-520.

[0050] The invention further provides a method of altering a plant phenotype of a plant, comprising altering expression in the plant of a polypeptide encoded by any one of SEQ ID NOs: 1-260.

[0051] In another aspect, the invention provides a polynucleotide comprising a nucleic acid selected from the group comprising of SEQ ID NOs: 521-772.

[0052] An aspect of the invention is a method of correlating gene expression in two different samples, comprising detecting a level of expression of one or more genes encoding a product encoded by a nucleic acid sequence selected from the group

consisting of SEQ ID NOs: 1-260 and conservative variants thereof in a first sample; detecting a level of expression of the one or more genes in a second sample; comparing the level of expression of the one or more genes in the first sample to the level of expression of the one or more genes in the second sample; and correlating a difference in expression level of the one or more genes between the first and second samples.

[0053] A further aspect of the invention is a method of correlating the possession of a plant phenotype to the level of gene expression in the plant of one or more genes comprising detecting a level of expression of one or more genes encoding a product encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260 and conservative variants thereof in a first plant possessing a phenotype; detecting a level of expression of the one or more genes in a second plant lacking the phenotype; comparing the level of expression of the one or more genes in the first plant to the level of expression of the one or more genes in the second plant; and correlating a difference in expression level of the one or more genes between the first and second plants to possession of the phenotype.

[0054] In a further aspect, the invention provides a method of correlating gene expression to a stage of the cell cycle, comprising detecting a level of expression of one or more genes encoding a product encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260 and conservative variants thereof in a first plant cell in a first stage of the cell cycle; detecting a level of expression of the one or more genes in a second plant cell in a second, different stage of the cell cycle; comparing the level of the expression of the one or more genes in the first plant cells to the level of expression of the one or more genes in the second plants cells; and correlating a difference in expression level of the one or more genes between the first and second samples to the first or second stage of the cell cycle.

[0055] An aspect of the invention is a combination for detecting expression of one or more genes, comprising two or more oligonucleotides, wherein each oligonucleotide is capable of hybridizing to a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260.

[0056] Another aspect of the invention is a combination for detecting expression of one or more genes, comprising two or more oligonucleotides, wherein each oligonucleotide is capable of hybridizing to a nucleic acid sequence encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260.

[0057] The invention further provides a microarray comprising a combination for detecting expression of one or more genes, comprising two or more oligonucleotides, wherein each oligonucleotide is capable of hybridizing to a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260 or wherein each oligonucleotide is capable of hybridizing to a nucleic acid sequence encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260, wherein each of said two or more oligonucleotides occupies a unique location on said solid support.

[0058] In another aspect, the invention provides a method for detecting one or more genes in a sample, comprising contacting the sample with two or more oligonucleotides, wherein each oligonucleotide is capable of hybridizing to a gene comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260 under standard hybridization conditions; and detecting the one or more genes of interest which are hybridized to the one or more oligonucleotides.

[0059] The invention also provides a method for detecting one or more nucleic acid sequences encoded by one or more genes in a sample, comprising contacting the sample with two or more oligonucleotides, wherein each oligonucleotide is capable of hybridizing to a nucleic acid sequence encoded by a gene comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-260 under standard hybridization conditions; and detecting the one or more nucleic acid sequences which are hybridized to the one or more oligonucleotides.

[0060] The invention further provides a kit for detecting gene expression comprising the microarray of the invention together with one or more buffers or reagents for a nucleotide hybridization reaction.

[0061] Other features, objects, and advantages of the present invention are apparent from the detailed description that follows. It should be understood, however, that the detailed description, while indicating preferred embodiments of the invention, are given by way of illustration only, not limitation. Various changes and modifications within the spirit and scope of the invention will be apparent to those skilled in the art from the detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0062] Figure 1: Exemplary microarray sampling parameters.

[0063] Figure 2: Plasmid map for pWVK202.

[0064] Figure 3: Plasmid map for pGrowth14.

[0065] Figure 4: Plasmid map for pGrowth15.

[0066] Figure 5: Plasmid map for pGrowth16.

[0067] Figure 6: Plasmid map for pGrowth18.

[0068] Figure 7: Plasmid map for pGrowth19.

[0069] Figure 8: Plasmid map for pGrowth20.

List of Tables

[0070] Table 1: shows genes having greater than doubled signal with any one sample as compared to the mean signal of the other three samples.

[0071] Table 2: identifies plasmid(s), genes, and Genesis ID numbers for constructs described in Example 17.

[0072] Table 3: Rooting medium for *Populus deltoids*.

[0073] Table 4: pGrowth information.

[0074] Table 5: shows genes having greater than doubled signal with any one sample as compared to the mean signal of the other three samples.

[0075] Table 6: Differentially expressed cDNAs.

[0076] Table 7: Consensus ID information.

[0077] Table 8: pGrowth information.

[0078] Table 9: *Eucalyptus grandis* cell cycle genes and proteins.

[0079] Table 10: *Pinus radiata* cell cycle genes and proteins.

[0080] Table 11: Annotated peptide sequences of the present invention.

[0081] Table 12: *Eucalyptus in silico* data.

[0082] Table 13: Pine *in silico* data.

[0083] Table 14: Oligo table.

[0084] Table 15: Peptide table.

[0085] Table 16: BLAST sequence alignment table.

DETAILED DESCRIPTION

[0086] The inventors have discovered novel isolated cell cycle genes and polynucleotides useful for identifying the multigenic factors that contribute to a phenotype and for manipulating gene expression to affect a plant phenotype. These genes, which are derived from plants of commercially important forestry genera, pine and eucalyptus, are involved in the plant cell cycle and are, at least in part, responsible for expression of phenotypic characteristics important in commercial wood, such as stiffness, strength, density, fiber dimensions, coarseness, cellulose and lignin content, and extractives content. Generally speaking, the genes and polynucleotides encode a protein which can be a cyclin, cyclin dependent kinase, cyclin dependent kinase inhibitor, histone acetyltransferase, histone deacetylase, peptidyl-prolyl cis-trans

isomerase, retinoblastoma-related protein, WEE1-like protein, or WD40 repeat protein, or a catalytic domain thereof, or a polypeptide having the same function, and the invention further includes such proteins and polypeptides.

[0087] The methods of the present invention for selecting cell cycle gene sequences to target for manipulation will permit better design and control of transgenic plants with more highly engineered phenotypes. The ability to control plant architecture and agronomically important traits in commercially important forestry species will be improved by the information obtained from the methods, such as which genes affect which phenotypes, which genes affect entry into which stage of the cell cycle, which genes are active in which stage of plant development, and which genes are expressed in which tissue at a given point in the cell cycle or plant development.

[0088] Unless indicated otherwise, all technical and scientific terms are used herein in a manner that conforms to common technical usage. Generally, the nomenclature of this description and the described laboratory procedures, including cell culture, molecular genetics, and nucleic acid chemistry and hybridization, respectively, are well known and commonly employed in the art. Standard techniques are used for recombinant nucleic acid methods, oligonucleotide synthesis, cell culture, tissue culture, transformation, transfection, transduction, analytical chemistry, organic synthetic chemistry, chemical syntheses, chemical analysis, and pharmaceutical formulation and delivery. Generally, enzymatic reactions and purification and/or isolation steps are performed according to the manufacturers' specifications. Absent an indication to the contrary, the techniques and procedures in question are performed according to conventional methodology disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING A LABORATORY MANUAL, 2d ed. (Cold Spring Harbor Laboratory Press, 1989), and CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1989). Specific scientific methods relevant to the present invention are discussed in more detail below. However, this discussion is provided as an example only, and does not limit the manner in which the methods of the invention can be carried out.

A. Plant Cell Cycle Genes and Proteins

1. Cell Cycle Genes, Polynucleotide and Polypeptide Sequences

[0089] One aspect of the present invention relates to novel plant cell cycle genes and polypeptides encoded by such genes. As used herein, the term "plant cell cycle genes" refers to genes encoding proteins that function during the plant cell cycle, and the term "plant cell cycle proteins" refers to proteins that function during the plant cell cycle. There are several known families of plant cell cycle proteins, including cyclin, cyclin dependent kinase, cyclin dependent kinase inhibitor, histone acetyltransferase, histone deacetylase, peptidyl-prolyl cis-trans isomerase, retinoblastoma-related protein, WEE1-like protein, and WD40 repeat protein. Although there is significant sequence homology within each gene and protein family, each member of each family can display different biochemical properties and altering the expression of at least one of these genes can result in a different plant phenotype.

[0090] The present invention provides novel plant cell cycle genes and polynucleotides and novel cell cycle proteins and polypeptides. In accordance with one embodiment of the invention, the novel plant cell cycle genes are the same as those expressed in a wild-type plant of a species of *Pinus* or *Eucalyptus.* Exemplary novel plant cell cycle gene sequences of the invention are set forth in Tables 9 and 10, which depict *Eucalyptus grandis* sequences and *Pinus radiata* sequences, respectively. Corresponding gene products, *i.e.*, oligonucleotides and polypeptides, are also listed in Tables 14, 15, and 16. The Sequence Listing in APPENDIX 1 provides the sequences of these aspects of the invention.

[0091] The sequences of the invention have cell cycle activity and encode proteins that are active in the cell cycle, such as proteins of the cell cycle families discussed above. As discussed in more detail below, manipulation of the expression of the cell cycle genes and polynucleotides, or manipulation of the activity of the encoded proteins and polypeptides, can result in a transgenic plant with a desired phenotype that differs from the phenotype of a wild-type plant of the same species.

[0092] Throughout this description, reference is made to cell cycle gene products. As used herein, a "cell cycle gene product" is a product encoded by a cell cycle gene, and includes both nucleotide products, such as RNA, and amino acid products, such as proteins and polypeptides. Examples of specific cell cycle genes of the invention include SEQ ID NOs: 1-260. Examples of specific cell cycle gene products of the invention include products encoded by any one of SEQ ID NOs: 1-260. Reference also is made herein to cell cycle proteins and cell cycle polypeptides. Examples of specific cell cycle proteins and polypeptides of the invention include polypeptides encoded by any of SEQ ID NOs: 1-260 or polypeptides comprising the amino acid sequence of any of SEQ ID NOs: 261-520. One aspect of the invention is directed to a subset of these cell cycle genes and cell cycle gene products, namely SEQ ID NOs: 1-12, 14-58, 60-62, 64-70, 72-75, 77-83, 85-86, 88-91, 93-119, 121-130, 132-148, 150-156, 158-191, 193-207, 209-218, 220-221, 223-231, 233-237, their respective conservative variants (as that term is defined below), and the nucleotide and amino acid products encoded thereby. Another aspect of the invention is directed to a subset of the cell cycle genes and cell cycle gene products, namely SEQ ID NOs: 1-12, 14, 16-26, 30-37, 40-41, 43-76, 78-103, 106, 108-113, 116-121, 124-125, 128-147, 150-152, 154-155, 161-162, 164-172, 174, 177-183, 185-191, 193-197, 200-204, 208-213, and 215-234 their respective conservative variants, and the nucleotide and amino acid products encoded thereby. A further aspect of the invention is directed to a subset of the cell cycle genes and cell cycle gene products, namely SEQ ID NOs: 1-12, 14, 16-26, 30-37, 40-41, 43-58, 60-62, 64-70, 72-75, 78-83, 85-86, 88-91, 93-103, 106, 108-113, 116-119, 121, 124-125, 128-130, 132-147, 150-152, 154-155, 161-162, 164-172, 174, 177-183, 185-191, 193-197, 200-204, 209-213, 215-218, 220-221, 223-231, and 233-234 their respective conservative variants, and the nucleotide and amino acid products encoded thereby.

[0093] The present invention also includes sequences that are complements, reverse sequences, or reverse complements to the nucleotide sequences disclosed herein.

[0094] The present invention also includes conservative variants of the sequences disclosed herein. The term "variant," as used herein, refers to a nucleotide or amino acid sequence that differs in one or more nucleotide bases or amino acid residues from the reference sequence of which it is a variant.

[0095] Thus, in one aspect, the invention includes conservative variant polynucleotides. As used herein, the term "conservative variant polynucleotide" refers to a polynucleotide that hybridizes under stringent conditions to an oligonucleotide probe that, under comparable conditions, binds to the reference gene the conservative variant is a variant of. Thus, for example, a conservative variant of SEQ ID NO: 1 hybridizes under stringent conditions to an oligonucleotide probe that, under comparable conditions, binds to SEQ ID NO: 1. One aspect of the invention provides conservative variant polynucleotides that exhibit at least about 75% sequence identity to their respective reference sequences.

[0096] "Sequence identity" has an art-recognized meaning and can be calculated using published techniques. See COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, *ed.* (Oxford University Press, 1988), BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, *ed.* (Academic Press, 1993), COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin & Griffin, *eds.*, (Humana Press, 1994), SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, Von Heinje *ed.*, Academic Press (1987), SEQUENCE ANALYSIS PRIMER, Gribskov & Devereux, *eds.* (Macmillan Stockton Press, 1991), and Carillo & Lipton, *SIAM J. Applied Math.* 48: 1073 (1988). Methods commonly employed to determine identity or similarity between two sequences include but are not limited to those disclosed in GUIDE TO HUGE COMPUTERS, Bishop, *ed.*, (Academic Press, 1994) and Carillo & Lipton, *supra.* Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include but are not limited to the GCG program package (Devereux *et al., Nucleic Acids Research* 12: 387 (1984)), BLASTP, BLASTN, FASTA (Atschul *et al., J. Mol. Biol.* 215: 403 (1990)), and FASTDB (Brutlag *et al., Comp. App. Biosci.* 6: 237 (1990)).

[0097] The invention includes conservative variant polynucleotides having a sequence identity that is greater than or equal to 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, or 60% to any one of SEQ ID NOs: 1 to 260. In such variants, differences between the variant and the reference sequence can occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

[0098] Additional conservative variant polynucleotides contemplated by and encompassed within the present invention include polynucleotides comprising sequences that differ from the polynucleotide sequences of SEQ ID NO: 1-260, or complements, reverse complements or reverse sequences thereof, as a result of deletions and/or insertions totaling less than 10% of the total sequence length.

[0099] The invention also includes conservative variant polynucleotides that, in addition to sharing a high degree of similarity in their primary structure (sequence) to SEQ ID NOs: 1 to 260, have at least one of the following features: (i) they contain an open reading frame or partial open reading frame encoding a polypeptide having substantially the same functional properties in the cell cycle as the polypeptide encoded by the reference polynucleotide, or (ii) they have nucleotide domains or encoded protein domains in common. The invention includes conservative variants of SEQ ID NOs: 1-260 that encode proteins having the enzyme or biological activity or binding properties of the protein encoded by the reference polynucleotide. Such conservative variants are functional variants, in that they have the enzymatic or binding activity of the protein encoded by the reference polynucleotide.

[0100] In accordance with the invention, polynucleotide variants can include a "shuffled gene" such as those described in *e.g.* U.S. Patent Nos. 6,500,639, 6,500,617 6,436,675, 6,379,964, 6,352,859 6,335,198 6,326,204, and 6,287,862. A variant of a nucleotide sequence of the present invention also can be a polynucleotide

modified as disclosed in U.S. Patent No. 6,132,970, which is incorporated herein by reference.

[0101] In accordance with one embodiment, the invention provides a polynucleotide that encodes a cell cycle protein from one of the following families: cyclin, cyclin dependent kinase, cyclin dependent kinase inhibitor, histone acetyltransferase, histone deacetylase, peptidyl-prolyl cis-trans isomerase, retinoblastoma-related protein, WEE1-like protein, or WD40 repeat protein. SEQ ID NOs: 1-260 provide examples of such polynucleotides.

[0102] In accordance with another embodiment, a polynucelotide of the invention encodes the catalytic or protein binding domain of a polypeptide encoded by any of SEQ ID NOs: 1-260 or of a polypeptide comprising any of SEQ ID NOs: 261-520. The catalytic and protein binding domains of the cell cycle proteins of the invention are known in the art. The conserved sequences of these proteins are shown in Entries 1-248 as underlined, bold, and/or italicized text.

[0103] The invention also encompasses as conservative variants polynucleotides that differ from the sequences discussed above but that, as a consequence of the degeneracy of the genetic code, encode a polypeptide which is the same as that encoded by a polynucleotide of the present invention. The invention also includes as conservative variants polynucleotides comprising sequences that differ from the polynucleotide sequences discussed above as a result of substitutions that do not affect the amino acid sequence of the encoded polypeptide sequence, or that result in conservative substitutions in the encoded polypeptide sequence.

[0104] The present invention also includes an isolated polypeptide encoded by a polynucleotide comprising any of SEQ ID NOs: 1-260 or any of the conservative variants thereof discussed above. The invention also includes polypeptides comprising SEQ ID NOs: 261-520 and conservative variants of these polypeptides. Another aspect of the invention include polypeptides comprising SEQ ID NOs: 261-272, 274-318, 320-322, 324-330, 332-335, 337-343, 345-346, 348-351, 353-379, 381-390, 392-408, 410-416, 418-451, 453-467, 469-478, 480-481, 483-491, and 493-494 and conservative variants thereof. A further aspect of

the invention includes polypeptides comprising SEQ ID NOs: 261-272, 274, 276-286, 289, 290-297, 300-301, 303-345, 347-363, 366, 368-373, 376-381, 384-385, 388-407, 410-412, 414-415, 420-422, 424-432, 434, 437-443, 445-451, 453-457, 460-464, 468-473, and 475-494 and conservative variants thereof. Another aspect of the invention includes polypeptides comprising SEQ ID NOs: 261-272, 274, 276-286, 290-297, 300-301, 303-318, 320-322, 324-330, 332-335, 337-343, 345, 348-351, 353-363, 366, 368-373, 376-381, 384-385, 388-390, 392-407, 410-412, 414-415, 421-422, 424-432, 434, 437-443, 445-451, 453-457, 460-464, 469-473, 475-478, 480-481, 483-491, and 493-494 and conservative variants thereof.

[0105] In accordance with the invention, a variant polypeptide or protein refers to an amino acid sequence that is altered by the addition, deletion or substitution of one or more amino acids.

[0106] The invention includes conservative variant polypeptides. As used herein, the term "conservative variant polypeptide" refers to a polypeptide that has similar structural, chemical or biological properties to the protein it is a conservative variant of. Guidance in determining which amino acid residues can be substituted, inserted, or deleted can be found using computer programs well known in the art such as Vector NTI Suite (InforMax, MD) software. In one embodiment of the invention, conservative variant polypeptides that exhibit at least about 75% sequence identity to their respective reference sequences.

[0107] Conservative variant protein includes an "isoform" or "analog" of the polypeptide. Polypeptide isoforms and analogs refers to proteins having the same physical and physiological properties and the same biological function, but whose amino acid sequences differs by one or more amino acids or whose sequence includes a non-natural amino acid.

[0108] Polypeptides comprising sequences that differ from the polypeptide sequences of SEQ ID NO: 261-520 as a result of amino acid substitutions, insertions, and/or deletions totaling less than 10% of the total sequence length are

2004311384 31 May 2011

contemplated by and encompassed within the present invention.

[0109] One aspect of the invention provides conservative variant polypeptides that have the same function in the cell cycle as the proteins of which they are variants, as determined by one or more appropriate assays, such as those described below. The invention includes variant polypeptides that function as cell cycle proteins, such as those having the biological activity of cyclin, cyclin dependent kinase, cyclin dependent kinase inhibitor, histone acetyltransferase, histone deacetylase, peptidyl-prolyl cis-trans isomerase, retinoblastoma-related protein, WEE1-like protein, and WD40 repeat protein, and are thus capable of modulating the cell cycle in a plant. As discussed above, the invention includes variant polynucleotides that encode polypeptides that function as cell cycle proteins.

[0110] The activities and physical properties of cell cycle proteins can be examined using any method known in the art. The following examples of assay methods are not exhaustive and are included to provide some guidance in examining the activity and distinguishing protein characteristics of cell cycle protein variants.

[0111] CDK activity can be assessed using roscovitine as described in Yamaguchi *et al., Proc. Natl. Acad. Sci. U.S.A.* 100:8019 (2003). CDK histone kinase activity can be assayed using autoradiography to detect histone H1 phosphorylation by CDK as described in Joubés *et al., Plant Physiol.* 121:857 (1999).

[0112] CKI activity can be assayed using a variation of the method described in Zhou *et al., Planta.* 6:604 (2003). The modified method can employ co-transformation or subsequent transformations to identify the interaction of CKI and cyclins *in vivo*. For example, in the first transformation pine tissue can be transformed using the method described in U.S. Patent Application Publication No. 2002/0100083 using geneticin selection to obtain transgenic plants possessing cycD3 and cdc2a homologs. The second transformation can be performed using alpha-methyltryptophan as a selectable marker to obtain transformants having an ICK1 homologue as described in U.S. Provisional Application No. 60/476,189. Tissue

capable of growing on both on geneticin and on alpha-methyltryptophan contains the ICK1 homologue and the cycD3 and cdc2a homologues. The CKI activity is determined by comparison of the phenotype of transformants having the cycD3 and cdc2a homologues to the transformants having ICK1 homologue and the cycD3 and cdc2a homologs.

[0113] Histone deacetylase activity can be assessed by complementation of the *Arabidopsis* mutants described in Tian *et al., Genetics* 165:399 (2003). Histone acetyltransferase activity can be assayed using anacardic acid as described in Balasubramanyam *et al., J. Biol. Chem.* 278:19134 (2003). Histone acetyltransferase also can be assayed using trichostatin A-treated plant lines as is described in Bhat *et al., Plant J.* 33:455 (2003). The plant lines described in Bhat *et al., supra,* also can be used to assay retinoblastoma-related proteins using the co-precipitation method described in Rossi *et al., Plant Mol. Biol.* 51:401 (2003).

[0114] Peptidyl-prolyl isomerase can be assayed as described in Edvardsson *et al., FEBS Lett.* 542:137 (2003). WD40 proteins can be evaluated based on the possession of the WD40 motif as well as their ability to interact with cdc2. WEE-1 can be assayed using any kinase activity assay known in the art.

2. Methods of Using Cell Cycle Genes, Polynucleotide and Polypeptide Sequences

[0115] The present invention provides methods of using plant cell cycle genes and conservative variants thereof. The invention includes methods and constructs for altering expression of plant cell cycle genes and/or gene products for purposes including, but not limited to (i) investigating function during the cell cycle and ultimate effect on plant phenotype and (ii) to effect a change in plant phenotype. For example, the invention includes methods and tools for modifying wood quality, fiber development, cell wall polysaccharide content, fruit ripening, and plant growth and yield by altering expression of one or more plant cell cycle genes.

[0116] The invention comprises methods of altering the expression of any of the cell cycle genes and variants discussed above. Thus, for example, the invention

comprises altering expression of a cell cycle gene present in the genome of a wild-type plant of a species of *Eucalyptus* or *Pinus*. In one embodiment, the cell cycle gene comprises a nucleotide sequence selected from SEQ ID NOs: 1-260, from the subset thereof comprising SEQ ID NOs: 1-12, 14-58, 60-62, 64-70, 72-75, 77-83, 85-86, 88-91, 93-119, 121-130, 132-148, 150-156, 158-191, 193-207, 209-218, 220-221, 223-231, and 233-237, from the subset thereof comprising SEQ ID NOs: 1-12, 14, 16-26, 30-37, 40-41, 43-76, 78-103, 106, 108-113, 116-121, 124-125, 128-147, 150-152, 154-155, 161-162, 164-172, 174, 177-183, 185-191, 193-197, 200-204, 208-213, and 215-234, from the subset thereof comprising SEQ ID NOs: 1-12, 14, 16-26, 30-37, 40-41, 43-58, 60-62, 64-70, 72-75, 78-83, 85-86, 88-91, 93-103, 106, 108-113, 116-119, 121, 124-125, 128-130, 132-147, 150-152, 154-155, 161-162, 164-172, 174, 177-183, 185-191, 193-197, 200-204, 209-213, 215-218, 220-221, 223-231, and 233-234, or the conservative variants thereof, as discussed above.

[0117] Techniques which can be employed in accordance with the present invention to alter gene expression, include, but are not limited to: (i) over-expressing a gene product, (ii) disrupting a gene's transcript, such as disrupting a gene's mRNA transcript; (iii) disrupting the function of a polypeptide encoded by a gene, or (iv) disrupting the gene itself. Over-expression of a gene product, the use of antisense RNAs, ribozymes, and the use of double-stranded RNA interference (dsRNAi) are valuable techniques for discovering the functional effects of a gene and for generating plants with a phenotype that is different from a wild-type plant of the same species.

[0118] Over-expression of a target gene often is accomplished by cloning the gene or cDNA into an expression vector and introducing the vector into recipient cells. Alternatively, over-expression can be accomplished by introducing exogenous promoters into cells to drive expression of genes residing in the genome. The effect of over-expression of a given gene on cell function, biochemical and/or physiological properties can then be evaluated by comparing plants transformed to over-express the gene to plants that have not been transformed to over-express the gene.

[0119] Antisense RNA, ribozyme, and dsRNAi technologies typically target RNA transcripts of genes, usually mRNA. Antisense RNA technology involves

2004311384 31 May 2011

expressing in, or introducing into, a cell an RNA molecule (or RNA derivative) that is complementary to, or antisense to, sequences found in a particular mRNA in a cell. By associating with the mRNA, the antisense RNA can inhibit translation of the encoded gene product. The use of antisense technology to reduce or inhibit the expression of specific plant genes has been described, for example in European Patent Publication No. 271988, Smith *et al., Nature*, 334:724-726 (1988); Smith *et. al., Plant Mol. Biol.*, 14:369-379 (1990)).

[0120] A ribozyme is an RNA that has both a catalytic domain and a sequence that is complementary to a particular mRNA. The ribozyme functions by associating with the mRNA (through the complementary domain of the ribozyme) and then cleaving (degrading) the message using the catalytic domain.

[0121] RNA interference (RNAi) involves a post-transcriptional gene silencing (PTGS) regulatory process, in which the steady-state level of a specific mRNA is reduced by sequence-specific degradation of the transcribed, usually fully processed mRNA without an alteration in the rate of *de novo* transcription of the target gene itself. The RNAi technique is discussed, for example, in Elibashir, *et al.*, *Methods Enzymol.* 26: 199 (2002); McManus & Sharp, *Nature Rev. Genetics* 3: 737 (2002); PCT application WO 01/75164; Martinez *et al.*, *Cell* 110: 563 (2002); Elbashir *et al., supra*; Lagos-Quintana *et al.*, *Curr. Biol.* 12: 735 (2002); Tuschl *et al.*, *Nat. Biotechnol.* 20:446 (2002); Tuschl, *Chembiochem.* 2: 239 (2001); Harborth *et al.*, *J. Cell Sci.* 114: 4557 (2001); *et al.*, *EMBO J.* 20:6877 (2001); Lagos-Quintana *et al.*, *Science.* 294: 8538 (2001); Hutvagner *et al., loc cit,* 834; Elbashir *et al.*, *Nature.* 411: 494 (2001).

[0122] The present invention provides a DNA construct comprising at least one polynucleotide of SEQ ID NOs: 1-260 or conservative variants thereof, such as the conservative variants discussed above. Any method known in the art can be used to generate the DNA constructs of the present invention. *See, e.g.* Sambrook *et al.*, *supra.*

[0123] The invention includes DNA constructs that optionally comprise a promoter. Any suitable promoter known in the art can be used. A promoter is a nucleic acid, preferably DNA, that binds RNA polymerase and/or other transcription regulatory elements. As with any promoter, the promoters of the invention facilitate or control the transcription of DNA or RNA to generate an mRNA molecule from a nucleic acid molecule that is operably linked to the promoter. The RNA can encode a protein or polypeptide or can encode an antisense RNA molecule or a molecule useful in RNAi. Promoters useful in the invention include constitutive promoters, inducible promoters, temporally regulated promoters and tissue-preferred promoters.

[0124] Examples of useful constitutive plant promoters include: the cauliflower mosaic virus (CaMV) 35S promoter, which confers constitutive, high-level expression in most plant tissues (Odel *et al. Nature* 313:810(1985)); the nopaline synthase promoter (An *et al. Plant Physiol.* 88:547 (1988)); and the octopine synthase promoter (Fromm *et al., Plant Cell* 1: 977 (1989)). It should be noted that, although the CaMV 35S promoter is commonly referred to as a constitutive promoter, some tissue preference can be seen. The use of CaMV 35S is envisioned by the present invention, regardless of any tissue preference which may be exhibited during use in the present invention.

[0125] Inducible promoters regulate gene expression in response to environmental, hormonal, or chemical signals. Examples of hormone inducible promoters include auxin-inducible promoters (Baumann *et al. Plant Cell* 11:323-334(1999)), cytokinin-inducible promoters (Guevara-Garcia, *Plant Mol. Biol.* 38:743-753(1998)), and gibberellin-responsive promoters (Shi *et al. Plant Mol. Biol.* 38:1053-1060(1998)). Additionally, promoters responsive to heat, light, wounding, pathogen resistance, and chemicals such as methyl jasmonate or salicylic acid, can be used in the DNA constructs and methods of the present invention.

[0126] Tissue-preferred promoters allow for preferred expression of polynucleotides of the invention in certain plant tissue. Tissue-preferred promoters are also useful for directing the expression of antisense RNA or siRNA in certain plant tissues, which can be useful for inhibiting or completely blocking the expression

of targeted genes as discussed above. As used herein, vascular plant tissue refers to xylem, phloem or vascular cambium tissue. Other preferred tissue includes apical meristem, root, seed, and flower. In one aspect, the tissue-preferred promoters of the invention are either "xylem-preferred," "cambium-preferred" or "phloem-preferred," and preferentially direct expression of an operably linked nucleic acid sequence in the xylem, cambium or phloem, respectively. In another aspect, the DNA constructs of the invention comprise promoters that are tissue-specific for xylem, cambium or phloem, wherein the promoters are only active in the xylem, cambium or phloem.

[0127] A vascular-preferred promoter is preferentially active in any of the xylem, phloem or cambium tissues, or in at least two of the three tissue types. A vascular-specific promoter is specifically active in any of the xylem, phloem or cambium, or in at least two of the three. In other words, the promoters are only active in the xylem, cambium or phloem tissue of plants. Note, however, that because of solute transport in plants, a product that is specifically or preferentially expressed in a tissue may be found elsewhere in the plant after expression has occurred.

[0128] In another embodiment, the promoter is under temporal regulation, wherein the ability of the promoter to initiate expression is linked to factors such as the stage of the cell cycle or the stage of plant development. For example, the promoter of a cyclin D2 gene may be expressed only during the G1 and early S-phase, and the promoters of particular cyclin genes may be expressed only within the primary vascular poles of the developing seedling.

[0129] Additionally, the promoters of particular cell cycle genes may be expressed only within the cambium in developing secondary vasculature. Within the cambium, particular cell cycle gene promoters may be expressed exclusively in the stem or in the root. Moreover, the cell cycle promoters may be expressed only in the spring (for early wood formation) or only in the summer.

[0130] A promoter may be operably linked to the polynucleotide. As used in this context, operably linked refers to linking a polynucleotide encoding a structural gene to a promoter such that the promoter controls transcription of the structural gene.

If the desired polynucleotide comprises a sequence encoding a protein product, the coding region can be operably linked to regulatory elements, such as to a promoter and a terminator, that bring about expression of an associated messenger RNA transcript and/or a protein product encoded by the desired polynucleotide. In this instance, the polynucleotide is operably linked in the 5'- to 3'- orientation to a promoter and, optionally, a terminator sequence.

[0131] Alternatively, the invention provides DNA constructs comprising a polynucleotide in an "antisense" orientation, the transcription of which produces nucleic acids that can form secondary structures that affect expression of an endogenous cell cycle gene in the plant cell. In another variation, the DNA construct may comprise a polynucleotide that yields a double-stranded RNA product upon transcription that initiates RNA interference of a cell cycle gene with which the polynucleotide is associated. A polynucleotide of the present invention can be positioned within a t-DNA, such that the left and right t-DNA border sequences flank or are on either side of the polynucleotide.

[0132] It should be understood that the invention includes DNA constructs comprising one or more of any of the polynucleotides discussed above. Thus, for example, a construct may comprise a t-DNA comprising one, two, three, four, five, six, seven, eight, nine, ten, or more polynucleotides.

[0133] The invention also includes DNA constructs comprising a promoter that includes one or more regulatory elements. Alternatively, the invention includes DNA constructs comprising a regulatory element that is separate from a promoter. Regulatory elements confer a number of important characteristics upon a promoter region. Some elements bind transcription factors that enhance the rate of transcription of the operably linked nucleic acid. Other elements bind repressors that inhibit transcription activity. The effect of transcription factors on promoter activity can determine whether the promoter activity is high or low, *i.e.* whether the promoter is "strong" or "weak."

[0134] A DNA construct of the invention can include a nucleotide sequence that serves as a selectable marker useful in identifying and selecting transformed plant cells or plants. Examples of such markers include, but are not limited to, a neomycin phosphotransferase (*nptII*) gene (Potrykus *et al., Mol. Gen. Genet.* 199:183-188 (1985)), which confers kanamycin resistance. Cells expressing the *nptII* gene can be selected using an appropriate antibiotic such as kanamycin or G418. Other commonly used selectable markers include a mutant EPSP synthase gene (Hinchee *et al., Bio/Technology* 6:915-922 (1988)), which confers glyphosate resistance; and a mutant acetolactate synthase gene (ALS), which confers imidazolinone or sulphonylurea resistance (European Patent Application 154,204, 1985).

[0135] The present invention also includes vectors comprising the DNA constructs discussed above. The vectors can include an origin of replication (replicons) for a particular host cell. Various prokaryotic replicons are known to those skilled in the art, and function to direct autonomous replication and maintenance of a recombinant molecule in a prokaryotic host cell.

[0136] In one embodiment, the present invention utilizes a pWVR8 vector as described in U.S. Application No. 60/476,222, filed June 6, 2003, or pART27 as described in Gleave, *Plant Mol. Biol*, 20:1203-27 (1992).

[0137] The invention also provides host cells which are transformed with the DNA constructs of the invention. As used herein, a host cell refers to the cell in which a polynucleotide of the invention is expressed. Accordingly, a host cell can be an individual cell, a cell culture or cells that are part of an organism. The host cell can also be a portion of an embryo, endosperm, sperm or egg cell, or a fertilized egg. In one embodiment, the host cell is a plant cell.

[0138] The present invention further provides transgenic plants comprising the DNA constructs of the invention. The invention includes transgenic plants that are angiosperms or gymnosperms. The DNA constructs of the present invention can be used to transform a variety of plants, both monocotyledonous (*e.g.* grasses, corn, grains, oat, wheat and barley), dicotyledonous (*e.g., Arabidopsis*, tobacco, legumes,

alfalfa, oaks, eucalyptus, maple), and Gymnosperms (*e.g.*, Scots pine; *see* Aronen, *Finnish Forest Res. Papers*, Vol. 595, 1996), white spruce (Ellis *et al.*, *Biotechnology* 11:84-89, 1993), and larch (Huang *et al.*, *In Vitro Cell* 27:201-207, 1991).

[0139] The plants also include turfgrass, wheat, maize, rice, sugar beet, potato, tomato, lettuce, carrot, strawberry, cassava, sweet potato, geranium, soybean, and various types of woody plants. Woody plants include trees such as palm oak, pine, maple, fir, apple, fig, plum and acacia. Woody plants also include rose and grape vines.

[0140] In one embodiment, the DNA constructs of the invention are used to transform woody plants, *i.e.*, trees or shrubs whose stems live for a number of years and increase in diameter each year by the addition of woody tissue. The invention includes methods of transforming plants including eucalyptus and pine species of significance in the commercial forestry industry such as plants selected from the group consisting of *Eucalyptus grandis* and its hybrids, and *Pinus taeda*, as well as the transformed plants and wood and wood pulp derived therefrom. Other examples of suitable plants include those selected from the group consisting of *Pinus banksiana, Pinus brutia, Pinus caribaea, Pinus clausa, Pinus contorta, Pinus coulteri, Pinus echinata, Pinus eldarica, Pinus ellioti, Pinus jeffreyi, Pinus lambertiana, Pinus massoniana, Pinus monticola, Pinus nigra, Pinus palustris, Pinus pinaster, Pinus ponderosa, Pinus radiata, Pinus resinosa, Pinus rigida, Pinus serotina, Pinus strobus, Pinus sylvestris, Pinus taeda, Pinus virginiana, Abies amabilis, Abies balsamea, Abies concolor, Abies grandis, Abies lasiocarpa, Abies magnifica, Abies procera, Chamaecyparis lawsoniona, Chamaecyparis nootkatensis, Chamaecyparis thyoides, Juniperus virginiana, Larix decidua, Larix laricina, Larix leptolepis, Larix occidentalis, Larix siberica, Libocedrus decurrens, Picea abies, Picea engelmanni, Picea glauca, Picea mariana, Picea pungens, Picea rubens, Picea sitchensis, Pseudotsuga menziesii, Sequoia gigantea, Sequoia sempervirens, Taxodium distichum, Tsuga canadensis, Tsuga heterophylla, Tsuga mertensiana, Thuja occidentalis, Thuja plicata, Eucalyptus alba, Eucalyptus bancroftii, Eucalyptus botryoides, Eucalyptus bridgesiana, Eucalyptus calophylla, Eucalyptus*

camaldulensis, Eucalyptus citriodora, Eucalyptus cladocalyx, Eucalyptus coccifera, Eucalyptus curtisii, Eucalyptus dalrympleana, Eucalyptus deglupta, Eucalyptus delagatensis, Eucalyptus diversicolor, Eucalyptus dunnii, Eucalyptus ficifolia, Eucalyptus globulus, Eucalyptus gomphocephala, Eucalyptus gunnii, Eucalyptus henryi, Eucalyptus laevopinea, Eucalyptus macarthurii, Eucalyptus macrorhyncha, Eucalyptus maculata, Eucalyptus marginata, Eucalyptus megacarpa, Eucalyptus melliodora, Eucalyptus nicholii, Eucalyptus nitens, Eucalyptus nova-angelica, Eucalyptus obliqua, Eucalyptus occidentalis, Eucalyptus obtusiflora, Eucalyptus oreades, Eucalyptus pauciflora, Eucalyptus polybractea, Eucalyptus regnans, Eucalyptus resinifera, Eucalyptus robusta, Eucalyptus rudis, Eucalyptus saligna, Eucalyptus sideroxylon, Eucalyptus stuartiana, Eucalyptus tereticornis, Eucalyptus torelliana, Eucalyptus urnigera, Eucalyptus urophylla, Eucalyptus viminalis, Eucalyptus viridis, Eucalyptus wandoo, and Eucalyptus youmanni.

[0141] As used herein, the term "plant" also is intended to include the fruit, seeds, flower, strobilus, etc. of the plant. A transformed plant of the current invention can be a direct transfectant, meaning that the DNA construct was introduced directly into the plant, such as through *Agrobacterium*, or the plant can be the progeny of a transfected plant. The second or subsequent generation plant can be produced by sexual reproduction, *i.e.*, fertilization. Furthermore, the plant can be a gametophyte (haploid stage) or a sporophyte (diploid stage).

[0142] As used herein, the term "plant tissue" encompasses any portion of a plant, including plant cells. Plant cells include suspension cultures, callus, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, seeds and microspores. Plant tissues can be grown in liquid or solid culture, or in soil or suitable media in pots, greenhouses or fields. As used herein, "plant tissue" also refers to a clone of a plant, seed, progeny, or propagule, whether generated sexually or asexually, and descendents of any of these, such as cuttings or seeds.

[0143] In accordance with one aspect of the invention, a transgenic plant that has been transformed with a DNA construct of the invention has a phenotype that is different from a plant that has not been transformed with the DNA construct.

[0144] As used herein, "phenotype" refers to a distinguishing feature or characteristic of a plant which can be altered according to the present invention by integrating one or more DNA constructs of the invention into the genome of at least one plant cell of a plant. The DNA construct can confer a change in the phenotype of a transformed plant by modifying any one or more of a number of genetic, molecular, biochemical, physiological, morphological, or agronomic characteristics or properties of the transformed plant cell or plant as a whole.

[0145] In one embodiment, transformation of a plant with a DNA construct of the present invention can yield a phenotype including, but not limited to any one or more of increased drought tolerance, herbicide resistance, reduced or increased height, reduced or increased branching, enhanced cold and frost tolerance, improved vigor, enhanced color, enhanced health and nutritional characteristics, improved storage, enhanced yield, enhanced salt tolerance, enhanced resistance of the wood to decay, enhanced resistance to fungal diseases, altered attractiveness to insect pests, enhanced heavy metal tolerance, increased disease tolerance, increased insect tolerance, increased water-stress tolerance, enhanced sweetness, improved texture, decreased phosphate content, increased germination, increased micronutrient uptake, improved starch composition, improved flower longevity, production of novel resins, and production of novel proteins or peptides.

[0146] In another embodiment, the affected phenotype includes one or more of the following traits: propensity to form reaction wood, a reduced period of juvenility, an increased period of juvenility, self-abscising branches, accelerated reproductive development or delayed reproductive development, as compared to a plant of the same species that has not been transformed with the DNA construct.

[0147] In a further embodiment, the phenotype that is different in the transgenic plant includes one or more of the following: lignin quality, lignin structure, wood composition, wood appearance, wood density, wood strength, wood stiffness, cellulose polymerization, fiber dimensions, lumen size, other plant components, plant cell division, plant cell development, number of cells per unit area, cell size, cell shape, cell wall composition, rate of wood formation, aesthetic appearance of wood,

formation of stem defects, average microfibril angle, width of the S2 cell wall layer, rate of growth, rate of root formation ratio of root to branch vegetative development, leaf area index, and leaf shape.

[0148] Phenotype can be assessed by any suitable means. The plants can be evaluated based on their general morphology. Transgenic plants can be observed with the naked eye, can be weighed and their height measured. The plant can be examined by isolating individual layers of plant tissue, namely phloem and cambium, which is further sectioned into meristematic cells, early expansion, late expansion, secondary wall formation, and late cell maturation. *See, e.g.,* Hertzberg, *supra.* The plants also can be assessed using microscopic analysis or chemical analysis.

[0149] Microscopic analysis includes examining cell types, stage of development, and stain uptake by tissues and cells. Fiber morphology, such as fiber wall thickness and microfibril angle of wood pulp fibers can be observed using, for example, microscopic transmission ellipsometry. *See* Ye and Sundström, *Tappi J.*, 80:181 (1997). Wood strength, density, and grain slope in wet wood and standing trees can be determined by measuring the visible and near infrared spectral data in conjunction with multivariate analysis. *See,* U.S. Patent Application Publication Nos. 2002/0107644 and 2002/0113212. Lumen size can be measured using scanning electron microscopy. Lignin structure and chemical properties can be observed using nuclear magnetic resonance spectroscopy as described in Marita *et al., J. Chem. Soc., Perkin Trans. I* 2939 (2001).

[0150] The biochemical characteristic of lignin, cellulose, carbohydrates and other plant extracts can be evaluated by any standard analytical method known including spectrophotometry, fluorescence spectroscopy, HPLC, mass spectroscopy, and tissue staining methods.

[0151] As used herein, "transformation" refers to a process by which a nucleic acid is inserted into the genome of a plant cell. Such insertion encompasses stable introduction into the plant cell and transmission to progeny. Transformation also refers to transient insertion of a nucleic acid, wherein the resulting transformant

transiently expresses the nucleic acid. Transformation can occur under natural or artificial conditions using various methods well known in the art. Transformation can be achieved by any known method for the insertion of nucleic acid sequences into a prokaryotic or eukaryotic host cell, including *Agrobacterium*-mediated transformation protocols, viral infection, whiskers, electroporation, microinjection, polyethylene glycol-treatment, heat shock, lipofection, and particle bombardment. Transformation can also be accomplished using chloroplast transformation as described in *e.g.* Svab *et al., Proc. Natl Acad. Sci.* 87:8526-30 (1990).

[0152] In accordance with one embodiment of the invention, transformation in *Eucalyptus* is performed as described in U.S. Patent Application No. 60/476,222 (*supra*) which is incorporated herein by reference in its entirety. In accordance with another embodiment, transformation of *Pinus* is accomplished using the methods described in U.S. Patent Application Publication No. 2002/0100083.

[0153] Another aspect of the invention provides methods of obtaining wood and/or making wood pulp from a plant transformed with a DNA construct of the invention. Methods of producing a transgenic plant are provided above and are known in the art. A transformed plant can be cultured or grown under any suitable conditions. For example, pine can be cultured and grown as described in U.S. Patent Application Publication No. 2002/0100083. *Eucalyptus* can be cultured and grown as in, for example, Rydelius, *et al.,* GROWING EUCALYPTUS FOR PULP AND ENERGY, presented at the Mechanization in Short Rotation, Intensive Culture Forestry Conference, Mobile, AL, 1994. Wood and wood pulp can be obtained from the plant by any means known in the art.

[0154] As noted above, the wood or wood pulp obtained in accordance with this invention may demonstrate improved characteristics including, but not limited to any one or more of lignin composition, lignin structure, wood composition, cellulose polymerization, fiber dimensions, ratio of fibers to other plant components, plant cell division, plant cell development, number of cells per unit area, cell size, cell shape, cell wall composition, rate of wood formation, aesthetic appearance of wood, formation of stem defects, rate of growth, rate of root formation ratio of root to branch

vegetative development, leaf area index, and leaf shape include increased or decreased lignin content, increased accessibility of lignin to chemical treatments, improved reactivity of lignin, increased or decreased cellulose content increased dimensional stability, increased tensile strength, increased shear strength, increased compression strength, increased shock resistance, increased stiffness, increased or decreased hardness, decreased spirality, decreased shrinkage, and differences in weight, density, and specific gravity.

B. Expression Profiling of Cell Cycle Genes

[0155] The present invention also provides methods and tools for performing expression profiling of cell cycle genes. Expression profiling is useful in determining whether genes are transcribed or translated, comparing transcript levels for particular genes in different tissues, genotyping, estimating DNA copy number, determining identity of descent, measuring mRNA decay rates, identifying protein binding sites, determining subcellular localization of gene products, correlating gene expression to a phenotype or other phenomenon, and determining the effect on other genes of the manipulation of a particular gene. Expression profiling is particularly useful for identifying gene expression in complex, multigenic events. For this reason, expression profiling is useful in correlating gene expression to plant phenotype and formation of plant tissues and the interconnection thereof to the cell cycle.

[0156] Only a small fraction of the genes of a plant's genome are expressed at a given time in a given tissue sample, and all of the expressed genes may not affect the plant phenotype. To identify genes capable of affecting a phenotype of interest, the present invention provides methods and tools for determining, for example, a gene expression profile at a given point in the cell cycle, a gene expression profile at a given point in plant development, and a gene expression profile a given tissue sample. The invention also provides methods and tools for identifying cell cycle genes whose expression can be manipulated to alter plant phenotype or to alter the biological activity of cell cycle gene products. In support of these methods, the invention also provides methods and tools that distinguish expression of different genes of the same family.

[0157] As used herein, "gene expression" refers to the process of transcription of a DNA sequence into an RNA sequence, followed by translation of the RNA into a protein, which may or may not undergo post-translational processing. Thus, the relationship between cell cycle stage and/or developmental stage and gene expression can be observed by detecting, quantitatively or qualitatively, changes in the level of an RNA or a protein. As used herein, the term "biological activity" includes, but is not limited to, the activity of a protein gene product, including enzyme activity.

[0158] The present invention provides oligonucleotides that are useful in these expression profiling methods. Each oligonucleotide is capable of hybridizing under a given set of conditions to a cell cycle gene or gene product. In one aspect of the invention, a plurality of oligonucleotides is provided, wherein each oligonucleotide hybridizes under a given set of conditions to a different cell cycle gene product. Examples of oligonucleotides of the present invention include SEQ ID NOs: 521-772. Each of the oligos of SEQ ID NOs 521-772 hybridizes under standard conditions to a different gene product of one of SEQ ID NOs: 1-260. The oligonucleotides of the invention are useful in determining the expression of one or more cell cycle genes in any of the above-described methods.

1. Cell, Tissue, Nucleic Acid, and Protein Samples

[0159] Samples for use in methods of the present invention may be derived from plant tissue. Suitable plant tissues include, but are not limited to, somatic embryos, pollen, leaves, stems, calli, stolons, microtubers, shoots, xylem, male strolbili, pollen cones, vascular tissue, apical meristem, vascular cambium, xylem, root, flower, and seed.

[0160] According to the present invention "plant tissue" is used as described previously herein. Plant tissue can be obtained from any of the plants types or species described supra.

[0161] In accordance with one aspect of the invention, samples are obtained from plant tissue at different stages of the cell cycle, from plant tissue at different

developmental stages, from plant tissue at various times of the year (*e.g.* spring versus summer), from plant tissues subject to different environmental conditions (*e.g.* variations in light and temperature) and/or from different types of plant tissue and cells. In accordance with one embodiment, plant tissue is obtained during various stages of maturity and during different seasons of the year. For example, plant tissue can be collected from stem dividing cells, differentiating xylem, early developing wood cells, differentiated spring wood cells, and differentiated summer wood cells. As another example, gene expression in a sample obtained from a plant with developing wood can be compared to gene expression in a sample obtained from a plant which does not have developing wood.

[0162] Differentiating xylem includes samples obtained from compression wood, side-wood, and normal vertical xylem. Methods of obtaining samples for expression profiling from pine and eucalyptus are known. *See, e.g.,* Allona *et al., Proc. Nat'l Acad. Sci.* 95:9693-8 (1998) and Whetton *et al., Plant Mol. Biol.* 47:275-91, and Kirst *et al.,* INT'L UNION OF FORESTRY RESEARCH ORGANIZATIONS BIENNIAL CONFERENCE, S6.8 (June 2003, Umea, Sweden).

[0163] In one embodiment of the invention, gene expression in one type of tissue is compared to gene expression in a different type of tissue or to gene expression in the same type of tissue in a difference stage of development. Gene expression can also be compared in one type of tissue which is sampled at various times during the year (different seasons). For example, gene expression in juvenile secondary xylem can be compared to gene expression in mature secondary xylem. Similarly, gene expression in cambium can be compared to gene expression in xylem. Furthermore, gene expression in apical meristems can be compared to gene expression in cambium.

[0164] In an alternative embodiment, differences in gene expression are determined as cells from different tissues advance during the cell cycle. In this method, the cells from the different tissues are synchronized and their gene expression is profiled. Methods of synchronizing the stage of cell cycle in a sample are known. These methods include, *e.g.*, cold acclimation, photoperiod, and aphidicoline. *See,*

e.g., Nagata *et al.*, *Int. Rev. Cytol.* 132:1-30 (1992), Breyne and Zabeau, *Curr. Opin. Plant Biol.* 4:136-42, 140 (2001). A sample is obtained during a specific stage of the cell cycle and gene expression in that sample is compared to a sample obtained during a different stage of the cell cycle. For example, tissue can be examined in any of the phases of the cell cycle, such as mitosis, G1, G0, S, and G2. In particular, one can examine the changes in gene expression at the G1, G2, and metaphase checkpoints.

[0165] In another embodiment of the invention, a sample is obtained from a plant having a specific phenotype and gene expression in that sample is compared to a sample obtained from a plant of the same species that does not have that phenotype. For example, a sample can be obtained from a plant exhibiting a fast rate of growth and gene expression can be compared with that of a sample obtained from a plant exhibiting a normal or slow rate of growth. Differentially expressed genes identified from such a comparison can be correlated with growth rate and, therefore, useful for manipulating growth rate.

[0166] In a further embodiment, a sample is obtained from clonally propagated plants. In one embodiment the clonally propagated plants are of the species *Pinus* or *Eucalyptus*. Individual ramets from the same genotype can be sacrificed at different times of year. Thus, for any genotype there can be at least two genetically identical trees sacrificed, early in the season and late in the season. Each of these trees can be divided into juvenile (top) to mature (bottom) samples. Further, tissue samples can be divided into, for example, phloem to xylem, in at least 5 layers of peeling. Each of these samples can be evaluated for phenotype and gene expression. *See* Entry 196.

[0167] Where cellular components may interfere with an analytical technique, such as a hybridization assay, enzyme assay, a ligand binding assay, or a biological activity assay, it may be desirable to isolate the gene products from such cellular components. Gene products, including nucleic acid and amino acid gene products, can be isolated from cell fragments or lysates by any method known in the art.

[0168] Nucleic acids used in accordance with the invention can be prepared by any available method or process, or by other processes as they become known in the art. Conventional techniques for isolating nucleic acids are detailed, for example, in Tijssen, LABORATORY TECHNIQUES IN BIOCHEMISTRY AND MOLECULAR BIOLOGY: HYBRIDIZATION WITH NUCLEIC ACID PROBES, chapter 3 (Elsevier Press, 1993), Berger and Kimmel, *Methods Enzymol.* 152:1 (1987), and GIBCO BRL & LIFE TECHNOLOGIES TRIZOL RNA ISOLATION PROTOCOL, Form No. 3786 (2000). Techniques for preparing nucleic acid samples, and sequencing polynucleotides from pine and eucalyptus are known. *See, e.g.,* Allona *et al., supra* and Whetton *et al., supra*, and U.S. Application No. 60/476,222.

[0169] A suitable nucleic acid sample can contain any type of nucleic acid derived from the transcript of a cell cycle gene, *i.e.*, RNA or a subsequence thereof or a nucleic acid for which an mRNA transcribed from a cell cycle gene served as a template. Suitable nucleic acids include cDNA reverse-transcribed from a transcript, RNA transcribed from that cDNA, DNA amplified from the cDNA, and RNA transcribed from the amplified DNA. Detection of such products or derived products is indicative of the presence and/or abundance of the transcript in the sample. Thus, suitable samples include, but are not limited to, transcripts of the gene or genes, cDNA reverse-transcribed from the transcript, cRNA transcribed from the cDNA, DNA amplified from the genes, and RNA transcribed from amplified DNA. As used herein, the category of "transcripts" includes but is not limited to pre-mRNA nascent transcripts, transcript processing intermediates, and mature mRNAs and degradation products thereof.

[0170] It is not necessary to monitor all types of transcripts to practice the invention. For example, the expression profiling methods of the invention can be conducted by detecting only one type of transcript, such as mature mRNA levels only.

[0171] In one aspect of the invention, a chromosomal DNA or cDNA library (comprising, for example, fluorescently labeled cDNA synthesized from total cell mRNA) is prepared for use in hybridization methods according to recognized methods in the art. See Sambrook *et al., supra.*

[0172] In another aspect of the invention, mRNA is amplified using, *e.g.*, the MessageAmp kit (Ambion). In a further aspect, the mRNA is labeled with a detectable label. For example, mRNA can be labeled with a fluorescent chromophore, such as CyDye (Amersham Biosciences).

[0173] In some applications, it may be desirable to inhibit or destroy RNase that often is present in homogenates or lysates, before use in hybridization techniques. Methods of inhibiting or destroying nucleases are well known. In one embodiment of the invention, cells or tissues are homogenized in the presence of chaotropic agents to inhibit nuclease. In another embodiment, RNase is inhibited or destroyed by heat treatment, followed by proteinase treatment.

[0174] Protein samples can be obtained by any means known in the art. Protein samples useful in the methods of the invention include crude cell lysates and crude tissue homogenates. Alternatively, protein samples can be purified. Various methods of protein purification well known in the art can be found in Marshak *et al.*, STRATEGIES FOR PROTEIN PURIFICATION AND CHARACTERIZATION: A LABORATORY COURSE MANUAL (Cold Spring Harbor Laboratory Press 1996).

2. Detecting Level of Gene Expression

[0175] For methods of the invention that comprise detecting a level of gene expression, any method for observing gene expression can be used, without limitation. Such methods include traditional nucleic acid hybridization techniques, polymerase chain reaction (PCR) based methods, and protein determination. The invention includes detection methods that use solid support-based assay formats as well as those that use solution-based assay formats.

[0176] Absolute measurements of the expression levels need not be made, although they can be made. The invention includes methods comprising comparisons of differences in expression levels between samples. Comparison of expression levels can be done visually or manually, or can be automated and done by a machine, using for example optical detection means. Subrahmanyam *et al.*, *Blood.* 97: 2457 (2001); Prashar *et al.*, *Methods Enzymol.* 303: 258 (1999). Hardware and software for

analyzing differential expression of genes are available, and can be used in practicing the present invention. *See, e.g.,* GenStat Software and *GeneExpress*® GX Explorer™ Training Manual, *supra*; Baxevanis & Francis-Ouellette, *supra.*

[0177] In accordance with one embodiment of the invention, nucleic acid hybridization techniques are used to observe gene expression. Exemplary hybridization techniques include Northern blotting, Southern blotting, solution hybridization, and S1 nuclease protection assays.

[0178] Nucleic acid hybridization typically involves contacting an oligonucleotide probe and a sample comprising nucleic acids under conditions where the probe can form stable hybrid duplexes with its complementary nucleic acid through complementary base pairing. For example, see PCT application WO 99/32660; Berger & Kimmel, *Methods Enzymol.* 152: 1 (1987). The nucleic acids that do not form hybrid duplexes are then washed away leaving the hybridized nucleic acids to be detected, typically through detection of an attached detectable label. The detectable label can be present on the probe, or on the nucleic acid sample. In one embodiment, the nucleic acids of the sample are detectably labeled polynucleotides representing the mRNA transcripts present in a plant tissue (*e.g.*, a cDNA library). Detectable labels are commonly radioactive or fluorescent labels, but any label capable of detection can be used. Labels can be incorporated by several approached described, for instance, in WO 99/32660, *supra*. In one aspect RNA can be amplified using the MessageAmp kit (Ambion) with the addition of aminoallyl-UTP as well as free UTP. The aminoallyl groups incorporated into the amplified RNA can be reacted with a fluorescent chromophore, such as CyDye (Amersham Biosciences)

[0179] Duplexes of nucleic acids are destabilized by increasing the temperature or decreasing the salt concentration of the buffer containing the nucleic acids. Under low stringency conditions (*e.g.*, low temperature and/or high salt) hybrid duplexes (*e.g.*, DNA:DNA, RNA:RNA or RNA:DNA) will form even where the annealed sequences are not perfectly complementary. Thus, specificity of hybridization is reduced at lower stringency. Conversely, at higher stringency (*e.g.*,

higher temperature and/or lower salt and/or in the presence of destabilizing reagents) hybridization tolerates fewer mismatches.

[0180] Typically, stringent conditions for short probes (*e.g.*, 10 to 50 nucleotide bases) will be those in which the salt concentration is at least about 0.01 to 1.0 M at pH 7.0 to 8.3 and the temperature is at least about 30°C. Stringent conditions can also be achieved with the addition of destabilizing agents such as formamide.

[0181] Under some circumstances, it can be desirable to perform hybridization at conditions of low stringency, *e.g.*, 6 × SSPE-T (0.9 M NaCl, 60 mM NaH_2PO_4, pH 7.6, 6 mM EDTA, 0.005% Triton) at 37°C, to ensure hybridization. Subsequent washes can then be performed at higher stringency (*e.g.*, 1× SSPE-T at 37°C) to eliminate mismatched hybrid duplexes. Successive washes can be performed at increasingly higher stringency (*e.g.*, down to as low as 0.25× SSPE-T at 37°C to 50°C) until a desired level of hybridization specificity is obtained.

[0182] In general, standard conditions for hybridization is a compromise between stringency (hybridization specificity) and signal intensity. Thus, in one embodiment of the invention, the hybridized nucleic acids are washed at successively higher stringency conditions and read between each wash. Analysis of the data sets produced in this manner will reveal a wash stringency above which the hybridization pattern is not appreciably altered and which provides adequate signal for the particular oligonucleotide probes of interest. For example, the final wash may be selected as that of the highest stringency that produces consistent results and that provides a signal intensity greater than approximately 10% of the background intensity.

a. Oligonucleotide Probes

[0183] Oligonucleotide probes useful in nucleic acid hybridization techniques employed in the present invention are capable of binding to a nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing via hydrogen bond formation. A probe can include natural bases (*i.e.*, A, G, U, C or T) or modified bases (7-deazaguanosine,

2004311384 31 May 2011

inosine, etc.). In addition, the nucleotide bases in the probes can be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, probes can be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages.

[0184] Oligonucleotide probes can be prepared by any means known in the art. Probes useful in the present invention are capable of hybridizing to a nucleotide product of cell cycle genes, such as one of SEQ ID NOs: 1-260. Probes useful in the invention can be generated using the nucleotide sequences disclosed in SEQ ID NOs: 1-260. The invention includes oligonucleotide probes having at least a 2, 10,15, 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, or 100 nucleotide fragment of a corresponding contiguous sequence of any one of SEQ ID NOs: 1-260. The invention includes oligonucleotides of less than 2, 1, 0.5, 0.1, or 0.05 kb in length. In one embodiment, the oligonucleotide is 60 nucleotides in length.

[0185] Oligonucleotide probes can be designed by any means known in the art. *See, e.g.,* Li and Stormo, Bioinformatics 17: 1067-76 (2001). Oligonucleotide probe design can be effected using software. Exemplary software includes ArrayDesigner, GeneScan, and ProbeSelect. Probes complementary to a defined nucleic acid sequence can be synthesized chemically, generated from longer nucleotides using restriction enzymes, or can be obtained using techniques such as polymerase chain reaction (PCR). PCR methods are well known and are described, for example, in Innis *et al. eds.*, PCR PROTOCOLS: A GUIDE TO METHODS AND APPLICATIONS, Academic Press Inc. San Diego, Calif. (1990). The probes can be labeled, for example, with a radioactive, biotinylated, or fluorescent tag. Optimally, the nucleic acids in the sample are labeled and the probes are not labeled. Oligonucleotide probes generated by the above methods can be used in solution or solid support-based methods.

[0186] The invention includes oligonucleotide probes that hybridize to a product of the coding region or a 3' untranslated region (3' UTR) of a cell cycle gene. In one embodiment, the oligonucleotide probe hybridizes to the 3'UTR of any one of SEQ ID NOs: 1-260. The 3' UTR is generally a unique region of the gene,

even among members of the same family. Therefore, the probes capable of hybridizing to a product of the 3' UTR can be useful for differentiating the expression of individual genes within a family where the coding region of the genes likely are highly homologous. This allows for the design of oligonucleotide probes to be used as members of a plurality of oligonucleotides, each capable of uniquely binding to a single gene. In another embodiment, the oligonucleotide probe comprises any one of SEQ ID NOs: 521-772. In another embodiment, the oligonucleotide probe consists of any one of SEQ ID NOs: 521-772.

a. Oligonucleotide Array Methods

[0187] One embodiment of the invention employs two or more oligonucleotide probes in combination to detect a level of expression of one or more cell cycle genes, such as the genes of SEQ ID NOs: 1-260. In one aspect of this embodiment, the level of expression of two or more different genes is detected. The two or more genes may be from the same or different cell cycle gene families discussed above. Each of the two or more oligonucleotides may hybridize to a different one of the genes.

[0188] One embodiment of the invention employs two or more oligonucleotide probes, each of which specifically hybridize to a polynucleotide derived from the transcript of a gene provided by SEQ ID NOs: 1-260. Another embodiment employs two or more oligonucleotide probes, at least one of which comprises a nucleic acid sequence of SEQ ID NOs: 521-772. Another embodiment employs two or more oligonucleotide probes, at least one of which consists of SEQ ID NOs: 521-772.

[0189] The oligonucleotide probes may comprise from about 5 to about 60, or from about 5 to about 500, nucleotide bases, such as from about 60 to about 100 nucleotide bases, including from about 15 to about 60 nucleotide bases.

[0190] One embodiment of the invention uses solid support-based oligonucleotide hybridization methods to detect gene expression. Solid support-based methods suitable for practicing the present invention are widely known and are

described, for example, in PCT application WO 95/11755; Huber *et al.*, *Anal. Biochem.* 299: 24 (2001); Meiyanto *et al.*, *Biotechniques.* 31: 406 (2001); Relogio *et al.*, *Nucleic Acids Res.* 30:e51 (2002). Any solid surface to which oligonucleotides can be bound, covalently or non-covalently, can be used. Such solid supports include filters, polyvinyl chloride dishes, silicon or glass based chips, etc.

[0191] One embodiment uses oligonucleotide arrays, *i.e.* microarrays, which can be used to simultaneously observe the expression of a number of genes or gene products. Oligonucleotide arrays comprise two or more oligonucleotide probes provided on a solid support, wherein each probe occupies a unique location on the support. The location of each probe may be predetermined, such that detection of a detectable signal at a given location is indicative of hybridization to an oligonucleotide probe of a known identity. Each predetermined location can contain more than one molecule of a probe, but each molecule within the predetermined location has an identical sequence. Such predetermined locations are termed features. There can be, for example, from 2, 10, 100, 1,000, 2,000 or 5,000 or more of such features on a single solid support. In one embodiment, each oligonucleotide is located at a unique position on an array at least 2, at least 3, at least 4, at least 5, at least 6, or at least 10 times.

[0192] Oligonucleotide probe arrays for detecting gene expression can be made and used according to conventional techniques described, for example, in Lockhart *et al.*, *Nat'l Biotech.* 14: 1675 (1996), McGall *et al.*, *Proc. Nat'l Acad. Sci. USA* 93: 13555 (1996), and Hughes *et al.*, *Nature Biotechnol.* 19:342 (2001). A variety of oligonucleotide array designs is suitable for the practice of this invention.

[0193] In one embodiment the one or more oligonucleotides include a plurality of oligonucleotides that each hybridize to a different gene expressed in a particular tissue type. For example, the tissue can be developing wood.

[0194] In one embodiment, a nucleic acid sample obtained from a plant can be amplified and, optionally labeled with a detectable label. Any method of nucleic acid amplification and any detectable label suitable for such purpose can be

used. For example, amplification reactions can be performed using, *e.g.* Ambion's MessageAmp, which creates "antisense" RNA or "aRNA" (complementary in nucleic acid sequence to the RNA extracted from the sample tissue). The RNA can optionally be labeled using CyDye fluorescent labels. During the amplification step, aaUTP is incorporated into the resulting aRNA. The CyDye fluorescent labels are coupled to the aaUTPs in a non-enzymatic reaction. Subsequent to the amplification and labeling steps, labeled amplified antisense RNAs are precipitated and washed with appropriate buffer, and then assayed for purity. For example, purity can be assay using a NanoDrop spectrophotometer. The nucleic acid sample is then contacted with an oligonucleotide array having, attached to a solid substrate (a "microarray slide"), oligonucleotide sample probes capable of hybridizing to nucleic acids of interest which may be present in the sample. The step of contacting is performed under conditions where hybridization can occur between the nucleic acids of interest and the oligonucleotide probes present on the array. The array is then washed to remove non-specifically bound nucleic acids and the signals from the labeled molecules that remain hybridized to oligonucleotide probes on the solid substrate are detected. The step of detection can be accomplished using any method appropriate to the type of label used. For example, the step of detecting can accomplished using a laser scanner and detector. For example, on can use and Axon scanner which optionally uses GenePix Pro software to analyze the position of the signal on the microarray slide.

[0195] Data from one or more microarray slides can analyzed by any appropriate method known in the art.

[0196] Oligonucleotide probes used in the methods of the present invention, including microarray techniques, can be generated using PCR. PCR primers used in generating the probes are chosen, for example, based on the sequences of SEQ ID NOs:1-260, to result in amplification of unique fragments of the cell cycle genes (*i.e.*, fragments that hybridize to only one polynucleotide of any one of SEQ ID NOs: 1-260 under standard hybridization conditions). Computer programs are useful in the design of primers with the required specificity and optimal hybridization properties. For example, Li and Stormo, *supra* at 1075, discuss a

method of probe selection using ProbeSelect which selects an optimum oligonucleotide probe based on the entire gene sequence as well as other gene sequences to be probed at the same time.

[0197] In one embodiment, oligonucleotide control probes also are used. Exemplary control probes can fall into at least one of three categories referred to herein as (1) normalization controls, (2) expression level controls and (3) negative controls. In microarray methods, one or more of these control probes may be provided on the array with the inventive cell cycle gene-related oligonucleotides.

[0198] Normalization controls correct for dye biases, tissue biases, dust, slide irregularities, malformed slide spots, *etc.* Normalization controls are oligonucleotide or other nucleic acid probes that are complementary to labeled reference oligonucleotides or other nucleic acid sequences that are added to the nucleic acid sample to be screened. The signals obtained from the normalization controls, after hybridization, provide a control for variations in hybridization conditions, label intensity, reading efficiency and other factors that can cause the signal of a perfect hybridization to vary between arrays. In one embodiment, signals (*e.g.*, fluorescence intensity or radioactivity) read from all other probes used in the method are divided by the signal from the control probes, thereby normalizing the measurements.

[0199] Virtually any probe can serve as a normalization control. Hybridization efficiency varies, however, with base composition and probe length. Preferred normalization probes are selected to reflect the average length of the other probes being used, but they also can be selected to cover a range of lengths. Further, the normalization control(s) can be selected to reflect the average base composition of the other probes being used. In one embodiment, only one or a few normalization probes are used, and they are selected such that they hybridize well (*i.e.*, without forming secondary structures) and do not match any test probes. In one embodiment, the normalization controls are mammalian genes.

[0200] Expression level controls probes hybridize specifically with constitutively expressed genes present in the biological sample. Virtually any constitutively expressed gene provides a suitable target for expression level control probes. Typically, expression level control probes have sequences complementary to subsequences of constitutively expressed "housekeeping genes" including, but not limited to certain photosynthesis genes.

[0201] "Negative control" probes are not complementary to any of the test oligonucleotides (*i.e.*, the inventive cell cycle gene-related oligonucleotides), normalization controls, or expression controls. In one embodiment, the negative control is a mammalian gene which is not complementary to any other sequence in the sample.

[0202] The terms "background" and "background signal intensity" refer to hybridization signals resulting from non-specific binding or other interactions between the labeled target nucleic acids (*i.e.*, mRNA present in the biological sample) and components of the oligonucleotide array. Background signals also can be produced by intrinsic fluorescence of the array components themselves.

[0203] A single background signal can be calculated for the entire array, or a different background signal can be calculated for each target nucleic acid. In a one embodiment, background is calculated as the average hybridization signal intensity for the lowest 5 to 10 percent of the oligonucleotide probes being used, or, where a different background signal is calculated for each target gene, for the lowest 5 to 10 percent of the probes for each gene. Where the oligonucleotide probes corresponding to a particular cell cycle gene hybridize well and, hence, appear to bind specifically to a target sequence, they should not be used in a background signal calculation. Alternatively, background can be calculated as the average hybridization signal intensity produced by hybridization to probes that are not complementary to any sequence found in the sample (*e.g.*, probes directed to nucleic acids of the opposite sense or to genes not found in the sample). In microarray methods, background can be calculated as the average signal intensity produced by regions of the array that lack any oligonucleotides probes at all.

c. PCR-Based Methods

[0204] In another embodiment, PCR-based methods are used to detect gene expression. These methods include reverse-transcriptase-mediated polymerase chain reaction (RT-PCR) including real-time and endpoint quantitative reverse-transcriptase-mediated polymerase chain reaction (Q-RTPCR). These methods are well known in the art. For example, methods of quantitative PCR can be carried out using kits and methods that are commercially available from, for example, Applied BioSystems and Stratagene®. See also Kochanowski, QUANTITATIVE PCR PROTOCOLS (Humana Press, 1999); Innis *et al., supra.*; Vandesompele *et al.*, *Genome Biol.* 3: RESEARCH0034 (2002); Stein, *Cell Mol. Life Sci.* 59: 1235 (2002).

[0205] Gene expression can also be observed in solution using Q-RTPCR. Q-RTPCR relies on detection of a fluorescent signal produced proportionally during amplification of a PCR product. See Innis *et al.*, *supra.* Like the traditional PCR method, this technique employs PCR oligonucleotide primers, typically 15-30 bases long, that hybridize to opposite strands and regions flanking the DNA region of interest. Additionally, a probe (*e.g.*, TaqMan®, Applied Biosystems) is designed to hybridize to the target sequence between the forward and reverse primers traditionally used in the PCR technique. The probe is labeled at the 5' end with a reporter fluorophore, such as 6-carboxyfluorescein (6-FAM) and a quencher fluorophore like 6-carboxy-tetramethyl-rhodamine (TAMRA). As long as the probe is intact, fluorescent energy transfer occurs which results in the absorbance of the fluorescence emission of the reporter fluorophore by the quenching fluorophore. As Taq polymerase extends the primer, however, the intrinsic 5' to 3' nuclease activity of Taq degrades the probe, releasing the reporter fluorophore. The increase in the fluorescence signal detected during the amplification cycle is proportional to the amount of product generated in each cycle.

[0206] The forward and reverse amplification primers and internal hybridization probe is designed to hybridize specifically and uniquely with one nucleotide derived from the transcript of a target gene. In one embodiment, the

2004311384 31 May 2011

selection criteria for primer and probe sequences incorporates constraints regarding nucleotide content and size to accommodate TaqMan® requirements.

[0207] SYBR Green® can be used as a probe-less Q-RTPCR alternative to the Taqman®-type assay, discussed above. ABI PRISM® 7900 SEQUENCE DETECTION SYSTEM USER GUIDE APPLIED BIOSYSTEMS, chap. 1-8, App. A-F. (2002).

[0208] A device measures changes in fluorescence emission intensity during PCR amplification. The measurement is done in "real time," that is, as the amplification product accumulates in the reaction. Other methods can be used to measure changes in fluorescence resulting from probe digestion. For example, fluorescence polarization can distinguish between large and small molecules based on molecular tumbling (see U.S. patent No. 5,593,867).

b. Protein Detection Methods

[0209] Proteins can be observed by any means known in the art, including immunological methods, enzyme assays and protein array/proteomics techniques.

[0210] Measurement of the translational state can be performed according to several protein methods. For example, whole genome monitoring of protein -- the "proteome" -- can be carried out by constructing a microarray in which binding sites comprise immobilized, preferably monoclonal, antibodies specific to a plurality of proteins having an amino acid sequence of any of SEQ ID NOs: 261-520 or proteins encoded by the genes of SEQ ID NOs:1-260 or conservative variants thereof. See Wildt *et al.*, *Nature Biotechnol.* 18: 989 (2000). Methods for making polyclonal and monoclonal antibodies are well known, as described, for instance, in Harlow & Lane, ANTIBODIES: A LABORATORY MANUAL (Cold Spring Harbor Laboratory Press, 1988).

[0211] Alternatively, proteins can be separated by two-dimensional gel electrophoresis systems. Two-dimensional gel electrophoresis is well-known in

the art and typically involves isoelectric focusing along a first dimension followed by SDS-PAGE electrophoresis along a second dimension. See, *e.g.*, Hames *et al*, , GEL ELECTROPHORESIS OF PROTEINS: A PRACTICAL APPROACH (IRL Press, 1990). The resulting electropherograms can be analyzed by numerous techniques, including mass spectrometric techniques, western blotting and immunoblot analysis using polyclonal and monoclonal antibodies, and internal and N-terminal micro-sequencing.

3. Correlating Gene Expression to Phenotype and Tissue Development

[0212] As discussed above, the invention provides methods and tools to correlate gene expression to plant phenotype. Gene expression may be be examined in a plant having a phenotype of interest and compared to a plant that does not have the phenotype or has a different phenotype. Such a phenotype includes, but is not limited to, increased drought tolerance, herbicide resistance, reduced or increased height, reduced or increased branching, enhanced cold and frost tolerance, improved vigor, enhanced color, enhanced health and nutritional characteristics, improved storage, enhanced yield, enhanced salt tolerance, enhanced resistance of the wood to decay, enhanced resistance to fungal diseases, altered attractiveness to insect pests, enhanced heavy metal tolerance, increased disease tolerance, increased insect tolerance, increased water-stress tolerance, enhanced sweetness, improved texture, decreased phosphate content, increased germination, increased micronutrient uptake, improved starch composition, improved flower longevity, production of novel resins, and production of novel proteins or peptides.

[0213] In another embodiment, the phenotype includes one or more of the following traits: propensity to form reaction wood, a reduced period of juvenility, an increased period of juvenility, self-abscising branches, accelerated reproductive development or delayed reproductive development.

[0214] In a further embodiment, the phenotype that is differs in the plants compares includes one or more of the following: lignin quality, lignin structure, wood composition, wood appearance, wood density, wood strength, wood stiffness, cellulose polymerization, fiber dimensions, lumen size, other plant components, plant

cell division, plant cell development, number of cells per unit area, cell size, cell shape, cell wall composition, rate of wood formation, aesthetic appearance of wood, formation of stem defects, average microfibril angle, width of the S2 cell wall layer, rate of growth, rate of root formation ratio of root to branch vegetative development, leaf area index, and leaf shape.

[0215] Phenotype can be assessed by any suitable means as discussed above.

[0216] In a further embodiment, gene expression can be correlated to a given point in the cell cycle, a given point in plant development, and in a given tissue sample. Plant tissue can be examined at different stages of the cell cycle, from plant tissue at different developmental stages, from plant tissue at various times of the year (*e.g.* spring versus summer), from plant tissues subject to different environmental conditions (*e.g.* variations in light and temperature) and/or from different types of plant tissue and cells. In accordance with one embodiment, plant tissue is obtained during various stages of maturity and during different seasons of the year. For example, plant tissue can be collected from stem dividing cells, differentiating xylem, early developing wood cells, differentiated spring wood cells, differentiated summer wood cells.

[0217] It will be apparent to those skilled in the art that various modifications and variations can be made in the methods and compositions of the present invention without departing from the spirit or scope of the invention. Thus, it is intended that the present invention cover the modifications and variations of this invention provided they come within the scope of the appended claims and their equivalents.

[0218] The following examples are given to illustrate the present invention. It should be understood, however, that the invention is not to be limited to the specific conditions or details described in these examples. Throughout the specification, any and all references to a publicly available document, including a U.S. patent, are specifically incorporated by reference.

[0219] EXAMPLES

Example 1

[0220] Example 1 illustrates a procedure for RNA extraction and purification, which is particularly useful for RNA obtained from conifer needle, xylem, cambium, and phloem.

[0221] Tissue is obtained from conifer needle, xylem, cambium or phloem. The tissue is frozen in liquid nitrogen and ground. The total RNA is extracted using Concert Plant RNA reagent (Invitrogen). The resulting RNA sample is extracted into phenol:chloroform and treated with DNase. The RNA is then incubated at 65 °C for 2 minutes followed by centrifugation at 4 °C for 30 minutes. Following centrifugation, the RNA is extracted into phenol at least 10 times to remove contaminants.

[0222] The RNA is further cleaned using RNeasy columns (Qiagen). The purified RNA is quantified using RiboGreen reagent (Molecular Probes) and purity assessed by gel electrophoresis.

[0223] RNA is then amplified using MessageAmp (Ambion). Aminoallyl-UTP and free UTP are added to the *in vitro* transcription of the purified RNA at a ratio of 4:1 aminoallyl-UTP-to-UTP. The aminoallyl-UTP is incorporated into the new RNA strand as it is transcribed. The amino-allyl group is then reacted with Cy dyes to attach the colorimetric label to the resulting amplified RNA using the Amersham procedure modified for use with RNA. Unincorporated dye is removed by ethanol precipitation. The labeled RNA is quantified spectrophotometrically (NanoDrop). The labeled RNA is fragmented by heating to 95 °C as described in Hughes *et al.*, *Nature Biotechnol.* 19:342 (2001).

Example 2

[0224] Example 2 illustrates how cell cycle genes important for wood development in *Pinus radiata* can be determined and how oligonucleotides which uniquely bind to those genes can be designed and synthesized for use on a microarray.

[0225] Pine trees of the species *Pinus radiata* are grown under natural light conditions. Tissue samples are prepared as described in, *e.g.*, Sterky *et al.*, *Proc. Nat'l Acad. Sci.* 95:13330 (1998). Specifically, tissue samples are collected from woody trees having a height of 5 meters. Tissue samples of the woody trees are prepared by taking tangential sections through the cambial region of the stem. The stems are sectioned horizontally into sections ranging from juvenile (top) to mature (bottom). The stem sections separated by stage of development are further separated into 5 layers by peeling into sections of phloem, differentiating phloem, cambium, differentiating xylem, developing xylem, and mature xylem. Tissue samples, including leaves, buds, shoots, and roots are also prepared from seedlings of the species *Pinus radiata.*

[0226] RNA is isolated and ESTs generated as described in Example 1 or Sterky *et al.*, *supra.* The nucleic acid sequences of ESTs derived from samples containing developing wood are compared with nucleic acid sequences of genes known to be involved in the plant cell cycle. ESTs from samples that do not contain developing wood are also compared with sequences of genes known to be involved in the plant cell cycle. An *in silico* hybridization analysis is performed using BLAST (NCBI). Sequences from among the known cell cycle genes that show hybridization *in silico* to ESTs made from samples containing developing wood, but that do not hybridize to ESTs from samples not containing developing wood are selected for further examination.

[0227] cDNA clones containing sequences that hybridize to the genes showing wood-preferred expression are selected from cDNA libraries using techniques well known in the art of molecular biology. Using the sequence information, oligonucleotides are designed such that each oligonucleotide is specific for only one cDNA sequence in the library. The oligonucleotide sequences are provided in Table 14. 60-mer oligonucleotide probes are designed using the method of Li and Stormo, *supra* or using software such as ArrayDesigner, GeneScan, and ProbeSelect.

[0228] The oligonucleotides are then synthesized *in situ* described in Hughes *et al., Nature Biotechnol.* 19:324 (2002) or as described in Kane *et al., Nucleic Acids Res.* 28:4552 (2000) and affixed to an activated glass slide (Sigma-Genosis, The Woodlands, TX) using a 5' amino linker. The position of each oligonucleotide on the slide is known.

Example 3

[0229] Example 3 illustrates how cell cycle genes important for wood development in *Eucalyptus grandis* can be determined and how oligonucleotides which uniquely bind to those genes can be designed and synthesized for use on a microarray.

[0230] Eucalyptus trees of the species *Eucalyptus grandis* are grown under natural light conditions. Tissue samples are prepared as described in, *e.g.,* Sterky *et al., Proc. Nat'l Acad. Sci.* 95:13330 (1998). Specifically, tissue samples are collected from woody trees having a height of 5 meters. Tissue samples of the woody trees are prepared by taking tangential sections through the cambial region of the stem. The stems are sectioned horizontally into sections ranging from juvenile (top) to mature (bottom). The stem sections separated by stage of development are further separated into 5 layers by peeling into sections of phloem, differentiating phloem, cambium, differentiating xylem, developing xylem, and mature xylem. Tissue samples, including leaves, buds, shoots, and roots are also prepared from seedlings of the species *Pinus radiata.*

[0231] RNA is isolated and ESTs generated as described in Example 1 or Sterky *et al., supra.* The nucleic acid sequences of ESTs derived from samples containing developing wood are compared with nucleic acid sequences of genes known to be involved in the plant cell cycle. ESTs from samples that do not contain developing wood are also compared with sequences of genes known to be involved in the plant cell cycle. An *in silico* hybridization analysis is performed as described in, for example, Audic and Claverie, *Genome Res.* 7:986 (1997). Sequences from among the known cell cycle genes that show hybridization *in silico* to ESTs made from

samples containing developing wood, but do not hybridize to ESTs from samples not containing developing wood are selected for further examination.

[0232] cDNA clones containing sequences that hybridize to the genes showing wood-preferred expression are selected from cDNA libraries using techniques well known in the art of molecular biology. Using the sequence information, oligonucleotides are designed such that each oligonucleotide is specific for only one cDNA sequence in the library. The oligonucleotide sequences are provided in Table 14. 60-mer oligonucleotide probes are designed using the method of Li and Stormo, *supra* or using software such as ArrayDesigner, GeneScan, and ProbeSelect.

[0233] The oligonucleotides are then synthesized *in situ* described in Hughes *et al., Nature Biotechnol.* 19:324 (2002) or as described in Kane *et al., Nucleic Acids Res.* 28:4552 (2000) and affixed to an activated glass slide (Sigma-Genosus, The Woodlands, TX) using a 5' amino linker. The position of each oligonucleotide on the slide is known.

Example 4

[0234] Example 4 illustrates how to detect expression of *Pinus radiata* cell cycle genes which are important in wood formation using an oligonucleotide microarray prepared as in Example 2. This is an example of a balanced incomplete block designed experiment carried out using aRNA samples prepared from mature-phase phloem (P), cambium (C), expanding xylem found in a layer below the cambium (X1) and differentiating, lignifying xylem cells found deeper in the same growth ring (X2). In this example, cell cycle gene expression is compared among the four samples, namely P, C, X1, and X2.

[0235] In the summer, plants of the species *Pinus radiata* are felled and the bark of the main stem is immediately pulled gently away to reveal the phloem and xylem. The phloem and xylem are then peeled with a scalpel into separate containers of liquid nitrogen. Needles (leaves) and buds from the trees are also harvested with a scalpel into separate containers of liquid nitrogen. RNA is subsequently isolated from

2004311384 31 May 2011

the frozen tissue samples as described in Example 1. Equal microgram quantities of total RNA are purified from each sample using RNeasy Mini columns (Qiagen, Valencia, CA) according to the manufacturer's instructions.

[0236] Amplification reactions are carried out for each of the P, C, X1, and X2 tissue samples. Amplification reactions are performed using Ambion's MessageAmp kit, a T7-based amplification procedure, following the manufacturer's instructions, except that labeled aaUTP is added to the reagent mix during in the amplification step. aaUTP is incorporated into the resulting antisense RNA formed during this step. CyDye fluorescent labels are coupled to the aaUTPs in a non-enzymatic reaction as described in Example 1. Labeled amplified antisense RNAs are precipitated and washed, and then assayed for purity using a NanoDrop spectrophotometer. These labeled antisense RNAs, corresponding to the RNA isolated from the P, C, X1, and X2 tissue samples, constitute the sample nucleic acids, which are referred to as the P, C, X1, and X2 samples.

[0237] Normalization control samples of known nucleic acids are added to each sample in a dilution series of 500, 200, 100, 50, 25 and 10 pg/μl for quantitation of the signals. Positive controls corresponding to specific genes showing expression in all tissues of pine, such as housekeeping genes, are also added to the plant sample.

[0238] Each of four microarray slides is incubated with 125 μL of a P, C, X1 or X2 sample under a coverslip at 42 °C for 16-18 hours. The arrays are washed in 1x SSC, 0.1% SDS for 10 minutes and then in 0.1 x SSC, 0.1% SDS for 10 minutes and the allowed to dry.

[0239] The array slides are scanned using an Axon laser scanner and analyzed using GenePix Pro software. Data from the microarray slides are subjected to microarray data analysis using GenStat SAS or Spotfire software. Outliers are removed and ratiometric data for each of the datasets are normalized using a global normalization which employs a cubic spline fit applied to correct for differential dye bias and spatial effects. A second transformation is performed to fit control signal

ratios to a mean $\log^2 = 0$ (*i.e.* 1:1 ratio). Normalized data are then subjected to a variance analysis.

[0240] Mean signal intensity for each signal at any given position on the microarray slide is determined for each of three of P, C, X1, and X2 sample microarray slides. This mean signal/probe position is compared to the signal at the same position on sample slide which was not used for calculating the mean. For example, a mean signal at a given position is determined for P, C, and X1 and the signal at that position in the X2 microarray slide is compared to the P, C, and X1 mean signal value.

[0241] Table 1 shows genes having greater than doubled signal with any one sample as compared to the mean signal of the other three samples.

[0242] Table 1.

Gene	PvCX12	PvX12	CvX12
WD40 repeat protein A	-1.24	-0.88	-1.07
CDC2	-1.09	-0.78	-0.92
CYCLIN	-1.08	-1	-0.26
WD-40 repeat protein B	-1.01	-0.87	-0.42
CDC2	-0.83	-0.49	-1.01

P = Phloem

C = Cambium

X1 = xylem layer-1

X2 = xylem layer-2

PvCX12 = Ratio of the signal for Phloem target versus mean signal for Cambium, Xylem1, and Xylem2 targets

[0243] The data shows that WD40 repeat protein A encodes a WD40 repeat protein is less highly expressed in cambium than in developing xylem, while WD40

repeat protein B encodes a WD40 repeat protein that is more highly expressed in phloem than in the other tissues.

[0244] Signal data are then verified with RT-PCR to confirm gene expression in the target tissue of the genes corresponding to the unique oligonucleotides in the probe.

Example 5

[0245] Example 5 demonstrates how one can correlate cell cycle gene expression with agronomically important wood phenotypes such as density, stiffness, strength, distance between branches, and spiral grain.

[0246] Mature clonally propagated pine trees are selected from among the progeny of known parent trees for superior growth characteristics and resistance to important fungal diseases. The bark is removed from a tangential section and the trees are examined for average wood density in the fifth annual ring at breast height, stiffness and strength of the wood, and spiral grain. The trees are also characterized by their height, mean distance between major branches, crown size, and forking.

[0247] To obtain seedling families that are segregating for major genes that affect density, stiffness, strength, distance between branches, spiral grain and other characteristics that may be linked to any of the genes affecting these characteristics, trees lacking common parents are chosen for specific crosses on the criterion that they exhibit the widest variation from each other with respect to the density, stiffness, strength, distance between branches, and spiral grain criteria. Thus, pollen from a plus tree exhibiting high density, low mean distance between major branches, and high spiral grain is used to pollinate cones from the unrelated plus tree among the selections exhibiting the lowest density, highest mean distance between major branches, and lowest spiral grain. It is useful to note that "plus trees" are crossed such that pollen from a plus tree exhibiting high density are used to pollinate developing cones from another plus tree exhibiting high density, for example, and pollen from a tree exhibiting low mean distance between major branches would be used to pollinate

developing cones from another plus tree exhibiting low mean distance between major branches.

[0248] Seeds are collected from these controlled pollinations and grown such that the parental identity is maintained for each seed and used for vegetative propagation such that each genotype is represented by multiple ramets. Vegetative propagation is accomplished using micropropagation, hedging, or fascicle cuttings. Some ramets of each genotype are stored while vegetative propagules of each genotype are grown to sufficient size for establishment of a field planting. The genotypes are arrayed in a replicated design and grown under field conditions where the daily temperature and rainfall are measured and recorded.

[0249] The trees are measured at various ages to determine the expression and segregation of density, stiffness, strength, distance between branches, spiral grain, and any other observable characteristics that may be linked to any of the genes affecting these characteristics. Samples are harvested for characterization of cellulose content, lignin content, cellulose microfibril angle, density, strength, stiffness, tracheid morphology, ring width, and the like. Samples are also examined for gene expression as described in Example 4. Ramets of each genotype are compared to ramets of the same genotype at different ages to establish age:age correlations for these characteristics.

Example 6

[0250] Example 6 demonstrates how the stage of plant development and responses to environmental conditions such as light and season can be correlated to cell cycle gene expression using microarrays prepared as in Example 4. In particular, the changes in gene expression associated with wood density are examined.

[0251] Trees of three different clonally propagated *Eucalyptus grandis* hybrid genotypes are grown on a site with a weather station that measures daily temperatures and rainfall. During the spring and subsequent summer, genetically identical ramets of the three different genotypes are first photographed with north-south orientation marks, using photography at sufficient resolution to show bark

characteristics of juvenile and mature portions of the plant, and then felled as in Example 4. The age of the trees is determined by planting records and confirmed by a count of the annual rings. In each of these trees, mature wood is defined as the outermost rings of the tree below breast height, and juvenile wood as the innermost rings of the tree above breast height. Each tree is accordingly sectored as follows:

NM - NORTHSIDE **MATURE**

SM - SOUTHSIDE **MATURE**

NT - NORTHSIDE **TRANSITION**

ST - SOUTHSIDE **TRANSITION**

NJ - NORTHSIDE **JUVENILE**

SJ - SOUTHSIDE **JUVENILE**

[0252] Tissue is harvested from the plant trunk as well as from juvenile and mature form leaves. Samples are prepared simultaneously for phenotype analysis, including plant morphology and biochemical characteristics, and gene expression analysis. The height and diameter of the tree at the point from which each sector was taken is recorded, and a soil sample from the base of the tree is taken for chemical assay. Samples prepared for gene expression analysis are weighed and placed into liquid nitrogen for subsequent preparation of RNA samples for use in the microarray experiment. The tissues are denoted as follows:

P - phloem

C - cambium

X1 - expanding xylem

X2 - differentiating and lignifying xylem

[0253] Thin slices in tangential and radial sections from each of the sectors of the trunk are fixed as described in Ruzin, Plant Microtechnique and Microscopy, Oxford University Press, Inc., New York, NY (1999) for anatomical examination and confirmation of wood developmental stage. Microfibril angle is examined at the different developmental stages of the wood, for example juvenile, transition and mature phases of *Eucalyptus grandis* wood. Other characteristics examined are the ratio of fibers to vessel elements and ray tissue in each sector. Additionally, the samples are examined for characteristics that change between juvenile and mature wood and between spring wood and summer wood, such as fiber morphology, lumen size, and width of the S2 (thickest) cell wall layer. Samples are further examined for measurements of density in the fifth ring and determination of modulus of elasticity using techniques well known to those skilled in the art of wood assays. *See, e.g., Wang, et al., Non-destructive Evaluations of Trees,* EXPERIMENTAL TECHNIQUES, pp. *28*-30 (2000).

[0254] For biochemical analysis, 50 grams from each of the harvest samples are freeze-dried and analyzed, using biochemical assays well known to those skilled in the art of plant biochemistry for quantities of simple sugars, amino acids, lipids, other extractives, lignin, and cellulose. *See, e.g.,* Pettersen & Schwandt, *J. Wood Chem. & Technol.* 11:495 (1991).

[0255] In the present example, the phenotypes chosen for comparison are high density wood, average density wood, and low density wood. Nucleic acid samples are prepared as described in Example 3, from trees harvested in the spring and summer. Gene expression profiling by hybridization and data analysis is performed as described in Examples 3 and 4.

[0256] Using similar techniques and clonally propagated individuals one can examine cell cycle gene expression as it is related to other complex wood characteristics such as strength, stiffness and spirality.

Example 7

[0257] Example 7 demonstrates the ability of the oligonucleotide probes of the invention to distinguish between highly homologous members of a family of cell cycle genes. Hybridization to a particular oligonucleotide on the array identifies a unique WD40 gene that is expressed more strongly in a genotype having a higher density wood than in observed in other genotypes examined. The WD40 gene is also expressed more strongly in mature wood than in juvenile wood and more strongly in summer wood than in spring wood. This gene is not found to be expressed at high levels either in leaves or buds.

[0258] The gene expression pattern is confirmed by RT-PCR. This gene, the putative "density-related" gene, is used for *in situ* hybridization of fixed radial sections. The density-related WD40 gene hybridizes most strongly to the vascular cambium in regions of the stem where the xylem is comprised primarily of fibers with few vessel elements and few xylem ray cells.

[0259] These results suggest that the WD40 gene product functions in radial cell division, which occurs in the cambium and results in diameter growth, rather than in axial cell division such as may be important in the apex or leaves. Such a gene would be difficult to identify by cDNA microarrays or other traditional hybridization means because the highly conserved regions present in the gene would result in confusing it with genes encoding enzymes having similar catalytic functions, but acting in axial or radial divisions. Furthermore, from the sequence similarity-based annotation suggesting a function of this gene product in cell division and the observation of this microarray hybridization pattern, confirmed by RT-PCR and *in silico* hybridization, this gene product functions specifically in developing secondary xylem to guide the cell division patterns of fibers, such that higher expression of this gene results in greater fiber production relative to vessel element or ray production. The fiber content is correlated with a principal components analysis (PCA) variable that accounts for at least 10% of the variation in basic density.

Example 8

[0260] Example 8 demonstrates how the use of oligonucleotide probes of the invention can be used to identify one wood "density related" WD40 repeat protein gene and its promoter from among the family of homologous genes. Further, this example demonstrates how a promoter sequence identified using this method is used to transform other hardwood species to result in increased diameter growth rates as compared to wild-type plants of the same species.

[0261] The sequence of the WD40 gene is used to probe a Genome Walker library in order to isolate 5' flanking sequences comprising a promoter region. The promoter region is then operably linked to a beta-glucuronidase reporter gene and cloned into a binary vector for transformation into *Eucalyptus* using the method described in U.S. Application Ser. No. 60/476,222. Regenerated transgenic tobacco and *Eucalyptus* plants are then sectioned and stained using X-gluc, demonstrating that the microarray data results in isolation of a promoter capable of highly cambial-specific expression solely in those portions of the stem that develop more fibers than vessel elements or xylem rays.

[0262] Using techniques well known to those skilled in the art of molecular biology, the promoter is then operably linked to a cell division promoting gene and this construct placed in a binary vector for transformation into hardwood plants such as Sweetgum and Populus, such that the cell division promoting gene is expressed more strongly than normally in the vascular cambium. This results in increased diameter growth rate in the transgenic hardwood plants relative to control hardwood plants.

Example 9

[0263] Example 9 demonstrates how a density related polypeptide can be linked to a tissue-preferred promoter and expressed in pine resulting in a plant with increased wood density.

[0264] A density-related polypeptide, which is more highly expressed during the early spring, is identified by the method described in Example 7. A DNA construct having the density-related polypeptide operably linked to a promoter is placed into an appropriate binary vector and transformed into pine using the method of Connett *et al.* (U.S. Patent Application Nos. 09/973,088 and 09/973,089). Pine plants are transformed as described in Connett *et al., supra,* and the transgenic pine plants are used to establish a forest planting. Increased density even in the spring wood (early wood) is observed in the transgenic pine plants relative to control pine plants which are not transformed with the density related DNA construct.

Example 10

[0265] Using techniques well known to those skilled in the art of molecular biology, the sequence of the putative density-related gene isolated in Example 7 is analyzed in genomic DNA isolated from alfalfa. This enables the identification of an orthologue in alfalfa whose sequence is then used to create an RNAi knockout construct. This construct is then transformed into alfalfa. *See, e.g.,* Austin *et al.,. Euphytica* 85, 381 1995. The regenerated transgenic plants show lower fiber content and increased ray cells content in the xylem. Such properties improved digestability which results in higher growth rates in cattle fed on this alfalfa as compared to wild-type alfalfa of the same species.

Example 11

[0266] Example 11 demonstrates how gene expression analysis can be used to find gene variants which are present in mature plants having a desirable phenotype. The presence or absence of such a variant can be used to predict the phenotype of a mature plant, allowing screening of the plants at the seedling stage. Although this example employs eucalyptus, the method used herein is also useful in breeding programs for pine and other tree species.

[0267] The sequence of a putative density-related gene is used to probe genomic DNA isolated from *Eucalyptus* that vary in density as described in previous examples. Non-transgenically produced *Eucalyptus* hybrids of different wood

phenotypes are examined. One hybrid exhibits high wood density and another hybrid exhibits lower wood density. A molecular marker in the 3' portion of the coding region is found which distinguishes a high-density gene variant from a lower density gene variant.

[0268] This molecular marker enables tree breeders to assay non-transgenic *Eucalyptus* hybrids for likely density profiles while the trees are still at seedling stage, whereas in the absence of the marker, tree breeders must wait until the trees have grown for multiple years before density at harvest age can be reliably predicted. This enables selective outplanting of the best trees at seedling stage rather than an expensive culling operation and resultant erosion at thinning age. This molecular marker is further useful in the breeding program to determine which parents will give rise to high density outcross progeny.

[0269] Molecular markers found in the 3' portion of the coding region of the gene that do not correspond to variants seen more frequently in higher or lower wood density non-transgenic *Eucalyptus* hybrid trees are also useful. These markers are found to be useful for fingerprinting different genotypes of *Eucalyptus*, for use in identity-tracking in the breeding program and in plantations.

Example 12

[0270] This Example describes microarrays for identifying gene expression differences that contribute to the phenotypic characteristics that are important in commercial wood, namely wood appearance, stiffness, strength, density, fiber dimensions, coarseness, cellulose and lignin content, extractives content and the like.

[0271] As in Examples 2-4, woody trees of genera that produce commercially important wood products, in this case *Pinus* and *Eucalyptus,* are felled from various sites and at various times of year for the collection and isolation of RNA from developing xylem, cambium, phloem, leaves, buds, roots, and other tissues. RNA is also isolated from seedlings of the same genera.

[0272] All contigs are compared to both the ESTs made from RNA isolated from samples containing developing wood and the sequences of the ESTs made from RNA of various tissues that do not contain developing wood. Contigs containing primarily ESTs that show more hybridization *in silico* to ESTs made from RNA isolated from samples containing developing wood than to ESTs made from RNA isolated from samples not containing developing wood are determined to correspond to possible novel genes particularly expressed in developing wood. These contigs are then used for BLAST searches against public domain sequences. Those contigs that hybridize with high stringency to no known genes or genes annotated as having only a "hypothetical protein" are selected for the next step. These contigs are considered putative novel genes showing wood-preferred expression.

[0273] The longest cDNA clones containing sequences hybridizing to the putative novel genes showing wood-preferred expression are selected from cDNA libraries using techniques well known to those skilled in the art of molecular biology. The cDNAs are sequenced and full-length gene-coding sequences together with untranslated flanking sequences are obtained where possible. Stretches of 45-80 nucleotides (or oligonucleotides) are selected from each of the sequences of putative novel genes showing wood-preferred expression such that each oligonucleotide probe hybridizes at high stringency to only one sequence represented in the ESTs made from RNA isolated from trees or seedlings of the same genus.

[0274] Oligomers are then chemically synthesized and placed onto a microarray slide as described in Example 3. Each oligomer corresponds to a particular sequence of a putative novel gene showing wood-preferred expression and to no other gene whose sequence is represented among the ESTs made from RNA isolated from trees or seedlings of the same genus.

[0275] Sample preparation and hybridization are carried out as in Example 4. The technique used in this example is more effective than use of a microarray using cDNA probes because the presence of a signal represents significant evidence of the expression of a *particular* gene, rather than of any of a number of genes that may contain similarities to the cDNA due to conserved functional domains or

common evolutionary history. Thus, it is possible to differentiate homologous genes, such as those in the same family, but which may have different functions in phenotype determination.

[0276] Thus hybridization data, gained using the method of Example 4, enable the user to identify which of the putative novel genes actually has a pattern of coordinate expression with known genes, a pattern of expression consistent with a particular developmental role, and/or a pattern of expression that suggests that the gene has a promoter that drives expression in a valuable way.

[0277] The hybridization data thus using this method can be used, for example, to identify a putative novel gene that shows an expression pattern particular to the tracheids with the lowest cellulose microfibril angle in developing spring wood (early wood). The promoter of this gene can also be isolated as in Example 8, and operably linked to a gene that has been shown as in Example 9 to be associated with late wood (summer wood). Transgenic pine plants containing this construct are generated using the methods of Example 9, and the early wood of these plants is then shown to display several characteristics of late wood, such as higher microfibril angle, higher density, smaller average lumen size, *etc.*

Example 13

[0278] Example 13 demonstrates the use of a cambium-specific promoter functionally linked to a cell cycle gene for increased plant biomass.

[0279] Cambium-specific cell cycle transcripts are identified via array analyses of different secondary vasculature layers as described in Example 4. Candidate promoters linked to the genes corresponding to these transcripts are cloned from pine genomic DNA using, *e.g.*, the BD Clontech GenomeWalker kit and tested in transgenic tobacco via a reporter assay(s) for cambium specificity/preference. The cambium-specific promoter overexpressing a cell cycle gene involved in secondary xylem cell division is used to increased wood biomass. A tandem cambium-specific promoter is constructed driving the cell cycle ORF. Boosted transcript levels of the candidate cell cycle gene result in an increased xylem biomass phenotype.

Example 14

Isolation and Characterization of cDNA Clones from *Eucalyptus grandis*

[0280] *Eucalyptus grandis* cDNA expression libraries were prepared from mature shoot buds, early wood phloem, floral tissue, leaf tissue (two independent libraries), feeder roots, structural roots, xylem or early wood xylem and were constructed and screened as follows.

[0281] Total RNA was extracted from the plant tissue using the protocol of Chang et al. (*Plant Molecular Biology Reporter* 11:113-116 (1993). mRNA was isolated from the total RNA preparation using either a Poly(A) Quik mRNA Isolation Kit (Stratagene, La Jolla, CA) or Dynal Beads Oligo $(dT)_{25}$ (Dynal, Skogen, Norway). A cDNA expression library was constructed from the purified mRNA by reverse transcriptase synthesis followed by insertion of the resulting cDNA clones in Lambda ZAP using a ZAP Express cDNA Synthesis Kit (Stratagene), according to the manufacturer's protocol. The resulting cDNAs were packaged using a Gigapack II Packaging Extract (Stratagene) using an aliquot (1 – 5 αl) from the 5 μl ligation reaction dependent upon the library. Mass excision of the library was done using XL1-Blue MRF' cells and XLOLR cells (Stratagene) with ExAssist helper phage (Stratagene). The excised phagemids were diluted with NZY broth (Gibco BRL, Gaithersburg, MD) and plated out onto LB-kanamycin agar plates containing X-gal and isopropylthio-beta-galactoside (IPTG).

[0282] Of the colonies plated and selected for DNA miniprep, 99% contained an insert suitable for sequencing. Positive colonies were cultured in NZY broth with kanamycin and cDNA was purified by means of alkaline lysis and polyethylene glycol (PEG) precipitation. Agarose gel at 1% was used to screen sequencing templates for chromosomal contamination. Dye primer sequences were prepared using a Turbo Catalyst 800 machine (Perkin Elmer/Applied Biosystems Division, Foster City, CA) according to the manufacturer's protocol.

[0283] DNA sequence for positive clones was obtained using a Perkin Elmer/Applied Biosystems Division Prism 377 sequencer. cDNA clones were sequenced first from the 5' end and, in some cases, also from the 3' end. For some clones, internal sequence was obtained using either Exonuclease III deletion analysis, yielding a library of differentially sized subclones in pBK-CMV, or by direct sequencing using gene-specific primers designed to identified regions of the gene of interest.

[0284] The determined cDNA sequences were compared with known sequences in the EMBL database using the computer algorithms FASTA and/or BLASTN. Multiple alignments of redundant sequences were used to build reliable consensus sequences. Based on similarity to known sequences from other plant species, the isolated polynucleotide sequences were identified as encoding transcription factors, as detailed herein. The predicted polypeptide sequences corresponding to the polynucleotide sequences are also depicted therein.

Example 15

Isolation and Characterization of cDNA Clones from *Pinus radiata*

[0285] *Pinus radiata* cDNA expression libraries (prepared from either shoot bud tissue, suspension cultured cells, early wood phloem (two independent libraries), fascicle meristem tissue, male strobilus, root (unknown lineage), feeder roots, structural roots, female strobilus, cone primordia, female receptive cones and xylem (two independent libraries) were constructed and screened as described above in Example 14.

[0286] DNA sequence for positive clones was obtained using forward and reverse primers on a Perkin Elmer/Applied Biosystems Division Prism 377 sequencer and the determined sequences were compared to known sequences in the database as described above.

[0287] Based on similarity to known sequences from other plant species, the isolated polynucleotide sequences were identified as encoding transcription factors, as

detailed herein. The predicted polypeptide sequences corresponding to the polynucleotide sequences are also depicted therein.

Example 16

5' RACE Isolation

[0288] To identify additional sequence 5' or 3' of a partial cDNA sequence in a cDNA library, 5' and 3' rapid amplification of cDNA ends (RACE) was performed. using the SMART RACE cDNA amplification kit (Clontech Laboratories, Palo Alto, Calif.). Generally, the method entailed first isolating poly(A) mRNA, performing first and second strand cDNA synthesis to generate double stranded cDNA, blunting cDNA ends, and then ligating of the SMART RACE. Adaptor to the cDNA to form a library of adaptor-ligated ds cDNA. Gene-specific primers were designed to be used along with adaptor specific primers for both 5' and 3' RACE reactions. Using 5' and 3' RACE reactions, 5' and 3' RACE fragments were obtained, sequenced, and cloned. The process may be repeated until 5' and 3' ends of the full-length gene were identified. A full-length cDNA may generated by PCR using primers specific to 5' and 3' ends of the gene by end-to-end PCR.

[0289] For example, to amplify the missing 5' region of a gene from first-strand cDNA, a primer was designed 5'→3' from the opposite strand of the template sequence, and from the region between ~100-200 bp of the template sequence. A successful amplification should give an overlap of ~100 bp of DNA sequence between the 5' end of the template and PCR product.

[0290] RNA was extracted from four pine tissues, namely seedling, xylem, phloem and structural root using the Concert Reagent Protocol (Invitrogen, Carlsbad, CA) and standard isolation and extraction procedures. The resulting RNA was then treated with DNase, using 10U/ul DNase I (Roche Diagnostics, Basel, Switzerland). For 100 μg of RNA, 9 μl 10x DNase buffer (Invitrogen, Carlsbad, CA), 10 μl of Roche DNase I and 90 μl of Rnase-free water was used. The RNA was then incubated at room temperature for 15 minutes and 1/10 volume 25mM EDTA is

added. A RNeasy mini kit (Qiagen, Venlo, The Netherlands) was used for RNA clean up according to manufacturer's protocol.

[0291] To synthesize cDNA, the extracted RNA from xylem, phloem, seedling and root was used and the SMART RACE cDNA amplification kit (Clontech Laboratories Inc, Palo Alto, CA) was followed according to manufacturer's protocol. For the RACE PCR, the cDNA from the four tissue types was combined. The master mix for PCR was created by combining equal volumes of cDNA from xylem, phloem, root and seedling tissues. PCR reactions were performed in 96 well PCR plates, with 1µl of primer from primer dilution plate (10mM) to corresponding well positions. 49 µl of master mix is aliquoted into the PCR plate with primers. Thermal cycling commenced on a GeneAmp 9700 (Applied Biosystems, Foster City, CA) at the following parameters:

94^0C (5 sec),

72^0C (3 min), 5 cycles;

94^0C (5 sec),

70^0C (10 sec),

72^0C (3 min), 5 cycles;

94^0C (5 sec),

68^0C (10 sec),

72^0C (3min), 25 cycles.

[0292] cDNA was separated on an agarose gel following standard procedures. Gel fragments were excised and eluted from the gel by using the Qiagen 96-well Gel Elution kit, following the manufacturer's instructions.

[0293] PCR products were ligated into pGEMTeasy (Promega, Madison, WI) in a 96 well plate overnight according to the following specifications: 60-80 ng

of DNA, 5 µl 2X rapid ligation buffer, 0.5 µl pGEMT easy vector, 0.1 µl DNA ligase, filled to 10 µl with water, and incubated overnight.

[0294] Each clone was transformed into *E.coli* following standard procedures and DNA was extracted from 12 clones picked by following standard protocols. DNA extraction and the DNA quality was verified on an 1% agarose gel. The presence of the correct size insert in each of the clones was determined by restriction digests, using the restriction endonuclease *Eco*RI, and gel electrophoresis, following standard laboratory procedures.

Example 17

Curation of an EST sequence.

[0295] During the production of cDNA libraries, the original transcripts or their DNA counterparts may have features that prevent them from coding for functional proteins. There may be insertions, deletions, base substitutions, or unspliced or improperly spliced introns. If such features exist, it is often possible to identify them so that they can be changed. Similar curation can be performed on any other sequences that have homology to sequences in the public databases.

[0296] After determination of the DNA sequence, BLAST analysis shows that it is related to an *Arabidopsis* gene on the publicly available *Arabidopsis* genome sequence). However, instead of coding for an approximately 240 amino acid polypeptide, the consensus being curated is predicted to code for a product of only 157 amino acid residues, suggesting an error in the DNA sequence. To identify where the genuine coding region might be, the DNA sequence to the end of each EST is translated in each of the three reading frames and the predicted sequences are aligned with the Arabidopsis gene's amino acid sequence. It is found that the DNA segment in one portion of the EST codes for a sequence with similarity to the carboxyl terminus of the Arabidopsis gene. Therefore, it appears that an unspliced intron is present in the EST.

[0297] Unspliced introns are a relatively minor issue with regard to use of a cloned sequence for overexpression of the gene of interest. The RNA resulting from transcription of the cDNA can be expected to undergo normal processing to remove the intron. Antisense and RNAi constructs are also expected to function to suppress the gene of interest. On other occasions, it may be desirable to identify the precise limits of the intron so that it can be removed. When the sequence in question has a published sequence that is highly similar, it may be possible to find the intron by aligning the two sequences and identifying the locations where the sequence identity falls off, aided by the knowledge that introns start with the sequence GT and end with the sequence AG.

[0298] When there is some doubt about the site of the intron because highly similar sequences are not available, the intron location can be verified experimentally. For example, DNA oligomers can be synthesized flanking the region where the suspected intron is located. RNA from the source species, either Pinus or Eucalyptus, is isolated and used as a template to make cDNA using reverse transcriptase. The selected primers are then used in a PCR reaction to amplify the correctly spliced DNA segment (predicted size of approximately 350 bp smaller than the corresponding segment of the original consensus) from the population of cDNAs. The amplified segment is then subjected to sequence analysis and compared to the consensus sequence to identify the differences.

[0299] The same procedure can be used when an alternate splicing event (partial intron remaining, or partial loss of an exon) is suspected. When an EST has a small change, such as insertion or deletion of a small number of bases, computer analysis of the EST sequence can still indicate its location when a translation product of the wrong size is predicted or if there is an obvious frameshift. Verification of the true sequence is done by synthesis of primers, production of new cDNA, and PCR amplification as described above.

Example 18

[0300] Transformation of *Populus deltoides* with constructs containing cell cycle genes.

[0301] Constructs made as described in the preceding example and shown in Table 2 below were each inoculated into *Agrobacterium* cultures by standard techniques.

[0302] Table 2 identifes plasmid(s), genes, and Genesis ID numbers for constructions described in Example 17.

[0303] Table 2

Plasmid(s)	Gene	Genesis ID
pGrw14	Cyclin A	prga001823
pGrw15	Cyclin A	prpe001264
pGrw16	Cyclin D	prxa004540
pGrw18	Cyclin D	prxl006271
PGrw19	Cyclin D	prpb019661
PGrw20	WEE1-like protein	prrd041233

[0304] *Populus deltoides* stock plant cultures were maintained on DKW medium (Driver and Kuniyuki, 1984, McGranahan et al. 1987, available commercially from Sigma/Aldrich) with 2.5 uM zeatin in a growth room with a 16h photoperiod. For transformation, petioles were excised aseptically using a sharp scalpel blade from the stock plants, cut into 4-6mm lengths, placed on DKW medium with 1 ug/ml BAP and 1 ug/ml NAA immediately after harvest, and incubated in a dark growth chamber (28 degrees) for 24 hours.

[0305] *Agrobacterium* cultures containing the desired constructs were grown to log phase, indicated by an OD600 between 0.8-1.0 A, then pelleted and resuspended in an equal volume of Agrobacterium Induction Medium (AIM), which contains Woody Plant Medium salts (Lloyd, G., and McCown, B., 1981. Woody plant

medium. Proc. Intern. Plant Prop. Soc. 30:421, available commercially from Sigma/Aldrich), 5 g/L glucose and 0.6 g/L MES at pH 5.8, with the addition of 1 ul of a 100 mM stock solution of acetosyringone per ml of AIM. The pellet was resuspended by vortexing. The bacterial cells were incubated for an hour in this medium at 28 degrees C in an environmental chamber, shaking at 100 rpm.

[0306] After the induction period, *Populus deltoides* explants were exposed to the Agrobacterium mixture for 15 minutes. The explants were then lightly blotted on sterile paper towels, replaced onto the same plant medium and cultured in the dark at 18-20 degrees C. After a three-day co-cultivation period, the explants were transferred to DKW medium in which the NAA concentration was reduced to 0.1 ug/ml and to which was added 400mg/L timentin to eradicate the Agrobacterium.

[0307] After 4 days on eradication medium, explants were transferred to small magenta boxes containing the same medium supplemented with timentin (400mg/L) as well as the selection agent geneticin (50 mg/L). Explants were transferred every two weeks to fresh selection medium. Calli that grow in the presence of selection were isolated and sub-cultured to fresh selection medium every three weeks. Calli were observed for the production of adventitious shoots.

[0308] Adventitious shoots were normally observed within two months from the initiation of transformation. These shoot clusters were transferred to DKW medium to which no NAA was added, and in which the BAP concentration was reduced to 0.5 ug.ml, for shoot elongation, typically for about 14 weeks. Elongated shoots were excised and transferred to BTM medium (Chalupa, Communicationes Instituti Forestalis Checosloveniae 13:7-39, 1983, available commercially from Sigma/Aldrich) at pH5.8, containing20 g/l sucrose and 5 g/l activated charcoal. See Table 3 below.

[0309] **Table 3. Rooting medium for *Populus deltoids*.**

BTM-1 Media Components	mg/L
NH_4NO_3	412
KNO_3	475
$Ca(NO_3)_2.4H_2O$	640
$CaCl_2.2H_2O$	440*
$MgSO_4.7H_2O$	370
KH_2PO_4	170
$MnSO_4.H_2O$	2.3
$ZnSO_4.7H_2O$	8.6
$CuSO_4.5H_2O$	0.25
$CoCl_2.6H_2O$	0.02
KI	0.15
H_3BO_3	6.2
$Na_2MoO_4.2H_2O$	0.25
$FeSO_4.7H_2O$	27.8
$Na_2EDTA.2H_2O$	37.3
Myo-inositol	100
Nicotinic acid	0.5
Pyridoxine HCl	0.5
Thiamine HCl	1
Glycine	2
Sucrose	20000
Activated Carbon	5000

[0310] After development of roots, typically four weeks, transgenic plants were propagated in the greenhouse by rooted cutting methods, or in vitro through axillary shoot induction for four weeks on DKW medium containing 11.4 uM zeatin, after which the multiplied shoots were separated and transferred to root induction medium. Rooted plants were transferred to soil for evaluation of growth in glasshouse and field conditions.

Example 19

[0311] Production of disproportionately large leaves mediated by ectopic expression of certain cyclin D genes

[0312] Approximately 100 explants of Populus deltoides per construct were transformed with pGRW16 and pGRW19, which contain genes that are normally show preferred expression in the vasculature, driven by a constitutive promoter (the Pinus radiata superubiquitin promoter). Upon regeneration, many of the ramets of many of the translines were observed to have disproportionately large leaves relative to control plants. The leaves were both longer and broader than those of control plants.

[0313] Disproportionately large leaves could be a very useful early indicator of growth potential. large leaf size and thus high growth potential. Lage leaf size can be a function of either increased numbers of leaf cells or increased leaf cell size or both.

Example 20

[0314] Production of unusual vascular development mediated by ectopic expression of a cyclin D gene.

[0315] Approximately 100 explants of *Populus deltoides* per construct were transformed with pGRW18. Multiple transgenic lines regenerated from this experiment showed a very unique pleiotropic phenotype. Leaves of these transgenic lines symmetrically folded on both sides of the midrib down the entire length of the leaf. Many petioles of these lines spiraled, and in many cases turned 360 degrees, in a right-handed fashion towards the leaf. The stem showed some thickening and slight bending near the middle.

[0316] One ramet of the transgenic line TDL002534 showing these phenotypes was sacrificed to investigate these aberrancies at the tissue level. Transverse sections of a curling petiole stained with toluidine blue revealed retardation of vascular development, but the presence of additional vascular cylinders

developing as indicated by the black arrows. The xylem and phloem within the vascular cylinders of the curling petiole appeared to be developmentally similar and spatially oriented correctly. Longitudinal sections of straight and curled petioles may offer an explanation for the spiraling phenomenon. Curled petioles showed more elongated cells on the outside turn of the curl and more compressed cells on the opposite side of the petiole.

[0317] Perhaps the most striking phenotype was identified in the leaves. As with the petioles, aberrant vascular development was noted, comprising additional forming vascular cylinders lateral to the larger midrib. In some sections almost fully-formed veins could be seen immediately adjacent to the midrib. In all instances where the folding phenotype was noted, this type of leaf configuration was associated with the phenotype.

[0318] The development of additional vascular cylinders in the space where normally a small number of vascular bundles or a single midrib are seen is indicative of unusual cell division activity at the level of early vascular development. Thus, this gene expressed under the control of a vascular-preferred promoter rather than a constitutive promoter could have utility in increasing cell division in later vascular development, creating additional wood.

Example 21

[0319] This example illustrates how polynucleotides important for wood development in P. radiata can be determined and how oligonucleotides which uniquely bind to those genes can be designed and synthesized for use on a microarray.

[0320] Open pollinated trees of approximately 16 years of age are selected from plantation-grown sites, in the United States for loblolly pine, and in New Zealand for radiata pine. Trees are felled during the spring and summer seasons to compare the expression of genes associated with these different developmental stages of wood formation. Trees are felled individually and trunk sections are removed from the bottom area approximately one to two meters from the base and within one to two meters below the live crown. The section removed from the basal end of the trunk

contains mature wood. The section removed from below the live crown contains juvenile wood. Samples collected during the spring season are termed earlywood or springwood, while samples collected during the summer season are considered latewood or summerwood (Larson et al., Gen. Tech. Rep. FPL-GTR-129. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. p. 42).

[0321] Tissues are isolated from the trunk sections such that phloem, cambium, developing xylem, and maturing xylem are removed. These tissues are collected only from the current year's growth ring. Upon tissue removal in each case, the material is immediately plunged into liquid nitrogen to preserve the nucleic acids and other components. The bark is peeled from the section and phloem tissue removed from the inner face of the bark by scraping with a razor blade. Cambium tissue is isolated from the outer face of the peeled section by gentle scraping of the surface. Developing xylem and lignifying xylem are isolated by sequentially performing more vigorous scraping of the remaining tissue. Tissues are transferred from liquid nitrogen into containers for long term storage at -70 until RNA extraction and subsequent analysis is performed.

Example 22

[0322] This example illustrates a procedure for RNA extraction and purification, which is particularly useful for RNA obtained from conifer needle, xylem, cambium, and phloem.

[0323] Tissue is obtained from conifer needle, xylem, cambium or phloem. The tissue is frozen in liquid nitrogen and ground. The total RNA is extracted using Concert Plant RNA reagent (Invitrogen). The resulting RNA sample is extracted into phenol:chloroform and treated with DNase. The RNA is then incubated at 65 °C for 2 minutes followed by centrifugation at 4 °C for 30 minutes. Following centrifugation, the RNA is extracted into phenol at least 10 times to remove contaminants.

[0324] The RNA is further cleaned using RNeasy columns (Qiagen). The purified RNA is quantified using RiboGreen reagent (Molecular Probes) and purity assessed by gel electrophoresis.

[0325] RNA is then amplified using MessageAmp (Ambion). Aminoallyl-UTP and free UTP are added to the in vitro transcription of the purified RNA at a ratio of 4:1 aminoallyl-UTP-to-UTP. The aminoallyl-UTP is incorporated into the new RNA strand as it is transcribed. The amino-allyl group is then reacted with Cy dyes to attach the colorimetric label to the resulting amplified RNA using the Amersham procedure modified for use with RNA. Unincorporated dye is removed by ethanol precipitation. The labeled RNA is quantified spectrophotometrically (NanoDrop). The labeled RNA is fragmented by heating to 95 °C as described in Hughes et al., *Nature Biotechnol.* 19:342 (2001).

Example 23

[0326] This Example illustrates how genes important for wood development in *P. radiata* can be determined and how oligonucleotides which uniquely bind to those genes can be designed and synthesized for use on a microarray.

[0327] Pine trees of the species *P. radiata* are grown under natural light conditions. Tissue samples are prepared as described in, *e.g.*, Sterky *et al.*, *Proc. Nat'l Acad. Sci.* 95:13330 (1998). Specifically, tissue samples are collected from woody trees having a height of 5 meters. Tissue samples of the woody trees are prepared by taking tangential sections through the cambial region of the stem. The stems are sectioned horizontally into sections ranging from juvenile (top) to mature (bottom). The stem sections separated by stage of development are further separated into 5 layers by peeling into sections of phloem, differentiating phloem, cambium, differentiating xylem, developing xylem, and mature xylem. Tissue samples, including leaves, buds, shoots, and roots are also prepared from seedlings of the species *P. radiata*.

[0328] RNA is isolated and ESTs generated as described in the Example above or Sterky *et al.*, *supra*. The nucleic acid sequences of ESTs derived from samples

containing developing wood are compared with nucleic acid sequences of genes known to be involved in polysaccharide synthesis. ESTs from samples that do not contain developing wood are also compared with sequences of genes known to be involved in the plant cell cycle. An *in silico* hybridization analysis is performed using BLAST (NCBI) as follows.

Example 24

***Eucalyptus* in silico data**

[0329] In silico gene expression can be used to determine the membership of the consensi EST libraries. For each library, a consensus is determined from the number of ESTs in any tissue class divided by the total number of ESTs in a class multiplied by 1000. These values provide a normalized value that is not biased by the extent of sequencing from a library. Several libraries were sampled for a consensus value, including reproductive, bud reproductive, bud vegetative, fruit, leaf, phloem, cambium, xylem, root, stem, sap vegetative, whole plant libraries.

[0330] As shown below, a number of the inventive sequences exhibit vascular-preferred expression (more than 50% of the hits by these sequences if the databases were searched at random would be in libraries made from developing vascular tissue) and thus are likely to be involved in wood-related developmental processes. The data are shown in Table 12.

Example 25

***Pinus* in silico data**

[0331] In silico gene expression can be used to determine the membership of the consensi EST libraries. For each library, a consensus is determined from the number of ESTs in any tissue class divided by the total number of ESTs in a class multiplied by 1000. These values provide a normalized value that is not biased by the extent of sequencing from a library. Several libraries were sampled for a consensus value, including needles, phloem, cambium, xylem, root, stem and, whole plant libraries.

[0332] As shown below, a number of the inventive sequences exhibit vascular-preferred expression (more than 50% of the hits by these sequences if the databases were searched at random would be in libraries made from developing vascular tissue) and thus are likely to be involved in wood-related developmental processes. The data are shown in Table 13.

Example 26

[0333] Sequences that show hybridization *in silico* to ESTs made from samples containing developing wood, but that do not hybridize to ESTs from samples not containing developing wood are selected for further examination.

[0334] cDNA clones containing sequences that hybridize to the genes showing wood-preferred expression are selected from cDNA libraries using techniques well known in the art of molecular biology. Using the sequence information, oligonucleotides are designed such that each oligonucleotide is specific for only one cDNA sequence in the library. The oligonucleotide sequences are provided in Table 14. 60-mer oligonucleotide probes are designed using the method of Li and Stormo, supra or using software such as ArrayDesigner, GeneScan, and ProbeSelect.

[0335] The oligonucleotides are then synthesized *in situ* described in Hughes *et al., Nature Biotechnol.* 19:324 (2002) or as described in Kane *et al., Nucleic Acids Res.* 28:4552 (2000) and affixed to an activated glass slide (Sigma-Genosis, The Woodlands, TX) using a 5' amino linker. The position of each oligonucleotide on the slide is known.

Example 27

[0336] This example illustrates how to detect expression of *Pinus radiata* genesof the instant application which are important in wood formation using an oligonucleotide microarray prepared as described above. This is an example of a balanced incomplete block designed experiment carried out using aRNA samples prepared from mature-phase phloem (P), cambium (C), expanding xylem found in a layer below the cambium (X1) and differentiating, lignifying xylem cells found

2004311384 31 May 2011

deeper in the same growth ring (X2). In this example, cell cycle gene expression is compared among the four samples, namely P, C, X1, and X2.

[0337] In the summer, plants of the species *Pinus radiata* are felled and the bark of the main stem is immediately pulled gently away to reveal the phloem and xylem. The phloem and xylem are then peeled with a scalpel into separate containers of liquid nitrogen. Needles (leaves) and buds from the trees are also harvested with a scalpel into separate containers of liquid nitrogen. RNA is subsequently isolated from the frozen tissue samples as described in Example 1. Equal microgram quantities of total RNA are purified from each sample using RNeasy Mini columns (Qiagen, Valencia, CA) according to the manufacturer's instructions.

[0338] Amplification reactions are carried out for each of the P, C, X1, and X2 tissue samples. Amplification reactions are performed using Ambion's MessageAmp kit, a T7-based amplification procedure, following the manufacturer's instructions, except that labeled aaUTP is added to the reagent mix during in the amplification step. aaUTP is incorporated into the resulting antisense RNA formed during this step. CyDye fluorescent labels are coupled to the aaUTPs in a non-enzymatic reaction as described in Example 1. Labeled amplified antisense RNAs are precipitated and washed, and then assayed for purity using a NanoDrop spectrophotometer. These labeled antisense RNAs, corresponding to the RNA isolated from the P, C, X1, and X2 tissue samples, constitute the sample nucleic acids, which are referred to as the P, C, X1, and X2 samples.

[0339] Normalization control samples of known nucleic acids are added to each sample in a dilution series of 500, 200, 100, 50, 25 and 10 pg/μl for quantitation of the signals. Positive controls corresponding to specific genes showing expression in all tissues of pine, such as housekeeping genes, are also added to the plant sample.

[0340] Each of four microarray slides is incubated with 125 μL of a P, C, X1 or X2 sample under a coverslip at 42 °C for 16-18 hours. The arrays are washed in 1x SSC, 0.1% SDS for 10 minutes and then in 0.1 x SSC, 0.1% SDS for 10 minutes and the allowed to dry.

[0341] The array slides are scanned using an Axon laser scanner and analyzed using GenePix Pro software. Data from the microarray slides are subjected to microarray data analysis using GenStat SAS or Spotfire software. Outliers are removed and ratiometric data for each of the datasets are normalized using a global normalization which employs a cubic spline fit applied to correct for differential dye bias and spatial effects. A second transformation is performed to fit control signal ratios to a mean $\log^2 = 0$ (*i.e.* 1:1 ratio). Normalized data are then subjected to a variance analysis.

[0342] Mean signal intensity for each signal at any given position on the microarray slide is determined for each of three of P, C, X1, and X2 sample microarray slides. This mean signal/probe position is compared to the signal at the same position on sample slide which was not used for calculating the mean. For example, a mean signal at a given position is determined for P, C, and X1 and the signal at that position in the X2 microarray slide is compared to the P, C, and X1 mean signal value.

[0343] Table 5 shows genes having greater than doubled signal with any one sample as compared to the mean signal of the other three samples.

[0344] **Table 5**

Gene	PvCX12	PvX12	CvX12
WD40 repeat protein A	-1.24	-0.88	-1.07
CDC2	-1.09	-0.78	-0.92
CYCLIN	-1.08	-1	-0.26
WD-40 repeat protein B	-1.01	-0.87	-0.42
CDC2	-0.83	-0.49	-1.01

P = Phloem

C = Cambium

X1 = xylem layer-1

X2 = xylem layer-2

PvCX12 = Ratio of the signal for Phloem target versus mean signal for Cambium, Xylem1, and Xylem2 targets

[0345] The data shows that WD40 repeat protein A encodes a WD40 repeat protein is less highly expressed in cambium than in developing xylem, while WD40 repeat protein B encodes a WD40 repeat protein that is more highly expressed in phloem than in the other tissues.

[0346] Signal data are then verified with RT-PCR to confirm gene expression in the target tissue of the genes corresponding to the unique oligonucleotides in the probe.

Example 28

[0347] This example illustrates how RNAs of tissues from multiple pine species, in this case both *P. radiata* and loblolly pine *P. taeda* trees, are selected for analysis of the pattern of gene expression associated with wood development in the juvenile wood and mature wood forming sections of the trees using the microarrays derived from *P. radiata* cDNA sequences described in Example 4.

[0348] Open pollinated trees of approximately 16 years of age are selected from plantation-grown sites, in the United States for loblolly pine, and in New Zealand for radiata pine. Trees are felled during the spring and summer seasons to compare the expression of genes associated with these different developmental stages of wood formation. Trees are felled individually and trunk sections are removed from the bottom area approximately one to two meters from the base and within one to two meters below the live crown. The section removed from the basal end of the trunk contains mature wood. The section removed from below the live crown contains juvenile wood. Samples collected during the spring season are termed earlywood or springwood, while samples collected during the summer season are considered latewood or summerwood. Larson *et al., Gen. Tech. Rep.* FPL-GTR-129. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. p. 42.

[0349] Tissues are isolated from the trunk sections such that phloem, cambium, developing xylem, and maturing xylem are removed. These tissues are collected only from the current year's growth ring. Upon tissue removal in each case,

the material is immediately plunged into liquid nitrogen to preserve the nucleic acids and other components. The bark is peeled from the section and phloem tissue removed from the inner face of the bark by scraping with a razor blade. Cambium tissue is isolated from the outer face of the peeled section by gentle scraping of the surface. Developing xylem and lignifying xylem are isolated by sequentially performing more vigorous scraping of the remaining tissue. Tissues are transferred from liquid nitrogen into containers for long term storage at -70 °C until RNA extraction and subsequent analysis is performed.

Example 29

[0350] This example illustrates procedures alternative to those used in the example above for RNA extraction and purification, particularly useful for RNA obtained from a variety of tissues of woody plants, and a procedure for hybridization and data analysis using the arrays described in Example 4.

[0351] RNA is isolated according to the protocol of Chang et al., Plant Mol. Biol. Rep. 11:113. DNA is removed using DNase I (Invitrogen, Carlsbad, CA) according to the manufacturer's recommendations. The integrity of the RNA samples is determined using the Agilent 2100 Bioanalyzer (Agilent Technologies, USA).

[0352] 10μg of total RNA from each tissue is reverse transcribed into cDNA using known methods.

[0353] In the case of *Pinus radiata* phloem tissue, it can be difficult to extract sufficient amounts of total RNA for normal labelling procedures. Total RNA is extracted and treated as previously described and 100ng of total RNA is amplified using the Ovation™ Nanosample RNA Amplification system from NuGEN™ (NuGEN, CA, USA). Similar amplification kits such as those manufactured by Ambion may alternatively be used. The amplified RNA is reverse transcribed into cDNA and labelled as described above.

[0354] Hybridization and stringency washes are performed using the protocol as described in the US Patent Application for "Methods and Kits for

Labeling and Hybridizing cDNA for Microarray Analysis" (supra) at 42 C. The arrays (slides) are scanned using a ScanArray 4000 Microarray Analysis System (GSI Lumonics, Ottawa, ON, Canada). Raw, non-normalized intensity values are generated using QUANTARRAY software (GSI Lumonics, Ottawa, ON, Canada).

[0355] A fully balanced, incomplete block experimental design (Kerr and Churchill, *Gen. Res.* 123:123, 2001) is used in order to design an array experiment that would allow maximum statistical inferences from analyzed data.

[0356] Gene expression data is analyzed using the SAS® Microarray Solution software package (The SAS Institute, Cary, NC, USA). Resulting data was then visualized using JMP® (The SAS Institute, Cary, NC, USA).

[0357] Analysis done for this experiment is an ANOVA approach with mixed model specification (Wolfinger *et al., J. Comp. Biol.* 8:625-637). Two steps of linear mixed models are applied. The first one, normalization model, is applied for global normalization at slide-level. The second one, gene model, is applied for doing rigorous statistical inference on each gene. Both models are stated in Models (1) and (2).

$$\log_2(Y_{ijkls}) = \theta_{ij} + D_k + S_l + DS_{kl} + \omega_{ijkls} \quad (1)$$

$$R^{(g)}_{ijkls} = \mu^{(g)}_{ij} + D^{(g)}_k + S^{(g)}_l + DS^{(g)}_{kl} + SS^{(g)}_{ls} + \varepsilon^{(g)}_{ijkls} \quad (2)$$

[0358] Y_{ijkls} represents the intensity of the s^{th} spot in the l^{th} slide with the k^{th} dye applying the j^{th} treatment for the i^{th} cell line. θ_{ij}, D_k, S_l, and D_{Skl} represent the mean effect of the jth treatment in the ith cell line, the kth dye effect, the l^{th} slide random effect, and the random interaction effect of the k^{th} dye in the l^{th} slide. ω_{ijkls} is the stochastic error term. represent the similar roles as θ_{ij}, D_k, S_l, and D_{Skl} except they are specific for the g^{th} gene. $R^{(g)}_{ijkls}$ represents the residual of the g^{th} gene from model (1). $\mu^{(g)}_{ij}, D^{(g)}_k, S^{(g)}_l$, and $DS^{(g)}_{kl}$ represent the similar roles as θ_{ij}, D_k, S_l, and DS_{kl} except they are specific for the g^{th} gene. $SS^{(g)}_{ls}$ represent the spot by slide random effect for the g^{th} gene. $\varepsilon^{(g)}_{ijkls}$ represent the stochastic error term. All random terms are assumed to be normal distributed and mutually independent within each model.

[0359] According to the analysis described above, certain cDNAs, some of which are shown in Table 6 below, are found to be differentially expressed.

[0360] Table 6

Gene corresponding to SEQ ID	Oligo ID	Gene_Family	Expression
162	Pra_000171_O_2	Peptidylprolyl isomerase	steady state RNA higher in xylem than cambium
164	Pra_001480_O_3	Peptidylprolyl isomerase	steady state RNA lower in xylem than cambium
control	Pra_000218_O_2	RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LARGE CHAIN (EC1.17.4.1).	steady state RNA lower in xylem than cambium
control	Pra_000193_O_2	PUTATIVE SURFACE PROTEIN.	steady state RNA lower in xylem than cambium

[0361] The involvement of these specific genes in wood development is inferred through the association of the up-regulation or down-regulation of genes to the particular stages of wood development. Both the spatial continuum of wood development across a section (phloem, cambium, developing xylem, maturing xylem) at a particular season and tree trunk position and the relationships of season and tree trunk position are considered when making associations of gene expression to the relevance in wood development.

Example 30

[0362] This example demonstrates how one can correlate polysaccharide gene expression with agronomically important wood phenotypes such as density, stiffness, strength, distance between branches, and spiral grain.

[0363] Mature clonally propagated pine trees are selected from among the progeny of known parent trees for superior growth characteristics and resistance to important fungal diseases. The bark is removed from a tangential section and the trees are examined for average wood density in the fifth annual ring at breast height,

stiffness and strength of the wood, and spiral grain. The trees are also characterized by their height, mean distance between major branches, crown size, and forking.

[0364] To obtain seedling families that are segregating for major genes that affect density, stiffness, strength, distance between branches, spiral grain and other characteristics that may be linked to any of the genes affecting these characteristics, trees lacking common parents are chosen for specific crosses on the criterion that they exhibit the widest variation from each other with respect to the density, stiffness, strength, distance between branches, and spiral grain criteria. Thus, pollen from a tree exhibiting high density, low mean distance between major branches, and high spiral grain is used to pollinate cones from the unrelated plus tree among the selections exhibiting the lowest density, highest mean distance between major branches, and lowest spiral grain. It is useful to note that "plus trees" are crossed such that pollen from a plus tree exhibiting high density are used to pollinate developing cones from another plus tree exhibiting high density, for example, and pollen from a tree exhibiting low mean distance between major branches would be used to pollinate developing cones from another plus tree exhibiting low mean distance between major branches.

[0365] Seeds are collected from these controlled pollinations and grown such that the parental identity is maintained for each seed and used for vegetative propagation such that each genotype is represented by multiple ramets. Vegetative propagation is accomplished using micropropagation, hedging, or fascicle cuttings. Some ramets of each genotype are stored while vegetative propagules of each genotype are grown to sufficient size for establishment of a field planting. The genotypes are arrayed in a replicated design and grown under field conditions where the daily temperature and rainfall are measured and recorded.

[0366] The trees are measured at various ages to determine the expression and segregation of density, stiffness, strength, distance between branches, spiral grain, and any other observable characteristics that may be linked to any of the genes affecting these characteristics. Samples are harvested for characterization of cellulose content, lignin content, cellulose microfibril angle, density, strength, stiffness,

tracheid morphology, ring width, and the like. RNA is then collected from replicated samples of trees showing divergent stiffness and density, or other characteristics, from genotypes that are otherwise as similar as possible in growth habit, in spring and fall so that early and late wood development is assayed. These samples are examined for gene expression similarly as described in above examples.

[0367] Table 7. Concensus ID Information.

Patent app	SEQ ID	Gene Family	Consensus_ID	Expression
–	control	Ribonucleoside-diphosphate reductase	pinusRadiata_000218	up in early spring xylem vs late summer xylem
Cell Cycle	168	Peptidylprolyl isomerase	pinusRadiata_001692	up in juvenile developing wood vs mature developing xylem
–	control	Nitrite transporter	pinusRadiata_016801	up mature developing xylem vs juvenile cambium

[0368] Ramets of each genotype are compared to ramets of the same genotype at different ages to establish age:age correlations for these characteristics.

Example 31

[0369] Example 8 demonstrates how responses to environmental conditions such as light and season alter plant phenotype and can be correlated to polysaccharide synthesis gene expression using microarrays. In particular, the changes in gene expression associated with wood density are examined.

[0370] Trees of three different clonally propagated *E. grandis* hybrid genotypes are grown on a site with a weather station that measures daily temperatures and rainfall. During the spring and subsequent summer, genetically identical ramets of the three different genotypes are first photographed with north-south orientation marks, using photography at sufficient resolution to show bark characteristics of juvenile and mature portions of the plant, and then felled. The age of the trees is

determined by planting records and confirmed by a count of the annual rings. In each of these trees, mature wood is defined as the outermost rings of the tree below breast height, and juvenile wood as the innermost rings of the tree above breast height. Each tree is accordingly sectored as follows:

NM - NORTHSIDE **MATURE**

SM - SOUTHSIDE **MATURE**

NT - NORTHSIDE **TRANSITION**

ST - SOUTHSIDE **TRANSITION**

NJ - NORTHSIDE **JUVENILE**

SJ - SOUTHSIDE **JUVENILE**

[0371] Tissue is harvested from the plant trunk as well as from juvenile and mature form leaves. Samples are prepared simultaneously for phenotype analysis, including plant morphology and biochemical characteristics, and gene expression analysis. The height and diameter of the tree at the point from which each sector was taken is recorded, and a soil sample from the base of the tree is taken for chemical assay. Samples prepared for gene expression analysis are weighed and placed into liquid nitrogen for subsequent preparation of RNA samples for use in the microarray experiment. The tissues are denoted as follows:

P - phloem

C - cambium

X1 - expanding xylem

X2 - differentiating and lignifying xylem

[0372] Thin slices in tangential and radial sections from each of the sectors of the trunk are fixed as described in Ruzin, PLANT MICROTECHNIQUE AND

MICROSCOPY, Oxford University Press, Inc., New York, NY (1999) for anatomical examination and confirmation of wood developmental stage. Microfibril angle is examined at the different developmental stages of the wood, for example juvenile, transition and mature phases of Eucalyptus grandis wood. Other characteristics examined are the ratio of fibers to vessel elements and ray tissue in each sector. Additionally, the samples are examined for characteristics that change between juvenile and mature wood and between spring wood and summer wood, such as fiber morphology, lumen size, and width of the S2 (thickest) cell wall layer. Samples are further examined for measurements of density in the fifth ring and determination of modulus of elasticity using techniques well known to those skilled in the art of wood assays. *See, e.g., Wang, et al., Non-destructive Evaluations of Trees,* EXPERIMENTAL TECHNIQUES, pp. *28*-30 (2000).

[0373] For biochemical analysis, 50 grams from each of the harvest samples are freeze-dried and analyzed, using biochemical assays well known to those skilled in the art of plant biochemistry for quantities of simple sugars, amino acids, lipids, other extractives, lignin, and cellulose. *See, e.g.,* Pettersen & Schwandt, *J. Wood Chem. & Technol.* 11:495 (1991).

[0374] In the present example, the phenotypes chosen for comparison are high density wood, average density wood, and low density wood. Nucleic acid samples are prepared as described in Example 3, from trees harvested in the spring and summer. Gene expression profiling by hybridization and data analysis is performed as described above.

[0375] Using similar techniques and clonally propagated individuals one can examine polysaccharide gene expression as it is related to other complex wood characteristics such as strength, stiffness and spirality.

Example 32

[0376] Example 32 demonstrates the use of a vascular-preferred promoter functionally linked to one of the genes of the instant application.

[0377] A vascular-preferred promoter is then linked to one of the genes in the instant application and used to transform tree species. Boosted transcript levels of the candidate gene in the xylem of the transformants results in an increased xylem biomass phenotype.

[0378] In another example, a vascular-preferred promoter such as any of those in ArborGen's November 2003 patent applications is then linked to an RNAi construct containing sequences from one of the genes in the instant application and used to transform a tree of the genus from which the gene was isolated. Reduced transcript levels of the candidate gene in the xylem of the transformants results in an increased xylem biomass phenotype.

Example 33

[0379] The vector pARB476 was developed using the following steps. The Bluescript vector (Stratagene, La Jolla, CA) was modified by adding the Superubiquitin 3'UTR and nos 3'terminator sequence at the KpnI and ClaI sites to produce the vector pARB005 (SEQ ID NO. 773). To this vector the *P. radiata* superubiquitin promoter with intron was added. The promoter/intron sequence was first amplified from the *P. radiata* superubiquitin sequence identifed in U.S. patent No. 6,380,459 using standard PCR techniques and the primers of SEQ ID NOS 774 and 775. The amplified fragment was then ligated into pARB005 using *XbaI* and *PstI* restriction digestion to produce the vector pARB119 (SEQ ID NO. 776).

[0380] The *poplus tremuloises* UDB Glucose binding domain gene (patent W0 OO71670, ptCelA Genbank number AF072131) was amplified using standard PCR techniques and primers including and ATG and a Cla1 site as part of the 5' primer and a TGA and a ClaI site as part of the 3' primer. The amplified fragment was then cloned into the ClaI site of pARB119 to produce the vector pARB476 (SEQ ID NO. 777).

[0381] The NotI cassette containing the *P. radiata* superubiquitin promoter with intron::UDP Glucose Binding domain::3'UTR: nos terminator from pARB476 was removed and cloned into the NotI site of pART29 to produce the vector pARB483. The binary vector pART29 is a modified pART27 vector (Gleave, *Plant Mol. Biol.* 20:1203-1207, 1992) that contains the Arabidopsis thaliana ubiquitin 3 (UBQ3) promoter instead of the nos5' promoter and no lacZ sequences.

[0382] SEQ ID 773

CGATGGGTGTTATTTGTGGATAATAAATTCGGGTGATGTTCAGTGTTTGTCGTATTTCTCACGAATAAA
TTGTGTTTATGTATGTGTTAGTGTTGTTTGTCTGTTTCAGACCCTCTTATGTTATATTTTTCTTTTCGT
CGGTCAGTTGAAGCCAATACTGGTGTCCTGGCCGGCACTGCAATACCATTTCGTTTAATATAAAGACTC
TGTTATCCGTGAGCTCGAATTTCCCCGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATC
CTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACA
TGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACG
CGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAG
ATCGCGGCCGCATTTAAATGGTACCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTGGCCG
TCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCC
CTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGA
ATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGA
CCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCG
CCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACC
TCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTC
GCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACC
CTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGC
TGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCG
GGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAG
ACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGT
CGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGT
AAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGAT
CCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGC
GGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTT
GGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGC
TGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCT
AACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGA
AGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATT
AACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGC
AGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCG
TGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAC
GACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAA
GCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATT
TAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTT
CCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAAT
CTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAAC
TCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTA
GTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGT
GGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGC
GCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACT
GAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCC
GGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTA
TAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAG
CCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACAT

GTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGC
TCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAA
ACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGC
GGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTAT
GCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCA
TGATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGGGCCGCTCTAGAACTAGTG
GATCCCCCGGGCTGCAGGAATTCGTCCAGCAGTTGTCTGGAGCTCCACCAGAAATCTGGAAGCTTAT

[0383] SEQ ID 774

AAATCTAGAGGTACCATTTAAATGCGGCCGCAAAACCCCTCACAAATACATAA

[0384] SEQ ID 775

TTTCTGCAGCTTGAAATTGAAATATGACTAACGAAT

[0385] SEQ ID 776

tctagaggtaccatttaaatgcggccgcaaaacccctcacaaatacataaaaaaaattctttatttaat
tatcaaactctccactacctttcccaccaaccgttacaatcctgaatgttggaaaaaactaactacatt
gatataaaaaaactacattacttcctaaatcatatcaaaattgtataaatatatccactcaaaggagtc
tagaagatccacttggacaaattgcccatagttggaaagatgttcaccaagtcaacaagatttatcaat
ggaaaaatccatctaccaaacttactttcaagaaaatccaaggattatagagtaaaaaatctatgtatt
attaagtcaaaaagaaaaccaaagtgaacaaatattgatgtacaagtttgagaggataagacattggaa
tcgtctaaccaggaggcggaggaattccctagacagttaaaagtggccggaatcccggtaaaaaagatt
aaaatttttttgtagagggagtgcttgaatcatgtttttttatgatggaaatagattcagcaccatcaaa
aacattcaggacacctaaaattttgaagtttaacaaaaataacttggatctacaaaaatccgtatcgga
ttttctctaaatataactagaattttcataactttcaaagcaactcctcccctaaccgtaaaacttttc
ctacttcaccgttaattacattccttaagagtagataaagaaataaagtaaataaaagtattcacaaac
caacaatttatttcttttatttacttaaaaaaacaaaaagtttatttatttttacttaaatggcataatg
acatatcggagatccctcgaacgagaatcttttatctccctggttttgtattaaaaagtaatttattgt
ggggtccacgcggagttggaatcctacagacgcgctttacatacgtctcgagaagcgtgacggatgtgc
gaccggatgaccctgtataacccaccgacacagccagcgcacagtatacacgtgtcatttctctattgg
aaaatgtcgttgttatccccgctggtacgcaaccaccgatggtgacaggtcgtctgttgtcgtgtcgcg
tagcgggagaagggtctcatccaacgctattaaatactcgccttcaccgcgttacttctcatcttttct
cttgcgttgtataatcagtgcgatattctcagagagcttttcattcaaaggtatggagttttgaagggc
tttactcttaacatttgtttttctttgtaaattgttaatggtggtttctgtgggggaagaatcttttgc
caggtccttttgggtttcgcatgtttatttgggttatttttctcgactatggctgacattactagggct
ttcgtgctttcatctgtgttttcttcccttaataggtctgtctctctggaatatttaatttcgtatgt
aagttatgagtagtcgctgtttgtaataggctcttgtctgtaaaggtttcagcaggtgtttgcgtttta
ttgcgtcatgtgtttcagaaggcctttgcagattattgcgttgtactttaatattttgtctccaacctt
gttatagtttccctccttgatctcacaggaacccttttcttctttgagcattttcttgtggcgttctgt
agtaatattttaattttgggcccgggttctgagggtaggtgattattcacagtgatgtgctttccctat
aaggtcctctatgtgtaagctgttagggtttgtgcgttactattgacatgtcacatgtcacatattttc
ttcctcttatccttcgaactgatggttctttttctaattcgtggattgctggtgccatattttatttct
attgcaactgtattttagggtgtctctttcttttttgatttcttgttaatatttgtgttcaggttgtaac
tatgggttgctagggtgtctgccctcttctttttgtgcttctttcgcagaatctgtccgttggtctgtat
ttgggtgatgaattatttattccttgaagtatctgtctaattagcttgtgatgatgtgcaggtatattc
gttagtcatatttcaatttcaagcgatcccccgggctgcaggaattcgtccagcagttgtctggagctc
caccagaaatctggaagcttatcgatgggtgttatttgtggataataaattcgggtgatgttcagtgtt
tgtcgtatttctcacgaataaattgtgtttatgtatgtgttagtgttgtttgtctgtttcagaccctct
tatgttatatttttctttttcgtcggtcagttgaagccaatactggtgtcctggccggcactgcaataccc
atttcgtttaatataaagactctgttatccgtgagctcgaatttccccgatcgttcaaacatttggcaa
taaagtttcttaagattgaatcctgttgccggtcttgcgatgattatcatataatttctgttgaattac
gttaagcatgtaataattaacatgtaatgcatgacgttatttatgagatgggtttttatgattagagtc
ccgcaattatacatttaatacgcgatagaaaacaaaatatagcgcgcaaactaggataaattatcgcgc
gcggtgtcatctatgttactagatcgcggccgcatttaaatggtacccaattcgccctatagtgagtcg
tattacgcgcgctcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaactt

aatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgccct
tcccaacagttgcgcagcctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggt
gtggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttc
ccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctccctttagggttc
cgatttagtgctttacggcacctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggcca
tcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttc
caaactggaacaacactcaaccctatctcggtctattcttttgatttataagggattttgccgatttcg
gcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgctt
acaatttaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacatt
caaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagta
tgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctc
acccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaac
tggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcactt
ttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgca
tacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatga
cagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaa
cgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatc
gttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatgg
caacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagact
ggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctg
ataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccct
cccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctg
agataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattg
atttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaa
tcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgag
atccttttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtt
tgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaata
ctgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcg
ctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaa
gacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttgg
agcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaag
ggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccag
ggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgt
gatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggcct
tttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccg
cctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaag
cggaagagcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacg
acaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattagg
caccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttc
acacaggaaacagctatgaccatgattacgccaagcgcgcaattaaccctcactaaagggaacaaaagc
tggggccgctctag

[0386] SEQ ID 777

TCTAGAGGTACCATTTAAATGCGGCCGCAAAACCCCTCACAAATACATAAAAAAAATTCTTTATTTAAT
TATCAAACTCTCCACTACCTTTCCCACCAACCGTTACAATCCTGAATGTTGGAAAAAACTAACTACATT
GATATAAAAAAACTACATTACTTCCTAAATCATATCAAAATTGTATAAATATATCCACTCAAAGGAGTC
TAGAAGATCCACTTGGACAAATTGCCCATAGTTGGAAAGATGTTCACCAAGTCAACAAGATTTATCAAT
GGAAAAATCCATCTACCAAACTTACTTTCAAGAAAATCCAAGGATTATAGAGTAAAAAATCTATGTATT
ATTAAGTCAAAAAGAAAACCAAAGTGAACAAATATTGATGTACAAGTTTGAGAGGATAAGACATTGGAA
TCGTCTAACCAGGAGGCGGAGGAATTCCCTAGACAGTTAAAAGTGGCCGGAATCCCGGTAAAAAAGATT
AAAATTTTTTTGTAGAGGGAGTGCTTGAATCATGTTTTTTATGATGGAAATAGATTCAGCACCATCAAA
AACATTCAGGACACCTAAAATTTTGAAGTTTAACAAAAATAACTTGGATCTACAAAAATCCGTATCGGA
TTTTCTCTAAATATAACTAGAATTTTCATAACTTTCAAAGCAACTCCTCCCCTAACCGTAAAACTTTTC
CTACTTCACCGTTAATTACATTCCTTAAGAGTAGATAAAGAAATAAAGTAAATAAAAGTATTCACAAAC
CAACAATTTATTTCTTTTATTTACTTAAAAAAACAAAAAGTTTATTTATTTTACTTAAATGGCATAATG
ACATATCGGAGATCCCTCGAACGAGAATCTTTTATCTCCCTGGTTTTGTATTAAAAAGTAATTTATTGT

GGGGTCCACGCGGAGTTGGAATCCTACAGACGCGCTTTACATACGTCTCGAGAAGCGTGACGGATGTGC
GACCGGATGACCCTGTATAACCCACCGACACAGCCAGCGCACAGTATACACGTGTCATTTCTCTATTGG
AAAATGTCGTTGTTATCCCCGCTGGTACGCAACCACCGATGGTGACAGGTCGTCTGTTGTCGTGTCGCG
TAGCGGGAGAAGGGTCTCATCCAACGCTATTAAATACTCGCCTTCACCGCGTTACTTCTCATCTTTTCT
CTTGCGTTGTATAATCAGTGCGATATTCTCAGAGAGCTTTTCATTCAAAGGTATGGAGTTTTGAAGGGC
TTTACTCTTAACATTTGTTTTTCTTTGTAAATTGTTAATGGTGGTTTCTGTGGGGGAAGAATCTTTTGC
CAGGTCCTTTTGGGTTTCGCATGTTTATTTGGGTTATTTTTCTCGACTATGGCTGACATTACTAGGGCT
TTCGTGCTTTCATCTGTGTTTTCTTCCCTTAATAGGTCTGTCTCTCTGGAATATTTAATTTTCGTATGT
AAGTTATGAGTAGTCGCTGTTTGTAATAGGCTCTTGTCTGTAAAGGTTTCAGCAGGTGTTTGCGTTTTA
TTGCGTCATGTGTTTCAGAAGGCCTTTGCAGATTATTGCGTTGTACTTTAATATTTTGTCTCCAACCTT
GTTATAGTTTCCCTCCTTTGATCTCACAGGAACCCTTTCTTCTTTGAGCATTTTCTTGTGGCGTTCTGT
AGTAATATTTTAATTTTGGGCCCGGGTTCTGAGGGTAGGTGATTATTCACAGTGATGTGCTTTCCCTAT
AAGGTCCTCTATGTGTAAGCTGTTAGGGTTTGTGCGTTACTATTGACATGTCACATGTCACATATTTTC
TTCCTCTTATCCTTCGAACTGATGGTTCTTTTTCTAATTCGTGGATTGCTGGTGCCATATTTTATTTCT
ATTGCAACTGTATTTTAGGGTGTCTCTTTCTTTTTGATTTCTTGTTAATATTTGTGTTCAGGTTGTAAC
TATGGGTTGCTAGGGTGTCTGCCCTCTTCTTTTGTGCTTCTTTCGCAGAATCTGTCCGTTGGTCTGTAT
TTGGGTGATGAATTATTTATTCCTTGAAGTATCTGTCTAATTAGCTTGTGATGATGTGCAGGTATATTC
GTTAGTCATATTTCAATTTCAAGCGATCCCCCGGGCTGCAGGAATTCGTCCAGCAGTTGTCTGGAGCTC
CACCAGAAATCTGGAAGCTTATCGATATGGATCAGTTCCCCAAGTGGAATCCTGTCAATAGAGAAACGT
ATATCGAAAGGCTGTCGGCAAGGTATGAAAGAGAGGGTGAGCCTTCTCAGCTTGCTGGTGTGGATTTTT
TCGTGAGTACTGTTGATCCGCTGAAGGAACCGCCATTGATCACTGCCAATACAGTCCTTTCCATCCTTG
CTGTGGACTATCCCGTCGATAAAGTCTCCTGCTACGTGTCTGATGATGGTGCAGCTATGCTTTCATTTG
AATCTCTTGTAGAAACAGCTGAGTTTGCAAGGAAGTGGGTTCCGTTCTGCAAAAAATTCTCAATTGAAC
CAAGAGCACCGGAGTTTTACTTCTCACAGAAAATTGATTACTTGAAAGACAAGGTTCAACCTTCTTTCG
TGAAAGAACGTAGAGCAATGAAAAGGGATTATGAAGAGTACAAAGTCCGAGTTAATGCCCTGGTAGCAA
AGGCTCAGAAAACACCTGAAGAAGGATGGACTATGCAAGATGGAACACCTTGGCCTGGGAATAACACAC
GTGATCACCCTGGCATGATTCAGGTCTTCCTTGGAAATACTGGAGCTCGTGACATTGAAGGAAATGAAC
TACCTCGTCTAGTATATGTCTCCAGGGAGAAGAGACCTGGCTACCAGCACCACAAAAAGGCTGGTGCAG
AAAATGCTCTGGTGAGAGTGTCTGCAGTACTCACAAATGCTCCCTACATCCTCAATGTTGATTGTGATC
ACTATGTAAACAATAGCAAGGCTGTTCGAGAGGCAATGTGCATCCTGATGGACCCACAAGTAGGTCGAG
ATGTATGCTATGTGCAGTTCCCTCAGAGGTTTGATGGCATAGATAAGAGTGATCGCTACGCCAATCGTA
ACGTAGTTTTCTTTGATGTTAACATGAAAGGGTTGGATGGCATTCAAGGACCAGTATACGTAGGAACTG
GTTGTGTTTTCAACAGGCAAGCACTTTACGGCTACGGGCCTCCTTCTATGCCCAGCTTACGCAAGAGAA
AGGATTCTTCATCCTGCTTCTCATGTTGCTGCCCCTCAAAGAAGAAGCCTGCTCAAGATCCAGCTGAGG
TATACAGAGATGCAAAAAGAGAGGATCTCAATGCTGCCATATTTAATCTTACAGAGATTGATAATTATG
ACGAGCATGAAAGGTCAATGCTGATCTCCCAGTTGAGCTTTGAGAAAACTTTTGGCTTATCTTCTGTCT
TCATTGAGTCTACACTAATGGAGAATGGAGGAGTACCCGAGTCTGCCAACTCACCAACACTCATCAAGG
AAGCAATTCATGTCATCGGCTGTGGCTATGAAGAGAAGACTGAATGGGGAAAAGAGATTGGTTGGATAT
ATGGGTCAGTCACTGAGGATATCTTAAGTGGCTTCAAGATGCACTGCCGAGGATGGAGATCAATTTACT
GCATGCCCGTAAGGCCTGCATTCAAAGGATCTGCACCCATCAACCTGTCTGATAGATTGCACCAGGTCC
TCCGATGGGCTCTTGGTTCTGTGGAAATTTTCTTTAGCAGACACTGTCCCCTCTGGTACGGGTTTGGAG
GAGGCCGTCTTAAATGGCTCCAAAGGCTTGCGTATATAAACACCATTGTGTACCCATGAATCGATGGGT
GTTATTTGTGGATAATAAATTCGGGTGATGTTCAGTGTTTGTCGTATTTCTCACGAATAAATTGTGTTT
ATGTATGTGTTAGTGTTGTTTGTCTGTTTCAGACCCTCTTATGTTATATTTTTCTTTTCGTCGGTCAGT
TGAAGCCAATACTGGTGTCCTGGCCGGCACTGCAATACCATTTCGTTTAATATAAAGACTCTGTTATCC
GTGAGCTCGAATTTCCCCGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCC
GGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGC
ATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAA
AACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCGCGGC
CGCATTTAAATGGTACCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTTTTA
CAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCC
AGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAA
TGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTT
CCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCC
AAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTG
ACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCG

GTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAA
CAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAATG
TGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAAC
CCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTA
TTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATG
CTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGA
GTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTAT
CCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGT
ACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAA
CCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTT
TTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATAC
CAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCG
AACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCAC
TTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTC
GCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGA
GTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGT
AACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGA
TCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAG
CGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCT
TGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTC
CGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCC
ACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTG
CCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGT
CGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACC
TACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG
GCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTG
TCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGA
AAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTC
CTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCA
GCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTC
TCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTG
AGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGG
CTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACG
CCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGGGCCGCTCTAG

[0387] Table 8. pGrowth Information.

CW AR	Plasmid(s)	Promoter	Gene	Genesis ID
88	pGrowth14	SUBIN	Cyclin A	prga001823
88	pGrowth15	SUBIN	Cyclin A	prpe001264
88	pGrowth16	SUBIN	Cyclin D	prxa004540
88	pGrowth18	SUBIN	Cyclin D	prxl006271
88	pGrowth19	SUBIN	Cyclin D	prpb019661
88	pGrowth20	SUBIN	WEE1-like protein	prrd041233

[0388] To make the growth100 plasmids, an acceptor vector (pWVK202) was built by first inserting the NotI-SUBIN::UDPGBD::nos term-NotI cassette from pARB483a into plasmid pWVK147 at NotI. Next, the UDPGBD gene was removed using restriction sites PstI and ClaI. A polylinker containing the restriction sites PstI,

2004311384 31 May 2011

NheI, AvrII, ScaI, and ClaI was inserted in place of the UDPGBD gene. Sites AvrII and NheI are both compatible with SpeI, a site found often in the plasmids provided by Genesis. ScaI is blunt, so any fragment can be blunted and then inserted at that position into the acceptor vector. Plasmids were received from Genesis and analyzed to determine which restriction sites would be most suitable for subcloning into the acceptor vector pWVK202. After the ligations were performed, the resulting products were checked by extensive restriction digest analysis to make sure that the desired plasmid had been created.

[0389] Table 9. *Eucalyptus grandis* Cell Cycle Genes and Proteins.

DNA SEQ ID NO	Protein SEQ ID NO	Sequence Identifier	Patent ORF start	Patent ORF stop
1	261	eucalyptusSpp_003910	387	1820
2	262	eucalyptusSpp_019213	99	1007
3	263	eucalyptusSpp_036800	120	1004
4	264	eucalyptusSpp_040260	23	937
5	265	eucalyptusSpp_041965	149	1033
6	266	eucalyptusSpp_002906	199	1116
7	267	eucalyptusSpp_001518	41	982
8	268	eucalyptusSpp_008078	291	2042
9	269	eucalyptusSpp_009826	107	2236
10	270	eucalyptusSpp_010364	82	1749
11	271	eucalyptusSpp_011523	151	1560
12	272	eucalyptusSpp_024358	82	1644
13	273	eucalyptusSpp_039125	626	2782
14	274	eucalyptusSpp_005362	13	1467
15	275	eucalyptusSpp_044857	113	1558
16	276	eucalyptusSpp_001743	187	1686
17	277	eucalyptusSpp_012405	238	1653
18	278	eucalyptusSpp_003739	235	1539
19	279	eucalyptusSpp_022338	158	1618
20	280	eucalyptusSpp_028605	205	1530
21	281	eucalyptusSpp_041006	174	1499
22	282	eucalyptusSpp_006643	94	1332
23	283	eucalyptusSpp_045338	176	1342
24	284	eucalyptusSpp_046486	150	1283
25	285	eucalyptusSpp_012070	101	367
26	286	eucalyptusSpp_006617	9	1352
27	287	eucalyptusSpp_007827	89	1486
28	288	eucalyptusSpp_008036	80	1477
29	289	eucalyptusSpp_001597	160	1062

DNA SEQ ID NO	Protein SEQ ID NO	Sequence Identifier	Patent ORF start	Patent ORF stop
30	290	eucalyptusSpp_001596	172	1077
31	291	eucalyptusSpp_005870	66	989
32	292	eucalyptusSpp_006901	111	1541
33	293	eucalyptusSpp_006902	116	1615
34	294	eucalyptusSpp_007440	155	1453
35	295	eucalyptusSpp_008994	228	2033
36	296	eucalyptusSpp_024580	110	1258
37	297	eucalyptusSpp_037831	50	1462
38	298	eucalyptusSpp_034958	176	739
39	299	eucalyptusSpp_022967	150	1529
40	300	eucalyptusSpp_008599	247	1971
41	301	eucalyptusSpp_009919	136	1644
42	302	eucalyptusSpp_015820	48	836
43	303	eucalyptusSpp_008327	49	822
44	304	eucalyptusSpp_004604	185	751
45	305	eucalyptusSpp_000966	103	621
46	306	eucalyptusSpp_001037	41	559
47	307	eucalyptusSpp_004603	127	693
48	308	eucalyptusSpp_005465	28	639
49	309	eucalyptusSpp_006571	135	812
50	310	eucalyptusSpp_006786	119	613
51	311	eucalyptusSpp_007057	38	562
52	312	eucalyptusSpp_008670	109	1872
53	313	eucalyptusSpp_009137	74	1159
54	314	eucalyptusSpp_010285	54	2045
55	315	eucalyptusSpp_010600	53	1879
56	316	eucalyptusSpp_011551	7	690
57	317	eucalyptusSpp_020743	83	601
58	318	eucalyptusSpp_023739	125	535
59	319	eucalyptusSpp_024103	55	573
60	320	eucalyptusSpp_031985	147	842
61	321	eucalyptusSpp_032025	167	487
62	322	eucalyptusSpp_032173	195	890
63	323	eucalyptusSpp_033340	68	586
64	324	eucalyptusSpp_009143	182	3265
65	325	eucalyptusSpp_000349	165	1145
66	326	eucalyptusSpp_000575	529	1569
67	327	eucalyptusSpp_000804	156	1136
68	328	eucalyptusSpp_000805	90	1073
69	329	eucalyptusSpp_000806	66	1049
70	330	eucalyptusSpp_002248	277	1512
71	331	eucalyptusSpp_003203	33	1076
72	332	eucalyptusSpp_003209	65	973
73	333	eucalyptusSpp_004429	82	1047

2004311384 31 May 2011

DNA SEQ ID NO	Protein SEQ ID NO	Sequence Identifier	Patent ORF start	Patent ORF stop
74	334	eucalyptusSpp_004607	43	1101
75	335	eucalyptusSpp_004682	142	1095
76	336	eucalyptusSpp_005786	61	1257
77	337	eucalyptusSpp_005887	193	1527
78	338	eucalyptusSpp_005981	109	1155
79	339	eucalyptusSpp_006766	71	1213
80	340	eucalyptusSpp_006769	109	1785
81	341	eucalyptusSpp_006907	364	2685
82	342	eucalyptusSpp_007518	96	1412
83	343	eucalyptusSpp_007717	116	1702
84	344	eucalyptusSpp_007718	46	1101
85	345	eucalyptusSpp_007741	23	1258
86	346	eucalyptusSpp_007884	404	2644
87	347	eucalyptusSpp_008258	107	2383
88	348	eucalyptusSpp_008465	243	1625
89	349	eucalyptusSpp_008616	126	1127
90	350	eucalyptusSpp_008690	257	1390
91	351	eucalyptusSpp_008708	178	1632
92	352	eucalyptusSpp_008850	290	2917
93	353	eucalyptusSpp_009072	148	1197
94	354	eucalyptusSpp_009465	140	1567
95	355	eucalyptusSpp_009472	376	1737
96	356	eucalyptusSpp_009550	69	1010
97	357	eucalyptusSpp_010284	149	1423
98	358	eucalyptusSpp_010595	365	2677
99	359	eucalyptusSpp_010657	24	923
100	360	eucalyptusSpp_012636	221	3598
101	361	eucalyptusSpp_012748	44	1447
102	362	eucalyptusSpp_012879	196	1314
103	363	eucalyptusSpp_015515	193	1668
104	364	eucalyptusSpp_015724	78	1634
105	365	eucalyptusSpp_016167	85	2826
106	366	eucalyptusSpp_016633	74	1246
107	367	eucalyptusSpp_017485	100	4377
108	368	eucalyptusSpp_018007	58	2439
109	369	eucalyptusSpp_020775	159	1064
110	370	eucalyptusSpp_023132	118	1665
111	371	eucalyptusSpp_023569	57	1628
112	372	eucalyptusSpp_023611	250	1566
113	373	eucalyptusSpp_024934	106	1434
114	374	eucalyptusSpp_025546	190	1917
115	375	eucalyptusSpp_030134	102	2942
116	376	eucalyptusSpp_031787	75	1079
117	377	eucalyptusSpp_034435	99	1148

DNA SEQ ID NO	Protein SEQ ID NO	Sequence Identifier	Patent ORF start	Patent ORF stop
118	378	eucalyptusSpp_034452	232	1806
119	379	eucalyptusSpp_035789	72	1124
120	380	eucalyptusSpp_035804	315	2069
121	381	eucalyptusSpp_043057	145	1968
122	382	eucalyptusSpp_046741	130	1488
123	383	eucalyptusSpp_047161	269	1693
235	495	eucalyptusSpp_006366	117	1580
236	496	eucalyptusSpp_017378	111	1700
252	512	eucalyptusSpp_045414	238	1648
253	513	eucalyptusSpp_044328	59	859
254	514	eucalyptusSpp_015615	44	1829
255	515	eucalyptusSpp_017239	109	1866
256	516	eucalyptusSpp_018643	212	1815
257	517	eucalyptusSpp_019127	207	1193
258	518	eucalyptusSpp_022624	6	2786
259	519	eucalyptusSpp_032424	213	1726
260	520	eucalyptusSpp_037472	101	2110

2004311384 31 May 2011

[0390] Table 10. *Pinus radiata* cell cycle genes and proteins.

DNA SEQ ID NO	Protein SEQ ID NO	Sequence Identifier	Patent ORF start	Patent ORF stop
124	384	pinusRadiata_001766	1163	2545
125	385	pinusRadiata_002927	152	1582
126	386	pinusRadiata_007642	389	1297
127	387	pinusRadiata_013714	38	946
128	388	pinusRadiata_016332	180	1088
129	389	pinusRadiata_021677	40	948
130	390	pinusRadiata_027562	229	1134
131	391	pinusRadiata_001504	105	2642
132	392	pinusRadiata_015211	187	2580
133	393	pinusRadiata_020421	220	1749
134	394	pinusRadiata_003187	438	1748
135	395	pinusRadiata_015661	240	1631
136	396	pinusRadiata_013874	252	1604
137	397	pinusRadiata_014615	261	1817
138	398	pinusRadiata_004578	167	1576
139	399	pinusRadiata_023387	183	1598
140	400	pinusRadiata_006970	98	1126
141	401	pinusRadiata_010322	148	894
142	402	pinusRadiata_022721	287	1363
143	403	pinusRadiata_023407	251	1348
144	404	pinusRadiata_001945	229	510
145	405	pinusRadiata_008233	92	409
146	406	pinusRadiata_008234	64	381
147	407	pinusRadiata_022054	68	349
148	408	pinusRadiata_012137	125	1849
149	409	pinusRadiata_012582	70	1602
150	410	pinusRadiata_015285	140	1465
151	411	pinusRadiata_017229	628	2565
152	412	pinusRadiata_020724	55	1818
153	413	pinusRadiata_004555	259	1710
154	414	pinusRadiata_004556	356	1807
155	415	pinusRadiata_005729	261	1298
156	416	pinusRadiata_007395	365	2251
157	417	pinusRadiata_009503	156	1454
158	418	pinusRadiata_011283	203	1348
159	419	pinusRadiata_012322	229	1644
160	420	pinusRadiata_018671	156	1454
161	421	pinusRadiata_023236	27	2222
162	422	pinusRadiata_000171	71	1759
163	423	pinusRadiata_000172	358	2040
164	424	pinusRadiata_001480	238	756

DNA SEQ ID NO	Protein SEQ ID NO	Sequence Identifier	Patent ORF start	Patent ORF stop
165	425	pinusRadiata_001481	285	803
166	426	pinusRadiata_001483	190	708
167	427	pinusRadiata_001484	156	674
168	428	pinusRadiata_001692	176	1912
169	429	pinusRadiata_005313	64	765
170	430	pinusRadiata_006362	93	881
171	431	pinusRadiata_006493	372	1070
172	432	pinusRadiata_006983	28	594
173	433	pinusRadiata_006984	34	648
174	434	pinusRadiata_007665	481	1611
175	435	pinusRadiata_012196	93	584
176	436	pinusRadiata_013382	250	1869
177	437	pinusRadiata_016461	84	422
178	438	pinusRadiata_017611	128	1213
179	439	pinusRadiata_019776	265	837
180	440	pinusRadiata_020659	38	781
181	441	pinusRadiata_022559	38	526
182	442	pinusRadiata_024188	37	1158
183	4443	pinusRadiata_027973	61	768
184	444	pinusRadiata_001353	421	2172
185	445	pinusRadiata_001978	163	1647
186	446	pinusRadiata_002810	192	1172
187	447	pinusRadiata_002811	131	1111
188	448	pinusRadiata_002812	149	1726
189	449	pinusRadiata_003514	948	2228
190	450	pinusRadiata_004104	332	1465
191	451	pinusRadiata_005595	232	1590
192	452	pinusRadiata_005754	207	1550
193	453	pinusRadiata_006463	221	1171
194	454	pinusRadiata_006665	221	3679
195	455	pinusRadiata_006750	269	1252
196	456	pinusRadiata_007030	214	1242
197	457	pinusRadiata_007854	119	2065
198	458	pinusRadiata_007917	186	1550
199	459	pinusRadiata_007989	244	3671
200	460	pinusRadiata_008506	163	1431
201	461	pinusRadiata_008692	155	1081
202	462	pinusRadiata_008693	537	1463
203	463	pinusRadiata_009170	284	1909
204	464	pinusRadiata_009408	610	1659
205	465	pinusRadiata_009522	241	1452
206	466	pinusRadiata_009734	223	1173
207	467	pinusRadiata_009815	251	1777
208	468	pinusRadiata_010670	367	1419

DNA SEQ ID NO	Protein SEQ ID NO	Sequence Identifier	Patent ORF start	Patent ORF stop
209	469	pinusRadiata_011297	284	1303
210	470	pinusRadiata_013098	684	1784
211	471	pinusRadiata_013172	336	2738
212	472	pinusRadiata_013589	81	1622
213	473	pinusRadiata_013608	399	1460
214	474	pinusRadiata_014299	207	1673
215	475	pinusRadiata_014498	263	1309
216	476	pinusRadiata_014548	232	2529
217	477	pinusRadiata_014610	56	2950
218	478	pinusRadiata_016090	193	2577
219	479	pinusRadiata_016722	187	1233
220	480	pinusRadiata_016785	51	1436
221	481	pinusRadiata_017094	525	2351
222	482	pinusRadiata_017527	152	1099
223	483	pinusRadiata_017591	470	4114
224	484	pinusRadiata_017769	196	2007
225	485	pinusRadiata_018047	214	1323
226	486	pinusRadiata_018414	68	2146
227	487	pinusRadiata_018986	874	3705
228	488	pinusRadiata_019479	360	1754
229	489	pinusRadiata_020144	185	1384
230	490	pinusRadiata_022480	241	1533
231	491	pinusRadiata_023079	230	1435
232	492	pinusRadiata_026739	101	2857
233	493	pinusRadiata_026951	43	1548
234	494	pinusRadiata_026951	206	1657
237	497	pinusRadiata_000888	144	1412
238	498	pinusRadiata_014166	793	1683
239	499	pinusRadiata_003189	415	2196
240	500	pinusRadiata_009356	109	1653
241	501	pinusRadiata_000065	343	1023
242	502	pinusRadiata_014197	417	2351
243	503	pinusRadiata_009081	69	641
244	504	pinusRadiata_013417	172	1623
245	505	pinusRadiata_005755	231	1768
246	506	pinusRadiata_006670	376	2943
247	507	pinusRadiata_007027	107	1498
248	508	pinusRadiata_007276	118	1425
249	509	pinusRadiata_007390	186	797
250	510	pinusRadiata_012648	387	2456
251	511	pinusRadiata_013171	359	2761

[0391] Table 11. Annotated Peptide Sequences of the Present Invention.

Entry	Sequence Description	Annotated Peptide Sequence
1	The amino acid sequence of SEQ ID 261. The conserved eukaryotic protein kinase domain is underlined and the serine/threonine protein kinases active-site signature is in bold.	MGDGSLGSGGRGNSGGGGGGGSRPEWLQQYDLIGKIGEG TYGLVFLARIKHPSTNRGKYIAIKKFKQSKDGDGVSPTA IREIMLLREISHENVVKLVNVHINPVDMSLYLAFDYADH DLYEIIRHHRDKVNQAINPYTVKSLLWQLLNGLNYLHSN WIIHRDLKPSNILVMGEGEEQGVVKIADFGLARVYQAPL KPLSDNGVVVTIWYRAPELLLGAKHYTSAVDMWAVGCIF AELLTLKPLFQGQEVKANPNPFQLDQLDKIFKVLGHPTQ EKWPMLVNLPHWQSDVQHIQRHKYDDNALGNVVRLSSKN ATFDLLSKMLEYDPQKRITAAQALEHEYFRMEPLPGRNA LVPSSPGDKVNYPTRPVDTTTDIEGTTSLQPSQSASSGN AVPGNMPGPHVVTNRPMPRPMHMVGMQRVPASGMAGYNL NPSGMGGGMNPSGIPMQRGVANQAQQSRRKDPGMGMGGY PPQQKQRRF
2	The amino acid sequence of SEQ ID 262. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are in bold.	MEKYQQLAKIGEGTYGIVYKAKDKKSGELLALKKIRLEA EDEGIPSTAIREISLLKQLQHPNIVRLYDVVHTEKKLTL VFEFLDQDLKKYLDACGDNGLEPYTVKSFLYQLLQGIAF CHEHRVLHRDLKPQNLLINMEGELKLADFGLARAFGIPV RNYTHEVVTLWYRAPDVLMGSRKYSTQVDIWSVGCIFAE MVNGRPLFPGSSEQDQLLRIFKTLGTPSLKTWPGMAELP DFKDNFPKYVVQSFKKICPKKLDKTGLDLLSRMLQYDPA KRISAEQAMGHPYFKDLKLRKPKAAGPGP
3	The amino acid sequence of SEQ ID 263. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are in bold.	MDQYEKIEKIGEGTYGVVYKAIDRSTNKTIALKKIRLEQ EDEGVPSTAIREISLLKEMQHGNIVKLQDVVHSERRLYL VFEYLDLDLKKHMDSCPEFSKDTHTIKMFLYQILRGISY CHSHRVLHRDLKPQNLLLDRRTNSLKLADFGLARAFGIP VRTFTHEVVTLWYRAPEILLGSRHYSTPVDVWSVGCIFA EMVNRRPLFPGDSEIDELFKIFRIMGTPNEDSWPGVTSL PDFKSTFPKWASQDLKTVTPTVDPAGIDLLSKMLCMDPR RRITAKVALEHEYFKDVGVIP
4	The amino acid sequence of SEQ ID 264. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are in bold.	MVMKSKLDKYEKLEKLGEGTYGVVYKAQDKTTKEIYALK KIRLESEDEGIPSTAIREIALLKELQHPNVVRIHDVIHT NKKLILVFEFVDYDLKKFLHNFDKGIDPKIVKSLLYQLV RGVAHCHQQKVLHRDLKPQNLLVSQEGILKLGDFGLARA FGIPVKNYTNEVVTLWYRAPDILLGSKNYSTSVDIWSIG CIFVEMLNQKPLFPGSSEQDQLKKIFKIMGTPDATKWPG IAELPDWKPENFEKYPGEPLNKVCPKMDPDGLDLLDKML KCNPSERIAAKNAMSHPYFKDIPDNLKKLYN
5	The amino acid sequence of SEQ ID 265. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are in bold.	MDQYEKVEKIGEGTYGVVYKAIDRLTNETIALKKIRLEQ EDEGVPSTAIREISLLKEMQHGNIVRLQDVVHSENRLYL VFEYLDLDLKKHMDSSPDFAKDPRLVKIFLYQILRGIAY CHSHRVLHRDLKPQNLLIDRRTNALKLADFGLARAFGIP VRTFTHEVVTLWYRAPEILLGSRHYSTPVDVWSVGCIFA EMVNQRPLFPGDSEIDELFKIFRILGTPNEDTWPGVTAL PDFKSAFPKWPAKNLQDMVPGLNSAGIDLLSKMLCLDPS KRITARSALEHEYFKDIGFVP
6	The amino acid sequence of SEQ ID 266. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are in bold.	MEKYEKLEKVGEGTYGKVYKAKDKATGQLVALKKTRLEM DEEGVPPTALREVSLLQLLSQSLYVVRLLSVEHVDGGSK RKAAAAAAAEGGGGEAHGGAVGGGKPMLYLVFEYLDTD LKKFIDSHRKGPNPRPVPAATVQNFLYQLLKGVAHCHSH GVLHRDLKPQNLLVDKEKGILKIADLGLGRAFTVPLKSY THEVFAFLAILLWRSEGESAADFDSXFRVSPVQVVTLWY RAPEVLLGSAHYSIGVDMWSVGCIFAEMVRRQALFPGDS EFQQLLHIFRLLGTPTEKQWPGVTTLRDWHVYPQWEPQN LARAVPSLGPDGVDLLSKMLKYDPAERISAKAALDHPFF DSLDKSQF

Entry	Sequence Description	Annotated Peptide Sequence
7	The amino acid sequence of SEQ ID 267. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are in bold.	MERPATAAVSAMEAFEKLEKVGEGTYGKVYRAREKATGK IVALKKTRLHEDEEGVPPTTLREISILRMLSRDPHIVRL MDVKQGQNKEGKTVLYLVFEYMETDLKKYIRGFRSSGES IPVNIVKSLMYQLCKGVAFCHGHGVLHRDLKPHNLLMDK KTLTLKIADLGLARAFTVPIKKYTHEILTLWYRAPEVLL GATHYSTAVDMWSVGCIFAELVTKQALFPGDSELQQLLH IFRLLGTPNEKMWPGVSSLMNWHEYPQWKPQSLSTAVPN LDKDGLDLLSQMLHYEPSRRISAKAAMEHPYFDDVNKTC L
8	The amino acid sequence of SEQ ID 268. The conserved eukaryotic protein kinase domain is underlined and the serine/threonine protein kinases active-site signature is in bold.	MGCVLGREVSSGIVTESKGRDSSEVETSKRDDSVAAKVE GEGKAEEVRTEETQKKEKVEDDQQSREQRRRSKPSTKLG NLPKHIRGEQVAAGWPSWLSDICGEALNGWIPRRANTFE KIDKIGQGTYSNVYKAKDLLTGKIVALKKVRFDNLEPES VRFMAREILILRHLDHPNVVKLEGLVTSRMSCSLYLVFE YMEHDLAGLAASPAIKFTEPQVKCYMHQLLSGLEHCHNR RVLHRDIKGSNLLIDNGGVLKIGDFGLASFYDPDHKHRM TSRVVTLWYRPPELLLGANDYGVGIDLWSAGCILAELLA GKPIMPGRTEVEQLHKIYKLCGSPSEEYWKKYKLPNATL FKPREPYRRCIRETFKDFPPSSLPLIETLLAIDPAERGT ATDALQSEFFRTEPYACEPSSLPQYPPSKEMDAKKRDDE ARRLRAASKGQADGSKKERTRDRRVRAVPAPEANAELQH NIDRRRLISHANAKSKSEKFPPPHQDGALGFPLGASHRF DPAVVPPDVPFTSTSFTSSKEHDQTWSGPLVDPPGAPRR KKHSAGGQRESSKLSMGTNKGRRADSHLKAYESKSIA
9	The amino acid sequence of SEQ ID 269. The conserved eukaryotic protein kinase domain is underlined and the serine/threonine protein kinase active-site signature is in bold.	MYSKSSAVDDSRESPKDRVSSSRRLSEVKTSRLDSSRRE NGFRARDKVGDVSVMLIDKKVNGSARFCDDQIEKKSDRL QKQRRERAEAAAAADHPGAGRVPKAVEGEQVAAGWPVWL SAVAGEAIKGWLPRRADTFEKLDKIGQGTYSSVYKARDV TNNKIVALKRVRFDNLDTESVKFMAREIHILRMLDHPNV IKLEGLITSRMSCSLYLVFEYMEHDLTGLASRPDVKFSE PQIKCYMKQLLSGLDHCHKHGVLHRDIKGSNLLIDNNGI LKIADFGLASVFDPHQTAPLTSRVVTLWYRPPELLLGAS RYGVEVDLWSTGCILGELYTGKPILPGKTEVEQLHKIFK LCGSPSDDYWRRLHLPHAAVFKPPQPYRRCVAEIFKELP PVALGLLETLISVDPSQRGTAAFALRSEFFTASPLPCDP SSLPKYPPSKEIDMKLREEEARRRGAAGGKNELEKRGTK DSRTNSAYYPNAGQLQVKQCHSNANGRSEIFGPYQEKTV SGFLVAPPKQARVSKETRKDYAEQPDRASFSGPLVPGPG FSKAGKELGHSITVSRNTNLSTLSSLVTSRTGDNKQKSG PLVSESANQASRYSGPIREMEPARKQDRRSHVRTNIDYR SREDGNSSTKEPALYGRGSAGNKIYVSGPLLVSSNNVDQ MLKEHDRRIQEHARRARFDKARVGNNHPQAAVDSKLVSV HDAG
10	The amino acid sequence of SEQ ID 270. The conserved eukaryotic protein kinase domain is underlined and the serine/threonine protein kinase active-site is in bold	MGCIPTIISDGRRRSAAPDKRRPRPRRSSSEGEAPPHAT AAGSEGGESARGAPGKERPEPAPRFVVRSPQGWPPWLVA AVGHAIGEFVPRCADSFRKLAKIGEGTYSNVYKARDLVT GKTVALKKVRFDNLEAESIKFMAREILVLTRLNHPNVIK LEGPVTSRMSSGLYLAFEYMEHDLSGIAARQNGKFTEPQ VKCFMRQLLSGLEHCHNHDVLHRDIKCSNLLIDNEGNLK IADFGLATFYDPERKQVMTNRVVTLWYRAPELLLGATSY GIGIDLWSAGCILAELLYGKPIMPGRTEVEQLHKIFKLC GSPSEAYWNKFKLPNANIFKPPQPYARCIAETFKDFPPS ALPLLETLLSIDPDERGTATTALNSEFFAAEPHACEPSS LPKYPPSKEMDLKLIKEKTRRDSSKRPSAIHGSRRDGIH DRAGRVIPAPEATAENQATLHRPRAMKKANPMSRSEKFP PAHMDGVVGSSANAWLSGPASNAAPDSRRHRSLNQNPSS SVGKASTGSSTTQETLKVAPELLQVGSSSLHPCHRMLVY GSNLTIRSK

Entry	Sequence Description	Annotated Peptide Sequence
11	The amino acid sequence of SEQ ID 271. The conserved protein kinase family domain is underlined, and the serine/threonine protein kinases active-site signature is in bold	MGCICAKQADRGPASPGSGILTGAGTGTGTRSSKIPSGL FEFEKSGVKEHGGRSGELRKLEEKGSLSKRLRLELGFSH RYVEAEQAAAGWPSWLTAVAGDAIQGLVPLKADSFEKLE KIGQGTYSSVFRARELANGRMVALKKVRFDNFQPESIQF MAREISILRRLDHPNIMKLEGIITSRMSNSIYLVFEYME HDLYGLISSPQVKFSDAQVKCYMKQLLSGIEHCHQHGVI HRDVKSSNILVNNEGILRIGDFGLANILNPKDRQQLTSH VVTLWYRPPELLMGSTSYGVTVDLWSVGCVFAELMFRKP ILRGRTEVEQLHKIFKLCGSPPDGYWKMCKVPQATMFRP RHAYECTLRERCKGIATSAMKLMETFLSIEPHKRGTASS ALISEYFRTVPYACDPSSLPKYPPNKEIDAKHREEARRK KARSRVREAEVGKRPTRIHRASQEQGFSSNIAPKEKRSY A
12	The amino acid sequence of SEQ ID 272. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are in bold.	MAVAAPGHLNVNESPSWGSRSVDCFEKLEQIGEGTYGQV YMAKEKKTGEIVALKKIRMDNEREGFPITAIREIKILKK LHHENVIKLKEIVTSPGPEKDEQGRPEGNKYKGGIYMVF EYMDHDLTGLADRPGMRFSVPQIKCYMRQLLTGLHYCHI NQVLHRDIKGSNLLIDNEGNLKLADFGLARSFSNDHNAN LTNRVITLWYRPPELLLGATKYGPAVDMWSVGCIFAELL HGKPIFPGKDEPEQLNKIFELCGAPDEINWPGVSKIPWY NNFKPTRPMKRRLREVFRHFDRHALELLERMLTLDPSQR ISAKDALDAEYFWADPLPCDPKSLPKYESSHEFQTKKKR QQQRQHEETAKRQKLQHPPQHPRLPPVQQSGQAHAQMRP GPNQLMHGSQPPVATGPPGHHYGKPRGPSGGAGRYPSSG NPGGGYNHPSRGGQGGSGGYNSGPYPPQGRAPPYGSSGM PGAGPRGGGGNNYGVGPSNYPQGGGGPYGGSGAGRGSNM MGGNRNQQYGWQQ
13	The amino acid sequence of SEQ ID 273. The conserved serine/threonine protein kinase domain is underlined, and the serine/threonine protein kinase active-site signature is in bold.	MGCICTKGILPAHYRIKDGGLKLSKSSKRSVGSLRRDEL AVSANGGGNDAADRLISSPHEVENEVEDRKNVDFNEKLS KSLQRRATMDVASGGHTQAQLKVGKVGGFPLGERGAQVV AGWPSWLTAVAGEAINGWVPRRADSFEKLEKIGQGTYSS VYRARDLETNTIVALKKVRFANMDPESVRFMAREIIIMR KLDHPNVMKLEGLITSRVSGSLYLVFEYMDHDLAGLAAT PSIKLTESQIKCYMQQLLRGLEYCHSHGVLHRDIKGSNL LVDNNGNLKIGDFGLATFFRTNQKQPLTSRVVTLWYRPP ELLLGSSDYGASVDLWSSGCILAELFAGKPIMPGRTEVE QLHKIFKLCGSPSEEYWKKSKLPHATIFKPQQPYKRCLL ETFKDFPSSALGLLDVLLAVEPECRGTASSALQNEFFTS NPLPSDPSSLPKYPSSKEFDARLRDEEARKHKATAGKAR GLESIRKGSKESKVVPTSNANADLKASIQKRQEQSNPRS TGEKPGGTTQNNFILSGQSAKPSLNGSTQIGNANEVEAL IVPDRELDSPRGGAELRRQRSFMQRRASQLSRFSNSVAV GGDSHLDCSREKGANTQWRDEGFVARCSHPDGGELAGKH DWSHHLLHRPISLFKKGGEHSRRDSIASYSPKKGRIHYS GPLLPSGDNLDEMLKEHERQIQNAVRKARLDKVKTKREY ADHGQTESLLCWANGR
14	The amino acid sequence of SEQ ID 274. The conserved protein kinase family domain is underlined and the serine/threonine protein kinases active-site signature is in bold.	MDPDPSPDPDPPKSWSIHTRREIIARYEILERVGSGAYS DVYRGRRLSDGLAVALKEVHDYQSAFREIEALQILRGSP HVVLLHEYFWREDEDAVLVLEFLRSDLAAVIADASRRPR DGGGGGAAALRAGEVKRWMLQVLEGVDACHRNSIVHRDL KPGNLLISEEGVLKIADFGQARILLDDGNVAPDYEPESF EERSSEQADILQQPETMEADTTCPEGQEQGAITREAYLR EVDEFKAKNPRHEIDKETSIFDGDTSCLATCTTSDIGED PFKGSYVYGAEEAGEDAQGCLTSCVGTRWFRAPELLYGS TDYGLEVDLWSLGCIFAELLTLEPLFPGISDIDQLSRIF NVLGNLSEEVWPGCTKLPDYRTISFCKIENPIGLESCLP NCSSDEVSLVRRLLCYDPAARATPMELLQDKYFTEEPLP VPISALQVPQSKNSHDEDSAGGWYDYNDMDSDSDFEDFG PLKFTPTSTGFSIQFP

Entry	Sequence Description	Annotated Peptide Sequence
15	The amino acid sequence of SEQ ID 275. The conserved serine/threonine protein kinase domain is underlined, and the serine/threonine protein kinase active-site signature is in bold.	MDPDPSPSPDPPKSWSIHTRREIIARYEILERVGSGAYS DVYRGRRLSDGLAVALKEVHDYQSAFREIEALQILRGSP HVVLLHEYFWREDEDAVLVLEFLRSDLAAVIADASRRPR GGGVAPLRAGEGKRWMLQVLEGVDACHRNSIVHRDLKPG NLLISEEGVLKIADFGQARILLDDGNVAPDYEPESFEER SSEQADILQQPETMEADTTCPEGQEQGAITREAYLREVD EFKAKNPRHEIDKETSIYDGDTSCLATCTTSDIGEDPFK GSYVYGAEEAGEDAQGSLTSCVGTRWFRAPELLYGSTDY GLEVDLWSLGCIFAELLTLEPLFPGISDIDQLSRIFNVL GNLSEEVWPGCTKLPDYRTISFCKIENPIGLESCLPNCS SDEVSLVRRLLCYDPAARATPMELLQDKYFTEEPLPVPI SALQVPQSKNSHDEDSAGGWYDYNDMDSDSDFEDFGPLK FTPTSTGFSIQFP
16	The amino acid sequence of SEQ ID 276. The conserved cyclin and cyclin C-terminal domains are underlined and the cyclins signature is in bold.	MSNQHRRSSFSSSTTSSLAKRHASSSSSSLENAGKAFAA AAVPSHLAKKRAPLGNLTNLKAGDGNSRSSSAPSTLVAN ATKLAKTRKGSSTSSSIMGLSGSALPRYASTKPSGVLPS VNPSIPRIEIAVDPMSCSMVVSPSRSDMQSVSLDESMST CESFKSPDVEYIDNEDVSAVDSIDRKTFSNLYISDAAAK TAVNICERDVLMEMETDEKIVNVDDNYSDPQLCATIACD IYQHLRASEAKKRPSTDFMDRVQKDITASMRAILIDWLV EVAEEYRLVPDTLYLTVNYIDRYLSGNVMNRQRLQLLGV ACMMIAAKYEEICAPQVEEFCYITDNTYFKEEVLQMESS VLNYLKFEMTAPTVKCFLRRFVRAAQGVNEVPSLQLECM ANYIAELSLLEYDMLCYAPSLVAASAIFLAKFVITPSKR PWDPTLQHYTLYQPSDLGNCVKDLHRLCFNNHGSTLPAI REKYSQHKYKYVAKKYCPPSIPPEFFHNLVY
17	The amino acid sequence of SEQ ID 277. The conserved cyclin and cyclin C-terminal domains are underlined.	MNKENAVGTKSEAPTIRITRSRSKALGTSTGMLPSSRPS FKQEQKRTVRANAKRSASDENKGTMVGNASKQHKKRTVL NDVTNIFCENSYSNCLNAAKAQTSRQGRKWSMKKDRDVH QSGAVQIMQEDVQAQFVEESSKIKVAESMEITIPDKWAK RENSEHSISMKDTVAESSRKPQEFICGEKSAALVQPSIV DIDSKLEDPQACTPYALDIYNYKRSTELERRPSTIYMET LQKDVTPNMRGILVDWLVEVSEEYKLVPDTLYLTVNLID RSLSQKFIEKQRLQLLGVTCMLIASKYEEICPPRVEEFC FITDNTYTSLEVLKMESRVLNLLHFQLSVPTVKTFLRRF VQAAQVSSEVPSVELEYLANYLAELTLVEYSFLKFLPSL MAASAVLLARWTLNQSDNPWNLTLEHYTKYKASELKAAV LALEDLQLNTSGSTLNAIREKYRQQKVNYSLLIHSKANH EIL
18	The amino acid sequence of SEQ ID 278. The conserved cyclin N- and C-terminal family domains are underlined and the cyclins signature is in bold.	MAGSDENNPGVVGGAHVQEGLRVGAGKMGAGNVQQRRAL SNINSNIIGAPPYPCAVNKRVLSEKNVNSENDLLNAAHR PITRQFAAQMAYKQQLRPEENKRTTQSVSNPSKSEDCAI LDVDDDKMADDFPVPMFVQHTEAMLEEIDRMEEVEMEDV AEEPVTDIDSGDKENQLAVVEYIDDLYMFYQKAEASSCV PPNYMDRQQDINERMRGILIDWLIEVHYKFELMDETLYL TVNLIDRFLAVQPVVKKKLQLVGVTAMLLACKYEEVSVP VVEDLILISDRAYSRKEVLEMERLMVNTLHFNMSVPTPY VFMRRFLKAAQSDKKLELLSFFIIELSLVEYDMLKFPPS LLAASAIYTALSTITRTKQWSTTCEWHTSYSEEQLLECA RLMVTFHQRAGSGKLTGVHRKYSTSKFGHAARTEPANFL LDFRL
19	The amino acid sequence of SEQ ID 279. The conserved cyclin and cyclin C-terminal domains are underlined.	MASRPIVPVQARGEAAIGGGAGKAAIGGGAGKQQKKNGA AEGRNRKALGDIGNLVTVRGIEGKVQPHRPITRSFCAQL LANAQAAAAAAENNKKQAVVNVNGAPSILDVPGAGKRAEP AAAAAAAVAKAAQKKVVKPKQKAEVIDLTSDSEERSRPR RSNNIMSLRRRKERNHREGICPLSLRSSLLEARLVDWLI EIHNKFDLMPETLYLTINIIDRFLSVKAVPRRELQLLGM GALFTASKYEEIWAPEVNDLVCIADRAYSHEQVLAMEKT ILGKLEWTLTVPTHYVFLVRFIKASLGDRKLENMVYFLA ELGVMNYATLTYCPSMVAASAVYAARCTLGLTPLWNDTL KLHTGFSESQLMDCARLLVGYHAKAKENKLQVVYKKYSS SQREGVALIPPAKALLCEGGGLSSSSSLASSS

Entry	Sequence Description	Annotated Peptide Sequence
20	The amino acid sequence of SEQ ID 280. The conserved cyclin and cyclin C-terminal domains are underlined and the cyclins signature is in bold.	MGLPDENNAALSKPTNLQVGGLEIGGRKFGQEIRQTRRA LSVINQNLVGDRAYPCHVVNKRGHSKRDAVCGKDQVDPV HRPLTRKFAAQTASTQQHCIEEAKKPRTAVQERNEFGDC IFVDVEDCQPSSENQPVPMFLEIPESRLDDDMEEVEMED IVEEEEEEPIMDIDGRDKKNPLAVVDYIEDIYANYRRTE NCSCVSANYMAQQADINEKMRSILIDWLIEVHDKFDLMH ETLFLTVNLIDRFLARQSVVRKKLQLVGLVAMLLACKYE EVSVPVVGDLILISDKAYTRKEVLEMESLMLNSLQFNMS VPTPYVFMRRFLKAAESDKKLEVLSFFLIELSLVEYEMV KFPPSLLAAAAIFTAQCTLYGFKQWTKTCEWHSNYTEDQ LLECARMMVGFHQKAATGKLTGVHRKYGTSKFGYTSKCE PANFLLGEMKNP
21	The amino acid sequence of SEQ ID 281. The conserved cyclin and cyclin C-terminal domains are underlined and the cyclins signature is in bold.	MGLPDENNAALSKPTNLQVGGLEIGGRKFGQEIRQTRRA LSVINQNLVGDRAYPCHVVNKRGHSKRDAVCGKDQVDPV HRPLTRKFAAQTASTQQHCIEEAKKPRTAVQERNEFGDC IFVDVEDCQPSSENQPVPMFLEIPESRLDDDMEEVEMED IVEEEEEEPIMDIDGRDKKNPLAVVDYIEDIYANYRRTE NCSCVSANYMAQQADINEKMRSILIDWLIEVHDKFDLMH ETLFLTVNLIDRFLARQSVVRKKLQLVGLVAMLLACKYE EVSVPVVGDLILISDKAYTRKEVLEMEKLMLNSLQFNMS VPTPYVFMRRFLKAAESDKKLEVLSFFLIELSLVEYEMV KFPPSLLAAAAIFTAQCTLYGFKQWTKTCEWHSNYTEDQ LLECARMMVGFHQKAATGKLTGVHRKYGTSKFGYTSKCE AANFLLGEMKNP
22	The amino acid sequence of SEQ ID 282. The conserved cyclin N- and C-terminal family domains are underlined.	MAMVQRQGHDPSSPQEQEDGPSSFLSDDALYCEEGRFEE DDGGGGGQVDGIPLFPSQPADRQQDSPWADEDGEEKEEE EAELQSLFSKERGARPELAKDDGGAVAARREAVEWMLMV RGVYGFSALTAVLAVDYLDRFLAGFRLQRDNRPWMTQLV AVACLALAAKVEETDVPLLVELQEVGDARYVFEAKTVQR MELLVLSTLGWEMHPVTPLSFVHHVARRLGASPHHGEFT HWAFLRRCERLLVAAVSDARSLKHLPSVLAAAAMLRVIE EVEPFRSSEYKAQLLSALHMSQEMVEDCCRFILGIAETA GDAVTSSLDSFLKRKRRCGHLSPRSPSGVIDASFSCDDE SNDSWATDPPSDPDDNDDLNPLPKKSRSSSPSSSPSSVP DKVLDLPFMNRIFEGIVNGSPI
23	The amino acid sequence of SEQ ID 283. The conserved cyclin and cyclin C-terminal domains are underlined.	MEASYQPHHHGHLRQHDPSSSQQEEQVPFDALYCSEEHW GEEDEEEGLASDGLLSEERDHRLLSPRALLDQDLLWEDE ELASLFSKEEPGGMRLNLENDPSLADARREAVEWIMRVH AHYAFSALTALLAVNYWDRFTCSFALQEDKPWMTQLSAV ACLSLAAKVEETQVPLLIDFQVEDSSPVFEAKNIQRMEL LVLSSLEWKMNPVTPLSFLDYMTRRLGLTGHLCWEFLRR CENVLLSVISDCRFTCYLPSVIAASTMLHVINGLKPRLD VEDQTQLLGILAMGMDKIDACYKLIDDDHALRSQRYSHN KRKFGSVPGSPRGVMELCFSSDGSNDSWSVAASVSSSPE PHSKKSRAGEEAEDRLLRGLEGEEDDPASADIFSFPH
24	The amino acid sequence of SEQ ID 284. The conserved cyclin and cyclin C-terminal domains are underlined.	MALQEEDTRRHYPTAPPFSPDGLYCEDETFGEDLADNAC EYAGGGARDGLCEIKDPTLPPSLLGQDLFWEDGELASLV SRETGTHPCWDELISDGSVALARKDAVGWILRVHGHYGF RPLTAMLAVNYLDRFFLSRSYQRDRPWISQLVAVACLSV AAKVEETQVPILLDLQVANAKFVFESRTIQRMELLLMST LDWRMNSVTPISFFDHILRRFGLTTNLHRQFFWMCERLL LSVVADVRLASFLPSVVATAAMLYVNKEIEPCICSEFLD QLLSLLKINEDRVNECYELILELSIDHPEILNYKHKRKR GSVPSSPSGVIDTSFSCDSSNDSWGVASSVSSSLEPRFK RSRFQDQQMGLPSVNVSSMGVLNSSY
25	The amino acid sequence of SEQ ID 285. The conserved cyclin-dependent kinases regulatory subunit domain is underlined and the cyclin-dependent kinases regulatory subunits signature 1 is in bold.	MGQIQYSEKYFDDTYEYRHVVLPPDVAKLLPKNRLLSEN EWRAIGVQQSRGWVHYAIHRPEPHIMLFRRPLNYQQQQE NQAQQNMLAK

Entry	Sequence Description	Annotated Peptide Sequence
26	The amino acid sequence of SEQ ID 286. The conserved chromo domain is underlined and the MOZ/SAS-like protein domain is in bold/italics.	MGSIDPPKAEQNGTAAAAVADPGQKPGAGDAMPPPPPVK HSNGTAAEPDVATKRRRMSVLPLEVGTRVMCRWRDGKYH PVKVIERRKLNPGDPNDYEYYVHYTEFNRRLDEWVKLEQ LDLNSVETVVDEKVEDKVTGLKMTRHQKRKIDETHVEGH EELDAASLREHEEFTKVKNIATIELGRYEIETWYFSPFP PEYNDCSKLYFCEFCLNFMKRKEQLQRHMKKCD***LKHPPG DEIYRSGTLSMFEVDGKKNKVYGQNLCYLAKLFLDHKTL YYDVDLFLFYVLCECDDRGCHMVGYFSKEKHSEESYNLA CILTLPPYQRKGYGKFLIAFSYELSKKEGKVGTPERPLS DLGLLSYKGYWTRVLLDILKKHKANISIKELSDMTAIKA DDILNTLQSLDLIQYRKGQHVICAD***PKVLDRHLKAAGRG GLEVDVSKLIWTPYREQG
27	The amino acid sequence of SEQ ID 292. The conserved histone deacetylase family domain is underlined.	MDTGGNSLPSGPDGVKRKVCYFYDPEVGNYYLLQHMQVL KPVPARDRDLCRFHADDYVAFLRSITPETQQDQLRQLKR FNVGEDCPVFDGLHSFCQTYAGGSVGGAVKLNHGLCDIA INWAGGLHHAKKCEASGFCYVNDIVLGILELLKQHERVL YVDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGD IRDIGYGKGKYYSLNVPLDDGIDDESYHSLFKPIIGKVM EVFKPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVRYM RSFNVPVLLLGGGGYTIRNVARCWCYETGVALGLEVDDK MPQHEYYEYFGPDYTLHVAPSNMENKNSRQLLEEIRSKL LENLSKLQHAPSVPFQERPPDTELPEADEDQEDPDERWD PDSDMDVDEDRKPLPSRVKRELIVEPEVKDQDSQKASID HGRGLDTTQEDNASIKVSDMNSMITDEQSVKMEQDNVNK PSEQIFPK
28	The amino acid sequence of SEQ ID 293. The conserved histone deacetylase family domain is underlined.	MDTGGNSLPSGPDGVKRKVCYFYDPEVGNYYYGQGHPMK PHRIRMTHALLAHYGLLQHMQVLKPVPARDRDLCRFHAD DYVAFLRSITPETQQDQLRQLKRFNVGEDCPVFDGLHSF CQTYAGGSVGGAVKLNHGLCDIAINWAGGLHHAKKCEAS GFCYVNDIVLGILELLKQHERVLYVDIDIHHGDGVEEAF YTTDRVMTVSFHKFGDYFPGTGDIRDIGYGKGKYYSLNV PLDDGIDDESYHSLFKPIIGKVMEVFKPGAVVLQCGADS LSGDRLGCFNLSIKGHAECVRYMRSFNVPVLLLGGGGYT IRNVARCWCYETGVALGLEVDDKMPQHEYYEYFGPDYTL HVAPSNMENKNSRQLLEDIRSKLLENLSKLQHAPSVPFQ ERPPDTELPEADEDQEDPDERWDPDSDMDVDEDRKPLPS RVKRELIVEPEVKDQDSQKASIDHGRGLDTTQEDNASIK VSDMNSMITDEQSVKMEQDNVNKPSEQIFPK
29	The amino acid sequence of SEQ ID 294. The conserved histone deacetylase domain is underlined.	MRPKDRISYFYDGDVGSVYFGPNHPMKPHRLCMTHHLVL SYELHTKMEIYRPHKAYPAELAQFHSPDYVEFLHRITPD TQHLFPNDLAKYNLGEDCPVFENLFEFCQIYAGGTIDAA RRLNNQLCDIAINWAGGLHHAKKCEASGFCYINDLVLGI LELLKYHARVLYIDIDVHHGDGVEEAFYFTDRVMTVSFH KFGDMFFPGTGDVKEIGGKEGKFYAINVPLKDGIDDTSF TRLFKAIISKVVETYQPGAIVLQCGADSLAGDRLGCFNL SIDGHSECVRFVKKFNLPLLVTGGGGYTKENVARCWVVE TGVLLDTELPNEIPENEYFKYFAPDYSLKIPRGNIVLEN LNSKSYLSAIKVQVLENLRNIQHAPSVQMQEVPPDFYIP DFDEDEQNPDERMDQHTQDKQIQRDDEYYDGDNDNDHNM DD
30	The amino acid sequence of SEQ ID 295. The conserved histone deacetylase family domain is underlined and the Zinc finger RanBP2-type profile is in bold.	MTVAEDFHVNNRSKMVSQATPESRLTGGEDDNSLHNQVD ELLCQELPERQVILEFEGTRPKPYFSDHNGGENSALGVR ATEDDLNSDVEAEEKQKEMTLEDMYKNDGTLYDDDEDDS DWEPVKRQVELMR**WFCTNCTMVNVEDVFLCDICGEHRDS** GILRHGFYASPFMQDVGAPSVEAEVQESREDHARSSPPS SSTVVGFDEKMLLHSEVEMKSHPHPERADRLQAIAASLA TAGIFPGRCRSLPVREITKEELQMVHSSEHVDAVEMTSH MFSSYFTPDTYANEHSARAARIAAGLCADLASTIISGRS KNGFALVRPPGHHAGIKHAMGFCLHNNAAVAALAAQGAG AKKVLIVDWDVHHGNGTQEIFDGNKSVLYISLHRHEGGN FYPGTGAAHEVGTMGAEGYCVNIPWSRRGVGDNDYVFAF HHIVLPIASAFAPDFTIISAGFDAARGDPLGCCDVTPAG YAQMTHMLSALSGGKLLVILEGGYNLRSISSSAVAVIKV LLGDSPISEIADAVPSKAGLRTVLEVLKIQRSYWPSLES IFWELQSQWGMFLVDNRRKQIRKRRRVLVPIWWKWGRKS VLYHLLNGHLHVKTKR

Entry	Sequence Description	Annotated Peptide Sequence
31	The amino acid sequence of SEQ ID 296. The conserved histone deacetylase family domain is underlined.	MAAAPSSPPTNRVDVFWHDGMLSHDTGRGVFDTGSDPGF LDVLEKHPENPDRVRNMVSILKRGPISPFISWHTATPAL ISQLLSFHSPEYINELVEADKNGGKVLCAGTFLNPGSWD AALLAAGNTLSAMKYVLDGKGKIAYALVRPPGHHAQPSQ ADGYCFLNNAGLAVRLALDSGCKRVVVVDIDVHYGNGTA EGFYQSSDVLTISLHMNHGSWGPSHPQSGSVDELGEDEG YGYNMNIPLPNGTGDRGYEYAVTELVVPAVESFKPEMVV LVVGQDSSAFDPNGRQCLTMDGYRAIGRTIRGLADRHSG GRILIVQEGGYHVTYSAYCLHATVEGILDLPDPLLADPI AYYPEDEAFPVKVVDSIKRYLVDKVPFLKEH
32	The amino acid sequence of SEQ ID 297. The conserved histone deacetylase family domain is underlined.	MVESSGGASLPSVGQDARKRRVSYFYEPTIGDYYYGQGH PMKPHRIRMAHNLIVHYYLHRRMEISRPFPAATTDIRRF HSEDYVTFISSVTPETVSDPAFSRQLKRFNVGEDCPVFD GIFGFCQASAGGSMGAAVKLNRGDSDIALNWAGGLHHAK KSEASGFCYVNDIVLGILELLKVHKRVLYVDIDVHHGDG VEEAFYTTDRVMTVSFHKFGDFFPGSGHIKDTGAGPGKN YALNVPLNDGIDDESFRGMFRPIIQKVMEVYQPDAVVLQ CGADSLSGDRLGCFNLSVKGHADCLRFLRSFNVPLMVLG GGGYTMRNVARCWCYETAVAVGVEPENDLPYNEYYEYFG PDYTLHVEPCSMENLNAPKDLERIRNMLLEQLSRIPHAP SVPFQMTPPITQEPEEAEEDMDERPKPRIWNGEDYESDA EEDKSQHRSSNADALHDENVEMRDSVGENSGDKTREDRS PS
33	The amino acid sequence of SEQ ID 299. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase family domain is underlined.	MAAIISCHHYHSCCSSLIASKWVGARIPTSCFGRSSTQS NNAASVRQFVTRCSSSPSSRGQWQPHQNGEKGRSFSLRE CAISIALAVGLVTGVPSLDMSTGNAYAASPALPDLSVLI SGPPIKDPEALLRYALPINNKAIREVQKPLEDITDSLKV AGLRALDSVERNVRQASRVLKQGKNLIVSGLAESKKDHG VELLDKLEAGMDELQQIVEDGNRDAVAGKQRELLNYVGG VEEDMVDGFPYEVPEEYKNMPLLKGRAAVDMKVKVKDNP NLEECVFRIVLDGYNAPVTAGNFVDLVERHFYDGMEIQR ADGFVVQTGDPEGPAESFIDPSTEKPRTIPLEIMVDGEK APVYGATLEELGLYKAQTKLPFNAFGTMAMARDEFEDNS ASSQIFWLLKESELTPSNANILDGRYAVFGYVTENQDFL ADLKVGDVIESVQVVSGLDNLANPSYKIAG
34	The amino acid sequence of SEQ ID 300. The conserved FKBP-type peptidylprolyl isomerase domains are underlined. The FKBP-type peptidyl-prolyl cis-trans isomerase signature 1 is in bold and the FKBP-type peptidyl-prolyl cis-trans isomerase signature 2 is in bold/italics.	MAGEDFDIPPADEMNEDFDLPDDDDDAPVMKAGDEKEIG KQGLKKKLVKEGDAWETPDNGDEVE**VHYTGTLLDGTQFD** SSRDRGTPFKFTLGQGQ***VIKGWDQGIKTMKKGENAIFTI*** ***PPELAYGE***AGSPPTIPPNATLQFDVELLSWTSVKDICKD GGIFKKILVEGEKWENPKDLDEVLVKYEFQLEDGTTIAR SDGVEFTVKEGHFCPAVAKAVKTMKKGEKVLLTVKPQYG FGEKGKPASGDEGAVPPNATLQITLELVSWKTVSEVTDD KKVIKKILKEGEGYERPNEGAVVEVKLIGKLQDGTVFVK KGHDDCEELFKFKIDEEQ***VVDGLDKAVMNMKKGEVALLT*** ***VAPEYAFG***SSESKQDLAVVPPSSTVYYEVELVSFVKDKE SWDMNTEEKIEAAGKKKEEGNVIFKAGKYAKASKRYEKA VKYIEYDTSFSEDEKKQAKALKVACNLNDAACKLKLKDY NQAEKLCTKVLELDSRNVKALYRRAQAYIELSDLDLAEF DIKKALEIDPHNRDVKLEYKVLKEKVKEFNKKDAKFYGN MFAKMSKLEPVEKTAAKEPEPMSIDSKA
35	The amino acid sequence of SEQ ID 301. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase family domain is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	*MSTVYVLEPPTKGKVVLNTTHGPLDVELWPKEAPKAVRN* FVQLCLEGYYDNTIFHRIIKDFLVQGGDPTGSGTGGESI YGDAFSDEFHSRLRFKHRGLVACANAGSPHSNGSQFFIT LDRCDWLDRKNTIFGKITGDSIYNLSGLAEVETDKSDRP LDPPPKIISVEVLWNPFEDIVPRAPVRSLVPTVPDVQNK EPKKKAVKKLNLLSFGEEAEEEEKALVVVKQKIKSSHDV LDDPRLLKEHIPSKQVDSYDSKTARDVQSVREALSSKKQ ELQKESGAEFSNSFREIADDEDDDDDDASFDARMRRQIL QKRKELGDLPPKPKPKSRDGISARKERETSISRDKDDDD DDDQPRVEKLSLKKKGIGSEARGERMANADADLQLLNDA ERGRQLQKQKKHRLRGREDEVLTKLETFKASVFGKPLAS SAKVGDGDGDLSDWRSVKLKFAPEPGKDRMTRNEDPNDY VVVDPLLEKGKEKFNRMQAKEKRRGREWAGKSLT

Entry	Sequence Description	Annotated Peptide Sequence
36	The amino acid sequence of SEQ ID 302. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase family domain is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MASAISMHSSGLLLLQGTNGKDVTEMGKAPASSRVANMQ QRKYGATCCVARGLTSRSHYASSLAFKQFSKTPSIKYDR MVEIKAMATDLGLQAKVTNKCFFDVEIGGEPAGRIVIGL FGDDVPKTVENFRALCTGEKGFGYKGCSFHRIIKDFMIQ GGDFTRGNGTGGKSIYGSTFEDENFALKHVGPGVLSMAN AGPSTNGSQFFICTVKTPWLDNRHVVFGQVVDGMDVVQK LESQETSRSDVPRQPCRIVNCGELPLDG
37	The amino acid sequence of SEQ ID 303. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined.	MAASFTALSNVGSLSSPRNGSEIRRFRPSCNVAASVRPP PLKAGLSASSSSSFSGSLRLIPLSSSPQRKSRPCSVRAS AEAAAAQSKVTNKVYLDISIGNPVGKLVGRIVIGLYGDD VPQTAENFRALCTGEKGFGYKGSTVHRVIKDFMIQGGDF DKGNGTGGKSIYGRTFKDENFKLSHVGPGVVSMANAGPN TNGSQFFICTVKTPWLDQRHVVFGQVLEGMDIVRLIESQ ETDRGDRPRKRVVVSDCGELPVV
38	The amino acid sequence of SEQ ID 304. The conserved FKBP-type peptidyl-prolyl cis-trans isomerase signature is underlined and the FKBP-type peptidyl-prolyl cis-trans isomerase signature 2 is in bold.	MAEAIDLTGDGGVMKTIVRRAKPDAVSPSETLPLVDVRY EGVLAETGEVFDSTHEDNTLFSFEIGKGSVISAWDTALR TMKVGEVAKITCKPEYAYGSTGSPPDIPPDATLIFEVEL VACKPCKGFSVTSVTEDKARLEELKKQREIAAATKEEEK KRREEAKAAAAARVQAKLDAKKGHGKGKGKAK
39	The amino acid sequence of SEQ ID 305. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase family domain is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MGNPKVFFDMSIGGQPAGRIVMELYADVVPRTAENFRAL CTGEKGAGRSGKPLHYKGSSFHRVIPGFMCQGGDFTAGN GTGGESIYGSKFADENFVKKHTGPGVLSMANAGPGTNGS QFFVCTAKTEWLDGKHVVFGQIVDGMDVVKAIEKVGSSS GRTSKPVVVADCGQLS
40	The amino acid sequence of SEQ ID 306. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MPNPKVFFDMTIGGAAAGRVVMELYADTTPRTAENFRAL CTGEKGVGRSKKPLHYKGSKFHRVIPSFMCQGGDFTAGN GTGGESIYGVKFADENFIKKHTGPGILSMANAGPGTNGS QFFICTTKTEWLDGKHVVFGKVVEGMEVVKAIEKVGSSS GRTSKPVVVADCGQLP
41	The amino acid sequence of SEQ ID 307. The conserved FKBP-type peptidyl-prolyl cis-trans isomerase signature is underlined and the FKBP-type peptidyl-prolyl cis-trans isomerase signature 2 is in bold.	MAEAIDLTGDGGVMKTIVRRAKPDAVSPSETLPLVDVRY EGVLAETGEVFDSTHEDNTLFSFEIGKGSVISAWDTALR TMKVGEVAKITCKPEYAYGSTGSPPDIPPDATLIFEVEL VACKPCKGFSVTSVTEDKARLEELKKQREIAAATKEEEK KRREEAKAAAAARVQAKLDAKKGHGKGKGKAK
42	The amino acid sequence of SEQ ID 308. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MATARSFFLCALLLLATLYLAQAKKSEDLKEVTHKVYFD VEIAGKPAGRIVMGLYGKAVPKTAENFRALCTGEKGTGK SGKPLHYKGSSFHRIIPSFMLQGGDFTLGDGRGGESIYG EKFADENFKLKHTGPGLLSMANAGPDTNGSQFFITTVTT SWLDGRHVVFGKVLSGMDVVYKVEAEGRQSGTPKSKVVI ADSGELPL
43	The amino acid sequence of SEQ ID 309. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase family domain is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MMRREISVLLQPRFVLAFLALAVLLLVFAFPFSRQRGDQ VEEEPEITHRVYLDVDIDGQHLGRIVIGLYGEVVPRTVE NFRALCTGEKGKSANGKKLHYKGTPFHRIISGFMIQGGD VIYGDGKGYESIYGGTFADENFRIKHSHAGIISMVNSGP DSNGSQFFITTVKASWLDGEHVVFGRVIQGMDTVYAIEG GAGTYNGKPRKKVIIADSGEIPKSKWDEER

Entry	Sequence Description	Annotated Peptide Sequence
44	The amino acid sequence of SEQ ID 310. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase family domain is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MWATAEGGPPEVTLETSMGSFTVELYFKHAPRTSRNFIE LSRRGYYDNVKFHRIIKDFIVQGGDPTGTGRGGESIYGK KFEDEIKPELKHTGAGILSMANAGPNTNGSQFFITLAPC PSLDGKHTIFGRVCRGMEIIKRLGSVQTDNNDRPIHDVK ILRTSVKD
45	The amino acid sequence of SEQ ID 311. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase family domain is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MSNPKVFFDILIGKMKAGRVVMELFADVTPKTAENFRAL CTGEKGIGRSGKPLHYKGSTFHRIIPNFMCQGGDFTRGN GTGGESIYGMKFADENFKIKHTGLGVLSMANAGPDTNGS QFFICTEKTPWLDGKHVVFGKVIDGYNVVKEMESVGSDS GSTRETVAIEDCGQLSEN
46	The amino acid sequence of SEQ ID 312. The conserved FKBP-type peptidylprolyl isomerase domains are underlined. The FKBP-type peptidyl-prolyl cis-trans isomerase signature 1 is in bold and the FKBP-type peptidyl-prolyl cis-trans isomerase signature 2 is in bold/italics. The TPR repeat is in italics.	MDDDFEFPASSNVENDDDDGMDMDDMGGDVPEEEDPVAS PAVLKVGEEREIGKAGFKKKLVKEGEGWETPSSGDEVEV HYTGTLLDGTKFDSSRDRGTPFKFKLGRGQ*VIKGWDEGI KTMKKGENAIFTIPPELAYGE*SGSPPTIPPNATLQFDVE LLSWSSVKDICKDGGILKKVLVEGEKWDNPKDLDEVFVK YEASLEDGTLISKSDGVEFTVGDGYFCAALAKAVKTMKK GEKVLLTVMPQYAFGETGRPASGDEAAVPPDASLQIMLE LVSWKTVSDVTKDKKVLKKTLKEGEGYERPNDGAAVQVR LCGKLQDGTVFVKKDDEEPFEFKIDEEQ*VIDGLDRAVKN MKKGEVALVTIQPEYAFG*PTESQQDLAVVPANSTVYYEV ELLSFVKEKESWEMNNQEKIEAAARKKEEGNAAFKAGKY VRASKRYEKAVRFIEYDSSFSDEEKQQAKTLKNTCNLND AACKLKLKDFKEAEKLCTKVLEGDGKN*VKALYRRAQAYI QLVDLDLAEQDIKKALEIDPNN*RDVKLEYKILKEKVREY NKRDAQFYGNMFAKMNKLEHSRTAGMGAKHEAAPMTIDS KA
47	The amino acid sequence of SEQ ID 313. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase family domain is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold. The TPR repeat is in bold/italics.	MAKPRCFMDISIGGELEGRIVGELYTDVAPKTAENFRAL CTGEKGIGPHTGAPLHYKGVRFHRVIKGFMVQGGDISAG DGTGGESIYGLKFEDENFDLKHERKGMLSMANSGPNTNG SQFFITTTRTSHLDGKHVVFGRVVKGMGVVRSVEHVTTA AGDCPTVDVVIADCGEIPAGADDGIRNFFKDGDTYPDWP ADLDESPAELSWWMDAVDSIKAFGNGSYKKQDYKMALRK YRKALRYLDICWEKEGIDEVESSSLRKTKSQIFTNSSAC KLKLCDLKGALLDAEFAVRDGENN*AKAYFRQGQAHMELN DIDAAAESFSKALELEPND*VGIKKELNAAKKKIFERREQ EKRAYRKMFL
48	The amino acid sequence of SEQ ID 314. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MTKRKNPLVFLDVSIDGDPVERIVIELFADTVPRTAENF RSLCTGEKGVGKTTGKPLHYKGSYFHRIIKGFMAQGGDF SNGNGTGGESIYGGKFADENFKLAHDGPGLLSMANGGPN TNGSQFFIIFKRQPHLDGKHVVFGKVMRGMEVVKKIEQV GSANGKPLQPVKIVDCGETSETGTQDAVVEEKSKSATLK AKKKRSARDSSSESRGKRRQRKSRKERTRKRRRYSSSDS YSSESSDSDSESYSSDTESESKSHSESSVSDSSSSDGRR RKRKSTKREKLRRQRGKDSRGEQKSARYDKKSRHKSADS SSDSESESSSRSRSRDDKKKSSRRESARSVSKLKDAEAN SPENLESPRDREIKKVEDNSSHEEGEFSPKNDVQHNGHG TDAKFGKYDDQRPRSDGSKKSSGSMRDSPKRLANSVPQG SPSSSPAHKASEPSSSIRARNPSRSPAPDGNSKRIRKGR GFTERFSYARRYRTPSPEDVTYRPYHYGRRNFHDRRNDR YSNYRSYSERSPHRRYRSPPRGRSPPRYQRRRSRSRSVS RSPGGNKGRYRGRDQSRSRSRSRSPRRGSSPANKQLP LSERLKSRLGTRVDEHSPRRRRSSSRSHDSSRSRSPDEV PDKHEGKAAPVSPARSRSSSPSGRGLVSYGDASPDSGIN

Entry	Sequence Description	Annotated Peptide Sequence
49	The amino acid sequence of SEQ ID 315. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined. The CCHC type zinc finger is in bold and the RNA-binding region RNP-1 (RNA recognition motif) is in bold/italics.	MSVLLVTSLGDIVVDLHADRCPLTCKNFLKLCRIKYYNG CVFHTVQKDFTAQTGDPTGTGTGGDSVYKFLYGDQARFF MDEIHLDLKHSKTGTVAMASGGENLNASQFYFTLRDDLD YLDGKHTVFGEVAEGLETLTRINEAYVDEKGRPYKNIRI RHTYILDDPFDDPPQLAELIPDASPEGKPKDEVVDDVRL EDDWVPLDEQLGPAQLEEAIRAKEAHSRAVVLESIGDIP DAEIKPPDNV*LFVCKLNPVTEDEDLHTIFSRFGTVVSAD VIRDFKTGDSLCYAFIEFENKDSCEQAYFKMDNALIDDR RIKV*DFSQSVAKLWSQFKRKDSQAAKG**KGCFKCGAPDHM ARECPG**SSTRQPLSKYILKEDNAQRGGDDSRYEMVFDED APESPSHGKKRRGRDDRDDRHKMSRQSVEETKFNDREGG HSVDKHRQSERSKHREDEMSRDSKASEAGRRRIDRDFPE EERDGEKYTESHRDRDGKRGDYRDYRKGRADVQTHGDRR GDENYRRKSAAYDDGHEGAGAARRKDSNDDHHAYRRGYG DSRKGTRDEDDDGRGRRDDPSYRRSSGHKDSSNGGREEQ KYRSGETDGKSHPERSHRGDRRR
50	The amino acid sequence of SEQ ID 316. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined.	MRPFNGGSSIACLVLVIAAGALAESQGPHLGSARVVFQT NYGDIEFGFFPGVAPRTVDHIFKLVRLGCYNTNHFFRVD KGFVAQVADVANGRTAPMNDEQRTEAEKTIVGEFSNVKH VRGILSMGRYDDPDSAQSSFSILLGDAPHLDGKYAIFGR VTKGDETLKKLEQLPTRREGMFVMPTERITILSSYYYDT GAESCEEENSTLRRRLAASAVEVERQRMKCFP
51	The amino acid sequence of SEQ ID 317. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRAL CTGEKGTGRSGKPLH**FKGSSFHRVIPGFMCQGG**DFTRGN GTGGESIYGEKFADENFVKKHTGPGILSMANAGPNTNGS QFFICTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGS GRTSKPVVIADSGQLA
52	The amino acid sequence of SEQ ID 318. The conserved FKBP-type peptidyl-prolyl cis-trans isomerase signature is underlined and the FKBP-type peptidyl-prolyl cis-trans isomerase signature 2 is in bold.	MRFTSITSAIALFAAAASALDKPLDIKVDKAVECSRKTK AGDKIQVHYRGTLEADGSEFDASYKRGQPLSFHVGKGQV IKGWDQGLLDMCPGEKRTLTIQPDWGYGSRGMGPIPANS VLIFETELVEIAGVAREEL
53	The amino acid sequence of SEQ ID 319. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined andThe cyclophilin-type peptidyl-prolyl cis-trans isomerase signature 2 is in bold.	MGNPKVFFDMSIGGQPAGRIVMELYADVVPRTAENFRAL CTGEKGAGRSGKPLH**YKGSSFHRVIPGFMCQGG**DFTAGN GTGGESIYGSKFADENFVKKHTGPGVLSMANAGPGTNGS QFFVCTAKTEWLDGKHVVFGQIVDGMDVVKAIEKVGSSS GRTSKPVVVADCGQLS
54	The amino acid sequence of SEQ ID 320. The conserved FKBP-type peptidyl-prolyl cis-trans isomerase signature is underlined and the Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MAVATRSRWVAMSVAWILVLFGTLALIQNRLSDTGASSD PKLVHRKVGEEKKKPDDLEEVTHKVFFDVEIGGKPAGRI VMGLFGKTVPKTVENFRALCTGEKGIGKSGKPLN**YKGSQ FHRIIPKFMIQGGDFTLGDGRGGESIYGNKFSDENFKLK** HTDAGRLSMTNAGPDTNGSQFFITTVTTSWLDGRHVVFG KVLSGMDVVHKIEAEGGQSGQPKSIVVISDSGELDL
55	The amino acid sequence of SEQ ID 321. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase domain is underlined	MAVTLHTNLGDIKCEIFCDEVPKAAEHNARGILSMANSG PNTNGSQFFIAYAKQPHLNGLYTIFGRVIHGFEVLDIME KTQTGPGDRPLAEIRLNRVTIHANPLAG

Entry	Sequence Description	Annotated Peptide Sequence
56	The amino acid sequence of SEQ ID 322. The conserved FKBP-type peptidyl-prolyl cis-trans isomerase signature is underlined and the Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MAVATRSRWVAMSVAWILVLFGTLALIQNRLSDTGASSD PKLVHRKVGEEKKKPDDLEEVTHKVFFDVEIGGKPAGRI VMGLFGKTVPKTVENFRALCTGEKGIGKSGKPLNYKGSQ FHRIIPKFMIQGGDFTLGDGRGGESIYGNKFSDENFKLK HTDAGRLSMANAGPDTNGSQFFITTVTTSWLDGRHVVFG KVLSGMDVVHKIEAEGGQSGQPKSIVVISDSGELDL
57	The amino acid sequence of SEQ ID 323. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined andThe cyclophilin-type peptidyl-prolyl cis-trans isomerase signature 2 is in bold.	MGNPKVFFDMSIGGQPAGRIVMELYADVVPRTAENFRAL CTGEKGAGRSGKPLHYKGSSFHRVIPGFMCQGGDFTAGN GTGGESIYGSKFADENFVKKHTGPGVLSMANAGPGTNGS QFFVCTAKTEWLDGKHVVFGQIVDGMDVVKAIEKVGSSS GRTSKPVVVADCGQLS
58	The amino acid sequence of SEQ ID 324. The conserved A-box of the Retinoblastoma-associated protein is underlined and the B-box of the Retinoblastoma-associated protein is in bold.	MSPVAANAMEEAAEPEVPAPVTPSKDDADTDAAVSRFLG FCKSKLGLAEGNCVQSSTLLRKTAHVLRSSGTVIGTGTA EEAERYWFAFVLYTVRRVGERKAEDEQNGSDETEVPLSR ILKASVLNLIDFFKEIPQFVIKAGAIVSGIYGANWDSRL EAREMQTNYVHLCILCKFYKRICGEFFILNDAKDDMKSA DSSTSDPVIMYQPFGWLLFLALRIHALSRFKDLVSSTNA LVSVLAILIIHLPTRFRKFSISDSSQLVKRSEKGVDLVG SLAYRYDTSEDEIKRTLEKANNVIAEILGITPPPASECK AENLENVDTDGLIYFGNLMEETSLSSILSTLEKIYEDAT RNDSEFDERVFINDDDSLLVSGSLSGAAINLTGAKRKYD SFASPAKTITRPLSPSRSPASHINGIIGGTNLRITATPV ATAMTTAKWLRTFVSPLPSKPSTDLQGFLASCDRDVTSD VIRRANIILEAIFPNSPIGERTVTGGLQNANLMDNMWAE QRRLEALKLYYRVLEAMCRAEAQILHSNNLTSLLTNERF HRCMLACSAELVLATHKTVTMLFPAVLERTGITAFDLSK VIESFVRHEETLPRELRRHLNTLEERLLENMVWERGSSM YNSLVVARPALAPEINRLGLLPEPMPSLDAIALLINFSS SGLPQSPVQKHEASPGQNGDIRSPKRISTEYRSVLVERN FTSPVKDRLLALSNIKSKLPPPPLQSAFASPTRPHPGGG GETCAETAIHIFFSKITKLAAVRINAMLERLQLSQQIKE GVYCLFQQILSQRTNLFFNRHIDQVILCCFYGVAKINQI NLTFREIIYNYRKQPQCKPQVFRNVFVDWSTRRNGKAGN EHVDIISFYNEIFIPSVKPLLVELGPTGATTRTNRTSEV GNKNDAQCPGSPKISSFPTLPDMSPKKVSASHNVYVSPL RSSKMDASISHSSKSYYACVGESTHAYQSPSKDLVAINS RLNGNRKVRGTLNFDDVDAGLVSDSMVANSLYLQNGSSM SSSTAKSSEK
59	The amino acid sequence of SEQ ID 325. The conserved G-protein beta WD-40 repeat domains are underlined.	MRPILMKGHERPLTFLKYNREGDLLFSCAKDHTPTVWFA DNGERLGTYRGHNGAVWCCDVSRDSMRLITGSADTTAKL WSVQNGTQLFTFNFDSPARSVDFSIGDKLAVITTDPFME LPSAIHVKRIARDPADQASESVLVLRGHQGRIARAVWGP LNKTIISAGEDAVIRIWDSETGKLLRESDKETGHKKAVT SLMKSVDGSHFVTGSQDKSAKLWDIRTLTLIKTYVTERP VNAVTMSPLLDHVVLGGGQDASAVTMTDHRAGKFEAKFF DKILQEEIGGVKGHFGPINALAFNPDGKSFSSGGEDGYV RLHHFDPDYFNIKI
60	The amino acid sequence of SEQ ID 326. The conserved G-protein beta domain is underlined and the WD-40 repeat domains are in bold	MDKKRTVVPLVCHGHSRPVVDLFYSPITPDGFFLISASK DSSPMLRNGETGDWIGTFEGHKGAVWSCCLDTNALRAAS GSADFSAKLWDALSGDELHSFEHKHIVRSCAFSEDTHLL LTGGVEKILRIFDLNRPDAPPREVDNSPGSIRTVAWLHS DQTILSSCTDIGGVRLWDVRSGKIVQTLETKSPVTSSEV SQDGRYITTADGSTVKFWDANHFGLVKSYNMPCNIESAS LEPKLGNKFIAGGEDMWVHIFDFHTGEEIGCNKGHHGPV HCVRFSPGGESYASGSEDGTIRIWQTGPANNVEGDANPS NGPVTGKAKVGADEVTRKVEDLQIGKEGKDWREG

Entry	Sequence Description	Annotated Peptide Sequence
61	The amino acid sequence of SEQ ID 327. The conserved G-protein beta WD-40 repeat domains are underlined.	MAEGLILKGTMRAHTDMVTAIAIPIDNSDMVVTSSRDKS IILWHLTKEEKVYGVPRRRLTGHSHFVQDVVLSSDGQFA LSGSWDGELRLWDLATGVSARRFVGHTKDVLSVAFSIDN RQIVSASRDRTIKLWNTLGECKYTIQEGEAHTDWVSCVR FSPNTLQPTIVSASWDRTIKVWNLTNCKLRNTLAGHNGY VNTVAVSPDGSLCASGGKDGVILLWDLAEGKRLYNLEAG AIIHSLCFSPNRYWLCAATENSIKIWDLESKSIVEDLRV DLKNEADKTDGTTTAASNKKVIYCTSLNWSADGSTLFSG YNDGVIRVWGTGRY
62	The amino acid sequence of SEQ ID 328. The conserved G-protein beta WD-40 repeat domains are underlined.	MAEGLHLKGTMKAHTDMVTAIAVPIDNADMIVTSSRDKS IILWHLTKEDKVYGVPRRRLTGHSHFVQDVVLSSDGQFA LSGSWDGELRLWDLATGVSARRFVGHTKDVLSVAFSIDN RQIVSASRDRTIKLWNTLGECKYTIQEGEAHNDWVSCVR FSPNTLQPTIVSASWDRTVKVWNLTNCKLRNTLQGHSGY VNTVAVSPDGSLCASGGKDGVILLWDLAEGKKLYSLEAG AIIHSLCFSPNRYWLCAATENSIKIWDLESKSIVEDLRV DLKNEADMSDGTTGAMSSNKKVIYCTSLNWSADGSTLFS GYNDGVIRVWGIGRY
63	The amino acid sequence of SEQ ID 329. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MAEGLHLKGTMKAHTDMVTAIAVPIDNADMIVTSSRDKS IILWHLTKEDKVYGVPRRRLTGHSHFVQDVVLSSDGQFA LSGSWDGELRLWDLATGVSARRFVGHTKDVLSVAFSIDN RQ**IVSASRDRTIKLWN**TLGECKYTIQEGEAHNDWVSCVR FSPNTLQPT**IVSASWDRTVKVWNL**TNCKLRNTLQGHSGY VNTVAVSPDGSLCASGGKDGVILLWDLAEGKKLYSLEAG AIIHSLCFSPNRYWLCAATENSIKIWDLESKSIVEDLRV DLKNEADMSDGTTGAMSSNKKVIYCTSLNWSADGSTLFS GYNDGVIRVWGIGRY
64	The amino acid sequence of SEQ ID 330. The conserved G-protein beta WD-40 repeat domains are underlined.	MSGVPAPPFATTTPENGTMSSNSPAFHRDSDDDDDQGEV FLDDSDIIHEVAVDDEDLPDADDEADEAEEADDSLHIFT GHNGEVYSLACSPTDATLVATGAGDDKGFLWRIGHGDWA VELQGHKDSISSLAFSLDGQLLASGSLDGVIQIWDVPSG NLKGTLDGPGGGIEWIRWHPKGHIILAGSEDSTVWMWNA DKMAYLNMFSGHGNSVTCGDFTPDGKTICTGSDDATLRI WNPKSGENIHVVKGHPYHAEGLTSMAISSDSGLAITGAK DGSVRIVNISSGRVVSSLDAHADSVEFVGLALSSPWAAT GSLDQKLIIWDLQHSSPRATCDHEDGVTCLSWVGASRFL ASGCVDGKVRVWDSLSGDCVRTFHGHSDAIQSLSVSANE EFLVSVSIDGTARVFEIAEFH
65	The amino acid sequence of SEQ ID 331. The conserved G-protein beta WD-40 repeat domains are underlined.	MGTSQHQLSSCLQLLPRRRGNKNLIFRRTMASGGAAAVA PPPGYKPYRHLKTLTGHVAAVSCVKFSNDGTLLASASLD KTLIIWSSAALSLLHRLVGHSEGVSDLAWSSDSHYICSA SDDRTLRIWSSRSPFDCLKTLRGHTDFVFCVNFNPQSSL IVSGSFDETIRIWEVKTGRCLNVIRAHSMPVTSVHFNRD GSLIVSGSHDGSCKIWDTKNGACLKTLIDDTVPAVSFAK FSPNGKFILVATLNDTLKLWNYATGKFLKIYTGHKNSVY CLTSTFSVTNGKYIVSGSEDRCICIWDLQGKNLIQKLEG HSDTVISVTCHPSENKIASAGLDSDRTVRIWLQDA
66	The amino acid sequence of SEQ ID 332. The conserved G-protein beta WD-40 repeat domains are underlined.	MPSQKIETGHQDIVHDVAMDYYGKRVATASSDTTIKIIG VSNSSGSQHLASLSGHKGPVWQVAWAHPKFGSILASCSY DGQVILWKEGNQNDWAQAHVFNDHKSSVNSIAWAPHELG LCLACGSSDGNISVFTARPDGGWDTTRIEQAHPVGVTSV SWAPSMAPGALVGSGLLDPVQKLASGGCDNTVKVWKLYN GTWKMDCFPALQMHSDWVRDVAWAPNLGLPKSTIASASQ DGTVVIWTVAKEGEQWQGKVLKDFKTPVWRVSWSLTGNL LAVADGNNNVTLWNEAVDGEWQQVTTVEP
67	The amino acid sequence of SEQ ID 333. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MKIAGLKSVENAHDESVWAAAWVPATESRPALLLTGSLD ETVKLWRPDELALERTNAGHFLGVVSVAAHPSGVIAASA SIDSFVRVFDVDTNATIATLEAPPSEVWQMQFDPKGTTL AVAGGGSASIKLWDTATWELNATLSIPRPEQPKPSEKGN KKFVLSVAWSPDGRRLACGSMDGTISIFDVARAKFLHHL EGHFMPVRSLVFSPVEPRLLFSASDDAHVHMYDSEGKSL VGSMSGHASWVLSVDVSPDGAALATGSSDRTVRLWDLSM RAAVQTMSNHSDQVWGVAFRPMAGAGVRAGGRLASVSDD KSISLYDYS

Entry	Sequence Description	Annotated Peptide Sequence
68	The amino acid sequence of SEQ ID 334. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MEIDLGNLAFDVDFHPSEQLVASGLITGDLLLYRYGDGS SPEKLLEVRAHGESCRAVRFINDGKAILTGSPDCSILAT DVETGSVVARVENAHEAAVNRLVNLTESTIATGDDNGCI KVWDTRQRSCCNTFSAHEDFISDMTFASDSMKLVVTSGD GTLSVCNLRSNKVQTRSEFSEDELLSVVIMKNGRKVVCG TQSGTLLLYSWGFFKDCSDRFVDLSPSSVDALLKLDEDR IIAGTENGLISLIGILPNRIIQPIAEHSDHPIERLAFSH DKKFLGSISHDQTLKLWDLNDILGSEDSPSSQAAIDDSD SDEMDVDANPPDSSKGNKKKHSGKGNDVGNANNFFADLG D
69	The amino acid sequence of SEQ ID 335. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MSQQPSVILATASYDHTIRFWEAKSGRCYRTIQYPDSQV NRLEITPHKRYLAVAGNPSIRLFDVNSNTPQPVMSFDSH TNNVMAVGFQYDGNWMYSGSEDGTVRIWDLRARGCQREY ESRGAVNTVVLHPNQTELISGDQNGNIRVWDLTANSCSC ELVPEVDTAVRSLTVMWDGSLVVAANNNGTCYVWRLLRG SQTMTNFEPLHKLQAHNGYILKCLLSPEFCEPHRYLATA SSDHTVKIWNVEGFTLEKTLIGHQRWVWDCVFSVDGAYL ITASSDTTARLWSMSTGQDIRVYQGHHKATTCCALHDGA EGSPG
70	The amino acid sequence of SEQ ID 336. The conserved G-protein beta WD-40 repeat domains are underlined.	MEDAMDMEVEVEVEAEEHSPSSSNPSGSSFRRFGLKNSI QTNFGSDYVFEITPKFDWSLMGVSLSSNAVKLYSPTTGQ YCGECRGHSDTVNGISFSGPSSPHVLHSCSSDGTIRAWD TRSFKEVSCISAGPSQEIFSFSFGGSSDSLLSAGCKSQI LFWDWRNKKQVACLEDSHVDDVTQVCFVPHHQNKLISAS VDGLICIFDTAGDINDDEHMESVINVGTSIGKVGIFGQT FEKLWCLTHIETLSVWDWKEGTNEANFEDARKLASDSWS LDHIDYFVDCHSAEEGEGLWVIGGTNAGTLGYFPVKYKG GAAIGSPEAVLGGGHSDVVRSVLPMSGMAGTTSKTRGIF GWTGGEDGRLCCWLSDDSSATSRSWMSSNLVLKSSRSHH KKNRHQPY
71	The amino acid sequence of SEQ ID 337. The conserved G-protein beta domain is underlined and the WD-40 repeat domains are in bold	MSQHQEYPMEYAADDYDVGEVEDDMYFHERVMGDSDTDE DEEYDHLDNKITDTSAADARRGKDIQGIPWERLSVTREK YRRTRIEQYKNYENVPQSGESSEKDCKPTRKGGNYYEFW RNTRSVKSTILHFQLRNLVWSTTKHDVYLMSHFSIIHWS SLTCKKTEVLDVYGHVAPREKHPGSLLEGFTQTQVSTLA VRDKLLIAGGFQGELICKNLDRPGVSYCCRTTYDDNAIT NAVEIYDYPSGAVHFMASNNDCGVRDFDMEKFELSRHFT FPWPVNHTSLSPDGKLLVIVGDNPEGIVVDSQRGKTIRP LQGHLDFSFASAWHPDGHIFATGNQDKTCRIWDIRNLSK SVAVLKGNLGAIRSIRFTSDGRFMAMAEPADFVHVYDVK SGYEKEQEIDFFGEISGVSFSPDTESLFVGVWDRTYGSL LQYNRCRNYSYLDSM
72	The amino acid sequence of SEQ ID 338. The conserved G-protein beta WD-40 repeat domains are underlined.	MGASSDPNPDVSDEHQKRSEIYTYEAPWHIYAMNWSVRR DKKYRLAIASLLDHPAAAAAVPNRVEIVQLDDSTGEIRA DPNLSFDHPYPATKAAFVPDKDCQRADLLATSSDFLRIW RIADDSSRVDLRSFLNGNKNSEFCRPLTSFDWNEAEPKR IGTSSIDTTCTIWDIERETVDTQLIAHDKEVYDIAWGGV SVFASVSADGSVRVFDLRDKEHSTIIYESSEPDTPLVRL GWNKQDPRYMATIIMDSAKVVVLDIRYPTMPVVELQRHQ ASVNAIAWAPHSSCHICTAGDDSQALIWDLSSMAQPVEG GLDPILAYTAGAEIEQLQWSSSQPDWVAIAFSLKLQ
73	The amino acid sequence of SEQ ID 339. The conserved G-protein beta WD-40 repeat domains are underlined.	MRGGGGGGDATGWDEDAYRESVLKEREVQTRTVFRAAFA PSPSPSPSPDAVVVASSDGSVASYSISACLSDHRLQSLR FADAKSQNVLEAEPACFLQGHDGPAYDVKFYGEGEDSLL LSCGDDGRIRGWMWRDITSSEAHDHSQGNSAKPVLDLVN PQSRGPWGALSPIPENNALAVDVKRGSIYAAAGDSCAYC WDVECGKIKTVFKGHSDYLHCIAARNSSSQIITGSEDGT ARIWDCRSGKCVQVIDPDKDHKKGFFASVSCLALDASES WLVCGRGRDLSVWSISASDCIAKISTNAPAQDVLFDDNQ ILLVGAEPLISRLDMNGAVLSQIHCAPQSVFSVSLHQSG VTAVGGYGGLVDVISQFGSHLCTFRCKCI

Entry	Sequence Description	Annotated Peptide Sequence
74	The amino acid sequence of SEQ ID 340. The conserved G-protein beta WD-40 repeat domains are underlined.	MEAPIIDPLQGDFPEVIEEYLEHGIMKCIAFNRRGTLLA AGCTDGSCIIWDFETRGVAKELRDKECTAAITSVCWSKY GHRILVSASDKSLILWDVLSGEKIAHTTLQHTVLQACLH PGSSTPSICLACPFSSAPMIVDLNTGSTTALPVLTADVS NGATPLSRNKTSDTSVTYSPCNACFNKHGDLVYAGTSKG EILIIDHKNVRVCAIVLVSGGAVIKNVVFSRNGQYMLTN SNDRLIRIYKNLLPPKDGLKMLDELNESFNESDDVEKLK AIGSKCLELLHEFQDSITRVQWKAPCFSGDGEWVIGGAA SRGEHKIYIWDRAGHLVKILEGPKEALMDLAWHPVHPII ISVSLTGLVYIWAKDYTENWSAFAPDFKELEENEEYVER EDEFDLVPETEKVKGLDVHEDDEVDVLTVERDSVFSDSD MSQEELCFLPAVPCLDIPEQQDKCVGSCSKLPDGNHSGS PLSVEAGQNGNASNHNSSPLEPMENSTADDTDGVRLKRK RKPSEKGLELQAEKVKKPVKPLKSSGRLSKTNKPVIDPD SSNGVYGDDGSD
75	The amino acid sequence of SEQ ID 341. The conserved G-protein beta WD-40 repeat domains are underlined.	MRGVSWPEDGNNPSTSSSSQRNQQQAHAPRAVSGHAASH PSASNIFKLLVQREVSPRSKHSSKKLWREASKCQPYPFQ QSCEAVRDVRQGLISWVESASLRHLSAKYCPLVPPPRST IAAAFSPDGKILASTHGDHTVKLIDSQTGSCLKVLRGHR RTPWVVRFHPLYPEILASGSLDHEVRLWDANTAECIGSR NFYRPIASIAFHARGELLAVASGHKLYIWHYNRRGETSS PTIVLRTQRSLRAVHFHPHAAPFLLTAEVNDLDSADSAM TLATSPGYLHYPPPTVYFADAHSHERSRLADELPLMPLP LLMWPSFTRDDGRVPLQRIDGDVGLNGQQRVDSSSSVRL WTYSTPSGQYELLLSPVESGNSPSMPEETGNNAFSSAVE AEVSQSAMDTVEDMEVQPEERNTQFFSFSDPRFWELPLL HGWLVGQTQAGPRSVRQSSPGDIETQSAFGEVASVSPIT SGVMPVSMDPSRFGGRSGSRYRSPGSRGVHVTGPNNDGP RDENDPQSVVSKLRSELAASLAAAASTELPCTVKLRIWP HDVKDPCAQLDLESCRLTIPHAVLCSEMGAHFSPCGRFL AACVACVLPHLESDPGLHGQVNQDVTGVATSPTRHPISA HQIMYELRIYSLEEATFGIVLASRPVRAAHCLTSIQFSP TSEHLLLAYGRRHSSLLKSIVIDGENTVPIYTILEVYRV SDMELVRVLPSAEDEVNVACFHPSVGGGLIYGTKEGKLR ILHYDSSHGLNLKSSGFLDENVPEVQTYALEC
76	The amino acid sequence of SEQ ID 342. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MDSAVAIAALSLVVGAAIALLFFGNYFRKRRSEVVAMAE ADLQPHPKNPSRPPPQPAAKKVHAKSHAHGADKDKNKRH HPLDLNTLKGHGDSVTGLCFASDGRSLATACADGVVRVF KLDDASNKSFKFLRINLPAGGHPTAVAFGDGVSSVIVAS QHLSGCSLYMYGEEKPTNLDSNKQQTKLPMPEIKWEHHK VHEQKAILTLSGAAANYDSGDGSTIIASCSEGTDIIIWH AKTGKILGNVDTNQLKNTMSAISPNGRFIAAAAFTADVK VWEIVYSKDGSVKGVTKVMQLKGHKSAVTWLCFTPNSEQ **IVTASKDGSIRIWNI**NVRYHLDEDTKTLKVFPIPLQDSS GTTLHYERLSLSPDGKILAATHGSMLQWLCIETGKVLDT AEKAHDGDITCMSWAPQSIPTGDKKVNVLATASGDKKVK LWAAPPLPS
77	The amino acid sequence of SEQ ID 343. The conserved G-protein beta WD-40 repeat domains are underlined.	MEVEPKKASKTFPVKPKLKPKPRTPSGKTPESKYWSSFK TTHPLDNLSFSVPSLAFSPSPPHLLAAAHSATVSLFSPH RTTISSFSDVVSSLSFRSDGQLLAASDLSGLIQVFDVRS RTPLRRLRSHARPVRFVRYPVLDKLHLVSGGDDALVKYW DVAGESVVSELRGHKDYVRCGDCSPADANCFVTGSYDHV VKLWDVRVRDGNRAATEVNHGSPVQDVIFLPSGSLVATA GGNSVKIWDLIGGGRMVYSMESHNKTVTSICVGTMGAQQ SGEEGVQLRILSVGLDGYMKVFDYSRMKVTHSMRFPAPL LSIGFSPDSNVRAIGTSNGILYVGKRKAKENAEGGANGI LGLGSVEEPRRRVLKPSFYRYFHRGQSEKPSEGDYLVMR PKKVKLAEHDKLLKKFQHKNALISVLGGNDPEKVVAVME ELVARRALLKCVLNLDADELGLILTFLHKNSTVPRYSSL LLGLAKKVIDLRLEDIRASDALKGHIRNLKRSVDEEIRI QEGLQEIQGMVSPLLRIAGRR

Entry	Sequence Description	Annotated Peptide Sequence
78	The amino acid sequence of SEQ ID 344. The conserved G-protein beta WD-40 repeat domains are underlined.	MQGGSSGVGYGLKYQARCISDVKADTDHTSFLTGTLSLK EENEVHLLRLSSGGTELICEGLFSHPSEIWDLSSCPFDQ RIFSTVFSTGESYGAAVWQIPELYGQLNSPQLEKIASLD AHSRKISCVLWWPSGRHDKLVSIDEENIFLWGLDCSKKS AQVQSQESAGMLHNLSGGAWDPHDVNTVAATCESSIQFW DLRTMKKANSLESVHARDLDYDMRKKHLLVTSEDESGVR VWDLRMPKAPIQEFPGHTHWTWAVRCNPDYEGLILSAGT DSAVNLWWSSTASSDELISERLIDSPTRKLDPLLHSYND YEDSVYGLAWSSREPWIFASLSYDGRVVVESVKPFLSRK
79	The amino acid sequence of SEQ ID 345. The conserved G-protein beta WD-40 repeat domains are underlined.	MAEEEGSAELEQQLEEEFAVWKKNTPILYDLLISHALEW PSLTVHWAPLLPQPSSSAAAAAGDPSLAAHRLVLGTHTS DGAPNFLILADALLPSSESDHCGDDAVLPKVEISQKIRV DGEVNRARFMPQNHNIVGAKTNGCEVYVFDCSKQAAKQH DGGFDPDLRLTGHDGEGYGLSWSPLKENYLLSASHDKKI CLWDISAAAQDKVLGAMHVFEAHEGAVGDASWHSKNDNL FGSAGDDCQLMIWDLRTNKAQQCVKAHEKEVNSVSFNSY NDWILATASSDTTVGLFDMRKLTTPLHVFSSHEGEVLQV EWDPNHEAVLASSSEDRRVMVWDLNRIGDEQQEGDASDG PAELLFSHGGHKAKISDFSWNKNEPWVISSVAEDNSVQV WQMAESICGDDDDMQAMEGYI
80	The amino acid sequence of SEQ ID 346. The conserved G-protein beta WD-40 repeat domains are underlined.	MGNYGEEDEDQYFDALEETASVSDRGSNSSDCCSSGSGL DENVLDSLGFEFWTKFPESVRARRNRFLMLTGLGIEANS VDKEDAFPPSCNEIEVYTCKVTRDDGAVQRSLDSYNCIS LLQSSTSIRSNQEVESLRGDSLLSSFRGRSKESDDLTEL CGMGCPESKRNAVSEFGSVSQGSIEELRRIVASSPLVHP LLHRKLEYERELIETKQKMGAGWLRKFGSATCISGRQGD TWSDPDDLEITAGMKMRRVRAHSSKKKYKELSSLYAAQE FLAHEGSISTMKFSMDGQYLASAGEDTVVRVWKVTEEDR SERVNVTVDPSCLYFALNESTQLASLNTNKEHIGKAKTF QRSSDSSCVILPLKVFQITEKPWHEFKGHNGEVLDLSWS SKGYLLSSSTDKTVRLWRVGCDRCQRVYSHNDYVTCISF NPVNENFFISGSIDGKVRIWNVFGGQVVAYIDCREIVSA VCYRSDGKGAIVGTMTGNCLFYSIKDNHLQMDAQVYLHG KKKSPGKRITGFQFPPNDPGKLMITSADSVIRVLSGLDV VCKLKGPRNSGGPMIATFTSDGKHVISASEDSNVYIWNY AGQDKTSSRVKKIWSCESFWSSNASVALPWCGIRTVPEA LAPPSRSEERRASCAENGENHHMLEEYFQKMPPYSPDCF SLSRGFFLELLPKGSATWPEEKLSDTSPPTVSSQAISKL EYKFLKSACHSVLSSAHMWGLVIVTAGWDGRIRTYHNYG LPVRS
81	The amino acid sequence of SEQ ID 347. The conserved G-protein beta WD-40 repeat domains are underlined.	MDIDFKEYRLRCELRGHEDDVRGVCVCGDGSIGTSSRDR TVRLWAPSAGERRKYEVARVLLGHKSFVGPLAWVPPSEE LPEGGIVSGGMDTLVMAWDLRNGEAQTLKGHQLQVTGIV LDGGDIVSASVDCTLIRWKNGQLTEHWEAHKAPIQAVIR LPSGELVTGSSDTTLKLWRGKTCTQTFVGHTDTVRGLAV MPDLGILSASHDGSIRLWAVSGECLMEMVDHTSIVYSVD SHASGLIVSGSEDRFAKIWKDGVCFQSIEHPGCVWDVKF LEDGDIVTACSDGTIRIWTNQEDRMANSTELELFDLELS SYKRSRKRVGGLKLEELPGLEALQVPGTSDGQTKVIREG DNGVAYAWNSTELKWDKIGEVVDGPEDSMNRPALDGVQY DYVFDVDIGDGEPTRKLPYNRSDNPYDTADKWLLKENLP LSYRQQIVEFILANSGQRDFNLDPSFRDPYTGSSAYVPG APSQLAAKQARPTFKHIPKKGMLVFDAAQFDGILKKINE FNNTLLSNQEKKNLSLTDIEISRLGAVVKILKDTSHYHS SKFADADFDLMLKLLESWPYEMMFPVIDIFRMVILHPDG ADGLLRHQEDKKDVLMESIKRATGNPSVPANFLTSIRAV TNLFKNSAYYSWLQKHRSEMLDAFSSCSSSSNKNLQLSY ATLLLNYAVLLIEKKDEEGQSQVLSAALELAENESLEVD ARYRALVAIGSLMLDGLVKRIALDFDVEHIAKAARTSKE AKIAEVGADIELLIKQS

Entry	Sequence Description	Annotated Peptide Sequence
82	The amino acid sequence of SEQ ID 348. The conserved G-protein beta domain is underlined and the WD-40 repeat domains are in bold	MEFTEAYKQSGPCCFSPNARFIAVAVDYRLVIRDTLSLK VVQLFSCLDKISYIEWALDSEYILCGLYKRPMIQAWSLI QPEWTCKIDEGPAGIAYARWSPDSRHILTTSDFQLRLTV WSLVNTACVHVQWPKHASKGVSFTRDGKFAAICTRHDCK DYINLLSCHNWEIMGVFAVDTLDLADIQWSPDDSAIVIW DSPLEYKVLVYSPDGRCLFKYQAYESGLGVKSVSWSPCG QFLAVGSYDQMLRVLSHLTWKTFAEFTHLSNVRAPCCAA IFKEVDEPLQIDMSELSLSDDYMQGNSGDAPEGHYRVRY DVTEVPITLPCQKPPADRPNPKQGIGLMSWSNDSQYICT RNDSMPTILWIWDMRHLELAAILVQKDPIRAAVWDPTGT RLVLCTGSSHLYMWTPSGAYCVSVPLSQFNITDLKWNSD GSCLLLKDKESFCCAAAPLPPDESSDYSSDD
83	The amino acid sequence of SEQ ID 349. The conserved G-protein beta WD-40 repeat domains are underlined.	MATIAALDDDMVRSMSIGAVFSDFVGKLNSLDFHRKDDI LVTAGEDDSVRLYDIANARLLKTTFHKKHGTDRVCFTHH PNSLICSSTKNLDTGESLRYISMYDNRSLRYFKGHKQRV VSLCMSPINDSFMSGSLDHSVRMWDLRVNACQGILRLRG RPTVAYDQQGLVFAVAMEGGAIKLFDSRSYDKGPFDAFL VGGDTSEVCDIKFSNDGKSVLLSTTNNNIYVLDAYAGDK QCGFNLEPSPSTPIEASFSPDGQYVVSGSGDGTLHAWNI SRRNEVACWNSHIGVASCLKWAPRRAMFVAASTVLTFWI PNSEPELASAKGEAGVPPEQV
84	The amino acid sequence of SEQ ID 350. The conserved G-protein beta WD-40 repeat domains are underlined and the beta G-protein (transducin) is in bold.	MSVAELKERHRAATETVNSLRERLKQKRVQLLDTDVAGY ARTQGKTPVTFGATDLVCCRTLQGHTGKVYSLDWTPERN RIVSVSQDGRFIVWNALTSQKTHAIRLPCAWVMTCAFAP NGQSVACGGLDSVCSIFNLNSPVDRDGNLPVSRMLSGHK GYVSSCQYVPDGDAHLITGSGDQTCVLWDITTGLRTSVF GGEFQSGHTADVLSVSINGSSPRIFVSGSCDSTARMWDT RVASRAVHTYHGHEGDVNAVKFFPDGNRFGTGSDDGTCR LFDIRTGHELQVYYQQRGIDEIPHVTSIAFSISGRLLIA GYSNGDCFVWDTLLAQVVLNLGSLQNSHEGRISCLGVSA DGSALCTGSWDTNLKIWAFGGIRRVT
85	The amino acid sequence of SEQ ID 351. The conserved G-protein beta domain is underlined and the WD-40 repeat domains are in bold	MKKRPRGASLDQAVVDIRRREVGGLSGLSFARRLAASEG LVLRLDIYNKLKGHRGCVNTVGFNLDGDIVISGSDDRHV KLWDWQTGKVKLSFDSGHLSNVFQAKIMPYTDDRSIVTC AADGQARHAQILEGGQVQTMLLAKHRGRAHKLAIDPGSP HIVYTCGEDGLVQRLDLRSNTARELFTCREVYGTHVEVV HLNAIAIDPRNPNLFVIGGSDEYARVYDIRNYKWNGSHN FGRSANYFCPSHLIGEAHVGITGLAFSGQSELLVSYNDE SIYLFTQEMGLGPDPLSASTKSVDSNSSEVTSPTAVNVD DNVTPQVYKGHRNCETVKGVGFFGPKCEYVVSGSDCGRI FIWKKKGGQLIRVMAADKHVVNCIEPHPHIPALASSGIE NDIKIWTPKAIERATLPMNVEQLKPKARGWMNRISSPRQ LLLQLYSLERWPEHGGETSSGLAAGQEELTELFFALSAN GNGSPDGGGDPSGPLL
86	The amino acid sequence of SEQ ID 352. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MSKRGYKLQEFVAHSSNVNCLSIGKKACRLFLTGGDDCK VNLWAIGKPNSLMSLCGHTNAVESVAFDSAEVLVLAGAS SGVIKLWDVEEAKLVRGLTGHRSNCTAMEFHPFGEFFAS GSTDTNLKIWDIRKKGCIHTYKGHTRGISTIRFSPDGRW VVSGGNDNVVKVWDLTAGKLLHDFKFHENHIRSIDFHPL EFLLATGSADRTVKFWDLETFELIGSSRPEAAGVRAIAF HPDGRTLFCGLEDSLKVYSWEPVICHDGVDMGWSTLADL CIHDGKLLGCSYYQSSVGVWVADASLIEPYGTNVKPQQK DSGDDEIEHQESRPSAKVGTTIRSTSIMRCASPDYETKD IKNIYVDTASGNPVSSQRVGTTNFAKVTQPLDFNDTPNL TLRRQGLVTETPDGLSGHVPSKSITQPKVVSRDSPDGKD SSRRESITFSRTKPGMLLRPAHSRRPSSTKYDVDRLSAC AEIGVLSSAKSGSESLVDSFLNIKVAPEDGARNGCEDNH SSVKNVSVESEKVLPLQTPKTEKCDQTVGFKEEINSVKF VNGVAVVPGRTRTLVEKFEKREKELNSTEDQTINTPENPT LDKTPPPSLAENEEKSDRLNIVERKATRMSSHMVTAEDR TPVTLVGSPEDQSTVMAPQRELPADESSKTPPLPVEDLE IHHGSNVSEDKATILSSQTVSEEDSKRSTLIRNFRRRDR FKSTEGRSPVMATQRKLPTDESGKTSSLPMEDLEIKGGL NVSEDKATSFSSRAPPREDRAHSALVRNVRKRDKFKSTN DTITVMVHQRGLSTDEASTVSVERVERRQLSNNVENPLN NLPPHSVPPTTTRGEPQYVGSESDSVNHEDVTELLLGNH EVFLSTLRSRLTKLQVV

Entry	Sequence Description	Annotated Peptide Sequence
87	The amino acid sequence of SEQ ID 353. The conserved G-protein beta WD-40 repeat domains are underlined.	MSTFLTGTALSNPNPNKSYEVVQPPNDSVSSLSFNPKAN FLVATSWDNQVRCWEIVRSGTSLGTTPKASISHDQPVLC STWKDDGTTVFSGGCDKQVKMWPLSGGQPMTVAMHDAPI KEISWIPEMNLLVTGSWDKTLRYWDTRQANPVHIQQLPE RCYALTVRHPLMVVGTADRNLIIYNLQSPQTEFKRISSP LKYQTRCLAAFPDQQGFLVGSIEGRVGVHHLDDSQQSKN FTFKCHREGSEIYSVNSLNFHPVHHTFATAGSDGAFNFW DKDSKQRLKAMSRCSQPIPCSTFNNDGSIFAYSACYDWS KGAENHNPATAKTYIFLHLPQESEVKGKPRLGTTGRK
88	The amino acid sequence of SEQ ID 354. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signatures are in bold.	MEVEAQQRDVNNVMCQLVDPEGTTLGPPMYLPQDVGPQQ LQQMVNKLLSNEDKLPYTFYISDQELVVPLESYLQKNKV SVEKVLSIVYQPQAIFRIRPVNRCSATIAGHSEAVLSVA FSPDGKQLASGSGDTTVRLWDLSTQTPMFTCKGHKNWVL SIAWSPDGKHLVSGSKAGEIQCWDPLTGQPSGNPLVGHK KWITGISWEPVHLSSPCRRFVSSSKDGDARIWDVTLRRC VICLSGHTLAVTCVKWGGDGVIYTGSQDCTIKVWETSQG KLIRELKGHGHWVNSLALSTEYVLRTGAFDHTGKQYSSA EEMKQVALERYKKMKGNAPERLVSGSDDFTMFLWEPSVS KHPKTRMTGHQQLVNHVYFSPDGQWVASASFDKSVKLWN GITGKFVAAFRGHVGPVYQISWSADSRLLLSGSKDSTLK IWDIRTKKLKRDLPGHADEVFAVDWSPDGEKVVSGGKDK VLKLWMG
89	The amino acid sequence of SEQ ID 355. The conserved G-protein beta WD-40 repeat domains are underlined.	MDAGSAHSSSNMKTQSRSPLQEQFLQRRNSRENLDRFIP NRSAMDFDYAHYMLTEGRKGKENPAVSSPSREAYRKQLA ETLNMNRTRILAFKNKPPTPVELIPHELTSAQPAKPTKT RRYIPQTSERTLDAPDLLDDYYLNLLDWGSSNVLSIALG NTVYLWNASDGSTSELVTIDDETGPVTSVSWAPDGRHIA VGLNNSDVQLWDSADNRLLRTLRGGHRSRVGSLAWNNHI LTTGGMDGLIVNNDVRVRSHIVDTYRGHTQEVCGLKWSA SGQQLASGGNDNILHIWDRSTASSNSPTQWLHRLEEHTA AVKALAWCPFQGNLLASGGGGGDRTIKFWNTHTGACLNS VDTGSQVCALLWNKNERELLSSHGFTQNQLTLWKYPSMV KIAELTGHTSRVLFMAQSPDGCTVASAAGDETLRFWNVF GVPEVAKPAPKANPEPFAHLNRIR
90	The amino acid sequence of SEQ ID 356. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MEEAIPFKNLPSREYQGHKKKVHSVAWNCTGTKLASGSV DQTARVWHIEPHGHGKVKDIELKGHTDSVDQLCWDPKHA DLIATASGDKTVRLWDARSGKCSQQAELSGENINITYKP DGTHVAVGNRDDELTILDVRKFKPIHKRKFNYEVNEIAW NMSGEMFFLTTGNGTVEVLAYPSLRPVDTLMAHTAGCYC IAIDPVGRYFAVGSADSLVSLWDISEMLCVRTFTKLEWP VRTISFNHTGDYVASASEDLFIDISNVQTGRTVHQIPCR AAMNSVEWNPKYNLLAYAGDDKNKYQADEGVFRIFGFES A
91	The amino acid sequence of SEQ ID 357. The conserved G-protein beta WD-40 repeat domains are underlined.	MGKDEEEMRGEIEERLINEEYKVWKKNTPFLYDLVITHA LEWPSLTVEWLPDREEPPGKDYSVQKLVLGTHTSENEPN YLMLAQVQLPLEDAENDARHYDDDRADVGGFGCANGKVQ IIQQINHDGEVNRARYMPQNSFIIATKTVSAEVYVFDYS KHPSKPPLDGACSPDLRLRGHSTEGYGLSWSKFKQGHLL SGSDDAQICLWDINATPKNKSLDAMQIFKVHEGVVEDVA WHLRHEYLFGSVGDDQYLLIWDLRTPSVTKPVQSVVAHQ SEVNCLAFNPFNEWVVATGSTDKTVKLFDLRKISTALHT FDAHKEEVFQVGWNPKNETILASCCLGRRLMVWDLSRID EEQTPEDAEDGPPELLFIHGGHTSKISDFSWNTCEDWVV ASVAEDNILQIWQMAENIYHDEDDVPGEESNKGS

Entry	Sequence Description	Annotated Peptide Sequence
92	The amino acid sequence of SEQ ID 358. The conserved G-protein beta WD-40 repeat domains are underlined.	MMRGFSCTEDGDAPSTSSTSPPPPPPPPHRQQMQAPRAS SSSSGQPTSRRSTGNVFKLLARREVSPRSKHSLKKFWGE ASECQLCPFQQSYEAVRDVRRSLISWVEAFSLQHLSAKY CPLMPPPRSTIAAAFSPDGKILASTHGDHTVKLIDSQTG SCLKVLRGHRRTPWVVRFHPLYPEILASGSLDHEVHLWD ANTAECIGSRNFYRPIASIAFHAQGDLLAVASGHKLYIW HYNRSGETSSPTIVLRTPRSLRAVHFHPHAAPFLLTAEV NDLDLTDSAMTLATSPGYLHYPPPTIYLADAHSNERSRL EDELPLMPSPLLMWPSFTRDDGRATLPHIGGDVGLSGQQ RVDSLSSGQYEFHPSPIEPSSSTSMHEEMGTDPFSSVRE SEVTQSAMNIVDNTEVQPEERSTYSFSFSDPRFWELPSV YGWLVGQTQAAPRTAPSPGALETASALGEVASVSPVRSE FMPGGMDQPRLGGRSGSGCRSSGSRMMRTAGLNDHPHDE NYPQSVVSKLRSELEASLAAAASTELPCTVKLRVWPYDM KDPCALFRSESCRLTIPHAVLCSEMGAHFSPCGRFFAAC VACVLPQLEADPVLHGQVDPDVTGVATSPTRHPVSAYQI MYELRIYSLEEATFGMVLASRSIRAAHCLTSIQFSPTSE HLLLAYGRRHNSLLKSIVIDGENTVPIYSILEVYRVSDM ELVRVLPSAEDEVNVACFHPSVGGGLVYGTKEGKLRILQ IDSSGGLNPKSTGFLDENMAEVPTYALEC
93	The amino acid sequence of SEQ ID 359. The conserved G-protein beta WD-40 repeat domains are underlined.	MGEGDLPRTEAGVLRGHEGAVLAARFNGDGNYCLSCGKD RTIRLWNPHRGIHIKTYKSHGREVRDVHCTSDNSKLISC GGDRQIFYWDVSTGRVIRRFRGHDSEVNAVKFNDYASVV VSAGYDRSVRAWDCRSHSTEPIQIINTFQDSVMSVCLTK TEIIGGSVDGTVRTFDIRIGREISDDLGQPVNCISMSND GNCILASCLDSTLRLVDRSAGELLQEYKGHTCKSYKLDC CLTNTDAHVAGGSEDGYVFFWDLVDASVISKFRAHSSVV TSVSYHPKEDCMITASVDGTIKVWKT
94	The amino acid sequence of SEQ ID 360. The conserved G-protein beta WD-40 repeat domains are underlined	MACIKGVGRSASVAMAPDGGYLATGTMAGTVDLSFSSSA SLEIFGLDFQSDDRDLPLIAESPSSERFNRLSWGKNGSG SDEFSLGLIAGGLVDGTIGLWNPLSLIRSEAGDKAIVGH LSRHKGPVRGLEFNVIAPNLLASGADDGEICIWDLAAPR EPSHFPPLRGSGSAAQGEISFLSWNSKVQHILASTSYNG TTVVWDLKKQKPVISFSDSVRRRCSVLQWNPDLATQLVV ASDEDSSPTLRLWDMRNIMSPVKEFAGHTRGVIAMSWCP NDSSYLVTCAKDNRTICWDTVTGEIVCELPAGSNWNFDV HWYPKIPGVISASSFDGKIGIYNVEGCSRYGVRENEFGA ATLRAPKWFKRPVGASFGFGGKVVSFHTRSTGGPSVNSS EVFVHDIITEQTLVSRSSEFEAAIQSGDRPSLRALCEKK SQHCESTDDQETWGFLKVLLEDDGTARSKLLAHLGFDIP TETNDGSQEDLSQQVNALGLEDVTADKVVQEDNNESMVF PTDNGEDFFNNLPSPRADTPVSTSADGFPTVNAAVEPSQ DEVDGLEESSDPSFDDSVQRALVVGDYKAAVALCMSANK LADALVIAHVGGASLWESTRDKYLKMSRLPYLKVVFAMV NNDLQSLVDTRPLKFWKETLAILCSFAQGEEWAMLCNSL ASKLMAAGNMLAATLCFICAGNIDKTVEIWSRSLATEHD GMSYMDLLQDLMEKTIVLALASGQKQFSASVCKLVEKYA EILASQGLLTTAMDYLKLLGTDDLSPELAVLRDRIAFSV EAEKGANISAFNGSQDPRGAVYGVDQSNYGMVDTSQHYY PEAAQPQVPHTVPGSPYGENYQQPFGSSFGKGYNTPMQY QAPSQASMFVPSEPPQNAQPSFVPTPVTSQPTTRSQFIP APPLALRNPEQYQQPTLGSHLYPGSVNPTFQPLPHAPGP VAPVPPQVSSVPGQNMPQAVAPTQMRGFMPVTNPGVVQN PGPISMQPATPIESAAAQPVVSPAAPPPTVQTADTSNVP APQKPVIATL
95	The amino acid sequence of SEQ ID 361. The conserved G-protein beta WD-40 repeat domains are underlined.	MKERGKGAGRSVDERYTQWKSLVPVLYDWLANHNLVWPS LSCRWGPQLEQATYKNRQRLYLSEQTDGSVPNTLVIANV EVVKPRVAAAEHISQFNEEARSPFVKKFKTIIHPGEVNR IRELPQNSKIVATHTDSPDVLIWDVETQPNRHAVLGAST SRPDLILTGHKDNAEFALAMSPTEPFVLSGGKDRYVVLW SIQDHISTLAADPGSAKSPGSAGTNNKQSSKAAGGNDKT GDSPSIEPRGVYLGHGDTVEDVTFCPSSAQEFCSVGDDS CLILWDARTGSSPAIKVEKAHHADLHCVDWNPHDVNLIL TGSADNTVRMFDRRNLTSGGVGSPVHTFEGHNAAVLCVQ WSPDKSSVFGSSAEDGILNIWDHEKIGRKIETVGSKVPN SPPGLFFRHAGHRDKVVDFHWNSSDPWTIVSVSDDGEST GGGGTLQIWRMIDLIYRPEEEVLAELDKFKSHILSCTS

Entry	Sequence Description	Annotated Peptide Sequence
96	The amino acid sequence of SEQ ID 362. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MAKIAPGCEPVAGTLTPSKKREYRVTNRLQEGKRPLYAV VFNFIDSRYFNVFATVGGNRVTVYQCLEGGVIAVLQSYI DEDKDESFYTVSWACNIDRTPFVVAGGINGIIRVIDAGN EKIHRSFVGHGDSINEIRTQPLNPSLIVSASKDESVRLW NVHTGICILIFAGAGGHRNEVLSVDFHPSDKYRIASCGM DNTVKIWSMKEFWTYVEKSFTWTDLPSKFPTKYVQFPVF IAPVHSNYVDCNRWLGDFVLSKSVDNEIVLWEPKMKEQS PGEGSVDILQKYPVPECDIWFIKFSCDFHYHSIAIGNRE GKIYVWELQSSPPVLIAKLSHPQSKSPIRQTAMSFDGST ILSCCEDGTIWRWDAITASTS
97	The amino acid sequence of SEQ ID 363. The conserved G-protein beta WD-40 repeat domains are underlined.	MNTAMHFGAGWRSIAEMGYTMSRLEIEPESCEDEKSLDG VGNSQGPNELPRCLDHELAHLTNLKSRPHEHLIRDFPGR RALPVSTVKMLAGRECNYSRRGRFSSADCCHMLSRYVPV NGPSPLDQMNSRAYVSQFSADGSLFVAGFQGSHIRIYNV DKGWKCQKNILTKSLRWTITDTSLSPDQRYLVYASMSPI VHIVDIGSAAMDSLANITEIHEGLDFSADSGPYSFGIFS VKFSTDGREVVAGSSDDSIYVYDLVANKLSLRIPAHESD VNTVCFADESGHIIYSGSDDTYCKVWDRRCLSARNKPAG VLMGHLEGITFIDSRGDGRYFISNGKDQTIKLWDIRKMG SDICRRGFRNFEWDYRWMDYPPRARDSKHPFDLSVATYK GHSVLRTLIRCYFSPVHSTGQKYIYTGSHDSCVYIYDVV TGAQVAALKHHKSPVRDCSWHPEYPMIVSSSWDGDIVKW EFFGNGETEIPAMKKRIRRRHLY
98	The amino acid sequence of SEQ ID 364. The conserved G-protein beta WD-40 repeat domains are underlined.	MEPQPQAPKKRGRKPKPKEDKKEEQLHQPPPPPPPQQQA APAPAPAATRSSTSGSAGGRDRRPQQQHAVDEKYARWKS LVPVLYDWLANHNLLWPSLSCRWGPQLEQATYKNRQRLY ISEQTDGSVPNTLVIANCEVVKPRVAAAEHVSQFNEEAR SPFIRKYKTIIHPGEVNRVRELPQNPNIVATHTDSPDVL IWDVESQPNRHAVYGATASRPNLILTGHQENAEFALAMC PAEPFVLSGGKDKTVVLWSIQDHITASATDQTTNKSPGS GGSIIKKTGEGNEETGNGPSVGPRGIYCGHEDTVEDVAF CPSTAQEFCSVGDDSCLILWDARVGTNPVAKVEKAHNGD LHCVDWNPHDNNLILTGSADNSVNMFDRRNLTSNGVGSP VYKFEGHKAAVLCVQWSPDKPSVFGSSAEDGLLNIWDYE RVDKKVDRAPNAPAGLFFQHAGHRDKIVDFHWNAADPWT MVSVSDDCDTAGGGGTLQIWRMSDLIYRPEEEVLAELEN FKAHVLECSKA
99	The amino acid sequence of SEQ ID 365. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold. The Utp21 specific WD40 associated putative domain is in italics.	MGIFEPYRAVGYITTGVPFSVQRLGTETFVTVSVGKAFQ VYNCAKLSLVLVGPQLPKKIRALASYREYTFAAYGSDIG IFKRAHQLATWSGHTAKVCLLLLFGEHILSVDVDGNAYI WAFKGMNYNLSPVGHILLDSNFTPSCIMHPDTYLNKVIL GSQEGPLQLWNISTKTKLYEFKGWNSSVSSCVSSPALDV VAVGCADGKIHVHNIRYDEELVTFSHSMRGSVTALSFST DGQPLLASGSSSGVVSIWNLDKRRLQSVIRDAHDGSIIS LHFFANEPVLMSSSADNSIKMWIFDTSDGDPRLLRFRSG HSAPPLCIRFYANGRHILSAGQDRAFRLFSVVQDQQSRE LSQRHVSKRAKKLKLKEEEIKLKPVIAFDVAEIRERDWC NVVTSHMDTPQAYVWRLQNFVIGEHILRPCPNKPTPVKA CMISACGNFAILGTAGGWIERFNLQSGISRGSYIDQLEG TNSAHDGEVVGVACDATNTLMISAGYAGDIKVWDFKGRE LKSRWEIGSSLVKISYHRLNGLLATVADDFIIRLFDAVA LRMVRKFEGHTDRITDLCFSEDGKWLLSSSMDGSLRIWD IILARQVDAVFVDVSITALSLSPNMDILATTHVDQNGVF LWVNQSMFSGDSDINLYASGKEVVTVKLPSVSSVEGSQV EESNEPTIRHSESKDVPSFRPSLEQIPDLVTLSLLPKSQ WQSLINLDIIKVRNKPVEPPKKPEKAPFFLPSIPSLSGE ILFKPSEMSDKGDMKADEDKSKITPEVPSSRFLQLLHSC SEAKNFSPFTTYIKGLSPSTLDLELRMLQIIDDDAVDAD ADDPQDVDKRQELLSIELLMDYFIHEISCRSNFEFVQAL VRLFLKIHGETIRRQSVLQNKAKVLLETQCSVWQRVDKL FQGARCMVAFLSNSQF

Entry	Sequence Description	Annotated Peptide Sequence
100	The amino acid sequence of SEQ ID 366. The conserved G-protein beta WD-40 repeat domains are underlined.	MEETKVTCGSWIRRPENVNLAVLGRSPRRRGSAALEIFA FDPKSTSLSSSPLVAHVIEEIEGDPLAIAVHPNGEDIVC FASSGSCLSFELSGQESNLKLLTKELPPLRGIGPQKCMA FSVDGSRFATGGVDGRLRILEWPSLRIILDEPKAHKSIR DLDFSLDSEFLATTSTDGSARIWKAEDGLPCTTLTRRSD EKIELCRFSKDGTKPFLFCTVQRGDKAVTGVWDISTWNK IGHKRLLRKPAVVMSISLDGKYLAQGSKDGDMCVVEVKK MEVSHWSKRLHLGTSLTSLEFCPIERVVITTSDEWGVLV TKLNVPADWKAWQVYLLLLGLFLASLVAFYIFYENSDSF WGFPLGKDQPARPKIGSVLGDPKSADDQNMWGEFGPLDM
101	The amino acid sequence of SEQ ID 367. The conserved G-protein beta WD-40 repeat domains are underlined.	MADPVEHQHQQHQQHQLQQQRRRGWRIQGGQYLGEISAL CFLHLPPPPLSLSSSPVLSLSSGLDSESRDRPACSFRFP SAGSGSQVSLFDLASGAMVRTFYVFRGIRVHGIVLGCAD FPGGSSSSSSTLDYVIAVYGERRVKLFRLSVRLGRGAGE GSGTVLSADLELVSAAPRLSHWVMDVRFLKENGTSEDEL QRCLTVAIGCSDNSIRLWDVDKCSFVLAVSSPERCLLYS MRLWGDNLEDLQVASGTIYNEILIWKVVPNHDAPSSNEL TEEGLTNSCAGNSVHECLRYEAYHICRLVGHEGSIFRIA WSSDGSKLVSVSDDRSARIWEVHCKVQYSEDAGEVGLLF GHSARVWDCYISDNLIVTAGEDCSCRVWGLDGQQHDVIK EHIGRGIWRCLYDPWSSLLVTGGFDSAIKVHKLDASLAE ASAKQSNIKDLSDGTELFTTHLPNSSGHSGHMDSKSEYV RCLSFSCEDVMYIATNHGYLYHAKLCNDGDLRWTELAQV SNEVQIICMELLPSNPYDPRIDADDWVAVGDGKGWTTVV RVVKNSDSPKVSTSFSWAAEMDRQLLGIHWCKSLGHRFI FTADPRGALKLWRFFEVSQSSSLYPENSPRISLIAEFKS DLGARIMCLDVAFESELLICGDLRGNLVLFPLLKDLLLD TFVVSAAKISPVNHFKGAHGISAVSSISVAHMSFNHIEL RSTGADGCICYMEYDKGLQSLNFVGMKQVKELSMIESVS TENESTGYRTSGSYASGFASTDFIIWNLVTEAKVLQVSC GGWRRPHSYYLGDVPEMKNCFAYVKDDIIYIRRHWIKDS KDKILPQNLRLQFHGREVHSLCFVTGDFQLRKNKQSSWI VTGCEDGTVRLTRYTQCTDNWSSSKLLGEHVGGSAVRSI CCVSNIHTTSSGTSVSDVKGIENLPKDIKGTLMEDECNP SLLISVGAKRVLTSWLLRRRKQDGKEDDVTDLQEAENSS LPSSAGSSTFSFQWLSTDMPVKYSVPSKKSGSIKKLIGV SDTNVRCKSL
102	The amino acid sequence of SEQ ID 368. The conserved G-protein beta WD-40 repeat domains are underlined.	MPYKLSATLSNHSSDVRAVASPSDDLILSASRDSTAISW FRQSPSSFTPASVIRAGSRFVNAIAYLPPTPRAPQGYAV VGGQDTVVNVFALGPGDKEEPEYTLVGHTDNVCALSVNS DDTIISGSWDKTAKVWKDFALVYDLKGHQQSVWAVLAMN EKEFLTASADRTIKYWVQHKTMQTYEGHRDAVRGLALIP DIGFASCSNDSEIRVWTMGGDVVYTLSGHTSFVYSLSVL PNGDLVSAGEDRSVRVWRDGECSQVIVHPAISVWAVSTM PNGDIISGSSDGVVRVFSESEKRWATASELKALEDQIAS QSLPSQQVGDVKKTDLPGPEALSVPGKKAGEVKMIRSGD VVEAHQWDSLASSWQKIGEVVDAIGSGRKQLHDGKEYDY VFDVDIQEGAPPLKLPYNVSENPYTAAQRFLEQNDLPTG YLDQVVKFIEQNTAGVKLGNDGYVDPFTGASRYQPATQS TSNTASSSYMDPFTGGSRHIAESAPSNVPQGSHATGIIP FSKPIFFKLANVSAMQAKMFQFDEVLRNEISTATLAMRP DEVIMVNETFTYLSKVVTSTSSARTSLGWIHIETIMQIL DRWPVPQRFPVIDLGRLVTAYCMNAFSGPGDLEKFFSCL FRTSEWTSITSGSKALTKAQETNVLLLFRTIANSLDGAP LNDMEWIKQIFRELAQTPQLVLNKSHRLALASVLFNFSC IGLKGPVPADVRTLHLTIILQVLRSPNDDPEVAYRTCVA LGNMLYSDKTRGTPRDAQSPSPTELKSAVAAIKGGFSDP RINDVHREIMSLI
103	The amino acid sequence of SEQ ID 369. The conserved G-protein beta domain is underlined and the WD-40 repeat domains are in bold	**MPPQKIESGHKDTVHDLAMDYYGKRLATASSDHTINVVG** **VSSSGSQHLATLIGHQGPVWQISWAHPKFGSLLASCSYD** **GRVIIWREGNPNEWTQAQVFEEHKSSVNSVAWAPHELGL** **CLACGSSDGNISVFTARQDGGWDTSRIDQAHPVGVTSVS** **WAPSTAPGALVGSGMMEPVQKLCSGGCDNTVKVWKLYNR** **VWKLDCFPVLQMHTDWVRDVAWAPNLGLPKSTIASASQD** **GRVIIWTLAKEGDQWQGKVLYDFRTPVWRVSWSLTGNIL** **AVADGNNNVSLWNEAVDGEWIQVSTVEP**

Entry	Sequence Description	Annotated Peptide Sequence
104	The amino acid sequence of SEQ ID 370. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MSAPMLEIEARDVVKIVLQFCKENSLHQTFQTLQSECQV SLNTVDSIETFVADINSGRWDAILPQVAQLKLPRNTLED LYEQIVLEMIELRELDTARAILRQTQAMGVMKQEQPERY LRLEHLLVRTYFDPNEAYQDSTKEKRRAQIAQALAAEVT VVPPSRLMALVGQALKWQQHQGLLPPGTQFDLFRGTAAM KQDVDDMYPTTLSHTIKFGTKSHAECARFSPDGQFLVSC SVDGFIEVWDYMSGKLKKDLQYQADETFMMHDDPVLCVD FSRDSEMLASGSQDGKIKVWRIRTGQCLRRLERAHSQGV TSVLFSRDGSQLLSTSFDGSARIHGLKSGKQLKEFRGHS SYVNDAIFSNDGSRVITASSDCTVKVWDVKTSDCLQTFK PPPPLRGGDASVNSVHLFPKNADHIVVCNKTSSIYIMTL QGQVVKSLSSGKREGGDFVAACVSPKGEWIYCVGEDRNL YCFSCQSGKLEHLMKVHEKDVIGVTHHPHRNLVATYSED STMKLWKP
105	The amino acid sequence of SEQ ID 371. The conserved G-protein beta WD-40 repeat domains are underlined.	MDLLQSYAEDNDGDLGRHSSPEPSPPRLLPSKSAAPKVD DTTLALTVAQTNQTLARPIDPSQHAVAFNPTYDQLWAPI CGPAHPYAKDGIAQGMRNHKLGFVEDAAIGSFLFDEQYN TFQRYGYAADPCASTGNEYVGDLDALKQNDGISVYNIRQ QEQKKYAEEYAKKKGEERGEGGREKAEVVSDKSTFHGKE ERDYQGRSWIAPPKDAKATNDHCYIPKRLVHTWSGHTKG VSAIRFFPKHGHLILSAGMDTKVKIWDVFNSGKCMRTYM GHSKAVRDISFCNDGTKFLTAGYDKNIKYWDTETGKVIS TFSTGKIPYVVKLHPDDEKQNILLAGMSDKKIVQWDMNT GQITQEYDQHLGAVNTITFVDDNRRFVTSSDDKSLRVWE FGIPVVIKYISEPHMHSMPSISLHPNTNWLAAQSLDNQI LIYSTRERFQLNKKKRFAGHIVAGYACQVNFSPDGRFVM SGDGEGRCWFWDWKSCKVFRTLKCHEGVCIGCEWHPLEQ SKVATCGWDGLIKYWD
106	The amino acid sequence of SEQ ID 372. The conserved G-protein beta WD-40 repeat domains are underlined.	MESNGNLEQTLQDGRIYRQLNSLIVAHLRDHNFPQAASA VALATMTPLNVEAPRNRLLELVAKGLAVEKGELLRGVSH AGTNDLGGSIPASYGLVPAPWTAIDFSSLRDTKGMSKSF TKHETRHLSDHKNVARCARFSTDGRFFATGSADTSIKLF EVSKIKQMMLPDSTDGAIRAVIRTFYDHTHPVNDLDFHP QNTVLISAAKDHTVKFFDYSKATAKRAFRVIQDTHNVRS VAFHPSGDFLLAGTDHPIPHLYDVNTFQCYLSANVPEFA VNAAINQVRYSSSGGMYVTASKDGTIRFWDGASANCVRS IAGAHGAAEVTSANFTKDQRYVLSCGKDSTVKLWEVGTG RLVKQYLGATHMQLRCQAVFNNTEEFVLSIDEPSNEIVV WDAMTAEKVARWPSNHNGPPRWIEHSPTEAAFVSCGTDR SIRFWKETH
107	The amino acid sequence of SEQ ID 373. The conserved G-protein beta WD-40 repeat domains are underlined.	MSNFQGEDGEYVADDFEAEDGDEELHGRESADPESDVDE IDTPSNRFTDTTADQARRGRDIQGIPWERLSITREKYRR TRLEQYKNYENVPQSGEKSGKDCTVTEKGNSFYEFRRNS RSVKSTILHFQLRNLVWATSKHDVYLMSNYSVVHWSSLT GKKSEVLNLAGHVAPNEKHPGSLLEGFTQTQVSTLAVKD RFLVAGGFQGELICKFLDRPGISFCSRTTYDDNAITNAV EIYVSPSGGIHFIASNNDCGVRDFDMENFELSKHFRFPW PVNHTSLSPDGKLLVIVGDDPEGILVDAKTGKTIMPLRG HLDFSFASEWHPDGVTFATGNQDKTCRIWDIRNLSKSIA VLKGNLGAIRSIRYTSDGRYMAIAEPADFVHVYDTKTGY KKEQEIDFFGEISGMSFSPDTESLFIGVWDRTYGSLLEY GRRRNFSYLDCLV
108	The amino acid sequence of SEQ ID 374. The conserved G-protein beta WD-40 repeat domains are underlined and the splicing factor motif is in bold.	MGVEEDLEDLNALAESTDAAVDGQAALASAVDSVTLQPA PPILPPVIPPPAVPVVAPVPTIPPVLRPLAPLPIRPPVL RPPAPKRDEAGSSDSDSDHDGTAAGSTAEYEIITEESRLV RERHEKAMQDLMMKRRGAALAVP**TNDKAVRARLRRLGEP MTLFGEREMERRDRLRMLMAKLDAEGQLEKLMKAHEDEE** AAASAAPEDVEEEMLQYPFYTEGSKALFNARIDIAKFSI TRAALRLERARRRRDDPDEDVDAEIDWALKKAESLSLHC SEIGDDRPLSGCSFSHDGKLLATCSMSGVAKLWDTCRMP QVNRVLTLKGHTERATDVAFSPVQNHIATASADRTAKLW NTEGTILKTFEGHLDRLGRIAFHPSGKYLGTTSFDKTWR LWDIESGEELLLQEGHSRSIYGIDFHRDGSLVASCGLDA LARVWDLRTGRSILALEGHVKPVLGVSFSPNGYHLATGG EDNTCRIWDLRKKKSLYTIPAHANLISEVKFEPQEGYFL VTASYDTTAKVWSARDFKPVKTLSVHEAKITSVDITADA SHIVTVSHDRTIKLWTSNDDVKEQAMDVD

Entry	Sequence Description	Annotated Peptide Sequence
109	The amino acid sequence of SEQ ID 375. The conserved G-protein beta WD-40 repeat domains are underlined, and the conserved Dip2/Utp12 domain is in bold.	MVKAYLRYEPAAAFGVIASVESNIAYDASGKHLLAPALE KVGVWHVRQGVCTKALAPSASSAAGPSLAVTAIASSPSS LIASGYADGSIRIWDFEKGSCETTLNGHKGAVSVLRYGK LGSLLASGSKDNDIILWDVVGETGLYRLRGHRDQVTDLV FLDSDKKLVSSSKDKYLRVWDLETQHCMQIVGGHHSEIW SLDTDPEERYLVTGSADPELRFYTVKNDSSDERSEADAS GGVGNGDLASHNKWDVLKQFGEIQRQSKDRVATVRFNKN GNLLACQAAGKLVEVFRVLDEAEAKRKAKRRLHRKREKK GADVNENGDSSRGIGEGHDTMVTVADVFKLLQTIRASKK ICSISFCPVAPKSSLATLALSLNNNLLEFHSIEADKTSK MLTIELQGHRSDVRSVTLSSDNTLLMSTSHNSVKIWNPS TGSCLRTIDSGYGLCGLIVPQNKHALIGTKDGAIEIFDV GSGTCIEVVEAHGGSIRSIVAIPNQNGFVTGSADHDIKF WEYGMKQKPGDNSKHLTVSNVRTLKMNDDVLVVAVSPDA QKIAVALLDCTVKVFFMDSLKLMHSLYGHRLPVLCLDIS SDGDLIVTGSADKNLMIWGLDFGDRHKSIFAHGDSIMAV QFVGNTHYMFSVGKDRLVKYWDADKFELLLTLEGHHADI WCLAISNRGDFLVTGSHDRSIRRWDRTEEPFFIEEEKEK RLEEMFESDLDNAFGNKYVPKEEIPEEGAVALAGKKTQE TLSATDSIIEALDIAEVELKRIAEHEEEKNNGKTAEFHP NYVMLGLSPSDFILRALSNVQTNDLEQTLLALPFSDALK LLSYLKDWTTYPDKVELVSRIATVLLQTHYNQLVSTPAA RPLLTTLKDILHKKVKECKDTIGFNLAAMDHLKQLMALR SDALFQDAKVKLLEIRSQLSKRLEERTDPREAKRRKKKQ KKSTNMHAWP
110	The amino acid sequence of SEQ ID 376. The conserved G-protein beta WD-40 repeat domains are underlined.	MGGVQAEREDKDKVSLELTEEILQSMEVGMTFRDYSGRI SSMDFHRASSYLVTASDDESIRLYDVASATCLKTINSKK YGVDLVSFTSHPMTVIYSSKNGWDESLRLLSLHDNKYLR YFKGHHDRVVSLSLCPRNECFISGSLDRTVLLWDQRAEK CQGLLRVQGRPATAYDDPGLVFAIAFGGCVRMFDARKYE KGPFEIFSVGGDVSDANVVKFSNDGRLMLLTTTDGHIHV LDSFRGTLLYTFNVKPTSSKSTLEASFSPEGMFVISGSG DGSVYAWSVRGGKEVASWLSTDTEPPVIKWAPGNLMFAT GSSELSFWIPDLSKLGAYVGRK
111	The amino acid sequence of SEQ ID 377. The conserved G-protein beta WD-40 repeat domains are underlined.	MAAFGAAPAGNHNPNKSSEVIQPPSDSVSSLCFSPRANH LVATSWDNQVRCWELTKNGASVTSVPKASMSHDQPVLCS AWKDDGTTVFSGGCDKQAKMWSLMSGGQPVTVAMHDAPI KEIAWIPEMNVLVTGSWDKTLKYWDTRQSNPVHTQQLPE RCYAMTVRYPLMVVGTADRNLIVFNLQNPQAEFKRFSSP LKYQTRCVAAFPDQQGFLVGSIEGRVGVHHLDDSQISKN FTFKCHRDNNDIYSVNSLNFHPVHHTFATAGSDGTFNFW DKDSKQRLKAMSRCSQPIPCSTFNNDGTIYAYSVCYDWS KGAENHNPATAKTYIFLHLPQESEVKAKPRVGTTNRK
112	The amino acid sequence of SEQ ID 378. The conserved G-protein beta WD-40 repeat domains are underlined.	MNCSISGEVPEEPVVSTKSGHVFERRLIERYVSDYGKCP VSGEPLTMDDVLPVKMGKIVKPRPLQAASIPGLLSIFQN EWDSLMLSNFALEQQLHTARQELSHALYQHDAACRVIAR LKKERDEARSLLALAERQIPMTASSDIAVNAPAMSNGRK ASLDEEPGYAGKKMRPGISASIIAEITDCNLALSQQRKK RQIPSTLAPVEDLERYTQLSSYPLHKTGKPGITSLDICH SKDIIATGGIDTSAVLFDRSSGQIMSTLSGHSKKVTSVN FDAQGDMVLTGSADKTVRIWQGSEDGSYNCRHILKDHTA EVQAITVHATNNYFATASLDNTWCFYEFSTGLCLTQVEG ASGSEGYTSAAFHPDGLILGTGTSNADVKIWDVKTQANV TTFSGHTGAITAISFSENGYFLATAAQDGVKLWDLRKLK NFRTFSAYDKDTGTNSVEFDHSGCYLGLAGSDIRVYQVA SVKSEWNCVKTFPDLSGTGKVTCVKFGPDSKYIAVGSMD HNLRIFGLPSEDGAMES
113	The amino acid sequence of SEQ ID 379. The conserved G-protein beta domain is underlined and the WD-40 repeat domains are in bold	MAAPGVETLKKEIKELKEKIAQHRLDTDGEQPLPAAAKS KSVPEVSAALKQRRILKGHFGKIYALHWSADSRHLVSAS QDGKLIIWNGFTTNKVHAIPLRSSWVMTCAYSPSGNLVA CGGLDNLCSVYKVPHGGNKESSSAQKTYGELAQHEGYLS CCRFIKDNEIVTSSGDSTCILWDVETKTPKAIFNDHTGD VMSLAVFDDKGVFVSGSCDATAKLWDHRVHKQCVMTFQG HESDINSVQFFPDGDAFGTGSDDSSCRLFDIRAYQQINK YSSDKILCGITSVAFSKTGKSLFAGYDDYNTYVWDTLSG NQVEVLTGHENRVSCLGVSEDGKALATGSWDTLLKIWA

Entry	Sequence Description	Annotated Peptide Sequence
114	The amino acid sequence of SEQ ID 380. The conserved G-protein beta WD-40 repeat domains are underlined.	MGGVEDESEPASKRMKLSSRVLRGLANGSSRTEPAAGSS LDLMARPLPIEGDEEVIGSKGVIKRVEFVRLIAKALYSL GYEKSGARLEEESGIPLQSSVVNLFMQQISDGLWDESVV TLHKIGLSDENLVKSASFLILEQKFLELLDQEKAMDALK TLRTEITPLCIKNSRVRELSSCIISPSSCGLLNQNKRNS TRARSRSELLEELQKLLPPAVIIPERRLEHLVEQALVLQ TDACMLHNSIDMEMSLYTDHQCGKEHIPCRTLQILQSHN DEVWLVQFSHNGKYLASASNDRSAIIWEVDENGSVSLKH KLTGHQKPISSVCWSPDDRQLLTCGVGETVRRWDVSSGE CLRVYEKAGHGLISCAWFPDGKWICYGVSDRSICMCDLE GKEIECWKGQRTLSISDLEITSDGKQIISICRETAILLL DREAKYERMIEENQTITSFSLSKDNRYLLVNLLNQEIHL WDIKGDFRLVAKYKGLKRSRFVIRSCFGGLKQAFVASGS EDSQVYIWHKGSGELIEPLPGHSGAVNCVSWNPANHHML ASASDDRTIRIWGLNELNTRHKGARPNGVHYCNGNGTS
115	The amino acid sequence of SEQ ID 381. The conserved G-protein beta WD-40 repeat domains are underlined.	MTQLAETYACMPSTERGRGILIAGNPKPGSNSVLYTNGR SVVILNLDNPLDISVYAEHAYPATVARFSPNGEWVASAD SSGAVRIWGAYNDHVLKKEFKVLSGRIDDLQWSPDGLRI VASGDGKGKSLVRAFMWDSGTNVGEFDGHSRRVLSCAFK PTRPFRIVTCGEDFLVNFYEGPPFKFKLSRRDHSNFVNC LRFSPDGNRFISVSSDKKGIIYDGKTGEKIGELSSDGGH TGSIYAVSWSPDSKQVITVSADKSAKIWDISEDGSGNLR KTLTSSGSGGVDDMLVGCLWQNNHLVTVSLGGTISIYTA GDLDKAPVSFSGHMKNVSSLSVLKGDPKVILSSSYDGLI IKWIQGIGFSGRVQRKESTQIKCLAAVDEEIVTSGYDNK VCRVSGSGDAEFIDIGCQPKDLSLALQCPEFALVSTDTG VVLLRGAKIVSTINLGFAVTASTVAPDGTEAIIGAQDGK LRIYSISGDTLTEEAVLEKHRGAISVIHYSPDLSMFASG DLNREAVVWDRASREVRLKNILYHTARINCLAWSPDSST VATGSLDTCVIIYEVDKPASNRLTIKGAHLGGVYGLAFT DDFSVVSSGEDACIRVWKINRQ
116	The amino acid sequence of SEQ ID 382. The conserved G-protein beta WD-40 repeat domains are underlined and the SOF1 protein domain is in bold.	MKVKVISRSTDEFTRERSQDLQRVFRNFDPNLRTQEKAV EYVRALNAAKLDKVFARPFVGAMDGHVDSVSCMAKNPNY LKGIFSGSMDGDIRLWDIASRRTVCQFPGHQGPVRGLAA STDGQILVSCGIDSTVRLWNVPVATLGESDGTHENLAKP LAVYVWKNAFWAVDHQWDGELFATAGAQVDIWNQNRSQP ISSFEWGTDTVISVRFNPGEPNVLATSGSDRSITLYDLR MSSPTRKVIMRTKTNAISWNPMEPMNFTAANEDCNCYSY DARKLEEAKCVHKDHVSAVMDIDYSPTGREFVTGSYDRT VRIFQYNGGHSREVYHTKRMQRVFCVKFSCDASYVISGS DDTNLRLWKAKASEQLGVVLPRERRKHEYHEAVKSRYKH LPEVKRIVRHRHLPKPIYKAGILRRTVNEADRRKEERRK AHSAPGSSSAEPLRKRRIIKEIE
117	The amino acid sequence of SEQ ID 383. The conserved G-protein beta WD-40 repeat domains are underlined.	MVRSIKNPKKAKRKNKGSKNGDGSSSSSSIPSMPTKVWQ PGVDKLEEGEELQCDPSAYNSLHAFHIGWPCLSFDIVRD TLGLVRTEFPHQVYFVAGTQAEKPTWNSIGIFKVSNITG KRRELVPSKPTDDADEESDSSDSDEDSDDEVGGSGTPIL QLRKVGHEGCVNRIRAMNQNPHICASWGDSGHVQIWDFS SHLNALAESEADVSQGASSVFNQAPLVKFGGHKDEGYAL DWSPLVPGRLVSGDCKNSIHLWEPTSGSTWNVDSTPFIG HAASVEDLQWSPTEENVFASCSVDGTIAIWDTRLGKTPA ASFKAHDADVNVISWNRLATCMLASGCDDGTFSIHDLRL LKEGDSVVAHFEYHKHPVTSIEWSPHEASTLAVSSADCQ LTIWDLSLEKDEEEAEFKAKTKEQVNAPEDLPPQLLFV HQGQKDLKELHWHAQIPGMIVSTAADGFNILMPSNIQST LPSDGA
118	The amino acid sequence of SEQ ID 384. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are in bold.	MERYKVIKELGDGTYGSVWKALNQQTHEIVAIKKMKRKY YIWEECINLREVKSLRKLNHPNIIKLKEVIRENNELFFI FEYMECNLYQIMKERSTPFSETAIIKFCYQILQGLSYMH RNGYFHRDLKPENLLVTSDLIKIADFGLAREVLTSPPYT DYVSTRWYRAPEVLLQSPTYTTAIDMWAVGAILAELFTL HPLFPGESELDEIYKICGVLGTPDYETWPDGMQLAAFRN FIFPQFLPVNLSVLIPHASPEAIDLITRLCSWDPQKRPT AEQALHHPFFRIGMSIPLSLGGHFQDNTCAAEVDTKFHS KKACKAWNGEKESSLECFLGLSLGLKPSLGHLGAMGSQG VGAVKQEVGSSPGCQSNPKQSLFQVLNSRAILPLFSSSP NLNVVPVKSSLPSAYTVNSQVMWPTIAGPPAAAVTVSTL QPSILGDFKIFGKSMGLASQYAGKEASPFS

Entry	Sequence Description	Annotated Peptide Sequence
119	The amino acid sequence of SEQ ID 385. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are boxed in bold.	MGEMGRGINNSSNNNNSNRPAWLQHYDLVGKIGEGTYGL VFLARSKLPNNRGLRIAIKKFKQSKDGDGVSPTAIREIM LLREFSHENVVKLVNVHINHVDMSLYLAFDYAEHDLYEI IRHHREKLNHHNINQYTVKSLLWQLLNGLNYLHSNWIVH RDLKPSNILVMGEGEEHGVVKIADFGLARIYQAPLKPLS DNGVVVTIWYRAPELLLGAKHYTSAVDMWAVGCIFAELI TLKPLFQGVEVKASPNPFQLDQLDKIFKVLGHPTIEKWP TLMNLPHWSKNLQQIQQHKYDNAGLHIGPIPAKSPAYDL LSKMLEYDPRKRITAAQALEHEYFRIDPQPGRNALVPSQ PGEKAINYPPRLVDANTDFDGTIAPQPSQVSSGNAPSGS IASAAVPAVRPLPQQMQLMGMQRMQNPGMAAFNLGAQAS MSGLNHNNIALQRGSSQQQAHQQVRRKEPNSGFPNTGYP PPPKSRRL
120	The amino acid sequence of SEQ ID 386. The conserved protein kinase family domain is underlined. The protein kinases ATP-binding region is in bold and the serine/threonine protein kinases active-site signature is in bold/italics.	MDKYEKLEKVGEGTYGKVYKARDKMTGQLVALKKTRLEM DEEGVPPSSLREISLLQMLSQSIYVVRLLCVEHVTKKGK PLLYLVFEYLDTDLKKFIDYRRSVNAGPLPQNVIQSFMY QLLKGVAHCHSHGVLHRDLKPQNLLVDKSKGLLKVGDLG LGRAFTVPLKCYTHEVVTLWYRAPEVLLGSTHYSTPVDI WSVGCIFAEMVRRQPLFPGDCEIQQLLHIFTLLGTPTEE MWPGVKRLRDWHEYPQWKPENLARAVPNLSPTGLDLISK MLQCDPAKRISAKAAMNHPYFDDLDKSQF
121	The amino acid sequence of SEQ ID 387. The conserved protein kinase family domain is underlined. The protein kinases ATP-binding region is in bold and the serine/threonine protein kinases active-site signature is in bold/italics.	MDGYEKMDKVGEGTYGKVYMARDKKTGQLVALKKTRLEN DGEGIPPTALREISLLQMLSQDIYIVRLLDVKHTENKLG KPLLYLVFEYMESDLKKYIDSYRRSHTKMPPSMIKSFMY QLCRGVAYCHSRGVMHRDLKPHNLLVDKEKGVLKIADLG LSRAFTVPVKKYTHEIVTLWYRAPEVLLGATHYSLPVDI WSVGCIFAEMSRMQALFTGDSEVQQLMNIFRFLGTPNEE VWPGVTKLKDWHIYPEWKPQDISHAVPDLEPSGLDLLSQ MLVYEPSKRISAKKALEHPYFDDLDKSQF
122	The amino acid sequence of SEQ ID 388. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are in bold.	MDAYEKLEKVGEGTYGKVYKAKDKNTGQLVALKKTRLES DDEGIPPTALREISLLQMLSQDIHIVRLLDVEHTENKNG KPLLYLVFEYMDSDLKKYIDGYRRSHTKVPPNIIKSFMY QLCQGVAYCHSRGVMHRDLKPHNLLVDKQRGVVKIADLG LGRAFTIPIKKYTHEIVTLWYRAPEVLLGATHYSTPVDI WSVGCIFAEMVRLQALFIGDSEVQQLFKIFSFLGTPNEE IWPGVTKFRDWHIYPQWKPQDISSAVPDLEPSGVDLLSK MLVYEPSKRISAKKALEHPYFDDLDKSQF
123	The amino acid sequence of SEQ ID 389. The conserved protein kinase family domain is underlined. The protein kinases ATP-binding region is in bold and the serine/threonine protein kinases active-site signature is in bold/italics.	MDSYEKLEKVGEGTYGKVYKAKDKKTGKLVALKKTRLEN DGEGIPPTALREISLLQMLSQDMNIVRLLDVEHTENKNG KPLLYLVFEYMDSDLKKYVDGYRRSHTKMPPKIIKSFMY QLCQGVAYCHSRGVMHRDLKPHNLLVDKQRGVLKIADLG LGRAFTVPIKKYTHEIVTLWYRAPEVLLGATHYSTPVDI WSVGCIFAEMSRMHALFCGDSEVQQLMSIFKFLGTPNEG VWPGVTKLKDWHIYPEWRPQDLSRAVPDLEPSGVDLLTK MLVYEPSKRISAKKALQHPYFDDLDKSQF
124	The amino acid sequence of SEQ ID 390. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are in bold.	MEKYEKLEKVGEGTYGKVYKGRDKRTGRLVALKKTPFHQ EEGIPPTAIREISLLKSLSQCIYIVKLLDVKASFNGKGK HVLFMVFEYADSDLKKHIDAHRQCNTKLSPRSIQSYMFQ LCKGIAYCHSHGVLHRDLKPQNILVDQKIGLLKIADLGL GRACTVPIKSYTFEVVTLWYRAPEVLLGAKRYSMALDIW SLGCIFAELCNLQALFAGDSQIQQLINIFRLLGTPNEQL WPGVTQLSDWHEFPQWRPQDLSKVVFNLDPNGVDLLSKM LQYDPAKRISAKEALDHPYFDSLDKSQF

Entry	Sequence Description	Annotated Peptide Sequence
125	The amino acid sequence of SEQ ID 391. The conserved eukaryotic protein kinase domain is underlined and the serine/threonine protein kinases active-site signatures are in bold.	MGCVCGKPSARAADYVESPAEKGASSNSRSSSMASRRLV APAVMDQGIDAENGHEGDYRTKLRGKQSNGADPVSLLSD DAEKQRHSRHHQHQQHHPIRPHHLRPQGEFVPNANSNPR FGNPPRHIEGEQVAAGWPAWLTAVAGEAIKGWIPRRADS FEKLDKIGQGTYSNVYKARDLDTGKIVALKKVRFDNLEP ESVRFMAREIQVLRRLDHPNVVKLEGLVTSRMSCSLYLV FEYMDHDLAGLAACPGIKFTEPQVKCYMQQLLRGLDHCH SRGVLHRDIKGSNLLIDNGGILKIADFGLATFFHPDQRQ PLTSRVVTLWYRPPELLLGATEYGVAVDLWSTGCILAEL LAGKPIMPGRTEVEQLHKIFKLCGSPSEDYWKKSKLPHA TIFKPQQPYKRCVAETFKDFPPSALALMEVLLAIEPADR GTATSALKSDFFTTKPLACDPSSLPKYPPSKEFDAKIRD EEARRQRAAGGRGRDAARRPSRESRAIPAPEANAELAIS IQKRRLSSQGPSKSKSEKFNPQQEDGAVGFPIEPPRPMH IGIDAGATSRMYSQQFGPSHSGPLSNQISSSIWGKNQKE DEIQMAPGRPSRSSKATISDFRKPGACAPQPGADLSHLS SLVATARSNAGIDTHKDRSGMWQHNRIDAIDGVHNNGKH EFLEVPEHPNRQDWTRFQQPESFKGLDNYHLQDLPATHH RKDERVASKEATMNWQGYGGQGGDKIHYSGPLLPPSGNI DEILKEHERHIQHAVRRARQDKGRPQRSNLSQNERKAFE HRSFVSGVNGNAGYSDLVNELPISVGSNRLKVSKTRGTE EIVELRELEREPLSSVMEKYEREHEM
126	The amino acid sequence of SEQ ID 392. The conserved eukaryotic protein kinase domain is underlined and serine/threonine protein kinases active-site signatures is in bold.	MGCVCAKQSDILGEPESPKVKGSNLASSRWSVSSETKQL PQHSDSGILHHQHYYHPRDESDEAKLKESNYGGSKRRTR QGRDPADLDMGIFVRTPSSQSEAELVAAGWPAWMAAFAG EAIHGWIPRRAESFEKLYKIGQGTYSNVYKARDLDNGKI VALKKVRFDSLDAESVRFMAREILVLRKLDHPNIVKLEG LVTSEVSSSLYLVFEYMEHDLAGLAACPGIKFTEPQVKC YMQQLLQGLDHCHRHGVLHRDIKGSNLLIDNGGILKIAD FGLATFFYPDQKQLLTSRVVTLWYRPPELLLGATDYGVA VDIWSAGCILAELLAGKPILPGRTEVEQLHKIFKLCGSP SEDYWKESKLPHATIFKPQHPYKSCIAEAFKDFSPSALA LLETLLAIEPGHRGEASGALKSEFFTTEPLSCDPSSLPK YPPSKEFDAKLRAQETRRQRDVGVRGHGSEAARRTSRLS RAGPTPNEGAELTALTQKQHSTSHATSNIGSEKPSTKKE DYTAGLHIDPPRPVNHSYETTGVSRAYDAIRGVAYSGPL SQTHVSGSTSGKKPKRDHVKGLSGQSSLQPSKPFIVSDS RSERIYEKSHVTDLSNHSRLAVGRNRDTTDPHKSLSTLM QQIQDGTLDGIDIGTHEYARAPVSSTKQKSAQLQRPSAL KYVDNVQLQNTRVGSRQSDERPANKESDMVSHRQGQRIH CSGPLLHPSANIEDLLQKHEQQIQQAVRRAHHGKREALS NKSSLPGKKPVDHRAWVSSGKGNKESPYFKGKGNKELSD LKGGPTAKVTNFRQKVM
127	The amino acid sequence of SEQ ID 393. The conserved protein kinase family domain is underlined. The protein kinases ATP-binding region is in bold and the serine/threonine protein kinases active-site signature is in bold/italics.	MAVANPGQLNLQEAPSWGSRSVNCFEKLEQIGEGTYGQV YMAKEIETGEIVALKKIRMDNEREGFPITAIREIKLLKK LQHENVIKLKEIVTSPGPEKDEQGKSDGNKYNGSIYMVF EYMDHDLTGLAERPGMRFSVPQIKCYMKQLLIGLHYCHI NQVLHRDIKGSNLLIDNNGILKLADFGLARSFCSDQNGN LTNRVITLWYRPPELLLGSTKYGPAVDMWSVGCIFAELL YGKPILPGKNEPEQLTKIFELCGSPDESNWPGVSKLPWY SNFKPQRQMKRRVRESFKNFDRHALDLVEKMLTLDPSQR ISAKDALDAEYFWTDPVPCAPSSLPRYEPSHDFQTKRKR QQQRQHDEMTKRQKISQHPPQQHVRLPPIQNAGQGHLPL RPGPNPTMHNPPPQFPVGPSHYTGGPRGAGGQNRHPQNI RPLHAAQGGGYNANRGYGGPPQQQGGGYPPHGMGNQGPR GGQFGGRGAGYSQGGPYGGPVGGRGPNVGGGNRGPQFWS EQ
128	The amino acid sequence of SEQ ID 394. The conserved eukaryotic protein kinase domain is underlined and the serine/threonine protein kinases active-site signature is in bold.	MQNMEDNVQSSWSLHGNKEICARYEILERVGSGTYSDVY RGRRKADGLIVALKEVHDYQSSWREIEALQRLCGCPNVV RLYEWFWRENEDAVLVLEFLPSDLYSVIKSGKNKGENGI PEAEVKAWMIQILQGLADCHANWVIHRDLKPSNLLISAD GILKLADFGQARILEEPEAIYEVEYELPQEDIVADAPGE RLMEEDDSVKGVRNEGEEDSSTAVETNFGDMAETANLDL SWKNEGDMVMQGFTSGVGTRWYRAPELLYGATIYGKEID LWSLGCILGELLILEPLFSGTSDIDQLSRLVKVLGTPTE ENWPGCSNLPDYRKLCFPGDGSPVGLKNHVPSCSDSVFS ILERLVCYDPAARLNAKEVLENKYFVEDPYPVLTHELRV PSPLREENNFSEDWAKWKDMEADSDLENIDEFNVVHSSD GFCIKFS

Entry	Sequence Description	Annotated Peptide Sequence
129	The amino acid sequence of SEQ ID 395. The conserved eukaryotic protein kinase domain is underlined and the protein kinases ATP-binding region and serine/threonine protein kinases active-site signatures are in bold.	MDLNQYPEDLNPELPEGTDNVDNPDNNKGSPVPSPHPPL KPLDPSERYRKGITLGQGTYGIVYKAFDTVTNKTVAVKK IHLGKAKEGVNVTALREIKLLKELSHPNIIQLIDAYPHK QNLHIVFEFMETDLEAVIKDRNLVFSPADIKSYLQMTLK GLAVCHKKWVLHRDMKPNNLLIAADGQLKLGDFGLARLF GSPDRKFTHQVFAVWYRAPELLFGAKQYGPAVDIWATGC IFAELLLRKPFLQGVSDLDQIGKIFAAFGTPRQSQWPDV ASLPDFVEFQFVPAPSLRSLFPMASEDALDLLSKMFTLD PKNRITAQQALEHRYFSSVPAPTRPDLLPKPSKVDSSRP PKHASPDGPVVLSPSKARRVMLFPNNLAGILPKQVSQST TGGTPIEFDMPTQKLREVCPRSRITESGKKHLKRKTMDM SAALDECAREQEGQEGKTILDPDHQRSAKKEKHM
130	The amino acid sequence of SEQ ID 396. The conserved cyclin N- and C-terminal family domains are underlined.	MAGGQENCVRITRARAACVSKASAPVIQSQVDEKKSRKR APKRAAVDDLAANASGSQPKRRAVLGDVTNLHAAATDCL STAEDQVDAPNPSIKGRARNKKKEARTSTKVVKDEIHPE SNPLADHSSNLSECQKPPAAKLAEQRSLRGVPSKAKQGG SSNSQSCSKHTDIDKDHTDPQMCTTYVEDIYEYLRNAEL KNRPSANFMETAQNDITPNMRAILVDWLVEVSEEYKLVP DTLYLTVSYIDRYLSANPTSRHKLQLLGVSCMLIASKYE EVCPPHVEEFCYITDNTYTRDEMLSMERKILIFLNFEMT KPTTKSFLRRFVRASQAGNKAPSLHMEFLANYLAELTLM ECSFLQYLPSLIAASTVFLSRLTLDFLTNPWNPTLAHYT GYKASQLKDCVMAIYNVQMNRKGSTLVAIREKYQQHKFK CVASLPPPPFIAERFFEDTPN
131	The amino acid sequence of SEQ ID 397. The conserved cyclin and cyclin C-terminal domains are underlined and the cyclins signature is in bold.	MTGTQASNVRITRARAAKSTLNNALPPLPPAQGKPRGKR AATESNISGFSVAAEPLKRRAVLSDVSNICKEAAAVDCL KKPKAVKVVSQNANAKGRGRGIPRNNKKITQEAEIKKET SPAICNVDDASAGNAIGDDKQNNNVNPLKEVQDNPKELN PIAEQISVHPHCKQSVEKPNEKEIVVSDNKAAIASLKQQ STLQSLRIPKQPKYSLKQGNPVPLANLHEDVGRSSCSDF IDIDSEYKDPQMCTAYVTDIYANMRVVELKRRPLPNFME TTQRDINANMRSVLIDWLVEVSEEYKLVPDTLYLTVSYI DRFLSANVVNRQRLQLLGVSCMLVASKYEEICAPPVEEF CYITDNTYKKEEVLEMEISVLNRLQYDLTTPTTKTFLRR FIRAAQASCKVSSLHLEFMGNYLAELTLVEYDFLKYLPS LIAAAAVFVARMTLDPMVHPWNSTLQHYTGYKVSDMRDC ICAIHDLQLNRKGCTLAAIREKYNQPKFKCVANLFPPPI ISPQFLIDNEV
132	The amino acid sequence of SEQ ID 398. The conserved cyclin and cyclin C-terminal domains are underlined and the cyclins signature is in bold.	MAAPNQNALLINNNNRRPLVDIGNLVGALNAQCNISKNG ARKRAFGDIGNLVEDLDAKCTISKYWVRKRPRTNFGVNA NKGASSSTQGQGIVVRGEQKAWDRIVWGNKQSCAIKMNA QHVTATQRGTAISISDIIDSSVQDGGIKAPSQLKARKQT VRTVTATLTARSEDSLRDVLEVPPGIDDGDRDNPLAVVE YVEDIYHFYRKIEVRSCVPPDYMTRQLEIKDSMRGVIID WLIEVHRTFLLMPETLYLTVNIIDRYLSIQSVTRNELQL MGITAMFIASKYEEISPPKINDLVYITKDAYTSKQIVNM EHTILNRLKFKLTVPTPYVFLVRFLKAAGPDKVMKNLAF FLVDLCLLHYKMIKYSPSMLAAAAVYTAQCTLKKHPYWN KTLILHIGYSEAHLRECAHLMADLHLKAEGSNLKSVYKK YSYPIFGSVAFLSPAKIPAGTVAAPAIDKCAHQIYLRNL R
133	The amino acid sequence of SEQ ID 399. The conserved cyclin N- and C-terminal family domains are underlined.	MFPNKQTQGLVQNKKMASKAAQPKAMVPPQRVPPAANNR RALGDIGNIVADVGGKCNVTKDGVNGKPLAQVSRPITRS FGAQLLAQAAANKGISAANNQTQVPVVIPKADVRGNKQR RTSKSKDIPPTTVVTNESDDCVIIEQAQRIKPTCNHNVG AVGNKEKPQLLTAKPKSLTASLTSRSAVALRGFRFDDEM TEAEEDPLPNIDVGDRDNQLAVVEYVEDIYKFYRRTEQM SCVPDYMPRQQEINPKMRAVLINWLIEVHYRFGLMPETL YLTTNLIDRYLATQLVSRSNYQLVGATAMLLASKYEEIW APEMNDFLDILENKFERKHVLVMEKAMLNKLKFHLTVPT PYVFLVRFLKAAASDEEMENLVFFLMELSLMQYVMIKFP PSMLAAAAVYTAQITLKKTTVWNDVLKRHTGYSEIDLKE CTRLMVAFHQSSEESKLNVVFKKYSMPEYDSVALIKPAK LPA

Entry	Sequence Description	Annotated Peptide Sequence
134	The amino acid sequence of SEQ ID 400. The conserved cyclin and cyclin C-terminal domains are underlined	MAPSFDCVANAYIESCEDQEKLRQNAQILAQSGENDVDE PVSMLVQRETHYMLPEDYLQRLRNRTLDVNVRREAVGWI LKVHSFYNFGAPTAYLAVNYLDRFLSRHRMPQGVKAWMI QLMAVACLSLAAKMEETQVPLPSDLQREDARFIFDARTI QRMELLILSTLQWGMRSITPFSFIDYFAYRAVQGHGHGH DATPKAVMSRAIELILSTTEEIDFMEYRPSAIAAAALLC AAEEVVPLQAVHYKRALSSSITDVDKDKMFGCYNLIQET IIEGGCYWTPMSLQSTEKTPVGVLDAAACLSNTPTSSYS VKPYASVTAAKRRKLNEICSALLVSQAHPC
135	The amino acid sequence of SEQ ID 401. The conserved cyclin and cyclin C-terminal domains are underlined.	MAANFWTSSHCKELLDAEKVGIVHPLDKDQGLTQEDVKI IKINMSNCIRTLAQYVKLRQRVVATAITYCRRVYTRKSF TEYDPQLVAPTCLYLASKAEESTVQAKLVIFYMKKYSKH RYEIKDMLEMEMKLLEALDYYLVIYHPYRPLIQFLQDAG LNDLKVTAWALVNDTYRTDLILTYPPYMIALACIYFACI MEEKDAQAWFEELRVDMNEIKNISMEIVDYYDNYRVIPD EKMNSALNKLPHRF
136	The amino acid sequence of SEQ ID 402. The conserved cyclin domain is underlined.	MAPALSSSYECLSHLLCAEDASNVVGCWDEDESKIFCEE EEGFGIQHFPDFPVPDDDEIRVLVRKESQYMPGKSYVQS YQNLGLDFTARQNAIGWILKVHGSYNFGPLTAYLSINYL DRFLSRNPLPKAKVWMLQLLSVACLSLAAKMEETQVPLL LDLQAEEPDFLFEPRTIQRMELLVLSTLEWRMLSVTPFS FVDYFLQGGGGRKPPPRAMVARANELIFNTHTVLDFLEH RPSAIAAAAVICAAEEVLPLEAAQYKETILSCSLVDKEW VFGSYNLIQEVLIEKFSTPKKAKSASSSIPQSPVGVLDA FCLSNNSNNTSLEASLSVNLYASVAAKRRKLNDYCNTWR MFQHSTC
137	The amino acid sequence of SEQ ID 403. The conserved cyclin domain is underlined.	MAPNCIDCAPSDLFCAEDAFGVVEWGDAETGSLYGDEDQ LHYNLDICDQHDEHLWDDGELVAFAEKETLYVPNPVEKN SAEAKARQDAVDWILKVHAHYGFGPVTAVLSINYLDRFL SANQLQQDKPWMTQLAAVACLSLAAKMDETEVPLLLDFQ VEEAKYIFESRTIQRMELLVLSTLEWRMSPVTPLSYIDH ASRMIGLENHHCWIFTMRCKEILLNTLRDAKFLGLLPSV VAAAIMLHVIKETELVNPCEYENRLLSAMKVNKDMCERC IGLLIAPESSSLGSFSLGLKRKSSTINIPVPGSPDGVLD ATFSCSSSSCGSGQSTPGSYDSNNSSILCISPAVIKKRK LNYEFCSDLHCLED
138	The amino acid sequence of SEQ ID 404. The conserved cyclin-dependent kinases regulatory subunit domain is underlined and the cyclin-dependent kinases regulatory subunits signature 1 is in bold.	MPQIQ**YSEKYTDDTYEYRHVVLPP**ETAKLLPKNRLLNEN EWRAIGVQQSRGWVHYAIHRPEPHIMLFRRPLNYQQNQQ QQAGAQSQPMGLKAQ
139	The amino acid sequence of SEQ ID 405. The conserved cyclin-dependent kinases regulatory subunit domain is underlined and the cyclin-dependent kinases regulatory subunits signature 1 is in bold.	MDQIE**YSEKYYDDTYEYRHVELPP**DVARLLPKNRLLTEN EWRGIGVQQSRGWVHYAIHCSEPHIMLFRRPLNYEQNHQ HPEPHIMLFRRPLNCQPNHQPQAHHPT
140	The amino acid sequence of SEQ ID 406. The conserved cyclin-dependent kinases regulatory subunit domain is underlined and the cyclin-dependent kinases regulatory subunits signature 1 is in bold.	MDQIE**YSEKYYDDTYEYRHVELPP**DVARLLPKNRLLTEN EWRGIGVQQSRGWVHYAIHCSEPHIMLFRRPLNYEQNHQ HPEPHIMLFRRPLNCQPNHQPQAHHPT
141	The amino acid sequence of SEQ ID 407. The conserved cyclin-dependent kinases regulatory subunit domain is underlined and the cyclin-dependent kinases regulatory subunits signature 1 is in bold.	MPQIQ**YSEKYYDDTYEYRHVVLPP**DVARLLPKNRLLNEN EWRGIGVQQSRGWVHYAIHRPEPHIMLFRRHLNYQQNQQ QQAQQQPAQAMGLQA

Entry	Sequence Description	Annotated Peptide Sequence
142	The amino acid sequence of SEQ ID 408. The conserved GCN5-related N-acetyltransferase family domain is underlined and the radical SAM family domain is in bold.	MALVETEPVTLIHPEEPKKFKKKPTPGRGGVISHGLTEE EARVKAIAEIVGAMVEGCRKGEDVDLNALKAAACRRYGL SRAPKLVEMIAALPDGERAAVLPKLKAKPVRTASGIAVV **AVMSKPHRCPHIATTGNICVYCPGGPDSDFEYSTQSYTG YEPTSMRAIRARYNPYVQTRSRIDQLKRLGHTVDKVEFI LMGGTFMSLPADYRDYFIRNLHDALSGHTSSNVEEAVCY SEHSATKCIGLTIETRPDYCLGPHLRQMLSYGCTRLEIG VQSTYEDVARDTNRGHTVAAVADCFCLAKDAGFKVVAHM MPDLPNVGVERDMESFREFFENPAFRADGLKIYPTLVIR** GTGLYELWKTGRYRNYPPEQLVDIIARVLALVPPWTRVY RVQRDIPMPLVTSGVEKGNLRELALARMDDLGLKCRDVR TREAGIQDIHHKIRPEVVELVRRDYCANEGWETFLSYED TRQDILVGLLRLRKCGHNTTCPELKGRCSIVRELHVYGT AVPVHGRDADKLQHQGYGTLLMEQAERIAWKEHRSIKIA VISGVGTRHYYRKLGYELEGPYMMKYLN
143	The amino acid sequence of SEQ ID 409. The conserved chromo domain is underlined and the MOZ/SAS-like protein domain is in bold.	MLGFRDLYTSICEHLQRASGRLPIIAAATSLISTPEIAA VEKENKAPNSVDKMGMGSADESGRFSTSNGQFMNMNNGV VKEEWKGGVPVVPSAPTTVPVITNVKLETPSSPDHDMAR KRKLGFLPLEVGTRVLCKWRDGKFHPVKIIERRKLPNGA TNDYEYYVHYTEFNRRLDEWVKLEQLELDSVETDADEKV DDKAGSLKMTRHQKRKIDETHVEGNEELDAASLREHEEF TKVKNITKIELGRYEIETWYFSPFPSEYNNCEKLYFCEF CLNFMKRKEQLQRHMRKCD**LKHPPGDEIYRSGTLSMFEV DGKKNKVYAQNLCYLAKLFLDHKTLYYDVDLFLFYILCE CDERGCHMVGYFSKEKHSEESYNLACILTLPPYQRKGYG KFLISFSYELSKKEGKVGTPERPLSDLGLLSYRGYWTRV LLDILKKHKSNISIKELSDMTAIKADDVLSTLQGLDLIQ YRKGQHAICAD**PKVLDRHLKAVGRGGLEVDVCKLIWTPY KEQ
144	The amino acid sequence of SEQ ID 410. The conserved MOZ/SAS-like protein domain is underlined.	MGSLDESTCSEEIRDEGKDSIRTKFKVESTVNNAQNGGN DNSKKKRAAGLPLEVGIRLLCKWRDSKLHPVKIIERRKL PNGFPQDYEYYVHYTEFNRRLDEWVKLEQFELDSVETDA DEKIEDKGGSLKMTRHQKRKIDEIHVEEGQGHEDFDPAS LREHEEFTKVKNIAKVELGRYEIETWYFSPFPPEYSHCE KLFFCEFCLNFMKRKEQLQRHMRKCDLKHPPGDEIYRNG TLSMFEVDGKKNKIYGQNLCYLAKLFLDHKTLYYDVDLF LFYVLCECDDRGCHVVGYFSKEKHSDEAYNLACILTLPP YQRKGYGKFLIAFSYELSKKEGKVGTPERPLSDLGLLSY RGYWTRILLDILKKQRGNISIKELSDMTAIKVEDVISTL QVLDLIQYRKGQHVICADPKVLDRHLKAAGIAGLEVDVS KLIWTPYKEQCG
145	The amino acid sequence of SEQ ID 411. The conserved bromo family domain is underlined.	MASAPMVGCDDSRDKHRWVESKVYMRKGHGKGSKGNAGF NAQNSTAQVRRENDNMGNSIADNGKSEAASEGLSSLSRK QITVNQDHPPNETSSMPAVGGLQNIDTHVTFKLEGCSKQ EIWELRKKLTNELEQVRGTFKKLEARELQLRGYSVSAGV NTSYSASQFSGNDMRNNGGKEVTSEVASGGAITPKQAQR ESNPPRQLSISLMENNQAASDMGEKGKRTPKANQYYRNS EFVLGKDKFPPAESKKSKSTGNKKISQSKVFSKETMQVG KEFMPQKSVNEVFKQCSLLLTKLMKHKYGWVFNLPVDAQ ALGLHDYHTIIKRPMDLGTVKSKLEKNLYNSPASFAEDV KLTFSNAMTYNPKGHEVHTMAEQLLQLFEERWKTIYEEH LDGKMRFGSGQGLGASSSTKKLPFQDSKKNIKKSEPAGG PSPPKPKSTNHHASRTPSAKKPKAKDPHKRDMTYEEKQK LSTNLQNLPQERLELIVQIIKKRNPSLCQHDEEIEVDID SFDTETLWELDRFVTNYKKSLSKNKKKALLADQAKRASE HGSARNKHPMIGRELPMNNKKGEQGEKVVEIDHMPPVNP PVVEVEKDGVYAKRSSSSSSSSDSGSSSSDSDSGSSSG SESDAYAATSPPAGSNTSARG

Entry	Sequence Description	Annotated Peptide Sequence
146	The amino acid sequence of SEQ ID 412. The conserved GCN5-related N-acetyltransferase family domain is underlined and the bromodomain is in bold.	MEGHSGALGFGQGFSRSSQSPNLSPSPSHSASASVTSSG QKRKRNEVEHAGVASNSTGMFAVPPSHIYSHLHPMSMSM PMPMHNSHPSSLSESRDGALTSNDDDDNLTGGNQSQLDS MSAGNTDGREDFDDEDDDDDDEEDDDEVEGDEEDQDHDP DADDDSDDGHDSMRTFTAARLDNGAPNSRNLKPKADAAG VAIAPTVKTEPILDTVKEEKVSGNNNNNSVSANNAQVAP SGSAVLLSAVKEEANKPTSTDHIQTSGAYCAREESLKRE EDADRLKFVCFGNDGIDQHMIWLIGLKNIFARQLPNMPK EYIVRLVMDRSHKSVMIIKQNQVVGGITYRPYLSQKFGE IAFCAITADEQVKGYGTRLMNHLKQHARDVDGLTHFLTY ADNNAVGYFIKQDFTKEIKLEKERWHGYIKDYDGGILME CKIDPKLPYTDLPAMIRWQRQTIDEKIRELSNCHIVYSG IDIQKKEAGIPRKPIKVEDIPGLKEAGWTTDQWGHSRFR LLNSPSEGLPNRQVLHAFMRSLHKAMVEHADAWPFKEPV DPRDVPDYYDIIKDPMDVKRMFTNARTYNTHETIYYKCA NR
147	The amino acid sequence of SEQ ID 413. The conserved histone deacetylase family domain is underlined.	MEESGNSLTSGPDGSKRRVSYFYDSDIGNYYYSQGHPMK PHRIRMAHSLIVHYALDEKMEVCRPNLLQSRELRVFHAD DYISFLQSVTPETQHEQLRQLKRFNVGEDCPVFDGLYNF CQTYAGGSVGAAIKLNNKEADIAINWSGGLHHAKKCEAS GFCYVNDIVLAILELLKVHQRVLYIDIDIHHGDGVEEAF YSTDRVMSVSFHKFGDYFPGTGHLKDVGYGKGKYYSLNV PLNDGIDDESYKNLFRPIIQKVMEIYQPEAVVLQCGADS LSGDRLGCFNLSVKGHADCVRFLRSFNVPLVLVGGGGYT IRNVARCWCYETAVAVGVEPQDKLPYNEYYEYFGPDYTL HVAPSNMENQNSAKELAKIRNTLLEQLKRIQHVPSVPFQ ERPPDTKFPEEDEEDYEKRPKGHKWGGEYFGSESDEEQK PQNRDIDISDKPGIRRQSPPNVEAAKKIKVEEEDGDIGI VNENDGAKWPLGEAG
148	The amino acid sequence of SEQ ID 414. The conserve histone deacetylase domain is underlined.	MEESGNSLTSGPDGSKRRVSYFYDSDIGNYYYSQGHPMK PHRIRMAHSLIVHYALDEKMEVCRPNLLQSRELRVFHAD DYISFLQSVTPETQHEQLRQLKRFNVGEDCPVFDGLYNF CQTYAGGSVGAAIKLNNKEADIAINWSGGLHHAKKCEAS GFCYVNDIVLAILELLKVHQRVLYIDIDIHHGDGVEEAF YSTDRVMSVSFHKFGDYFPGTGHLKDVGYGKGKYYSLNV PLNDGIDDESYKNLFRPIIQKVMEIYQPEAVVLQCGADS LSGDRLGCFNLSVKGHADCVRFLRSFNVPLVLVGGGGYT IRNVARCWCYETAVAVGVEPQDKLPYNEYYEYFGPDYTL HVAPSNMENQNSAKELAKIRNTLLEQLKRIQHVPSVPFQ ERPPDTKFPEEDEEDYEKRPKGHKWGGEYFGSESDEEQK PQNRDIDISDKPGIRRQSPPNVEAAKKIKVEEEDGDIGI VNENDGAKWPLGEAG
149	The amino acid sequence of SEQ ID 416. The conserved histone deacetylase family domain is underlined.	MMETGGNSLPSGPDGVKRKVAYFYDPEVGNYYYGQGHPM KPHRIRMTHALLVQYGLHKEMQILKPYPARDRDLCRFHA DDYVAFLRGITPETIQDQVKALKRFNVGDDCPVFDGLYQ YCQTYAGGSVGGAVKLNHKLCDIAINWAGGLHHAKKCEA SGFCYVNDIVLAILELLKYHKRVLYVDIDIHHGDGVEEA FYTTDRVMTVSFHKFGDYFPGTGDIRDIGCGKGKYYAVN VPLDDGIDDESFQSLFKPIIQQVMLVYNPEAIVLQCGAD SLSGDRLGCFNLSVKGHAECVRYMRSFNVPLLMVGGGGY TVRNVARCWCYETGVAVGVEIDDKMPQHEYYEYFGPDYT VHVAPSNMENKNTKQYLDKIRSKILENINSLPCAPSAQF QVQPPDTDFPELEEEDYDERTRSHKWDGASCDSDSENGD LKHRNHDVEESAFPRHNLANISYNTKIKLEGVGTGGLDM AAGTDTKKNDESFEAMDYESGEELRQDHFASTINASQPC DPALLTGVQNQLQSTDTVKPIEQSGNAPGIPPPSVATVS TGTRPSSISRTSSLNSMSSVKQGSILGPNPPQGLNASGL QFPVPTSNSPIRQGGSYSITVQAPDKQGLQNHMKGPQNM PGNS

Entry	Sequence Description	Annotated Peptide Sequence
150	The amino acid sequence of SEQ ID 417. The conserved histone deacetylase family domain is underlined.	MPPKDRVAYFYDGDVGSVYFGPNHPMKPHRLCMTHHLVL SYELHKKMEIYRPHKAYPVELAQFHSADYVEFLHRITPD TQHLFTKELVKYNMGEDCPVFENLFEFCQIYAGGTIDAA HRLNNQICDIAINWSGGLHHAKKCEASGFCYINDLVLGI LELLKHHARVLYVDIDVHHGDGVEEAFYFTDRVMTVSFH KYGDMFFPGTGDVKEVGEREGKYYAINVPLKDGIDDASF TRLFKTIITKVVDIYQPGAIVLQCGADSLAGDRLGCFNL SIDGHAQCVRIVKKFNLPLLVTGGGGYTKENVARCWSVE TGVLLDTELPNEIPDNDYIKYFAPDYSLKINTAGNMENL NSKTYLSAIKVQVMENLRAIQHAPSVQMHEVPPDFYIPD IDEDELNPDERMDQHTQDRQIQRDDEYYDGDNDIDHDME EAS
151	The amino acid sequence of SEQ ID 418. The conserved histone deacetylase family domain is underlined.	MDSSKSEEANILHVFWHEGMLNHDLGTGVFDTLEDPGFL EVLEKHPENADRVRNMLSILRKGPIAPYTEWHTGRAAYL SELYSFHRPDYVDMLAKTSTAGGKTLCHGTRLNPGSWEA ALLAAGTTLEAMRYILDGHGKLSYALVRPPGHHAQPTQA DGYCFLNNAGLAVELAVASGCKRVAVVDIDVHYGNGTAE GFYERDDVLTISLHMNHGSWGPSHPQTGFHDEVGRGKGL GFNLNVPLPNGTGDKGYEHAMHELVVPAISKFMPEMIVL VIGQDSSAFDPNGRECLTMEGYRKIGQIMRQQADQFSGG RLVVVQEGGYHITYAAYCLHATLEGVLCLPHPLLSDPIA YYPEHDIYSERVTFIKNYWQGIISTTDKRN
152	The amino acid sequence of SEQ ID 419. The conserved histone deacetylase family domain is underlined.	MEESGNALVSGPDGSKRRVTYFYDADIGNYYYGQGHPMK PHRMRMAHNLIVHYGLHQRMEVCRPHLAQSKDIRAFHTD DYIHFLSSVAPDTQQEQLRQLKRFNVGEDCPVFDGLFNF CQSSAGGSIGAALKLNRKDADIAINWAGGLHHAKKCEAS GFCYVNDIVLGILELLKVHQRVLYIDIDIHHGDGVEEAF YTTDRVMTVSFHKFGDYFPGTGHIKDVGYGKGKYYALNV PLNDGIDDESYKHLFRPIIQKVMEVYQPEAVVLQCGADS LSGDRLGCFNLSVKGHADCVRFVRSFNIPLMLVGGGGYT IRNVARCWCYETAVAVGVEPQDKLPYNEYYEYFGPDYTL YVAPSNMENLNTEKDLEKMRNVLLEQLSKIQHTPSVPFQ ERPPDTEFNDEEEEDMEKRSKCRIWDGEYVGSEPEEDGK LPRFDADTYERSVLKHENKRLVPVSNVEPLKRIKQEEDG AAV
153	The amino acid sequence of SEQ ID 421. The conserved histone deacetylase family domain is underlined.	MDLNLVSHGEEEEGVRRRKVGIVYDERMCKHATPEDQPH PEQPDRIRVIWDKLNSAGVLHKCVMVEAKEASEEQLAGV HSRKHIEVMKSIGTARYNKKKRDKLAASYSSIYFSQGSS EAALLAAGSVVEISEKVASGELDAGVAIVRPPGHHAEAD KAMGFCLFNNIAIAAKHLVHERPELGVQKVLIVDWDVHH GNGTQHMFWTDPHVLYFSVHRFDAGTFYPGGDDGFYDKI GEGKGAGYNINVPWEQGKCGDADYLAVWDHVLVPVAKSY DPDMVLISGGFDAALGDPLGGCRLTPYGYSLMTKKLMEF AGGKIVLALEGGYNLKSLADSFLACVEALLKDGPSRSSV LTHPFGSTWRVIQAVRKELSSFWPALNEELQLPRLLKDA SESFDKLSSSSSDESSASEDEKKFAEVTSIMEVSPDPSS ILALTAEDIAQPLAGLKIEEAGTDSQRSSDHTLLDLTND DTQKLKQFEGEIFVMIGDEESVPSASSSKDQNESTVVLS KSNIKAHSWRLTFSSIYVWYASYGSNMWNPRFLCYIEGG QVEGMAKRCCGSEDKLLLKGYSGKLFLIECFLGDHTQIH GVQEECPFLIQIVVIRVKRMSACIK
154	The amino acid sequence of SEQ ID 422. The conserved FKBP-type peptidyl-prolyl cis-trans isomerase signature is underlined and the FKBP-type peptidyl-prolyl cis-trans isomerase signatures 1 and 2 are in bold.	MADEDLDLSDVGEVEDEPGEEIESTPPLAVGQEKEINSL ALKKKLLKVGTRWETPENGDEV**TVHYTGTLPDGTKFDSS** RDRGEPFTFKLGQGQVIKGWDQGIVTMKKGERALFTIPP ELAYGSSGVRPTIPPNATLQFDVELLSWTNIVDVCNDGG ILKRIISEGEKYERPKDPDEVTVKYEAKLEDGTLVAKSP EEGVEFYVNDGHFCPAIAKAVKTMKRGESVILTIKPTYA FGERGKDAEEGFAAIPPNATLTTSLELVSFKAVIAVTED KKVIKKILKEADGYDKPSDGTVVQIRYTAKLQDGTIFEK KGYEGEEPFQFVVDEEQVIAGLDKAVETMKTGEIALITI GAEYGFGNFETQRDLAVIPPNSTLIYEVEMISFTKEKES WDMDTTEKIEASKQKKEQGNSLFKVGKYQRAAKKYEKAA KYIEHDSSFSAEEKKQSKVLKVSCNLNHAACRLKLKDFK EAVKLCSKVLELESQNVKALYRRAQAYIETADLDLAEFD IKKALEIEPQNREVQLEYKILKQKQIEYNKKDAKLYGNM FAKLNKLEAFEGKVLS

Entry	Sequence Description	Annotated Peptide Sequence
155	The amino acid sequence of SEQ ID 423. The conserved FKBP-type peptidyl-prolyl cis-trans isomerase family domains are underlined. The FKBP-type peptidyl-prolyl cis-trans isomerase signatures 1 and 2 are in bold. The TPR repeat is in bold/italics.	MADEGLELSDVAEVEDEPGEEFESAPPLVVGQEKELNSS GLKKKLLKAGTRCETPENGDEVTVHYTGTLLDGTKFDSS RDRGEPFTFNIGQGQVIKGWDQGIVTMKKREHALFTIPP ELAYGASGMPPTIPPNATLQFDVELLSWTNIVDVCKDGG ILKRIISDGEKYERPKDPDEVTVKYEAKLEDGMLVAKSP EEGVEFYVNDGNFCPAIVKAVKTMKKGENVTLTIKPAYA FGEQGKDAEEGFAAIPPNATITINLQLVSFKAVKEVTED KKVIKKILKEADGYDKPSDGTVVQIRYTAKLQDGTIFEK KGYAGEEPFQFVVDEEQVIAGLDKAVETMKTGEVALITI GPEYGFGNIETQRDLAVIPPYSTLIYEVEMVSFTKEKES WDMNTTENIEASKQKKEQGNSLFKVGKYLRAAKKYDKAA KYIEHDNSFSAEEKKQSKVLKVSCNLNHAACCLKLKDFK KAVKLCSKVLELESQNVKALYRRAQAYIETADLDLAEFD IKKALEIEPQNREVRLEYLILKQKQIEYNKKDAKLYGNM FARQNKLEAIEGKD
156	The amino acid sequence of SEQ ID 424. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature 2 is in bold.	MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRAL CTGEKGTGRSGKPLHFKGSSFHRVIPGFMCQGGDFTRGN GTGGESIYGEKFADENFVKKHTGPGILSMANAGPNTNGS QFFICTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGS GRTSKPVVIADSGQLA
157	The amino acid sequence of SEQ ID 425. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined andThe cyclophilin-type peptidyl-prolyl cis-trans isomerase signature 2 is in bold.	MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRAL CTGEKGNGRSGKPLHFKGSSFHRVIPGFMCQGGDFTRGN GTGGESIYGEKFADENFVKKHTGPGILSMANAGPNTNGS QFFICTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGS GRTSKPVVIADSGQLA
158	The amino acid sequence of SEQ ID 426. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined andThe cyclophilin-type peptidyl-prolyl cis-trans isomerase signature 2 is in bold.	MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRAL CTGEKGTGRSGKPLHFKGSSFHRVIPGFMCQGGDFTRGN GTGGESIYGEKFADENFVKKHTGPGILSMANAGPNTNGS QFFICTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGS GRTSKPVVIADSGQLA
159	The amino acid sequence of SEQ ID 427. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined andThe cyclophilin-type peptidyl-prolyl cis-trans isomerase signature 2 is in bold.	MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRAL CTGEKGTGRSGKPLHFKGSSFHRVIPGFMCQGGDFTRGN GTGGESIYGEKFADENFVKKHTGPGILSMANAGPNTNGS QFFICTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGS GRTSKPVVIADSGQLA
160	The amino acid sequence of SEQ ID 428. The conserved FKBP-type peptidyl-prolyl cis-trans isomerase signature is underlined and the FKBP-type peptidyl-prolyl cis-trans isomerase signature 1 is in bold and underlined. The TPR repeat is in bold/italics.	MADDFELPESAGMMENEDFGDTVFKVGEEKEIGKQGLKK LLVKEGGSWETPETGDEVEVHYTGTLLDGTKFDSSRDRG TPFKFKLGQGQVIKGWDQGIATMKKGENAVFTIPPDLAY GESGSQPTIPPNATLKFDVELLSWASVKDICKDGGIFKK IIKEGEKWEHPKEADEVLVKYEARLEDGTVVSKSEEGVE FYVKDGYFCPAFAIAVKTMKKGEKVLLTVKPQYGFGHQG REAIGNDVARSTNATLLVDLELVSWKVVDEVTDDKKVLK KILKQGEGYERPNDGAVVKVKYTGKLEDGTIFEEKGSDE EPFEFMAGEEQVVDGLDRAVMTMKKGEVALVSVAAEYGY QTEIKTDLAVVPPKSTLIYEVELVSFVKEKESWDMNTAE KIEAAGKKKEEGNALFKVGKYFRASKKYEKATKYIEYDT SFSEEEKKQSKPLKVTCNLNNAACKLKLKDYTQAEKLCT KVLEVESQNVKALYRRAQAYIQTADLELAELDIKKALEI DPNNRDVKLEYRALKEKQKEYNKKEAKFYGNMFARMSKL EELESRKSGSQKVETANKEEGSDAMAVDGESA

Entry	Sequence Description	Annotated Peptide Sequence
161	The amino acid sequence of SEQ ID 429. The conserved FKBP-type peptidylprolyl isomerase domain is underlined.	MAASLTPLGAGLAYATIYDQAKVRKLEPTKRSLIALCQH SDSQHRRFITRKYHVNVQILNRRDAIRLIGLAAGLCIDL SLMYDARGAGLPPQENAKLCDTTCEKELENAPMITTESG LQYKDIKIGNGPSPPIGFQVAANYVAMVPSGQVFDSSLD KGQPYIFRVGSGQVIKGLDEGLLSMKVGGKRRLYIPGPL AFPKGLNSAPGRPRVAPSSPVIFDVSLEFIPGLESEEE
162	The amino acid sequence of SEQ ID 430. The conserved FKBP-type peptidylprolyl isomerase domain is underlined and the Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MSAASLSADMAIRGTILGKTALHVLGPQVVSQCRQPVMF KCPPHTLRKMRFSAQDLQSKNFYSGFTPFKSVFISTSKR SWQAGSARAMSQDAAFQSKVTTKCFLDIEIGGDPAGRIV LGLFGEDVPKTAENFRALCTGEKGFGYKGSSFHRIIKDF MLQGGDFDRGDGTGGKSIYGRTFEDENFKLAHVGPGVLS MANAGPNTNGSQFFICTVKTPWLDKRHVVFGQVIEGMEI VKKLESEETNRTDRPKRPCRIVDCGELP
163	The amino acid sequence of SEQ ID 431. The conserved FKBP-type peptidylprolyl isomerase domain is underlined.	MGRIKPQTLLQQSKKKKVPGRISVSTIIVCNLIIIFLMF SLVGIYRQRAKRNRATSRSDGDEEMENFGRSKINSVPHQ AIVNTTKGLITLELFGKSSAHTVEKFVEWSERGYFNGLP FYRVIKHFVIQVGDPKFAGNREDWTVGGQLNVQLEFSPK HEAFMLGTSKLEDQGDGFELFITTAPIPDLNDKLNVFGR VIKGQDVVQEIEEVDTDEHFQPKSPIIINDVRLKDEL
164	The amino acid sequence of SEQ ID 432. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MARQSTLLLFWSLVFLGAIVFTQAKHEELEEVTHKVYFD VDIAGKPAGRVVIGLFGKAVPKTVENFRALCTGEKGVGK SGKPLHYKGSFFHRIIPSFMIQGGDFTLGDGRGGESIYG TKFADENFKLKHTGPVFITTVTTDWLDGRHVVFGKIISG MDVVYKVEAEGRQSGQPKRKVKIADSGELSMD
165	The amino acid sequence of SEQ ID 434. The conserved FKBP-type peptidyl-prolyl cis-trans isomerase signature is underlined and the TPR repeat is in bold.	MEMDEIQEQSQPQSSEKQDISQESDTGNDKTINAEKITS ENAEVEEDDMLPPKVNTEVEVLHDKVTKQIIKEGSGNKP SRNSTCFLHYRAWAESTMHKFQDTWQEQQPLELVLGREK KELSGFAIGVAGMKAGERALLHVDWQLGYGEEGNFSFPN VPPRANLIYEAELIGFEEAKEGKARSDMTVEERIEAADR RRQQGNELFKEDKLAEAMQQYEMALAYMGDDFMFQLFGK YKDMANAVKNPCHLNMAQCLLKLNRYEEAIGQCNMVLAE DEKNIKALFRRGKARATLGQTDDAREDFQKVRKFSPEDK AVIRELRLLAEHDKQVYQKQKEMFKGLFGQKPEQKPKKL HWFVVFWQWLLSMIRTIFRMRSKTD
166	The amino acid sequence of SEQ ID 435. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MAGAGEGTPEVTLETSMGPITVELYHKHAPKTCRNFLEL SRRGYYNNVKFHRVIKDFMVQGGDPTGTGRGGESIYGPR FEDEITRDLKHTGAGILSMANAGPNTNGSQFFISLAPTP WLDEKHTIFGRVCKGMDVVKRLGNVQTDKNDRPIHDVKI LRTTVKD
167	The amino acid sequence of SEQ ID 436. The conserved TPR repeat domain is underlined.	MMDPELMRLAQEQMSKISPDELMKMQRQIMANPDLMRMA SENMKNLKPEDIRFAAEQMKNVRKEEMAEISERISRASP EEIEAMKARANLQSAYQLQVAQNLKDQGNQLHARMKYSE AAEKYLQARNNLTGIPFSEAKSLLLASSSNLMSCYLKTG QYEECVQTGSEVLAYDAMNVKALYRRGQAYKQIGKLELA VADLRKAVEVSPEDETIAQALREASTELMEKGGTQDQNG PRIEEIIEEEAVQPTAEKYPQSAPMVTSVTEDVSDDEQG SEDQNGFSRDSFQATNAPDGQMYAESLRNLTENPDMLRT MQSLMKNVDPDSLVALSGGKLSPDMVKTVSGMFGRMSPE EIQNMMKMSSTLSRQNPSTSSRFDDITRGHSNMDSSPQS VSVDNDLFEENQNRVGESSTNLSSSAAFSGMPNFSAEMQ EQVRNQMNDPATRQMFTSMIQNMSPEMMASMSEQFGVKL SPEDAVKAQNAMASLSPNDLDRLMNWATRLQTAIDYARK IKNWILGRPGLIFAISMLLLAIILHRFGYIGD
168	The amino acid sequence of SEQ ID 437. The conserved FKBP-type peptidylprolyl isomerase domain is underlined and the Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MGVEKEILRPGNGPKPRPGQSVTVHCTGYGKNEDLSQKF WSTKDPGQKPFTFTIGQGRVIKGWDEGVLDMQLGEIFKL RCSPDYGYGSNGFPAWGIRPNSVLVFEIEVLSVN

Entry	Sequence Description	Annotated Peptide Sequence
169	The amino acid sequence of SEQ ID 438. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase family domain is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MPNPRCYLDITIGEELEGRILVELYSDVVPKTAENFRAL CTGEKGIGPHTGVPLHYKGLPFHRVIKGFMIQGGDISAQ NGTGGESIYGLKFDDENFQLKHERRGMLSMANSGPNTNG SQFFITTTRTSHLDGKHVVFGKVIKGMGVVRGIEHTPTE SNDRPSLDVVISDCGEIPEGSDDGIANFFKDGDLYPDWP ADLDEKSAEISWWMNAVDSAKCFGNENYKKGDYKMALRK YRKALRYLDICWEKEEIDEEKSNHLRKTKSQIFTNSSAC KLKLGDLKGALLDTEFAMRDGEDNVKALFRQGQAYMALK DVDSAVASFKKALQLEPNDAGIRKELAVATKMINDRRDQ ERRAYARMFQ
170	The amino acid sequence of SEQ ID 439. The conserved FKBP-type peptidylprolyl isomerase domain is underlined and the Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MGDVIDLNGDGGVLKTIIRSAKPGAMQPTEDLPNVDVHY EGTLADTGEVFDTTREDNTLFSFELGKGTVIKAWDIAVK TMKVGEVARITCKPEYAYGSAGSPPDIPENATLIFEVEL VACKPRKGSTFGSVSDEKARLEELKKQREIAAASKEEEK KRREEAKATAAARVQAKLEAKKGQGRGKGKSKGK
171	The amino acid sequence of SEQ ID 440. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is underlined.	MGLGLKIASASFLPIFNIMATRSLCILLVCFIPVLAHVL SLQDPELGTVRVYFQTTYGDIEFGFFPHVAPKTVEHIYK LVRLGCYNSNHFFRVDKGFVAQVADVVGGREVPLNSEQR KEGEKTIVGEFSEVKHVRGILSMGRYSDPDSASSSFSIL LGNAPHLDGQYAVFGKVTKGDDTLKRLEEVPTRQEGIFV MPLERIRILSTYYYDTNERESNLTCDHEVSILKRRLVES AYEIEYQRRKCLP
172	The amino acid sequence of SEQ ID 441. The conserved FKBP-type peptidylprolyl isomerase domain is underlined and the Cyclophilin-type peptidyl-prolyl cis-trans isomerase signatures are in bold.	MASKRSLRTMNVWPTLPPLVLLLLLCFSSMSSSVVAKKS DVSELQIGVKHKPKSCDIQAHKGDRIKVHYRGSLTDGTV FDSSFERGDPIEFELGSGQVIKGWDQGLLGMCVGEKRKL RIPSKLGYGAQGSPPKIPGGATLIFDTELVAVNGKGISN DGDSDL
173	The amino acid sequence of SEQ ID 442. The conserved FKBP-type peptidylprolyl isomerase domain is underlined and the Cyclophilin-type peptidyl-prolyl cis-trans isomerase signatures are in bold.	MSGAPAERPISYFDITIGGKPIGRIVFSLYADLVPKTAE NFRALCTGEKGIGKSGKPLCYAGSGFHRVIKGFMCQGGD FTAGNGTGGESIYGEKFEDEAFPVKHTKPFLLSMANAGK DTNGSQFFITVSQTPHLDDKHVVFGEVIKGKSIVRAIEN YPTASGDVPTSPIIISACGVLSPDDPSLAASEETIGDSY EDYPEDDDSDVQNPEVALDIARKIRELGNKLFKEGQIEL ALKKYLKSIRYLDVHPVLPDDSPPELKDSYDALLAPLLL NSALAALRTQPADAQTAVKNATRALERLELSDADKAKAL YRRASAHVILKQEDEAEEDLVAASQLSPEDMAISSKLKE VKDEKKKKREKEKKAFKKMFSS
174	The amino acid sequence of SEQ ID 443. The conserved FKBP-type peptidylprolyl isomerase domain is underlined.	MASSLRSSLFSSWALDSKSVCSLFNLNPGKMGLPSISTP LNWRTCCCSHSSELLELNEGLQSSRRKTVMGLSTVIALS LVYCDEVGAVSTSKRALRSQKVPEDEYTTLPNGLKYYDL KVGSGTEAVKGSRVAVHYVAKWKGITFMTSRQGMGITGG TPYGFDVGASERGAVLKGLDLGVQGMRVGGQRILIVPPE LAYGNTGIQEIPPNATLEFDVELISIKQSPFGSSVKIVE G
175	The amino acid sequence of SEQ ID 444. The conserved G-protein beta WD-40 repeat domains are underlined.	MGAIEDEEPPLKRLKVSSPGLRRGLEEEAPSLSVGSVSI LMAKSLSLEEGETVGSKGLIRRVEFVRIITQALYSLGYQ KAGALLEEESGILLQSSNVALFRKQILDGKWDESVVTLR GIDQVEVEGNTLKAASFLILQQKFFELLDKGNIPEAMKT LRLEISPMQLNTKRVHELASCIVFPSRCEELGYSKQGNP KSSQRMKVLQEIQQLLPPSIMIPEKRLERLVEQALNVQR EACIFHNSLDPALSLYTDHQCGRDQIPTTTLQVLESHKN EVWFLQFSNNGKYLASASKDCSAIIWEITEGDSFSMKHR LSAHQKPVSFVAWSPDDKLLLTCGIEEVVKLWNVETGEC KLTYDKANSGFTSCGWFPDGERFISGGVDKCIYIWDLEG KELDSWKGQGMPKISDLAVTSDGKEIISICGDNAIVMYN LDTKTERLIEEESGITSLCVSKDSRFLLLNLANQEIHLW DIGARSKLLLKYKGHRQGRYVIRSCFGGSDLAFVVSGSE DSQVYIWHRGNGELLAVLPGHSGTVNCVSWNPVNPHVFA SASDDYTIRIWGVNRNTFRSKNASSSNGVVHLANGGP

Entry	Sequence Description	Annotated Peptide Sequence
176	The amino acid sequence of SEQ ID 445. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MPGTTAGAGIEPIEPQSLKKLSLKSLKRSFDLFASLHGE PQPPDQRSQRIRIACKVRAEYEVVKNLPTLPQREVGSSV SNSNVGETHSSLTTNQAQGFPTDTSGDLSKDEGKEITSI AVHLQPQTGLIDGKAGAIAGTSTAISSVGSSDRYQPSAA IMKRLPSKWPRPIWHPPWKNYRVISGHLGWVRSVAFDPG NEWFCTGSADRTIKIWEVATGKLKLTLTGHIEQIRGLAV SSRHPYLFSAGDDKQVKCWDLEYNKAIRSYHGHLSGVYC LALHPTLDILCTGGRDSVCRVWDIRTKAQIFALSGHENT VCSVFTQAIDPQVVTGSHDTTIKLWDLAAGKTMSTLTYH KKSVRAIAKHPFEHTFASASADNIKKFKLPKGEFLHNML SQQKTIVNAMAINEDNVLVSAGDNGSLWFWDWKSGHNFQ QAQTIVQPGSLDSEAGIYALQYDITGSRLVSCEADKTIK MWKEDETATPESHPINFKAPKDIRRF
177	The amino acid sequence of SEQ ID 446. The conserved G-protein beta WD-40 repeat domains are underlined.	MRPILMKGHERPLTFLKYNRDGDLLFSCAKDHTPTVWYG HNGERLGTYRGHNGAVWCCDVSRDSTRLITSSADQTAKL WNVETGAQLFSFNFESPARAVDLAIGDKLVVITTDPFME LPSAIHIKRIEKDLSKQTADSVLTITGIKGRINRAVWGP LNSTIISGGEDSVVRIWDSETGKLLRESDKETGHQKPIT SLCKSADGSHFLTGSLDKSARLWDIRTLTLIKTYVTERP VNAVAISPLLDHVVIGGGQEASHVTTTDRRAGKFEAKFF HKILEEEIGGVKGHFGPINSLAFNPDGRSFASGGEDGYV RLHHFDPDYFHIKM
178	The amino acid sequence of SEQ ID 447. The conserved G-protein beta WD-40 repeat domains are underlined.	MRPILMKGHERPLTFLKYNRDGDLLFSCAKDHTPTVWYG HNGERLGTYRGHNGAVWCCDVSRDSTRLITSSADQTAKL WNVETGNQLFSFNFESPARAVDLAIGDKLVVITTDPFME LPSAIHIKRIEKDLSKQTADSVLTITGIKGRINRAVWGP LNSTIISGGEDSVVRIWDSETGKLLRESDKETGHQKAIT SLCKSADGSHFLTGSLDKSARLWDIRTLTLIKTYVTERP VNAVAISPLLDHVVIGGGQEASHVTTTDRRAGKFEAKFF HKILEEEIGGVKGHFGPINSLAFNPDGRSFASGGEDGYV RLHHFDPDYFHIKM
179	The amino acid sequence of SEQ ID 448. The conserved G-protein beta WD-40 repeat domains are underlined.	MAENNVGDFIPLDRQEYPSKPAPGAVDSSFWKSFKKKEV SRQIAGVTCINFCPEPPHDFAVTSSTRVHIYDGKSCELK KTITKFKDVAYSGVFRSDGQIIAAGGETGVIQVFNAKSQ MVLRQLKGHGRPVRVVRYSPQDKLHLLSGGDDSMVKWWD ITTQEELLNLEGHKDYVRCGAASPSSVNLWATGSYDHTV RLWDLRNSKTVLQLKHGKPLEDVLFFPSGGLLATAGGNV VKVWDILGGGRPIHTMETHQKTVMAMCISKVPRSGQALG DAPSRLVTASLDGYMKVFDLDHFKVTHSARYPAPILSMG ISSLCRTMAVGTSSGLLFIRQRKGQIEDKIHSDSSGLQV NPVNDEKDSAVLKPNQYRYYLRGRSEKPSEGDYVVKRMA KVYFQEYDKDLRHFNHSKALVSALKAADSKGTVAVIEEL VARKRLIQTLSILNLDELELLINFLSRFILVPKYSRFLI SLTDRVLDARAVDLGKSENLKKQIADLKGIVVQELRVQQ SMQELQGIIEPLIRASAR
180	The amino acid sequence of SEQ ID 449. The conserved C-x8-C-x5-C-x3-H type zinc finger is underlined and in italics and the conserved Cys and His residues in bold, The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold (non-italics).	MDVETSGKPTGNKRTYTRLPRQVCVFWQEGRCTRESCNF LHVDEPGSVKRGGATNGFAPKRSYNGSDERDTLAAGPPG GSRRNISARWGRGRGGIFISDERQKI*RNKVCNYWLAGNC QRGEECKYLHSF*VMGSDVKFLTQLSGHVKAIRGIAFPSD SGKLYSGGQDKKVIVWDCQTGQGTDIPLNDEVGCLMSEG PWIFVGLPNAVKAWNILTSTELSLVGPRGQVHALAVGNG MLFAGTHDGSILAWKFSPASNTFEPAASLVGHTQAVVSL VSGADRLYSGSGMDKTIRVWDLGTFQCLQTLRDHTSVVMS LLCWDQFLLSCSLDNTVKVWVATSSGALEVTYTHNEEHG VLALCGMNDEQAKPVLLCSCNDNTVRLYDLPSFSERGRI FSRNEVRTFQIAPGGLFFTGDATGELKVWNWATQKS
181	The amino acid sequence of SEQ ID 450. The conserved G-protein beta WD-40 repeat domains are underlined.	MSVQELRERHAAATAKVNALRERIKAKRLQLLDTDVATY ASSNGRTPISFSFTDLVCCRTLQGHTGKVYSLDWTSEKN RIVSASQDGRLIVWNALTSQKTHAIKLPCAWVMTCAFSP SGQAVACGGLDSVCSIFQLNNQLDRDGHLPVSRILSGHR SYVSSCQYVPDGDTHVITGSGDRTCIQWDVTTGQRIAIF GGEFPLGHTADVMSVSISAANPKEFVSGSCDTTTRLWDT RIASRAIRTFHGHEADVNTVKFFPDGLRFGSGSDDGTCR LFDIRTGHQLQVYRQPPRENQSPTVTAIAFSFSGRLLFA GYSNGDCFVWDTILEKVVLNLGELQNTHNGRISCLGLSA DGSALCTGSWDKNLKIWAFGGHRKIV

Entry	Sequence Description	Annotated Peptide Sequence
182	The amino acid sequence of SEQ ID 451. The conserved G-protein beta WD-40 repeat domains are underlined.	MKVKIISRSTDEFTRERSNDLQRVFRNFDPNLHTQARAQ EYVRALNAAKLDKIFAKPFLAAMSGHIDGISAMAKSPRH LKSIFSGSVDGDIRLWDIAARRTVQQFPGHRGAVRGLTV STEGGRLISCGDDCTVRLWDIPVAGIGESSYGSENVQKP LATYVGKNSFRAVDYQWDSNVFATGGAQVDIWDHDRSEP TNSFAWGSDTVISVRFNPAEKDIFATTASDRSIVLYDLR MASPLNKLIMQTRNNAIAWNPREPMNFTAANEDCNCYSY DMRRMNISTCVHQDHVSAVMDIDYSPSGREFVTGSYDRT VRIFPYNAGHSREIYHTKRMQRVFCVKFSGDATYVVSGS DDANIRLWKAKASEQLGVLLPRERKRHEYLDAVKERFKH LPEIKRIERHRHLPKPIYKAALLRHTVNAAAKRKEERKR AHSAPGSVVTNPLRKKRIVAQLE
183	The amino acid sequence of SEQ ID 452. The conserved G-protein beta WD-40 repeat domains are underlined.	MDHYYQDDFDYLVDDEMVDFADDVEDDVRTRRRSDIDSD SENDFDLNNKSPDTTALQAKRGKDIQGIPWNRLNFTREK YRETRLQQYKNYENLPRPRRSRNLDKECTNFERGSSFYD FRHNTRSVKATIVHFQLRNLVWATSKHNVYLMQNYSIMH WSSLKQKGEEVLNVAGPIVPSVKHPGSSPQGLTRVQVSA MSVKDNLVVAGGFQGELICKYLDKPGVSFCTKISHDENG ITNAVEIYNDASGATRLMTANNDLAVRVFDTEKFTVLER FSFPWSVNHTSVSPDGKLVAVLGDNADCLLADCKTGKTV GTLRGHLDYSFAAAWHPDGYILATGNQDTTCRLWDVRKL SSSLAVLKGRMGAIRSIRFSSDGRFMAMAEPADFVHLYD TRQNYTKSQEIDLFGEIAGISFSPDTEAFFVGVADRTYG SLLEFNRRRMNYYLDSIL
184	The amino acid sequence of SEQ ID 453. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signatures are in bold.	MAEALVLRGTMEGHTDAVTAIATPIDNSDMI**VSSSR**DKS ILLWNLTKEPEKYGVPRRRLTGHSHFVQDVVISSDGQFA LSGSWDSELRLWDLNTGLTTRRFVGHTKDVLSVAFSIDN RQI**VSASRDRTIKLWN**TLGECKYTIQPDAEGHSNWISCV RFSPSATNPT**IVSCSWDRTVKVWNL**TNCKLRNTLVGHGG YVNTAAVSPDGSLCASGGKDGVTMLWDLAEGKRLYSLDA GDIIYALCFSPNRYWLCAATQQCVKIWDLESKSIVADLR PDFIPNKKAQIPYCTSLSWSADGSTLFSGYTDGKIRVWG IGHV
185	The amino acid sequence of SEQ ID 454. The conserved G-protein beta WD-40 repeat domains are underlined.	MAAIKSTSRSASVAFAPDAPLLAAGTMAGAIDLSFSSLA NLEIFKLDFQSDDPELPVVGECPSNERLNRLSWGSAGGS FGIIAGGLVDGTINIWNPATLINSEDNGDALIARLEQHT GPVRGLEFNTISTNLLASGAEDGELCIWDLANPTAPTHF PPLKGVGSGAQGEISFLAWNRKVQHILASTSYSGTTVVW DLRRQKPIISFPDATRRRCSVLQWNPDASTQLIVASDDD NSPTLRAWDLRNTISPYKEFVGHSRGVIAMSWCPSDSLF LLTCAKDNRTLCWDTGSGEIVCELPAGANWNFDVQWSPK IPGILSTSSFDGKIGIHNIEACSRNVSGEVEFGGAIVRG GPSALLKAPKWLERPAGVSFGFGGKLASFRPSTVAQAAD HRHSEVFIHNLVTEDNLVIRSTEFEAAIADGEKVSLRAL CDRKAEESQSDEEKETWNFLRVMFEDEGTARTKLLEHLG FKVQSEENGDLQETHSSKIDDIGSEIGKTLTLDDKTEED VLPQLKGGQDAAIPQDNGEDFFDNLHSPKEEVSLSHVGN DFVGEKDKDMVVNGAEIEHETEDLTEYSDWNEAIQHSLV VGDYKGAVLQCLSANRMADALIIAHLGGNSLWEKTRDEY LKKAKSSYLKVVSAMVNNDLTGLVNSRPLKSWKETLAML CTYSQREEWTVLCDMLASRLIAAGNVMAATLCYICAGNI EKTVEIWSRSLKYDYDGRSFVDHLQDVMEKTVVLALATG QKRVSPSLSKLVENYAELLASQGLLTTAMEYLKLLGTEE SSHELSILRDRLYLSGTDNKVEASSFPFETRQDLTESQY NMHQTGFGAPETQKNYQENVHQVLPSGSYTDNYQPTANT HYIAGYQPAPQQQPSFQNYFTPASYQPAPSPNVFYPSQV SQAEQSNFAPPVNQPPMKTFVPSTPPILRNVDQYQTPSL NPQLYQGVSSATVETHPYQTGAPASVSVGTTPGQPSVVP NFMVPGPVTAPTVTPRGFMPVTTPTQHPLGSANPPVQPQ SPQSSQVQSV
186	The amino acid sequence of SEQ ID 455. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MAGAADSQLQTLSERDSTPNFKNLHTREYAAHKKKVHSV AWNCTGTKLASGSVDQTARVWNIEPHGHSKTKDLELKGH ADSVDQLCWDPKHSELL**ATASGDRTVRLWDA**RSGKCSQQ VELSGENINITFKPDGTHIAVGNRDDELTIIDVRKFKPL HKRKFSYEVNEIAWNTTGELFFLTTGNGTVEVLSYPSLQ VLHTLVAHTAGCYCIAIDPIGRYFAVGSADALVSLWDLS EMLCVRTFTKLEWPVRTISFNHDGQYIASASEDLFIDIA DVQTGRTVHQISCRAAMNSVEWNPKYNLLAFAGDDKNKY MQDEGVFRVFGFETP

Entry	Sequence Description	Annotated Peptide Sequence
187	The amino acid sequence of SEQ ID 456. The conserved G-protein beta WD-40 repeat domains are underlined.	MAATSPVGAGSGRELANPPTDGISNLRFSNHSDHLLVSS WDRKVRLYDASANSLKGQFVHGGPVLDCCFHDDASGFSG SADNTVRRYDFSTRKEDILGRHEAPVRCVEYSYAAGQVI TGSWDKTLKCWDPRGASGQEKTLVGTYSQLERVYSMSLV GHRLVVATAGRHINVYDLRNMSQPEQRRESSLKYQTRCV RCYPNGTGFALSSVEGRVAMEFFDLSEAGQAKKYAFKCH RKSEAGRDTVYPVNAIAFHPIYGTFATGGCDGYVNVWDG NNKKRLYQYSKYPTSIAALSFSRDGRLLAVASSYTFEEG EKPHEPDAVFVRSVNEAEVKPKPKVYAAPP
188	The amino acid sequence of SEQ ID 457. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MASDDEEGFKNEEAPGVVDEAEVQEGLRACFPLSFGKQE KKQAPLESIHSATKRPEDPRPRRQLGPPRPPPSILAEQE DSDRFVGPPRPPQFVRDDNDDGEAEIMIGPPRPPAQYSD DHDNEETIGPPKPSYLEKGEETDQMVGPSKRGSDDETSG DSDDGDDAVDFRVPLSNEIVLRGHTKVVSALAIDQTGSR VLTGSYDYSVRMYDFQGMTSQLKSFRQLEPAEGHQVRSL SWSPTSDRFLCVTGSAQAKIFDRDGLTLGEFVKGDMYLR DLKNTKGHISGLTCGEWHPKEKQT**ILTCSEDGSLRIWDV** NDFNTQKQVIKPKLAKPGRVPVTACAWGRDGKCIAGGVG DGSIQVWNLKPGWGSRPDLYVAKGHDDDITGLQFSADGN ILLTRSTDETLKVWDLRKAITPLQVFRDLPNNYAQTNVA FSPDERLIFTGTSVERDGNSGGLLCFYDRQTLELVLRIG VSPVHSVVRCTWHPRHNQVFATVGDKKEGGAHILYDPAL SERGALVCVARAPRKKSLDDFEAKPVIHNPHALPLFRDE PSRKRQREKARMDPMKSQRPDLPVTGPGFGGRVGSTKGS LLTQYLLKEGGLIKETWMEEDPREAILKYADVAAKDPKF IAPAYAQTQPETVFAETDSEEEQK
189	The amino acid sequence of SEQ ID 458. The conserved G-protein beta WD-40 repeat domains are underlined.	MKERGQSHAGQPSVDERYTQWKSLVPVLYDWLANHNLVW PSLSCRWGPQMHQATYKNSQRLYLSEQTDGTVPNTLVIA TCEVVKPRVAAAEHISQFNEEARSPFVKKFKTIIHPGEV NRIRELPQNSKIVATHTDGPDVLIWDVDTQPNRQATLGA ADSRPDLVLTGHKDNAEFALAMSPSAPFVLSGGKDKCVL LWSIQDHISAATEPSSAKASKTPSSAHGEKVPKIPSIGP RGVYKGHKDTVEDVQFCPSNAQEFCSVGDDSALILWDAR NGNEPVIKVEKAHNADLHCVDWNPHDENLILTGSADNSV RMFDRRNLTSSGVGSPVHKFEGHSAPVLCVQWCPDKASV FGSAAEDSYLNVWDYEKVGKNVGKKTPPGLFFQHAGHRD KVVDFHWNSFDPWTIVSVSDDGESTGGGGTLQIWRMSDL IYRPEDEVLAELERFRAHILSCQNK
190	The amino acid sequence of SEQ ID 459. The conserved G-protein beta WD-40 repeat domains are underlined. The Lissencephaly type-1-like homology motif is in bold and the CTLH, C-terminal to LisH motif is in italics.	MSS**LSRELVFLILQFLDEEKFKESVHKLEQESG***FFFNMK YFDEKAQAGEWDEVERYLSGFTKVDDNRYSMKIFFEIRK QKYLEALDRQDRAK*AVDILVKDLKVFSTFNEELYKEITQ LLTLDNFRENEQLSKYGDTKSARTIMMSELKKLIEANPL FREKLIYPNLKASRLRTLINQSLNWQHQLCKNPRPNPDI KTLFTDHACGPPNGARTPTQPTASLGVLPKATTFTPIGP HGPFPSSSTATSGLASWMSNPNMVTSPQAPVAVGPSVPV PPNQATLLKRPRTPPGSSSVVDYQTADSEQLIKRLRPVS QSIDEATYPGPTLRVPWSTDDLPKTLARALNEPYPVTSI DFHPSQQTFLLVGTKNGEITLWEVGSREKLATRSFKIWD NANCSNHLEAAFVKDSSVSINRVLWSPDGTLIGIAFTKH LVHTYTFQGLDLRQHLEIDAHVGGVNDLAFSHPNKQLCV VTCGDDKMIKVWDAVTGRKLYNFEGHDAPVYSVCPHHKE NIQFIFSTAVDGKIKAWLYDHLGSRVDYDAPGHSCTTMM YSADGTRLFSCGTSKEGESFLVEWNESEGAIKRTYSGLR KKGSGVVQFDTTQNHFLAVGDEHLIKFWDMDSTNMLTSC DAEGGLLNLPRLRFNKEGSLLAVTTVNGIKILANADGQK LLKTMENRTFDLPSRAHIDAASATSSPATGRMERIERTS SANTVSGINGVDPAQSSEKLRLSDDLSEKTKIWKLTEIT DSIQCRCITLPENAAEPASKVSRLLYTNSGVGLLALGSN AVHKLWKWNRSEQNPSGKATASVHPQRWQPTSGLLMTND ITDINPEEAVPCIALSKNDSYVMSASGGKVSLFNMMTFK VMTTFMPPPPASTFLAFHPQDNNIIAIGMEDSTIHIYNV RVDEVKTKLKGHQKRITGLAFSSTQNILVSSGADAQLCV WNTETWEKRKSKTIQMPVGKTVSGDTRVQFHSDQLHILV VHETQLAIYDAYKLERQYQWVPQDALSAPILYATYSCNR QLIYATFSDG

Entry	Sequence Description	Annotated Peptide Sequence
191	The amino acid sequence of SEQ ID 460. The conserved G-protein beta WD-40 repeat domains are underlined.	MAKDEEEFRGEMEERLVNEEYKIWKKNTPFLYDLVITHA LEWPSLTVQWLPDREEPPGKDYSVQKMILGTHTSDNEPN YLMLAQVQLPLEDAENDARQYDDERGEIGGFGCANGKVQ VIQQINHDGEVNRARYMPQNPFIIATKTVSAEVYVFDYS KHPSKPPQDGGCHPDLRLRGHNTEGYGLSWSPFKHGHLL SGSDDAQICLWDINVPAKNKVLEAQQIFKVHEGVVEDVA WHLRHEYLFGSVGDDRHLLIWDLRTSATNKPLHSVVAHQ GEVNCLAFNPFNEWVLATGSADRTVKLFDLRKISSALHT FSCHKEEVFQIGWSPKNETILASCSADRRLMVWDLSRID EFQTPEDALDGPPELLFIHGGHTSKISDFSWNPCEDWVI ASVAEDNILQIWQMAENIYHDEEDDMPPEEVV
192	The amino acid sequence of SEQ ID 461. The conserved G-protein beta WD-40 repeat domains are underlined.	MSPGVKQTGSQKFESGHQDVVHDVTMDYYGKRIATCSAD RTIKLFGLNASDTPSLLASLTGHEGPVWQVAWAHPKFGS MLASCSYDGRVIIWREGQQENEWSQVQVFKEHEASVNSI SWAPNELGLCLACGSSDGSITVFTCREDGSWDKTKIDQA HQVGVTAVSWAPASAPGSLVGQPSDPIQKLVSGGCDNTA KVWKFYNGSWKLDCFPPLQMHTDWVRDVAWAPNLGLPKS TIASCSQDGKVVIWTQGKEGDKWEGRILNDFKIPVWRVN WSLTGNILAVADGNNSVTLWKEAVDGDWNQVTTVQ
193	The amino acid sequence of SEQ ID 462. The conserved G-protein beta WD-40 repeat domains are underlined.	MSSGVKQTGSQKFESGHQDVVHDVTMDYYGKRIATCSAD RTIKLFGMNTSDTPTLLASLTGHEGPVWQVAWAHPKFGS MLASCSYDRRVIIWREGQQENEWSQVQVFKEHEASVNSI SWAPHELGLCLACGSSDGSITVFTGREDGSWDKTKIDQA HQVGVTAVSWAPASAPGSLVGQPSDPVQKLVSGGCDNTA KVWKFYNGSWKLDCFPPLQMHTDWVRDVAWAPNLGLPKS TIASCSQDGRVVIWTQGKEGDKWEGKILNDFKTPVWRIS WSLTGNILAVADGNNNVTLWKEAVDGEWNQVTTVQ
194	The amino acid sequence of SEQ ID 463. The conserved G-protein beta WD-40 repeat domains are underlined.	MKKRSRPSNGHLSTAAKNKSRKTAPITKDPFFDSAHNRN KSKGKGKSRGKGEEIFSSDEDDDAIGRDAPAEEEEEIAE EERETADEKRLRVAKAYLDKIRAITKANEEDNEEEAGED EETEAERRGKRDSLVAEILQQEQLEESGRVQRQLASRVV TPSKLVECRVVKRHKQSVTAVALTEDDLRGFSASKDGTI IHWDVETGASEKYEWPSQAVSVSSSNEVSKTQKGKGSKK QGSKHVLSMAVSSDGRYLATGGLDRYIHLWDTRTQKHIQ AFRGHRGAVSCLAFRQGTQQLISGSFDRTIKLWSAEDRA YMDTLYGHQSEILAVDCLRKERVLSVGRDHTLRLWKVPE ETQLVFRGHAASLECCCFINNEDFLSGSDDGSIELWSML RKKPVFMAKNAHGHAIVENLSEDTSTREEPDEEVTTRQL PNGNSIGNGMTNQMGITPSVESWVGAVTVCRGTDLAASG AGNGVVRLWAIENSSKSLRALHDIPLTGFVNSLTFARSG RFLIAGVGQEPRLGRWGRIQAARNGVTLCPIELS
195	The amino acid sequence of SEQ ID 464. The conserved G-protein beta WD-40 repeat domains are underlined.	MAATFGTINTATSPHNPNKSFEIVQPPNDSISSLSFSPK ANYLVATSWDNQVRCWEVLQTGASMPKAAMSHDQPVLCS TWKDDGTAVFSAGCDKQAKMWPLLTGGQPVTVAMHDAPI KDIAWIPEMNLLATGSWDKTLKYWDTRQSNPVHTQQLPE RCFALSVRHPLMVVGTADRNLIIFNLQNPQTEFKRISSP LKYQTRCVAAFPDKQGFLVGSIEGRVGVHHVEEAQQSKN FTFKCHRDSNDIYAVNSLNFHPVHQTFATAGSDGAFNFW DKDSKQRLKAMARSNQPIPCSTFNSDGSLYAYAVSYDWS KGAENHNPATAKHHILLHVPQESEIKGKPRVTTSGRK
196	The amino acid sequence of SEQ ID 465. The conserved G-protein beta WD-40 repeat domains are underlined.	MVVMDKGTHQTNEDESESEFIDEDDVIDEISIDEEDLPD ADVEGEDVQEDNKRSEPDENSSSLDDAIHTFEGHEDTLF AVACSPVDATWVASGGGDDKAFMWRIGHATPFFELKGHT DSVVALSFSNDGLLLASGGLDGVVRIWDASTGNLIHVLD GPGGGIEWVRWHPKGHLVLAGSEDYSTWMWNADLGKCLS VYTGHCESVTCGDFTPDGKAICTGSADGSLRVWNPQTQE SKLTVKGYPYHTEGLTCLSISSDSTLVVSGSTDGSVHVV NIKNGKVVASLVGHSGSIECVRFSPSLTWVATGGMDKKL MIWELQSSSLRCTCQHEEGVMRLSWSLSSQHIITSSLDG IVRLWDSRSGVCERVFEGHNDSIQDMVVTVDQRFILTGS DDTTAKVFEIGAF

Entry	Sequence Description	Annotated Peptide Sequence
197	The amino acid sequence of SEQ ID 466. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MPVFRTAFNGYAVKFSPFVETRLAVATAQNFGIIGNGRQ HVLELTPNGIVEVCAFDSSDGLYDCTWSEANENLVVSAS GDGSVKIWDIALPPVANPIRSLEEHAREVYSVDWNLVRK DCFLSASWDDTIRLWTIDRPQSMRLFKEHTYCIYAAVWN PRHADVFASASGDCTVRIWDVREPNATIIIPAHEHEILS CDWNKYNDCMLVTGSVDKLIKVWDIRTYRTPMTVLEGHT YAIRRVKFSPHQESLIASCSYDMTTCMWDYRAPEDALLA RYDHHTEFAVGIDISVLVEGLLASTGWDETVYVWQHGMD PRAC
198	The amino acid sequence of SEQ ID 467. The conserved G-protein beta WD-40 repeat domains are underlined.	MDSRNRRSRLNLPPGMSPSSLHLETTAGSPGLSRVNSSP STPSPSRTTTYSDRFIPSRTGSRLNGFALIDKQPQPLPS PTRSAAEGRDDASSSSASAYSTLLRNELFGEDVVGPATP ATPEKSTGLYGGSRDSIKSPMSPSRNLFRFKNDHGGNSP GSPYSASTVGSEGLFSSNVGTPPKPARKITRSPYKVLDA PALQDDFYLNLVDWSSNNVLAVGLGTCVYLWSACTSKVT KLCDLGVNDSVCSVGWTPQGTHLAVGTNIGEVQIWDTSR CKKVRTMGGHCTRAGALAWSSYILSSGSRDRNILHRDIR VQDDFIRKLVGHKSEVCGLKWSYDDRELASGGNDNQLLV WNQQSAQPLLRFNEHTAAVKAIAWSPHQHGILASGGGTA DRCLRFWNTATDTRLNCVDTGSQVCNLVWCKNVNELVST HGYSQNQIMVWRYPSMSKLATLTGHTLRVLYLAISPDGQ TIVTGAGDETLRFWSIFPSPKSQSAVHDSGLWSLGRTHI R
199	The amino acid sequence of SEQ ID 468. The conserved G-protein beta WD-40 repeat domains are underlined.	MEKKKVVVPIVCHGHSRPIVDLFYSPVTPDGLFLISASK DSSTMLRNGETGDWIGTFEGHKGAVWSCCLDNRALRAAS GSADFSAKIWDALTGDELHCFVHKHIVRACAFSESTSLL LTGGHEKILRIFDLNRPDAPPKEVDNSPGSIRTVAWLHS DQTILSSNSDAGGVRLWDLRTEKIVRVLETKSPVTSAEV SQDGRYITTADGNSVKFWDANHFGMVKSYTMPCMVESAS LEPTMGNMFVAGGEDMWVRLFDFHTGEEIACNKGHHGPV HCVRFAPGGESYSSGSEDGTIRIWQTLNMNSEENESYGV NGLSGKVRVGVDDVVQKVEGFQITADGHLNDKPEKPNP
200	The amino acid sequence of SEQ ID 469. The conserved G-protein beta WD-40 repeat domains are underlined.	MERYSQGTQKKSEIYTYEAPWQIYGMNWSVRKDKKFRLG IGSFLEEYNNRVEIIELDEESGEFKSDPRLAFDHPYPTT KIMFVPDKECQRPDLLATTGDYLRIWQVCEDRVEPKSLL NNNKNSEFCAPLTSFDWNDADPKRIGTSSIDTTCTIWDI EKEVVDTQLIAHDKEVYDIAWGEVGVFASVSADGSVRVF DLRDKEHSTIIYESSQPETPLLRLGWNKQDPRFIATILM DSCKVVILDIRFPTLPVAELQRHQASVNTIAWAPHSPCH ICTAGDDSQALIWELSSVSQPLVEGGGLDPILAYTAAAE INQLQWSSMQPDWVAIAFSNEVQILRV
201	The amino acid sequence of SEQ ID 470. The conserved G-protein beta WD-40 repeat domains are underlined.	MQSENNLDESLHLREVQFLQGHTDTVWAVAWNPVTGIDG APSMLASCSGDKTVRIWENTHTLNSTSPSWACKAVLEET HTRTVRSCAWSPNGKLLATASFDATTAIWENVGGEFECI ASLEGHENEVKSVSWSASGMLLATCGRDKSVWIWDVQPG NEFECVSVLQGHTQDVKMVQWHPNRDILVSASYDNSIKV WAEDGDGDDWACMQTLGNSVSGHTSTVWAVSFNSSGDRM VSCSDDLTLMVWDTSINPAERSGNAGPWKHLCTISGYHD RTIFSVHWSRSGLIASGASDDCIRLFS

Entry	Sequence Description	Annotated Peptide Sequence
202	The amino acid sequence of SEQ ID 471. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MKRAYKLQEFVAHASNVNCLKIGKKSSRVLVTGGEDHKV NMWAIGKPNAILSLSGHSSAVESVTFDSAEALVVAGAAS GTIKLWDLEEAKIVRTLTGHRSNCISVDFHPFGEFFASG SLDTNLKIWDIRRKGCIHTYKGHTRGVNSIRFSPDGRWV VSGGEDNIVKLWDLTAGKLMHDFKCHEGQIQCMDFHPQE FLLATGSADRTVKFWDLETFELIGSAGPETTGVRAMIFN PDGRTLLTGLHESLKVFSWEPLRCYDAVDVGWSKLADLN IHEGKLLGCSYNQSCVGVWVVDISRVGPYAAGNVSRTNG HNEAKLASSGHPSVQQLDNNLKTNMARLSLSHSTESGIK EPKTTTSLTTTEGLSSTPQRAGIAFSSKNLPASSGPPSY VSTPKKNSTSRVQPTTNFQTLSRPDIVPVIVPRSNSLRP ETTSDVKKEMNNFGRVVPSTVSTKSTDVIKSGSNRDESD KIDSINQKRMTGNDKTDLNIARAEQHVSSRLDNTNTSSV VCDGNQPAARWIGAAKFRRNSPVDPVVSPHDRSPTFPWS ATDDGVTCQPDRQVTAPELSKRVVEPGRARALVASWETR EKALTADTPVLVSGRPPTSPGVDMNSFIPRGSHGTSESD LTVSDDNSAIEELMQQHNAFTSILQARLTKLQVIRRFWQ RNDLKGAIDATGKMGDHSVSADVISVLIERSEIFTLDIC TVILPLLTRLLQSETDRHLTVAMETLLVLVKTFGDVIRA TISATPTIGVDLQAEQRLERCNLCYVELENIKQILVPLI RRGGAVAKSAQELSLALQEV
203	The amino acid sequence of SEQ ID 472. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MSTLEIEARDVIKIVLQFCKENSLHQTFQTLQNECQVSL NTVDSLETFVADINSGRWDVILPQVAQLKLPRKKLEDLY EQIVLEMIELRELDTARAILRQTQAMGFMKQEQPERYLR LEHLLVRTYFDPREAYHESSKEKRRSQIAQALASEVTVV PPSRLMALIGQSLKWQQHQGLLPPGTQFDLFRGTAAVKA DEEEMYPTTLAHTIKFGKQSHPECARFSPDGQYLVSCSV DGFIEVWDYISGKLKKDLQYQADDSFMMHDDAVLCVDFS RDSEMLASGSQDGKIKVWRIRTGQCLRRLERAHSQGVTS LSFSRDGSQLLSTSFDSTARIHGLKSGKALKEFRGHTSY VNDAIFTSDGGRVITASSDCTVKVWDVKTTDCIQTFKPP PPLKGGDVSVNSVHLFPKNSEHIVVCNKASSIYIMTLQG QVVKSFSSGKREGGDFVAACISPKGEWIYCVGEDRNIYC FSQQSGKLEHLMKAHDKDIIGVTPHPHRNLLVTYSEDST MKIWKP
204	The amino acid sequence of SEQ ID 473. The conserved G-protein beta WD-40 repeat domains are underlined.	MDIELEDQPFDLDFHPSAPIVAVALITGRLQLFRYVDIS SEPERLWTVTAHTESCRAARFINAGSSVLTASPDCSILA TNVETGQPVARLDNAHGAAINCLTNLTESTIASGDENGI IKVWDTRQNSCCNKFKAHEDYISDMEFVPDTMQLLGTSG DGTLSVCNLRKNKVHARSEFSEDELLSVALMKNGKKVVC GSQEGVLLLYSWGYFKDCSDRFVGHPHSVDALLKLDEDT VLTGSSDGIIRVVSILPNKMIGVIGEHSSYPIERLAFSH DRNVLGSASHDQILKLWDIHYLHEDDEPETNKQEAVNDE NVDMDLDVDTEKRPRGSKRKKRAEKGQTSSQKQSSDFFA DI
205	The amino acid sequence of SEQ ID 474. The conserved G-protein beta WD-40 repeat domains are underlined.	MDRIQQIPHTCVARKINLPLGMSKESLALNLPANLAPTM SPPSITYSDRFIPSRKASNFEEFALPDKTSPSPNSAGGQ SSSTNGEGRDDACAAYSALLRTELFPATPDKTEGCRRPV IGSPSGNVFRFKSQQCKSQSPFSLCPVGEDGDLSETGAV ARKTTRKIPRSPFKVLDAPALQDDFYLNLVDWSSHNILA VGLSACVYLWSASSSKVTKLCDLGLDDNVCSVAWTQRGT YLAVGTNNGGVQIWDAAHCKQVRTMEGHCTRVGTLAWNS HILSSGGRDRNILQRDIRAQDDFVSKFSGHKSEVCGLKW SYDNRELASGGNDNQLFVWNQQSQQPVLKYNEHTAAVKA IAWSPHQHGLLASGGGTADRCIRFWNTATNTSLNCVDTG SQVCNLVWSKNVNELVSTHGYSQNQIIVWRYPTMSKLAT LTGHTLRVLYLAISPDGQTIVTGAGDETLRFWNVFPSSK TQQNTIRDMGVWSSGRTHIR
206	The amino acid sequence of SEQ ID 475. The conserved G-protein beta WD-40 repeat domains are underlined.	MAGGQGEGEEKVDKLSMELTEDVMKSMEIGAVFKDYNGK INSLDFHRTNNYLVTASDDEAIRLFDTASATWQKTSYSK KYGVDLICFTNHQTSVLYSSKNGWDESLRHLSLMDNKYL RYFKGHHDRVVSLCMSPKGECFMSGSLDRTVLLWDLRID KCQGLIRVRGRPAVAYDEQGLVFAISNEGGLIKMFDARL YDKGPFDTFVVEGDKSEASGIKFSNDGKLILLSTMDSNI HVLDAYQGTTVHSFSVEAVPNGGEAVPNGGTLEASFSPD GKFVISGSGNGNIHAWSVNSGKEVACWTTEGVIPAVVKW APRRLMFASGSSVLSLWVPDLSKLASLTGSNSNSAY

Entry	Sequence Description	Annotated Peptide Sequence
207	The amino acid sequence of SEQ ID 476. The conserved G-protein beta WD-40 repeat domains are underlined.	MHRVGSTGNTSNSSRPRREKRLTYVLNDANDSRHCSGIN CLVISKLSLLGGNDYLFSGSRDGTLKRWELADDSAVCSA TFESHVDWVNDAVLTGETLVSCSSDTTLKTWRPFSDGVC TRTLRQHSDYVTCLAAASKNSNIVASGGLGREVFIWDIE AAMAPVSRTSEAMDDDTSNGVLSSGNSVLSTTVRSTNAT NSASLHTSQLQGYTPIAAKGHKESVYALAMNDVGTLLVS GGTEKVVRVWDPRSGAKQMKLRGHTDNVRALILDSTGRF CLSGSSDSIIRLWDLGQQRCVHSYAVHTDSVWALASTPN FSHVYSGGRDLSLYLTDLTTRESLLLCMEKHPLLRLTLQ DDSIWVATTDSSLHRWPAEGQNPPKMFQRGGSFLAGNLS FTRARACLEGSAPVPVNTQPSFVIPGSPGIVQHEILNNR RHVLTKDAEGTVKLWEITRGAVLDDYGKVSFEEKKEELF EMVSIPAWFTMDTRLGSMSVHLDTPQCFTAEMYAVDLNV PDAPEEQKINLAQETLRGLLAHWLSRRRQRLATQASANG DFPAGQENALRNHISSRIDVHDDAETHIAGILPAFDFST TSPPSIITEGSQGGPWRKKITDLDGTEDEKDFPWWCLEC VLHGRLSPRESLKCSFYLHPYEGTTVQVLTQGKLSAPRI LRIQKVINYVLEKMVLDRPLDSSNSETTFTPGLSGNQSH AAVVGDGSLRSGARVWQQKAKPLVEILCNNQVLSPDMSL ATVRTYIWKKPDDLYLYYRLVQNR
208	The amino acid sequence of SEQ ID 477. The conserved G-protein beta WD-40 repeat domains are underlined.	MMKGKTIQMQAAHQNHDGETSVACVLWDWHAKHLITAGA DNTILIHSYPSSSSSKPITLRHHKNAVTALAINSNVRSL ASGSVDHSVKLYSYPGGEFQSNVTRFTLPIRSLAFNKSG ELLAAAGDDEGIKLISTIDNSIARVLKGHNGPVTSISFD PKNEFLASSDSDGTVIYWELSTGKPVHTLKKIAPNTTSN PTSLNQISWRPDGEMLAVPGRKSEVSMYDRDTAEKLFSL KGGHSDTICSLAWSPNGKYIATAGTDRQVMVWDADRRQD IDKQRFDNPICSVAWKPSDNALAVIDVLGRFGVWESPIA SHMKSPADGAERYDNMEDEEPLMARYEEELEDSVSGSLN EIINDDDDDDEMGKIPRKILQKKPSVKVEKGKEESNAKA FKSGQDSFKLKSAMQEAFQPGATQRQSGKRNFLAYNMLG SVITFDNDGFSHIEVDFHDIGKGCRVPSMTDYFGFTMAS LSESGSVFGSPQKGEKNPSTLMYRPFSSWANNSEWSMRF PMGEEVKAVALGSGWVAAVTSLNFLRVFSEGGLQKFVLS MDGPVVTAAGYENLLVVVSHASNPLLSGDQVLSFTVYDI SQKTCPLSGRLPLSPGSHLTWLGFSEEGLLSSYDSEGNL RVFTNDYNGCWVPIFSAARERKSETESIWMVGLNSTQVF CVVCKLPDTYPQVAPKPVLSVLNLSLPLACSDLGADDLE NEYLRGSLLLSQMQKKAEDAVACGRESNMEEDSIFKMEA ALDRCLLRLIANCCKGDKLVRATELARLLSLEKSLQGAI KLVSAMKLPMLAERFNTILEEKILQENMETISCRRLTSE AQDMDTPISISVKQVSYGANLGDSPFLPNRQVEPKHSTP VFSKPDTKIEVDTSEAIAKGCDAQNGNIKSGDAEVQPAS HNDSIQKPSNPFAKASNTSANQAVQRNASLLSSIKQMKT ATENEGKRKERARSGSLPQKPAKQSKIS
209	The amino acid sequence of SEQ ID 478. The conserved G-protein beta WD-40 repeat domains are underlined.	MKQKRKGHQVDDPKYSVQTPQEDDTPNESGPASEEVESS DEEGGNSSNIEDDIIYSSSEEDPVVSSDYEEDEDAESDA EGVTAEQELEGDIDNALQNYMGTLTVLSNFHGENLKNAE GEDTSGDDDDEEEMPKRAEESDSPEDENDERPKRAEESD FSEDEDEERPKRAEESDSSEDEVPSRNTVGDVPLRWYKD EQHIGYDIKGKKIKKQPKKDQLDSFLASTDDSSDWRKVY DEYNDEEVELTKDEIKFISRLRKGTIPHADVNPYEPYVD WFDWKDKGHPLSNAPEPKRRFIPSKWEAKKVVKLVRAIR KGWITFQKAEEKPRFYLMWGDDLKPSEKMANGLSYIPAP KPKLPGHEESYNPPPEYIPTQEEINSYQLMYEEDRPKFI PKRFDSLRNVPAYDRFLSEIFERCLDLYLCPRTRKKRIN IDPESLIPKLPKPKDLQPFPSICFLEYKGHTGAVSCISP ESSGQWLASGSKDGTVRIWEVETARCLKVWDIGRPIQHI AWNPVSQLSILAVAVDEEVLVLNTGLGSEDSQEKVAELL HVKSKPVSADDLGDNTSLTKWIKHEKFDGIKLTHLKPVH LISWHHKGDYFATVAPDGNTRAVLVHQLSKQQTQNPFKK MQGRVVHVLFHPSRAIFFVATKTHVRVYDLVKQQLVKRL VTGLHEVSSMAVHHKGDNLLVGSKEGKVCWFDMDLSTQP YKTLKNHSKDIHSVAFHDSYPLFASCSDDCKAYVFYGLV YSDLLQNPLIVPLKVLQGHQSVNGMGVLDCQFHPKQPWL FTAGADSVVKLYCN

Entry	Sequence Description	Annotated Peptide Sequence
210	The amino acid sequence of SEQ ID 479. The conserved G-protein beta WD-40 repeat domains are underlined.	MMSLKRGFEESLVPAKRQKTELSTVTYGDGPRRTSSLES PIMLLTGHHAAIYTMKFNPTGTVIASGSHEREIFLWNVH GDCKNFMVLKGHKNAVLDLHWTTDGCQIISASPDKTLRA WDVETGKQIKKMAEHSSFVNSCCPSRRGPPLVVSGSDDG TAKLWDLRHRGAIQTFPDKYQITAVGFSDAADKIYSGGI DNEIKVWDLRRGEVTMRLQGHTDTITGMQLSSDGSYLLT NSMDCSLRIWDMRPYAPQNRCVKILTGHQHNFEKNLLKC SWSSDGSKVTAGSADRMVYIWDTTTRRILYKLPGHTGSV NETGFHPTQPIIGSCSSDKQIYLGEIEPNVGYQAVI
211	The amino acid sequence of SEQ ID 480. The conserved G-protein beta WD-40 repeat domains are underlined.	MEFSDTYKHTGPCCFSPDARYLAIAVDYRLVIRDVVTLK VVQLYSCMDKISNIEWALDSEYILCGLYKRAMVQAWSLS QPEWTCKIDEGPAGIAHARWSPDSRHIITTSDFQLRLTV WSLVNTACIHIQWPKHASKGVSFTQDGKFAAIATRRDCK DYVNLLSCHTWEVMGTFTVDTIDLADLEWSPNDSAIVVW DSPLEYKVLIYSPDGRCLFKYQAYDSWLGVKTVAWSPCS QFLAVGSYDQTLRTLNHLTWKPFAEFVHVSTVRGPASAV VFKEVEEPWNLDVSGLHLNDDNAHDIQDGKPAEGHSRVR YKVVEFPVNVSSQKHPVDKPNPKQGIGLLAWSRDSQYLF TRNDNMPTALWIWDICRLELAALLIQKEPIRAAAWDPVY PRVALCTGSSHLYMWTPSGACCVNIPLPQFVVSDLKWNP DGTSMLLKDRESFCCTFVPMLPEFNDDETNEE
212	The amino acid sequence of SEQ ID 481. The conserved G-protein beta WD-40 repeat domains are underlined.	MAKLIETHSCVPSTERGRGILIAGDAKTNSIIYCNGRSV IMRNLDNPLEASVYGEHSYPATVARFSPNGEWVASGDTS GTVRIWGRGSDHTLKYEYKALAGRIDDLEWSADGQRIVV CGDSKGKSMVRAFMWDSGTNVGEFDGHSRRVLSCSFKPT RPFRVATCGEDFLVNFYEGPPFRFKTSHRDHSNYVNCVR FAPDGSKFITVGSDRKGVIFDGKMGEKIGELSKEGGHTG SIYAASWSPDSKQVLTVSADKSAKIWEISETGNGTVKKT LTFGSQGGADDMLVGCLWLNDYLITVSLGGIVSLLSAVD PDKPPKTISGHMKSINAIALSLQSGQSEVCSSSYDGVIV RWILGVGYAGRVERKDSTQIKCLATIEGELVTCGFDNKV RRVPLLSEQHKESEPIDIGAQPKDLDVAVGCPELTFVST DAGIIIIRASKIVSTTNVGYAVTAAAISPDGTEAVVGGQ DGKLRVYSIKGDTLLEESVLERHRGPINAIRFSPDGSMF ASGDLNREAVVWDRITREVKLKNMVYHTARINCIAWSPD SSKVATGSLDTCILIYEVGKPASSRITIKGAHLGGVYGL AFSDQSTVISAGEDACVRVWSLP
213	The amino acid sequence of SEQ ID 482. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MPQPSVILATAGYDHTVRFWEATSGRCYRTLQYPDSQVN HLEITPDKQYLAAAGNPHIRLFEVNSNNPQPVISYDSHT NNVTAVGFQCDGKWMYSGSEDGTVKIWDLRAPGFQREYE SRAAVNTVVLHPNQTELISGDQNGNIRVWDLNANSCSCE LVPEDTAVRSLTVMWDGSLVVAANNHGTCYVWRLMRGTQ TMTNFEPLHKLQAHNSYILKCLLSPEFCEHHRYLATTSS DQTVKIWNVDGFTLERTLTGHQRWVWDCVFSVDGAFLVT ASSDSTARLWDLSTGEAIRTYQGHHKATVCCALHDGTDG ASC

Entry	Sequence Description	Annotated Peptide Sequence
214	The amino acid sequence of SEQ ID 483. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold. The coatomer WD associated region is in bold/italics.	MLTKFETKSNRVKGLSFHPKRPWILASLHSGVIQLWDYR MGTLIDKFDEHDGPVRGVHFHKTQPLFVSGGDDYKIKVW NYKMRQCLFTFVGHLDYIRTVHFHNEYPWIVSASDDQTI RLWNWQSRVCISVLTGHNHYVMSASFHPKEDLVVSASLD QTVRVWDISGLRKKTVSPADDLSRLAQMNTDLFGGGDVV VKYVLEGHDRGVNWAAFHTSLPLIVSGADDRQVKLWRMN DTKAWEVDTLRGHTNNVSCVIFHARQDIIVSNSEDKSIR VWDMSKRTSVQTFRREHDRFWILAAHPEMNLLAAGHDSG MIVFKLERERPAYVVYGGSLLYVK***DRYLRTYEFATQKDN PLIPIRKPGSIGPNQGPRSLSYSPTENAILICSDADGGA YELYAVPKDSHGRSDTVQEAKKGLGGSAVFVARNRFAVL DKNHNQVTIKNLKNEVTKKFDLPVTADALFYAGTGNLLC RSEDSVFLFDMQQRTVLGEIQTPNVRYVVWSNDMENVAL LSKHTIIIASKKLSSTCSLHETIRVKSGAWDDNGIFMYS TLNHIKYCLPNGDSGIIKTLDVPVYITKVSGKSLYCLDR DGKNRVIQIDITECLFKLALSKKKYDYVINMIRNSQLCG QAIIAYLQQKGFPEVALHFVRDERTRFNLAVESGNIEIA VASAKEIDEKDHWYRLGVEALRQGNAGIVEYAYQRTKNF ERLSFLYLITGNLDKLSKMLRIAEMKNDVMGQFHNALYL GDIQERIKILEESGHLHLAYATASLHGLADIADRLAADL GGN***IPVLPPGKKSSLLMPPAPILHGGDWPLLRVTKGIFE GGLENSTSAAYEEEDEEAAADWGEDIDIENIEGENGEAT VLDDQEVKGGEDDEGGWDMEDLELPPDVAAANVGTNQKT LFVAPTLGMPVSQIWMQKSSLAGEHAAAGNFETALRLLT RQLGIKNFSPLKPLFLELYMGSHTFLPSFASVPAFSLAL QRGWSESASPNIRGPPALVYRLSVLEEKLTVAYRATTEG RFSEALRLFL
215	The amino acid sequence of SEQ ID 484. The conserved G-protein beta WD-40 repeat domains are underlined.	MDLLQNYQDDSEDSNPELRNHPPLEDATATSAPAGVENE TSSSPDSSPLRLALPAKSCAPDVDETLMALGVPGSEKKN NHNKPIDPTQHSVTFNPSYDQLWAPLYGPAHPYAKDGIA QGMRNHKLGFVEDSAIEPFMFDEQYNTFHRYGYAADPSA SLGSHIVGDLESLKKNDGASVYNLPKREHKRQKLEKKMI QKDENEEEEKEVGEEVDNPSTEEWLKKNRKSPWAGKKEG LQTELTEEQKKYAQEHAEKKGDREKGEKVEIVDKTTFHG KEERDYQGRSWIDPPKDAKATNDHCYIPKRWVHTWSGHT KGVSAIRFFPKYGHLLLSAGMDTKVKIWDVFNSGKCMRT YMGHSKAVRDISFSNDGSRFLSAGYDRNIKLWDTETGKV ISTFSTGKIPYVVKLHPDEDKQNVLLAGMSDKKIVQWDM NSGEITQEYDQHLGAVNTITFVDNNRRFVTSSDDKSLRV WEFGIPVVIKYISEPHMHSMPSISLHPNTNWLAAQSLDN QILIYSTRERFQLNKKKRFAGHIAAGYACQVNFSPDGRF VMSGDGEGRCWFWDWKTCKVFRTLKCHDNVCIGCEWHPL EQSKVATCGWDGMIKYWD
216	The amino acid sequence of SEQ ID 485. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MARKGLGTDPAIGSLMSSKKRKEYKVTNRFQEGKRPLYA IAFNFIDARYHNIFATAGGTRVTIYQCLEGGAISVLQAY VDDDKDESFYTLSWACDVNGSPLLVAGGHNGIIRVLDVA NEKVHKSFVGHGDSVNEIRTQALKPSL**ILSASKDESVRL** WNVQTGICILIFAGAGGHRNEVLSVDFHPSDVYRIASCG MDNTVKIWSMKEFWTYVEKSFTWTDLPSKFPTKYVQFPV FIAAVHSNYVDCTRWLGNFILSKSVDNEVVLWEPYSKEQ STSDGVVDILQKYPVPECDIWFIKFSCDFHYNSMAVGNR EGKVYVWELQSSPPNLIARLSHAHCKNPIRQTAISHDGS TILCCCDDGSMWRWDVVQ

Entry	Sequence Description	Annotated Peptide Sequence
217	The amino acid sequence of SEQ ID 486. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MESGAGGSVGARVPSAKPEMLQQPPYSNGDDDNDMERGT APVPSSNPNTVSKWELDKDFLCPICMQTMKDAFLTACGH SFCYMCIMTHLNNKSNCPCCSLYLTNNQLFPNFLLNKLL KKTSACQMASTASPVENLCLSLQQGAEVSVKELDFLLTL LAEKKRKMEQEEAETNMEILLDFLQRLRQQKQAELNEVQ ADLHYIKDDILALEKRRLELSRARERYSRKLHMLLDDPM DTTLGHAAIDDGNNVRTAFVRGGQGDAISGKFQQKKAEI KAQASSQGMQKRANFCHSDSQVLPTLSGLTIARKRRVLA QFDDLQECYLQKRRRWATQLRKQCDGGLRKERDGNSISR EGYHAGLEEFQSILTTFTRYSRLRVISELRHGDLFHSAN IVSSIEFDRDDELFATAGVSRRIKVFDFATVVNEPADVH CPVVEMSTRSKLSCLSWNKCIKSQIASSDYEGIVTVWDV NTRQSVMMYEEHEKRAWSVDFSRTEPTRLISGSDDGKVK VWCTRQETSVLNIDMKANICCVKYNPGSSYYVAVGSADH HIHYYDLRNPSVPLYEFNGHRKTVSYVKFISTNELASAS TDSTLRLWDVRDNCLVRTFKGHTNEKNFVGLTVNSEYIA CGSETNGVFVYHKAISKPAAWHQFGSPDLDDSDDDTSHF ISAVCWKSESPTMLAANSQGTIKVLVLAP
218	The amino acid sequence of SEQ ID 487. The conserved G-protein beta WD-40 repeat domains are underlined.	MANYVDSKKNFKCVPALQQFYTGGPFRLSSDGSFLVCAC NDEVKVVDLATGSVKNTLEGDSELIVALALTPDNKYLFS ASRSTQIKFWDLSSATCKRTWKAHNGPVADMACDASGGL LATAGADRSILVWDVDGGYCTHSFRGHQGVVTTVIFHPD PHCLLLFSGSDDATVRIWDLVAKKCISVLEKHFSTVTSL AISENGWNLLSAGRDKVVNIWDLRDYHCRATIPTYEPLE AVCVLPTGSRLVSVMNQSRALPENRKKSGAAPVYFLTVG ERGIVRIWYSEGALCLYEQKSSDAIISSDKDELKGGFVS AVLLPLTQGVMCVTADQRFLFYNLDESDEGKCDLKVSKR LIGYNEEIVDLKFLGDEEKFLAVATNLEQVRMYDLSSMT CVYELSGHTDIVLCLDTVVFSGHSLLASGSKDHTVRIWD TESKSCICVAAGHMGAVGAVAFSKKAKNFFVSGSSDRTI KVWSFASVLDFGGISKSIKLSSQAAVAAHDKDINSVAVA PNDSLICTGSQDRTARIWRLPDLVPVLVLRGHKRGVWCV EFSPVDQCVMTASGDKTIKIWALSDGSCLKTFEGHTASV LRASFLTRGTQFVSSGADGLLKLWTIKSNECIATFDQHE DKIWAMAVGKKTEMLATGGSDSLVNLWHDCTTTDEEEAL LKEEEAALKDQELLNALADTDYVKAIQLAFELRRPYKLL NVFTELYSKGHAQDQIQKVIRELGNEELRLLLEYVREWN TKPKFAHVAQFVLFQLFNVLPPKEIIEVQGISELLEGLI PYAQRHYSRIDRLMRSTFLLDYTLSSMSVLSPTETDLSS SNLLARTADPLHAQIDQFHPTHFPEPNLTPIQSLLDSGN TDSVEVTARRAKKKRVSGNDSEKTTVAEVKIGDMENAFD EPDVADQGSSRKHKPASSKKRKSIAVGNASIKRIASGNA VTIALQV
219	The amino acid sequence of SEQ ID 488. The conserved G-protein beta WD-40 repeat domains are underlined.	MESSCSSMNSNRHSTEKRCLRPLQKQGASMNKHSSDRFI PARGSIDLDVARFMVTQKQKDNNDIHALSPSPSPSKKAY QKEMADTLLKNAGAADNNCRILSFNGKSSTVSQGSQENV LANLSISRRARRYIPQSADRTLDAPDLLDDYYLNLLDWS STNVLSTALGNTVYLWDASNSSISELLIADEEEGPVTSV SWAPDGSQIAVGLNNSVVQLWDSQSNKKLRALKGHHDRV GALSWNGPILTTGGLDGIIINHDVRTRDHIVQTYKGHTQ EVCGLKWSPSGQQLASGGNDNLLYIWDKSMASHNPSSQY FHQLDEHCAAVKALAWCPFQTNLLASGGGTSDGSIKFWN TQTGACLNTVDTHSQVCSLLWNRHERELLSSHGLNQNQL TLWKYPSMVKITELTGHTARVLHMAQSPDGYTVASAAAD ETLKFWQVFGAPDASKKTKTKDTKGAFNMFHMHIR
220	The amino acid sequence of SEQ ID 489. The conserved G-protein beta WD-40 repeat domains are underlined.	MLDEIVADEEEEFNIWKKNTPLLYDVVITHALEWPSLTV QWLPDRHQSPTKDYSLQKMIVGTHTSGDEPNYLMIAEVQ MPLQYSEDGNVGGFESTEAKVHIIQQINHEGEVNRAQYM PQNSFIIATKTVSSDVYVFDYTKHSSNAPQERVCNPELI LKGHTNEGYSLSWSPLKEGQLLSGSNDAQICFWDINAAS GRKVVEAKQIFKVHEGAVEDVSWHLKHEYLFGSVGDDCH LLIWDTRTAAPNKPQHSVVAHESEVNSLAFNPFNEWLLA TGSADKTVKLFDLRKLSCSLHTFSNHTEEVFQIEWSPMN ETILASSGGDRRLMVWDLRRIGDEQTSEDAEDGPPELIF IHGGHTSKISDFSWNLHDDWLIASVAEDNILQIWQMAEN IYHDDADIL

Entry	Sequence Description	Annotated Peptide Sequence
221	The amino acid sequence of SEQ ID 490. The conserved G-protein beta WD-40 repeat domains are underlined.	MTKEDHGESRDEMGERMVNEEYKLWKKNTPFLYDLVITH ALEWPSLTVQWLPPSCKQQQDIIKDDDIDHPNTQMVILG THTSDNEPNYLILAEVQLHDGTEDEDGDGDVKRPQDKMK PGTSGGAMGKVRILQQINHQKEVNRARYMPQKPTIIATK TVNADVYVFDYSKHPSKPPQEGRCNPELRLQGHESEGYG LSWSPLKEGHLLSASDDAQICLWDITAATKAPKVVEANQ IFRYHDGPVEDVAWHAIHDHLFGSVGDDHHLLLWDIRND SEKPLHIVEAHQAEVNCLAFNPFNEWIVATGSADRTVAL HDIRKLDKVLHTCAHHMEEVFQIGWSPQNGAILASCGSD RRLMVWDLSRIGDEQNPEDAEEAPPELLFIHGGHTSKIS DFSWNPAEEWVIASVAEDNILQVWQMSEHIYNDDNDSPT A
222	The amino acid sequence of SEQ ID 491. The conserved G-protein beta WD-40 repeat domains are underlined.	MAMAMGDENAADPVEEFNIWKKNTPFLYDLVITHALEWP SLTVQWLPDRHQSSTADYSLQKMIVGTHTSEDEPNYLMI AEVQIPLQNSEDNIIGGFESTEAKVQIIQKINHEGEVNK ARYMPQNSFVIATKTVSSDVYVFDYSKHPSKAPQERVCN PELILKGHSNEGYGLSWSPLKEGYLLSGSNDAQICLWDI NAAFGKKVLEANQIFKVHEGAVGDVSWHLKHEYLFGSVG DDCHLLIWDMRTAAPNKPQQSVIAHQSEVNSLAFNPFNE WLLATGSMDKTVKLFDLRKLSCSLHTFSNHTDQVFQIEW SPMNETILASSGADRRLMVWDLARIGETPEDEEDGPPEL LFVHGGHTSKISDFSWNLNDDRVIASVAEDNILQIWQMA ENIYHDDEDML
223	The amino acid sequence of SEQ ID 492. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MGLFEPFRALGYITDGVPFAVQRRGIETFVTLSVGKAWQ IYNCAKLIPVLVGPQMDKKIRALACWRDFTFAATGHDIA VFRRAHQVATWSGHKAKVTLLLSFGQHVLSVDLEGCLFI WAVAEVNQNKPPIGQIQLGEKFSPSCIMHPDTYLNKVLI GSEEGTLQLWNVNTRKKLYEFKGWGSSIRCCVSSPALDV VGIGCSDGKIHVHNLRYDEEIVTFMHSTRGAVTALSFRT DGQPLLAAGGSSGVISIWNLEKKKLQSVIKDAHDSSVCS LHFFANEPVLMSSATDNSIKMWIFDTTDGEARLLKYRSG HSAPPMCIRYYGKGRHILSAGQDRAFRIFSVIQDQQSRE LSQGHVGKRAKKLKVKDEEIKLPPVIAFDAAEIRERDWC NVVTCHLDDPCAYTWRLQNFVIGEHILKPCLEDPTPVKS CSISACGNFAVLGTEGGWLERFNLQSGISRGTYIDIGEK RQCAHNGAVVGLACDATNTLLISGGYNGDIKVWDFKGRE LKFRWEIEVPLIKIVYHPGNGILATAADDMILRLFDVTA MRLVRIFVGHMDRVTDLCFSGDGKWLLSSSMDGTIRVWD IISSRQLNAMHMDSAVTALSLSPGMDMLATTHVGHNGIY LWANRMIYSKATDIEPFISGKQVVKVSMPTVSSKRESEE GDEKRTIVAESNVNKSDVSGSLIGDSYSAQLTPELVTLA LLPKAQWQSLVNLDIIKMRNKPIEPPKKPEKAPFFLPSL PTLSGERIFIPSSMNGDGDQDETRNDKTVFEARGKKLGG ESLSFMQLLQSCAKIKDFTTFTNYLKGLSPSAVDMELRL LQIVDNENISETEHSVELQGIGMLLDYFVNEVSCNNNFE FVQALIRLFLKIHGETIRCQVSLQEKARKLLEIQSSTWE RLDTSFQNARCMITFLSSSQF
224	The amino acid sequence of SEQ ID 493. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MIAAVCWVPKGVAKVLPDSAEPPTQEEIQELLKCNVVAE SDDNEDSDEESEEMDTETDKNTDAVAKALAAANALGSQS SDFQRQHKVDDIANGLKELDMDHYDDEDEGIDIFGSGSL GNCYYPANDMDPYLVEQDDDDEDEIEDMTIKPSDLIILS ARNEDDVSHLEVWIYEEETEEGGSNMYVHHDIILPAFPL SLAWLDCNLKGGEKGNFVAVGTMQPEIELWDLDVLDEVE PAVVLGGAVKDEASGKTTKLKKKKKNKQAVNFKEGSHTD AVLGLAWNMEYRNVLASASADKSVKIWDIVAEKCEHTMQ PHTDKVQAVAWNPNQATVLLSGSFDRSVIMMDMRAPTHS GIRWPVPADVESLAWDPHTDHSFMVSAEDGTVRGFDIRA AASTADFDGKPMFILHAHDKAVCAISYNPAAPSLLTTGS TDKMVKLWDITNNQPSCIASTNPNVGAVFSAAFSKNSPF LLATGGSKGILHVWDTLDNSEVARRFGKFRPQN

Entry	Sequence Description	Annotated Peptide Sequence
225	The amino acid sequence of SEQ ID 494. The conserved eukaryotic protein kinase domain is underlined.	MIMDENEFCDIFSLRKRLCLLSSQEGEEEEELEAMSQLD AGEFTVTGNEEVVAIAEDDVNTGILSQDLFSSQDYCTPS QPQDSTDLDSKDKAPCPLSPVKSTIQRKRCRPELLSNPP DSIQFSFQRLERVRSEESIQSSSQQLARVRSEVSSSDDF KTPKITASGQKNYVSQSALALRARVMSPPCIKNPYLDEN EELNEKIQRSTRRSPACVTPIQSGACLSRYRADFHELEE IGRGNFSRVYKALNRLDGCCYAVKCSQSELRLDTERKVA LMEVQSLAALGPHKNIVGYHTAWFENDHLYIQMELCDHN LTTANDRGILRTDTDFLEAVYQIAQALEFIHGRGVAHLD VKPENIYVRDGTYKLGDFGRATLINGTLHVEEGDARYMS REILNDNYEHLDKVDMFSLGATFFELLMRKQYPGSGKRI DRDTEIKIPILPGFSIYFQKLLQDLVSNDPGKRPSAKDV LKNPIFNKVRGAKEV
226	The amino acid sequence of SEQ ID 495. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MLAPALEMEPVEPQSLKKLSFKSLKRALDLFSPVHGQIA PPDPESKKMRISYKLNFEYGGGSGSEDQVPKRKESGAAQ NQGQQAAGASNALALPGPEGSKIPPMEKSQNALTVGPSL RPQGLNDVGLHGKGTAIISASGSSDRNLSTSAIMERLPS RWPRPVWHPPWKNYRVISGHLGWVRSIAFDPSNQWFCTG SADRTIKIWDLASGRLKLTLTGHIEQIRGLAVSSKHTYM FSAGDDKQVKCWDLEQNKVIRSYHGHLSGVYCLALHPTI DILLTGGRDSVCRVWDIRSKMQIFALSGHDNTVCSVFAR PTDPQ**VVTGSHDTTIKFWD**LRHGKTMTTLTNHKKSVRAM AQHPKENCFASASADNIKKFQLPRGEFLHNMLSQQKTII NTMAVNEEGVMATGGDNGSLWFWDWKSGHNFQQAHTIVQ PGSLESEAGIYALSYDLTGSRLVSCEADKTIKMWKEDEL ATPETHPLNFKPPKDIRRF
227	The amino acid sequence of SEQ ID 496. The conserved G-protein beta WD-40 repeat domains are underlined.	MEEAAKEQSAGSGKPKLLRYGLRSAAKPKEDKKEEQLHQ PPPPPPQQQAAPAPAPAATRSSTSGSAGGRDRRPQQQH AVDEKYARWKSLVPVLYDWLANHNLLWPSLSCRWGPQLE QATYKNRQRLYISEQTDGSVPNTLVIANCEVVKPRVAAA EHVSQFNEEARSPFIRKYKTIIHPGEVNRIRELPQNPNI VATHTDSPDVLIWDVESQPNRHAVYGATASRPNLILTGH QENAEFALAMCPAEPFVLSGGKDKTVVLWSIQDHITASA TDQTTNKSPGSGGSIIKKTGEGNEETGNGPSVGPRGIYC GHEDTVEDVAFCPSTAQEFCSVGDDSCLILWDARIGTNP VAKVEKAHNGDLHCVDWNPHDNNLILTGSADNSVNMFDR RNLTSNGVGSPVYKFEGHKAAVLCVQWSPDKPSVFGSSA EDGLLNIWDYERVDKKVDRAPNAPAGLFFQHAGHRDKIV DFHWNTADPWTMVSVSDDCDTAGGGGTLQIWRMSDLIYR PEEEVLAELENFKAHVLECSKA
228	The amino acid sequence of SEQ ID 497. The conserved G-protein beta WD-40 repeat domains are underlined.	MAKDEEEFRGEMEERLVNEEYKIWKKNTPFLYDLVITHA LEWPSLTVQWLPDREEPPGKDYSVQKMILGTHTSDNEPN YLMLAQVQLPLEDAENDARQYDDERGEIGGFGCANGKVQ VIQQINHDGEVNRARYMPQNPFIIATKTVSAEVYVFDYS KHPSKPPQDGGCHPDLRLRGHNTEGYGLSWSPFKHGHLL SGSDDAQICLWDINVPAKNKVLEAQQIFKVHEGVVEDVA WHLRHEYLFGSVGDDRHLLIWDLRTSATNKPLHSVVAHQ GEVNCLAFNPFNEWVLATGSADRTVKLFDLRKISSALHT FSCHKEEVFQIGWSPKNETILASCSADRRLMVWDLSRID EFQTPEDALDGPPELLFIHGGHTSKISDFSWNPCEDWVI ASVAEDNILQIWQMAENIYHDEEDDMPPEEVV
229	The amino acid sequence of SEQ ID 498. The conserved cyclin-dependent kinase inhibitor domain is underined.	MGKYMRKGKGVGEVAVMEVSQGSLGVRTRARTLAAASSQ KDHRRLGASKSVTTKHQSSAPPASPCVESSMHTCYLELR SRKLEKFSRCYHSAHGATSHGESKRSLSLSEPSRLAVSE EARVASDKSSHRVLQQQSSVAHSRNNSATFSHNAKPAKA AQRKERRDDDHTSARPSEAPHEDEDGMEVEASFGENVMD LDSRERRTRETTPSSYTRDVETMETPGSTTRPPSNAGRR RFQTEGGHGTRNQFHVPTTNEIEEFFAGAEQQEQRRFTD RYNYDPVSDSPLPGRFEWVRLRP
230	The amino acid sequence of SEQ ID 499. The conserved serine/threonine protein kinase domain is underlined, and the serine/threonine protein kinase active-site signature is in bold.	MQNMEENVQSSWSLHGNKEICARYEILKRVSSGTYLDVY RGRRKEDGLIVALKEVHDYQSSWREIEALQRLCGCPNVV RLYEVILEFLTSDLYSVIKSAKNKGENGIPEAEVKAWMI QILQGLANCHANW**VIHRDLKPSNMLI**SAYGILKLADFGS MSFLKRAIYEVEYELPQEDILADAPGERLMDEDDSVKGV WNEGEEDSSTAVETNFDDMAETANLDLSWKNEGDMVMQG FTSGVGTRWYRAPDFLYGATIYGKEIDLWSLGCILGELL ILEPLFSGTSNIDQLSRLVKVLGLQQKKNWPGCSNLPDY RKLCFPGDGSPVGLKNHVPNCSDNMFSILERLVCYDPAA RLNAKEIVENKYFVEDPYPVLTHELRVPSPLREENNFSE DWAKWKDMEVDSDLENIDEFNVVHSSDGFCIKFS

Entry	Sequence Description	Annotated Peptide Sequence
231	The amino acid sequence of SEQ ID 502. The conserved histone deacetylase family domain is underlined	MADVPESLQQEKDEQGTDKNCCDGKFQKEIDIDDMEEEY NESSIDDEEENLSDNVATNNMGTIPQGQACMAVTVEGIE HANSVGCGRNGREGSEEVTAAEDMGHVSIENIREQGRNR KSSEQLLALYEQEGLLEDDEDDDDVDWEPFEGVTVQMKW YCTNCTMANSDDSVHCDSCGEHRNSDILRQGFLASPYLP AESPSSSDVPDERLEESKCVMTTLTPSISPMIGVCCSSL QSERRTVVGFDERMLLHSEIQMETYPHPERPDRLRAIAA SLRAAGLFPGKCFSIPAREATCEELQTIHSLEHVNAVES TSCGMLSHLSPDTYANEHSSLAARLAAGLCADLAKAIMT GQAQNGFALVRPPGHHAGVKDSMGFCLHNNAAIAVSASR VVGAKKVLIVDWDVHHGNGTQEIFEADQSVLYISLHRHG EGFYPGSGAVTEVGSSKGEGYSVNIPWKCGGVGDNDYIF AFQHAVLPIAEQFEPDLTIISAGFDAAKGDPLGRCEVTP DGFAHMAQMLSCLSKGKMLVILEGGYNLRSISASATAVI KVLLGDNPKALPIDIQPSKGGLQTLLEVFEIQSKYWSSL KGHDQKLRSQWEAQYGSKKRKVIRKRHMHIVGGPVWWKW GRKRVVYYHWFARVSSRKHL
232	The amino acid sequence of SEQ ID 503. The conserved cyclophilin-type peptidyl-prolyl cis-trans isomerase family domain is underlined and the cyclophilin-type peptidyl-prolyl cis-trans isomerase signature is in bold.	MASGAGAAGVVEWHQKPPNPKNPVVFFDVTIGTIPAGRI KMELFADIVPRTAENFRQFCTGEYRKAGIPIGYKGCHFH RVIKDFMIQAGDFVKGDGSGCISIYGSKFEDENFIAKHT GPGLLSMANSGPNTNGCQFFLTCAKCDWLDNKHVVFGRV LGEGLLVLRKIENVQTGQHNRPKLPCVIAECGEM
233	The amino acid sequence of SEQ ID 505. The conserved G-protein beta WD-40 repeat domain is underlined.	MDHYYQDDFDYLVDDEMVDFADDVEDDVRTRRRSDIDSD SENDFDSNNKSPDTTALQAKRGKDIQGIPWNRLNFTREK YRETRLQQYKNYENLPRPRRSRNLDKECTNFERGSSFYD FRHNTRSVKATIVHFQLRNLVWATSKHNVYLMQNYSIMH WSSLKQKGEEVLNVAGPIIPSVKHPGSSPQGLTRVQVSA MSVKDNLVVAGGFQGELICKYLDKPGVSFCTKISHDENG ITNAVEIYNDASGATRLMTANNDLAVRVFDTEKFTVLER FSFPWSVNHTSVSPDGKLVAVLGDNADCLLADCKTGKTV GTLRGHLDYSFAAAWHPDGYILATGNQDTTCRLWDVRKL SSSLAVLKGRMGAIRSIRFSSDGRFMAMAEPADFVHLYD TRQNYTKSQEIDLFGEIAGISFSPDTEAFFVGVADRTYG SLLEFNRRRMNYYLDSIL
234	The amino acid sequence of SEQ ID 506. The conserved G-protein beta WD-40 repeat domains are underlined.	MDCSGDEEEEQFFESLEEMLSPSDSGSEAADNETGCRNA DARSKYEIWKRAPSSIQERRQRFLVRMGLANPSELGNQV NSTSAESTCSTETANIPNGIERLRENSGAVLRTAGSSGR KTHCKNVINIGLREGSVRSSSSSNGTPDVGEDNGEFGGT IFSRSGGTWECMCKIKNLDSGKEFVVDELGQDGLWNKLR EVGTDRQLTMDEFERSLGLSPLVQELMRRESGVAQADCN GVHHHDAEISSSKRRSWLKALKSAAYSMRRPKEDQSNYD SERSGRRSGSFDVPWGKPQWTKVRHYRKRYKEFTALYMG QEIEAHEGSIWTMKFSLDGRYLASAGQDCVIHVREVIES MRTFGADTPDLYASSAYFSMNGLQELVPLSIEDHANKMK RGKIIGSKKSSNSDCIVLPNKVFQLSEEPVCSFHGHLLD VFDLSWSPSQYLLSSSMDKTVRLWKLGHESCLKVFSHND IVTCIQFNPVDERYFISGSLDGKARIWSIPDRQVVDWSD LREMVTAVCYTPDGQGGLVGSIKGSCRFYNTSGNKLQLE NQLNVRSKKKKSSGKKITGFQFAPGGDSQKVLITSADSR VRVYNGSELVCKYKGFRNTCSQISASFAPNGQHFVCASE DSRVYIWNHESPRGSGARHEKSSWSHEHFLSQGVSVAIP WSGMKLQPPVWNSPEFMLGQRHNLLSLQGGKDVGCQNGL LSREAGEGQESETPLHYISQVSHSCGSQNMVDRDGQDDL SRYSACISDSRLSSFMAFPESPGNPDDLNSKVFFSDSSS KGSATWPEEKLPPTRKQSRSNSTSSHYDTLKTHLGNTIQ GQSGASAAVAWGLVIVTAGHGGEIRSFQNYGLPVRL

Entry	Sequence Description	Annotated Peptide Sequence
235	The amino acid sequence of SEQ ID 507. The conserved G-protein beta WD-40 repeat domain is underlined.	MPSIPAIGEFTVCEINRELLTTKDESDTQAKDAYAKILG LVFPPISFQIEEGFGSASRQQFDQDLDREDTIVTPSTSE GTNALQEGGLLLKGVSVLKNILASSFGPIFSPNDTKVLK KVELLQGISWHRHKHILAFISGSNQVTVHDFQDPEWRES SLLVSESQRGIEALEWRPNGGTTLSVACRGGICIWSASY PGSVAPVRSGVASFLGTSTRGSSVRWTLVDFLQIPGGKA VTALSWSPTGRLLASASREDSSFTIWDVAQGVGTPLRRG LGGISLLKWSPTGDYLFSAKPNGTFYLWETNTWTLEQWS SSGGCVISATWGPDGRMLFMAFSESTTLGSLHFAGRPPS LDAHLLPMELPEIGSITGGFGNIEKMAWDGCGERLAVSY TGGDLMYVGLIAIYDTRRTPFISASLVGFIRGPGEQVKP LAFAFHDKFKQGPLLSVCWSSGLCCTYPLIFRAH
236	The amino acid sequence of SEQ ID 508. The conserved G-protein beta WD-40 repeat domains are underlined.	MEEENAKHTEETRQVQVRFTTKLQPALRVPTTSIAIPAH LTRYGLSDIVNTLLGNDKPQPFDFLVESELVRTSLEKLL LIKGISAEKILNIEYILAVVPPKQEEPSLHDDWVSVVDG SYPNFIFSGSFDSIGRIWKGEGLCTHVLEGHRDAITSAA FIMPSDSSDSFINLATASKDRTLRLWQFKPNEHMTNGKM VRPYKLLKGHTSSVQTVSACPRRNLICSGSWDCSIKIWQ TAGEMDIESNAGSVKKRKLEDSTEQIISQIEASRTLEGH SQCVSSVVWLEKDTIYSASWDHSVRSWDVETGVNSLTVG CRKALHCLSIGGEGSALIAAGGADSVLRIWDPRMPGTFT PILQLSSHKSWITACKWHPKSRHHLISASHDGTLKLWDV RSKVPLTTLEAHKDKVLCADWWKEDCVISGGADSTLQIF SNLNLT
237	The amino acid sequence of SEQ ID 509. The conserved RING-type zinc finger is underlined.	MNRLRSKRNHILELRLGQSEPEKEATLASNRSRGTNAPI VVEDDDDVVVSSPRSFALARSSVSQRSSRIPIVNEEDLE LRLGLAVTGRTSAEHNPRRRHGRVPPNKPIVLCDDAGEA DQSSSKKRRTGQQLSSDVQSDESKEVKLTCAICISTMEE ETSTICGHIFCKKCITNAIHRWKRCPTCRKKLAINNIHR IYISSSTG
238	The amino acid sequence of SEQ ID 510. The conserved G-protein beta WD-40 repeat domains are underlined and the splicing factor motif is in bold.	MEEPPPPAVLPSSEDTSIVSSHSFVNAPPTVPVGLDASI PQISTPGINQPGLTIPVPPEAAPLTASLVAASAGMPPAV VPSFVRPAIVAHPSVMPPPSMPLAALPMPVASAVPVAAP HFPPSTPNDNSITPSMPVPTPIVASSSVPPSVTIPGIAP *LPFIAPIPVPSSRPVAPSPFMPPARPLGASVSVAMDVDN* TDEQDQDADNKGESPSSSPDHPEDPSAAEYEITEESRKV RERQEQAIQELLLRRRAYALAVP**TNDSSVRARLRRLNEP** **ITLFGEREMERRDRLRALMAKLDAEGQLEKLMKVQEEE**E AAANVDAEEVQEMEGPQVYPFYTEGSQELLKARTEITKF SLPRAVSRLQRARRKREDPDEDEDEELKCVLQQSAQINM DCSEIGDDRPLSGCAFSSDGTLLATSAWSGVTKLWSVPN INKVATLKGHTERVTDVAFSPTNCHLATACADRTAMLWN SEGVLMKTYEGHLDRLARLAFHPSGLYLGTASFDKTWRL WDVNTGIELLLQEGHSRSVYGIAFQCDGSLAATCGLDGL ARIWDLRTGRSILALEGHVKPVLGIDFSPNGYHLATGSE DHTCRIWDLRKRQSVYIIPAHSHLVSQVKFEPQEGYFLV TASYDSTAKVWSARDFKSIKVLAGHEAKVTSVDITADGQ YIATVSHDRTIKLWSSKNSTNDMNIG
239	The amino acid sequence of SEQ ID 511. The conserved G-protein beta WD-40 repeat domains are underlined and the Trp-Asp (WD) repeats signature is in bold.	MKRAYKLQEFVAHASNVNCLKIGKKSSRVLVTGGEDHKV NMWAIGKPNAILSLSGHSSAVESVTFDSAEALVVAGAAS GTIKLWDLEEAKIVRTLTGHRSNCISVDFHPFGEFFASG SLDTNLKIWDIRRKGCIHTYKGHTRGVNSIRFSPDGRWV VSGGEDNIVKLWDLTAGKLMHDFKCHEGQIQCMDFHPQE FLLATGSADRTVKFWDLETFELIGSAGPETTGVRAMIFN PDGRTLLTGLHESLKVFSWEPLRCYDAVDVGWSKLADLN IHEGKLLGCSYNQSCVGVWVVDISRVGPYAAGNVSRTNG HNEAKLASSGHPSVQQLDNNLKTNMARLSLSHSTESGIK EPKTTTSLTTTEGLSSTPQRAGIAFSSKNLPASSGPPSY VSTPKKNSTSRVQPTTNFQTLSRPDIVPVIVPRSNSLRP ETTSDAKKEMNNFGRVVPSTVSTKSTDVIKSGSNRDESD KIDSINQKRMTGNDKTDLNIARAEQHVSSRLDNTNTSSV VCDGNQPAARWIGAAKFRRNSPVDPVVSPHDRSPTFPWS ATDDGVTCQPDRQVTAPELSKRVVEPGRARALVASWETR EKALTADTPVLVSGRPPTSPGVDMNSFIPRGSHGTSESD LTVSDDNSAIEELMQQHNAFTSILQARLTKLQVIRRFWQ RNDLKGAIDATGKMGDHSVSADVISVLIERSEIFTLDIC TVILPLLTRLLQSETDRHLTVAMETLLVLVKTFGDVIRA TISATPTIGVDLQAEQRLERCNLCYVELENIKQILVPLI RRGGAVAKSAQELSLALQEV

Entry	Sequence Description	Annotated Peptide Sequence
240	The amino acid sequence of SEQ ID 512. The conserved cyclin N- and C-terminal family domains are underlined.	MAGSDENNPGVVGGAHVQEGLRVGAGKMGAGNVQQRRAL SNINSNIIGAPPYPCAVNKRVLSEKNVNSENDLLNAAHR PITRQFAAQMAYKQQLRPEENKRTTQSVSNPSKSEDCAI LDVDDDKMADDFPVPMFVQHTEAMLEEIDRMEEVEMEDV AEEPVTDIDSGDKENQLAVVEYIDDLYMFYQKAEASSCV PPNYMDRQQDINERMRGILIDWLIEVHYKFELMDETLYL TVNLIDRFLAVQPVVKKKLQLVGVTAMLLACKYEEVSVP VVEDLILISDRAYSRKEVLEMERLMVNTLHFNMSVPTPY VFMRRFLKAAQSDKKLELLSFFIIELSLVEYDMLKFPPS LLAASAIYTALSTITRTKQWSTTCEWHTSYSEEQLLECA RLMVTFHQRAGSGKLTGVHRKYSTSKFGHAARTEPANFL LDFRL
241	The amino acid sequence of SEQ ID 513. The conserved cyclin-dependent kinase inhibitor domain is underlined.	MQAPREGKSAAAIVGMGKYMKKSKAIPRDVSLLEASPRS PSATGVRTRAKTLASRRLRRASQRRPPPPAAAAAAAAPS LDASPCPFSYLQLRSRRLRRPRLAPSPEARIDEGPAGSG SRGSRDASCSARTASSSGGVEGEGACVGRGDRGNGGECV RDAAVDASYGENDLEIEDRDRSTRESTPCSLIRDSNANT PPGSTTRQQSSCTAHRTQMSILRSIPTSDEMEEFFAYAE QRQQRSFIEKYNFDIVKDRPLPGRFEWVQVIP
242	The amino acid sequence of SEQ ID 514. The conserved GCN5-related N-acetyltransferase family domain is underlined and the bromodomain is in bold.	MDGHSSHLAAQNRSRGSQTPSPSHSAASASATSSIHLKR KLSAANASAASAAAAAAAAAAAADDHAPPFPPSSISADT RDGALTSNDDLESISARGGGAGDDSDDDSDDEEEDDGDN DGGSSLRTFTAARLENVGPAAARNRKIKAESNATVKVEK EDSAKDGGNGAGVGALGPAATSGAGSGSGTVPKEDAVKI FTENLQASGAYSAREENLKREEEAGRLKFECLSNDGVDD HMVWLIGLKNIFARQLPNMPKEYIVRLVMDRNHKSVMVI RRNLVVGGITYRPYASQKFGEIAFCAIKADEQVKGYGTR LMNHLKQHARDVDGLTHFLTYADNNAVGYFIKQGFTKEI YLDKDRWHGYIKDYDGGILMECKIDPKLPYTDLSTMVRR QRQAIDEKIRELSNCHIVYQGIDFQKRDAGVPQNTIKME DIPGLREAGWTPDQWGYSRFRGLSDQKRLTFFIRQLLKV LNDHSDAWPFKEPVDAREVPDYYDIIKDPMDLKTMTKRV ESEQYYVTLEMFIADVKRMFANARTYNSPDTIYFKIATR LEAHFQSKVQSNLQSGAGKIQQ
243	The amino acid sequence of SEQ ID 515. The conserved TPR repeat domain is underlined	MFNGMMDPELFKLAQEQMNRMSPAELAKIQQQMMSNPEL MRMASESMKNMRPEDLRQAAEQLKHVRPEEMAEIGEKMA NASPEEIAAVRARADAQMTYEINAAKILKKEGNELHSQG RFKDASQKYLRAKNNLKGIPSSEGKNLLLACSLNLMSCY LKTRQYEECIKEGSEALACEEKNLKAFYRRGQAYRELGQ LKDAVSDLRKAHEISPDDETIAQVLRDTEESLTKEGGSA PRGVVIEEITEEDETLASVNHESPSEYSEKRHQESEDAH KGPINGDIMGQMTNSESLKALKGDPDAIRSFQNFISNAD PTTLAAMGAGNAGEVSPDLIKTASSMIGKMSAEELQKMI QLASSFPGENPYVTRNSDSNSNSFGNGSIPNVSPDMLKT ASDMMSKMSPDDLQRMFEMASSSRGKDPSLDANHASSSS GANLAANLNHILGESEPSSSYHIPSSSRNISSSPLSNFP SSPGDMQEQIRNQMKDPAMRQMFTSMMKNMSPEMMANMG KQFGLELSPEDAAKAQEAMSSLSPEMLDKMMRWADRAQR GVETAKKTKNWLLGRPGMILAICMLLLAVILHRLGFIGS
244	The amino acid sequence of SEQ ID 516. The conserved G-protein beta WD-40 repeat domains are underlined.	MIAAISWVPRGASKAVPEVAEPPSKEEIEEILKSGVVER SGDSDGEEDDENMDAVASEKADEVSTALSAADALGRISK VTKAGSGFEDIADGLRELDMDNYDEEDEDVKLFSTGLGD LYYPSNDMDPYLKDKDDDDDTEEIEDLSIKPMDSLIVCA RTDDEVNLLEVYLLEPSLSDESNMYVHHEVVISEFPLCT AWLDCPIKGGDKGNFIAVGSMEPAIEIWDLDIIDAVEPC LVLGGQEELKKKKKKGKKASIKYKEGSHTDSVLGLAWNK EFRNILASASADRQVKIWDVAAGKCNITMEHHTDKVQAV AWNHHAPQVLLSGSFDHSVVMKDGRIPSHSGYRWSVTAD VESLAWDPHSEHFFVVSLEDGTVRGFDVRAAISNSASQS LPSFTLHAHEKAVSTISYNPAAPNLLATGSTDKMVKLWD LSNNQPSCIASRNPKAGAVFSVSFSEDSPLLLAIGGSKG RLEVWDTSSDAAVSRRFGKHGKPKTAEPGS

Entry	Sequence Description	Annotated Peptide Sequence
245	The amino acid sequence of SEQ ID 517. The conserved Zn-finger, RING domain is underlined, and the SPX, N-terminal is in bold	**MKFCKKYQEYMQGQEGKKLPGLGFKKLKKILKRCRRRDS LHSQKALQAVQNPRTCPAHCSVCDGSFFPSLLEEMSAVL GCFNKQAQKLLELHLASGFQKYLMWFKGKLRGNHVALIQ EGKDLVTYALINAIAIRKILKKYDKIHLSTQGQAFKSQV** QRMHMEILQSPWLCELIAFHINVRETKANSGKGHALFEG CSLVVDDGKPSLSCELFDSIKLDIDLTCSICLDTVFDSV SLTCGHIYCYMCACSAASVTIVDGLKAAEPKEKCPLCRE ARVFEGAVHLDELNILLSRSCPEYWAERLQTERVERVRQ AKEHWESQCRAFMGVE
246	The amino acid sequence of SEQ ID 518. The conserved G-protein beta WD-40 repeat domains are underlined.	MVSTQSTRENPSIFFPPPLKPWLLPVVLSLSLSRQLGMA AAAAASLPFKKNYRSSQALQQFYAGGPFAVSSDGSFIAC NCGDSIKIVDSSNASLRPSIDCGSDTITALSLSPDGKLL FSAGHSRQIRVWDLSTSTCLRSWKGHDGPVMSMACPVSG GLLATGGADRKVMVWDVDGGFCTHFFKGHDGVVSTVLFH PDSNRSLLFSGSDDGTIRVWDLLAKKCASTLRGHDSTVT SLAFSEDGLTLLAAGRDKVVSLWDLHNYACKKTIPMYEV LESVCVIHSGTVLASQLGLDDQLKVTKESAQNIHFITVG ERGILRIWKSEGSVCLFKQEHSDVTVISDEDDSRSGFTA AVMLPLDQGLLCVTADQQFLFYYPEKHPEGIFSLTLCRR LVGYNEEIVDMKFLGEEENFLAVATNLEQVRVYELASMS CSYVLAGHTETVLCLDTCISSSGRTLIVTGSKDNSVRLW DSESRHCIGVGVGHMGAVGAVAFSRKRQDFFVSGSSDRT LKVWSLDGISEDGVDSTNLKAKAVVAAHDKDINSVAVAP NDSLVCSGSQDRTACVWRLPDLVSVVVLKGHKRGIWSVE FSPVDQCVLTASGDKTVKIWAISDGSCLKTFEGHVSSVL RASFLTRGTQFVSCGADGLVKLWTVRTNECIATYDQHSD KVWALAVGKKTEMLATGGSDAVVNLWYDSTASDKEDAFR KEEEGVLKGQELENAVSDADYTKAIELALELRRPHKLFE LFSELCRTREVGDRVERILSALSGEEVCLLLEYIREWNA KPKLCHVAQSVLSQVFRILSPTEIVEIKGIGELLEGLIP YSQRHFSRIDRLVRSTYLLDYTLTGMSVIEPEADRSAVN DGSPDKSGLEKLEDGLLGENVGEEKIQNKEELESSAYKK RKLPRSKDRSKKKSKNVVYADAAAISFRA
247	The amino acid sequence of SEQ ID 519. The conserved G-protein beta WD-40 repeat domains are underlined.	MDSAPRRKSGGINLPSGMSETSLRLDGFSGSSSSFRAIS NLTSPSKSSSISDRFIPCRSSSRLHTFGLVERGSPVKEG GNEAYSRLLKAELFGSDFGSLSPAGQGSPMSPSKNMLRF KTESSGPNSPFSPSILRQDSGFSSEASTPPKPPRKVPKT PHKVLDAPSLQDDFYLNLVDWSSQNTLAVGLGTCVYLWS ASNSKVTKLCDLGPNDGVCAVQWTREGSYISIGTSLGQV QIWDGTQCKRVRTMGGHQTRTGVLAWNSRILASGSRDRV ILQHDLRVPNEFIGKLVGHKSEVCGLKWSHDDRELASGG NDNQLLVWNQHSQQPVLKLTEHTAAVKAIAWSPHQNGLL ASGGGTADRCIRFWNTTNGHQTSSVDTGSQVCNLAWSKN VNELVSTHGYSQNQIMVWKYPSMAKVATLTGHSLRVLYL AMSPDGQTIVTGAGDETLRFWNVFPSAKAPAPVKDTGLW SLGRTHIR
248	The amino acid sequence of SEQ ID 520. The conserved G-protein beta WD-40 repeat domains are underlined.	MEDEAEIYDGVRAQFPLTFGKQSKPQTSLESVHSATRRG GPAPAPAPASSSSLPSTTSPSAAGGAGKSSGLPSLSSSS TAWLEGLRAGNPRAGREAGIGSRGGDGEDGGRAMIGPPR PPPGFSANDDGGGEDDDDDGDGVMVGPPPPPPGNLGDGD DDEEEEEAMIGPPRPPVVDSDEEEEEEEEENRYRLPLSN EIVLKGHNKIVSALAVDPTGSRVLSGSYDYTVRMFDFQG MNSRLSSFRDFEPVEGHQVRNLSWSPTADRFLCVTGSAQ AKIYDRDGLTLGEFVKGDMYIRDLKNTKGHITGLTWGEW HPKTKETILTSSEDGSLRIWDVNDFKSQKQVIKPKLARP GRVPVTTCTWDREGKCIAGGIGDGSIQIWNLKPGWGSRP DIHVEQAHADDITGLKFSSDGKILLTRSFDDSLKVWDLR LMKNPLKVFEDLPNHYAQTNIACSPDEQLFLTGTSVERE STIGGLLCFFDRSKLELVSRIGISPTCSVVQCAWHPRLN QIFATSGDKSQGGTHVLYDPTLSERGALVCVARAPRKKS VDDFELKPVIHNPHALPLFRDQPSRKRQREKILKDPLKS HKPELPMNGPGHGGRVGASKGSLLTQYLLKQGGMIKETW MDEDPREAILKHADAAEKNPKFTRAYAETQPDPVFAKSD SEDEDK

[0392] **Table 12. Eucalyptus *in silico* Data.**

SEQ ID	ConsID eucSpp	Family	1	2	3	4	5	6	7	8	9	10	11	12
1	3910	Cyclin-dependant protein kinase		0.25				0.11		0.20				0.73
2	19213	Cyclin-dependant protein kinase								0.59				0.64
3	36800	Cyclin-dependant protein kinase							0.11	0.36				
4	40260	Cyclin-dependant protein kinase				0.85								
5	41965	Cyclin-dependant protein kinase					0.35							0.86
6	2906	Cyclin-dependant protein kinase		0.93										0.81
7	1518	Cyclin-dependant protein kinase		0.08	0.28	0.08		0.06	0.11					
8	8078	Cyclin-dependant protein kinase					0.17							3.20
9	9826	Cyclin-dependant protein kinase			0.36	0.23			0.15	0.04		0.24		0.43
10	10364	Cyclin-dependant protein kinase						0.11	1.52					0.13
11	11523	Cyclin-dependant protein kinase				0.15		0.06	0.15					2.40
12	24358	Cyclin-dependant protein kinase		0.76					0.07	0.04		0.24		
13	39125	Cyclin-dependant protein kinase				0.23								
14	5362	Cyclin-dependant protein kinase		0.68				0.06		0.08				1.17
15	44857	Cyclin-dependant protein kinase		0.68				0.06		0.08				1.17
16	1743	Cyclin A			0.19		2.10	0.06	0.15					

SEQ ID	ConsID eucSpp	Family	1	2	3	4	5	6	7	8	9	10	11	12
17	12405	Cyclin A						0.06	0.59			2.84		
18	3739	Cyclin B		0.42	1.99	0.08								2.33
19	22338	Cyclin B												0.86
20	28605	Cyclin B						0.39		0.04		0.47		
21	41006	Cyclin B										0.71		
22	6643	Cyclin D		0.85			0.83	0.06	1.06	0.08				0.26
23	45338	Cyclin D												2.03
24	46486	Cyclin D							0.30					
25	12070	Cyclin-dependent kinase regulatory subunit	0.24		0.82			0.06	0.26					0.92
26	6617	Histone acetyltransferase		0.08				0.06	0.04	0.55			0.51	0.26
27	7827	Histone acetyltransferase			2.27			0.11		0.04				
28	8036	Histone acetyltransferase					1.16							
30	1596	Histone deacetylase		0.17	0.16	0.08	2.98	0.88	0.26	0.98				0.71
31	5870	Histone deacetylase			0.19		0.17			0.12				5.43
32	6901	Histone deacetylase	1.21	0.08				2.01	1.16	0.08				
33	6902	Histone deacetylase		0.08				0.11	1.21					0.47
34	7440	Histone deacetylase	0.48		1.23	0.15		0.22	0.48	0.20				2.02
35	8994	Histone deacetylase			0.09				0.15					
36	24580	Histone deacetylase		0.42					1.22					
37	37831	Histone deacetylase		0.08				0.22		0.40		1.19		0.12
38	34958	MAT1 CDK-activating kinase assembly factor							0.15	0.23				
39	22967	Peptidyl-prolyl cis-trans isomerase			0.72									0.69
40	8599	Peptidyl-prolyl cis-trans isomerase			0.46	0.08	0.50	0.17	0.51	0.28				3.01
41	9919	Peptidyl-prolyl cis-trans isomerase			0.51		0.35	0.06	0.15	0.43				4.24
42	15820	Peptidyl-prolyl cis-trans							0.04					6.78

SEQ ID	ConsID eucSpp	Family	1	2	3	4	5	6	7	8	9	10	11	12
		isomerase												
43	8327	Peptidyl-prolyl cis-trans isomerase						0.06		0.04				6.86
44	4604	Peptidyl-prolyl cis-trans isomerase												0.68
45	966	Peptidyl-prolyl cis-trans isomerase		0.59	1.02	0.54	0.69	0.50	0.93	0.59		0.95		18.65
46	1037	Peptidyl-prolyl cis-trans isomerase		0.59										
47	4603	Peptidyl-prolyl cis-trans isomerase		0.17				0.17	1.24	0.04				0.34
48	5465	Peptidyl-prolyl cis-trans isomerase			1.21	0.08	0.66	0.11	0.29	0.16				6.99
49	6571	Peptidyl-prolyl cis-trans isomerase		0.51		0.08			0.41	0.08				1.14
50	6786	Peptidyl-prolyl cis-trans isomerase		0.42			0.33	0.06	0.41	0.04				
51	7057	Peptidyl-prolyl cis-trans isomerase		0.42				0.11	0.04					
52	8670	Peptidyl-prolyl cis-trans isomerase			1.56			0.39		0.20				0.12
53	9137	Peptidyl-prolyl cis-trans isomerase							0.04	0.59				
54	10285	Peptidyl-prolyl cis-trans isomerase					0.60	1.16	0.04	0.04				0.45
55	10600	Peptidyl-prolyl cis-trans isomerase			0.16		0.17	0.06						0.46
56	11551	Peptidyl-prolyl cis-trans isomerase				0.08		0.06	0.04	0.08				1.89
57	20743	Peptidyl-prolyl cis-trans isomerase								0.76				
58	23739	Peptidyl-prolyl cis-trans isomerase		0.59										

SEQ ID	ConsID eucSpp	Family	1	2	3	4	5	6	7	8	9	10	11	12
60	31985	Peptidyl-prolyl cis-trans isomerase					1.99							
61	32025	Peptidyl-prolyl cis-trans isomerase					0.99							
62	32173	Peptidyl-prolyl cis-trans isomerase					1.99							
64	9143	Retinoblastoma related protein			0.90				0.15					
65	349	WD40 repeat protein	0.24		0.34	0.08	0.17	0.22	0.33	0.08			0.25	2.24
66	575	WD40 repeat protein		0.25	0.94	0.31	0.34	0.11		0.16		0.47		1.87
67	804	WD40 repeat protein				0.15	0.34	0.39	0.33	0.39				1.82
68	805	WD40 repeat protein	0.97	0.51	4.66	0.23	0.17	0.77	0.33	1.07		0.24		4.43
69	806	WD40 repeat protein					0.83			0.04				
70	2248	WD40 repeat protein		0.08		0.08	1.92	0.06		0.08				0.91
71	3203	WD40 repeat protein		0.34	0.18	0.15	0.17	0.11	0.30	0.04				0.72
72	3209	WD40 repeat protein		0.08		0.15	0.17			0.12				0.61
73	4429	WD40 repeat protein		0.08	1.16	0.08								0.13
74	4607	WD40 repeat protein		0.76		0.54		0.06	0.07					
75	4682	WD40 repeat protein		0.08	0.28	0.23			1.13	0.08				0.12
76	5786	WD40 repeat protein		0.08				0.06	0.46	0.08				0.13
77	5887	WD40 repeat protein		1.61	1.23	0.08		0.06	0.15	0.28				1.41
78	5981	WD40 repeat protein		0.08										0.37
79	6766	WD40 repeat protein	0.24	0.08	1.31		0.51	0.06	0.74	0.51				0.28
80	6769	WD40 repeat protein		0.93			0.17			0.12				2.28
81	6907	WD40 repeat protein		0.25			0.17	0.06	0.45	0.32		0.47		1.67
82	7518	WD40 repeat protein			0.91			0.28	0.15	0.55				0.59

SEQ ID	ConsID eucSpp	Family	1	2	3	4	5	6	7	8	9	10	11	12
83	7717	WD40 repeat protein			0.47									0.38
84	7718	WD40 repeat protein	0.24		1.88	0.08		0.22		0.04				0.92
85	7741	WD40 repeat protein			1.42			0.11				0.47		
86	7884	WD40 repeat protein			1.33	0.15						0.24		
87	8258	WD40 repeat protein	0.72		0.19	0.23	0.87		0.15	0.08				0.08
88	8465	WD40 repeat protein			0.47	0.08	1.75							
89	8616	WD40 repeat protein			0.57	0.08	0.69			0.16				0.13
90	8690	WD40 repeat protein			0.26	0.08	0.35	1.39	0.34	0.32		2.13		0.80
91	8708	WD40 repeat protein			0.57					0.04				
92	8850	WD40 repeat protein			0.09			0.06		0.27				2.03
93	9072	WD40 repeat protein			1.21		0.17							0.48
94	9465	WD40 repeat protein	0.24		0.72		0.33		0.15					
95	9472	WD40 repeat protein			0.36		1.99	0.11	0.61					6.90
96	9550	WD40 repeat protein			0.90			0.11	1.78					
97	10284	WD40 repeat protein	0.24	0.08			1.82		1.22	0.16		0.47		0.28
98	10595	WD40 repeat protein			0.16		0.17	0.11	6.52					0.85
99	10657	WD40 repeat protein						0.06		0.12				
100	12636	WD40 repeat protein						0.06						0.65
101	12748	WD40 repeat protein			1.50	0.08		0.06	1.67	0.04				0.38
102	12879	WD40 repeat protein				0.08	0.33	0.06	0.04	0.08				2.00
103	15515	WD40 repeat protein					0.35		0.30					
104	15724	WD40 repeat protein		0.25	0.33	0.15			0.47	0.04				0.39
105	16167	WD40 repeat protein	0.24				0.52							
106	16633	WD40 repeat protein			1.96					0.12				0.42

SEQ ID	ConsID eucSpp	Family	1	2	3	4	5	6	7	8	9	10	11	12
107	17485	WD40 repeat protein			0.65									
108	18007	WD40 repeat protein								0.12				
109	20775	WD40 repeat protein					0.17			0.08				
110	23132	WD40 repeat protein												2.42
111	23569	WD40 repeat protein							0.91					0.91
112	23611	WD40 repeat protein							4.15					
113	24934	WD40 repeat protein		0.34						0.04				
114	25546	WD40 repeat protein			0.09									
115	30134	WD40 repeat protein							0.07					
116	31787	WD40 repeat protein			0.19									1.19
117	34435	WD40 repeat protein					0.35			0.08				
118	34452	WD40 repeat protein			1.44					0.20				0.25
119	35789	WD40 repeat protein								0.20				
120	35804	WD40 repeat protein							0.19	0.27				0.08
121	43057	WD40 repeat protein							0.30					0.57
122	46741	WD40 repeat protein							0.46					
123	47161	WD40 repeat protein							1.78					
235	6366	WD40 repeat protein		0.08	0.68	0.23	0.93	0.11	0.36	0.83		0.24		0.94
236	17378	WD40 repeat protein			0.65					0.12				0.08
252	45414	Cyclin B												3.13
253	44328	Cyclin-dependant kinase inhibitor												0.38
254	15615	Histone acetyltransferase							0.22	0.04				
255	17239	Peptidyl-prolyl cis-trans isomerase				0.08	0.50			0.08				

SEQ ID	ConsID eucSpp	Family	1	2	3	4	5	6	7	8	9	10	11	12
256	18643	WD40 repeat protein								0.04				0.90
257	19127	WD40 repeat protein								0.04				0.89
258	22624	WD40 repeat protein												1.16
259	32424	WD40 repeat protein					0.50							
260	37472	WD40 repeat protein								0.08				0.17

In Table 12, the following numbers 1-12 represent the following tissues: 1 is bud reproductive; 2 is bud vegetative; 3 is cambium; 4 is fruit; 5 is leaf; 6 is phloem; 7 is reproductive; 8 is root; 9 is sap vegetative; 10 is stem; 11 is whole; and 12 is xylem.

[0393] **Table 13: Pine *in silico* data.**

SEQ ID	ConsID pinus Radiata	Family	1	2	3	4	5	6	7	8	9	10	11	12
124	1766	Cyclin-dependant protein kinase				1.02	0.05	1.58	0.15	0.22	0.22	0.18	2.16	4.91
125	2927	Cyclin-dependant protein kinase	0.16					0.19	0.11	0.14	0.04	0.36	0.38	0.17
126	7642	Cyclin-dependant protein kinase			0.22	0.21	0.05				0.07			
127	13714	Cyclin-dependant protein kinase						0.11	0.11					
128	16332	Cyclin-dependant protein kinase						0.54	0.26		0.14		0.04	0.91
129	21677	Cyclin-dependant protein kinase					0.05	0.14						0.17
130	27562	Cyclin-dependant protein kinase												0.41
131	1504	Cyclin-dependant protein kinase	0.16					0.36		0.35	0.21	0.54	0.09	0.65
132	15211	Cyclin-dependant protein kinase						0.13	0.15				0.19	0.19
133	20421	Cyclin-dependant protein kinase									0.04		0.05	0.95
134	3187	Cyclin-dependant protein kinase						0.34	0.15		0.04	0.18	0.38	

SEQ ID	ConsID pinus Radiata	Family	1	2	3	4	5	6	7	8	9	10	11	12
135	15661	Cyclin-dependant protein kinase									0.04		0.13	
136	13874	Cyclin A	0.31					0.27	0.15				0.05	
137	14615	Cyclin A	0.16					0.15						
138	4578	Cyclin B	0.47	0.14				0.13	0.22				0.74	0.38
139	23387	Cyclin B						0.29	0.26					0.17
140	6970	Cyclin D		0.14						0.27	0.04			
141	10322	Cyclin D	0.16			0.19		0.06			0.14		1.12	1.36
142	22721	Cyclin D						0.27		0.36				
143	23407	Cyclin D						0.15	0.26					0.31
144	1945	Cyclin-dependent kinase regulatory subunit		0.28	0.55	0.41	0.16	1.62	5.02	0.22	0.72		0.39	3.06
145	8233	Cyclin-dependent kinase regulatory subunit				0.21								
146	8234	Cyclin-dependent kinase regulatory subunit	0.16		0.11									
147	22054	Cyclin-dependent kinase regulatory subunit					0.05		0.22				0.18	

SEQ ID	ConsID pinus Radiata	Family	1	2	3	4	5	6	7	8	9	10	11	12
148	12137	Histone acetyltransferase						0.06					1.51	0.19
149	12582	Histone acetyltransferase						0.64	0.15	1.09			0.33	0.63
150	15285	Histone acetyltransferase				0.21					0.12		0.70	0.14
151	17229	Histone acetyltransferase											0.94	0.16
152	20724	Histone acetyltransferase									0.04		0.19	0.19
153	4555	Histone deacetylase	0.16	0.14				0.97		0.14			0.89	0.89
154	4556	Histone deacetylase												0.14
155	5729	Histone deacetylase	0.31	0.28	0.22	0.58	0.22	2.00	0.48	0.07	0.04		2.73	1.46
156	7395	Histone deacetylase		0.14		0.14			0.19	0.93	0.04		0.14	1.33
157	9503	Histone deacetylase			0.11					0.14				
158	11283	Histone deacetylase				0.19			0.15				0.96	1.35
159	12322	Histone deacetylase	0.16					0.06	0.11		0.04		0.05	0.29
161	23236	Histone deacetylase						0.13			0.11			
162	171	Peptidyl-prolyl cis-trans isomerase									0.07			0.46

SEQ ID	ConsID pinus Radiata	Family	1	2	3	4	5	6	7	8	9	10	11	12
163	172	Peptidyl-prolyl cis-trans isomerase						0.19			0.11	0.18	0.11	0.46
164	1480	Peptidyl-prolyl cis-trans isomerase	2.51	4.20	0.88	2.97	1.58	3.53	7.36	1.33	2.74	0.72	6.62	10.14
168	1692	Peptidyl-prolyl cis-trans isomerase	0.16		0.22		0.65	0.61	0.26		0.29	0.18	1.28	0.34
169	5313	Peptidyl-prolyl cis-trans isomerase		0.14						0.07			0.37	0.17
170	6362	Peptidyl-prolyl cis-trans isomerase		0.14	0.33	0.05		0.06	0.60		0.04		2.92	0.68
171	6493	Peptidyl-prolyl cis-trans isomerase		0.42	0.11	0.21			0.11		0.04		0.25	0.32
172	6983	Peptidyl-prolyl cis-trans isomerase				0.61		0.13					0.04	
174	7665	Peptidyl-prolyl cis-trans isomerase			0.11	0.39	0.05	0.62					0.25	
175	12196	Peptidyl-prolyl cis-trans isomerase						0.19	0.15	0.14	0.16			

SEQ ID	ConsID pinus Radiata	Family	1	2	3	4	5	6	7	8	9	10	11	12
176	13382	Peptidyl-prolyl cis-trans isomerase				0.25		0.06		0.07	0.04		0.87	0.15
177	16461	Peptidyl-prolyl cis-trans isomerase				0.19		0.15	0.15		0.04		0.04	0.74
178	17611	Peptidyl-prolyl cis-trans isomerase				0.24	0.11	0.27	0.41				0.99	
179	19776	Peptidyl-prolyl cis-trans isomerase						0.13		0.07	0.16		0.05	0.61
180	20659	Peptidyl-prolyl cis-trans isomerase								0.15			0.19	
181	22559	Peptidyl-prolyl cis-trans isomerase							0.11	0.14				0.20
182	24188	Peptidyl-prolyl cis-trans isomerase									0.23			
183	27973	Peptidyl-prolyl cis-trans isomerase											1.01	
184	1353	WD40 repeat protein			0.44		0.05	0.73			0.11	1.07	0.70	1.32
185	1978	WD40 repeat protein		0.14		0.05		0.44	0.11	0.21	0.27	0.36	1.46	0.82

SEQ ID	ConsID pinus Radiata	Family	1	2	3	4	5	6	7	8	9	10	11	12
186	2810	WD40 repeat protein		0.42		0.79	0.11	0.39		0.27		0.36	1.69	1.03
187	2811	WD40 repeat protein									0.14		0.09	0.14
188	2812	WD40 repeat protein							0.15			0.18	0.04	0.16
189	3514	WD40 repeat protein				0.63		0.06		0.14		0.18	0.48	0.56
190	4104	WD40 repeat protein		0.14		0.25		0.27	0.37	0.36	0.19	0.18	0.39	0.53
191	5595	WD40 repeat protein		0.14		0.25			0.15	0.14	0.07		0.23	
192	5754	WD40 repeat protein	0.31	0.14				0.06		0.07	0.16		0.10	0.16
193	6463	WD40 repeat protein	0.16	0.56	0.22	0.43		0.81	0.53	0.21	0.08		1.00	0.70
194	6665	WD40 repeat protein	0.31	0.28		0.45	0.44	0.96			0.07		3.37	2.68
195	6750	WD40 repeat protein		0.14		0.59	0.05		0.37	0.42	0.04		0.18	0.52
196	7030	WD40 repeat protein	0.31			0.40	0.54	0.45	0.37		0.07		1.58	3.41
197	7854	WD40 repeat protein			0.11					0.14			0.05	
198	7917	WD40 repeat protein			0.22	0.39		0.13	0.15				0.18	0.56
199	7989	WD40 repeat protein			0.11						0.04		0.11	
200	8506	WD40 repeat protein	0.47		0.33		0.11	0.86	0.19	1.28	0.04		1.23	3.12
201	8692	WD40 repeat protein				0.21		0.06	0.11		0.15		0.10	0.87

SEQ ID	ConsID pinus Radiata	Family	1	2	3	4	5	6	7	8	9	10	11	12
202	8693	WD40 repeat protein			0.11	0.80		0.25		0.14	0.18		0.53	0.31
203	9170	WD40 repeat protein	0.16		0.11	0.05							0.05	
204	9408	WD40 repeat protein			0.33		0.05	0.41	0.15	0.14			0.41	0.33
205	9522	WD40 repeat protein			0.11								0.18	
206	9734	WD40 repeat protein			0.11	0.05	0.11	0.15		0.07	0.25		0.11	
207	9815	WD40 repeat protein			0.11								0.18	0.14
208	10670	WD40 repeat protein				0.40	0.16	0.11			0.16		0.34	0.31
209	11297	WD40 repeat protein				0.53			0.15		0.16		0.05	
210	13098	WD40 repeat protein				0.19	0.11	0.54	0.31	0.14	0.26		1.85	0.14
211	13172	WD40 repeat protein									0.04			
212	13589	WD40 repeat protein					0.11	0.06		0.21			0.05	0.37
213	13608	WD40 repeat protein						0.11			0.04		0.59	0.33
214	14299	WD40 repeat protein	0.16			0.05				1.09			0.38	
215	14498	WD40 repeat protein				0.21							0.44	0.30
216	14548	WD40 repeat protein	0.16								0.11		0.11	0.82
217	14610	WD40 repeat protein	0.16					0.27						

SEQ ID	ConsID pinus Radiata	Family	1	2	3	4	5	6	7	8	9	10	11	12
218	16090	WD40 repeat protein						0.43			0.04		0.37	0.85
219	16722	WD40 repeat protein											0.10	
220	16785	WD40 repeat protein				0.05		0.13					0.38	0.50
221	17094	WD40 repeat protein						0.29	0.15				0.24	0.81
222	17527	WD40 repeat protein									0.04		0.10	
223	17591	WD40 repeat protein						0.14					0.10	
224	17769	WD40 repeat protein											0.39	
225	18047	WD40 repeat protein				0.05	0.22	0.98	0.15	2.68	0.07		0.19	0.80
226	18414	WD40 repeat protein					0.16	0.15			0.34		0.23	0.19
227	18986	WD40 repeat protein						0.41					0.15	
228	19479	WD40 repeat protein					0.05						0.28	0.32
229	20144	WD40 repeat protein				0.43			0.29				0.05	
230	22480	WD40 repeat protein							0.15	0.27				
231	23079	WD40 repeat protein						0.13			0.04			
232	26739	WD40 repeat protein							0.15				0.18	
233	26951	WD40 repeat protein				0.21								0.20

SEQ ID	ConsID pinus Radiata	Family	1	2	3	4	5	6	7	8	9	10	11	12
234	26529	WEE1-like protein									0.04		0.18	
237	888	WD40 repeat protein							0.11				0.18	
238	14166	Cyclin-dependant kinase inhibitor	0.16			0.05							0.05	
239	3189	Cyclin-dependant protein kinase						0.06						
240	9356	Histone acetyltransferase			0.11				0.22					0.46
241	65	Histone deacetylase	0.16				0.22	0.27	0.22				0.24	0.34
242	14197	Histone deacetylase	0.16				0.33						0.05	
243	9081	Peptidyl-prolyl cis-trans isomerase			0.11		0.05		0.29		0.26		0.69	
244	13417	Peptidyl-prolyl cis-trans isomerase						0.06					0.59	
245	5755	WD40 repeat protein	0.16											
246	6670	WD40 repeat protein		0.14		0.05								
247	7027	WD40 repeat protein		0.14				0.15					1.30	0.15
248	7276	WD40 repeat protein		0.14					0.11				0.05	
249	7390	WD40 repeat protein	0.31	0.14			0.11		0.44				1.29	0.38

SEQ ID	ConsID pinus Radiata	Family	1	2	3	4	5	6	7	8	9	10	11	12
250	12648	WD40 repeat protein				0.05		0.06					0.05	0.94
251	13171	WD40 repeat protein				0.19		0.63					0.19	0.34

Table 13, the following numbers 1-12 represent the following tissues: 1 is bud reproductive; 2 is bud vegetative; 3 is callus; 4 is cambium; 5 is meristem vegetative; 6 is phloem; 7 is reproductive female; 8 is reproductive male; 9 is root; 10 is vascular; 11 is whole; and 12 is xylem.

[0394] **Table 14. Oligo Table.**

Oligo SEQ ID	Oligo ID	Microarray Oligo Seq
521	Euc_003910_O_4	GATTTTAAGTAACTCAATTAGCAGTTCCAACATTAAACCATTATTATTACCCCTTTTATC
522	Euc_019213_O_1	CTCAAAAAGTACTTGGATGCGTGCGGTGACAACGGACTCGAACCGTACACTGTCAAATCT
523	Euc_036800_O_4	TTGTCAAGTTGCAGGACGTAGTGCACAGTGAGAGGCGTCTATATCTAGTTTTTGAGTACT
524	Euc_040260_O_1	GAAGAAATTATATAACTAGATACAAGGTTAGCTAGGTATATAATAGCGGTACAAGTCTTT
525	Euc_041965_O_1	GGACAAATCAAGTAGAACTTCTCTCGGCAGCATCAGTTTTTCTAATCCATGCCTTGTTGC
526	Euc_002906_O_1	CTCAGTTCTGATAATGCCTCGGATATATGGCCGAGTGTTCGCTGGACGGCCTCTTATGTT
527	Euc_001518_O_3	GGAGATTCTGAACTGCAACAGCTCCTACACATTTTCAGACTGTTGGGTACTCCAAATGAA
528	Euc_008078_O_2	GACTGGTAAAATCGTTGCACTAAAAAAGGTCCGGTTTGACAACTTGGAACCTGAAAGCGT
529	Euc_009826_O_4	AAACACCAATCTATCAACACTGTCGAGTTTAGTCACTAGTAGAACCGGAGATAACAAACA
530	Euc_010364_O_1	CTATGATCCTGAGCGCAAGCAAGTTATGACCAATAGAGTCGTTACACTATGGTACCGAGC
531	Euc_011523_O_1	TGTTGTGAAGGTAGTTATAGCCATCGATTAGACAGTGATTAAAGTAGTACCCGTGCCAAT
532	Euc_024358_O_2	CCACATACAAGAGTTGTTACGCTACACATCCTATACCATCAAAGGAACGTTGGAATGCCA
533	Euc_039125_O_3	TATGATCGACACAAGCATTTTGTGTTGGAGCCTCAGCTAATTGTATGTCATCGAGTACTT
534	Euc_005362_O_3	AAAATTTTTGCTACGGATAATGTTGTGAGGCGAGGCAGTCGAAATTACGGAGGTTGACTT
535	Euc_044857_O_1	ATGCAGGGATCAAATTTGTGAGTACTACGTAAAATTTTGCTACGGAGGCGAGGCAGTCGA
536	Euc_001743_O_1	GAAGAATACAGGCTCGTACCTGATACACTGTACCTGACTGTTAACTACATAGATCGGTAT
537	Euc_012405_O_1	TCCACCCTAAATGCGATACGTGAAAAGTATAGACAACAGAAGGTAAACTATTCATTACTG
538	Euc_003739_O_2	AGGCTTCTAGTTGCGTTCCCCCAAACTACATGGATCGGCAGCAGGATATTAATGAGCGGA
539	Euc_022338_O_2	GAGAAAAATGACAGATTGATATCGATGATGATGACTGTCGTGTCATCAGTAGTGTGCTTT
540	Euc_028605_O_5	TTTCCAATTGTAGTTCGTCTTTTATTGTAACAATAAATTGATAGATACTGATTCGAAATA
541	Euc_041006_O_1	ACATTTATGCTAACTATAGGAGAACGGAGAATTGTAGCTGCGTCTCTGCTAACTACATGG
542	Euc_006643_O_1	TTCTGGCTTAAAGGCTATTCTTTGTGCACAATGACCTGAGGGAGGTCTCGACAGACCACT
543	Euc_045338_O_1	TTCATCCGGGTCCTGGTTATCATACTCTTATATATGTTGGGGAATAACGGTTCATATGTT
544	Euc_046486_O_3	GGGTGTGCTTAATAGTTCTTATTAGTCTTAGCTTATTATCTTTGATTGGACATGCTATAA
545	Euc_012070_O_2	CTTGCTAAGTAGACATGTTATATTTCTAATGCTTTGAGAACAATATTACAGTATAATTAG
546	Euc_006617_O_2	AATCATCGACTAGACCGATGGTCAAAGTGGTAATCATGTAATTAAACGCGTTTGTCATTG
547	Euc_007827_O_2	ATGGAAAAATCTATGGATATGAAGGATTGAAGATATCCGTCTGGGTAAGCTGTGTATCAT
548	Euc_008036_O_3	TTATGATTTGAGAAAACCCTTGCAGGCTGCGATTTGCGGATCATGACAGCATAGTTTTGC
549	Euc_001596_O_2	GTTTTGTTGTGAGGGCTTGGTAGGTTTTCATTATATTGTAATGTCGACGACAGAGATTTT
550	Euc_005870_O_3	CCAATTAATGTTACTGCTCAAGCTGACGTACCTGCGAAAAAAGCACCAGTGACTGCTAAT
551	Euc_006901_O_3	TGATGTCAAAACGTAGCTCTTTTTTGTGTGAGCTATCCTGCTAAATTAAACCTCAGCAAA
552	Euc_006902_O_1	ACATGAGTATTATGAATACTTCGGTCCTGACTATACACTTCATGTTGCTCCGAGTAACAT
553	Euc_007440_O_2	GAATTGGCGATCACAATCTACTGTAGTCAATACTCAAGTGGGAGGTGTAAATAGATTCCA
554	Euc_008994_O_1	GATCATGTGTAATCAGTATATCAGGTTAGAAACAGTACTCTTGAGCTTAGCGGGCACTGT
555	Euc_024580_O_2	TCCTGTGAAGGTGGTCGACTCAATCAAAAGGTACCTTGTAGATAAGGTACCTTTTCTCAA
556	Euc_037831_O_5	GCATTTTATACGACGGATAGAGTCATGACCGTATCTTTCCATAAGTTTGGGGACTTCTTC
557	Euc_034958_O_3	CCTCGTTTCTTTGCGGTTCGGACGCATCATGGATGTATCTCCAAAGAGTAATCTGTCGAT
558	Euc_022967_O_2	AATTCAGATCTATTAGTGAAAGTTGGCATGAGTCTCAATCTTAGGGGAATACAGTACGGA
559	Euc_008599_O_3	TGATATGAGTATCATAACTCGGATGGTGACAACTTTGTACTACGGTCGGCACCGGTAGAT
560	Euc_009919_O_1	CATATACAATCTTAGTGGATTAGCTGAGGTCGAAACTGACAAGAGTGATCGCCCGTTGGA

Oligo SEQ ID	Oligo ID	Microarray Oligo Seq
561	Euc_015820_0_2	CATGGCTAACGCTGGCCCTAGCACTAATGGGAGCCAATTTTTCATATGCACTGTAAAGAC
562	Euc_008327_0_2	AACAAAGTCTACCTTGACATTAGCATCGGTAACCCTGTCGGGAAACTAGTCGGAAGAATT
563	Euc_004604_0_2	TGTGCTTGGATATACTGTATAAGCATTCTATATTATGCTTGTTGGCTTCGTTTTGAGGGA
564	Euc_000966_0_1	TTAACGTCGACCGCTTCTCTGCCCCTTGAATTTTCCCGAGAAAACCAGGAACCTGCCAAA
565	Euc_001037_0_1	TGTTGAATACGATGTATTATAATGTTGGTGTCTTGGTGAAATACAGAATTATGCTTGCGT
566	Euc_004603_0_2	ATCGCTGTGGCTGATCTCGTCGCTCCGGCTTTTCATAAAAATCATGGCTGAGGCAATCGA
567	Euc_005465_0_2	CTCGCAACCCTATATCTCGCTCAGGCGAAGAAGTCTGAGGATTTGAAAGAGGTGACTCAC
568	Euc_006571_0_1	TGTTTTTGGGTACACGCAGTTAGGATAACTAGCATGAAAGCCCGATCCCGCATATACAGG
569	Euc_006786_0_2	GAGGACTAGCCGGAACTTCATCGAACTCTCTCGGAGGGGTTACTACGATAACGTCAAGTT
570	Euc_007057_0_1	GATGGCTAGCACTGTGTAGAAAGGTGAATTTAAAGTACTTGTCTACACTGCTTATTAAAT
571	Euc_008670_0_2	TGAGACTGTCTTGGCGTGTATTTTGGAATAAACTATTATCACGTTTTGTTAAATATAATA
572	Euc_009137_0_3	TTACAAAATGGCTCTCAGAAAGTATCGAAAGGCCCTGCGCTATCTGGATATCTGCTGGGA
573	Euc_010285_0_2	AATTTTATGTTTGCTACTGCTTAGTGCTTAATGGACTTGCGTAGGTATTCAAATTACAGA
574	Euc_010600_0_1	TGGAACCGTGGTATCGGCTGACGTTATCCGTGATTTTAAGACTGGAGATAGTTTATGCTA
575	Euc_011551_0_2	CTTTGATGTATCCTCAGTGTACTGCTTTTAGCTATGTATAGATCGAGTCAACTCATTGAA
576	Euc_020743_0_3	TTTTTATTATTTACCTTCGCCTTTACGCTGCATACGTTAATAGGTTATTATTTCCTTCAA
577	Euc_023739_0_1	ATTTGTCCATGACAATCGTAGTCGAAGACACGATACGCTCTTAGATGGTACGGAAATCTG
578	Euc_031985_0_2	TGAATAGAGATAACTTTTCTGAGTGTGAATTGGATATTACGTTGCAAATAGCCGAATGAA
579	Euc_032025_0_2	GCTTTAGGTTAGGGATCCCTGTAAGCTGATGATAGATATTGGAGATGGTACTTGTAAGAT
580	Euc_032173_0_1	TGTTGTGTTTGGAAAGGTGCTGTCTGGGATGGATGTTGTCCACAAGATTGAGGCTGAAGG
581	Euc_009143_0_1	GGAAAGCGGGGAATGAGCATGTGGATATTATCTCTTTCTACAATGAAATATTCATTCCTT
582	Euc_000349_0_1	CATCAGGACGTTGACTCTAATTAAGACATATGTGACAGAGCGCCCTGTTAATGCGGTTAC
583	Euc_000575_0_2	CTTTAGGTTTGATCTGTCTGTTTTGTCTATCCTGCGAGTTTCGAGCATGTGCGTGTGTGA
584	Euc_000804_0_1	CAGCCCCAATAGATACTGGCTCTGTGCCGCTACTGAGAACAGTATTAAAATCTGGGACCT
585	Euc_000805_0_2	AAGAATGAAGCTGATATGAGTGATGGAACTACGGGGGCCATGAGCTCAAATAAGAAGGTC
586	Euc_000806_0_1	TGACTACAATTAGCACCTCACCATTATCGAACTGTATAATTGTGCTTGCCTGCTATTATT
587	Euc_002248_0_4	TTGAAGCGGAAATATATATTTATGCTACTACATAAGTAATGTACTACTTGACAAGATGAG
588	Euc_003203_0_1	TACTCGATGTGGTATAGAATTTATCCAATGTACTCCTAAATGTAGATACATCGTGTATTG
589	Euc_003209_0_2	GCTTCGTCTGATACCACTATCAAGATAATAGGCGTGAGCAATAGCTCTGGATCACAGCAC
590	Euc_004429_0_4	GGTCGGCTTGCTAGTGTATCTGATGACAAGAGCATATCACTCTATGATTACTCATGAAGG
591	Euc_004607_0_3	GAAAGGAGAAAAGCATGGAGATCGATCTCGGAAACCTCGCATTCGACGTCGATTTTCATC
592	Euc_004682_0_1	GATTCAGTACCCGGATTCGCAAGTCAACCGGTTGGAGATAACTCCACATAAGCGGTACCT
593	Euc_005786_0_1	TTCCATGTATCAAGCCGCATCAATGTTTGTCGCTGCAATTAACATGTGTGCAGTCGATCC
594	Euc_005887_0_2	TTCAGCGCATTGTGTAAATGTAGATAGGTGATATATTTCTCGTTGCAATGTAGGGTAAGA
595	Euc_005981_0_2	TCCAATAATCACATTTACCATCAACAGGCATCAGCAACATACTGTTGTAGTGTAATTAAT
596	Euc_006766_0_1	GGGCATTCTGACTACCTGCACTGTATAGCTGCACGGAACTCTTCTAGTCAGATTATAACA
597	Euc_006769_0_1	AATCGTCTGGTAGATTGTCAAAAACTAATAAACCTGTGATTGATCCGGATTCTAGTAATG
598	Euc_006907_0_2	AGTTGAGGATTCTCCACTATGACAGCTCTCATGGCTTGAATCTAAAGTCATCTGGTTTTC
599	Euc_007518_0_1	GAACAATCATTCTGTAGAACACTAGAGTCTATATGCTTGACTGTATCGGTTAATTAATTC
600	Euc_007717_0_1	AGATAGCGATAGAGTTATACTGCATGTACTGAGGTAAATGTTTTGATTACTCCACCCAAT
601	Euc_007718_0_1	AAGAATTGTTAGGAGGTGTATACTTTCTGTAACTGTATTCAATGAGCATACACCTGACGG
602	Euc_007741_0_2	CAACTCATATAATGACTGGATTCTGGCAACCGCGTCTTCAGACACAACAGTTGGACTATT
603	Euc_007884_0_1	AGTGTAAAAGGATGCCCCTAATAGATTATATGCCAAGTGTAGTATATATAATAGTGCTTT

Oligo SEQ ID	Oligo ID	Microarray Oligo Seq
604	Euc_008258_O_2	AAGAATCTACAGTTGTCTTATGCTACTCTATTACTCAATTATGCTGTGCTATTGATTGAG
605	Euc_008465_O_4	TCTGAATACATACTTTGTGGTCTCTATAAAAGACCAATGATACAGGCATGGTCATTAATT
606	Euc_008616_O_5	TAAATCTTCTCATGTGCCTGGCGTAAATTTTGCAGTTATTACTAGACCAAGATAGTTTCA
607	Euc_008690_O_4	ACATGGATTCGATCAATCGCCACATGACAACTAAAACAAGCGGTTCACGTGATTGTAATT
608	Euc_008708_O_4	AGATGAGTATGCTCGGGTGTATGATATTCGCAATTACAAGTGGAATGGATCGCATAATTT
609	Euc_008850_O_5	TCTTTGATTCTGTTGTATGGTGTATCTTATTGTATCTTCTATCTGCCCCCCATGTAATTC
610	Euc_009072_O_1	TTCGTTGTGTAGTACTGGGAGTTACTACTTGTATGTATGTAAATCATGTGGCGTCTGTCC
611	Euc_009465_O_1	GGAGATGTGTAATATGTCTGAGCGGTCACACTCTAGCTGTTACATGCGTAAAGTGGGGAG
612	Euc_009472_O_3	CCACCGTTGCGTAACTCGAATAGCCGGATTTTCGTTTTCGTTTTTATTTCCCCGTTAATT
613	Euc_009550_O_1	TGAGATGCTCTGTGTGAGGACTTTTACGAAACTTGAATGGCCCGTAAGGACAATAAGCTT
614	Euc_010284_O_3	TGGGTTGTTGCGACGGGTTCTACAGATAAGACTGTTAAGTTATTTGATCTACGCAAGATC
615	Euc_010595_O_1	GCAGAGGTGCCTACATATGCTTTAGAATGCTAGTAGCTTGGAAGTGCAACACGCTCGTGA
616	Euc_010657_O_1	AGTAAAGTTTAACGACTATGCATCTGTCGTAGTATCAGCCGGCTATGATCGTTCAGTGCG
617	Euc_012636_O_2	CGTTAGGATAGTCTTTAAAGGAGTTGGTGATTATTGATTTCCACCCAATATATGTAGCGT
618	Euc_012748_O_2	GAGCAAGCTACTTACAAAAATCGACAGCGTCTTTACCTATCTGAACAGACAGATGGCAGT
619	Euc_012879_O_2	TCCTTCCGACAAGTACCGTATTGCAAGTTGTGGTATGGACAATACGGTTAAAATCTGGTC
620	Euc_015515_O_1	TTTCACTCGATGACGGTTGGCCGGATAAATAATCGCTTATATAGTCCTAATAAGTTCCAT
621	Euc_015724_O_3	ATATGTAGGTGGTAGAGGTGTGGATATTGCATAGACCGAACCTCCGCAGGTCCGCATTCT
622	Euc_016167_O_1	CCATTGAACTACTTATGGATTACTTTATACATGAAATATCATGCCGGAGTAATTTTGAGT
623	Euc_016633_O_3	AGCATTAGAGACCTGGATTTTAGTCTAGATTCAGAGTTTTTGGCTACGACATCTACTGAT
624	Euc_017485_O_3	AAAGGTTTATCCCTCATTGGATTTGATATATAAACTGAGAGTGTTTTGCCCCCCATTAAA
625	Euc_018007_O_1	GTACAGCGTGTATTTCTTGTTACGATACTTGAGGGGTTAGAGGCACCTACGAATTAGGAA
626	Euc_020775_O_3	ATATCCTTATGAATGAAGTTTGGATGATAAGTGGCGCCAGACTTTCTACTCACCCTTTTT
627	Euc_023132_O_3	TGATCACATCGTTGTTTGCAATAAGACGTCATCAATTTATATCATGACTCTACAGGGACA
628	Euc_023569_O_2	TTTTCCCAGTGTACTGCGAGAGTGATGCTACATAAGTTTACTCTTGTGTCTAACTTTTCC
629	Euc_023611_O_1	AGATTCTACAGATGGCGCTATACGAGCTGTTATACGGACATTTTATGACCATACACATCC
630	Euc_024934_O_3	TGCTACGGGAAACCAGGACAAAACTTGTAGGATTTGGGACATACGAAACTTATCTAAGTC
631	Euc_025546_O_1	CAAGTCATATAGTTACAGTGTCGCATGACAGAACAATTAAGCTCTGGACTAGTAACGACG
632	Euc_030134_O_2	TGCCACATCGTAACCATCATAGCACTTATCATCTAATTATGGTGAAAGGGAGTTATATAT
633	Euc_031787_O_5	GTTTATACTTATAAACAACAGAGAGACAACTGTACAGGTGTTGTAAACACTCCCAGTGTG
634	Euc_034435_O_1	CTGTGTTTTAGCCCGAGGGCCAATCACTTAGTTGCTACTTCGTGGGATAATCAGGTACGG
635	Euc_034452_O_3	GCAAAGTAGAGTTTAAGTTTCGTTGTGCTTGGACCGGAAAACTCACATGCTTAGAGTTTA
636	Euc_035789_O_5	AAGATTTGGGCATAACTTGTATGAACTTTTTCTGTTGTCGACACTGTAATTACACGAGCT
637	Euc_035804_O_4	AAACAGATGCATGTATGCTTCATAACTCTATAGATATGGAAATGTCACTGTACACTGATC
638	Euc_043057_O_2	TTATTGGTGCACAGGACGGAAAATTGCGCATATATTCTATTTCAGGTGATACATTAACAG
639	Euc_046741_O_1	AGGCACAGACACTTGCCTAAACCAATATACAAGGCAGGTATTCTAAGGCGCACCGTGAAT
640	Euc_047161_O_4	CATGCGAAGGTTTCTGGGAATTTTCAGTAGAAAATTCGGTCGTGGCGGCCATCCTCGATA
641	Pra_001766_O_1	TTAAGCTGATAGCTTTAGTTCCTACGTGGAATGTATAAATGCACCATTGTCCATAAGGCA
642	Pra_002927_O_2	GGATGCTCTGGTTACATGACTACTCCTTAGGGAATCAGTCAGACATTTTAAATAACTTCC
643	Pra_007642_O_2	TCATTAAGCGGTACTGGCAGAGGACATGTCTATTTATACAAGCAAATGGTCCTATTGGCT
644	Pra_013714_O_1	ATGTTGGTCAGACCTCAAATATTGTACTCCCCACACTAGGGAGCATTTACGGTGAATATA
645	Pra_016332_O_1	TCCTCTCGACCCTTAGAGTCCTCTGCGAATCTTGTTGTTAGTTACTGTGTACGCTGTAAC
646	Pra_021677_O_3	AAGCATGTTTTGAATTTATGGTGGTGGCATGTGGATATTTGAACTTGGTTGAGAAAAATT

Oligo SEQ ID	Oligo ID	Microarray Oligo Seq
647	Pra_027562_O_2	CATTCCTATTGAAGGGTCAACCTTTAATTTTGGCTAGCAGGACTGTATAGGATTATATGC
648	Pra_001504_O_2	TTATTGTATTTTAGATTCTTGATGGCCATCTAAACTTCTGGCTGCTTGGTGCAACATTGA
649	Pra_015211_O_2	ATAGCTAATGATTCCATGCTATCCATGGTATCTACTTCACGATAATAAAGGTCTTAGTCC
650	Pra_020421_O_2	CACCTAATAGGCCTGAGTATTGCTCACCACTATGCTGATATGGGGAGCAATAACGTTAGT
651	Pra_003187_O_2	TTTCTTTTCACTTTGTACTAATGATCATTGTGACCACAAAATCTTTATACACAATACAGA
652	Pra_015661_O_1	CTTGTCACTATCCTCATATTGATATCACCTCGTGTATGTTGTGGGGTGGCAAAATTACTT
653	Pra_013874_O_1	TATTTTAACTCAGCGACTTACCAGCCTAGTAAGCAATGGGGAGCTTGCATGTATTAGTTT
654	Pra_014615_O_1	ATTCGTCCTGGTCCTTTAGGACATGTACTTATGTCCATGCAAGTGCTTCTTGCCTAAGCT
655	Pra_004578_O_2	TTCTAGGCGATATATATCGCCGTAACTTTGGATGTGTTAAGAATATAGGGGATCATTAGC
656	Pra_023387_O_3	AGTTGCAGAGTGTGTAGCAACTGATGAGCATAGTTGTTATGTTTCTCAACTCAGTTGCAC
657	Pra_006970_O_1	AAGAAACTCATACACTGGACAGGCCAACCTTCCAAATATGTGTTTAGAAAACCTTTGTCT
658	Pra_010322_O_1	AAGGGGTGCTATCCATATCTAGAATCTACCATGCTCAATGAGGTATCTTCATTAGTATAC
659	Pra_022721_O_1	ATCTAATGCTAGTTTATTGATTTCTATGATCCAAGACCTCGTCATAGATCAAGTGCCTAG
660	Pra_023407_O_1	TTGTTATTAAATACCATTCAATATGCTTATGATTCATGAATGCTTAAGAGATTCTGCTGC
661	Pra_001945_O_2	GCTTCTAAACTGTAGAAGCCTGTTATCTTTAGACTCGTGGTTATGTGAACTACTTTTACA
662	Pra_008233_O_1	GGCTGTGGGGATTCGAGCCTGATGGTTATGCACTGTGGCCAGCAAGATGTTGAAGTTTTA
663	Pra_008234_O_4	GCCTGATGGTTATGCACTGTAAGTGATCTGATTTGATTAACTATTTTATCAATTAATTTT
664	Pra_022054_O_2	ATGGTCATTATCCGAGATAGTGCGCTTTGTCATGGGAAAATGACTATTGAATGTGAGTTT
665	Pra_012137_O_2	TTTTCTGGTGCATCCTTAACACAGCTTGGTTACATGGTGAATTACAGTATTTGAAGGAGT
666	Pra_012582_O_2	AGATTTAATGCCACTTAGGTGATCGGTGACCCACTTGTACATATAGATGTTGGCGATGTT
667	Pra_015285_O_2	AAGAAATTCATCAATTCTTTGAAATTATTGTTCCCTTTTGATGCGGCCCCTTTCTGGAGG
668	Pra_017229_O_1	TAAAGTATATTTTAGCCGCTGTTGTTGTAAATTTATGTTTTTCATTGCTATCAACATTTA
669	Pra_020724_O_2	GGTTTTCCTATAAGATGTATGAATTCGCACTGTGGTGCAATTTTATGAATTAAACTCAAA
670	Pra_004555_O_1	TTTACTATTCCGTCTGGGCTTAGAGATGTACGTTAATTGGTCATTTAAGACGACTCAGTT
671	Pra_004556_O_5	TCAAATCTAGTCAATATCCGTGTTGAGCTAAACAAGCGCTGAAAGTTTGCTCGAATCAGC
672	Pra_005729_O_2	AGAAAGTTGTGTACTAATTTGTATTGTAACGTCCATTTATCCAACGAGTCCTCCATTCAT
673	Pra_007395_O_3	CAGTACTGTATTCGAAGATCCTGAAAATTTACTAAAACAAATGGAATATCAACAACCTAG
674	Pra_009503_O_1	TTGCTCTATATAATTTGTGCTCGTGTGTGTACTTGAAGATCCATCCTCACATAGTCCAAT
675	Pra_011283_O_1	GTGTGTATAGTTTTATAACACTCTATGGTATCACTACCACTATGGGCCTGTTTAGTCCAA
676	Pra_012322_O_3	GAAGCAGAATCAGCTTTGACCAGTATTTAGTGTCTTGTATACAATTCTTGTTTCAGTGAA
677	Pra_023236_O_3	AAATCAAGATTAAAATCCGAAACCAAGGCTAACCAGCAAACTGTGAGGTGTACATTGTTG
678	Pra_000171_O_2	TTCCAAGCAGAAGGGCACATGTTGTGACATCAAGTAGTAGATTGTTCTGCAGATTCTGGT
679	Pra_000172_O_1	GTTAATGTAATACATTTAGTTTTTAGATAACTGTTAATGTGTAGTAAAGCACTAGGAAGA
680	Pra_001480_O_3	GAGGCTTCAAAGGTTTTTGTGTCTTTTCTAGTTATTATAAACGCTTCATAGGTTCCTAGG
681	Pra_001692_O_2	GAAGATTGTAAGTTGGGTGAACTTTTTTACCACGCTAGGTTGATCTATTTTAAGACTCTT
682	Pra_005313_ORF_01	AAAATAGCTGCGCGTACCACAAAGGTGACAAACGCCGGATTTCTCTTATCAGACTTGTCA
683	Pra_006362_O_1	TTTAATTATCATAGTTTTATTCCGGCTATCTTGATCATTCACGGAAGTCCCGAGAGTCAA
684	Pra_006493_O_3	GTGGAGTGAACGTGGTTACTTCAATGGATTACCCTTCTATCGTGTCATTAAACACTTTGT
685	Pra_006983_O_1	GCTAACTCTTCTAGTTGAGATCTCCATCAATTAATGGATACAAACATTGAGTTTCACTTT
686	Pra_007665_O_1	GGATCACTACTGGATTCCGTTACATTAGTTATTGCAAGTTGGTTATTATGTACGTTTATA
687	Pra_012196_O_1	ATGAACAAATGCAATTACCCTGTTTTATTCTATCCCGCTTTAATTAATATTGGTCATGTT
688	Pra_013382_O_1	TTTGCTTGTGGATTGTACTGTGGTACATGGTATAAATCTATAGGCTATGTCGATTATTTT
689	Pra_016461_O_1	ATATAAGATATAAGATATTGCCAGCAAACTATTTGACAGGTTATTTAATAAAGTGTGCTA

Oligo SEQ ID	Oligo ID	Microarray Oligo Seq
690	Pra_017611_O_1	TTTTAAATGTGGACAGAGGCACTATAAGAATGCGAAATATCGTCGGAGCACGACTAATTG
691	Pra_019776_O_1	ATAGACTAGTTCTACAAAGCCCTAGGATGATGGACTTCATTTCTTTTGCATTAAGATGAA
692	Pra_020659_O_1	GATTTCTTATGGGGTTGGAACATTCCTCGCTGCCTTCTGGTAATATTAGGTTATGCGTTT
693	Pra_022559_O_3	AATTGAGGTTGACTGTGTACTTCTCCAGTGGACAGGAGAAAGCGATAAAATTCAAACGTT
694	Pra_024188_O_5	AAGGAAGGGCAAATAGAGCTCGCGCTCAAGAAATACCTTAAATCGATACGGTATTTGGAT
695	Pra_027973_O_2	TAATTTAAGAGCTATGAAACAACTACCTTTTGGAATGGTTTTGTTTTTAGCATCCCAATT
696	Pra_001353_O_1	TTGTAAATTATGCTGGTTCCATATGGGGGTTAATCAGTATCCTGGTTATTTGTGACACCA
697	Pra_001978_O_3	GTTGTGAACTATCAATAGACGGGGATGGTCCTTTTTAGCTGCTCCTTAAGCAGCTCAAAT
698	Pra_002810_O_2	TCAATTCCGGTCATATGTAGACGACTATAATGTTGTTTGTGTCCTATAACTATAGTGTTG
699	Pra_002811_O_1	CATTTTACACCCTATAACAAAATATAGTGTCATAAGTTTACACCAGGTAACAACTCTATA
700	Pra_002812_O_3	ATGGAGAGTTTTATTCATTACATGAAAGAGTATGTCACCTTTCGTGCTCCATCTATTGAT
701	Pra_003514_O_1	TTTCACGTCCTGTATACTCACTCAAGCAACTTTAGGATGAAGAGCTAAAGTATATCAAAG
702	Pra_004104_O_2	AATGCACTCTTTATAAAGTGGGATGAGGTATGTGTTTCCTTCCTATTGGCTAACCTGAAT
703	Pra_005595_O_1	ATTGGGCAATCGTTATTGATTTTACCTATCGCTATCTCACTGTCCGCCAATTTAGTGTAA
704	Pra_005754_O_1	TTTCAGCGGATATAAAGTCTTCCAACTTGTAAACCGGTGCTGTGAAGATTAAAAGTCCTT
705	Pra_006463_O_1	GCTTTAGAGGCAATGGTAGATTATGAAGTCAACACCAGGGAGTTTGACCGTTTGGGACAT
706	Pra_006665_O_1	CATTCAATTTGACATTGGAGTTTCAAGGCATTCCAAGGATAGCATGTACACAAGTTGAAT
707	Pra_006750_O_1	CATAAAATTACTATGGAAGTTGGATCATTATCTATGCCATAGTGGAGTAGAACTAGATTT
708	Pra_007030_O_1	CTCTTGATTCTAGAATCTAAACTACTACCTTGCGGACATGACTGAGCATCTCTCTAACAG
709	Pra_007854_O_1	CAGGGTTGTGCTAGTTTAACATTTTAACTTAATGTAATCATGTAAGCTTTAGAGAGGTGG
710	Pra_007917_O_1	GTAAATGTTTACATTGAGGTCATGCATGAGTGTTAATTACGCTTTCACTACTGTTCACTT
711	Pra_007989_ORF_02	AATTAAAGCTTGGTTGTATGATCATTTGGGATCGAGAGTAGATTATGATGCTCCTGGGCA
712	Pra_008506_O_1	TTATCTAGCTAGAAGTTGTGAAATTAAGAGGGATGTGAGGATTGGGTTATAACTAGTGTA
713	Pra_008692_ORF_02	AATGAATCAGGCATTAAAGCGGGAATCATTTATGACTTGGCAACCTGAAAATTCTATTAA
714	Pra_008693_O_2	TTCTTGACGTTTTAATATGGTATGGTATTAAATTTGGAAGGCCTATTCGATTGTTTGCAA
715	Pra_009170_O_1	TTCTTATAACCTGTACGATTGCCGATATATCACCAATTTTGCTGATTTTAATCTGAGTTT
716	Pra_009408_O_1	CAATTTCATATTCGGGTTCAATGTAGTGCCTCTCATTTTAGGGTGATAGCATGAGTTTTT
717	Pra_009522_O_1	TCCACAAGTTAACATAGGTAACTATCGACTGAAGTGAACTGGGGGGCAGAAGCTAACTAT
718	Pra_009734_O_2	TTTAGATAGCCATTTACATTTTACTTATTATTGGACTTGTAAAGATTTTTGTACCCTTGT
719	Pra_009815_O_4	TTGCTGAAATATTTCAAGCTGAAAGTTATGATTCTGGCCAAGAAGTCTACTGAAAATTTG
720	Pra_010670_O_2	AAACATAAGTTTGGCCCAGATTCGGTTTATCATAAAATCTGGCTGCATATAAGGTGTCAG
721	Pra_011297_O_1	ATGTTCTAGAATTTGTCTAAGCTAGCTACTGGTGTTTAACTGATATGGAAAACTTTTGCC
722	Pra_013098_O_2	TTTGGGGAGTACTTTAGTCAATAAAAGTGAAGTGAATCATGATATAAAGGGTTTAAGTAA
723	Pra_013172_O_2	AGAAGTTACTAATTTGTAGATAAATTCTAACGAAGGTGATGATAGCATACACGTAATGAA
724	Pra_013589_O_2	GAATTTTGATGGTAGCGTATGGTTGAAGGAAAACTTGGATATATCATGTAAACATTTTTC
725	Pra_013608_O_1	TTAATGAACCGCTTTTTCCTTGAGAGGCTATGAATGCCTGTAGAACTAATCCTTTAAGTA
726	Pra_014299_O_2	TTTCTCTAACACTATATTTTCTGGTATGACCGCTCTACATTGTATATTAACCCTTGCAAA
727	Pra_014498_O_1	TATATTCACTGTGCTGGGATTATCCTCTCCCCTTTTTGACCCACTGTTGTGTGTATTTGA
728	Pra_014548_O_1	GAGCATACAGCGTTATCTTTGAGACGAGTCATCAATGATAATATCCTCGTAAAAGGTTAC
729	Pra_014610_O_2	TTTATTCAATTACGACGGATTCAGTTGGCCTTTTGTAACATTCAAGTATCCATCTATCAC
730	Pra_016090_O_2	ATGTTCAGGGGTATTAAAAATTCAGAGGATAAATTTCCTCACTCTCAAGTGTTAGATGGT
731	Pra_016722_O_2	CAAAGTCTAGACGTTAATGTTTTGGAACTCTTTTTTCGAATTTGTGCCTATTGAATCACT
732	Pra_016785_O_3	TATAAATATATTGTACTGGGGATCCAAGACATGGCAATATATGTCGAGATTTTCATTTTC

Oligo SEQ ID	Oligo ID	Microarray Oligo Seq
733	Pra_017094_O_3	CTTTTGCATGAGTTCAAATGTCTTTGTGACATATTGTCTTGAACCACCGAGGATATATCA
734	Pra_017527_O_2	GTTTGTATGTCCAATAGATTATAACCTATTTACTGTGACACTATTCTTCACACCCATGTC
735	Pra_017591_ORF_O2	AGATCTAGTTGTTTCAGCATCGTTGGACCAAACTGTTCGTGTATGGGATATAAGTGGCCT
736	Pra_017769_O_2	TGCCGTATCAAAAGATTGGTACTTCCTTATGGACACACAAGATCGTAAGCATGGCTGAAT
737	Pra_018047_O_2	TTGATGGCCACATGAGTTGTTTATACAAGTCGTTGTTTTATGAGAGAACCTTCTTCAGAT
738	Pra_018414_O_1	ATTTCTATAGTGCCATATGCTTGTCGGTTGTCATTGACCTCTAATAGAATAGCCAGAGTA
739	Pra_018986_O_1	TTCACGGCAGTTGAACTAGTCATAGTGGAATATTATTTAAATGGTGTATTCTAGTCACAT
740	Pra_019479_ORF_O1	TGCAGGCGCTCTATAGTTCTGTTCTCTAGCATGAAGTGTGTATTTTATCTATTGTGGACC
741	Pra_020144_O_1	TGTCTTTAATCTTCAGGGTTCGTTACTAACAATTGAGCTCAAATCTCTATTCTGACCAGC
742	Pra_022480_O_1	CATTTATAGAGTTGTGCAAAATCACCCATAATGCTATGAATTGACAGGTGACTGTAATCT
743	Pra_023079_O_2	GGAGAAAATTTCCTATCCCTTTGTGGGTGTGTGAAAAACGAAATATAGAGGAACAATGTG
744	Pra_026739_O_2	ACCAATCATTTATTTGCAGTGTAGTTGATATGAAGGGAGAAATATGACAGTTGGTTTCAA
745	Pra_026951_O_2	AAGTTAATGTTCTCATAGGTTATTCATTGGAGTTGTCTCGTATGTACGCTGTGCCGTAGT
746	Pra_026529_O_2	CTCATAAATTGAGGCTTGCCTACGTTAATTGTTATATATGGAGAGCCATGCTAATTGTTA
747	Euc_006366_O_2	GCAGATCATGTAATTGTATCTCAAATTATAGTATCCGTATTCTGTACAAATGCTCCGGAA
748	Euc_017378_O_1	TCTTTACGCAGATGGTGACTGAAGCTGGTTCCGAGATCGGCATATGTAGCTGGTAGAGGT
749	Pra_000888_O_1	TTCACATTGAGGGTTGCCGTCGGTATTCGCCGATGATATCCTGTTTTACGCGCAACAGTT
750	Pra_014166_O_1	TCATTATTTAGGGTGCAGGCTGTATAAAATGTTGTAAATTGTAGTATCAATGTGTACAAT
751	Pra_003189_O_1	GCATTCACCACGACAGTAAAGTAATCATTATGATTACTAATGTATTGCTTTCATGGGGTG
752	Pra_009356_O_4	AAAGGGTATATTTTGTCTCATGTTGGGGTGATAATTCTCCCTGAAAGTCTCCAAAATATA
753	Pra_000065_ORF_O_2	AAATTTCCGGTTGCCATAGTCTAGTGGGGTGAGGGTTCATTCTAGGGGATTTATTGTGTT
754	Pra_014197_ORF_O1	GCAGTGATAAAGGTACTTCTTGGTGATAATCCTAAAGCCTTACCCATGGATATCCAGCCT
755	Pra_009081_O_2	TTCTTTAACAAGGTAAAAATCCCCCCCCTTGGCATGTAGCTCAATTAGTTGTAATGGAACT
756	Pra_013417_O_1	AGTTGTAAACAGTGTAATAAGGAGCAGAAGTTGTGATAGCTTTTAGGAACGATAGACTTT
757	Pra_005755_O_1	TGAACCAATTCTTGTATATTAGATATGTAACATGTATGAATGTCCATAGAGCAGAGCTTT
758	Pra_006670_O_2	AGCCAGGCACGCTTAACTAAATTTCGTTTAGTTCACCATGACTATTCGTTGAACTTAATG
759	Pra_007027_O_1	CAAAACCCCTTGTAGGGTGGACTTCTGTTGTATCCAATTTTTATGGCATAATTAGCTAGT
760	Pra_007276_O_1	AATTTGGTGATTATTCCTTACCATATCGTACTGTACAGATACGGTAAGGTCGAAATATAT
761	Pra_007390_ORF_O1	CATGCCGTGATCGGTCGATTGCATTAAGTGCTGCAAGGATCAAATAGTGGCACTGTCATG
762	Pra_012648_ORF_O1	CAAACATAAATAAGGTTGCTACTTTAAAGGGACATACGGAACGAGTTACTGATGTGGCAT
763	Pra_013171_O_2	ATTTATGGATGAGGTACTCCTTATGAATATCTTCAAACTAAGAAATAACTATATATGCAA
764	Euc_045414_O_2	CTTGGTTTTTGTTGAGCTTTCTATTTCAAGCAATTTGTGATTGGGGGGTTCTGCATTCTT
765	Euc_044328_O_2	ATGTCTAAAGAGCCGTGATCTATGAGTAGATTAGAAACCGCCTTTTTAGTTGCAAACGCC
766	Euc_015615_O_2	TTGCAACAAGGTATACTTAGTCAGTCCTTGTTATGTATGTCTTTTGTCAACCCTTCAGGG
767	Euc_017239_O_3	GGCGGAATCCCTTTGTTCTTTCGAGCTTTACGTGACAAGTCGGCCAGAAAGCAGTAGCAT
768	Euc_018643_O_3	TTGATGTACGAGCCGCTATATCTAATTCTGCCTCCCAGTCACTGCCAAGTTTTACTCTTC
769	Euc_019127_O_5	GTCTTGCATGTCAGCTATTATACAGTCCTGTTTATAGTCCTGTGATGTAATAAAAAGCTG
770	Euc_022624_O_3	AAGTAGGAGATCGTGTAGAGAGAATACTTTCTGCTCTCAGCGGCGAAGAGGTTTGTCTGC
771	Euc_032424_O_1	AATTGTGAGTAGAATAGGAGAAACTTTTGTACAAGATTAATACGTGTGGCATAATAAGAT
772	Euc_037472_O_1	TGATGTGCAGTTTACATTATTATGGTTCGAGTATTATTTAGCTGCCCTATCTTAAGTCAT

[0395] Table 15. Peptide Table.

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
261	CDK type A	MGDGSLGSGGRGNSGGGGGGGSRPEWLQQYDLIGKIGEGTYGLVFLARIKHPST NRGKYIAIKKFKQSKDGDGVSPTAIREIMLLREISHENVVKLVNVHINPVDMSL YLAFDYADHDLYEIIRHHRDKVNQAINPYTVKSLLWQLLNGLNYLHSNWIIHRD LKPSNILVMGEGEEQGVVKIADFGLARVYQAPLKPLSDNGVVVTIWYRAPELLL GAKHYTSAVDMWAVGCIFAELLTLKPLFQGQEVKANPNPFQLDQLDKIFKVLGH PTQEKWPMLVNLPHWQSDVQHIQRHKYDDNALGNVVRLSSKNATFDLLSKMLEY DPQKRITAAQALEHEYFRMEPLPGRNALVPSSPGDKVNYPTRPVDTTTDIEGTT SLQPSQSASSGNAVPGNMPGPHVVTNRPMPRPMHMVGMQRVPASGMAGYNLNPS GMGGGMNPSGIPMQRGVANQAQQSRRKDPGMGMGGYPPQQKQRRF	387	1820
262	CDK type A	MEKYQQLAKIGEGTYGIVYKAKDKKSGELLALKKIRLEAEDEGIPSTAIREISL LKQLQHPNIVRLYDVVHTEKKLTLVFEFLDQDLKKYLDACGDNGLEPYTVKSFL YQLLQGIAFCHEHRVLHRDLKPQNLLINMEGELKLADFGLARAFGIPVRNYTHE VVTLWYRAPDVLMGSRKYSTQVDIWSVGCIFAEMVNGRPLFPGSSEQDQLLRIF KTLGTPSLKTWPGMAELPDFKDNFPKYVVQSFKKICPKKLDKTGLDLLSRMLQY DPAKRISAEQAMGHPYFKDLKLRKPKAAGPGP	99	1007
263	CDK type A	MDQYEKIEKIGEGTYGVVYKAIDRSTNKTIALKKIRLEQEDEGVPSTAIREISL LKEMQHGNIVKLQDVVHSERRLYLVFEYLDLDLKKHMDSCPEFSKDTHTIKMFL YQILRGISYCHSHRVLHRDLKPQNLLLDRRTNSLKLADFGLARAFGIPVRTFTH EVVTLWYRAPEILLGSRHYSTPVDVWSVGCIFAEMVNRRPLFPGDSEIDELFKI FRIMGTPNEDSWPGVTSLPDFKSTFPKWASQDLKTVTPTVDPAGIDLLSKMLCM DPRRRITAKVALEHEYFKDVGVIP	120	1004
264	CDK type A	MVMKSKLDKYEKLEKLGEGTYGVVYKAQDKTTKEIYALKKIRLESEDEGIPSTA IREIALLKELQHPNVVRIHDVIHTNKKLILVFEFVDYDLKKFLHNFDKGIDPKI VKSLLYQLVRGVAHCHQQKVLHRDLKPQNLLVSQEGILKLGDFGLARAFGIPVK NYTNEVVTLWYRAPDILLGSKNYSTSVDIWSIGCIFVEMLNQKPLFPGSSEQDQ LKKIFKIMGTPDATKWPGIAELPDWKPENFEKYPGEPLNKVCPKMDPDGLDLLD KMLKCNPSERIAAKNAMSHPYFKDIPDNLKKLYN	23	937
265	CDK type A	MDQYEKVEKIGEGTYGVVYKAIDRLTNETIALKKIRLEQEDEGVPSTAIREISL LKEMQHGNIVRLQDVVHSENRLYLVFEYLDLDLKKHMDSSPDFAKDPRLVKIFL YQILRGIAYCHSHRVLHRDLKPQNLLIDRRTNALKLADFGLARAFGIPVRTFTH EVVTLWYRAPEILLGSRHYSTPVDVWSVGCIFAEMVNQRPLFPGDSEIDELFKI FRILGTPNEDTWPGVTALPDFKSAFPKWPAKNLQDMVPGLNSAGIDLLSKMLCL DPSKRITARSALEHEYFKDIGFVP	149	1033
266	CDK type B-1	MEKYEKLEKVGEGTYGKVYKAKDKATGQLVALKKTRLEMDEEGVPPTALREVSL LQLLSQSLYVVRLLSVEHVDGGSKRKPMLYLVFEYLDTDLKKFIDSHRKGPNPR PVPAATVQNFLYQLLKGVAHCHSHGVLHRDLKPQNLLVDKEKGILKIADLGLGR AFTVPLKSYTHEVVTLWYRAPEVLLGSAHYSIGVDMWSVGCIFAEMVRRQALFP GDSEFQQLLHIFRLLGTPTEKQWPGVTTLRDWHVYPQWEPQNLARAVPSLGPDG VDLLSKMLKYDPAERISAKAALDHPFFDSLDKSQF	199	1116

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
267	CDK type B-2	MERPATAAVSAMEAFEKLEKVGEGTYGKVYRAREKATGKIVALKKTRLHEDEEG VPPTTLREISILRMLSRDPHIVRLMDVKQGQNKEGKTVLYLVFEYMETDLKKYI RGFRSSGESIPVNIVKSLMYQLCKGVAFCHGHGVLHRDLKPHNLLMDKKTLTLK IADLGLARAFTVPIKKYTHEILTLWYRAPEVLLGATHYSTAVDMWSVGCIFAEL VTKQALFPGDSELQQLLHIFRLLGTPNEKMWPGVSSLMNWHEYPQWKPQSLSTA VPNLDKDGLDLLSQMLHYEPSRRISAKAAMEHPYFDDVNKTCL	41	982
268	CDK type C	MGCVLGREVSSGIVTESKGRDSSEVETSKRDDSVAAKVEGEGKAEEVRTEETQK KEKVEDDQQSREQRRRSKPSTKLGNLPKHIRGEQVAAGWPSWLSDICGEALNGW IPRRANTFEKIDKIGQGTYSNVYKAKDLLTGKIVALKKVRFDNLEPESVRFMAR EILILRHLDHPNVVKLEGLVTSRMSCSLYLVFEYMEHDLAGLAASPAIKFTEPQ VKCYMHQLLSGLEHCHNRRVLHRDIKGSNLLIDNGGVLKIGDFGLASFYDPDHK HRMTSRVVTLWYRPPELLLGANDYGVGIDLWSAGCILAELLAGKPIMPGRTEVE QLHKIYKLCGSPSEEYWKKYKLPNATLFKPREPYRRCIRETFKDFPPSSLPLIE TLLAIDPAERGTATDALQSEFFRTEPYACEPSSLPQYPPSKEMDAKKRDDEARR LRAASKGQADGSKKERTRDRRVRAVPAPEANAELQHNIDRRRLISHANAKSKSE KFPPPHQDGALGFPLGASHRFDPAVVPPDVPFTSTSFTSSKEHDQTWSGPLVDP PGAPRRKKHSAGGQRESSKLSMGTNKGRRADSHLKAYESKSIA	291	2042
269	CDK type C	MYSKSSAVDDSRESPKDRVSSSRRLSEVKTSRLDSSRRENGFRARDKVGDVSVM LIDKKVNGSARFCDDQIEKKSDRLQKQRRERAEAAAAADHPGAGRVPKAVEGEQ VAAGWPVWLSAVAGEAIKGWLPRRADTFEKLDKIGQGTYSSVYKARDVTNNKIV ALKRVRFDNLDTESVKFMAREIHILRMLDHPNVIKLEGLITSRMSCSLYLVFEY MEHDLTGLASRPDVKFSEPQIKCYMKQLLSGLDHCHKHGVLHRDIKGSNLLIDN NGILKIADFGLASVFDPHQTAPLTSRVVTLWYRPPELLLGASRYGVEVDLWSTG CILGELYTGKPILPGKTEVEQLHKIFKLCGSPSDDYWRRLHLPHAAVFKPPQPY RRCVAEIFKELPPVALGLLETLISVDPSQRGTAAFALRSEFFTASPLPCDPSSL PKYPPSKEIDMKLREEEARRRGAAGGKNELEKRGTKDSRTNSAYYPNAGQLQVK QCHSNANGRSEIFGPYQEKTVSGFLVAPPKQARVSKETRKDYAEQPDRASFSGP LVPGPGFSKAGKELGHSITVSRNTNLSTLSSLVTSRTGDNKQKSGPLVSESANQ ASRYSGPIREMEPARKQDRRSHVRTNIDYRSREDGNSSTKEPALYGRGSAGNKI YVSGPLLVSSNNVDQMLKEHDRRIQEHARRARFDKARVGNNHPQAAVDSKLVSV HDAG	107	2236
270	CDK type C	MGCIPTIISDGRRRSAAPDKRRPRPRRSSSEGEAPPHATAAGSEGGESARGAPG KERPEPAPRFVVRSPQGWPPWLVAAVGHAIGEFVPRCADSFRKLAKIGEGTYSN VYKARDLVTGKTVALKKVRFDNLEAESIKFMAREILVLTRLNHPNVIKLEGPVT SRMSSGLYLAFEYMEHDLSGIAARQNGKFTEPQVKCFMRQLLSGLEHCHNHDVL HRDIKCSNLLIDNEGNLKIADFGLATFYDPERKQVMTNRVVTLWYRAPELLLGA TSYGIGIDLWSAGCILAELLYGKPIMPGRTEVEQLHKIFKLCGSPSEAYWNKFK LPNANIFKPPQPYARCIAETFKDFPPSALPLLETLLSIDPDERGTATTALNSEF FAAEPHACEPSSLPKYPPSKEMDLKLIKEKTRRDSSKRPSAIHGSRRDGIHDRA GRVIPAPEATAENQATLHRPRAMKKANPMSRSEKFPPAHMDGVVGSSANAWLSG PASNAAPDSRRHRSLNQNPSSSVGKASTGSSTTQETLKVAPELLQVGSSSLHPC HRMLVYGSNLTIRSK	82	1749

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
271	CDK type C	MGCICAKQADRGPASPGSGILTGAGTGTGTRSSKIPSGLFEFEKSGVKEHGGRS GELRKLEEKGSLSKRLRLELGFSHRYVEAEQAAAGWPSWLTAVAGDAIQGLVPL KADSFEKLEKIGQGTYSSVFRARELANGRMVALKKVRFDNFQPESIQFMAREIS ILRRLDHPNIMKLEGIITSRMSNSIYLVFEYMEHDLYGLISSPQVKFSDAQVKC YMKQLLSGIEHCHQHGVIHRDVKSSNILVNNEGILRIGDFGLANILNPKDRQQL TSHVVTLWYRPPELLMGSTSYGVTVDLWSVGCVFAELMFRKPILRGRTEVEQLH KIFKLCGSPPDGYWKMCKVPQATMFRPRHAYECTLRERCKGIATSAMKLMETFL SIEPHKRGTASSALISEYFRTVPYACDPSSLPKYPPNKEIDAKHREEARRKKAR SRVREAEVGKRPTRIHRASQEQGFSSNIAPKEKRSYA	151	1560
272	CDK type C	MAVAAPGHLNVNESPSWGSRSVDCFEKLEQIGEGTYGQVYMAKEKKTGEIVALK KIRMDNEREGFPITAIREIKILKKLHHENVIKLKEIVTSPGPEKDEQGRPEGNK YKGGIYMVFEYMDHDLTGLADRPGMRFSVPQIKCYMRQLLTGLHYCHINQVLHR DIKGSNLLIDNEGNLKLADFGLARSFSNDHNANLTNRVITLWYRPPELLLGATK YGPAVDMWSVGCIFAELLHGKPIFPGKDEPEQLNKIFELCGAPDEINWPGVSKI PWYNNFKPTRPMKRRLREVFRHFDRHALELLERMLTLDPSQRISAKDALDAEYF WADPLPCDPKSLPKYESSHEFQTKKKRQQQRQHEETAKRQKLQHPPQHPRLPPV QQSGQAHAQMRPGPNQLMHGSQPPVATGPPGHHYGKPRGPSGGAGRYPSSGNPG GGYNHPSRGGQGGSGGYNSGPYPPQGRAPPYGSSGMPGAGPRGGGGNNYGVGPS NYPQGGGGPYGGSGAGRGSNMMGGNRNQQYGWQQ	82	1644
273	CDK type C	MGCICTKGILPAHYRIKDGGLKLSKSSKRSVGSLRRDELAVSANGGGNDAADRL ISSPHEVENEVEDRKNVDFNEKLSKSLQRRATMDVASGGHTQAQLKVGKVGGFP LGERGAQVVAGWPSWLTAVAGEAINGWVPRRADSFEKLEKIGQGTYSSVYRARD LETNTIVALKKVRFANMDPESVRFMAREIIIMRKLDHPNVMKLEGLITSRVSGS LYLVFEYMDHDLAGLAATPSIKLTESQIKCYMQQLLRGLEYCHSHGVLHRDIKG SNLLVDNNGNLKIGDFGLATFFRTNQKQPLTSRVVTLWYRPPELLLGSSDYGAS VDLWSSGCILAELFAGKPIMPGRTEVEQLHKIFKLCGSPSEEYWKKSKLPHATI FKPQQPYKRCLLETFKDFPSSALGLLDVLLAVEPECRGTASSALQNEFFTSNPL PSDPSSLPKYPSSKEFDARLRDEEARKHKATAGKARGLESIRKGSKESKVVPTS NANADLKASIQKRQEQSNPRSTGEKPGGTTQNNFILSGQSAKPSLNGSTQIGNA NEVEALIVPDRELDSPRGGAELRRQRSFMQRRASQLSRFSNSVAVGGDSHLDCS REKGANTQWRDEGFVARCSHPDGGELAGKHDWSHHLLHRPISLFKKGGEHSRRD SIASYSPKKGRIHYSGPLLPSGDNLDEMLKEHERQIQNAVRKARLDKVKTKREY ADHGQTESLLCWANGR	626	2782
274	CDK type D	MDPDPSPDPDPPKSWSIHTRREIIARYEILERVGSGAYSDVYRGRRLSDGLAVA LKEVHDYQSAFREIEALQILRGSPHVVLLHEYFWREDEDAVLVLEFLRSDLAAV IADASRRPRDGGGGGAAALRAGEVKRWMLQVLEGVDACHRNSIVHRDLKPGNLL ISEEGVLKIADFGQARILLDDGNVAPDYEPESFEERSSEQADILQQPETMEADT TCPEGQEQGAITREAYLREVDEFKAKNPRHEIDKETSIFDGDTSCLATCTTSDI GEDPFKGSYVYGAEEAGEDAQGCLTSCVGTRWFRAPELLYGSTDYGLEVDLWSL GCIFAELLTLEPLFPGISDIDQLSRIFNVLGNLSEEVWPGCTKLPDYRTISFCK IENPIGLESCLPNCSSDEVSLVRRLLCYDPAARATPMELLQDKYFTEEPLPVPI SALQVPQSKNSHDEDSAGGWYDYNDMDSDSDFEDFGPLKFTPTSTGFSIQFP	13	1467

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
275	CDK type D	MDPDPSPSPDPPKSWSIHTRREIIARYEILERVGSGAYSDVYRGRRLSDGLAVA LKEVHDYQSAFREIEALQILRGSPHVVLLHEYFWREDEDAVLVLEFLRSDLAAV IADASRRPRGGGVAPLRAGEGKRWMLQVLEGVDACHRNSIVHRDLKPGNLLISE EGVLKIADFGQARILLDDGNVAPDYEPESFEERSSEQADILQQPETMEADTTCP EGQEQGAITREAYLREVDEFKAKNPRHEIDKETSIYDGDTSCLATCTTSDIGED PFKGSYVYGAEEAGEDAQGSLTSCVGTRWFRAPELLYGSTDYGLEVDLWSLGCI FAELLTLEPLFPGISDIDQLSRIFNVLGNLSEEVWPGCTKLPDYRTISFCKIEN PIGLESCLPNCSSDEVSLVRRLLCYDPAARATPMELLQDKYFTEEPLPVPISAL QVPQSKNSHDEDSAGGWYDYNDMDSDSDFEDFGPLKFTPTSTGFSIQFP	113	1558
276	Cyclin A	MSNQHRRSSFSSSTTSSLAKRHASSSSSSLENAGKAFAAAAVPSHLAKKRAPLG NLTNLKAGDGNSRSSSAPSTLVANATKLAKTRKGSSTSSSIMGLSGSALPRYAS TKPSGVLPSVNPSIPRIEIAVDPMSCSMVVSPSRSDMQSVSLDESMSTCESFKS PDVEYIDNEDVSAVDSIDRKTFSNLYISDAAAKTAVNICERDVLMEMETDEKIV NVDDNYSDPQLCATIACDIYQHLRASEAKKRPSTDFMDRVQKDITASMRAILID WLVEVAEEYRLVPDTLYLTVNYIDRYLSGNVMNRQRLQLLGVACMMIAAKYEEI CAPQVEEFCYITDNTYFKEEVLQMESSVLNYLKFEMTAPTVKCFLRRFVRAAQG VNEVPSLQLECMANYIAELSLLEYDMLCYAPSLVAASAIFLAKFVITPSKRPWD PTLQHYTLYQPSDLGNCVKDLHRLCFNNHGSTLPAIREKYSQHKYKYVAKKYCP PSIPPEFFHNLVY	187	1686
277	Cyclin A	MNKENAVGTKSEAPTIRITRSRSKALGTSTGMLPSSRPSFKQEQKRTVRANAKR SASDENKGTMVGNASKQHKKRTVLNDVTNIFCENSYSNCLNAAKAQTSRQGRKW SMKKDRDVHQSGAVQIMQEDVQAQFVEESSKIKVAESMEITIPDKWAKRENSEH SISMKDTVAESSRKPQEFICGEKSAALVQPSIVDIDSKLEDPQACTPYALDIYN YKRSTELERRPSTIYMETLQKDVTPNMRGILVDWLVEVSEEYKLVPDTLYLTVN LIDRSLSQKFIEKQRLQLLGVTCMLIASKYEEICPPRVEEFCFITDNTYTSLEV LKMESRVLNLLHFQLSVPTVKTFLRRFVQAAQVSSEVPSVELEYLANYLAELTL VEYSFLKFLPSLMAASAVLLARWTLNQSDNPWNLTLEHYTKYKASELKAAVLAL EDLQLNTSGSTLNAIREKYRQQKVNYSLLIHSKANHEIL	238	1653
278	Cyclin B	MAGSDENNPGVVGGAHVQEGLRVGAGKMGAGNVQQRRALSNINSNIIGAPPYPC AVNKRVLSEKNVNSENDLLNAAHRPITRQFAAQMAYKQQLRPEENKRTTQSVSN PSKSEDCAILDVDDDKMADDFPVPMFVQHTEAMLEEIDRMEEVEMEDVAEEPVT DIDSGDKENQLAVVEYIDDLYMFYQKAEASSCVPPNYMDRQQDINERMRGILID WLIEVHYKFELMDETLYLTVNLIDRFLAVQPVVKKKLQLVGVTAMLLACKYEEV SVPVVEDLILISDRAYSRKEVLEMERLMVNTLHFNMSVPTPYVFMRRFLKAAQS DKKLELLSFFIIELSLVEYDMLKFPPSLLAASAIYTALSTITRTKQWSTTCEWH TSYSEEQLLECARLMVTFHQRAGSGKLTGVHRKYSTSKFGHAAARTEPANFLLDF RL	235	1539

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
279	Cyclin B	MASRPIVPVQARGEAAIGGGAGKAAIGGGAGKQQKKNGAAEGRNRKALGDIGNL VTVRGIEGKVQPHRPITRSFCAQLLANAQAAAAAAENNKKQAVVNVNGAPSILDV PGAGKRAEPAAAAAAAVAKAAQKKVVKPKQKAEVIDLTSDSERAIEAKKKQQHH EPTKKEGEKSSRRNMPTLTSVLTARSKAACGMTKKPKEKVVDIDAGDAHNELAA FEYIEDIYTYYKEAENESLPRNYMSSQPEINEKMRAILVDWLIEIHNKFDLMPE TLYLTINIIDRFLSVKAVPRRELQLLGMGALFTASKYEEIWAPEVNDLVCIADR AYSHEQVLAMEKTILGKLEWTLTVPTHYVFLVRFIKASLGDRKLENMVYFLAEL GVMNYATLTYCPSMVAASAVYAARCTLGLTPLWNDTLKLHTGFSESQLMDCARL LVGYHAKAKENKLQVVYKKYSSSQREGVALIPPAKALLCEGGGLSSSSSLASSS	158	1618
280	Cyclin B	MGLPDENNAALSKPTNLQVGGLEIGGRKFGQEIRQTRRALSVINQNLVGDRAYP CHVVNKRGHSKRDAVCGKDQVDPVHRPLTRKFAAQTASTQQHCIEEAKKPRTAV QERNEFGDCIFVDVEDCQPSSENQPVPMFLEIPESRLDDDMEEVEMEDIVEEEE EEPIMDIDGRDKKNPLAVVDYIEDIYANYRRTENCSCVSANYMAQQADINEKMR SILIDWLIEVHDKFDLMHETLFLTVNLIDRFLARQSVVRKKLQLVGLVAMLLAC KYEEVSVPVVGDLILISDKAYTRKEVLEMESLMLNSLQFNMSVPTPYVFMRRFL KAAESDKKLEVLSFFLIELSLVEYEMVKFPPSLLAAAAIFTAQCTLYGFKQWTK TCEWHSNYTEDQLLECARMMVGFHQKAATGKLTGVHRKYGTSKFGYTSKCEPAN FLLGEMKNP	205	1530
281	Cyclin B	MGLPDENNAALSKPTNLQVGGLEIGGRKFGQEIRQTRRALSVINQNLVGDRAYP CHVVNKRGHSKRDAVCGKDQVDPVHRPLTRKFAAQTASTQQHCIEEAKKPRTAV QERNEFGDCIFVDVEDCQPSSENQPVPMFLEIPESRLDDDMEEVEMEDIVEEEE EEPIMDIDGRDKKNPLAVVDYIEDIYANYRRTENCSCVSANYMAQQADINEKMR SILIDWLIEVHDKFDLMHETLFLTVNLIDRFLARQSVVRKKLQLVGLVAMLLAC KYEEVSVPVVGDLILISDKAYTRKEVLEMEKLMLNSLQFNMSVPTPYVFMRRFL KAAESDKKLEVLSFFLIELSLVEYEMVKFPPSLLAAAAIFTAQCTLYGFKQWTK TCEWHSNYTEDQLLECARMMVGFHQKAATGKLTGVHRKYGTSKFGYTSKCEAAN FLLGEMKNP	174	1499
282	Cyclin D	MAMVQRQGHDPSSPQEQEDGPSSFLSDDALYCEEGRFEEDDGGGGGQVDGIPLF PSQPADRQQDSPWADEDGEEKEEEEAELQSLFSKERGARPELAKDDGGAVAARR EAVEWMLMVRGVYGFSALTAVLAVDYLDRFLAGFRLQRDNRPWMTQLVAVACLA LAAKVEETDVPLLVELQEVGDARYVFEAKTVQRMELLVLSTLGWEMHPVTPLSF VHHVARRLGASPHHGEFTHWAFLRRCERLLVAAVSDARSLKHLPSVLAAAAMLR VIEEVEPFRSSEYKAQLLSALHMSQEMVEDCCRFILGIAETAGDAVTSSLDSFL KRKRRCGHLSPRSPSGVIDASFSCDDESNDSWATDPPSDPDDNDDLNPLPKKSR SSSPSSSPSSVPDKVLDLPFMNRIFEGIVNGSPI	94	1332
283	Cyclin D	MEASYQPHHHGHLRQHDPSSSQQEEQVPFDALYCSEEHWGEEDEEEGLASDGLL SEERDHRLLSPRALLDQDLLWEDEELASLFSKEEPGGMRLNLENDPSLADARRE AVEWIMRVHAHYAFSALTALLAVNYWDRFTCSFALQEDKPWMTQLSAVACLSLA AKVEETQVPLLIDFQVEDSSPVFEAKNIQRMELLVLSSLEWKMNPVTPLSFLDY MTRRLGLTGHLCWEFLRRCENVLLSVISDCRFTCYLPSVIAASTMLHVINGLKP RLDVEDQTQLLGILAMGMDKIDACYKLIDDDHALRSQRYSHNKRKFGSVPGSPR GVMELCFSSDGSNDSWSVAASVSSSPEPHSKKSRAGEEAEDRLLRGLEGEEDDP ASADIFSFPH	176	1342

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
284	Cyclin D	MALQEEDTRRHYPTAPPFSPDGLYCEDETFGEDLADNACEYAGGGARDGLCEIK DPTLPPSLLGQDLFWEDGELASLVSRETGTHPCWDELISDGSVALARKDAVGWI LRVHGHYGFRPLTAMLAVNYLDRFFLSRSYQRDRPWISQLVAVACLSVAAKVEE TQVPILLDLQVANAKFVFESRTIQRMELLLMSTLDWRMNSVTPISFFDHILRRF GLTTNLHRQFFWMCERLLLSVVADVRLASFLPSVVATAAMLYVNKEIEPCICSE FLDQLLSLLKINEDRVNECYELILELSIDHPEILNYKHKRKRGSVPSSPSGVID TSFSCDSSNDSWGVASSVSSSLEPRFKRSRFQDQQMGLPSVNVSSMGVLNSSY	150	1283
285	Cyclin-dependent kinase regulatory subunit	MGQIQYSEKYFDDTYEYRHVVLPPDVAKLLPKNRLLSENEWRAIGVQQSRGWVH YAIHRPEPHIMLFRRPLNYQQQQENQAQQNMLAK	101	367
286	Histone acetyltransferase	MGSIDPPKAEQNGTAAAAVADPGQKPGAGDAMPPPPPVKHSNGTAAEPDVATKR RRMSVLPLEVGTRVMCRWRDGKYHPVKVIERRKLNPGDPNDYEYYVHYTEFNRR LDEWVKLEQLDLNSVETVVDEKVEDKVTGLKMTRHQKRKIDETHVEGHEELDAA SLREHEEFTKVKNIATIELGRYEIETWYFSPFPPEYNDCSKLYFCEFCLNFMKR KEQLQRHMKKCDLKHPPGDEIYRSGTLSMFEVDGKKNKVYGQNLCYLAKLFLDH KTLYYDVDLFLFYVLCECDDRGCHMVGYFSKEKHSEESYNLACILTLPPYQRKG YGKFLIAFSYELSKKEGKVGTPERPLSDLGLLSYKGYWTRVLLDILKKHKANIS IKELSDMTAIKADDILNTLQSLDLIQYRKGQHVICADPKVLDRHLKAAGRGGLE VDVSKLIWTPYREQG	9	1352
287	Histone acetyltransferase	MAQKHSTAPDPAAEPKKRRRVGFSGIDAGVDPNGCFKVYLVSREEEVGAPDSFC LDPVDLSHFFEEEDGKIYGYEGLKISVWVSCVSFHSYAEIAFESKSDGGKGITD LNTALKNMFGETLVDNKDDFLQTFSKETQFIRSTVSAGEILKHKHSDDHVNDSV SNLKVGSDVEAVRMLMGDMTAGHLYSRLVPLVLLLVDGSSPIDVTDSSWELYLL IQKTSDQQGNFHDRLLGFAAVYRFYHYPDSSRLRLGQILVLPLYQRKGYGRYLL EVLNNVAIADDVYDFTIEEPVDNLQHLRTCIDVQRLLSFDKVQQAVNSTVSQLK QGKLSKKTYIPRLLPPPSVVEDARKRFKINKKQFLQCWEILVYLGLDPADKSIQ DYFSVISNRVRADILGKDSETAGKKVIEVPSDFDPEMSFVMHRAKAGGEANGIQ VEDNQNKQEEQLQQLIDERLKDIKLIAEKVTQK	89	1486
288	Histone acetyltransferase	MAQKHSTAPDPAAEPKKRRRVGFSGIDAGVDPNGCFKVYLVSREEEVGAPDSFC LDPVDLSHFFEEEDGKIYGYEGLKISVWVSCVSFHSYAEIAFESKSDGGKGITD LNTALKNMFGETLVDNKDDFLQTFSKETQFIRSTVSAGEILKHKHSDGHVNDSV SNLKVGSDVEAVRMLMGDMTAGHLYSRLVPLVLLLVDGSNPIDVTDSSWELYLL IQKTSDQQGNFHDRLLGFAAVYRFYHYPDSLRLRLGQILVLPLYQRKGYGHYLL EVLNNVAIADDVYDFTIEEPVDNLQHLRTCIDVQRLLSFDKVQQAVNSTVSQLK QGKLSKKTYIPRLLPPPSVVEDARKRFKINKKQFLQCWEILVYLGLDPADKSIQ DYFSVISNRVRADILGKDSETAGKKVIEVPSDFDPEMSFVLHRAKAGGETNGIQ VEDNQNKQEEQLQQLIDERLKDIKLIAQKVSRK	80	1477

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
289	Histone deacetylase	MALPMEFWGVEVKAGQPLKVNPGNAKILHLSQASLGECKSSKGNESVPLHVKFG DQKLVLGTLSTENFPQLAFDLVFEKEFELSHNWKSGSVYFCGYKSVVHDDDDEF SDLESDSEEEDLPMIGVENGKVAAQASAKTATASANASKVESSGKQKARIPQPM KVDEDDSDEDDDDEDEDESDEEGVDGEADSDEEEDESDEEETPKKAEIGKKRAA DSATKTPVPAKKSKLPTPQKTDGKKGGHTATPHPAKQAGKNPANSANKSQSPKS AGQVSCKSCSKTFNSDGALQSHSKAKHGGK;	160	1062
290	Histone deacetylase	MEFWGVEVKAGQPLKVNPGNAKILHLSQASLGECKSSKGNESVPLHVKFGDQKL VLGTLSTENFPQLAFDLVFEKEFELSHNWKSGSVYFCGYKSVVHDDDDEFSDLE SDSEEEDLPMIGVENGKVAAQASAKTATASANASKVESSGKQKASIPQPMKVDE DDSDEDDDEDDDDEDESDEGVDGEADSDEEEDESDEEETPKKAEIGKKRAADSA TKTPVPAKKSKLPTPQKTDGKKGGHTATPHPAKQAGKNPANSANKSQSPKSAGQ VSCKSCSKTFNSDGALQSHSKAKHGGK	172	1077
291	Histone deacetylase	MEFWGVEVKSGEPLNVEPGAETVVHLSQACLGETKEKTKESVLLYVHIGVQKLV LGTLSADKFPQIPFDLVFEKSFKLSHNWKNGSVFFSGYKTLLPCGSDADSPYSD SDTDEGLPINVTAQADVPAKKAPVTANANAAKPNLASAKQKVKIVESNEDGKNE GDDDEDADVSSDDDAEDDSGDEDMVDGGDESSDEDDDDSEEGESSEEEEPKAQP SKKRPADSVLKTPASDKKSKLETPQKTDGKKASEHVATPYPSKQAGKAIASKGQ AKQQTPNSNEFSCKPCNRSFKSDQALQSHNKAKHGGS	66	989
292	Histone deacetylase	MDTGGNSLPSGPDGVKRKVCYFYDPEVGNYYLLQHMQVLKPVPARDRDLCRFHA DDYVAFLRSITPETQQDQLRQLKRFNVGEDCPVFDGLHSFCQTYAGGSVGGAVK LNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLGILELLKQHERVLYVDIDI HHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDIGYGKGKYYSLNVPLDDG IDDESYHSLFKPIIGKVMEVFKPGAVVLQCGADSLSGDRLGCFNLSIKGHAECV RYMRSFNVPVLLLGGGGYTIRNVARCWCYETGVALGLEVDDKMPQHEYYEYFGP DYTLHVAPSNMENKNSRQLLEEIRSKLLENLSKLQHAPSVPFQERPPDTELPEA DEDQEDPDERWDPDSDMDVDEDRKPLPSRVKRELIVEPEVKDQDSQKASIDHGR GLDTTQEDNASIKVSDMNSMITDEQSVKMEQDNVNKPSEQIFPK	111	1541
293	Histone deacetylase	MDTGGNSLPSGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYG LLQHMQVLKPVPARDRDLCRFHADDYVAFLRSITPETQQDQLRQLKRFNVGEDC PVFDGLHSFCQTYAGGSVGGAVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVN DIVLGILELLKQHERVLYVDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGT GDIRDIGYGKGKYYSLNVPLDDGIDDESYHSLFKPIIGKVMEVFKPGAVVLQCG ADSLSGDRLGCFNLSIKGHAECVRYMRSFNVPVLLLGGGGYTIRNVARCWCYET GVALGLEVDDKMPQHEYYEYFGPDYTLHVAPSNMENKNSRQLLEDIRSKLLENL SKLQHAPSVPFQERPPDTELPEADEDQEDPDERWDPDSDMDVDEDRKPLPSRVK RELIVEPEVKDQDSQKASIDHGRGLDTTQEDNASIKVSDMNSMITDEQSVKMEQ DNVNKPSEQIFPK	116	1615

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
294	Histone deacetylase	MRPKDRISYFYDGDVGSVYFGPNHPMKPHRLCMTHHLVLSYELHTKMEIYRPHK AYPAELAQFHSPDYVEFLHRITPDTQHLFPNDLAKYNLGEDCPVFENLFEFCQI YAGGTIDAARRLNNQLCDIAINWAGGLHHAKKCEASGFCYINDLVLGILELLKY HARVLYIDIDVHHGDGVEEAFYFTDRVMTVSFHKFGDMFFPGTGDVKEIGGKEG KFYAINVPLKDGIDDTSFTRLFKAIISKVVETYQPGAIVLQCGADSLAGDRLGC FNLSIDGHSECVRFVKKFNLPLLVTGGGGYTKENVARCWVVETGVLLDTELPNE IPENEYFKYFAPDYSLKIPRGNIVLENLNSKSYLSAIKVQVLENLRNIQHAPSV QMQEVPPDFYIPDFDEDEQNPDERMDQHTQDKQIQRDDEYYDGDNDNDHNMDDS	155	1453
295	Histone deacetylase	MTVAEDFHVNNRSKMVSQATPESRLTGGEDDNSLHNQVDELLCQELPERQVILE FEGTRPKPYFSDHNGGENSALGVRATEDDLNSDVEAEEKQKEMTLEDMYKNDGT LYDDDEDDSDWEPVKRQVELMRWFCTNCTMVNVEDVFLCDICGEHRDSGILRHG FYASPFMQDVGAPSVEAEVQESREDHARSSPPSSSTVVGFDEKMLLHSEVEMKS HPHPERADRLQAIAASLATAGIFPGRCRSLPVREITKEELQMVHSSEHVDAVEM TSHMFSSYFTPDTYANEHSARAARIAAGLCADLASTIISGRSKNGFALVRPPGH HAGIKHAMGFCLHNNAAVAALAAQGAGAKKVLIVDWDVHHGNGTQEIFDGNKSV LYISLHRHEGGNFYPGTGAAHEVGTMGAEGYCVNIPWSRRGVGDNDYVFAFHHI VLPIASAFAPDFTIISAGFDAARGDPLGCCDVTPAGYAQMTHMLSALSGGKLLV ILEGGYNLRSISSSAVAVIKVLLGDSPISEIADAVPSKAGLRTVLEVLKIQRSY WPSLESIFWELQSQWGMFLVDNRRKQIRKRRRVLVPIWWKWGRKSVLYHLLNGH LHVKTKR	228	2033
296	Histone deacetylase	MAAAPSSPPTNRVDVFWHDGMLSHDTGRGVFDTGSDPGFLDVLEKHPENPDRVR NMVSILKRGPISPFISWHTATPALISQLLSFHSPEYINELVEADKNGGKVLCAG TFLNPGSWDAALLAAGNTLSAMKYVLDGKGKIAYALVRPPGHHAQPSQADGYCF LNNAGLAVRLALDSGCKRVVVVDIDVHYGNGTAEGFYQSSDVLTISLHMNHGSW GPSHPQSGSVDELGEDEGYGYNMNIPLPNGTGDRGYEYAVTELVVPAVESFKPE MVVLVVGQDSSAFDPNGRQCLTMDGYRAIGRTIRGLADRHSGGRILIVQEGGYH VTYSAYCLHATVEGILDLPDPLLADPIAYYPEDEAFPVKVVDSIKRYLVDKVPF LKEH	110	1258
297	Histone deacetylase	MVESSGGASLPSVGQDARKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHNLIV HYYLHRRMEISRPFPAATTDIRRFHSEDYVTFISSVTPETVSDPAFSRQLKRFN VGEDCPVFDGIFGFCQASAGGSMGAAVKLNRGDSDIALNWAGGLHHAKKSEASG FCYVNDIVLGILELLKVHKRVLYVDIDVHHGDGVEEAFYTTDRVMTVSFHKFGD FFPGSGHIKDTGAGPGKNYALNVPLNDGIDDESFRGMFRPIIQKVMEVYQPDAV VLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSFNVPLMVLGGGGYTMRNVARC WCYETAVAVGVEPENDLPYNEYYEYFGPDYTLHVEPCSMENLNAPKDLERIRNM LLEQLSRIPHAPSVPFQMTPPITQEPEEAEEDMDERPKPRIWNGEDYESDAEED KSQHRSSNADALHDENVEMRDSVGENSGDKTREDRSPS	50	1462
298	MAT1 CDK-activating kinase assembly factor	MVVPSSNPHNREMAIRRRMASTFNKREDDFPSLREYNDYLEEVEEMTFNLIEGV DVPTIEAKIAKYQEENAEQIMINRAKKAEEFAAALAASKGLPPQTDPDGALNSQ AGLSVGTQGQYAPAIAGGQPRPTGMAPQPVPLGTGLDIHGYDDEEMIKLRAERG GRAGGWSIELSKKRALEEAFGSLWL	176	739

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
299	Peptidylprolyl isomerase	MAAIISCHHYHSCCSSLIASKWVGARIPTSCFGRSSTQSNNAASVRQFVTRCSS SPSSRGQWQPHQNGEKGRSFSLRECAISIALAVGLVTGVPSLDMSTGNAYAASP ALPDLSVLISGPPIKDPEALLRYALPINNKAIREVQKPLEDITDSLKVAGLRAL DSVERNVRQASRVLKQGKNLIVSGLAESKKDHGVELLDKLEAGMDELQQIVEDG NRDAVAGKQRELLNYVGGVEEDMVDGFPYEVPEEYKNMPLLKGRAAVDMKVKVK DNPNLEECVFRIVLDGYNAPVTAGNFVDLVERHFYDGMEIQRADGFVVQTGDPE GPAESFIDPSTEKPRTIPLEIMVDGEKAPVYGATLEELGLYKAQTKLPFNAFGT MAMARDEFEDNSASSQIFWLLKESELTPSNANILDGRYAVFGYVTENQDFLADL KVGDVIESVQVVSGLDNLANPSYKIAG;	150	1529
300	Peptidylprolyl isomerase	MAGEDFDIPPADEMNEDFDLPDDDDDAPVMKAGDEKEIGKQGLKKKLVKEGDAW ETPDNGDEVEVHYTGTLLDGTQFDSSRDRGTPFKFTLGQGQVIKGWDQGIKTMK KGENAIFTIPPELAYGEAGSPPTIPPNATLQFDVELLSWTSVKDICKDGGIFKK ILVEGEKWENPKDLDEVLVKYEFQLEDGTTIARSDGVEFTVKEGHFCPAVAKAV KTMKKGEKVLLTVKPQYGFGEKGKPASGDEGAVPPNATLQITLELVSWKTVSEV TDDKKVIKKILKEGEGYERPNEGAVVEVKLIGKLQDGTVFVKKGHDDCEELFKF KIDEEQVVDGLDKAVMNMKKGEVALLTVAPEYAFGSSESKQDLAVVPPSSTVYY EVELVSFVKDKESWDMNTEEKIEAAGKKKEEGNVIFKAGKYAKASKRYEKAVKY IEYDTSFSEDEKKQAKALKVACNLNDAACKLKLKDYNQAEKLCTKVLELDSRNV KALYRRAQAYIELSDLDLAEFDIKKALEIDPHNRDVKLEYKVLKEKVKEFNKKD AKFYGNMFAKMSKLEPVEKTAAKEPEPMSIDSKA;	247	1971
301	Peptidylprolyl isomerase	MSTVYVLEPPTKGKVVLNTTHGPLDVELWPKEAPKAVRNFVQLCLEGYYDNTIF HRIIKDFLVQGGDPTGSGTGGESIYGDAFSDEFHSRLRFKHRGLVACANAGSPH SNGSQFFITLDRCDWLDRKNTIFGKITGDSIYNLSGLAEVETDKSDRPLDPPPK IISVEVLWNPFEDIVPRAPVRSLVPTVPDVQNKEPKKKAVKKLNLLSFGEEAEE EEKALVVVKQKIKSSHDVLDDPRLLKEHIPSKQVDSYDSKTARDVQSVREALSS KKQELQKESGAEFSNSFREIADDEDDDDDDASFDARMRRQILQKRKELGDLPPK PKPKSRDGISARKERETSISRDKDDDDDDDQPRVEKLSLKKKGIGSEARGERMA NADADLQLLNDAERGRQLQKQKKHRLRGREDEVLTKLETFKASVFGKPLASSAK VGDGDGDLSDWRSVKLKFAPEPGKDRMTRNEDPNDYVVVDPLLEKGKEKFNRMQ AKEKRRGREWAGKSLT;	136	1644
302	Peptidylprolyl isomerase	MASAISMHSSGLLLLQGTNGKDVTEMGKAPASSRVANMQQRKYGATCCVARGLT SRSHYASSLAFKQFSKTPSIKYDRMVEIKAMATDLGLQAKVTNKCFFDVEIGGE PAGRIVIGLFGDDVPKTVENFRALCTGEKGFGYKGCSFHRIIKDFMIQGGDFTR GNGTGGKSIYGSTFEDENFALKHVGPGVLSMANAGPSTNGSQFFICTVKTPWLD NRHVVFGQVVDGMDVVQKLESQETSRSDVPRQPCRIVNCGELPLDG;	48	836
303	Peptidylprolyl isomerase	MAASFTALSNVGSLSSPRNGSEIRRFRPSCNVAASVRPPPLKAGLSASSSSSFS GSLRLIPLSSSPQRKSRPCSVRASAEAAAAQSKVTNKVYLDISIGNPVGKLVGR IVIGLYGDDVPQTAENFRALCTGEKGFGYKGSTVHRVIKDFMIQGGDFDKGNGT GGKSIYGRTFKDENFKLSHVGPGVVSMANAGPNTNGSQFFICTVKTPWLDQRHV VFGQVLEGMDIVRLIESQETDRGDRPRKRVVVSDCGELPVV;	49	822

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
304	Peptidylprolyl isomerase	MAEAIDLTGDGGVMKTIVRRAKPDAVSPSETLPLVDVRYEGVLAETGEVFDSTH EDNTLFSFEIGKGSVISAWDTALRTMKVGEVAKITCKPEYAYGSTGSPPDIPPD ATLIFEVELVACKPCKGFSVTSVTEDKARLEELKKQREIAAATKEEEKKRREEA KAAAAAARVQAKLDAKKGHGKGKGKAK;	185	751
305	Peptidylprolyl isomerase	MGNPKVFFDMSIGGQPAGRIVMELYADVVPRTAENFRALCTGEKGAGRSGKPLH YKGSSFHRVIPGFMCQGGDFTAGNGTGGESIYGSKFADENFVKKHTGPGVLSMA NAGPGTNGSQFFVCTAKTEWLDGKHVVFGQIVDGMDVVKAIEKVGSSSGRTSKP VVVADCGQLS	103	621
306	Peptidylprolyl isomerase	MPNPKVFFDMTIGGAAAGRVVMELYADTTPRTAENFRALCTGEKGVGRSKKPLH YKGSKFHRVIPSFMCQGGDFTAGNGTGGESIYGVKFADENFIKKHTGPGILSMA NAGPGTNGSQFFICTTKTEWLDGKHVVFGKVVEGMEVVKAIEKVGSSSGRTSKP VVVADCGQLP	41	559
307	Peptidylprolyl isomerase	MAEAIDLTGDGGVMKTIVRRAKPDAVSPSETLPLVDVRYEGVLAETGEVFDSTH EDNTLFSFEIGKGSVISAWDTALRTMKVGEVAKITCKPEYAYGSTGSPPDIPPD ATLIFEVELVACKPCKGFSVTSVTEDKARLEELKKQREIAAATKEEEKKRREEA KAAAAAARVQAKLDAKKGHGKGKGKAK	127	693
308	Peptidylprolyl isomerase	MATARSFFLCALLLLATLYLAQAKKSEDLKEVTHKVYFDVEIAGKPAGRIVMGL YGKAVPKTAENFRALCTGEKGTGKSGKPLHYKGSSFHRIIPSFMLQGGDFTLGD GRGGESIYGEKFADENFKLKHTGPGLLSMANAGPDTNGSQFFITTVTTSWLDGR HVVFGKVLSGMDVVYKVEAEGRQSGTPKSKVVIADSGELPL	28	639
309	Peptidylprolyl isomerase	MMRREISVLLQPRFVLAFLALAVLLLVFAFPFSRQRGDQVEEEPEITHRVYLDV DIDGQHLGRIVIGLYGEVVPRTVENFRALCTGEKGKSANGKKLHYKGTPFHRII SGFMIQGGDVIYGDGKGYESIYGGTFADENFRIKHSHAGIISMVNSGPDSNGSQ FFITTVKASWLDGEHVVFGRVIQGMDTVYAIEGGAGTYNGKPRKKVIIADSGEI PKSKWDEER	135	812
310	Peptidylprolyl isomerase	MWATAEGGPPEVTLETSMGSFTVELYFKHAPRTSRNFIELSRRGYYDNVKFHRI IKDFIVQGGDPTGTGRGGESIYGKKFEDEIKPELKHTGAGILSMANAGPNTNGS QFFITLAPCPSLDGKHTIFGRVCRGMEIIKRLGSVQTDNNDRPIHDVKILRTSV KD	119	613
311	Peptidylprolyl isomerase	MSNPKVFFDILIGKMKAGRVVMELFADVTPKTAENFRALCTGEKGIGRSGKPLH YKGSTFHRIIPNFMCQGGDFTRGNGTGGESIYGMKFADENFKIKHTGLGVLSMA NAGPDTNGSQFFICTEKTPWLDGKHVVFGKVIDGYNVVKEMESVGSDSGSTRET VAIEDCGQLSEN	38	562

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
312	Peptidylprolyl isomerase	MDDDFEFPASSNVENDDDDGMDMDDMGGDVPEEEDPVASPAVLKVGEEREIGKA GFKKKLVKEGEGWETPSSGDEVEVHYTGTLLDGTKFDSSRDRGTPFKFKLGRGQ VIKGWDEGIKTMKKGENAIFTIPPELAYGESGSPPTIPPNATLQFDVELLSWSS VKDICKDGGILKKVLVEGEKWDNPKDLDEVFVKYEASLEDGTLISKSDGVEFTV GDGYFCAALAKAVKTMKKGEKVLLTVMPQYAFGETGRPASGDEAAVPPDASLQI MLELVSWKTVSDVTKDKKVLKKTLKEGEGYERPNDGAAVQVRLCGKLQDGTVFV KKDDEEPFEFKIDEEQVIDGLDRAVKNMKKGEVALVTIQPEYAFGPTESQQDLA VVPANSTVYYEVELLSFVKEKESWEMNNQEKIEAAARKKEEGNAAFKAGKYVRA SKRYEKAVRFIEYDSSFSDEEKQQAKTLKNTCNLNDAACKLKLKDFKEAEKLCT KVLEGDGKNVKALYRRAQAYIQLVDLDLAEQDIKKALEIDPNNRDVKLEYKILK EKVREYNKRDAQFYGNMFAKMNKLEHSRTAGMGAKHEAAPMTIDSKA	109	1872
313	Peptidylprolyl isomerase	MAKPRCFMDISIGGELEGRIVGELYTDVAPKTAENFRALCTGEKGIGPHTGAPL HYKGVRFHRVIKGFMVQGGDISAGDGTGGESIYGLKFEDENFDLKHERKGMLSM ANSGPNTNGSQFFITTTRTSHLDGKHVVFGRVVKGMGVVRSVEHVTTAAGDCPT VDVVIADCGEIPAGADDGIRNFFKDGDTYPDWPADLDESPAELSWWMDAVDSIK AFGNGSYKKQDYKMALRKYRKALRYLDICWEKEGIDEVESSSLRKTKSQIFTNS SACKLKLCDLKGALLDAEFAVRDGENNAKAYFRQGQAHMELNDIDAAAESFSKA LELEPNDVGIKKELNAAKKKIFERREQEKRAYRKMFL	74	1159
314	Peptidylprolyl isomerase	MTKRKNPLVFLDVSIDGDPVERIVIELFADTVPRTAENFRSLCTGEKGVGKTTG KPLHYKGSYFHRIIKGFMAQGGDFSNGNGTGGESIYGGKFADENFKLAHDGPGL LSMANGGPNTNGSQFFIIFKRQPHLDGKHVVFGKVMRGMEVVKKIEQVGSANGK PLQPVKIVDCGETSETGTQDAVVEEKSKSATLKAKKKRSARDSSSESRGKRRQR KSRKERTRKRRRYSSSDSYSSESSDSDSESYSSDTESESKSHSESSVSDSSSSD GRRRKRKSTKREKLRRQRGKDSRGEQKSARYDKKSRHKSADSSSDSESESSSRS RSRDDKKKSSRRESARSVSKLKDAEANSPENLESPRDREIKKVEDNSSHEEGEF SPKNDVQHNGHGTDAKFGKYDDQRPRSDGSKKSSGSMRDSPKRLANSVPQGSPS SSPAHKASEPSSSIRARNPSRSPAPDGNSKRIRKGRGFTERFSYARRYRTPSPE DVTYRPYHYGRRNFHDRRNDRYSNYRSYSERSPHRRYRSPPRGRSPPRYQRRRS RSRSVSRSPGGNKGRYRGRDQSRSRSRSRSPRRGSSPANKQLPLSERLKSRL GTRVDEHSPRRRRSSSRSHDSSRSRSPDEVPDKHEGKAAPVSPARSRSSSPSGR GLVSYGDASPDSGIN	54	2045
315	Peptidylprolyl isomerase	MSVLLVTSLGDIVVDLHADRCPLTCKNFLKLCRIKYYNGCVFHTVQKDFTAQTG DPTGTGTGGDSVYKFLYGDQARFFMDEIHLDLKHSKTGTVAMASGGENLNASQF YFTLRDDLDYLDGKHTVFGEVAEGLETLTRINEAYVDEKGRPYKNIRIRHTYIL DDPFDDPPQLAELIPDASPEGKPKDEVVDDVRLEDDWVPLDEQLGPAQLEEAIR AKEAHSRAVVLESIGDIPDAEIKPPDNVLFVCKLNPVTEDEDLHTIFSRFGTVV SADVIRDFKTGDSLCYAFIEFENKDSCEQAYFKMDNALIDDRRIKVDFSQSVAK LWSQFKRKDSQAAKGKGCFKCGAPDHMARECPGSSTRQPLSKYILKEDNAQRGG DDSRYEMVFDEDAPESPSHGKKRRGRDDRDDRHKMSRQSVEETKFNDREGGHSV DKHRQSERSKHREDEMSRDSKASEAGRRRIDRDFPEEERDGEKYTESHRDRDGK RGDYRDYRKGRADVQTHGDRRGDENYRRKSAAYDDGHEGAGAARRKDSNDDHHA YRRGYGDSRKGTRDEDDDGRGRRDDPSYRRSSGHKDSSNGGREEQKYRSGETDG KSHPERSHRGDRRR	53	1879

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
316	Peptidylprolyl isomerase	MRPFNGGSSIACLVLVIAAGALAESQGPHLGSARVVFQTNYGDIEFGFFPGVAP RTVDHIFKLVRLGCYNTNHFFRVDKGFVAQVADVANGRTAPMNDEQRTEAEKTI VGEFSNVKHVRGILSMGRYDDPDSAQSSFSILLGDAPHLDGKYAIFGRVTKGDE TLKKLEQLPTRREGMFVMPTERITILSSYYYDTGAESCEEENSTLRRRLAASAV EVERQRMKCFP	7	690
317	Peptidylprolyl isomerase	MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRALCTGEKGTGRSGKPLH FKGSSFHRVIPGFMCQGGDFTRGNGTGGESIYGEKFADENFVKKHTGPGILSMA NAGPNTNGSQFFICTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGSGRTSKP VVIADSGQLA	83	601
318	Peptidylprolyl isomerase	MRFTSITSAIALFAAAASALDKPLDIKVDKAVECSRKTKAGDKIQVHYRGTLEA DGSEFDASYKRGQPLSFHVGKGQVIKGWDQGLLDMCPGEKRTLTIQPDWGYGSR GMGPIPANSVLIFETELVEIAGVAREEL	125	535
319	Peptidylprolyl isomerase	MGNPKVFFDMSIGGQPAGRIVMELYADVVPRTAENFRALCTGEKGAGRSGKPLH YKGSSFHRVIPGFMCQGGDFTAGNGTGGESIYGSKFADENFVKKHTGPGVLSMA NAGPGTNGSQFFVCTAKTEWLDGKHVVFGQIVDGMDVVKAIEKVGSSSGRTSKP VVVADCGQLS	55	573
320	Peptidylprolyl isomerase	MAVATRSRWVAMSVAWILVLFGTLALIQNRLSDTGASSDPKLVHRKVGEEKKKP DDLEEVTHKVFFDVEIGGKPAGRIVMGLFGKTVPKTVENFRALCTGEKGIGKSG KPLNYKGSQFHRIIPKFMIQGGDFTLGDGRGGESIYGNKFSDENFKLKHTDAGR LSMTNAGPDTNGSQFFITTVTTSWLDGRHVVFGKVLSGMDVVHKIEAEGGQSGQ PKSIVVISDSGELDL	147	842
321	Peptidylprolyl isomerase	MAVTLHTNLGDIKCEIFCDEVPKAAEHNARGILSMANSGPNTNGSQFFIAYAKQ PHLNGLYTIFGRVIHGFEVLDIMEKTQTGPGDRPLAEIRLNRVTIHANPLAG	167	487
322	Peptidylprolyl isomerase	MAVATRSRWVAMSVAWILVLFGTLALIQNRLSDTGASSDPKLVHRKVGEEKKKP DDLEEVTHKVFFDVEIGGKPAGRIVMGLFGKTVPKTVENFRALCTGEKGIGKSG KPLNYKGSQFHRIIPKFMIQGGDFTLGDGRGGESIYGNKFSDENFKLKHTDAGR LSMANAGPDTNGSQFFITTVTTSWLDGRHVVFGKVLSGMDVVHKIEAEGGQSGQ PKSIVVISDSGELDL	195	890
323	Peptidylprolyl isomerase	MGNPKVFFDMSIGGQPAGRIVMELYADVVPRTAENFRALCTGEKGAGRSGKPLH YKGSSFHRVIPGFMCQGGDFTAGNGTGGESIYGSKFADENFVKKHTGPGVLSMA NAGPGTNGSQFFVCTAKTEWLDGKHVVFGQIVDGMDVVKAIEKVGSSSGRTSKP VVVADCGQLS	68	586

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
324	Retinoblastoma related protein	MSPVAANAMEEAAEPEVPAPVTPSKDDADTDAAVSRFLGFCKSKLGLAEGNCVQ SSTLLRKTAHVLRSSGTVIGTGTAEEAERYWFAFVLYTVRRVGERKAEDEQNGS DETEVPLSRILKASVLNLIDFFKEIPQFVIKAGAIVSGIYGANWDSRLEAREMQ TNYVHLCILCKFYKRICGEFFILNDAKDDMKSADSSTSDPVIMYQPFGWLLFLA LRIHALSRFKDLVSSTNALVSVLAILIIHLPTRFRKFSISDSSQLVKRSEKGVD LVGSLAYRYDTSEDEIKRTLEKANNVIAEILGITPPPASECKAENLENVDTDGL IYFGNLMEETSLSSILSTLEKIYEDATRNDSEFDERVFINDDDSLLVSGSLSGA AINLTGAKRKYDSFASPAKTITRPLSPSRSPASHINGIIGGTNLRITATPVATA MTTAKWLRTFVSPLPSKPSTDLQGFLASCDRDVTSDVIRRANIILEAIFPNSPI GERTVTGGLQNANLMDNMWAEQRRLEALKLYYRVLEAMCRAEAQILHSNNLTSL LTNERFHRCMLACSAELVLATHKTVTMLFPAVLERTGITAFDLSKVIESFVRHE ETLPRELRRHLNTLEERLLENMVWERGSSMYNSLVVARPALAPEINRLGLLPEP MPSLDAIALLINFSSSGLPQSPVQKHEASPGQNGDIRSPKRISTEYRSVLVERN FTSPVKDRLLALSNIKSKLPPPPLQSAFASPTRPHPGGGGETCAETAIHIFFSK ITKLAAVRINAMLERLQLSQQIKEGVYCLFQQILSQRTNLFFNRHIDQVILCCF YGVAKINQINLTFREIIYNYRKQPQCKPQVFRNVFVDWSTRRNGKAGNEHVDII SFYNEIFIPSVKPLLVELGPTGATTRTNRTSEVGNKNDAQCPGSPKISSFPTLP DMSPKKVSASHNVYVSPLRSSKMDASISHSSKSYYACVGESTHAYQSPSKDLVA INSRLNGNRKVRGTLNFDDVDAGLVSDSMVANSLYLQNGSSMSSSTAKSSEKPE S	182	3265
325	WD40 repeat protein	MRPILMKGHERPLTFLKYNREGDLLFSCAKDHTPTVWFADNGERLGTYRGHNGA VWCCDVSRDSMRLITGSADTTAKLWSVQNGTQLFTFNFDSPARSVDFSIGDKLA VITTDPFMELPSAIHVKRIARDPADQASESVLVLRGHQGRIARAVWGPLNKTII SAGEDAVIRIWDSETGKLLRESDKETGHKKAVTSLMKSVDGSHFVTGSQDKSAK LWDIRTLTLIKTYVTERPVNAVTMSPLLDHVVLGGGQDASAVTMTDHRAGKFEA KFFDKILQEEIGGVKGHFGPINALAFNPDGKSFSSGGEDGYVRLHHFDPDYFNI KI	165	1145
326	WD40 repeat protein	MDKKRTVVPLVCHGHSRPVVDLFYSPITPDGFFLISASKDSSPMLRNGETGDWI GTFEGHKGAVWSCCLDTNALRAASGSADFSAKLWDALSGDELHSFEHKHIVRSC AFSEDTHLLLTGGVEKILRIFDLNRPDAPPREVDNSPGSIRTVAWLHSDQTILS SCTDIGGVRLWDVRSGKIVQTLETKSPVTSSEVSQDGRYITTADGSTVKFWDAN HFGLVKSYNMPCNIESASLEPKLGNKFIAGGEDMWVHIFDFHTGEEIGCNKGHH GPVHCVRFSPGGESYASGSEDGTIRIWQTGPANNVEGDANPSNGPVTGKAKVGA DEVTRKVEDLQIGKEGKDWREG	529	1569
327	WD40 repeat protein	MAEGLILKGTMRAHTDMVTAIAIPIDNSDMVVTSSRDKSIILWHLTKEEKVYGV PRRRLTGHSHFVQDVVLSSDGQFALSGSWDGELRLWDLATGVSARRFVGHTKDV LSVAFSIDNRQIVSASRDRTIKLWNTLGECKYTIQEGEAHTDWVSCVRFSPNTL QPTIVSASWDRTIKVWNLTNCKLRNTLAGHNGYVNTVAVSPDGSLCASGGKDGV ILLWDLAEGKRLYNLEAGAIIHSLCFSPNRYWLCAATENSIKIWDLESKSIVED LRVDLKNEADKTDGTTTAASNKKVIYCTSLNWSADGSTLFSGYNDGVIRVWGTG RY	156	1136

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
328	WD40 repeat protein	MAEGLHLKGTMKAHTDMVTAIAVPIDNADMIVTSSRDKSIILWHLTKEDKVYGV PRRRLTGHSHFVQDVVLSSDGQFALSGSWDGELRLWDLATGVSARRFVGHTKDV LSVAFSIDNRQIVSASRDRTIKLWNTLGECKYTIQEGEAHNDWVSCVRFSPNTL QPTIVSASWDRTVKVWNLTNCKLRNTLQGHSGYVNTVAVSPDGSLCASGGKDGV ILLWDLAEGKKLYSLEAGAIIHSLCFSPNRYWLCAATENSIKIWDLESKSIVED LRVDLKNEADMSDGTTGAMSSNKKVIYCTSLNWSADGSTLFSGYNDGVIRVWGI GRY	90	1073
329	WD40 repeat protein	MAEGLHLKGTMKAHTDMVTAIAVPIDNADMIVTSSRDKSIILWHLTKEDKVYGV PRRRLTGHSHFVQDVVLSSDGQFALSGSWDGELRLWDLATGVSARRFVGHTKDV LSVAFSIDNRQIVSASRDRTIKLWNTLGECKYTIQEGEAHNDWVSCVRFSPNTL QPTIVSASWDRTVKVWNLTNCKLRNTLQGHSGYVNTVAVSPDGSLCASGGKDGV ILLWDLAEGKKLYSLEAGAIIHSLCFSPNRYWLCAATENSIKIWDLESKSIVED LRVDLKNEADMSDGTTGAMSSNKKVIYCTSLNWSADGSTLFSGYNDGVIRVWGI GRY	66	1049
330	WD40 repeat protein	MSGVPAPPFATTTPENGTMSSNSPAFHRDSDDDDDQGEVFLDDSDIIHEVAVDD EDLPDADDEADEAEEADDSLHIFTGHNGEVYSLACSPTDATLVATGAGDDKGFL WRIGHGDWAVELQGHKDSISSLAFSLDGQLLASGSLDGVIQIWDVPSGNLKGTL DGPGGGIEWIRWHPKGHIILAGSEDSTVWMWNADKMAYLNMFSGHGNSVTCGDF TPDGKTICTGSDDATLRIWNPKSGENIHVVKGHPYHAEGLTSMAISSDSGLAIT GAKDGSVRIVNISSGRVVSSLDAHADSVEFVGLALSSPWAATGSLDQKLIIWDL QHSSPRATCDHEDGVTCLSWVGASRFLASGCVDGKVRVWDSLSGDCVRTFHGHS DAIQSLSVSANEEFLVSVSIDGTARVFEIAEFH	277	1512
331	WD40 repeat protein	MGTSQHQLSSCLQLLPRRRGNKNLIFRRTMASGGAAAVAPPPGYKPYRHLKTLT GHVAAVSCVKFSNDGTLLASASLDKTLIIWSSAALSLLHRLVGHSEGVSDLAWS SDSHYICSASDDRTLRIWSSRSPFDCLKTLRGHTDFVFCVNFNPQSSLIVSGSF DETIRIWEVKTGRCLNVIRAHSMPVTSVHFNRDGSLIVSGSHDGSCKIWDTKNG ACLKTLIDDTVPAVSFAKFSPNGKFILVATLNDTLKLWNYATGKFLKIYTGHKN SVYCLTSTFSVTNGKYIVSGSEDRCICIWDLQGKNLIQKLEGHSDTVISVTCHP SENKIASAGLDSDRTVRIWLQDA	33	1076
332	WD40 repeat protein	MPSQKIETGHQDIVHDVAMDYYGKRVATASSDTTIKIIGVSNSSGSQHLASLSG HKGPVWQVAWAHPKFGSILASCSYDGQVILWKEGNQNDWAQAHVFNDHKSSVNS IAWAPHELGLCLACGSSDGNISVFTARPDGGWDTTRIEQAHPVGVTSVSWAPSM APGALVGSGLLDPVQKLASGGCDNTVKVWKLYNGTWKMDCFPALQMHSDWVRDV AWAPNLGLPKSTIASASQDGTVVIWTVAKEGEQWQGKVLKDFKTPVWRVSWSLT GNLLAVADGNNNVTLWNEAVDGEWQQVTTVEP	65	973
333	WD40 repeat protein	MKIAGLKSVENAHDESVWAAAWVPATESRPALLLTGSLDETVKLWRPDELALER TNAGHFLGVVSVAAHPSGVIAASASIDSFVRVFDVDTNATIATLEAPPSEVWQM QFDPKGTTLAVAGGGSASIKLWDTATWELNATLSIPRPEQPKPSEKGNKKFVLS VAWSPDGRRLACGSMDGTISIFDVARAKFLHHLEGHFMPVRSLVFSPVEPRLLF SASDDAHVHMYDSEGKSLVGSMSGHASWVLSVDVSPDGAALATGSSDRTVRLWD LSMRAAVQTMSNHSDQVWGVAFRPMAGAGVRAGGRLASVSDDKSISLYDYS	82	1047

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
334	WD40 repeat protein	MEIDLGNLAFDVDFHPSEQLVASGLITGDLLLYRYGDGSSPEKLLEVRAHGESC RAVRFINDGKAILTGSPDCSILATDVETGSVVARVENAHEAAVNRLVNLTESTI ATGDDNGCIKVWDTRQRSCCNTFSAHEDFISDMTFASDSMKLVVTSGDGTLSVC NLRSNKVQTRSEFSEDELLSVVIMKNGRKVVCGTQSGTLLLYSWGFFKDCSDRF VDLSPSSVDALLKLDEDRIIAGTENGLISLIGILPNRIIQPIAEHSDHPIERLA FSHDKKFLGSISHDQTLKLWDLNDILGSEDSPSSQAAIDDSDSDEMDVDANPPD SSKGNKKKHSGKGNDVGNANNFFADLGD	43	1101
335	WD40 repeat protein	MSQQPSVILATASYDHTIRFWEAKSGRCYRTIQYPDSQVNRLEITPHKRYLAVA GNPSIRLFDVNSNTPQPVMSFDSHTNNVMAVGFQYDGNWMYSGSEDGTVRIWDL RARGCQREYESRGAVNTVVLHPNQTELISGDQNGNIRVWDLTANSCSCELVPEV DTAVRSLTVMWDGSLVVAANNNGTCYVWRLLRGSQTMTNFEPLHKLQAHNGYIL KCLLSPEFCEPHRYLATASSDHTVKIWNVEGFTLEKTLIGHQRWVWDCVFSVDG AYLITASSDTTARLWSMSTGQDIRVYQGHHKATTCCALHDGAEGSPG	142	1095
336	WD40 repeat protein	MEDAMDMEVEVEVEAEEHSPSSSNPSGSSFRRFGLKNSIQTNFGSDYVFEITPK FDWSLMGVSLSSNAVKLYSPTTGQYCGECRGHSDTVNGISFSGPSSPHVLHSCS SDGTIRAWDTRSFKEVSCISAGPSQEIFSFSFGGSSDSLLSAGCKSQILFWDWR NKKQVACLEDSHVDDVTQVCFVPHHQNKLISASVDGLICIFDTAGDINDDEHME SVINVGTSIGKVGIFGQTFEKLWCLTHIETLSVWDWKEGTNEANFEDARKLASD SWSLDHIDYFVDCHSAEEGEGLWVIGGTNAGTLGYFPVKYKGGAAIGSPEAVLG GGHSDVVRSVLPMSGMAGTTSKTRGIFGWTGGEDGRLCCWLSDDSSATSRSWMS SNLVLKSSRSHHKKNRHQPY	61	1257
337	WD40 repeat protein	MSQHQEYPMEYAADDYDVGEVEDDMYFHERVMGDSDTDEDEEYDHLDNKITDTS AADARRGKDIQGIPWERLSVTREKYRRTRIEQYKNYENVPQSGESSEKDCKPTR KGGNYYEFWRNTRSVKSTILHFQLRNLVWSTTKHDVYLMSHFSIIHWSSLTCKK TEVLDVYGHVAPREKHPGSLLEGFTQTQVSTLAVRDKLLIAGGFQGELICKNLD RPGVSYCCRTTYDDNAITNAVEIYDYPSGAVHFMASNNDCGVRDFDMEKFELSR HFTFPWPVNHTSLSPDGKLLVIVGDNPEGIVVDSQRGKTIRPLQGHLDFSFASA WHPDGHIFATGNQDKTCRIWDIRNLSKSVAVLKGNLGAIRSIRFTSDGRFMAMA EPADFVHVYDVKSGYEKEQEIDFFGEISGVSFSPDTESLFVGVWDRTYGSLLQY NRCRNYSYLDSM	193	1527
338	WD40 repeat protein	MGASSDPNPDVSDEHQKRSEIYTYEAPWHIYAMNWSVRRDKKYRLAIASLLDHP AAAAAVPNRVEIVQLDDSTGEIRADPNLSFDHPYPATKAAFVPDKDCQRADLLA TSSDFLRIWRIADDSSRVDLRSFLNGNKNSEFCRPLTSFDWNEAEPKRIGTSSI DTTCTIWDIERETVDTQLIAHDKEVYDIAWGGVSVFASVSADGSVRVFDLRDKE HSTIIYESSEPDTPLVRLGWNKQDPRYMATIIMDSAKVVVLDIRYPTMPVVELQ RHQASVNAIAWAPHSSCHICTAGDDSQALIWDLSSMAQPVEGGLDPILAYTAGA EIEQLQWSSSQPDWVAIAFSLKLQ	109	1155

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
339	WD40 repeat protein	MRGGGGGGDATGWDEDAYRESVLKEREVQTRTVFRAAFAPSPSPSPSPDAVVVA SSDGSVASYSISACLSDHRLQSLRFADAKSQNVLEAEPACFLQGHDGPAYDVKF YGEGEDSLLLSCGDDGRIRGWMWRDITSSEAHDHSQGNSAKPVLDLVNPQSRGP WGALSPIPENNALAVDVKRGSIYAAAGDSCAYCWDVECGKIKTVFKGHSDYLHC IAARNSSSQIITGSEDGTARIWDCRSGKCVQVIDPDKDHKKGFFASVSCLALDA SESWLVCGRGRDLSVWSISASDCIAKISTNAPAQDVLFDDNQILLVGAEPLISR LDMNGAVLSQIHCAPQSVFSVSLHQSGVTAVGGYGGLVDVISQFGSHLCTFRCK CI	71	1213
340	WD40 repeat protein	MEAPIIDPLQGDFPEVIEEYLEHGIMKCIAFNRRGTLLAAGCTDGSCIIWDFET RGVAKELRDKECTAAITSVCWSKYGHRILVSASDKSLILWDVLSGEKIAHTTLQ HTVLQACLHPGSSTPSICLACPFSSAPMIVDLNTGSTTALPVLTADVSNGATPL SRNKTSDTSVTYSPCNACFNKHGDLVYAGTSKGEILIIDHKNVRVCAIVLVSGG AVIKNVVFSRNGQYMLTNSNDRLIRIYKNLLPPKDGLKMLDELNESFNESDDVE KLKAIGSKCLELLHEFQDSITRVQWKAPCFSGDGEWVIGGAASRGEHKIYIWDR AGHLVKILEGPKEALMDLAWHPVHPIIISVSLTGLVYIWAKDYTENWSAFAPDF KELEENEEYVEREDEFDLVPETEKVKGLDVHEDDEVDVLTVERDSVFSDSDMSQ EELCFLPAVPCLDIPEQQDKCVGSCSKLPDGNHSGSPLSVEAGQNGNASNHNSS PLEPMENSTADDTDGVRLKRKRKPSEKGLELQAEKVKKPVKPLKSSGRLSKTNK PVIDPDSSNGVYGDDGSD	109	1785
341	WD40 repeat protein	MRGVSWPEDGNNPSTSSSSQRNQQQAHAPRAVSGHAASHPSASNIFKLLVQREV SPRSKHSSKKLWREASKCQPYPFQQSCEAVRDVRQGLISWVESASLRHLSAKYC PLVPPPRSTIAAAFSPDGKILASTHGDHTVKLIDSQTGSCLKVLRGHRRTPWVV RFHPLYPEILASGSLDHEVRLWDANTAECIGSRNFYRPIASIAFHARGELLAVA SGHKLYIWHYNRRGETSSPTIVLRTQRSLRAVHFHPHAAPFLLTAEVNDLDSAD SAMTLATSPGYLHYPPPTVYFADAHSHERSRLADELPLMPLPLLMWPSFTRDDG RVPLQRIDGDVGLNGQQRVDSSSSVRLWTYSTPSGQYELLLSPVESGNSPSMPE ETGNNAFSSAVEAEVSQSAMDTVEDMEVQPEERNTQFFSFSDPRFWELPLLHGW LVGQTQAGPRSVRQSSPGDIETQSAFGEVASVSPITSGVMPVSMDPSRFGGRSG SRYRSPGSRGVHVTGPNNDGPRDENDPQSVVSKLRSELAASLAAAASTELPCTV KLRIWPHDVKDPCAQLDLESCRLTIPHAVLCSEMGAHFSPCGRFLAACVACVLP HLESDPGLHGQVNQDVTGVATSPTRHPISAHQIMYELRIYSLEEATFGIVLASR PVRAAHCLTSIQFSPTSEHLLLAYGRRHSSLLKSIVIDGENTVPIYTILEVYRV SDMELVRVLPSAEDEVNVACFHPSVGGGLIYGTKEGKLRILHYDSSHGLNLKSS GFLDENVPEVQTYALEC	364	2685
342	WD40 repeat protein	MDSAVAIAALSLVVGAAIALLFFGNYFRKRRSEVVAMAEADLQPHPKNPSRPPP QPAAKKVHAKSHAHGADKDKNKRHHPLDLNTLKGHGDSVTGLCFASDGRSLATA CADGVVRVFKLDDASNKSFKFLRINLPAGGHPTAVAFGDGVSSVIVASQHLSGC SLYMYGEEKPTNLDSNKQQTKLPMPEIKWEHHKVHEQKAILTLSGAAANYDSGD GSTIIASCSEGTDIIIWHAKTGKILGNVDTNQLKNTMSAISPNGRFIAAAAFTA DVKVWEIVYSKDGSVKGVTKVMQLKGHKSAVTWLCFTPNSEQIVTASKDGSIRI WNINVRYHLDEDTKTLKVFPIPLQDSSGTTLHYERLSLSPDGKILAATHGSMLQ WLCIETGKVLDTAEKAHDGDITCMSWAPQSIPTGDKKVNVLATASGDKKVKLWA APPLPS	96	1412

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
343	WD40 repeat protein	MEVEPKKASKTFPVKPKLKPKPRTPSGKTPESKYWSSFKTTHPLDNLSFSVPSL AFSPSPPHLLAAAHSATVSLFSPHRTTISSFSDVVSSLSFRSDGQLLAASDLSG LIQVFDVRSRTPLRRLRSHARPVRFVRYPVLDKLHLVSGGDDALVKYWDVAGES VVSELRGHKDYVRCGDCSPADANCFVTGSYDHVVKLWDVRVRDGNRAATEVNHG SPVQDVIFLPSGSLVATAGGNSVKIWDLIGGGRMVYSMESHNKTVTSICVGTMG AQQSGEEGVQLRILSVGLDGYMKVFDYSRMKVTHSMRFPAPLLSIGFSPDSNVR AIGTSNGILYVGKRKAKENAEGGANGILGLGSVEEPRRRVLKPSFYRYFHRGQS EKPSEGDYLVMRPKKVKLAEHDKLLKKFQHKNALISVLGGNDPEKVVAVMEELV ARRALLKCVLNLDADELGLILTFLHKNSTVPRYSSLLLGLAKKVIDLRLEDIRA SDALKGHIRNLKRSVDEEIRIQEGLQEIQGMVSPLLRIAGRR	116	1702
344	WD40 repeat protein	MQGGSSGVGYGLKYQARCISDVKADTDHTSFLTGTLSLKEENEVHLLRLSSGGT ELICEGLFSHPSEIWDLSSCPFDQRIFSTVFSTGESYGAAVWQIPELYGQLNSP QLEKIASLDAHSRKISCVLWWPSGRHDKLVSIDEENIFLWGLDCSKKSAQVQSQ ESAGMLHNLSGGAWDPHDVNTVAATCESSIQFWDLRTMKKANSLESVHARDLDY DMRKKHLLVTSEDESGVRVWDLRMPKAPIQEFPGHTHWTWAVRCNPDYEGLILS AGTDSAVNLWWSSTASSDELISERLIDSPTRKLDPLLHSYNDYEDSVYGLAWSS REPWIFASLSYDGRVVVESVKPFLSRK	46	1101
345	WD40 repeat protein	MAEEEGSAELEQQLEEEFAVWKKNTPILYDLLISHALEWPSLTVHWAPLLPQPS SSAAAAAGDPSLAAHRLVLGTHTSDGAPNFLILADALLPSSESDHCGDDAVLPK VEISQKIRVDGEVNRARFMPQNHNIVGAKTNGCEVYVFDCSKQAAKQHDGGFDP DLRLTGHDGEGYGLSWSPLKENYLLSASHDKKICLWDISAAAQDKVLGAMHVFE AHEGAVGDASWHSKNDNLFGSAGDDCQLMIWDLRTNKAQQCVKAHEKEVNSVSF NSYNDWILATASSDTTVGLFDMRKLTTPLHVFSSHEGEVLQVEWDPNHEAVLAS SSEDRRVMVWDLNRIGDEQQEGDASDGPAELLFSHGGHKAKISDFSWNKNEPWV ISSVAEDNSVQVWQMAESICGDDDDMQAMEGYI	23	1258
346	WD40 repeat protein	MGNYGEEDEDQYFDALEETASVSDRGSNSSDCCSSGSGLDENVLDSLGFEFWTK FPESVRARRNRFLMLTGLGIEANSVDKEDAFPPSCNEIEVYTCKVTRDDGAVQR SLDSYNCISLLQSSTSIRSNQEVESLRGDSLLSSFRGRSKESDDLTELCGMGCP ESKRNAVSEFGSVSQGSIEELRRIVASSPLVHPLLHRKLEYERELIETKQKMGA GWLRKFGSATCISGRQGDTWSDPDDLEITAGMKMRRVRAHSSKKKYKELSSLYA AQEFLAHEGSISTMKFSMDGQYLASAGEDTVVRVWKVTEEDRSERVNVTVDPSC LYFALNESTQLASLNTNKEHIGKAKTFQRSSDSSCVILPLKVFQITEKPWHEFK GHNGEVLDLSWSSKGYLLSSSTDKTVRLWRVGCDRCQRVYSHNDYVTCISFNPV NENFFISGSIDGKVRIWNVFGGQVVAYIDCREIVSAVCYRSDGKGAIVGTMTGN CLFYSIKDNHLQMDAQVYLHGKKKSPGKRITGFQFPPNDPGKLMITSADSVIRV LSGLDVVCKLKGPRNSGGPMIATFTSDGKHVISASEDSNVYIWNYAGQDKTSSR VKKIWSCESFWSSNASVALPWCGIRTVPEALAPPSRSEERRASCAENGENHHML EEYFQKMPPYSPDCFSLSRGFFLELLPKGSATWPEEKLSDTSPPTVSSQAISKL EYKFLKSACHSVLSSAHMWGLVIVTAGWDGRIRTYHNYGLPVRS	404	2644

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
347	WD40 repeat protein	MDIDFKEYRLRCELRGHEDDVRGVCVCGDGSIGTSSRDRTVRLWAPSAGERRKY EVARVLLGHKSFVGPLAWVPPSEELPEGGIVSGGMDTLVMAWDLRNGEAQTLKG HQLQVTGIVLDGGDIVSASVDCTLIRWKNGQLTEHWEAHKAPIQAVIRLPSGEL VTGSSDTTLKLWRGKTCTQTFVGHTDTVRGLAVMPDLGILSASHDGSIRLWAVS GECLMEMVDHTSIVYSVDSHASGLIVSGSEDRFAKIWKDGVCFQSIEHPGCVWD VKFLEDGDIVTACSDGTIRIWTNQEDRMANSTELELFDLELSSYKRSRKRVGGL KLEELPGLEALQVPGTSDGQTKVIREGDNGVAYAWNSTELKWDKIGEVVDGPED SMNRPALDGVQYDYVFDVDIGDGEPTRKLPYNRSDNPYDTADKWLLKENLPLSY RQQIVEFILANSGQRDFNLDPSFRDPYTGSSAYVPGAPSQLAAKQARPTFKHIP KKGMLVFDAAQFDGILKKINEFNNTLLSNQEKKNLSLTDIEISRLGAVVKILKD TSHYHSSKFADADFDLMLKLLESWPYEMMFPVIDIFRMVILHPDGADGLLRHQE DKKDVLMESIKRATGNPSVPANFLTSIRAVTNLFKNSAYYSWLQKHRSEMLDAF SSCSSSSNKNLQLSYATLLLNYAVLLIEKKDEEGQSQVLSAALELAENESLEVD ARYRALVAIGSLMLDGLVKRIALDFDVEHIAKAARTSKEAKIAEVGADIELLIK QS	107	2383
348	WD40 repeat protein	MEFTEAYKQSGPCCFSPNARFIAVAVDYRLVIRDTLSLKVVQLFSCLDKISYIE WALDSEYILCGLYKRPMIQAWSLIQPEWTCKIDEGPAGIAYARWSPDSRHILTT SDFQLRLTVWSLVNTACVHVQWPKHASKGVSFTRDGKFAAICTRHDCKDYINLL SCHNWEIMGVFAVDTLDLADIQWSPDDSAIVIWDSPLEYKVLVYSPDGRCLFKY QAYESGLGVKSVSWSPCGQFLAVGSYDQMLRVLSHLTWKTFAEFTHLSNVRAPC CAAIFKEVDEPLQIDMSELSLSDDYMQGNSGDAPEGHYRVRYDVTEVPITLPCQ KPPADRPNPKQGIGLMSWSNDSQYICTRNDSMPTILWIWDMRHLELAAILVQKD PIRAAVWDPTGTRLVLCTGSSHLYMWTPSGAYCVSVPLSQFNITDLKWNSDGSC LLLKDKESFCCAAAPLPPDESSDYSSDD	243	1625
349	WD40 repeat protein	MATIAALDDDMVRSMSIGAVFSDFVGKLNSLDFHRKDDILVTAGEDDSVRLYDI ANARLLKTTFHKKHGTDRVCFTHHPNSLICSSTKNLDTGESLRYISMYDNRSLR YFKGHKQRVVSLCMSPINDSFMSGSLDHSVRMWDLRVNACQGILRLRGRPTVAY DQQGLVFAVAMEGGAIKLFDSRSYDKGPFDAFLVGGDTSEVCDIKFSNDGKSVL LSTTNNNIYVLDAYAGDKQCGFNLEPSPSTPIEASFSPDGQYVVSGSGDGTLHA WNISRRNEVACWNSHIGVASCLKWAPRRAMFVAASTVLTFWIPNSEPELASAKG EAGVPPEQV	126	1127
350	WD40 repeat protein	MSVAELKERHRAATETVNSLRERLKQKRVQLLDTDVAGYARTQGKTPVTFGATD LVCCRTLQGHTGKVYSLDWTPERNRIVSVSQDGRFIVWNALTSQKTHAIRLPCA WVMTCAFAPNGQSVACGGLDSVCSIFNLNSPVDRDGNLPVSRMLSGHKGYVSSC QYVPDGDAHLITGSGDQTCVLWDITTGLRTSVFGGEFQSGHTADVLSVSINGSS PRIFVSGSCDSTARMWDTRVASRAVHTYHGHEGDVNAVKFFPDGNRFGTGSDDG TCRLFDIRTGHELQVYYQQRGIDEIPHVTSIAFSISGRLLIAGYSNGDCFVWDT LLAQVVLNLGSLQNSHEGRISCLGVSADGSALCTGSWDTNLKIWAFGGIRRVT	257	1390

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
351	WD40 repeat protein	MKKRPRGASLDQAVVDIRRREVGGLSGLSFARRLAASEGLVLRLDIYNKLKGHR GCVNTVGFNLDGDIVISGSDDRHVKLWDWQTGKVKLSFDSGHLSNVFQAKIMPY TDDRSIVTCAADGQARHAQILEGGQVQTMLLAKHRGRAHKLAIDPGSPHIVYTC GEDGLVQRLDLRSNTARELFTCREVYGTHVEVVHLNAIAIDPRNPNLFVIGGSD EYARVYDIRNYKWNGSHNFGRSANYFCPSHLIGEAHVGITGLAFSGQSELLVSY NDESIYLFTQEMGLGPDPLSASTKSVDSNSSEVTSPTAVNVDDNVTPQVYKGHR NCETVKGVGFFGPKCEYVVSGSDCGRIFIWKKKGGQLIRVMAADKHVVNCIEPH PHIPALASSGIENDIKIWTPKAIERATLPMNVEQLKPKARGWMNRISSPRQLLL QLYSLERWPEHGGETSSGLAAGQEELTELFFALSANGNGSPDGGGDPSGPLL	178	1632
352	WD40 repeat protein	MSKRGYKLQEFVAHSSNVNCLSIGKKACRLFLTGGDDCKVNLWAIGKPNSLMSL CGHTNAVESVAFDSAEVLVLAGASSGVIKLWDVEEAKMVRGLTGHRSNCTAMEF HPFGEFFASGSTDTNLKIWDIRKKGCIHTYKGHTRGISTIRFSPDGRWVVSGGN DNVVKVWDLTAGKLLHDFKFHENHIRSIDFHPLEFLLATGSADRTVKFWDLETF ELIGSSRPEAAGVRAIAFHPDGRTLFCGLEDSLKVYSWEPVICHDGVDMGWSTL ADLCIHDGKLLGCSYYQSSVGVWVADASLIEPYGTNVKPQQKDSGDDEIEHQES RPSAKVGTTIRSTSIMRCASPDYETKDIKNIYVDTASGNPVSSQRVGTTNFAKV TQPLDFNDTPNLTLRRQGLVTETPDGLSGHVPSKSITQPKVVSRDSPDGKDSSR RESITFSRTKPGMLLRPAHSRRPSSTKYDVDRLSACAEIGVLSSAKSGSESLVD SFLNIKVAPEDGARNGCEDNHSSVKNVSVESEKVLPLQTPKTEKCDQTVGFKEE INSVKFVNGVAVVPGRTRTLVEKFEKREKLNSTEDQTINTPENPTLDKTPPPSL AENEEKSDRLNIVERKATRMSSHMVTAEDRTPVTLVGSPEDQSTVMAPQRELPA DESSKTPPLPVEDLEIHHGSNVSEDKATILSSQTVSEEDSKRSTLIRNFRRRDR FKSTEGRSPVMATQRKLPTDESGKTSSLPMEDLEIKGGLNVSEDKATSFSSRAP PREDRAHSALVRNVRKRDKFKSTNDTITVMVHQRGLSTDEASTVSVERVERRQL SNNVENPLNNLPPHSVPPTTTRGEPQYVGSESDSVNHEDVTELLLGNHEVFLST LRSRLTKLQVV	290	2917
353	WD40 repeat protein	MSTFLTGTALSNPNPNKSYEVVQPPNDSVSSLSFNPKANFLVATSWDNQVRCWE IVRSGTSLGTTPKASISHDQPVLCSTWKDDGTTVFSGGCDKQVKMWPLSGGQPM TVAMHDAPIKEISWIPEMNLLVTGSWDKTLRYWDTRQANPVHIQQLPERCYALT VRHPLMVVGTADRNLIIYNLQSPQTEFKRISSPLKYQTRCLAAFPDQQGFLVGS IEGRVGVHHLDDSQQSKNFTFKCHREGSEIYSVNSLNFHPVHHTFATAGSDGAF NFWDKDSKQRLKAMSRCSQPIPCSTFNNDGSIFAYSACYDWSKGAENHNPATAK TYIFLHLPQESEVKGKPRLGTTGRK	148	1197
354	WD40 repeat protein	MEVEAQQRDVNNVMCQLVDPEGTTLGPPMYLPQDVGPQQLQQMVNKLLSNEDKL PYTFYISDQELVVPLESYLQKNKVSVEKVLSIVYQPQAIFRIRPVNRCSATIAG HSEAVLSVAFSPDGKQLASGSGDTTVRLWDLSTQTPMFTCKGHKNWVLSIAWSP DGKHLVSGSKAGEIQCWDPLTGQPSGNPLVGHKKWITGISWEPVHLSSPCRRFV SSSKDGDARIWDVTLRRCVICLSGHTLAVTCVKWGGDGVIYTGSQDCTIKVWET SQGKLIRELKGHGHWVNSLALSTEYVLRTGAFDHTGKQYSSAEEMKQVALERYK KMKGNAPERLVSGSDDFTMFLWEPSVSKHPKTRMTGHQQLVNHVYFSPDGQWVA SASFDKSVKLWNGITGKFVAAFRGHVGPVYQISWSADSRLLLSGSKDSTLKIWD IRTKKLKRDLPGHADEVFAVDWSPDGEKVVSGGKDKVLKLWMG	140	1567

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
355	WD40 repeat protein	MDAGSAHSSSNMKTQSRSPLQEQFLQRRNSRENLDRFIPNRSAMDFDYAHYMLT EGRKGKENPAVSSPSREAYRKQLAETLNMNRTRILAFKNKPPTPVELIPHELTS AQPAKPTKTRRYIPQTSERTLDAPDLLDDYYLNLLDWGSSNVLSIALGNTVYLW NASDGSTSELVTIDDETGPVTSVSWAPDGRHIAVGLNNSDVQLWDSADNRLLRT LRGGHRSRVGSLAWNNHILTTGGMDGLIVNNDVRVRSHIVDTYRGHTQEVCGLK WSASGQQLASGGNDNILHIWDRSTASSNSPTQWLHRLEEHTAAVKALAWCPFQG NLLASGGGGGDRTIKFWNTHTGACLNSVDTGSQVCALLWNKNERELLSSHGFTQ NQLTLWKYPSMVKIAELTGHTSRVLFMAQSPDGCTVASAAGDETLRFWNVFGVP EVAKPAPKANPEPFAHLNRIR	376	1737
356	WD40 repeat protein	MEEAIPFKNLPSREYQGHKKKVHSVAWNCTGTKLASGSVDQTARVWHIEPHGHG KVKDIELKGHTDSVDQLCWDPKHADLIATASGDKTVRLWDARSGKCSQQAELSG ENINITYKPDGTHVAVGNRDDELTILDVRKFKPIHKRKFNYEVNEIAWNMSGEM FFLTTGNGTVEVLAYPSLRPVDTLMAHTAGCYCIAIDPVGRYFAVGSADSLVSL WDISEMLCVRTFTKLEWPVRTISFNHTGDYVASASEDLFIDISNVQTGRTVHQI PCRAAMNSVEWNPKYNLLAYAGDDKNKYQADEGVFRIFGFESA	69	1010
357	WD40 repeat protein	MGKDEEEMRGEIEERLINEEYKVWKKNTPFLYDLVITHALEWPSLTVEWLPDRE EPPGKDYSVQKLVLGTHTSENEPNYLMLAQVQLPLEDAENDARHYDDDRADVGG FGCANGKVQIIQQINHDGEVNRARYMPQNSFIIATKTVSAEVYVFDYSKHPSKP PLDGACSPDLRLRGHSTEGYGLSWSKFKQGHLLSGSDDAQICLWDINATPKNKS LDAMQIFKVHEGVVEDVAWHLRHEYLFGSVGDDQYLLIWDLRTPSVTKPVQSVV AHQSEVNCLAFNPFNEWVVATGSTDKTVKLFDLRKISTALHTFDAHKEEVFQVG WNPKNETILASCCLGRRLMVWDLSRIDEEQTPEDAEDGPPELLFIHGGHTSKIS DFSWNTCEDWVVASVAEDNILQIWQMAENIYHDEDDVPGEESNKGS	149	1423
358	WD40 repeat protein	MMRGFSCTEDGDAPSTSSTSPPPPPPPPHRQQMQAPRASSSSSGQPTSRRSTGN VFKLLARREVSPRSKHSLKKFWGEASECQLCPFQQSYEAVRDVRRSLISWVEAF SLQHLSAKYCPLMPPPRSTIAAAFSPDGKILASTHGDHTVKLIDSQTGSCLKVL RGHRRTPWVVRFHPLYPEILASGSLDHEVHLWDANTAECIGSRNFYRPIASIAF HAQGDLLAVASGHKLYIWHYNRSGETSSPTIVLRTPRSLRAVHFHPHAAPFLLT AEVNDLDLTDSAMTLATSPGYLHYPPPTIYLADAHSNERSRLEDELPLMPSPLL MWPSFTRDDGRATLPHIGGDVGLSGQQRVDSLSSGQYEFHPSPIEPSSSTSMHE EMGTDPFSSVRESEVTQSAMNIVDNTEVQPEERSTYSFSFSDPRFWELPSVYGW LVGQTQAAPRTAPSPGALETASALGEVASVSPVRSEFMPGGMDQPRLGGRSGSG CRSSGSRMMRTAGLNDHPHDENYPQSVVSKLRSELEASLAAAASTELPCTVKLR VWPYDMKDPCALFRSESCRLTIPHAVLCSEMGAHFSPCGRFFAACVACVLPQLE ADPVLHGQVDPDVTGVATSPTRHPVSAYQIMYELRIYSLEEATFGMVLASRSIR AAHCLTSIQFSPTSEHLLLAYGRRHNSLLKSIVIDGENTVPIYSILEVYRVSDM ELVRVLPSAEDEVNVACFHPSVGGGLVYGTKEGKLRILQIDSSGGLNPKSTGFL DENMAEVPTYALEC	365	2677

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
359	WD40 repeat protein	MGEGDLPRTEAGVLRGHEGAVLAARFNGDGNYCLSCGKDRTIRLWNPHRGIHIK TYKSHGREVRDVHCTSDNSKLISCGGDRQIFYWDVSTGRVIRRFRGHDSEVNAV KFNDYASVVVSAGYDRSVRAWDCRSHSTEPIQIINTFQDSVMSVCLTKTEIIGG SVDGTVRTFDIRIGREISDDLGQPVNCISMSNDGNCILASCLDSTLRLVDRSAG ELLQEYKGHTCKSYKLDCCLTNTDAHVAGGSEDGYVFFWDLVDASVISKFRAHS SVVTSVSYHPKEDCMITASVDGTIKVWKT	24	923
360	WD40 repeat protein	MACIKGVGRSASVAMAPDGGYLATGTMAGTVDLSFSSSASLEIFGLDFQSDDRD LPLIAESPSSERFNRLSWGKNGSGSDEFSLGLIAGGLVDGTIGLWNPLSLIRSE AGDKAIVGHLSRHKGPVRGLEFNVIAPNLLASGADDGEICIWDLAAPREPSHFP PLRGSGSAAQGEISFLSWNSKVQHILASTSYNGTTVVWDLKKQKPVISFSDSVR RRCSVLQWNPDLATQLVVASDEDSSPTLRLWDMRNIMSPVKEFAGHTRGVIAMS WCPNDSSYLVTCAKDNRTICWDTVTGEIVCELPAGSNWNFDVHWYPKIPGVISA SSFDGKIGIYNVEGCSRYGVRENEFGAATLRAPKWFKRPVGASFGFGGKVVSFH TRSTGGPSVNSSEVFVHDIITEQTLVSRSSEFEAAIQSGDRPSLRALCEKKSQH CESTDDQETWGFLKVLLEDDGTARSKLLAHLGFDIPTETNDGSQEDLSQQVNAL GLEDVTADKVVQEDNNESMVFPTDNGEDFFNNLPSPRADTPVSTSADGFPTVNA AVEPSQDEVDGLEESSDPSFDDSVQRALVVGDYKAAVALCMSANKLADALVIAH VGGASLWESTRDKYLKMSRLPYLKVVFAMVNNDLQSLVDTRPLKFWKETLAILC SFAQGEEWAMLCNSLASKLMAAGNMLAATLCFICAGNIDKTVEIWSRSLATEHD GMSYMDLLQDLMEKTIVLALASGQKQFSASVCKLVEKYAEILASQGLLTTAMDY LKLLGTDDLSPELAVLRDRIAFSVEAEKGANISAFNGSQDPRGAVYGVDQSNYG MVDTSQHYYPEAAQPQVPHTVPGSPYGENYQQPFGSSFGKGYNTPMQYQAPSQA SMFVPSEPPQNAQPSFVPTPVTSQPTTRSQFIPAPPLALRNPEQYQQPTLGSHL YPGSVNPTFQPLPHAPGPVAPVPPQVSSVPGQNMPQAVAPTQMRGFMPVTNPGV VQNPGPISMQPATPIESAAAQPVVSPAAPPPTVQTADTSNVPAPQKPVIATLTR LYNETSEALGGSRANPAKKREIEDNSRKIGALFAKLNSGDISKNAADKLVQLCQ ALDNGDYSTALQIQVLLTTSEWDECNFWLATLKRMIKTRQNVRLS	221	3598
361	WD40 repeat protein	MKERGKGAGRSVDERYTQWKSLVPVLYDWLANHNLVWPSLSCRWGPQLEQATYK NRQRLYLSEQTDGSVPNTLVIANVEVVKPRVAAAEHISQFNEEARSPFVKKFKT IIHPGEVNRIRELPQNSKIVATHTDSPDVLIWDVETQPNRHAVLGASTSRPDLI LTGHKDNAEFALAMSPTEPFVLSGGKDRYVVLWSIQDHISTLAADPGSAKSPGS AGTNNKQSSKAAGGNDKTGDSPSIEPRGVYLGHGDTVEDVTFCPSSAQEFCSVG DDSCLILWDARTGSSPAIKVEKAHHADLHCVDWNPHDVNLILTGSADNTVRMFD RRNLTSGGVGSPVHTFEGHNAAVLCVQWSPDKSSVFGSSAEDGILNIWDHEKIG RKIETVGSKVPNSPPGLFFRHAGHRDKVVDFHWNSSDPWTIVSVSDDGESTGGG GTLQIWRMIDLIYRPEEEVLAELDKFKSHILSCTS	44	1447
362	WD40 repeat protein	MAKIAPGCEPVAGTLTPSKKREYRVTNRLQEGKRPLYAVVFNFIDSRYFNVFAT VGGNRVTVYQCLEGGVIAVLQSYIDEDKDESFYTVSWACNIDRTPFVVAGGING IIRVIDAGNEKIHRSFVGHGDSINEIRTQPLNPSLIVSASKDESVRLWNVHTGI CILIFAGAGGHRNEVLSVDFHPSDKYRIASCGMDNTVKIWSMKEFWTYVEKSFT WTDLPSKFPTKYVQFPVFIAPVHSNYVDCNRWLGDFVLSKSVDNEIVLWEPKMK EQSPGEGSVDILQKYPVPECDIWFIKFSCDFHYHSIAIGNREGKIYVWELQSSP PVLIAKLSHPQSKSPIRQTAMSFDGSTILSCCEDGTIWRWDAITASTS	196	1314

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
363	WD40 repeat protein	MNTAMHFGAGWRSIAEMGYTMSRLEIEPESCEDEKSLDGVGNSQGPNELPRCLD HELAHLTNLKSRPHEHLIRDFPGRRALPVSTVKMLAGRECNYSRRGRFSSADCC HMLSRYVPVNGPSPLDQMNSRAYVSQFSADGSLFVAGFQGSHIRIYNVDKGWKC QKNILTKSLRWTITDTSLSPDQRYLVYASMSPIVHIVDIGSAAMDSLANITEIH EGLDFSADSGPYSFGIFSVKFSTDGREVVAGSSDDSIYVYDLVANKLSLRIPAH ESDVNTVCFADESGHIIYSGSDDTYCKVWDRRCLSARNKPAGVLMGHLEGITFI DSRGDGRYFISNGKDQTIKLWDIRKMGSDICRRGFRNFEWDYRWMDYPPRARDS KHPFDLSVATYKGHSVLRTLIRCYFSPVHSTGQKYIYTGSHDSCVYIYDVVTGA QVAALKHHKSPVRDCSWHPEYPMIVSSSWDGDIVKWEFFGNGETEIPAMKKRIR RRHLY	193	1668
364	WD40 repeat protein	MEPQPQAPKKRGRKPKPKEDKKEEQLHQPPPPPPPQQQAAPAPAPAATRSSTSG SAGGRDRRPQQQHAVDEKYARWKSLVPVLYDWLANHNLLWPSLSCRWGPQLEQA TYKNRQRLYISEQTDGSVPNTLVIANCEVVKPRVAAAEHVSQFNEEARSPFIRK YKTIIHPGEVNRVRELPQNPNIVATHTDSPDVLIWDVESQPNRHAVYGATASRP NLILTGHQENAEFALAMCPAEPFVLSGGKDKTVVLWSIQDHITASATDQTTNKS PGSGGSIIKKTGEGNEETGNGPSVGPRGIYCGHEDTVEDVAFCPSTAQEFCSVG DDSCLILWDARVGTNPVAKVEKAHNGDLHCVDWNPHDNNLILTGSADNSVNMFD RRNLTSNGVGSPVYKFEGHKAAVLCVQWSPDKPSVFGSSAEDGLLNIWDYERVD KKVDRAPNAPAGLFFQHAGHRDKIVDFHWNAADPWTMVSVSDDCDTAGGGGTLQ IWRMSDLIYRPEEEVLAELENFKAHVLECSKA	78	1634
365	WD40 repeat protein	MGIFEPYRAVGYITTGVPFSVQRLGTETFVTVSVGKAFQVYNCAKLSLVLVGPQ LPKKIRALASYREYTFAAYGSDIGIFKRAHQLATWSGHTAKVCLLLLFGEHILS VDVDGNAYIWAFKGMNYNLSPVGHILLDSNFTPSCIMHPDTYLNKVILGSQEGP LQLWNISTKTKLYEFKGWNSSVSSCVSSPALDVVAVGCADGKIHVHNIRYDEEL VTFSHSMRGSVTALSFSTDGQPLLASGSSSGVVSIWNLDKRRLQSVIRDAHDGS IISLHFFANEPVLMSSSADNSIKMWIFDTSDGDPRLLRFRSGHSAPPLCIRFYA NGRHILSAGQDRAFRLFSVVQDQQSRELSQRHVSKRAKKLKLKEEEIKLKPVIA FDVAEIRERDWCNVVTSHMDTPQAYVWRLQNFVIGEHILRPCPNKPTPVKACMI SACGNFAILGTAGGWIERFNLQSGISRGSYIDQLEGTNSAHDGEVVGVACDATN TLMISAGYAGDIKVWDFKGRELKSRWEIGSSLVKISYHRLNGLLATVADDFIIR LFDAVALRMVRKFEGHTDRITDLCFSEDGKWLLSSSMDGSLRIWDIILARQVDA VFVDVSITALSLSPNMDILATTHVDQNGVFLWVNQSMFSGDSDINLYASGKEVV TVKLPSVSSVEGSQVEESNEPTIRHSESKDVPSFRPSLEQIPDLVTLSLLPKSQ WQSLINLDIIKVRNKPVEPPKKPEKAPFFLPSIPSLSGEILFKPSEMSDKGDMK ADEDKSKITPEVPSSRFLQLLHSCSEAKNFSPFTTYIKGLSPSTLDLELRMLQI IDDDAVDADADDPQDVDKRQELLSIELLMDYFIHEISCRSNFEFVQALVRLFLK IHGETIRRQSVLQNKAKVLLETQCSVWQRVDKLFQGARCMVAFLSNSQF	85	2826

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
366	WD40 repeat protein	MEETKVTCGSWIRRPENVNLAVLGRSPRRRGSAALEIFAFDPKSTSLSSSPLVA HVIEEIEGDPLAIAVHPNGEDIVCFASSGSCLSFELSGQESNLKLLTKELPPLR GIGPQKCMAFSVDGSRFATGGVDGRLRILEWPSLRIILDEPKAHKSIRDLDFSL DSEFLATTSTDGSARIWKAEDGLPCTTLTRRSDEKIELCRFSKDGTKPFLFCTV QRGDKAVTGVWDISTWNKIGHKRLLRKPAVVMSISLDGKYLAQGSKDGDMCVVE VKKMEVSHWSKRLHLGTSLTSLEFCPIERVVITTSDEWGVLVTKLNVPADWKAW QVYLLLLGLFLASLVAFYIFYENSDSFWGFPLGKDQPARPKIGSVLGDPKSADD QNMWGEFGPLDM	74	1246
367	WD40 repeat protein	MADPVEHQHQQHQQHQLQQQRRRGWRIQGGQYLGEISALCFLHLPPPPLSLSSS PVLSLSSGLDSESRDRPACSFRFPSAGSGSQVSLFDLASGAMVRTFYVFRGIRV HGIVLGCADFPGGSSSSSSTLDYVIAVYGERRVKLFRLSVRLGRGAGEGSGTVL SADLELVSAAPRLSHWVMDVRFLKENGTSEDELQRCLTVAIGCSDNSIRLWDVD KCSFVLAVSSPERCLLYSMRLWGDNLEDLQVASGTIYNEILIWKVVPNHDAPSS NELTEEGLTNSCAGNSVHECLRYEAYHICRLVGHEGSIFRIAWSSDGSKLVSVS DDRSARIWEVHCKVQYSEDAGEVGLLFGHSARVWDCYISDNLIVTAGEDCSCRV WGLDGQQHDVIKEHIGRGIWRCLYDPWSSLLVTGGFDSAIKVHKLDASLAEASA KQSNIKDLSDGTELFTTHLPNSSGHSGHMDSKSEYVRCLSFSCEDVMYIATNHG YLYHAKLCNDGDLRWTELAQVSNEVQIICMELLPSNPYDPRIDADDWVAVGDGK GWTTVVRVVKNSDSPKVSTSFSWAAEMDRQLLGIHWCKSLGHRFIFTADPRGAL KLWRFFEVSQSSSLYPENSPRISLIAEFKSDLGARIMCLDVAFESELLICGDLR GNLVLFPLLKDLLLDTFVVSAAKISPVNHFKGAHGISAVSSISVAHMSFNHIEL RSTGADGCICYMEYDKGLQSLNFVGMKQVKELSMIESVSTENESTGYRTSGSYA SGFASTDFIIWNLVTEAKVLQVSCGGWRRPHSYYLGDVPEMKNCFAYVKDDIIY IRRHWIKDSKDKILPQNLRLQFHGREVHSLCFVTGDFQLRKNKQSSWIVTGCED GTVRLTRYTQCTDNWSSSKLLGEHVGGSAVRSICCVSNIHTTSSGTSVSDVKGI ENLPKDIKGTLMEDECNPSLLISVGAKRVLTSWLLRRRKQDGKEDDVTDLQEAE NSSLPSSAGSSTFSFQWLSTDMPVKYSVPSKKSGSIKKLIGVSDTNVRCKSLLP DSEALQSKVSAVDKNEDDWRYLAVTAFLVRHSGSRLIVCFIIVACSDATLAIRA LVLPYRLWFDVALMVPLSSPVLSLQHVIIGRCQLPDENVQIGNVYYVVISGATDG SIAFWDLTESVEAFMRRLSNIHLEKFMDCQKRPRTGRGSQGGRWWRSLSKIACK EQPINDPVTAKAIKELNRKLTGGVACGSSSSMLDASPELDSNAANSSFEIIEVN PFHVLNGVHQSGVNCLHVCETKHGQSSDGRFLYQLVSGGDDQALHLLKFEVLVQ PPVQVPDVPNSDIRNSILVEEFLLDEQNQKTKCTIEFISQEKIASAHNSAVKGV WTDGTWVFSTGLDQRVRCWISKDRGTPTELAHFIISVPEPEALDARSICWDQYQ IAVAGRGMQMIEFHVPSSEIR	100	4377

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
368	WD40 repeat protein	MPYKLSATLSNHSSDVRAVASPSDDLILSASRDSTAISWFRQSPSSFTPASVIR AGSRFVNAIAYLPPTPRAPQGYAVVGGQDTVVNVFALGPGDKEEPEYTLVGHTD NVCALSVNSDDTIISGSWDKTAKVWKDFALVYDLKGHQQSVWAVLAMNEKEFLT ASADRTIKYWVQHKTMQTYEGHRDAVRGLALIPDIGFASCSNDSEIRVWTMGGD VVYTLSGHTSFVYSLSVLPNGDLVSAGEDRSVRVWRDGECSQVIVHPAISVWAV STMPNGDIISGSSDGVVRVFSESEKRWATASELKALEDQIASQSLPSQQVGDVK KTDLPGPEALSVPGKKAGEVKMIRSGDVVEAHQWDSLASSWQKIGEVVDAIGSG RKQLHDGKEYDYVFDVDIQEGAPPLKLPYNVSENPYTAAQRFLEQNDLPTGYLD QVVKFIEQNTAGVKLGNDGYVDPFTGASRYQPATQSTSNTASSSYMDPFTGGSR HIAESAPSNVPQGSHATGIIPFSKPIFFKLANVSAMQAKMFQFDEVLRNEISTA TLAMRPDEVIMVNETFTYLSKVVTSTSSARTSLGWIHIETIMQILDRWPVPQRF PVIDLGRLVTAYCMNAFSGPGDLEKFFSCLFRTSEWTSITSGSKALTKAQETNV LLLFRTIANSLDGAPLNDMEWIKQIFRELAQTPQLVLNKSHRLALASVLFNFSC IGLKGPVPADVRTLHLTIILQVLRSPNDDPEVAYRTCVALGNMLYSDKTRGTPR DAQSPSPTELKSAVAAIKGGFSDPRINDVHREIMSLI	58	2439
369	WD40 repeat protein	MPPQKIESGHKDTVHDLAMDYYGKRLATASSDHTINVVGVSSSGSQHLATLIGH QGPVWQISWAHPKFGSLLASCSYDGRVIIWREGNPNEWTQAQVFEEHKSSVNSV AWAPHELGLCLACGSSDGNISVFTARQDGGWDTSRIDQAHPVGVTSVSWAPSTA PGALVGSGMMEPVQKLCSGGCDNTVKVWKLYNRVWKLDCFPVLQMHTDWVRDVA WAPNLGLPKSTIASASQDGRVIIWTLAKEGDQWQGKVLYDFRTPVWRVSWSLTG NILAVADGNNNVSLWNEAVDGEWIQVSTVEP	159	1064
370	WD40 repeat protein	MSAPMLEIEARDVVKIVLQFCKENSLHQTFQTLQSECQVSLNTVDSIETFVADI NSGRWDAILPQVAQLKLPRNTLEDLYEQIVLEMIELRELDTARAILRQTQAMGV MKQEQPERYLRLEHLLVRTYFDPNEAYQDSTKEKRRAQIAQALAAEVTVVPPSR LMALVGQALKWQQHQGLLPPGTQFDLFRGTAAMKQDVDDMYPTTLSHTIKFGTK SHAECARFSPDGQFLVSCSVDGFIEVWDYMSGKLKKDLQYQADETFMMHDDPVL CVDFSRDSEMLASGSQDGKIKVWRIRTGQCLRRLERAHSQGVTSVLFSRDGSQL LSTSFDGSARIHGLKSGKQLKEFRGHSSYVNDAIFSNDGSRVITASSDCTVKVW DVKTSDCLQTFKPPPPLRGGDASVNSVHLFPKNADHIVVCNKTSSIYIMTLQGQ VVKSLSSGKREGGDFVAACVSPKGEWIYCVGEDRNLYCFSCQSGKLEHLMKVHE KDVIGVTHHPHRNLVATYSEDSTMKLWKP	118	1665
371	WD40 repeat protein	MDLLQSYAEDNDGDLGRHSSPEPSPPRLLPSKSAAPKVDDTTLALTVAQTNQTL ARPIDPSQHAVAFNPTYDQLWAPICGPAHPYAKDGIAQGMRNHKLGFVEDAAIG SFLFDEQYNTFQRYGYAADPCASTGNEYVGDLDALKQNDGISVYNIRQQEQKKY AEEYAKKKGEERGEGGREKAEVVSDKSTFHGKEERDYQGRSWIAPPKDAKATND HCYIPKRLVHTWSGHTKGVSAIRFFPKHGHLILSAGMDTKVKIWDVFNSGKCMR TYMGHSKAVRDISFCNDGTKFLTAGYDKNIKYWDTETGKVISTFSTGKIPYVVK LHPDDEKQNILLAGMSDKKIVQWDMNTGQITQEYDQHLGAVNTITFVDDNRRFV TSSDDKSLRVWEFGIPVVIKYISEPHMHSMPSISLHPNTNWLAAQSLDNQILIY STRERFQLNKKKRFAGHIVAGYACQVNFSPDGRFVMSGDGEGRCWFWDWKSCKV FRTLKCHEGVCIGCEWHPLEQSKVATCGWDGLIKYWD	57	1628

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
372	WD40 repeat protein	MESNGNLEQTLQDGRIYRQLNSLIVAHLRDHNFPQAASAVALATMTPLNVEAPR NRLLELVAKGLAVEKGELLRGVSHAGTNDLGGSIPASYGLVPAPWTAIDFSSLR DTKGMSKSFTKHETRHLSDHKNVARCARFSTDGRFFATGSADTSIKLFEVSKIK QMMLPDSTDGAIRAVIRTFYDHTHPVNDLDFHPQNTVLISAAKDHTVKFFDYSK ATAKRAFRVIQDTHNVRSVAFHPSGDFLLAGTDHPIPHLYDVNTFQCYLSANVP EFAVNAAINQVRYSSSGGMYVTASKDGTIRFWDGASANCVRSIAGAHGAAEVTS ANFTKDQRYVLSCGKDSTVKLWEVGTGRLVKQYLGATHMQLRCQAVFNNTEEFV LSIDEPSNEIVVWDAMTAEKVARWPSNHNGPPRWIEHSPTEAAFVSCGTDRSIR FWKETH	250	1566
373	WD40 repeat protein	MSNFQGEDGEYVADDFEAEDGDEELHGRESADPESDVDEIDTPSNRFTDTTADQ ARRGRDIQGIPWERLSITREKYRRTRLEQYKNYENVPQSGEKSGKDCTVTEKGN SFYEFRRNSRSVKSTILHFQLRNLVWATSKHDVYLMSNYSVVHWSSLTGKKSEV LNLAGHVAPNEKHPGSLLEGFTQTQVSTLAVKDRFLVAGGFQGELICKFLDRPG ISFCSRTTYDDNAITNAVEIYVSPSGGIHFIASNNDCGVRDFDMENFELSKHFR FPWPVNHTSLSPDGKLLVIVGDDPEGILVDAKTGKTIMPLRGHLDFSFASEWHP DGVTFATGNQDKTCRIWDIRNLSKSIAVLKGNLGAIRSIRYTSDGRYMAIAEPA DFVHVYDTKTGYKKEQEIDFFGEISGMSFSPDTESLFIGVWDRTYGSLLEYGRR RNFSYLDCLV	106	1434
374	WD40 repeat protein	MGVEEDLEDLNALAESTDAAVDGQAALASAVDSVTLQPAPPILPPVIPPPAVPV VAPVPTIPPVLRPLAPLPIRPPVLRPPAPKRDEAGSSDSDSDHDGTAAGSTAEY EITEESRLVRERHEKAMQDLMMKRRGAALAVPTNDKAVRARLRRLGEPMTLFGE REMERRDRLRMLMAKLDAEGQLEKLMKAHEDEEAAASAAPEDVEEEMLQYPFYT EGSKALFNARIDIAKFSITRAALRLERARRRRDDPDEDVDAEIDWALKKAESLS LHCSEIGDDRPLSGCSFSHDGKLLATCSMSGVAKLWDTCRMPQVNRVLTLKGHT ERATDVAFSPVQNHIATASADRTAKLWNTEGTILKTFEGHLDRLGRIAFHPSGK YLGTTSFDKTWRLWDIESGEELLLQEGHSRSIYGIDFHRDGSLVASCGLDALAR VWDLRTGRSILALEGHVKPVLGVSFSPNGYHLATGGEDNTCRIWDLRKKKSLYT IPAHANLISEVKFEPQEGYFLVTASYDTTAKVWSARDFKPVKTLSVHEAKITSV DITADASHIVTVSHDRTIKLWTSNDDVKEQAMDVD	190	1917

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
375	WD40 repeat protein	MVKAYLRYEPAAAFGVIASVESNIAYDASGKHLLAPALEKVGVWHVRQGVCTKA LAPSASSAAGPSLAVTAIASSPSSLIASGYADGSIRIWDFEKGSCETTLNGHKG AVSVLRYGKLGSLLASGSKDNDIILWDVVGETGLYRLRGHRDQVTDLVFLDSDK KLVSSSKDKYLRVWDLETQHCMQIVGGHHSEIWSLDTDPEERYLVTGSADPELR FYTVKNDSSDERSEADASGGVGNGDLASHNKWDVLKQFGEIQRQSKDRVATVRF NKNGNLLACQAAGKLVEVFRVLDEAEAKRKAKRRLHRKREKKGADVNENGDSSR GIGEGHDTMVTVADVFKLLQTIRASKKICSISFCPVAPKSSLATLALSLNNNLL EFHSIEADKTSKMLTIELQGHRSDVRSVTLSSDNTLLMSTSHNSVKIWNPSTGS CLRTIDSGYGLCGLIVPQNKHALIGTKDGAIEIFDVGSGTCIEVVEAHGGSIRS IVAIPNQNGFVTGSADHDIKFWEYGMKQKPGDNSKHLTVSNVRTLKMNDDVLVV AVSPDAQKIAVALLDCTVKVFFMDSLKLMHSLYGHRLPVLCLDISSDGDLIVTG SADKNLMIWGLDFGDRHKSIFAHGDSIMAVQFVGNTHYMFSVGKDRLVKYWDAD KFELLLTLEGHHADIWCLAISNRGDFLVTGSHDRSIRRWDRTEEPFFIEEEKEK RLEEMFESDLDNAFGNKYVPKEEIPEEGAVALAGKKTQETLSATDSIIEALDIA EVELKRIAEHEEEKNNGKTAEFHPNYVMLGLSPSDFILRALSNVQTNDLEQTLL ALPFSDALKLLSYLKDWTTYPDKVELVSRIATVLLQTHYNQLVSTPAARPLLTT LKDILHKKVKECKDTIGFNLAAMDHLKQLMALRSDALFQDAKVKLLEIRSQLSK RLEERTDPREAKRRKKKQKKSTNMHAWP	102	2942
376	WD40 repeat protein	MGGVQAEREDKDKVSLELTEEILQSMEVGMTFRDYSGRISSMDFHRASSYLVTA SDDESIRLYDVASATCLKTINSKKYGVDLVSFTSHPMTVIYSSKNGWDESLRLL SLHDNKYLRYFKGHHDRVVSLSLCPRNECFISGSLDRTVLLWDQRAEKCQGLLR VQGRPATAYDDPGLVFAIAFGGCVRMFDARKYEKGPFEIFSVGGDVSDANVVKF SNDGRLMLLTTTDGHIHVLDSFRGTLLYTFNVKPTSSKSTLEASFSPEGMFVIS GSGDGSVYAWSVRGGKEVASWLSTDTEPPVIKWAPGNLMFATGSSELSFWIPDL SKLGAYVGRK	75	1079
377	WD40 repeat protein	MAAFGAAPAGNHNPNKSSEVIQPPSDSVSSLCFSPRANHLVATSWDNQVRCWEL TKNGASVTSVPKASMSHDQPVLCSAWKDDGTTVFSGGCDKQAKMWSLMSGGQPV TVAMHDAPIKEIAWIPEMNVLVTGSWDKTLKYWDTRQSNPVHTQQLPERCYAMT VRYPLMVVGTADRNLIVFNLQNPQAEFKRFSSPLKYQTRCVAAFPDQQGFLVGS IEGRVGVHHLDDSQISKNFTFKCHRDNNDIYSVNSLNFHPVHHTFATAGSDGTF NFWDKDSKQRLKAMSRCSQPIPCSTFNNDGTIYAYSVCYDWSKGAENHNPATAK TYIFLHLPQESEVKAKPRVGTTNRK	99	1148
378	WD40 repeat protein	MNCSISGEVPEEPVVSTKSGHVFERRLIERYVSDYGKCPVSGEPLTMDDVLPVK MGKIVKPRPLQAASIPGLLSIFQNEWDSLMLSNFALEQQLHTARQELSHALYQH DAACRVIARLKKERDEARSLLALAERQIPMTASSDIAVNAPAMSNGRKASLDEE PGYAGKKMRPGISASIIAEITDCNLALSQQRKKRQIPSTLAPVEDLERYTQLSS YPLHKTGKPGITSLDICHSKDIIATGGIDTSAVLFDRSSGQIMSTLSGHSKKVT SVNFDAQGDMVLTGSADKTVRIWQGSEDGSYNCRHILKDHTAEVQAITVHATNN YFATASLDNTWCFYEFSTGLCLTQVEGASGSEGYTSAAFHPDGLILGTGTSNAD VKIWDVKTQANVTTFSGHTGAITAISFSENGYFLATAAQDGVKLWDLRKLKNFR TFSAYDKDTGTNSVEFDHSGCYLGLAGSDIRVYQVASVKSEWNCVKTFPDLSGT GKVTCVKFGPDSKYIAVGSMDHNLRIFGLPSEDGAMES	232	1806

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
379	WD40 repeat protein	MAAPGVETLKKEIKELKEKIAQHRLDTDGEQPLPAAAKSKSVPEVSAALKQRRI LKGHFGKIYALHWSADSRHLVSASQDGKLIIWNGFTTNKVHAIPLRSSWVMTCA YSPSGNLVACGGLDNLCSVYKVPHGGNKESSSAQKTYGELAQHEGYLSCCRFIK DNEIVTSSGDSTCILWDVETKTPKAIFNDHTGDVMSLAVFDDKGVFVSGSCDAT AKLWDHRVHKQCVMTFQGHESDINSVQFFPDGDAFGTGSDDSSCRLFDIRAYQQ INKYSSDKILCGITSVAFSKTGKSLFAGYDDYNTYVWDTLSGNQVEVLTGHENR VSCLGVSEDGKALATGSWDTLLKIWA	72	1124
380	WD40 repeat protein	MGGVEDESEPASKRMKLSSRVLRGLANGSSRTEPAAGSSLDLMARPLPIEGDEE VIGSKGVIKRVEFVRLIAKALYSLGYEKSGARLEEESGIPLQSSVVNLFMQQIS DGLWDESVVTLHKIGLSDENLVKSASFLILEQKFLELLDQEKAMDALKTLRTEI TPLCIKNSRVRELSSCIISPSSCGLLNQNKRNSTRARSRSELLEELQKLLPPAV IIPERRLEHLVEQALVLQTDACMLHNSIDMEMSLYTDHQCGKEHIPCRTLQILQ SHNDEVWLVQFSHNGKYLASASNDRSAIIWEVDENGSVSLKHKLTGHQKPISSV CWSPDDRQLLTCGVGETVRRWDVSSGECLRVYEKAGHGLISCAWFPDGKWICYG VSDRSICMCDLEGKEIECWKGQRTLSISDLEITSDGKQIISICRETAILLLDRE AKYERMIEENQTITSFSLSKDNRYLLVNLLNQEIHLWDIKGDFRLVAKYKGLKR SRFVIRSCFGGLKQAFVASGSEDSQVYIWHKGSGELIEPLPGHSGAVNCVSWNP ANHHMLASASDDRTIRIWGLNELNTRHKGARPNGVHYCNGNGTS	315	2069
381	WD40 repeat protein	MTQLAETYACMPSTERGRGILIAGNPKPGSNSVLYTNGRSVVILNLDNPLDISV YAEHAYPATVARFSPNGEWVASADSSGAVRIWGAYNDHVLKKEFKVLSGRIDDL QWSPDGLRIVASGDGKGKSLVRAFMWDSGTNVGEFDGHSRRVLSCAFKPTRPFR IVTCGEDFLVNFYEGPPFKFKLSRRDHSNFVNCLRFSPDGNRFISVSSDKKGII YDGKTGEKIGELSSDGGHTGSIYAVSWSPDSKQVITVSADKSAKIWDISEDGSG NLRKTLTSSGSGGVDDMLVGCLWQNNHLVTVSLGGTISIYTAGDLDKAPVSFSG HMKNVSSLSVLKGDPKVILSSSYDGLIIKWIQGIGFSGRVQRKESTQIKCLAAV DEEIVTSGYDNKVCRVSGSGDAEFIDIGCQPKDLSLALQCPEFALVSTDTGVVL LRGAKIVSTINLGFAVTASTVAPDGTEAIIGAQDGKLRIYSISGDTLTEEAVLE KHRGAISVIHYSPDLSMFASGDLNREAVVWDRASREVRLKNILYHTARINCLAW SPDSSTVATGSLDTCVIIYEVDKPASNRLTIKGAHLGGVYGLAFTDDFSVVSSG EDACIRVWKINRQ	145	1968
382	WD40 repeat protein	MKVKVISRSTDEFTRERSQDLQRVFRNFDPNLRTQEKAVEYVRALNAAKLDKVF ARPFVGAMDGHVDSVSCMAKNPNYLKGIFSGSMDGDIRLWDIASRRTVCQFPGH QGPVRGLAASTDGQILVSCGIDSTVRLWNVPVATLGESDGTHENLAKPLAVYVW KNAFWAVDHQWDGELFATAGAQVDIWNQNRSQPISSFEWGTDTVISVRFNPGEP NVLATSGSDRSITLYDLRMSSPTRKVIMRTKTNAISWNPMEPMNFTAANEDCNC YSYDARKLEEAKCVHKDHVSAVMDIDYSPTGREFVTGSYDRTVRIFQYNGGHSR EVYHTKRMQRVFCVKFSCDASYVISGSDDTNLRLWKAKASEQLGVVLPRERRKH EYHEAVKSRYKHLPEVKRIVRHRHLPKPIYKAGILRRTVNEADRRKEERRKAHS APGSSSAEPLRKRRIIKEIE	130	1488

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
383	WD40 repeat protein	MVRSIKNPKKAKRKNKGSKNGDGSSSSSSIPSMPTKVWQPGVDKLEEGEELQCD PSAYNSLHAFHIGWPCLSFDIVRDTLGLVRTEFPHQVYFVAGTQAEKPTWNSIG IFKVSNITGKRRELVPSKPTDDADEESDSSDSDEDSDDEVGGSGTPILQLRKVG HEGCVNRIRAMNQNPHICASWGDSGHVQIWDFSSHLNALAESEADVSQGASSVF NQAPLVKFGGHKDEGYALDWSPLVPGRLVSGDCKNSIHLWEPTSGSTWNVDSTP FIGHAASVEDLQWSPTEENVFASCSVDGTIAIWDTRLGKTPAASFKAHDADVNV ISWNRLATCMLASGCDDGTFSIHDLRLLKEGDSVVAHFEYHKHPVTSIEWSPHE ASTLAVSSADCQLTIWDLSLEKDEEEEAEFKAKTKEQVNAPEDLPPQLLFVHQG QKDLKELHWHAQIPGMIVSTAADGFNILMPSNIQSTLPSDGA	269	1693
384	CDK type A	MERYKVIKELGDGTYGSVWKALNQQTHEIVAIKKMKRKYYIWEECINLREVKSL RKLNHPNIIKLKEVIRENNELFFIFEYMECNLYQIMKERSTPFSETAIIKFCYQ ILQGLSYMHRNGYFHRDLKPENLLVTSDLIKIADFGLAREVLTSPPYTDYVSTR WYRAPEVLLQSPTYTTAIDMWAVGAILAELFTLHPLFPGESELDEIYKICGVLG TPDYETWPDGMQLAAFRNFIFPQFLPVNLSVLIPHASPEAIDLITRLCSWDPQK RPTAEQALHHPFFRIGMSIPLSLGGHFQDNTCAAEVDTNFHSKKACKGRGMGEK ESSLECFLGLSLGLKPSLGHLGAMGSQGVGAVKQEVGSSPGCQSNPKQSLFQVL NSRAILPLFSSSPNLNVVPVKSSLPSAYTVNSQVMWPTIAGPPAAAVTVSTLQP SILGDFKIFGKSMGLASQYAGKEASPFS	1163	2545
385	CDK type A	MGEMGRGINNSSNNNNSNRPAWLQHYDLVGKIGEGTYGLVFLARSKLPNNRGLR IAIKKFKQSKDGDGVSPTAIREIMLLREFSHENVVKLVNVHINHVDMSLYLAFD YAEHDLYEIIRHHREKLNHHNINQYTVKSLLWQLLNGLNYLHSNWIVHRDLKPS NILVMGEGEEHGVVKIADFGLARIYQAPLKPLSDNGVVVTIWYRAPELLLGAKH YTSAVDMWAVGCIFAELITLKPLFQGVEVKASPNPFQLDQLDKIFKVLGHPTIE KWPTLMNLPHWSKNLQQIQQHKYDNAGLHIGPIPAKSPAYDLLSKMLEYDPRKR ITAAQALEHEYFRIDPQPGRNALVPSQPGEKAINYPPRLVDANTDFDGTIAPQP SQVSSGNAPSGSIASAAVPAVRPLPQQMQLMGMQRMQNPGMAAFNLGAQASMSG LNHNNIALQRGSSQQQAHQQVRRKEPNSGFPNTGYPPPPKSRRL	152	1582
386	CDK type B-1	MDKYEKLEKVGEGTYGKVYKARDKMTGQLVALKKTRLEMDEEGVPPSSLREISL LQMLSQSIYVVRLLCVEHVTKKGKPLLYLVFEYLDTDLKKFIDYRRSVNAGPLP QNVIQSFMYQLLKGVAHCHSHGVLHRDLKPQNLLVDKSKGLLKVGDLGLGRAFT VPLKCYTHEVVTLWYRAPEVLLGSTHYSTPVDIWSVGCIFAEMVRRQPLFPGDC EIQQLLHIFTLLGTPTEEMWPGVKRLRDWHEYPQWKPENLARAVPNLSPTGLDL ISKMLQCDPAKRISAKAAMNHPYFDDLDKSQF;	389	1297
387	CDK type B-1	MDGYEKMDKVGEGTYGKVYMARDKKTGQLVALKKTRLENDGEGIPPTALREISL LQMLSQDIYIVRLLDVKHTENKLGKPLLYLVFEYMESDLKKYIDSYRRSHTKMP PSMIKSFMYQLCRGVAYCHSRGVMHRDLKPHNLLVDKEKGVLKIADLGLSRAFT VPVKKYTHEIVTLWYRAPEVLLGATHYSLPVDIWSVGCIFAEMSRMQALFTGDS EVQQLMNIFRFLGTPNEEVWPGVTKLKDWHIYPEWKPQDISHAVPDLEPSGLDL LSQMLVYEPSKRISAKKALEHPYFDDLDKSQF	38	946

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
388	CDK type B-1	MDAYEKLEKVGEGTYGKVYKAKDKNTGQLVALKKTRLESDDEGIPPTALREISL LQMLSQDIHIVRLLDVEHTENKNGKPLLYLVFEYMDSDLKKYIDGYRRSHTKVP PNIIKSFMYQLCQGVAYCHSRGVMHRDLKPHNLLVDKQRGVVKIADLGLGRAFT IPIKKYTHEIVTLWYRAPEVLLGATHYSTPVDIWSVGCIFAEMVRLQALFIGDS EVQQLFKIFSFLGTPNEEIWPGVTKFRDWHIYPQWKPQDISSAVPDLEPSGVDL LSKMLVYEPSKRISAKKALEHPYFDDLDKSQF	180	1088
389	CDK type B-1	MDSYEKLEKVGEGTYGKVYKAKDKKTGKLVALKKTRLENDGEGIPPTALREISL LQMLSQDMNIVRLLDVEHTENKNGKPLLYLVFEYMDSDLKKYVDGYRRSHTKMP PKIIKSFMYQLCQGVAYCHSRGVMHRDLKPHNLLVDKQRGVLKIADLGLGRAFT VPIKKYTHEIVTLWYRAPEVLLGATHYSTPVDIWSVGCIFAEMSRMHALFCGDS EVQQLMSIFKFLGTPNEGVWPGVTKLKDWHIYPEWRPQDLSRAVPDLEPSGVDL LTKMLVYEPSKRISAKKALQHPYFDDLDKSQF	40	948
390	CDK type B-1	MEKYEKLEKVGEGTYGKVYKGRDKRTGRLVALKKTPFHQEEGIPPTAIREISLL KSLSQCIYIVKLLDVKASFNGKGKHVLFMVFEYADSDLKKHIDAHRQCNTKLSP RSIQSYMFQLCKGIAYCHSHGVLHRDLKPQNILVDQKIGLLKIADLGLGRACTV PIKSYTFEVVTLWYRAPEVLLGAKRYSMALDIWSLGCIFAELCNLQALFAGDSQ IQQLINIFRLLGTPNEQLWPGVTQLSDWHEFPQWRPQDLSKVVFNLDPNGVDLL SKMLQYDPAKRISAKEALDHPYFDSLDKSQF	229	1134
391	CDK type C	MGCVCGKPSARAADYVESPAEKGASSNSRSSSMASRRLVAPAVMDQGIDAENGH EGDYRTKLRGKQSNGADPVSLLSDDAEKQRHSRHHQHQQHHPIRPHHLRPQGEF VPNANSNPRFGNPPRHIEGEQVAAGWPAWLTAVAGEAIKGWIPRRADSFEKLDK IGQGTYSNVYKARDLDTGKIVALKKVRFDNLEPESVRFMAREIQVLRRLDHPNV VKLEGLVTSRMSCSLYLVFEYMDHDLAGLAACPGIKFTEPQVKCYMQQLLRGLD HCHSRGVLHRDIKGSNLLIDNGGILKIADFGLATFFHPDQRQPLTSRVVTLWYR PPELLLGATEYGVAVDLWSTGCILAELLAGKPIMPGRTEVEQLHKIFKLCGSPS EDYWKKSKLPHATIFKPQQPYKRCVAETFKDFPPSALALMEVLLAIEPADRGTA TSALKSDFFTTKPLACDPSSLPKYPPSKEFDAKIRDEEARRQRAAGGRGRDAAR RPSRESRAIPAPEANAELAISIQKRRLSSQGPSKSKSEKFNPQQEDGAVGFPIE PPRPMHIGIDAGATSRMYSQQFGPSHSGPLSNQISSSIWGKNQKEDEIQMAPGR PSRSSKATISDFRKPGACAPQPGADLSHLSSLVATARSNAGIDTHKDRSGMWQH NRIDAIDGVHNNGKHEFLEVPEHPNRQDWTRFQQPESFKGLDNYHLQDLPATHH RKDERVASKEATMNWQGYGGQGGDKIHYSGPLLPPSGNIDEILKEHERHIQHAV RRARQDKGRPQRSNLSQNERKAFEHRSFVSGVNGNAGYSDLVNELPISVGSNRL KVSKTRGTEEIVELRELEREPLSSVMEKYEREHEM	105	2642
392	CDK type C	MGCVCAKQSDILGEPESPKVKGSNLASSRWSVSSETKQLPQHSDSGILHHQHYY HPRDESDEAKLKESNYGGSKRRTRQGRDPADLDMGIFVRTPSSQSEAELVAAGW PAWMAAFAGEAIHGWIPRRAESFEKLYKIGQGTYSNVYKARDLDNGKIVALKKV RFDSLDAESVRFMAREILVLRKLDHPNIVKLEGLVTSEVSSSLYLVFEYMEHDL AGLAACPGIKFTEPQVKCYMQQLLQGLDHCHRHGVLHRDIKGSNLLIDNGGILK IADFGLATFFYPDQKQLLTSRVVTLWYRPPELLLGATDYGVAVDIWSAGCILAE LLAGKPILPGRTEVEQLHKIFKLCGSPSEDYWKESKLPHATIFKPQHPYKSCIA EAFKDFSPSALALLETLLAIEPGHRGEASGALKSEFFTTEPLSCDPSSLPKYPP SKEFDAKLRAQETRRQRDVGVRGHGSEAARRTSRLSRAGPTPNEGAELTALTQK	187	2580

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
		QHSTSHATSNIGSEKPSTKKEDYTAGLHIDPPRPVNHSYETTGVSRAYDAIRGV AYSGPLSQTHVSGSTSGKKPKRDHVKGLSGQSSLQPSKPFIVSDSRSERIYEKS HVTDLSNHSRLAVGRNRDTTDPHKSLSTLMQQIQDGTLDGIDIGTHEYARAPVS STKQKSAQLQRPSALKYVDNVQLQNTRVGSRQSDERPANKESDMVSHRQGQRIH CSGPLLHPSANIEDLLQKHEQQIQQAVRRAHHGKREALSNKSSLPGKKPVDHRA WVSSGKGNKESPYFKGKGNKELSDLKGGPTAKVTNFRQKVM		
393	CDK type C	MAVANPGQLNLQEAPSWGSRSVNCFEKLEQIGEGTYGQVYMAKEIETGEIVALK KIRMDNEREGFPITAIREIKLLKKLQHENVIKLKEIVTSPGPEKDEQGKSDGNK YNGSIYMVFEYMDHDLTGLAERPGMRFSVPQIKCYMKQLLIGLHYCHINQVLHR DIKGSNLLIDNNGILKLADFGLARSFCSDQNGNLTNRVITLWYRPPELLLGSTK YGPAVDMWSVGCIFAELLYGKPILPGKNEPEQLTKIFELCGSPDESNWPGVSKL PWYSNFKPQRQMKRRVRESFKNFDRHALDLVEKMLTLDPSQRISAKDALDAEYF WTDPVPCAPSSLPRYEPSHDFQTKRKRQQQRQHDEMTKRQKISQHPPQQHVRLP PIQNAGQGHLPLRPGPNPTMHNPPPQFPVGPSHYTGGPRGAGGQNRHPQNIRPL HAAQGGGYNANRGYGGPPQQQGGGYPPHGMGNQGPRGGQFGGRGAGYSQGGPYG GPVGGRGPNVGGGNRGPQFWSEQ	220	1749
394	CDK type D	MQNMEDNVQSSWSLHGNKEICARYEILERVGSGTYSDVYRGRRKADGLIVALKE VHDYQSSWREIEALQRLCGCPNVVRLYEWFWRENEDAVLVLEFLPSDLYSVIKS GKNKGENGIPEAEVKAWMIQILQGLADCHANWVIHRDLKPSNLLISADGILKLA DFGQARILEEPEAIYEVEYELPQEDIVADAPGERLMEEDDSVKGVRNEGEEDSS TAVETNFGDMAETANLDLSWKNEGDMVMQGFTSGVGTRWYRAPELLYGATIYGK EIDLWSLGCILGELLILEPLFSGTSDIDQLSRLVKVLGTPTEENWPGCSNLPDY RKLCFPGDGSPVGLKNHVPSCSDSVFSILERLVCYDPAARLNAKEVLENKYFVE DPYPVLTHELRVPSPLREENNFSEDWAKWKDMEADSDLENIDEFNVVHSSDGFC IKFS	438	1748
395	CDK type D	MDLNQYPEDLNPELPEGTDNVDNPDNNKGSPVPSPHPPLKPLDPSERYRKGITL GQGTYGIVYKAFDTVTNKTVAVKKIHLGKAKEGVNVTALREIKLLKELSHPNII QLIDAYPHKQNLHIVFEFMETDLEAVIKDRNLVFSPADIKSYLQMTLKGLAVCH KKWVLHRDMKPNNLLIAADGQLKLGDFGLARLFGSPDRKFTHQVFAVWYRAPEL LFGAKQYGPAVDIWATGCIFAELLLRKPFLQGVSDLDQIGKIFAAFGTPRQSQW PDVASLPDFVEFQFVPAPSLRSLFPMASEDALDLLSKMFTLDPKNRITAQQALE HRYFSSVPAPTRPDLLPKPSKVDSSRPPKHASPDGPVVLSPSKARRVMLFPNNL AGILPKQVSQSTTGGTPIEFDMPTQKLREVCPRSRITESGKKHLKRKTMDMSAA LDECAREQEGQEGKTILDPDHQRSAKKEKHM	240	1631
396	Cyclin A	MAGGQENCVRITRARAACVSKASAPVIQSQVDEKKSRKRAPKRAAVDDLAANAS GSQPKRRAVLGDVTNLHAAATDCLSTAEDQVDAPNPSIKGRARNKKKEARTSTK VVKDEIHPESNPLADHSSNLSECQKPPAAKLAEQRSLRGVPSKAKQGGSSNSQS CSKHTDIDKDHTDPQMCTTYVEDIYEYLRNAELKNRPSANFMETAQNDITPNMR AILVDWLVEVSEEYKLVPDTLYLTVSYIDRYLSANPTSRHKLQLLGVSCMLIAS KYEEVCPPHVEEFCYITDNTYTRDEMLSMERKILIFLNFEMTKPTTKSFLRRFV RASQAGNKAPSLHMEFLANYLAELTLMECSFLQYLPSLIAASTVFLSRLTLDFL TNPWNPTLAHYTGYKASQLKDCVMAIYNVQMNRKGSTLVAIREKYQQHKFKCVA SLPPPPFIAERFFEDTPN	252	1604

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
397	Cyclin A	MTGTQASNVRITRARAAKSTLNNALPPLPPAQGKPRGKRAATESNISGFSVAAE PLKRRAVLSDVSNICKEAAAVDCLKKPKAVKVVSQNANAKGRGRGIPRNNKKIT QEAEIKKETSPAICNVDDASAGNAIGDDKQNNNVNPLKEVQDNPKELNPIAEQI SVHPHCKQSVEKPNEKEIVVSDNKAAIASLKQQSTLQSLRIPKQPKYSLKQGNP VPLANLHEDVGRSSCSDFIDIDSEYKDPQMCTAYVTDIYANMRVVELKRRPLPN FMETTQRDINANMRSVLIDWLVEVSEEYKLVPDTLYLTVSYIDRFLSANVVNRQ RLQLLGVSCMLVASKYEEICAPPVEEFCYITDNTYKKEEVLEMEISVLNRLQYD LTTPTTKTFLRRFIRAAQASCKVSSLHLEFMGNYLAELTLVEYDFLKYLPSLIA AAAVFVARMTLDPMVHPWNSTLQHYTGYKVSDMRDCICAIHDLQLNRKGCTLAA IREKYNQPKFKCVANLFPPPIISPQFLIDNEV	261	1817
398	Cyclin B	MAAPNQNALLINNNNRRPLVDIGNLVGALNAQCNISKNGARKRAFGDIGNLVED LDAKCTISKYWVRKRPRTNFGVNANKGASSSTQGQGIVVRGEQKAWDRIVWGNK QSCAIKMNAQHVTATQRGTAISISDIIDSSVQDGGIKAPSQLKARKQTVRTVTA TLTARSEDSLRDVLEVPPGIDDGDRDNPLAVVEYVEDIYHFYRKIEVRSCVPPD YMTRQLEIKDSMRGVIIDWLIEVHRTFLLMPETLYLTVNIIDRYLSIQSVTRNE LQLMGITAMFIASKYEEISPPKINDLVYITKDAYTSKQIVNMEHTILNRLKFKL TVPTPYVFLVRFLKAAGPDKVMKNLAFFLVDLCLLHYKMIKYSPSMLAAAAVYT AQCTLKKHPYWNKTLILHIGYSEAHLRECAHLMADLHLKAEGSNLKSVYKKYSY PIFGSVAFLSPAKIPAGTVAAPAIDKCAHQIYLRNLR	167	1576
399	Cyclin B	MFPNKQTQGLVQNKKMASKAAQPKAMVPPQRVPPAANNRRALGDIGNIVADVGG KCNVTKDGVNGKPLAQVSRPITRSFGAQLLAQAAANKGISAANNQTQVPVVIPK ADVRGNKQRRTSKSKDIPPTTVVTNESDDCVIIEQAQRIKPTCNHNVGAVGNKE KPQLLTAKPKSLTASLTSRSAVALRGFRFDDEMTEAEEDPLPNIDVGDRDNQLA VVEYVEDIYKFYRRTEQMSCVPDYMPRQQEINPKMRAVLINWLIEVHYRFGLMP ETLYLTTNLIDRYLATQLVSRSNYQLVGATAMLLASKYEEIWAPEMNDFLDILE NKFERKHVLVMEKAMLNKLKFHLTVPTPYVFLVRFLKAAASDEEMENLVFFLME LSLMQYVMIKFPPSMLAAAAVYTAQITLKKTTVWNDVLKRHTGYSEIDLKECTR LMVAFHQSSEESKLNVVFKKYSMPEYDSVALIKPAKLPA	183	1598
400	Cyclin D	MAPSFDCVANAYIESCEDQEKLRQNAQILAQSGENDVDEPVSMLVQRETHYMLP EDYLQRLRNRTLDVNVRREAVGWILKVHSFYNFGAPTAYLAVNYLDRFLSRHRM PQGVKAWMIQLMAVACLSLAAKMEETQVPLPSDLQREDARFIFDARTIQRMELL ILSTLQWGMRSITPFSFIDYFAYRAVQGHGHGHDATPKAVMSRAIELILSTTEE IDFMEYRPSAIAAAALLCAAEEVVPLQAVHYKRALSSSITDVDKDKMFGCYNLI QETIIEGGCYWTPMSLQSTEKTPVGVLDAAACLSNTPTSSYSVKPYASVTAAKR RKLNEICSALLVSQAHPC	98	1126
401	Cyclin D	MAANFWTSSHCKELLDAEKVGIVHPLDKDQGLTQEDVKIIKINMSNCIRTLAQY VKLRQRVVATAITYCRRVYTRKSFTEYDPQLVAPTCLYLASKAEESTVQAKLVI FYMKKYSKHRYEIKDMLEMEMKLLEALDYYLVIYHPYRPLIQFLQDAGLNDLKV TAWALVNDTYRTDLILTYPPYMIALACIYFACIMEEKDAQAWFEELRVDMNEIK NISMEIVDYYDNYRVIPDEKMNSALNKLPHRF	148	894

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
402	Cyclin D	MAPALSSSYECLSHLLCAEDASNVVGCWDEDESKIFCEEEEGFGIQHFPDFPVP DDDEIRVLVRKESQYMPGKSYVQSYQNLGLDFTARQNAIGWILKVHGSYNFGPL TAYLSINYLDRFLSRNPLPKAKVWMLQLLSVACLSLAAKMEETQVPLLLDLQAE EPDFLFEPRTIQRMELLVLSTLEWRMLSVTPFSFVDYFLQGGGGRKPPPRAMVA RANELIFNTHTVLDFLEHRPSAIAAAAVICAAEEVLPLEAAQYKETILSCSLVD KEWVFGSYNLIQEVLIEKFSTPKKAKSASSSIPQSPVGVLDAFCLSNNSNNTSL EASLSVNLYASVAAKRRKLNDYCNTWRMFQHSTC	287	1363
403	Cyclin D	MAPNCIDCAPSDLFCAEDAFGVVEWGDAETGSLYGDEDQLHYNLDICDQHDEHL WDDGELVAFAEKETLYVPNPVEKNSAEAKARQDAVDWILKVHAHYGFGPVTAVL SINYLDRFLSANQLQQDKPWMTQLAAVACLSLAAKMDETEVPLLLDFQVEEAKY IFESRTIQRMELLVLSTLEWRMSPVTPLSYIDHASRMIGLENHHCWIFTMRCKE ILLNTLRDAKFLGLLPSVVAAAIMLHVIKETELVNPCEYENRLLSAMKVNKDMC ERCIGLLIAPESSSLGSFSLGLKRKSSTINIPVPGSPDGVLDATFSCSSSSCGS GQSTPGSYDSNNSSILCISPAVIKKRKLNYEFCSDLHCLED	251	1348
404	Cyclin-dependent kinase regulatory subunit	MPQIQYSEKYTDDTYEYRHVVLPPETAKLLPKNRLLNENEWRAIGVQQSRGWVH YAIHRPEPHIMLFRRPLNYQQNQQQQAGAQSQPMGLKAQ	229	510
405	Cyclin-dependent kinase regulatory subunit	MDQIEYSEKYYDDTYEYRHVELPPDVARLLPKNRLLTENEWRGIGVQQSRGWVH YAIHCSEPHIMLFRRPLNYEQNHQHPEPHIMLFRRPLNCQPNHQPQAHHPT	92	409
406	Cyclin-dependent kinase regulatory subunit	MDQIEYSEKYYDDTYEYRHVELPPDVARLLPKNRLLTENEWRGIGVQQSRGWVH YAIHCSEPHIMLFRRPLNYEQNHQHPEPHIMLFRRPLNCQPNHQPQAHHPT	64	381
407	Cyclin-dependent kinase regulatory subunit	MPQIQYSEKYYDDTYEYRHVVLPPDVARLLPKNRLLNENEWRGIGVQQSRGWVH YAIHRPEPHIMLFRRHLNYQQNQQQQAQQQPAQAMGLQA	68	349

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
408	Histone acetyltransferase	MALVETEPVTLIHPEEPKKFKKKPTPGRGGVISHGLTEEEARVKAIAEIVGAMV EGCRKGEDVDLNALKAAACRRYGLSRAPKLVEMIAALPDGERAAVLPKLKAKPV RTASGIAVVAVMSKPHRCPHIATTGNICVYCPGGPDSDFEYSTQSYTGYEPTSM RAIRARYNPYVQTRSRIDQLKRLGHTVDKVEFILMGGTFMSLPADYRDYFIRNL HDALSGHTSSNVEEAVCYSEHSATKCIGLTIETRPDYCLGPHLRQMLSYGCTRL EIGVQSTYEDVARDTNRGHTVAAVADCFCLAKDAGFKVVAHMMPDLPNVGVERD MESFREFFENPAFRADGLKIYPTLVIRGTGLYELWKTGRYRNYPPEQLVDIIAR VLALVPPWTRVYRVQRDIPMPLVTSGVEKGNLRELALARMDDLGLKCRDVRTRE AGIQDIHHKIRPEVVELVRRDYCANEGWETFLSYEDTRQDILVGLLRLRKCGHN TTCPELKGRCSIVRELHVYGTAVPVHGRDADKLQHQGYGTLLMEQAERIAWKEH RSIKIAVISGVGTRHYYRKLGYELEGPYMMKYLN	125	1849
409	Histone acetyltransferase	MLGFRDLYTSICEHLQRASGRLPIIAAATSLISTPEIAAVEKENKAPNSVDKMG MGSADESGRFSTSNGQFMNMNNGVVKEEWKGGVPVVPSAPTTVPVITNVKLETP SSPDHDMARKRKLGFLPLEVGTRVLCKWRDGKFHPVKIIERRKLPNGATNDYEY YVHYTEFNRRLDEWVKLEQLELDSVETDADEKVDDKAGSLKMTRHQKRKIDETH VEGNEELDAASLREHEEFTKVKNITKIELGRYEIETWYFSPFPSEYNNCEKLYF CEFCLNFMKRKEQLQRHMRKCDLKHPPGDEIYRSGTLSMFEVDGKKNKVYAQNL CYLAKLFLDHKTLYYDVDLFLFYILCECDERGCHMVGYFSKEKHSEESYNLACI LTLPPYQRKGYGKFLISFSYELSKKEGKVGTPERPLSDLGLLSYRGYWTRVLLD ILKKHKSNISIKELSDMTAIKADDVLSTLQGLDLIQYRKGQHAICADPKVLDRH LKAVGRGGLEVDVCKLIWTPYKEQ	70	1602
410	Histone acetyltransferase	MGSLDESTCSEEIRDEGKDSIRTKFKVESTVNNAQNGGNDNSKKKRAAGLPLEV GIRLLCKWRDSKLHPVKIIERRKLPNGFPQDYEYYVHYTEFNRRLDEWVKLEQF ELDSVETDADEKIEDKGGSLKMTRHQKRKIDEIHVEEGQGHEDFDPASLREHEE FTKVKNIAKVELGRYEIETWYFSPFPPEYSHCEKLFFCEFCLNFMKRKEQLQRH MRKCDLKHPPGDEIYRNGTLSMFEVDGKKNKIYGQNLCYLAKLFLDHKTLYYDV DLFLFYVLCECDDRGCHVVGYFSKEKHSDEAYNLACILTLPPYQRKGYGKFLIA FSYELSKKEGKVGTPERPLSDLGLLSYRGYWTRILLDILKKQRGNISIKELSDM TAIKVEDVISTLQVLDLIQYRKGQHVICADPKVLDRHLKAAGIAGLEVDVSKLI WTPYKEQCG	140	1465
411	Histone acetyltransferase	MASAPMVGCDDSRDKHRWVESKVYMRKGHGKGSKGNAGFNAQNSTAQVRRENDN MGNSIADNGKSEAASEGLSSLSRKQITVNQDHPPNETSSMPAVGGLQNIDTHVT FKLEGCSKQEIWELRKKLTNELEQVRGTFKKLEARELQLRGYSVSAGVNTSYSA SQFSGNDMRNNGGKEVTSEVASGGAITPKQAQRESNPPRQLSISLMENNQAASD MGEKGKRTPKANQYYRNSEFVLGKDKFPPAESKKSKSTGNKKISQSKVFSKETM QVGKEFMPQKSVNEVFKQCSLLLTKLMKHKYGWVFNLPVDAQALGLHDYHTIIK RPMDLGTVKSKLEKNLYNSPASFAEDVKLTFSNAMTYNPKGHEVHTMAEQLLQL FEERWKTIYEEHLDGKMRFGSGQGLGASSSTKKLPFQDSKKNIKKSEPAGGPSP PKPKSTNHHASRTPSAKKPKAKDPHKRDMTYEEKQKLSTNLQNLPQERLELIVQ IIKKRNPSLCQHDEEIEVDIDSFDTETLWELDRFVTNYKKSLSKNKKKALLADQ AKRASEHGSARNKHPMIGRELPMNNKKGEQGEKVVEIDHMPPVNPPVVEVEKDG VYAKRSSSSSSSSSDSGSSSSDSDSGSSSGSESDAYAATSPPAGSNTSARG	628	2565

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
412	Histone acetyltransferase	MEGHSGALGFGQGFSRSSQSPNLSPSPSHSASASVTSSGQKRKRNEVEHAGVAS NSTGMFAVPPSHIYSHLHPMSMSMPMPMHNSHPSSLSESRDGALTSNDDDDNLT GGNQSQLDSMSAGNTDGREDFDDEDDDDDDEEDDDEVEGDEEDQDHDPDADDDS DDGHDSMRTFTAARLDNGAPNSRNLKPKADAAGVAIAPTVKTEPILDTVKEEKV SGNNNNNSVSANNAQVAPSGSAVLLSAVKEEANKPTSTDHIQTSGAYCAREESL KREEDADRLKFVCFGNDGIDQHMIWLIGLKNIFARQLPNMPKEYIVRLVMDRSH KSVMIIKQNQVVGGITYRPYLSQKFGEIAFCAITADEQVKGYGTRLMNHLKQHA RDVDGLTHFLTYADNNAVGYFIKQDFTKEIKLEKERWHGYIKDYDGGILMECKI DPKLPYTDLPAMIRWQRQTIDEKIRELSNCHIVYSGIDIQKKEAGIPRKPIKVE DIPGLKEAGWTTDQWGHSRFRLLNSPSEGLPNRQVLHAFMRSLHKAMVEHADAW PFKEPVDPRDVPDYYDIIKDPMDVKRMFTNARTYNTHETIYYKCANR	55	1818
413	Histone deacetylase	MEESGNSLTSGPDGSKRRVSYFYDSDIGNYYYSQGHPMKPHRIRMAHSLIVHYA LDEKMEVCRPNLLQSRELRVFHADDYISFLQSVTPETQHEQLRQLKRFNVGEDC PVFDGLYNFCQTYAGGSVGAAIKLNNKEADIAINWSGGLHHAKKCEASGFCYVN DIVLAILELLKVHQRVLYIDIDIHHGDGVEEAFYSTDRVMSVSFHKFGDYFPGT GHLKDVGYGKGKYYSLNVPLNDGIDDESYKNLFRPIIQKVMEIYQPEAVVLQCG ADSLSGDRLGCFNLSVKGHADCVRFLRSFNVPLVLVGGGGYTIRNVARCWCYET AVAVGVEPQDKLPYNEYYEYFGPDYTLHVAPSNMENQNSAKELAKIRNTLLEQL KRIQHVPSVPFQERPPDTKFPEEDEEDYEKRPKGHKWGGEYFGSESDEEQKPQN RDIDISDKPGIRRQSPPNVEAAKKIKVEEEDGDIGIVNENDGAKWPLGEAG	259	1710
414	Histone deacetylase	MEESGNSLTSGPDGSKRRVSYFYDSDIGNYYYSQGHPMKPHRIRMAHSLIVHYA LDEKMEVCRPNLLQSRELRVFHADDYISFLQSVTPETQHEQLRQLKRFNVGEDC PVFDGLYNFCQTYAGGSVGAAIKLNNKEADIAINWSGGLHHAKKCEASGFCYVN DIVLAILELLKVHQRVLYIDIDIHHGDGVEEAFYSTDRVMSVSFHKFGDYFPGT GHLKDVGYGKGKYYSLNVPLNDGIDDESYKNLFRPIIQKVMEIYQPEAVVLQCG ADSLSGDRLGCFNLSVKGHADCVRFLRSFNVPLVLVGGGGYTIRNVARCWCYET AVAVGVEPQDKLPYNEYYEYFGPDYTLHVAPSNMENQNSAKELAKIRNTLLEQL KRIQHVPSVPFQERPPDTKFPEEDEEDYEKRPKGHKWGGEYFGSESDEEQKPQN RDIDISDKPGIRRQSPPNVEAAKKIKVEEEDGDIGIVNENDGAKWPLGEAG	356	1807
415	Histone deacetylase	MEFWGVEVKPGEALTCDPGDERYLHMSQAAIGDKEGAKENERVSLYVHVDGKKF VLGTLSRGKCDQIGLDLVFEKEFKLSHTSQTGSVFVSGYTTVDHEALDGFPDDE DLESSEDEEEELAQITTLTAKENGGKTGAKPVKPESKSSVTDKAAAKGKPEVKP PVKKQEDDSDSDEDEDEDEDEDEDDDDEDDEDMKDASASDDGDEEDDSDEESDD DEEEDEETPKPAAGKKRPMPASDNKSPATDKKAKITTPAGGQKPGADKGKKTEH IATPYPKHGAKGPASGVKGKETPLGSKQTPGSKVKNSSTPESGKKSGQFKCQSC SRDFATEGALSSHNAAKHGGK	261	1298

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
416	Histone deacetylase	MMETGGNSLPSGPDGVKRKVAYFYDPEVGNYYYGQGHPMKPHRIRMTHALLVQY GLHKEMQILKPYPARDRDLCRFHADDYVAFLRGITPETIQDQVKALKRFNVGDD CPVFDGLYQYCQTYAGGSVGGAVKLNHKLCDIAINWAGGLHHAKKCEASGFCYV NDIVLAILELLKYHKRVLYVDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPG TGDIRDIGCGKGKYYAVNVPLDDGIDDESFQSLFKPIIQQVMLVYNPEAIVLQC GADSLSGDRLGCFNLSVKGHAECVRYMRSFNVPLLMVGGGGYTVRNVARCWCYE TGVAVGVEIDDKMPQHEYYEYFGPDYTVHVAPSNMENKNTKQYLDKIRSKILEN INSLPCAPSAQFQVQPPDTDFPELEEEDYDERTRSHKWDGASCDSDSENGDLKH RNHDVEESAFPRHNLANISYNTKIKLEGVGTGGLDMAAGTDTKKNDESFEAMDY ESGEELRQDHFASTINASQPCDPALLTGVQNQLQSTDTVKPIEQSGNAPGIPPP SVATVSTGTRPSSISRTSSLNSMSSVKQGSILGPNPPQGLNASGLQFPVPTSNS PIRQGGSYSITVQAPDKQGLQNHMKGPQNMPGNS	365	2251
417	Histone deacetylase	MPPKDRVAYFYDGDVGSVYFGPNHPMKPHRLCMTHHLVLSYELHKKMEIYRPHK AYPVELAQFHSADYVEFLHRITPDTQHLFTKELVKYNMGEDCPVFENLFEFCQI YAGGTIDAAHRLNNQICDIAINWSGGLHHAKKCEASGFCYINDLVLGILELLKH HARVLYVDIDVHHGDGVEEAFYFTDRVMTVSFHKYGDMFFPGTGDVKEVGEREG KYYAINVPLKDGIDDASFTRLFKTIITKVVDIYQPGAIVLQCGADSLAGDRLGC FNLSIDGHAQCVRIVKKFNLPLLVTGGGGYTKENVARCWSVETGVLLDTELPNE IPDNDYIKYFAPDYSLKINTAGNMENLNSKTYLSAIKVQVMENLRAIQHAPSVQ MHEVPPDFYIPDIDEDELNPDERMDQHTQDRQIQRDDEYYDGDNDIDHDMEEAS	156	1454
418	Histone deacetylase	MDSSKSEEANILHVFWHEGMLNHDLGTGVFDTLEDPGFLEVLEKHPENADRVRN MLSILRKGPIAPYTEWHTGRAAYLSELYSFHRPDYVDMLAKTSTAGGKTLCHGT RLNPGSWEAALLAAGTTLEAMRYILDGHGKLSYALVRPPGHHAQPTQADGYCFL NNAGLAVELAVASGCKRVAVVDIDVHYGNGTAEGFYERDDVLTISLHMNHGSWG PSHPQTGFHDEVGRGKGLGFNLNVPLPNGTGDKGYEHAMHELVVPAISKFMPEM IVLVIGQDSSAFDPNGRECLTMEGYRKIGQIMRQQADQFSGGRLVVVQEGGYHI TYAAYCLHATLEGVLCLPHPLLSDPIAYYPEHDIYSERVTFIKNYWQGIISTTD KRN	203	1348
419	Histone deacetylase	MEESGNALVSGPDGSKRRVTYFYDADIGNYYYGQGHPMKPHRMRMAHNLIVHYG LHQRMEVCRPHLAQSKDIRAFHTDDYIHFLSSVAPDTQQEQLRQLKRFNVGEDC PVFDGLFNFCQSSAGGSIGAALKLNRKDADIAINWAGGLHHAKKCEASGFCYVN DIVLGILELLKVHQRVLYIDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGT GHIKDVGYGKGKYYALNVPLNDGIDDESYKHLFRPIIQKVMEVYQPEAVVLQCG ADSLSGDRLGCFNLSVKGHADCVRFVRSFNIPLMLVGGGGYTIRNVARCWCYET AVAVGVEPQDKLPYNEYYEYFGPDYTLYVAPSNMENLNTEKDLEKMRNVLLEQL SKIQHTPSVPFQERPPDTEFNDEEEEDMEKRSKCRIWDGEYVGSEPEEDGKLPR FDADTYERSVLKHENKRLVPVSNVEPLKRIKQEEDGAAV	229	1644

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
420	Histone deacetylase	MPPKDRVAYFYDGDVGSVYFGPNHPMKPHRLCMTHHLVLSYELHKKMEIYRPHK AYPVELAQFHSADYVEFLHRITPDTQHLFTKELVKYNMGEDCPVFENLFEFCQI YAGGTIDAAHRLNNQICDIAINWSGGLHHAKKCEASGFCYINDLVLGILELLKH HARVLYVDIDVHHGDGVEEAFYFTDRVMTVSFHKYGDMFFPGTGDVKEVGEREG KYYAINVPLKDGIDDASFTRLFKTIITKVVDIYQPGAIVLQCGADSLAGDRLGC FNLSIDGHAQCVRIVKKFNLPLLVTGGGGYTKENVARCWSVETGVLLDTELPNE IPDNDYIKYFAPDYSLKINTAGNMENLNSKTYLSAIKVQVMENLRAIQHAPSVQ MHEVPPDFYIPDIDEDELNPDERMDQHTQDRQIQRDDEYYDGDNDIDHDMEEAS	156	1454
421	Histone deacetylase	MDLNLVSHGEEEEGVRRRKVGIVYDERMCKHATPEDQPHPEQPDRIRVIWDKLN SAGVLHKCVMVEAKEASEEQLAGVHSRKHIEVMKSIGTARYNKKKRDKLAASYS SIYFSQGSSEAALLAAGSVVEISEKVASGELDAGVAIVRPPGHHAEADKAMGFC LFNNIAIAAKHLVHERPELGVQEVLIVDWDVHHGNGTQHMFWTDPHVLYFSVHR FDAGTFYPGGDDGFYDKIGEGKGAGYNINVPWEQGKCGDADYLAVWDHVLVPVA KSYDPDMVLISGGFDAALGDPLGGCRLTPYGYSLMTKKLMEFAGGKIVLALEGG YNLKSLADSFLACVEALLKDGPSRSSVLTHPFGSTWRVIQAVRKELSSFWPALN EELQLPRLLKDASESFDKLSSSSSDESSASEDEKKIAEVTSIMEVSPDPSSILA LTAEDIAQPLAGLKIEEAGTDSQRSSDHTLLDLTNDDTQKLKQFEGEIFVMIGD EESVPSASSSKDQNESTVVLSKSNIKAHSWRLTFSSIYVWYASYGSNMWNPRFL CYIEGGQVEGMAKRCCGSEDKTPPQRIQWKVVPHRMFFGRSYTNTWGSGGVSFL DPNCSDTSEAHVCLYKITLAQFNDLLLQENNLNCGTEHPLVDLSSIDAIRNGNS ILELIKDSWYGTLIYLGMEGGLPIVTFTCSVCDVEKFKHGQLPLCPPSSRYENI LIRGLVQGKKLSEDDATAYIRAASTSPLL	27	2222
422	Peptidylprolyl isomerase	MADEDLDLSDVGEVEDEPGEEIESTPPLAVGQEKEINSLALKKKLLKVGTRWET PENGDEVTVHYTGTLPDGTKFDSSRDRGEPFTFKLGQGQVIKGWDQGIVTMKKG ERALFTIPPELAYGSSGVRPTIPPNATLQFDVELLSWTNIVDVCNDGGILKRII SEGEKYERPKDPDEVTVKYEAKLEDGTLVAKSPEEGVEFYVNDGHFCPAIAKAV KTMKRGESVILTIKPTYAFGERGKDAEEGFAAIPPNATLTTSLELVSFKAVIAV TEDKKVIKKILKEADGYDKPSDGTVVQIRYTAKLQDGTIFEKKGYEGEEPFQFV VDEEQVIAGLDKAVETMKTGEIALITIGAEYGFGNFETQRDLAVIPPNSTLIYE VEMISFTKEKESWDMDTTEKIEASKQKKEQGNSLFKVGKYQRAAKKYEKAAKYI EHDSSFSAEEKKQSKVLKVSCNLNHAACRLKLKDFKEAVKLCSKVLELESQNVK ALYRRAQAYIETADLDLAEFDIKKALEIEPQNREVQLEYKILKQKQIEYNKKDA KLYGNMFAKLNKLEAFEGKVLS	71	1759

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
423	Peptidylprolyl isomerase	MADEGLELSDVAEVEDEPGEEFESAPPLVVGQEKELNSSGLKKKLLKAGTRCET PENGDEVTVHYTGTLLDGTKFDSSRDRGEPFTFNIGQGQVIKGWDQGIVTMKKR EHALFTIPPELAYGASGMPPTIPPNATLQFDVELLSWTNIVDVCKDGGILKRII SDGEKYERPKDPDEVTVKYEAKLEDGMLVAKSPEEGVEFYVNDGNFCPAIVKAV KTMKKGENVTLTIKPAYAFGEQGKDAEEGFAAIPPNATITINLQLVSFKAVKEV TEDKKVIKKILKEADGYDKPSDGTVVQIRYTAKLQDGTIFEKKGYAGEEPFQFV VDEEQVIAGLDKAVETMKTGEVALITIGPEYGFGNIETQRDLAVIPPYSTLIYE VEMVSFTKEKESWDMNTTENIEASKQKKEQGNSLFKVGKYLRAAKKYDKAAKYI EHDNSFSAEEKKQSKVLKVSCNLNHAACCLKLKDFKKAVKLCSKVLELESQNVK ALYRRAQAYIETADLDLAEFDIKKALEIEPQNREVRLEYLILKQKQIEYNKKDA KLYGNMFARQNKLEAIEGKD	358	2040
424	Peptidylprolyl isomerase	MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRALCTGEKGTGRSGKPLH FKGSSFHRVIPGFMCQGGDFTRGNGTGGESIYGEKFADENFVKKHTGPGILSMA NAGPNTNGSQFFICTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGSGRTSKP VVIADSGQLA	238	756
425	Peptidylprolyl isomerase	MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRALCTGEKGNGRSGKPLH FKGSSFHRVIPGFMCQGGDFTRGNGTGGESIYGEKFADENFVKKHTGPGILSMA NAGPNTNGSQFFICTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGSGRTSKP VVIADSGQLA	285	803
426	Peptidylprolyl isomerase	MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRALCTGEKGTGRSGKPLH FKGSSFHRVIPGFMCQGGDFTRGNGTGGESIYGEKFADENFVKKHTGPGILSMA NAGPNTNGSQFFICTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGSGRTSKP VVIADSGQLA	190	708
427	Peptidylprolyl isomerase	MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRALCTGEKGTGRSGKPLH FKGSSFHRVIPGFMCQGGDFTRGNGTGGESIYGEKFADENFVKKHTGPGILSMA NAGPNTNGSQFFICTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGSGRTSKP VVIADSGQLA	156	674
428	Peptidylprolyl isomerase	MADDFELPESAGMMENEDFGDTVFKVGEEKEIGKQGLKKLLVKEGGSWETPETG DEVEVHYTGTLLDGTKFDSSRDRGTPFKFKLGQGQVIKGWDQGIATMKKGENAV FTIPPDLAYGESGSQPTIPPNATLKFDVELLSWASVKDICKDGGIFKKIIKEGE KWEHPKEADEVLVKYEARLEDGTVVSKSEEGVEFYVKDGYFCPAFAIAVKTMKK GEKVLLTVKPQYGFGHQGREAIGNDVARSTNATLLVDLELVSWKVVDEVTDDKK VLKKILKQGEGYERPNDGAVVKVKYTGKLEDGTIFEEKGSDEEPFEFMAGEEQV VDGLDRAVMTMKKGEVALVSVAAEYGYQTEIKTDLAVVPPKSTLIYEVELVSFV KEKESWDMNTAEKIEAAGKKKEEGNALFKVGKYFRASKKYEKATKYIEYDTSFS EEEKKQSKPLKVTCNLNNAACKLKLKDYTQAEKLCTKVLEVESQNVKALYRRAQ AYIQTADLELAELDIKKALEIDPNNRDVKLEYRALKEKQKEYNKKEAKFYGNMF ARMSKLEELESRKSGSQKVETANKEEGSDAMAVDGESA	176	1912
429	Peptidylprolyl isomerase	MAASLTPLGAGLAYATIYDQAKVRKLEPTKRSLIALCQHSDSQHRRFITRKYHV NVQILNRRDAIRLIGLAAGLCIDLSLMYDARGAGLPPQENAKLCDTTCEKELEN APMITTESGLQYKDIKIGNGPSPPIGFQVAANYVAMVPSGQVFDSSLDKGQPYI FRVGSGQVIKGLDEGLLSMKVGGKRRLYIPGPLAFPKGLNSAPGRPRVAPSSPV IFDVSLEFIPGLESEEE	64	765

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
430	Peptidylprolyl isomerase	MSAASLSADMAIRGTILGKTALHVLGPQVVSQCRQPVMFKCPPHTLRKMRFSAQ DLQSKNFYSGFTPFKSVFISTSKRSWQAGSARAMSQDAAFQSKVTTKCFLDIEI GGDPAGRIVLGLFGEDVPKTAENFRALCTGEKGFGYKGSSFHRIIKDFMLQGGD FDRGDGTGGKSIYGRTFEDENFKLAHVGPGVLSMANAGPNTNGSQFFICTVKTP WLDKRHVVFGQVIEGMEIVKKLESEETNRTDRPKRPCRIVDCGELP	93	881
431	Peptidylprolyl isomerase	MGRIKPQTLLQQSKKKKVPGRISVSTIIVCNLIIIFLMFSLVGIYRQRAKRNRA TSRSDGDEEMENFGRSKINSVPHQAIVNTTKGLITLELFGKSSAHTVEKFVEWS ERGYFNGLPFYRVIKHFVIQVGDPKFAGNREDWTVGGQLNVQLEFSPKHEAFML GTSKLEDQGDGFELFITTAPIPDLNDKLNVFGRVIKGQDVVQEIEEVDTDEHFQ PKSPIIINDVRLKDEL	372	1070
432	Peptidylprolyl isomerase	MARQSTLLLFWSLVFLGAIVFTQAKHEELEEVTHKVYFDVDIAGKPAGRVVIGL FGKAVPKTVENFRALCTGEKGVGKSGKPLHYKGSFFHRIIPSFMIQGGDFTLGD GRGGESIYGTKFADENFKLKHTGPVFITTVTTDWLDGRHVVFGKIISGMDVVYK VEAEGRQSGQPKRKVKIADSGELSMD	28	594
433	Peptidylprolyl isomerase	MARQSTLLLFWSLVFLGAIVFTQAKHEELEEVTHKVYFDVDIAGKPAGRVVIGL FGKAVPKTVENFRALCTGEKGVGKSGKPLHYKGSFFHRIIPSFMIQGGDFTLGD GRGGESIYGTKFADENFKLKHTGPGFLSMANAGPDTNGSQFFITTVTTDWLDGR HVVFGKIISGMDVVYKVEAEGRQSGQPKRKVKIADSGELSMD	34	648
434	Peptidylprolyl isomerase	MEMDEIQEQSQPQSSEKQDISQESDTGNDKTINAEKITSENAEVEEDDMLPPKV NTEVEVLHDKVTKQIIKEGSGNKPSRNSTCFLHYRAWAESTMHKFQDTWQEQQP LELVLGREKKELSGFAIGVAGMKAGERALLHVDWQLGYGEEGNFSFPNVPPRAN LIYEAELIGFEEAKEGKARSDMTVEERIEAADRRRQQGNELFKEDKLAEAMQQY EMALAYMGDDFMFQLFGKYKDMANAVKNPCHLNMAQCLLKLNRYEEAIGQCNMV LAEDEKNIKALFRRGKARATLGQTDDAREDFQKVRKFSPEDKAVIRELRLLAEH DKQVYQKQKEMFKGLFGQKPEQKPKKLHWFVVFWQWLLSMIRTIFRMRSKTD	481	1611
435	Peptidylprolyl isomerase	MAGAGEGTPEVTLETSMGPITVELYHKHAPKTCRNFLELSRRGYYNNVKFHRVI KDFMVQGGDPTGTGRGGESIYGPRFEDEITRDLKHTGAGILSMANAGPNTNGSQ FFISLAPTPWLDEKHTIFGRVCKGMDVVKRLGNVQTDKNDRPIHDVKILRTTVK D	93	584
436	Peptidylprolyl isomerase	MMDPELMRLAQEQMSKISPDELMKMQRQIMANPDLMRMASENMKNLKPEDIRFA AEQMKNVRKEEMAEISERISRASPEEIEAMKARANLQSAYQLQVAQNLKDQGNQ LHARMKYSEAAEKYLQARNNLTGIPFSEAKSLLLASSSNLMSCYLKTGQYEECV QTGSEVLAYDAMNVKALYRRGQAYKQIGKLELAVADLRKAVEVSPEDETIAQAL REASTELMEKGGTQDQNGPRIEEIIEEEAVQPTAEKYPQSAPMVTSVTEDVSDD EQGSEDQNGFSRDSFQATNAPDGQMYAESLRNLTENPDMLRTMQSLMKNVDPDS LVALSGGKLSPDMVKTVSGMFGRMSPEEIQNMMKMSSTLSRQNPSTSSRFDDIT RGHSNMDSSPQSVSVDNDLFEENQNRVGESSTNLSSSAAFSGMPNFSAEMQEQV RNQMNDPATRQMFTSMIQNMSPEMMASMSEQFGVKLSPEDAVKAQNAMASLSPN DLDRLMNWATRLQTAIDYARKIKNWILGRPGLIFAISMLLLAIILHRFGYIGD	250	1869

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
437	Peptidylprolyl isomerase	MGVEKEILRPGNGPKPRPGQSVTVHCTGYGKNEDLSQKFWSTKDPGQKPFTFTI GQGRVIKGWDEGVLDMQLGEIFKLRCSPDYGYGSNGFPAWGIRPNSVLVFEIEV LSVN	84	422
438	Peptidylprolyl isomerase	MPNPRCYLDITIGEELEGRILVELYSDVVPKTAENFRALCTGEKGIGPHTGVPL HYKGLPFHRVIKGFMIQGGDISAQNGTGGESIYGLKFDDENFQLKHERRGMLSM ANSGPNTNGSQFFITTTRTSHLDGKHVVFGKVIKGMGVVRGIEHTPTESNDRPS LDVVISDCGEIPEGSDDGIANFFKDGDLYPDWPADLLDEKSAEISWWMNAVDSAK CFGNENYKKGDYKMALRKYRKALRYLDICWEKEEIDEEKSNHLRKTKSQIFTNS SACKLKLGDLKGALLDTEFAMRDGEDNVKALFRQGQAYMALKDVDSAVASFKKA LQLEPNDAGIRKELAVATKMINDRRDQERRAYARMFQ	128	1213
439	Peptidylprolyl isomerase	MGDVIDLNGDGGVLKTIIRSAKPGAMQPTEDLPNVDVHYEGTLADTGEVFDTTR EDNTLFSFELGKGTVIKAWDIAVKTMKVGEVARITCKPEYAYGSAGSPPDIPEN ATLIFEVELVACKPRKGSTFGSVSDEKARLEELKKQREIAAASKEEEKKRREEA KATAAARVQAKLEAKKGQGRGKGKSKGK	265	837
440	Peptidylprolyl isomerase	MGLGLKIASASFLPIFNIMATRSLCILLVCFIPVLAHVLSLQDPELGTVRVYFQ TTYGDIEFGFFPHVAPKTVEHIYKLVRLGCYNSNHFFRVDKGFVAQVADVVGGR EVPLNSEQRKEGEKTIVGEFSEVKHVRGILSMGRYSDPDSASSSFSILLGNAPH LDGQYAVFGKVTKGDDTLKRLEEVPTRQEGIFVMPLERIRILSTYYYDTNERES NLTCDHEVSILKRRLVESAYEIEYQRRKCLP	38	781
441	Peptidylprolyl isomerase	MASKRSLRTMNVWPTLPPLVLLLLLCFSSMSSSVVAKKSDVSELQIGVKHKPKS CDIQAHKGDRIKVHYRGSLTDGTVFDSSFERGDPIEFELGSGQVIKGWDQGLLG MCVGEKRKLRIPSKLGYGAQGSPPKIPGGATLIFDTELVAVNGKGISNDGDSDL	38	526
442	Peptidylprolyl isomerase	MSGAPAERPISYFDITIGGKPIGRIVFSLYADLVPKTAENFRALCTGEKGIGKS GKPLCYAGSGFHRVIKGFMCQGGDFTAGNGTGGESIYGEKFEDEAFPVKHTKPF LLSMANAGKDTNGSQFFITVSQTPHLDDKHVVFGEVIKGKSIVRAIENYPTASG DVPTSPIIISACGVLSPDDPSLAASEETIGDSYEDYPEDDDSDVQNPEVALDIA RKIRELGNKLFKEGQIELALKKYLKSIRYLDVHPVLPDDSPPELKDSYDALLAP LLLNSALAALRTQPADAQTAVKNATRALERLELSDADKAKALYRRASAHVILKQ EDEAEEDLVAASQLSPEDMAISSKLKEVKDEKKKKREKEKKAFKKMFSS	37	1158
443	Peptidylprolyl isomerase	MASSLRSSLFSSWALDSKSVCSLFNLNPGKMGLPSISTPLNWRTCCCSHSSELL ELNEGLQSSRRKTVMGLSTVIALSLVYCDEVGAVSTSKRALRSQKVPEDEYTTL PNGLKYYDLKVGSGTEAVKGSRVAVHYVAKWKGITFMTSRQGMGITGGTPYGFD VGASERGAVLKGLDLGVQGMRVGGQRILIVPPELAYGNTGIQEIPPNATLEFDV ELISIKQSPFGSSVKIVEG	61	768

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
444	WD40 repeat protein	MGAIEDEEPPLKRLKVSSPGLRRGLEEEAPSLSVGSVSILMAKSLSLEEGETVG SKGLIRRVEFVRIITQALYSLGYQKAGALLEEESGILLQSSNVALFRKQILDGK WDESVVTLRGIDQVEVEGNTLKAASFLILQQKFFELLDKGNIPEAMKTLRLEIS PMQLNTKRVHELASCIVFPSRCEELGYSKQGNPKSSQRMKVLQEIQQLLPPSIM IPEKRLERLVEQALNVQREACIFHNSLDPALSLYTDHQCGRDQIPTTTLQVLES HKNEVWFLQFSNNGKYLASASKDCSAIIWEITEGDSFSMKHRLSAHQKPVSFVA WSPDDKLLLTCGIEEVVKLWNVETGECKLTYDKANSGFTSCGWFPDGERFISGG VDKCIYIWDLEGKELDSWKGQGMPKISDLAVTSDGKEIISICGDNAIVMYNLDT KTERLIEEESGITSLCVSKDSRFLLLNLANQEIHLWDIGARSKLLLKYKGHRQG RYVIRSCFGGSDLAFVVSGSEDSQVYIWHRGNGELLAVLPGHSGTVNCVSWNPV NPHVFASASDDYTIRIWGVNRNTFRSKNASSSNGVVHLANGGP	421	2172
445	WD40 repeat protein	MPGTTAGAGIEPIEPQSLKKLSLKSLKRSFDLFASLHGEPQPPDQRSQRIRIAC KVRAEYEVVKNLPTLPQREVGSSVSNSNVGETHSSLTTNQAQGFPTDTSGDLSK DEGKEITSIAVHLQPQTGLIDGKAGAIAGTSTAISSVGSSDRYQPSAAIMKRLP SKWPRPIWHPPWKNYRVISGHLGWVRSVAFDPGNEWFCTGSADRTIKIWEVATG KLKLTLTGHIEQIRGLAVSSRHPYLFSAGDDKQVKCWDLEYNKAIRSYHGHLSG VYCLALHPTLDILCTGGRDSVCRVWDIRTKAQIFALSGHENTVCSVFTQAIDPQ VVTGSHDTTIKLWDLAAGKTMSTLTYHKKSVRAIAKHPFEHTFASASADNIKKF KLPKGEFLHNMLSQQKTIVNAMAINEDNVLVSAGDNGSLWFWDWKSGHNFQQAQ TIVQPGSLDSEAGIYALQYDITGSRLVSCEADKTIKMWKEDETATPESHPINFK APKDIRRF	163	1647
446	WD40 repeat protein	MRPILMKGHERPLTFLKYNRDGDLLFSCAKDHTPTVWYGHNGERLGTYRGHNGA VWCCDVSRDSTRLITSSADQTAKLWNVETGAQLFSFNFESPARAVDLAIGDKLV VITTDPFMELPSAIHIKRIEKDLSKQTADSVLTITGIKGRINRAVWGPLNSTII SGGEDSVVRIWDSETGKLLRESDKETGHQKPITSLCKSADGSHFLTGSLDKSAR LWDIRTLTLIKTYVTERPVNAVAISPLLDHVVIGGGQEASHVTTTDRRAGKFEA KFFHKILEEEIGGVKGHFGPINSLAFNPDGRSFASGGEDGYVRLHHFDPDYFHI KM	192	1172
447	WD40 repeat protein	MRPILMKGHERPLTFLKYNRDGDLLFSCAKDHTPTVWYGHNGERLGTYRGHNGA VWCCDVSRDSTRLITSSADQTAKLWNVETGNQLFSFNFESPARAVDLAIGDKLV VITTDPFMELPSAIHIKRIEKDLSKQTADSVLTITGIKGRINRAVWGPLNSTII SGGEDSVVRIWDSETGKLLRESDKETGHQKAITSLCKSADGSHFLTGSLDKSAR LWDIRTLTLIKTYVTERPVNAVAISPLLDHVVIGGGQEASHVTTTDRRAGKFEA KFFHKILEEEIGGVKGHFGPINSLAFNPDGRSFASGGEDGYVRLHHFDPDYFHI KM	131	1111

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
448	WD40 repeat protein	MAENNVGDFIPLDRQEYPSKPAPGAVDSSFWKSFKKKEVSRQIAGVTCINFCPE PPHDFAVTSSTRVHIYDGKSCELKKTITKFKDVAYSGVFRSDGQIIAAGGETGV IQVFNAKSQMVLRQLKGHGRPVRVVRYSPQDKLHLLSGGDDSMVKWWDITTQEE LLNLEGHKDYVRCGAASPSSVNLWATGSYDHTVRLWDLRNSKTVLQLKHGKPLE DVLFFPSGGLLATAGGNVVKVWDILGGGRPIHTMETHQKTVMAMCISKVPRSGQ ALGDAPSRLVTASLDGYMKVFDLDHFKVTHSARYPAPILSMGISSLCRTMAVGT SSGLLFIRQRKGQIEDKIHSDSSGLQVNPVNDEKDSAVLKPNQYRYYLRGRSEK PSEGDYVVKRMAKVYFQEYDKDLRHFNHSKALVSALKAADSKGTVAVIEELVAR KRLIQTLSILNLDELELLINFLSRFILVPKYSRFLISLTDRVLDARAVDLGKSE NLKKQIADLKGIVVQELRVQQSMQELQGIIEPLIRASAR	149	1726
449	WD40 repeat protein	MDVETSGKPTGNKRTYTRLPRQVCVFWQEGRCTRESCNFLHVDEPGSVKRGGAT NGFAPKRSYNGSDERDTLAAGPPGGSRRNISARWGRGRGGIFISDERQKIRNKV CNYWLAGNCQRGEECKYLHSFVMGSDVKFLTQLSGHVKAIRGIAFPSDSGKLYS GGQDKKVIVWDCQTGQGTDIPLNDEVGCLMSEGPWIFVGLPNAVKAWNILTSTE LSLVGPRGQVHALAVGNGMLFAGTHDGSILAWKFSPASNTFEPAASLVGHTQAV VSLVSGADRLYSGSMDKTIRVWDLGTFQCLQTLRDHTSVVMSLLCWDQFLLSCS LDNTVKVWVATSSGALEVTYTHNEEHGVLALCGMNDEQAKPVLLCSCNDNTVRL YDLPSFSERGRIFSRNEVRTFQIAPGGLFFTGDATGELKVWNWATQKS	948	2228
450	WD40 repeat protein	MSVQELRERHAAATAKVNALRERIKAKRLQLLDTDVATYASSNGRTPISFSFTD LVCCRTLQGHTGKVYSLDWTSEKNRIVSASQDGRLIVWNALTSQKTHAIKLPCA WVMTCAFSPSGQAVACGGLDSVCSIFQLNNQLDRDGHLPVSRILSGHRSYVSSC QYVPDGDTHVITGSGDRTCIQWDVTTGQRIAIFGGEFPLGHTADVMSVSISAAN PKEFVSGSCDTTTRLWDTRIASRAIRTFHGHEADVNTVKFFPDGLRFGSGSDDG TCRLFDIRTGHQLQVYRQPPRENQSPTVTAIAFSFSGRLLFAGYSNGDCFVWDT ILEKVVLNLGELQNTHNGRISCLGLSADGSALCTGSWDKNLKIWAFGGHRKIV	332	1465
451	WD40 repeat protein	MKVKIISRSTDEFTRERSNDLQRVFRNFDPNLHTQARAQEYVRALNAAKLDKIF AKPFLAAMSGHIDGISAMAKSPRHLKSIFSGSVDGDIRLWDIAARRTVQQFPGH RGAVRGLTVSTEGGRLISCGDDCTVRLWDIPVAGIGESSYGSENVQKPLATYVG KNSFRAVDYQWDSNVFATGGAQVDIWDHDRSEPTNSFAWGSDTVISVRFNPAEK DIFATTASDRSIVLYDLRMASPLNKLIMQTRNNAIAWNPREPMNFTAANEDCNC YSYDMRRMNISTCVHQDHVSAVMDIDYSPSGREFVTGSYDRTVRIFPYNAGHSR EIYHTKRMQRVFCVKFSGDATYVVSGSDDANIRLWKAKASEQLGVLLPRERKRH EYLDAVKERFKHLPEIKRIERHRHLPKPIYKAALLRHTVNAAAKRKEERKRAHS APGSVVTNPLRKKRIVAQLE	232	1590
452	WD40 repeat protein	MDHYYQDDFDYLVDDEMVDFADDVEDDVRTRRRSDIDSDSENDFDLNNKSPDTT ALQAKRGKDIQGIPWNRLNFTREKYRETRLQQYKNYENLPRPRRSRNLDKECTN FERGSSFYDFRHNTRSVKATIVHFQLRNLVWATSKHNVYLMQNYSIMHWSSLKQ KGEEVLNVAGPIVPSVKHPGSSPQGLTRVQVSAMSVKDNLVVAGGFQGELICKY LDKPGVSFCTKISHDENGITNAVEIYNDASGATRLMTANNDLAVRVFDTEKFTV LERFSFPWSVNHTSVSPDGKLVAVLGDNADCLLADCKTGKTVGTLRGHLDYSFA AAWHPDGYILATGNQDTTCRLWDVRKLSSSLAVLKGRMGAIRSIRFSSDGRFMA MAEPADFVHLYDTRQNYTKSQEIDLFGEIAGISFSPDTEAFFVGVADRTYGSLL EFNRRRMNYYLDSIL	207	1550

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
453	WD40 repeat protein	MAEALVLRGTMEGHTDAVTAIATPIDNSDMIVSSSRDKSILLWNLTKEPEKYGV PRRRLTGHSHFVQDVVISSDGQFALSGSWDSELRLWDLNTGLTTRRFVGHTKDV LSVAFSIDNRQIVSASRDRTIKLWNTLGECKYTIQPDAEGHSNWISCVRFSPSA TNPTIVSCSWDRTVKVWNLTNCKLRNTLVGHGGYVNTAAVSPDGSLCASGGKDG VTMLWDLAEGKRLYSLDAGDIIYALCFSPNRYWLCAATQQCVKIWDLESKSIVA DLRPDFIPNKKAQIPYCTSLSWSADGSTLFSGYTDGKIRVWGIGHV	221	1171
454	WD40 repeat protein	MAAIKSTSRSASVAFAPDAPLLAAGTMAGAIDLSFSSLANLEIFKLDFQSDDPE LPVVGECPSNERLNRLSWGSAGGSFGIIAGGLVDGTINIWNPATLINSEDNGDA LIARLEQHTGPVRGLEFNTISTNLLASGAEDGELCIWDLANPTAPTHFPPLKGV GSGAQGEISFLAWNRKVQHILASTSYSGTTVVWDLRRQKPIISFPDATRRRCSV LQWNPDASTQLIVASDDDNSPTLRAWDLRNTISPYKEFVGHSRGVIAMSWCPSD SLFLLTCAKDNRTLCWDTGSGEIVCELPAGANWNFDVQWSPKIPGILSTSSFDG KIGIHNIEACSRNVSGEVEFGGAIVRGGPSALLKAPKWLERPAGVSFGFGGKLA SFRPSTVAQAADHRHSEVFIHNLVTEDNLVIRSTEFEAAIADGEKVSLRALCDR KAEESQSDEEKETWNFLRVMFEDEGTARTKLLEHLGFKVQSEENGDLQETHSSK IDDIGSEIGKTLTLDDKTEEDVLPQLKGGQDAAIPQDNGEDFFDNLHSPKEEVS LSHVGNDFVGEKDKDMVVNGAEIEHETEDLTEYSDWNEAIQHSLVVGDYKGAVL QCLSANRMADALIIAHLGGNSLWEKTRDEYLKKAKSSYLKVVSAMVNNDLTGLV NSRPLKSWKETLAMLCTYSQREEWTVLCDMLASRLIAAGNVMAATLCYICAGNI EKTVEIWSRSLKYDYDGRSFVDHLQDVMEKTVVLALATGQKRVSPSLSKLVENY AELLASQGLLTTAMEYLKLLGTEESSHELSILRDRLYLSGTDNKVEASSFPFET RQDLTESQYNMHQTGFGAPETQKNYQENVHQVLPSGSYTDNYQPTANTHYIAGY QPAPQQQPSFQNYFTPASYQPAPSPNVFYPSQVSQAEQSNFAPPVNQPPMKTFV PSTPPILRNVDQYQTPSLNPQLYQGVSSATVETHPYQTGAPASVSVGTTPGQPS VVPNFMVPGPVTAPTVTPRGFMPVTTPTQHPLGSANPPVQPQSPQSSQVQSVTA ATTPPPTIQNVDTSNVAAEIRPVIGTLRRLYDETSEALGGARANPAKRREIEDN SRKIGSLFAKLNSGDISSNAASKLVHLCQALESRDYATAFQIQVGLTTSDWDEC SFWLAALKRMIKVKQNMR	221	3679
455	WD40 repeat protein	MAGAADSQLQTLSERDSTPNFKNLHTREYAAHKKKVHSVAWNCTGTKLASGSVD QTARVWNIEPHGHSKTKDLELKGHADSVDQLCWDPKHSELLATASGDRTVRLWD ARSGKCSQQVELSGENINITFKPDGTHIAVGNRDDELTIIDVRKFKPLHKRKFS YEVNEIAWNTTGELFFLTTGNGTVEVLSYPSLQVLHTLVAHTAGCYCIAIDPIG RYFAVGSADALVSLWDLSEMLCVRTFTKLEWPVRTISFNHDGQYIASASEDLFI DIADVQTGRTVHQISCRAAMNSVEWNPKYNLLAFAGDDKNKYMQDEGVFRVFGF ETP	269	1252
456	WD40 repeat protein	MAATSPVGAGSGRELANPPTDGISNLRFSNHSDHLLVSSWDRKVRLYDASANSL KGQFVHGGPVLDCCFHDDASGFSGSADNTVRRYDFSTRKEDILGRHEAPVRCVE YSYAAGQVITGSWDKTLKCWDPRGASGQEKTLVGTYSQLERVYSMSLVGHRLVV ATAGRHINVYDLRNMSQPEQRRESSLKYQTRCVRCYPNGTGFALSSVEGRVAME FFDLSEAGQAKKYAFKCHRKSEAGRDTVYPVNAIAFHPIYGTFATGGCDGYVNV WDGNNKKRLYQYSKYPTSIAALSFSRDGRLLAVASSYTFEEGEKPHEPDAVFVR SVNEAEVKPKPKVYAAPP	214	1242

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
457	WD40 repeat protein	MASDDEEGFKNEEAPGVVDEAEVQEGLRACFPLSFGKQEKKQAPLESIHSATKR PEDPRPRRQLGPPRPPPSILAEQEDSDRFVGPPRPPQFVRDDNDDGEAEIMIGP PRPPAQYSDDHDNEETIGPPKPSYLEKGEETDQMVGPSKRGSDDETSGDSDDGD DAVDFRVPLSNEIVLRGHTKVVSALAIDQTGSRVLTGSYDYSVRMYDFQGMTSQ LKSFRQLEPAEGHQVRSLSWSPTSDRFLCVTGSAQAKIFDRDGLTLGEFVKGDM YLRDLKNTKGHISGLTCGEWHPKEKQTILTCSEDGSLRIWDVNDFNTQKQVIKP KLAKPGRVPVTACAWGRDGKCIAGGVGDGSIQVWNLKPGWGSRPDLYVAKGHDD DITGLQFSADGNILLTRSTDETLKVWDLRKAITPLQVFRDLPNNYAQTNVAFSP DERLIFTGTSVERDGNSGGLLCFYDRQTLELVLRIGVSPVHSVVRCTWHPRHNQ VFATVGDKKEGGAHILYDPALSERGALVCVARAPRKKSLDDFEAKPVIHNPHAL PLFRDEPSRKRQREKARMDPMKSQRPDLPVTGPGFGGRVGSTKGSLLTQYLLKE GGLIKETWMEEDPREAILKYADVAAKDPKFIAPAYAQTQPETVFAETDSEEEQK	119	2065
458	WD40 repeat protein	MKERGQSHAGQPSVDERYTQWKSLVPVLYDWLANHNLVWPSLSCRWGPQMHQAT YKNSQRLYLSEQTDGTVPNTLVIATCEVVKPRVAAAEHISQFNEEARSPFVKKF KTIIHPGEVNRIRELPQNSKIVATHTDGPDVLIWDVDTQPNRQATLGAADSRPD LVLTGHKDNAEFALAMSPSAPFVLSGGKDKCVLLWSIQDHISAATEPSSAKASK TPSSAHGEKVPKIPSIGPRGVYKGHKDTVEDVQFCPSNAQEFCSVGDDSALILW DARNGNEPVIKVEKAHNADLHCVDWNPHDENLILTGSADNSVRMFDRRNLTSSG VGSPVHKFEGHSAPVLCVQWCPDKASVFGSAAEDSYLNVWDYEKVGKNVGKKTP PGLFFQHAGHRDKVVDFHWNSFDPWTIVSVSDDGESTGGGGTLQIWRMSDLIYR PEDEVLAELERFRAHILSCQNK	186	1550
459	WD40 repeat protein	MSSLSRELVFLILQFLDEEKFKESVHKLEQESGFFFNMKYFDEKAQAGEWDEVE RYLSGFTKVDDNRYSMKIFFEIRKQKYLEALDRQDRAKAVDILVKDLKVFSTFN EELYKEITQLLTLDNFRENEQLSKYGDTKSARTIMMSELKKLIEANPLFREKLI YPNLKASRLRTLINQSLNWQHQLCKNPRPNPDIKTLFTDHACGPPNGARTPTQP TASLGVLPKATTFTPIGPHGPFPSSSTATSGLASWMSNPNMVTSPQAPVAVGPS VPVPPNQATLLKRPRTPPGSSSVVDYQTADSEQLIKRLRPVSQSIDEATYPGPT LRVPWSTDDLPKTLARALNEPYPVTSIDFHPSQQTFLLVGTKNGEITLWEVGSR EKLATRSFKIWDNANCSNHLEAAFVKDSSVSINRVLWSPDGTLIGIAFTKHLVH TYTFQGLDLRQHLEIDAHVGGVNDLAFSHPNKQLCVVTCGDDKMIKVWDAVTGR KLYNFEGHDAPVYSVCPHHKENIQFIFSTAVDGKIKAWLYDHLGSRVDYDAPGH SCTTMMYSADGTRLFSCGTSKEGESFLVEWNESEGAIKRTYSGLRKKGSGVVQF DTTQNHFLAVGDEHLIKFWDMDSTNMLTSCDAEGGLLNLPRLRFNKEGSLLAVT TVNGIKILANADGQKLLKTMENRTFDLPSRAHIDAASATSSPATGRMERIERTS SANTVSGINGVDPAQSSEKLRLSDDLSEKTKIWKLTEITDSIQCRCITLPENAA EPASKVSRLLYTNSGVGLLALGSNAVHKLWKWNRSEQNPSGKATASVHPQRWQP TSGLLMTNDITDINPEEAVPCIALSKNDSYVMSASGGKVSLFNMMTFKVMTTFM PPPPASTFLAFHPQDNNIIAIGMEDSTIHIYNVRVDEVKTKLKGHQKRITGLAF SSTQNILVSSGADAQLCVWNTETWEKRKSKTIQMPVGKTVSGDTRVQFHSDQLH ILVVHETQLAIYDAYKLERQYQWVPQDALSAPILYATYSCNRQLIYATFSDGNI GVYDAEILRPRCRIAPTTYLSSGTSSSTSLPLVVAAHPHEPNQFAIGLSDGAVQ VLEPSESEGKWGVSPPPENGVVPAVVAGPSTSNQGSEQAPR	244	3671

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
460	WD40 repeat protein	MAKDEEEFRGEMEERLVNEEYKIWKKNTPFLYDLVITHALEWPSLTVQWLPDRE EPPGKDYSVQKMILGTHTSDNEPNYLMLAQVQLPLEDAENDARQYDDERGEIGG FGCANGKVQVIQQINHDGEVNRARYMPQNPFIIATKTVSAEVYVFDYSKHPSKP PQDGGCHPDLRLRGHNTEGYGLSWSPFKHGHLLSGSDDAQICLWDINVPAKNKV LEAQQIFKVHEGVVEDVAWHLRHEYLFGSVGDDRHLLIWDLRTSATNKPLHSVV AHQGEVNCLAFNPFNEWVLATGSADRTVKLFDLRKISSALHTFSCHKEEVFQIG WSPKNETILASCSADRRLMVWDLSRIDEFQTPEDALDGPPELLFIHGGHTSKIS DFSWNPCEDWVIASVAEDNILQIWQMAENIYHDEEDDMPPEEVV	163	1431
461	WD40 repeat protein	MSPGVKQTGSQKFESGHQDVVHDVTMDYYGKRIATCSADRTIKLFGLNASDTPS LLASLTGHEGPVWQVAWAHPKFGSMLASCSYDGRVIIWREGQQENEWSQVQVFK EHEASVNSISWAPNELGLCLACGSSDGSITVFTCREDGSWDKTKIDQAHQVGVT AVSWAPASAPGSLVGQPSDPIQKLVSGGCDNTAKVWKFYNGSWKLDCFPPLQMH TDWVRDVAWAPNLGLPKSTIASCSQDGKVVIWTQGKEGDKWEGRILNDFKIPVW RVNWSLTGNILAVADGNNSVTLWKEAVDGDWNQVTTVQ	155	1081
462	WD40 repeat protein	MSSGVKQTGSQKFESGHQDVVHDVTMDYYGKRIATCSADRTIKLFGMNTSDTPT LLASLTGHEGPVWQVAWAHPKFGSMLASCSYDRRVIIWREGQQENEWSQVQVFK EHEASVNSISWAPHELGLCLACGSSDGSITVFTGREDGSWDKTKIDQAHQVGVT AVSWAPASAPGSLVGQPSDPVQKLVSGGCDNTAKVWKFYNGSWKLDCFPPLQMH TDWVRDVAWAPNLGLPKSTIASCSQDGRVVIWTQGKEGDKWEGKILNDFKTPVW RISWSLTGNILAVADGNNNVTLWKEAVDGEWNQVTTVQ	537	1463
463	WD40 repeat protein	MKKRSRPSNGHLSTAAKNKSRKTAPITKDPFFDSAHNRNKSKGKGKSRGKGEEI FSSDEDDDAIGRDAPAEEEEEIAEEERETADEKRLRVAKAYLDKIRAITKANEE DNEEEAGEDEETEAERRGKRDSLVAEILQQEQLEESGRVQRQLASRVVTPSKLV ECRVVKRHKQSVTAVALTEDDLRGFSASKDGTIIHWDVETGASEKYEWPSQAVS VSSSNEVSKTQKGKGSKKQGSKHVLSMAVSSDGRYLATGGLDRYIHLWDTRTQK HIQAFRGHRGAVSCLAFRQGTQQLISGSFDRTIKLWSAEDRAYMDTLYGHQSEI LAVDCLRKERVLSVGRDHTLRLWKVPEETQLVFRGHAASLECCCFINNEDFLSG SDDGSIELWSMLRKKPVFMAKNAHGHAIVENLSEDTSTREEPDEEVTTRQLPNG NSIGNGMTNQMGITPSVESWVGAVTVCRGTDLAASGAGNGVVRLWAIENSSKSL RALHDIPLTGFVNSLTFARSGRFLIAGVGQEPRLGRWGRIQAARNGVTLCPIEL S	284	1909
464	WD40 repeat protein	MAATFGTINTATSPHNPNKSFEIVQPPNDSISSLSFSPKANYLVATSWDNQVRC WEVLQTGASMPKAAMSHDQPVLCSTWKDDGTAVFSAGCDKQAKMWPLLTGGQPV TVAMHDAPIKDIAWIPEMNLLATGSWDKTLKYWDTRQSNPVHTQQLPERCFALS VRHPLMVVGTADRNLIIFNLQNPQTEFKRISSPLKYQTRCVAAFPDKQGFLVGS IEGRVGVHHVEEAQQSKNFTFKCHRDSNDIYAVNSLNFHPVHQTFATAGSDGAF NFWDKDSKQRLKAMARSNQPIPCSTFNSDGSLYAYAVSYDWSKGAENHNPATAK HHILLHVPQESEIKGKPRVTTSGRK	610	1659

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
465	WD40 repeat protein	MVVMDKGTHQTNEDESESEFIDEDDVIDEISIDEEDLPDADVEGEDVQEDNKRS EPDENSSSLDDAIHTFEGHEDTLFAVACSPVDATWVASGGGDDKAFMWRIGHAT PFFELKGHTDSVVALSFSNDGLLLASGGLDGVVRIWDASTGNLIHVLDGPGGGI EWVRWHPKGHLVLAGSEDYSTWMWNADLGKCLSVYTGHCESVTCGDFTPDGKAI CTGSADGSLRVWNPQTQESKLTVKGYPYHTEGLTCLSISSDSTLVVSGSTDGSV HVVNIKNGKVVASLVGHSGSIECVRFSPSLTWVATGGMDKKLMIWELQSSSLRC TCQHEEGVMRLSWSLSSQHIITSSLDGIVRLWDSRSGVCERVFEGHNDSIQDMV VTVDQRFILTGSDDTTAKVFEIGAF	241	1452
466	WD40 repeat protein	MPVFRTAFNGYAVKFSPFVETRLAVATAQNFGIIGNGRQHVLELTPNGIVEVCA FDSSDGLYDCTWSEANENLVVSASGDGSVKIWDIALPPVANPIRSLEEHAREVY SVDWNLVRKDCFLSASWDDTIRLWTIDRPQSMRLFKEHTYCIYAAVWNPRHADV FASASGDCTVRIWDVREPNATIIIPAHEHEILSCDWNKYNDCMLVTGSVDKLIK VWDIRTYRTPMTVLEGHTYAIRRVKFSPHQESLIASCSYDMTTCMWDYRAPEDA LLARYDHHTEFAVGIDISVLVEGLLASTGWDETVYVWQHGMDPRAC	223	1173
467	WD40 repeat protein	MDSRNRRSRLNLPPGMSPSSLHLETTAGSPGLSRVNSSPSTPSPSRTTTYSDRF IPSRTGSRLNGFALIDKQPQPLPSPTRSAAEGRDDASSSSASAYSTLLRNELFG EDVVGPATPATPEKSTGLYGGSRDSIKSPMSPSRNLFRFKNDHGGNSPGSPYSA STVGSEGLFSSNVGTPPKPARKITRSPYKVLDAPALQDDFYLNLVDWSSNNVLA VGLGTCVYLWSACTSKVTKLCDLGVNDSVCSVGWTPQGTHLAVGTNIGEVQIWD TSRCKKVRTMGGHCTRAGALAWSSYILSSGSRDRNILHRDIRVQDDFIRKLVGH KSEVCGLKWSYDDRELASGGNDNQLLVWNQQSAQPLLRFNEHTAAVKAIAWSPH QHGILASGGGTADRCLRFWNTATDTRLNCVDTGSQVCNLVWCKNVNELVSTHGY SQNQIMVWRYPSMSKLATLTGHTLRVLYLAISPDGQTIVTGAGDETLRFWSIFP SPKSQSAVHDSGLWSLGRTHIR	251	1777
468	WD40 repeat protein	MEKKKVVVPIVCHGHSRPIVDLFYSPVTPDGLFLISASKDSSTMLRNGETGDWI GTFEGHKGAVWSCCLDNRALRAASGSADFSAKIWDALTGDELHCFVHKHIVRAC AFSESTSLLLTGGHEKILRIFDLNRPDAPPKEVDNSPGSIRTVAWLHSDQTILS SNSDAGGVRLWDLRTEKIVRVLETKSPVTSAEVSQDGRYITTADGNSVKFWDAN HFGMVKSYTMPCMVESASLEPTMGNMFVAGGEDMWVRLFDFHTGEEIACNKGHH GPVHCVRFAPGGESYSSGSEDGTIRIWQTLNMNSEENESYGVNGLSGKVRVGVD DVVQKVEGFQITADGHLNDKPEKPNP	367	1419
469	WD40 repeat protein	MERYSQGTQKKSEIYTYEAPWQIYGMNWSVRKDKKFRLGIGSFLEEYNNRVEII ELDEESGEFKSDPRLAFDHPYPTTKIMFVPDKECQRPDLLATTGDYLRIWQVCE DRVEPKSLLNNNKNSEFCAPLTSFDWNDADPKRIGTSSIDTTCTIWDIEKEVVD TQLIAHDKEVYDIAWGEVGVFASVSADGSVRVFDLRDKEHSTIIYESSQPETPL LRLGWNKQDPRFIATILMDSCKVVILDIRFPTLPVAELQRHQASVNTIAWAPHS PCHICTAGDDSQALIWELSSVSQPLVEGGGLDPILAYTAAAEINQLQWSSMQPD WVAIAFSNEVQILRV	284	1303

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
470	WD40 repeat protein	MQSENNLDESLHLREVQELQGHTDTVWAVAWNPVTGIDGAPSMLASCSGDKTVR IWENTHTLNSTSPSWACKAVLEETHTRTVRSCAWSPNGKLLATASFDATTAIWE NVGGEFECIASLEGHENEVKSVSWSASGMLLATCGRDKSVWIWDVQPGNEFECV SVLQGHTQDVKMVQWHPNRDILVSASYDNSIKVWAEDGDGDDWACMQTLGNSVS GHTSTVWAVSFNSSGDRMVSCSDDLTLMVWDTSINPAERSGNAGPWKHLCTISG YHDRTIFSVHWSRSGLIASGASDDCIRLFSESTDDSVTPVDGTSYKLILKKEKA HSMDVNSVQWHPSEPQLLASASDDGRIKIWEVTRINGLANSH	684	1784
471	WD40 repeat protein	MKRAYKLQEFVAHASNVNCLKIGKKSSRVLVTGGEDHKVNMWAIGKPNAILSLS GHSSAVESVTFDSAEALVVAGAASGTIKLWDLEEAKIVRTLTGHRSNCISVDFH PFGEFFASGSLDTNLKIWDIRRKGCIHTYKGHTRGVNSIRFSPDGRWVVSGGED NIVKLWDLTAGKLMHDFKCHEGQIQCMDFHPQEFLLATGSADRTVKFWDLETFE LIGSAGPETTGVRAMIFNPDGRTLLTGLHESLKVFSWEPLRCYDAVDVGWSKLA DLNIHEGKLLGCSYNQSCVGVWVVDISRVGPYAAGNVSRTNGHNEAKLASSGHP SVQQLDNNLKTNMARLSLSHSTESGIKEPKTTTSLTTTEGLSSTPQRAGIAFSS KNLPASSGPPSYVSTPKKNSTSRVQPTTNFQTLSRPDIVPVIVPRSNSLRPETT SDVKKEMNNFGRVVPSTVSTKSTDVIKSGSNRDESDKIDSINQKRMTGNDKTDL NIARAEQHVSSRLDNTNTSSVVCDGNQPAARWIGAAKFRRNSPVDPVVSPHDRS PTFPWSATDDGVTCQPDRQVTAPELSKRVVEPGRARALVASWETREKALTADTP VLVSGRPPTSPGVDMNSFIPRGSHGTSESDLTVSDDNSAIEELMQQHNAFTSIL QARLTKLQVIRRFWQRNDLKGAIDATGKMGDHSVSADVISVLIERSEIFTLDIC TVILPLLTRLLQSETDRHLTVAMETLLVLVKTFGDVIRATISATPTIGVDLQAE QRLERCNLCYVELENIKQILVPLIRRGGAVAKSAQELSLALQEV	336	2738
472	WD40 repeat protein	MSTLEIEARDVIKIVLQFCKENSLHQTFQTLQNECQVSLNTVDSLETFVADINS GRWDVILPQVAQLKLPRKKLEDLYEQIVLEMIELRELDTARAILRQTQAMGFMK QEQPERYLRLEHLLVRTYFDPREAYHESSKEKRRSQIAQALASEVTVVPPSRLM ALIGQSLKWQQHQGLLPPGTQFDLFRGTAAVKADEEEMYPTTLAHTIKFGKQSH PECARFSPDGQYLVSCSVDGFIEVWDYISGKLKKDLQYQADDSFMMHDDAVLCV DFSRDSEMLASGSQDGKIKVWRIRTGQCLRRLERAHSQGVTSLSFSRDGSQLLS TSFDSTARIHGLKSGKALKEFRGHTSYVNDAIFTSDGGRVITASSDCTVKVWDV KTTDCIQTFKPPPPLKGGDVSVNSVHLFPKNSEHIVVCNKASSIYIMTLQGQVV KSFSSGKREGGDFVAACISPKGEWIYCVGEDRNIYCFSQQSGKLEHLMKAHDKD IIGVTPHPHRNLLVTYSEDSTMKIWKP	81	1622
473	WD40 repeat protein	MDIELEDQPFDLDFHPSAPIVAVALITGRLQLFRYVDISSEPERLWTVTAHTES CRAARFINAGSSVLTASPDCSILATNVETGQPVARLDNAHGAAINCLTNLTEST IASGDENGIIKVWDTRQNSCCNKFKAHEDYISDMEFVPDTMQLLGTSGDGTLSV CNLRKNKVHARSEFSEDELLSVALMKNGKKVVCGSQEGVLLLYSWGYFKDCSDR FVGHPHSVDALLKLDEDTVLTGSSDGIIRVVSILPNKMIGVIGEHSSYPIERLA FSHDRNVLGSASHDQILKLWDIHYLHEDDEPETNKQEAVNDENVDMDLDVDTEK RPRGSKRKKRAEKGQTSSQKQSSDFFADI	399	1460

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
474	WD40 repeat protein	MDRIQQIPHTCVARKINLPLGMSKESLALNLPANLAPTMSPPSITYSDRFIPSR KASNFEEFALPDKTSPSPNSAGGQSSSTNGEGRDDACAAYSALLRTELFPATPD KTEGCRRPVIGSPSGNVFRFKSQQCKSQSPFSLCPVGEDGDLSETGAVARKTTR KIPRSPFKVLDAPALQDDFYLNLVDWSSHNILAVGLSACVYLWSASSSKVTKLC DLGLDDNVCSVAWTQRGTYLAVGTNNGGVQIWDAAHCKQVRTMEGHCTRVGTLA WNSHILSSGGRDRNILQRDIRAQDDFVSKFSGHKSEVCGLKWSYDNRELASGGN DNQLFVWNQQSQQPVLKYNEHTAAVKAIAWSPHQHGLLASGGGTADRCIRFWNT ATNTSLNCVDTGSQVCNLVWSKNVNELVSTHGYSQNQIIVWRYPTMSKLATLTG HTLRVLYLAISPDGQTIVTGAGDETLRFWNVFPSSKTQQNTIRDMGVWSSGRTH IR	207	1673
475	WD40 repeat protein	MAGGQGEGEEKVDKLSMELTEDVMKSMEIGAVFKDYNGKINSLDFHRTNNYLVT ASDDEAIRLFDTASATWQKTSYSKKYGVDLICFTNHQTSVLYSSKNGWDESLRH LSLMDNKYLRYFKGHHDRVVSLCMSPKGECFMSGSLDRTVLLWDLRIDKCQGLI RVRGRPAVAYDEQGLVFAISNEGGLIKMFDARLYDKGPFDTFVVEGDKSEASGI KFSNDGKLILLSTMDSNIHVLDAYQGTTVHSFSVEAVPNGGEAVPNGGTLEASF SPDGKFVISGSGNGNIHAWSVNSGKEVACWTTEGVIPAVVKWAPRRLMFASGSS VLSLWVPDLSKLASLTGSNSNSAY	263	1309
476	WD40 repeat protein	MHRVGSTGNTSNSSRPRREKRLTYVLNDANDSRHCSGINCLVISKLSLLGGNDY LFSGSRDGTLKRWELADDSAVCSATFESHVDWVNDAVLTGETLVSCSSDTTLKT WRPFSDGVCTRTLRQHSDYVTCLAAASKNSNIVASGGLGREVFIWDIEAAMAPV SRTSEAMDDDTSNGVLSSGNSVLSTTVRSTNATNSASLHTSQLQGYTPIAAKGH KESVYALAMNDVGTLLVSGGTEKVVRVWDPRSGAKQMKLRGHTDNVRALILDST GRFCLSGSSDSIIRLWDLGQQRCVHSYAVHTDSVWALASTPNFSHVYSGGRDLS LYLTDLTTRESLLLCMEKHPLLRLTLQDDSIWVATTDSSLHRWPAEGQNPPKMF QRGGSFLAGNLSFTRARACLEGSAPVPVNTQPSFVIPGSPGIVQHEILNNRRHV LTKDAEGTVKLWEITRGAVLDDYGKVSFEEKKEELFEMVSIPAWFTMDTRLGSM SVHLDTPQCFTAEMYAVDLNVPDAPEEQKINLAQETLRGLLAHWLSRRRQRLAT QASANGDFPAGQENALRNHISSRIDVHDDAETHIAGILPAFDFSTTSPPSIITE GSQGGPWRKKITDLDGTEDEKDFPWWCLECVLHGRLSPRESLKCSFYLHPYEGT TVQVLTQGKLSAPRILRIQKVINYVLEKMVLDRPLDSSNSETTFTPGLSGNQSH AAVVGDGSLRSGARVWQQKAKPLVEILCNNQVLSPDMSLATVRTYIWKKPDDLY LYYRLVQNR	232	2529

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
477	WD40 repeat protein	MMKGKTIQMQAAHQNHDGETSVACVLWDWHAKHLITAGADNTILIHSYPSSSSS KPITLRHHKNAVTALAINSNVRSLASGSVDHSVKLYSYPGGEFQSNVTRFTLPI RSLAFNKSGELLAAAGDDEGIKLISTIDNSIARVLKGHNGPVTSISFDPKNEFL ASSDSDGTVIYWELSTGKPVHTLKKIAPNTTSNPTSLNQISWRPDGEMLAVPGR KSEVSMYDRDTAEKLFSLKGGHSDTICSLAWSPNGKYIATAGTDRQVMVWDADR RQDIDKQRFDNPICSVAWKPSDNALAVIDVLGRFGVWESPIASHMKSPADGAER YDNMEDEEPLMARYEEELEDSVSGSLNEIINDDDDDDEMGKIPRKILQKKPSVK VEKGKEESNAKAFKSGQDSFKLKSAMQEAFQPGATQRQSGKRNFLAYNMLGSVI TFDNDGFSHIEVDFHDIGKGCRVPSMTDYFGFTMASLSESGSVFGSPQKGEKNP STLMYRPFSSWANNSEWSMRFPMGEEVKAVALGSGWVAAVTSLNFLRVFSEGGL QKFVLSMDGPVVTAAGYENLLVVVSHASNPLLSGDQVLSFTVYDISQKTCPLSG RLPLSPGSHLTWLGFSEEGLLSSYDSEGNLRVFTNDYNGCWVPIFSAARERKSE TESIWMVGLNSTQVFCVVCKLPDTYPQVAPKPVLSVLNLSLPLACSDLGADDLE NEYLRGSLLLSQMQKKAEDAVACGRESNMEEDSIFKMEAALDRCLLRLIANCCK GDKLVRATELARLLSLEKSLQGAIKLVSAMKLPMLAERFNTILEEKILQENMET ISCRRLTSEAQDMDTPISISVKQVSYGANLGDSPFLPNRQVEPKHSTPVFSKPD TKIEVDTSEAIAKGCDAQNGNIKSGDAEVQPASHNDSIQKPSNPFAKASNTSAN QAVQRNASLLSSIKQMKTATENEGKRKERARSGSLPQKPAKQSKIS	56	2950
478	WD40 repeat protein	MKQKRKGHQVDDPKYSVQTPQEDDTPNESGPASEEVESSDEEGGNSSNIEDDII YSSSEEDPVVSSDYEEDEDAESDAEGVTAEQELEGDIDNALQNYMGTLTVLSNF HGENLKNAEGEDTSGDDDDEEEMPKRAEESDSPEDENDERPKRAEESDFSEDED EERPKRAEESDSSEDEVPSRNTVGDVPLRWYKDEQHIGYDIKGKKIKKQPKKDQ LDSFLASTDDSSDWRKVYDEYNDEEVELTKDEIKFISRLRKGTIPHADVNPYEP YVDWFDWKDKGHPLSNAPEPKRRFIPSKWEAKKVVKLVRAIRKGWITFQKAEEK PRFYLMWGDDLKPSEKMANGLSYIPAPKPKLPGHEESYNPPPEYIPTQEEINSY QLMYEEDRPKFIPKRFDSLRNVPAYDRFLSEIFERCLDLYLCPRTRKKRINIDP ESLIPKLPKPKDLQPFPSICFLEYKGHTGAVSCISPESSGQWLASGSKDGTVRI WEVETARCLKVWDIGRPIQHIAWNPVSQLSILAVAVDEEVLVLNTGLGSEDSQE KVAELLHVKSKPVSADDLGDNTSLTKWIKHEKFDGIKLTHLKPVHLISWHHKGD YFATVAPDGNTRAVLVHQLSKQQTQNPFKKMQGRVVHVLFHPSRAIFFVATKTH VRVYDLVKQQLVKRLVTGLHEVSSMAVHHKGDNLLVGSKEGKVCWFDMDLSTQP YKTLKNHSKDIHSVAFHDSYPLFASCSDDCKAYVFYGLVYSDLLQNPLIVPLKV LQGHQSVNGMGVLDCQFHPKQPWLFTAGADSVVKLYCN	193	2577
479	WD40 repeat protein	MMSLKRGFEESLVPAKRQKTELSTVTYGDGPRRTSSLESPIMLLTGHHAAIYTM KFNPTGTVIASGSHEREIFLWNVHGDCKNFMVLKGHKNAVLDLHWTTDGCQIIS ASPDKTLRAWDVETGKQIKKMAEHSSFVNSCCPSRRGPPLVVSGSDDGTAKLWD LRHRGAIQTFPDKYQITAVGFSDAADKIYSGGIDNEIKVWDLRRGEVTMRLQGH TDTITGMQLSSDGSYLLTNSMDCSLRIWDMRPYAPQNRCVKILTGHQHNFEKNL LKCSWSSDGSKVTAGSADRMVYIWDTTTRRILYKLPGHTGSVNETGFHPTQPII GSCSSDKQIYLGEIEPNVGYQAVI	187	1233

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
480	WD40 repeat protein	MEFSDTYKHTGPCCFSPDARYLAIAVDYRLVIRDVVTLKVVQLYSCMDKISNIE WALDSEYILCGLYKRAMVQAWSLSQPEWTCKIDEGPAGIAHARWSPDSRHIITT SDFQLRLTVWSLVNTACIHIQWPKHASKGVSFTQDGKFAAIATRRDCKDYVNLL SCHTWEVMGTFTVDTIDLADLEWSPNDSAIVVWDSPLEYKVLIYSPDGRCLFKY QAYDSWLGVKTVAWSPCSQFLAVGSYDQTLRTLNHLTWKPFAEFVHVSTVRGPA SAVVFKEVEEPWNLDVSGLHLNDDNAHDIQDGKPAEGHSRVRYKVVEFPVNVSS QKHPVDKPNPKQGIGLLAWSRDSQYLFTRNDNMPTALWIWDICRLELAALLIQK EPIRAAAWDPVYPRVALCTGSSHLYMWTPSGACCVNIPLPQFVVSDLKWNPDGT SMLLKDRESFCCTFVPMLPEFNDDETNEE	51	1436
481	WD40 repeat protein	MAKLIETHSCVPSTERGRGILIAGDAKTNSIIYCNGRSVIMRNLDNPLEASVYG EHSYPATVARFSPNGEWVASGDTSGTVRIWGRGSDHTLKYEYKALAGRIDDLEW SADGQRIVVCGDSKGKSMVRAFMWDSGTNVGEFDGHSRRVLSCSFKPTRPFRVA TCGEDFLVNFYEGPPFRFKTSHRDHSNYVNCVRFAPDGSKFITVGSDRKGVIFD GKMGEKIGELSKEGGHTGSIYAASWSPDSKQVLTVSADKSAKIWEISETGNGTV KKTLTFGSQGGADDMLVGCLWLNDYLITVSLGGIVSLLSAVDPDKPPKTISGHM KSINAIALSLQSGQSEVCSSSYDGVIVRWILGVGYAGRVERKDSTQIKCLATIE GELVTCGFDNKVRRVPLLSEQHKESEPIDIGAQPKDLDVAVGCPELTFVSTDAG IIIIRASKIVSTTNVGYAVTAAAISPDGTEAVVGGQDGKLRVYSIKGDTLLEES VLERHRGPINAIRFSPDGSMFASGDLNREAVVWDRITREVKLKNMVYHTARINC IAWSPDSSKVATGSLDTCILIYEVGKPASSRITIKGAHLGGVYGLAFSDQSTVI SAGEDACVRVWSLP	525	2351
482	WD40 repeat protein	MPQPSVILATAGYDHTVRFWEATSGRCYRTLQYPDSQVNHLEITPDKQYLAAAG NPHIRLFEVNSNNPQPVISYDSHTNNVTAVGFQCDGKWMYSGSEDGTVKIWDLR APGFQREYESRAAVNTVVLHPNQTELISGDQNGNIRVWDLNANSCSCELVPEDT AVRSLTVMWDGSLVVAANNHGTCYVWRLMRGTQTMTNFEPLHKLQAHNSYILKC LLSPEFCEHHRYLATTSSDQTVKIWNVDGFTLERTLTGHQRWVWDCVFSVDGAF LVTASSDSTARLWDLSTGEAIRTYQGHHKATVCCALHDGTDGASC	152	1099

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
483	WD40 repeat protein	MLTKFETKSNRVKGLSFHPKRPWILASLHSGVIQLWDYRMGTLIDKFDEHDGPV RGVHFHKTQPLFVSGGDDYKIKVWNYKMRQCLFTFVGHLDYIRTVHFHNEYPWI VSASDDQTIRLWNWQSRVCISVLTGHNHYVMSASFHPKEDLVVSASLDQTVRVW DISGLRKKTVSPADDLSRLAQMNTDLFGGGDVVVKYVLEGHDRGVNWAAFHTSL PLIVSGADDRQVKLWRMNDTKAWEVDTLRGHTNNVSCVIFHARQDIIVSNSEDK SIRVWDMSKRTSVQTFRREHDRFWILAAHPEMNLLAAGHDSGMIVFKLERERPA YVVYGGSLLYVKDRYLRTYEFATQKDNPLIPIRKPGSIGPNQGPRSLSYSPTEN AILICSDADGGAYELYAVPKDSHGRSDTVQEAKKGLGGSAVFVARNRFAVLDKN HNQVTIKNLKNEVTKKFDLPVTADALFYAGTGNLLCRSEDSVFLFDMQQRTVLG EIQTPNVRYVVWSNDMENVALLSKHTIIIASKKLSSTCSLHETIRVKSGAWDDN GIFMYSTLNHIKYCLPNGDSGIIKTLDVPVYITKVSGKSLYCLDRDGKNRVIQI DITECLFKLALSKKKYDYVINMIRNSQLCGQAIIAYLQQKGFPEVALHFVRDER TRFNLAVESGNIEIAVASAKEIDEKDHWYRLGVEALRQGNAGIVEYAYQRTKNF ERLSFLYLITGNLDKLSKMLRIAEMKNDVMGQFHNALYLGDIQERIKILEESGH LHLAYATASLHGLADIADRLAADLGGNIPVLPPGKKSSLLMPPAPILHGGDWPL LRVTKGIFEGGLENSTSAAYEEEDEEAAADWGEDIDIENIEGENGEATVLDDQE VKGGEDDEGGWDMEDLELPPDVAAANVGTNQKTLFVAPTLGMPVSQIWMQKSSL AGEHAAAGNFETALRLLTRQLGIKNFSPLKPLFLELYMGSHTFLPSFASVPAFS LALQRGWSESASPNIRGPPALVYRLSVLEEKLTVAYRATTEGRFSEALRLFLNI LHTIPVIVVDSRKEIDEVKELIGIAKEYVLGLRMEVKRKEIRDDAVRQQELAAY FTHCNLQKAHLKLALLNAMGISYRCKNYNTAANFARRLLETDPSSNHATKARQV LQVCERNLQDATQLNYDFRNPFVVCGATFTPIYRGQKEVSCPYCMARFVPDIAG KLCSICDLAIVGSDASGLFCFATQTR	470	4114
484	WD40 repeat protein	MDLLQNYQDDSEDSNPELRNHPPLEDATATSAPAGVENETSSSPDSSPLRLALP AKSCAPDVDETLMALGVPGSEKKNNHNKPIDPTQHSVTFNPSYDQLWAPLYGPA HPYAKDGIAQGMRNHKLGFVEDSAIEPFMFDEQYNTFHRYGYAADPSASLGSHI VGDLESLKKNDGASVYNLPKREHKRQKLEKKMIQKDENEEEEKEVGEEVDNPST EEWLKKNRKSPWAGKKEGLQTELTEEQKKYAQEHAEKKGDREKGEKVEIVDKTT FHGKEERDYQGRSWIDPPKDAKATNDHCYIPKRWVHTWSGHTKGVSAIRFFPKY GHLLLSAGMDTKVKIWDVFNSGKCMRTYMGHSKAVRDISFSNDGSRFLSAGYDR NIKLWDTETGKVISTFSTGKIPYVVKLHPDEDKQNVLLAGMSDKKIVQWDMNSG EITQEYDQHLGAVNTITFVDNNRRFVTSSDDKSLRVWEFGIPVVIKYISEPHMH SMPSISLHPNTNWLAAQSLDNQILIYSTRERFQLNKKKRFAGHIAAGYACQVNF SPDGRFVMSGDGEGRCWFWDWKTCKVFRTLKCHDNVCIGCEWHPLEQSKVATCG WDGMIKYWD	196	2007
485	WD40 repeat protein	MARKGLGTDPAIGSLMSSKKRKEYKVTNRFQEGKRPLYAIAFNFIDARYHNIFA TAGGTRVTIYQCLEGGAISVLQAYVDDDKDESFYTLSWACDVNGSPLLVAGGHN GIIRVLDVANEKVHKSFVGHGDSVNEIRTQALKPSLILSASKDESVRLWNVQTG ICILIFAGAGGHRNEVLSVDFHPSDVYRIASCGMDNTVKIWSMKEFWTYVEKSF TWTDLPSKFPTKYVQFPVFIAAVHSNYVDCTRWLGNFILSKSVDNEVVLWEPYS KEQSTSDGVVDILQKYPVPECDIWFIKFSCDFHYNSMAVGNREGKVYVWELQSS PPNLIARLSHAHCKNPIRQTAISHDGSTILCCCDDGSMWRWDVVQ	214	1323

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
486	WD40 repeat protein	MESGAGGSVGARVPSAKPEMLQQPPYSNGDDDNDMERGTAPVPSSNPNTVSKWE LDKDFLCPICMQTMKDAFLTACGHSFCYMCIMTHLNNKSNCPCCSLYLTNNQLF PNFLLNKLLKKTSACQMASTASPVENLCLSLQQGAEVSVKELDFLLTLLAEKKR KMEQEEAETNMEILLDFLQRLRQQKQAELNEVQADLHYIKDDILALEKRRLELS RARERYSRKLHMLLDDPMDTTLGHAAIDDGNNVRTAFVRGGQGDAISGKFQQKK AEIKAQASSQGMQKRANFCHSDSQVLPTLSGLTIARKRRVLAQFDDLQECYLQK RRRWATQLRKQCDGGLRKERDGNSISREGYHAGLEEFQSILTTFTRYSRLRVIS ELRHGDLFHSANIVSSIEFDRDDELFATAGVSRRIKVFDFATVVNEPADVHCPV VEMSTRSKLSCLSWNKCIKSQIASSDYEGIVTVWDVNTRQSVMMYEEHEKRAWS VDFSRTEPTRLISGSDDGKVKVWCTRQETSVLNIDMKANICCVKYNPGSSYYVA VGSADHHIHYYDLRNPSVPLYEFNGHRKTVSYVKFISTNELASASTDSTLRLWD VRDNCLVRTFKGHTNEKNFVGLTVNSEYIACGSETNGVFVYHKAISKPAAWHQF GSPDLDDSDDDTSHFISAVCWKSESPTMLAANSQGTIKVLVLAP	68	2146
487	WD40 repeat protein	MANYVDSKKNFKCVPALQQFYTGGPFRLSSDGSFLVCACNDEVKVVDLATGSVK NTLEGDSELIVALALTPDNKYLFSASRSTQIKFWDLSSATCKRTWKAHNGPVAD MACDASGGLLATAGADRSILVWDVDGGYCTHSFRGHQGVVTTVIFHPDPHCLLL FSGSDDATVRIWDLVAKKCISVLEKHFSTVTSLAISENGWNLLSAGRDKVVNIW DLRDYHCRATIPTYEPLEAVCVLPTGSRLVSVMNQSRALPENRKKSGAAPVYFL TVGERGIVRIWYSEGALCLYEQKSSDAIISSDKDELKGGFVSAVLLPLTQGVMC VTADQRFLFYNLDESDEGKCDLKVSKRLIGYNEEIVDLKFLGDEEKFLAVATNL EQVRMYDLSSMTCVYELSGHTDIVLCLDTVVFSGHSLLASGSKDHTVRIWDTES KSCICVAAGHMGAVGAVAFSKKAKNFFVSGSSDRTIKVWSFASVLDFGGISKSI KLSSQAAVAAHDKDINSVAVAPNDSLICTGSQDRTARIWRLPDLVPVLVLRGHK RGVWCVEFSPVDQCVMTASGDKTIKIWALSDGSCLKTFEGHTASVLRASFLTRG TQFVSSGADGLLKLWTIKSNECIATFDQHEDKIWAMAVGKKTEMLATGGSDSLV NLWHDCTTTDEEEALLKEEEAALKDQELLNALADTDYVKAIQLAFELRRPYKLL NVFTELYSKGHAQDQIQKVIRELGNEELRLLLEYVREWNTKPKFAHVAQFVLFQ LFNVLPPKEIIEVQGISELLEGLIPYAQRHYSRIDRLMRSTFLLDYTLSSMSVL SPTETDLSSSNLLARTADPLHAQIDQFHPTHFPEPNLTPIQSLLDSGNTDSVEV TARRAKKKRVSGNDSEKTTVAEVKIGDMENAFDEPDVADQGSSRKHKPASSKKR KSIAVGNASIKRIASGNAVTIALQV	874	3705
488	WD40 repeat protein	MESSCSSMNSNRHSTEKRCLRPLQKQGASMNKHSSDRFIPARGSIDLDVARFMV TQKQKDNNDIHALSPSPSPSKKAYQKEMADTLLKNAGAADNNCRILSFNGKSST VSQGSQENVLANLSISRRARRYIPQSADRTLDAPDLLDDYYLNLLDWSSTNVLS TALGNTVYLWDASNSSISELLIADEEEGPVTSVSWAPDGSQIAVGLNNSVVQLW DSQSNKKLRALKGHHDRVGALSWNGPILTTGGLDGIIINHDVRTRDHIVQTYKG HTQEVCGLKWSPSGQQLASGGNDNLLYIWDKSMASHNPSSQYFHQLDEHCAAVK ALAWCPFQTNLLASGGGTSDGSIKFWNTQTGACLNTVDTHSQVCSLLWNRHERE LLSSHGLNQNQLTLWKYPSMVKITELTGHTARVLHMAQSPDGYTVASAAADETL KFWQVFGAPDASKKTKTKDTKGAFNMFHMHIR	360	1754

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
489	WD40 repeat protein	MLDEIVADEEEEFNIWKKNTPLLYDVVITHALEWPSLTVQWLPDRHQSPTKDYS LQKMIVGTHTSGDEPNYLMIAEVQMPLQYSEDGNVGGFESTEAKVHIIQQINHE GEVNRAQYMPQNSFIIATKTVSSDVYVFDYTKHSSNAPQERVCNPELILKGHTN EGYSLSWSPLKEGQLLSGSNDAQICFWDINAASGRKVVEAKQIFKVHEGAVEDV SWHLKHEYLFGSVGDDCHLLIWDTRTAAPNKPQHSVVAHESEVNSLAFNPFNEW LLATGSADKTVKLFDLRKLSCSLHTFSNHTEEVFQIEWSPMNETILASSGGDRR LMVWDLRRIGDEQTSEDAEDGPPELIFIHGGHTSKISDFSWNLHDDWLIASVAE DNILQIWQMAENIYHDDADIL	185	1384
490	WD40 repeat protein	MTKEDHGESRDEMGERMVNEEYKLWKKNTPFLYDLVITHALEWPSLTVQWLPPS CKQQQDIIKDDDIDHPNTQMVILGTHTSDNEPNYLILAEVQLHDGTEDEDGDGD VKRPQDKMKPGTSGGAMGKVRILQQINHQKEVNRARYMPQKPTIIATKTVNADV YVFDYSKHPSKPPQEGRCNPELRLQGHESEGYGLSWSPLKEGHLLSASDDAQIC LWDITAATKAPKVVEANQIFRYHDGPVEDVAWHAIHDHLFGSVGDDHHLLLWDI RNDSEKPLHIVEAHQAEVNCLAFNPFNEWIVATGSADRTVALHDIRKLDKVLHT CAHHMEEVFQIGWSPQNGAILASCGSDRRLMVWDLSRIGDEQNPEDAEEAPPEL LFIHGGHTSKISDFSWNPAEEWVIASVAEDNILQVWQMSEHIYNDDNDSPTA	241	1533
491	WD40 repeat protein	MAMAMGDENAADPVEEFNIWKKNTPFLYDLVITHALEWPSLTVQWLPDRHQSST ADYSLQKMIVGTHTSEDEPNYLMIAEVQIPLQNSEDNIIGGFESTEAKVQIIQK INHEGEVNKARYMPQNSFVIATKTVSSDVYVFDYSKHPSKAPQERVCNPELILK GHSNEGYGLSWSPLKEGYLLSGSNDAQICLWDINAAFGKKVLEANQIFKVHEGA VGDVSWHLKHEYLFGSVGDDCHLLIWDMRTAAPNKPQQSVIAHQSEVNSLAFNP FNEWLLATGSMDKTVKLFDLRKLSCSLHTFSNHTDQVFQIEWSPMNETILASSG ADRRLMVWDLARIGETPEDEEDGPPELLFVHGGHTSKISDFSWNLNDDRVIASV AEDNILQIWQMAENIYHDDEDML	230	1435
492	WD40 repeat protein	MGLFEPFRALGYITDGVPFAVQRRGIETFVTLSVGKAWQIYNCAKLIPVLVGPQ MDKKIRALACWRDFTFAATGHDIAVFRRAHQVATWSGHKAKVTLLLSFGQHVLS VDLEGCLFIWAVAEVNQNKPPIGQIQLGEKFSPSCIMHPDTYLNKVLIGSEEGT LQLWNVNTRKKLYEFKGWGSSIRCCVSSPALDVVGIGCSDGKIHVHNLRYDEEI VTFMHSTRGAVTALSFRTDGQPLLAAGGSSGVISIWNLEKKKLQSVIKDAHDSS VCSLHFFANEPVLMSSATDNSIKMWIFDTTDGEARLLKYRSGHSAPPMCIRYYG KGRHILSAGQDRAFRIFSVIQDQQSRELSQGHVGKRAKKLKVKDEEIKLPPVIA FDAAEIRERDWCNVVTCHLDDPCAYTWRLQNFVIGEHILKPCLEDPTPVKSCSI SACGNFAVLGTEGGWLERFNLQSGISRGTYIDIGEKRQCAHNGAVVGLACDATN TLLISGGYNGDIKVWDFKGRELKFRWEIEVPLIKIVYHPGNGILATAADDMILR LFDVTAMRLVRIFVGHMDRVTDLCFSGDGKWLLSSSMDGTIRVWDIISSRQLNA MHMDSAVTALSLSPGMDMLATTHVGHNGIYLWANRMIYSKATDIEPFISGKQVV KVSMPTVSSKRESEEGDEKRTIVAESNVNKSDVSGSLIGDSYSAQLTPELVTLA LLPKAQWQSLVNLDIIKMRNKPIEPPKKPEKAPFFLPSLPTLSGERIFIPSSMN GDGDQDETRNDKTVFEARGKKLGGESLSFMQLLQSCAKIKDFTTFTNYLKGLSP SAVDMELRLLQIVDNENISETEHSVELQGIGMLLDYFVNEVSCNNNFEFVQALI RLFLKIHGETIRCQVSLQEKARKLLEIQSSTWERLDTSFQNARCMITFLSSSQF	101	2857

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
493	WD40 repeat protein	MIAAVCWVPKGVAKVLPDSAEPPTQEEIQELLKCNVVAESDDNEDSDEESEEMD TETDKNTDAVAKALAAANALGSQSSDFQRQHKVDDIANGLKELDMDHYDDEDEG IDIFGSGSLGNCYYPANDMDPYLVEQDDDDEDEIEDMTIKPSDLIILSARNEDD VSHLEVWIYEEETEEGGSNMYVHHDIILPAFPLSLAWLDCNLKGGEKGNFVAVG TMQPEIELWDLDVLDEVEPAVVLGGAVKDEASGKTTKLKKKKKNKQAVNFKEGS HTDAVLGLAWNMEYRNVLASASADKSVKIWDIVAEKCEHTMQPHTDKVQAVAWN PNQATVLLSGSFDRSVIMMDMRAPTHSGIRWPVPADVESLAWDPHTDHSFMVSA EDGTVRGFDIRAAASTADFDGKPMFILHAHDKAVCAISYNPAAPSLLTTGSTDK MVKLWDITNNQPSCIASTNPNVGAVFSAAFSKNSPFLLATGGSKGILHVWDTLD NSEVARRFGKFRPQN	43	1548
494	WEE1-like protein	MIMDENEFCDIFSLRKRLCLLSSQEGEEEEELEAMSQLDAGEFTVTGNEEVVAI AEDDVNTGILSQDLFSSQDYCTPSQPQDSTDLDSKDKAPCPLSPVKSTIQRKRC RPELLSNPPDSIQFSFQRLERVRSEESIQSSSQQLARVRSEVSSSDDFKTPKIT ASGQKNYVSQSALALRARVMSPPCIKNPYLDENEELNEKIQRSTRRSPACVTPI QSGACLSRYRADFHELEEIGRGNFSRVYKALNRLDGCCYAVKCSQSELRLDTER KVALMEVQSLAALGPHKNIVGYHTAWFENDHLYIQMELCDHNLTTANDRGILRT DTDFLEAVYQIAQALEFIHGRGVAHLDVKPENIYVRDGTYKLGDFGRATLINGT LHVEEGDARYMSREILNDNYEHLDKVDMFSLGATFFELLMRKQYPGSGKRIDRD TEIKIPILPGFSIYFQKLLQDLVSNDPGKRPSAKDVLKNPIFNKVRGAKEV	206	1657
495	WD40 repeat protein	MLAPALEMEPVEPQSLKKLSFKSLKRALDLFSPVHGQIAPPDPESKKMRISYKL NFEYGGGSGSEDQVPKRKESGAAQNQGQQAAGASNALALPGPEGSKIPPMEKSQ NALTVGPSLRPQGLNDVGLHGKGTAIISASGSSDRNLSTSAIMERLPSRWPRPV WHPPWKNYRVISGHLGWVRSIAFDPSNQWFCTGSADRTIKIWDLASGRLKLTLT GHIEQIRGLAVSSKHTYMFSAGDDKQVKCWDLEQNKVIRSYHGHLSGVYCLALH PTIDILLTGGRDSVCRVWDIRSKMQIFALSGHDNTVCSVFARPTDPQVVTGSHD TTIKFWDLRHGKTMTTLTNHKKSVRAMAQHPKENCFASASADNIKKFQLPRGEF LHNMLSQQKTIINTMAVNEEGVMATGGDNGSLWFWDWKSGHNFQQAHTIVQPGS LESEAGIYALSYDLTGSRLVSCEADKTIKMWKEDELATPETHPLNFKPPKDIRR F	117	1580
496	WD40 repeat protein	MEEAAKEQSAGSGKPKLLRYGLRSAAKPKEDKKEEQLHQPPPPPPPQQQAAPAP APAATRSSTSGSAGGRDRRPQQQHAVDEKYARWKSLVPVLYDWLANHNLLWPSL SCRWGPQLEQATYKNRQRLYISEQTDGSVPNTLVIANCEVVKPRVAAAEHVSQF NEEARSPFIRKYKTIIHPGEVNRIRELPQNPNIVATHTDSPDVLIWDVESQPNR HAVYGATASRPNLILTGHQENAEFALAMCPAEPFVLSGGKDKTVVLWSIQDHIT ASATDQTTNKSPGSGGSIIKKTGEGNEETGNGPSVGPRGIYCGHEDTVEDVAFC PSTAQEFCSVGDDSCLILWDARIGTNPVAKVEKAHNGDLHCVDWNPHDNNLILT GSADNSVNMFDRRNLTSNGVGSPVYKFEGHKAAVLCVQWSPDKPSVFGSSAEDG LLNIWDYERVDKKVDRAPNAPAGLFFQHAGHRDKIVDFHWNTADPWTMVSVSDD CDTAGGGGTLQIWRMSDLIYRPEEEVLAELENFKAHVLECSKA	111	1700

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
497	WD40 repeat protein	MAKDEEEFRGEMEERLVNEEYKIWKKNTPFLYDLVITHALEWPSLTVQWLPDRE EPPGKDYSVQKMILGTHTSDNEPNYLMLAQVQLPLEDAENDARQYDDERGEIGG FGCANGKVQVIQQINHDGEVNRARYMPQNPFIIATKTVSAEVYVFDYSKHPSKP PQDGGCHPDLRLRGHNTEGYGLSWSPFKHGHLLSGSDDAQICLWDINVPAKNKV LEAQQIFKVHEGVVEDVAWHLRHEYLFGSVGDDRHLLIWDLRTSATNKPLHSVV AHQGEVNCLAFNPFNEWVLATGSADRTVKLFDLRKISSALHTFSCHKEEVFQIG WSPKNETILASCSADRRLMVWDLSRIDEFQTPEDALDGPPELLFIHGGHTSKIS DFSWNPCEDWVIASVAEDNILQIWQMAENIYHDEEDDMPPEEVV	144	1412
498	Cyclin-dependant kinase inhibitor	MGKYMRKGKGVGEVAVMEVSQGSLGVRTRARTLAAASSQKDHRRLGASKSVTTK HQSSAPPASPCVESSMHTCYLELRSRKLEKFSRCYHSAHGATSHGESKRSLSLS EPSRLAVSEEARVASDKSSHRVLQQQSSVAHSRNNSATFSHNAKPAKAAQRKER RDDDHTSARPSEAPHEDEDGMEVEASFGENVMDLDSRERRTRETTPSSYTRDVE TMETPGSTTRPPSNAGRRRFQTEGGHGTRNQFHVPTTNEIEEFFAGAEQQEQRR FTDRYNYDPVSDSPLPGRFEWVRLRP	793	1683
499	CDK type D	MQNMEENVQSSWSLHGNKEICARYEILKRVSSGTYLDVYRGRRKEDGLIVALKE VHDYQSSWREIEALQRLCGCPNVVRLYEVILEFLTSDLYSVIKSAKNKGENGIP EAEVKAWMIQILQGLANCHANWVIHRDLKPSNMLISAYGILKLADFGSMSFLKR AIYEVEYELPQEDILADAPGERLMDEDDSVKGVWNEGEEDSSTAVETNFDDMAE TANLDLSWKNEGDMVMQGFTSGVGTRWYRAPDFLYGATIYGKEIDLWSLGCILG ELLILEPLFSGTSNIDQLSRLVKVLGLQQKKNWPGCSNLPDYRKLCFPGDGSPV GLKNHVPNCSDNMFSILERLVCYDPAARLNAKEIVENKYFVEDPYPVLTHELRV PSPLREENNFSEDWAKWKDMEVDSDLENIDEFNVVHSSDGFCIKFS	415	2196
500	Histone acetyltransferase	MAPVKRIEPEKTKANEGKPKRRKVAFAIDTGIEANDCISLHLVSTPEEMRDAEG VEDQSLSFNPEYMQHFVGEHGKIYGYKGLKIDVWLNALSFHAYVDIQYESKVEE GKSEKEATDLTDIMKRIFGRGLVEDRNAFIQSFSSNSQSIESMIHNEGERIATR EILTDKGLSAQGDSERLGVSNEIFRLELSDPQIREWHARLEPLVLLFVEGSQPI EQDDPKWEMYIRVQRESLSGGSAVCRLLGFCTVYRFYHYPDTTRLRISQILVFP PYQGKGHGLLLLEAVNKTAVSRDSYDVTVEEPSESLQELRDCMDTIRLLSFEPV MPAVKSAVQKLKEANPSDKGAADHCLEGNVNNETVTTSSTKPKNKSGWFPPPGL VEEVRKHLKISKKQFKRCWEILLYLNLDRSDSQCEDKYHISLMEQIMSELFDKS SEKSAKGKRVIDIDNEYDNSKTFIMVRTRNPGNGEGFLPEALEGGMEVSQEDQL KSLFEERLEEIAQIAEKVPSLCKALQMP	109	1653
501	Histone deacetylase	MPEDRKKILEALAAKRKAEAESGEKKKRQKSSLNPAKPVSKPVSKPVGGIGSKG KSTSAPISSTKAKSKHKEEVKAKRVTKMDRYETDEDDESEEEEDLDSESDDDEL SDEDSEDDIKSKSVKKLPPQSKGKAPVKGISSSNGKGRDEKGKGIMKDKGKAKA KVEESSSDAEGDSDDDGGDLSDDPLQEVDPSNILPSKTRRRASQPTNYQFANMS GDDDDDDDSD	343	1023

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
502	Histone deacetylase	MADVPESLQQEKDEQGTDKNCCDGKFQKEIDIDDMEEEYNESSIDDEEENLSDN VATNNMGTIPQGQACMAVTVEGIEHANSVGCGRNGREGSEEVTAAEDMGHVSIE NIREQGRNRKSSEQLLALYEQEGLLEDDEDDDDVDWEPFEGVTVQMKWYCTNCT MANSDDSVHCDSCGEHRNSDILRQGFLASPYLPAESPSSSDVPDERLEESKCVM TTLTPSISPMIGVCCSSLQSERRTVVGFDERMLLHSEIQMETYPHPERPDRLRA IAASLRAAGLFPGKCFSIPAREATCEELQTIHSLEHVNAVESTSCGMLSHLSPD TYANEHSSLAARLAAGLCADLAKAIMTGQAQNGFALVRPPGHHAGVKDSMGFCL HNNAAIAVSASRVVGAKKVLIVDWDVHHGNGTQEIFEADQSVLYISLHRHGEGF YPGSGAVTEVGSSKGEGYSVNIPWKCGGVGDNDYIFAFQHAVLPIAEQFEPDLT IISAGFDAAKGDPLGRCEVTPDGFAHMAQMLSCLSKGKMLVILEGGYNLRSISA SATAVIKVLLGDNPKALPIDIQPSKGGLQTLLEVFEIQSKYWSSLKGHDQKLRS QWEAQYGSKKRKVIRKRHMHIVGGPVWWKWGRKRVVYYHWFARVSSRKHL	417	2351
503	Peptidylprolyl isomerase	MASGAGAAGVVEWHQKPPNPKNPVVFFDVTIGTIPAGRIKMELFADIVPRTAEN FRQFCTGEYRKAGIPIGYKGCHFHRVIKDFMIQAGDFVKGDGSGCISIYGSKFE DENFIAKHTGPGLLSMANSGPNTNGCQFFLTCAKCDWLDNKHVVFGRVLGEGLL VLRKIENVQTGQHNRPKLPCVIAECGEM	69	641
504	Peptidylprolyl isomerase	MAKLVSSVCAFSCQQRHPHSRPRFLSNRDHYNHYHNHSHYHNVCYFPPMMMMQQ QLQKQKRMTTKTITSLFKCNSSNHTLLKGLKEFMGFKFRLQAAMLSCEMSILGR VFAIFFIVHQAAAPFPFNHFDNWLVPPASAVLYSPNTKVPRTGEVALRKSIPAN PAMKSIQDFLEDIYYLLRFPQRKPYGTMEGDVKSALQIAINEKDSILGSVPLDM KERGLQLYNFLIDGQGGLQVLIEYIKEKDPDKVSVNLSSSLDTIAQLELLQAPG LPYLLPEEYQQYPRLNGRATIEFTMEKGDNSMFSVSSGGGLQKTATIQVVLDGY SAPLTAGNFTKLVIDGAYNGLKLKTTEQAVISDNERAEAGFNLPIEILPAGGFE PLYRTTLSVQDGELPVLPLSVYGAIAMAHNTISEDYSSPSQFFFYLYDKRNAGL GGLSFDEGQFSVFGYTTVGKEILPQLKTGDIIKSAKLVDGFDHLVLPSSST	172	1623
505	WD40 repeat protein	MDHYYQDDFDYLVDDEMVDFADDVEDDVRTRRRSDIDSDSENDFDSNNKSPDTT ALQAKRGKDIQGIPWNRLNFTREKYRETRLQQYKNYENLPRPRRSRNLDKECTN FERGSSFYDFRHNTRSVKATIVHFQLRNLVWATSKHNVYLMQNYSIMHWSSLKQ KGEEVLNVAGPIIPSVKHPGSSPQGLTRVQVSAMSVKDNLVVAGGFQGELICKY LDKPGVSFCTKISHDENGITNAVEIYNDASGATRLMTANNDLAVRVFDTEKFTV LERFSFPWSVNHTSVSPDGKLVAVLGDNADCLLADCKTGKTVGTLRGHLDYSFA AAWHPDGYILATGNQDTTCRLWDVRKLSSSLAVLKGRMGAIRSIRFSSDGRFMA MAEPADFVHLYDTRQNYTKSQEIDLFGEIAGISFSPDTEAFFVGVADRTYGSLL EFNRRRMNYYLDSIL	231	1768

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
506	WD40 repeat protein	MDCSGDEEEEQFFESLEEMLSPSDSGSEAADNETGCRNADARSKYEIWKRAPSS IQERRQRFLVRMGLANPSELGNQVNSTSAESTCSTETANIPNGIERLRENSGAV LRTAGSSGRKTHCKNVINIGLREGSVRSSSSSNGTPDVGEDNGEFGGTIFSRSG GTWECMCKIKNLDSGKEFVVDELGQDGLWNKLREVGTDRQLTMDEFERSLGLSP LVQELMRRESGVAQADCNGVHHHDAEISSSKRRSWLKALKSAAYSMRRPKEDQS NYDSERSGRRSGSFDVPWGKPQWTKVRHYRKRYKEFTALYMGQEIEAHEGSIWT MKFSLDGRYLASAGQDCVIHVREVIESMRTFGADTPDLYASSAYFSMNGLQELV PLSIEDHANKMKRGKIIGSKKSSNSDCIVLPNKVFQLSEEPVCSFHGHLLDVFD LSWSPSQYLLSSSMDKTVRLWKLGHESCLKVFSHNDIVTCIQFNPVDERYFISG SLDGKARIWSIPDRQVVDWSDLREMVTAVCYTPDGQGGLVGSIKGSCRFYNTSG NKLQLENQLNVRSKKKKSSGKKITGFQFAPGGDSQKVLITSADSRVRVYNGSEL VCKYKGFRNTCSQISASFAPNGQHFVCASEDSRVYIWNHESPRGSGARHEKSSW SHEHFLSQGVSVAIPWSGMKLQPPVWNSPEFMLGQRHNLLSLQGGKDVGCQNGL LSREAGEGQESETPLHYISQVSHSCGSQNMVDRDGQDDLSRYSACISDSRLSSF MAFPESPGNPDDLNSKVFFSDSSSKGSATWPEEKLPPTRKQSRSNSTSSHYDTL KTHLGNTIQGQSGASAAVAWGLVIVTAGHGGEIRSFQNYGLPVRL	376	2943
507	WD40 repeat protein	MPSIPAIGEFTVCEINRELLTTKDESDTQAKDAYAKILGLVFPPISFQIEEGFG SASRQQFDQDLDREDTIVTPSTSEGTNALQEGGLLLKGVSVLKNILASSFGPIF SPNDTKVLKKVELLQGISWHRHKHILAFISGSNQVTVHDFQDPEWRESSLLVSE SQRGIEALEWRPNGGTTLSVACRGGICIWSASYPGSVAPVRSGVASFLGTSTRG SSVRWTLVDFLQIPGGKAVTALSWSPTGRLLASASREDSSFTIWDVAQGVGTPL RRGLGGISLLKWSPTGDYLFSAKPNGTFYLWETNTWTLEQWSSSGGCVISATWG PDGRMLFMAFSESTTLGSLHFAGRPPSLDAHLLPMELPEIGSITGGFGNIEKMA WDGCGERLAVSYTGGDLMYVGLIAIYDTRRTPFISASLVGFIRGPGEQVKPLAF AFHDKFKQGPLLSVCWSSGLCCTYPLIFRAH	107	1498
508	WD40 repeat protein	MEEENAKHTEETRQVQVRFTTKLQPALRVPTTSIAIPAHLTRYGLSDIVNTLLG NDKPQPFDFLVESELVRTSLEKLLLIKGISAEKILNIEYILAVVPPKQEEPSLH DDWVSVVDGSYPNFIFSGSFDSIGRIWKGEGLCTHVLEGHRDAITSAAFIMPSD SSDSFINLATASKDRTLRLWQFKPNEHMTNGKMVRPYKLLKGHTSSVQTVSACP RRNLICSGSWDCSIKIWQTAGEMDIESNAGSVKKRKLEDSTEQIISQIEASRTL EGHSQCVSSVVWLEKDTIYSASWDHSVRSWDVETGVNSLTVGCRKALHCLSIGG EGSALIAAGGADSVLRIWDPRMPGTFTPILQLSSHKSWITACKWHPKSRHHLIS ASHDGTLKLWDVRSKVPLTTLEAHKDKVLCADWWKEDCVISGGADSTLQIFSNL NLT	118	1425
509	WD40 repeat protein	MNRLRSKRNHILELRLGQSEPEKEATLASNRSRGTNAPIVVEDDDDVVVSSPRS FALARSSVSQRSSRIPIVNEEDLELRLGLAVTGRTSAEHNPRRRHGRVPPNKPI VLCDDAGEADQSSSKKRRTGQQLSSDVQSDESKEVKLTCAICISTMEEETSTIC GHIFCKKCITNAIHRWKRCPTCRKKLAINNIHRIYISSSTG	186	797

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
510	WD40 repeat protein	MEEPPPPAVLPSSEDTSIVSSHSFVNAPPTVPVGLDASIPQISTPGINQPGLTI PVPPEAAPLTASLVAASAGMPPAVVPSFVRPAIVAHPSVMPPPSMPLAALPMPV ASAVPVAAPHFPPSTPNDNSITPSMPVPTPIVASSSVPPSVTIPGIAPLPFIAP IPVPSSRPVAPSPFMPPARPLGASVSVAMDVDNTDEQDQDADNKGESPSSSPDH PEDPSAAEYEITEESRKVRERQEQAIQELLLRRRAYALAVPTNDSSVRARLRRL NEPITLFGEREMERRDRLRALMAKLDAEGQLEKLMKVQEEEEAAANVDAEEVQE MEGPQVYPFYTEGSQELLKARTEITKFSLPRAVSRLQRARRKREDPDEDEDEEL KCVLQQSAQINMDCSEIGDDRPLSGCAFSSDGTLLATSAWSGVTKLWSVPNINK VATLKGHTERVTDVAFSPTNCHLATACADRTAMLWNSEGVLMKTYEGHLDRLAR LAFHPSGLYLGTASFDKTWRLWDVNTGIELLLQEGHSRSVYGIAFQCDGSLAAT CGLDGLARIWDLRTGRSILALEGHVKPVLGIDFSPNGYHLATGSEDHTCRIWDL RKRQSVYIIPAHSHLVSQVKFEPQEGYFLVTASYDSTAKVWSARDFKSIKVLAG HEAKVTSVDITADGQYIATVSHDRTIKLWSSKNSTNDMNIG	387	2456
511	WD40 repeat protein	MKRAYKLQEFVAHASNVNCLKIGKKSSRVLVTGGEDHKVNMWAIGKPNAILSLS GHSSAVESVTFDSAEALVVAGAASGTIKLWDLEEAKIVRTLTGHRSNCISVDFH PFGEFFASGSLDTNLKIWDIRRKGCIHTYKGHTRGVNSIRFSPDGRWVVSGGED NIVKLWDLTAGKLMHDFKCHEGQIQCMDFHPQEFLLATGSADRTVKFWDLETFE LIGSAGPETTGVRAMIFNPDGRTLLTGLHESLKVFSWEPLRCYDAVDVGWSKLA DLNIHEGKLLGCSYNQSCVGVWVVDISRVGPYAAGNVSRTNGHNEAKLASSGHP SVQQLDNNLKTNMARLSLSHSTESGIKEPKTTTSLTTTEGLSSTPQRAGIAFSS KNLPASSGPPSYVSTPKKNSTSRVQPTTNFQTLSRPDIVPVIVPRSNSLRPETT SDAKKEMNNFGRVVPSTVSTKSTDVIKSGSNRDESDKIDSINQKRMTGNDKTDL NIARAEQHVSSRLDNTNTSSVVCDGNQPAARWIGAAKFRRNSPVDPVVSPHDRS PTFPWSATDDGVTCQPDRQVTAPELSKRVVEPGRARALVASWETREKALTADTP VLVSGRPPTSPGVDMNSFIPRGSHGTSESDLTVSDDNSAIEELMQQHNAFTSIL QARLTKLQVIRRFWQRNDLKGAIDATGKMGDHSVSADVISVLIERSEIFTLDIC TVILPLLTRLLQSETDRHLTVAMETLLVLVKTFGDVIRATISATPTIGVDLQAE QRLERCNLCYVELENIKQILVPLIRRGGAVAKSAQELSLALQEV	359	2761
512	Cyclin B	MAGSDENNPGVVGGAHVQEGLRVGAGKMGAGNVQQRRALSNINSNIIGAPPYPC AVNKRVLSEKNVNSENDLLNAAHRPITRQFAAQMAYKQQLRPEENKRTTQSVSN PSKSEDCAILDVDDDKMADDFPVPMFVQHTEAMLEEIDRMEEVEMEDVAEEPVT DIDSGDKENQLAVVEYIDDLYMFYQKAEASSCVPPNYMDRQQDINERMRGILID WLIEVHYKFELMDETLYLTVNLIDRFLAVQPVVKKKLQLVGVTAMLLACKYEEV SVPVVEDLILISDRAYSRKEVLEMERLMVNTLHFNMSVPTPYVFMRRFLKAAQS DKKLELLSFFIIELSLVEYDMLKFPPSLLAASAIYTALSTITRTKQWSTTCEWH TSYSEEQLLECARLMVTFHQRAGSGKLTGVHRKYSTSKFGHAARTEPANFLLDF RL	238	1648
513	Cyclin-dependant kinase inhibitor	MQAPREGKSAAAIVGMGKYMKKSKAIPRDVSLLEASPRSPSATGVRTRAKTLAS RRLRRASQRRPPPPAAAAAAAAPSLDASPCPFSYLQLRSRRLRRPRLAPSPEAR IDEGPAGSGSRGSRDASCSARTASSSGGVEGEGACVGRGDRGNGGECVRDAAVD ASYGENDLEIEDRDRSTRESTPCSLIRDSNANTPPGSTTRQQSSCTAHRTQMSI LRSIPTSDEMEEFFAYAEQRQQRSFIEKYNFDIVKDRPLPGRFEWVQVIP	59	859

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
514	Histone acetyltransferase	MDGHSSHLAAQNRSRGSQTPSPSHSAASASATSSIHLKRKLSAANASAASAAAA AAAAAAAADDHAPPFPPSSISADTRDGALTSNDDLESISARGGGAGDDSDDDSD DEEEDDGDNDGGSSLRTFTAARLENVGPAAARNRKIKAESNATVKVEKEDSAKD GGNGAGVGALGPAATSGAGSGSGTVPKEDAVKIFTENLQASGAYSAREENLKRE EEAGRLKFECLSNDGVDDHMVWLIGLKNIFARQLPNMPKEYIVRLVMDRNHKSV MVIRRNLVVGGITYRPYASQKFGEIAFCAIKADEQVKGYGTRLMNHLKQHARDV DGLTHFLTYADNNAVGYFIKQGFTKEIYLDKDRWHGYIKDYDGGILMECKIDPK LPYTDLSTMVRRQRQAIDEKIRELSNCHIVYQGIDFQKRDAGVPQNTIKMEDIP GLREAGWTPDQWGYSRFRGLSDQKRLTFFIRQLLKVLNDHSDAWPFKEPVDARE VPDYYDIIKDPMDLKTMTKRVESEQYYVTLEMFIADVKRMFANARTYNSPDTIY FKIATRLEAHFQSKVQSNLQSGAGKIQQ	44	1829
515	Peptidylprolyl isomerase	MFNGMMDPELFKLAQEQMNRMSPAELAKIQQQMMSNPELMRMASESMKNMRPED LRQAAEQLKHVRPEEMAEIGEKMANASPEEIAAVRARADAQMTYEINAAKILKK EGNELHSQGRFKDASQKYLRAKNNLKGIPSSEGKNLLLACSLNLMSCYLKTRQY EECIKEGSEALACEEKNLKAFYRRGQAYRELGQLKDAVSDLRKAHEISPDDETI AQVLRDTEESLTKEGGSAPRGVVIEEITEEDETLASVNHESPSEYSEKRHQESE DAHKGPINGDIMGQMTNSESLKALKGDPDAIRSFQNFISNADPTTLAAMGAGNA GEVSPDLIKTASSMIGKMSAEELQKMIQLASSFPGENPYVTRNSDSNSNSFGNG SIPNVSPDMLKTASDMMSKMSPDDLQRMFEMASSSRGKDPSLDANHASSSSGAN LAANLNHILGESEPSSSYHIPSSSRNISSSPLSNFPSSPGDMQEQIRNQMKDPA MRQMFTSMMKNMSPEMMANMGKQFGLELSPEDAAKAQEAMSSLSPEMLDKMMRW ADRAQRGVETAKKTKNWLLGRPGMILAICMLLLAVILHRLGFIGS	109	1866
516	WD40 repeat protein	MIAAISWVPRGASKAVPEVAEPPSKEEIEEILKSGVVERSGDSDGEEDDENMDA VASEKADEVSTALSAADALGRISKVTKAGSGFEDIADGLRELDMDNYDEEDEDV KLFSTGLGDLYYPSNDMDPYLKDKDDDDDTEEIEDLSIKPMDSLIVCARTDDEV NLLEVYLLEPSLSDESNMYVHHEVVISEFPLCTAWLDCPIKGGDKGNFIAVGSM EPAIEIWDLDIIDAVEPCLVLGGQEELKKKKKKGKKASIKYKEGSHTDSVLGLA WNKEFRNILASASADRQVKIWDVAAGKCNITMEHHTDKVQAVAWNHHAPQVLLS GSFDHSVVMKDGRIPSHSGYRWSVTADVESLAWDPHSEHFFVVSLEDGTVRGFD VRAAISNSASQSLPSFTLHAHEKAVSTISYNPAAPNLLATGSTDKMVKLWDLSN NQPSCIASRNPKAGAVFSVSFSEDSPLLLAIGGSKGRLEVWDTSSDAAVSRRFG KHGKPKTAEPGS	212	1815
517	WD40 repeat protein	MKFCKKYQEYMQGQEGKKLPGLGFKKLKKILKRCRRRDSLHSQKALQAVQNPRT CPAHCSVCDGSFFPSLLEEMSAVLGCFNKQAQKLLELHLASGFQKYLMWFKGKL RGNHVALIQEGKDLVTYALINAIAIRKILKKYDKIHLSTQGQAFKSQVQRMHME ILQSPWLCELIAFHINVRETKANSGKGHALFEGCSLVVDDGKPSLSCELFDSIK LDIDLTCSICLDTVFDSVSLTCGHIYCYMCACSAASVTIVDGLKAAEPKEKCPL CREARVFEGAVHLDELNILLSRSCPEYWAERLQTERVERVRQAKEHWESQCRAF MGVE	207	1193

Protein SEQ ID	Target	Patent PEPTIDE Sequence	Patent ORF start	Patent ORF stop
518	WD40 repeat protein	MVSTQSTRENPSIFFPPPLKPWLLPVVLSLSLSRQLGMAAAAAASLPFKKNYRS SQALQQFYAGGPFAVSSDGSFIACNCGDSIKIVDSSNASLRPSIDCGSDTITAL SLSPDGKLLFSAGHSRQIRVWDLSTSTCLRSWKGHDGPVMSMACPVSGGLLATG GADRKVMVWDVDGGFCTHFFKGHDGVVSTVLFHPDSNRSLLFSGSDDGTIRVWD LLAKKCASTLRGHDSTVTSLAFSEDGLTLLAAGRDKVVSLWDLHNYACKKTIPM YEVLESVCVIHSGTVLASQLGLDDQLKVTKESAQNIHFITVGERGILRIWKSEG SVCLFKQEHSDVTVISDEDDSRSGFTAAVMLPLDQGLLCVTADQQFLFYYPEKH PEGIFSLTLCRRLVGYNEEIVDMKFLGEEENFLAVATNLEQVRVYELASMSCSY VLAGHTETVLCLDTCISSSGRTLIVTGSKDNSVRLWDSESRHCIGVGVGHMGAV GAVAFSRKRQDFFVSGSSDRTLKVWSLDGISEDGVDSTNLKAKAVVAAHDKDIN SVAVAPNDSLVCSGSQDRTACVWRLPDLVSVVVLKGHKRGIWSVEFSPVDQCVL TASGDKTVKIWAISDGSCLKTFEGHVSSVLRASFLTRGTQFVSCGADGLVKLWT VRTNECIATYDQHSDKVWALAVGKKTEMLATGGSDAVVNLWYDSTASDKEDAFR KEEEGVLKGQELENAVSDADYTKAIELALELRRPHKLFELFSELCRTREVGDRV ERILSALSGEEVCLLLEYIREWNAKPKLCHVAQSVLSQVFRILSPTEIVEIKGI GELLEGLIPYSQRHFSRIDRLVRSTYLLDYTLTGMSVIEPEADRSAVNDGSPDK SGLEKLEDGLLGENVGEEKIQNKEELESSAYKKRKLPRSKDRSKKKSKNVVYAD AAAISFRA	6	2786
519	WD40 repeat protein	MDSAPRRKSGGINLPSGMSETSLRLDGFSGSSSSFRAISNLTSPSKSSSISDRF IPCRSSSRLHTFGLVERGSPVKEGGNEAYSRLLKAELFGSDFGSLSPAGQGSPM SPSKNMLRFKTESSGPNSPFSPSILRQDSGFSSEASTPPKPPRKVPKTPHKVLD APSLQDDFYLNLVDWSSQNTLAVGLGTCVYLWSASNSKVTKLCDLGPNDGVCAV QWTREGSYISIGTSLGQVQIWDGTQCKRVRTMGGHQTRTGVLAWNSRILASGSR DRVILQHDLRVPNEFIGKLVGHKSEVCGLKWSHDDRELASGGNDNQLLVWNQHS QQPVLKLTEHTAAVKAIAWSPHQNGLLASGGGTADRCIRFWNTTNGHQTSSVDT GSQVCNLAWSKNVNELVSTHGYSQNQIMVWKYPSMAKVATLTGHSLRVLYLAMS PDGQTIVTGAGDETLRFWNVFPSAKAPAPVKDTGLWSLGRTHIR	213	1726
520	WD40 repeat protein	MEDEAEIYDGVRAQFPLTFGKQSKPQTSLESVHSATRRGGPAPAPAPASSSSLP STTSPSAAGGAGKSSGLPSLSSSSTAWLEGLRAGNPRAGREAGIGSRGGDGEDG GRAMIGPPRPPPGFSANDDGGGEDDDDDGDGVMVGPPPPPPGNLGDGDDDEEEE EAMIGPPRPPVVDSDEEEEEEEEENRYRLPLSNEIVLKGHNKIVSALAVDPTGS RVLSGSYDYTVRMFDFQGMNSRLSSFRDFEPVEGHQVRNLSWSPTADRFLCVTG SAQAKIYDRDGLTLGEFVKGDMYIRDLKNTKGHITGLTWGEWHPKTKETILTSS EDGSLRIWDVNDFKSQKQVIKPKLARPGRVPVTTCTWDREGKCIAGGIGDGSIQ IWNLKPGWGSRPDIHVEQAHADDITGLKFSSDGKILLTRSFDDSLKVWDLRLMK NPLKVFEDLPNHYAQTNIACSPDEQLFLTGTSVERESTIGGLLCFFDRSKLELV SRIGISPTCSVVQCAWHPRLNQIFATSGDKSQGGTHVLYDPTLSERGALVCVAR APRKKSVDDFELKPVIHNPHALPLFRDQPSRKRQREKILKDPLKSHKPELPMNG PGHGGRVGASKGSLLTQYLLKQGGMIKETWMDEDPREAILKHADAAEKNPKFTR AYAETQPDPVFAKSDSEDEDK	101	2110

[0396] Table 16. BLAST Sequence Alignment Table.

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
1	CDK type A	eucalyptusSpp_003910	Q9FRN5	PUTATIVE SERINE/THREO NINE KINASE	0	367	492
2	CDK type A	eucalyptusSpp_019213	O44000	CDC2-LIKE PROTEIN KINASE TPK2	e-160	217	290
3	CDK type A	eucalyptusSpp_036800	Q40789	PROTEIN KINASE P34CDC2	0	259	294
4	CDK type A	eucalyptusSpp_040260	Q27168	CDC2	e-156	208	304
5	CDK type A	eucalyptusSpp_041965	Q43361	CDC2PA mRNA. SPTREMBL	e-159	274	294
6	CDK type B-1	eucalyptusSpp_002906	Q9FYT9	Cyclin-dependent kinase B1-1	e-159	269	305
7	CDK type B-2	eucalyptusSpp_001518	Q9FSH4	B2-TYPE CYCLIN DEPENDENT KINASE	0	270	315

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
8	CDK type C	eucalyptusSpp_008078	Q9LDC1	CRK1 protein	0	415	558
9	CDK type C	eucalyptusSpp_009826	Q9LNN0	F8L10.9 protein. SPTREMBL	0	392	716
10	CDK type C	eucalyptusSpp_010364	Q8GZA7	Putative cyclin-dependent protein kinase.	e-172	309	499
11	CDK type C	eucalyptusSpp_011523	Q8W2N0	Cyclin-dependent kinase CDC2C	e-165	273	405
12	CDK type C	eucalyptusSpp_024358	P93320	CDC2MSC PROTEIN	0	448	523
13	CDK type C	eucalyptusSpp_039125	O80540	F14J9.26 protein	0	418	743
14	CDK type D	eucalyptusSpp_005362	O80345	CDK-activating kinase 1AT (Cdk-activating kinase CAK1At)	e-180	305	483
15	CDK type D	eucalyptusSpp_044857	O80345	CDK-activating kinase 1AT (Cdk-activating kinase CAK1At)	e-177	302	477

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
16	Cyclin A	eucalyptusSpp_001743	Q39879	MITOTIC CYCLIN A2-TYPE	0	360	508
17	Cyclin A	eucalyptusSpp_012405	Q39878	MITOTIC CYCLIN A2-TYPE	e-179	278	470
18	Cyclin B	eucalyptusSpp_003739	Q9LDM4	F2D10.10 (F5M15.6)	e-148	288	466
19	Cyclin B	eucalyptusSpp_022338	P93557	Mitotic cyclin	e-168	310	476
20	Cyclin B	eucalyptusSpp_028605	Q40337	B-like cyclin. SPTREMBL	e-158	300	439
21	Cyclin B	eucalyptusSpp_041006	Q40337	B-like cyclin	e-158	300	439
22	Cyclin D	eucalyptusSpp_006643	Q9SXN7	NtcycD3-1 protein	1E-73	177	404
23	Cyclin D	eucalyptusSpp_045338	Q8LK74	Cyclin D3.1 protein. SPTREMBL	e-101	190	332

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
24	Cyclin D	eucalyptusSpp_046486	Q9ZRX7	CYCLIN D3.2 PROTEIN	e-126	196	373
25	Cyclin-dependent kinase regulatory subunit	eucalyptusSpp_012070	CAB69358	SEQUENCE 1 FROM PATENT WO9841642	8E-64	83	88
26	Histone acetyltransf erase	eucalyptusSpp_006617	O80378	181 (Fragment)	0	371	395
27	Histone acetyltransf erase	eucalyptusSpp_007827	Q9FJT8	Histone acetyltransf erase HAT B	e-148	260	465
28	Histone acetyltransf erase	eucalyptusSpp_008036	Q9FJT8	Histone acetyltransf erase HAT B. SPTREMBL	e-149	262	465
30	Histone deacetylase	eucalyptusSpp_001596	Q9M4T5	Putative histone deacetylase HD2	7E-76	156	305
31	Histone deacetylase	eucalyptusSpp_005870	Q9M4T4	Putative histone deacetylase HD2c (AT5g03740/F 17C15_160)	7E-66	144	318
32	Histone deacetylase	eucalyptusSpp_006901	HDAC_ARATH	Histone deacetylase (HD)	0	405	499

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blastx overlap
33	Histone deacetylase	eucalyptusSpp_006902	AAM13152	HISTONE DEACETYLASE	0	427	499
34	Histone deacetylase	eucalyptusSpp_007440	Q8W508	HISTONE DEACETYLASE	0	369	428
35	Histone deacetylase	eucalyptusSpp_008994	Q8LD93	Histone deacetylase, putative	0	354	536
36	Histone deacetylase	eucalyptusSpp_024580	Q94EJ2	Atlg08460/T27G7_7 (HDA8). SPTREMBL	e-165	274	373
37	Histone deacetylase	eucalyptusSpp_037831	Q9FML2	Histone deacetylase. SPTREMBL	0	356	464
38	MAT1 CDK-activating kinase assembly factor	eucalyptusSpp_034958	Q8LES8	Hypothetical protein	4E-47	101	190
39	Peptidylprolyl isomerase	001209EGXC004488HT	TL40_SPIOL	Peptidyl-prolyl cis-trans isomerase, chloroplast precursor	0	329	392
40	Peptidylprolyl isomerase	010310EGXD012820HT	Q9FJL3	PEPTIDYLPROLYL ISOMERASE	0	453	579

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
41	Peptidylprol yl isomerase	010310EGXD013036HT	O82646	HYPOTHETICAL 57.1 KDA PROTEIN (EC 5.2.1.8)	0	302	521
42	Peptidylprol yl isomerase	010316EGXF999037HT	BAB39983	PUTATIVE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE, CHLOROPLAST	e-115	146	172
43	Peptidylprol yl isomerase	010324EGXF002118HT	AAK32894	AT5G13120/T1 9L5_80	e-122	179	264
44	Peptidylprol yl isomerase	011019EGKA001923HT	AAM14253	HYPOTHETICAL 20.3 KDA PROTEIN	e-108	146	188
45	Peptidylprol yl isomerase	eucalyptusSpp_000966	Q8L5T1	Peptidylprol yl isomerase (Cyclophilin) (EC 5.2.1.8)	1E-91	155	170
46	Peptidylprol yl isomerase	eucalyptusSpp_001037	Q8VX73	CYCLOPHILIN (EC 5.2.1.8)	e-120	155	169
47	Peptidylprol yl isomerase	eucalyptusSpp_004603	AAM14253	HYPOTHETICAL 20.3 KDA PROTEIN.	e-108	146	188
48	Peptidylprol yl isomerase	eucalyptusSpp_005465	Q9SP02	Cyclophilin ROC7 (EC 5.2.1.8) (AT5g58710/m zn1_160) (Pepti...	2E-93	172	204

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
49	Peptidylprol yl isomerase	eucalyptusSpp_006571	O49605	EC 5.2.1.8 (Cyclophilin -like protein) (Peptidyl- prolyl	9E-98	169	224
50	Peptidylprol yl isomerase	eucalyptusSpp_006786	Q93VG0	Cyclophilin (EC 5.2.1.8) (Peptidyl- prolyl cis- trans	5E-82	142	164
51	Peptidylprol yl isomerase	eucalyptusSpp_007057	Q38901	Cytosolic cyclophilin (EC 5.2.1.8) (Peptidyl- prolyl	3E-84	144	172
52	Peptidylprol yl isomerase	eucalyptusSpp_008670	Q9FJL3	PEPTIDYLPROL YL ISOMERASE	0	423	596
53	Peptidylprol yl isomerase	eucalyptusSpp_009137	Q9C566	Cyclophilin- 40 (EC 5.2.1.8) (Expressed protein)	e-168	285	361
54	Peptidylprol yl isomerase	eucalyptusSpp_010285	Q9LY75	Cyclophylin- like protein (EC 5.2.1.8) (Peptidyl- prolyl	e-160	345	658
55	Peptidylprol yl isomerase	eucalyptusSpp_010600	Q93YQ8	HYPOTHETICAL 50.1 KDA PROTEIN (FRAGMENT)	0	346	475
56	Peptidylprol yl isomerase	eucalyptusSpp_011551	Q9ZVG4	T2P11.13 PROTEIN	e-115	154	192

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
57	Peptidylprol yl isomerase	eucalyptusSpp_020743	Q8VXA5	PUTATIVE CYCLOSPORIN A-BINDING PROTEIN	e-125	161	172
58	Peptidylprol yl isomerase	eucalyptusSpp_023739	FK21_NEUCR	FK506-binding protein precursor (FKBP-21)	3E-49	74	112
60	Peptidylprol yl isomerase	eucalyptusSpp_031985	Q8L8W5	Cyclophilin-like protein (EC 5.2.1.8) (Peptidyl-prolyl	1E-82	155	229
61	Peptidylprol yl isomerase	eucalyptusSpp_032025	Q9LPC7	F22M8.7 protein (EC 5.2.1.8) (Peptidyl-prolyl cis-trans	1E-45	99	160
62	Peptidylprol yl isomerase	eucalyptusSpp_032173	Q8L8W5	Cyclophilin-like protein (EC 5.2.1.8) (Peptidyl-prolyl	4E-83	156	229
64	Retinoblasto ma related protein	eucalyptusSpp_009143	Q9SLZ4	Retinoblasto ma-related protein	0	704	1008
65	WD40 repeat protein	eucalyptusSpp_000349	AAK49947	TGF-BETA RECEPTOR-INTERACTING PROTEIN 1	0	291	326
66	WD40 repeat protein	eucalyptusSpp_000575	Q9LW17	WD-40 repeat protein-like (Expressed protein)	e-168	282	341

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
67	WD40 repeat protein	eucalyptusSpp_000804	GBLP_SOYBN	Guanine nucleotide-binding protein beta subunit-like	0	291	326
68	WD40 repeat protein	eucalyptusSpp_000805	GBLP_MEDSA	Guanine nucleotide-binding protein beta	e-171	291	327
69	WD40 repeat protein	eucalyptusSpp_000806	GBLP_MEDSA	Guanine nucleotide-binding protein beta subunit-like	e-171	291	327
70	WD40 repeat protein	eucalyptusSpp_002248	AAL86002	HYPOTHETICAL 43.8 KDA PROTEIN	0	261	388
71	WD40 repeat protein	eucalyptusSpp_003203	Q9SY00	Putative WD-repeat protein (AT4G02730/T 5J8_2)	e-144	236	317
72	WD40 repeat protein	eucalyptusSpp_003209	AAM14986	HYPOTHETICAL 32.6 KDA PROTEIN	e-160	259	302
73	WD40 repeat protein	eucalyptusSpp_004429	Q9SZQ5	HYPOTHETICAL 34.3 KDA PROTEIN	0	260	322
74	WD40 repeat protein	eucalyptusSpp_004607	AAC27402	EXPRESSED PROTEIN	0	253	356

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
75	WD40 repeat protein	eucalyptusSpp_004682	AAK00964	HYPOTHETICAL 35.3 KDA PROTEIN	0	264	313
76	WD40 repeat protein	eucalyptusSpp_005786	Q944S2	At2g47790/F17A22.18 (Expressed protein). SPTREMBL	e-155	264	396
77	WD40 repeat protein	eucalyptusSpp_005887	Q94AB4	AT3g13340/MDC11_13	0	332	446
78	WD40 repeat protein	eucalyptusSpp_005981	Q8L4X6	WD-repeat protein GhTTG2. SPTREMBL	0	315	348
79	WD40 repeat protein	eucalyptusSpp_006766	Q8L4M1	Putative WD-40 repeat protein	e-137	234	369
80	WD40 repeat protein	eucalyptusSpp_006769	Q9LJC6	RETINOBLASTOMA-BINDING PROTEIN-LIKE	0	372	566
81	WD40 repeat protein	eucalyptusSpp_006907	Q94C94	Hypothetical protein.	0	446	812
82	WD40 repeat protein	eucalyptusSpp_007518	Q93ZN5	AT4G00090/F6N15_8	0	311	436

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
83	WD40 repeat protein	eucalyptusSpp_007717	O82266	At2g47990 protein (Hypothetica l 58.9 kDa protein)	e-180	327	528
84	WD40 repeat protein	eucalyptusSpp_007718	Q8RWD8	Hypothetical protein. SPTREMBL	e-173	278	350
85	WD40 repeat protein	eucalyptusSpp_007741	Q8LA40	Putative WD-40 repeat protein, MSI2	e-158	269	409
86	WD40 repeat protein	eucalyptusSpp_007884	Q9FHY2	Similarity to unknown protein	e-149	316	765
87	WD40 repeat protein	eucalyptusSpp_008258	Q9LHN3	EMB|CAB63739 .1 (AT3G18860/M CB22_3)	0	524	758
88	WD40 repeat protein	eucalyptusSpp_008465	Q9FLS2	WD-repeat protein-like	0	366	460
89	WD40 repeat protein	eucalyptusSpp_008616	Q9LYK6	Hypothetical protein	e-148	252	321
90	WD40 repeat protein	eucalyptusSpp_008690	Q9SW94	G PROTEIN BETA SUBUNIT	0	326	376

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
91	WD40 repeat protein	eucalyptusSpp_008708	Q8L862	Hypothetical protein	e-167	297	487
92	WD40 repeat protein	eucalyptusSpp_008850	O22725	F11P17.7 protein. SPTREMBL	0	402	853
93	WD40 repeat protein	eucalyptusSpp_009072	Q9SAJ0	F23A5.2(form 2) (mRNA export protein, putative)	e-176	288	350
94	WD40 repeat protein	eucalyptusSpp_009465	Q9FLX9	NOTCHLESS PROTEIN HOMOLOG	0	384	475
95	WD40 repeat protein	eucalyptusSpp_009472	Q9SZA4	WD-REPEAT PROTEIN-LIKE PROTEIN	0	374	457
96	WD40 repeat protein	eucalyptusSpp_009550	Q9FKT5	Gb|AAF54217.1 (Hypothetical protein)	e-167	275	313
97	WD40 repeat protein	eucalyptusSpp_010284	O22466	WD-40 repeat protein MSI1	0	397	423
98	WD40 repeat protein	eucalyptusSpp_010595	Q94C94	Hypothetical protein	0	419	789

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
99	WD40 repeat protein	eucalyptusSpp_010657	Q94AH2	HYPOTHETICAL 33.1 KDA PROTEIN	0	243	298
100	WD40 repeat protein	eucalyptusSpp_012636	Q8L611	Hypothetical protein	0	756	1133
101	WD40 repeat protein	eucalyptusSpp_012748	AAD10151	PUTATIVE WD-40 REPEAT PROTEIN, MSI4	0	375	469
102	WD40 repeat protein	eucalyptusSpp_012879	Q8VZY6	FERTILIZATIO N-INDEPENDENT ENDOSPERM PROTEIN	0	291	377
103	WD40 repeat protein	eucalyptusSpp_015515	Q8LPI5	Putative WD-repeat protein. SPTREMBL	0	360	493
104	WD40 repeat protein	eucalyptusSpp_015724	O22607	WD-40 repeat protein MSI4	0	395	522
105	WD40 repeat protein	eucalyptusSpp_016167	Q93YS7	Putative WD-repeat membrane protein	0	663	917
106	WD40 repeat protein	eucalyptusSpp_016633	Q9SUY6	HYPOTHETICAL 43.8 KDA PROTEIN	e-174	240	384

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
107	WD40 repeat protein	eucalyptusSpp_017485	Q8RXC4	Hypothetical 144.7 kDa protein	0	650	1348
108	WD40 repeat protein	eucalyptusSpp_018007	O94289	WD repeat-containing protein	e-129	302	794
109	WD40 repeat protein	eucalyptusSpp_020775	Q8W403	Sec13p	e-150	242	304
110	WD40 repeat protein	eucalyptusSpp_023132	AAK52092	WD-40 REPEAT PROTEIN	0	458	515
111	WD40 repeat protein	eucalyptusSpp_023569	Q9XIJ3	T10O24.21. SPTREMBL	0	404	576
112	WD40 repeat protein	eucalyptusSpp_023611	Q8L4J2	Cleavage stimulation factor 50K chain (Cleavage stimulation	e-174	301	438
113	WD40 repeat protein	eucalyptusSpp_024934	Q94AB4	AT3g13340/MD C11_13. WD-repeat protein-like SPTREMBL	0	343	444
114	WD40 repeat protein	eucalyptusSpp_025546	O22212	Hypothetical 61.8 kDa Trp-Asp repeats containing protein	0	352	566

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
115	WD40 repeat protein	eucalyptusSpp_030134	Q9LVF2	Genomic DNA, chromosome 3, P1 clone: MIL23	0	677	946
116	WD40 repeat protein	eucalyptusSpp_031787	AAL91206	WD REPEAT PROTEIN-LIKE	0	264	329
117	WD40 repeat protein	eucalyptusSpp_034435	Q9SAJ0	F23A5.2(form 2) (mRNA export protein, putative). SPTREMBL	e-178	290	349
118	WD40 repeat protein	eucalyptusSpp_034452	Q94BR4	Hypothetical protein (Putative pre-mRNA splicing factor	0	381	525
119	WD40 repeat protein	eucalyptusSpp_035789	P93563	Guanine nucleotide-binding protein beta subunit	3E-88	171	356
120	WD40 repeat protein	eucalyptusSpp_035804	Q9FNN2	WD-repeat protein-like. SPTREMBL	0	356	589
121	WD40 repeat protein	eucalyptusSpp_043057	Q9LV35	WD40-repeat protein. SPTREMBL	0	472	610
122	WD40 repeat protein	eucalyptusSpp_046741	Q93VK1	AT4g28450/F2 009_130	0	363	452

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
123	WD40 repeat protein	eucalyptusSpp_047161	Q9ZUN8	Putative WD-40 repeat protein	0	350	473
124	CDK type A	pinusRadiata_001766	Q9M3W7	PUTATIVE CDC2-RELATED PROTEIN KINASE CRK2. 459 e-128	e-128	237	436
125	CDK type A	pinusRadiata_002927	Q9FRN5	PUTATIVE SERINE/THREO NINE KINASE	0	349	470
126	CDK type B-1	990309PRCA009171HT	Q9FYT8	Cyclin-dependent kinase B1-2	e-145	244	303
127	CDK type B-1	pinusRadiata_013714	Q9FYT8	CYCLIN-DEPENDENT KINASE B1-2	e-174	222	304
128	CDK type B-1	pinusRadiata_016332	Q9FYT8	CYCLIN-DEPENDENT KINASE B1-2	e-178	228	304
129	CDK type B-1	pinusRadiata_021677	Q9FYT8	CYCLIN-DEPENDENT KINASE B1-2	e-176	229	304
130	CDK type B-1	pinusRadiata_027562	Q9FYT8	Cyclin-dependent kinase B1-2	e-118	211	304

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
131	CDK type C	pinusRadiata_001504	Q9LNN0	F8L10.9 protein	0	434	790
132	CDK type C	pinusRadiata_015211	Q9LNN0	F8L10.9 protein	0	371	746
133	CDK type C	pinusRadiata_020421	P93320	Cdc2MsC protein	0	318	432
134	CDK type D	pinusRadiata_003187	O80345	CDK-ACTIVATING KINASE 1AT (CDK-ACTIVATING KINASE CAK1AT)	e-137	226	485
135	CDK type D	pinusRadiata_015661	Q947K6	CDK-ACTIVATING KINASE.	0	266	407
136	Cyclin A	pinusRadiata_013874	Q96226	Cyclin	e-108	223	474
137	Cyclin A	pinusRadiata_014615	CAC27333	PUTATIVE A-LIKE CYCLIN (FRAGMENT)	0	332	390
138	Cyclin B	pinusRadiata_004578	O65064	Probable G2/mitotic-specific cyclin (Fragment)	9E-87	162	217

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overlap
139	Cyclin B	pinusRadiata_023387	O04389	B-like cyclin	2E-98	220	466
140	Cyclin D	pinusRadiata_006970	P93103	CYCLIN-D LIKE PROTEIN	1E-75	135	293
141	Cyclin D	pinusRadiata_010322	CAC17049	SEQUENCE 33 FROM PATENT WO0065040	e-131	171	254
142	Cyclin D	pinusRadiata_022721	P93103	CYCLIN-D LIKE PROTEIN	1E-76	137	289
143	Cyclin D	pinusRadiata_023407	Q9SMD5	CYCD3,2 PROTEIN	8E-90	139	278
144	Cyclin-dependent kinase regulatory subunit	pinusRadiata_001945	Q947Y1	PUTATIVE CYCLIN-DEPENDENT KINASE REGULATORY SUBUNIT	5E-55	74	86
145	Cyclin-dependent kinase regulatory subunit	pinusRadiata_008233	CAB69358	SEQUENCE 1 FROM PATENT WO9841642	4E-49	65	86
146	Cyclin-dependent kinase regulatory subunit	pinusRadiata_008234	CAB69358	SEQUENCE 1 FROM PATENT WO9841642	4E-49	65	86

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
147	Cyclin-dependent kinase regulatory subunit	pinusRadiata_022054	CAB69358	SEQUENCE 1 FROM PATENT WO9841642	8E-55	70	82
148	Histone acetyltransf erase	pinusRadiata_012137	Q9FK40	Histone acetyltransf erase (AT5g50320/M XI22_3)	0	496	555
149	Histone acetyltransf erase	pinusRadiata_012582	O80378	181 (Fragment) SPTREMBL	0	354	402
150	Histone acetyltransf erase	pinusRadiata_015285	O80378	181 (Fragment)	0	342	401
151	Histone acetyltransf erase	pinusRadiata_017229	Q9LNC4	F9P14.9 protein	e-118	268	585
152	Histone acetyltransf erase	pinusRadiata_020724	Q9AR19	Histone acetyltransf erase GCN5 (Expressed protein)	e-177	355	639
153	Histone deacetylase	pinusRadiata_004555	AAM13152	HISTONE DEACETYLASE	0	331	488
154	Histone deacetylase	pinusRadiata_004556	AAM13152	HISTONE DEACETYLASE	0	331	488

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
155	Histone deacetylase	pinusRadiata_005729	Q9M4U5	Histone deacetylase 2 isoform b	9E-62	154	348
156	Histone deacetylase	pinusRadiata_007395	AAM13152	HISTONE DEACETYLASE	0	335	426
157	Histone deacetylase	pinusRadiata_009503	Q8W508	Histone deacetylase	0	365	427
158	Histone deacetylase	pinusRadiata_011283	AAM19887	AT1G08460/T2 7G7_7	0	255	366
159	Histone deacetylase	pinusRadiata_012322	Q9FML2	HISTONE DEACETYLASE (PUTATIVE HISTONE DEACETYLASE)	0	327	435
161	Histone deacetylase	pinusRadiata_023236	Q8RX28	Putative histone deacetylase	e-144	238	390
162	Peptidylprol yl isomerase	pinusRadiata_000171	Q9FJL3	PEPTIDYLPROL YL ISOMERASE	0	364	549
163	Peptidylprol yl isomerase	pinusRadiata_000172	Q38949	FK506 BINDING PROTEIN FKBP62 (ROF1)	0	365	552

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
164	Peptidylprol yl isomerase	pinusRadiata_001480	Q8VXA5	PUTATIVE CYCLOSPORIN A-BINDING PROTEIN	e-125	161	172
168	Peptidylprol yl isomerase	pinusRadiata_001692	FKB7_WHEAT	70 kDa peptidylprol yl isomerase (EC 5.2.1.8)	0	418	553
169	Peptidylprol yl isomerase	pinusRadiata_005313	AAB64339	FKBP-TYPE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE	1E-97	135	175
170	Peptidylprol yl isomerase	pinusRadiata_006362	BAB39983	PUTATIVE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE, CHLOROPLA... 290 3e-77	3E-77	129	168
171	Peptidylprol yl isomerase	pinusRadiata_006493	Q9C835	Hypothetical 26.4 kDa protein (EC 5.2.1.8) (Peptidyl-prol...	2E-62	128	235
172	Peptidylprol yl isomerase	pinusRadiata_006983	AAK96784	CYCLOPHILIN	e-103	151	204
174	Peptidylprol yl isomerase	pinusRadiata_007665	Q9LDC0	FKBP-like protein (Genomic DNA, chromosome 3, P1 clone:	e-138	239	378
175	Peptidylprol yl isomerase	pinusRadiata_012196	Q93VG0	Cyclophilin (EC 5.2.1.8) (Peptidyl-prolyl cis-trans	4E-74	132	160

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overlap
176	Peptidylprolyl isomerase	pinusRadiata_013382	Q9C588	HYPOTHETICAL 60.2 KDA PROTEIN	0	288	581
177	Peptidylprolyl isomerase	pinusRadiata_016461	O04287	IMMUNOPHILIN	9E-66	88	109
178	Peptidylprolyl isomerase	pinusRadiata_017611	Q9C566	Cyclophilin-40 (EC 5.2.1.8) (Expressed protein)	e-163	276	360
179	Peptidylprolyl isomerase	pinusRadiata_019776	AAM14253	HYPOTHETICAL 20.3 KDA PROTEIN	e-110	146	190
180	Peptidylprolyl isomerase	pinusRadiata_020659	AAO63961	Hypothetical protein SPTREMBL	7E-85	159	227
181	Peptidylprolyl isomerase	pinusRadiata_022559	AAK43974	PUTATIVE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE	2E-73	113	153
182	Peptidylprolyl isomerase	pinusRadiata_024188	Q9P3X9	PEPTIDYL-PROLYL CIS-TRANS ISOMERASE (EC 5.2.1.8)	e-122	210	379
183	Peptidylprolyl isomerase	pinusRadiata_027973	Q9SR70	T22K18.11 protein (AT3g10060/T22K18_11)	3E-69	125	171

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
184	WD40 repeat protein	pinusRadiata_001353	Q9FNN2	WD-repeat protein-likeSPTREMBL	0	317	590
185	WD40 repeat protein	pinusRadiata_001978	PRL1_ARATH	PP1/PP2A phosphatases pleiotropic regulator PRL1	0	341	502
186	WD40 repeat protein	pinusRadiata_002810	AAK49947	TGF-BETA RECEPTOR-INTERACTING PROTEIN 1	0	273	326
187	WD40 repeat protein	pinusRadiata_002811	AAK49947	TGF-BETA RECEPTOR-INTERACTING PROTEIN 1	0	273	326
188	WD40 repeat protein	pinusRadiata_002812	AAM15129	HYPOTHETICAL 58.9 KDA PROTEIN	e-127	225	521
189	WD40 repeat protein	pinusRadiata_003514	Q9FJ94	Similarity to myosin heavy chain kinaseSPTREMBL	e-137	242	445
190	WD40 repeat protein	pinusRadiata_004104	GBB_ORYSA	Guanine nucleotide-binding protein beta subunit	0	294	378
191	WD40 repeat protein	pinusRadiata_005595	Q9FTT9	PUTATIVE DKFZP564O0463 PROTEIN	0	320	459

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overlap
192	WD40 repeat protein	pinusRadiata_005754	Q94JT6	At1g78070/F28K19_28SPTREMBL	e-168	294	451
193	WD40 repeat protein	pinusRadiata_006463	GBLP_MEDSA	Guanine nucleotide-binding protein beta subunit-like ... 538 e-152	e-152	261	324
194	WD40 repeat protein	pinusRadiata_006665	AAM20553	HYPOTHETICAL 119.9 KDA PROTEIN. 1229 0.0	0	655	1169
195	WD40 repeat protein	pinusRadiata_006750	AAM13119	HYPOTHETICAL 35.4 KDA PROTEIN. 560 e-158	e-158	264	312
196	WD40 repeat protein	pinusRadiata_007030	Q9LJN8	MITOTIC CHECKPOINT PROTEIN. 595 e-169	e-169	284	335
197	WD40 repeat protein	pinusRadiata_007854	Q8H919	Putative WD domain containing protein	0	429	644
198	WD40 repeat protein	pinusRadiata_007917	AAD10151	PUTATIVE WD-40 REPEAT PROTEIN, MSI4	0	353	462
199	WD40 repeat protein	pinusRadiata_007989	Q9LRZ0	Genomic DNA, chromosome 3, TAC clone:K20I9	0	480	687

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
200	WD40 repeat protein	pinusRadiata_008506	MSI1_LYCES	WD-40 repeat protein MSI1	0	364	420
201	WD40 repeat protein	pinusRadiata_008692	Q8W403	Sec13p	e-134	218	301
202	WD40 repeat protein	pinusRadiata_008693	Q8W403	Sec13p	e-137	222	301
203	WD40 repeat protein	pinusRadiata_009170	Q9M0V4	U3 snoRNP-associated-like protein. SPTREMBL	e-127	244	524
204	WD40 repeat protein	pinusRadiata_009408	Q9SAJ0	F23A5.2(FORM 2). 602 e-171	e-171	282	350
205	WD40 repeat protein	pinusRadiata_009522	Q8RXQ4	Hypothetical 43.8 kDa protein	e-129	231	395
206	WD40 repeat protein	pinusRadiata_009734	AAO27452	Peroxisomal targeting signal type 2 receptor. SPTREMBL	e-142	227	317
207	WD40 repeat protein	pinusRadiata_009815	AAM20433	CELL CYCLE SWITCH PROTEIN	0	326	500

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
208	WD40 repeat protein	pinusRadiata_010670	AAN72058	Expressed protein	e-157	264	345
209	WD40 repeat protein	pinusRadiata_011297	AAM13100	WD REPEAT PROTEIN ATAN11	e-157	262	337
210	WD40 repeat protein	pinusRadiata_013098	AAM13153	HYPOTHETICAL 39.1 KDA PROTEIN. 487 e-136	e-136	229	352
211	WD40 repeat protein	pinusRadiata_013172	Q8H0T9	Hypothetical protein	0	437	860
212	WD40 repeat protein	pinusRadiata_013589	AAK52092	WD-40 REPEAT PROTEIN	0	448	512
213	WD40 repeat protein	pinusRadiata_013608	AAC27402	EXPRESSED PROTEIN	e-141	202	358
214	WD40 repeat protein	pinusRadiata_014299	Q9XED5	Cell cycle switch proteinSPTRE MBL	0	335	488
215	WD40 repeat protein	pinusRadiata_014498	Q9FH64	WD REPEAT PROTEIN-LIKE	e-152	206	329

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
216	WD40 repeat protein	pinusRadiata_014548	Q93ZS6	HYPOTHETICAL 82.2 KDA PROTEIN	0	505	763
217	WD40 repeat protein	pinusRadiata_014610	Q9M298	Hypothetical 104.7 kDa protein	0	450	922
218	WD40 repeat protein	pinusRadiata_016090	Q9SIY9	Putative WD-40 repeat proteinSPTREMBL	0	442	802
219	WD40 repeat protein	pinusRadiata_016722	O22826	Putative splicing factorSPTREMBL	e-159	257	310
220	WD40 repeat protein	pinusRadiata_016785	AAG60193	PUTATIVE WD40 PROTEIN	0	344	464
221	WD40 repeat protein	pinusRadiata_017094	Q9LV35	WD40-REPEAT PROTEIN	0	406	604
222	WD40 repeat protein	pinusRadiata_017527	Q9AYE4	Hypothetical 35.3 kDa protein	e-154	254	314
223	WD40 repeat protein	pinusRadiata_017591	O80706	F8K4.21 protein	0	905	1218

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
224	WD40 repeat protein	pinusRadiata_017769	Q9XIJ3	T10O24.21	0	446	607
225	WD40 repeat protein	pinusRadiata_018047	Q8VZY6	FERTILIZATIO N-INDEPENDENT ENDOSPERM PROTEIN	0	285	373
226	WD40 repeat protein	pinusRadiata_018414	Q947M8	COPI	0	455	638
227	WD40 repeat protein	pinusRadiata_018986	Q9LFE2	WD40-repeat protein	0	518	886
228	WD40 repeat protein	pinusRadiata_019479	Q9SZA4	WD-repeat protein-like protein	e-156	276	454
229	WD40 repeat protein	pinusRadiata_020144	Q8W514	MSI TYPE NUCLEOSOME/C HROMATIN ASSEMBLY FACTOR C	0	288	413
230	WD40 repeat protein	pinusRadiata_022480	Q8W514	MSI type nucleosome/c hromatin assembly factor C	e-167	287	426
231	WD40 repeat protein	pinusRadiata_023079	Q8W514	MSI type nucleosome/c hromatin assembly factor C. SPTREMBL	e-169	283	397

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
232	WD40 repeat protein	pinusRadiata_026739	Q93YS7	Putative WD-repeat membrane protein. SPTREMBL	0	591	918
233	WD40 repeat protein	pinusRadiata_026951	Q93VS5	AT4g18900/F13C5_70 (Hypothetical protein)	e-163	290	503
234	WEE1-like protein	pinusRadiata_026529	Q9SRY9	F22D16.3 PROTEIN	e-122	209	451
235	WD40 repeat protein	eucalyptusSpp_006366	Q8LF96	PRL1 protein	0	374	492
236	WD40 repeat protein	eucalyptusSpp_017378	O22607	WD-40 repeat protein MSI4	0	371	453
237	WD40 repeat protein	pinusRadiata_000888	O22466	WD-40 repeat protein MSI1	0	364	420
238	Cyclin-dependant kinase inhibitor	pinusRadiata_014166	Q9FKB5	GENOMIC DNA, CHROMOSOME 5, TAC CLONE:K24G6 (CYCLIN-DEPENDENT	5E-42	114	304
239	CDK type D	pinusRadiata_003189	Q9M5G4	CDK-activating kinase	8E-21	56	100

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
240	Histone acetyltransf erase	pinusRadiata_009356	Q9FJT8	Histone acetyltransf erase HAT B	7E-85	187	510
241	Histone deacetylase	pinusRadiata_000065	Q9LPW6	F13K23.8 protein.	5E-18	71	209
242	Histone deacetylase	pinusRadiata_014197	Q8GXJ1	Putative histone deacetylase	e-170	308	519
243	Peptidylprol yl isomerase	pinusRadiata_009081	Q9ZRQ9	Cyclophilin (EC 5.2.1.8) (Peptidyl-prolyl cis-trans	e-106	185	190
244	Peptidylprol yl isomerase	pinusRadiata_013417	Q8H4T0	Putative peptidyl-prolycis-trans isomerase protein	e-140	235	345
245	WD40 repeat protein	pinusRadiata_005755	Q9SKW4	F5J5.6.	e-143	144	319
246	WD40 repeat protein	pinusRadiata_006670	Q9LDG7	WD-40 repeat protein-like (MJK13.13 protein)	e-163	393	960
247	WD40 repeat protein	pinusRadiata_007027	Q8GWR1	Hypothetical protein.	e-157	276	470

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
248	WD40 repeat protein	pinusRadiata_007276	Q9LF27	Hypothetical 47.3 kDa protein	e-138	235	428
249	WD40 repeat protein	pinusRadiata_007390	Q94AH4	PUTATIVE RING ZINC FINGER PROTEIN. 91 3e-17	3E-17	53	158
250	WD40 repeat protein	pinusRadiata_012648	O22212	Hypothetical 61.8 kDa Trp-Asp repeats containing protein	0	324	561
251	WD40 repeat protein	pinusRadiata_013171	Q8H0T9	Hypothetical protein.	0	437	860
252	Cyclin B	eucalyptusSpp_045414	Q9LDM4	F2D10.10 (F5M15.6)	e-142	255	423
253	Cyclin-dependant kinase inhibitor	eucalyptusSpp_044328	Q9FKB5	GENOMIC DNA, CHROMOSOME 5, TAC CLONE:K24G6 (CYCLIN-DEPENDENT	1E-54	121	260
254	Histone acetyltransf erase	eucalyptusSpp_015615	Q9AR19	Histone acetyltransf erase GCN5 (Expressed protein)	0	390	563
255	Peptidylprol yl isomerase	eucalyptusSpp_017239	Q8GWM6	Hypothetical protein	0	364	591

SEQ ID	Target	Patent Identifier	BlastX top hit	Gene name	BlastX e value	BlastX identities	Blast X overl ap
256	WD40 repeat protein	eucalyptusSpp_018643	Q93VS5	AT4g18900/F1 3C5_70 (Hypothetica l protein)	0	229	327
257	WD40 repeat protein	eucalyptusSpp_019127	Q9SRX9	F22D16.14 protein. SPTREMBL	e-131	232	337
258	WD40 repeat protein	eucalyptusSpp_022624	Q9LFE2	WD40-repeat protein	0	594	868
259	WD40 repeat protein	eucalyptusSpp_032424	Q8LPL5	Cell cycle switch protein	0	255	327
260	WD40 repeat protein	eucalyptusSpp_037472	Q9SK69	Putative WD-40 repeat protein (AT2G20330/F 11A3.12)	0	461	677

2004311384 02 Mar 2011

Interpretation of this specification

It will therefore be understood that the invention could take many forms and be put to many different uses. All such forms and uses are embodied within the spirit and scope of the invention, which is to be understood as not being limited to the particular details of the embodiments or the examples discussed previously, but which extends to each novel feature and combination of features disclosed in or evident from this specification (including in the accompanying claims and drawings). All of these different combinations constitute various alternative aspects of the invention.

Throughout this specification, unless the context requires otherwise, the word "comprise" is (and variants, variations or other forms of that word, such as "comprises" or "comprising" are) to be understood as implying the inclusion of a stated element, feature or integer or group of elements, features or integers, but not the exclusion of any other element, feature or integer or group of elements, features or integers. Further, wherever used in this specification, the term "includes" is not a term of limitation, and is not be taken as excluding the presence of other any element, feature or integer or group of elements, features or integers.

It is further to be understood that any discussion in this specification of background or prior art documents, devices, acts, information, knowledge or use ('**Background Information**') is included solely to explain the context of the invention. Any discussions of such Background Information is not be taken as an admission in any jurisdiction that any such Background Information constitutes prior art, part of the prior art base or the common general knowledge in the field of the invention on or before the priority date of the appended claims or any amended claims later introduced into this specification.

2004311384 02 Mar 2011

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A DNA construct comprising a polynucleotide which comprises:

   (i) the sequence of SEQ ID NO. 136 operably linked in a sense or an antisense orientation to a promoter; or

   (ii) a sequence that encodes the protein sequence of SEQ ID NO. 396, operably linked in sense or antisense orientation to a promoter,

   wherein the promoter in (i) or (ii) is selected from the group consisting of:

   (a) a constitutive promoter;

   (b) an inducible promoter;

   (c) a regulatable promoter;

   (d) a temporally regulated promoter, and

   (e) a tissue-preferred promoter.

2. A DNA construct as claimed in paragraph (ii) of claim 1, in which the sequence that encodes the protein sequence of SEQ ID NO. 396 is operably linked in a sense orientation to the promoter.

3. A DNA construct as claimed in paragraph (ii) of claim 1, in which the sequence that encodes the protein sequence of SEQ ID NO. 396 is operably linked in an antisense orientation to the promoter.

4. A DNA construct as claimed in any one of claims 1 to 3, in which an RNA transcript of the polynucleotide is complementary to a nucleic acid sequence of SEQ ID NO. 136.

5. A plant cell comprising the DNA construct of any one of claims 1 to 4.

6. A plant cell as claimed in claim 5, in which the plant cell is in a plant, and where the phenotype of the plant is different from a plant of the

same species which does not comprise the plant cell, wherein the difference in phenotype is in any one or more of:

(a) lignin quality;

(b) lignin structure;

(c) wood composition;

(d) wood appearance;

(e) wood density;

(f) wood strength;

(g) wood stiffness;

(h) cellulose polymerization;

(i) fibre dimensions;

(j) lumen size;

(k) other plant components;

(l) plant cell division;

(m) plant cell development;

(n) number of cells per unit area;

(o) cell size;

(p) cell shape;

(q) cell wall composition;

(r) rate of wood formation;

(s) formation of stem defects;

(t) average microfibril angle;

(u) width of the S2 cell wall layer;

(v) rate of growth;

(w) rate of root formation ratio of root to branch vegetative development;

(x) leaf area index; and

2004311384 02 Mar 2011

(y) leaf shape.

7. A transgenic plant containing at least one plant cell according to either of claims 5 or 6, where the plant is of a species of *Eucalyptus* or *Pinus*.

8. A transgenic plant as claimed in claim 7, in which the plant exhibits one or more traits selected from the group consisting of:

(a) increased drought tolerance;

(b) herbicide resistance;

(c) reduced or increased height;

(d) reduced or increased branching;

(e) enhanced cold and frost tolerance;

(f) improved vigour;

(g) enhanced colour;

(h) enhanced health and nutritional characteristics;

(i) improved storage;

(j) enhanced yield;

(k) enhanced salt tolerance;

(l) enhanced resistance of the wood to decay;

(m) enhanced resistance to fungal diseases;

(n) altered attractiveness to insect pests;

(o) enhanced heavy metal tolerance;

(p) increased disease tolerance;

(q) increased insect tolerance;

(r) increased water-stress tolerance;

(s) enhanced sweetness, improved texture;

(t) decreased phosphate content;

(u) increased germination;

2004311384 11 Dec 2011

(v) increased micronutrient uptake;

(w) improved starch composition;

(y) improved flower longevity;

(z) production of novel resins;

(za) production of novel proteins or peptides;

(zb) reduced period of juvenility;

(zc) an increased period of juvenility;

(zd) propensity to form reaction wood;

(ze) self-abscising branches; and

(zf) accelerated reproductive development or delayed reproductive development as compared to a plant of the same species that has not been transformed with the DNA construct.

9. A DNA construct as claimed in paragraph (i) of claim 1 in which the sequence of SEQ ID NO: 136 is operably linked in a sense orientation to the promoter.

10. A DNA construct as claimed in paragraph (i) of claim 1 in which the sequence of SEQ ID NO: 136 is operably linked in an antisense orientation to the promoter.

11 A DNA construct as claimed in any one of claims 1 to 4, 9 or 10, substantially as described in this specification, and with reference to the examples given.

12. A plant cell as claimed in either of claims 5 or 6, substantially as described in this specification, and with reference to the examples given.

13. A transgenic plant as claimed in either of claims 7 or 8, substantially as described in this specification, and with reference to the examples given.

2004311384 11 Dec 2011

14. A DNA construct as claimed in either of claims 9 or 10, in which an RNA transcript of the polynucleotide is complementary to a nucleic acid sequence of SEQ ID NO. 136.

15. A plant cell comprising the DNA construct of either of claims 9 or 10.

Dated: 10 December 2011

ARBORGEN, LLC

By its Patent Attorneys

KNIGHTSBRIDGE PATENT ATTORNEYS

FIGURE 1. Exemplary Microarray Sampling Parameters.

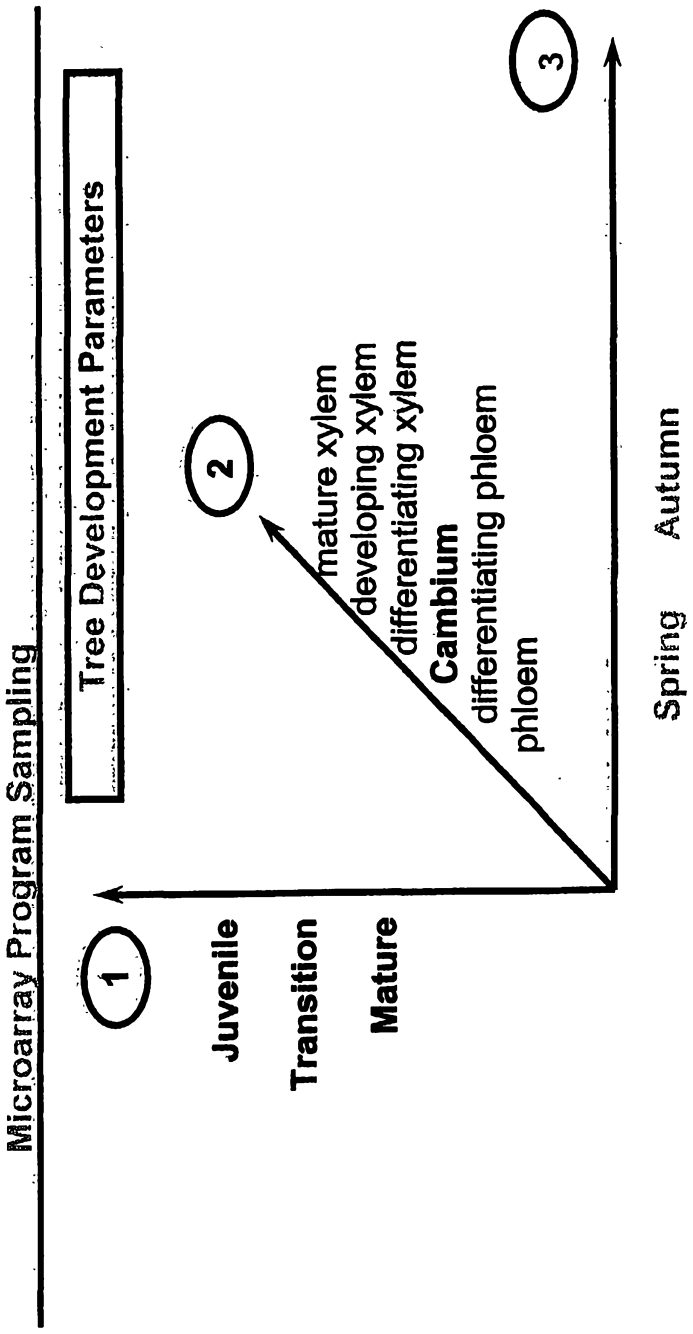


FIGURE 2

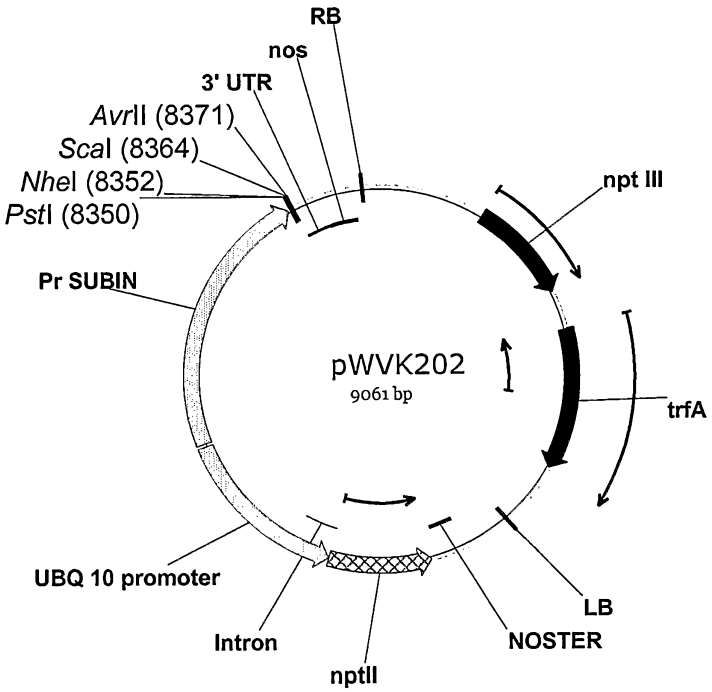

RB
nos
3' UTR
AvrII (8371)
ScaI (8364)
NheI (8352)
PstI (8350)
Pr SUBIN
npt III
pWVK202
9061 bp
trfA
UBQ 10 promoter
Intron
nptII
NOSTER
LB

FIGURE 3

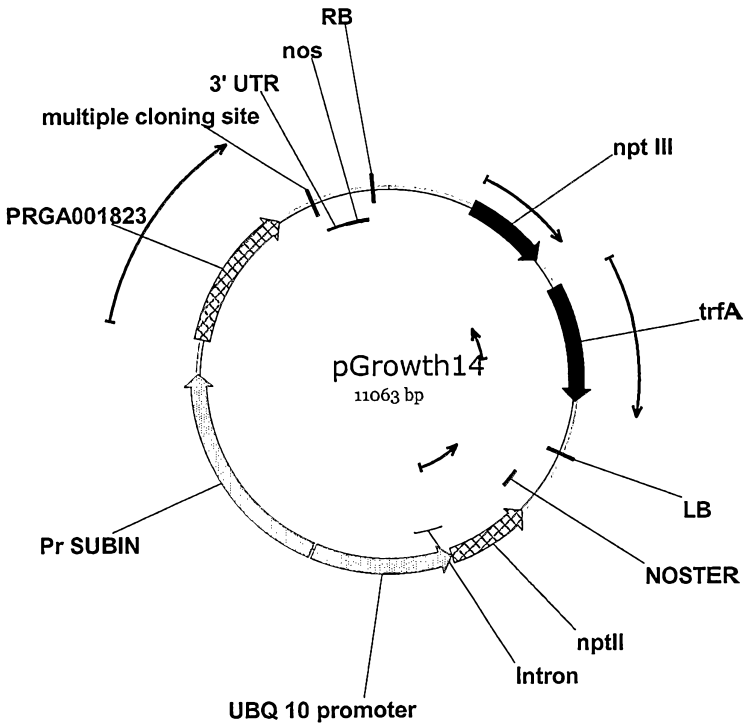

RB
nos
3' UTR
multiple cloning site
npt III
PRGA001823
trfA
pGrowth14
11063 bp
LB
Pr SUBIN
NOSTER
nptII
Intron
UBQ 10 promoter

FIGURE 4

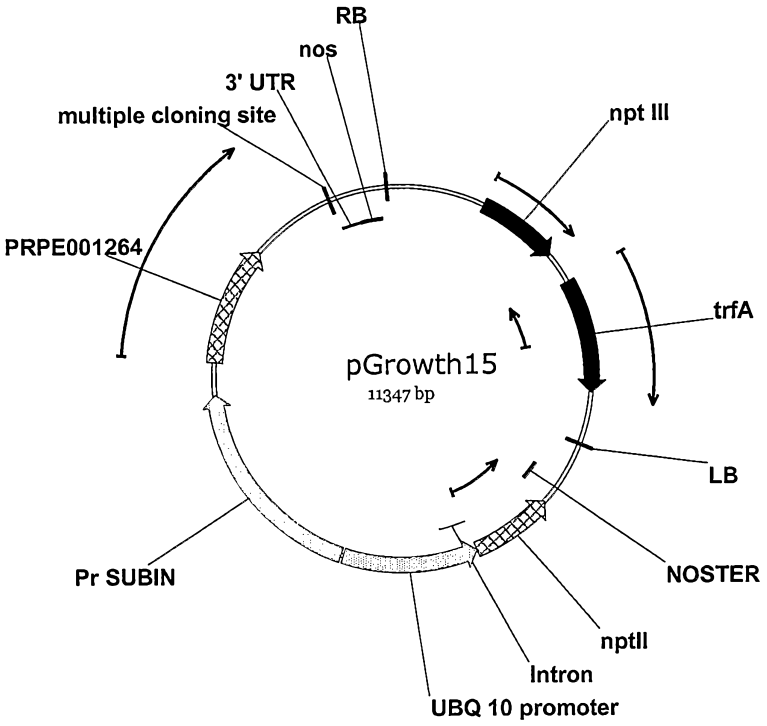

RB
nos
3' UTR
multiple cloning site
npt III
PRPE001264
trfA
pGrowth15
11347 bp
LB
NOSTER
Pr SUBIN
nptII
Intron
UBQ 10 promoter

FIGURE 5

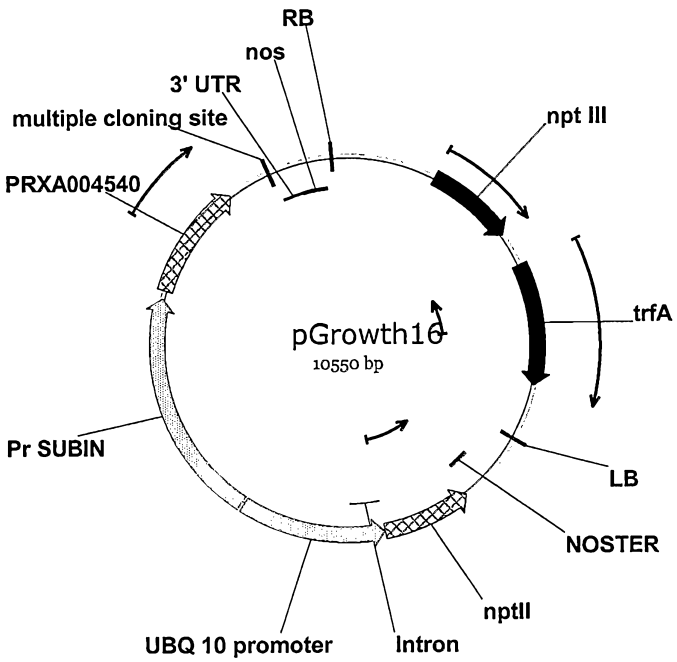

RB
nos
3' UTR
multiple cloning site
PRXA004540
npt III
trfA
pGrowth16
10550 bp
Pr SUBIN
LB
NOSTER
nptII
UBQ 10 promoter
Intron

FIGURE 6

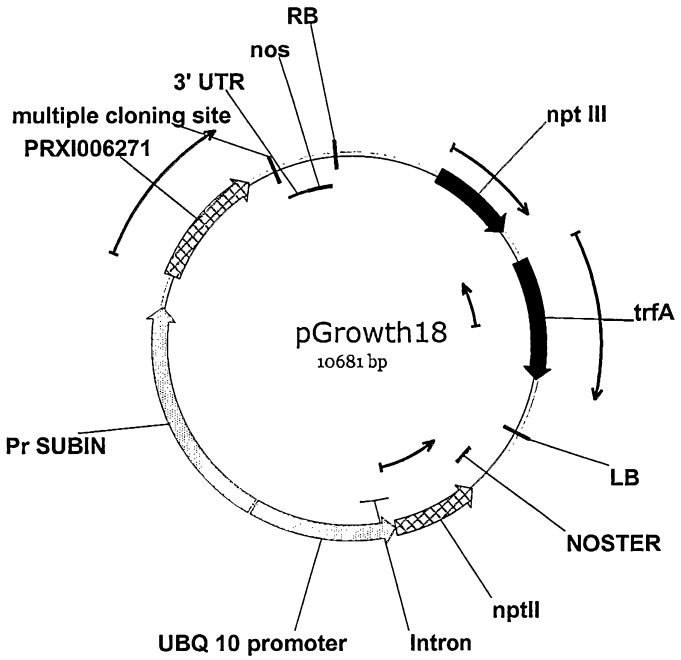

RB
nos
3' UTR
multiple cloning site
PRXI006271
npt III
trfA
pGrowth18
10681 bp
Pr SUBIN
LB
NOSTER
nptII
UBQ 10 promoter
Intron

FIGURE 7

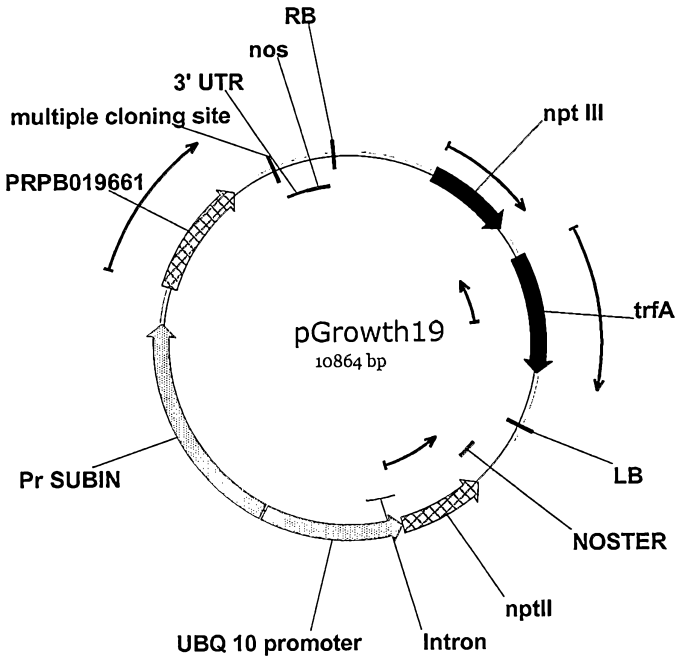
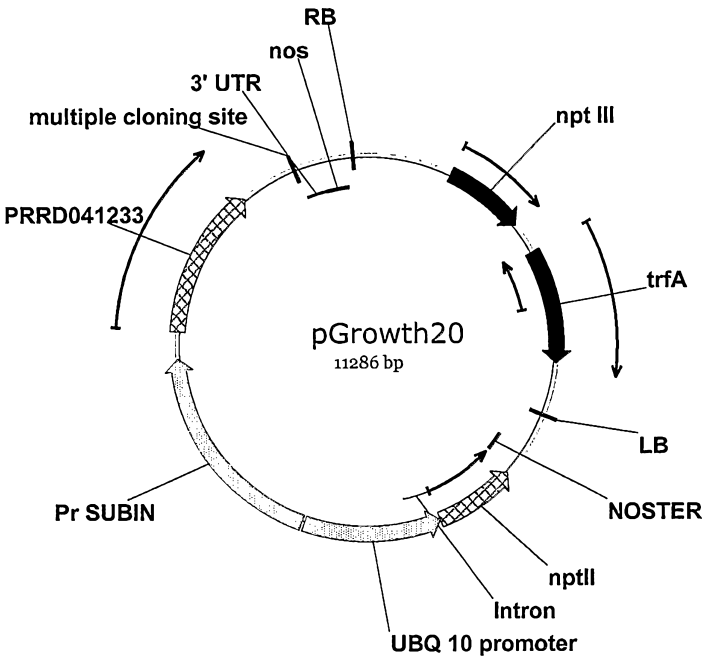

RB
nos
3' UTR
multiple cloning site
PRPB019661
npt III
trfA
pGrowth19
10864 bp
LB
NOSTER
nptII
Intron
UBQ 10 promoter
Pr SUBIN

FIGURE 8


RB
nos
3' UTR
multiple cloning site
npt III
PRRD041233
trfA
pGrowth20
11286 bp
LB
NOSTER
Pr SUBIN
nptII
Intron
UBQ 10 promoter

SEQ ID NO: 1
CAAAAGGGAAACAAGGATTTCAACTCCTGAGATTTTAAGTAACTCAATTAGCAGTTCCAA
CATTAAACCATTATTATTACCCCTTTTATCAGAGCTTTTTGACCTGGTTTAGGAAAATTCCA
AGAACTCAGCCAAATCTTCGACCTTTTTGGGTCTGCAAACCTTGCAAAGAAAGCCCAACT
CAGCTGGAGGCTGGAGATGAACATTTCCATGGCGGATTCCTGCCAAGATTCAATTTTGAT
GACGCAGAGCGCGAACCACCACCTTCCCACCGCCTAATTACTCCCTTTCCCTCTCTCTC
TCTCTCTCTCTCTCTCTATAGGAATTTCGGTCTTCTTCGAAGGGGCAGTCGCACTGTATT
GGGGAAAATCTCTCCAGTGGCGAGGATGGGAGACGGGAGCCTCGGCTCCGGCGGCAG
GGGCAACAGCGGCGGCGGCGGCGGCGGCGGCAGCCGGCCGGAGTGGCTGCAGCAGT
ACGATCTGATCGGCAAGATCGGCGAGGGCACCTACGGCCTCGTTTTCCTCGCGCGAAT
CAAGCATCCATCCACCAATCGCGGCAAGTACATCGCCATCAAGAAGTTCAAGCAGTCCA
AAGACGGCGACGGCGTCTCGCCCACCGCGATCCGCGAAATCATGCTGCTTCGTGAAAT
CTCGCATGAGAATGTTGTGAAGCTCGTGAACGTTCACATTAACCCCGTGGACATGTCGTT
GTATCTGGCTTTTGATTACGCAGATCATGATCTTTATGAAATTATCAGACATCACAGAGAC
AAGGTTAACCAGGCCATCAATCCCTACACAGTTAAGTCATTGCTTTGGCAGCTGCTTAAT
GGACTGAATTATCTCCACAGTAATTGGATCATACATCGAGATCTGAAGCCATCAAATATT
CTGGTCATGGGTGAAGGAGAGGAGCAAGGCGTCGTGAAAATTGCAGACTTTGGACTTG
CCAGGGTCTACCAAGCTCCTTTGAAGCCACTATCTGATAATGGGGTTGTGGTAACTATCT
GGTATCGGGCACCAGAATTGCTGCTTGGTGCAAAGCACTATACAAGCGCTGTTGATATG
TGGGCTGTCGGATGCATTTTTGCTGAGCTTTTGACATTGAAGCCATTGTTTCAAGGTCAA
GAAGTGAAAGCCAATCCTAATCCTTTCCAACTTGATCAACTTGACAAGATTTTCAAGGTCT
TGGGTCATCCAACGCAGGAGAAATGGCCTATGCTTGTGAATCTGCCTCATTGGCAATCA
GATGTGCAACATATTCAGAGACACAAGTATGATGACAATGCGCTTGGTAATGTTGTACGT
CTCTCTAGCAAAAACGCTACATTTGACCTCCTGTCAAAGATGCTAGAGTATGATCCTCAA
AAGCGTATAACAGCAGCTCAAGCCTTGGAACATGAATATTTTCGTATGGAACCTCTTCCT
GGGCGCAATGCACTTGTACCCAGCTCACCTGGAGATAAAGTGAATTATCCTACTCGTCC
TGTAGATACAACTACAGATATTGAGGGGACAACTAGTCTTCAACCTTCACAATCGGCATC
ATCTGGAAATGCTGTTCCTGGAAATATGCCCGGTCCTCATGTTGTGACAAATAGGCCCAT
GCCTCGTCCTATGCATATGGTCGGTATGCAAAGGGTGCCAGCTTCAGGAATGGCCGGTT
ACAATCTAAATCCTAGTGGCATGGGTGGTGGAATGAATCCTAGTGGCATCCCCATGCAG
CGGGGAGTTGCGAACCAGGCCCAACAGTCTCGAAGGAAGGACCCAGGAATGGGAATGG
GTGGATACCCTCCGCAACAGAAACAAAGGCGCTTCTAAGGGATCTAATTTACTGCGAGA
AATTTCGAGACCTTTCAGAAATTATGTTTGCTCTAGTGATACTGGCTGCAGCATCATGCTT
GCTTTGGTGACACGATGCAAAGAACTAGACTGCATCTTAAGGAAGGAAAGGCAACTATC
GACTTCATTAATGGGCTCATCCTCTGGGGGCTTAGAGGCTTCATCGGCATTGGTGCCGA
CTATGGCAGGATGGAATAATCAACGGCTGTTCTGCTAAATGCATTGGAAGTTGCTGCAG
ATTCACAGTGATAGATATCAATGGAGAATTGGATTCCAGCATCATAGCACTCGCAATCTT
TAGGGATCTGCCTCAATTTTTTGCTGAGGTTATAACTCGCTCTTTCAGAAATGTCATCGAT
TATTGATGGGCATGTCCATCAATCCGACACGTTTCCCTAATGTACATCAATTCTGCCAAA
AAAAAAA

SEQ ID NO: 2
CTGATCTCGTCGCCGCCGGCCGCCCGTCAGAGTCGCTCCCGCCGCGGCGCGCACTCC
TCCGAACCGTTCGAATTGTTTGTCCGTTGAAAACCCGCGAGATGGAGAAGTACCAACAG
CTCGCAAAGATCGGGGAGGGCACGTACGGGATCGTGTACAAGGCCAAAGACAAGAAGA
GTGGTGAACTGCTCGCTTTGAAGAAGATTCGTTTGGAGGCGGAGGATGAGGGCATTCCC
TCCACTGCCATCCGCGAGATTTCCCTCCTCAAACAGCTCCAACACCCAAACATCGTCCG
GTTGTACGATGTCGTTCACACAGAGAAGAAGCTGACACTGGTGTTCGAGTTCTTGGACC
AGGACCTCAAAAAGTACTTGGATGCGTGCGGTGACAACGGACTCGAACCGTACACTGTC
AAATCTTTCTTGTACCAACTGCTGCAAGGCATCGCCTTCTGCCACGAGCACCGCGTGCT
CCACCGCGATCTCAAACCGCAGAATTTGCTCATCAACATGGAGGGTGAGTTGAAGCTGG
CCGACTTCGGTCTGGCGCGCGCCTTCGGTATTCCCGTCCGGAATTACACGCACGAAGT
CGTCACCCTCTGGTACCGCGCCCCCGACGTGCTGATGGGATCGCGGAAGTACTCCACG
CAAGTCGACATTTGGAGCGTCGGCTGCATCTTCGCTGAGATGGTGAACGGTCGGCCGC
TGTTTCCCGGCTCGAGCGAGCAGGACCAACTGCTGCGCATCTTCAAAACGCTCGGGACT
CCGTCCCTCAAAACGTGGCCGGGAATGGCGGAGCTGCCGGACTTTAAGGACAATTTCC
CCAAGTATGTCGTGCAGAGCTTCAAGAAGATCTGTCCAAAGAAGCTCGACAAGACCGGC
CTCGATCTGTTGTCGCGGATGTTGCAGTACGACCCCGCCAAGCGCATCTCCGCCGAGC
AGGCGATGGGCCACCCGTACTTCAAAGACCTCAAGTTGAGGAAACCGAAGGCGGCCGG
ACCGGGACCCTAAAAAAAAGCCACCCAACTCACCCAATTCTTCTTCTTTTTCCATCTTTTT

TTTTCGTTGTAGTCAAGTTGGTGGTCGGTTGTCGGGAAAATTTATTTCATTGTTTTTTTCT
TCTCGAGATTTTTCGGTTCTGTGGGTACCGTTTCCAAACAAAAAAACTTTTGCCTGCCGT
GTCTTCACTTTGCTTCCCTCTGCTATTTAGCTGTGAACTTTCTTTTTTTTCTCATTCGCTTC
TTTGGGCGCGTGTGTTCAGTGGGGATAAGTGTTCTCGTTTTTCTCTGTACATACTCGCGC
TTCTACGCAGCAAATCACGAAGAGACCGTGTTTACTGAGCAAATATTTTTTTCTAATAAAA
AAAAA

SEQ ID NO: 3
GCCAACGCCAACACCACCACCAGCGGCGGCAACAGCAGCAGCAGCAACAGCAACCGG
CCGCCATTACTGCCCCCTCCAGGTTGCTGTTGCATTGCGCGGCATCTTTGAACTGGAGT
TGAATGGATCAGTATGAGAAGATCGAGAAGATCGGGGAAGGAACTTATGGCGTGGTCTA
TAAGGCTATTGATCGCTCCACCAATAAGACAATTGCTCTGAAGAAAATTCGTTTGGAACA
GGAAGATGAAGGAGTTCCGAGTACTGCAATCAGAGAAATCTCTCTCTTGAAAGAAATGCA
GCATGGAAACATTGTCAAGTTGCAGGACGTAGTGCACAGTGAGAGGCGTCTATATCTAG
TTTTTGAGTACTTGGACTTGGATTTGAAAAAGCACATGGATTCATGTCCAGAATTTTCTAA
GGACACCCACACAATAAAAATGTTCCTTTATCAGATCCTGCGTGGCATTTCCTATTGCCA
TTCTCATAGAGTTTTACATCGAGATTTGAAGCCCCAGAATTTGCTGCTAGATCGTCGTAC
TAATTCATTGAAGCTAGCTGACTTTGGGCTGGCCAGGGCTTTTGGGATTCCTGTTAGGAC
ATTTACCCATGAGGTGGTGACTTTATGGTATAGAGCTCCTGAGATACTCCTTGGATCCCG
CCATTACTCAACGCCCGTTGACGTGTGGTCTGTGGGTTGTATATTTGCAGAGATGGTGA
ACCGGCGACCACTATTTCCTGGGGACTCTGAGATTGATGAATTGTTTAAGATTTTCAGAA
TAATGGGCACGCCAAATGAAGATTCATGGCCCGGAGTGACCTCTTTGCCTGATTTTAAAT
CAACCTTTCCTAAGTGGGCTTCACAGGACCTAAAAACTGTTACGCCAACTGTTGATCCAG
CTGGCATCGATCTTCTTTCTAAAATGCTGTGCATGGATCCTAGAAGAAGAATAACTGCCA
AGGTCGCTCTTGAGCACGAATACTTCAAGGACGTCGGTGTCATACCGTGAATATCATTG
CCTTACTTACTCGGTGAGACATGTCTATATGGAGTTTGTAGCATATTAATGGGTTTTGATT
GTGTGAAATGTGAATCTTTTCAATTCTTAACCGCATCTTCTTTCTCGAGTCAAATACAAGT
CCTGTTGAACTGGATTTTGGATGTTTTAGCCAGCAAAGTTGATCTCTTAAAAAAAAAA

SEQ ID NO: 4
CTAACTCTCATCTAATTCAAAAATGGTTATGAAAAGTAAACTGGACAAGTACGAGAAGCTT
GAAAAGCTTGGTGAGGGTACTTACGGTGTTGTATACAAGGCCCAAGACAAGACGACCAA
GGAAATATATGCTCTTAAGAAGATCCGTTTGGAGTCTGAAGATGAAGGTATTCCAAGTAC
CGCAATTAGAGAGATTGCTCTTTTGAAAGAACTCCAACATCCTAATGTTGTCAGGATCCA
TGATGTTATCCACACCAATAAGAAGCTCATTCTGGTCTTTGAGTTTGTCGACTATGATCTT
AAAAAGTTTCTTCACAACTTTGATAAAGGTATAGACCCCAAGATTGTAAAGAGTCTTCTTT
ACCAACTGGTTAGAGGAGTGGCTCACTGCCATCAACAGAAAGTCCTTCATAGAGATCTTA
AACCTCAAAATCTGCTCGTCAGTCAGGAAGGTATTCTAAAGCTTGGTGATTTTGGTCTTG
CCAGAGCCTTTGGTATTCCAGTCAAGAACTATACAAATGAAGTCGTTACTCTATGGTACA
GAGCTCCTGATATTCTTTTGGGATCTAAGAACTACTCTACTTCTGTTGATATTTGGAGTAT
TGGGTGCATCTTTGTTGAAATGCTCAACCAAAAACCTCTCTTCCCCGGAAGCTCTGAACA
AGATCAACTTAAGAAGATCTTCAAGATTATGGGAACCCCCGATGCCACTAAATGGCCTGG
AATCGCTGAGCTACCTGACTGGAAGCCTGAAAACTTCGAGAAATATCCTGGAGAACCTTT
GAACAAAGTTTGTCCCAAGATGGACCCTGATGGTTTGGATTTACTTGATAAAATGTTGAA
ATGCAACCCAAGTGAAAGAATTGCTGCTAAGAACGCTATGAGTCATCCCTACTTCAAGGA
CATTCCTGATAACTTGAAGAAATTATATAACTAGATACAAGGTTAGCTAGGTATATAATAG
CGGTACAAGTCTTTTTTAAAGTAATTTACAAAAAGATACTTGATTATCTTAGATAAGAATTA
TATGTCTGCTAAACTTGGAATATGTTGGGTGAGAATACATATTCTATAATTATATAACAATT
TATATGTCGAAAATTATTATCTAAGCTAATAAATAACTATTAAAAAAAAAA

SEQ ID NO: 5
GGGAAGGGAGGTACCAAAATAAAAAGGGAGGGCAGAGACGCGCACGCTCCCCGACCAA
AGAAGACACGCCCCCGCCCAGCGCTGGCGCTAAACACCCGCCATTTCTGCAACCAGGT
ATCGCGCCGGAAGGACTTTTGACCTGGGCTGGATGGATCAGTACGAGAAAGTGGAGAA
GATCGGTGAAGGAACATATGGTGTGGTGTATAAGGCCATTGATCGGCTCACCAATGAGA
CTATAGCTTTGAAGAAGATACGCTTGGAGCAGGAAGACGAAGGAGTACCGAGCACCGC
GATCCGAGAAATCTCTCTCCTGAAGGAAATGCAGCATGGGAATATAGTCAGGTTGCAGG
ATGTAGTGCATAGTGAGAATAGGCTGTATCTAGTTTTTGAGTATCTGGACTTGGATTTGA
AGAAGCACATGGATTCATCTCCAGACTTTGCCAAGGATCCTCGTCTGGTAAAAATATTTC
TTTATCAAATACTGCGTGGAATAGCATATTGTCACTCACATAGGGTGCTCCACAGAGATT

TAAAGCCTCAGAATCTGTTAATTGATCGGCGTACAAATGCCTTAAAGCTTGCTGACTTTG
GACTTGCAAGAGCTTTTGGTATTCCTGTTAGGACTTTTACACATGAGGTGGTGACTTTAT
GGTACAGAGCACCAGAAATACTTCTTGGTTCTCGTCACTATTCTACCCCGGTTGATGTCT
GGTCAGTTGGTTGTATATTCGCTGAGATGGTGAACCAACGTCCTTTATTTCCTGGAGATT
CTGAAATTGATGAACTGTTCAAGATCTTCAGGATCTTGGGCACGCCAAATGAGGACACAT
GGCCGGGAGTAACTGCCTTGCCAGATTTCAAGTCTGCCTTTCCTAAGTGGCCTGCGAAG
AATCTACAAGATATGGTTCCAGGTCTCAATTCAGCTGGAATCGATCTTCTATCTAAAATGC
TCTGCTTGGACCCCAGCAAAAGAATTACAGCCAGGAGTGCTCTCGAGCATGAATACTTC
AAGGACATCGGATTTGTACCCTGATTATCTTTTGCCATTTGGTCAAGATTGTTATATATTG
GCCTGTGTAGGTGATTCTGCTCTTCTTTTAAGGTGTGAGTTGTGCGTGATCCCACATATT
CACATGTTGCAGTTTCTCTGAAGCCTTGGACAAATCAAGTAGAACTTCTCTCGGCAGCAT
CAGTTTTTCTAATCCATGCCTTGTTGCTTGCTGATCCAAAAAAAAAA

SEQ ID NO: 6
ATCTCCCTCCTCTCGACAAAATTCAAATTCCTCGCCGACGCCGACCACCACTTTCTCTCT
CTCTCCCGCTCAACGTTCCAAAAACACCAACCGCCCCATTTCTAGAGAGAGAAACAAGA
GAGAGAAGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCGATCGG
AGATCGGAGATCGGCGGCGGCGATGGAGAAGTACGAGAAGCTGGAGAAGGTCGGGGA
GGGGACGTACGGCAAGGTGTACAAGGCCAAGGACAAGGCGACGGGGCAGCTCGTGGC
CCTCAAGAAGACGCGCCTCGAGATGGACGAGGAGGGCGTCCCCCCCACCGCCCTCCG
CGAGGTCTCCCTCCTCCAGCTCCTCTCCCAGTCCCTCTACGTCGTCCGCCTCCTCTCCG
TCGAGCACGTCGACGGCGGCTCCAAGCGCAAGCCGATGCTGTACCTGGTGTTCGAGTA
CCTGGACACCGATCTCAAGAAGTTCATCGACTCGCACCGCAAGGGGCCAAACCCTAGG
CCCGTCCCCGCCGCGACCGTGCAGAACTTCCTCTACCAGCTCCTCAAGGGCGTCGCCC
ACTGCCACAGCCACGGCGTGCTCCACCGCGATCTCAAGCCCCAGAACCTGCTCGTCGA
CAAGGAGAAGGGGATCCTGAAGATCGCCGACCTCGGGCTCGGCCGCGCCTTCACCGTC
CCGCTCAAGAGTTACACGCATGAGGTCGTCACTCTCTGGTACAGAGCGCCTGAGGTGTT
GCTGGGATCCGCTCACTATTCGATCGGCGTGGACATGTGGTCTGTCGGGTGTATCTTCG
CTGAGATGGTGAGAAGGCAAGCCTTATTTCCTGGGGACTCTGAGTTTCAGCAACTGCTT
CACATATTCAGGCTATTGGGAACCCCAACTGAGAAGCAATGGCCAGGAGTTACCACTTT
GAGGGATTGGCATGTTTATCCACAATGGGAACCTCAAAACTTGGCAAGAGCAGTTCCAT
CCCTTGGACCAGATGGGGTGGACCTTCTGTCGAAAATGCTCAAATATGATCCTGCCGAA
AGGATCTCGGCTAAAGCAGCACTTGATCATCCCTTTTTTGACAGTCTCGACAAGTCTCAG
TTCTGATAATGCCTCGGATATATGGCCGAGTGTTCGCTGGACGGCCTCTTATGTTCTGAT
GTGGTCCTTGTATTTACCTTAAAAGTTACTTTCACTGCTTAAAGGGTGCTTCTGTTTGCTA
GGCACTCGACGCAGCCGTGTGTCATCATCTCGGTGATGAAGTATCATGTAACCAATATC
AATGAATCTATTTTTTGTCCTTTCAAAAAAAAAA

SEQ ID NO: 7
GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAATGGAGAGACCGGCGAC
GGCGGCGGTGTCGGCGATGGAGGCGTTCGAGAAGCTGGAGAAGGTCGGGGAAGGGAC
GTACGGGAAGGTGTACAGGGCCCGGGAGAAGGCGACGGGCAAGATCGTCGCCCTGAA
GAAGACCCGCCTCCACGAGGACGAGGAGGGCGTCCCTCCCACCACCCTCCGCGAGAT
CTCCATCCTCCGCATGCTCTCCCGCGACCCTCACATCGTCAGGTTGATGGATGTCAAGC
AAGGTCAGAACAAAGAAGGCAAGACGGTACTTTACCTGGTGTTTGAGTACATGGAAACT
GATTTGAAGAAGTACATCCGCGGTTTCCGCAGCTCTGGAGAGAGCATCCCTGTCAACAT
TGTTAAGAGTTTGATGTACCAACTCTGCAAGGGCGTGGCTTTCTGCCATGGGCATGGGG
TCTTGCATAGGGACCTTAAGCCTCACAATCTCCTCATGGACAAAAAGACGTTGACTCTTA
AAATAGCGGATCTTGGACTTGCCAGAGCTTTCACAGTCCCAATAAAAAAATACACTCATG
AGATACTGACTCTTTGGTATAGAGCTCCTGAAGTTCTTTTGGGGGCTACTCACTATTCAA
CCGCGGTTGACATGTGGTCTGTCGGCTGTATATTTGCTGAGTTAGTCACCAAGCAAGCA
CTCTTCCCTGGAGATTCTGAACTGCAACAGCTCCTACACATTTTCAGACTGTTGGGTACT
CCAAATGAAAAGATGTGGCCAGGGGTTAGCAGTTTGATGAACTGGCATGAGTATCCACA
GTGGAAACCTCAAAGCCTGTCCACCGCGGTTCCCAATCTGGACAAGGATGGGTTGGATC
TGCTTTCTCAAATGCTGCATTATGAGCCCTCGAGGAGGATTTCTGCAAAGGCAGCCATG
GAGCATCCTTACTTCGATGACGTCAACAAGACTTGCCTATGAGACTGCGCAAAGTGTGTA
GATTAGGAACCGGGATTTGGAATTTTAGCGAACTGAAAAGACATGTACTGAAAACTAATT
GCTTGCCTCAAGCATGACCTTTTGTGCACCATGTGCTTCTGTCAATTAATATCTTTTCTAG
GATGATATGGTTAATGCATCCTCACGCTTATATATATATCTCAATTTGTGTCAATTGATGTT
GCAAGCATGTTATACGAAAAAAAAAA

SEQ ID NO: 8
GGCAGAATTTCTTGCTTCCTCCCACCCCAGTTCCTTCTTTTTTACCGAGCAAGATTGACC
GGGGAGGACACACGTTTCACGCCTCGCTGGAGGAGCTTGCACCCTGACCGGAGCAAGG
TGGGTAGATTTGACTTGCTGATGCGGAGATTGAGCTGCTTGTGCTTGGGAGTCGCCTTG
TTCTTGCGGTGAGTGATCCGTCGGAGGATATTGGTGGCTGGTGGCTGGTGGGGTTAAG
AGGCCGTACACGAGCTGTATGGGATATGTTCAGAGATGAAAATTAGAAGCGAGACATGG
GTTGCGTGCTTGGTCGAGAGGTGTCGTCCGGCATAGTCACAGAGTCCAAAGGGCGGGA
TAGTTCGGAGGTCGAGACCAGCAAGAGGGATGATTCGGTCGCTGCGAAGGTAGAGGGA
GAGGGTAAAGCCGAGGAGGTGCGGACCGAGGAGACCCAGAAGAAGGAGAAGGTCGAA
GATGATCAGCAGTCGCGGGAGCAAAGGAGACGGTCCAAGCCAAGCACGAAGCTAGGCA
ACCTGCCAAAGCATATTCGAGGAGAGCAGGTTGCTGCTGGATGGCCCTCGTGGCTCTC
GGACATATGCGGGGAGGCTCTCAATGGCTGGATTCCTCGAAGAGCGAACACGTTTGAG
AAAATTGACAAGATTGGACAAGGGACATATAGCAATGTGTACAAAGCCAAGGATTTGTTG
ACTGGTAAAATCGTTGCACTAAAAAAGGTCCGGTTTGACAACTTGGAACCTGAAAGCGTG
AGGTTTATGGCACGAGAAATTCTCATTCTACGGCATCTGGATCATCCTAATGTTGTGAAG
TTGGAGGGTTTAGTTACCTCACGAATGTCCTGCAGCTTGTACCTGGTATTTGAGTACATG
GAGCATGATTTGGCTGGATTAGCAGCAAGTCCAGCAATCAAATTCACTGAGCCACAGGT
CAAGTGTTACATGCATCAGCTGCTCTCTGGACTAGAGCACTGTCACAATCGCCGTGTGC
TTCACCGTGATATTAAGGGCTCAAATCTACTGATTGACAATGGAGGAGTTCTTAAAATTG
GTGATTTTGGGCTGGCTTCATTCTATGATCCAGACCACAAGCATCGAATGACAAGTCGG
GTGGTCACTCTGTGGTATCGTCCCCCTGAACTTCTACTTGGGGCCAATGATTATGGTGT
GGGTATAGACTTGTGGAGTGCTGGTTGTATACTAGCCGAGTTGTTGGCGGGGAAGCCCA
TCATGCCTGGTCGAACAGAGGTAGAGCAACTCCACAAGATATACAAATTATGTGGCTCAC
CTTCAGAGGAGTATTGGAAGAAATATAAGTTGCCAAATGCAACATTGTTCAAGCCTCGAG
AACCTTATAGAAGATGCATCAGGGAAACATTTAAGGACTTCCCCCCATCATCTTTGCCAC
TTATTGAAACTCTTCTGGCAATCGATCCAGCGGAACGCGGGACAGCCACTGATGCTTTG
CAAAGCGAATTCTTCAGGACGGAGCCATATGCTTGTGAACCGTCTAGCCTCCCACAGTA
TCCCCCCAGTAAGGAAATGGATGCTAAAAAACGTGATGACGAAGCTCGACGGTTGAGAG
CAGCTAGTAAAGGACAAGCAGATGGTTCTAAGAAAGAACGGACACGTGATCGACGTGTC
AGGGCTGTCCCTGCTCCAGAAGCCAATGCCGAGCTTCAGCATAACATTGATAGAAGACG
ACTCATTTCCCACGCAAATGCGAAGAGCAAAAGCGAAAAGTTCCCTCCGCCACATCAGG
ATGGAGCGCTTGGTTTCCCATTGGGAGCTTCTCATCGTTTTGATCCGGCTGTTGTCCCTC
CCGATGTCCCATTTACCTCTACCTCATTTACTTCTTCGAAAGAGCATGATCAAACATGGTC
AGGCCCACTGGTCGACCCTCCCGGTGCTCCAAGGCGAAAGAAGCACAGTGCAGGCGGA
CAACGAGAGTCTTCAAAATTATCTATGGGGACCAATAAAGGGAGAAGAGCAGATTCTCAT
TTGAAAGCATACGAAAGCAAAAGCATAGCTTAAACGATCTACCGGACGGAAGACCATTTG
TTTTACAAACACGGAAAAAGTTAGTTCGAGGGTAAACTTCACTCGGTATTTGTTATCCATG
TCAATGTTTATTTTTTTGTCCCATGGAGGATTTAAATTTTTGCCGTTACCGTTTTTTGGTTT
GTTGGGATGGTCAGTCAATGTTGGTGCTGTAAATTTTTGCTTCAGTGGCTTCTTGGTACC
AAGTATTCCATATTTTTTACACTCTGAAGAATTGAAATAAGTATTGACAAAAAAAAAA

SEQ ID NO: 9
ATAGAAAGAGGATTGGAGTTCTTGAGAATTGTGAGATGGGTATTGGGTGTTGAGTGA
TTCTGTAAAGATCAATTCTTTTTAGAGTGATTCGTTGATTGTTGCAACATGGGTTGCATGT
ATTCTAAGAGTTCAGCGGTTGATGACAGCCGAGAGAGTCCCAAAGACAGAGTATCGTCT
AGTAGGCGATTATCAGAGGTGAAAACTTCAAGATTGGATTCATCGAGGAGAGAGAATGG
GTTCAGGGCAAGAGATAAGGTGGGTGATGTGAGTGTTATGTTGATCGATAAGAAAGTGA
ACGGGTCTGCTAGGTTTTGCGATGATCAGATTGAGAAGAAGAGTGATCGTCTTCAGAAG
CAGAGGAGAGAGAGAGCGGAAGCAGCTGCGGCGGCTGATCACCCTGGCGCTGGTCGG
GTCCCCAAGGCGGTAGAGGGCGAGCAGGTGGCAGCCGGGTGGCCGGTGTGGCTGTCT
GCGGTGGCAGGAGAGGCCATCAAGGGATGGCTTCCGCGACGGGCGGACACTTTTGAAA
AGCTGGACAAAATTGGCCAAGGAACTTACAGTAGCGTTTACAAGGCACGTGATGTCACT
AATAACAAAATTGTGGCTCTGAAAAGAGTGCGGTTCGACAACCTCGATACTGAGAGCGT
CAAGTTCATGGCGAGGGAAATCCATATCTTGCGTATGCTTGATCATCCTAATGTTATAAA
GCTGGAAGGCTTGATAACTTCAAGGATGTCCTGTAGCCTGTACCTTGTTTTCGAGTACAT
GGAGCACGATCTTACTGGACTTGCGTCGCGGCCTGATGTAAAATTTTCTGAGCCACAGA
TCAAATGTTACATGAAGCAGCTTCTAAGCGGTCTCGATCATTGTCATAAACATGGGGTCC
TACACCGGGACATAAAGGGCTCGAACCTTCTCATTGACAACAATGGCATCTTGAAGATTG
CGGATTTTGGTTTAGCCAGCGTCTTTGATCCTCATCAGACCGCTCCGCTGACAAGCCGG

GTGGTAACTTTATGGTACCGACCACCTGAACTTTTGCTTGGAGCTTCTCGCTATGGAGTT
GAGGTGGATTTGTGGAGTACCGGCTGTATACTAGGTGAACTCTATACTGGCAAGCCTAT
ATTGCCAGGGAAAACAGAGGTGGAGCAATTGCATAAAATTTTTAAGCTTTGCGGTTCACC
GTCTGATGATTACTGGAGAAGATTGCATCTTCCCCATGCAGCTGTTTTCAAGCCTCCACA
ACCTTATCGACGATGTGTTGCGGAGATATTCAAAGAACTCCCTCCAGTTGCTTTGGGCCT
CTTAGAGACCTTAATCTCTGTAGATCCTTCACAAAGAGGGACAGCAGCTTTTGCTCTCAG
GAGTGAGTTCTTTACAGCAAGTCCCCTTCCTTGTGATCCTTCAAGTCTGCCCAAATATCC
ACCAAGCAAAGAAATTGACATGAAATTGCGGGAGGAGGAAGCAAGACGGCGTGGAGCA
GCTGGAGGAAAGAACGAGCTTGAAAAGAGAGGAACCAAAGATTCACGAACAAATTCTGC
ATATTATCCTAATGCAGGACAATTGCAGGTCAAACAATGCCATTCCAATGCAAATGGCCG
AAGTGAAATTTTCGGCCCTTATCAAGAGAAAACTGTGTCTGGGTTCTTGGTTGCACCCCC
TAAACAGGCGCGAGTTTCCAAGGAAACAAGAAAAGATTACGCAGAGCAACCAGATAGAG
CTTCATTCTCGGGACCACTAGTTCCAGGTCCCGGATTCTCCAAGGCTGGAAAAGAACTT
GGCCATTCAATCACTGTCTCAAGAAACACCAATCTATCAACACTGTCGAGTTTAGTCACT
AGTAGAACCGGAGATAACAAACAAAAATCCGGTCCTTTAGTGTCAGAATCAGCAAACCAA
GCAAGCAGATATTCAGGACCCATAAGGGAGATGGAGCCTGCGAGGAAGCAGGACCGTA
GGAGTCATGTGCGGACGAATATTGATTATCGTTCGAGAGAAGATGGAAATTCCAGTACC
AAAGAACCTGCTCTGTACGGACGTGGGTCTGCAGGAAACAAGATCTACGTCTCAGGCCC
GTTGCTCGTCTCATCAAACAATGTGGATCAGATGCTCAAAGAGCATGACCGTCGAATCCA
AGAACATGCCAGGAGAGCACGGTTTGATAAGGCTAGAGTGGGCAATAATCATCCCCAAG
CAGCCGTCGACTCTAAGCTCGTGTCCGTGCACGATGCCGGGTAGATTCCCTGCAAAGA
GCTACTAAATGTCAGATATATCTTCAAAGCTGGGCAAGCTTACTTGCCTATGCGAAGGGA
AGCGCATATAGAGGCTTAGCAGGGTAGCTCATCCATTTGGTTTGGGATTTATCTTATATG
TATAGTCGGCATCTTTCTCTGCTCATTCTTATTTATTTTCATGGTAGCGAGGTGGGATCTC
TGGTAAGGAAGAGTCGGCCACCTTTTAGTACACTATGTAATACATAGTTTACGACAGGAT
ATAGCAGAATTAGAATAAGGTCAATCTAGAAACTGATTTGTTTCTGAATGGCGCGCCGCG
CACGCAGCAAGCGCTTGTAACCAGTGTTAAGAGAAAAAATGCATCGCATCGTCACTTGA
ACGGCCTCTTCGGCTTTAAGCGGTTTGCTTGGCGAGAGCAAGTAAGACGTTTTCAATTGT
TCTAAAAAAAAAAAA

SEQ ID NO: 10
GCTGGAGTGCCGACCGAAGCACCTCCGAGTCTTCCTCCTCCCCTCCACCGCGTGCTGC
GGTGGCTGACCGCATGATACCAGATGGGTTGCATCCCCACCATCATCTCCGACGGCCG
CCGCCGCTCCGCCGCCCCCGACAAACGCCGCCCCCGCCCTCGCCGGAGCTCGAGCGA
GGGGGAGGCGCCGCCCCACGCCACCGCCGCGGGCTCTGAGGGGGGTGAGTCCGCCC
GCGGTGCTCCGGGGAAGGAGAGGCCGGAGCCGGCGCCGCGTTTCGTCGTCAGGAGCC
CGCAGGGCTGGCCGCCGTGGCTGGTCGCCGCGGTGGGGCATGCCATTGGCGAGTTCG
TCCCTCGCTGCGCCGATAGCTTCCGGAAGCTCGCCAAGATTGGGGAAGGCACGTACAG
CAACGTGTACAAGGCGAGGGATCTGGTGACGGGGAAGACGGTGGCGCTGAAGAAGGT
CAGGTTCGATAATCTGGAGGCGGAGAGCATCAAGTTCATGGCGCGGGAGATTCTGGTG
TTGACGAGGCTCAACCACCCTAACGTGATCAAGCTCGAGGGGCCGGTCACTTCGCGCA
TGTCGTCGGGCCTTTACTTGGCTTTCGAGTACATGGAACATGATCTCTCTGGAATCGCG
GCCCGCCAGAATGGCAAGTTCACTGAGCCTCAGGTCAAATGTTTTATGAGGCAGCTGTT
GTCAGGGCTTGAGCATTGCCATAATCATGATGTCCTGCACCGTGATATTAAGTGCTCCAA
TTTACTTATTGACAATGAAGGAAACCTGAAAATAGCTGATTTTGGACTGGCCACATTCTAT
GATCCTGAGCGCAAGCAAGTTATGACCAATAGAGTCGTTACACTATGGTACCGAGCACC
GGAGCTTTTGCTTGGGGCCACCAGCTATGGAATTGGCATTGACCTCTGGAGTGCAGGCT
GCATTTTGGCGGAGTTACTGTATGGAAAGCCGATTATGCCAGGACGGACAGAGGTAGAG
CAACTGCATAAGATTTTCAAATTATGTGGCTCACCATCTGAAGCATATTGGAACAAATTCA
AATTGCCAAATGCCAATATTTTCAAGCCGCCACAGCCTTATGCTCGTTGTATAGCAGAAA
CATTCAAGGACTTTCCACCATCTGCTCTGCCTCTCCTTGAGACTCTGCTTTCGATAGACC
CTGATGAACGGGGCACCGCCACAACAGCACTAAATAGTGAATTCTTTGCTGCTGAACCT
CATGCATGTGAACCATCAAGCCTGCCAAAGTATCCCCCAAGTAAAGAGATGGATCTAAA
GTTGATCAAGGAAAAGACGAGAAGAGATTCAAGCAAGAGACCCAGTGCAATCCATGGTT
CCAGAAGAGATGGAATTCATGATCGAGCTGGGAGGGTAATTCCAGCTCCAGAAGCCACT
GCAGAGAATCAGGCAACTCTCCATAGGCCGAGAGCTATGAAGAAAGCCAACCCCATGAG
CAGAAGCGAGAAATTCCCACCAGCCCATATGGATGGAGTGGTTGGATCTTCTGCTAACG
CATGGCTATCTGGTCCAGCATCAAATGCCGCACCTGATTCCCGTCGACACCGTTCACTT
AATCAGAACCCATCAAGTTCAGTTGGAAAGGCTTCAACTGGATCCTCCACAACCCAGGA
AACGTTGAAGGTGGCTCCCGAGTTGTTGCAGGTGGGAAGCTCATCCTTGCATCCATGCC

ATCGGATGCTTGTTTATGGATCCAATCTCACGATCAGAAGTAAATAGTTGGTGCCTCTGG
CTCTGGATGCTTTCTGTAAATGGGAAAAGAAGAGCAGCACACAGACTCGATAATCTTTTT
GTTGGCTAGGATCCTTTAGATTCTTTTTCCCCACACAAGTTCTGTATCTTTTTTCAGAGAG
TAACCATTTGTTGGACAAAAGAATAAAGATTCTGCATTGTCAATGTCAACCTATTCCTTCT
GTTAAAAAAAAAA

SEQ ID NO: 11
GGGGTTTCCTTACTTCTTCACAGATGTAATGCACGTTCCTTCCTGCAAGAAGGTGGGGT
GTATCCTCCCCCCCGTACATGGTCGTTTCCTGATGCGGGTCAATCACCGTTTCATGTAAT
TCGATCCAATGCGGTCGGTGCATCGCAGCTGATGGGCTGCATATGCGCAAAGCAGGCG
GACCGGGGACCAGCGTCGCCTGGTTCAGGGATCTTGACTGGAGCGGGTACCGGTACC
GGTACCAGGTCGTCGAAGATCCCATCCGGTCTGTTTGAATTCGAGAAGAGCGGTGTGAA
GGAACACGGGGGTCGTAGCGGGGAATTGAGGAAGTTGGAGGAGAAGGGTTCATTGAGC
AAGAGATTGAGGTTGGAGTTGGGATTTTCACATCGGTATGTGGAAGCCGAACAAGCGGC
TGCAGGATGGCCTTCTTGGCTCACCGCCGTTGCTGGTGACGCAATTCAGGGATTGGTTC
CTCTTAAAGCAGATTCTTTCGAGAAATTGGAGAAGATAGGACAAGGCACATACAGCAGTG
TTTTCCGAGCAAGGGAGTTGGCAAATGGAAGGATGGTTGCTCTTAAGAAAGTCCGGTTC
GACAACTTTCAACCCGAGAGCATTCAATTTATGGCACGAGAGATATCGATCCTCCGCAG
GCTTGATCATCCGAATATCATGAAACTGGAGGGCATAATCACTTCTCGAATGTCGAACAG
CATTTATCTTGTGTTCGAGTACATGGAGCATGACCTTTACGGGCTAATATCTTCTCCTCA
GGTCAAGTTCAGCGATGCCCAGGTCAAATGCTATATGAAGCAGCTTCTGTCAGGAATAG
AGCATTGCCACCAGCACGGAGTGATTCATCGAGACGTAAAGTCTTCGAACATTCTGGTG
AACAATGAGGGGATTCTTAGAATAGGAGATTTCGGACTGGCTAACATTCTTAATCCAAAG
GACAGGCAACAACTCACCAGTCATGTCGTTACCTTATGGTACCGTCCTCCTGAGCTTCTC
ATGGGTTCCACAAGCTACGGCGTGACTGTGGATCTGTGGAGCGTTGGCTGTGTTTTTGC
AGAACTTATGTTTCGAAAGCCTATTCTCAGAGGGAGAACAGAGGTTGAGCAATTGCACAA
GATTTTTAAACTCTGTGGTTCCCCGCCCGATGGCTACTGGAAAATGTGTAAGGTACCCCA
GGCCACCATGTTTAGACCCCGCCATGCTTATGAATGCACTCTGCGAGAGAGATGCAAAG
GTATTGCGACCAGTGCAATGAAGCTGATGGAGACTTTTCTTTCCATAGAACCGCACAAGC
GTGGAACTGCTTCAAGTGCGCTCATATCTGAGTATTTCAGGACAGTACCATATGCATGTG
ATCCATCAAGCTTGCCCAAGTATCCTCCTAACAAAGAAATTGATGCTAAACATCGAGAAG
AGGCACGGAGGAAAAAGGCTCGTTCTAGAGTACGAGAAGCCGAAGTAGGGAAAAGGCC
AACAAGAATCCATAGAGCTTCGCAAGAGCAGGGGTTTTCCAGTAACATAGCTCCAAAAG
AGAAGAGAAGCTATGCCTGAGAGCTTGGATTCAAACCTCTTCCTTGATAGAGATGCTCAC
AAGCACCTTGCTTCTTCGCCATCAGGATGTCTCGGTTTTATAGTCGAATGCTCTTGTAAT
CCAACAATTTCAATGTTGATTGCCCAGGAGTTGCAAGGCCAGGCTCAGTTTGTTCGGAG
AAATCATGTGAGAAGTAACCTGAAGGGAGAAATTGATGTTCCATGTAGAGAGCCTCTGAA
GTCATCTTGCGATGCCTCGGCGACTTCTGAAGTGGCGGACACGTCTCAAGGACCGAGTA
TTTTATCGGGTCCAACACAGATCAGTGCGTCAAATGGCTTTGCATGGGCGAGAAGGCGA
AAGGATGATTCTGTTCTGAAAAGGTCATATAGTAGGTCTAGCGCCAGAAGTCAAGTTAGT
GCACTCGATTCATCAAGCATCGTGTTCGAACCCCATGACACGGATGCAAACGGAAGCAC
CATGCATGAACTGCGGAAACTCGGAATCGAGCTTGATCAGGATGATTCCTTTGAAGTACT
TGTGTTGAGGGCTAATACATCCCGATCTGAAGTTGAATCGATGAAGTGAGTGAATGAGAA
TGATACTTGAAGTTCGGCATCGACAAAGTGGTTACAGACATTGACAAAGTAGTTCTAGTC
ACCGTTGTTGTGAAGGTAGTTATAGCCATCGATTAGACAGTGATTAAAGTAGTACCCGTG
CCAATGAAGTGCCGTCCAAATGCTCAGCCACACTGCGGTCGACAGAATTCGATTTCTGG
TTGGCTACGAGTTGCTGATATCGCGACATGGCTGGGGCAGATGCAGTGAAAGAATTACC
CACAACATGCAGGATAGGGTGTTTGTAACACGCTTGATTTTCAACTCGAATGCTCTTCAT
ATGAATGTAACCAATTTCCTCTAGGATTTCTCTATGGACTAGGATGTAACCAGGCCTCCC
CTCAAAAGTAGATGAAACGCTTTATGCTTTATCTTTCTTTTCCTTGGAATGTACGCCTTAC
TTCCTTTAAAAAAAAAA

SEQ ID NO: 12
AGGGTGACCTAAATCGTTTATACCTTGCTCAGGCTAGTGAGAGAAAGAGAGAGAGATAG
AGAGAGAGAGAGAGACAGAGAGATGGCCGTAGCAGCGCCGGGCCACCTCAACGTCAAT
GAGTCCCCGTCGTGGGGATCGCGGAGCGTCGACTGCTTCGAGAAGCTCGAGCAGATCG
GCGAAGGCACTTACGGTCAAGTTTACATGGCTAAAGAGAAAAAAACTGGTGAAATCGTG
GCTCTGAAGAAGATTCGGATGGACAATGAGAGAGAAGGGTTTCCAATAACTGCCATACG
TGAGATCAAAATTTTGAAGAAACTCCATCATGAAAATGTGATAAAGCTGAAGGAAATCGT
GACTTCTCCAGGTCCTGAAAAAGATGAGCAAGGAAGGCCAGAGGGAAACAAGTATAAGG

GTGGCATCTACATGGTTTTTGAATACATGGATCATGATCTAACTGGCCTTGCCGACCGTC
CAGGGATGAGATTTTCTGTTCCCCAAATAAAGTGCTATATGAGACAGCTTTTGACAGGGC
TTCATTATTGTCATATCAATCAAGTTCTTCATCGTGATATAAAAGGGTCTAATCTTCTTATT
GACAATGAGGGAAACCTGAAGCTTGCAGACTTTGGCCTCGCTCGATCTTTCTCAAATGAT
CACAATGCAAACCTCACCAACCGTGTTATAACATTGTGGTATAGACCTCCGGAGTTGCTG
CTTGGGGCAACAAAATATGGTCCAGCTGTTGATATGTGGTCTGTGGGTTGCATCTTTGCC
GAACTTCTCCATGGGAAGCCAATTTTTCCTGGAAAAGATGAGCCAGAGCAATTGAACAAG
ATTTTTGAGCTTTGTGGAGCTCCGGATGAAATTAATTGGCCTGGTGTCTCGAAGATTCCC
TGGTACAACAACTTCAAGCCAACCCGGCCAATGAAGAGACGCCTCAGGGAGGTATTCAG
ACATTTTGACCGGCATGCTTTGGAGTTACTAGAAAGAATGTTAACACTGGATCCTTCTCA
GAGAATTTCTGCCAAGGATGCGTTGGATGCAGAGTATTTCTGGGCTGACCCATTGCCTT
GTGATCCAAAGAGTTTGCCTAAATATGAGTCATCTCATGAGTTCCAGACGAAGAAGAAGC
GCCAGCAGCAGAGGCAACACGAGGAAACAGCCAAGCGCCAGAAACTGCAGCATCCTCC
TCAGCATCCCCGGCTGCCACCTGTTCAGCAGTCAGGGCAAGCGCATGCACAAATGAGG
CCAGGACCTAACCAACTCATGCATGGTTCCCAGCCCCCCGTGGCGACAGGCCCGCCGG
GGCACCATTATGGAAAGCCCCGTGGACCATCGGGAGGAGCTGGGAGGTATCCTTCCAG
TGGAAACCCAGGCGGTGGGTACAATCACCCAAGTCGCGGTGGTCAAGGAGGTAGTGGA
GGTTACAACAGCGGGCCATATCCTCCTCAGGGGCGAGCTCCACCGTATGGTTCGAGTG
GAATGCCTGGTGCTGGACCTCGAGGTGGCGGGGGTAATAACTATGGTGTTGGTCCATC
GAATTATCCTCAAGGTGGTGGTGGTCCTTACGGTGGGTCTGGTGCTGGTCGTGGGTCAA
ACATGATGGGCGGTAACCGGAATCAGCAGTACGGCTGGCAGCAGTGAATTTCCACTGGT
CGCTGCTATTGTGGGCAAAAGGTTTTTGAATTGCGAGTGTAGTTTTTTATGATGACACTAT
TTCTGCGAAAGAGCAAAGAGGATCCACATACAAGAGTTGTTACGCTACACATCCTATACC
ATCAAAGGAACGTTGGAATGCCAGCATTGCTGTTTCATGATTGATACGTGTATGTAAAAT
CAATGTGGAAGCATTGTAATCTTGGCTTTGTCGAGGCTTTGCTTGAATAATGGTTTCCAA
TCGGTTTCAATGACGGCGTACAGACCCACTTGTTTCGGGAAAAAAAAA

SEQ ID NO: 13
TGGCATTGGATTTGATTAGCGAGTTGACGCAACAATCTGTCCCAACCCTCTGTCTCTCAC
ATCCACCGGCTCTTCGCCTTTTCAGGGTCCAGGCTTGATCCGCAATTCTCTCCCTCTCTT
TTCCTCTCTTTCTCTCTCTCTCTCCCTCCTTCCTTGCCCCTCCCGCCTCACATTTCGAA
GCAAAAACCAACTTGGAGCAGACGTTATGGTTTTTTAGACATTTCCCATTTAGGATTTTCC
TCGTGGCAGCTCGGAAGTTCGGACGTTTGGTTCCCATCCTTGTTCTTGTTCCAAGGCATT
GGGGGTTTCTCAGCATAGGAGGGAAGTAGGGACGGTCATTTGCCACCTCGAAGTTAATT
TGGGTTGCAAATAAGGACGGAGAACTTTGGTTGATTTTCTGGATCAGGGTTTTTCCTCAT
TCGTTTTTGCCATAGGATTAAGCATTTCTTGGTTTTCGCATTGAAGCAATTCACCAAGGGA
AAGTTGTAGAGATTGATAGGAATTAAGATCCTCTGGTTTCTTTTGGGAATGCTGATCGAT
GACAGTCTCAATTTTAGAAAAGAGGGGTTTCTCTCGGAACCGGTACTTTGATTTGACGAA
GAGGGCAAATTTGAAGTGGGATAATGGGTTGCATTTGCACCAAAGGAATTCTTCCCGCC
CATTACAGGATTAAAGACGGGGGTTTGAAGCTTAGCAAATCTTCAAAAAGGTCTGTTGGA
TCCTTAAGGAGAGATGAACTGGCGGTTTCAGCCAATGGTGGAGGCAACGATGCTGCAGA
TCGACTGATATCAAGCCCACATGAGGTTGAGAATGAAGTGGAGGATAGGAAAAACGTTG
ATTTCAATGAGAAACTGTCAAAATCCCTTCAAAGACGAGCTACAATGGACGTGGCAAGTG
GGGGACACACACAAGCACAGTTGAAAGTGGGTAAAGTAGGTGGCTTTCCCCTTGGTGAG
AGAGGAGCACAAGTGGTAGCCGGATGGCCTTCTTGGCTGACAGCAGTGGCCGGAGAAG
CTATTAACGGATGGGTGCCTCGAAGGGCCGATTCATTTGAGAAGCTGGAAAAGATTGGG
CAAGGTACTTACAGCAGTGTATATAGAGCACGTGACCTGGAAACGAATACAATTGTAGCC
TTGAAAAAGGTTCGATTTGCTAATATGGACCCAGAGAGTGTTCGATTCATGGCAAGGGA
GATAATTATCATGCGAAAGCTTGACCATCCAAATGTCATGAAGCTGGAGGGTTTGATAAC
TTCAAGGGTTTCTGGCAGTCTGTATCTTGTGTTTGAATATATGGACCATGATCTTGCTGG
CCTCGCTGCCACTCCGAGTATAAAGCTCACTGAATCTCAGATCAAATGTTACATGCAACA
ACTACTTCGTGGACTTGAATATTGCCACAGCCATGGTGTTCTACACCGTGACATAAAAGG
CTCTAACTTGTTAGTTGACAACAATGGCAATCTCAAAATTGGAGACTTTGGATTGGCAAC
TTTTTTCCGGACCAATCAAAAGCAGCCTCTAACGAGTCGCGTAGTCACTCTCTGGTACCG
ACCTCCTGAGCTGTTGCTTGGTTCTTCAGATTATGGAGCTTCTGTGGATTTGTGGAGTTC
TGGTTGCATCCTAGCTGAATTGTTTGCTGGGAAGCCCATAATGCCTGGAAGAACAGAGG
TGGAGCAATTGCACAAAATTTTCAAACTTTGTGGGTCTCCGTCAGAAGAATATTGGAAGA
AGTCAAAATTGCCACATGCAACCATTTTCAAGCCTCAGCAACCTTACAAGCGTTGTCTCC
TGGAGACATTTAAGGATTTTCCTTCATCTGCATTGGGCCTGCTAGATGTTCTTCTTGCAG
TTGAACCAGAATGCCGTGGAACGGCTTCCTCAGCCCTTCAGAATGAGTTCTTCACATCCA

ATCCTCTTCCAAGTGATCCGTCAAGTTTGCCAAAGTATCCATCAAGCAAGGAGTTTGATG
CCAGACTTCGAGATGAGGAAGCCAGAAAACATAAAGCTACTGCCGGTAAAGCTCGTGGT
CTTGAATCAATTAGAAAGGGTTCGAAAGAGTCTAAAGTTGTGCCTACATCAAATGCCAAT
GCTGATTTAAAGGCATCCATCCAGAAGCGACAAGAGCAATCGAATCCCAGAAGTACCGG
TGAGAAACCAGGAGGAACGACCCAGAACAATTTCATTCTATCTGGACAATCAGCAAAACC
CAGTCTTAATGGATCAACACAAATTGGAAATGCAAATGAGGTTGAGGCTTTGATTGTGCC
AGACCGAGAACTCGATTCTCCAAGAGGTGGGGCCGAGTTGAGAAGGCAGAGATCTTTCA
TGCAAAGAAGAGCATCGCAGTTATCCAGGTTTTCTAATTCTGTTGCAGTGGGAGGTGATT
CGCATCTCGATTGTAGTAGGGAGAAGGGTGCTAACACCCAATGGAGGGATGAGGGTTTT
GTTGCCAGGTGTAGTCATCCAGATGGTGGTGAATTAGCGGGAAAGCATGACTGGTCACA
TCATTTGTTACATAGGCCGATATCTTTGTTTAAGAAGGGGGGAGAGCACTCCCGAAGGG
ATTCCATTGCGAGTTATTCTCCCAAAAAGGGCCGAATCCACTATTCTGGACCTTTGCTTC
CCTCAGGAGATAACCTCGATGAAATGCTCAAGGAGCATGAGCGGCAAATACAAAATGCC
GTCCGGAAAGCTCGCCTTGATAAGGTCAAGACGAAGAGAGAGTATGCCGATCACGGGC
AGACGGAGTCACTTCTCTGTTGGGCAAATGGTAGGTGATTGTAAGGATGACAAGGAGTA
AGTGCATTTCTCTCTTGAAGCAAGGAGGCCCATGAGACGGCAACAATGTCGACAGGAAG
CAGAATTCCTATCCAGAAGCGGTAAAAGATATGATCGACACAAGCATTTTGTGTTGGAGC
CTCAGCTAATTGTATGTCATCGAGTACTTTTTCGAAGCACTTGATGATGCTCCTTGTGTC
GTATGAAATCCGAAGCAATCGGATTTGGTAAAATCAGGTGGCGGTCTTGTTTGGTGCTC
CAACAATGGAAAGGGCTGTATATTGTCAATGTTTACTCTGATTGTTGGAAGGATTCTCTG
TACAGTGATGGCCCATCCCTATCTATTCATTGACTGTTCCCTGAAAAAAAAAA

SEQ ID NO: 14
CTTCGAGCTCCGATGGACCCGGACCCGAGCCCGGACCCGGACCCGCCGAAGAGCTGG
AGCATCCACACCCGGCGGGAGATCATCGCCCGGTACGAGATCCTGGAGCGCGTCGGCT
CCGGCGCCTACTCCGACGTCTACCGCGGCCGCCGCCTCTCCGACGGCCTCGCCGTCG
CCCTCAAGGAGGTCCACGACTACCAGTCCGCCTTCCGCGAGATCGAGGCCCTCCAGAT
CCTCCGCGGCTCCCCCCACGTCGTCCTACTCCACGAGTACTTCTGGCGGGAGGACGAG
GACGCCGTCCTCGTGCTCGAGTTCCTCCGCAGCGACCTCGCCGCCGTGATCGCCGACG
CCAGCAGGCGGCCGAGGGACGGCGGCGGCGGGGGGGCGGCGGCGCTGCGGGCCGG
CGAGGTCAAGCGGTGGATGCTGCAGGTCTTGGAAGGGGTCGATGCTTGTCACCGGAAC
TCGATCGTTCATCGCGACTTGAAGCCCGGGAATTTGCTCATATCGGAGGAGGGAGTGCT
TAAGATTGCTGATTTTGGGCAGGCAAGGATACTCCTGGATGATGGAAATGTTGCTCCAG
ACTATGAGCCTGAATCATTCGAAGAGAGATCATCGGAACAGGCTGATATCCTTCAGCAG
CCAGAAACTATGGAGGCAGATACCACATGCCCTGAAGGTCAAGAGCAGGGAGCTATCAC
TAGGGAGGCATACCTCAGAGAGGTGGACGAATTCAAGGCTAAAAACCCTAGGCATGAAA
TCGACAAGGAAACAAGCATATTTGATGGCGATACTTCTTGTCTGGCCACATGCACGACCA
GTGACATTGGAGAAGATCCTTTTAAAGGTTCCTATGTTTATGGGGCCGAAGAAGCTGGA
GAAGATGCACAAGGCTGTCTCACATCTTGTGTTGGGACACGCTGGTTCAGAGCACCTGA
ACTGCTCTATGGGTCCACAGACTATGGGCTCGAGGTTGATCTCTGGTCACTGGGATGCA
TTTTTGCTGAGCTTTTGACTCTGGAACCCCTTTTCCCTGGGATTTCCGATATCGACCAAC
TTAGTAGAATCTTCAATGTTTTGGGCAACCTGAGCGAGGAAGTCTGGCCAGGCTGTACG
AAACTTCCAGATTATAGAACAATTTCATTCTGCAAAATCGAAAACCCCATCGGTTTGGAAT
CCTGCCTGCCGAACTGCTCAAGTGATGAAGTCTCTTTAGTTCGGCGACTTCTTTGCTACG
ATCCAGCTGCAAGAGCCACTCCCATGGAACTGTTGCAAGACAAGTACTTCACTGAAGAA
CCACTTCCAGTTCCAATTAGTGCGCTTCAGGTGCCCCAGTCAAAGAATAGCCATGATGA
GGACTCTGCTGGTGGTTGGTATGACTACAATGACATGGACTCGGATTCTGATTTTGAGG
ACTTTGGCCCTTTGAAGTTCACACCTACAAGTACTGGTTTCTCTATACAGTTCCCCTAGTC
ATCGTGAGAATGCAGGGATCAAATTTGTGAGTACTTGCTAAAATTTTTGCTACGGATAAT
GTTGTGAGGCGAGGCAGTCGAAATTACGGAGGTTGACTTCTTTCGACATTCATATGCTTG
CTGTTATAGACAAGGGGGACGTGTCTGTAAAAGACCATGAGAGGGTTATATTCTCGGTG
TCCTTTTGTCTTGTAAATTTCCTTCACCAGAGACACAAATTCCCATCGAAAAAAAAAA

SEQ ID NO: 15
CCAAAAAAGAAGAAAAGAGAAGAAGAGACGTTCCGTCTGACTGGTCTCGCTGCAAAAGC
TCCGCCGGGAGTCCCCTCCACGCCGCCGTCACTCTCTCCCTCTTTGAGCTCCGATGGA
CCCGGACCCGAGCCCGAGCCCGGACCCGCCGAAGAGCTGGAGCATCCACACCCGGCG
GGAGATCATCGCCCGGTACGAGATCCTGGAGCGCGTCGGCTCCGGCGCCTACTCCGAC
GTCTACCGCGGCCGCCGCCTCTCCGACGGCCTCGCCGTCGCCCTCAAGGAGGTCCAC
GACTACCAGTCCGCCTTCCGCGAGATCGAGGCCCTCCAGATCCTCCGCGGCTCCCCCC

ACGTCGTCCTACTCCACGAGTACTTCTGGCGGGAGGACGAGGACGCCGTCCTCGTGCT
CGAGTTCCTCCGCAGCGACCTCGCCGCCGTGATCGCGGACGCCAGCAGGCGGCCGAG
GGGCGGCGGGGTGGCGCCGCTGCGGGCCGGCGAGGGCAAGCGGTGGATGCTGCAGG
TCTTGGAAGGGGTCGATGCTTGTCACCGGAACTCGATCGTTCATCGCGACTTGAAGCCC
GGGAATTTGCTCATATCGGAGGAGGGAGTGCTTAAGATTGCTGATTTTGGGCAGGCAAG
GATACTCCTGGATGATGGAAATGTTGCTCCAGACTATGAGCCTGAATCATTCGAAGAGA
GATCATCGGAACAGGCTGATATCCTTCAGCAGCCAGAAACTATGGAGGCAGATACCACA
TGTCCTGAAGGTCAAGAGCAGGGAGCTATCACTAGGGAGGCATACCTCAGAGAGGTGG
ATGAATTCAAGGCTAAAAATCCTAGGCATGAAATCGACAAGGAAACAAGCATATATGATG
GCGATACTTCTTGTCTGGCCACATGCACGACCAGTGACATTGGAGAAGATCCTTTTAAAG
GTTCCTATGTTTATGGGGCCGAAGAGGCTGGAGAAGATGCACAAGGCTCTCTCACATCT
TGTGTTGGGACACGCTGGTTCAGAGCACCTGAACTGCTCTATGGGTCCACAGACTACGG
GCTCGAGGTTGATCTCTGGTCACTGGGATGCATTTTTGCTGAGCTTTTGACTCTGGAACC
CCTTTTCCCTGGGATTTCCGATATCGACCAACTTAGTAGAATCTTCAATGTTTTGGGCAA
CCTGAGTGAGGAAGTCTGGCCAGGCTGTACGAAACTTCCAGACTATAGAACAATTTCATT
CTGCAAAATCGAAAACCCCATCGGTTTGGAATCCTGCCTGCCGAACTGCTCAAGTGATG
AAGTCTCTTTAGTTCGGCGACTTCTTTGCTACGATCCAGCTGCAAGAGCCACTCCCATGG
AACTGTTGCAAGACAAGTACTTCACTGAAGAACCACTTCCAGTTCCAATTAGTGCGCTTC
AGGTGCCCCAGTCAAAGAATAGCCATGATGAGGACTCTGCTGGTGGTTGGTATGACTAC
AATGACATGGACTCGGATTCTGATTTTGAGGACTTTGGCCCTTTGAAGTTCACACCTACA
AGTACTGGTTTCTCTATACAGTTCCCCTAGTCATCGTGAGAATGCAGGGATCAAATTTGT
GAGTACTACGTAAAATTTTGCTACGGAGGCGAGGCAGTCGAAATTACGGAGGTTGACTT
CTTTCGACATTCATATGCTTGCTGTTATAGACAAGGGGGACGTGTCTGTAAAAGACCATG
AGAGGGTTATATTCTCGGTGTCCTTTTTTCTTGTAAATTTCCTTCACCAGAGACACAAATT
CCCATCGACTTGTCAAAAAAAAAA

SEQ ID NO: 16
GGAATCGAGAAAAGTCCCTCTCCCTCTCTCCATCTCTCCATCTCTCCTTCCTTTTTTG
GAAGAAAGGTTGTGGGGGATACATCATCGCCGTTGGGAAACGGAGAGGGGGTTACCGG
AGCGGGCGGGGGGGAGATAGCGATTCGTGTTGAACTGCTGGTCTCGCCGCCGCCGCCG
CCGCTCCTCGAGATGTCGAACCAGCACCGGCGCTCCTCCTTCTCGTCCTCCACGACGTC
GTCCCTCGCCAAGCGCCACGCCTCCTCCTCCTCCTCCTCCTTGGAGAACGCCGGGAAG
GCCTTCGCCGCCGCCGCCGTCCCGTCGCACCTCGCCAAGAAGCGGGCCCCCCTCGGC
AACCTGACCAACCTCAAGGCCGGCGATGGCAATTCCCGCAGCTCATCGGCACCATCCA
CTTTGGTGGCTAATGCGACGAAACTGGCAAAGACGAGGAAGGGATCTTCTACTTCCAGC
TCCATCATGGGCCTCTCGGGAAGTGCTTTACCGAGATATGCTAGCACGAAACCCAGTGG
AGTTCTTCCTAGCGTTAATCCTTCCATTCCAAGAATAGAGATAGCCGTTGATCCCATGTC
GTGCAGCATGGTTGTTTCGCCCAGTAGATCTGACATGCAATCGGTTTCGCTGGATGAGA
GCATGTCCACCTGCGAGTCTTTCAAGAGTCCCGACGTTGAGTATATCGACAATGAGGAT
GTTTCGGCAGTTGATTCTATCGACAGGAAGACATTTAGCAATCTTTATATCTCAGATGCT
GCAGCAAAAACAGCGGTTAACATTTGCGAGAGAGATGTACTCATGGAAATGGAAACAGA
TGAGAAGATTGTCAATGTTGATGACAACTACTCGGATCCGCAACTCTGTGCAACCATTGC
TTGTGACATTTACCAGCACTTACGTGCATCTGAGGCCAAGAAGAGACCTTCCACTGATTT
TATGGATAGAGTACAAAAGGATATAACTGCCAGCATGCGTGCCATACTAATTGATTGGCT
TGTGGAGGTGGCTGAAGAATACAGGCTCGTACCTGATACACTGTACCTGACTGTTAACT
ACATAGATCGGTATCTTTCAGGAAATGTGATGAATAGACAACGTCTGCAGCTGCTTGGTG
TTGCTTGCATGATGATAGCTGCCAAGTACGAGGAGATCTGTGCACCTCAGGTGGAAGAG
TTCTGTTATATTACCGACAATACGTACTTCAAGGAGGAGGTATTGCAAATGGAATCTTCT
GTGTTGAATTACTTGAAGTTTGAAATGACTGCTCCCACTGTCAAGTGCTTTTTAAGACGAT
TTGTTCGTGCTGCGCAAGGCGTGAATGAGGTTCCATCTTTGCAACTGGAGTGCATGGCC
AACTATATTGCAGAACTCTCTCTACTAGAGTACGATATGCTTTGTTATGCTCCATCTCTTG
TAGCTGCATCAGCAATATTTCTGGCCAAATTTGTCATCACCCCTTCGAAGAGACCATGGG
ACCCGACGTTGCAGCATTACACTCTCTACCAGCCTTCCGATCTGGGAAACTGTGTCAAG
GATCTGCATCGGCTGTGTTTTAACAACCATGGTTCGACCTTACCAGCAATCAGGGAGAA
GTACAGCCAGCATAAGTACAAATACGTGGCCAAGAAGTATTGCCCTCCCTCGATACCTC
CGGAGTTTTTCCACAATCTTGTCTATTAGCCGTTGACGACACTCTGCTCCAATGCTCCGT
TGGAACCTATCCTCCGTTTGCGCACCAGCAACTCTTACGTGCTTTCGCCCCATAATCCG
GTGATTGTCAAAGAGTGTTGATCTGTATTCACTTCGTAAGATTTTAGGATCTCTGTCCCTT
TCCCCATCTTCGTTGCTGAAGTGTAAATCCCCTCCTCTCTTTCGAGTTCCGTCATTTTATT
TTGCACCGTTGATGATGTTGTACTCTGGTGCACGCGGTTGATTTTAGCCCGGCTGCATG

GCCGTACGGAGGATTAGTAGCTTCTTTATGATGAAATGTCGTGGCAGATTTCTTTTTTTCT
TTCTGTTCTGTTTTGATTGCTCTTTTTTCCTCCATAGAATGAGCTCACTTCTGCGCAGAAG
TTGGTGAAGATACCGGTAAGTTGACTGCTAATGTGAAATTTGAGCTCCCAAAAAAAAAA

SEQ ID NO: 17
GCTTTCTCTTTCCTCCTGGCGAGCTTTCTCTCTCTACATCTCTCTATGCGGACCTCCTCC
GAGGGGTCGCGGAGCTCCAGCACAACAGCGTTACTTGTTCTTTACACTTGGCTTTTCGA
ATTTCATATCGAGCATCTGCAGAATGTTTGTTCCCTCTTTGTCATGATACAAGGATAGTGG
GTGCAATGAGGAGCCATTTGATTTGATGACACGTACTTCCTTGGGAGGAGTGCAAATAT
GAACAAAGAAAATGCAGTTGGAACAAAAAGTGAAGCACCCACTATCCGAATTACCCGATC
AAGGTCTAAAGCATTGGGCACGTCAACAGGGATGCTCCCATCCTCAAGGCCCTCCTTTA
AACAGGAACAAAAGCGTACTGTCCGTGCAAATGCCAAGAGATCAGCATCAGACGAGAAC
AAAGGAACTATGGTTGGAAATGCTAGCAAGCAGCACAAAAAGAGAACAGTACTTAATGAT
GTTACTAACATCTTTTGTGAGAATTCATACTCGAATTGTCTCAATGCTGCCAAAGCTCAGA
CCAGCAGGCAGGGTAGAAAGTGGTCTATGAAGAAGGACAGAGATGTGCATCAAAGTGG
AGCTGTCCAAATAATGCAGGAGGATGTCCAAGCGCAATTTGTAGAAGAGTCGTCCAAAA
TAAAAGTGGCAGAGTCTATGGAAATCACTATCCCAGACAAATGGGCAAAACGAGAAAATT
CAGAGCATAGTATCTCGATGAAGGACACTGTAGCAGAGTCTTCAAGGAAACCACAAGAA
TTTATATGTGGTGAGAAGTCAGCAGCACTAGTTCAACCAAGTATCGTAGATATAGATTCA
AAACTCGAGGATCCTCAAGCTTGCACCCCATATGCACTTGACATATACAACTATAAACGC
AGCACAGAGCTTGAACGAAGACCTTCAACTATATATATGGAAACCTTGCAGAAAGATGTC
ACTCCAAACATGCGTGGAATTCTAGTTGACTGGCTTGTTGAGGTTTCTGAAGAATATAAA
CTAGTTCCTGATACTCTTTACCTCACTGTGAATCTCATAGACCGATCTCTCTCACAAAAGT
TTATTGAGAAGCAAAGACTCCAATTGCTTGGTGTTACTTGCATGCTAATTGCCTCGAAGT
ACGAGGAAATCTGTCCACCACGAGTGGAAGAGTTTTGCTTTATCACGGACAATACCTATA
CAAGCCTAGAGGTATTGAAAATGGAGAGTCGAGTTCTAAATCTTCTGCATTTTCAGTTGT
CCGTTCCCACCGTTAAAACATTTCTGAGGAGATTTGTTCAAGCTGCCCAAGTATCGAGCG
AGGTCCCTTCTGTTGAACTAGAATATTTGGCAAATTATTTGGCGGAGTTAACTCTCGTTG
AGTATAGTTTCTTAAAGTTTCTGCCTTCCCTTATGGCTGCATCTGCTGTACTTCTTGCCAG
ATGGACACTCAATCAATCTGACAACCCTTGGAATCTAACTCTAGAACACTACACAAAGTA
CAAGGCATCAGAGTTAAAAGCTGCGGTTCTTGCATTGGAAGATTTGCAGCTCAACACCA
GTGGTTCCACCCTAAATGCGATACGTGAAAAGTATAGACAACAGAAGGTAAACTATTCAT
TACTGATCCATAGCAAGGCCAATCATGAGATATTATAGACCACACATTTTTGCTAAACTGT
TGGAGTTCATAGAACTAGACGTTTTCCATTGACACTTGATCAACTGATCATGGCAGTTCA
AATGCGTGGCGACTTTAACTCCCTCTGAACCTGCACTTTCACTCTTCCAAAGATTTAACTA
CAGCTGATCTTCCTGAAGGGCTAACTTGAGAATAATCAAATGGAGTGACTTCCCACATTG
CGGCACTTCCAGGGTGAATTCGAGCTCCACTAATAGGATGTCTACTTACCATTTTTTGTC
GACCGAGGTTTTGTTGTCCACATTAATGTTCTCATTGTTGTAGCATTTATTTACGGTCTGG
TGCTTCCAGAACAGTCCCTTAATGATGAGGAACTCTCAGATCATGACCCCATACCAGGC
GGCAAAACATTGCGCCGCTCGGCAATAAACCTCGGAGTTTTCCTAGTCTCTAGGAACTC
GTACAGTAACCCCATTTCTTTTGCTTGCCCGAGTGTCAAGCTTTCAGATGTACTGTTTAGT
CACATTACTTCTCCGAAAGAAGATTTGGAGATCATTCTCTTTGTACTTCAATGAATTTAAC
AGAACGATACTGGAATCTCTCTGTTACTCGTTCAAAAAAAAAA

SEQ ID NO: 18
GCTACTGCCCCACACACACTCTCTCTCTCTCTCTCTCCTTTTCCCCCAAATCAATAAGA
AGAGACGACGACTGTGTAGTAGTGAGGGAAGCGGCCGCAAACCGGGCGGTACTGAAAT
TTTGATGGAGGGTTTCGGGGGAATCTCTCTCTGCTGCTGATCTCCGGGAGGGCTGGTTC
GTCTTCGGCTTCGGGAGGGGGAGGAGCGGGAGATCGTGATTTCGAAGGAGGCAGAGAT
GGCCGGATCGGACGAGAACAACCCGGGCGTTGTCGGAGGTGCGCACGTTCAAGAGGG
CTTGCGGGTCGGAGCGGGGAAGATGGGTGCGGGGAATGTTCAGCAGAGACGAGCTCT
GAGCAACATCAACAGCAACATCATCGGGGCTCCTCCTTATCCATGCGCGGTCAACAAAA
GGGTCTTGTCCGAAAAAAATGTCAACTCTGAAAACGATCTCCTCAACGCTGCTCATCGGC
CAATTACTAGGCAGTTTGCTGCTCAGATGGCTTACAAGCAGCAACTTAGACCTGAGGAG
AACAAGAGGACGACCCAATCAGTCTCAAATCCCAGCAAATCTGAAGATTGTGCCATCTTA
GATGTGGATGACGACAAGATGGCTGATGACTTTCCGGTGCCAATGTTTGTGCAACACAC
CGAAGCAATGTTAGAAGAAATTGATCGGATGGAGGAGGTTGAGATGGAAGATGTAGCTG
AAGAACCTGTCACGGACATTGACAGCGGTGATAAAGAGAACCAGTTGGCTGTTGTTGAG
TACATTGATGACCTATACATGTTCTATCAGAAAGCCGAGGCTTCTAGTTGCGTTCCCCCA
AACTACATGGATCGGCAGCAGGATATTAATGAGCGGATGAGAGGTATACTAATTGACTG

GCTGATTGAGGTTCATTACAAGTTTGAATTGATGGATGAGACCTTGTATCTTACGGTCAA
TCTCATCGATAGATTCTTAGCTGTTCAACCTGTAGTGAAGAAAAAACTCCAGCTAGTAGG
GGTAACAGCCATGCTTCTGGCATGCAAATATGAAGAGGTCTCAGTTCCAGTAGTGGAGG
ATCTCATTCTGATTTCAGACAGGGCTTATAGCAGGAAAGAAGTTCTAGAAATGGAGCGGT
TGATGGTGAATACTTTGCACTTTAACATGTCAGTGCCTACTCCTTATGTTTTCATGAGGAG
ATTTCTTAAAGCCGCTCAATCTGACAAGAAGCTCGAGCTCTTGTCATTCTTCATCATCGA
GCTTTCCCTGGTTGAATATGACATGCTGAAGTTCCCACCCTCTTTATTAGCTGCTTCTGC
AATCTACACTGCTCTGAGTACAATTACCAGAACTAAACAGTGGAGTACAACATGTGAATG
GCACACCAGCTACTCAGAAGAACAGCTTCTGGAATGCGCCAGATTGATGGTGACTTTCC
ATCAGAGGGCTGGATCAGGGAAGCTCACTGGTGTGCACCGGAAGTATAGCACATCCAA
GTTTGGTCATGCTGCGAGAACTGAGCCTGCTAACTTTCTCTTGGACTTCCGTTTGTAGTG
GGTTGCGTGTGTCGTTGTGTACCTTGCCTAAATCAAACATAAAAAACTGACTTTTGGGCC
AGGGTGGGGGGCTTAAAAATATAGCAACATAAGTTGCCCTGCGAGTGTGTGTATAGCTG
ATTCATTTCAGCGTTTTATCTGAACATATGATGGTGATTCATTCCTTGTATCCCTTGAACA
TATATGATGCCGATTCATTTCAGCGTTTTATTTGCACCTGACAGTGGATAGCTGTTGTTTG
TTGACACTGTTCCTTAGTTGCATTTATTCATCGATTTTCATTGTTGCCAATGATGATGTTAC
TAGCAATAACATCCCTTGAACTTAAAAAAAAAA

SEQ ID NO: 19
AAGCAGTGGTATCAACGCAGAGTACGCGGGGATTCTCGTTTCTTCCTGCACAGGATCTC
ACTGTGAGAGAGACCAGAGGAGAAGGCACAGGTTCTCTCGTGGGTAGGGTTTCGCCTC
GCCAATCAGATTCGGAGGAGCTTTTTGTTCTTTGATCGAAATGGCTTCAAGACCCATTGT
CCCCGTGCAAGCGAGAGGCGAGGCAGCGATAGGAGGAGGAGCGGGTAAGGCAGCGAT
AGGAGGAGGAGCGGGTAAGCAGCAGAAGAAGAACGGCGCAGCAGAAGGAAGAAACCG
CAAAGCCCTCGGCGACATTGGCAATCTCGTCACTGTTCGAGGCATCGAAGGCAAGGTTC
AGCCTCATCGCCCCATCACAAGAAGCTTCTGCGCACAACTACTGGCAAATGCACAAGCT
GCCGCGGCTGCCGAAAACAACAAGAAACAAGCTGTAGTTAACGTGAATGGAGCGCCGT
CCATTCTTGATGTACCTGGAGCAGGCAAAAGAGCCGAGCCTGCGGCGGCGGCGGCGG
CGGCGGTGGCAAAAGCTGCTCAAAAGAAAGTCGTTAAACCGAAACAAAAGGCTGAAGTT
ATCGATTTAACGTCGGATTCGGAAAGAGCGATCGAGGCCAAGAAGAAGCAACAACATCA
TGAGCCTACGAAGAAGGAAGGAGAGAAATCATCGAGAAGGAATATGCCCACTCTCACTT
CGGTCCTCACTGCTCGAAGCAAGGCAGCTTGTGGGATGACTAAAAAACCCAAAGAGAAG
GTCGTCGACATCGACGCGGGTGACGCCCACAACGAGTTGGCTGCATTCGAATATATCGA
AGACATCTACACTTACTACAAGGAGGCTGAGAATGAGAGCCTGCCCCGTAACTACATGT
CGTCCCAGCCGGAGATAAACGAGAAGATGAGGGCGATCCTGGTGGACTGGCTGATCGA
GATCCACAACAAGTTCGACCTCATGCCCGAGACCCTCTACCTCACCATCAACATCATCGA
CCGGTTCCTCTCGGTGAAGGCAGTCCCGAGGCGGGAGCTCCAGCTGCTGGGCATGGG
CGCTCTCTTCACCGCCTCCAAGTACGAAGAGATCTGGGCTCCCGAGGTGAACGATCTGG
TGTGCATCGCGGACAGAGCCTACAGCCATGAGCAAGTGCTGGCGATGGAGAAGACGAT
CCTGGGGAAGCTGGAGTGGACGCTCACCGTACCCACCCACTACGTCTTCCTCGTTCGAT
TCATCAAGGCGTCGCTCGGCGATCGCAAATTGGAGAACATGGTGTACTTCCTGGCCGAG
CTCGGGGTGATGAACTACGCGACGCTCACGTACTGCCCCTCCATGGTGGCGGCCTCGG
CGGTGTACGCGGCGCGGTGCACTCTCGGCCTAACCCCGCTATGGAACGACACCCTCAA
GCTTCACACCGGCTTCTCCGAATCCCAGCTCATGGACTGCGCGAGGCTGCTGGTGGGA
TACCACGCGAAGGCGAAAGAGAACAAGCTGCAGGTGGTGTACAAGAAGTACTCGAGTTC
TCAGAGAGAAGGCGTGGCCCTGATCCCGCCGGCCAAGGCTCTGCTTTGCGAAGGCGGT
GGTCTGTCGTCTTCTTCGTCGTTGGCGTCCTCTTCTTGACGAAAGGCTCTGCTATAACTA
CAAGAGAGAGAGGAGAGACGAGTGCACACAGCTTTTGAGTTTGTTCTCGGCCAGAGAAA
AATGACAGATTGATATCGATGATGATGACTGTCGTGTCATCAGTAGTGTGCTTTTTTTTTA
CAGTCACGGATTGGTGACTTTTGTTGTTGACTGGAATTCTTCTCCTTTCACTTGTTTGAAG
ATTCGTTTAATTCGTTGAAGGAAACATTGGCTGCTGTAAAAAAAAAAAAAAA

SEQ ID NO: 20
CTCTCTCTCTCTCTCTCTCTCTCCTGCTCGTAAGATCAAAGGAGGTCGACGACGCTCTCT
CTCTCTCTCTCTAGGTCTCTGGGTGGCGGTGGGTGTTCGGGAGGGAGAGAGAAGGGGG
AGAGGTAGAGGGGGAGAGAGAAGGGAGAGATTGTGTTCTGAGGAGTTTATTGTTCTTCT
TACCAGATAAGTGAGGCAAAACGAAAAATGGGTTTGCCTGATGAGAACAATGCTGCACT
GAGCAAACCCACGAATCTCCAAGTTGGAGGGTTGGAGATCGGAGGCAGGAAATTCGGG
CAGGAGATTAGGCAGACCAGGAGGGCGTTGAGCGTGATTAACCAGAATTTAGTTGGAGA
TCGTGCTTACCCATGTCATGTCGTCAACAAGAGAGGCCACTCCAAAAGAGATGCAGTGT

GCGGGAAGGATCAGGTGGATCCAGTTCACAGACCACTTACAAGGAAGTTCGCAGCGCA
AACGGCGAGCACGCAGCAACATTGCATCGAGGAAGCAAAGAAGCCAAGAACGGCAGTT
CAAGAGAGAAATGAATTTGGAGACTGCATATTTGTGGACGTGGAGGACTGCCAACCATC
TTCAGAAAATCAACCAGTCCCTATGTTTCTGGAAATCCCAGAATCAAGATTGGATGATGA
CATGGAGGAAGTGGAAATGGAAGATATAGTGGAGGAGGAAGAGGAAGAGCCAATTATG
GACATCGATGGTCGAGATAAGAAGAACCCACTTGCAGTGGTTGACTACATTGAAGACATT
TATGCTAACTATAGGAGAACGGAGAATTGTAGCTGCGTCTCTGCTAACTACATGGCACAA
CAAGCTGACATTAATGAGAAGATGAGGTCCATCCTCATTGACTGGCTTATAGAGGTGCAC
GACAAATTCGATCTCATGCACGAGACTTTGTTCCTCACGGTCAATCTCATAGACAGATTC
TTGGCTCGGCAATCTGTCGTGAGAAAGAAGCTCCAGTTGGTGGGGTTGGTGGCCATGTT
ATTAGCATGCAAGTATGAAGAAGTCTCTGTTCCTGTTGTAGGAGACTTGATTCTAATATC
GGACAAGGCTTACACCAGAAAGGAAGTTCTTGAAATGGAAAGTTTGATGCTCAACTCGTT
GCAGTTCAACATGTCAGTGCCTACTCCGTATGTTTTCATGCGACGGTTTCTCAAGGCTGC
TGAATCTGATAAAAAGCTTGAGGTGCTATCCTTCTTCTTGATTGAGCTCTCACTGGTGGA
GTACGAGATGGTCAAGTTCCCGCCATCTCTTCTAGCCGCCGCTGCGATCTTTACAGCTC
AGTGCACCCTCTATGGTTTCAAACAGTGGACCAAGACTTGTGAATGGCACAGCAACTAC
ACAGAAGACCAGCTCCTAGAATGTGCAAGGATGATGGTAGGTTTCCATCAAAAAGCTGC
AACAGGGAAGCTCACAGGAGTACATAGAAAGTACGGTACATCAAAGTTCGGTTACACAT
CAAAATGTGAACCTGCTAACTTTCTATTAGGAGAGATGAAGAATCCATAGCAGGAGCTCA
GTGTACAGCTAACCTTTTTTTCACTGAGAAAAATTGGGTGAAAAATAAATTTTTAGCAACT
TGGGTTGCGCTTGGGAATTGGTAAAAAGGAGAAAATTCTCTAGATTTAAGCCTAACCCTC
GGTTGTTAGTCTCTTACAGCCAATTGTTTCAGATATCTCATCATTGTTTGTGATTGATTGA
ACACTAGACTAATGTTTTTCCTTTCCAATTGTAGTTCGTCTTTTATTGTAACAATAAATTGA
TAGATACTGATTCGAAATAATATTTCTCATAAAAAAAAAA

SEQ ID NO: 21
GTAAGATCAAAGGAGGTCGACGACGCTCTCTCTCTCTCTCTAGGTCTCTGGGTGGCGGT
GGGTGTTCGGGAGGGAGAGAGAAGGGGGAGAGGTAGAGGGGGAGAGAGAAGGGAGA
GATTGTGTTCTGAGGAGTTTATTGTTCTTCTCACCAGATAAGTGAGGCAAAACGAAAAAT
GGGTTTGCCTGATGAGAACAATGCTGCACTGAGCAAACCCACGAATCTCCAAGTTGGAG
GGTTGGAGATCGGAGGCAGGAAATTCGGGCAGGAGATTAGGCAGACCAGGAGGGCGTT
GAGCGTGATTAACCAGAATTTAGTTGGAGATCGTGCTTACCCATGTCATGTCGTCAACAA
GAGAGGCCACTCCAAAAGAGATGCAGTGTGCGGGAAGGATCAGGTGGATCCAGTTCAC
AGACCACTTACAAGGAAGTTCGCAGCGCAAACGGCGAGCACGCAGCAACATTGCATCG
AGGAAGCAAAGAAGCCAAGAACGGCAGTTCAAGAGAGAAATGAATTTGGAGACTGCATA
TTTGTGGACGTGGAGGACTGCCAACCATCTTCAGAAAATCAACCAGTCCCTATGTTTCTG
GAAATCCCAGAATCAAGATTGGATGATGACATGGAGGAAGTGGAAATGGAAGATATAGT
GGAGGAGGAAGAGGAAGAGCCAATTATGGACATCGATGGTCGAGATAAGAAGAACCCA
CTTGCAGTGGTTGACTACATTGAAGACATTTATGCTAACTATAGGAGAACGGAGAATTGT
AGCTGCGTCTCTGCTAACTACATGGCACAACAAGCTGACATTAATGAGAAGATGAGGTC
CATCCTCATTGACTGGCTTATAGAGGTGCACGACAAATTCGATCTCATGCACGAGACTTT
GTTCCTCACGGTCAATCTCATAGACAGATTCTTGGCTCGGCAATCTGTCGTGAGAAAGAA
GCTCCAGTTGGTGGGGTTGGTGGCCATGTTATTAGCATGCAAGTATGAAGAAGTCTCTG
TTCCTGTTGTAGGAGACTTGATTCTAATATCGGACAAAGCTTACACCAGAAAAGAAGTTC
TTGAAATGGAAAAGTTGATGCTCAACTCGTTGCAGTTCAACATGTCAGTGCCTACTCCGT
ATGTTTTCATGCGACGGTTTCTCAAGGCTGCTGAATCTGATAAAAAGCTTGAGGTGCTAT
CCTTCTTCTTGATTGAGCTCTCACTGGTGGAGTACGAGATGGTCAAGTTCCCGCCATCTC
TTCTAGCCGCCGCTGCGATCTTTACAGCTCAGTGCACCCTCTATGGTTTCAAACAGTGGA
CCAAGACTTGTGAATGGCACAGCAACTACACAGAAGACCAGCTCCTAGAATGTGCAAGG
ATGATGGTAGGTTTCCATCAAAAAGCTGCAACAGGGAAGCTCACAGGAGTACATAGAAA
GTACGGTACATCAAAGTTCGGTTACACATCAAAATGTGAAGCTGCTAACTTTCTATTAGG
AGAGATGAAGAATCCATAGCAGGAGCTCAGTGTACAGCTAACCTTTTTTTCACTGAGAAA
AATTGGGTGAAAAATAAATTTTTAGCAACTTGGGTTGCGCTTGGGAATTGGTAAAAAGGA
GAAAATTCTCTAGATTTAAGCCTAACCCTCGGTTGTGAGTCTCTTACAGCCAATTGTTTCA
GATATCTCATCATTGTTTGTGATTGATTGAACACTAGACTAATGTTTTTCCTTTCCAATTGT
AGTTCGTCTTTTATTGTAACAATAAATTGATAGATACTGATTCGAAAAAAAAAAA

SEQ ID NO: 22
GAGACGTTGGCTTTCTGGCTTAAAGGCTATTCTTTGTGCACAATGACCTGAGGGAGGTC
TCGACAGACCACTTCTTCTCCGCCAAAGAAGAAGATGGCGATGGTACAGCGACAAGGTC

ACGACCCATCATCGCCGCAGGAGCAAGAAGACGGTCCTTCCTCCTTCTTGTCCGACGAT
GCTCTCTACTGTGAAGAAGGCAGATTCGAAGAAGACGACGGCGGCGGCGGCGGCCAAG
TTGACGGAATTCCACTCTTCCCCTCACAGCCGGCGGATCGACAGCAAGACTCGCCGTG
GGCAGACGAAGACGGCGAGGAGAAGGAGGAGGAGGAGGCGGAGCTGCAGTCGCTCTT
CTCCAAGGAGCGCGGAGCGAGGCCGGAGCTCGCGAAAGACGACGGGGGCGCCGTCG
CGGCGCGGCGGGAGGCCGTGGAGTGGATGCTGATGGTGAGGGGCGTCTACGGGTTCT
CGGCGCTCACGGCGGTGCTGGCCGTCGATTACCTCGACCGGTTCCTCGCCGGGTTCCG
CCTGCAGCGGGACAACAGGCCGTGGATGACGCAGCTGGTGGCCGTCGCGTGCCTCGC
CCTGGCCGCCAAGGTGGAGGAGACCGACGTCCCTCTCCTCGTAGAACTCCAAGAGGTC
GGGGACGCGAGGTACGTGTTCGAGGCGAAGACGGTGCAGCGGATGGAGCTCCTGGTG
CTGTCGACGCTCGGGTGGGAGATGCACCCGGTGACGCCGCTGTCGTTCGTCCACCACG
TCGCCAGGAGGCTCGGCGCGAGCCCCCACCATGGGGAGTTCACCCACTGGGCGTTCCT
CCGCCGCTGCGAGCGGCTCCTCGTCGCCGCCGTCTCCGATGCGAGATCGCTGAAGCAT
CTCCCGTCGGTCCTGGCCGCGGCGGCGATGCTGCGCGTCATCGAGGAGGTCGAGCCG
TTTCGTTCCTCGGAGTACAAGGCCCAGCTCTTGAGTGCCCTCCACATGAGCCAGGAAAT
GGTGGAAGACTGTTGCAGATTCATTCTGGGAATAGCAGAGACCGCGGGCGATGCCGTG
ACCTCGTCCCTCGACAGCTTCCTGAAGCGCAAGCGTCGCTGTGGTCACCTTAGCCCGA
GGAGCCCGAGCGGGGTCATCGACGCCTCGTTCAGCTGCGACGACGAGTCGAACGACTC
GTGGGCCACCGACCCGCCATCCGATCCGGACGACAACGACGATCTGAACCCTCTACCG
AAGAAGAGCAGGTCGTCGTCGCCGTCCTCCTCCCCCTCCTCGGTGCCAGACAAGGTGT
TGGACTTGCCCTTCATGAACAGGATCTTTGAGGGCATCGTCAACGGCAGTCCTATCTGA
TCGTCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAGAATTTGTATCGACCC
TTTTCAATTAAATCAAAGTGAAGAAAATGTGAAGTGAAAGATGAGAGCTTTGCGTTGAAG
AAACGGGAAGGGTCTGCGCTTACGTATGCATGTCTTTTTTTGGCGCTCCCTCTCGGTCT
CTTCAATGATCTTGAAGTGTCCCTTTCACTTCAGAATTTGCTTCATGTATGGGACATGGAC
AGGAGATATATATATTATGTCACCATTACAATAAAAAATTACAATTTTTCGGAAAAAAAA

SEQ ID NO: 23
AGATGGCTCCGAGAGCAGGCCCATTCTTGGACCTCTCTTTCTTCTTTATCTACACCGTCT
CCATTGCCCCTCTCATGGAGTAATGATCTTGTCGTTCATCGACCCTTCTTCTTCTAAGAA
GAAGCAGAAGCAGAAGCAGAAGCAGAAGCAGGAGGAGGAGGAGCAGAAGGAGGAGAT
GGAGGCCAGTTATCAACCCCACCACCATGGTCATCTTCGTCAACACGACCCATCGAGCT
CTCAACAAGAAGAGCAAGTCCCTTTCGACGCCCTTTACTGCTCGGAGGAGCACTGGGGA
GAAGAAGACGAAGAAGAGGGATTGGCGAGCGATGGGCTCTTGTCGGAAGAGAGAGATC
ACAGATTGCTGAGCCCGCGAGCCTTGCTCGATCAGGACCTGCTCTGGGAAGACGAGGA
GCTGGCCTCCCTCTTCTCCAAAGAGGAGCCGGGCGGCATGCGCTTGAATCTGGAGAAC
GACCCGTCTTTGGCCGATGCTCGCCGCGAGGCCGTGGAGTGGATAATGAGAGTCCACG
CGCACTACGCGTTCTCCGCTCTCACGGCTCTGCTCGCGGTGAACTACTGGGATAGGTTC
ACGTGCAGCTTCGCCTTGCAGGAAGACAAGCCGTGGATGACTCAGCTCTCCGCCGTCG
CTTGCCTCTCTCTCGCCGCCAAAGTGGAGGAGACCCAAGTGCCTCTTCTCATCGATTTC
CAAGTCGAGGACAGTAGCCCCGTCTTTGAGGCGAAGAACATACAGAGAATGGAGCTCCT
GGTGCTCTCGTCGCTCGAATGGAAGATGAATCCAGTGACCCCGCTATCGTTTCTTGATTA
CATGACAAGGAGACTAGGGCTGACGGGCCATCTGTGCTGGGAGTTTCTTAGAAGGTGC
GAGAACGTCCTCCTCTCTGTAATCTCAGATTGCAGATTCACGTGTTATCTTCCTTCAGTG
ATAGCTGCTTCCACAATGCTGCACGTAATCAATGGCCTAAAGCCTCGTCTCGATGTCGAA
GACCAAACCCAGCTCCTGGGAATCCTAGCAATGGGCATGGACAAGATCGATGCTTGCTA
TAAGCTCATCGACGACGACCACGCATTGAGAAGCCAGAGATATTCCCACAACAAGCGCA
AGTTTGGATCGGTCCCCGGGAGCCCCAGAGGAGTAATGGAATTGTGCTTCAGCTCCGAT
GGCTCCAACGATTCTTGGTCCGTGGCGGCCTCGGTATCGTCCTCCCCGGAGCCTCACT
CCAAGAAGAGCAGAGCCGGCGAGGAAGCCGAAGACCGGCTTCTGCGGGGGCTCGAGG
GCGAGGAAGATGACCCCGCGAGCGCGGATATCTTCAGCTTCCCTCACTAGCGGGCACT
TCATCCGGGTCCTGGTTATCATACTCTTATATATGTTGGGGAATAACGGTTCATATGTTTC
ATGTAATGCGCAAGTTAAATTGCCAACCATCTCTCGGTTCCAGCCTAATTCCCCAGAACG
ATTGAAAACACACCAGATCGGAAGGTTGCAAGAACCATTCTCGACTGGGATTGCTGGGT
TCCTAGGCTTTTGAGATTTGGAATGGAGATGGTTGGCATTTTATCCGGATTATAACTTTAG
GGATTATGAACAAAATGAAGAGAAGAAATCGTTGGGAAGGATCCTGCTGTTCCTGTTGTT
CCTTTTTTCCCTGTACGGTCCCCCCTCCTTCTTTCCCTATTTTCTGTATCGAAGCACTTCT
GCAGCAGTTCAGCTATAGGTAGTACTTTGAAATAAAAGAGCTCATCATGAAGGGTAAGCA
TTTTTGTTTCAAAAAAAAAAA

SEQ ID NO: 24
TCTCTCTCTCCCTCTCTCTCTCTCTCCATCCCACATTCTTGAAAGAGACACATCAGAAG
CGCAGACATGAATCTGGGCACCGGCTCGTTCTCTTTCCTGCTCCATCTATAATGAGCAAA
GCCTTCGCTCTCTTCTCCGCAGAAGGGCATGGCTCTGCAGGAGGAGGACACCCGCCGC
CACTACCCGACCGCTCCTCCGTTCTCGCCCGACGGCCTCTACTGCGAGGACGAGACCT
TCGGTGAAGATCTGGCCGACAATGCCTGCGAATACGCCGGCGGGGGAGCCCGGGATG
GCCTCTGCGAGATAAAGGACCCGACTTTGCCTCCGTCGTTGCTCGGGCAGGACTTGTTT
TGGGAAGATGGCGAGCTCGCCTCCCTGGTTTCGAGAGAGACCGGGACGCACCCCTGCT
GGGACGAGCTAATCTCTGATGGGTCAGTGGCGCTTGCTCGGAAGGATGCGGTCGGGTG
GATTTTGAGGGTCCATGGGCATTACGGGTTCCGTCCGTTGACTGCTATGTTGGCCGTGA
ACTACTTGGACAGGTTCTTTTTGAGCCGGAGTTACCAGAGGGACAGGCCTTGGATAAGC
CAGCTCGTAGCTGTGGCTTGTCTCTCCGTTGCTGCCAAAGTGGAGGAAACCCAAGTGCC
TATTCTCCTTGACCTGCAAGTGGCTAATGCGAAATTTGTGTTTGAATCGAGGACGATTCA
GAGAATGGAGCTCTTGCTGATGTCTACACTTGATTGGAGAATGAACTCGGTGACTCCGA
TTTCGTTCTTTGATCACATCCTTAGGAGGTTTGGCTTGACGACTAATTTGCACAGGCAGT
TTTTCTGGATGTGCGAGCGTTTACTTCTCTCAGTGGTTGCAGATGTGAGGCTTGCAAGTT
TTCTTCCGTCAGTTGTTGCCACGGCTGCAATGTTGTATGTTAACAAGGAGATAGAACCGT
GTATATGCAGTGAATTCCTGGACCAGTTACTGAGCCTGCTCAAGATCAATGAGGACCGA
GTAAATGAGTGCTATGAACTCATTCTTGAATTGTCAATCGACCATCCTGAGATCCTCAACT
ACAAGCACAAGCGTAAGAGGGGATCAGTACCCAGTAGCCCCAGTGGTGTGATTGATACT
TCTTTCAGCTGTGACAGTTCGAATGATTCGTGGGGTGTGGCATCGTCTGTTTCTTCTTCA
CTGGAGCCTCGGTTCAAGAGGAGCAGATTCCAGGATCAACAGATGGGCCTGCCATCTGT
GAATGTTTCATCCATGGGTGTGCTTAATAGTTCTTATTAGTCTTAGCTTATTATCTTTGATT
GGACATGCTATAAAGCAGTCTTTCCCCTGCTTCAAAAAAAAAA

SEQ ID NO: 25
GGTTCCTCCTCCTCCTCCTCTCTCCATTTCCGCTCTCGATTCCTCTCCTTCCGCTAAACC
CCGCAGCTTCGGAATTCCCGCTCGGGTTTTCGCTGAGAAGATGGGCCAGATCCAGTACT
CCGAGAAGTACTTCGACGACACCTACGAGTACAGGCATGTGGTTCTCCCTCCTGATGTG
GCCAAGCTTCTCCCGAAGAACCGCCTCCTTTCTGAAAATGAATGGCGTGCCATTGGAGT
GCAGCAGAGTCGTGGTTGGGTTCACTATGCAATTCATCGCCCTGAACCACACATAATGC
TATTCAGGAGGCCTCTGAATTACCAGCAGCAGCAAGAGAACCAGGCCCAGCAAAACATG
CTTGCTAAGTAGACATGTTATATTTCTAATGCTTTGAGAACAATATTACAGTATAATTAGG
GTTGGAAGCTTTAGTAAATGTTAGGATGTTTTGAAACTTGTCATTGTAATTGGCAGCAATT
CTCCTCTTTGGAGAATCTATCGTGGGACTTTGTTTAAAAAAAAAAA

SEQ ID NO: 26
CCAGATCCATGGGTTCGATTGACCCGCCAAAAGCCGAACAGAACGGCACCGCGGCCGC
CGCCGTCGCCGATCCCGGCCAGAAGCCCGGCGCCGGAGACGCCATGCCGCCGCCGCC
GCCCGTCAAGCACTCCAACGGGACCGCGGCGGAGCCCGATGTTGCGACGAAGAGGAG
GAGGATGAGCGTGCTTCCCCTCGAGGTGGGCACGCGCGTGATGTGCCGCTGGCGAGA
CGGCAAGTATCACCCCGTGAAGGTCATCGAGCGGCGGAAGCTGAATCCCGGAGATCCT
AACGACTACGAGTACTATGTTCATTACACGGAATTCAACCGGAGGCTCGATGAATGGGT
GAAACTTGAACAGCTTGATCTGAATTCTGTAGAGACTGTGGTCGATGAAAAAGTGGAGG
ACAAGGTGACGGGGTTAAAAATGACACGTCACCAAAAGCGGAAGATTGATGAGACTCAT
GTCGAGGGGCATGAGGAGCTTGATGCTGCCAGCTTGCGTGAACATGAAGAATTCACCAA
AGTGAAGAATATCGCAACCATAGAGCTTGGAAGATACGAGATCGAGACATGGTACTTCT
CCCCCTTTCCACCAGAATACAATGATTGTTCGAAGCTCTACTTTTGTGAGTTTTGCCTCAA
TTTTATGAAGCGCAAGGAACAGCTTCAAAGGCACATGAAGAAGTGTGATCTTAAACATCC
CCCTGGAGATGAGATATACAGGAGCGGTACATTGTCAATGTTTGAGGTTGATGGCAAGA
AGAATAAGGTTTACGGGCAGAATCTTTGCTATTTGGCAAAACTCTTTCTAGATCACAAGA
CCCTGTATTATGATGTAGATCTTTTTCTATTTTATGTCTTGTGCGAATGTGATGATCGAGG
TTGCCACATGGTTGGATATTTCTCCAAGGAAAAGCATTCCGAGGAATCCTACAATTTGGC
TTGCATCCTCACACTCCCACCATATCAGAGGAAAGGCTATGGAAAGTTCTTAATAGCTTT
TTCATATGAACTTTCCAAAAAAGAAGGAAAAGTCGGCACGCCAGAGAGGCCTCTGTCTG
ATCTTGGGCTGTTGAGCTACAAAGGATACTGGACAAGGGTTCTGTTGGACATTTTGAAAA
AGCATAAAGCAAATATTTCTATCAAGGAGCTCAGTGACATGACAGCAATAAAGGCTGATG
ATATTTTGAACACCCTTCAGAGCCTAGACTTAATTCAGTATAGAAAAGGGCAGCACGTCA
TATGTGCAGACCCGAAAGTCTTGGATCGTCATCTGAAAGCTGCTGGACGAGGTGGCCTA
GAGGTTGACGTCAGCAAACTGATCTGGACTCCATACAGAGAGCAAGGTTGATATTAACT

GGCAAGTTAGGATAGGCATTTCCTTTGGCTTGTATCATTACACAGTTTGTACAAAATGCA
CAGCTTGTAGTATCTCAGGCGGAATAAGCCTATGTTTCTAAAGTCGCTGACCTCACGGTC
CAATACTTTGGCCTTTGATTTGGTAGTTTTGTGGAAAATCATCGACTAGACCGATGGTCA
AAGTGGTAATCATGTAATTAAACGCGTTTGTCATTGCTCGGCAGGAA

SEQ ID NO: 27
CGCACGAGGCTCTCTCTCTCTCTCTCTCTCTGCGCGCAATTTCCCGCCTTTCCTCCTC
CCATTTCCCGCTCGGAGAAACCCTAACGATGGCGCAGAAGCACAGCACCGCCCCCGAT
CCGGCGGCCGAACCCAAGAAGCGGCGGCGCGTCGGCTTCTCCGGCATCGATGCTGGA
GTCGACCCGAACGGCTGCTTCAAAGTGTACCTCGTGTCCAGGGAAGAAGAAGTGGGGG
CTCCGGATAGCTTCTGCCTTGATCCAGTTGACTTAAGTCACTTCTTTGAGGAGGAAGATG
GAAAAATCTATGGATATGAAGGATTGAAGATATCCGTCTGGGTAAGCTGTGTATCATTTC
ATTCATATGCAGAGATTGCCTTTGAGAGCAAATCGGATGGAGGAAAAGGAATCACAGAT
CTGAACACTGCTCTTAAGAATATGTTTGGTGAAACTCTCGTGGATAATAAAGATGACTTC
CTTCAAACTTTTTCCAAGGAGACCCAATTTATTAGGTCTACAGTTTCAGCCGGGGAGATT
TTAAAGCATAAACACTCTGACGATCATGTCAACGATTCCGTTAGTAATCTAAAAGTTGGTT
CTGATGTCGAGGCTGTGCGCATGCTGATGGGCGATATGACAGCAGGACACCTGTATAGT
CGGTTGGTCCCTCTTGTTCTGCTTCTTGTAGACGGTAGCAGTCCTATTGATGTAACGGAT
TCAAGTTGGGAGCTATATCTTCTTATTCAGAAGACAAGTGATCAGCAAGGCAATTTTCAT
GATAGGCTTCTTGGTTTTGCTGCTGTATATCGTTTCTATCACTATCCTGATAGTTCGCGAC
TGCGGCTTGGTCAGATTCTAGTATTACCACTTTACCAGCGCAAAGGCTATGGCCGCTATC
TCCTGGAGGTGCTTAACAATGTCGCAATAGCTGATGATGTTTATGACTTCACAATAGAAG
AGCCAGTGGATAATCTTCAACACTTGCGAACATGTATCGATGTGCAACGGCTTCTGAGCT
TTGACAAAGTCCAGCAGGCAGTAAATTCAACAGTATCTCAGTTGAAGCAAGGCAAACTAT
CAAAGAAAACCTACATCCCTCGGTTGTTGCCTCCTCCTAGTGTGGTTGAGGATGCCAGG
AAGCGTTTCAAAATCAACAAGAAGCAGTTTCTTCAATGTTGGGAGATTCTAGTCTATCTTG
GGCTTGATCCCGCCGATAAGAGCATACAGGATTATTTTTCCGTCATTTCAAACCGTGTCA
GGGCAGACATTTTAGGAAAAGACTCTGAGACTGCCGGAAAGAAAGTGATTGAAGTACCG
AGTGATTTTGATCCAGAGATGTCCTTCGTCATGCATAGGGCAAAAGCAGGCGGTGAAGC
TAATGGTATCCAAGTGGAGGACAACCAGAACAAGCAAGAAGAGCAGCTGCAGCAGTTAA
TCGATGAAAGATTGAAGGACATCAAGCTGATCGCTGAGAAGGTAACTCAGAAATGATCA
CAGAAAAAAGGTCATGTAATACAAATGTTGTAGCCCTTGGAATGGAATTAACAGAGGTCC
TGTAGATTGACTGAGGTGGGCATTGCCCTTTTAGCTTATGATTTGAGAAACCCTTGCAGG
CTGCGATTAGCGGATCATGAGAGCATAGTTTTGCTTATGCTCAATCCCTAGTTTTGGGTC
CAATTTCATTAGAGGTAGATCATTTCCTGTTTTTATAACCTTCGTGTAAGTTAGGGAAGGA
GCTGCGGTCGCGGTCCTGGCATTTTCAAGCGCCTGTCTCTCGGTTCAAAAAAAAAAA

SEQ ID NO: 28
CTCTTCCTCCCTCCCTCTCTCTCTGCGCGCAATTTCCCGCCTTTCCTCCTCCCATTTCCC
GCTCGGAGAAACCCTAACGATGGCGCAGAAGCACAGCACCGCCCCCGATCCGGCGGC
CGAGCCCAAGAAGCGGCGGCGCGTCGGCTTCTCCGGCATCGATGCTGGAGTCGACCC
GAACGGCTGCTTCAAAGTGTACCTCGTGTCCAGGGAAGAGGAAGTGGGGGCTCCAGAT
AGCTTCTGCCTTGATCCAGTTGACTTAAGTCACTTCTTTGAGGAGGAAGATGGAAAAATC
TATGGATACGAAGGATTGAAGATATCCGTCTGGGTAAGCTGTGTATCATTTCATTCATAT
GCAGAGATTGCCTTTGAGAGCAAATCGGATGGAGGAAAAGGAATCACAGATCTGAACAC
TGCTCTTAAGAATATGTTTGGTGAAACTCTCGTGGATAATAAAGATGACTTCCTTCAAACT
TTTTCCAAGGAGACCCAATTTATTAGGTCTACAGTTTCAGCTGGGGAGATTTTAAAGCAT
AAACACTCTGACGGTCATGTCAACGATTCTGTTAGTAATCTAAAAGTTGGTTCTGATGTC
GAGGCTGTGCGCATGCTGATGGGCGATATGACAGCAGGACACCTGTATAGTCGGTTGG
TCCCTCTTGTTCTGCTTCTTGTAGACGGTAGCAATCCTATTGATGTAACGGATTCAAGTT
GGGAGCTATATCTTCTTATTCAGAAGACAAGTGATCAGCAAGGCAATTTTCATGATAGGC
TTCTTGGTTTTGCTGCTGTATATCGTTTCTATCACTATCCTGATAGTTTGCGACTGCGGCT
TGGTCAGATCCTAGTATTACCACTTTACCAGCGCAAAGGCTATGGCCACTATCTCCTGGA
GGTGCTTAACAATGTCGCTATAGCTGATGATGTTTATGACTTCACAATAGAAGAGCCAGT
GGATAATCTTCAACACTTGCGAACATGTATCGATGTGCAACGGCTTCTGAGCTTTGACAA
AGTCCAGCAGGCAGTAAATTCAACAGTATCTCAGTTGAAGCAAGGCAAACTATCAAAGAA
AACCTACATCCCTCGGTTGTTGCCTCCTCCTAGTGTGGTTGAGGATGCCAGGAAGCGTT
TCAAAATCAACAAGAAACAGTTTCTTCAATGTTGGGAGATTCTAGTCTATCTTGGTCTTGA
TCCAGCCGATAAGAGCATACAGGATTATTTTTCCGTCATTTCAAACCGTGTCAGGGCGGA
CATTTTAGGAAAAGACTCTGAGACTGCCGGGAAGAAAGTGATTGAAGTGCCGAGTGATT

TTGATCCAGAGATGTCGTTCGTCTTGCATAGGGCAAAAGCAGGTGGCGAAACTAATGGT
ATCCAAGTGGAGGACAACCAGAACAAGCAAGAAGAGCAGCTGCAGCAGTTAATCGATGA
AAGATTGAAGGACATCAAGCTGATCGCTCAGAAGGTATCTCGGAAATGATCACAGAAAAA
AGGTCATGTAATACAAATGTTGTAGCCCTTGGAATGGAATTACCAGAGGTCCTGTAGATT
GACTGAGGTGGGCATTGCCCTTTTAGCTTATGATTTGAGAAAACCCTTGCAGGCTGCGA
TTTGCGGATCATGACAGCATAGTTTTGCTTATGCTCAATCCCTAGTTTTGGGTCCAATTTC
ATTAGAGGTAGATCATTTCCTGTTTTTATAACCTTCGTGTAAGTTAGGGAAGGTGCTGCAT
TCGCGGTCCTGGCATTTTCAAGTGCCTATCTCTTGGTTCAAAAAAAAAA

SEQ ID NO: 29
CAACGAGCGAAAATCGCTATAAATTCCCTCCTAAACCCTCGCCTCACCCTTCGTCCCTGA
ACCCTCGATCTCTCTCTAGCTAGCTTTTGGCGCCCGGCAGAGCTTCCTCTCCTCCCCTC
GCCTCGACTCGCCTCCTCTCTTCGAGCTAGAAAGCTCTCGATGGCCCTTCCGATGGAGT
TCTGGGGAGTTGAAGTGAAGGCTGGACAGCCCCTTAAAGTCAACCCTGGCAATGCTAAA
ATCTTGCATCTTTCTCAGGCATCACTTGGTGAATGCAAGAGTAGCAAAGGAAATGAATCA
GTGCCTCTCCATGTGAAGTTTGGCGATCAGAAGCTTGTTTTAGGAACTCTCTCCACGGA
GAACTTCCCTCAATTAGCATTCGACTTGGTTTTTGAGAAAGAGTTTGAACTATCTCACAAC
TGGAAAAGTGGAAGTGTCTACTTTTGTGGATACAAGTCTGTCGTACATGATGATGATGAT
GAATTTTCTGATTTGGAGAGTGATTCAGAAGAGGAAGATCTTCCGATGATTGGTGTGGAA
AATGGAAAGGTTGCAGCACAAGCATCAGCTAAGACTGCTACTGCCAGTGCTAATGCTAG
CAAGGTTGAATCATCGGGAAAGCAAAAGGCCCGCATTCCACAACCAATGAAAGTTGATG
AGGATGACAGTGATGAGGATGATGACGACGAAGATGAAGATGAATCTGATGAGGAGGG
AGTTGATGGTGAGGCTGATTCTGATGAGGAGGAAGATGAAAGTGATGAGGAAGAGACTC
CAAAGAAGGCTGAAATAGGCAAGAAGAGAGCTGCGGATTCTGCAACTAAGACACCTGTC
CCTGCCAAGAAGTCAAAGTTACCTACTCCACAGAAGACAGATGGTAAGAAAGGTGGCCA
TACAGCAACTCCTCACCCTGCAAAACAGGCTGGAAAGAATCCTGCCAACAGTGCCAACA
AGTCGCAAAGCCCCAAATCAGCTGGCCAAGTCTCTTGCAAATCATGTAGCAAGACGTTC
AATTCAGACGGTGCTCTTCAGTCTCATTCAAAGGCTAAGCATGGTGGCAAGTAAATTAAA
GAACCAAGATAACCATCCGGAAATAGTCGGATGTTTTTTGTTCGTCATCTTCTGTCCCCC
CAAAAGTAGTGTTTGAGGTGTATCTCGGTGGAGGTTTTGTTGTGAGGGCTTGGTAGGTTT
TCATTCTATTGTAATGTCGGCGACGGAGATTTTGGATGGGGATTCTTTATTCTAATCAGC
AGTAATTAGTCATTGGGTTGTCAACTTTTCCCTCTTTCCTTCACTTGGTCGTGGTACATTT
TGCATGTATTGTGATGGATTTTTGACTAAAAAAAAAA

SEQ ID NO: 30
CCTAGCTCGTTCCCCGCCCTGGCAACGAGCGAAAATCGCTATAAATTCCCTCCTAAACC
CTCGCCTCGCTCTTCCTCCCTGAACCCTCGATCTCTCTCTAGCTAGCTTTTGGCGCCCG
GCAGAGCTTCCTCTCCTCGCCTCGCCTCCTCTCTTCGAGCTAGAAAGCTCTCGATGGCC
CTTCCGATGGAGTTCTGGGGAGTTGAAGTGAAGGCTGGACAGCCCCTTAAAGTCAACCC
TGGCAATGCTAAAATCTTGCATCTTTCTCAGGCATCACTTGGCGAATGCAAGAGTAGCAA
AGGAAATGAATCAGTGCCTCTCCATGTGAAGTTTGGCGATCAGAAGCTTGTTTTAGGAAC
TCTCTCCACGGAGAACTTCCCTCAATTAGCATTCGACTTGGTTTTTGAGAAAGAGTTTGA
ACTATCTCACAACTGGAAAAGTGGAAGTGTCTACTTTTGTGGATACAAGTCTGTCGTACA
TGATGATGATGATGAATTTTCTGATTTGGAGAGTGATTCAGAAGAGGAAGATCTTCCGAT
GATTGGTGTGGAAAATGGAAAGGTTGCAGCACAAGCATCAGCTAAGACTGCTACTGCCA
GTGCTAATGCTAGCAAGGTTGAATCATCGGGAAAGCAAAAGGCCAGCATTCCACAACCA
ATGAAAGTTGATGAGGATGACAGTGATGAGGATGATGATGAGGACGATGACGACGAAGA
TGAATCTGATGAGGGAGTTGATGGTGAGGCTGATTCTGATGAGGAGGAAGATGAAAGTG
ATGAGGAAGAGACTCCAAAGAAGGCTGAAATAGGCAAGAAGAGAGCTGCGGATTCTGCA
ACTAAGACACCTGTCCCTGCCAAGAAGTCAAAGTTACCTACTCCACAGAAGACAGATGG
TAAGAAAGGTGGCCATACAGCAACTCCTCACCCTGCAAAACAGGCTGGAAAGAATCCTG
CCAACAGTGCCAACAAGTCGCAAAGCCCCAAATCAGCTGGCCAAGTCTCTTGCAAATCA
TGTAGCAAGACGTTCAATTCAGACGGTGCTCTTCAGTCTCATTCAAAGGCTAAGCATGGT
GGCAAGTAAATTAAAGAACCAAGATAACCATCCGGAAATAGTCGGATGTTTTTTGTTCGT
CATCTTCTGTCCCCCCAAAAGTAGTGTTTGAGGTGTATCTCGGAGGAGGTTTTGTTGTGA
GGGCTTGGTAGGTTTTCATTATATTGTAATGTCGACGACAGAGATTTTGGATGGGGATTC
TTTATTCTAATCAGCAGTAATTAGTCATTGGGTTGTCAACTTTTCCCTCTTTCCTTCACTTG
GTCGTGGTACATTTTGCATGTATTGTGATGGATTTTTGACTAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAA

SEQ ID NO: 31
GAAAATTCATCGTCTTCCAGAAAATTCACCGCACAGTCCCCTCTCGAGCTTCTCCGCCGT
CAGCAATGGAGTTCTGGGGTGTGGAAGTCAAAAGTGGGGAACCCCTCAATGTCGAACC
CGGGGCTGAAACAGTTGTGCATTTGTCACAGGCCTGTCTCGGAGAAACCAAGGAAAAAA
CAAAGGAATCTGTACTTCTGTATGTCCATATTGGGGTTCAGAAACTCGTCCTTGGAACTC
TCTCTGCTGATAAGTTTCCCCAGATACCTTTTGATCTGGTGTTTGAGAAAAGTTTTAAGCT
TTCTCATAATTGGAAAAATGGAAGCGTCTTCTTCAGTGGATATAAGACGCTACTTCCATGT
GGATCCGACGCTGACAGTCCATATTCTGACTCCGACACTGATGAGGGTCTTCCAATTAAT
GTTACTGCTCAAGCTGACGTACCTGCGAAAAAAGCACCAGTGACTGCTAATGCCAATGC
AGCAAAGCCCAACCTGGCTTCTGCTAAGCAAAAGGTAAAGATTGTTGAATCAAATGAGGA
TGGGAAAAATGAAGGGGACGATGATGAAGATGCTGATGTATCTTCGGATGATGATGCTG
AGGATGATTCTGGTGATGAGGACATGGTCGATGGAGGTGATGAAAGTAGCGATGAAGAT
GATGATGACAGTGAGGAAGGGGAAAGTTCTGAAGAGGAAGAGCCTAAGGCTCAACCAA
GTAAGAAGAGACCTGCTGATTCTGTACTGAAAACCCCTGCGTCAGATAAGAAATCAAAAC
TGGAAACTCCTCAAAAGACTGATGGCAAGAAGGCCAGTGAGCATGTAGCAACTCCATAC
CCTTCTAAGCAAGCTGGGAAGGCGATTGCCAGTAAAGGTCAAGCTAAGCAACAAACACC
AAACTCCAACGAGTTCTCGTGCAAACCTTGCAACCGGTCATTTAAGTCGGATCAGGCTCT
TCAATCACACAACAAAGCGAAGCATGGTGGAAGCTAAAACTTTGGGAAGTCCAAAGACA
CTGGCGATTTTGTCCGACGGAGGTGGGTGATGCCAATCGAAGAGAGCTCTGCAACCAG
GATGTTCTTTTATTATTAGTTGTGGCTTTTCATGACTCCATTCTGGAACGAACTATGGGTC
TTAACGATCTGTAACGGCTTAAAACTTTTGAAGTTGAAATACTCGTGTCCTTCAGTTAGCT
GTTCGCATTTTGAGCCCAAAAAAA

SEQ ID NO: 32
GCGACGAGCGAGGGAGAGGGGAAGAGAAGGAGAGAGAGAGTCCGTGAAAAGCTTCAA
AAATCATTACTTTGGGCGGAGACATCTCACACAGGAAACAGAAGTGACGAGTGATGGAT
ACTGGAGGGAATTCTCTCCCGTCTGGGCCTGATGGGGTGAAGCGGAAGGTGTGCTACT
TCTATGACCCTGAAGTCGGCAATTACTACTTGCTCCAGCATATGCAGGTCCTCAAGCCC
GTCCCCGCGAGAGATAGGGATCTCTGCCGCTTCCATGCCGACGATTACGTCGCCTTTCT
TAGGAGTATAACCCCCGAGACGCAGCAGGATCAGCTGAGGCAGCTCAAGAGGTTCAAT
GTTGGAGAGGATTGTCCTGTGTTCGATGGCCTTCACTCTTTCTGTCAGACTTATGCCGGA
GGCTCGGTTGGCGGTGCCGTGAAGTTGAATCACGGTCTCTGCGATATCGCTATCAACTG
GGCAGGTGGGCTGCATCATGCTAAGAAGTGCGAGGCTTCTGGATTTTGCTATGTCAATG
ACATCGTGCTTGGCATATTGGAGCTCCTTAAACAGCATGAGCGCGTTTTGTATGTGGATA
TTGATATTCACCATGGGGACGGTGTTGAAGAGGCTTTCTACACAACTGATAGGGTCATGA
CCGTTTCATTTCACAAATTTGGTGATTACTTTCCCGGTACTGGCGATATACGTGACATAG
GATATGGCAAAGGGAAATATTACTCCCTCAATGTTCCATTGGATGATGGGATTGACGACG
AAAGCTACCATTCTCTATTCAAACCAATAATTGGAAAAGTAATGGAAGTGTTCAAACCCG
GTGCTGTGGTCCTTCAATGTGGCGCTGACTCTTTGTCAGGAGACCGATTAGGATGCTTC
AATCTATCAATCAAAGGACATGCAGAATGTGTCAGATACATGAGATCATTCAATGTGCCA
GTGTTGCTGCTAGGAGGCGGTGGTTACACCATTCGTAATGTTGCTCGTTGCTGGTGCTA
TGAGACTGGAGTGGCCCTTGGACTTGAAGTTGATGATAAGATGCCACAACATGAGTATTA
TGAATACTTCGGTCCTGACTATACACTTCATGTTGCTCCGAGTAACATGGAAAACAAAAA
TTCGCGACAATTGCTTGAGGAGATAAGATCGAAGCTTCTCGAAAATCTCTCTAAGCTCCA
GCATGCACCCAGTGTCCCATTCCAGGAAAGACCCCCTGATACTGAGCTCCCAGAGGCTG
ATGAAGATCAAGAAGATCCGGATGAAAGATGGGATCCAGACTCTGACATGGATGTTGAC
GAAGATCGCAAGCCCCTGCCTAGCAGAGTAAAGAGAGAACTAATAGTAGAACCTGAGGT
CAAAGACCAGGATTCCCAGAAAGCATCTATAGATCATGGAAGAGGGCTTGACACAACAC
AGGAGGATAATGCATCTATAAAGGTCTCTGATATGAATTCTATGATAACAGATGAACAGA
GCGTGAAAATGGAGCAAGACAATGTGAACAAACCATCTGAGCAAATATTCCCTAAATAGA
TTGTCCATTGTTGGTAATTGACTTGTAATGCTTCTGTGGCAAATTCTTAAAGCTTATGATG
TCAAAACGTAGCTCTTTTTTGTGTGAACTATCCTGCTAAATTAAACCTCAGCAAATACTGC
TTCCTTGTTTTATAGATCCAGACATTTCGATTTTGTGTAAAGTTGACAGTTAAGCTATTGA
ACTCCCAGTACGGTTCGTTAGATTTTAAACTGATCAATGAAATCTTCATAGTTTTATGATA
AAAAAAAAA

SEQ ID NO: 33
GCGAGCGAGGGAGAGGGGGAGAGGGGGAGAGAGAGAGTCCGTGAAAAGCTTCAAAAA
TCATTACTTTGGGCGGAGACATCTCACACAGAATCCGAGGAAACAGAAGTGACGAGTGA
TGGATACTGGAGGGAATTCTCTCCCGTCTGGGCCTGATGGGGTGAAGCGGAAGGTGTG

CTACTTCTATGACCCGGAAGTCGGCAATTACTACTACGGCCAAGGGCATCCGATGAAGC
CGCATAGAATCCGCATGACGCACGCCCTGCTGGCGCACTATGGCTTGCTCCAGCATATG
CAGGTCCTCAAGCCCGTCCCCGCGAGAGATAGGGATCTCTGCCGCTTCCATGCCGACG
ATTACGTCGCCTTTCTTAGGAGTATAACCCCCGAGACGCAGCAGGATCAGCTGAGGCAG
CTCAAGAGGTTCAATGTTGGAGAGGATTGTCCTGTGTTCGATGGCCTTCACTCTTTCTGT
CAGACTTATGCCGGAGGCTCGGTTGGCGGTGCCGTGAAGTTGAATCACGGTCTCTGCG
ATATCGCTATCAACTGGGCAGGTGGGCTGCATCATGCTAAGAAGTGCGAGGCTTCTGGA
TTTTGCTATGTCAATGACATCGTGCTTGGCATATTGGAGCTCCTTAAACAGCATGAGCGC
GTTTTGTATGTGGATATTGATATTCACCATGGGGACGGTGTTGAAGAGGCTTTCTACACA
ACTGATAGGGTCATGACCGTTTCATTTCACAAATTTGGTGATTACTTTCCTGGTACTGGC
GATATACGTGACATAGGATATGGCAAAGGGAAATATTACTCCCTCAATGTTCCATTGGAT
GATGGGATTGACGACGAAAGCTACCATTCTCTGTTCAAACCAATAATTGGAAAAGTAATG
GAAGTGTTCAAACCCGGTGCTGTGGTCCTTCAATGTGGCGCTGACTCTCTGTCAGGAGA
CCGATTAGGATGCTTCAATCTATCAATCAAAGGACATGCAGAATGTGTCAGATACATGAG
ATCATTCAATGTGCCAGTGTTGCTGCTAGGAGGCGGTGGTTACACCATTCGTAATGTTGC
TCGTTGCTGGTGCTATGAGACTGGAGTGGCCCTTGGACTTGAAGTTGATGATAAGATGC
CACAACATGAGTATTATGAATACTTCGGTCCTGACTATACACTTCATGTTGCTCCGAGTAA
CATGGAAAACAAAAATTCGCGACAATTGCTTGAGGATATAAGATCGAAGCTTCTCGAAAA
TCTCTCTAAGCTCCAGCATGCACCCAGTGTCCCATTCCAGGAAAGACCCCCTGATACTG
AGCTCCCAGAGGCTGATGAAGATCAAGAAGATCCGGATGAAAGATGGGATCCAGACTCT
GACATGGATGTTGACGAAGATCGCAAGCCCCTGCCTAGCAGAGTAAAGAGAGAACTAAT
AGTAGAACCTGAGGTCAAAGACCAGGATTCCCAGAAAGCATCTATAGATCATGGAAGAG
GGCTTGACACAACACAGGAGGATAATGCATCTATAAAGGTCTCTGATATGAATTCTATGA
TAACAGATGAACAGAGCGTGAAAATGGAGCAAGACAATGTGAACAAACCATCTGAGCAA
ATATTCCCTAAATAGATTGTCCATTGTTGGTAATTGACTTGTAATATGCTTCTGTGGCAAA
TTCTTAAAGCTTATGATGTCAAAACGTAGCTCTTTTTTGTGTGAACTATCATGCTAAATTAA
ACCTCAGCAAATACTGCTTCAAAAAAAAAA

SEQ ID NO: 34
GGAGAAGGCAAGGCAACCCACCAGCCAAGCGGACGGAGAGAGGGTCTGCTCCTCCTC
CTTCCGTCGCAGCACCACAGCACCTTCACGCCAATGGTGCGGCCTCGCTAGCGATACA
CCGTCTCCGCCGCTTCGTTCGCAGGACGGGAAGCTCAGGATGCGACCCAAGGACAGGA
TTTCGTACTTCTACGACGGAGATGTGGGCAGCGTCTATTTCGGTCCGAATCATCCGATG
AAGCCGCACCGGCTCTGCATGACCCACCATCTTGTCCTCTCTTACGAGCTTCACACGAA
GATGGAAATTTACCGGCCGCACAAAGCCTACCCCGCTGAGCTCGCCCAGTTCCACTCTC
CTGATTATGTCGAGTTTTTGCACCGGATTACACCCGACACCCAGCACCTGTTCCCAAATG
ATTTGGCAAAATATAATCTTGGAGAAGATTGTCCTGTCTTTGAAAACTTGTTCGAGTTTTG
TCAAATTTATGCTGGTGGCACGATAGATGCTGCAAGAAGATTAAATAATCAACTCTGTGA
TATAGCTATCAACTGGGCTGGAGGATTACATCATGCCAAGAAGTGTGAAGCATCTGGATT
CTGCTATATAAATGACCTAGTTTTGGGAATATTGGAGCTGCTGAAATATCATGCACGTGT
TTTATATATTGACATAGATGTGCATCATGGTGATGGCGTAGAAGAAGCCTTTTACTTCACT
GACAGGGTAATGACTGTTAGCTTTCACAAATTCGGGGATATGTTCTTTCCAGGAACTGGT
GATGTTAAGGAAATAGGAGGAAAAGAAGGAAAGTTCTATGCTATCAATGTCCCACTCAAG
GACGGAATAGACGATACTAGCTTCACTCGACTTTTCAAGGCTATCATTTCGAAGGTCGTC
GAAACATACCAACCTGGTGCTATTGTACTCCAATGCGGAGCAGATTCACTTGCTGGGGA
CCGTTTGGGGTGTTTCAATCTCTCTATTGATGGACATTCTGAGTGTGTAAGGTTCGTCAA
GAAATTCAATTTGCCATTATTGGTTACTGGTGGTGGAGGATACACAAAAGAGAATGTTGC
TCGATGTTGGGTTGTTGAAACAGGAGTTCTCTTAGATACAGAACTTCCAAATGAGATTCC
TGAAAACGAGTATTTCAAGTACTTTGCTCCTGATTATTCATTGAAGATTCCTCGTGGAAAT
ATCGTACTCGAGAATCTAAATAGCAAGTCCTACCTCAGTGCAATCAAAGTGCAAGTGTTA
GAAAACCTTCGTAATATTCAACATGCCCCAAGTGTACAAATGCAGGAGGTTCCACCCGAT
TTCTATATTCCCGATTTTGACGAAGACGAGCAGAATCCTGATGAACGAATGGATCAGCAT
ACCCAGGACAAGCAAATCCAACGAGATGATGAATATTATGATGGGGACAATGATAACGA
CCATAACATGGATGATTCATGATGCCAAGCTAAGCAATATGTAATTTTTTTTTTAAGGATA
CGACATTTGAATTGGCGATCACAATCTACTGTAGTCAATACTCAAGTGGGAGGTGTAAAT
AGATTCCATCTGAATTTTGTGAAGCAGATATTAGTTCCTACTTTTGTGAAAAAAAAAAAAA
AAAA

SEQ ID NO: 35

GGCACCAAACATCGAGGCACAACAGCATCGTCCCCCGCGTCTCCGTCGCGACACCCTC
CTCGTCCGGTGGCGGTGGTGGCGGTGCGGAGTTTAAAAGCGTCGCGGTTTCCCTCGAC
ACGACCGAGCCTCCCCCACAACCTCACTCGCCCGCTCGCTCGCGATCGACGGCGCGCG
CTCGCCGGCTTCCGCCTTCCTCTCCGGCGATCTCGGCATGATGTATTAAAGGAATGACA
GTGGCTGAAGATTTCCATGTAAACAACAGGTCTAAAATGGTTTCCCAGGCAACTCCAGAA
AGTCGTCTCACAGGTGGAGAAGATGATAACAGTCTGCACAATCAAGTTGATGAGCTCCT
CTGTCAGGAATTACCAGAAAGACAAGTGATATTGGAGTTTGAAGGCACCAGGCCTAAGC
CTTATTTCAGTGATCATAACGGGGGAGAAAATAGTGCATTAGGTGTGAGAGCCACCGAA
GATGATTTGAATTCTGATGTGGAGGCTGAAGAAAAGCAGAAAGAAATGACTTTAGAAGAT
ATGTACAAAAATGATGGCACCTTATATGATGACGATGAGGATGACAGTGATTGGGAACC
GGTGAAAAGGCAAGTGGAACTGATGAGGTGGTTTTGCACCAACTGTACAATGGTGAACG
TTGAAGATGTCTTTCTCTGCGATATATGTGGGGAACATAGAGATTCAGGTATCTTGAGAC
ATGGCTTTTATGCATCTCCATTCATGCAAGATGTGGGTGCTCCTAGCGTTGAAGCAGAAG
TACAGGAATCACGTGAAGATCATGCTAGAAGTTCTCCGCCAAGCAGTTCAACTGTTGTG
GGTTTTGATGAGAAGATGCTGCTACACTCAGAGGTGGAGATGAAGTCACATCCTCACCC
TGAAAGGGCAGACCGTCTTCAAGCTATTGCTGCTAGCCTTGCTACTGCCGGCATATTTC
CTGGAAGATGCCGTTCGCTTCCTGTGAGAGAAATTACAAAAGAAGAGCTGCAGATGGTC
CATTCTTCAGAGCATGTTGATGCTGTTGAAATGACAAGCCATATGTTTTCTAGCTACTTTA
CTCCTGACACATACGCGAATGAACATTCAGCGCGTGCTGCTAGGATTGCAGCCGGTTTG
TGTGCTGATCTCGCTTCAACAATTATTTCTGGACGTTCGAAGAATGGTTTTGCTCTGGTT
CGTCCTCCTGGTCATCATGCTGGTATCAAACATGCCATGGGGTTCTGCCTCCACAATAAT
GCAGCAGTTGCTGCACTAGCAGCACAAGGTGCGGGGGCGAAGAAAGTGCTTATAGTTG
ACTGGGATGTTCACCATGGAAATGGCACGCAAGAGATTTTTGATGGAAATAAATCGGTCT
TATACATATCACTACATCGACATGAAGGAGGAAACTTTTATCCCGGTACAGGTGCTGCCC
ACGAGGTCGGTACCATGGGTGCTGAAGGATACTGTGTGAATATTCCATGGAGCCGTCGA
GGAGTCGGTGACAATGATTATGTTTTTGCATTTCATCATATAGTGCTTCCTATAGCTTCCG
CATTTGCTCCTGATTTCACCATCATATCAGCTGGATTTGATGCTGCGAGGGGTGATCCCC
TAGGGTGCTGTGATGTCACTCCGGCTGGCTACGCACAGATGACACACATGTTGAGTGCT
CTTTCTGGTGGGAAGCTGCTTGTCATTCTAGAGGGAGGTTACAATCTTCGCTCAATCTCC
TCTTCTGCTGTGGCAGTTATTAAGGTGTTACTGGGTGACTCACCTATATCTGAAATTGCG
GATGCGGTGCCCTCGAAAGCTGGCTTGCGTACTGTGTTGGAAGTCCTGAAGATACAAAG
GAGCTACTGGCCTAGTCTTGAATCCATTTTTTGGGAATTGCAGTCACAATGGGGAATGTT
TCTTGTTGATAACAGAAGAAAACAGATCAGAAAGAGACGACGGGTGTTGGTGCCAATAT
GGTGGAAATGGGGTCGGAAAAGTGTGTTGTATCATCTCCTAAACGGTCATCTTCATGTGA
AAACGAAGCGGTGAATGTCTAGGCTCTAGTCTGTGTATTCTTCCTTTGTGCAACTTTGTA
TACCATCAGCTGGCAGTCCTCTTTTATCCGTGGAAGACTCTCATGGTGATTTTGACAACC
ATTTTTACATTGTAAATCTAGGTTAACCATCATTAGAGTAGAGCTCCTTCTGATCTACTAC
TGTTAAGAGAGGAGTGGTCAGCGGCCTTCATTTTATGTTTAAATGTTCAAATGTTGATAAT
TGCAACTTTTTTGGGAGCTTGCTTTTGCATTAAGCTCGGATTGGATCGAGAAGCATGCCG
ATTTTGTTGGGGCAGTCAGATGTATGTGATCATGTGTAATCAGTATATCAGGTTAGAAAC
AGTACTCTTGAGCTTAGCGGGCACTGTTCTTCGCTGCTTCAGAGGTTTAAGTCCTCAAGG
TCAGGGAAAATGTTTGTGGTTTCGCTTCTTTGAGTTCATAACTGTCTTTTTGGCTCAAAAA
AAAAA

SEQ ID NO: 36
ACAGGCAAAAAGCATAAGCTGGCAAGAGCTTCCTCCTGCTGAGTTACAGCATCCTTT
AGGATCAGAAGAGGGGAAACATCATCGGCTGCTTCGCTATTCAACTCTCGAATGGCTGC
TGCTCCTTCTTCTCCTCCGACGAACCGGGTCGACGTGTTCTGGCACGACGGGATGCTGA
GCCACGACACCGGTCGGGGCGTGTTCGACACGGGATCCGACCCGGGCTTCTTGGACGT
GCTCGAGAAGCACCCGGAGAACCCCGACAGGGTCCGGAACATGGTCTCGATCCTCAAG
CGAGGACCCATCTCTCCCTTCATCTCCTGGCACACTGCAACGCCTGCTCTCATCTCTCA
GCTCCTCTCTTTTCACTCTCCGGAATACATAAATGAACTAGTGGAAGCTGATAAAAACGG
AGGGAAAGTTCTTTGCGCTGGAACTTTCTTGAACCCAGGGTCATGGGATGCCGCGCTTC
TTGCTGCTGGTAACACACTTTCTGCTATGAAATACGTACTTGATGGTAAGGGGAAGATTG
CATATGCACTAGTGAGGCCACCTGGTCACCATGCTCAGCCTTCTCAAGCCGATGGATAT
TGCTTCCTAAACAATGCCGGTCTAGCAGTTCGATTGGCATTGGACTCTGGGTGCAAAAG
GGTTGTCGTTGTTGATATAGATGTGCACTATGGAAATGGAACAGCGGAGGGATTTTACCA
ATCCAGTGACGTTCTCACCATCTCTCTTCACATGAATCACGGGTCCTGGGGTCCATCTCA
TCCACAAAGTGGATCTGTCGATGAGCTTGGTGAGGATGAAGGATACGGGTATAATATGA
ACATTCCGTTGCCAAATGGAACCGGAGACAGGGGATACGAGTATGCTGTGACTGAGCTG

GTTGTGCCAGCTGTGGAGAGTTTCAAACCTGAAATGGTTGTTCTTGTTGTTGGCCAAGAC
TCTAGCGCGTTCGATCCAAATGGAAGGCAGTGCTTGACAATGGATGGATATCGGGCGAT
TGGTCGAACAATTCGTGGCCTCGCGGATAGGCACAGTGGAGGCCGCATTCTCATTGTCC
AAGAAGGTGGATATCATGTTACCTACTCAGCTTATTGCCTTCATGCCACGGTGGAAGGCA
TTCTTGATCTTCCAGATCCATTATTAGCTGATCCAATTGCTTATTACCCAGAAGATGAGGC
TTTTCCTGTGAAGGTGGTCGACTCAATCAAAAGGTACCTTGTAGATAAGGTACCTTTTCT
CAAGGAACATTGATACAGAGATGTTGCGATATAACTTTGCTTTATTGGCACGGTGTCCCC
AAATCTAGTTTGAATATTTTGTTTTTTTAGGACTTCATCTGAGATAGAAGTAGAAGCTAATA
TGAAGAAAGTAGTGATTTCTATATTTTATTTTTTGGGTGAGAA

SEQ ID NO: 37
AGCGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGGCGCGCGGCGGAGATGGTCGA
GAGCAGCGGCGGCGCGTCGCTGCCGTCGGTGGGGCAGGACGCGCGGAAGCGGCGCG
TGAGCTACTTCTACGAGCCGACGATCGGCGACTACTACTACGGGCAGGGGCACCCGAT
GAAGCCCCACAGGATCCGGATGGCGCACAACCTCATCGTCCACTACTACCTCCACCGC
CGCATGGAGATCAGCCGCCCCTTCCCCGCCGCCACCACCGACATCCGCCGCTTCCACT
CCGAGGACTACGTCACCTTCATCTCCTCCGTCACCCCCGAGACCGTCTCCGACCCGGC
CTTCTCCCGCCAACTCAAGCGCTTCAACGTCGGCGAGGACTGCCCCGTCTTCGACGGC
ATCTTCGGGTTCTGCCAGGCCTCCGCCGGCGGTTCCATGGGCGCCGCCGTCAAGCTCA
ACCGCGGCGACTCCGACATCGCCCTCAATTGGGCCGGCGGGCTGCATCACGCCAAGAA
GTCCGAGGCTTCTGGGTTTTGCTATGTCAACGATATCGTGCTTGGGATTCTCGAGCTCCT
CAAGGTCCACAAGCGTGTTCTCTACGTGGATATCGATGTTCACCATGGGGATGGGGTGG
AGGAGGCATTTTATACGACGGATAGAGTCATGACCGTATCTTTCCATAAGTTTGGGGACT
TCTTCCCTGGAAGCGGGCACATTAAGGACACCGGGGCAGGGCCTGGGAAGAACTACGC
CCTCAATGTGCCACTGAATGATGGCATAGATGATGAAAGTTTCCGTGGTATGTTTCGCCC
CATTATCCAAAAGGTTATGGAAGTATATCAACCGGATGCCGTTGTTCTTCAGTGTGGGGC
AGATTCACTGTCTGGAGACCGATTGGGATGCTTCAATTTGTCTGTGAAAGGTCATGCAGA
TTGCCTCCGTTTTCTAAGATCTTTTAACGTTCCTTTAATGGTCTTGGGTGGGGGAGGTTA
TACAATGAGAAATGTCGCTCGTTGCTGGTGTTACGAGACAGCTGTTGCGGTGGGAGTAG
AACCTGAAAATGACTTGCCTTACAATGAGTATTACGAGTACTTTGGCCCAGATTATACTCT
TCATGTCGAGCCATGTAGCATGGAGAATCTTAATGCACCAAAAGATTTGGAGAGAATCAG
GAATATGTTATTAGAGCAGCTATCGAGAATCCCACATGCACCAAGTGTACCTTTTCAAAT
GACACCGCCCATTACACAAGAGCCAGAAGAGGCAGAGGAAGATATGGATGAAAGGCCA
AAGCCACGCATTTGGAATGGTGAGGATTATGAATCGGATGCTGAAGAAGACAAGAGTCA
GCATAGATCATCAAATGCTGATGCTTTGCATGATGAAAATGTTGAAATGAGGGACAGTGT
TGGTGAAAACAGTGGGGACAAAACAAGGGAAGATCGTTCTCCATCTTGACAAAGATCAG
CAAAGATCAGCTCTAACTCTTGTGTACTTTTGGGTCTTTCTGATATGCGATGCATGTAATT
CTTTCTAGGCCAGGAGATAACCTTATCAAGCCAATAGAGTCTAATGGTTGGCCACACTGG
ATGATTGCTGATAGACGAGGCGGGTTATGTTTATCAGGTGTCACAGTGCATGTTGGAATA
GGAAGACTGAATTGCTGAGTGTAAATGGGATGTTTTCTTATGTCTCCAGCTGATGTGGTT
TGTTTCATTGGCTCATGTGAGATGGATTTACTTAGTTAAAAAAAAAA

SEQ ID NO: 38
CTGCGCTTCATTTGCCATTCGTTTTCTCCTCTGCCTCGTTTCTTTGCGGTTCGGACGCAT
CATGGATGTATCTCCAAAGAGTAATCTGTCGATTCCGTGGACAAGATCCACAGCTGCTTC
TTGAGGTCATGCTATATTGGAACCAGCTTAAGGTCAGTTGAGAGGACTTCAACCAATGGT
AGTTCCTAGCAGTAACCCCCACAACAGGGAAATGGCCATTAGGAGGAGGATGGCCAGC
ACATTCAATAAAAGAGAAGATGATTTTCCATCCTTGAGAGAGTACAATGACTACTTGGAA
GAAGTGGAGGAAATGACATTCAATCTGATTGAAGGAGTGGATGTGCCGACTATTGAAGC
AAAGATTGCAAAGTACCAGGAAGAAAATGCTGAGCAGATTATGATTAATCGTGCGAAAAA
GGCTGAAGAGTTTGCTGCGGCATTGGCTGCAAGCAAGGGGCTGCCTCCTCAAACTGAT
CCAGATGGGGCTCTTAACTCTCAAGCAGGATTGAGTGTTGGGACACAGGGACAATACGC
TCCTGCAATTGCTGGAGGACAGCCACGGCCAACAGGCATGGCTCCGCAGCCGGTGCCA
CTTGGGACTGGCCTCGACATCCATGGATATGATGATGAAGAAATGATAAAGCTCCGAGC
TGAAAGAGGAGGCAGAGCGGGAGGGTGGAGCATAGAGCTAAGCAAGAAAAGGGCTCTT
GAAGAAGCATTTGGAAGCCTTTGGTTATAATTCAGTTAGTTCTAGCACTGTTTGTGGGTG
GGAGCATTCTAAGATGGTTCTTAAGGGAGGAAGCGAGAATTGCAATTGTTATGATTGGTC
TGACTAAGAAGTGTTTGCAATTATCCGCGTGAAGATCATGTCTTTCTGGTTCTTGCATGG
AGAGCATTTGGAGCAGCAGCAGCAGCCTTGCTGCTCTCTGGCAATTTTAATGTGACATT
GAAGGTGTAAAACTGTTACCTCTAATACGATCATACTTGGAATAAAAAAAAAA

SEQ ID NO: 39
CTTCCTCTCAGACCGTCCCAGAAGCCTTATCTTCATTGGCTTCTCTCAGCTCCTGTACCA
ATTTCGTGTCTCTCTTGTTTTCTTCCGACACACCCGTTTCCCATGTCTTTGTAACACCTCG
AAAGCTCTTATCTTCTGCGCTCTCTCCCATGGCAGCAATCATATCCTGCCACCACTACCA
CTCCTGCTGCTCCTCCCTCATCGCCTCGAAGTGGGTTGGCGCCAGAATACCCACAAGTT
GTTTCGGCCGTTCGAGTACGCAGAGCAACAATGCGGCGTCGGTCAGGCAGTTCGTCAC
CCGGTGTTCTTCAAGCCCGAGCAGTCGCGGCCAATGGCAGCCTCACCAGAACGGAGAG
AAAGGGAGATCATTTTCCTTGAGAGAATGTGCGATATCTATTGCTTTGGCGGTTGGATTG
GTGACGGGAGTGCCTTCGTTGGATATGTCTACGGGCAATGCTTATGCTGCTAGTCCTGC
TCTGCCTGATCTCTCTGTCCTAATATCTGGCCCACCGATTAAAGACCCCGAGGCTTTGCT
AAGATATGCTCTTCCGATTAATAATAAAGCCATAAGGGAAGTTCAGAAGCCGCTCGAGGA
CATCACGGATAGTCTCAAGGTTGCAGGACTCAGAGCTCTTGACTCCGTGGAGAGAAATG
TGAGGCAAGCATCTCGGGTGCTCAAGCAAGGTAAAAATCTGATTGTATCTGGCTTAGCT
GAGTCAAAGAAGGATCACGGAGTGGAGTTGCTTGACAAGCTGGAAGCTGGAATGGATG
AGCTCCAACAAATTGTGGAGGATGGAAATAGAGATGCTGTAGCGGGCAAACAGAGGGAA
CTGCTAAACTATGTTGGAGGTGTTGAAGAGGATATGGTGGATGGCTTCCCTTATGAAGTT
CCTGAGGAATACAAGAATATGCCTCTCTTGAAAGGTAGGGCTGCTGTGGATATGAAGGT
CAAGGTCAAGGACAATCCCAACCTGGAGGAGTGTGTGTTTCGAATAGTCTTAGATGGTT
ACAATGCACCTGTGACTGCTGGGAATTTCGTGGATTTGGTAGAGAGGCACTTCTACGAC
GGCATGGAAATCCAGAGAGCCGATGGATTTGTTGTTCAGACAGGTGATCCCGAAGGTCC
TGCTGAGAGTTTTATTGATCCTAGCACAGAGAAACCCCGGACGATACCTTTGGAGATCAT
GGTGGATGGGGAGAAAGCGCCAGTATATGGAGCAACTCTTGAAGAGCTTGGCCTCTACA
AGGCTCAAACAAAGCTTCCATTTAATGCATTTGGGACAATGGCCATGGCCAGAGATGAG
TTTGAGGACAATTCCGCATCAAGTCAGATATTTTGGCTGCTAAAGGAAAGTGAACTGACC
CCTAGTAACGCGAATATATTAGATGGGCGGTATGCAGTTTTCGGTTATGTGACAGAAAAT
CAGGATTTCTTAGCAGATCTTAAGGTCGGTGATGTCATAGAGTCGGTGCAAGTCGTCTC
GGGCCTGGATAATCTGGCGAATCCCAGCTACAAGATTGCTGGCTAGAGTTTTCCTTTTCT
CTCCGTTTCCCAATATCGTTCGTTTGGACACTGCACATGCCAGCTTGTATGACTCCAACG
TGTAGTTGTTTAAGGCTAATGTTCTTCGTATGATTCTTGTTATTTATGAGAGATAATGTGTT
GTCCACATTGAGAATGTATTCTTTTTTTGGGATTGCATTGATATCAAATTCAGATCTATTAG
TGAAAGTTGGCATGAGTCTCAATCTTAGGGGAATACAGTACGGAGAATGAATTTCTGGAT
GAAAAAAAAAA

SEQ ID NO: 40
GGTCATCTCTCTCTCTTCCTCGAGACTTCGCCCCCCGCAAAACCCTTGCGCTCCGAACT
CTCTAGAACCGCCTCCCTCTCTCCGCCCTTATAAATCTCTCCCTCTCTCCCCCCCCCCGC
CGCGCCTCCGCTCGCTTCCCGACACCCCCACGGAGACGCCATCCCCAGCTCCGCTTCG
CGATCCGTTTGCAGCCTCGCGGTCGTCGGCAGCCGTCCCTCTCTCTCGGGCGAGAGAC
TTCCGGCGAACATGGCCGGCGAGGACTTCGACATTCCGCCCGCCGACGAGATGAACGA
GGACTTCGACCTCCCCGACGACGACGACGACGCCCCCGTCATGAAGGCCGGCGATGA
GAAGGAGATCGGCAAGCAGGGCCTCAAGAAGAAGCTCGTCAAGGAAGGCGACGCCTG
GGAGACTCCCGATAATGGCGACGAAGTTGAAGTGCATTACACCGGGACGCTCCTGGAT
GGTACTCAATTCGATTCTAGCCGGGACCGCGGGACTCCTTTCAAGTTCACTCTTGGCCA
AGGGCAAGTGATAAAGGGATGGGATCAAGGTATCAAGACAATGAAGAAGGGGGAAAAT
GCAATTTTCACCATACCTCCAGAGCTAGCTTATGGCGAGGCTGGCTCACCTCCAACTATA
CCTCCTAATGCAACTCTGCAATTTGATGTTGAACTACTTTCTTGGACGAGTGTCAAGGAT
ATTTGCAAGGATGGTGGTATTTTCAAGAAAATTCTGGTGGAAGGCGAAAAATGGGAAAAT
CCTAAAGATCTGGACGAAGTATTAGTCAAGTATGAGTTTCAGCTGGAAGATGGTACCACC
ATTGCAAGGTCTGATGGAGTGGAATTCACTGTTAAAGAAGGACATTTTTGCCCAGCAGTT
GCTAAAGCTGTCAAGACAATGAAAAAGGGAGAGAAAGTTCTTTTGACTGTCAAGCCACAA
TATGGATTTGGCGAGAAAGGTAAGCCAGCTTCTGGTGATGAGGGTGCTGTTCCACCTAA
TGCCACTCTTCAAATAACGTTAGAGTTGGTATCGTGGAAGACTGTGTCTGAAGTGACCGA
TGACAAGAAAGTAATCAAAAAGATTCTCAAGGAAGGGGAAGGGTATGAGAGGCCTAACG
AAGGAGCTGTAGTTGAAGTGAAATTGATTGGGAAGTTGCAAGATGGCACGGTATTCGTG
AAGAAGGGCCACGATGATTGTGAGGAGTTATTCAAGTTCAAGATCGACGAAGAACAAGT
AGTGGACGGGCTTGACAAAGCTGTGATGAACATGAAAAAGGGAGAGGTTGCTTTGCTGA
CTGTTGCACCTGAGTATGCTTTTGGATCTTCTGAGTCTAAGCAGGATCTGGCTGTGGTAC
CTCCTAGTTCGACTGTGTATTATGAGGTGGAGTTGGTCTCTTTTGTCAAGGACAAGGAGT
CGTGGGACATGAATACTGAGGAGAAGATTGAGGCTGCAGGTAAGAAGAAAGAAGAAGG

AAATGTTATATTTAAGGCAGGGAAGTATGCGAAGGCTTCCAAAAGATATGAAAAGGCTGT
GAAGTACATAGAATACGATACCTCCTTTAGTGAAGATGAGAAGAAACAAGCGAAGGCACT
GAAAGTTGCTTGCAATCTGAATGATGCAGCATGCAAGCTGAAACTCAAAGATTACAATCA
GGCTGAGAAACTATGCACCAAGGTTTTGGAGCTGGATAGCAGGAATGTGAAGGCTCTTT
ATCGGAGGGCGCAGGCTTACATTGAGCTGTCTGATCTTGATTTGGCTGAATTTGACATCA
AGAAGGCGCTAGAGATTGATCCTCACAATAGGGATGTGAAACTGGAGTATAAAGTATTGA
AGGAGAAGGTCAAGGAGTTCAACAAAAAGGATGCCAAGTTCTATGGCAATATGTTTGCC
AAAATGAGCAAACTAGAACCTGTTGAGAAGACTGCAGCTAAGGAGCCTGAGCCTATGAG
CATAGACAGCAAGGCGTGAGAGAGAAGCACCTAATTCCGTTTCTTTTCTTTCGGTTCATT
TGGGTCCTTCTCTACATGCAAGGAATTCTCTATTCCGCAGCAATGTAACTGGGATTATTTT
ATCGGATTGAAAAACATATAATGTCGGTACAGTGTGTCGTGTGGTTTAGGTCAGTTATGT
TGCGGTTGGAGTTTTCAAAAGTTCTTGTACCATTGAAAATTCTCTTGCGCAGTTTGGAGA
GGTTATTGTCTCTCAATCGATCTGTCTGTTTACTTGTCCAGTTGATATGAGTATCATAACT
CGGATGGTGACAACTTTGTACTACGGTCGGCACCGGTAGATGATGGTGCATATGGGAGA
CGAGGTACGGTTATAAAAAAAAAA

SEQ ID NO: 41
GGGACCGGCGGCCGACGGCGTCGTTAGGGGTGTTCTTCTTCCCCAGCCGAGCTCCAGA
AAGACCTGGAAATCGCGGGCGGCGGAAGATTCGCCTTCGTCGAGATCAGCGCGGGGA
GCTCCGGCTTCGAGTAGAAGATGTCGACCGTGTACGTGTTGGAGCCGCCGACGAAGGG
GAAGGTGGTGCTGAACACGACGCACGGCCCGCTCGACGTCGAGCTCTGGCCGAAGGA
AGCGCCCAAGGCCGTACGGAACTTCGTCCAGCTCTGCCTCGAGGGCTACTACGACAAC
ACCATCTTCCACCGCATCATCAAGGACTTCCTCGTCCAAGGCGGCGATCCCACCGGCTC
CGGCACAGGTGGCGAAAGTATTTATGGGGATGCGTTTTCTGATGAGTTTCATTCACGTTT
GAGGTTCAAGCACAGAGGTTTAGTGGCTTGTGCCAATGCGGGATCACCACACTCCAATG
GGAGTCAGTTTTTTATCACATTGGATCGATGTGACTGGCTTGACCGGAAAAATACCATTT
TTGGAAAGATAACTGGCGATTCCATATACAATCTTAGTGGATTAGCTGAGGTCGAAACTG
ACAAGAGTGATCGCCCGTTGGATCCACCTCCCAAAATTATCTCAGTGGAGGTACTCTGG
AACCCTTTCGAAGATATTGTTCCTAGAGCACCAGTGAGGTCTTTGGTTCCAACTGTACCT
GATGTTCAAAATAAAGAACCAAAGAAAAAGGCCGTAAAAAAGCTGAACTTACTTTCATTTG
GAGAGGAAGCTGAAGAGGAGGAGAAGGCATTGGTTGTTGTTAAACAGAAGATCAAGAGC
AGTCACGATGTGCTTGATGATCCTCGCTTGCTGAAAGAACATATACCAAGTAAACAAGTG
GACTCATATGACAGCAAAACCGCAAGGGATGTCCAATCTGTTAGAGAAGCTTTAAGCTCA
AAGAAACAGGAGCTGCAGAAAGAGTCCGGAGCTGAATTTTCAAATTCATTCAGGGAAATT
GCCGATGATGAAGATGACGACGACGACGATGCGAGCTTTGATGCAAGAATGCGTAGACA
AATACTTCAGAAAAGAAAGGAGCTGGGTGATCTCCCTCCAAAGCCAAAGCCAAAGTCGC
GTGATGGGATTTCTGCCAGAAAGGAGCGGGAGACATCTATTTCAAGGGACAAAGATGAT
GATGATGATGATGATCAACCAAGGGTGGAAAAGCTCTCCTTGAAGAAAAAAGGAATAGG
ATCTGAAGCCAGAGGTGAGCGAATGGCTAATGCTGATGCAGATTTGCAGCTGTTGAATG
ATGCTGAACGAGGTAGACAGTTGCAAAAACAGAAAAAGCACCGCCTTCGAGGACGTGAA
GATGAGGTACTTACTAAGCTCGAAACGTTCAAGGCGTCAGTATTTGGAAAACCTTTGGCG
TCAAGTGCCAAGGTTGGAGATGGTGATGGTGATTTATCTGACTGGAGATCAGTGAAGTT
GAAGTTTGCTCCTGAGCCTGGCAAGGACCGAATGACTCGTAATGAAGACCCAAATGATT
ATGTTGTTGTGGACCCTCTCTTGGAGAAGGGAAAAGAGAAATTCAACAGGATGCAGGCC
AAAGAAAAACGAAGAGGACGGGAATGGGCGGGCAAATCTCTCACTTGATCCTATTTATC
CCCATTTGCATGTCTCATGAATATAACCCACTCCGTGAGCCACCGGGCGATGGTGCCAG
GTTCATGCTTGCAGAGCCATTCGACCAACAAACTTTGGACAAACTAAACCAGCAATTCTA
GAAGTGCCATGTTAGAGTTGGTAGCTCTCTTCATTGAAGAGATCGACCTGTTGACATCCT
CTTTCTGCATCACTAGGACACAAATCAACATGTACATAAGTCTCTCCTTTGAGAGAGGCG
GATATTCTAGTAGGACCGTTCTCGAAGTGGGACTGTTCTTGCTGAGAGTTTTCTTAGTTC
TTTCCTTTTGTTTTCTCTCTCAACATTATGTTTTGCTGGTCTCATTTCGACCGAAGCGCTT
CTGGAAGAAAACAATACGGCCGGAAGACTTGGAAAACAAGGGGAGTGATCTCCGTCTCT
GCCCACTAACGCCGAGCATGACAGGAAGGAAAGGCGTTGGATTCTGACCAACCCCACC
ACGATGTTCTCGATTCCTTCTAGACGTTTGACGATGGGGGCTGACTGCAGACCTGTCCG
AGCCTGCTCATAATTTGGTTGGCTATTCTTGAAGGAGAGCCATCACATGGTCCTTGTTGG
CTGTGTCGGGTCCATTGGAGTTTTGATCGTCTTGTGTCGTATTAGTTGTCATCATCATGA
GAATTTGGGAATCTGACTCTTCTCATTCTTGAATCGCAACGATCAATTTCTATGCATAGAA
GAGAGAATTATTTTTAAAAAAAAAA

SEQ ID NO: 42

CCGGATAACCGCGTCTTATCTGAAACCGCCGGTCCAAACCACCATCAATGGCGTCCGCA
ATCTCCATGCACAGCTCTGGCCTCCTCCTCCTCCAGGGTACCAACGGTAAAGATGTTAC
TGAAATGGGTAAGGCACCTGCAAGCTCTCGAGTTGCTAATATGCAGCAAAGGAAGTATG
GGGCGACATGCTGTGTGGCCAGAGGATTAACATCTAGATCTCACTATGCCTCCAGCTTG
GCATTCAAACAATTTTCTAAAACCCCCTCCATCAAATATGACAGGATGGTTGAGATCAAG
GCTATGGCCACAGACTTGGGGCTGCAGGCGAAAGTAACAAACAAATGCTTCTTTGACGT
GGAAATTGGCGGAGAGCCTGCTGGTAGAATAGTAATAGGCCTCTTTGGAGATGATGTTC
CAAAAACTGTCGAGAACTTCCGTGCTTTATGCACAGGTGAAAAGGGATTTGGCTACAAAG
GATGCTCCTTCCATCGTATCATTAAGGATTTTATGATCCAGGGAGGGGATTTCACTAGAG
GAAATGGAACTGGAGGAAAGAGCATCTACGGTTCAACTTTTGAAGATGAGAACTTTGCAT
TGAAGCATGTCGGACCTGGAGTACTGAGCATGGCTAACGCTGGCCCTAGCACTAATGG
GAGCCAATTTTTCATATGCACTGTAAAGACTCCATGGTTGGACAATCGCCACGTTGTGTT
TGGACAAGTCGTCGACGGGATGGATGTTGTGCAGAAACTTGAATCTCAGGAGACAAGTC
GATCGGACGTGCCTCGGCAGCCATGCAGAATTGTGAACTGTGGGGAACTTCCCCTAGAT
GGTTGATTCACATGAGCTGGCTTGTTGCAATGCGGAGCACAAATTTTGGTAATGTGTTCC
GGTTTTGTCATGTTTCATGCGACAGCTCTTCGTTGTGTCAGTATTTTTTAACTGGATGTTC
ATGTATGCTTAGAGATTCCATTGGTGATTCTGACTTCAGTGGTAGCCCATGAAATTTGGG
TTCTGTGTTAGAATTCTTTCGTCAAGGAAAGTGAATTTTCAAAAAAAAAA

SEQ ID NO: 43
GGGAAAAAGGAAAAATATCTTCGTTCTCTTCCCTCTTCATCGCCTTCCATGGCGGCTTCT
TTCACAGCTCTGTCCAATGTCGGCTCGCTTTCCAGCCCGAGGAACGGCTCGGAGATTAG
ACGGTTCCGCCCCTCCTGCAACGTCGCCGCCAGCGTCCGGCCGCCTCCGTTGAAGGCC
GGCTTGTCGGCGTCGTCCTCGAGCTCTTTCTCCGGCTCTCTGCGCCTGATTCCCCTCAG
CTCTTCTCCTCAGAGGAAATCCCGTCCGTGCTCGGTTCGAGCAAGTGCTGAGGCTGCG
GCCGCTCAGTCGAAAGTCACCAACAAAGTCTACCTTGACATTAGCATCGGTAACCCTGT
CGGGAAACTAGTCGGAAGAATTGTGATTGGATTATACGGCGACGATGTCCCGCAAACAG
CAGAGAATTTCCGCGCATTATGCACAGGAGAGAAGGGCTTTGGATACAAAGGTTCAACG
GTCCACCGTGTCATCAAAGATTTTATGATTCAGGGAGGTGACTTTGACAAAGGAAATGGA
ACAGGGGGTAAAAGCATATATGGCCGTACTTTTAAAGATGAAAACTTTAAGTTGTCTCAT
GTCGGACCTGGGGTGGTCAGCATGGCAAATGCAGGGCCCAACACCAATGGTAGCCAGT
TCTTTATCTGCACTGTGAAGACACCGTGGCTGGACCAGAGGCATGTTGTGTTTGGGCAA
GTGCTGGAAGGCATGGACATTGTCAGGCTCATTGAGTCACAAGAAACAGACCGAGGAGA
CCGTCCCAGGAAAAGAGTGGTCGTCTCTGACTGCGGTGAACTTCCCGTGGTGTGATATG
CCTTGATTGGTTGGGTTCTGCAAGTTAGTTCTGGCTAAAATCTGAGTCATGCTCTCCCTC
GAATTTTCTTCTTTGTTAATTTCTCTGTTTTGTTGAGAGAGCAAACTTTGGGAAGAGCCAG
GTCGCGATTGGTAACGCTCAACTATTGTAGTGTACCGACGGAATTTTCCTTAAATTGGAA
GCACAGAATGACTGCTCCTGAGTCTCTTCTTTAAAAAAAAAA

SEQ ID NO: 44
GGCCGATCGCTGTGGCTGATCTCGTCGCTCCGGTTGTAAAGCATGTCGAGCACATCATC
TCCTCCTTCTTATTCTTCTTCTTGGATCAAAGCATGTTCAGCACATCATGACCCAACAGAA
GCCCTTTGGCTTGCATCTGTAACTCAGTTCCCCTTTCCATACAAGCAGGCTTTTCATAAA
AATCATGGCTGAGGCAATCGATTTGACGGGTGATGGAGGGGTCATGAAGACAATTGTGC
GACGAGCAAAACCGGATGCAGTCTCTCCTTCAGAGACCCTTCCTCTTGTTGATGTTCGTT
ATGAAGGAGTACTTGCTGAAACTGGTGAAGTCTTTGACTCAACACATGAAGACAATACTC
TATTCTCCTTTGAGATCGGAAAGGGCTCGGTGATCAGTGCTTGGGACACTGCATTGAGA
ACTATGAAGGTTGGGGAGGTTGCAAAAATCACATGTAAGCCAGAATATGCCTATGGCAG
CACGGGTTCTCCACCCGATATCCCACCAGATGCAACCCTTATTTTTGAGGTGGAGTTAGT
TGCATGCAAACCGTGCAAGGGCTTTTCAGTGACCAGTGTCACAGAAGACAAGGCTAGGC
TTGAGGAGCTGAAGAAGCAAAGGGAGATAGCTGCCGCAACCAAAGAGGAAGAGAAGAA
GAGGAGGGAAGAAGCTAAAGCTGCAGCTGCTGCTCGTGTTCAAGCCAAGCTGGATGCT
AAGAAAGGTCACGGGAAGGGAAAGGGAAAAGCAAAATGAACAGCTTTATAAGGAATACT
CAAGTAGATCATAACTCAGTTTTCCTTCATATGCCTTTTAAACTTGGCTAGGATCTTAACC
TTGAGTATTATGGAATATAAGTTAAGACCAGACTTCCCATACTTGATGGTGTGCTTGGATA
TACTGTATAAGCATTCTATATTATGCTTGTTGGCTTCGTTTTGAGGGAGCAAAACTGACTA
GAAAGGAGAAAGTATGTAGTACATCGACTCTGATGAATGGCATTTTGATTACGGTAGCTT
TCTCAAAAAAAAAA

SEQ ID NO: 45

GGAATTCATTGACGACGACAAGTTAACGTCGACCGCTTCTCTGCCCCTTGAATTTTCCCG
AGAAAACCAGGAACCTGCCAAATATCTCTCTGAAAGATCTCCATGGGTAACCCGAAGGT
GTTCTTCGACATGTCGATCGGCGGCCAGCCGGCCGGCCGGATCGTGATGGAGCTCTAC
GCCGACGTGGTGCCGCGCACGGCGGAGAACTTCCGCGCGCTCTGCACCGGGGAGAAG
GGCGCCGGCCGCTCCGGGAAGCCCCTCCACTACAAGGGCTCGAGCTTCCACCGCGTG
ATCCCGGGGTTCATGTGCCAGGGCGGCGACTTCACCGCCGGGAACGGGACCGGCGGC
GAGTCGATCTACGGCTCCAAGTTCGCCGACGAGAACTTCGTCAAGAAGCACACCGGCC
CGGGTGTCCTGTCCATGGCGAACGCCGGCCCGGGGACCAACGGCTCCCAGTTCTTCGT
CTGCACCGCCAAGACCGAGTGGCTCGACGGCAAGCACGTCGTGTTCGGGCAGATCGTG
GACGGGATGGACGTGGTGAAGGCCATCGAGAAGGTGGGGTCCAGCTCCGGCAGGACC
TCGAAGCCCGTCGTCGTCGCGGACTGCGGCCAGCTTTCCTAGATCGGGCGGTCCCCGT
CGGTCGCCGGGATCTCCCCCCCCCCCCCCCTCCCCGTGTGCGTGTGGGATCTTATCTG
ATCTCCAGTCCGTTTCCGAGCATGGTGTTTTAGGGCTTTCCTTTTTTTTTCGGCTTTTAGC
GTGTGGGGTGTTCGGCCAGATCTGGTTATGGGTCTCTCCGCGGATCCTTGTTGTCTGTA
GGATCGTCGAGACCTTTTATGATGGTTATGTTGAACTGCCAGTCGGCCATTAACTGCTGG
AAAAATAAGATAGGGAGCTTTTCTTGAAATGCGACCCTCTTTTTACCCTGTAAAAAAAAAA

SEQ ID NO: 46
TCTCGATCTCGAACTCTCTACCTGTTGATCCGTGTTGACCATGCCGAACCCGAAGGTCTT
CTTCGACATGACGATCGGCGGGGCGGCCGCCGGGCGGGTCGTGATGGAGCTGTATGC
GGACACGACCCCACGCACCGCGGAGAACTTCCGGGCGCTCTGCACCGGCGAGAAGGG
GGTCGGGAGGAGCAAGAAGCCGCTCCACTACAAGGGCTCCAAGTTCCACCGGGTCATA
CCTAGCTTCATGTGCCAGGGAGGTGACTTCACGGCCGGGAATGGGACCGGAGGGGAGT
CCATATACGGAGTGAAGTTTGCAGATGAGAACTTCATAAAGAAGCACACGGGGCCGGGG
ATCCTGTCCATGGCCAACGCAGGGCCAGGGACGAACGGGTCCCAATTCTTCATATGCAC
TACGAAGACGGAGTGGCTCGACGGGAAGCACGTGGTGTTTGGGAAGGTGGTGGAGGG
AATGGAAGTGGTCAAGGCCATCGAGAAGGTTGGGTCCTCTTCCGGTCGCACATCCAAGC
CCGTCGTGGTTGCTGATTGCGGCCAGCTTCCTTGAGCCATAATAATAGAGATCTCTCTCT
CACACACAATGGCTTCTTCATATAATAAAAAGTGTCACCTATGAAGTGTTCAATGTTGTCG
ATCTTTTGTATGTGCTCGTTTGAGGTTTGATTACTTCTTGGTCGTGCACTCCTGAAAAGAA
GAGAAGATCTCACCCGTTTTTCGTTTTCAATGGGATAATCTCGTATTACTTGGCTGAAAAT
AAGGTGAAACATTTACTTGCCTTGCAAGAGGGAAACTTTTGACAAGGTTTGGAACATAAA
TTGTTGAATACGATGTATTATAATGTTGGTGTCTTGGTGAAATACAGAATTATGCTTGCGT
GGAAAAAAAAAA

SEQ ID NO: 47
GCAGCAGCGTAGCAGCTTTTATGGCCGCCGGTCGCCGATCGCTCAACTCCTTCGCTCCT
CCTTCTCCTCCTTCGTCTTCGCCGATCGCTGTGGCTGATCTCGTCGCTCCGGCTTTTCAT
AAAAATCATGGCTGAGGCAATCGATTTGACGGGTGATGGAGGGGTCATGAAGACAATTG
TGCGACGAGCAAAACCGGATGCAGTCTCTCCTTCAGAGACCCTTCCTCTTGTTGATGTTC
GTTATGAAGGAGTACTTGCTGAAACTGGTGAAGTCTTTGACTCAACACATGAAGACAATA
CTCTATTCTCCTTTGAGATCGGAAAGGGCTCGGTGATCAGTGCTTGGGACACTGCATTG
AGAACTATGAAGGTTGGGGAGGTTGCAAAAATCACATGTAAGCCAGAATATGCCTATGG
CAGCACGGGTTCTCCACCCGATATCCCACCAGATGCAACCCTTATTTTTGAGGTGGAGT
TAGTTGCATGCAAACCGTGCAAGGGCTTTTCAGTGACCAGTGTCACAGAAGACAAGGCT
AGGCTTGAGGAGCTGAAGAAGCAAAGGGAGATAGCTGCCGCAACCAAAGAGGAAGAGA
AGAAGAGGAGGGAAGAAGCTAAAGCTGCAGCTGCTGCTCGTGTTCAAGCCAAGCTGGA
TGCTAAGAAAGGTCACGGGAAGGGAAAGGGAAAAGCAAAATGAACAGCTTTATAAGGAA
TACTCAAGTAGATCATAACTCAGTTTTCCTTCATATGCCTTTTAAACTTGGCTAGGATCTT
AACCTTGAGTATTATGGAATATAAGTTAAGACCAGACTTCCCATAAAAAAAAAA

SEQ ID NO: 48
CTGCAAGCGTGCTCGGAGATCGTGGAGATGGCGACGGCGAGATCGTTTTTCCTCTGTG
CTCTGCTCCTCCTCGCAACCCTATATCTCGCTCAGGCGAAGAAGTCTGAGGATTTGAAA
GAGGTGACTCACAAAGTTTATTTTGATGTGGAGATAGCTGGAAAACCGGCAGGTCGAAT
TGTCATGGGTCTATATGGAAAAGCAGTTCCAAAGACTGCTGAAAACTTCAGGGCGCTGT
GTACAGGTGAGAAAGGCACTGGAAAGAGTGGAAAACCTCTTCACTACAAGGGAAGCAGT
TTCCACCGGATCATTCCAAGCTTCATGCTACAGGGAGGTGACTTCACTCTTGGTGATGG
CAGAGGTGGCGAGTCTATTTATGGAGAGAAGTTTGCTGATGAAAACTTTAAGTTGAAGCA
CACTGGACCAGGGCTTTTGTCTATGGCGAACGCTGGTCCTGACACGAACGGTTCACAGT

TTTTTATCACAACTGTGACGACTAGTTGGCTTGATGGGAGACACGTCGTGTTCGGGAAA
GTGTTGTCTGGGATGGATGTGGTTTACAAAGTTGAAGCTGAGGGAAGACAGAGCGGCAC
CCCCAAAAGCAAAGTCGTGATAGCAGACAGTGGTGAACTCCCACTCTGATCGAACTGAT
GCCAACATGTTTACCATGACTAGCTCTTAGGCAGCAGTTAAGAGATAAATGCAAAAACAG
TGTAGGTTATGTGAAACCGAAAACTGGATTTGGAGTTATAGCCGTGCTCTTCTTTTTGGG
CATTTCGGTTGTTCTAATGTGTCTGGAGTGAATATTTTTCAATATGATGTAATTTTGCCTTT
GTAAAAAAAAAA

SEQ ID NO: 49
TCGACATTGAGGATCCGGGAGCCGACTCCCTTCGCTCGATGGAGGTTCGAGGGGTCGC
GGCCTGACGACGATCGAGGGCGCCCGCCTTCCTATGCCATCTTCTCATCTTCTTCATCT
TCTTCCGCGGCTCGGGAATGATGCGTCGAGAGATCTCGGTCCTGCTCCAGCCTCGCTTC
GTCCTCGCCTTCCTCGCCCTCGCCGTCCTCCTCCTCGTCTTCGCCTTCCCCTTCTCGAG
ACAGAGAGGAGACCAAGTAGAAGAAGAACCTGAAATTACCCACAGAGTATACTTGGACG
TTGACATTGATGGACAACATTTAGGTAGAATTGTGATTGGATTATATGGCGAGGTGGTAC
CAAGAACTGTAGAAAATTTCCGGGCTTTGTGCACAGGGGAAAAGGGCAAGAGTGCCAAT
GGAAAAAAACTCCACTACAAAGGAACACCTTTTCACCGTATAATATCTGGCTTCATGATC
CAAGGGGGAGATGTAATCTACGGTGATGGAAAAGGATATGAATCCATATACGGTGGCAC
CTTTGCTGATGAGAATTTTAGGATAAAGCATTCCCACGCAGGTATCATTTCCATGGTGAA
TTCTGGCCCTGACTCTAATGGATCACAATTCTTTATTACCACAGTCAAAGCCAGCTGGTT
GGATGGAGAGCATGTTGTTTTCGGCAGGGTTATTCAAGGCATGGACACTGTCTATGCAA
TCGAAGGTGGAGCTGGAACATACAACGGAAAGCCCAGAAAGAAAGTCATAATCGCTGAC
TCTGGAGAGATTCCGAAAAGTAAGTGGGACGAGGAAAGGTGACCACTTTTTGTTTTTGG
GTACACGCAGTTAGGATAACTAGCATGAAAGCCCGATCCCGCATATACAGGACTTGGAG
GCGAGTTCTCTTCTTTGGTCGACATTCTTTGCCCCGTCTTTTGTCCAGAAACCTGGCTAC
CCAGCTAACTCCAACGCTCATATTCACCATGGTACAGTTCTTGGAAATGCCATGGGTCTG
TACCCGACTTCTACGATGTATTCTTTGTAAAGATGAATGGGTTCACTCAACTTTTTTTCTG
GCCTTTTGTCTTGTGTCTGCTGTTCCATTCATGAGCAGTCTACAGGAGTTTGGTCTCATC
ACCCAGTCCACGACGTATGAGTCCTGATAAGAATTT

SEQ ID NO: 50
GGGAGCAGCGCATTTCATCAGACATTTGGTGAGCGACACCATCTCCCCTGCAAGAAAAG
AGAACCTCTCGCGGCTCCTCTCCCGGCGTGGCCCTCGTAGCTGGCTTTGTGCGTAGAA
GATGTGGGCGACCGCGGAAGGAGGTCCTCCCGAGGTCACTCTCGAGACCTCCATGGGT
TCTTTCACTGTCGAGCTATACTTCAAGCATGCGCCGAGGACTAGCCGGAACTTCATCGA
ACTCTCTCGGAGGGGTTACTACGATAACGTCAAGTTCCACCGAATCATCAAGGACTTCAT
CGTGCAAGGTGGGGATCCCACCGGGACTGGAAGGGGTGGAGAATCGATCTACGGTAAA
AAGTTTGAGGATGAGATAAAACCAGAACTGAAGCACACGGGTGCAGGTATCCTGTCCAT
GGCAAATGCTGGACCAAACACCAACGGCAGCCAGTTCTTCATCACCTTGGCGCCATGCC
CTTCACTTGATGGAAAACACACAATATTTGGACGAGTATGCAGAGGAATGGAGATTATCA
AAAGACTTGGGAGTGTCCAAACAGACAACAATGATAGACCGATCCATGATGTGAAGATAT
TGCGGACATCAGTAAAAGATTGATTTCGAGGGAGTTAATTTTCACTGACTGGGTGTTCTG
TGATACCGTCTTTCAACTGCACATTGATCATGGAGCCATATCTGTCCAACATTTGTAATAT
TGTGCAATGGCGTTAGTATCATCTGATGCATGTTATCGCCGGTCACTTCTCTTCGACAAA
AGATACATTATGCAATGCGTTGATGCTTTGTCAGGATTATGTCAGCTTGGGACAAATTTG
GGGCAATTTCATGGCATATGAACTCAACAGGATCT

SEQ ID NO: 51
GAACTTCGATTGCAGTTCACAAGAGGTTGTGGCTAGGATGTCGAATCCAAAGGTTTTCTT
CGACATATTGATTGGTAAGATGAAAGCAGGCAGGGTTGTGATGGAGCTCTTTGCAGACG
TCACGCCCAAGACCGCCGAAAATTTCCGCGCACTGTGCACCGGGGAGAAGGGGATCGG
CAGATCTGGGAAACCGCTGCATTACAAGGGATCAACCTTTCACCGTATAATCCCAAACTT
CATGTGTCAAGGCGGGGATTTCACTAGGGGCAACGGAACCGGAGGGGAGTCTATATAC
GGAATGAAATTTGCCGATGAGAACTTCAAGATTAAGCACACCGGCCTGGGAGTGCTGTC
GATGGCCAACGCGGGACCCGACACCAATGGCTCTCAGTTCTTCATATGCACCGAGAAGA
CCCCGTGGCTGGATGGGAAGCACGTCGTGTTTGGGAAGGTCATAGACGGTTATAACGT
GGTTAAGGAGATGGAGAGCGTGGGTTCTGATAGCGGAAGCACGAGGGAAACGGTTGCA
ATCGAAGACTGCGGTCAGCTATCAGAGAATTGATGGCTAGCACTGTGTAGAAAGGTGAA
TTTAAAGTACTTGTCTACACTGCTTATTAAATCATTGTGGATGAGCCCCTTAAAATTTCTTA
TGATG

SEQ ID NO: 52
GAAGAAGGAGGAGGAGCCGAAGTGCCTGCATTCCGAGAGCCGAAGTGCGTTGGCTTCC
CTTGCGATCGTTCCTCCGTTTCGTGATCCGACCCGAAGACCCCGCGCGCCATGGACGA
CGACTTCGAGTTCCCGGCCTCGAGCAACGTCGAGAACGACGACGACGACGGCATGGAC
ATGGACGACATGGGCGGGGACGTCCCCGAGGAGGAGGACCCGGTGGCGAGCCCCGC
CGTCCTCAAGGTCGGCGAGGAGAGGGAGATCGGGAAGGCCGGGTTCAAGAAGAAGCT
CGTCAAGGAAGGCGAAGGGTGGGAGACCCCGAGCTCCGGCGACGAGGTCGAAGTGCA
TTACACGGGGACTCTGCTCGACGGGACCAAGTTCGACTCGAGCCGCGATCGTGGGACG
CCGTTCAAGTTCAAGCTCGGTCGAGGTCAAGTGATCAAGGGATGGGATGAAGGGATCAA
GACTATGAAGAAGGGTGAGAATGCCATATTCACCATCCCTCCGGAACTTGCCTATGGTG
AATCGGGATCCCCTCCTACCATTCCTCCTAATGCCACTCTTCAATTTGATGTGGAGTTGC
TCTCATGGAGTAGTGTGAAAGATATATGCAAGGATGGAGGGATCCTCAAGAAAGTTCTTG
TGGAAGGAGAGAAGTGGGATAATCCCAAGGACCTAGATGAGGTTTTTGTTAAATATGAA
GCAAGCCTCGAAGATGGGACGCTCATTTCGAAATCAGATGGGGTGGAGTTCACTGTGG
GAGATGGATATTTCTGCGCTGCGTTGGCCAAGGCTGTAAAGACAATGAAGAAAGGAGAG
AAGGTTTTGCTGACAGTGATGCCACAATATGCATTTGGGGAGACTGGCAGACCGGCTTC
TGGGGATGAGGCTGCTGTTCCCCCTGACGCTAGCCTCCAAATTATGCTCGAGCTAGTCT
CTTGGAAGACCGTTTCTGATGTAACCAAAGACAAGAAGGTCTTGAAGAAAACTTTGAAGG
AAGGAGAGGGGTATGAGCGCCCGAATGATGGGGCAGCTGTTCAAGTGAGACTCTGCGG
TAAACTCCAAGATGGAACGGTCTTTGTGAAGAAGGATGACGAGGAGCCTTTCGAGTTCA
AAATAGACGAGGAACAAGTGATCGATGGACTTGACAGAGCTGTCAAAAACATGAAGAAA
GGGGAAGTGGCATTGGTGACAATCCAGCCGGAGTATGCCTTTGGCCCGACCGAATCTC
AACAAGATTTGGCTGTTGTTCCTGCCAACTCCACTGTGTATTATGAGGTCGAGCTGTTAT
CATTCGTTAAGGAGAAAGAATCATGGGAGATGAACAATCAGGAGAAAATTGAAGCTGCA
GCAAGGAAGAAGGAAGAGGGGAATGCTGCATTTAAAGCTGGAAAATATGTAAGGGCTTC
GAAGAGATATGAGAAGGCAGTGAGGTTTATCGAATATGATTCATCATTTAGTGATGAGGA
GAAGCAGCAGGCAAAGACCCTAAAAAACACCTGCAACCTTAACGATGCAGCTTGCAAAT
TGAAACTCAAAGACTTCAAGGAGGCAGAGAAACTGTGCACCAAGGTATTGGAAGGTGAT
GGTAAAAATGTGAAGGCACTCTACCGGAGGGCTCAAGCTTACATACAACTTGTTGATTTG
GATCTTGCAGAGCAGGACATCAAAAAGGCATTAGAGATTGACCCTAACAACAGGGATGT
GAAACTCGAGTACAAGATACTGAAGGAAAAAGTGAGGGAATATAACAAGAGAGATGCCC
AGTTCTATGGGAATATGTTTGCCAAGATGAACAAACTGGAGCATTCTAGAACAGCGGGC
ATGGGGGCAAAGCATGAGGCAGCACCTATGACGATAGATAGCAAAGCGTAGATTGTACT
CGTGAATAGCTAATGGTGATGTATGTTTGGAGTCCGCTTGAGTAATGTGTACCCACATCA
ACTGTCCATGTATTCCTTAACAGACGCTAAAAACTCGTCTATCGACTTGAGACTGTCTTG
GCGTGTATTTTGGAATAAACTATTATCACGTTTTGTTAAATATAATACCAAAACTCCGTGC
TGTGAATAAAAAAAAAA

SEQ ID NO: 53
TTTTGGTTTGTTTCCTTCTTTCCCGCCGTGGAGGAGGAGAAGAGTCTAGAGAGAGAGAG
AGAGAGAGAGGGAGATGGCGAAGCCGAGGTGCTTCATGGACATCAGCATCGGAGGGG
AGCTCGAAGGCAGGATCGTCGGCGAGCTCTACACCGACGTCGCCCCCAAGACGGCCGA
GAATTTCAGGGCCCTCTGCACCGGCGAGAAGGGCATTGGCCCTCACACCGGCGCCCCC
CTCCACTACAAGGGGGTTCGCTTTCATCGTGTTATCAAAGGATTCATGGTGCAAGGTGG
AGATATCTCTGCTGGTGATGGTACCGGGGGAGAATCTATCTATGGCTTGAAATTTGAGG
ATGAAAATTTTGATCTGAAGCATGAAAGGAAGGGAATGTTATCAATGGCAAACTCTGGCC
CAAACACAAATGGCTCCCAGTTCTTCATCACAACAACCCGCACTTCGCATCTGGACGGA
AAGCATGTTGTGTTTGGGAGGGTAGTTAAAGGAATGGGAGTGGTCCGATCAGTCGAGCA
TGTTACAACTGCTGCTGGTGACTGTCCAACTGTTGATGTTGTAATCGCTGACTGTGGAGA
AATTCCTGCTGGGGCGGATGATGGCATTAGAAATTTTTTTAAGGATGGCGATACTTACCC
AGACTGGCCTGCAGATCTTGATGAGAGTCCTGCTGAACTTTCTTGGTGGATGGATGCTG
TAGATTCTATCAAGGCATTCGGGAATGGAAGTTATAAGAAACAAGATTACAAAATGGCTC
TCAGAAAGTATCGAAAGGCCCTGCGCTATCTGGATATCTGCTGGGAGAAAGAAGGGATT
GATGAAGTTGAGAGTTCATCTTTGAGGAAAACAAAGTCACAGATATTCACCAATAGTTCT
GCTTGTAAATTGAAACTATGTGATCTCAAGGGAGCATTGTTAGATGCAGAATTTGCTGTT
CGTGATGGAGAAAACAATGCAAAAGCTTATTTTCGACAGGGACAGGCTCATATGGAACTT
AATGATATCGATGCTGCAGCGGAAAGCTTCTCTAAGGCGTTGGAGTTGGAGCCAAATGA
TGTTGGGATCAAGAAAGAGCTTAATGCTGCCAAGAAGAAGATTTTTGAGAGGCGCGAAC
AGGAGAAGAGGGCATATCGGAAGATGTTTCTATAGCAGATTCTCTACTATTGCAGAGGAA

AGCACAAAGATTCTTGGCAAACCAACTGTATAGATTGGAGGCAGGGATCAGATCAAGATT
TTTATTTGCAGTTAAATAGAGTTGAAATTCATTAGGATAAGAGTTTTTGTTTTACACCCGG
GGGGGGGACCATCGTCGTTGCCCAAATGGCATGACCAGGACATTTTGGAACTGGACATT
TTTAGTCACTTCTCCACGTCGTTTCGGTTTAACTCTAATGGGAGCGTTGGGTGTTCAGGG
CAAAAAAAAAA

SEQ ID NO: 54
TCTCGCTCTAAACTCCAAGTTCTAAACTTGAAGCATTCTTGAAGTTGCGGGACATGACTA
AGAGAAAGAACCCTCTCGTCTTCTTAGATGTGTCGATTGATGGGGATCCTGTGGAGAGA
ATTGTTATTGAGCTCTTTGCTGACACTGTTCCCAGGACAGCAGAAAATTTCCGATCACTC
TGCACAGGAGAGAAAGGAGTGGGGAAAACTACTGGCAAGCCTCTGCACTACAAAGGAT
CGTACTTTCATCGAATTATTAAAGGGTTCATGGCCCAAGGTGGTGATTTTTCAAATGGAA
ATGGCACTGGTGGAGAAAGTATATATGGAGGGAAGTTTGCTGATGAGAATTTCAAACTAG
CACATGATGGACCTGGTCTTCTCTCTATGGCAAATGGTGGGCCAAATACCAATGGGTCC
CAGTTCTTCATAATATTCAAGCGCCAACCTCACCTTGATGGGAAGCATGTTGTATTTGGG
AAGGTTATGAGGGGAATGGAAGTTGTGAAGAAAATTGAACAGGTAGGAAGTGCCAATGG
AAAACCCCTCCAACCGGTGAAAATTGTAGATTGCGGTGAGACTTCTGAAACAGGAACAC
AAGATGCTGTTGTGGAGGAAAAAAGTAAATCTGCAACACTAAAAGCAAAAAAGAAACGAT
CTGCAAGAGATTCGTCTTCTGAATCCCGAGGAAAAAGACGACAGAGAAAATCCCGCAAG
GAGAGAACGAGGAAACGAAGGAGATACTCTTCATCTGATTCCTACAGCTCTGAGAGCTC
TGATAGCGATTCTGAATCTTATTCCTCAGACACTGAGTCAGAATCCAAATCACACTCTGA
GTCATCGGTATCTGATTCAAGCTCAAGTGATGGAAGGCGCAGGAAAAGGAAGTCAACAA
AAAGAGAAAAACTTCGTCGGCAAAGAGGGAAGGATAGTCGTGGGGAGCAAAAGAGTGC
CAGATATGACAAAAAATCAAGGCACAAGAGTGCAGATAGTTCAAGTGATTCAGAAAGTGA
AAGTTCTAGCCGTAGTAGGAGTAGAGATGACAAGAAGAAATCTTCTCGACGTGAGTCTG
CTCGAAGTGTTAGCAAGTTAAAAGATGCCGAAGCAAATTCTCCTGAGAATTTAGAATCAC
CAAGGGATCGTGAAATCAAGAAAGTTGAGGACAACTCATCACATGAAGAAGGTGAATTTT
CACCAAAGAATGATGTTCAGCATAATGGACATGGTACTGATGCAAAATTTGGTAAATATG
ATGATCAACGTCCCCGTTCAGATGGTTCAAAGAAGTCCAGCGGAAGCATGAGGGACAGC
CCCAAAAGGTTGGCTAACAGTGTCCCTCAGGGAAGTCCATCAAGCAGTCCTGCACATAA
AGCCTCTGAACCTTCTTCCTCTATTCGTGCCCGCAATCCATCAAGAAGCCCTGCGCCAG
ATGGTAATTCGAAGCGCATTAGAAAGGGCCGTGGCTTCACAGAACGTTTCTCCTATGCC
CGTCGATATCGTACCCCATCTCCAGAAGATGTAACATATAGGCCCTACCATTATGGTAGA
AGAAACTTTCATGACAGGAGGAATGATAGGTATTCAAATTACAGAAGCTATTCTGAGCGC
TCACCACACAGAAGATACAGAAGCCCCCCAAGAGGCAGGAGCCCTCCAAGATACCAAC
GAAGAAGAAGTCGAAGCAGGAGTGTGTCTCGCAGCCCGGGTGGTAATAAAGGCCGATA
CCGGGGTCGGGACCAGAGCCGGAGCAGGAGCAGGAGCAGGAGTCGGAGCCCTAGAC
GTGGTTCTAGTCCAGCGAATAAGCAATTGCCACTGAGCGAGAGGTTGAAATCACGTCTT
GGGACCAGAGTCGATGAGCATTCTCCACGCAGAAGAAGGTCCTCCTCGAGAAGCCATG
ATTCCTCAAGATCTAGATCCCCTGATGAAGTACCAGACAAGCATGAGGGCAAAGCTGCC
CCTGTATCTCCTGCAAGGTCTCGATCCAGCTCCCCTTCTGGGAGGGGGTTGGTCTCCTA
TGGCGACGCCAGTCCTGATTCTGGGATCAACTAAAGGGTCTTCCACTTCACATCATCAC
GATTGGCCGAGAGATTTGATGATCTGTGGAGTACCTGCGCCCAACGTTTTGAGTCAGGT
GGACTCGCATGGGGTGTGCAACCGTTGGAATGAACCAAGAGCGGCTTGGACCAGCCTC
ACTCGGTCGGTATTATGTACTCGGTAAATGAAATTTAATGACATAGCTGCTGGATTTTTAC
GAGGTCATCTTAAACGTTCTGAGAATTTCCTGAGATCATGGCCCTCTCATTTGCATCTTAT
AAAACTTTAGTAGACTTGGATTGAAAAAAAAA

SEQ ID NO: 55
TCGATACCTCCTCGGAGCCGGAGGTTTCTGTCGGGGGGAGGCCGGGGCGAACATGTC
GGTGCTGCTCGTGACGAGTTTAGGGGACATCGTGGTGGACTTGCACGCCGATAGGTGC
CCCTTGACCTGCAAGAACTTCCTCAAGCTCTGCAGGATCAAGTACTACAATGGGTGTGT
GTTTCACACTGTGCAGAAGGATTTCACAGCACAGACAGGCGACCCCACTGGGACTGGAA
CAGGCGGCGATTCTGTTTATAAATTTCTTTATGGTGATCAGGCTCGCTTTTTCATGGATG
AGATTCATCTTGATCTTAAACACTCCAAGACAGGAACGGTTGCCATGGCAAGCGGAGGA
GAAAATCTCAATGCTTCTCAGTTCTATTTCACATTGCGGGATGATTTGGATTATCTCGATG
GAAAGCACACGGTGTTTGGTGAAGTTGCAGAAGGCCTTGAGACATTAACTAGGATTAAT
GAAGCCTATGTAGATGAGAAGGGGAGACCCTATAAGAACATCAGGATCAGGCACACATA
TATATTGGATGACCCTTTTGATGACCCTCCCCAGCTAGCTGAGTTGATTCCTGATGCTTC
TCCTGAAGGAAAGCCGAAAGATGAGGTGGTCGATGATGTGCGACTTGAAGATGATTGGG

TACCTTTGGATGAGCAACTAGGCCCGGCTCAGCTTGAGGAGGCTATTCGTGCTAAGGAA
GCACACTCCCGTGCAGTTGTGCTTGAGAGTATTGGAGATATTCCTGATGCTGAGATTAAG
CCTCCAGATAATGTGCTCTTTGTTTGTAAACTGAATCCAGTTACCGAAGATGAGGACTTG
CATACAATCTTTTCACGATTTGGAACCGTGGTATCGGCTGACGTTATCCGTGATTTTAAG
ACTGGAGATAGTTTATGCTATGCATTCATAGAGTTCGAGAACAAAGATTCATGTGAGCAG
GCATATTTTAAGATGGACAATGCCTTGATTGATGATAGGAGGATAAAGGTGGATTTCAGC
CAGAGTGTTGCAAAACTTTGGTCCCAGTTCAAACGAAAAGATAGCCAAGCAGCTAAAGG
AAAAGGTTGCTTTAAATGTGGGGCTCCTGATCACATGGCCAGGGAATGTCCTGGAAGCT
CCACTAGACAGCCACTTTCAAAGTACATTCTGAAGGAGGACAACGCACAGAGGGGTGGA
GATGACTCTAGGTATGAAATGGTCTTTGATGAAGATGCACCAGAAAGTCCAAGTCATGGA
AAGAAAAGGCGAGGACGAGATGACCGGGATGATAGACATAAGATGAGCAGGCAAAGTG
TAGAAGAAACAAAATTCAATGATAGAGAGGGAGGGCACTCAGTGGATAAGCATAGGCAG
AGTGAAAGAAGTAAGCATAGAGAGGATGAGATGAGTCGAGACAGTAAAGCTAGTGAGGC
TGGTAGGAGAAGGATTGACAGGGATTTTCCCGAAGAGGAAAGAGGATGGAGAGAAATATA
CGGAGAGTCACAGAGACAGAGATGGTAAGAGAGGTGATTATCGAGATTATAGAAAGGGA
AGAGCAGATGTTCAGACTCATGGAGACAGGAGAGGTGATGAAAACTATAGGAGAAAGAG
TGCTGCTTATGATGATGGCCATGAGGGTGCGGGAGCTGCTAGGAGAAAAGATTCCAATG
ATGATCACCATGCATATAGGAGAGGATATGGAGATTCTAGGAAGGGAACTAGAGATGAA
GATGACGATGGACGAGGAAGGAGGGATGACCCTAGTTATAGGAGAAGCAGTGGACACA
AGGATAGCTCCAACGGTGGAAGGGAAGAGCAGAAGTATAGAAGTGGAGAGACAGATGG
CAAAAGTCATCCGGAGAGAAGCCATCGAGGTGATAGGAGGAGATGAAGGGCGGGCCAG
AGTACAGAGACATTTCTTCTACTGGAGAGCTCCCTTGTAGGGATGGGTGCCTAATGCTG
TCCCGTCGGGGTGCTGCATGACTTTAAATAGCTGATTCTGACCATATCCTAAAATTCTGG
TTTTGATTGGGTCATTTTGCACTCAAAAAAAAAA

SEQ ID NO: 56
GCGAGCATGAGGCCTTTCAATGGCGGATCCTCCATCGCTTGCCTCGTCCTGGTGATCGC
CGCCGGCGCGCTGGCCGAGTCGCAAGGGCCCCACCTCGGATCGGCTCGCGTCGTTTT
CCAGACTAATTACGGGGACATCGAGTTCGGGTTCTTTCCGGGCGTCGCTCCGAGGACG
GTGGATCACATCTTCAAGCTCGTCCGCCTCGGGTGCTATAACACCAATCACTTCTTCCG
GGTGGATAAAGGTTTTGTTGCCCAAGTCGCTGATGTTGCCAACGGGAGAACCGCGCCCA
TGAATGATGAGCAGAGAACGGAAGCCGAGAAAACTATCGTTGGGGAGTTCAGTAATGTC
AAGCATGTTAGGGGCATTCTTTCCATGGGGAGATATGACGATCCAGACAGTGCACAATC
CTCTTTCTCAATACTTCTCGGAGATGCTCCGCATCTTGATGGCAAGTATGCTATATTTGGT
AGAGTTACCAAAGGTGACGAGACATTGAAAAAGCTTGAGCAGCTACCTACTCGCCGTGA
AGGGATGTTTGTAATGCCAACTGAACGCATCACAATTCTATCATCATATTATTATGATACT
GGGGCTGAGAGTTGTGAAGAGGAGAATTCAACTTTGAGGCGCAGGCTTGCTGCTTCAG
CTGTTGAGGTAGAGAGACAGAGGATGAAATGCTTCCCGTGATGCTTTCTAGATCATTGG
AAGGAAGTGGTCTGATTTCCTTATCATTCATCCATTGTGCTTTGATGTATCCTCAGTGTAC
TGCTTTTAGCTATGTATAGATCGAGTCAACTCATTGAAGATATGGGGAAGGAAGATTCAA
CTGTTCCATTCATTCTGTAAGTTTAACAAAAAGCATGAAGATGAAGGTTACTTTGTGAGAA
CATTCTTGTCCACTAAGCTGACATTGTTAGTGGATTTACTTGAGAGGTGAGAAATGAAAG
ATCATAGACATTTTCACTTCCAAAGTTGGGAACTCAAGCCATTTATCCATATATATATTGTT
CCTGAATGCTGTTTTTTCAGCCAAAAAAAAAA

SEQ ID NO: 57
AGAAAGTAAGGTTGATCAATTTGGCGGAAAGTCAGGCTTTTGGTGCAGCTTTTTGCAGAA
TCTCAGAGGTTTCGAACTCAGGATGCCGAACCCAAAGGTTTTCTTTGATATGCAGGTCG
GCGGTGCCCCAGCCGGCCGGATCGTGATGGAGCTCTATGCGGATGTGGTGCCAAAGAC
GGCTGAGAACTTCCGCGCGTTGTGTACCGGCGAGAAAGGCACCGGCCGCTCAGGCAAG
CCTCTGCACTTCAAGGGTTCGTCGTTCCACCGTGTGATCCCAGGGTTCATGTGCCAGGG
CGGTGACTTCACAAGGGGCAACGGTACCGGTGGAGAGTCGATCTACGGCGAGAAGTTT
GCCGATGAGAACTTTGTAAAGAAGCACACGGGGCCTGGCATCCTCTCCATGGCCAACG
CCGGCCCTAACACTAACGGCTCCCAGTTCTTCATCTGTACCGCCCAGACCTCGTGGCTG
GATGGTAAGCATGTCGTATTTGGTCAAGTTGTAGAGGGCTTGGAGGTCGTGCGCGATAT
CGAGAAGGTTGGATCTGGATCTGGCAGAACTTCAAAGCCGGTTGTCATTGCCGACTCTG
GACAGCTCGCTTGAATTTTTATTATTTACCTTCGCCTTTACGCTGCATACGTTAATAGGTT
ATTATTTCCTTCAACCATTACGCTGCATAGGTTGTTAGCGTATTGTTTCCCTTTACCATTA
CGCTGCATGACTCCCTAGGGTTTGTCAGCATAGGCGTTTTAAGGGTTTTTTGCATCTTTC
TACTCAAGATAGTCGCTGCATAAGTTCCTAGGGTTTGTCAGCAAAGAGGCTTCAAAGGTT

TTTGTGTCTTTTCTAGTTATTATAAACGCTTCATAGGTTCCTAGGGTTTGTCAGCATATAG
ATTTCAAGGGTTTTTTTAGATCTTTGGAGTTGAGATAAACGCTATGGCAATAACCCAGTAA
TGTTTGTTTTTATCATATGAAATTTTTACATCTGGAGTTGCATTCGCAGTAAAAAAAAAAA

SEQ ID NO: 58
CTTGCTTGCGGCTCGGGATCCGATCAACAACACACGTTTCAACAGATACCGACGATTCC
TGTCCTTAGAAACACGTTCGACATCTTCTATCTCTTGTCGAAATCACACAAATCAAAATCC
AATCATGCGTTTCACCAGCATCACCAGCGCCATTGCGCTCTTCGCCGCCGCTGCTTCGG
CGCTCGACAAGCCGTTGGATATCAAGGTTGACAAGGCGGTCGAGTGCTCGCGCAAGAC
CAAGGCCGGCGACAAGATCCAAGTGCACTACCGCGGCACCCTCGAGGCAGACGGCAG
CGAGTTCGACGCCTCGTACAAGCGCGGCCAACCGCTCAGCTTCCACGTCGGCAAGGGC
CAGGTGATCAAGGGATGGGACCAGGGCCTCCTCGACATGTGCCCCGGCGAGAAGCGC
ACTCTCACCATCCAGCCCGACTGGGGCTACGGCAGCAGAGGCATGGGACCCATTCCCG
CCAACAGCGTTCTGATCTTCGAGACCGAGCTCGTTGAGATTGCTGGTGTTGCCAGAGAG
GAGCTTTAGATGGGTAATTTGTCCATGACAATCGTAGTCGAAGACACGATACGCTCTTAG
ATGGTACGGAAATCTGATTTGAGCTTTCTCCAAAAAAAAAA

SEQ ID NO: 59
TGAATTTTCCCGAGAAAACCAGGAACCTGCCAAATATCTCTCTGAAAGATCTCCATGGGT
AACCCGAAGGTGTTCTTCGACATGTCGATCGGCGGCCAGCCGGCCGGCCGGATCGTGA
TGGAGCTCTACGCCGACGTGGTGCCGCGCACGGCGGAGAACTTCCGCGCGCTCTGCA
CCGGGGAGAAGGGCGCCGGCCGCTCCGGGAAGCCCCTCCACTACAAGGGCTCGAGCT
TCCACCGCGTGATCCCGGGGTTCATGTGCCAGGGCGGCGACTTCACCGCCGGGAACG
GGACCGGCGGCGAGTCGATCTACGGCTCCAAGTTCGCCGACGAGAACTTCGTCAAGAA
GCACACCGGCCCGGGTGTCCTGTCCATGGCGAACGCCGGCCCGGGGACCAACGGCTC
CCAGTTCTTCGTCTGCACCGCCAAGACCGAGTGGCTCGACGGCAAGCACGTCGTGTTC
GGGCAGATCGTGGACGGGATGGACGTGGTGAAGGCCATCGAGAAGGTGGGGTCCAGC
TCCGGCAGGACCTCGAAGCCCGTCGTCGTCGCCGACTGCGGCCAGCTTTCCTAGATCG
GGCGGCCCCCGTCGGTCGCCGGGATCTCCCCCCCCCCCCCTCCCCGTGTGCGTGTGG
GATCTTATCTGATCTCCAGTCCGTTTCCGAGCATGGTTTTTTAGGGCTTTCCTTTTTTTTT
TCGGCTTTTAGCGTGTGGGGTGTTCGGCCAGATCTGGTTATGGGTCTCTCCGCGGATCC
TTGTTGTCTGTAGGATCGTCGAGACCTTTTATGTTGGTTATGTTGAACTGCCAGTCGGCC
ATTAACTGCTGGAAAAATAAGATAGGGAGCTTTTCTTGAAAAAAAAAA

SEQ ID NO: 60
ATCTCCCATGGCTGCTGACGACTGACCCCCCGTCAAAGCGCAGAGAAGAGACAGCGAC
CGCGCACCCGCAAGGCCCTCCGCCCTCTCCGGCGAGTCACGGGAGAAAGACGGGAAA
TTTAAGCGGGGGAGGAGAGCGAATTCGGGGAATGGCCGTCGCGACGAGGTCGAGATG
GGTCGCCATGTCGGTGGCGTGGATTCTGGTCTTGTTCGGAACCCTGGCTCTCATCCAGA
ACCGATTAAGCGACACCGGAGCTTCGTCTGATCCGAAACTTGTTCATCGCAAAGTTGGT
GAGGAAAAGAAGAAGCCGGACGATTTGGAAGAAGTGACTCACAAGGTTTTTTTCGATGT
CGAGATCGGAGGAAAACCAGCGGGTCGAATTGTCATGGGCCTCTTTGGCAAAACCGTC
CCCAAAACAGTCGAAAACTTCCGAGCTCTTTGCACGGGGGAGAAAGGGATTGGAAAGAG
TGGGAAGCCCCTGAATTACAAAGGGAGCCAATTCCACAGGATCATTCCTAAATTTATGAT
CCAGGGCGGTGACTTCACTCTTGGAGATGGGAGAGGGGGAGAGTCAATCTATGGGAAC
AAATTTTCTGATGAAAACTTCAAACTGAAGCATACAGATGCAGGGCGCCTTTCGATGACA
AATGCTGGGCCAGACACAAATGGGTCACAATTTTTCATTACCACTGTGACAACTAGCTGG
TTGGATGGCCGCCATGTTGTGTTTGGAAAGGTGCTGTCTGGGATGGATGTCGTCCACAA
GATTGAGGCTGAAGGGGGACAGAGCGGCCAACCTAAAAGTATCGTTGTCATTTCAGACA
GCGGAGAATTAGATCTATGAACACTTCCCTGCACACTCCCCTCCAGGTTGCACTTGACAA
GTTTGACCTCATAGCTTACTCGACACGAATGTAGGGCTCAGGCATCGTGCACTGCTTTG
ATGCAAATGTTTTTTTTTCTTTAATGGTGAAGAAGAAAATGTGAGGCAAGCTTGTTGATTT
GATATTGGGACATACTTGAAAGAGCTTGGATCGATCATAGAAGCCAGCCATGAATAGAG
ATAACTTTTCTGAGTGTGAATTGGATATTACGTTGCAAATAGCCGAATGAAATGTAGAAAA
CAGGATGTTTGTTCTGAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 61
GGGAAGAATCCCGTCCGGTGGAAACCCGAACGCCGTTCGTTCCCTCTCGACGAGTCGC
CGGCGAAGCTCAGCTCGTCGAAACGAGCAAAGAATCTCTCCTCCTCCTCCTCCTCCTCC
TCCTCCTCCCGCCGCCGCCGCCTCCGCCGCCGCAGCAGCAGCAGTCGAGATGGCCGT

CACTCTGCACACGAATCTCGGCGACATCAAGTGCGAGATCTTCTGCGACGAGGTCCCCA
AAGCTGCCGAGCACAATGCAAGGGGAATACTCTCAATGGCTAATAGCGGGCCTAACACC
AATGGAAGTCAATTTTTCATCGCATATGCAAAACAACCACATCTGAATGGATTATACACCA
TCTTTGGCAGAGTGATTCACGGGTTTGAGGTCCTAGATATCATGGAAAAGACTCAAACAG
GGCCAGGGGACAGACCTCTTGCGGAGATCAGACTCAATCGTGTCACAATTCATGCAAAT
CCTCTTGCTGGGTAGCCTTTCTTGATGGGTTTTTTTTTTTCTTATCTTGATAGGACTCTTTT
ACAGTTTTGCTTTAGGTTAGGGATCCCTGTAAGCTGATGATAGATATTGGAGATGGTACT
TGTAAGATTCAATTACATCGAACTACTTCAAACCTTGTATTGGAGCATTCCAAATACAAAA
GATCGAAAGATGCTTGTTGATATGTGCCGGTGGCTCATCTGTTGTGCTGGTTGCATGAT
GGTCTTTGTTTGCGATGTATTCGCCGTTCATCTCAAAAAAAAAA

SEQ ID NO: 62
AGTCGAAGACGAAAGAGGAAAAGCAAGCTCCTCGAAGAGAGCTCGAGCATCTCCCATG
GCTGCTGACGACTGACCCCCCGTCAAAGCGCAGAGAAGAGACAGCGACCGCCCACCCG
CAAGGCCCTCCGCCCTCTCCGGCGAGTCACGGGAGAAAGACGGGAAATTTAAGCGGGG
GAGGAGAGCGAATTCGGGGAATGGCCGTCGCGACGAGGTCGAGATGGGTCGCCATGT
CGGTGGCGTGGATTCTGGTCTTGTTCGGAACCCTAGCTCTCATCCAGAACCGATTAAGC
GATACCGGAGCTTCGTCTGATCCGAAACTTGTTCATCGCAAAGTTGGTGAGGAAAAGAA
GAAGCCGGACGATTTGGAAGAAGTGACTCACAAGGTTTTTTTCGATGTCGAGATCGGAG
GAAAACCAGCGGGTCGAATTGTCATGGGCCTCTTTGGCAAAACCGTCCCCAAAACAGTT
GAAAACTTCCGAGCTCTTTGCACGGGGGAGAAAGGGATTGGAAAGAGTGGGAAGCCCC
TGAATTACAAAGGGAGCCAATTCCACAGGATCATTCCTAAATTTATGATCCAGGGCGGTG
ACTTCACTCTTGGAGACGGGAGAGGGGGAGAGTCAATCTATGGGAACAAATTTTCTGAT
GAAAACTTCAAACTGAAGCATACAGATGCAGGGCGCCTTTCGATGGCAAATGCTGGGCC
AGACACAAATGGGTCACAATTTTTCATTACCACTGTGACAACTAGCTGGTTGGATGGCCG
CCATGTTGTGTTTGGAAAGGTGCTGTCTGGGATGGATGTTGTCCACAAGATTGAGGCTG
AAGGGGGACAGAGCGGCCAACCTAAAAGTATCGTTGTCATTTCAGACAGCGGAGAATTA
GATCTATGAACACTTCCCTGCACACTCCCCTCCAGGTTGCACTTGACAAGTTTGACCTCA
TAGCTTACTCGACACGAATGTAGGGCTCAGGCATCGTGCACTGCTTTGATGCAAATGTTT
TTTTTTCTTTAATGGTGAAGAAGAAAATGTGAGGCAAGCTTGTTGATTTGATATTGGGACA
TACTTGAAAGAGCTTGGATCGATCATAGAAGCCAGCCAGAAT

SEQ ID NO: 63
CTTCTCTGCCCCTTGAATTTTCCCGAGAAAACCAGGAACCTGCCAAATATCTCTCTGAAA
GATCTCCATGGGTAACCCGAAGGTGTTCTTCGACATGTCGATCGGCGGCCAGCCGGCC
GGCCGGATCGTGATGGAGCTCTACGCCGACGTGGTGCCGCGCACGGCGGAGAACTTC
CGCGCGCTCTGCACCGGGGAGAAGGGCGCCGGCCGCTCCGGGAAGCCCCTCCACTAC
AAGGGCTCGAGCTTCCACCGCGTGATCCCGGGGTTCATGTGCCAGGGCGGCGACTTCA
CCGCCGGGAACGGGACCGGCGGCGAGTCGATCTACGGCTCCAAGTTCGCCGACGAGA
ACTTCGTCAAGAAGCACACCGGCCCGGGTGTCCTGTCCATGGCGAACGCCGGCCCGGG
GACCAACGGCTCCCAGTTCTTCGTCTGCACCGCCAAGACCGAGTGGCTCGACGGCAAG
CACGTCGTGTTCGGGCAGATCGTGGACGGGATGGACGTGGTGAAGGCCATCGAGAAGG
TGGGGTCCAGCTCCGGCAGGACCTCGAAGCCCGTCGTCGTCGCCGACTGCGGCCAGC
TTTCCTAGATCGGGCGGCCCCCGTCGGTCGCCGGGATCTCCCCCCCCCCCCCTCCCCG
TGTGCGTGTGGGATCTTATCTGATCTCCAGTCCGTTTCCGAGCATGGTTTTTTAGGGCTT
TCCTTTTTTTTTTCGGCTTTTAGCGTGTGGGGTGTTCGGCCAGATCTGGTTATGGGTCTCT
CCGCGGATCCTTGTTGTTTGTAGGATCGTCGAGACCTTTTATGTTGGTTATGTTGAACTG
CCAGTCGGCCATTAACTGCTGGAAAAATAAGATAGGGAGCTTTTCTTGAAATGCGACCCT
CTTTTTACCCTGTAAAAAAAAAA

SEQ ID NO: 64
AGAAGAGATTGCTGGTTTTGGCAGAGGGCGGAGGCGCAGATAAGAAGAGGGAGGCGAT
AGAGCAGCTCCTTCGCGAAAAATCCTCTTCATCGCTTCGCTTCGGGTTGGGGCGGAGGT
CGGCTCCTGTTAGGGTTTCGGCCGGCCAGATCCGCGGGCGGATCTCTCCTTTCTCCGG
GTTTGGATGAGCCCGGTCGCTGCGAACGCCATGGAGGAGGCGGCGGAACCCGAGGTC
CCAGCTCCCGTCACTCCCAGCAAGGACGACGCCGACACCGACGCCGCCGTGTCCCGCT
TCTTGGGATTCTGCAAGAGTAAATTGGGCCTGGCCGAGGGAAATTGCGTGCAGTCTTCG
ACGCTGCTGAGGAAAACCGCACATGTGCTGCGGTCGAGTGGGACTGTTATCGGTACTG
GAACGGCGGAAGAAGCAGAACGCTATTGGTTCGCTTTTGTGTTGTACACTGTTAGAAGG
GTAGGCGAGAGGAAAGCCGAGGATGAGCAAAATGGATCGGACGAGACTGAGGTCCCTT

TGTCCCGGATATTGAAAGCTTCAGTACTAAACCTCATCGACTTTTTCAAAGAGATTCCGC
AATTTGTTATCAAGGCTGGGGCAATCGTAAGCGGCATATATGGTGCTAATTGGGACAGC
AGGCTAGAGGCCAGAGAGATGCAGACAAACTACGTGCATTTATGCATTTTGTGCAAGTTT
TATAAACGTATATGCGGTGAATTCTTCATTCTAAACGATGCAAAAGATGATATGAAGTCCG
CTGATTCAAGCACTTCTGACCCCGTGATTATGTACCAACCTTTCGGATGGTTGCTGTTTT
TGGCGCTTAGAATACATGCCTTGAGCCGATTTAAGGACCTTGTTTCTAGCACAAATGCTT
TGGTCTCTGTACTGGCTATTCTAATAATTCATCTTCCAACCCGCTTCAGAAAATTCAGCAT
TTCTGACTCATCACAACTTGTTAAGAGAAGTGAGAAGGGGGTTGACCTTGTTGGATCACT
CGCTTATCGTTACGATACCTCAGAAGATGAGATCAAGAGAACATTGGAAAAAGCCAATAA
CGTGATAGCAGAAATTTTGGGGATAACTCCGCCTCCAGCTTCAGAGTGCAAAGCAGAGA
ATCTGGAGAATGTTGACACAGATGGTTTGATCTACTTTGGAAATCTCATGGAGGAAACAT
CTTTGTCATCCATTTTAAGTACTTTAGAAAAAATTTATGAAGATGCAACTCGTAATGACAG
TGAATTCGATGAGAGGGTATTTATAAATGACGACGACAGCTTGCTTGTATCGGGTAGCTT
GTCTGGGGCCGCCATTAATTTAACTGGTGCCAAGAGGAAATATGACTCATTTGCCTCCC
CTGCAAAGACTATCACAAGACCACTCTCTCCCAGCCGCTCTCCAGCATCTCACATCAAC
GGTATAATTGGTGGCACTAATTTGAGGATCACTGCTACTCCTGTGGCCACTGCTATGACA
ACTGCCAAGTGGCTTCGGACGTTCGTGTCCCCACTCCCATCAAAACCTTCGACTGATCT
ACAGGGATTTCTAGCGTCATGCGATAGGGATGTGACCAGTGATGTGATACGTAGGGCCA
ATATAATTTTGGAGGCCATCTTTCCAAACAGTCCTATTGGCGAGCGTACTGTAACTGGAG
GCTTGCAAAATGCTAATCTCATGGACAACATGTGGGCTGAGCAACGAAGGCTGGAAGCC
CTTAAGCTTTACTATCGGGTTTTGGAAGCTATGTGTAGAGCAGAGGCACAAATTTTACAT
TCGAATAATTTGACCTCTTTGTTAACAAATGAGAGGTTTCATAGGTGTATGCTTGCATGTT
CCGCTGAGTTGGTTCTTGCAACACATAAGACTGTGACGATGTTGTTTCCTGCAGTGCTAG
AGCGAACAGGTATTACAGCTTTTGACCTTAGTAAGGTGATAGAGAGTTTTGTTAGACATG
AGGAAACTCTTCCCAGAGAATTAAGAAGGCATCTGAATACATTAGAAGAGCGACTTCTAG
AGAACATGGTGTGGGAGAGAGGTTCTTCAATGTACAATTCCTTGGTAGTGGCAAGACCA
GCTCTTGCTCCAGAGATAAATCGGCTAGGCCTATTACCAGAACCAATGCCGTCCTTGGA
TGCCATTGCCTTGCTTATTAATTTTTCTTCCAGTGGATTGCCCCAGTCACCGGTGCAAAA
GCACGAGGCTTCTCCTGGTCAGAATGGGGATATCAGGTCTCCCAAGAGAATTAGTACAG
AATACCGGAGTGTATTAGTTGAAAGAAACTTCACTTCACCAGTAAAAGATCGACTGTTAG
CCTTGAGCAATATCAAGTCAAAGCTACCGCCACCTCCACTTCAGTCTGCATTTGCCAGTC
CAACTCGACCACATCCGGGAGGTGGAGGGGAAACATGTGCAGAAACCGCAATCCATATA
TTCTTTAGCAAGATTACTAAGCTGGCAGCCGTCAGAATTAATGCCATGCTTGAAAGGCTA
CAACTCTCTCAGCAGATAAAGGAAGGTGTTTATTGCCTATTTCAGCAAATACTCAGTCAG
CGAACTAATCTCTTCTTTAATCGCCATATTGATCAAGTAATACTCTGCTGTTTCTACGGTG
TTGCGAAGATCAATCAAATAAACCTGACCTTTAGGGAGATCATTTACAACTATAGGAAGC
AACCCCAGTGCAAGCCACAAGTGTTCCGCAATGTTTTTGTTGATTGGTCAACCCGGCGG
AATGGGAAAGCGGGGAATGAGCATGTGGATATTATCTCTTTCTACAATGAAATATTCATT
CCTTCTGTGAAACCGCTGCTCGTTGAACTTGGGCCCACAGGAGCAACTACTAGAACAAA
CCGGACTTCTGAAGTTGGCAATAAAAATGATGCTCAATGCCCTGGGTCACCTAAGATATC
TTCTTTTCCAACTCTCCCGGATATGTCCCCTAAGAAAGTATCTGCATCCCACAATGTCTAC
GTATCTCCACTGAGATCATCTAAGATGGACGCTTCAATCTCTCATAGCTCGAAAAGCTAT
TATGCTTGTGTTGGAGAGAGCACCCATGCCTACCAAAGCCCTTCCAAAGACCTTGTTGC
CATCAACAGTCGCCTGAATGGTAACCGAAAGGTGAGAGGCACGCTTAATTTTGATGATG
TTGACGCTGGCCTCGTCAGTGACTCTATGGTGGCCAACAGCCTCTACCTTCAGAATGGG
AGCAGTATGTCTTCATCAACTGCTAAATCATCTGAGAAACCAGAATCATAAAGTTAAATGA
TCCCCAGTCCCATTCCCATGCATGTAAATTCCAGTTTTCTTGTTCCCCTCAACTGGCTAAA
TCAGGAAGTGGAAGCGTGCAGTGAGTTATTTGTATAATAGCTGCCCTTTCCTTATCTTCT
GTCCATGAGACTGCTGCAAGCAAGTTACTTTAGCAGTTTAGTTTAATAAAATCCAATCACC
TTCACATGGTGAAAAAAAAAA

SEQ ID NO: 65
CCCACACCGCTCGGCTAGGGTTTCTCTCCAACAGCGCACGGACAGGAAGAGCAGAGAC
CCGCCGGCGGTTACACCCTCCTCCGCCGCCCCCCGCCGAATCGAGGCCGTGGCAGCC
CCACCGGAGCTCCGAGAAGCGAGGGCGTTTTCCGATTCGATCGCCGACCATGAGGCCG
ATTCTGATGAAGGGCCACGAGAGGCCGCTGACGTTCCTCAAGTACAACCGGGAGGGCG
ACCTCCTCTTCTCCTGCGCCAAGGACCACACCCCCACCGTCTGGTTCGCCGACAACGG
CGAGCGCCTCGGCACCTACCGCGGCCACAACGGCGCCGTCTGGTGCTGCGACGTCTC
CCGGGATTCGATGAGGCTGATTACCGGAAGCGCGGACACGACGGCGAAGCTGTGGAG
CGTGCAGAACGGGACGCAGCTGTTCACGTTCAACTTCGATTCGCCGGCGAGGTCGGTG

GATTTCTCCATCGGCGATAAGCTCGCGGTGATCACCACCGATCCTTTCATGGAGCTGCC
TTCGGCTATTCACGTTAAGCGCATTGCCAGAGACCCTGCCGACCAGGCTAGTGAGTCTG
TCCTCGTCCTTCGGGGCCATCAAGGAAGAATAGCTAGAGCTGTCTGGGGGCCTCTGAAC
AAGACTATCATAAGTGCTGGAGAGGATGCTGTAATTCGCATTTGGGACTCTGAGACTGG
GAAGCTTCTCAGAGAGTCAGACAAGGAAACTGGCCATAAAAAGGCAGTAACTTCACTTAT
GAAATCCGTTGATGGTTCCCATTTCGTCACTGGTTCACAAGATAAATCTGCCAAGCTATG
GGACATCAGGACGTTGACTCTAATTAAGACATATGTGACAGAGCGCCCTGTTAATGCGG
TTACAATGTCTCCGCTTCTTGATCATGTGGTCCTTGGAGGTGGCCAAGATGCTTCGGCT
GTGACCATGACTGATCATCGTGCTGGAAAGTTTGAAGCCAAGTTTTTTGACAAGATTCTT
CAAGAAGAAATTGGAGGTGTGAAGGGACACTTTGGGCCTATTAATGCTTTGGCTTTCAAT
CCTGATGGGAAAAGTTTCTCAAGTGGAGGTGAGGATGGTTATGTTAGGTTGCATCACTTT
GACCCGGATTACTTCAACATCAAGATTTAGCTTTTGAGAAAATAGTTTTGCAGTTTTGAAA
ATTTGCTTCTGTATTGGCTCTTGATCATTGTGGAGAGAGTATTGTGACAAAGAAACAACT
GCCTTCTCGGAGGTTGGGGGAAGTACACATTGTGGAAAGATATCGTTGTCTGTATGTTG
GGGCGAGTTTTATGGGAAAAGGCTCTGCCTTCAGAGTGGCGGAATTTTTCCAGAATGTT
TATTATCATGTTATTTTCCTATCTTAGATTGGAAATTTTGTTTTCAAGGTTAGGTTGTCTAA
ATTATGAGGATCTTAATGAAAGGAGCTATTTCCTAGCAAAAAAAAAA

SEQ ID NO: 66
GAACGAGGAAAGCCCGCGGATTTGGACAAGTATAAGAACCGGGCTGCCATTGCAGCGA
GCTTCTCTGCGTCTCGCCCCTTCGAGTCCGGCTCCGTTGATCTTTCGGTTTTTTTTTTGC
CCAATTTTCGGTTTTTTTTTTTGGGTTTTCCATAGGTTTTCTTTGTTCTTTGGTCGGGCTTC
GATCCGCATTGAACCTTCTTAGCAGATTAGTCAGTGAAGTGCGAACACCCACGAATTTTT
TTTTTTTTTTTCCTTGGCAAAAGATCGGATCCCAGTGTTTTGGGATCAATCCCTCTGTAAA
TCGCCGCAGGAAAGGAGGAGAATTTCAAAGCATTTTCGTGTTTTCGATTTACCCTCTTGT
AAGATAGCTAGTGATCAGCTCTTTCCGTGCTCAAATCGTTGGTTGCTTTTTTGAGGTTTC
CTGGATCTTGGGTCAGATTGAATTTTAGGGCTTTCTTCTCTTTTGTTATTTTGGGGGTGCT
TACCCTCTTTTGGCGGTTCCTCGTAATTTATTTCTTGTGGTTCGAAATGGACAAGAAGAG
GACGGTGGTGCCGCTCGTGTGCCACGGGCATTCGCGGCCGGTGGTCGATTTGTTTTAC
AGCCCGATCACGCCGGATGGCTTCTTCCTCATCAGTGCGAGCAAAGATTCTAGCCCAAT
GCTGCGAAATGGAGAGACTGGAGATTGGATTGGAACCTTTGAGGGACATAAAGGTGCAG
TCTGGAGCTGCTGCCTGGACACTAATGCTCTACGTGCTGCATCCGGCTCTGCAGATTTC
TCCGCGAAACTGTGGGATGCATTGTCCGGGGATGAATTGCACTCTTTTGAACACAAGCA
CATTGTCCGGTCATGCGCCTTCTCAGAGGACACTCATCTCTTGCTGACCGGCGGTGTCG
AGAAAATTCTTCGAATTTTTGACTTGAACCGACCAGATGCTCCTCCCAGAGAAGTTGACA
ATTCACCAGGTTCAATAAGAACTGTTGCATGGCTCCATAGTGATCAGACCATATTAAGTT
CCTGTACTGATATTGGTGGCGTAAGGTTATGGGATGTAAGGAGTGGGAAAATTGTTCAAA
CTTTGGAGACAAAATCCCCTGTCACTAGTTCTGAAGTGAGTCAGGATGGTCGCTATATTA
CCACAGCTGATGGTTCGACCGTTAAATTCTGGGATGCAAACCACTTTGGGTTGGTGAAG
AGCTACAACATGCCCTGCAATATCGAGTCAGCCTCATTGGAACCAAAGCTTGGGAATAAA
TTTATTGCTGGTGGAGAAGACATGTGGGTTCACATTTTTGATTTCCATACTGGAGAAGAG
ATTGGATGCAACAAGGGTCATCACGGTCCTGTCCATTGCGTGCGATTTTCACCCGGTGG
GGAATCCTACGCATCTGGATCAGAAGACGGCACTATAAGAATTTGGCAGACTGGGCCTG
CAAATAATGTCGAGGGTGATGCAAATCCAAGCAACGGCCCAGTAACTGGTAAAGCAAAA
GTCGGGGCAGATGAAGTTACTCGTAAAGTGGAGGACTTACAAATTGGCAAAGAAGGGAA
AGATTGGCGAGAAGGATAATGCTGCAGACGCATGATGTGCTCTTTAGGTTTGATCTGTCT
GTTTTGTCTATCCTGCGAGTTTCGAGCATGTGCGTGTGTGATGCTTGTGCTTTGCAAATA
AAGTGGACATGCTCTTGATAAGTTTCTTTCCTCTGCCATCCTTTTAAATGAATCTCCCTTG
GAAGCTCGCTTTCTCCTCTTGTTCCTCTTGTAACAGCAGACACCGCCGAAGTTTGGAGAC
TTCATTTTTATGCTTTCCGGATTGTGCTTAATGAAGAAGAATCTGGAATTTTGCTAAAAAA
AAAAAAAAA

SEQ ID NO: 67
CCTAATTCCCAGGGGAAGGGGAGTATATAGCCTCGCTTCCATCCAGCCATTCATCTCTG
GTCGCGCCTTTTGCCTCCAGCTCGCCGAAGAAGAAAGCCCCCCCGGCAGAACCGCCGC
CCCCCTCCTCCTCCTCCTCGCCGGAACCACCGTGAGCAATGGCGGAGGGACTAATCCT
GAAGGGCACGATGAGGGCCCACACCGACATGGTGACGGCCATCGCCATCCCGATCGAC
AACTCCGACATGGTCGTCACCTCCTCCCGCGACAAGTCCATCATCCTCTGGCACCTCAC
CAAGGAGGAGAAGGTCTACGGCGTCCCCCGCCGCCGGCTCACCGGCCACTCCCACTTC
GTCCAGGACGTCGTCCTCTCCTCCGACGGCCAGTTCGCCCTCTCCGGATCCTGGGACG

GCGAGCTCCGCCTTTGGGACCTCGCCACCGGCGTATCCGCACGCAGGTTTGTGGGCCA
CACCAAGGATGTGCTTTCTGTGGCATTCTCCATCGACAATCGCCAGATCGTGTCGGCAT
CTCGTGATCGCACCATCAAGCTGTGGAATACATTAGGAGAGTGCAAGTACACAATTCAG
GAGGGCGAAGCTCATACTGATTGGGTCAGCTGCGTGAGGTTCAGTCCCAACACACTGCA
ACCCACCATTGTGTCTGCTTCATGGGATAGAACCATCAAGGTCTGGAACTTGACCAATTG
CAAGCTTAGAAACACCTTGGCGGGGCACAATGGCTATGTGAATACAGTAGCTGTATCTC
CTGATGGGTCGCTCTGCGCTAGTGGTGGCAAGGATGGAGTGATTTTGTTGTGGGACTTG
GCAGAGGGCAAGAGGCTGTATAATCTGGAGGCTGGTGCTATAATTCACTCTCTTTGCTT
CAGCCCCAATAGATACTGGCTCTGTGCCGCTACTGAGAACAGTATTAAAATCTGGGACC
TCGAAAGCAAGAGTATCGTCGAAGATTTGAGGGTTGATTTGAAAAACGAAGCTGACAAAA
CTGATGGAACCACGACCGCTGCTTCAAATAAGAAGGTTATATACTGCACGAGCTTGAACT
GGAGTGCAGATGGAAGCACCTTGTTCAGTGGATACAACGATGGCGTGATCAGAGTTTGG
GGCACTGGGCGTTACTAATTCGAGGAACTGCCAGTCGGCCTTTTCTTGATCTTTTCTTGA
AGACTTAGGATGATGTCTTGAAAGTGTTGCAAGTAATTTTTGTTTCACATGACGTGCGGT
TGTTTCGGTTTTGTGGCCAAGTAATCTTTGGTTTTTTTCCCATAATTTTTAAAATGTACGCT
GGTCCTCAATGTAGTTGATTTGTAGCTGATTCTGTTCCAAAAAAAAAA

SEQ ID NO: 68
CTCTTCTCTTCCTCTCCGTCGCCGCTCTCTCTCTCCTCCTCCTCCTCCTCCTCCCGCAGC
AGCAGAGAAACCCTAATCGAAGAGCAGACATGGCGGAGGGGCTGCACCTGAAGGGGAC
GATGAAGGCCCACACGGACATGGTGACGGCGATCGCGGTCCCCATCGACAACGCCGAC
ATGATCGTGACCTCCTCCCGCGACAAGTCCATCATCCTCTGGCACCTCACCAAGGAGGA
CAAGGTCTACGGCGTCCCCCGCCGCCGCCTCACCGGCCACTCCCACTTCGTCCAGGAC
GTCGTCCTCTCCTCCGACGGCCAGTTCGCCCTCTCCGGCTCCTGGGACGGCGAGCTCC
GCCTCTGGGACCTCGCCACCGGCGTCTCCGCCCGCAGGTTCGTGGGCCACACCAAGG
ATGTGCTCTCGGTGGCCTTCTCGATCGACAACCGGCAGATTGTGTCGGCGTCTCGTGAC
CGCACGATCAAGCTGTGGAACACTCTGGGAGAGTGCAAGTACACCATTCAGGAAGGTGA
GGCTCACAATGATTGGGTGAGCTGTGTCAGGTTCAGTCCCAACACCCTTCAGCCGACCA
TCGTGTCCGCATCTTGGGACCGCACTGTGAAAGTGTGGAATTTGACCAACTGCAAGCTG
AGGAACACCCTGCAGGGGCATTCTGGCTATGTGAACACCGTGGCTGTGTCTCCTGATGG
GTCACTTTGTGCGAGCGGTGGCAAGGATGGAGTGATATTGCTTTGGGACTTGGCAGAAG
GGAAGAAGCTGTACTCACTGGAGGCTGGTGCCATTATTCACTCACTTTGCTTCAGTCCAA
ATAGGTACTGGCTCTGCGCAGCCACAGAGAACAGCATTAAGATCTGGGATCTGGAGAGC
AAGAGCATTGTCGAGGACTTGAGGGTTGATTTGAAGAATGAAGCTGATATGAGTGATGG
AACTACGGGGGCCATGAGCTCAAATAAGAAGGTCATCTACTGCACGAGCTTGAACTGGA
GCGCTGATGGAAGCACCTTGTTCAGTGGTTACAATGATGGCGTGATCAGAGTCTGGGGA
ATCGGGCGTTATTAGGAAATTTTTGACGATCGCTACTGCTGAACTGTTTTTCAAAGATTCC
CTTTTTCGTATGTCTTTTTAGTGGTTTAAAGAACCTTTTAGGATCTTGTTTAAGATGCTTGG
AGCAATATATGCCTGCCTCTACTCTTATGTTTAGAAGATTTGGCCTTTGGGCAGTCATATT
TTTTTGAGAACAGAGATATAAAAAAAAAA

SEQ ID NO: 69
CGCTCTCTCTCTCTCTCCTCCTCCTCCTCCCGCAGCAGCAGAGAAACCCTAAGCGAAGA
GCAGACATGGCGGAGGGGCTGCACCTGAAGGGGACGATGAAGGCCCACACGGACATG
GTGACGGCGATCGCCGTCCCCATCGACAACGCCGACATGATCGTGACCTCCTCCCGCG
ACAAGTCCATCATCCTCTGGCACCTCACCAAGGAGGACAAGGTCTACGGCGTCCCCCG
CCGCCGCCTCACCGGCCACTCCCACTTCGTCCAGGACGTCGTCCTCTCCTCCGACGGC
CAGTTCGCCCTCTCCGGCTCCTGGGACGGCGAGCTCCGCCTCTGGGACCTCGCCACCG
GCGTCTCCGCCCGCAGGTTCGTGGGCCACACCAAGGATGTGCTCTCGGTGGCCTTCTC
CATCGACAACCGGCAGATTGTGTCGGCGTCTCGCGACCGCACGATCAAGCTGTGGAAC
ACTCTGGGAGAGTGCAAGTACACCATTCAGGAAGGTGAGGCTCACAATGATTGGGTGAG
CTGTGTCAGGTTCAGTCCCAACACCCTTCAGCCGACCATCGTGTCCGCATCGTGGGACC
GCACTGTGAAAGTGTGGAATTTGACCAACTGCAAGCTGAGGAATACCCTGCAGGGGCAT
TCTGGCTATGTGAACACCGTGGCTGTGTCTCCTGATGGGTCACTTTGCGCAAGCGGTGG
CAAGGATGGAGTGATATTGCTTTGGGACTTGGCAGAAGGGAAGAAGCTGTACTCACTGG
AGGCTGGTGCCATTATTCACTCACTTTGCTTCAGTCCAAATAGGTACTGGCTCTGCGCAG
CCACAGAGAACAGCATTAAGATCTGGGATCTGGAGAGCAAGAGCATTGTTGAGGACTTG
AGGGTTGATTTGAAGAATGAAGCTGATATGAGTGATGGAACTACGGGGGCCATGAGCTC
AAATAAGAAGGTCATCTACTGCACGAGCTTGAACTGGAGCGCTGATGGAAGCACCTTGT
TCAGTGGTTACAATGATGGCGTGATCAGAGTCTGGGGAATCGGGCGTTATTAGGAAATT

TTTGACGATCGCTACTGCTGAACTGTTTTTCAAAGATTCCCTTTTTCGTATGTCTTTTTAGT
GGTTTAAAGAACCTTTTAGGATCTTGTTTAAGATGCTTGGAGCAATATATGCCTGCCTCTA
CTCTTATGTTTAGAAGATTTGGCCTTTGGGCAGTCATATTTTTTGAGAACAGAGATATAAT
ACCTGTTTTTACGGCAAAAAAAAAA

SEQ ID NO: 70
TTTTTTTTTTGGTTTTTCTTTCAATACTTATCTTTTATTCATCAAAGCGAGCTGAGGCGGAG
AGGGAGGGAGGACATGGTTAAAGTTCAAAAGCTAAAGTACAACAGCGTTTACATGGAAG
ACACTTATTAGAGAGACCCATCCATGAGTGTAAGCTTCATGACCGAGTCTCCGGCACGA
CAAAAACCGCCGACGCGATCTTCTACTTCTTGCCACCGGTTGCCTTCTTCTTAGGCTGAG
CCGGCTTCGGCTCGGCCTTCTTGGGGGCCGGCTTGGATGTCAGGGGTTCCAGCGCCG
CCATTCGCGACAACCACGCCGGAGAACGGGACGATGAGCAGCAACTCGCCGGCCTTCC
ACCGCGATTCCGACGACGACGACGACCAGGGCGAGGTCTTCCTCGACGACTCCGACAT
CATCCACGAGGTCGCCGTCGACGACGAAGATCTTCCTGACGCCGACGATGAGGCAGAC
GAGGCAGAGGAAGCGGATGACTCTTTGCACATATTCACAGGCCACAATGGGGAAGTGTA
CAGCCTTGCATGCAGTCCCACAGATGCAACACTAGTGGCAACTGGGGCTGGAGATGATA
AAGGGTTTCTGTGGAGGATAGGTCATGGAGATTGGGCTGTTGAGCTCCAAGGTCATAAG
GATTCCATCTCTAGTTTAGCGTTTAGTCTGGATGGGCAGTTGCTTGCATCTGGAAGCCTT
GATGGAGTCATACAGATTTGGGATGTTCCATCTGGAAATCTTAAAGGCACCCTTGATGGA
CCTGGAGGGGGCATAGAGTGGATCAGGTGGCATCCCAAGGGACACATAATATTAGCAG
GTTCGGAGGATTCCACTGTTTGGATGTGGAATGCTGACAAGATGGCCTACTTGAATATGT
TTTCAGGGCATGGTAACAGCGTAACGTGTGGAGATTTTACTCCTGATGGTAAAACAATTT
GTACCGGCTCGGATGATGCAACATTGAGAATTTGGAATCCCAAGAGTGGGGAAAACATT
CATGTTGTAAAAGGTCATCCATATCATGCCGAAGGACTAACAAGCATGGCAATAAGCTCT
GATTCAGGTCTTGCTATTACTGGTGCCAAAGATGGATCTGTTCGCATTGTCAATATATCA
AGTGGAAGGGTTGTTAGTTCTCTGGATGCTCATGCAGATTCTGTCGAATTTGTGGGACTG
GCTCTAAGCTCCCCGTGGGCTGCAACTGGAAGCTTGGATCAGAAGCTCATTATATGGGA
TCTCCAGCATTCTTCTCCCCGCGCCACCTGTGATCACGAGGATGGAGTGACTTGCTTGA
GTTGGGTCGGTGCATCAAGATTTTTGGCCTCGGGTTGTGTTGATGGCAAAGTACGAGTG
TGGGATAGCCTTTCTGGTGATTGCGTAAGAACATTTCATGGGCATTCCGATGCTATTCAG
TCCCTGTCGGTGTCTGCCAATGAGGAGTTCCTTGTTTCTGTTTCAATTGATGGAACTGCT
AGGGTTTTTGAGATTGCAGAGTTTCACTAGGGATAACACCGGATGCACGATATTGTTTTT
GTCTGTACATTTTCTGTACTCTAAAACTATGATCTTTTCTTTTTGTAGCAGAAACACCCCC
CCCCCACCCCCCCACAAACCCCCCACAAAAAAGAAAAGCATATTATCTTTTTGAAGCGGA
AATATATATTTATGCTACTACATAAGTAATGTACTACTTGACAAGATGAGGAAAGTATTTG
TTTCGCGGTGTGATTTAGTTGATCAATTTACATTTTAAAAAAAAAA

SEQ ID NO: 71
CAAAACTTCAACGAAGTTCCGTTGATTAAGGAATGGGAACCTCTCAGCATCAATTATCTT
CATGCCTCCAGCTTCTTCCAAGACGCCGAGGAAATAAGAACCTCATTTTCAGACGGACG
ATGGCCAGCGGTGGCGCCGCCGCCGTCGCCCCGCCGCCGGGCTACAAGCCCTACCGC
CACCTGAAGACCCTGACCGGCCACGTCGCCGCCGTCTCCTGCGTCAAGTTCTCCAACG
ACGGCACCCTCCTGGCCTCCGCCTCCCTCGACAAAACCCTAATCATCTGGTCCTCCGCC
GCCCTCTCCCTCCTCCACCGCCTCGTCGGCCACTCCGAGGGCGTCTCCGACCTCGCCT
GGTCCTCCGACTCCCACTACATCTGCTCCGCCTCCGACGACCGGACCCTCCGCATCTG
GTCCTCCCGCTCCCCCTTCGACTGCCTCAAGACCCTGCGCGGCCACACCGACTTCGTCT
TCTGCGTCAACTTCAACCCGCAGTCCAGCCTCATCGTCTCCGGGTCGTTCGACGAGACG
ATCCGCATCTGGGAGGTCAAGACCGGCCGGTGCCTCAACGTGATCCGGGCCCACTCCA
TGCCCGTCACCTCCGTCCACTTCAACCGCGACGGCTCCCTCATCGTCTCCGGGAGCCA
CGACGGGTCGTGCAAGATCTGGGACACCAAGAACGGGGCCTGCCTGAAGACGCTGATC
GACGACACGGTGCCCGCCGTCTCCTTCGCCAAGTTCTCGCCCAACGGCAAGTTCATCCT
CGTCGCCACTCTCAATGACACCCTCAAGCTGTGGAACTATGCAACTGGGAAGTTCTTGA
AGATTTACACGGGTCATAAGAACAGCGTCTACTGTTTAACTTCTACATTCTCTGTTACGAA
TGGGAAGTACATTGTTAGCGGCTCAGAGGATCGGTGCATCTGCATATGGGATCTCCAAG
GGAAAAACCTAATTCAGAAACTCGAAGGCCACTCTGATACGGTCATTTCCGTCACGTGC
CATCCATCGGAGAACAAGATTGCATCTGCTGGCCTTGATTCCGATAGAACTGTGAGAATT
TGGCTTCAAGATGCTTAACTGCTTAGGCTGTGATGGATAGATGGATTTCAAGCTGTTTCC
CTGGATTTTGGGGAGGGTAATTTGCAGGTGAAGTGGGCTCGAGCAAGCAAAGTGCTCTT
CCTTGTTCTGATGTGGGTTCAAGCTTTGCCATCCTCGCATGTGATGGATACAATAATCTA
CCAAAGGCTACTTAGACCGACAAGCAGGATGAATGGGTGAAGTCCAAAATAATTGGAAC

ACTTTTGGGGATAAGTTGTTACTAGATTATGATCCAGCATAGATACTCGATGTGGTATAG
AATTTATCCAATGTACTCCTAAATGTAGATACATCGTGTATTGATTTCTCCCCATGTTTGA
GGATACTCAGCCTGAATTACGTATCTAATGGGTTGTCCTGAAAGTGTAAGTCCTACGTCA
TTCCTGATAAGGCTCTTTTTCTGTTTCTCATAGGTGCCCCATTGGTTCTGAAGAAACACTT
CCATCAAAATAAGTTTCATTGAATGTGTACGGGTTCTTTTCTTAGTAACAGCTCCATCTAT
TTGTAAAAAAAAAA

SEQ ID NO: 72
CTGATTCGCAGATTCGCGCAGACGAGGAAATTAGTGGTCAAGAGGAGAAGAGAACTTAG
TCATCATGCCTTCTCAAAAGATCGAGACAGGTCACCAAGACATAGTCCATGATGTAGCTA
TGGATTACTATGGAAAGCGTGTGGCAACAGCTTCGTCTGATACCACTATCAAGATAATAG
GCGTGAGCAATAGCTCTGGATCACAGCACCTTGCTTCATTGAGTGGTCATAAAGGCCCT
GTTTGGCAGGTTGCTTGGGCACACCCTAAATTTGGATCAATCCTTGCTTCTTGCTCATAT
GATGGACAAGTTATCTTATGGAAGGAGGGTAATCAAAATGATTGGGCTCAAGCTCATGTT
TTCAATGATCACAAGTCATCCGTGAATTCCATCGCTTGGGCGCCTCACGAACTAGGTCTC
TGCTTGGCTTGTGGTTCATCTGATGGGAATATCTCTGTCTTTACTGCCCGACCTGATGGT
GGTTGGGACACAACTAGGATAGAGCAAGCTCACCCTGTTGGTGTCACTTCTGTTTCATG
GGCCCCGTCCATGGCTCCGGGTGCTCTAGTTGGATCTGGTTTGCTAGATCCTGTTCAGA
AGCTGGCCTCTGGAGGGTGTGACAATACAGTGAAGGTGTGGAAGCTTTATAATGGAACT
TGGAAGATGGATTGCTTCCCTGCCCTTCAGATGCACTCTGATTGGGTCAGAGATGTGGC
TTGGGCACCTAATTTGGGGCTTCCGAAGTCCACAATTGCTAGTGCTTCTCAAGATGGGA
CTGTTGTAATATGGACTGTGGCCAAGGAAGGAGAGCAATGGCAGGGTAAAGTTTTGAAG
GATTTCAAGACTCCAGTTTGGCGGGTTTCCTGGTCGCTTACTGGAAATTTGTTGGCAGTT
GCTGATGGGAATAACAACGTAACTTTGTGGAATGAGGCAGTGGATGGTGAGTGGCAACA
AGTTACTACAGTTGAGCCATAGATTTGGAGTTTGTCTGTTTTGTTATCTACTTTAATGTGT
TTTGCCTTGCCTAGGACCTCTTTGATAGACTGTTATTGTTATGGTGTTTCTCTTTTTGATTT
TAGAGTTTGTGAATACAATTATTTTAACATAATCTTTTCTTGGGAGATGAATGAAGGGGTA
TTATCAAAAAAAAAA

SEQ ID NO: 73
CCAACCAAACCCCAACACTCTCCAGCTCTACTCTCTCTCTCTACTTTCTCTCTCCATCTCC
TTCGCCGGAATTCGAATGAGATGAAGATCGCGGGCCTCAAGTCCGTCGAGAACGCCCA
CGACGAGTCGGTGTGGGCCGCCGCGTGGGTGCCGGCGACGGAGTCCCGGCCGGCGC
TGCTCCTCACCGGCTCCCTCGACGAGACCGTGAAGCTGTGGCGGCCCGACGAGCTCGC
CCTCGAGCGGACCAACGCCGGCCACTTCCTCGGCGTCGTCTCCGTCGCCGCCCACCCC
TCCGGCGTCATCGCCGCCTCGGCCTCCATCGACAGCTTCGTCCGGGTCTTCGACGTGG
ACACCAACGCCACGATCGCCACGCTCGAGGCACCGCCGTCGGAAGTCTGGCAGATGCA
GTTCGATCCCAAGGGCACCACTCTAGCGGTGGCAGGTGGAGGGAGCGCATCAATCAAG
CTTTGGGACACTGCCACATGGGAACTGAATGCAACCCTCTCGATTCCTCGTCCAGAACA
GCCCAAACCCTCCGAAAAAGGCAACAAGAAGTTTGTCCTCTCTGTCGCGTGGAGTCCTG
ATGGCAGAAGGCTTGCCTGCGGTTCAATGGATGGTACAATCTCTATATTCGACGTGGCT
CGGGCCAAGTTTCTGCACCACCTGGAAGGCCACTTCATGCCAGTGCGATCTCTCGTATT
TTCCCCAGTTGAGCCACGGCTGCTTTTTTCTGCATCCGACGATGCTCATGTGCACATGTA
TGATTCCGAAGGTAAGTCCTTGGTGGGGTCCATGTCTGGCCATGCTAGCTGGGTATTGA
GTGTTGATGTCAGCCCGGATGGAGCAGCACTTGCGACAGGTTCAAGTGACAGGACTGT
GAGGCTGTGGGATCTCAGTATGAGGGCTGCCGTTCAGACGATGAGTAACCATTCGGATC
AAGTTTGGGGGGTTGCCTTTCGACCAATGGCGGGTGCTGGTGTCAGAGCTGGTGGTCG
GCTTGCTAGTGTATCTGATGACAAGAGCATATCACTCTATGATTACTCATGAAGGTATATC
TGCAAATTGTGAAAAGATAAGTTCTTGGAGCCACTGATCGTTAACCTTTTGTTTGAATAGT
CTGCAAGAGATTTTCTTGGTTTCAGATCATTTTTATTTGTATAATTAGCTGACTATGATGAC
TTTCGGAGAAATGGCTCAATATAGGTCTGATCTGATGGGTGCCGTGTATTAACTTCGAGG
TCAATTTGATACCCGATATTTCTGAGAGGTTTTTTTGTGAAGTTATATTGACTATGGGGGT
TACTATCACTAATCAGCTCAGCTGTTGATATAAAAAAAAAA

SEQ ID NO: 74
GCAAGATCGATTGCTCTGTAGAGCAAAGGAAAGGAGAAAAGCATGGAGATCGATCTCGG
AAACCTCGCATTCGACGTCGATTTTCATCCATCGGAGCAGCTCGTCGCCTCCGGCCTCA
TCACCGGCGACCTCCTCCTGTACCGCTACGGCGACGGCTCCTCGCCGGAAAAGTTGCT
GGAAGTGCGAGCGCACGGCGAGTCTTGTCGGGCTGTTCGGTTTATCAACGATGGGAAA
GCGATTCTGACCGGTTCTCCGGATTGCTCGATTCTCGCGACGGATGTGGAGACGGGAT

CCGTCGTCGCTCGAGTCGAAAATGCTCACGAGGCTGCTGTCAATAGGTTGGTCAATTTG
ACGGAGTCTACTATTGCCACGGGAGATGATAATGGGTGCATCAAGGTTTGGGACACCAG
ACAACGTTCTTGCTGCAACACATTTAGTGCCCATGAAGATTTCATATCGGACATGACATTT
GCATCTGATTCCATGAAACTTGTTGTCACAAGTGGAGATGGGACTCTATCTGTTTGCAAT
CTTCGAAGTAATAAAGTCCAAACTCGGTCCGAGTTCTCTGAAGATGAGTTACTATCTGTT
GTCATTATGAAGAATGGAAGGAAAGTTGTTTGTGGAACACAAAGTGGAACATTATTACTA
TATTCATGGGGATTCTTCAAGGATTGCAGTGATCGCTTTGTGGATCTCTCCCCCAGCTCA
GTTGATGCGTTGCTAAAGCTTGATGAGGATAGGATCATTGCAGGAACTGAGAACGGACT
TATCAGTCTGATAGGAATATTACCCAACAGAATCATCCAACCGATTGCAGAGCATTCAGA
TCATCCTATTGAGCGCCTTGCCTTCTCTCATGATAAAAAGTTTCTTGGCAGCATATCGCAT
GATCAGACATTGAAGCTCTGGGATTTGAATGACATACTAGGGTCTGAAGATTCTCCATCA
AGTCAAGCAGCCATAGACGACAGTGATAGCGATGAGATGGATGTGGATGCGAATCCTCC
CGATTCTAGCAAAGGGAACAAGAAGAAGCATTCAGGCAAAGGAAATGATGTTGGCAATG
CCAACAACTTTTTCGCCGACTTAGGCGATTGATGGCTCCAATCTTCATATGCACATGACA
AGTTGAGGTTCATACAAACATATCAAGGCTCGATTCTAGTCGGCACATTTGGCAAGCAAT
TCGTACATGTTTTCATTGACAGTGTGTTCTCAGGGGGGAATTTTGACAGGACTCCATGAC
CAGGGGCTTCTTATTGTTTATCTGACATTCCCAGGGATGTAATACCCCATCTTAAACTCAA
ACAAGTTTATGTATTTTGAGTGTTCTAGGACAACTATATATACAAAGTAGAAATCCAAATT
GCAAAAGGCATATTCCTCCTAAAAAAAAAA

SEQ ID NO: 75
GGGGAAAGTCAAGCTGTCTTTCATCATGGTCATTTGTCCGATCACGATCGCCCGTTCCG
GCGATCGAGACGGCGACTTCAGCTGCTGATTAGCCACGTTCGATTCCGTCTGATTTGGT
CGGAAAAAAAAAGGGGTTTGAAAATGAGTCAGCAGCCGTCGGTGATCCTCGCGACCGC
CAGCTACGATCACACCATCCGATTCTGGGAGGCCAAGAGCGGGCGATGCTACAGAACG
ATTCAGTACCCGGATTCGCAAGTCAACCGGTTGGAGATAACTCCACATAAGCGGTACCT
GGCTGTGGCGGGAAATCCCAGCATAAGATTGTTTGATGTCAACTCAAACACCCCTCAAC
CGGTGATGAGTTTCGACTCCCATACCAATAATGTCATGGCTGTGGGGTTTCAGTATGATG
GAAATTGGATGTACTCAGGCTCTGAAGATGGAACTGTTAGAATATGGGACCTGAGAGCT
CGTGGTTGCCAAAGAGAATATGAAAGCCGTGGAGCTGTAAATACAGTTGTTCTGCACCC
AAATCAGACTGAACTAATATCAGGAGACCAAAATGGAAATATCCGTGTATGGGATTTGAC
AGCGAATTCATGCAGCTGTGAGTTGGTACCAGAGGTGGATACAGCGGTCAGATCCTTAA
CAGTTATGTGGGATGGGAGTCTGGTGGTTGCTGCAAATAATAATGGAACATGTTATGTTT
GGCGCTTGTTGCGTGGGAGTCAGACAATGACCAACTTTGAGCCACTTCATAAACTACAA
GCACATAATGGATATATTCTTAAATGTCTTCTTTCACCTGAGTTTTGTGAACCCCACAGAT
ATTTGGCTACTGCTTCTTCTGACCATACCGTCAAGATTTGGAATGTTGAAGGTTTCACTCT
AGAAAAGACTTTGATAGGACATCAACGCTGGGTGTGGGACTGCGTTTTCTCGGTAGATG
GTGCTTATCTTATAACAGCTTCCTCGGACACAACAGCAAGACTCTGGTCCATGTCGACTG
GTCAAGATATCAGGGTGTATCAAGGACATCATAAAGCAACTACTTGCTGTGCCCTCCACG
ATGGTGCGGAAGGGTCTCCAGGCTGATATTGGAAAAACGCTAGGGCTTAAACCGCATCA
GTTCGCAATTCCCCTAACGGAGGCCGACTGTTTTCCTTTTGAGTAAAAAAGTTCGTTGTT
GTCCGAACAATCACATTCTTCGCAGATTAAAGTGCACATAAGTTGCTAGCGTTGGACTAT
CCAAGTTGTCTTTTGAATATCGAAATGATGAATTTCCGAGGGAAACAGGGCAAGAGATGA
AAGCTTGAAGCTTTTTAGCACGGCCAGCTCATGACTCGTGGATCTCAGATGTCCCAAGTT
TTCCGGTCTTTCTACTTTATTGCAGAAAGTGGTGGGTTCACTGAATTGCTTGTAAATATAC
TCAATAATCAATGTAGTTGTTCGTCGATAAAAAAAAAA

SEQ ID NO: 76
CCTCCTCGAAGCAGGCAGCTGAAACTACTAGCTGGACGGAGCTCGGGAGAGAGAGCAA
AGATGGAAGATGCCATGGACATGGAAGTGGAAGTGGAAGTGGAAGCGGAAGAACACTC
TCCTTCCAGCTCGAATCCGAGCGGCTCTTCCTTCCGCAGATTCGGGCTCAAGAACTCGA
TCCAGACCAACTTCGGCAGCGACTACGTCTTCGAAATTACTCCCAAGTTTGATTGGTCGC
TGATGGGGGTATCGTTGTCGTCCAACGCGGTGAAGCTGTATTCCCCAACGACGGGTCA
GTACTGCGGAGAGTGCAGGGGGCACTCCGATACCGTCAATGGCATTTCGTTCTCGGGA
CCGTCGAGTCCCCACGTCTTGCACTCTTGCTCTTCCGATGGCACCATCCGAGCTTGGGA
CACCAGGTCTTTTAAAGAGGTTTCTTGCATAAGTGCCGGGCCATCGCAGGAGATCTTCA
GCTTTTCGTTCGGTGGTTCAAGTGACAGTCTTCTTTCTGCTGGATGTAAATCTCAGATAC
TCTTTTGGGATTGGAGAAACAAAAAGCAGGTTGCATGTTTAGAAGATTCTCATGTGGATG
ATGTTACCCAGGTTTGCTTTGTCCCTCACCATCAAAATAAGCTTATTTCAGCTTCAGTGGA
TGGGCTGATATGTATATTTGATACGGCTGGAGACATCAATGATGATGAGCATATGGAATC

GGTGATTAATGTGGGAACTTCAATTGGCAAAGTGGGGATATTTGGACAGACATTTGAAAA
GCTCTGGTGTTTGACACATATTGAGACCTTAAGTGTTTGGGACTGGAAAGAGGGAACAA
ATGAAGCCAACTTTGAAGATGCCCGAAAATTAGCATCTGATAGCTGGTCACTAGATCATA
TTGATTATTTTGTTGACTGTCACTCAGCTGAAGAAGGTGAAGGTTTGTGGGTGATTGGTG
GCACAAATGCCGGGACTTTAGGATACTTCCCTGTAAAGTACAAGGGAGGGGCAGCAATA
GGATCCCCAGAGGCCGTTCTTGGAGGTGGCCACTCGGATGTGGTTAGGAGCGTTTTGC
CTATGTCGGGCATGGCAGGGACAACTTCCAAAACCCGAGGCATTTTTGGATGGACGGGT
GGCGAGGATGGTCGCTTGTGTTGCTGGCTTTCTGATGACTCTTCCGCTACAAGCCGATC
GTGGATGTCAAGCAATCTGGTCTTGAAGTCATCGAGAAGTCACCACAAGAAAAATAGGC
ACCAACCTTACTAGTTTTTTTGTCAGATTTGCTTCTTTTATTTCCATGTATCAAGCCGCATC
AATGTTTGTCGCTGCAATTAACATGTGTGCAGTCGATCCCGTATTTGATTGATTTTTTGTA
ATATATTACAGAGCGACCCGTGTTCTGTACTAAAAAAAAAA

SEQ ID NO: 77
GAATCGACGAACAACGCATCGCCGATGCAGCGAAATCCTCTGGATCATTGATTCTCCCT
CCCTTCCCCCACGAGTTAAAAGAAAAATCAATCAAGGAATTTCTACAGGGAACGAACAAA
AATTCAACGAGCTCCCGGATTCGTTATTGCCTCGCCCCGCCCCGCCCGTCGATTTCGCG
GTTGGTTGGGTTACATGTCTCAGCACCAAGAGTATCCGATGGAATACGCAGCAGATGAT
TACGATGTGGGAGAAGTAGAAGATGATATGTACTTCCATGAAAGAGTAATGGGCGATTCA
GACACCGATGAGGATGAGGAATATGATCATCTGGATAATAAGATTACCGATACCTCTGCT
GCTGATGCTCGGAGAGGTAAAGATATCCAGGGGATTCCTTGGGAAAGACTGAGCGTCAC
TCGTGAAAAATACAGAAGAACTAGAATAGAGCAGTACAAGAACTATGAAAATGTACCACA
ATCCGGAGAAAGTTCGGAAAAGGATTGCAAACCCACAAGAAAAGGTGGAAACTATTATG
AGTTTTGGCGAAACACAAGATCTGTGAAGTCCACAATCCTTCATTTTCAATTGAGAAACTT
GGTTTGGTCCACGACAAAGCATGATGTTTACCTTATGTCGCACTTCTCGATCATTCACTG
GTCGTCATTGACTTGCAAGAAGACTGAAGTGCTTGATGTTTATGGACATGTAGCTCCTCG
TGAGAAACACCCTGGAAGTCTCTTAGAGGGTTTTACACAGACACAAGTCAGTACTCTTGC
AGTACGGGATAAATTACTGATTGCTGGTGGTTTCCAGGGGGAGCTTATCTGTAAGAACTT
AGATCGTCCTGGTGTTAGCTACTGTTGCAGAACAACTTATGATGACAATGCTATCACTAA
TGCGGTTGAGATTTACGATTATCCCAGTGGCGCCGTACATTTCATGGCATCGAATAATGA
CTGTGGAGTCAGAGATTTCGACATGGAGAAATTTGAGCTCTCCAGACATTTCACCTTTCC
TTGGCCGGTGAATCATACTTCTCTGAGTCCTGATGGTAAGCTCCTCGTCATTGTTGGTGA
CAACCCTGAGGGGATAGTGGTGGATTCTCAAAGAGGAAAGACCATTAGGCCACTACAAG
GGCACTTGGATTTCTCATTTGCATCTGCATGGCATCCGGATGGTCACATATTTGCCACTG
GAAACCAAGACAAGACTTGCCGTATCTGGGACATTCGCAACTTATCAAAGTCCGTTGCTG
TACTCAAGGGCAACCTTGGAGCCATCCGGTCCATTCGGTTCACCTCCGATGGGCGATTT
ATGGCCATGGCAGAGCCTGCTGATTTTGTCCATGTCTATGATGTGAAGAGTGGGTATGA
GAAGGAGCAGGAGATCGATTTTTTTGGCGAGATCTCTGGTGTGTCTTTCAGCCCTGACA
CTGAATCGCTATTTGTTGGGGTGTGGGATCGCACCTATGGCAGTCTCCTTCAGTATAACC
GATGCAGGAATTATTCATATCTCGACTCCATGTAGTGAGGAGATGCCTGCACCGTTTCTT
AACTAAATAACTGCAATTATATAGGATTCTTGAAATATGGCTAAAGTGGCTTCAGCGCATT
GTGTAAATGTAGATAGGTGATATATTTCTCGTTGCAATGTAGGGTAAGAGAACAAGTATG
GTGTTACTTGGTGGATCATTGTCTTCCTTCTTTAGGTCATAGAATGCTTGCTCTTCATTGG
AAGGTTGAGCGAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGCTTTATGGTGATGAAG
TGTTCTGTTGTTTGCAGAATAATGGAGATGTATTTAATAAAGGTTGGTGGTGTTTGGGTG
AAGATTGAAGCAAACTCTTCTCATTGCCAATAAGAACCCTATTCACTTTCTCTGTCAA

SEQ ID NO: 78
CTCTCTCTCGGCGGCTCCACCTCAAAACCCTATCCCGAAACCTCCTCGCCGGGGAGAG
CTCGAATCGAGCGCCGTCGAGTGCCGATCCGACCGGGAGGGACGAAGGAAATGGGGG
CGAGCAGCGACCCGAACCCGGACGTCTCCGACGAGCACCAGAAGCGGTCGGAGATCTA
CACGTACGAGGCGCCGTGGCACATCTACGCCATGAACTGGAGCGTCCGCCGCGACAAG
AAGTACCGGCTCGCCATCGCCAGCCTCCTCGACCACCCGGCCGCCGCCGCCGCCGTC
CCCAACCGCGTCGAGATCGTCCAGCTCGACGACTCCACCGGCGAGATCCGGGCCGACC
CGAACCTCTCCTTCGACCACCCTTACCCGGCCACCAAGGCGGCCTTCGTCCCGGACAA
GGACTGCCAGCGCGCCGACCTGCTCGCCACCTCCAGCGACTTCCTCCGCATCTGGCGC
ATCGCCGACGACTCCTCCCGCGTCGACCTCCGCTCCTTCCTCAACGGCAACAAGAACAG
CGAGTTCTGCCGCCCCCTCACCTCCTTCGACTGGAACGAGGCGGAGCCCAAGCGCATC
GGCACCTCCAGCATCGACACCACCTGCACCATCTGGGACATCGAGCGCGAGACCGTCG
ACACCCAGCTCATCGCCCACGACAAGGAGGTCTACGACATCGCCTGGGGCGGCGTCAG

CGTCTTCGCCTCCGTCTCCGCCGACGGCTCCGTCCGCGTCTTCGACCTCCGCGACAAG
GAGCACTCCACCATCATCTACGAGAGCTCCGAGCCCGACACCCCGCTCGTGCGGCTCG
GGTGGAACAAGCAGGACCCCCGCTACATGGCCACCATCATCATGGACAGCGCCAAGGT
CGTGGTGCTCGACATCCGCTACCCTACCATGCCGGTCGTCGAGCTGCAGAGGCACCAG
GCCAGCGTGAACGCCATCGCGTGGGCGCCGCACAGCTCCTGCCACATTTGCACGGCG
GGGGACGACTCGCAGGCACTCATCTGGGACCTGTCATCAATGGCGCAGCCTGTGGAGG
GCGGCCTGGACCCGATTCTGGCCTACACTGCAGGGGCGGAGATTGAGCAGCTGCAGTG
GTCGTCCTCGCAGCCTGATTGGGTCGCGATCGCCTTCTCTTTGAAACTCCAATAATCACA
TTTACCATCAACAGGCATCAGCAACATACTGTTGTAGTGTAATTAATTTAATGGAAAAGTT
CATTTGTGTGGTGGGCCTTGGAAAGTGACTTATTGGTGCCCTGTTCATTGGACTTGAGTG
ACTGGTGAAACTGATGTTGCTGATGGCTGTTGATGGTAAATGTGATTATTGGAGTTTCAT
AAAAAAAAAAAAAAA

SEQ ID NO: 79
GTCAACACTCAAAAGCAGCAGCTCAAGTGGCCAGCGCTTCCCGCGGATTAGAGAAATCC
GCAGATCGGCGATGCGCGGCGGCGGCGGCGGCGGCGACGCGACGGGCTGGGACGAG
GACGCCTACCGGGAGAGCGTCCTGAAGGAGAGGGAAGTCCAGACCCGGACCGTGTTC
CGCGCCGCCTTCGCCCCCTCGCCGAGCCCTAGCCCCAGCCCCGACGCCGTCGTCGTC
GCCTCCAGCGACGGCTCCGTCGCCTCCTACTCCATCTCCGCCTGCCTGTCCGACCACA
GATTGCAGTCGTTGCGGTTTGCTGACGCGAAGTCTCAAAACGTTCTAGAGGCGGAACCC
GCTTGTTTCCTTCAGGGGCATGATGGACCTGCCTATGATGTCAAATTTTATGGCGAGGGT
GAAGATTCTCTTCTACTGAGTTGTGGTGATGATGGACGGATTCGGGGGTGGATGTGGAG
AGACATTACGAGCTCAGAGGCGCATGACCATTCACAAGGAAACAGTGCAAAACCAGTAC
TTGATTTGGTGAATCCCCAAAGTAGAGGTCCTTGGGGTGCCCTTTCACCAATCCCCGAG
AACAATGCCCTGGCTGTAGATGTTAAGCGGGGATCCATTTATGCTGCAGCTGGTGATTC
TTGTGCATATTGTTGGGATGTGGAATGTGGTAAAATAAAAACTGTTTTCAAAGGGCATTCT
GACTACCTGCACTGTATAGCTGCACGGAACTCTTCTAGTCAGATTATAACAGGTTCTGAG
GATGGGACAGCAAGAATATGGGATTGCAGAAGTGGTAAGTGTGTCCAAGTAATTGATCC
AGACAAAGACCATAAGAAAGGTTTCTTCGCAAGTGTTAGTTGTCTTGCTCTTGATGCAAG
TGAGAGTTGGTTGGTCTGTGGTCGGGGTCGGGATTTATCTGTCTGGAGTATATCTGCTT
CTGATTGCATAGCAAAAATCTCAACCAATGCTCCTGCACAGGATGTCTTATTTGATGATAA
TCAAATATTACTGGTTGGAGCTGAACCTTTGATCAGTCGTTTAGACATGAACGGGGCGGT
TCTTTCTCAAATACATTGTGCACCTCAGTCGGTGTTTTCTGTTTCTCTGCATCAATCTGGT
GTAACTGCAGTCGGGGGTTATGGAGGCTTAGTCGACGTAATCTCCCAGTTCGGAAGCCA
CCTCTGCACATTTCGCTGCAAATGCATATGAATGAAAAACTGCATCGACATGATATTCTT
GTAACGAACCTGCAGTGCTAAGCCCATGCCTCGGAGGAAATATGGTTGCTCCAAAGAAC
CTTGTGCATTTGCCTTAGTTAACCTTATCTGGCAAGGAGTTGACCTGACTTTTGCTAAAAT
CTGAGGTTTTGGATCTGTGGGCTGCTCTTTTCTTTCACTAAGTTTGACAAAGAGGGTGCT
TGCTTGCAGCCTGAACATGCCGTGAAATGTAGTAGCGTAGAGAAAAAAGAGAGCTCGAC
TTCGTTATCGCAGGAGACTTTTGTGTGGACCATTATACTCCTTTTGATAGATTTTTCCTTC
TAAAATTTTGTATCTTGTAAAAGATTTGGTGATTTCAACGTTTAGAAGGCTAAT

SEQ ID NO: 80
TCTCTCTCTCCTCTCCCTCCCAACCTTTCTCTCTCCTCCCGACCTCGACTGCGGATGGCG
TCGTCGCCGGAGTCACGCGCTGCCTGCTGAAGCGATACGGCGGCGACCATGGAGGCT
CCAATCATCGATCCGCTGCAGGGAGATTTCCCTGAAGTGATCGAGGAGTACTTGGAGCA
TGGCATCATGAAGTGCATCGCCTTCAACCGCCGCGGTACCCTCCTCGCCGCTGGATGC
ACCGATGGAAGTTGCATTATATGGGACTTTGAGACCCGGGGCGTCGCAAAAGAGCTCCG
AGACAAAGAATGCACTGCTGCAATTACGAGTGTCTGCTGGTCCAAGTATGGTCACCGCA
TACTTGTGTCTGCTTCTGATAAGTCTCTTATTCTTTGGGATGTTCTAAGTGGAGAGAAGAT
AGCACATACGACTCTTCAGCATACCGTTCTACAAGCTTGTCTTCATCCTGGTTCCTCAAC
TCCATCTATTTGCCTGGCCTGCCCTTTCTCATCTGCTCCCATGATTGTTGACTTGAACACT
GGAAGCACAACAGCTCTTCCAGTTTTGACTGCTGATGTGAGCAATGGAGCCACTCCTCT
GTCCCGCAACAAGACATCAGATACATCAGTAACTTACTCTCCATGCAATGCCTGCTTTAA
CAAGCACGGGGATCTGGTTTATGCAGGAACCTCGAAGGGAGAAATACTTATAATTGATC
ACAAGAATGTCAGAGTGTGTGCCATCGTCCTTGTTTCTGGTGGCGCAGTTATAAAAAATG
TCGTGTTTAGTAGAAATGGACAGTACATGTTGACAAATTCCAATGACCGGCTGATTAGAA
TCTATAAAAATCTTCTACCTCCTAAAGATGGATTAAAGATGCTTGATGAATTGAATGAGAG
CTTCAATGAATCAGATGATGTGGAGAAATTGAAGGCTATTGGATCGAAGTGCTTAGAACT
TCTCCATGAGTTTCAAGATTCCATCACTAGGGTCCAGTGGAAAGCACCTTGTTTTAGTGG

TGATGGAGAGTGGGTGATTGGTGGTGCTGCAAGCAGAGGAGAGCACAAAATTTATATAT
GGGATAGGGCCGGGCATCTTGTTAAAATTCTTGAAGGCCCAAAAGAAGCATTGATGGAT
TTGGCATGGCATCCTGTTCATCCCATTATCATTTCTGTTTCTTTGACTGGCTTGGTTTATA
TATGGGCTAAAGACTATACTGAAAATTGGAGTGCATTTGCTCCAGATTTCAAAGAGCTCG
AGGAGAATGAGGAGTATGTCGAACGAGAAGATGAATTTGATCTGGTGCCTGAAACTGAA
AAGGTAAAAGGACTAGATGTTCATGAAGATGATGAAGTTGATGTTTTGACTGTGGAGAGG
GATTCAGTTTTTAGTGACTCAGATATGTCCCAGGAGGAACTATGCTTTTTGCCTGCAGTT
CCTTGTCTCGATATTCCTGAGCAGCAAGACAAGTGTGTTGGCAGTTGTTCAAAGTTGCCA
GATGGTAACCATTCTGGATCGCCACTTTCAGTAGAAGCTGGTCAGAATGGAAATGCAAG
CAACCATAATTCAAGCCCTCTCGAACCTATGGAGAATTCAACTGCTGATGACACAGACGG
AGTGCGTTTAAAAAGAAAACGCAAACCTTCGGAGAAGGGGTTGGAATTGCAGGCTGAGA
AGGTCAAGAAACCTGTGAAACCTTTAAAATCGTCTGGTAGATTGTCAAAAACTAATAAAC
CTGTGATTGATCCGGATTCTAGTAATGGTGTATATGGAGATGATGGTTCTGATTGATTTG
ATTAAATTGCAGTTCTCATTTCTTCAAAAAAAAAA

SEQ ID NO: 81
AAAAAACAGACGACCAGAGGAAGATTTCTCTTTCCTGATTCTAGGGTTTCCTTCCCCTTT
CGTCTTCTTCGTCTTCCGCTCCATCAATCGGCTCCGACACCCGCAGGCACCGAGGCAGC
TTCGTCGCCCCCGCGGCCGAAGAGGACGGGTCAAGGTTGCAAATTTTGCCCAATCCCG
CGCAGATTCCGCGCCCCTCGCCGGCGTCGCGTCGTGTCGGAGGAGGAAAATGGCACG
TCCGAGCAGGGCGGCCCGGAAATGGGGAGCGTGCTGATCGCCCGAGAGGGATCGGGC
GCGGCGGCGAGAGCTGGATTTGGGTTCCCTCGCAGGAGAATTGTGGATGCATGAGGCT
GCTCCGAGGATATAATGAGGGGTGTCTCGTGGCCCGAAGACGGCAATAATCCTTCGACC
TCGAGTTCTTCGCAGCGCAATCAGCAGCAGGCGCATGCGCCCCGAGCTGTTTCTGGGC
ACGCTGCAAGTCATCCAAGTGCTAGCAATATCTTTAAACTCCTAGTGCAAAGAGAGGTCT
CTCCACGGTCAAAACATTCATCAAAGAAATTGTGGAGAGAAGCTTCAAAGTGCCAGCCCT
ATCCATTCCAGCAAAGTTGTGAAGCAGTGAGAGATGTGAGACAGGGTCTCATATCATGG
GTGGAGTCAGCGTCACTGCGGCATTTGTCTGCCAAATATTGTCCACTTGTGCCTCCTCC
AAGATCAACAATTGCAGCTGCCTTTAGTCCTGATGGAAAAATACTTGCTTCTACACATGG
AGATCATACGGTGAAACTCATTGATTCTCAGACAGGAAGTTGCTTGAAGGTGTTAAGGG
GTCATCGGAGGACACCATGGGTGGTGAGATTCCATCCGCTGTATCCGGAGATCCTTGCA
AGTGGCAGTCTGGATCATGAAGTTCGCCTGTGGGATGCAAATACCGCAGAGTGCATAGG
ATCACGCAATTTTTACCGTCCTATTGCATCAATTGCATTCCATGCTCGAGGAGAGCTTCTT
GCTGTTGCATCTGGTCACAAGTTGTACATATGGCACTACAATAGGAGAGGAGAGACATC
ATCCCCAACTATCGTTCTGAGGACACAGCGATCTCTTAGAGCCGTGCATTTCCACCCACA
TGCAGCTCCATTTCTTTTAACAGCTGAGGTCAATGACCTTGACTCAGCAGATTCAGCTAT
GACTCTTGCAACTTCTCCTGGTTATTTGCACTACCCTCCTCCTACTGTATATTTTGCAGAT
GCTCATTCTCATGAAAGATCTAGGTTGGCGGATGAACTGCCTCTCATGCCTTTGCCATTA
TTGATGTGGCCTTCCTTTACTAGAGACGATGGAAGAGTACCCTTGCAGCGCATAGATGG
GGATGTTGGTCTTAATGGACAGCAAAGGGTAGATTCATCTTCTTCAGTGCGCCTTTGGAC
ATATTCAACCCCATCAGGGCAGTATGAGCTCCTTCTGTCTCCAGTTGAAAGTGGCAACTC
TCCTTCCATGCCTGAGGAAACGGGAAATAATGCTTTCTCAAGTGCAGTGGAAGCTGAAG
TAAGTCAATCTGCAATGGATACTGTGGAGGATATGGAAGTGCAACCTGAAGAGAGAAAT
ACTCAATTTTTCTCTTTTAGCGACCCCAGATTTTGGGAGCTGCCTTTATTGCATGGATGGT
TGGTTGGTCAGACCCAAGCTGGTCCACGAAGTGTACGTCAATCAAGTCCTGGAGATATT
GAAACTCAATCTGCTTTTGGTGAGGTTGCAAGTGTTTCACCAATCACATCTGGAGTGATG
CCAGTTAGCATGGACCCGTCACGGTTTGGTGGAAGATCTGGTTCTAGATATCGCTCTCC
TGGGTCTCGGGGGGTGCACGTGACTGGACCTAATAATGATGGACCACGAGATGAAAAC
GATCCTCAATCTGTTGTTAGTAAACTCAGGTCTGAACTTGCAGCCTCGCTGGCTGCAGCA
GCATCTACGGAGTTACCCTGCACTGTGAAGCTTAGAATATGGCCACACGACGTAAAAGA
TCCTTGTGCACAGCTTGATTTAGAAAGTTGCCGCTTAACTATTCCACATGCTGTTCTATGC
AGTGAAATGGGAGCCCATTTTTCTCCATGTGGGAGATTTTTAGCTGCCTGTGTGGCATGT
GTGCTGCCTCATTTGGAATCTGATCCCGGATTACATGGTCAAGTCAATCAGGATGTCACA
GGAGTGGCAACCTCACCTACAAGACACCCAATTTCTGCTCATCAAATCATGTATGAGCTA
AGGATATATTCCTTAGAGGAGGCAACATTTGGAATTGTGCTTGCTTCACGACCAGTAAGA
GCCGCTCATTGTTTAACCTCCATTCAGTTCTCTCCGACATCTGAACACTTGTTACTTGCCT
ACGGCCGTCGCCATAGTTCACTTCTTAAGAGTATTGTCATTGATGGAGAGAATACGGTGC
CTATTTACACCATTTTGGAGGTATACAGAGTTTCTGACATGGAACTTGTGAGAGTTCTTCC
AAGTGCAGAGGATGAAGTCAATGTTGCATGCTTTCACCCTTCAGTTGGAGGTGGCCTCA
TTTATGGAACGAAGGAAGGCAAGTTGAGGATTCTCCACTATGACAGCTCTCATGGCTTGA

ATCTAAAGTCATCTGGTTTTCTTGATGAAAATGTGCCAGAGGTGCAGACTTATGCTTTGG
AATGTTAGTAGCCAGCAACTGCAACATGTTTGTGATACTTACTGGTCAACCTCAAAGATC
TTGGGTCGAGGGCAGCTGTACCTATATCTTCACAACTGTAGACCCATCTGCGGCCAGGC
TCGTGGTTCCTCAATCAATTACTAGTTAAACGCAATTCTGTACCAACTACCATTCTTTGCT
GTTTGTGGGGATTTATGTTATCAACTGCAAGAGCATTAGGAAGTTCATCTAACTTGAGAT
AGATGGGTTAGTGCGAGTTGGAAATGTCGGGATTCGTGTTGTATATTTGTTTGTCCGCTT
TGATATTTACGGACTTTTGTAAAAAAAAAA

SEQ ID NO: 82
CGAAGTAAACCTCTCGGAACCTGAATCTGCTCTCAAAATCTCGTCGGTGCCGGAGAAAG
TTTCCCCGATCGATCGCCTCCTCCGCCGAGGAGCGCATGGATTCCGCCGTCGCGATCG
CGGCGCTGTCGCTCGTGGTCGGAGCCGCGATCGCCCTCCTTTTCTTCGGCAATTACTTC
CGCAAGCGAAGATCCGAGGTCGTGGCCATGGCCGAGGCGGACCTCCAGCCGCATCCC
AAGAACCCGTCGCGGCCTCCGCCGCAGCCCGCCGCCAAGAAGGTCCACGCCAAGTCC
CATGCTCACGGCGCCGATAAGGATAAAAACAAGAGGCATCATCCTTTGGATCTGAATACT
TTAAAAGGTCACGGGGATTCAGTCACAGGGCTATGTTTCGCTTCTGATGGACGAAGTTT
GGCAACAGCTTGTGCTGATGGTGTTGTGAGGGTGTTCAAGTTGGACGATGCCTCAAACA
AAAGTTTCAAATTTCTGAGGATTAACTTGCCTGCTGGAGGCCATCCAACTGCTGTTGCAT
TTGGTGATGGTGTATCATCAGTAATTGTGGCATCTCAACATTTGTCTGGTTGTTCTCTGTA
CATGTATGGAGAAGAGAAGCCTACAAATTTGGACAGCAATAAGCAGCAGACTAAGCTAC
CTATGCCTGAAATCAAGTGGGAGCATCACAAAGTTCATGAGCAGAAAGCAATATTGACCC
TTTCTGGGGCTGCTGCAAATTATGACAGCGGTGATGGGAGTACAATAATTGCATCTTGTT
CAGAAGGAACTGACATCATAATTTGGCATGCAAAAACTGGGAAGATTTTGGGGAATGTC
GATACAAATCAATTGAAAAACACTATGTCTGCTATCTCACCCAATGGGCGATTTATTGCTG
CTGCTGCTTTTACTGCTGATGTTAAGGTCTGGGAGATTGTTTACTCAAAAGATGGTTCTG
TGAAGGGAGTTACAAAAGTCATGCAGCTTAAGGGACACAAGAGTGCAGTGACTTGGTTA
TGCTTCACTCCAAACTCAGAACAAATAGTTACAGCATCAAAGGATGGCTCTATAAGAATTT
GGAACATCAACGTTCGATACCACCTTGATGAGGATACGAAGACTTTAAAGGTGTTTCCAA
TCCCATTGCAAGATTCAAGTGGTACTACTTTACACTATGAGCGCCTCAGTCTATCCCCTG
ATGGAAAGATACTGGCAGCAACCCATGGTTCAATGTTGCAGTGGCTGTGCATTGAAACT
GGAAAGGTTTTGGACACAGCTGAAAAAGCTCATGACGGTGACATCACTTGCATGTCCTG
GGCACCACAGAGTATTCCAACAGGCGATAAAAAAGTTAATGTTTTGGCGACGGCCAGTG
GTGATAAAAAAGTGAAGCTTTGGGCAGCCCCACCACTTCCCTCATAGATGAGTTAATACA
GGAGGCTGAAATTCAGCATCAGAACAATGTGAAGCTTGCTCCTACGAATACGCTTTGCC
ATTTGAGATGCATTATGGGGTTCTAAATTATATCTGAGCCACTGTCACTCTTCACAAGAAA
GCCTCAGAGTGTCCTATGGAACGTTCCAACACTGTCCATTCTGACAGCATCCGGACATA
CAGTTCATACTTCGACGTTGTTAGCCTTTCCTTAAAAAATTTTGCCGATAAACTTGGCCTA
ATGCAGAATCCAACAGGAAGAGTTTTCTGTATTCCTTCTGTTTATTTGTCTTCTCTTTGGA
GGGATATTGAGAAGGTTGTTGTTGAGATGGTGGCTGAACTGCGAGGCGATGACCCAGTA
TAATCTATCTGAGCCGGACTGATCTGGACGGCTCAATCAAGTGCTTTTCCTTTTAATTTG
CTTGCTTTTTGGGACAGGCTATTACACTTTCAAATTTGCTTGTGTGGCTGAAGCACCATAT
TCTACTAAGTTAGTCAGCCTCAGAGTGGGAGGACGAACCTATGCGGATGCGCATTGGCA
AATCTGAACAATCATTCTGTAGAACACTAGAGTCTATATGCTTGACTGTATCGGTTAATTA
ATTCAAGATGACCACGATATGTTGTTGCACATGTGATCGTCAAAACTGGTCGTTCTTCAA
AAAAAAAA

SEQ ID NO: 83
AGAAGAGGCCACCTAAAACCCTAGAAGCCAGAGACTGCAAACCTCATAAAACCTCCAAA
GCGCGCACCAAGAAACCCCCCGCAGAAGAAGAGCGCGAGCTTTCGCATTCGCATCAAT
GGAGGTGGAACCCAAGAAGGCGTCCAAGACCTTCCCGGTGAAGCCGAAGCTGAAGCCG
AAGCCGAGAACCCCTTCTGGCAAGACCCCAGAATCCAAGTACTGGTCCTCCTTCAAGAC
CACCCACCCCCTCGACAACCTCTCCTTCTCCGTCCCTTCCCTCGCCTTCTCCCCTTCCC
CTCCCCACCTCCTCGCCGCCGCCCACTCCGCCACCGTCTCCCTCTTCTCCCCCCACCG
CACCACCATCTCCTCCTTCTCCGACGTCGTCTCCTCCCTCTCCTTCCGCTCCGACGGCC
AACTCCTCGCCGCCTCCGACCTCTCCGGCCTCATCCAGGTCTTCGACGTCCGCTCCCG
CACCCCGCTCCGCCGCCTCCGCTCCCACGCCCGCCCTGTCAGGTTCGTCCGCTACCCG
GTCCTCGACAAGCTCCACCTGGTGTCCGGGGGCGATGACGCCCTCGTCAAGTACTGGG
ACGTGGCTGGGGAGAGCGTGGTGTCGGAGCTTAGAGGGCATAAGGACTACGTCCGGTG
CGGGGACTGTTCGCCCGCCGACGCGAATTGCTTCGTCACCGGGTCTTATGATCACGTG
GTGAAGCTCTGGGATGTGAGGGTGAGAGATGGGAACAGGGCGGCGACGGAGGTGAAT

CACGGGTCGCCGGTGCAGGATGTGATCTTCTTGCCGTCGGGGAGTTTGGTCGCCACTG
CGGGAGGGAATAGCGTGAAGATTTGGGACTTGATTGGAGGAGGGAGGATGGTTTATTC
GATGGAGAGTCACAACAAGACCGTGACCTCCATTTGTGTGGGTACGATGGGGGCGCAG
CAGAGCGGGGAAGAAGGTGTGCAGCTGAGGATCTTGAGCGTTGGTCTTGACGGGTATA
TGAAGGTGTTTGATTATTCTAGAATGAAGGTCACACATTCGATGAGGTTCCCGGCACCTC
TGCTGTCGATCGGGTTTTCGCCAGATAGCAACGTGAGAGCCATTGGGACTTCGAATGGT
ATTTTGTATGTGGGAAAGAGAAAGGCGAAGGAAAATGCGGAGGGTGGTGCCAATGGAAT
CTTAGGGTTGGGCAGTGTGGAGGAGCCGCGGAGGCGGGTCTTGAAGCCCTCGTTCTAT
AGGTACTTCCACAGAGGTCAGAGTGAGAAGCCATCCGAGGGAGATTATTTGGTTATGAG
GCCGAAGAAGGTGAAGTTGGCTGAGCATGATAAACTCTTGAAGAAGTTCCAGCATAAGA
ATGCTCTCATTTCAGTTTTAGGTGGAAATGATCCTGAAAAGGTGGTGGCTGTGATGGAG
GAATTGGTGGCTCGGAGGGCACTGCTAAAGTGCGTCCTGAACTTGGATGCGGATGAATT
GGGTCTGATTTTGACGTTCTTGCATAAGAACTCGACTGTGCCTAGATATTCAAGCTTGTT
GCTGGGGCTGGCGAAGAAGGTTATTGACTTGAGGCTCGAAGACATAAGAGCATCCGAT
GCCTTGAAGGGTCATATTAGGAACCTCAAGCGCTCAGTTGATGAGGAGATTCGGATACA
AGAGGGGTTGCAAGAGATTCAGGGTATGGTATCTCCTTTACTAAGGATTGCAGGCAGAA
GATAGCGATAGAGTTATACTGCATGTACTGAGGTAAATGTTTTGATTACTCCACCCAATT
GCGTTGGTTCTCTACTTTTTCCCCTTGAGGGATGAAATATGAGGAGATTCGGATACAAGA
GGGGTTGCAAGAGATTCAGGGTATGGTATCTCCTTTACTAAGGATTGCAGGCAGAAGAT
AGCGATAGAGTTATACTGCATGTACTGAGGTAAATGTTTTGATTACTCCACCCAATTGCG
TTGGTTCTCTACTTTTTCCCCTTGAGGGATGAAATTTTGCTGCAATGTATGAGTTTCACTA
AATTTATGGAACACTTATGTTTTTATACGGAGATGGTTGGAGTTGAAGGTCCATCTTCCG
GGTTTTATCTGTAACTGGATGCTCAAGTAAAACAATTTTTTTTCCCTTTAAAAAAAAAA

SEQ ID NO: 84
CGAGTTCAGCTCAAGCGCCATCTCCGATCGCCGGTCACCGCATCGATGCAGGGAGGAT
CGTCGGGCGTCGGCTATGGCCTCAAGTACCAGGCCAGGTGCATCTCCGACGTGAAGGC
CGACACCGACCACACCAGCTTCCTCACCGGAACCCTCAGCCTCAAAGAAGAGAACGAG
GTCCATTTGCTGAGGCTCTCGTCGGGCGGCACGGAGCTGATCTGCGAGGGCTTGTTCT
CGCACCCGAGCGAGATTTGGGACCTCTCTTCCTGTCCCTTCGATCAGCGCATCTTCTCT
ACTGTTTTCTCCACCGGTGAATCTTATGGAGCTGCTGTGTGGCAGATTCCTGAGTTATAT
GGACAGTTAAATTCTCCTCAGTTGGAAAAAATTGCCTCGCTTGATGCTCATTCACGCAAG
ATCAGTTGTGTTCTTTGGTGGCCGTCTGGAAGGCATGACAAGTTGGTTAGCATTGATGA
GGAAAACATCTTCTTATGGGGTTTAGATTGTTCGAAAAAGTCGGCTCAGGTCCAATCACA
GGAGTCTGCTGGCATGCTGCACAACCTCTCTGGTGGAGCGTGGGATCCACATGATGTAA
ATACTGTCGCTGCAACCTGTGAATCATCAATCCAATTTTGGGATCTTCGGACTATGAAGA
AAGCAAATTCACTAGAATCTGTCCATGCTCGTGACCTAGACTATGACATGCGAAAGAAGC
ACTTACTTGTTACCTCCGAGGATGAATCTGGTGTACGTGTATGGGATCTTAGAATGCCTA
AAGCTCCTATTCAAGAGTTTCCTGGTCACACACACTGGACTTGGGCTGTTAGGTGTAATC
CTGATTACGAAGGACTGATTCTGAGTGCAGGTACAGACTCAGCTGTAAATTTGTGGTGGT
CATCTACTGCGAGCAGTGATGAGCTGATATCTGAGAGGCTAATTGACTCGCCTACCAGA
AAGCTTGATCCATTGCTCCATTCATACAATGACTATGAAGACAGTGTTTATGGCCTGGCA
TGGAGTTCACGAGAACCTTGGATATTTGCATCGTTGTCATACGATGGGAGGGTGGTTGT
AGAATCAGTGAAGCCCTTCCTGTCCAGAAAATGACTTCAGGGACACTAAATTAGGTTCAT
TGCCCCTTTGCGGTCCTGGCCTTTTTACATTTGGAGATGCTACGAAGTATTTCTTCTTGC
TTTTTCGGTTGGTTGGGGACTTATAAATTCTCATTATAGATTTTTTTCTCCAATCATTGATT
TTGAACAAGGTATGGTTGAGAGATCATGTTTCTTCTTCCCAGCCCTGGAAAGCTGTTTTA
TCTTTGGATCATTCGAGTGAAACTTCCAACCGATTAAGGCAAGAATTGTTAGGAGGTGTA
TACTTTCTGTAACTGTATTCAATGAGCATACACCTGACGGAAAAAAAAAAAATCATCGAATC
ACTGTAAAACTTCCTTCACAAATTTTTAAAAAAAAAA

SEQ ID NO: 85
GGGAGAGAGAGAATCCAGGTCCATGGCGGAAGAGGAGGGAAGCGCGGAGCTGGAGCA
GCAGCTGGAGGAAGAGTTCGCGGTGTGGAAGAAGAACACCCCGATCCTCTACGACCTC
CTCATCTCCCACGCCCTCGAGTGGCCCTCCCTCACCGTCCACTGGGCCCCCCTCCTTCC
CCAGCCCTCCTCCTCCGCCGCCGCCGCCGCCGGCGACCCTTCCCTCGCCGCCCACCG
CCTCGTCCTCGGCACCCACACCTCCGACGGCGCCCCCAACTTCCTCATCCTCGCCGAC
GCCCTCCTCCCTTCTTCCGAATCCGATCATTGCGGCGACGACGCCGTGCTCCCCAAGGT
GGAGATATCGCAGAAGATTCGCGTGGATGGGGAAGTGAATAGGGCGCGTTTCATGCCA
CAAAATCACAATATTGTGGGCGCTAAGACAAATGGTTGCGAGGTTTATGTTTTTGATTGTT

CTAAGCAGGCTGCGAAGCAGCACGATGGTGGGTTTGATCCGGATTTGAGGTTGACGGG
TCATGATGGAGAAGGCTATGGATTGTCCTGGAGCCCCCTGAAAGAAAACTACCTTTTAAG
CGCGTCCCATGATAAAAAAATTTGTCTCTGGGATATATCTGCTGCCGCTCAGGATAAAGT
GCTCGGTGCAATGCATGTTTTCGAGGCTCACGAGGGTGCGGTCGGCGATGCATCTTGG
CACTCGAAGAATGACAACTTATTTGGGTCTGCTGGTGATGATTGTCAATTGATGATCTGG
GACTTGCGAACGAATAAAGCACAACAATGTGTCAAAGCACATGAGAAAGAGGTGAATTCT
GTATCTTTCAACTCATATAATGACTGGATTCTGGCAACCGCGTCTTCAGACACAACAGTT
GGACTATTTGACATGCGGAAGCTGACCACGCCATTGCATGTCTTTAGTAGCCATGAAGG
AGAGGTCTTGCAAGTAGAATGGGATCCTAATCACGAGGCTGTGTTAGCTTCTTCCTCAGA
GGACAGAAGGGTGATGGTCTGGGACCTCAACAGGATCGGAGATGAACAGCAAGAAGGA
GATGCAAGTGATGGCCCTGCCGAACTGCTCTTTTCTCATGGGGGTCATAAAGCTAAGAT
CTCAGATTTCTCATGGAACAAAAACGAACCATGGGTTATCTCAAGTGTGGCTGAAGACAA
CTCTGTTCAGGTTTGGCAAATGGCTGAGAGTATCTGTGGAGATGACGACGACATGCAGG
CCATGGAAGGATATATATGAGATTTGGTACAATTGCTTCATCGAAATTTCTCGCTCTGGC
TTTGTTTCGGGGCCAACAGAGAATTTTTGCTGAATTTACTAATTCTACAATGCTTCACGAA
GCATTTTCTTTCAGTCCTTTTAGCAGCTAAGATGGAAGATAAACTCACTGGCATGATCCG
TTTGGTGTAAATCAGCTTATTTGCGAAAAAAAAAA

SEQ ID NO: 86
AGAGAGGGAGACTTCCACCAAAAGGACAAACTAGAAGAATCACCTCCTCCAACCTCCGG
GACCTTACCTTCTCCTCCCTCCTCCTTGAGAGCACCATCGCTATTATCGAGGCCATTTTT
TTCGTTTGCCTCTTTTCCAATTCGTCTTCTCTCTCTCTCTCTCTCTCTCTCTGTGTTCTTTT
TTTATATGTTTAAATCTCGCCTTCATTTTTATCCCACTGTAATGGATTCCTTGCTTCCTCAT
GCGGGACTTCGGTAGACAGAGGAGGGAGAGGGAGAGAGGGAGAGGGAGCGGGGTGG
AATTTTCAGGAGTGGAAATCCCCAGGGCCGTTTCGATTTCTGGGTTTCGACCCGAATCC
CATTTTTGATTTCCAGGTTCGTATGGGTTGAAAGCTAAAAGGGTCATGGGGAATTACGGT
GAAGAGGACGAGGATCAGTATTTTGATGCCCTGGAGGAGACGGCCTCTGTTTCAGATCG
GGGTTCAAACAGCTCTGATTGTTGTAGTTCAGGTTCCGGGCTTGATGAGAACGTTTTGGA
TAGTTTAGGGTTCGAGTTTTGGACTAAATTTCCCGAGAGCGTGAGGGCGCGCCGCAATC
GGTTCTTGATGTTGACGGGGTTGGGAATAGAGGCAAATTCAGTGGACAAGGAGGATGCA
TTCCCGCCATCTTGTAACGAAATCGAGGTGTATACGTGTAAAGTAACAAGAGATGATGGG
GCGGTTCAAAGATCTTTGGATTCATATAATTGTATTTCACTGTTGCAGTCATCCACATCCA
TTCGGTCGAATCAAGAAGTTGAATCATTGCGTGGGGACTCGCTCCTCAGCAGTTTCAGG
GGTAGAAGTAAGGAGTCTGATGATTTGACGGAATTGTGTGGGATGGGTTGTCCAGAGTC
AAAGAGAAATGCGGTGAGTGAATTTGGTTCAGTAAGTCAGGGAAGCATTGAAGAACTGA
GGAGAATAGTTGCATCTTCTCCTTTGGTTCATCCGCTGTTGCATAGGAAATTAGAGTATG
AAAGGGAATTGATTGAAACCAAACAAAAAATGGGAGCTGGTTGGTTGAGGAAATTTGGCT
CTGCCACATGCATTAGTGGAAGACAGGGTGACACATGGTCAGACCCTGATGATCTTGAA
ATAACGGCAGGGATGAAAATGCGGAGAGTTCGTGCACATTCATCTAAGAAGAAATACAA
GGAATTGTCTTCCCTTTATGCTGCACAGGAATTTCTGGCACATGAGGGGTCAATTTCGAC
AATGAAGTTCAGTATGGATGGGCAATACTTGGCAAGTGCAGGTGAAGACACAGTAGTGA
GAGTTTGGAAGGTGACTGAGGAGGATAGATCAGAGAGGGTCAACGTTACTGTGGACCCT
TCCTGTTTATATTTTGCTCTGAATGAATCCACGCAGTTGGCGTCGCTCAATACGAATAAA
GAGCACATTGGTAAAGCAAAAACATTTCAGAGATCATCAGACTCATCGTGTGTCATTTTA
CCCTTAAAGGTCTTCCAGATAACAGAGAAGCCTTGGCACGAGTTCAAGGGACATAATGG
CGAGGTCTTAGACCTCTCCTGGTCTAGTAAAGGGTATTTGCTATCCTCCTCTACTGATAA
GACCGTGCGGTTGTGGAGGGTTGGATGTGATAGATGCCAGAGAGTATATTCTCACAATG
ATTATGTTACTTGTATAAGTTTCAACCCTGTTAATGAGAACTTCTTCATCAGCGGGTCCAT
AGATGGAAAAGTACGCATCTGGAATGTGTTTGGTGGTCAGGTTGTTGCTTACATCGATTG
TAGGGAGATTGTCTCTGCAGTGTGCTATCGATCCGATGGAAAGGGAGCCATCGTGGGTA
CCATGACAGGCAATTGCCTTTTTTATAGTATTAAAGATAATCATTTGCAAATGGATGCACA
AGTATACCTACATGGTAAGAAGAAGTCTCCTGGCAAGAGAATAACTGGCTTTCAGTTTCC
TCCCAATGATCCTGGCAAACTAATGATCACATCTGCTGATTCTGTAATTCGAGTATTAAGC
GGGCTAGATGTAGTCTGCAAACTCAAAGGCCCTCGAAATTCAGGGGGACCAATGATTGC
CACTTTCACTTCAGACGGGAAACACGTGATCTCAGCAAGTGAAGATTCAAATGTCTACAT
CTGGAACTACGCCGGTCAGGACAAGACTTCATCCCGGGTGAAGAAGATATGGTCTTGCG
AGAGTTTCTGGTCCAGCAATGCCTCTGTTGCTTTACCCTGGTGCGGCATAAGAACTGTG
CCTGAAGCACTTGCACCTCCTTCACGAAGTGAAGAGAGAAGAGCGAGTTGTGCAGAAAA
CGGGGAGAATCATCATATGCTGGAGGAATATTTCCAGAAAATGCCCCCATATTCCCCGG
ACTGTTTCTCTCTCAGCCGCGGATTCTTCCTGGAGCTCTTGCCGAAGGGATCCGCGACT

TGGCCGGAGGAGAAATTATCTGATACTAGCCCACCTACGGTCTCATCTCAAGCAATTTCT
AAACTGGAGTACAAGTTCCTGAAGAGCGCGTGCCACAGCGTGTTGAGTTCTGCTCACAT
GTGGGGCCTGGTGATCGTGACTGCAGGTTGGGACGGAAGGATCAGGACATATCACAAC
TATGGATTACCTGTACGTTCTTGAAATCCCGATTCACTAAAATTATGAAGGCGGTGAGTT
CAGTCGTCCGCCTTTTCTTAGTTTTACCCTACGAATACATGGTGATGCTCACGCTCAAGG
GGGGTTTGCCTCAAGTGTAAAAGGATGCCCCTAATAGATTATATGCCAAGTGTAGTATAT
ATAATAGTGCTTTTGTCACACACCAAAAAAAAAAAA

SEQ ID NO: 87
GAAGAAGTAAAGCCTTGACGCTGGTTTTTGGTCAGCTCCGTTTCCCTCACTCATCAACTC
TCTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCCGTGCGCCATGGACATCGACTTC
AAGGAGTACCGACTTCGCTGCGAGCTGCGCGGCCACGAGGACGATGTCCGGGGCGTAT
GCGTGTGCGGGGACGGCAGCATCGGGACCTCGTCGCGGGATCGGACGGTGAGGCTGT
GGGCTCCGAGCGCCGGCGAGAGGCGCAAGTACGAGGTGGCGAGGGTGCTGTTAGGGC
ACAAGAGCTTCGTGGGTCCCCTGGCGTGGGTTCCGCCCAGCGAGGAGCTTCCGGAGG
GCGGGATCGTGTCCGGCGGGATGGACACTCTCGTGATGGCTTGGGATTTGAGGAATGG
AGAGGCGCAGACGTTGAAGGGCCATCAGTTGCAGGTCACCGGCATCGTGTTGGACGGC
GGCGACATTGTTTCTGCCTCTGTTGATTGTACCTTAATAAGATGGAAGAATGGCCAGCTT
ACGGAGCACTGGGAGGCTCATAAGGCCCCCATACAAGCAGTCATAAGGTTGCCCTCCG
GGGAGCTTGTTACAGGTTCAAGTGACACAACCTTAAAACTCTGGAGAGGAAAGACGTGC
ACTCAGACTTTTGTTGGGCATACAGATACCGTTCGAGGCTTGGCAGTGATGCCTGATCT
GGGCATTCTATCTGCATCACATGATGGGTCTATCAGGTTATGGGCAGTGAGCGGTGAAT
GTCTAATGGAGATGGTTGATCATACTTCTATAGTCTATTCTGTTGATTCACATGCTTCGGG
CCTCATAGTCAGTGGTAGTGAAGATCGTTTTGCCAAAATATGGAAAGATGGAGTTTGCTT
TCAGAGCATTGAGCATCCTGGTTGTGTGTGGGACGTCAAATTTTTGGAAGATGGAGACA
TCGTGACAGCATGTTCTGATGGGACCATACGTATATGGACAAATCAAGAAGATAGGATG
GCTAACTCAACTGAGCTCGAGTTATTTGATTTAGAACTATCTTCTTACAAGAGAAGCAGG
AAAAGGGTTGGAGGACTGAAATTGGAAGAGTTGCCCGGGTTAGAGGCTTTACAAGTGCC
AGGAACTAGTGATGGCCAGACAAAAGTAATCAGAGAAGGAGACAATGGGGTGGCATATG
CATGGAATTCAACTGAACTAAAATGGGATAAGATTGGTGAAGTAGTTGATGGACCAGAAG
ATAGCATGAACCGCCCAGCCTTAGATGGTGTTCAATATGATTATGTATTTGATGTTGATAT
TGGAGATGGTGAGCCTACACGTAAGTTGCCCTACAATCGATCAGATAATCCATACGACA
CTGCCGATAAGTGGCTTCTCAAGGAGAACCTTCCTCTTTCCTATCGCCAACAGATAGTGG
AGTTTATACTTGCAAATTCTGGGCAGAGGGACTTCAATCTTGATCCATCATTTCGTGATC
CCTATACTGGTTCCAGTGCATATGTTCCCGGCGCACCTTCACAGCTGGCAGCTAAACAA
GCGAGACCTACTTTCAAGCATATACCCAAGAAAGGAATGCTGGTCTTTGATGCGGCTCA
GTTTGATGGAATTCTTAAAAAGATCAATGAGTTTAATAATACTTTGCTCTCTAATCAGGAA
AAGAAGAACTTATCATTGACAGATATAGAGATTTCCAGATTGGGAGCAGTTGTCAAAATTT
TAAAGGACACGTCACATTATCATTCTAGCAAATTTGCCGATGCTGACTTTGATTTGATGTT
GAAGTTGCTAGAATCATGGCCATATGAAATGATGTTTCCTGTTATTGATATTTTCAGGATG
GTAATCCTGCATCCAGATGGGGCAGATGGACTTCTGAGGCATCAAGAGGACAAGAAGGA
TGTTCTTATGGAATCAATCAAGAGAGCTACTGGAAATCCTTCAGTTCCTGCAAATTTCTTA
ACTAGCATCCGGGCTGTGACTAATCTATTCAAGAATTCAGCATACTACAGCTGGTTGCAG
AAGCATCGTAGTGAGATGCTTGATGCGTTTTCGAGCTGCTCTTCTTCTTCAAACAAGAAT
CTACAGTTGTCTTATGCTACTCTATTACTCAATTATGCTGTGCTATTGATTGAGAAGAAGG
ATGAAGAAGGTCAATCTCAAGTTCTTTCAGCAGCACTTGAGCTTGCAGAAAATGAATCTT
TAGAAGTTGATGCCAGATATCGAGCTTTAGTAGCTATTGGATCATTGATGCTTGATGGTC
TTGTTAAAAGAATCGCTTTGGATTTTGATGTTGAGCACATTGCCAAAGCTGCAAGGACTT
CTAAAGAAGCCAAGATTGCTGAAGTAGGAGCAGATATCGAACTCCTTATAAAGCAGAGTT
GAATGACTGTTATTTGAGCTGTGCAAATATTCAACCGCTCGTGATTCCGTCGTGGGACTG
CACACCTGTTATACAGAAGTCTTTCAGAACAGCATTCATGCTACCGTGGAGGAACAATGG
GTCACTTGGATTGAGAGATGGCCACATTTGGAAGGCATTGATTGGCCTTAAAGTATCATT
GTTGGGAATCGTATACTGTGTAGGAGATTACATGATAACAGATTGAGATTGCACAGGTCT
ACGATTTTCTTTCTGCCAGCAAAAGCTGAGACTTTTGTATGGACTAAAAACTCGCCATAA
GAAGCATGGAAACGCGGGGGGGCGGAAAACCAGTGCCGTGTGCCCTGCAATGCATACG
TCTTTCACTGTAAAAAAAAAA

SEQ ID NO: 88
CAAAAGTTGCAGCTTCCTGCAAATGGAATGGCCGAATCAGCTGCGAGGAAGCAAGCTCT
CGCGGCCACCTTCGATGCATCAGACCGAGCCGTAGTCGGAGATCCACTCGCCTCCGAC

ACCTCCCGGCACGCCTCGATTCTCCTCGCGTGCCCCCGATCCCCTCTTCAATCGTGAAT
CGGGAGCTCGGGAACTTGCTCTTCCTTCGTCTGTTGCCGCGTGATTCGGGGTTGCGTTT
CAGCGGAATGGAGTTCACGGAAGCGTACAAGCAGTCGGGCCCTTGCTGCTTCTCTCCC
AACGCGCGCTTCATTGCCGTCGCTGTCGATTATCGCCTCGTCATACGCGACACATTGTC
GCTCAAGGTTGTGCAGTTGTTTTCATGCTTGGATAAAATAAGCTACATAGAATGGGCCCT
TGATTCTGAATACATACTTTGTGGTCTCTATAAAAGACCAATGATACAGGCATGGTCATTA
ATTCAACCTGAATGGACATGTAAAATAGATGAAGGTCCAGCTGGCATTGCTTATGCTAGG
TGGAGCCCGGATAGTCGGCACATACTGACAACATCTGATTTCCAACTCCGCTTGACAGT
TTGGTCATTGGTCAACACAGCATGTGTACACGTGCAGTGGCCAAAGCATGCTTCCAAGG
GGGTGTCTTTCACTCGAGATGGAAAGTTCGCTGCAATATGCACAAGACATGATTGCAAG
GACTACATCAATCTGCTTTCTTGTCATAATTGGGAGATAATGGGTGTTTTTGCTGTTGATA
CTTTAGACTTAGCTGATATCCAGTGGTCACCGGATGATAGTGCTATAGTGATATGGGATT
CACCTCTTGAATACAAGGTTTTGGTATACTCACCAGATGGTAGGTGTCTGTTTAAGTATC
AAGCCTATGAAAGTGGATTGGGAGTGAAAAGCGTTTCTTGGTCTCCTTGTGGACAATTTC
TAGCCGTTGGTAGCTATGACCAGATGTTAAGGGTCTTGAGCCACCTTACTTGGAAAACTT
TTGCAGAATTCACGCATCTATCTAATGTTCGTGCTCCATGTTGTGCTGCCATCTTTAAGG
AGGTGGATGAGCCCTTGCAAATTGACATGTCTGAATTGTCTTTAAGTGATGATTATATGC
AAGGCAATTCTGGAGATGCTCCAGAGGGACATTACAGAGTCAGATATGATGTTACAGAA
GTTCCAATTACTTTGCCTTGCCAGAAGCCTCCAGCGGACAGACCGAATCCCAAACAAGG
AATTGGTCTGATGTCATGGAGCAATGACAGCCAGTATATATGCACTCGCAATGATAGTAT
GCCGACCATTCTTTGGATCTGGGACATGCGCCATCTTGAACTTGCTGCCATCTTGGTTCA
AAAGGATCCTATCCGAGCTGCAGTTTGGGACCCTACAGGCACTCGCCTAGTCCTTTGCA
CAGGAAGCTCGCACCTGTACATGTGGACTCCTTCAGGTGCCTACTGTGTCAGCGTCCCC
CTATCTCAGTTTAATATAACAGATTTGAAGTGGAATTCGGATGGGAGCTGTCTTTTGCTCA
AGGACAAGGAGTCATTTTGCTGTGCTGCAGCACCACTGCCACCAGACGAATCTAGTGAT
TATAGCTCAGATGATTGAGGCATCGATGTGTACTCTCCAGAGTGACATGGACTGCAAACT
GCAGTTGTTCGTAGGACTGTATATTTAGGCATGTTAGATCCACCGGATAGAGAATAGAGA
TTGACTTTTTGGTGTTGTATAATTCTTTACAGTTGGGATGATTAGCTTCTGGGAGGTAAAT
TGATGACAATCGATTGGCCGCTCCTCATCTTTTCCTTTTGAACTGTAGAAAGCCGGACTT
TTGGGTAATGGTCAATCTGTTTGCCTGTGTTCGATAGAAGGGCAAAAATCTGTACGCTGA
TATGAATCGCGGTGAGAGCC

SEQ ID NO: 89
GTCACCTCGAAGTGCTGCGCTCAAACTAAACTCCCCCAGGCCACCGTCGCCGCGTCTC
CTCTCGTTCGCTCGCGCGGGGGGCGCCCACCAGAGCAGCCGCGGAAGAACGCGATCG
ATCGACGGCGATGGCAACGATCGCGGCGCTTGACGACGACATGGTCCGCAGCATGTCG
ATCGGAGCCGTTTTCTCCGACTTCGTTGGGAAATTAAATTCGCTTGATTTTCACCGTAAG
GATGATATTTTGGTCACGGCCGGCGAGGATGATTCAGTGAGATTGTATGATATTGCAAAT
GCCAGGTTGCTCAAGACCACCTTTCACAAGAAACATGGCACTGATCGTGTGTGCTTTACT
CATCATCCAAACTCCCTCATATGCTCTTCAACCAAAAACTTAGATACTGGAGAGTCCTTAA
GATATATATCAATGTATGATAACCGAAGCCTCCGCTACTTTAAGGGACACAAACAGAGAG
TTGTTTCCCTGTGCATGTCTCCAATCAATGATAGCTTCATGTCAGGCTCTCTTGACCACA
GCGTGAGAATGTGGGATCTCCGTGTGAATGCTTGCCAGGGAATTCTGCGTCTACGTGGC
AGACCTACTGTTGCATATGACCAGCAGGGTCTTGTCTTTGCTGTAGCTATGGAAGGGGG
TGCTATCAAATTGTTTGATTCACGCTCTTACGACAAGGGCCCCTTTGATGCCTTTTTAGTT
GGTGGAGATACATCTGAGGTCTGTGATATCAAGTTCAGCAATGATGGAAAATCAGTGCTA
TTATCAACCACAAACAACAACATCTATGTTCTCGATGCATATGCAGGAGATAAGCAATGT
GGATTCAATTTAGAACCATCTCCAAGTACACCGATAGAGGCAAGTTTTTCACCAGATGGC
CAGTACGTTGTATCAGGCTCAGGAGATGGAACTTTGCATGCATGGAATATCAGCAGGCG
AAATGAGGTAGCATGTTGGAATAGCCATATTGGCGTCGCATCATGCTTGAAATGGGCTC
CTCGTCGGGCCATGTTCGTAGCGGCATCTACTGTCCTCACTTTCTGGATTCCAAATTCCG
AGCCCGAGCTTGCTTCTGCCAAAGGCGAGGCCGGAGTTCCACCAGAACAAGTATAACTG
TGAAGCCTCCATTTACAGCTCTAGACCCTGCCATGAGCTTCTTTAAGCGAGCAATAGAGT
TAATAATATGTTGTAAATCTTCTCATGTGCCTGGCGTAAATTTTGCAGTTATTACTAGACC
AAGATAGTTTCATGCTCCACAGAACAAATTGACACCCCCACCTGTACTTTAAAGTCACAA
TTTTCAAAAGCTTCTGTTTTCTTAAAAAAAAAA

SEQ ID NO: 90
GCTTCGTCGAGCTGTCCAGACGCGTCCCAAAATCAGTTGCTCTCGATTCTCCTCCCGCG
ATCTCCTTCCCTTCCTTCCCTTCCCCCGTCTGTACATCTCGAAGCCCTAACCTCCGAAGC

TCCTTCCGATGCACCGCCGCCACCGCCGCCGTCATCACCGCCGCGACTGCTCCCGCCG
ATGAGCCGAGCCGGGCGATGACGGCTCCTGCCTAGGGGAAGAGGGAAACACTCTCGC
GCGCGCGCCCCCGACGACGACGATGTCCGTGGCGGAGCTGAAGGAGCGCCACAGGGC
CGCGACCGAGACCGTCAATTCTCTCCGGGAGCGCCTCAAGCAGAAGCGAGTGCAGCTG
CTCGACACCGACGTGGCCGGGTACGCGAGGACTCAGGGGAAAACTCCGGTCACTTTTG
GGGCGACGGATCTGGTGTGCTGCAGGACCTTGCAGGGCCACACTGGCAAGGTGTACTC
ACTAGATTGGACTCCTGAAAGGAATCGCATTGTCAGCGTGTCTCAAGATGGGAGATTTAT
AGTGTGGAATGCTCTGACCAGTCAGAAAACTCATGCCATAAGGCTTCCTTGTGCATGGG
TAATGACATGTGCTTTTGCACCGAATGGTCAGTCTGTTGCCTGTGGCGGTCTTGACAGC
GTATGCTCTATCTTCAACCTCAATTCCCCCGTTGACAGGGATGGGAACCTACCGGTATCA
AGAATGCTCAGTGGGCATAAAGGTTATGTGTCATCATGTCAGTATGTACCAGACGGAGAT
GCTCACCTGATTACAGGATCTGGTGACCAAACATGTGTCTTGTGGGATATTACCACGGG
TCTGAGAACTTCTGTCTTTGGAGGAGAATTTCAATCAGGGCACACTGCTGATGTTTTAAG
TGTGTCAATCAATGGATCGAGCCCAAGAATATTTGTTTCTGGTTCATGTGATTCAACTGCT
AGGATGTGGGATACTCGTGTTGCAAGTCGAGCCGTTCATACATATCATGGACATGAAGG
CGATGTAAATGCCGTAAAGTTCTTTCCAGATGGAAATAGATTTGGAACTGGGTCTGATGA
CGGTACTTGCAGGCTCTTTGACATCAGAACAGGGCATGAACTTCAAGTGTACTATCAACA
GCGTGGCATCGATGAGATCCCACATGTCACCTCCATTGCATTTTCCATTTCGGGGAGGC
TGCTAATTGCTGGATACTCAAATGGTGACTGCTTTGTGTGGGATACATTATTGGCCCAGG
TTGTGTTGAACTTGGGATCGTTGCAGAACTCACACGAAGGTCGGATCAGCTGTTTGGGT
GTGTCTGCTGATGGAAGTGCCTTGTGTACTGGCAGTTGGGATACAAACCTAAAGATTTG
GGCTTTTGGAGGGATTCGGAGGGTGACTTAATTTCTGCAGGAGCCGAACCCTTTGATTT
CTTCGAGGTCTCTCATTTCATGTTACTGCGTCCGTTTCGTTACATGCCCATAATTTGTGGA
AAAATAGATTGAGAAGCCCTATCCTGTAATTGCATCGCTGTCATTCTCATGATGAAAAAG
AAAATGACATGGATTCGATCAATCGCCACATGACAACTAAAACAAGCGGTTCACGTGATT
GTAATTTTATCCTCATTGCTTCTGTGTGCTTGCTTTGTTCCTAACGTCATTCTCATGAACG
TCTTGTAAAAAAAAAA

SEQ ID NO: 91
GCGGCGAGAGCATCGAGCACTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCCT
CTTCTTCATCACTTGTTCGACTCGGATTGGTCGATCAGCCCTCCCGTCGCGCCGCCGAC
CGTCCCTCGCTCCCCTCCGTTCGCGCGCTCCACCGGCTGATCGGGTCGTCCCGCGGG
GATGAAGAAGAGGCCGAGAGGAGCAAGCCTCGATCAGGCGGTCGTCGATATACGGCGG
CGAGAGGTCGGCGGCCTCTCCGGTCTGAGCTTCGCTCGCCGCCTCGCCGCTTCCGAG
GGTCTTGTGTTGAGGCTCGATATCTATAATAAACTGAAAGGGCATAGGGGATGTGTCAAC
ACTGTCGGTTTCAATCTTGATGGTGACATAGTAATATCGGGCTCGGACGACAGGCATGT
GAAGCTCTGGGATTGGCAAACCGGGAAAGTCAAGTTGTCATTCGATTCGGGTCATCTAA
GCAATGTCTTCCAAGCCAAGATCATGCCATACACTGATGATAGGAGCATTGTCACTTGTG
CAGCTGATGGACAGGCAAGGCATGCTCAGATTCTGGAGGGGGGACAAGTTCAGACAAT
GTTGTTAGCCAAGCATCGGGGAAGGGCTCATAAGTTGGCTATAGATCCTGGAAGTCCGC
ATATAGTTTACACATGTGGTGAAGATGGATTAGTTCAACGTCTTGATCTCAGGAGTAATA
CTGCCAGAGAACTTTTCACTTGCCGAGAAGTATATGGGACTCATGTGGAAGTTGTTCATT
TGAATGCAATTGCAATCGATCCAAGGAACCCAAATCTATTTGTGATTGGTGGGTCAGATG
AGTATGCTCGGGTGTATGATATTCGCAATTACAAGTGGAATGGATCGCATAATTTTGGTC
GATCTGCTAACTACTTTTGTCCTTCTCATCTCATTGGTGAGGCGCATGTGGGAATAACAG
GCTTAGCCTTCTCTGGTCAGAGTGAACTCCTCGTCTCTTACAATGATGAATCCATCTATC
TCTTCACCCAGGAGATGGGACTGGGTCCTGATCCGCTCTCTGCCTCCACCAAGTCCGTG
GACAGTAATTCCAGTGAGGTGACATCTCCCACAGCTGTGAATGTTGATGACAATGTTACC
CCTCAAGTCTACAAAGGGCACAGGAACTGTGAGACAGTGAAGGGTGTTGGCTTTTTTGG
GCCTAAATGTGAATATGTGGTGAGCGGGTCTGACTGTGGCCGCATATTCATTTGGAAGA
AAAAAGGGGGACAGCTTATCCGTGTCATGGCAGCTGATAAACACGTTGTAAACTGTATTG
AGCCTCATCCTCATATCCCTGCTCTGGCTAGTAGCGGAATAGAAAATGACATCAAAATCT
GGACTCCCAAGGCTATTGAAAGAGCTACTCTACCAATGAATGTCGAACAGCTAAAGCCA
AAGGCTAGGGGGTGGATGAACAGAATATCGTCGCCCCGGCAACTGTTGTTGCAGCTATA
TTCCCTGGAAAGGTGGCCGGAGCACGGTGGAGAGACTTCATCTGGCTTGGCTGCTGGG
CAGGAGGAACTCACGGAACTCTTTTTCGCACTGAGCGCTAACGGCAACGGGTCTCCGG
ATGGTGGTGGTGATCCTTCTGGTCCACTCCTTTAGGCGACTGATGTGGTGCGTTCAGAA
GACGACTAGCAACTGTGCATAATAATGTAGGGCAAAACCGTTTTGCTTCCCCATCCACAC
TCTTTTTTTCTGTTTCTTCTTCGGCTTTCTTCGGCTTCTCTTCCTCCTCTTCATGGATGTAT

ATGTTCATCCACATAACTTGTTCATGCTTGTTTATATTCTTTTATTTATTTTTCTCCTCCTTT
TCATGGATTAAAAAAAAAA

SEQ ID NO: 92
GAAGACTGTCTTCCTCAAGAAAAAGCAAAGCGGGGAATCAAGAAAGCAAGGCCTTCCCG
TTCTTGGCCAGCACTGTCGAAATCTCACGCGCGAATTTCTCGCGAATCCGTCGTCCTTC
CTTCGTCCCTTCCGAATCTGTGAATCCGATCCCCATCTTTCGTCTTGTCCGGAAGAGCGT
GAGTTTGTTGTTCAGAGCTTCCTTCGGGCTGGAGGAAAGTCTCGGGTTCCGGCTGCTTA
ATCTTGGGGGTTCTGGGTTCTGGGTTTTGGCTCCCGAGTGGCTGGGAGGAAGATGTCC
AAGCGGGGCTACAAGCTGCAGGAATTTGTGGCCCATTCGTCCAATGTCAACTGTCTCAG
TATTGGAAAGAAGGCATGCCGGCTTTTTCTCACTGGTGGAGATGACTGCAAGGTCAATC
TATGGGCCATCGGCAAGCCGAATTCCTTAATGAGCCTTTGTGGTCATACAAATGCTGTAG
AGTCCGTAGCCTTTGATTCTGCAGAAGTGTTGGTGCTTGCTGGAGCTTCTTCTGGTGTCA
TTAAGCTGTGGGACGTGGAAGAAGCAAAGATGGTTCGTGGTCTTACTGGACACAGATCC
AATTGCACCGCTATGGAATTCCATCCATTTGGAGAGTTCTTTGCATCTGGTTCCACTGAC
ACAAATCTGAAGATATGGGATATCAGGAAGAAGGGATGTATACACACGTACAAGGGTCA
TACTCGAGGCATTAGCACCATCAGATTCTCTCCTGATGGTCGCTGGGTTGTTTCAGGGG
GAAATGATAATGTTGTGAAGGTGTGGGATCTAACTGCTGGAAAGCTTTTGCATGATTTTA
AGTTCCATGAAAATCATATCCGATCTATAGATTTCCATCCCTTGGAGTTCCTACTTGCTAC
AGGTTCGGCAGATAGAACGGTTAAATTCTGGGACTTGGAAACGTTTGAACTAATCGGATC
TTCCAGACCTGAGGCCGCAGGAGTACGTGCAATTGCCTTCCATCCTGATGGGAGGACCT
TGTTCTGTGGTTTGGAGGATAGCTTAAAGGTTTACTCATGGGAGCCTGTAATCTGCCATG
ATGGTGTTGACATGGGATGGTCAACCCTTGCTGATCTTTGTATTCATGATGGAAAACTCT
TGGGTTGCTCATATTACCAAAGTTCTGTTGGTGTTTGGGTAGCGGATGCGTCGCTTATTG
AACCGTATGGAACTAATGTAAAGCCTCAGCAGAAGGATAGTGGGGATGATGAAATTGAA
CACCAAGAAAGTCGTCCCTCGGCTAAAGTTGGGACCACCATAAGATCAACTTCAATCATG
CGCTGCGCCTCTCCGGATTATGAAACCAAAGACATAAAGAATATATATGTGGATACCGCT
AGTGGTAATCCTGTTTCTTCACAGCGTGTTGGTACCACAAACTTTGCAAAAGTGACTCAA
CCGCTGGATTTTAACGACACTCCTAATTTGACACTACGGAGGCAGGGTTTGGTAACAGA
AACGCCAGATGGATTAAGTGGGCATGTTCCTAGTAAATCTATTACTCAACCGAAAGTTGT
TAGTAGGGACAGTCCTGATGGAAAAGACTCATCTCGAAGAGAATCTATTACTTTTTCAAG
GACTAAACCGGGCATGTTGCTCAGGCCTGCTCATTCGAGGAGGCCGTCGAGCACCAAA
TATGATGTCGACAGGTTATCAGCATGTGCTGAAATTGGAGTGCTTAGCAGTGCAAAAAGT
GGTTCTGAAAGTCTTGTAGATTCATTTTTGAACATCAAAGTTGCACCTGAAGATGGAGCA
AGAAATGGCTGTGAAGACAATCATTCGAGTGTCAAGAATGTCTCTGTGGAATCTGAGAAG
GTGCTACCACTGCAGACACCTAAAACAGAAAAGTGTGACCAAACTGTCGGTTTCAAAGAA
GAAATCAACTCTGTCAAGTTTGTCAATGGAGTTGCAGTTGTGCCAGGCAGGACACGCAC
TTTAGTTGAGAAGTTTGAAAAAAGAGAAAAATTGAACAGCACTGAAGATCAAACAATCAAT
ACACCTGAAAATCCCACGTTAGATAAAACCCCTCCTCCCTCTCTTGCAGAGAACGAGGAA
AAGAGTGACAGATTAAACATCGTTGAACGAAAGGCAACAAGAATGTCTTCTCATATGGTA
ACCGCTGAGGATAGAACACCTGTTACCCTTGTTGGCAGTCCTGAAGATCAGTCAACTGT
AATGGCTCCTCAAAGAGAATTACCAGCTGATGAATCAAGTAAAACTCCTCCATTGCCGGT
TGAGGATTTGGAGATTCATCATGGATCAAATGTCAGTGAAGATAAGGCGACTATTTTGTC
TTCTCAGACAGTATCAGAGGAGGATAGCAAACGCTCTACCCTTATTCGCAACTTCCGTAG
AAGAGACAGATTTAAAAGTACTGAAGGTCGATCCCCTGTAATGGCTACTCAAAGAAAACT
ACCAACAGATGAATCAGGCAAAACTTCTTCTTTGCCAATGGAGGATTTGGAAATCAAAGG
TGGATTGAATGTGAGTGAAGATAAGGCAACTAGTTTCTCTTCTCGTGCACCACCGAGGG
AGGATAGAGCACACTCCGCTCTTGTTCGCAATGTTCGAAAAAGAGACAAATTTAAAAGTA
CTAATGATACAATTACTGTAATGGTTCATCAAAGAGGACTATCGACAGATGAAGCATCAA
CTGTCTCAGTTGAGAGGGTTGAAAGGAGACAATTATCTAACAATGTGGAGAACCCGCTA
AATAATTTGCCCCCTCACTCAGTACCTCCAACGACCACGAGAGGGGAACCTCAATATGT
GGGGAGTGAGTCAGACTCTGTTAACCATGAGGATGTCACTGAACTTCTACTGGGAAATC
ATGAAGTTTTCTTAAGTACTCTCCGTTCTCGCTTGACAAAACTACAGGTAGTTTAGCTCTT
TTTGAATGTTTACGTGGTTGAGTAAGGCTAATTTGTATATTGCATGTCCCTGTATACTTTC
GTTTTGAATGTTGGCCTCCTTACACTTTGGCTATTCCTCTAATAACTAGTCTTTTATTCCC
CACCACAATCGTACAACTTGCCAGGTAGGGTTAGCTAAGGGTTGCAATTATGTTGGCAC
CATTTCTAATACTTGAAAGACATAGTATTAGCAGAAAGTATTGCAGTTTCAGCAAAGTGTG
ATGACCAAATGCTTCTAAATTCTAATGCGTAAATCTAATCACCAAATGGATGGTTCAGAAG
TCATAGAAATGGCATTTTGACCGACTATATATTACTTTTCTTAGCTGCATTTCTACTTGGTA
GTTGCCTTGACAATTTTCTCTTGTATTTCCAGGTAGTCAGGCATTTCTTTGAAAGGAAAGA

TATGAAAGGTGCTATTAACGCATTGAGGAGGTTGCCGGATCATTCTGTTCAAGCAGATGT
AATCAGCATCCTTACGGAAAAAATGGAGATGCTGAACTTAGATTTATTTTCTTGCTTGCTT
CCTGTGCTTGTGGGTTTACTGGATAGCAAGGTAGAAAGGCATTCATGTGTGTCTCTCGA
GATGCTACTAAAGCTTGTAGCAGTTTTTGGTTCAGTAGTTCGGTCAACAGTTTCTGCACC
TCCGACTGTTGGTGTTGATCTACATGCTGAGCAAAGGCTAGAGTGCTGCAACCAGTGTT
TTACAGAGCTGCAAAAGATCCAGAAAATTATACCTATGCTTGTGAGGAGAGGTGGTGTAA
TTGCAAGAAGCGCTCAGGAATTGAATCTAGTCCTTCAACAGTGCTGATGGACCATACCAT
CTGCATCCCTTTTCTGCTTACCTGCAGAGTCCAGTTGGATGACATCTAACAGAATACTCG
TGTACATGGACACCGTCTGTCCTTAACTCGGAGTTTGGCTTTACGGCTACTGATTTTGCC
GCTTCCATCGATCAGTTTTGCTTGGTTTCTCATGAGGCAAATGCCCGAGCACTATGCGCA
TGGAAAATTAAGCTCGAAGACTCAAGAGGCCTTGTTCGGGGCATGACTCCATGATATTC
CGAAAGAGTAGGATCTCTCAATTCTTTGATTCTGTTGTATGGTGTATCTTATTGTATCTTC
TATCTGCCCCCCATGTAATTCTTCATGTGTACTTCTCGCTCCCGTGTAATTCATGTGTACT
TCTCGATGGTATAGTGAAAATCCACTTTGTTCAAAAAAAAAA

SEQ ID NO: 93
CTGGACCTCAGTAGAAAGCACGAACCCGTCTTCTTCTCCCTCTTCTTCTTCTCCTCCTCC
TTCCATCGGCGAGCGAGAGAAGACGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAATC
GGAGGAACCTCCATGCGCATTGGCGAACGATGTCCACCTTCCTGACCGGCACCGCTCTT
TCCAATCCCAACCCGAACAAGTCGTACGAGGTCGTTCAACCTCCGAATGACTCCGTCTC
GAGCCTGAGCTTCAACCCCAAGGCCAACTTCTTGGTGGCCACTTCCTGGGACAACCAGG
TCAGGTGCTGGGAGATTGTGCGGAGTGGAACAAGCCTCGGCACTACGCCAAAGGCATC
CATATCTCATGACCAGCCAGTATTGTGCTCAACTTGGAAGGATGATGGAACAACTGTCTT
TTCTGGCGGCTGTGACAAGCAAGTCAAAATGTGGCCGTTGTCAGGTGGCCAACCAATGA
CTGTGGCCATGCATGACGCGCCAATAAAAGAGATTTCGTGGATACCTGAAATGAACCTTT
TAGTTACTGGAAGTTGGGACAAGACATTGAGATACTGGGACACGAGACAAGCCAATCCA
GTTCATATTCAACAATTGCCTGAGCGTTGCTATGCCCTTACTGTGAGACATCCCCTCATG
GTTGTTGGTACTGCAGATAGGAATCTGATAATCTACAATTTGCAAAGTCCTCAGACTGAA
TTCAAGAGAATCAGTTCGCCCCTCAAGTATCAGACCAGATGTCTCGCTGCCTTCCCTGAT
CAACAAGGCTTTCTGGTTGGCTCTATTGAAGGAAGAGTTGGGGTACATCATCTAGATGAT
TCACAACAAAGCAAGAATTTCACCTTCAAATGCCACAGAGAGGGCAGTGAGATATACTCA
GTCAATTCATTGAACTTCCATCCTGTACATCACACCTTCGCAACTGCTGGTTCTGACGGC
GCCTTCAATTTTTGGGACAAGGATAGCAAACAGAGACTGAAGGCAATGTCAAGGTGCAG
TCAGCCCATTCCCTGCAGTACCTTCAATAATGACGGGTCTATTTTCGCATACTCGGCGTG
CTATGATTGGAGTAAGGGTGCAGAGAACCATAATCCTGCAACAGCCAAAACTTACATTTT
CCTGCATTTACCACAGGAATCTGAGGTTAAAGGCAAACCACGACTGGGGACTACTGGAA
GAAAATGACTCCTAATTTCTAGTCGCAAAGCCCCATCATCATATTTCCAGTTGGAAGAGT
CGTATTCCCTGTCGGAAGGACGAGGGATTTGTACATGTATAGACTTTCCAGCTGCTTCTC
CCAAAAAATTTAGCAGTTGACCTGTTTTTGGCCGAAGTCAAGGATTTCGTTGTGTAGTAC
TGGGAGTTACTACTTGTATGTATGTAAATCATGTGGCGTCTGTCCAGTTACTATAGATGG
GAGTCCCCTGGGTGGATTTTGTGAAATATATTTTGGTGATCGGTTGTCTTTTCGTTGTATC
CACAAGACTACAGAAATTACATGCTAAAAAAAAAA

SEQ ID NO: 94
CTTCCTTCCCTGTCACAAGACCCTGCCTTCCTTCTCTGTCTATCTACGCACTCCTGGCAT
CATCGTATCGATTAAAGCTTCGATTTTTATCGGGATTCGAGGAGAAAAGATTGGATTTTTG
CTGCTGTGAGGGAAGAAGATGGAGGTGGAAGCGCAGCAGCGAGACGTTAACAACGTGA
TGTGCCAACTGGTGGATCCAGAAGGAACGACCTTGGGTCCTCCCATGTACCTTCCCCAG
GACGTTGGCCCTCAGCAGCTTCAGCAGATGGTCAACAAACTCCTCAGTAACGAGGACAA
ATTGCCGTATACTTTTTACATATCGGACCAGGAGCTCGTTGTCCCTCTTGAATCTTACTTA
CAGAAAAACAAAGTTTCTGTGGAGAAGGTGTTGTCCATAGTCTATCAGCCACAAGCCATT
TTCCGAATTCGTCCTGTAAATCGCTGTTCAGCAACAATTGCTGGTCACTCAGAAGCTGTT
CTGTCTGTGGCCTTTAGTCCTGATGGGAAGCAACTGGCTAGTGGTTCAGGAGATACCAC
AGTCCGATTGTGGGACCTAAGTACTCAGACCCCGATGTTTACATGCAAAGGGCACAAGA
ATTGGGTCCTCTCCATTGCATGGTCGCCTGATGGCAAGCATCTTGTAAGTGGAAGCAAG
GCTGGGGAAATCCAATGTTGGGATCCGTTGACCGGGCAACCATCAGGCAATCCACTTGT
TGGCCACAAGAAATGGATCACAGGTATATCTTGGGAACCAGTCCATCTGAGTTCACCAT
GCCGTCGCTTTGTGAGTTCTAGTAAAGATGGTGATGCACGCATATGGGATGTAACACTAA
GGAGATGTGTAATATGTCTGAGCGGTCACACTCTAGCTGTTACATGCGTAAAGTGGGGA
GGAGACGGTGTTATATATACTGGCTCTCAGGACTGTACAATCAAAGTCTGGGAAACTTCC

CAGGGGAAGCTGATAAGGGAATTGAAGGGTCATGGACATTGGGTTAACTCCCTTGCTCT
GAGCACTGAGTATGTTCTTCGGACTGGAGCTTTCGATCACACTGGCAAACAGTATTCATC
TGCTGAAGAAATGAAGCAAGTTGCATTAGAAAGGTACAAGAAAATGAAAGGCAATGCTCC
TGAAAGATTGGTCTCTGGATCTGATGATTTTACCATGTTTTTGTGGGAACCTTCTGTCAG
CAAGCACCCAAAAACGCGAATGACTGGTCATCAACAGCTTGTGAATCATGTCTACTTTTC
ACCCGATGGGCAATGGGTGGCTAGTGCTTCGTTTGATAAATCTGTGAAGTTGTGGAATG
GTATTACTGGGAAGTTTGTTGCTGCATTCAGGGGACATGTCGGGCCTGTATATCAAATAA
GTTGGTCTGCGGATAGTAGACTTCTTTTAAGTGGGAGCAAAGACTCTACTCTGAAGATCT
GGGATATTCGCACAAAGAAGTTGAAACGAGATCTTCCAGGTCATGCGGATGAGGTTTTT
GCGGTCGATTGGAGTCCAGATGGTGAGAAAGTAGTTTCTGGTGGTAAAGATAAGGTGTT
GAAGCTATGGATGGGTTAATGCATTGGGAGTTATACAAGGCTTATCAGACCATTCCATAT
GCAAATGAAATGAGATGGCAGTTCAAATGCTTTCATCCGAAGTGAAGAATTGATTGTTTG
ATAGCCCATATTACGGCTGAAGACACAAAGTGCACACTCAACTATCAGGTACTGCAAGTT
CGGATGCAATCAGCCACAAAGTGCGCACTCAACTATCAGGTACTGGAAGTTAGGATGCA
ATTTGGTTGCGTAAAGCCCTCTTGCATCAGGATGGTAGTCGCAGCCGATATGCATTGTCT
ACTTATGATGCAGCTGATATTCTCAAAGGTTTTGTTAATCTAAGAAACACTGTTCAATGCT
GTGTTAATTGAGAGTATCATGAGTCACCGATATTTTGTTCATATAAGTAATTGGCAGAATC
AGTCTTTGATATTCATATCCAGTTATAGTGAGCCCTTCATTTTTCTATTGGCTTTATACATA
ATGGCAGTCAAGTATGTTGGCCAGAGAGGAGCCTTCTTGGATTAACCCCTAGCTTGAGT
CATTAAATCAATTTATCGCACAAGTTTCAATAAAAAAAAAA

SEQ ID NO: 95
CATCTCTCTCTCTCTCTCTCTTTCTCTCTGTGATTGCCTGCGGATCGGCTCTCTGCAG
GCGCTTGAATTGTGAAGCGCAGTCTCTTGATCGCATTTTGCCAGGGCTCTCGCGAAGAT
TTCGCGTCTCCGTGCCTCTTTACTTAAAGATCTGATTTTTGTCTCGAATCACCTTTGTTTG
TGTTGCAGGGTTCCTTCTTTTATTTTTATATTCTTTTGCGAGGTTTCTATCTCTCGCTCCAT
TCGCAAAGAACTTTCCCTCCGCCACCGTTGCGTAACTCGAATAGCCGGATTTTCGTTTTC
GTTTTTATTTCCCCGTTAATTTTCTGCATATCGCCTTTCTTGTGCTTCCCTTTTTTTTTGGT
GTTTGTGTAGTATGGATGCAGGATCTGCGCACTCTTCATCGAATATGAAGACGCAATCCC
GTTCCCCGCTTCAAGAACAGTTTCTGCAGAGGAGGAATTCTCGCGAAAACCTGGACAGA
TTTATACCCAATCGCTCAGCCATGGATTTCGATTATGCGCACTACATGCTGACCGAAGGG
AGGAAGGGGAAGGAGAATCCGGCGGTGAGTTCCCCCTCAAGAGAAGCCTACCGGAAGC
AACTGGCGGAAACGCTCAACATGAACCGGACTCGAATCCTGGCTTTCAAGAACAAACCC
CCGACCCCCGTGGAGTTGATTCCGCATGAACTCACTTCTGCTCAGCCAGCCAAGCCAAC
AAAAACCCGCCGATACATTCCTCAGACCTCGGAGAGGACACTGGATGCTCCTGACCTTT
TGGACGACTACTACCTCAACTTGTTGGACTGGGGAAGCAGCAATGTACTTTCAATTGCTT
TGGGGAATACAGTTTATCTCTGGAATGCTTCGGATGGCTCTACCTCTGAACTTGTCACGA
TCGACGACGAAACTGGCCCTGTAACCAGTGTCAGCTGGGCCCCAGATGGTCGCCACAT
TGCTGTCGGCTTGAACAATTCTGATGTACAGCTATGGGATTCTGCTGATAATAGACTGCT
AAGGACTTTGAGAGGCGGCCACAGATCTCGGGTTGGGTCTCTGGCATGGAACAATCACA
TCCTTACAACGGGAGGAATGGATGGACTGATCGTCAATAATGATGTGAGAGTGAGATCT
CACATTGTCGACACCTACAGGGGTCACACCCAGGAGGTTTGTGGGTTGAAATGGTCGGC
ATCAGGGCAACAATTAGCTAGTGGAGGGAACGATAACATCCTCCACATCTGGGACAGGT
CCACGGCATCTTCCAATTCACCCACTCAGTGGCTTCACCGCCTTGAAGAGCACACTGCC
GCTGTTAAAGCCCTTGCCTGGTGTCCTTTCCAGGGGAATTTGCTGGCTTCCGGTGGAGG
TGGAGGTGATCGGACCATTAAGTTTTGGAACACCCACACCGGTGCGTGCTTGAACTCTG
TGGATACAGGTTCCCAGGTCTGTGCTCTGCTATGGAACAAGAACGAGAGAGAGTTGCTT
AGCTCTCATGGATTCACTCAGAATCAGCTCACCCTTTGGAAGTACCCTTCAATGGTAAAG
ATTGCAGAGCTGACTGGCCATACTTCCAGAGTGCTATTCATGGCCCAGAGTCCGGATGG
TTGCACTGTAGCATCAGCGGCAGGGGATGAAACACTGAGGTTTTGGAATGTGTTTGGTG
TTCCGGAAGTGGCTAAACCAGCTCCAAAAGCTAATCCGGAGCCTTTTGCTCACTTGAATC
GCATTCGCTAAATGAAAAGTCATCTTTGCTGGTTATGTCTATCGTTTTGATCTTCAAGAAT
GCCAAACAATGGGATTCACAAGTCAAGAAGATCGCCACATTGAACAAGTATTGCGAATAG
TTGAAGACAAAATATGCGCCTATACAAAACCATAGTAGCAGTTGCCGTGAGTGTTCCCAA
GTGTATATTTGTACATTTTATATAGTCAAACCATAATTTTTTTTTCCAATGGATGTAGTCAC
TGATGCGTGACCTGCCATGTATATTTTTTGTGAACTAAAATTTTGACTATTGTCACTCCTC
TTTTAGAGCAAACAATATGTTTTTGTAGCAAAAAAAAAAAA

SEQ ID NO: 96

TACTTTCTTCCCCTCCTGCGTCGCTAGCAGCGCGAGAGCAAGCTTCCGCAAGCTTTCGA
TTCGCTCGCATGGAGGAGGCGATCCCCTTCAAGAATCTCCCAAGCAGAGAATACCAAGG
TCACAAGAAGAAGGTGCATTCGGTGGCCTGGAACTGCACGGGGACGAAGCTCGCGTCC
GGCTCCGTCGATCAGACCGCTCGCGTCTGGCACATCGAGCCTCACGGCCATGGTAAGG
TTAAAGATATTGAATTGAAAGGGCACACGGACAGTGTGGATCAATTATGTTGGGATCCGA
AGCATGCAGATTTAATTGCAACAGCATCAGGCGACAAGACTGTACGGCTGTGGGATGCA
CGAAGTGGAAAATGCTCGCAGCAAGCAGAACTCAGTGGGGAAAACATCAACATAACCTA
CAAACCTGATGGAACACATGTGGCAGTTGGTAATAGGGATGATGAGCTAACGATTTTGG
ATGTTCGGAAGTTTAAGCCAATTCACAAGCGGAAGTTCAATTATGAGGTAAATGAAATAG
CATGGAACATGTCAGGAGAGATGTTCTTTTTGACAACAGGAAATGGTACTGTTGAAGTAC
TAGCATACCCATCTCTTCGACCTGTTGATACCCTTATGGCTCACACTGCTGGTTGCTATT
GTATTGCAATTGACCCAGTTGGAAGGTATTTTGCTGTTGGAAGTGCAGATTCCTTAGTGA
GCCTGTGGGATATCTCTGAGATGCTCTGTGTGAGGACTTTTACGAAACTTGAATGGCCC
GTAAGGACAATAAGCTTCAATCATACTGGAGATTACGTTGCTTCTGCCAGTGAAGATCTG
TTCATTGACATTTCAAATGTTCAAACTGGACGGACAGTGCACCAAATTCCATGCCGAGCT
GCGATGAACAGTGTGGAATGGAATCCTAAATACAATTTACTCGCATATGCTGGGGATGAC
AAAAACAAGTACCAGGCAGATGAAGGTGTTTTCCGGATATTTGGATTTGAGAGTGCATGA
GATTTCAGAGCAGGGCGAAATGATTTTGCGCATGCACAAGAGTTTCCAATTTTGTGTTCG
CTTAGGATTTTTTGTACACATTCGAAATCTCGTGTATGGGACGGATCTATCTTTCTTGCAA
GGAGTTCTTTGCACGTCTGTCTATTTACGAAGGAAGAGCGTGTTCTTTCTTTGATCTTAC
CAAGAACTTATTTACTTCACAAAACTTTTTTTTTTAAAAAAAAAA

SEQ ID NO: 97
GGCTCCGTCCCGAAAATCTCGAAAGCTTTCAACCGTCCCCCGCCTAGGGTTTCCACTTC
TCTTGGCGGGAACCCGAAAAATCCCCCGTTGAAAATCCTCGAGAGTCCAGAATTCCCCC
CGGCGGCGAGCGGCGGCGGCGGCGGCGGCGATGGGGAAGGACGAGGAGGAGATGAG
GGGGGAGATCGAGGAGCGGCTGATCAACGAGGAGTACAAGGTGTGGAAGAAGAACAC
GCCGTTCCTGTACGACCTGGTCATCACCCACGCCCTCGAGTGGCCCTCCCTCACCGTC
GAGTGGCTCCCCGACCGGGAGGAGCCCCCCGGGAAGGACTACTCCGTCCAGAAGCTC
GTCCTCGGCACCCACACCTCCGAGAACGAGCCCAACTACCTCATGCTCGCCCAGGTCC
AGCTCCCCCTCGAGGACGCCGAGAACGACGCCCGCCACTACGACGACGACCGCGCCG
ACGTCGGCGGCTTCGGCTGCGCCAACGGCAAGGTTCAAATTATACAGCAAATAAATCAT
GACGGAGAGGTGAACCGAGCCCGTTACATGCCACAAAACTCCTTTATCATAGCAACCAA
GACAGTTAGTGCAGAAGTTTATGTTTTTGATTACAGCAAGCACCCATCTAAGCCTCCACT
AGATGGTGCGTGCAGTCCTGACTTAAGGTTGAGGGGCCACAGTACTGAAGGCTATGGAT
TGTCGTGGAGTAAGTTCAAGCAAGGACACCTACTCAGTGGTTCAGATGATGCTCAGATAT
GCTTATGGGACATCAATGCAACTCCTAAAAACAAGTCTCTCGATGCCATGCAAATCTTCA
AGGTTCATGAAGGTGTTGTTGAAGATGTCGCGTGGCATCTCAGACATGAATATTTATTTG
GTTCAGTTGGGGATGATCAATACCTGCTAATATGGGATCTACGGACTCCATCAGTAACCA
AGCCTGTCCAATCTGTAGTGGCTCATCAGAGCGAGGTTAACTGTCTAGCATTCAACCCCT
TTAATGAATGGGTTGTTGCGACGGGTTCTACAGATAAGACTGTTAAGTTATTTGATCTAC
GCAAGATCAGCACTGCACTTCACACCTTTGATGCTCACAAGGAAGAAGTTTTCCAAGTTG
GGTGGAATCCGAAAAATGAGACAATCTTAGCTTCATGCTGCCTTGGCAGACGACTTATG
GTGTGGGATCTTAGCAGGATTGATGAAGAGCAGACGCCGGAGGATGCTGAAGATGGAC
CGCCCGAGTTGCTTTTCATTCATGGTGGCCATACGAGTAAAATTTCAGACTTTTCATGGA
ACACCTGTGAAGACTGGGTTGTAGCCAGCGTGGCTGAAGACAACATCCTTCAAATATGG
CAGATGGCCGAGAATATCTACCATGATGAGGATGATGTTCCAGGAGAAGAGTCCAATAA
AGGATCTTGACCTAATCATTAACAATTCAGAAGTTTAGGAGAAAGCCGTAGTAACTTGAA
CCCCTATTTATCTTTACCTGTACCTACGCGCACCATTATCTTTGATCTTGCAATGCTGTGC
TAGATGATGAGTTTTTACTGTCTTTAGAATTACCATTTGTAAAGGATTCACTTCTCTTTTTA
ATTATTTTTTTGAACTCCGAAAAAAAAAA

SEQ ID NO: 98
AAAAATAAAAGCGACGGTTTGTGATTTTCGAACAAAAAGTATTTTGGGGAGACATTTCCC
TGAATCCCTCAACGGTTTGCCTCGCCCCTGTTCGGAAAATCCGAGGGAAAAATAGGTTC
GAAACCTGCGTCGCCTCAAAGATGGGATCGAACCGGAATGATTAGACCCTTGATTAGGG
TTCTCTCACATTCGATCCCCTGAAATTATCATCATCATCATCATCATCATCAAATCGA
ATCCATTTTCAGGATCATTTTGAGTTTCGAGTTCGATCTCCATTGCCGATCCGATCTCGG
GAGGACGAGAAGGGACGCGGCGCAAAGATTTCGGTGCGTGAAGGAGGTTGTTCCGGG
GACGACAAATGATGAGGGGCTTCTCGTGTACGGAAGATGGCGATGCGCCTTCGACCTC

GAGTACTTCGCCGCCGCCGCCGCCGCCGCCGCCCCATCGACAGCAGATGCAGGCGCC
TCGAGCTTCTTCTTCTTCTTCTGGGCAGCCGACGAGCCGGCGGAGCACTGGCAATGTCT
TTAAACTCTTAGCACGAAGAGAGGTCTCTCCACGTTCAAAACATTCTTTAAAGAAGTTTTG
GGGAGAAGCCTCAGAGTGTCAGCTCTGTCCCTTCCAGCAAAGTTATGAAGCAGTAAGAG
ATGTAAGACGAAGTCTCATCTCATGGGTTGAGGCATTTTCACTGCAACATTTGTCTGCCA
AATATTGCCCCCTTATGCCTCCTCCAAGATCAACGATTGCTGCAGCCTTTAGTCCTGATG
GGAAAATACTTGCTTCTACACATGGAGATCATACTGTCAAACTTATTGATTCTCAGACGG
GGAGTTGTTTAAAGGTGTTAAGGGGTCATAGGAGGACACCGTGGGTGGTAAGGTTCCAC
CCCTTGTACCCAGAGATCCTTGCAAGTGGTAGTTTGGATCACGAAGTTCACCTGTGGGA
TGCAAATACTGCAGAGTGCATAGGATCTCGCAATTTTTACCGTCCTATTGCATCCATTGC
GTTCCACGCCCAAGGAGATCTTCTTGCTGTTGCATCTGGTCACAAGTTGTACATTTGGCA
CTACAATAGAAGTGGAGAGACATCATCACCAACTATTGTTTTGAGGACACCACGATCTCT
TAGGGCTGTGCATTTCCATCCACATGCAGCTCCATTTCTTTTAACAGCTGAGGTCAATGA
CCTTGACTTGACAGATTCGGCTATGACCCTGGCAACTTCTCCCGGTTATTTGCACTACCC
TCCTCCTACTATATATCTTGCAGATGCTCATTCTAATGAAAGATCTAGGTTGGAAGATGAA
CTGCCTCTCATGCCTTCACCATTGTTGATGTGGCCTTCCTTTACTAGAGATGATGGAAGA
GCAACCTTGCCACACATAGGTGGGGATGTTGGTCTTAGTGGACAGCAGAGGGTAGACTC
GTTGTCTTCAGGGCAGTATGAATTCCATCCATCTCCAATTGAACCTAGCAGCTCTACTTC
CATGCATGAAGAGATGGGAACTGATCCTTTCTCTAGTGTAAGGGAATCTGAAGTAACTCA
GTCTGCAATGAACATTGTGGACAATACGGAAGTGCAACCTGAAGAGAGAAGTACTTATA
GTTTCTCTTTTAGCGACCCAAGGTTTTGGGAGCTGCCTTCGGTGTATGGATGGTTAGTTG
GTCAAACCCAAGCTGCCCCTCGAACTGCACCAAGTCCTGGAGCTCTTGAAACTGCATCT
GCTCTTGGTGAGGTAGCGAGTGTTTCACCTGTCAGATCTGAATTTATGCCAGGTGGCAT
GGACCAGCCACGGCTTGGTGGAAGATCTGGGTCTGGATGTCGGTCTTCTGGGTCTCGG
ATGATGCGCACAGCTGGACTTAATGATCACCCACATGATGAGAACTATCCTCAATCTGTT
GTTAGTAAACTCAGGTCAGAACTTGAAGCCTCGCTGGCTGCAGCAGCATCTACAGAGTT
GCCATGCACGGTGAAGCTTAGAGTATGGCCATACGACATGAAAGATCCTTGTGCACTTTT
TCGTTCAGAAAGTTGCCGCTTAACTATTCCTCATGCAGTTCTATGCAGTGAAATGGGCGC
CCATTTTTCTCCTTGTGGCAGATTCTTTGCTGCCTGTGTTGCATGTGTGCTACCTCAGTT
GGAAGCTGATCCTGTATTACACGGCCAGGTCGATCCCGATGTCACTGGAGTGGCAACCT
CACCTACTAGACACCCTGTTTCTGCTTATCAAATCATGTATGAGCTTCGGATTTATTCCTT
GGAGGAGGCAACATTTGGAATGGTGCTTGCATCACGATCCATAAGAGCTGCTCATTGTT
TAACCTCCATTCAGTTCTCTCCAACATCTGAACACTTACTGCTCGCCTACGGCCGTCGAC
ATAATTCGCTTCTTAAAAGTATAGTCATTGATGGAGAGAATACAGTGCCTATTTACAGTAT
ATTGGAGGTCTACAGAGTTTCTGATATGGAACTTGTGAGAGTTCTTCCCAGTGCGGAAGA
TGAAGTAAATGTTGCATGCTTTCACCCTTCAGTTGGAGGTGGCCTGGTTTATGGAACAAA
GGAAGGGAAGTTAAGGATTCTCCAAATTGACAGTTCTGGGGGCTTAAATCCAAAATCAAC
TGGTTTTCTTGATGAAAACATGGCAGAGGTGCCTACATATGCTTTAGAATGCTAGTAGCT
TGGAAGTGCAACACGCTCGTGAAATTTACTGCTCGACCTTGAAGCAGTCTTTTGAAGCAT
ATTATGTCAAAACTGGGAAGAAGTTCATCAATTATGACATCTGAGGTAAACTCAAAAAAAA
AA

SEQ ID NO: 99
CTTCGAGAGAGAGAGAGAGAGATGGGCGAGGGCGATCTCCCCCGGACGGAGGCGG
GCGTGCTCCGGGGACACGAGGGCGCGGTGCTGGCCGCCCGGTTCAACGGCGACGGCA
ACTACTGCCTCAGCTGCGGCAAGGACCGCACCATCCGCCTCTGGAACCCCCACCGCGG
CATCCACATCAAGACCTACAAGTCCCACGGCCGCGAGGTCCGCGACGTCCACTGCACC
TCGGACAATTCAAAGTTGATATCTTGCGGTGGTGACCGGCAGATATTCTATTGGGATGTG
TCGACTGGCCGAGTCATTAGGAGATTCCGTGGGCATGATAGCGAGGTGAATGCAGTAAA
GTTTAACGACTATGCATCTGTCGTAGTATCAGCCGGCTATGATCGTTCAGTGCGTGCTTG
GGATTGCAGATCTCATAGCACTGAGCCGATTCAGATTATCAACACGTTTCAAGACAGTGT
GATGTCTGTTTGCTTAACAAAAACTGAAATTATTGGTGGCAGCGTTGATGGTACTGTTCG
AACATTTGACATTCGTATTGGTAGAGAAATATCTGACGACTTGGGGCAACCTGTCAACTG
TATTTCAATGTCAAATGATGGTAACTGCATATTGGCGAGTTGTTTAGATTCAACTTTGCGT
CTTGTAGATAGGTCTGCCGGTGAGCTATTGCAAGAATATAAGGGCCATACTTGCAAGTC
CTACAAATTGGATTGCTGTCTTACTAATACTGATGCACATGTGGCTGGTGGATCTGAGGA
CGGTTATGTCTTCTTCTGGGATTTGGTCGATGCATCAGTGATATCCAAATTTCGAGCTCA
TTCCTCTGTGGTAACAAGTGTAAGTTATCATCCAAAGGAAGACTGTATGATCACTGCCTC
TGTGGATGGCACAATTAAGGTGTGGAAGACATGAAACTGTTCAGCTGAGCCTGAGGGAT
CATTCCTTTGTCTGAGATACGATATAGGGAATTCACCGAGTCAATTCTTAGGATATCTTGA

ATCAGACCATTTGTACTTACTTTTGCCTCAGGCAATGCATCTTTTCATGCTTCATTAGGAG
GTAATATCTTTTGGGTAGCAAAACTTTTTGCAGAATTTCGGAAGCTGAGGTTCATGTATAT
TATAGTCTAGCTCGGGAGGATTGCACTTCCAGCAAATTTCTTAGGGGTTGTGTACCATTG
ATGGGATGGATTGATAATCTATGAGCTGAATGTGCATTTCGATTTCTTGTCTGTGCCAGA
GAAAAAAAAAA

SEQ ID NO: 100
GACACGCCCAAAATCCACTCCTTCTCCCTCCCTTCCCTTCCCTTCCCTCTCATGCACAAG
ATCAGATCACGCAAACCCTAGATCACCACCGCCAGATCACCACCGCCATCGCCCCATCT
GAGATTCATTCTTCCCATTCCCCGATTGATCCGATTGATTTTATACTGTACCAATAAATCC
GTAATCTTGAGAGCTTTTGCGCTCTGTGCAATTCCCGGCGATGGCGTGCATAAAGGGGG
TGGGCCGGTCGGCGTCGGTTGCCATGGCCCCCGACGGCGGGTACTTGGCGACTGGGA
CCATGGCCGGGACTGTGGATCTCTCCTTCAGCTCCTCCGCCAGCCTCGAGATCTTCGGC
CTCGACTTCCAGTCCGACGATCGGGATCTCCCCCTCATCGCCGAATCCCCCAGCTCCGA
GCGCTTCAACCGGCTCTCCTGGGGCAAGAACGGATCCGGCTCCGACGAATTCTCCCTG
GGCCTGATCGCCGGTGGCCTCGTCGACGGGACCATTGGCCTTTGGAACCCGCTCTCCC
TGATCCGTTCTGAGGCTGGTGATAAGGCAATCGTAGGACACCTTTCTAGGCATAAAGGA
CCTGTTCGTGGTCTTGAGTTTAACGTCATTGCACCAAACTTACTTGCTTCTGGGGCTGAT
GATGGTGAAATTTGCATTTGGGATTTGGCTGCACCAAGAGAACCTTCTCATTTCCCTCCT
CTTAGGGGTAGTGGTTCTGCTGCTCAAGGTGAAATTTCATTTTTATCTTGGAATAGCAAA
GTTCAACATATATTAGCCTCCACTTCCTATAACGGGACAACAGTTGTATGGGATCTCAAG
AAACAAAAACCAGTCATAAGTTTTTCAGATTCAGTTAGAAGGCGTTGCTCAGTTTTGCAAT
GGAACCCTGACCTGGCTACTCAACTTGTTGTTGCATCAGATGAAGATAGTTCTCCTACTT
TAAGGCTTTGGGATATGAGAAACATAATGTCACCTGTTAAAGAATTTGCGGGGCATACAA
GAGGTGTTATTGCAATGTCCTGGTGTCCTAATGATAGTTCCTATTTGGTCACCTGTGCTA
AAGACAATAGAACCATATGCTGGGATACAGTTACTGGAGAGATTGTTTGTGAATTGCCTG
CTGGATCCAACTGGAATTTTGATGTGCATTGGTACCCTAAGATACCCGGAGTTATATCAG
CATCTTCATTTGATGGAAAGATTGGCATCTACAATGTTGAGGGTTGTAGCCGCTATGGTG
TCAGGGAAAATGAATTCGGGGCTGCAACTCTGAGAGCTCCAAAATGGTTTAAACGACCT
GTTGGCGCATCCTTTGGTTTTGGAGGCAAGGTGGTTTCATTTCATACTCGGTCAACAGGA
GGTCCTTCAGTCAATTCCTCAGAGGTTTTTGTACACGATATCATTACAGAACAAACTTTGG
TGAGCCGCTCGTCTGAATTTGAAGCTGCAATCCAAAGTGGGGATAGACCTTCATTGAGA
GCATTATGTGAAAAGAAGTCTCAACATTGCGAATCTACAGATGACCAAGAAACATGGGGA
TTCTTGAAAGTTTTGTTAGAAGATGATGGAACCGCAAGGTCAAAGCTACTTGCTCACCTT
GGTTTCGATATTCCTACGGAGACAAATGATGGTTCACAAGAGGATCTCTCCCAGCAAGTT
AATGCTCTTGGGCTCGAGGATGTGACTGCAGATAAAGTAGTGCAGGAGGACAACAACGA
AAGCATGGTGTTTCCTACTGATAATGGCGAAGATTTTTTTAACAATCTTCCTAGTCCCAGG
GCCGATACACCAGTATCAACTTCTGCTGATGGCTTTCCTACTGTGAATGCTGCCGTGGAA
CCATCGCAAGACGAAGTAGATGGACTTGAGGAGAGCTCTGACCCATCGTTTGATGATAG
TGTTCAACGTGCTTTGGTTGTTGGAGACTACAAGGCTGCTGTTGCATTGTGCATGTCTGC
TAATAAATTGGCTGATGCTTTGGTTATTGCACATGTTGGAGGTGCCTCCCTATGGGAGAG
TACTCGTGATAAGTACCTCAAAATGAGCCGCTTACCTTACTTGAAGGTTGTTTTTGCAATG
GTGAATAATGATCTCCAGAGCCTTGTGGATACCAGGCCCCTCAAATTCTGGAAAGAGAC
GCTCGCTATTCTATGTAGTTTTGCACAGGGGGAGGAGTGGGCGATGCTCTGCAACTCTC
TAGCCTCAAAACTTATGGCAGCTGGTAACATGTTGGCAGCGACACTCTGTTTCATTTGTG
CTGGGAATATTGATAAAACAGTGGAAATTTGGTCGAGAAGCCTGGCAACTGAACATGAT
GGGATGTCCTACATGGACCTTCTTCAGGATTTGATGGAAAAGACTATTGTTCTTGCATTG
GCTAGCGGCCAAAAACAATTCAGCGCTTCAGTGTGCAAGCTTGTGGAAAAATATGCAGA
GATTTTGGCTAGTCAAGGGCTTTTGACAACAGCGATGGACTATTTGAAGCTGTTGGGGA
CAGATGATTTGTCACCTGAACTCGCGGTCTTGAGAGATCGTATCGCCTTCTCTGTAGAAG
CTGAAAAAGGAGCCAACATCTCTGCTTTTAATGGCTCTCAAGATCCAAGAGGTGCAGTAT
ATGGTGTCGATCAATCCAATTATGGCATGGTTGATACTTCTCAGCACTATTATCCGGAAG
CTGCACAACCACAGGTGCCGCATACTGTTCCTGGTAGTCCTTATGGTGAAAACTATCAG
CAGCCTTTTGGCTCTTCATTTGGGAAAGGATACAATACTCCCATGCAATATCAGGCTCCT
TCACAGGCGTCTATGTTTGTTCCATCCGAGCCACCTCAGAATGCCCAGCCAAGTTTTGTT
CCAACTCCTGTCACCAGTCAGCCCACGACGAGATCTCAATTTATACCAGCGCCTCCTCTT
GCCCTGAGGAACCCAGAGCAGTATCAGCAGCCAACATTGGGTTCTCATTTGTATCCGGG
AAGTGTTAATCCTACTTTTCAACCTTTGCCTCATGCACCTGGTCCGGTTGCCCCTGTGCC
ACCACAAGTGAGCTCTGTTCCTGGCCAGAATATGCCACAAGCTGTGGCTCCTACTCAAA
TGCGGGGGTTTATGCCAGTTACTAATCCGGGGGTTGTTCAAAATCCCGGGCCAATTTCC

ATGCAGCCGGCTACTCCCATTGAATCAGCAGCTGCACAACCAGTTGTCTCACCTGCTGC
ACCCCCACCTACTGTGCAGACTGCAGATACTTCAAACGTTCCTGCCCCCCAGAAACCTG
TCATCGCAACATTGACGAGACTCTACAATGAGACATCAGAAGCATTGGGAGGCTCACGT
GCAAATCCTGCCAAGAAGCGTGAGATAGAAGATAATTCAAGGAAAATAGGTGCCTTGTTT
GCAAAGCTCAACAGTGGAGACATATCTAAAAACGCAGCTGATAAGCTTGTCCAGCTATGT
CAGGCTTTAGACAACGGTGATTATAGCACTGCTCTGCAAATACAGGTACTTCTCACGACT
AGCGAGTGGGATGAGTGCAATTTCTGGTTGGCAACGTTGAAGCGGATGATCAAGACAAG
GCAGAATGTGAGACTAAGTTGAGAAGAGGCCAATTCTTCTCAACGCATCTCCTTTGTTCT
TTTTTCTCTCCTTAGCCTTGTGTGGTGGATTCTGTAGGGGAATTGATCATTTCTTTGGAAA
GCATGGCTCTCTTAGATTGCTACTGGAGTAATTTTTGGTCTTTTGTTATAGAATGTCTTTT
TTCATTGATGCTGACTTCTGAAAGCCGTTAGGATAGTCTTTAAAGGAGTTGGTGATTATT
GATTTCCACCCAATATATGTAGCGTCAAGGAATTAGAGAGCTTTTTACTAGCCCTGGTGG
ATGTTTTTGACACTATTTTTCACTATAAGTTACTCTCCCAGGCTGGTCTCTTCTGAGATGC
CCGACGTCTGTTCATGACTTAGTGTTATAGTGATCATTTATGAATTTTTGGTATGCTTTTG
GTGAGTGATAGCTTGTCAAAGAGAGAGAGAGAGGCTTGTCACAAGAAAACAATCAATTAT
CTTGTTGGTGTGGGGCTGAAGATTCTTCGTTTTGTTTGTATTTTCTTTTTCCATTATCTGG
ATAAGGAAGACAGCAGCAAATGTTGTCGCTGAAAAATAACCTTGCTTTTGACACTGTTGA
GAGAAGACTCGGTAGGATGTATTATTGGATGCTGTTATTCCAAAAAAAAAA

SEQ ID NO: 101
GGAAACGCCGAGAGAGGGAAAGGAAGGAAGGAAGGAGAAGGAGATGAAGGAGAGAGG
GAAGGGGGCGGGGAGATCGGTGGACGAGAGGTACACCCAGTGGAAGTCGCTCGTCCC
CGTCCTCTACGACTGGCTCGCCAACCACAACCTCGTCTGGCCCTCCCTCTCCTGCAGAT
GGGGTCCACAGCTCGAGCAAGCTACTTACAAAAATCGACAGCGTCTTTACCTATCTGAA
CAGACAGATGGCAGTGTTCCAAATACACTTGTTATTGCAAATGTCGAGGTTGTCAAACCA
AGAGTTGCAGCTGCAGAGCATATCTCTCAGTTTAATGAAGAAGCGCGCTCTCCCTTTGTC
AAGAAGTTCAAAACTATTATACATCCTGGGGAGGTAAACAGGATCAGGGAGCTGCCCCA
AAACAGTAAGATAGTAGCAACTCACACCGATAGCCCTGATGTCCTCATCTGGGATGTTGA
GACCCAGCCTAATCGCCATGCTGTTTTAGGTGCTTCAACTTCACGTCCAGATCTGATACT
CACCGGGCATAAAGATAATGCTGAGTTTGCTCTTGCAATGTCCCCAACCGAACCTTTCGT
TCTTTCTGGAGGTAAAGACAGATATGTAGTTTTATGGAGCATTCAAGATCATATCTCAACT
TTGGCTGCAGATCCTGGGTCTGCAAAATCTCCAGGGTCTGCTGGAACCAACAATAAGCA
GTCTTCTAAGGCTGCTGGTGGCAATGATAAGACTGGGGACAGCCCTTCAATTGAACCTC
GTGGAGTCTACCTGGGGCATGGCGACACTGTTGAGGATGTGACTTTCTGCCCGTCGAGT
GCACAGGAGTTCTGTAGTGTGGGGGATGATTCTTGTCTCATCCTATGGGATGCAAGAAC
TGGCTCTTCTCCAGCTATCAAGGTTGAGAAAGCTCATCATGCTGATCTCCATTGTGTTGA
TTGGAATCCTCATGATGTAAACCTGATTCTAACCGGGTCAGCTGATAATACAGTTCGCAT
GTTTGACCGCCGAAACCTCACTTCGGGGGGAGTTGGATCACCTGTCCATACTTTTGAGG
GTCACAATGCTGCTGTGCTCTGCGTACAGTGGTCTCCAGACAAGTCATCTGTCTTTGGAA
GTTCTGCTGAGGATGGCATTTTAAATATCTGGGATCATGAAAAGATTGGTAGAAAGATAG
AGACCGTGGGTTCAAAAGTACCCAATTCTCCTCCAGGGTTATTCTTCCGTCATGCTGGG
CACAGGGACAAGGTTGTTGACTTCCATTGGAATTCATCCGATCCCTGGACAATTGTTAGT
GTTTCTGATGATGGTGAAAGTACTGGTGGAGGTGGTACACTGCAGATATGGCGGATGAT
CGATTTGATTTACAGACCTGAGGAAGAAGTTCTGGCTGAGCTGGACAAATTCAAGTCTCA
CATTCTCTCTTGCACCTCTTGAACTTCTCATAATAAGAAGAAAGATCAGCAAGGTTGATGA
CTAACTTGAAGTGGGAAGCAAGTTAACCGTGTTCGATTGTTTTGTGTAAAAAGGAGCTCA
TCATCCTCGAGTTTCATCCTTTTTTATGATGCTAGGGATTATTTAGTGCGTCTGCTTGAGG
TTTAGTTTTGTCTCGGTGTGGACATGGAGTCTGACAAATGTGTTCTCCAGCATTCCTTAA
GGCCATGGCTGAGAGGCTCGTTTTGGTTGCTTCATTGGATTATGAAGAGATTGCTTAATA
CTTAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 102
CTAATTTCAAGGTTCTCTTTCCTTTCTTGCAGACAACACTCGCCCCTCGCTTCTGGTCTTT
TCCAAAGCAATAAAATTCCCTCGTTCGCCGTTCGGCGTCTTCACCGGCGGCGCGCGCC
GCACTGCGTACCCACCGGCTGTCGCGTTCTCGCGGATCGAACTCGAGGAAAAGGCATC
GGCGGCGGATCGGGGCAAATGGCGAAGATCGCGCCCGGGTGCGAACCGGTGGCGGG
GACGCTGACCCCGTCGAAGAAGAGGGAGTACAGGGTCACCAACAGGCTCCAGGAGGG
GAAGCGTCCCCTCTACGCCGTCGTCTTCAACTTCATCGACTCCCGCTACTTCAACGTCTT
CGCCACCGTCGGCGGCAACCGGGTTACTGTTTATCAGTGTCTCGAAGGTGGAGTAATAG
CTGTGTTGCAGTCATACATTGATGAAGATAAGGACGAGTCGTTTTACACGGTCAGCTGG

GCGTGCAATATCGATAGAACCCCGTTTGTGGTGGCGGGAGGAATCAATGGTATCATCCG
TGTAATTGATGCTGGCAATGAGAAGATACACAGGAGTTTTGTAGGCCATGGGGATTCAAT
AAATGAAATCAGGACTCAACCATTGAACCCATCCCTCATCGTGTCTGCTAGCAAAGATGA
ATCCGTTAGGCTCTGGAACGTTCATACGGGAATTTGTATCCTGATATTTGCTGGAGCTGG
GGGTCATCGCAATGAAGTTTTGAGTGTGGACTTCCATCCTTCCGACAAGTACCGTATTGC
AAGTTGTGGTATGGACAATACGGTTAAAATCTGGTCAATGAAAGAGTTCTGGACATACGT
GGAGAAGTCATTTACATGGACAGATCTTCCATCGAAGTTTCCCACCAAATACGTGCAGTT
TCCAGTTTTCATAGCTCCAGTTCATTCAAACTATGTTGACTGCAACAGGTGGCTTGGTGA
TTTTGTTCTGTCAAAGAGTGTTGACAACGAGATTGTGCTTTGGGAACCCAAAATGAAGGA
ACAATCTCCGGGAGAGGGGTCGGTGGATATCCTTCAGAAATATCCAGTTCCAGAGTGTG
ACATTTGGTTCATCAAATTTTCCTGTGACTTTCATTATCACTCAATTGCTATAGGAAATAG
GGAAGGGAAGATCTACGTATGGGAGCTGCAGAGTAGCCCTCCTGTTCTAATTGCAAAGT
TGTCTCATCCCCAATCAAAATCCCCAATCAGACAGACTGCCATGTCATTTGATGGGAGCA
CAATCCTGAGCTGCTGTGAGGATGGTACTATATGGCGCTGGGATGCAATTACGGCATCA
ACATCCTAAGCCTTCCATGGCAGATGGACTGGAGAACTCCGTTAGTAATTATGAATCCCT
CTTGTGTGGGCATGTTCCCCACCATGTATCAGCTGAATGGGAGCTGCTTCAACCTCTTAT
CTCGATGGAGACTCGAATAGCATCACCGCACAGATGCAAGCGGACAACTGCTTTTTCGT
AACGAAGAAAGCAAGTGGATGATTTGGTTGTGCATCAGTCTGAACGATTTATGAAGTTAC
TTTTTGGTGTCAAATGTACTCTCCGTGAATCATTTCACTTCGCAAACTGGGATTTGTACCC
TAGAAACATCCAGTTTAATCTACCTTAACTTCCCAGAAAAAAAAAA

SEQ ID NO: 103
ATTTCGGGAGGAGTAAAAAAGGAAAATGAAAAGAGAGGGAGAGACAGAGAGGTCGGGC
ATTATCCATTCAGCATAGGAGGCGGGGGAGAAGGAGACAAAGCCCATGAGATCATCAG
GATTGGTGGTGGGATTACTGTTTAACTGCGACATGGATCGAGCAATGTCGCTGTTCACT
CTGTTGCCACTAGCGGTATGAATACGGCAATGCATTTTGGTGCTGGTTGGCGATCGATT
GCTGAGATGGGGTATACGATGAGCAGACTAGAGATTGAGCCTGAGTCGTGTGAGGACG
AGAAGAGCTTGGATGGGGTTGGTAACAGCCAGGGACCGAATGAGTTGCCGAGATGCTT
GGATCATGAGTTGGCGCATTTGACGAATCTGAAGTCGAGGCCCCATGAACATTTGATCC
GAGATTTCCCTGGGAGGCGGGCTCTGCCTGTTTCCACCGTTAAGATGCTGGCGGGTCG
AGAGTGTAATTATTCACGAAGAGGGAGGTTCTCCTCCGCTGATTGCTGTCACATGCTGA
GCAGATATGTGCCTGTTAATGGTCCTTCGCCCCTGGATCAGATGAATAGTCGAGCTTATG
TTTCGCAATTTTCAGCTGATGGTTCTCTATTTGTTGCTGGCTTTCAGGGTAGCCACATTAG
AATTTATAATGTTGATAAAGGATGGAAATGTCAGAAGAACATTCTTACCAAGAGTTTACGG
TGGACGATCACTGATACATCTCTTTCTCCTGACCAACGTTACCTTGTGTATGCCAGTATG
TCACCCATCGTCCATATTGTTGACATCGGCTCCGCTGCTATGGATTCTCTTGCAAACATC
ACGGAGATCCATGAGGGTTTGGATTTTTCCGCTGACAGTGGACCATATTCTTTTGGAATC
TTCTCTGTTAAATTTTCTACCGATGGACGAGAAGTCGTCGCTGGAAGCAGCGACGATTCT
ATATATGTCTATGATCTTGTGGCAAATAAGCTTTCCCTCAGAATTCCAGCACATGAGTCTG
ATGTGAACACAGTATGCTTTGCTGATGAAAGTGGTCATATAATTTATTCTGGGAGTGATG
ATACATACTGCAAGGTGTGGGATAGACGTTGCCTGAGTGCCAGAAATAAACCTGCAGGA
GTTCTAATGGGACACCTTGAAGGCATTACGTTCATTGATAGCCGTGGTGATGGTCGTTAT
TTCATATCAAATGGCAAAGATCAGACGATCAAACTTTGGGATATCCGGAAAATGGGCTCT
GATATCTGTCGTCGAGGCTTTAGGAATTTCGAATGGGATTACAGATGGATGGACTACCCA
CCCCGGGCTAGGGATTCGAAACACCCTTTTGATCTGTCAGTGGCAACATATAAAGGCCA
TTCGGTGTTGCGTACTCTTATTCGGTGCTACTTCTCCCCAGTACATAGCACTGGTCAAAA
GTATATCTACACTGGATCCCATGATTCCTGTGTTTATATCTATGATGTGGTGACTGGAGC
TCAAGTTGCGGCCCTCAAGCACCATAAATCGCCGGTCAGAGACTGCAGTTGGCACCCG
GAGTACCCGATGATTGTGAGCTCTTCTTGGGATGGGGATATTGTGAAATGGGAATTCTTT
GGGAACGGAGAAACTGAGATCCCGGCGATGAAGAAGAGGATCCGGAGGCGGCATTTGT
ATTAAAAGATGCTTACTATTTTTCACTCGATGACGGTTGGCCGGATAAATAATCGCTTATA
TAGTCCTAATAAGTTCCATCTCATGTACAGATAAGGTGTATCATTCAAGCTATAGTCCCAA
ACGACACGCGATTCGTGTGAATTTCGTTGTACATAATCGACCTTTCGAATGTAATTGTGA
CTGTGAGCCATTGTTGCTTGTTACAAGCTTCCAAAAAAAAAA

SEQ ID NO: 104
ATTCACTGGCCTAACTACTACCCGACCGGGAGCAACCGAAGCGCACGGTGCCCGTCGA
GCGCGCTCTCTCTCCGGCGATGGAGCCGCAGCCGCAGGCCCCGAAGAAGCGCGGCCG
GAAGCCGAAGCCGAAGGAGGACAAGAAGGAGGAGCAGCTCCACCAGCCGCCGCCGCC
GCCGCCGCCGCAGCAGCAGGCGGCTCCGGCGCCGGCACCGGCGGCCACCAGGTCGT

CGACGTCGGGGTCGGCGGGGGGCCGGGACCGGAGGCCGCAGCAGCAGCACGCGGTC
GACGAGAAGTACGCGCGGTGGAAGTCCCTCGTCCCAGTCCTCTACGACTGGCTCGCCA
ACCACAACCTCCTCTGGCCTTCTCTCTCTTGCCGGTGGGGCCCGCAACTCGAGCAAGC
GACTTATAAGAATCGGCAGCGGCTCTACATTTCTGAGCAGACTGATGGCAGTGTTCCAA
ATACTTTGGTGATAGCAAACTGTGAAGTTGTGAAACCTAGAGTTGCGGCTGCAGAGCAC
GTGTCCCAGTTTAATGAGGAAGCTCGCTCTCCCTTCATAAGGAAGTACAAGACAATTATA
CATCCTGGAGAGGTTAACAGAGTCAGGGAACTTCCTCAGAATCCCAATATTGTGGCAAC
TCACACTGACAGCCCAGATGTTCTCATTTGGGATGTGGAATCTCAGCCTAACCGGCATG
CTGTCTATGGAGCTACAGCTTCTCGTCCAAATCTGATTTTAACTGGACATCAAGAGAATG
CTGAATTTGCCCTTGCAATGTGTCCAGCTGAACCCTTTGTTCTCTCTGGAGGGAAGGATA
AGACGGTGGTTTTGTGGAGTATCCAAGACCATATAACAGCATCTGCAACAGATCAAACAA
CTAATAAATCTCCAGGATCTGGAGGATCCATCATTAAGAAGACTGGGGAAGGTAACGAG
GAAACTGGAAATGGCCCTTCTGTTGGACCACGAGGAATCTACTGTGGACATGAGGATAC
TGTTGAAGATGTGGCTTTTTGTCCATCCACTGCACAGGAATTTTGTAGCGTTGGTGATGA
TTCATGCCTTATATTGTGGGATGCACGAGTTGGGACTAATCCCGTTGCTAAGGTCGAGAA
GGCACATAATGGTGACCTCCATTGTGTGGATTGGAATCCCCATGACAACAACCTAATCTT
AACTGGGTCGGCAGATAACTCTGTTAACATGTTTGATCGGCGAAATCTCACTTCTAATGG
AGTTGGTTCACCAGTCTACAAGTTTGAGGGGCATAAGGCAGCTGTTCTTTGTGTGCAGT
GGTCTCCAGACAAGCCTTCCGTCTTTGGGAGTTCTGCTGAAGACGGTCTCTTGAACATTT
GGGATTATGAGAGGGTTGATAAAAAGGTTGACAGGGCTCCAAATGCTCCTGCGGGATTG
TTTTTCCAGCATGCTGGTCACAGGGACAAAATTGTCGACTTCCACTGGAACGCAGCTGAT
CCATGGACTATGGTTAGCGTATCTGATGACTGTGATACTGCTGGAGGAGGTGGTACATT
GCAGATATGGCGAATGAGCGATCTGATCTACAGGCCGGAAGAAGAGGTTTTGGCTGAG
CTGGAGAATTTCAAGGCGCATGTACTGGAATGCTCGAAGGCATGAGAGTGCCTCGAGAA
CAGGCCCTCCGGGTCTCAAAACACTAACTAGACAAAGCGGGTTTTTGCTGGCTGTTACT
GCTGTAAAATCTGTAGGTACTTAGCCATGGTTTAGACTCATCTGTGAGCGCCAGGACTCC
CCTCTTTACGCAGATGGTGACTGAAGCTGGTTCCGAGATCGGCATATGTAGCTGGTAGA
GGTGTGGATATTGCATAGACCGAACCTCCGCAGGTCCGCATTCTCGAGTGAGAAACAGA
GATAAATTTTAAGGGGGTTCCCAAAAAAAAAA

SEQ ID NO: 105
TCCTTACCCGCCCCCGAACTCGCGAAGTTCCCAACTTCAAATCCTTCGGTCACCACCAC
CGCCACCTCCTCCGTCGCGGGAACGATGGGGATATTCGAGCCGTACAGAGCGGTCGGA
TACATCACCACCGGCGTCCCCTTCTCCGTCCAGCGCCTCGGCACCGAGACCTTCGTCAC
CGTCAGCGTCGGCAAGGCTTTCCAAGTGTACAACTGCGCGAAGCTGAGCTTGGTGCTTG
TTGGTCCCCAATTGCCAAAGAAAATACGGGCCCTCGCATCGTACCGTGAATACACTTTTG
CTGCTTATGGAAGTGACATTGGCATATTCAAGCGTGCTCACCAGTTAGCAACTTGGAGC
GGGCATACTGCTAAGGTTTGTTTGCTGCTGCTGTTTGGAGAGCACATATTGAGTGTAGAT
GTTGATGGCAATGCATACATATGGGCGTTCAAAGGAATGAATTACAACCTTTCCCCAGTG
GGACACATCTTGTTGGATAGCAATTTTACTCCTAGCTGTATAATGCATCCAGACACTTACT
TAAATAAGGTTATTCTGGGAAGCCAGGAAGGGCCACTGCAGCTTTGGAACATAAGCACT
AAGACAAAGCTCTATGAGTTCAAGGGATGGAATTCCTCTGTTAGCAGTTGTGTTTCATCA
CCTGCTTTAGATGTTGTTGCGGTTGGCTGTGCTGATGGCAAGATTCACGTTCACAATATT
CGCTATGATGAAGAATTGGTTACATTTTCTCATTCGATGAGAGGTTCTGTGACCGCCTTA
TCTTTCAGCACAGATGGGCAGCCTCTTCTAGCTTCTGGTAGTTCATCTGGTGTTGTCAGC
ATATGGAATCTCGATAAAAGAAGGCTCCAGTCAGTCATAAGAGATGCTCATGACGGTTCT
ATAATTTCTCTCCACTTCTTTGCTAATGAGCCTGTGCTAATGAGTTCATCTGCAGATAACT
CAATTAAGATGTGGATTTTCGACACAAGTGATGGAGATCCTCGTCTATTACGCTTTCGAA
GTGGTCATAGTGCTCCTCCACTCTGCATAAGGTTCTATGCTAATGGTAGGCATATACTAT
CAGCTGGTCAGGATCGTGCCTTCCGGCTTTTCTCGGTTGTCCAGGATCAGCAAAGTCGA
GAGCTATCTCAACGTCATGTGTCCAAGCGAGCAAAAAAATTGAAGTTAAAGGAAGAAGA
GATAAAGCTGAAGCCTGTTATAGCATTTGATGTTGCTGAAATTAGGGAGCGGGACTGGT
GCAACGTAGTGACAAGTCATATGGATACCCCGCAGGCATACGTGTGGAGACTTCAAAAT
TTTGTCATAGGGGAGCATATTCTGAGACCATGCCCCAACAAGCCCACACCTGTTAAGGC
TTGCATGATTAGTGCATGTGGAAATTTTGCCATTTTGGGAACAGCAGGTGGTTGGATTGA
GCGATTCAACCTTCAATCAGGAATTAGTCGAGGAAGTTACATAGACCAGTTAGAAGGAAC
AAACAGTGCTCATGATGGTGAAGTGGTTGGAGTGGCTTGTGATGCCACAAATACTCTCAT
GATAAGTGCCGGTTATGCTGGGGACATCAAGGTCTGGGATTTCAAGGGACGTGAATTGA
AATCTAGATGGGAAATTGGTTCTTCTTTGGTAAAAATATCTTACCATCGTCTGAATGGTCT
TTTGGCTACGGTTGCAGATGATTTTATCATTCGCCTTTTTGATGCTGTCGCCCTAAGAAT

GGTCCGAAAGTTTGAAGGTCATACTGATCGAATTACAGATTTGTGCTTTAGTGAGGATGG
GAAATGGCTCTTGTCTTCCAGTATGGATGGCAGTCTAAGGATTTGGGATATTATTTTGGC
CAGACAAGTAGATGCTGTATTTGTTGACGTTTCCATCACAGCTTTATCTCTGTCACCAAAT
ATGGATATTTTAGCCACGACACATGTTGATCAAAATGGGGTCTTTCTCTGGGTTAACCAA
TCAATGTTTTCTGGAGATTCTGACATCAATTTGTATGCAAGTGGGAAAGAAGTTGTGACC
GTCAAGTTGCCATCAGTATCATCAGTGGAAGGTTCTCAAGTTGAAGAATCTAATGAGCCT
ACTATTAGACATTCAGAGTCTAAAGATGTCCCTTCCTTCCGGCCTTCGCTTGAGCAAATT
CCAGATCTAGTTACCCTCTCACTTTTGCCAAAGAGCCAGTGGCAGAGCTTGATTAATCTA
GACATTATAAAGGTTCGCAATAAGCCAGTTGAGCCCCCCAAGAAACCTGAGAAAGCTCC
TTTCTTTTTGCCTTCTATTCCATCTCTTTCTGGAGAAATACTATTTAAGCCAAGTGAGATGT
CTGATAAGGGAGACATGAAAGCTGATGAAGACAAATCTAAGATAACACCTGAAGTGCCAT
CATCACGGTTTCTGCAACTGCTTCATTCATGTTCAGAGGCGAAAAACTTTTCACCCTTTAC
CACCTACATCAAAGGGTTATCTCCCTCAACTTTGGATTTGGAACTTCGGATGCTTCAGAT
AATAGATGATGATGCTGTTGACGCTGATGCTGATGATCCACAAGATGTTGACAAGAGACA
AGAGTTGCTTTCCATTGAACTACTTATGGATTACTTTATACATGAAATATCATGCCGGAGT
AATTTTGAGTTTGTACAAGCTCTAGTCAGGTTGTTTTTGAAGATACATGGGGAGACAATTA
GACGCCAGTCAGTGTTGCAGAATAAAGCAAAGGTCCTTTTGGAGACTCAGTGCTCTGTA
TGGCAGAGAGTGGATAAGTTATTTCAAGGTGCTAGATGCATGGTTGCCTTTCTTAGCAAC
TCACAGTTTTAGCAAAAGGAAGATGTTATAGACGGAAAAAGTAGCTGGATCAACACTGAG
GCATGAGTTTAAATGTTGGATGGGTTCCTCTTAGAGCTATAAATTCGCTCATTTTGTAACT
GAGCAATCATCCTGTATCACTTCTTTCAGAAATTTTCAATGGAAGCCTTTTTTTAAAAAAAA
AA

SEQ ID NO: 106
CCGTTTGCGCGTGCTTTCTCTCTCTCTCTCTCTCTCCTCTGCTCACTTCTTCGGGAAC
TTCGATTCCGGCCATGGAGGAGACGAAGGTGACGTGCGGGTCGTGGATCCGCCGCCC
GGAGAACGTTAACCTCGCCGTGCTGGGGAGGTCGCCGCGGCGCCGCGGCTCCGCCGC
GCTCGAGATCTTCGCCTTCGATCCCAAGTCCACTTCCCTCTCCTCATCCCCGCTGGTTG
CTCATGTGATCGAGGAAATTGAAGGCGATCCGTTGGCGATCGCGGTTCACCCCAACGGA
GAGGATATCGTGTGCTTCGCGAGCTCCGGTAGTTGCTTATCATTTGAATTATCTGGTCAA
GAATCAAATTTGAAGCTGTTGACCAAGGAGCTTCCTCCTCTTAGAGGGATTGGTCCTCAA
AAATGCATGGCCTTTAGTGTTGATGGATCTAGATTTGCCACTGGTGGAGTGGATGGTCG
TCTAAGAATCCTGGAATGGCCTAGTCTGCGCATCATTCTAGATGAACCAAAAGCACACAA
AAGCATTAGAGACCTGGATTTTAGTCTAGATTCAGAGTTTTTGGCTACGACATCTACTGAT
GGATCGGCTAGAATATGGAAGGCCGAAGATGGTTTGCCTTGCACTACTTTGACTCGCAG
ATCTGACGAGAAGATTGAATTGTGTCGGTTCTCTAAGGATGGAACTAAACCCTTTCTATT
CTGTACTGTTCAAAGAGGTGATAAAGCTGTTACTGGTGTTTGGGACATAAGTACATGGAA
CAAGATTGGGCACAAACGACTCCTCAGAAAGCCTGCTGTTGTAATGTCCATCAGTTTGGA
TGGAAAATATCTTGCTCAGGGAAGCAAAGATGGTGATATGTGCGTTGTTGAAGTGAAGAA
AATGGAGGTTAGCCATTGGAGCAAGAGGCTACACCTGGGAACCTCCCTTACATCACTAG
AATTCTGTCCCATTGAAAGGGTTGTGATTACTACTTCTGATGAGTGGGGTGTGTTAGTGA
CCAAGTTAAATGTTCCTGCCGATTGGAAAGCATGGCAAGTGTATCTACTACTTTTGGGTT
TGTTTTTGGCGTCTCTTGTTGCGTTCTACATCTTCTATGAAAACTCCGACTCATTCTGGGG
CTTCCCTCTTGGAAAGGATCAACCTGCAAGACCTAAGATCGGCAGTGTTCTGGGGGATC
CCAAGTCTGCTGATGATCAAAATATGTGGGGTGAATTTGGGCCATTGGATATGTGAGCAA
ACCGCCGGTTGTGTCGTGCTGCAAGTCTCCAGAAGGAGCTGCAGCAGCCAGTGGTCTG
AAGAACTCCATTTGATTATTGTTACACAAAGGCAATTTGTTTATAGGTCTTTTTGGAACAA
AATGTCCGCCTGGGCGTGAACAAGCGGCATGGGACTGACACCGTCGCTTTCATTAAAGA
AGTCTCATCAAAGTCGAACTAATACGAATTTTGATTTTAAAAAAAAAA

SEQ ID NO: 107
CTCTCATGAACCTCACGAAGCGCGCGCACACCAGCAGCTCCGCGAGCACGAGCGGCGA
GCACCACGGCAAGAGCCAGAGCTCCGCATTTCGCAGCGCAAATGGCGGACCCCGTAGA
GCACCAGCACCAGCAGCACCAGCAGCACCAGCTTCAGCAGCAGCGGCGGCGCGGATG
GCGGATTCAGGGCGGGCAGTACCTCGGGGAAATCTCCGCTCTCTGCTTCCTCCACCTC
CCTCCTCCTCCCCTCTCCCTCTCCTCCTCGCCGGTACTCTCTCTCTCTTCTGGACTCGAC
TCGGAATCGCGAGACCGACCGGCGTGCTCCTTCCGTTTCCCGTCTGCAGGTTCCGGCT
CGCAGGTGTCGCTGTTCGACCTGGCGTCCGGGGCCATGGTGCGGACGTTCTACGTCTT
CCGGGGGATCCGCGTGCACGGGATCGTCCTCGGCTGCGCCGATTTCCCCGGCGGATC
GTCCTCGTCTTCTTCGACGCTCGATTACGTCATCGCCGTGTACGGTGAGAGGAGAGTGA

AGCTGTTCCGCCTGTCCGTGCGGCTCGGAAGGGGAGCTGGCGAGGGGAGCGGAACTG
TGCTGTCCGCGGATTTGGAGCTGGTTTCCGCGGCGCCGAGGTTGAGCCACTGGGTTAT
GGATGTTCGCTTCTTGAAGGAGAATGGAACTTCTGAGGATGAGCTGCAGAGGTGTCTTA
CCGTTGCTATAGGATGCAGTGACAACTCCATACGCCTTTGGGACGTCGATAAATGTAGC
TTCGTTCTTGCAGTTTCTTCCCCTGAGAGATGCCTTCTGTACTCCATGCGGTTGTGGGGT
GACAATCTTGAAGATCTCCAAGTTGCATCTGGAACAATTTACAATGAGATTTTGATCTGGA
AAGTGGTTCCCAATCATGATGCTCCATCTTCAAATGAGCTCACAGAAGAAGGCCTGACAA
ACTCTTGTGCTGGCAACAGCGTCCACGAATGTCTTCGCTATGAAGCCTATCACATCTGTA
GACTTGTTGGTCATGAAGGTTCAATATTTAGAATCGCATGGTCCTCTGATGGCTCCAAAC
TCGTCTCTGTATCTGATGATCGTAGTGCACGTATTTGGGAAGTTCATTGCAAGGTGCAAT
ATTCTGAAGATGCGGGGGAGGTTGGATTGCTTTTTGGACACAGTGCTCGAGTTTGGGAT
TGCTATATATCTGATAATTTGATTGTTACTGCTGGCGAGGATTGCTCATGTCGTGTGTGG
GGACTGGATGGACAACAGCATGATGTCATCAAAGAGCATATTGGAAGGGGTATATGGCG
GTGTCTCTATGATCCATGGTCTTCACTCCTTGTCACTGGTGGTTTTGACTCTGCAATTAAA
GTGCATAAACTGGATGCTTCTTTAGCTGAGGCTTCCGCAAAACAGTCCAACATAAAAGAC
TTGAGTGACGGAACTGAGCTATTTACTACACATCTTCCGAATTCATCAGGTCATAGCGGA
CACATGGACAGCAAAAGTGAGTATGTCCGATGCTTGAGCTTCTCATGTGAAGATGTGAT
GTATATTGCCACTAATCACGGTTATCTCTATCATGCTAAATTATGCAATGATGGCGATCTA
AGGTGGACTGAACTTGCTCAAGTAAGTAATGAGGTGCAAATTATTTGTATGGAACTGTTG
CCCTCAAATCCATATGATCCTCGGATTGATGCTGATGATTGGGTTGCTGTTGGAGACGGT
AAAGGATGGACAACAGTTGTCAGAGTTGTGAAGAATTCTGATTCTCCTAAAGTGAGCACC
AGCTTCTCTTGGGCAGCTGAAATGGATCGACAGCTCTTGGGAATCCATTGGTGCAAATC
ACTAGGACATAGGTTCATTTTCACTGCTGACCCTAGAGGAGCCTTGAAACTTTGGAGATT
CTTCGAGGTTTCACAATCTAGTTCTCTTTATCCAGAAAACAGTCCACGGATTTCCTTGATT
GCAGAGTTCAAATCAGATTTGGGTGCTCGAATTATGTGCTTGGATGTGGCATTTGAGAGC
GAGCTGCTGATCTGTGGGGATCTACGGGGTAATCTGGTTCTGTTCCCTTTATTGAAGGA
CCTGTTGCTGGATACCTTTGTCGTGTCAGCAGCTAAAATCTCTCCAGTAAACCATTTTAAA
GGAGCCCATGGCATCTCAGCAGTTTCCAGCATTTCAGTGGCTCACATGAGTTTCAATCAC
ATTGAATTACGTTCTACTGGAGCTGATGGATGCATATGCTACATGGAATATGACAAAGGT
CTGCAATCTTTAAATTTTGTAGGGATGAAGCAGGTGAAAGAATTAAGCATGATTGAATCT
GTTTCCACTGAAAATGAATCTACCGGTTACAGAACAAGTGGCAGTTATGCATCTGGTTTT
GCATCCACAGATTTTATAATATGGAACCTAGTAACTGAGGCTAAGGTCCTGCAAGTTTCA
TGTGGTGGTTGGCGGCGTCCACATTCTTATTATCTGGGTGATGTACCAGAGATGAAGAA
CTGCTTTGCTTATGTCAAGGATGATATTATTTACATCCGTAGACACTGGATAAAGGACTCA
AAGGACAAGATACTCCCTCAAAATCTACGTTTGCAGTTCCATGGAAGGGAGGTGCATTCT
TTATGCTTTGTCACTGGAGATTTCCAGCTCCGAAAAAATAAACAATCAAGTTGGATTGTGA
CTGGTTGTGAGGATGGGACAGTGAGGCTGACTAGGTATACTCAGTGCACTGACAATTGG
TCTTCATCCAAATTACTTGGAGAGCATGTTGGTGGATCAGCTGTCAGATCAATATGTTGC
GTTTCAAATATCCATACAACCTCATCAGGTACTAGCGTGTCTGATGTAAAGGGTATAGAG
AACCTCCCAAAGGATATAAAGGGAACACTTATGGAGGATGAATGTAATCCATCATTGCTT
ATTTCTGTTGGTGCAAAGCGTGTTCTGACTTCTTGGCTGCTGAGAAGAAGAAAACAGGAT
GGGAAAGAAGATGACGTTACTGATTTACAAGAAGCCGAAAATAGTTCATTGCCTTCATCA
GCTGGATCCTCTACATTTTCATTCCAATGGCTTTCCACGGACATGCCAGTTAAGTACTCA
GTGCCTTCCAAAAAATCAGGAAGCATTAAAAAGCTAATCGGTGTCTCTGACACCAATGTA
AGATGTAAATCACTTCTGCCTGATAGTGAGGCTCTGCAGTCAAAAGTATCTGCAGTTGAT
AAGAATGAAGATGACTGGAGATACCTAGCTGTCACTGCTTTTCTCGTTAGACATTCTGGT
TCTAGGTTAATTGTCTGTTTTATTATCGTTGCTTGTTCAGATGCTACACTTGCAATACGAG
CTCTTGTCTTACCCTATCGTCTATGGTTTGATGTTGCTCTGATGGTTCCTCTATCATCACC
AGTTCTGTCATTGCAGCATGTCATCATTGGAAGATGTCAATTACCAGATGAAAATGTGCA
GATTGGGAATGTATATGTCGTGATTAGCGGAGCCACTGATGGGAGCATTGCTTTCTGGG
ATCTGACTGAAAGTGTTGAAGCCTTTATGAGGCGGTTGTCAAATATCCATCTAGAAAAGT
TCATGGATTGTCAGAAACGGCCACGGACTGGGAGAGGAAGCCAGGGTGGAAGATGGTG
GAGATCCCTGAGCAAAATTGCTTGTAAGGAACAACCAATCAATGATCCTGTTACTGCGAA
AGCTATAAAGGAACTAAACAGGAAATTGACTGGCGGTGTTGCCTGCGGGTCTTCATCTT
CTATGCTAGATGCCTCCCCTGAGTTAGATAGCAATGCAGCTAATTCCTCATTTGAAATTAT
TGAAGTAAATCCTTTCCATGTTCTTAATGGTGTTCACCAATCTGGCGTGAATTGCCTTCAT
GTTTGCGAAACAAAACATGGCCAAAGTTCCGATGGTCGTTTTCTGTACCAGTTAGTCAGT
GGTGGTGATGATCAAGCACTTCATCTTCTTAAATTTGAGGTGTTGGTGCAGCCCCCAGTC
CAAGTTCCAGATGTTCCAAATTCAGACATCAGAAATTCTATACTCGTTGAGGAATTTCTTC
TTGATGAGCAGAACCAGAAAACCAAGTGTACAATTGAATTCATTTCTCAGGAAAAAATTG

CTTCTGCTCACAACTCTGCTGTAAAAGGCGTTTGGACAGACGGCACTTGGGTTTTCTCGA
CTGGTCTTGATCAGCGTGTCAGGTGCTGGATTAGCAAGGATCGTGGTACACCAACGGAG
CTTGCTCATTTTATCATTAGTGTGCCAGAGCCAGAAGCATTGGATGCAAGATCCATTTGC
TGGGACCAGTACCAAATAGCAGTAGCTGGAAGAGGGATGCAGATGATCGAATTTCATGT
GCCTTCTTCTGAGATTCGGTAACGACAATCAAAGGTTTTCCCTTTTTTTCTTTCTTTGCCC
ACTAAAAGTTCTTCAGGGCAAAGCCAAAGGTTTATCCCTCATTGGATTTGATATATAAACT
GAGAGTGTCTTGCACACCATTAAAATGGCCCATCAACAGTGAGTTGCATCTAAAAAAAAA
AAAAAAAAAAAAAA

SEQ ID NO: 108
CTGATTTGTCCAATTTGAACATCACATCTTTGTATTCGTACTTGACACTCCTTCGACATGC
CATATAAACTTTCAGCAACGCTATCGAACCACTCATCTGATGTTCGTGCTGTCGCGTCGC
CTTCAGATGATTTGATCCTTTCTGCGTCGCGTGATTCTACTGCGATTTCATGGTTCAGAC
AATCGCCATCTTCCTTCACCCCCGCCTCCGTCATACGAGCAGGCTCCAGGTTCGTCAAT
GCGATAGCCTACCTGCCACCGACACCGAGAGCCCCTCAAGGGTATGCGGTAGTGGGCG
GTCAAGACACCGTCGTTAACGTCTTCGCTCTGGGTCCCGGCGACAAGGAAGAGCCCGA
ATACACACTCGTCGGCCATACCGACAACGTCTGTGCCCTCAGCGTGAACTCGGACGATA
CCATCATTTCTGGATCATGGGACAAGACAGCCAAAGTTTGGAAGGATTTCGCTTTGGTTT
ATGACTTGAAAGGTCACCAACAGTCTGTCTGGGCTGTGCTGGCTATGAACGAAAAAGAG
TTTCTGACCGCTTCTGCGGATAGAACCATCAAGTACTGGGTCCAGCACAAAACAATGCA
GACATATGAAGGGCACCGAGACGCTGTGCGGGGGTTGGCGCTAATTCCTGACATCGGT
TTCGCTTCGTGTTCCAATGACAGCGAAATTCGGGTGTGGACAATGGGAGGGGACGTGGT
GTATACGTTATCGGGACACACGTCTTTTGTATACAGTCTGTCTGTTCTGCCGAACGGCGA
TTTGGTCTCTGCAGGAGAAGATCGCTCTGTACGGGTGTGGCGTGACGGAGAGTGCTCT
CAGGTTATCGTTCATCCAGCCATTTCGGTGTGGGCGGTATCTACGATGCCGAACGGTGA
CATCATCAGTGGATCTAGTGATGGAGTAGTTCGGGTCTTCAGTGAATCTGAGAAGCGTT
GGGCTACCGCCAGTGAACTCAAGGCACTTGAAGACCAAATTGCAAGCCAATCCCTTCCG
TCACAGCAAGTTGGTGACGTCAAAAAGACCGATTTGCCAGGGCCTGAAGCCCTTTCCGT
TCCAGGGAAAAAGGCTGGGGAAGTGAAGATGATTCGGAGTGGAGACGTCGTCGAAGCA
CATCAGTGGGATAGCTTGGCCTCTAGCTGGCAGAAAATAGGCGAGGTTGTGGATGCAAT
TGGCTCAGGACGTAAACAACTGCACGACGGCAAGGAGTATGACTATGTGTTTGATGTTG
ATATCCAAGAAGGTGCTCCTCCATTGAAACTACCTTATAATGTTTCCGAAAATCCCTATAC
CGCTGCTCAGCGTTTCCTCGAACAGAACGACTTGCCTACGGGCTATCTCGATCAAGTGG
TGAAGTTTATAGAACAAAACACTGCAGGAGTCAAGCTTGGCAATGATGGTTATGTTGACC
CATTTACAGGGGCATCTCGATACCAGCCAGCGACACAGTCCACTTCGAATACGGCGTCG
TCGTCTTACATGGATCCTTTTACGGGTGGATCCCGCCACATTGCAGAGTCAGCTCCTTCC
AATGTACCACAGGGCTCTCATGCAACTGGGATCATTCCGTTTTCTAAGCCTATATTCTTC
AAACTGGCCAATGTATCTGCAATGCAAGCCAAGATGTTCCAATTTGACGAAGTTCTTCGC
AATGAAATATCCACAGCGACTCTCGCGATGCGTCCTGATGAAGTGATCATGGTCAACGA
GACATTTACATATCTTTCTAAAGTGGTTACTTCCACATCCTCGGCACGGACTTCTCTCGG
ATGGATCCATATTGAAACAATCATGCAAATACTGGACAGGTGGCCTGTTCCTCAAAGATT
TCCGGTCATAGACCTTGGTCGTCTGGTAACGGCATACTGTATGAATGCTTTTTCTGGTCC
CGGCGACCTCGAAAAGTTCTTCAGTTGCCTATTCAGGACATCGGAGTGGACTTCCATCA
CCTCTGGAAGCAAGGCGCTGACCAAGGCGCAAGAAACAAATGTACTACTTCTATTCCGC
ACGATTGCGAACTCGTTAGACGGTGCACCTTTAAACGACATGGAGTGGATCAAGCAGAT
ATTTAGGGAATTGGCACAGACACCGCAATTGGTTCTCAACAAGTCCCATCGGTTGGCAC
TAGCCTCCGTTTTGTTCAACTTTTCGTGCATCGGCCTCAAAGGCCCTGTCCCTGCGGAC
GTGAGGACATTGCACCTGACTATAATTTTGCAGGTGTTGCGATCTCCAAACGATGACCCC
GAGGTTGCTTATCGGACCTGTGTTGCCCTGGGAAACATGCTCTACTCAGACAAAACGCG
AGGCACACCGCGAGACGCACAATCGCCATCACCAACTGAACTCAAGAGTGCCGTTGCC
GCAATCAAAGGGGGATTCTCGGATCCAAGGATAAACGATGTTCATAGAGAAATCATGTC
CCTCATCTGAGCGCGTTCAACCTGTTGCAAAGGTTTGGGGGCTGTACAGCGTGTATTTC
TTGTTACGATACTTGAGGGGTTAGAGGCACCTACGAATTAGGAATGTTGGATTGTCGCTG
TCAAGCGCTCGACAGCATCAAAAAAAAAA

SEQ ID NO: 109
TTTTCCAAGTCAAAATCTCTCTCTCTCTCTCTCTCTCTCTGTGCAGCAAGGCACGGCA
GCTCTGGGGGCGGACGCCGGCGGAGCAAAGCGGCGACCAGGGGAGCGGCTGCAAGA
CGACGTCCACGATCTCCATCCTCTGTAAGAATTACTTGAAAGATGCCTCCACAGAAGATT
GAAAGTGGACACAAGGACACGGTCCATGATCTGGCGATGGACTACTATGGTAAGAGACT

GGCGACAGCGTCGTCGGACCACACCATTAACGTTGTTGGAGTCAGCAGCTCGGGATCT
CAGCATCTTGCCACGCTGATCGGCCATCAAGGACCCGTTTGGCAGATATCTTGGGCTCA
TCCGAAATTCGGGTCTTTGCTTGCTTCGTGCTCGTATGACGGACGGGTTATCATATGGA
GAGAAGGTAATCCAAATGAGTGGACGCAGGCACAAGTATTTGAAGAGCACAAATCATCG
GTCAATTCCGTTGCTTGGGCTCCTCATGAGCTCGGCCTTTGTCTGGCTTGCGGTTCATC
CGATGGGAATATTTCGGTGTTCACTGCTAGACAGGACGGCGGTTGGGATACTTCAAGGA
TCGATCAGGCTCATCCGGTCGGGGTCACCTCCGTGTCGTGGGCCCCGTCGACTGCCCC
TGGTGCACTTGTTGGTTCTGGCATGATGGAACCTGTTCAGAAGCTCTGCTCAGGTGGTT
GTGATAACACCGTGAAGGTGTGGAAGCTCTACAACAGGGTCTGGAAGTTGGACTGCTTT
CCGGTGCTTCAGATGCACACTGACTGGGTGAGGGACGTTGCCTGGGCACCCAACTTGG
GCCTTCCGAAATCGACCATCGCAAGCGCATCGCAGGACGGGAGGGTCATCATATGGAC
CTTGGCCAAGGAAGGGGATCAATGGCAGGGGAAGGTTTTGTACGATTTTAGGACTCCGG
TTTGGAGGGTCTCGTGGTCCCTGACCGGTAACATTTTGGCGGTGGCAGACGGGAATAAC
AATGTCTCCCTTTGGAACGAAGCCGTTGATGGTGAGTGGATCCAGGTTTCAACAGTCGA
GCCATAGGATTCAGCTTGTGGTTGTCCTGTCGAACTGCTTTTAGCTGCGCAATGTGCATA
GAGAACGATTTGCCAACATAGTAGAACCCGCGAACATGCTTGGGTGTACTTCATAGATG
AATTTATGTGCTTTTTGGGCCGGTATAGATGCTTTAAATTGTTCTTGCTTTGCGGCTAGAT
ATCCTTATGAATGAAGTTTGGATGATAAGTGGCGCCAGACTTTCTACTCACCCTTTTTTGT
CAGCCATGTTTTGAGCGATGCGGTGACCAGTTTGGCCTGAATAGTTGGTCTTAATGGCTT
GTGGATTTGCACCGGCTGTTTACCATTGAAACTTGTATGACCTATTGCTGATGAACCTCT
TAATCTGCTCAGAAGGGTGATCGTTTAGAAAAAAAAAA

SEQ ID NO: 110
GCTTCCATGGCGCACCCCGAATTCTACTGATCGAATTCGCCTTCTCTTCTCACTGCAAAA
CCCTAAAACGCACGCCCTCTCCCTCTCCCTCTCCCTCTCCCTCTCACAGCTGCCAAAAT
GTCTGCTCCTATGCTGGAAATCGAGGCCCGCGACGTGGTCAAGATCGTGCTCCAGTTCT
GCAAGGAGAACTCGCTGCACCAGACGTTCCAGACCTTGCAGAGCGAGTGCCAGGTCTC
CCTCAACACCGTCGACAGCATCGAGACGTTCGTCGCCGACATCAACAGCGGCAGGTGG
GACGCCATACTGCCTCAGGTCGCCCAGCTTAAGCTCCCCAGGAATACGTTGGAGGATCT
GTATGAGCAGATTGTATTGGAAATGATTGAACTTCGTGAGTTGGATACTGCTCGTGCAAT
TCTAAGGCAAACTCAGGCAATGGGTGTCATGAAGCAGGAGCAACCTGAAAGATACTTGC
GACTTGAGCATCTCCTTGTTAGAACATATTTTGATCCCAATGAGGCTTACCAAGACTCCA
CAAAAGAAAAACGACGTGCACAAATTGCCCAAGCTCTTGCTGCGGAAGTTACTGTAGTG
CCACCTTCAAGATTGATGGCCCTAGTGGGGCAGGCGCTCAAGTGGCAGCAGCACCAAG
GATTACTACCTCCGGGAACCCAATTTGATCTATTTCGAGGAACTGCTGCTATGAAGCAAG
ATGTGGATGATATGTACCCTACAACTCTTTCACACACCATTAAGTTTGGAACCAAAAGTCA
TGCAGAGTGCGCTAGGTTCTCGCCAGATGGACAGTTTCTTGTTTCTTGCTCTGTCGATG
GATTCATTGAGGTCTGGGATTACATGAGTGGGAAGCTCAAAAAGGATCTTCAGTATCAG
GCTGATGAGACCTTCATGATGCATGATGATCCCGTTCTTTGTGTTGATTTTAGTAGAGAC
TCGGAGATGCTTGCTTCTGGGTCGCAAGATGGCAAGATCAAAGTTTGGCGAATAAGAAC
AGGTCAATGCTTGCGTCGTCTTGAACGTGCACATTCTCAGGGTGTCACAAGTGTCCTCTT
TTCCCGTGATGGCAGTCAGTTACTCAGTACCTCCTTTGATGGCTCAGCCAGAATCCATG
GCCTTAAATCTGGGAAGCAGCTGAAAGAGTTTCGAGGCCACTCATCTTATGTGAATGATG
CAATATTCAGCAATGATGGCAGTCGTGTTATTACCGCCTCAAGTGATTGTACTGTAAAGG
TCTGGGATGTTAAGACTTCAGACTGTCTTCAAACATTTAAGCCTCCACCTCCATTGAGGG
GAGGTGATGCTTCTGTTAATTCTGTTCATCTCTTCCCAAAGAATGCTGATCACATCGTTGT
TTGCAATAAGACGTCATCAATTTATATCATGACTCTACAGGGACAGGTGGTGAAGAGTCT
TTCATCTGGTAAAAGAGAGGGTGGAGATTTTGTGGCAGCCTGTGTATCACCAAAAGGTG
AATGGATTTACTGTGTCGGTGAAGACAGGAATTTGTACTGTTTCAGCTGCCAGTCTGGGA
AATTAGAGCATCTTATGAAGGTTCATGAAAAGGATGTAATCGGCGTAACACACCATCCCC
ACCGTAATCTCGTTGCAACATACAGCGAAGATAGCACGATGAAACTATGGAAGCCTTGA
CGTATTATTTCTCATCTTGCGACTTTGCAATTGTTAGACTGTTGTAGATGTAATTGTCATTT
CATTTATTGACCGGTGAAAGATTCCGAGTGTATTTTAGGGAAAGAGCAGTTATAATCTCC
TCATTGGCCACCAAAAAAAAAA

SEQ ID NO: 111
GAATCCCCGAAATCGCACTGGAAAATTCCACTGCGCTGCTTCTCCGAATCGCCCCCATG
GATCTCCTGCAATCTTACGCGGAGGATAACGACGGCGATCTCGGCCGTCACTCCTCGCC
GGAACCCTCGCCGCCCCGCCTCCTCCCCTCCAAATCCGCCGCCCCGAAGGTTGACGAC
ACCACGCTCGCCCTCACCGTCGCGCAGACCAACCAAACCCTAGCCCGCCCCATCGACC

CGTCCCAGCACGCCGTCGCTTTCAACCCCACCTACGACCAGCTCTGGGCCCCAATCTG
CGGCCCCGCCCACCCCTACGCCAAGGACGGCATCGCCCAGGGCATGCGCAACCACAA
GCTCGGCTTCGTCGAGGACGCCGCCATCGGCTCCTTCCTCTTCGACGAGCAGTACAAC
ACGTTCCAGAGGTACGGGTACGCCGCCGACCCCTGCGCCTCCACGGGGAACGAGTAC
GTCGGCGACCTCGACGCGCTCAAGCAGAACGACGGCATCTCGGTGTACAATATCCGCC
AGCAGGAGCAGAAGAAGTATGCTGAAGAGTACGCGAAGAAGAAGGGCGAGGAGAGGG
GCGAGGGCGGCAGGGAGAAGGCGGAGGTGGTTTCCGATAAGAGCACATTCCATGGTAA
AGAGGAGAGGGATTATCAGGGCAGGTCGTGGATCGCGCCGCCCAAGGACGCCAAGGC
AACGAATGATCATTGTTATATACCCAAGAGATTGGTGCACACCTGGAGTGGGCACACCA
AGGGCGTGTCTGCCATTAGGTTCTTCCCCAAGCATGGTCACTTGATCCTCTCAGCTGGG
ATGGATACTAAGGTGAAGATCTGGGACGTGTTCAATTCGGGTAAGTGTATGAGGACTTA
CATGGGTCATTCGAAAGCTGTTAGGGACATATCGTTTTGTAATGACGGGACTAAGTTCTT
GACTGCCGGTTATGATAAGAACATCAAATATTGGGATACTGAAACCGGGAAGGTTATCTC
TACTTTTTCCACTGGGAAAATTCCTTATGTGGTTAAGTTGCATCCGGACGATGAAAAGCA
GAATATTTTGTTGGCGGGAATGAGCGATAAAAAGATTGTTCAGTGGGATATGAATACGGG
GCAGATTACACAGGAATATGATCAGCATTTGGGTGCGGTGAATACGATTACCTTTGTTGA
TGATAACCGGAGGTTTGTGACGTCAAGCGATGATAAGTCCCTTCGTGTTTGGGAATTTG
GGATCCCTGTGGTTATTAAGTATATAAGTGAGCCTCATATGCATTCTATGCCCTCGATCT
CGCTGCATCCAAATACGAATTGGCTTGCAGCACAGAGTCTGGACAATCAGATTCTTATTT
ATAGCACTAGGGAAAGATTTCAGCTCAATAAGAAGAAGAGGTTTGCTGGCCACATTGTG
GCTGGCTACGCCTGTCAAGTTAACTTCTCACCAGATGGTCGTTTTGTTATGTCTGGAGAT
GGAGAGGGTAGATGCTGGTTCTGGGACTGGAAGAGTTGCAAAGTCTTCAGAACTCTGAA
ATGTCATGAGGGGGTATGTATTGGGTGTGAGTGGCATCCACTGGAACAAAGTAAAGTAG
CAACCTGTGGCTGGGATGGCTTAATCAAGTATTGGGACTGATCTTCTTGCACCTTTCATG
CAAAGGCGTTTGATACTTCTTTGCTTCTCACATTGTGGACACTGAACAGCCTGTGGAACA
CAAGGCATACTTTCCGAGGTCAATCAGTTGGAATCTGCTCAGACTAGGGAATCTGCATG
ACAGTGGTGTACTTACAAACTCCTTTTGTAACTAGATGCAAAATCTTCACCAGCCTGTTGT
ACTTCAGTTTATTTTGCTAGCAGAACATTCACCACCAATTTTCCCAGTGTACTGCGAGAGT
GATGCTACATAAGTTTACTCTTGTGTCTAACTTTTCCAATTAGATCCAGATTGAGTTGACA
TAAAAAAAAAA

SEQ ID NO: 112
GGGGCAAACACCATGGAAAGATTTTGAACCTGCTAGAAAATGGCGGCAGTCATTCCGCA
GGCAATTAGGGTTTCAGCATACCTCCGGAAACGTCCGCAAGAATCAATCATCTCTCCCTC
GCCATGCGAACGGAGCTCGATTATAGGGTGTGATCGTCCCAATTTCGAAAGCAAAAGAC
TTGTCTTTAATCGAGAAATTGGGGCTTTTTAGCTGGCGGGTGGAAGTTAGGGTTTTGATT
CGGTGCAGAAGATGGAAAGCAACGGCAATTTGGAGCAGACTTTGCAAGATGGGAGGAT
ATACAGGCAGCTCAATTCGCTCATCGTCGCTCATCTTCGCGACCACAACTTCCCGCAGG
CGGCAAGTGCGGTTGCTCTAGCAACAATGACGCCCTTGAATGTTGAAGCCCCAAGAAAT
AGGCTTCTTGAGCTGGTTGCCAAGGGTCTTGCAGTGGAAAAGGGTGAACTATTGAGAGG
TGTTTCTCATGCTGGGACAAATGATCTGGGTGGATCAATACCTGCTTCATATGGCTTGGT
TCCAGCTCCGTGGACCGCTATTGATTTCAGTTCTCTGCGAGACACAAAGGGCATGTCCA
AGAGTTTCACTAAACATGAGACCAGGCATCTTTCAGACCACAAGAATGTTGCCAGATGTG
CAAGGTTTAGCACTGATGGAAGGTTTTTTGCAACAGGAAGTGCAGACACTTCAATTAAGC
TCTTTGAGGTCTCAAAAATAAAGCAAATGATGCTACCAGATTCTACAGATGGCGCTATAC
GAGCTGTTATACGGACATTTTATGACCATACACATCCTGTAAATGACTTGGATTTTCATCC
TCAAAATACTGTCCTGATATCTGCAGCCAAAGACCATACAGTAAAGTTCTTTGATTATTCA
AAGGCTACAGCAAAGAGAGCATTCAGAGTTATTCAGGATACTCACAATGTACGTTCAGTT
GCTTTCCATCCTTCTGGCGACTTTCTTTTGGCTGGGACTGATCATCCAATTCCACACTTG
TATGATGTCAACACATTCCAATGCTATCTTTCTGCAAATGTCCCAGAATTTGCAGTTAATG
CAGCGATAAACCAGGTAAGATATTCATCTAGTGGCGGCATGTATGTTACAGCATCTAAAG
ATGGTACTATACGATTTTGGGATGGGGCATCGGCCAACTGCGTTCGCTCCATTGCTGGT
GCACATGGGGCAGCTGAAGTAACTAGTGCCAACTTCACGAAGGATCAGAGATATGTACT
CTCTTGTGGGAAGGACTCTACTGTGAAACTCTGGGAAGTTGGCACAGGAAGATTGGTTA
AACAATATCTTGGAGCCACTCACATGCAGTTGCGATGCCAGGCTGTCTTCAACAATACAG
AAGAGTTTGTTCTATCCATTGATGAACCGAGCAATGAGATAGTGGTCTGGGATGCCATGA
CAGCAGAAAAAGTGGCAAGATGGCCATCCAACCATAATGGCCCTCCTCGTTGGATTGAG
CACTCACCCACAGAAGCAGCTTTTGTATCCTGCGGAACTGACAGATCGATCCGATTCTG
GAAGGAAACTCACTAGGGTGATCTATGGAGATGATCGTGATATTTGCAGCCTCGGACAT
CTTAGAAGTGATCACAATTTGATTCAGAAGGATGATGGCCCCATACACTGCGCCTTCCAC

AAATTCTTTCCGTCTTTGTAAGGAACTTGTGAGGTATTCGAGGGGACATTGTCAATAACTT
ATATCCTACCTCTAATGGAATT

SEQ ID NO: 113
ATTTCCGGTTTCCAGATCGATCGCCCAGCCCCGTCGCACCGCCCCCCGCCCTCTCTCTC
TCTCCAATCCGCGCCGAGGTGAATCGCGTTCATTGTTGTGATAAAGATGTCCAATTTCCA
AGGGGAGGATGGTGAGTATGTAGCAGATGACTTCGAAGCAGAAGATGGCGATGAAGAG
CTCCATGGCAGAGAATCGGCAGACCCAGAATCTGATGTTGATGAAATAGACACTCCAAG
TAATAGATTTACCGATACTACTGCTGATCAAGCAAGAAGAGGAAGAGATATTCAGGGAAT
TCCTTGGGAAAGGCTTAGCATCACCAGGGAGAAGTACCGGCGGACTAGGCTAGAACAAT
ATAAGAATTATGAAAATGTTCCTCAGTCTGGAGAGAAATCGGGGAAGGATTGCACAGTTA
CAGAGAAAGGTAACTCATTTTATGAGTTTAGACGGAACTCAAGATCTGTCAAATCTACTAT
TCTTCATTTCCAGTTGAGGAATTTGGTTTGGGCAACATCAAAGCACGATGTTTACTTGAT
GTCAAACTATTCTGTCGTCCATTGGTCTTCATTGACGGGCAAGAAGTCTGAAGTTCTTAA
TCTTGCAGGACATGTGGCACCAAATGAGAAACATCCTGGAAGCTTGTTGGAAGGATTTA
CGCAGACTCAAGTCAGTACTCTAGCAGTAAAGGATAGATTTCTAGTTGCAGGCGGGTTT
CAGGGAGAGCTCATATGCAAGTTTTTAGATCGACCTGGAATTAGCTTCTGTTCCAGGACA
ACCTATGATGATAATGCCATAACAAATGCGGTTGAGATATACGTCAGCCCCAGTGGTGG
AATTCACTTTATAGCATCGAACAATGACTGTGGAGTCAGAGACTTTGATATGGAAAACTTT
GAGCTGTCTAAACATTTTCGTTTCCCTTGGCCTGTGAATCATACTTCTTTGAGTCCAGATG
GGAAGCTTCTTGTCATAGTTGGCGATGATCCTGAGGGTATTTTGGTTGATGCAAAAACTG
GAAAAACAATCATGCCGTTGCGCGGGCATTTGGATTTCTCCTTTGCATCGGAGTGGCAC
CCAGATGGTGTCACTTTTGCTACGGGAAACCAGGACAAAACTTGTAGGATTTGGGACAT
ACGAAACTTATCTAAGTCAATCGCTGTTTTGAAAGGCAACCTTGGAGCCATACGGTCTAT
CCGCTATACATCCGATGGCCGGTATATGGCGATAGCAGAACCTGCAGATTTTGTCCATG
TCTACGACACGAAAACGGGGTACAAGAAAGAGCAGGAGATAGACTTCTTCGGCGAGATA
TCCGGCATGTCGTTCAGTCCCGACACAGAATCGCTCTTCATTGGAGTGTGGGACCGAAC
ATATGGTAGCCTCCTCGAGTATGGCCGGCGACGGAACTTCTCGTACCTCGACTGCCTCG
TCTGAATAGAATGTGCGGTCACTGTTAATAATTTTCTTTCGACTGTGCAGCTTTTTGGGTG
ATAAATGTGTGAGTACGTGGACAGAGCATGGTGCACAGGGGAGTTCCATTTTGTGTGGG
TGAAATGTTTGTATATATGGGTCGAGTAGAATCTTGATGTCTTATATGATTTTAGTTGGCA
CAG

SEQ ID NO: 114
GCTACTGCACTGCGCTCTCGGCGCCCTCCAAACGACCGAGGAAAGCTTCAGCTGAAAAA
AGCTTCGGCTACAATCCCTTCAGGTTCGGCTTGATATATATGGTTTTCTGCTAGCGACGC
GCGCCTGGAGTGTTGATTCGAGTAGGACCCGTATCGCGTCGATTGCCTAGCGTTCCGTC
AATTCGCCGTCATGGGTGTTGAAGAGGATCTCGAAGATTTGAATGCCCTCGCGGAGTCT
ACCGATGCCGCCGTCGATGGCCAGGCGGCTCTGGCTTCCGCCGTCGATAGCGTTACCC
TTCAGCCGGCCCCGCCAATACTTCCCCCCGTAATCCCTCCTCCCGCCGTGCCTGTGGTT
GCTCCCGTTCCGACAATTCCTCCGGTCCTTCGACCTCTGGCGCCTCTACCCATCCGCCC
TCCCGTTTTGAGGCCTCCCGCACCCAAAAGAGATGAGGCTGGAAGTAGCGACTCGGATT
CAGACCATGATGGTACCGCTGCCGGGTCGACTGCAGAGTATGAGATTACGGAAGAGAG
TAGGCTGGTGCGGGAGCGGCATGAAAAGGCGATGCAGGATCTGATGATGAAGCGGCGC
GGTGCTGCTCTGGCGGTGCCCACTAATGATAAGGCTGTGCGTGCCCGTCTTCGTCGGC
TCGGAGAACCTATGACCCTCTTTGGGGAAAGAGAGATGGAGAGGCGGGACCGGCTGCG
GATGCTTATGGCTAAATTAGATGCGGAAGGACAGCTCGAGAAGCTCATGAAAGCCCACG
AGGATGAGGAGGCTGCTGCTTCCGCTGCACCGGAGGATGTCGAGGAAGAGATGCTTCA
GTATCCATTCTATACTGAAGGATCGAAAGCTCTTTTCAATGCTAGAATAGATATTGCAAAA
TTTTCGATCACAAGGGCTGCACTCCGTCTTGAACGTGCAAGAAGGAGAAGGGATGACCC
TGATGAGGATGTGGATGCAGAGATAGATTGGGCATTGAAAAAGGCAGAGAGCTTGTCCT
TGCATTGCAGTGAGATTGGCGATGATCGACCACTTTCAGGCTGCTCTTTCTCCCACGAT
GGAAAGTTGCTTGCCACATGTTCCATGAGTGGAGTTGCTAAGTTATGGGACACATGTCG
CATGCCTCAAGTGAATAGAGTGTTGACATTGAAGGGTCACACAGAACGTGCTACTGATG
TGGCCTTTTCTCCGGTGCAAAATCATATAGCAACTGCTTCTGCTGACCGGACTGCAAAGT
TATGGAATACTGAAGGAACTATCTTGAAGACTTTTGAGGGGCATCTGGATCGCCTTGGTC
GTATTGCATTTCATCCATCAGGGAAGTACCTGGGTACGACTAGCTTTGACAAGACATGGA
GGTTGTGGGATATAGAAAGTGGTGAGGAACTACTTCTTCAAGAAGGCCATAGCAGAAGT
ATCTATGGGATAGACTTCCACCGGGATGGATCCTTAGTGGCATCTTGCGGACTGGATGC
TCTTGCACGTGTTTGGGACCTCCGCACCGGTAGAAGCATCCTTGCTTTGGAAGGACATG

TTAAGCCGGTTCTGGGAGTCAGCTTTTCACCCAATGGATATCATTTAGCCACCGGTGGT
GAAGATAATACCTGTCGTATTTGGGATTTAAGAAAGAAAAAGTCATTGTATACTATTCCAG
CTCATGCAAATCTTATATCAGAAGTGAAATTTGAGCCTCAAGAAGGATATTTTCTGGTGAC
TGCATCATATGATACTACGGCAAAGGTCTGGTCGGCCCGGGACTTTAAGCCAGTGAAAA
CTTTGTCAGTGCATGAAGCCAAAATAACATCAGTGGATATCACTGCGGATGCAAGTCATA
TAGTTACAGTGTCGCATGACAGAACAATTAAGCTCTGGACTAGTAACGACGACGTGAAG
GAACAGGCCATGGATGTTGACTGAATTCAATTTTTGTTAGCATGTGCTGTTTTAGCGGGA
GCACGGAAGTTGATCCTGTCAAGCGTTTGTAAACATTTTGGCAGATCTTTCCCCGTTTAA
AAGCGTTTCTAGCTTGCTTGTACTATAGTTTTTCAATTGGCTGCCCGTAAGGTATTCAAAA
GTACCAAACCATCCTAGCAGTAGTTTTTGCCCTTGACATGTTTTGGCCTGTGATCCAAGG
AAAGTTCTGATGTGGAGTGTCTTACGTCTTACGAACAAGGAGAGGACACTTTATGAGGTA
TCGCTTCCTCATTTTTCATATCTATCGTTATTGCCATTAGAATCAAAGCCCATTTGCTTGA
AAAAAAAAA

SEQ ID NO: 115
ATAGTTTCAACAGCCTGCGAAGCTCTTCCCCCCCTGAACCAGACACAGTCTCTCTCTCTC
TCTCTCTCTCTTCTCTCGCGCCACGGCAGCTGCCAAATGCGATGGTGAAGGCTTACCTG
AGGTACGAACCGGCGGCGGCCTTCGGGGTGATCGCGTCGGTGGAGTCCAACATCGCG
TACGACGCCTCCGGCAAGCACCTCCTCGCCCCGGCGCTCGAGAAGGTCGGCGTCTGG
CACGTGCGGCAGGGCGTCTGCACCAAAGCCCTAGCCCCTTCCGCCTCCTCCGCCGCCG
GACCCTCCCTCGCCGTCACGGCCATCGCCTCCTCTCCTTCGTCTCTGATTGCGAGTGGA
TATGCTGATGGTAGCATACGGATATGGGATTTTGAGAAAGGTTCTTGCGAGACGACACT
GAATGGCCATAAAGGGGCTGTCTCTGTACTCAGATATGGCAAGCTTGGATCTTTACTCG
CATCTGGAAGCAAGGACAACGACATTATATTGTGGGATGTAGTCGGAGAGACTGGTCTT
TACCGCTTACGAGGGCATCGCGATCAGGTTACCGACCTTGTCTTTCTAGATTCTGACAAG
AAACTCGTTAGTTCCTCCAAAGACAAGTATCTCAGAGTGTGGGATCTTGAAACACAGCAC
TGTATGCAGATTGTTGGTGGTCATCACAGTGAAATCTGGTCCCTGGACACTGATCCAGAA
GAGAGATATCTCGTCACAGGGTCTGCAGATCCAGAACTTCGATTTTACACTGTTAAGAAT
GATTCATCTGATGAACGATCTGAAGCAGATGCAAGTGGGGGTGTGGGCAATGGTGACTT
AGCTTCTCATAACAAATGGGATGTACTAAAACAATTTGGCGAAATTCAGCGACAAAGCAA
GGATAGAGTTGCAACAGTGAGATTCAACAAGAATGGGAATTTGCTGGCTTGTCAAGCGG
CAGGTAAACTTGTAGAGGTGTTCCGTGTACTAGATGAAGCTGAAGCAAAGCGTAAAGCA
AAACGCAGGCTTCATAGGAAGAGGGAGAAGAAAGGGGCAGATGTGAATGAAAATGGAG
ATTCCAGTCGTGGTATTGGAGAAGGACACGACACCATGGTGACGGTTGCTGATGTTTTT
AAGCTCCTCCAGACTATTCGAGCTAGCAAAAAGATTTGTTCTATTTCTTTCTGCCCTGTAG
CTCCTAAGAGTTCGTTGGCCACACTGGCATTGTCATTAAACAATAATCTTTTGGAATTTCA
CTCTATTGAAGCTGATAAAACTAGTAAAATGCTAACTATTGAACTACAAGGACACCGCTCT
GATGTCAGAAGCGTCACCCTTAGTTCTGATAATACCCTTCTTATGTCTACCAGTCACAAC
TCGGTAAAGATTTGGAACCCGAGTACAGGTTCCTGCTTGCGAACAATTGACTCTGGGTAT
GGGCTTTGTGGTTTAATTGTTCCTCAGAATAAGCACGCACTTATTGGAACAAAAGATGGA
GCCATAGAAATATTTGATGTTGGAAGTGGCACTTGTATTGAAGTGGTAGAAGCTCATGGA
GGCTCTATCCGGTCAATTGTGGCTATACCAAACCAAAATGGTTTTGTCACTGGCAGTGCA
GATCACGACATTAAATTCTGGGAGTATGGTATGAAGCAGAAACCTGGTGATAATTCCAAG
CACCTAACTGTGTCAAATGTCAGAACCCTGAAGATGAATGATGATGTTCTTGTGGTTGCT
GTGAGCCCAGACGCTCAAAAAATTGCTGTTGCACTATTAGACTGCACAGTGAAGGTTTTC
TTCATGGATTCTCTGAAGCTTATGCATTCCTTATATGGGCACAGGCTTCCTGTGCTATGC
TTGGATATCTCGTCAGATGGAGATCTAATTGTAACTGGCTCTGCAGATAAAAATCTGATG
ATATGGGGATTGGATTTTGGTGACCGCCATAAATCCATTTTTGCACATGGAGATAGTATC
ATGGCAGTGCAGTTTGTGGGCAACACACATTATATGTTTAGTGTAGGTAAAGATCGGCTT
GTAAAATACTGGGATGCTGACAAATTCGAGCTTCTTTTGACGCTTGAGGGACATCATGCT
GATATTTGGTGCCTTGCAATTAGCAACCGTGGTGATTTTTTGGTCACTGGATCTCATGAT
CGTTCAATACGCCGTTGGGATCGTACGGAAGAACCATTCTTCATTGAGGAAGAAAAGGA
GAAGAGGTTGGAGGAGATGTTTGAGTCTGACCTTGATAATGCATTTGGGAACAAGTATGT
ACCCAAGGAAGAAATTCCAGAAGAGGGTGCTGTGGCCTTAGCAGGGAAGAAAACTCAAG
AGACACTTTCAGCAACTGATTCAATTATTGAGGCGTTGGATATAGCAGAAGTGGAGTTGA
AACGAATTGCCGAACATGAGGAGGAGAAAAACAATGGAAAGACTGCAGAATTCCACCCA
AATTACGTGATGCTGGGGCTTTCTCCCTCTGACTTTATTCTTCGTGCTCTTTCGAACGTTC
AAACTAATGACCTTGAGCAGACATTACTGGCTTTACCTTTCTCAGATGCTTTGAAGCTCCT
ATCTTACCTGAAGGATTGGACAACATACCCTGATAAGGTTGAGCTTGTTTCAAGGATTGC
TACAGTGCTTTTACAGACACATTACAATCAATTAGTTTCAACCCCGGCTGCCAGGCCTTT

GTTGACTACACTAAAGGACATTCTTCACAAGAAAGTCAAGGAATGTAAGGACACGATTGG
ATTTAATCTTGCAGCAATGGACCATCTTAAGCAATTGATGGCCTTAAGATCAGATGCACT
CTTTCAAGATGCCAAAGTGAAGTTGCTGGAAATTCGCTCACAACTTTCCAAACGACTGGA
AGAAAGGACGGACCCGAGAGAAGCAAAGCGCAGGAAGAAGAACAGAAGAAATCCACT
AACATGCATGCCTGGCCATGAGGTCTTGGTAATGAGAATTCGCTAGTTGAAGAATTCGG
GAATTTTTTGCCACATCGTAACCATCATAGCACTTATCATCTAATTATGGTGAAAGGGAGT
TATATATATGTCAGTTTTGGCGGATGTTGAGTGTATAGGATGTGAAATAGCAAATGATAAT
CTCTCTTCTATCTTTTGGGCAATTGAACTTTTCATTCCCATAAAAAAAAAAA

SEQ ID NO: 116
GCCCCCGGACGTCTCCAGACCTCTCGCTGCGTCTGCACAGCATCTCCGTCGGAAGTTC
CCGTGCGACAATCGACATGGGCGGCGTGCAAGCCGAGAGAGAAGACAAGGACAAGGTC
TCTCTGGAGCTCACCGAGGAGATCCTCCAGAGCATGGAAGTCGGCATGACGTTCAGAG
ACTATAGTGGTAGAATCAGTTCCATGGATTTTCACCGGGCCTCGAGCTATCTAGTGACAG
CCAGCGATGATGAGTCTATTCGCCTTTATGATGTGGCAAGCGCAACATGTCTGAAGACA
ATTAATAGCAAAAAGTATGGGGTTGATTTGGTCTCATTCACTTCTCATCCGATGACTGTTA
TATACTCCTCAAAGAATGGTTGGGATGAGTCACTGAGGCTGTTATCCTTGCATGACAACA
AGTACCTGCGCTACTTTAAAGGTCACCATGACAGGGTTGTCTCCCTTAGCTTGTGCCCAC
GCAATGAATGCTTTATCTCTGGTTCTCTGGATCGCACTGTTTTACTCTGGGATCAACGAG
CTGAGAAGTGTCAGGGTCTTTTACGTGTACAAGGAAGGCCTGCCACAGCTTATGATGAT
CCGGGCCTGGTGTTTGCAATTGCTTTTGGAGGATGTGTTAGGATGTTTGATGCTCGCAA
ATATGAAAAAGGTCCTTTCGAAATCTTTTCTGTTGGGGGAGATGTGTCGGATGCAAACGT
TGTCAAGTTTAGCAATGACGGAAGGCTTATGCTTTTGACTACCACAGATGGGCATATACA
TGTTCTCGACTCATTCCGAGGCACACTGTTATACACTTTCAATGTCAAGCCAACATCAAG
CAAGTCCACCTTGGAGGCATCTTTTAGCCCTGAAGGAATGTTTGTCATTTCTGGTTCTGG
AGATGGCAGTGTCTATGCGTGGAGCGTAAGGGGCGGGAAAGAGGTCGCAAGTTGGTTA
AGTACTGACACGGAGCCTCCTGTCATAAAATGGGCCCCAGGAAACCTCATGTTCGCAAC
AGGATCGTCGGAATTATCATTCTGGATTCCTGATCTCTCTAAATTGGGAGCTTATGTTGG
AAGAAAGTAGAAGGGTCGGGATAGTCATTGCTTAGAGGCAACGCAGTTGCAAGTCTACC
ATTCCCACCAATTATAATTTGGAAAGAGTTTAAGATGCTGACTTATCCAGGAAAAGGTTGT
TTATACTTATAAACAACAGAGAGACAACTGTACAGGTGTTGTAAACACTCCCAGTGTGAG
GGTAATTTTGGAAGTTGTGCCTTAAAAAAAAAA

SEQ ID NO: 117
ACTTGAAAAATCTCTACTTTTTTTTCCCTCTTCTGAGAGAGAAGTTCTCCGCCTCGAAGCC
GCGACGGCGCCCGCCCTCCTCCTCGCTCCGCCGCTGCAATGGCCGCTTTCGGAGCAG
CTCCCGCCGGGAACCATAACCCTAACAAGTCTTCCGAGGTGATTCAACCTCCCAGCGAT
TCCGTTTCGAGCCTGTGTTTTAGCCCGAGGGCCAATCACTTAGTTGCTACTTCGTGGGAT
AATCAGGTACGGTGCTGGGAACTTACGAAGAATGGGGCTTCTGTAACTAGCGTGCCCAA
GGCGTCGATGTCCCATGACCAACCGGTACTTTGTTCAGCTTGGAAGGATGATGGAACGA
CTGTCTTTTCTGGCGGCTGTGATAAACAAGCAAAAATGTGGTCTTTAATGTCTGGAGGTC
AGCCAGTGACAGTTGCCATGCATGACGCACCCATTAAGGAGATTGCTTGGATCCCAGAG
ATGAATGTCTTAGTTACAGGAAGCTGGGACAAGACCCTGAAGTACTGGGACACGAGGCA
GTCAAATCCAGTACATACTCAACAGCTCCCAGAGCGCTGCTATGCAATGACAGTGAGAT
ATCCCCTGATGGTTGTTGGCACTGCAGATAGGAATCTTATTGTCTTTAATCTGCAGAATC
CTCAGGCTGAGTTCAAGAGATTTTCTTCACCCCTAAAGTACCAGACAAGATGTGTTGCAG
CTTTCCCTGACCAACAAGGTTTCCTGGTTGGATCCATCGAGGGGAGGGTTGGCGTTCAT
CACCTGGATGATTCCCAGATCAGTAAAAACTTCACATTTAAGTGTCATAGAGATAACAAC
GACATTTATTCTGTCAACTCCTTGAATTTCCATCCAGTGCATCATACCTTTGCCACTGCTG
GATCTGATGGTACTTTTAACTTTTGGGACAAGGATAGTAAGCAGAGACTTAAGGCAATGT
CAAGATGCAGCCAACCAATACCTTGTAGCACCTTCAATAATGACGGCACAATATATGCAT
ATTCGGTCTGTTATGACTGGAGCAAGGGGGCAGAAAATCACAATCCTGCTACAGCAAAG
ACCTACATCTTCTTGCACTTACCACAGGAATCGGAGGTCAAAGCAAAGCCACGGGTTGG
AACAACTAACAGGAAGTGAAGTTGCTGATTAGTTTAGAATATGAAGTTTTGAGGAAAACA
GTGGCAGTGAAGAAAATGGTGCTATTCATTTGCCAGAAATTGTAGAATAGGAATGGTCCT
TCCCTCTCATTGTGGTTTTTTGAGTCATAGTAATCCATTGTTGATATTTTTTTTTTCCTTTTT
TTGCTCATTTTCATATACCAGTAACTTTTGCGGTCTCTGATAGGCCAAGTTCTATTAACTG
TGGTAGATGGGAAGAGTGGCGCTTATTTCAGTACTGTTTTTGTCAAGTCATGTCTATCAC
ATTGATGTTTGAAGATTAAATACATTTTGTACTTTTCCGGAAAAAAAAAA

SEQ ID NO: 118
GACACGCGCCCTCTCTCTCTCCCTCTCTCTCCCCCTCCCTCTCTCCCTCCCCTTGCCGC
TGCTCTCCAAAATCTTCCACCAAATCTCTTCTTCTTCAAAGACGCAGCAGCAGCAGCAGC
AGCAGCAGCCGCCCCAGCAGCCCTCGCTACACGCTCCCTCTCTCGCCCTCCCCCAAGC
CGCGTGCGGCGGAAACCGAAGCCCTAGCTCTCGCTCCCTCTCCTCTCTCGCGAAATGA
ACTGTTCCATATCCGGCGAGGTGCCGGAGGAGCCCGTCGTCTCGACCAAGTCCGGCCA
CGTCTTCGAGAGGCGCCTAATCGAGCGATACGTTTCGGATTATGGAAAGTGCCCAGTTT
CCGGGGAACCACTTACCATGGATGATGTCCTCCCTGTGAAAATGGGAAAGATAGTCAAG
CCCAGACCTTTGCAGGCTGCCAGCATACCTGGGTTGCTCAGCATTTTCCAAAATGAATG
GGACAGCTTGATGCTTTCCAACTTTGCACTGGAACAACAACTGCACACAGCTAGGCAAG
AGCTAAGTCATGCATTGTACCAGCATGATGCTGCTTGCCGTGTTATTGCAAGACTTAAGA
AGGAAAGGGATGAAGCCCGATCATTACTTGCACTGGCTGAAAGACAGATACCCATGACA
GCATCTTCTGATATTGCAGTGAATGCTCCTGCCATGAGCAATGGGAGAAAAGCTTCTCTA
GATGAAGAGCCAGGCTATGCCGGGAAGAAAATGCGACCTGGTATTTCTGCAAGCATCAT
TGCTGAGATAACGGATTGTAACTTAGCACTTTCTCAACAGCGTAAAAAGCGACAGATTCC
CTCTACACTGGCACCTGTTGAGGATTTGGAGAGGTACACCCAACTTTCTAGTTATCCACT
GCACAAAACTGGCAAACCAGGCATTACATCTCTTGATATTTGTCATTCCAAGGATATCATT
GCAACTGGTGGGATTGACACATCTGCTGTACTTTTTGATCGATCATCTGGACAGATCATG
TCTACACTATCAGGGCACTCAAAGAAGGTTACTAGTGTGAATTTTGATGCCCAAGGTGAT
ATGGTTTTAACTGGTTCTGCAGATAAGACTGTGCGGATTTGGCAAGGTTCTGAAGATGG
GAGCTATAACTGTCGGCATATATTGAAAGATCATACTGCTGAGGTGCAAGCAATCACAGT
TCATGCTACAAATAACTACTTTGCAACAGCATCTCTCGATAATACATGGTGCTTTTATGAG
TTTTCAACTGGTTTATGTCTTACTCAGGTTGAAGGCGCTTCAGGATCTGAGGGTTATACA
TCTGCAGCTTTCCATCCCGATGGTCTTATCCTGGGTACTGGCACCTCGAATGCCGATGTT
AAAATATGGGATGTAAAAACACAGGCGAATGTTACAACATTTTCTGGCCATACCGGGGCA
ATTACTGCCATATCTTTCTCTGAGAATGGATACTTCCTTGCGACTGCAGCTCAGGATGGG
GTTAAGCTGTGGGATCTGCGGAAGTTGAAGAACTTTCGCACGTTTTCGGCATATGACAAA
GACACCGGGACAAATTCTGTTGAATTTGATCATAGTGGATGTTATCTTGGACTTGCCGGC
TCAGATATAAGAGTATACCAAGTTGCCAGTGTAAAATCAGAGTGGAATTGTGTCAAGACT
TTCCCTGATCTATCTGGGACAGGTAAAGTGACATGCGTTAAGTTCGGTCCGGATTCAAAA
TACATTGCTGTCGGATCGATGGATCACAATTTGCGAATCTTTGGATTGCCTTCGGAGGAC
GGTGCTATGGAATCATGAAGTGTTGGGATCACATCTAGCAGGTAGGATCAGCTTTTGTG
AGAGCAAAGTAGAGTTTAAGTTTCGTTGTGCTTGGACCGGAAAACTCACATGCTTAGAGT
TTAAGCTTTGTGCCAGCTGAATCAGAGTATGAAACATCCTGTACTGCTGCAGTGAACGAT
TAGCGGCCCTTTGTCTAACATGATTACGATTCTTTGTTTTTCTTTTCGAATTTTGCCTTCTT
GAGCACTAAGTCGTGTATACTATTGTTGTACAAGATTTTGGGACTGACCCATTCTCTAAAA
AAAAAA

SEQ ID NO: 119
AAAAATTCTCTCTTCAGTGTTTTTCTTTGTGCTCCCTGCACCTTTGCACATTAACCTTTGG
GAAGGCAATAATGGCGGCGCCTGGAGTCGAGACTTTGAAGAAAGAGATTAAGGAGCTTA
AAGAGAAAATCGCACAACATAGGCTCGACACTGATGGGGAGCAACCATTGCCCGCTGCT
GCCAAGTCTAAGTCAGTGCCTGAAGTTTCTGCAGCATTGAAGCAGAGACGTATTCTGAA
GGGTCACTTTGGCAAAATCTATGCTTTGCACTGGTCCGCAGACTCCCGACACCTCGTTA
GCGCCTCACAGGATGGCAAACTCATTATTTGGAACGGCTTCACGACCAACAAAGTGCAC
GCGATCCCACTGCGGTCGTCTTGGGTCATGACCTGCGCCTATTCTCCCAGTGGCAACCT
CGTTGCCTGTGGTGGCTTGGACAACCTGTGCTCGGTGTACAAGGTTCCGCATGGAGGA
AACAAGGAATCTTCTTCGGCTCAGAAGACGTATGGTGAACTCGCACAACATGAAGGATA
CCTTTCTTGCTGCCGATTCATCAAGGACAACGAGATTGTGACGTCCTCCGGCGACTCTA
CTTGCATTTTGTGGGATGTGGAAACCAAGACGCCCAAGGCCATCTTCAATGATCACACA
GGAGACGTCATGTCTCTCGCTGTTTTTGACGACAAGGGCGTTTTCGTTTCTGGCTCTTGC
GACGCTACCGCCAAGCTGTGGGATCACCGCGTGCACAAGCAATGCGTTATGACCTTCCA
AGGCCATGAATCCGACATCAACAGTGTGCAGTTCTTCCCGGACGGCGACGCTTTCGGCA
CTGGCTCTGACGACTCTTCTTGCCGTCTCTTTGACATTCGCGCCTACCAACAAATTAACA
AATATTCCAGTGACAAGATCTTGTGTGGAATCACATCCGTCGCCTTCTCGAAGACTGGCA
AGAGCTTGTTTGCTGGCTATGATGACTACAACACTTATGTTTGGGACACTCTGAGTGGAA
ATCAGGTTGAGGTTCTGACCGGACACGAGAATCGTGTGAGCTGCTTGGGTGTCAGCGAA
GATGGCAAAGCACTAGCCACTGGCAGCTGGGATACATTGCTCAAGATTTGGGCATAACT
TGTATGAACTTTTTCTGTTGTCGACACTGTAATTACACGAGCTCCCCTTCTTTTGCTGTGT

ATGTCGAGGTTGTTTGTTGCTATTTGATGGATTGTCCGATAAAGCTGACAACACGAAAAA
AAAAA

SEQ ID NO: 120
AAGCTGTTTCTTCTTTTTTCCCCTCGCCTCCTGAGCACAGGAAGGATCGTATCTTTTTTCG
ACCCTCTTGCCGCTTCCTTCTCGTGCGCCCCCAATTCGCGTTCATCTGCGGCTGTATTAT
CAAATGACTCTGGTGGATGGTAGGTGAAGAGTTTATAACCCACAAACCCTACCCGTCCTA
CAGCAGCCAACATTCTCCTTACTCTATCTCCGGAGTAGTCCAAAGCAACTTGTCCATTAC
TTTGGATCATAAACTTCTCCTAAGATCTGGCCTTTTACAAGCTCAATAATTTATGGGCTGA
GGTGCACTTTTCATGGGAGGTGTTGAGGATGAGAGTGAACCAGCCTCAAAACGCATGAA
GTTATCATCCAGAGTTTTGAGAGGTCTTGCAAACGGTTCATCTCGTACAGAGCCTGCAGC
TGGCTCTTCACTAGATTTAATGGCTCGGCCCCTACCAATTGAAGGGGACGAAGAAGTTAT
TGGTTCAAAGGGTGTGATCAAAAGAGTTGAATTTGTACGACTTATAGCGAAGGCATTATA
CTCCCTCGGTTATGAAAAAAGTGGTGCTCGCTTGGAGGAAGAGTCTGGGATACCGTTGC
AGTCCTCTGTGGTAAATTTGTTCATGCAACAGATATCTGATGGGCTTTGGGATGAAAGCG
TGGTGACGTTGCATAAAATTGGTCTCTCTGATGAAAATTTAGTAAAGTCAGCCTCTTTCTT
GATATTGGAGCAGAAATTCTTAGAACTTCTGGATCAGGAAAAAGCTATGGATGCTCTGAA
GACGTTAAGGACGGAGATCACACCTCTTTGCATAAAAAATAGTAGGGTACGTGAGCTCT
CGTCGTGCATCATCTCTCCATCATCATGTGGGCTTCTTAACCAGAATAAAAGAAATAGTA
CAAGAGCAAGGTCCCGTTCAGAGCTTCTGGAGGAATTGCAAAAATTACTTCCTCCAGCA
GTTATTATTCCAGAAAGAAGGTTGGAACATCTGGTGGAGCAGGCCCTTGTCCTGCAAAC
AGATGCATGTATGCTTCATAACTCTATAGATATGGAAATGTCACTGTACACTGATCATCAA
TGTGGTAAAGAACACATCCCCTGTCGAACTTTGCAGATTTTACAATCACATAATGATGAA
GTTTGGCTTGTGCAATTTTCACATAATGGGAAATATTTAGCTTCTGCATCCAATGATCGAT
CAGCAATCATTTGGGAGGTTGATGAGAATGGCAGCGTCTCATTGAAGCATAAATTGACTG
GTCACCAGAAGCCGATTTCTTCTGTCTGTTGGAGTCCAGATGACCGACAGCTTCTCACTT
GTGGGGTTGGGGAGACAGTGAGGCGCTGGGATGTCTCTTCTGGTGAATGCCTTCGTGT
TTATGAGAAAGCTGGCCATGGCCTCATTTCATGTGCGTGGTTTCCAGATGGAAAATGGAT
ATGTTATGGTGTTAGTGATCGGAGCATATGCATGTGCGACTTGGAGGGGAAAGAGATTG
AATGCTGGAAAGGGCAGAGAACTCTTAGTATTTCTGATTTGGAAATTACTAGCGATGGAA
AGCAGATCATAAGTATATGTAGAGAAACTGCAATACTTTTACTTGACCGGGAAGCAAAAT
ATGAAAGAATGATAGAAGAAAATCAAACGATAACTTCTTTCTCATTGTCAAAGGATAATAG
ATACTTGCTTGTTAATCTCTTGAATCAAGAGATCCATCTTTGGGATATAAAAGGGGATTTC
AGGCTGGTTGCAAAGTACAAAGGTCTTAAGCGCAGTCGGTTTGTAATCAGGTCTTGTTTC
GGTGGACTCAAACAGGCCTTTGTTGCCAGTGGTAGTGAAGACTCACAGGTTTACATTTG
GCACAAAGGCTCAGGTGAGCTGATCGAGCCATTGCCAGGTCACTCAGGAGCTGTGAATT
GCGTGAGCTGGAACCCAGCAAACCACCACATGTTGGCATCGGCCAGTGACGACCGTAC
CATCCGGATATGGGGCTTGAATGAGCTAAACACGAGGCACAAGGGTGCACGCCCCAAT
GGTGTCCACTACTGTAATGGCAATGGCACCAGCTGAAGAAGAAGAAGAAAGATGCATGG
CTCTAGCTGAGAGCCCTGATGCTGCTTTGTTAGCCATTGCTTTGGTGTAAATACGTATGT
TCATCTAAAAAGCAGACACATTCATGTCAGACCAAGCATGTGAATCTTCAGCAAACTTAC
TGGTAATAAATTTTTACTATTTCATCATAAAAAAAAAA

SEQ ID NO: 121
CCTTTCAAAATTCCCTCTCTCTTTCTTTTCTCTCTCTAGCCAGATCTCATCTCCTTCTTCCC
CTTTCCCCTTTCATCTCTGCATCTGTACATGCCCTAATTTCTCTCTCCCTCTCTCTCTCTC
TCTTGTTTCTCTCTCTAGAAGATGACGCAGCTGGCGGAGACCTACGCGTGCATGCCCTC
GACGGAGCGCGGCCGCGGGATCCTCATCGCCGGCAACCCGAAGCCCGGGTCCAACTC
CGTCCTCTACACCAACGGCCGATCCGTCGTCATCCTCAACCTCGACAACCCGCTCGACA
TCTCCGTCTACGCCGAGCATGCCTACCCCGCCACCGTCGCCCGCTTCTCCCCCAACGG
CGAGTGGGTCGCCTCCGCCGACTCCTCCGGCGCCGTCCGCATCTGGGGCGCCTACAA
CGACCACGTCCTCAAGAAGGAGTTCAAGGTCCTGTCCGGCCGGATCGACGATCTCCAG
TGGTCCCCCGACGGCCTCAGGATCGTGGCTTCCGGCGATGGGAAGGGCAAATCGCTCG
TCCGCGCGTTTATGTGGGACTCAGGCACCAATGTGGGAGAATTCGATGGCCATTCACGT
CGAGTTCTGAGCTGTGCTTTTAAGCCAACTCGGCCCTTTCGCATTGTGACTTGTGGAGA
GGATTTTTTGGTGAACTTTTATGAAGGACCGCCTTTTAAATTCAAGCTGTCTCGCAGGGA
TCATTCCAACTTTGTCAACTGCTTGAGATTTTCTCCAGATGGCAACAGGTTCATTAGCGT
GAGCTCTGATAAAAAGGGAATCATCTATGATGGTAAGACTGGTGAGAAGATAGGAGAGC
TGTCATCTGATGGTGGTCACACAGGTAGCATTTATGCTGTCAGTTGGAGTCCTGATAGTA
AGCAGGTTATAACTGTGTCTGCTGACAAGTCAGCGAAGATATGGGACATTTCTGAGGAT

GGCAGTGGTAACCTAAGGAAAACATTGACTTCTTCTGGTTCAGGCGGGGTTGATGATAT
GTTAGTCGGTTGTCTTTGGCAAAACAACCACCTAGTCACTGTCTCTCTTGGTGGCACAAT
CAGCATATACACAGCAGGTGATCTTGATAAAGCACCTGTCTCTTTTTCTGGACACATGAA
GAACGTCTCTTCCTTATCTGTGCTCAAAGGTGATCCAAAAGTGATCCTCTCCAGCAGCTA
TGATGGTCTTATAATAAAATGGATTCAAGGAATTGGATTTAGTGGCAGAGTACAGAGGAA
AGAGTCTACACAAATCAAATGTTTGGCTGCAGTAGATGAGGAGATTGTAACCTCTGGATA
TGATAATAAGGTATGCAGAGTTTCTGGCAGTGGAGATGCAGAATTCATTGACATTGGCTG
TCAACCCAAGGACTTGAGTCTTGCCCTTCAGTGTCCTGAATTTGCTTTGGTTTCAACCGA
TACTGGAGTTGTCTTGCTCCGCGGTGCAAAAATTGTGTCAACCATCAATCTTGGGTTTGC
TGTGACAGCATCAACTGTCGCACCAGATGGAACTGAAGCAATTATTGGTGCACAGGACG
GAAAATTGCGCATATATTCTATTTCAGGTGATACATTAACAGAGGAAGCTGTTCTGGAGA
AACACAGAGGTGCTATTAGTGTCATACACTACTCGCCTGATTTATCCATGTTTGCTTCTG
GGGATCTGAACAGAGAGGCTGTTGTTTGGGATCGTGCCTCAAGAGAGGTGAGGCTGAA
GAATATCTTGTACCACACAGCCCGCATCAACTGTTTGGCTTGGTCCCCTGATAGCAGCA
CAGTAGCAACTGGATCACTTGACACTTGTGTCATCATCTATGAAGTTGACAAGCCAGCAT
CTAACCGTCTAACCATAAAAGGAGCTCACTTGGGTGGGGTTTATGGATTAGCTTTCACCG
ATGACTTCAGTGTGGTCAGTTCTGGTGAGGATGCGTGCATTCGTGTCTGGAAGATAAAC
AGACAGTGATCTTGGGGAGATTATGTGAAGTTTCAGAAATCGCATCTACTTGAGGGTAGA
TTTGCTGTAAGAAGTGGGAGGCTTTTTTTCGGCTGCGTCAACCATTAGAGTCCGAGCTG
CAGGAGTCACTGCCTCTTATCTTTTAAGAGAGGTTTTTTCATCATTGAGGCTTGTGTTTGT
AAGTGTTAGCAGATAAATCTTTGCTGTATTTTCTCTTCTCCTCCTTTTCTCCTCGAGATTTT
TATTGTGGGAATGTTTTTCAAGTTTTACTGTATTCTACTGTTTATTGCTTG

SEQ ID NO: 122
TTTTAGTTCGCTTCTCCAGCGCCCCATTCTCTTTTTAGGGTTTTTCGGCGATCCAGAGAA
ACGGTCACCGCCGGCGAGAGCTTCTGAAACCGGCTTACTTCTGGCCGAGCAATATCGG
CAGACGCCGCGATGAAGGTGAAAGTGATATCGCGTTCCACGGATGAGTTCACCCGAGA
GCGCAGCCAGGACCTCCAGAGGGTATTCCGCAACTTCGACCCCAACCTTCGGACCCAA
GAGAAGGCGGTGGAGTACGTCCGGGCTCTGAATGCAGCTAAATTGGACAAGGTTTTCG
CAAGGCCCTTTGTTGGAGCCATGGATGGACACGTCGATTCGGTATCGTGTATGGCCAAG
AACCCGAATTACTTGAAGGGAATATTTTCTGGCTCTATGGATGGAGATATCCGCCTTTGG
GACATTGCTTCGAGGCGAACAGTATGTCAATTTCCTGGTCATCAAGGCCCTGTCAGAGG
CTTGGCGGCATCTACAGATGGTCAAATTCTGGTATCCTGCGGAATTGATAGCACGGTTC
GATTGTGGAATGTTCCTGTAGCTACTCTTGGGGAGTCTGATGGCACACATGAGAACTTG
GCAAAGCCACTGGCAGTTTATGTATGGAAGAATGCATTCTGGGCAGTTGATCACCAATG
GGATGGCGAACTTTTTGCTACAGCTGGTGCTCAAGTAGATATTTGGAATCAAAACAGGTC
TCAGCCAATAAGCAGTTTCGAATGGGGAACGGATACCGTCATATCTGTAAGATTCAACCC
TGGAGAACCTAATGTATTGGCAACATCAGGGAGTGACCGCAGCATAACCCTCTACGATT
TGCGTATGTCGTCCCCAACAAGAAAAGTTATCATGAGGACAAAAACTAATGCTATTTCTT
GGAACCCAATGGAGCCAATGAACTTCACTGCTGCAAATGAAGATTGCAATTGCTACAGCT
ACGATGCTAGAAAGTTGGAAGAAGCCAAATGTGTACACAAGGATCATGTTTCTGCTGTGA
TGGATATTGATTACTCTCCCACTGGCCGGGAATTTGTAACTGGATCCTATGATAGAACAG
TAAGAATTTTCCAGTATAATGGAGGTCACAGTAGGGAAGTTTATCATACAAAAAGAATGC
AAAGGGTATTTTGTGTCAAGTTCAGTTGCGATGCGAGTTATGTCATATCCGGGAGTGATG
ACACCAACTTGAGGCTTTGGAAAGCTAAAGCCTCTGAACAGCTGGGAGTTGTTCTTCCAA
GAGAGCGCCGCAAGCATGAGTATCATGAAGCTGTCAAGAGTCGATACAAGCACCTCCCT
GAAGTGAAGCGTATCGTGAGGCACAGACACTTGCCTAAACCAATATACAAGGCAGGTAT
TCTAAGGCGCACCGTGAATGAAGCAGATAGAAGGAAAGAAGAAAGAAGAAAAGCACACA
GTGCCCCTGGATCTAGTTCAGCAGAGCCATTACGTAAAAGAAGAATCATCAAAGAAATTG
AGTGATGCTTCTTCTTTTAGCGTGTTGTGCTGCATGGCATAATATTTGCATTTCTCAAATC
TGTATTGTTGGATGGGGCTCTATCTTTCTCAAGAGGCTGCTTTTGAGTGAAAAGTTAACT
TCACCTGAGAGATGAGATAGGGCGACGATGCACAAGCTCCTTTCAGTCTTAACAATTTTG
GTAGAATACGGCGGTCCTTTCCTATGAAGGGTGGCCCGAGAGTTAGCCAATGTTTGGTC
ACGCCAGCTTTTGCCTGGAGCATCACAAGTAGTCAGTTTCTGATCTATTGGCGATTTATC
AAGACATGATTTGTTTGTATCATAAGCTTTGATCTTTAACTTGAGTTCAAAAAAAAAAAA

SEQ ID NO: 123
AAGCATAAACCCTAGCAATCAGACATCCCAAAAACCCTCTCTGCTCAACCCACTCGATCC
TCTTCCCCAGGGATCCCGGGAACCCTAAGAGAGCCGCCGGCTCCAACTACCGTCCATG
TACGATCCATACTAGATCTGACCAACTGTCCGCCGTTCACCGGGGCTTCGGTCGCCGGC

TCCCCTGTCATCAGGTCACCAACTTCCTTGACCTTCCGGCACTGTTGCCAGTCGCGTTTC
TCTGCACGCGTACCATCGAACGACCACGAGGATGGTTCGAAGCATAAAGAACCCCAAGA
AAGCCAAAAGAAAGAACAAGGGATCAAAGAACGGAGATGGGTCGTCTTCGTCTTCCTCA
ATTCCTTCAATGCCTACGAAGGTGTGGCAGCCAGGCGTGGACAAGTTGGAGGAAGGAG
AAGAGCTTCAGTGTGACCCGTCTGCTTACAATTCCCTTCATGCCTTCCACATCGGCTGGC
CTTGCCTCAGTTTTGACATTGTCCGTGATACACTGGGTTTGGTTCGCACTGAATTTCCAC
ATCAAGTTTACTTTGTTGCTGGAACTCAGGCAGAGAAACCTACTTGGAACTCAATTGGGA
TATTTAAAGTAAGCAACATTACTGGTAAAAGGCGTGAACTGGTGCCATCCAAACCTACTG
ATGATGCCGATGAGGAGAGTGACAGCAGTGATAGTGATGAAGATAGCGACGATGAAGTT
GGTGGATCTGGGACGCCAATTTTGCAGCTAAGAAAAGTAGGCCATGAAGGATGTGTTAA
TCGAATAAGAGCTATGAATCAAAATCCCCATATCTGTGCATCTTGGGGAGACTCTGGACA
TGTGCAAATATGGGACTTCAGCTCCCACTTGAATGCATTGGCTGAATCAGAAGCAGACG
TCAGCCAGGGAGCTTCCTCAGTTTTTAATCAGGCTCCATTAGTGAAGTTTGGCGGCCAC
AAAGATGAAGGTTATGCCTTAGATTGGAGTCCTCTTGTACCTGGGAGGCTCGTATCTGG
GGACTGCAAGAATAGCATCCATCTTTGGGAACCGACATCCGGTTCAACATGGAATGTTG
ATTCCACTCCTTTTATTGGACATGCTGCAAGTGTTGAAGATCTTCAATGGAGCCCCACAG
AAGAAAATGTCTTTGCCTCTTGTTCAGTCGATGGAACTATTGCGATATGGGATACCCGTT
TAGGGAAGACACCAGCTGCTTCTTTTAAGGCACATGATGCTGATGTGAATGTGATCTCAT
GGAACAGGCTGGCTACCTGTATGTTGGCATCTGGATGTGATGACGGGACATTTTCAATT
CATGATCTTAGATTACTCAAGGAAGGTGATTCTGTGGTTGCTCATTTCGAGTATCATAAAC
ACCCGGTTACCTCAATCGAATGGAGCCCACATGAAGCCTCCACATTGGCAGTATCATCG
GCTGACTGCCAGCTCACAATCTGGGATCTTTCCTTGGAAAAGGATGAGGAAGAGGAGGC
AGAGTTTAAAGCCAAAACGAAAGAGCAAGTGAATGCCCCAGAGGATTTACCTCCGCAGC
TCCTCTTTGTTCACCAGGGACAAAAGGACTTGAAAGAACTGCATTGGCATGCTCAGATTC
CGGGAATGATTGTATCTACTGCAGCAGACGGCTTTAACATCCTGATGCCCTCGAACATAC
AGAGCACTCTTCCTTCAGATGGTGCCTGAAATTTTGACCGAGACCGAACAATACCCAACA
CCCGGCCTAATTTTTCAAAGAACAGAGGCCTAAGGGCTGTCCAACCATGGAGTCGCACA
AACCGCAGTGCAAATTTCTGTGTACATTGTGATGCAATGATGAGCAAGTTTGAGACAGTA
GATGATGCAGCAACATATCGTGCTCATAACTGCGAGTTGCGTCTTTTTTTTTTTTTTTTC
TTTTTTCGTTTGTACGGTTCGTCTTCCTTGCTATGCTTAGCTTTGTTGTCCCGCCCTTGTA
ATGTTTTCCCCATGCGAAGGTTTCTGGGAATTTTCAGTAGAAAATTCGGTCGTGGCGGCC
ATCCTCGATATTTTGGAATGTTTGGTAACTTGAGAGATTTTTTTATCATGAGCATGGGCAT
AAGTTTAATGCACATACGGATACTTTAGAGTCAAAAAAAAAAA

SEQ ID NO: 124
GGGTCATGGTCATCGCTCTCATACTGCCCAGGTGAAGGAAAGCGCTAAGCATAAGCATA
TACCATAAGCATATAAGCAAATGGGTCCTATCCAAAATGCAAGAGGTTGGAAGCCTTGGA
GGTCATGGTCACCGCTCTCATACTGCCCAGGTGAAGGAAAGCGCTAAGCATAAGCATAT
ACCTTTGGCACCTCCTCCTTTAATGGTCATGCAAATAATCACGGTGAAGGAAGCTGAATC
TATATGAAGAGATATTGGCTTGCATAGTCACGGTGAAGGAAGCTGAATATATATGAAGAG
ATATTGGCTTGCATAGTCACGGTGAAGGAAGCTGAATCTATATGAAGAGATATTGGCCTC
TCTAGATTGTCTTGGGTTTGAACATTCTATTATAAGCAATCCATGCAGTGGAACTCCTTCA
ATGGTCATGCAAATGGTCACGGTGGAGGAAGCTGAATCTATACAAAGGGATATTGGCTT
CTTTATATTGGATTGGGTTTGAACATTCTATTATAAGCAATCCATGCATAGAGCTTCAGAA
CATTAAAAGCTTACAATCTTATCCCATATCCTTTATATCTATCCTCTCCCTCTTTAATTTCG
CCCCTGTAATTCCTTTTTCTTTCTCTGGTCTTTTGATCTAGTGTTTGAGGAGAATGGTTTA
GAGCAATTAGGCATGGAGAATTTCATTCGCTTCCATGGAGAGCTTCACGGGGATTTCAG
AGCTTTAGATCGTTTCCGGTTTCAGGCTATAATTATAAGGTATTTTTTTTAAATAGATCGA
GCTATTTGAAAAGGCCATGGAGTGGGGATTTAATCACATTACAACCCAGACTCCTCTTCA
GGTTAAATTGCATGCCACTGCATTGGAAATTGCAGTGGAAGAGTTGCCAGATAAACTGGT
TTTTAGTGGATGATTGTCGAGGTGCTGCAGCTTATAAAGGCTGCACTCAATGATGCCGAT
CTACTCTTTCCTTCTAGAACTCCAGAGTTACTGTTATTCTTGTTTTACTAGGGTGGACTTT
TGTGGTTGTGCTTCCTTTTATCTCAGTGGATGGTACCAGGTCATGGGTTTAAAGAATTAT
CTGCAGAGTGTTCGCTGTCAATTTTCAAGCAGTGTTTTACCAGCTTGTTGATCTCCTATTT
GATTATCTAGACAAAGCCATGGAGCGTTACAAGGTCATAAAAGAGTTGGGGGATGGCAC
TTATGGAAGTGTATGGAAGGCTTTGAATCAGCAGACACATGAGATAGTAGCTATAAAAAA
AATGAAGAGGAAATATTATATCTGGGAAGAGTGTATCAATCTCCGGGAGGTCAAGTCTCT
GCGGAAGTTGAACCATCCCAACATTATCAAGCTGAAGGAGGTCATTAGGGAAAATAATG
AACTCTTTTTCATATTTGAATACATGGAATGTAATCTGTATCAAATAATGAAAGAGAGATCT
ACTCCTTTTTCGGAAACAGCAATTATCAAATTTTGCTATCAAATACTGCAAGGGTTATCCT

ATATGCACAGGAATGGTTATTTTCACCGAGACTTAAAACCAGAGAATTTGTTGGTAACTA
GCGACTTGATTAAAATTGCGGATTTTGGGCTGGCAAGGGAAGTTCTCACTAGCCCACCT
TATACAGATTATGTTTCAACAAGATGGTATCGTGCTCCAGAAGTCCTGCTGCAATCTCCG
ACATACACTACTGCAATTGACATGTGGGCAGTTGGGGCTATACTGGCGGAACTCTTCAC
TTTGCATCCTCTCTTTCCTGGTGAAAGTGAACTGGATGAAATTTACAAGATCTGTGGTGT
GCTTGGCACTCCAGATTATGAGACTTGGCCTGACGGCATGCAGCTTGCAGCTTTTAGGA
ACTTCATCTTTCCCCAGTTTCTACCAGTCAATCTTTCAGTTCTTATTCCCCATGCTAGCCC
AGAAGCAATCGATTTGATTACGCGTCTATGTTCTTGGGACCCTCAGAAGAGGCCAACAG
CAGAGCAGGCACTGCATCATCCTTTCTTCCGTATTGGCATGTCTATTCCTCTTTCTCTTG
GGGGACATTTCCAAGATAATACATGTGCAGCAGAGGTAGATACAAATTTTCATTCCAAAA
AGGCATGCAAGGGGCGTGGAATGGGGGAAAAAGAATCAAGCTTGGAATGCTTCCTTGG
TTTGTCTTTGGGGCTCAAGCCAAGCCTTGGTCATCTGGGCGCAATGGGATCTCAAGGTG
TGGGAGCAGTGAAGCAGGAAGTGGGGTCCTCTCCAGGGTGTCAGAGTAATCCAAAGCA
GTCCCTATTTCAGGTTTTAAACTCAAGAGCAATTCTACCACTATTTTCCTCAAGCCCCAAC
CTGAATGTGGTTCCAGTCAAGTCCTCTCTACCTTCAGCATATACAGTAAACAGTCAAGTC
ATGTGGCCAACAATAGCAGGTCCACCTGCTGCTGCAGTTACTGTTTCTACTCTACAGCCA
AGCATACTCGGCGATTTTAAGATCTTTGGAAAATCCATGGGGTTGGCTTCACAGTATGCG
GGAAAGGAAGCTTCCCCTTTCAGCTAGATAATCAAATTTAGTTCATGCTTTCTGAACAAA
GTAAATTTTTAATGAGTGAAATTCACCATTTTAAGCTGATAGCTTTAGTTCCTACGTGGAA
TGTATAAATGCACCATTGTCCATAAGGCAAGAGCTTTAATCTGGATTCCAGTATATATACC
CAGTTTGGTTTCTAAATTCATTTGTAAATAGTTCCAACCGAGAGAGGATATTTTTGAGGTA
GATTGGCAATGGAATTACTGCTTGCCATTAAAAAAAAAA

SEQ ID NO: 125
TGTTCAATTTAATTGGATGTGAAAGTCATTTGCAGGTTCTTTGCTTAACGAACGATCGTGT
CACGTTAGTGTCAGCGCTGTGTGAAGAATGAGTTTGAGGCGAATATATAAATGGAATTAT
AAACGCCAGCTGAACTGAAAGCATTCAGTCATGGGCGAGATGGGTAGGGGCATCAACAA
TAGCAGCAATAACAATAATAGCAACAGGCCGGCATGGCTCCAGCATTATGATTTGGTGG
GCAAAATCGGGGAGGGCACCTACGGCCTTGTATTTCTCGCGCGGAGCAAATTACCCAAT
AACAGAGGTCTCCGCATTGCAATTAAAAAGTTCAAGCAATCTAAGGATGGGGATGGCGT
TTCCCCCACTGCTATCAGAGAGATAATGCTACTCAGGGAGTTTTCCCATGAAAATGTCGT
TAAACTCGTGAATGTGCATATAAATCACGTAGACATGTCTCTGTACTTGGCATTTGACTAT
GCGGAACATGATCTTTATGAGATTATCCGACATCATCGGGAGAAGCTAAATCATCATAAC
ATTAATCAATACACCGTTAAGTCACTACTCTGGCAATTATTGAATGGATTGAATTATCTTC
ACAGTAATTGGATTGTACATCGTGACTTAAAGCCATCTAATATCCTGGTGATGGGTGAAG
GTGAAGAGCACGGGGTTGTGAAAATAGCTGATTTTGGGCTTGCCAGGATATACCAGGCT
CCCTTAAAACCTTTATCTGATAATGGGGTTGTTGTTACGATATGGTATCGAGCACCGGAG
TTACTCTTAGGGGCCAAACATTACACGAGTGCTGTTGATATGTGGGCTGTAGGATGTATT
TTTGCAGAACTGATAACACTGAAACCACTGTTTCAAGGTGTGGAGGTCAAAGCTTCACCA
AATCCTTTTCAGCTTGACCAACTTGACAAGATATTTAAGGTCTTAGGACATCCAACAATAG
AAAAGTGGCCAACTCTAATGAATCTACCACATTGGTCAAAGAATTTGCAACAAATTCAACA
GCACAAATATGACAATGCAGGGTTGCATATTGGTCCCATTCCTGCAAAAAGCCCTGCTTA
TGATCTTCTTTCAAAAATGCTTGAGTATGATCCCCGCAAGCGTATCACAGCAGCACAAGC
TTTGGAACATGAGTATTTTCGAATTGATCCTCAACCAGGACGCAATGCACTTGTTCCCAG
CCAGCCTGGCGAGAAAGCTATTAATTATCCACCTCGTTTGGTCGACGCCAATACAGATTT
TGATGGAACAATCGCTCCTCAGCCTTCTCAGGTATCTTCTGGGAATGCACCATCTGGTTC
AATTGCATCAGCTGCAGTACCTGCTGTTAGACCACTTCCTCAGCAAATGCAGCTAATGGG
TATGCAAAGAATGCAGAATCCAGGCATGGCTGCTTTCAACTTAGGTGCACAGGCAAGCA
TGTCAGGACTCAATCACAACAATATTGCTTTGCAACGTGGTTCTTCTCAACAACAAGCTC
ATCAACAGGTTCGGAGAAAAGAACCCAATAGTGGGTTTCCAAATACAGGATATCCACCAC
CACCAAAATCAAGGCGCTTATAAGGATGAACTTCATGTGGAAAGTTGTATAAAATGCATG
TGAAGTCACTTCATATGTCCCTTGGTACGGCTGCCAAGCACTTCATCCTCTAATCAGTTG
GTGTCTTGGATGTTGTCAATCCAACATGTGAGGCCAGGCCTTGATGTTGGCCATATATG
GTGCTGATTGGTGTTCAATGGCTGGACTAAGAGACTCTTGTCCAAATTTAACAAGAGGCT
ATATTCTCTCTCCCAGAGAGGTCCTTGCCTATGCCAGAAATTTCAAAATGATGTCAGCCA
ATTCATCAGGTTATATTTCTGAATTAGCAGAAGCATTACATGAATACCTGTTACCTGATGG
AAACTAATTTTTACCTGATGGAATGGAGTGATCAGATTAGGCATCCTTTAGGGATTTGTAT
GGATTTAAGAACGCAGATCCTTGGATGCTCTGGTTACATGACTACTCCTTAGGGAATCAG
TCAGACATTTTAAATAACTTCCATATCTTTGAAGGTTCTTTTCTCTTCAGGCAGTAGATAC
CCAATTATATACCAACCTTGTATTAATTGCTGAGCATTCAGCATGACTTCATTTTTCTTCA

GACGTTTAGGTATCATATATGTTTTGACAATAAATAATGTGCGTATGATTCTTTGGATGCC
CTGGATATAAGACCAGTTATAAGACTACCTACCCTGGGGTAACTTGTCATGCATCATTAA
CCTTGATATCTTTGAAGGTTCTTTTCATCTTGGGCAATGGGACCCCAATCATATAACAAGT
TATGATTAGTTGTGAAGCAAAAAAAAAA

SEQ ID NO: 126
GGTGGAATTTCCATTCAACACTGGAAAAACTCCAAATCAAAGGTCAGATATCATAACCTA
TATGCTACAATGGGATTATATAGATTAATAATGAATAATACAGCCCTGAACCCAATTCTGT
TCAAAATTTGAAACAATGCTAGCTGTTGGATGGAGATTGGGTATTTAAATATGGGCCGTC
AGATGGGGTTTGAAACAACCACAGCAACGGTTCCATTGAGTTTATTCAGATCCATTCCAT
CGTCTTACTTTTTTTTTAGCTCAGTGGGGTGGGCAGGGTGTGTATTGTGGGCTCGTTGGT
GGGGTGAATTGATTACAGAGGAAGCCCAGGAAGGCACTGGAAGCATTTCAATATAAATA
ATCTCTGCAATTGATTGAATCGGCCCGCCATGGACAAGTATGAAAAGCTTGAGAAGGTC
GGGGAGGGCACCTATGGGAAGGTGTACAAGGCAAGAGACAAGATGACTGGACAGCTCG
TTGCTCTCAAGAAGACTCGCCTTGAGATGGACGAGGAAGGCGTCCCACCCAGTTCTCTT
CGTGAAATCTCCCTCCTGCAAATGCTGTCTCAGAGCATATATGTTGTTCGGTTGCTTTGT
GTGGAGCATGTGACGAAGAAGGGAAAACCACTGCTTTACCTAGTCTTCGAGTACCTTGA
TACAGATCTGAAGAAATTCATCGACTATCGACGCAGTGTCAATGCTGGTCCTCTGCCGCA
AAATGTTATTCAGAGTTTCATGTATCAACTGTTGAAAGGTGTAGCTCATTGTCACAGCCAT
GGAGTGTTGCACAGGGATTTAAAGCCACAGAATCTATTGGTCGATAAAAGCAAAGGCTTA
CTTAAAGTTGGAGATCTGGGACTTGGGAGGGCTTTTACTGTGCCTTTAAAGTGTTACACC
CATGAGGTTGTAACCCTATGGTACAGAGCTCCAGAGGTGTTGTTGGGATCAACTCACTAT
TCCACACCTGTGGACATTTGGTCTGTAGGATGTATTTTTGCTGAAATGGTGAGAAGACAA
CCACTTTTCCCTGGGGATTGTGAAATACAACAACTGCTTCATATCTTCACGTTGCTTGGA
ACCCCAACTGAGGAAATGTGGCCTGGAGTAAAACGTCTAAGGGACTGGCATGAGTATCC
TCAGTGGAAACCTGAGAACCTTGCTCGGGCAGTTCCAAATCTATCACCAACTGGTCTCG
ATCTTATCAGTAAAATGCTGCAGTGCGATCCTGCAAAGAGGATTTCAGCGAAGGCAGCT
ATGAATCACCCTTACTTTGATGATCTGGACAAGTCTCAATTCTGAAGTATTCAAATTTCAC
TATTTATGGGTGTTCAAGGATGCCAGAGACTTTTAGGTGATCATAGTTAAGGAACCGTTC
CTGCCAATTTGGAGGTTTAAGCAGCCATAGTAAAATAGTTTTTCTTGCTATAGATTGCAAG
TCTCATCTGTTGTTTGCAAAAGCAAACAGAAATTCCAATTTTTGCAGTTGATTCTACTCAG
CCTTTCATTCATTTTTTCATTAAGCGGTACTGGCAGAGGACATGTCTATTTATACAAGCAA
ATGGTCCTATTGGCTGTTTAAAACAGTTCTATTTAACTCGATCAAAATCTGACTTATTTTGA
AATTCTTCTAAAAACCAAA

SEQ ID NO: 127
GTTCGTCCCACCCATATCTCTGAACACCAGAGTAGCAATGGATGGTTATGAGAAAATGGA
TAAGGTGGGAGAAGGAACTTATGGGAAGGTGTACATGGCCAGGGACAAGAAAACAGGG
CAACTGGTCGCCCTTAAAAAGACCAGGCTAGAGAACGATGGTGAGGGAATTCCTCCCAC
TGCCCTCCGGGAGATTTCTCTTCTGCAGATGCTTTCTCAGGATATCTACATTGTAAGGCT
GTTAGATGTGAAGCACACTGAGAACAAGCTTGGGAAGCCCCTTCTGTACTTAGTTTTTGA
ATACATGGAGTCTGATCTCAAGAAGTACATCGATAGCTATCGCCGCAGCCACACTAAAAT
GCCTCCTAGTATGATTAAGAGCTTCATGTACCAGCTGTGCCGTGGAGTTGCTTATTGCCA
CAGCCGTGGCGTGATGCACAGGGACTTAAAGCCTCATAATCTGCTGGTGGACAAGGAAA
AGGGCGTGTTAAAAATAGCAGATCTTGGACTGAGTAGGGCATTTACTGTTCCTGTCAAAA
AGTACACTCATGAGATTGTGACCCTTTGGTACAGGGCTCCTGAAGTTCTCTTGGGGGCT
ACTCACTACTCATTGCCTGTTGACATCTGGTCTGTTGGCTGTATATTCGCTGAAATGTCC
CGAATGCAAGCTCTCTTTACTGGGGATTCTGAAGTACAACAACTTATGAACATTTTCAGG
TTCCTAGGAACTCCGAATGAAGAAGTGTGGCCAGGAGTGACTAAATTGAAGGACTGGCA
TATCTACCCCGAATGGAAACCTCAAGATATAAGCCATGCTGTTCCAGACTTGGAACCAAG
CGGTTTAGATCTGTTGTCTCAAATGCTGGTTTATGAGCCATCCAAGCGAATCTCAGCTAA
GAAGGCACTAGAACATCCTTATTTTGATGACCTGGATAAATCACAATTCTGAGTCCCTTTA
AGGATTTAAGCAATTCAGGATAGTTTGTAGGTATCACCATTATTATATTGATTGGTTAATT
GATTTGTGCTTTTCTGGGGGTTCTATAAATGGCCTGTTGTCTGAGAAATAATCTCTGCAA
ACTTGTTGGCCGGTAAATAAGTGTTTTGTTGTGTTTGCACAAGCGAACGGACAATGTTGG
TCAGACCTCAAATATTGTACTCCCCACACTAGGGAGCATTTACGGTGAATATAATTTTTCA
TATTGTGTGTAAAAAAAAAA

SEQ ID NO: 128

CCACTTTCGAAAAACCCGTTTCAAGCCTTTCACGAAAGTCCAACGGTCAGAAAATTCAAA
ATGACTGTTTGAGGCAGAGCCAATCTAGGACCACACTCCATTTATATATGCCCTCTGCTC
CTCTCGACCCTTAGAGTCCTCTGCGAATCTTGTTGTTAGTTACTGTGTACGCTGTAACAA
TGGATGCCTATGAGAAGTTGGAGAAGGTGGGAGAAGGAACCTATGGGAAGGTGTACAA
GGCCAAGGACAAGAACACAGGGCAATTGGTCGCCCTTAAGAAAACAAGGCTGGAGAGC
GACGATGAGGGTATTCCTCCCACCGCTCTCCGCGAGATTTCCCTTCTGCAGATGCTTTC
TCAGGACATCCACATTGTCAGGCTGTTGGATGTAGAACACACAGAGAACAAGAACGGGA
AACCCCTTCTTTATTTGGTTTTTGAATACATGGACTCTGATCTCAAGAAGTACATTGATGG
TTATCGCCGAAGCCACACAAAAGTCCCTCCCAATATTATTAAGAGCTTCATGTACCAGTT
GTGCCAGGGGGTTGCTTACTGCCACAGTCGTGGTGTAATGCACAGAGACCTGAAGCCT
CACAATCTCCTGGTGGATAAGCAAAGAGGTGTAGTAAAAATAGCAGATCTTGGCCTCGG
AAGAGCATTTACAATTCCTATTAAGAAATATACACATGAGATTGTCACTCTCTGGTACAGG
GCTCCTGAGGTTCTTCTGGGGGCTACTCACTACTCTACACCTGTAGACATCTGGTCTGTT
GGTTGTATATTCGCTGAAATGGTCAGATTGCAAGCCCTGTTTATTGGAGATTCTGAAGTA
CAACAACTTTTCAAGATTTTCAGTTTTCTAGGAACTCCCAATGAAGAAATCTGGCCAGGA
GTGACTAAATTCAGGGACTGGCATATCTATCCTCAATGGAAACCCCAAGATATAAGCTCT
GCTGTTCCAGACTTGGAACCAAGTGGTGTAGACCTGTTGTCTAAAATGCTGGTTTATGAG
CCATCCAAACGAATATCAGCTAAAAAGGCATTGGAACATCCCTATTTTGATGATCTGGAT
AAATCTCAGTTCTGAGATCGCTTTTAAGGATTTAAGTAATTCAGGATAGTATATATAAATTA
CTGCATGCATTGGTGGTTTTGGAAATGTCCCATTTTCTACGAATTCATCTTTGCAAAGGAA
AATATATTAGCCAGTATGTGTTTCTTTTAATGTGTATGGGCAGGCAAACAAGCAATGTTTG
GTTGCTGTTGTGGACCTTTTAAATATTGTATTCCCACACTGGAGAGTAACTTTCTTTGATG
ATCTGGATAAATCTCAGTTCTGAGATCGCTTTTAAGGATTTAAGTAATTCAGGATAGTATA
TATAAATTACTGCATGCATTGGTGGTTTTGGAAATGTCCCATTTTCTACGAATTCATCTTT
GCAAAGGCAAATATATTAGCCAGTATGTGTTTCTTTTAATTTGTATGGGCAGGCAAACAA
GCAATGTTTGGTTGCTGTTGTGGACCTTTTAAATATTGTATTCCCACACTGGAGAGTAACT
TTCTTTAAAATTAATGTTAAGATGTGATATTAAAAAAAAAA

SEQ ID NO: 129
GATCGATCTATATCTCTGTGCTATAAACAAAGCAAAACAATGGATTCTTATGAGAAACTGG
AGAAAGTGGGAGAAGGAACCTATGGGAAGGTCTACAAGGCTAAGGACAAGAAAACAGG
GAAACTGGTTGCCCTTAAAAAGACCAGGCTGGAGAATGATGGCGAGGGAATTCCTCCCA
CAGCCCTTCGCGAGATTTCTCTCCTGCAGATGCTTTCTCAAGATATGAACATTGTCAGGC
TGCTGGATGTGGAACACACTGAGAACAAGAATGGGAAGCCGCTTCTGTACTTGGTTTTT
GAATACATGGACTCTGATCTCAAGAAGTACGTTGATGGTTATCGCCGCAGCCACACAAAA
ATGCCCCCCAAGATTATCAAGAGCTTCATGTACCAGTTGTGCCAGGGGGTTGCCTACTG
TCACAGTCGCGGTGTGATGCACAGGGACTTGAAGCCTCACAACCTGCTGGTCGACAAG
CAAAGGGGTGTGCTAAAAATAGCAGATCTTGGCCTGGGAAGGGCTTTCACAGTTCCTAT
CAAGAAATACACACATGAGATTGTGACCCTTTGGTACAGGGCTCCTGAAGTGCTCTTGG
GGGCTACTCACTACTCTACACCTGTTGACATTTGGTCTGTTGGCTGTATATTTGCTGAAA
TGTCCCGAATGCATGCTCTGTTCTGTGGAGATTCTGAAGTGCAACAACTTATGAGCATTT
TCAAGTTTCTAGGAACTCCAAATGAAGGAGTTTGGCCGGGAGTGACCAAATTGAAGGAC
TGGCATATCTATCCTGAATGGAGACCTCAAGATTTAAGTCGTGCTGTTCCAGACTTGGAA
CCAAGTGGGGTAGACCTATTGACTAAAATGCTGGTTTATGAGCCCTCCAAAAGAATCTCA
GCTAAGAAGGCATTGCAACATCCTTATTTTGATGACCTGGATAAATCTCAATTCTGAGATT
CCTTTTAAGGATTTAGGCATTTTAAGGATTTGTCATATTTGGGGGTTTTGGAGATCTTCCA
TTTCTGAGATTTTCATCTTTGCACAAAGGCAAACATAAGCCCATATAAATAACTTGAAGTG
TTTGCACAAGCAAACAAGTAGTGGGAGCTTTCCCAAATATTGTTTTCCCACACCTGGGAG
CTTTTGTCATGAGCATTTATGCTCAGATTTAACATGGCCCTTCATGCTTAAACCTGTTGTT
GTTTGAAAATAATTTTAGAAAATGTGAAGTTGAAGCATGTTTTGAATTTATGGTGGTGGCA
TGTGGATATTTGAACTTGGTTGAGAAAAATTGAAACATCTTTGTTAGGGAAAAAAAAAA

SEQ ID NO: 130
TAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGGTATTCC
AGATATCTATATGTACGGAACACCCTCCCCGCCGAGAAAATTAAGGATGATCGTCGACT
GCGTAAGAGGCGGCAAAATTCGAAATTTGTTTATATCATCGCTATTCAGATGCATGCCTG
CACTAGCTTACGAGAAATAAATTTCTTAGCTCTGCATCAACAAACTAAACATGGAGAAATA
CGAGAAGTTAGAGAAGGTAGGGGAGGGAACCTATGGTAAAGTGTACAAGGGAAGAGAC
AAACGCACTGGAAGACTGGTGGCCCTCAAGAAAACCCCCTTTCACCAGGAAGAGGGCAT
TCCTCCCACTGCCATTCGGGAGATTTCTCTTCTCAAAAGCCTCTCGCAATGCATATACAT

TGTCAAGTTGTTGGATGTAAAGGCTTCATTTAATGGCAAAGGAAAGCACGTACTGTTTAT
GGTATTTGAATATGCAGATTCTGATCTCAAAAAGCACATTGATGCACACCGCCAATGCAA
TACCAAGTTGTCTCCAAGGTCTATTCAGAGCTATATGTTCCAATTATGTAAGGGTATTGCC
TATTGCCACAGCCACGGGGTGCTCCACAGGGATCTGAAGCCACAGAATATTTTGGTAGA
TCAAAAAATAGGGTTGCTGAAAATTGCAGATCTTGGACTTGGAAGAGCTTGCACAGTACC
TATCAAGAGCTATACTTTTGAGGTTGTTACTCTTTGGTACAGAGCTCCTGAAGTGCTGTT
GGGTGCCAAGCGCTATTCTATGGCATTAGACATATGGTCTCTTGGTTGCATCTTTGCTGA
ACTATGTAATCTGCAAGCACTTTTTGCTGGAGATTCTCAAATACAGCAGCTTATAAACATA
TTCAGGTTGCTGGGAACTCCTAATGAACAGCTATGGCCAGGTGTGACCCAGCTAAGCGA
CTGGCATGAATTTCCTCAATGGAGGCCTCAAGATCTTTCCAAAGTCGTGTTCAATCTGGA
TCCAAATGGTGTGGATCTTCTTTCTAAAATGTTGCAGTATGATCCTGCGAAGAGGATCTC
GGCAAAAGAAGCACTAGACCATCCATACTTCGACAGTTTAGACAAATCACAATTTTGACA
TCAGTTTTTTTATCCATGGAGTTGATTGGAAGCGGTTAAAAGGCATGCATCTGAGTGGTC
AGCATAAGCACATAATAAAGGAGAACAATAGGTAACTTCTGTACTGTTCATGAAACTTAAT
GGGAGGAAAGGTTTTCCAGTGTTTTCTGACATCTCGATTTCAGCACTGAATTGTGGCATT
ATATAAATTCAGGGAGGTTCATAGAGTTGCATTTTTGCATCTCTCAAGGCTATTTTTAGAA
AGTAGTCATTCCTATTGAAGGGTCAACCTTTAATTTTGGCTAGCAGGACTGTATAGGATT
ATATGCATACAAGATGACAAATAATTGTTATAAATCCACAATGTGAACAGTTATTTTTGAT
GGTGCCAATGAAGCTGTGCGCCTATGGAAAAAGCATACAGGGCGACAAATAATTATTAC
AAGTTCACAATGTAAGACACTTATTTTTGATGGCTACTGAAGCTGTGCGCCTACTAGAAA
TCGTTGATATGCTACACATGGCCAGAAAGCTTACCATAAGATTTCTTTTTTGGTTAAGATT
CTGGAATGGTTTAGGAATCTCTAGCTAGGGAGCTAAGCCTTAAATATGCTGGCAATTGTT
GTAATTGTGAACAAACACAGGCAGTTTCCACATTATTATTGTCATTTATCTTTCAATTCCC
ATTTGGTCTTCAATATTATTTAGTAATTTCCTTGTGTAAACTACGTTGTGGTGATCTTCTAC
CTGGCAGTCATTAGGGATATAGAAGTCAACAGGCTAAGGCCGGATTCTGCATATCTTTTT
CCATCTTAGAATGTTATTAAACTTGGCCTCATCCAAAGGACTCGGTATATTTGAGGGAAT
CGCAAGGACCCTGCCCCTAGTGGGAAGAGCTTTTGGTTGGTGAAAGATGCCATTCCCAA
CTGCAGAGTGATGTAAGGTTTGAATCACTTAGTGAAAGATTCTCATTAACTTAAAAAAAAA
A

SEQ ID NO: 131
ATTTGTTTTCTCTTTTGGGTTTAAATTCCATGTTTTCGAGGATTTTTTGGGGGTTTTAGGG
TTGTCAGACGAGCGCTTTTGGAGTGGGTTTTGTAATTATTCTAATGGGGTGCGTTTGCGG
CAAACCCTCTGCGAGGGCCGCGGATTATGTAGAGAGCCCTGCAGAGAAGGGGGCATCC
TCCAATAGCCGATCTTCTTCAATGGCGTCTCGGCGGTTGGTAGCCCCCGCTGTCATGGA
CCAGGGCATTGATGCCGAGAACGGACACGAGGGGGATTATAGGACTAAATTGAGAGGA
AAACAGAGCAATGGTGCTGACCCAGTTTCATTGTTGTCGGACGACGCTGAAAAGCAACG
GCACTCTCGGCACCATCAGCATCAGCAGCATCATCCTATTCGTCCTCATCATCTGCGGC
CACAGGGGGAATTCGTTCCCAACGCCAATTCTAATCCCAGGTTCGGGAATCCTCCCAGG
CACATCGAAGGCGAGCAGGTTGCTGCAGGATGGCCAGCCTGGCTTACTGCTGTGGCTG
GCGAAGCGATCAAAGGCTGGATCCCACGCCGGGCCGACTCTTTTGAGAAGCTCGATAA
GATTGGACAAGGAACTTACAGCAATGTGTATAAAGCACGTGATTTAGATACTGGAAAAAT
TGTTGCCCTAAAGAAGGTGCGGTTTGACAATTTGGAGCCTGAAAGTGTGCGCTTTATGG
CCAGAGAGATACAGGTTTTACGTAGACTTGACCATCCAAATGTAGTGAAGTTGGAAGGAT
TGGTCACATCAAGGATGTCCTGCAGCTTGTACCTTGTATTTGAATACATGGATCATGACC
TTGCTGGCCTTGCTGCATGTCCCGGTATAAAGTTCACAGAACCGCAGGTGAAATGTTATA
TGCAACAATTGCTTCGAGGTCTTGATCATTGCCATAGCCGTGGTGTGCTACACCGTGATA
TCAAGGGTTCAAATCTTTTGATTGACAATGGCGGTATTCTAAAGATAGCTGACTTTGGCC
TAGCTACCTTCTTTCACCCTGATCAGAGGCAGCCCTTGACAAGTCGTGTTGTAACACTTT
GGTATCGACCTCCGGAACTTTTACTGGGTGCTACAGAGTATGGAGTTGCCGTGGATTTG
TGGAGCACAGGTTGCATACTTGCAGAGTTGCTTGCTGGAAAACCTATCATGCCAGGAAG
AACAGAGGTAGAACAACTGCACAAAATTTTTAAATTGTGTGGTTCACCATCTGAAGATTAT
TGGAAAAAGTCAAAATTGCCACATGCAACTATCTTCAAGCCACAGCAGCCATATAAACGC
TGTGTTGCAGAGACGTTTAAAGATTTTCCACCATCGGCTCTGGCACTGATGGAGGTTCTT
CTTGCCATAGAACCTGCTGATCGTGGAACTGCCACTTCAGCATTAAAGAGTGATTTCTTT
ACCACCAAACCACTCGCTTGTGATCCTTCAAGTTTGCCAAAGTACCCACCAAGCAAGGA
GTTTGATGCAAAAATTCGTGATGAGGAGGCAAGAAGGCAAAGAGCAGCGGGAGGAAGA
GGGCGTGATGCAGCTAGGCGCCCATCACGAGAATCAAGAGCAATTCCAGCACCAGAAG
CAAATGCTGAATTAGCAATTTCCATACAGAAGCGGCGCTTAAGTTCACAAGGGCCTTCTA
AAAGCAAGAGTGAGAAATTCAATCCCCAGCAGGAAGATGGTGCTGTGGGATTTCCTATT

GAGCCTCCAAGGCCTATGCATATTGGCATTGATGCAGGTGCCACTTCTCGCATGTATTCT
CAACAATTTGGGCCTTCTCATTCTGGTCCATTATCAAATCAAATTTCTAGTTCAATATGGG
GAAAGAATCAGAAAGAGGACGAGATACAAATGGCTCCAGGTCGTCCATCGCGGTCCTCA
AAAGCCACAATATCTGATTTCAGAAAACCAGGGGCCTGTGCACCCCAACCTGGAGCAGA
TTTGTCACATTTATCCAGTTTAGTCGCAACAGCAAGAAGTAATGCTGGTATAGATACACAT
AAGGACCGTAGTGGCATGTGGCAACATAATCGTATTGATGCAATAGATGGTGTACATAAT
AATGGGAAGCATGAATTTCTTGAAGTTCCAGAACATCCAAACAGACAAGATTGGACTCGG
TTCCAGCAGCCAGAATCATTTAAAGGTTTAGATAATTATCACTTGCAGGATCTGCCAGCA
ACTCACCATCGCAAAGATGAAAGGGTTGCTAGTAAAGAAGCTACCATGAACTGGCAGGG
TTATGGTGGTCAAGGGGGGGACAAAATACATTACTCGGGCCCATTGCTTCCGCCTTCTG
GAAATATTGATGAAATTTTGAAAGAGCATGAACGACACATTCAGCATGCTGTGCGTAGAG
CTCGGCAGGACAAGGGTAGACCACAGAGAAGCAATTTATCACAGAACGAGAGGAAAGC
ATTTGAACACAGAAGTTTTGTTTCTGGGGTGAATGGAAATGCAGGGTATTCTGATCTTGT
AAATGAATTGCCCATTTCAGTAGGTAGTAATAGGTTGAAAGTGAGCAAGACAAGAGGGA
CTGAAGAAATAGTTGAGCTGAGGGAGTTGGAGAGGGAACCCCTCTCATCAGTAATGGAG
AAGTATGAAAGAGAGCATGAAATGTGATACCTCCTGTGTAGTCTTGTTTTTCTTCATTTTA
GGTTCTGCAGGTGAGATAGAAGAGATGAGACTGTTCTGCTGTGCATATACCTCAACTGG
TTAATTGCATCTCAGAAAGAACTTTCAACATCTGGTTTTGTATGAAAGAGGCAGCGGAGT
TGAGAGCTTAATTTAGTTTAAGCTCAAGGTTCAGAGAATTCTTGCAGGACTGGGACCTGG
GGTATTTGAAAAACAACATATAGGTTTTCCCTATGAAAATGTACGGAATTTATTGTATTTTA
GATTCTGGAGGGCCATCTAAACTTCTGGCTGCTTGGTGCAACATTGATGGACTGATATAG
TTAAGGTCTTGACCTTTGTTGCATTTTGCTTTGCATGCAACCATGGTTGTAGATCTCATGC
ATGTACGAATGTAGAGGACCTGGTCTCAAAATGTTATTTGCTGCTTTGTTTTACCCTTGAA
GTGAAATGCAGGCAAGCTTCACCAAGGTATTCCCCCTAATATCTTTTCAGAACTTTTGGA
TTAAAAAAAAAAA

SEQ ID NO: 132
TCTTTCCATACTCAATGTTTTTTCCATACTATGGAATTGAATTCTGTTGAAAAATATTGCTAA
ATACAAAATAGTATTAGTAATAATTATTTTAAATTCCGCTTTTATTCGTATTTATAACTATTA
GCCAGCAGAGCCGAACGCATGCAGGAGTCTCTCTCCTTTTGCCTTCGAAGGTTTCGCTC
AGATGGGCTGCGTCTGTGCCAAACAATCCGACATTCTCGGTGAACCAGAATCTCCCAAG
GTCAAGGGTTCGAATCTCGCCTCCAGCAGGTGGTCGGTCTCCTCCGAAACAAAACAACT
GCCGCAACATTCTGATTCTGGAATCCTGCATCATCAGCATTATTACCACCCTCGAGACGA
ATCCGACGAAGCCAAATTGAAAGAGAGCAACTATGGTGGATCGAAGAGGAGAACAAGGC
AGGGAAGGGATCCCGCTGACTTGGATATGGGCATCTTCGTCCGCACTCCTTCCAGCCAA
TCAGAGGCCGAGCTGGTGGCAGCTGGATGGCCGGCCTGGATGGCAGCTTTTGCAGGG
GAGGCCATCCATGGCTGGATCCCTCGCAGGGCGGAGTCCTTCGAGAAATTGTACAAGAT
TGGACAAGGGACTTACAGTAATGTGTATAAAGCTCGTGATCTTGATAATGGAAAAATTGT
TGCCCTGAAGAAGGTACGTTTTGACAGTTTGGATGCTGAAAGTGTGCGATTTATGGCAC
GAGAAATACTGGTTTTACGCAAACTTGATCATCCAAATATTGTCAAATTGGAAGGACTTGT
TACTTCAGAGGTATCCTCTAGTCTGTACCTTGTATTTGAGTACATGGAGCATGACCTTGC
TGGACTTGCTGCTTGCCCGGGGATCAAGTTCACTGAACCACAGGTTAAATGTTATATGCA
ACAATTACTTCAAGGACTTGATCACTGTCACAGACATGGTGTACTCCATCGTGATATCAA
GGGTTCAAACCTTTTAATTGACAATGGAGGCATTTTAAAGATAGCTGACTTTGGTCTAGC
AACTTTCTTTTATCCTGATCAGAAACAGCTCCTGACAAGTCGTGTTGTAACACTTTGGTAC
CGGCCTCCAGAACTTTTGCTTGGTGCTACAGATTATGGAGTTGCTGTGGATATATGGAGT
GCTGGTTGCATACTTGCTGAACTGCTTGCTGGCAAGCCTATCTTGCCCGGAAGAACAGA
GGTGGAACAACTGCACAAAATATTTAAATTGTGTGGATCACCATCTGAGGACTATTGGAA
GGAGTCAAAATTACCACATGCAACCATATTCAAGCCACAACACCCTTACAAAAGTTGCAT
TGCTGAGGCTTTCAAAGATTTCTCTCCATCAGCTTTGGCCTTGTTAGAAACTCTCCTTGCT
ATAGAACCTGGTCATCGTGGAGAAGCAAGTGGGGCCCTTAAGAGTGAATTCTTTACAAC
GGAGCCGCTTTCTTGTGATCCATCAAGCTTACCTAAATACCCGCCAAGCAAAGAGTTTGA
TGCAAAATTGCGTGCTCAAGAAACAAGAAGGCAAAGAGATGTGGGTGTGAGAGGTCATG
GATCTGAGGCAGCAAGGAGAACGTCCCGACTATCTAGAGCAGGTCCAACACCAAATGAA
GGTGCTGAATTAACAGCATTAACTCAGAAGCAGCATTCGACTTCTCATGCAACTTCAAAC
ATTGGAAGTGAAAAACCAAGCACTAAGAAGGAAGATTACACTGCTGGATTGCATATCGAT
CCTCCAAGGCCTGTCAATCATTCTTATGAAACAACTGGTGTTTCACGTGCATATGATGCA
ATTCGTGGGGTTGCTTATTCTGGCCCATTGTCACAGACACATGTAAGTGGTTCAACATCA
GGAAAGAAGCCAAAAAGAGATCATGTAAAGGGACTTTCAGGTCAATCATCTTTGCAACCA
TCAAAACCTTTTATAGTTTCTGACTCAAGATCAGAGAGAATCTATGAAAAAAGCCATGTAA

CTGATTTGTCAAATCATTCAAGACTAGCAGTAGGAAGAAACCGTGATACTACAGACCCAC
ACAAAAGTTTGAGTACTCTGATGCAACAAATCCAGGATGGTACATTAGATGGAATAGATA
TTGGCACACATGAATATGCAAGGGCTCCAGTTTCTTCAACAAAGCAAAAATCAGCTCAAT
TGCAAAGACCGTCAGCATTGAAATATGTAGATAATGTTCAACTTCAGAATACACGTGTAG
GAAGTCGCCAAAGTGATGAAAGACCTGCCAATAAAGAATCTGATATGGTATCTCATCGTC
AGGGGCAGAGAATTCACTGCTCGGGACCTCTGCTGCACCCATCTGCCAACATTGAAGAC
CTTTTACAAAAGCATGAGCAACAAATCCAACAGGCTGTACGCAGAGCACACCATGGTAAA
CGTGAAGCTCTAAGTAACAAATCATCTCTCCCTGGAAAGAAACCAGTGGACCATAGAGCT
TGGGTTTCTTCTGGAAAAGGAAACAAAGAATCACCATATTTTAAAGGAAAAGGGAACAAA
GAATTGTCAGATCTTAAAGGGGGACCAACCGCCAAAGTAACAAACTTTAGGCAGAAGGT
AATGTAAAGTATAGCTAAGGAAATTGCAGATGAAGGGATTCAGAAAGAGAACCCCTCCA
GTCAGGCACAAAAAGATATGAACAAAGAAAAATACTTGCTACATGTCTTCTAGGGTCATA
TTCTGGTCTCTCTAGTTGCTGACGTCAATTTATGCATGGATTGGTTGATTTGGGATGGGA
GTTTTTTATTTTCTAAGCACTGGGATTGCTTCAGCGGGCAAATCAATAATCACGTACCTTA
TAAAATAGGGTATCCTCAATTTTTTTTTCTTTATTTATAACTGCGAGGGTTTATGGGATCTT
TTAACTCTGCAGAAAACTTATACAGGGAGTTTCAAACCATCAGATAAGTCTTTTGCATTTA
AGATTATTGATCACCTTCTAAAGAAGTCACAATTGTCTTGCAACTGCCCATTAAAACGTTG
AAGGACTATTTGGTTTTATCTCGAGATCTCTGTGATGCAACAACTGACTATATGGTCACTC
TTATGTTGCAGCGAATGCGAATGGCTCTAGCCACATTGTAATGATCTGTTTCAGTATTTTC
TGTACAGTGGAAGAATAAAGTTATGCCAAAGTTCTTTCAAGAACTTCCATTAGACTTTTGT
ATTTGTTCATACATGATTCGATAGAATAAGAAATCATGCATAAGAGACTTTCTTTTAAGGG
AAACATCTATGCCTGGAAGATTGAGATGAAGGGTTCTATTTTAATAGATATGGGATTGGT
TGCAAATGTAAGGAGCCAGGTTCTTCCATTGGCCTGGAGGATTTTGTATGGGACCCTTA
CATGCTGCTGATGAAGACCTATCACTTGGCTATGCTTGGCATTAAAACGAACTGAAAAGA
GAGGCGGCACTGAGGAGTCAAGCCACATCTGAGAAAAGACTTTAAATTCAGCATGCACT
CATTGTGAAGATTGTTATTATGGGTGAGTCCTAAATGGTTATGAACGCTGTATTTGTCTGC
AAGGATAGCGTTCTTGACAAGGGCCAAATCTTTCCATGTTTGGTTTACAGAAGTTTGTATT
TTAATTTTGTTCATTGCATTCTGTTGCAGTGTTTCAATTGAATCCGTGTCAGAGGAAGCAG
CACAATATGTAGGTTTTGTGGAGTTAATCTTTTAGTTTTAGGAAGCAAAGTTTTCTACATTT
CAGAGTAGGGCTTTCCCTCTACAGTTTTGAGTGTAGTGTGTTTCTTTTATATCTTGCAATC
AAAGAAATCATAGCTAATGATTCCATGCTATCCATGGTATCTACTTCACGATAATAAAGGT
CTTAGTCCACAAGTTTGATGTGGTATACAATGGAGCAAGAGAATTGCACAATGAATATCA
CATGACCTTGATTATTTCTAATGAACAATAAATGGTATGGTGGTTTTCATTATGAAAACCT
CAAAACTTTTATATGGAGTATCTTCATAACATTTTAAATGAGCAATAAATGATATGCTTCAA
AACTTTTATAAAAAAAAAA

SEQ ID NO: 133
GTCCTTCATTGTTTAAAGCCGATTGATCTATATTTGTTAAGAACAAAGAAGATTCTGTACT
TCTGCATATTTATACAAATGCTGGAGCTCAAGTATTGATTGCATCGGAGACAACCCATCT
GTTGAATAGGCTGTGAATGTGCCAGCGGACCTTCCCTCGCTTGAATTTTTTCGTTTAGCA
GTTGAGTCTGAGGCCGATTTGGAAGAAAGCCGAGGAGGATGGCAGTCGCAAATCCTGG
TCAGCTGAATCTGCAGGAGGCGCCCTCATGGGGTTCTCGCAGCGTCAATTGCTTTGAAA
AGCTTGAACAGATTGGAGAAGGCACATATGGGCAAGTTTACATGGCCAAGGAGATCGAG
ACTGGGGAAATTGTTGCCCTGAAGAAAATTCGTATGGACAATGAAAGAGAAGGGTTCCC
AATAACCGCCATTCGGGAAATCAAGCTTCTGAAGAAGTTGCAGCATGAAAATGTTATCAA
GTTAAAGGAAATTGTGACCTCTCCAGGTCCTGAGAAGGATGAACAAGGGAAATCAGATG
GTAATAAGTACAATGGAAGCATTTACATGGTCTTTGAATATATGGATCATGATCTGACTGG
TTTAGCTGAGAGACCAGGAATGCGCTTTAGTGTTCCCCAGATTAAGTGTTACATGAAGCA
ATTGTTAATCGGGCTGCACTATTGTCATATCAACCAAGTTTTGCACCGGGACATCAAAGG
ATCTAATCTGCTGATCGATAATAATGGGATCCTAAAGCTTGCCGATTTTGGCCTGGCAAG
ATCATTCTGCAGTGACCAGAATGGAAACCTGACCAATAGAGTAATAACATTGTGGTATAG
GCCCCCAGAGTTGCTGCTAGGCTCAACAAAATATGGTCCAGCTGTTGACATGTGGTCAG
TGGGATGTATATTTGCGGAGCTTTTATATGGAAAACCTATTTTACCAGGAAAGAATGAGC
CAGAGCAACTCACTAAAATTTTTGAGCTTTGTGGATCACCGGATGAGTCCAACTGGCCG
GGTGTTTCCAAGCTTCCATGGTACAGCAATTTCAAGCCGCAAAGGCAAATGAAGAGGCG
TGTTAGGGAATCTTTTAAAAATTTTGACAGACATGCTTTGGATCTTGTGGAAAAAATGCTC
ACTTTGGATCCTAGTCAGAGAATAAGTGCAAAGGATGCACTTGATGCTGAATATTTCTGG
ACAGATCCGGTTCCTTGTGCACCAAGCAGCTTGCCAAGGTATGAACCATCACATGATTTT
CAAACAAAGAGGAAACGGCAGCAACAAAGGCAACATGATGAAATGACCAAGAGACAGAA
AATTTCACAGCATCCTCCACAACAGCATGTTAGACTACCTCCTATACAAAATGCTGGACA

AGGTCATCTTCCTCTAAGGCCAGGTCCAAATCCAACTATGCATAATCCACCTCCCCAGTT
TCCAGTTGGACCTAGTCATTATACGGGAGGACCTAGGGGAGCTGGTGGACAGAATAGG
CATCCTCAGAATATACGCCCACTCCATGCAGCTCAAGGGGGAGGATATAATGCAAATCG
AGGATATGGGGGCCCTCCTCAACAGCAAGGAGGAGGATATCCCCCACATGGGATGGGA
AATCAAGGACCAAGAGGTGGACAATTTGGGGGTAGAGGTGCAGGCTATTCTCAAGGAG
GTCCTTATGGTGGACCTGTCGGGGGCCGAGGTCCAAATGTTGGTGGAGGCAATCGAGG
TCCACAGTTTTGGTCGGAACAGTGAATCGATTGTGGAGTACACATTTAGCTTGCATGTGA
TACTATTAAATGCAAGTGTCTTTAGTTTGAAATTGTGAAATCATCCGAGCATGGCCGGATA
CACCTTTTTAAATTCAGGAATTATTCACTTGATCAGTACGTTCTGCTCACCCCTACAGCAA
AGCTACAGTTTTTTCTTACATCTAAGATTTTTCTTCTTTTTAGCGGGGTTACAAACCCAAC
AGGTGGTGTTCCTTATGGCTGTCAAGCATACATGCAGCAGCTCATTCAGCCTACTGATTT
GGGTGGTGCATATGCTCATTACTCCATTACAGGGATTCAAATGCCAACCATGTCACCACC
TACCACCTAATAGGCCTGAGTATTGCTCACCACTATGCTGATATGGGGAGCAATAACGTT
AGTAAATTTGTGAATTTTGAAGATTTTCTTCTAGCAAGTCATATTGTTGGGGCCTGTTAAA
TGAACTTTCTGGTGTATCCGGACAATTTAGGGATGTGTGTTTCGATTACTACCAAATGTA
CATTAGAATGCCATTTTCTTGATTTAGGTTGTTTTTTTAAGAAGTGACCAAATTCACTTTGA
CATTGTCTTGGCTAATCATGATGACTGATATTCTCTCTGTATTATTTAAAGACAACTCTCC
AGAAATATGACTCGAAGGGATTCCGTAAGAAAAAAAAAA

SEQ ID NO: 134
CTTACAGTTTTGACATTCTCATAACGGGAATGTTCGAACTGTTAATACCATTTAATCACCG
AGTGACTTTAACGCCATTTCAGTCAGCAAAGATTTCTCAATTTGTTTATCAGAATTGGGTC
GGTTCGGTCATAACTCCTGCATTTCTTGTTTGTGTTTGATCTATAAGGCTACTCAGATGTT
AAATATATTTGATTCAGTCTACTCTGTTCTGAGGGCAAATGAAACTCTCCGGGCTTATGAA
GGTTGGTTTGAATTGAGGAATTTGGGATATATAAACACTGGCAATGTTCTCCGGACCATG
GAATTGGGGGTTTCTCCATAGGGTGTTTCAGGGTTTTGAGACTTGAATACCTCGAAAAGA
TTTGATGCGCGGGTTTTGAGAGGTAAAATCTCGAAAAGATTTGATGGGAACAGGTCTTCC
GGAGATTTGAGGAATGCAAAACATGGAGGACAATGTGCAGAGCAGCTGGAGTCTCCAC
GGCAACAAAGAAATATGTGCCCGTTACGAGATTCTAGAGAGGGTCGGCAGCGGAACTTA
TTCGGATGTGTACAGAGGGCGGAGAAAGGCGGACGGTCTAATAGTAGCCCTTAAGGAG
GTGCATGACTATCAAAGCTCTTGGAGAGAGATTGAGGCACTGCAGAGGCTTTGTGGGTG
CCCCAATGTTGTGAGGCTCTATGAGTGGTTTTGGAGAGAAAATGAAGACGCAGTTCTGG
TTTTGGAGTTTCTGCCTTCCGATCTCTATTCTGTCATTAAGTCTGGTAAGAATAAGGGAGA
AAATGGTATTCCCGAGGCCGAGGTTAAGGCCTGGATGATTCAGATCTTGCAAGGGCTGG
CTGACTGCCATGCCAATTGGGTTATCCATCGTGACCTAAAGCCCTCTAATCTGCTGATTT
CGGCTGATGGAATTCTCAAGCTCGCTGATTTTGGACAGGCAAGGATACTTGAAGAGCCT
GAAGCGATCTATGAAGTAGAGTATGAACTTCCTCAAGAGGATATCGTTGCTGATGCCCCA
GGAGAAAGGTTGATGGAGGAAGATGATAGTGTGAAGGGAGTGCGGAATGAAGGGGAGG
AGGATTCATCCACTGCAGTTGAAACTAATTTTGGTGATATGGCAGAAACTGCGAATTTGG
ATTTGAGCTGGAAAAATGAAGGTGATATGGTGATGCAGGGATTCACATCTGGTGTTGGA
ACTCGATGGTACAGAGCTCCGGAGTTGCTCTATGGAGCAACGATCTATGGAAAAGAAAT
TGATTTGTGGTCGTTAGGTTGCATTCTGGGGGAGCTCTTGATTTTAGAACCTCTCTTTTCT
GGGACTTCAGACATTGATCAACTTAGCAGGTTGGTTAAAGTTCTTGGGACTCCAACAGAA
GAAAATTGGCCTGGATGCTCCAATCTTCCTGATTATAGGAAACTTTGTTTTCCTGGTGAT
GGAAGCCCCGTTGGTCTGAAGAACCATGTCCCCAGTTGCTCAGACAGCGTGTTTTCTAT
TTTGGAAAGACTTGTTTGCTATGACCCTGCAGCTAGGCTAAATGCTAAAGAGGTACTTGA
GAATAAGTATTTTGTTGAGGATCCTTATCCTGTCCTTACCCATGAATTGAGAGTTCCCTCA
CCTCTGAGGGAAGAAAACAATTTTTCAGAGGATTGGGCGAAATGGAAGGATATGGAAGC
AGATTCTGACTTGGAAAACATTGATGAGTTCAATGTTGTTCACTCAAGTGATGGTTTCTGC
ATTAAATTTTCATAAACTCTAATTGGCTGATCATGTTTCATCTTCCTCAAATTGCCTTTTAT
ATGGATTAGTATTCTAAATTAGCTTTGGGGGCTGCATGTCTTCAAGCATTCACCACGACA
GTAAAGTAATCATTATGATTACTGATGTATTGCTGTCACTGGGTGAGAAACATTATAGCCA
TGAACCTGCTGTTAAATTTGTAGAATTGTCTTATACTTGCTCAAAATAGTGACTGCATGTT
TGAGAATAAATTTCTCTGCAAAGACCAACTCCAAATTCTTTCTTTTCACTTTGTACTAATGA
TCATTGTGACCACAAAATCTTTATACACAATACAGAAAATGCAAATTTAGAGAAAAAAAAA
A

SEQ ID NO: 135
CTCCGGCACAGAAAGCCAAGATGTTGATCGCCTCCCACGCTCGCACTTGACCGTTCTCA
CGAGAAATCTCATGTCCATTCCTTCCGTGCGAGCCGACAGAGGACCCCAAGAGAAGCAC

AATATTTTTGCTATAACTAATAATGCCATTAACATCCGCTGTTGGGACGCCCGGAAGTAA
ATGATCCATCGTTGCGCGAAGAGCAAAGTCAGTTAAAGCGCTAACAACCATATCAGTAAC
GATGGACTTGAACCAGTACCCTGAAGACTTGAACCCGGAACTACCCGAGGGTACAGACA
ACGTAGATAATCCTGATAATAACAAAGGCTCCCCTGTTCCTTCTCCCCATCCTCCTCTGA
AGCCACTCGATCCCTCTGAACGGTACCGCAAAGGTATCACCCTTGGGCAAGGTACCTAT
GGGATCGTCTACAAGGCCTTCGATACTGTGACAAACAAAACTGTGGCAGTAAAGAAAAT
CCATCTGGGCAAAGCCAAAGAAGGTGTTAATGTAACTGCTTTAAGAGAAATCAAATTATT
GAAGGAGCTTTCACATCCAAATATAATTCAGCTAATAGATGCATATCCACACAAACAGAAT
TTACATATTGTCTTTGAATTCATGGAGACTGATCTTGAGGCTGTTATAAAAGACAGAAACT
TAGTTTTCTCCCCAGCTGATATAAAATCCTATTTGCAAATGACGTTGAAGGGGCTTGCTG
TTTGCCATAAAAAATGGGTTTTGCACAGGGATATGAAACCAAATAACTTGTTGATTGCTG
CTGACGGCCAACTGAAGTTAGGAGATTTTGGCCTGGCTCGCTTGTTTGGTAGTCCAGAC
CGCAAATTTACTCATCAGGTTTTTGCAGTTTGGTACAGGGCACCGGAGTTACTGTTTGGA
GCTAAACAGTATGGTCCAGCTGTAGATATTTGGGCAACAGGTTGTATTTTTGCAGAACTT
CTGCTCCGAAAACCGTTTCTTCAGGGTGTGAGTGATCTCGACCAGATAGGGAAGATTTTT
GCAGCCTTTGGAACTCCTAGACAGTCACAGTGGCCAGATGTGGCCTCTCTTCCAGATTT
TGTTGAATTCCAGTTTGTTCCTGCACCATCGCTACGTTCTTTATTTCCCATGGCTAGTGAA
GATGCCTTGGATCTTCTGTCAAAGATGTTTACGTTGGATCCAAAAAATAGGATTACTGCA
CAGCAAGCATTGGAACACAGATATTTTTCTTCTGTGCCTGCTCCCACGAGGCCAGATTTG
CTTCCCAAACCCTCTAAGGTGGACTCATCAAGGCCCCCAAAGCATGCATCTCCAGATGG
TCCTGTAGTGCTGTCTCCTTCTAAGGCAAGGCGAGTAATGCTCTTTCCAAATAATTTGGC
TGGAATTCTACCAAAGCAGGTTTCCCAATCTACAACTGGAGGGACACCAATTGAATTTGA
CATGCCGACACAAAAGCTACGTGAAGTTTGCCCTAGGTCCAGAATTACTGAATCTGGCA
AGAAACATTTGAAGAGGAAAACAATGGACATGTCTGCTGCATTGGATGAATGCGCAAGG
GAGCAAGAAGGGCAAGAAGGTAAAACCATTTTGGACCCCGACCATCAGCGTTCTGCCAA
AAAAGAGAAGCATATGTAATTTCACTTGATTTCACTTTGGGTGATCATATTGTTAGGCAAG
GAGACTAGGCATTTTCCTTGTCACTATCCTCATATTGATATCACCTCGTGTATGTTGTGG
GGTGGCAAAATTACTTCCATTATTGTTGATATCCTTTATGTTGAGAAGGTGGATTCATTTT
TAAGCAGTGGTCTCAAAAAATGACCCTGCAAGAACTGGAACTGCCAAATTTGTAAGCAAT
ATACTGTTGAGAAAAATGTTTTGCATCATCTTGAGTTCTCATTATATGACAAATTTCGGAG
GAGAAAAAGATTGGTTTCTAAAAAAAAAA

SEQ ID NO: 136
GTAAAGATCCTCTATCTCTTCCCCTATTTGTTCCTCGCCACAGTACGTTATTTCGGCGAAA
TTCAGATGATCGGTCTCAACTATTACAGTGACCGCCATTACTGTCTGCTGTTCCTGTGAA
TGACAAAAGGTTCTTGATGTCAAGAAGTTAAGAAGTCGTGGAAACGGATTACGCGTTAAT
GCTGTTTACAAGGAAGGATTGCTAGTGCTCCGTCTGTTAGCGGATCCTGTTCTTGGTGG
GGAGATTAAAGATGGCAGGGGGACAGGAAAATTGCGTCCGTATTACTCGAGCTCGAGCT
GCCTGTGTTTCGAAAGCTTCTGCCCCTGTGATTCAATCTCAAGTAGATGAGAAGAAATCC
AGGAAAAGAGCTCCTAAGCGAGCTGCAGTGGATGATCTTGCTGCCAATGCAAGCGGTTC
TCAGCCTAAAAGGCGCGCAGTTTTAGGCGATGTCACAAATCTTCATGCTGCTGCTACCG
ATTGCTTGTCAACAGCGGAAGACCAGGTTGATGCTCCCAATCCTAGCATTAAAGGGAGG
GCGCGTAACAAGAAGAAAGAGGCAAGGACTTCAACGAAAGTAGTTAAAGATGAAATTCA
TCCGGAATCGAATCCTCTTGCTGATCATAGTTCGAATCTTTCGGAATGTCAAAAGCCTCC
CGCTGCAAAATTAGCAGAGCAACGATCATTAAGAGGCGTACCTTCAAAGGCAAAGCAAG
GGGGAAGCTCCAATTCACAAAGCTGTTCGAAGCACACAGACATAGACAAAGACCACACA
GATCCTCAAATGTGTACTACTTATGTCGAAGATATTTACGAGTATTTACGAAATGCAGAGT
TGAAGAATAGGCCTTCGGCAAATTTCATGGAGACTGCACAGAATGATATCACCCCAAATA
TGCGAGCAATTCTCGTTGATTGGCTGGTAGAGGTTAGCGAGGAGTATAAGCTTGTCCCT
GACACACTATACTTGACAGTATCTTACATTGATCGTTATCTGTCCGCAAATCCAACTAGTA
GGCACAAACTGCAACTTCTGGGAGTCTCTTGCATGCTAATTGCCTCGAAATACGAGGAA
GTCTGCCCTCCTCATGTGGAAGAATTCTGTTATATCACAGATAATACATACACAAGAGAC
GAGATGTTGTCAATGGAGAGAAAAATTTTGATCTTCTTGAACTTTGAGATGACAAAGCCT
ACAACAAAGAGTTTTCTCAGGCGTTTTGTACGAGCTTCTCAAGCAGGGAATAAGGCCCC
GAGTTTGCATATGGAGTTTTTAGCCAATTATTTGGCAGAGTTGACACTTATGGAATGCAG
TTTTCTCCAATACCTACCATCCTTGATAGCAGCCTCAACTGTATTTCTGTCTCGATTGACA
CTTGATTTTCTAACAAATCCATGGAATCCAACTTTGGCGCATTATACAGGATACAAGGCAT
CACAGCTTAAGGACTGTGTAATGGCAATCTATAATGTACAAATGAATCGAAAAGGTAGTA
CATTGGTTGCAATAAGGGAAAAATATCAACAACACAAGTTCAAATGTGTAGCAAGCTTGC
CTCCACCTCCTTTTATTGCTGAAAGGTTCTTTGAAGACACCCCAAACTAAAGCTATCCATC

ATTGGTTGTGCACAAGCGTTTTGCTGAAGCTTGTTTGGAAGTACAATCTCATGTAATAGTT
GTAATTGAATCCGGAGTCTTTTACTTATTGTTGGTTAATTTCCATTTAGATGTGTCAATTTG
TTCAGACCGTCAGTGCCCATTCTTTTGGTGCTCATATTTTAACTCAGCGACTTACCAGCC
TAGTAAGCAATGGGGAGCTTGCATGTATTAGTTTTGAGCACTCAAAGGAAGTGCCCATTT
ATCAAACTGTATCATATTTTTTGGAAATTATTACGTGGAAACAGACGACTCCATGTGGTCA
TGGAGTAATGATTAACAAAAAAAAAA

SEQ ID NO: 137
GAAAGGAGGGAGTGGTACGAGCTTCGGAAGAGGAGCTTCCAACCCGACACCAAATCTT
CTCACGGGGCTCCTCTATATGTACAGCAAACACTGGCGACTCTTTCCCTCAGTTCCGTTG
GTCCTACTGACATTAGCTGGAACAGGCTGATCGGCCATGGTGATCCTTTGGGCGGGCTG
CAAGAACTGGTAAACTTCCGCGGCTGGATCTCTGAATATTGCTGTTCGATTTACACATAT
TCAAAGGTGTAATAAAAATAACGATGACTGGAACCCAGGCATCCAATGTTCGCATTACAC
GGGCGAGGGCTGCAAAGTCCACGTTGAACAATGCCCTTCCTCCTCTTCCGCCAGCCCAA
GGTAAACCGAGAGGAAAACGTGCAGCAACGGAGTCAAATATTTCCGGATTCAGTGTTGC
TGCTGAACCCCTTAAACGACGTGCAGTGTTATCAGATGTGTCCAATATCTGTAAGGAAGC
GGCCGCGGTTGATTGCTTGAAAAAACCAAAAGCAGTTAAGGTTGTATCTCAGAATGCCAA
TGCCAAGGGAAGAGGCCGAGGAATTCCTCGGAATAACAAGAAGATTACTCAAGAGGCAG
AGATTAAAAAAGAGACTTCACCAGCTATTTGTAATGTTGACGATGCATCTGCTGGTAATG
CAATCGGGGATGATAAGCAAAACAATAATGTGAACCCTCTAAAAGAAGTACAGGACAATC
CTAAAGAACTTAATCCAATTGCTGAGCAAATTTCAGTACACCCACATTGCAAACAGTCAG
TGGAAAAACCGAATGAGAAAGAGATTGTGGTAAGTGATAATAAAGCAGCAATTGCTTCTC
TGAAGCAACAATCAACATTGCAGTCTCTACGTATACCAAAACAACCAAAATATTCTTTGAA
GCAAGGTAATCCTGTTCCTCTGGCAAACCTTCATGAAGATGTGGGACGGTCTAGCTGTT
CGGATTTCATTGACATTGATTCGGAATACAAAGATCCTCAGATGTGTACAGCCTATGTTA
CAGACATCTATGCAAATATGCGAGTTGTAGAGCTAAAGAGACGGCCCTTGCCCAATTTTA
TGGAAACCACTCAGCGTGATATCAATGCAAACATGAGGAGTGTACTCATTGATTGGCTTG
TTGAGGTATCAGAGGAGTACAAGCTTGTCCCAGACACACTTTATTTAACCGTATCTTATAT
TGACCGGTTCCTCTCTGCAAATGTAGTCAATAGACAAAGGCTACAGCTTTTAGGGGTTTC
TTGCATGCTTGTTGCCTCGAAGTATGAAGAAATTTGTGCTCCTCCAGTAGAAGAATTTTG
CTATATTACCGATAATACCTACAAAAAGGAAGAAGTGTTGGAAATGGAAATTAGTGTTTTG
AACCGCTTGCAATATGATCTGACAACACCAACCACCAAGACCTTTCTCAGGCGTTTCATC
CGAGCAGCTCAAGCATCATGTAAGGTTTCAAGCTTGCATTTGGAATTCATGGGCAACTAT
CTTGCAGAGCTAACACTTGTAGAATATGATTTTCTCAAGTATCTACCATCTCTAATAGCTG
CAGCAGCTGTATTTGTGGCAAGAATGACCCTTGACCCTATGGTTCACCCTTGGAATTCTA
CTCTGCAACATTACACAGGTTATAAAGTATCAGACATGAGAGATTGTATTTGCGCCATAC
ACGACTTGCAATTGAACAGAAAAGGTTGCACATTGGCTGCAATAAGGGAGAAGTACAAC
CAACCTAAGTTCAAATGTGTGGCGAACTTGTTTCCACCTCCAATAATTTCTCCTCAGTTCC
TCATAGATAATGAGGTGTAGTTGATTCTGAATATGTGGCACTCATCTTCATTTGGCTACAT
CTCAAATGGATGCAGATAGCTGTGTTTTTCAGATTTGTGAATTGGGAATTCATCACTGAAA
AAGAATTCGTCCTGGTCCTTTAGGACATGTACTTATGTCCATGCAAGTGCTTCTTGCCTA
AGCTATGTTGATATTTCACCTTTCTTTAGATAAGGTGCTTGGTTGATGCTGTCATATAAAT
GGCCATGTCTACCATGCACATAAATTTAAACATTTTCTTGAAAAATCAAGATTATGACTAT
GGGTATCAGTGACCACTTTTCACTTGTATTTTATATGCAGAGTTAATCTGGAGGATCATTA
TCCATGAGAGTGGCCCATTCTGATTTTAAACAATGAACTTACCTGGTTGTCTATAACATTT
TTAGTTAATATTGAAATCATGGCTTAAAAAAAAAA

SEQ ID NO: 138
CTTGAACAGGAACAGACAGATTGTTTCAAGGCCACATGCAGCAGCAAACATAGAGATTG
TTTCAAAGCCACCTGCAGCAGCAAACAGACAGATTGTTTCAAAGCTACCTTCAGCAGCAA
ACATACAGATTGTTTCAAAGCCACCCGCAGCAGCAAAGGGTCCTGCAATGGCAGCACCT
AACCAGAATGCGTTATTAATCAACAACAACAATAGGAGGCCCTTAGTAGACATTGGCAAT
CTAGTGGGAGCTTTGAACGCCCAATGCAATATCAGCAAAAATGGTGCCAGGAAGAGGGC
CTTCGGAGACATTGGCAATCTCGTGGAAGATCTGGACGCTAAATGCACTATCAGCAAATA
TTGGGTTAGAAAGAGGCCTAGAACAAACTTTGGTGTCAATGCTAACAAGGGAGCCAGTT
CTAGCACTCAGGGACAGGGTATTGTTGTCAGGGGAGAACAAAAAGCATGGGACAGGATA
GTCTGGGGAAACAAACAAAGCTGTGCTATTAAAATGAATGCTCAACACGTCACTGCTACT
CAGAGGGGTACTGCCATTTCAATCAGTGATATCATAGATTCAAGTGTACAAGATGGGGG
AATAAAAGCACCTAGCCAGCTGAAAGCAAGGAAACAAACAGTACGAACAGTGACTGCAA
CACTAACCGCTCGAAGCGAGGATTCATTGAGGGACGTCCTTGAGGTTCCTCCTGGCATT

GATGATGGAGACCGAGACAACCCACTGGCAGTTGTTGAATATGTTGAAGATATATACCAC
TTCTACCGGAAGATTGAGGTCAGGAGCTGTGTACCTCCAGATTATATGACAAGGCAGCT
TGAAATCAAGGACTCCATGAGGGGAGTTATAATTGATTGGTTAATAGAGGTTCATCGGAC
ATTTCTACTGATGCCAGAGACATTGTATCTCACTGTAAACATCATCGATAGATACCTCTCT
ATTCAGAGTGTTACAAGGAATGAATTACAATTAATGGGTATCACTGCAATGTTTATTGCTT
CCAAATATGAAGAAATCTCTCCTCCAAAGATCAACGATTTAGTCTACATCACGAAAGACG
CATACACATCAAAACAGATAGTGAATATGGAGCATACAATATTAAATCGACTGAAATTTAA
GCTTACGGTTCCCACTCCATATGTATTCCTTGTCAGGTTTCTTAAGGCAGCTGGTCCAGA
CAAAGTGATGAAAAATCTGGCATTCTTCTTGGTTGACCTGTGTTTGCTTCATTACAAAATG
ATCAAATACAGTCCATCAATGCTTGCTGCTGCTGCTGTCTACACTGCTCAATGCACTTTG
AAGAAGCATCCATATTGGAATAAAACACTTATCCTTCACATTGGCTACTCGGAAGCACAC
CTCAGGGAATGCGCACACTTAATGGCTGATTTACACCTGAAGGCAGAAGGAAGTAACCT
CAAAAGTGTTTATAAGAAGTATTCATATCCAATTTTTGGCTCTGTTGCTTTTTTAAGCCCT
GCCAAGATACCTGCGGGAACAGTTGCAGCTCCAGCAATTGATAAATGTGCACACCAGAT
TTACTTGAGGAACCTTCGCTGAAGGTTGCCCATTCTTCCGATCGGTCGAAAACCTATCTC
TGGTGATGTCGCTAAGAGAATGTTTTCGGTTGTGTGTATGTAAAATTGTAGATAGCCCTT
GAACAAGACATGTAAATTTATGGCTGATTATTTTTCATACATTGAGATTCTCATTATATTCT
GAACAACTAGTTTTCATTAGCAAGATAAATGCTAGATGCTTTGAACTGGCATAGACTGCT
AACATCCCCGGCCTTCTTAATTTCTAGGCGATATATATCACCGTAACTTTGGATTTGTAAA
GAATATAGGGGATCATTAGCTTTATGTAATTATATTTTCAAGTATGATTCTTCATTCAACAA
TCTTTGTTGATGGATACTTTGGTTCATAACATCTCTATAAGTGAATATGGAAGCCAATAGG
CTTGTAGTTCAAATGTAATTTTGCCAGTCTCGAAAATTAAATAGGCC

SEQ ID NO: 139
CTAGAACAGTGAAGAAATTTGCGAGCACAAGATCCCTCGTTTCTGTGGTCGTTGTGATTT
CAGTTTATTTGGTTGCCAGAGCTGCCACCCCTGGGGCTCTGGGATACTGCCTGATTCAA
CCTTGTTATACAGGGTTTTATAACTTGTTTCTTGCGATTAGCCAGCAAATTTAATTTGCTTA
AATGTTCCCCAACAAGCAAACCCAAGGGCTTGTACAGAACAAGAAGATGGCTTCTAAGG
CAGCCCAACCAAAAGCAATGGTCCCTCCCCAGAGGGTTCCTCCAGCTGCAAACAATAGG
AGGGCTCTGGGAGATATTGGTAACATTGTGGCAGATGTCGGTGGTAAATGCAATGTCAC
CAAGGACGGAGTCAATGGGAAACCTCTTGCTCAGGTCTCTCGCCCTATTACAAGGAGTT
TTGGTGCTCAGTTACTTGCACAGGCAGCTGCAAACAAGGGGATCTCCGCTGCCAATAAC
CAAACTCAGGTACCGGTTGTGATTCCAAAGGCAGATGTGCGGGGAAACAAGCAAAGGA
GGACTAGCAAGAGCAAGGACATTCCTCCGACAACTGTAGTTACCAATGAATCAGATGATT
GTGTTATTATTGAACAAGCACAGAGAATAAAACCAACTTGTAATCACAATGTTGGAGCTG
TAGGGAACAAAGAGAAACCTCAATTACTAACAGCAAAGCCTAAATCACTAACGGCATCAC
TTACATCTAGAAGTGCAGTTGCGTTACGTGGGTTTAGATTTGATGATGAAATGACTGAGG
CAGAGGAGGACCCCTTGCCAAATATTGATGTGGGGGACCGCGATAACCAGTTAGCGGT
GGTTGAATATGTTGAGGATATATACAAGTTCTATCGTAGAACAGAGCAAATGAGCTGTGT
GCCAGATTATATGCCCAGACAACAAGAGATTAATCCCAAAATGAGAGCAGTACTAATAAA
TTGGTTAATTGAGGTACATTACAGATTTGGGTTGATGCCAGAGACTTTGTATCTTACCACA
AATCTTATTGATAGATATCTTGCTACTCAACTTGTATCGAGGAGCAACTATCAATTAGTGG
GTGCCACTGCAATGCTTCTAGCTTCAAAGTATGAGGAAATTTGGGCTCCAGAGATGAATG
ACTTTCTTGACATCTTAGAAAATAAATTTGAGAGGAAACATGTACTGGTTATGGAGAAGG
CAATGCTAAACAAACTGAAATTCCATCTCACAGTCCCAACTCCATATGTATTCCTTGTCCG
GTTCCTCAAAGCTGCTGCCTCTGATGAAGAGATGGAAAATTTGGTATTCTTCTTGATGGA
ACTGAGCTTGATGCAATATGTGATGATAAAGTTCCCCCCATCAATGCTCGCTGCTGCTGC
AGTCTACACCGCTCAAATCACTTTGAAAAAGACAACTGTCTGGAACGATGTACTTAAACG
CCATACAGGCTACTCTGAAATAGACCTCAAGGAGTGTACAAGACTGATGGTGGCATTCC
ATCAGAGTTCAGAAGAAAGCAAATTGAATGTAGTTTTTAAAAAGTACTCAATGCCAGAGTA
TGACTCCGTGGCACTTATCAAACCTGCCAAGCTCCCTGCATAGGCTTTCGATTGGAGTT
GCAGTTGCAGAGTGTGTAGCAACTGATGAGCATAGTTGTTATGTTTCTCAACTCAGTTGC
ACATCCAGCAACTGAGATTGATATGTCAATCAAAATTTTTAACATAGAAGATCATTCATTC
TCTTGATATTCCCCTTTAGGTGGATGGATTTTTTGGTCATTCTTTGTAACTGTACATAGTT
GTGGCAAAATCATGATTAAGAGAAGATATAGCCCACTTCATTTTTGGCATAGGCCTGCCC
ACATGGGAAGTGGTGGCTTAAAGCCAAGGGATTATTACTTGTCAATTGTAATAAATGCTT
GAGGGCATTTATGTGTTTTTTAATTTAAAAATTTGAATTATAGAACACTGAAATTGATTTCT
CAAAAAAAAAA

SEQ ID NO: 140

GGGCTTTTGGGATTCGGAGATGTGAAGAAAATGAGGTGGGGAGGGAGTAGATAACCAG
AAGAAGATGAACAAGATCTGGATTTGATACCAGCCAGCCATGGCGCCAAGCTTCGATTG
TGTGGCCAATGCTTACATCGAATCCTGTGAAGACCAGGAGAAGCTCAGACAAAATGCCC
AAATCTTGGCCCAGTCTGGTGAGAACGATGTTGATGAACCAGTCTCCATGTTGGTACAG
AGAGAGACCCATTACATGCTTCCAGAGGACTATTTGCAGAGGCTCCGGAATCGAACTCT
CGACGTCAATGTTCGTCGAGAGGCCGTGGGCTGGATTCTGAAGGTGCATTCTTTCTACA
ATTTCGGAGCTCCAACGGCCTACCTGGCTGTCAATTACCTGGATCGATTTCTCTCGAGG
CACAGAATGCCGCAAGGTGTGAAGGCTTGGATGATTCAGCTCATGGCCGTGGCCTGCC
TGTCTCTCGCGGCTAAGATGGAGGAGACCCAGGTTCCACTCCCTTCGGACTTGCAGAG
GGAGGACGCCAGATTCATCTTTGATGCTCGCACTATTCAAAGAATGGAGCTTCTGATCTT
GAGTACATTGCAATGGGGAATGCGCTCCATCACGCCATTTTCATTCATCGATTACTTTGC
TTACAGGGCTGTTCAAGGTCATGGCCATGGCCATGATGCTACTCCAAAAGCTGTGATGT
CCAGAGCCATTGAACTCATTCTCAGTACCACAGAGGAAATAGATTTCATGGAGTATAGGC
CATCGGCAATAGCTGCAGCAGCCCTGTTATGTGCAGCTGAAGAGGTGGTGCCTCTGCAA
GCAGTCCATTACAAAAGGGCTTTGTCCTCCTCCATTACTGATGTTGATAAAGACAAAATG
TTTGGATGTTACAATCTGATTCAAGAGACCATCATAGAAGGAGGCTGCTACTGGACTCCA
ATGTCCCTACAATCCACTGAGAAAACTCCAGTGGGAGTGCTGGATGCAGCAGCCTGTTT
AAGCAATACACCCACCTCAAGCTACAGCGTGAAGCCATATGCCTCTGTTACAGCAGCCA
AGAGGAGAAAATTGAATGAAATCTGCAGTGCATTGTTGGTATCTCAAGCCCATCCCTGTT
GACACAGTACGAGGTCCCCAGCTTCGGTTGTCGAGATTGGAGAGTGGGATTTGGTAGG
GTTTTAGTGGAGCAGCAGATGCAAAACCCTATATAGATATATAATCGACTGTAAATGGGG
TATAGAGTACTTAGCCTGCAAATTGCTGTCCCTTCCTCGCCTATTTGACCAACTGGTTTG
GCCCAGACTGGATTTTGTCTTGTTTTAGAGAAGGGTTTAATTCAGTTTGAATAATGGAGG
ATGATGATGACAGATCACCCATAATATCAGTGAAGAAACTCATACACTGGACAGGCCAAC
CTTCCAAATATGTGTTTAGAAAACCTTTGTCTTCTTTTTATGTCATACACTCATCCAATCAA
ACAGGCCCACCTTTCCATAAAAAAAAAAAAAAAAAA

SEQ ID NO: 141
GTTTATGCAATTCCAATATTTATCACATTCACCTCATATCATATATTTGCGTATTGTGACTT
CTATACAACAAAAAACAGGTACCAGAGCAGGCACGAATAACCTGCAGGTCTATCTATCTT
GTTTTTTTTAAGACGGGGTTGCAGAATGGCAGCAAATTTTTGGACGTCTTCTCACTGCAA
GGAACTTCTGGACGCGGAGAAGGTAGGAATCGTCCATCCGCTCGATAAGGACCAGGGG
CTTACTCAAGAAGATGTCAAAATTATAAAGATCAATATGTCAAATTGTATACGAACATTGG
CTCAATATGTGAAGCTGCGACAAAGGGTAGTGGCAACAGCTATTACATATTGCAGACGT
GTTTATACCAGAAAGAGCTTCACAGAATACGATCCGCAGTTGGTAGCTCCTACTTGCCTG
TACCTAGCATCCAAAGCAGAAGAAAGCACTGTTCAAGCTAAACTAGTAATCTTCTATATG
AAAAAATACAGTAAGCATCGTTATGAAATCAAGGATATGCTCGAGATGGAAATGAAGCTT
CTCGAAGCTCTTGATTACTATTTAGTCATCTACCATCCATACCGTCCTCTAATTCAGTTCT
TGCAAGATGCAGGTTTGAACGATTTGAAAGTTACGGCTTGGGCTCTAGTAAATGATACAT
ACAGGACGGATCTGATTTTGACTTATCCCCCTTATATGATAGCCCTGGCATGCATCTATT
TTGCATGCATTATGGAAGAAAAGGATGCACAGGCATGGTTTGAGGAACTTCGGGTTGAC
ATGAATGAGATCAAGAACATATCAATGGAAATAGTGGACTACTATGATAACTACCGAGTC
ATCCCAGACGAAAAGATGAATTCAGCTTTAAATAAACTCCCACACAGATTTTAAAGTATCA
AGCATACCTTTGTATTATGGTGAAACCCATTTGAAAGCATTCTATTTCTTTTTCTCATTTCC
ATTGTCATGGGAAATATAGAAGACTCGGGCTCTTGTAAAAGTTGGGAAGGGGTGCTATC
CATATCTAGAATCTACCATGCTCAATGAGGTATCTTCATTAGTATACTAATTGAATTCCAG
TTCATCCCTGTGTTATTCAAGCTTTATGGTAATTTAAAAATTACTGCGTAGTGGAAGCTTG
AAAGCCTAGTTATTCAAGTGACAATATTTAAAAAGGACAATCTTATGTATTTTGTATTTCTG
TATTTTACATATAAGAACACATAATCTATAATCATCAGATAAAATATTATTGCGTTGAACAT
TAAAAAAAAAAA

SEQ ID NO: 142
CTGGAGTTGTCTGCATTCTCGGAGCTTCAATCAATCCATGCTCCATCTCCTGTGCAGTTT
GAAGTCTGTGAAACCATCTGTTGAATGTGGGATTCAATAGGGCATTTGGGCACTTGTTCT
GCTAATTTGGGGGGGCCTTGGCGGCGGTGTTATGGAGCGTAAACTGATGTGCAAAGGG
CCTTAAACTATGCTGCCATTCCTGTATTGTTAGCACACTGAAAGGGGTTTTCATGGGTTTT
GATTAGATGAGTAGCAGTTGTAAAACATTGGTCTGTGCTGCTAAATTATGGCCCCTGCAC
TATCCTCCAGCTATGAATGCCTGTCCCATTTGCTCTGTGCGGAAGATGCCAGCAATGTAG
TGGGCTGCTGGGATGAGGATGAGAGTAAAATCTTTTGTGAGGAAGAAGAGGGTTTTGGG
ATTCAGCATTTCCCTGATTTCCCAGTTCCAGATGATGATGAAATAAGGGTCTTAGTGAGG

AAGGAAAGCCAGTATATGCCTGGGAAGTCCTATGTACAGTCCTATCAGAACCTCGGACT
GGATTTCACCGCAAGGCAAAACGCCATCGGATGGATTCTCAAGGTCCATGGCTCTTACA
ACTTTGGTCCACTGACTGCCTATCTGTCTATAAACTATCTAGATCGGTTTCTGTCCAGGA
ATCCACTACCGAAAGCCAAGGTCTGGATGCTGCAGCTCCTATCTGTGGCTTGCTTGTCT
CTGGCTGCTAAAATGGAGGAAACCCAAGTTCCTTTACTCCTGGACTTGCAGGCTGAAGA
GCCCGACTTCCTCTTCGAGCCGCGTACTATCCAAAGAATGGAGCTTCTGGTTCTGAGCA
CTCTGGAATGGCGGATGCTCTCTGTTACACCGTTTTCGTTCGTCGATTACTTCTTGCAGG
GTGGAGGGGGCAGGAAGCCACCGCCGAGAGCTATGGTGGCGCGAGCCAACGAGCTCA
TATTCAACACACACACAGTGTTGGATTTCCTTGAGCACCGACCATCTGCCATAGCTGCTG
CAGCTGTTATCTGTGCAGCCGAGGAGGTTTTGCCCCTGGAAGCAGCCCAGTACAAGGA
GACCATCCTCTCCTGCTCTCTTGTAGACAAGGAATGGGTGTTCGGCTCTTATAATCTGAT
ACAAGAAGTCCTGATTGAGAAATTCTCGACGCCTAAGAAGGCAAAATCAGCATCCTCCTC
AATCCCGCAGAGCCCTGTTGGAGTGCTGGACGCCTTCTGTTTGAGCAACAACAGCAACA
ATACTTCACTGGAAGCTAGCTTAAGCGTTAACCTCTACGCCTCTGTGGCTGCCAAACGCA
GGAAGCTCAACGATTACTGCAATACATGGCGAATGTTTCAGCACAGCACCTGTTAATCAA
ACATGAGATTGAAGGCCAAGGACTCAAAAACCAGTCCAGCATTTGTGTGCTGTTTTAGAG
GGTTTTAATACTTCAAGTCCTCTGTAGAATTTGATACCTCTGGACGCTACAGTGATCAGG
GATTTGATCTAATGCTAGTTTATTGATTTCTATGATCCAAGACCTCGTCATAGATCAAGTG
CCTAGTTTATTGATTTTTATGACACAGTCGAAGCTTTTCTAAGCTTTTCCGAGACTCCTAA
CATATGTAACAACAGAATTCGTGAGCCAAGAGAATTTGTTGAATTCGTTCTCTATACTTAT
TTTGTAACTCTGGACGGACCAGTTAAATTATAAATCCTGGAAGTTTTGATCAATTATATTC
AAAATCAAAAAAATCCTGCGTTGAAGGTTCTGTGAAAAAAAAAA

SEQ ID NO: 143
GCTAACCTGTAGATATCCTTTGCTTCCCATTCCCTCTCATGGATTTGGATTTCCCCTGATT
TGATTCAGGACAAAAGCCCTTGTGGCTCCCCTACAAATATGTATTTGCATATCTGAAATTT
CCCTGGCCATAATAGCTGAGAGGCTTTTTCTTTTTTTTCTTTCATTCAATTTCATGGCGGG
GATTCCTGAGTTGGGTTGTTGATGAATAACAGGATCAGCAGCTCAGGGGCCCGGGGCTT
GTAGGGGGATGGCACCCAACTGCATAGACTGTGCCCCTAGTGATCTGTTTTGCGCGGA
GGATGCTTTTGGAGTTGTGGAATGGGGCGATGCAGAGACTGGAAGTTTGTATGGAGATG
AGGATCAGCTGCATTATAATTTGGACATTTGTGACCAGCATGATGAGCATTTGTGGGATG
ACGGTGAACTTGTAGCTTTTGCAGAAAAAGAGACCCTCTATGTTCCTAACCCAGTTGAGA
AAAACAGTGCTGAAGCTAAAGCTAGGCAGGATGCTGTGGATTGGATTTTGAAGGTTCAT
GCACATTATGGCTTTGGTCCTGTGACTGCAGTGCTCTCAATAAACTATCTTGATCGGTTT
TTGTCTGCAAATCAATTACAGCAAGATAAGCCATGGATGACTCAACTGGCAGCTGTGGCT
TGCCTCTCCCTCGCTGCCAAGATGGATGAGACAGAGGTTCCCCTTCTCCTGGACTTTCA
GGTTGAGGAGGCTAAGTATATATTCGAATCTCGCACCATTCAGAGAATGGAATTACTGGT
GCTTAGTACCCTTGAATGGCGAATGAGTCCTGTGACACCTCTTTCCTACATTGATCATGC
CAGTCGTATGATTGGGTTGGAGAATCACCATTGTTGGATTTTCACAATGCGCTGCAAGGA
GATACTGTTGAATACACTCAGAGATGCAAAGTTTTTGGGCCTTCTGCCCTCTGTTGTAGC
TGCTGCAATAATGCTGCATGTGATCAAGGAAACAGAGCTTGTTAATCCATGTGAGTACGA
GAATCGCCTGCTCAGTGCCATGAAAGTTAACAAGGACATGTGTGAAAGATGCATAGGAC
TACTCATAGCCCCTGAATCATCATCCTTGGGCAGTTTCTCTTTGGGTTTGAAAAGAAAGA
GCAGCACCATCAATATTCCTGTTCCTGGCAGCCCAGATGGAGTGCTGGACGCTACCTTT
AGCTGCAGCAGCAGCAGCTGTGGTAGCGGACAGAGCACCCCAGGGTCATATGATTCCA
ATAACTCCAGCATTCTTTGCATCTCACCAGCGGTGATAAAGAAGAGAAAGCTTAATTATG
AGTTTTGTAGCGATCTTCATTGTTTGGAGGATTAGTAGATTACCATAAATACAAGTGCTCC
TCTATGAAGAACGCTTTGGAAGCAAGCTTTGCTATTAAAACTTCATTTCAACTATACCTAT
CCAGAATTAAGTGCAAGTTCAGGCCTATAATAGGCAACTATTATAGCTCTATGCCATTAT
GTCTCGCATGGGAGAAACATATTGTTATTAAATACCATTCAATATGCTTATGATTCATGAA
TGCTTAAGAGATTCTGCTGCTGTGTGAAAAAAAAAA

SEQ ID NO: 144
TTCTGCCCTTGCGAGCTGCTTAGTTTATTCCTCGTTAACTTCGTCCTTTCGACCTTTTACG
CGCCGTCTCGTTCGCCCAATCTCAGGTTCTCGCTCACTAATTTTCGTTTCAAAGGGTTTT
ATCAGAGCTTTCCACCGCTGTTTGCAGGGCATTTCAAGTTCTCAGGGCGGAATTCGACC
TGTCTGTTATCCGCCAGATCATTCAAGCCTTTAGTTCGCTCGACCAGGATGCCTCAAATT
CAGTACTCAGAGAAATACACTGATGATACCTATGAATACAGACATGTGGTTCTCCCTCCG
GAAACTGCCAAATTGCTTCCCAAGAACCGACTTCTCAATGAGAATGAATGGCGAGCCATT
GGAGTTCAGCAGTCTCGTGGATGGGTGCACTATGCCATTCATCGTCCTGAGCCACACAT

CATGTTATTCAGAAGACCTTTGAATTACCAGCAAAACCAGCAGCAACAGGCTGGGGCTC
AATCTCAACCTATGGGTTTGAAAGCCCAGTGAGTTTTATTGTGGGTTGTTGAAAGCAGTT
TCAATGTTCTGTTTGAAACTAATCAGAATAGGTTCTCCAGGGTGTTTGACTTTTTCCTTTG
CAGGTAGTTCTGCAGTTTTAGTATATTAGGGTGATGACTTTCTTTATCAAGGCTAGTCTGT
TGTTTAGTTAATACGGTTGACAATGAATGTCTAGTACATATTTTTGTGAACTATTATGAACT
ATTGCTTCTAAACTGTAGAAGCCTGTTATCTTTAGACTCGTGGTTATGTGAACTACTTTTA
CAGTAAA

SEQ ID NO: 145
GGCGTGTCAATGTTTCAATACCAAATAGCATTTAGAATTATACTGCAGCGTTTTATCTGAT
ACACGAAGGATTTTTAAATTTGCCTGTAAAATGGATCAAATAGAGTATTCTGAGAAATACT
ACGACGATACCTATGAGTACAGGCATGTTGAGCTTCCGCCTGATGTTGCCCGGCTACTT
CCCAAGAATCGCCTTCTAACCGAGAATGAATGGCGAGGAATCGGGGTTCAGCAGTCTCG
TGGGTGGGTGCACTATGCTATTCACTGCTCTGAACCACACATTATGTTATTCAGAAGGCC
TTTGAATTACGAGCAAAACCACCAGCACCCTGAGCCACACATTATGTTATTCAGAAGGCC
GTTGAACTGCCAGCCAAACCACCAGCCACAAGCACATCATCCAACATAGGCTGTGGGGA
TTCGAGCCTGATGGTTATGCACTGTGGCCAGCAAGATGTTGAAGTTTTAGCTGAGTAATT
TGAAAGTTCCTTTTTTCCTTTTCACCATAGCTATTATTTGTGTACGTATTTCCCAGGCTATG
TACAGATTTAAATTGAAATCTAGCCATGACTATGGGCCTTGAGATATGACTATTGTATATG
TATCCCCTATCTTGTGAATTGTGAAATTTATATGTTTTTTTCTTCTGCGAATGGTTCAAAAT
TAGACGAACAGAAATTTTGTTCCCAGTACAGTAAAGCCCAAATGCAAGCCCAAATGCAAG
AGGGCGGACGCAGGCGTTGTGGGTTACATCATGTCATGTAATTTGTCTGATCAAGTTCTA
AGGCTGGCTTGGTATCAATGAACTTTTAACTTCTAATTTTGAAGACATACATTTATTCTTTT
AATTTGATGTCTTTCTGTGAGTAAAAAA

SEQ ID NO: 146
GTTTTATCTGATACACGAAGGATTTTTAAATTTGTAAGTGCTCAGTTTTTGCAGGCCTGTA
AAATGGATCAAATAGAGTATTCTGAGAAATACTACGACGATACCTATGAGTACAGGCATG
TTGAGCTTCCGCCTGATGTTGCCCGGCTACTTCCCAAGAATCGCCTTCTAACCGAGAAT
GAATGGCGAGGAATCGGGGTTCAGCAGTCTCGTGGGTGGGTGCACTATGCTATTCACT
GCTCTGAACCACACATTATGTTATTCAGAAGGCCTTTGAATTACGAGCAAAACCACCAGC
ACCCTGAGCCACACATTATGTTATTCAGAAGGCCGTTGAACTGCCAGCCAAACCACCAG
CCACAAGCACATCATCCAACATAGGCTGTGGGGATTCGAGCCTGATGGTTATGCACTGT
AAGTGATCTGATTTGATTAACTATTTTATCAATTAATTTTTCATTTGAAATTCAACCAATCAA
TTTCTCTCGACCAATTCCATCTCTGTGCCTGATTCGTAGGTGGCCAGCAAGATGTTGAAG
TTTTAGCTGAGTAATTTGAAAGTTCCTTTTTTCCTTTTCACCATAGCTATTATTTGTGTACG
TATTTCCCAGGCTATGTACAGATTTAAATTGAAATCTAGCCATGACTATGGGCCTTGAGAT
ATGACTATTGTATATGTATCCCCTATCTTGTGAATTGTGAAATTTATATGTTTTTTTCTTCT
GCGAATGGTTCAAAATTAGACGAACAGAAATTTTGTTCCCAGTACAGTAAAGCCCAAATG
CAAGCCCAAATGCAAGAGGGCGGACGCAGGCGTTGTGGGTTACATCATGTCATGTAATT
TGTCTGATCAAGTTCTAAGGCTGGCTTGGTATCAATGAACTTTTAACTTCTAATTTTGAAG
ACATACATTTATTCTTTTAAAAAAAAAAA

SEQ ID NO: 147
CTTTTGCAGTTTACCGTAGTCTTTCAGCGTTTCTACTTCAAAGAGACTATTTTTGCAGGCC
GGAAAGATGCCCCAAATACAGTATTCTGAGAAATACTACGACGATACCTATGAGTACAGG
CATGTTGTGCTTCCGCCTGACGTTGCCCGACTTCTTCCCAAGAATCGCCTGTTGAACGA
GAATGAGTGGCGAGGAATTGGGGTGCAACAGTCTCGTGGGTGGGTGCACTATGCTATC
CACCGTCCTGAGCCACACATTATGCTATTCAGAAGGCATTTGAATTACCAGCAAAACCAA
CAGCAGCAGGCACAACAGCAACCAGCACAGGCTATGGGGCTTCAAGCCTGATTGGTTTT
ATGTTGTGGCAAGCGTTGACAAAATCTCCTTGTTAAGGATTCTTGGATTCTTCCCGAGCA
AGCAGGCTTTCGTACGTAGTGTTAGATGGTTGATTTTTCCATCATAAAAGTAATTAGAATT
AATGTGAGAGTCCCATTTGCATTTAAAAAATGGTCATTATCCGAGATAGTGCGCTTTGTC
ATGGGAAAATGACTATTGAATGTGAGTTTCCTTTGCCCAGGGAATATGTAAAGGGATTTT
GAATTCTCATGTGAAAACGGCTATTACATGTGAAAATGGCTGTTATATGTTCAAAGTGATT
TCGAATTGATTTTTGTTACTTCAGGAATGAAATTATTTCTATCTTTTTTTTGCTTTTTTAAAAA
AAAAA

SEQ ID NO: 148

GGCTACGCGAACTGAACAGTGACAAAATATCACTCGGACGCCATGCTCTGCAAGTCATT
CAAAACTGTGTAAAACCGTAAACCATCCCTCTGTCCTCAAACTGTAGGTCGCCTCGATTG
CTGTCATGGCGCTGGTAGAAACGGAGCCTGTAACCTTAATACACCCGGAGGAGCCCAAA
AAATTCAAGAAAAAGCCCACTCCGGGCCGCGGCGGTGTGATTTCCCATGGGCTAACAGA
GGAAGAGGCAAGGGTTAAAGCCATAGCCGAGATAGTCGGAGCGATGGTGGAAGGATGC
CGGAAGGGCGAGGACGTCGACCTGAACGCCCTCAAGGCAGCAGCATGCCGGCGCTAC
GGGCTATCGAGGGCGCCGAAGCTGGTGGAGATGATCGCAGCTCTGCCCGATGGCGAG
AGGGCAGCGGTGCTTCCCAAGCTCAAGGCCAAGCCGGTGCGTACGGCCTCGGGCATC
GCCGTTGTGGCTGTGATGTCGAAGCCACATCGTTGTCCGCACATCGCCACCACTGGAAA
TATATGCGTGTACTGCCCCGGGGGCCCCGACTCAGATTTCGAATACAGCACGCAGTCTT
ATACTGGCTACGAGCCCACGAGCATGCGTGCCATTCGAGCCAGATATAATCCATATGTT
CAGACAAGGAGTCGGATAGATCAATTGAAGCGACTGGGTCATACTGTGGACAAGGTGGA
GTTCATCTTGATGGGAGGGACTTTCATGTCTTTGCCAGCTGACTATCGCGATTATTTCAT
TCGAAATCTGCATGATGCTTTATCTGGTCACACCTCTTCAAATGTGGAGGAAGCTGTGTG
CTATTCAGAGCACAGTGCCACTAAATGTATTGGCTTGACTATTGAAACGAGACCTGATTA
CTGTCTTGGACCTCATCTAAGACAAATGTTATCATATGGTTGCACACGTCTAGAAATTGG
AGTCCAGAGTACTTATGAGGATGTTGCACGTGACACAAATAGGGGTCACACTGTGGCTG
CAGTTGCTGATTGCTTTTGCCTGGCAAAGGATGCTGGTTTCAAGGTGGTAGCACATATGA
TGCCTGACTTGCCAAATGTTGGTGTGGAGCGAGATATGGAAAGTTTTCGGGAGTTCTTT
GAGAACCCCGCATTTAGGGCTGATGGTCTTAAGATCTATCCCACCCTAGTAATTCGTGGA
ACTGGACTTTATGAACTCTGGAAAACTGGCAGGTACCGGAATTATCCACCAGAGCAGCT
TGTGGACATAATTGCAAGGGTTTTAGCTTTGGTACCACCCTGGACCCGAGTCTATCGTGT
TCAGAGGGACATTCCAATGCCGCTCGTAACATCTGGTGTTGAAAAGGGCAATCTTCGGG
AGCTGGCTTTGGCACGTATGGATGACTTGGGACTTAAATGTCGTGATGTTAGAACCCGT
GAGGCTGGTATTCAGGATATTCATCATAAAATAAGACCGGAAGTAGTGGAGCTTGTACGA
CGTGACTATTGTGCTAATGAAGGTTGGGAGACATTCCTTTCATATGAAGATACACGGCAG
GACATTTTAGTTGGTTTATTGCGTTTGCGCAAATGTGGCCACAACACAACATGTCCAGAG
CTCAAGGGTAGGTGTTCAATTGTTCGTGAGCTTCATGTTTATGGAACTGCCGTGCCTGTT
CATGGCCGTGATGCAGACAAATTGCAACACCAGGGTTATGGGACGCTTCTTATGGAACA
AGCAGAGAGGATTGCTTGGAAGGAACATAGGTCAATCAAAATAGCTGTCATATCAGGTG
TAGGAACACGCCATTATTATCGAAAGCTTGGGTATGAGCTTGAAGGGCCGTACATGATG
AAGTATTTGAACTAGTTGTACATAGTAAGAGTGAAATCTTGTTTCATATCATTGGATATCG
TCCCTTCTTGGGTGCAATATGCATTATTTTCTGGTGCATCCTTAACACAGCTTGGTTACAT
GGTGAATTACAGTATTTGAAGGAGTCAAGAATGTTCTACCTGTTCCTGATTACAAGATTTA
GATGCTTTATGGAAAGCAGTATCATTTAGATGTAAAAGGATTGAGGCCAATTATACCTTGT
AACATAACATGTTATTTTTGACATTTATTCCTAATAGATTAATTTTCCTGCATTTAATATATA
TATATATATAATTTTATGAGATTTTTCACTTTCTTCAAAAAAAAAA

SEQ ID NO: 149
CTCTCCTGAAAGCAAACATCTTGAAAGCAAACATCTCGGAAGAAGCCAAGCTCAATTGGA
TGTACACAGATGCTCGGATTCAGGGATTTATATACATCTATATGCGAGCATCTGCAGAGA
GCCTCGGGCAGACTCCCAATTATTGCAGCGGCCACTAGTTTGATCAGCACCCCTGAAAT
AGCAGCAGTAGAAAAAGAAAACAAGGCGCCAAATTCTGTTGACAAGATGGGCATGGGCT
CTGCGGACGAAAGTGGGAGGTTCAGTACAAGCAATGGCCAATTTATGAATATGAACAAT
GGGGTTGTCAAGGAGGAGTGGAAAGGCGGGGTTCCAGTTGTCCCGTCAGCGCCCACGA
CAGTCCCTGTAATCACAAACGTTAAACTCGAAACCCCATCGTCGCCCGATCATGACATG
GCGAGGAAGAGGAAATTAGGGTTCCTTCCGCTGGAGGTCGGTACTCGGGTGCTCTGCA
AATGGCGAGACGGCAAATTCCACCCTGTCAAAATCATCGAGCGTCGCAAGCTGCCCAAC
GGAGCCACCAACGACTACGAGTATTATGTCCATTACACTGAATTCAATAGACGACTTGAT
GAATGGGTGAAGCTTGAACAACTTGAACTGGACTCTGTGGAGACTGATGCGGATGAGAA
GGTTGATGACAAGGCAGGAAGCTTGAAAATGACCCGCCATCAAAAAAGAAAGATAGATG
AAACTCATGTGGAGGGTAATGAAGAGCTTGATGCCGCGAGTCTTCGTGAGCACGAAGAA
TTTACAAAAGTGAAGAACATTACAAAAATAGAGCTTGGAAGATATGAGATTGAGACATGG
TACTTTTCACCATTTCCATCTGAATACAATAATTGCGAGAAGCTGTATTTCTGTGAATTTT
GCCTCAATTTTATGAAGCGGAAGGAGCAACTTCAACGACATATGAGGAAGTGTGATCTTA
AACATCCACCTGGAGATGAGATTTATCGAAGTGGAACACTTTCTATGTTTGAGGTCGATG
GAAAGAAGAACAAAGTGTATGCACAGAATCTCTGCTATTTGGCCAAGTTGTTCTTAGACC
ACAAGACACTTTACTACGATGTGGATTTGTTCCTATTTTATATTCTTTGTGAATGTGATGA
GCGGGGATGTCACATGGTTGGATATTTTTCAAAGGAGAAGCATTCCGAGGAGTCATATA
ATTTGGCATGTATACTTACACTTCCTCCATATCAACGGAAAGGATATGGAAAGTTCCTAAT

TTCCTTTTCATATGAGTTATCCAAGAAGGAAGGAAAAGTTGGGACCCCAGAGCGCCCCTT
GTCTGATCTTGGATTATTAAGCTACAGAGGTTACTGGACAAGGGTCCTGCTGGACATTTT
GAAAAAACACAAGAGCAATATTTCCATCAAGGAACTTAGTGATATGACAGCTATCAAGGC
AGACGATGTGCTCAGCACTTTACAAGGCTTAGATCTTATTCAGTATCGGAAAGGACAGCA
TGCTATCTGTGCAGATCCTAAGGTTTTGGATCGACACTTAAAAGCTGTAGGGCGAGGTG
GTTTGGAAGTTGATGTTTGTAAACTTATTTGGACCCCCTACAAGGAACAATGATGAAATAT
CTGTTTTTTTTCTTGTTTTTTTTTTCCTGTTCAATATGCCCCTTTCTAAATGAAGGGCACTG
TAATATTATCGATTATTCCTTGTTTTGTACCCCTAGAATCTGGGTAGGTGATCAGTAAACG
TTGTACATAGATGGTGGCGACGTTTTTTATTTGAAATTTTAGATTTTTTGAAAGAAGCAAC
TGGAATTTATCGAGAATAGGAATCATCCAGATTTAATGCCACTTAGGTGATCGGTGACCC
ACTTGTACATATAGATGTTGGCGATGTTTTTTTCCTGGAAGATGCAGCTGGAATTATAAAG
GACGGGATTCATCCACAAAAAAAAAA

SEQ ID NO: 150
ATTTATGTGTGCACGCTAAGTTCCAACGATACAGTCCTTTGAAGCACGTGGAAAGAAACT
CGGCCTCTGCAGTGGACATAAACCCAAACCCTAGTGAGTTTTAGGGTTTCATGTAGAGA
AGGTTTGAAATCTGTTGGCGATGGGATCCTTGGATGAATCCACTTGCAGCGAAGAGATC
AGAGATGAAGGCAAGGATTCAATTAGGACTAAATTTAAAGTCGAGTCCACTGTAAATAAT
GCACAAAACGGAGGCAATGACAATTCAAAGAAGAAAAGAGCGGCCGGCCTTCCCCTTGA
AGTTGGCATTCGTCTCCTCTGCAAGTGGAGGGATTCGAAACTGCACCCTGTTAAAATCAT
TGAGCGTCGTAAGCTTCCTAATGGGTTTCCTCAAGATTACGAGTATTACGTTCACTACAC
TGAATTCAATAGAAGGCTTGATGAGTGGGTGAAACTTGAGCAATTTGAACTTGATTCTGT
GGAGACTGATGCTGATGAGAAAATTGAGGACAAGGGAGGAAGCTTGAAAATGACTCGCC
ACCAGAAACGCAAAATTGATGAAATCCACGTTGAAGAGGGTCAGGGTCATGAGGATTTT
GATCCTGCTAGCCTTCGAGAGCATGAGGAGTTTACGAAAGTTAAGAACATAGCAAAGGT
AGAGCTTGGGAGGTATGAGATTGAGACGTGGTACTTTTCACCTTTCCCTCCTGAATACAG
CCATTGTGAGAAGTTATTCTTTTGCGAATTTTGTCTCAATTTCATGAAGAGGAAAGAACAG
CTTCAAAGACATATGAGGAAGTGTGATCTGAAGCATCCACCTGGAGATGAAATATATCGC
AATGGAACCCTCTCCATGTTTGAGGTTGATGGAAAGAAGAACAAGATATATGGGCAGAA
CCTCTGCTATCTGGCAAAGTTATTCCTCGATCACAAGACACTTTATTATGATGTTGATTTG
TTCCTTTTTTACGTTCTTTGTGAATGCGATGATCGTGGATGCCATGTTGTGGGATACTTTT
CAAAGGAGAAGCATTCTGATGAAGCATATAATTTGGCCTGCATTCTTACTCTTCCCCCAT
ATCAGAGAAAAGGATATGGAAAGTTCCTCATAGCTTTTTCATATGAATTGTCCAAAAAAGA
AGGGAAGGTTGGGACCCCAGAGCGCCCTCTGTCTGATCTTGGGTTATTAAGTTACAGAG
GATACTGGACAAGGATCCTTCTTGACATTCTTAAGAAACAAAGGGGTAACATTTCAATCA
AGGAACTTAGTGACATGACAGCTATCAAGGTAGAAGATGTTATCAGTACACTGCAAGTGT
TGGATCTTATTCAGTATCGGAAAGGACAGCATGTTATATGTGCCGATCCGAAGGTTTTAG
ATCGTCACCTGAAAGCCGCTGGAATAGCTGGGCTGGAAGTTGATGTATCTAAGCTTATCT
GGACTCCCTACAAAGAGCAATGTGGATGATTAATTTTTTAGACATGTTTCTTCATGGTTAT
GCTGAGCGAAATGAGAAAATCTAGCAATTCCATTCTTTTTTGCATCCTTTGGAAAGGTAG
CTAGAGAAAATTGCTCCTGGTGAAGAGGCAGTTGTATATTTGTTTCCATGGTGGTTTTGA
TATTCTTTTTCAAATCTGAGAACATCTCTTACTCATTATTGCGAAAGTTTAAGGTTTCATCA
TGCATCCCTGCGTGCAACAAGTTGTGATAATTTAAGGCTGAATGAACAGCATTCCAGCAC
ACAAGTGTCACTTAAAGAAATTCATCAATTCTTTGAAATTATTGTTCCCTTTTGATGCGGC
CCCTTTCTGGAGGGGCAATCAATGTCTGCGTTGCCACTCTTTTGATGAGGATTGTAAAGG
AACGGTACCATTGCCAATAAAGTAATGAAAATTTTCTAACGAAAAAAAAAA

SEQ ID NO: 151
ACGACTCCCCGTTCTTCGTCAGATTTACTTCATGGTTTTTCTCAAATCCATCCGAAGCAGT
AGCTGCTTATCAGTCTGCAACAAACCCTAAGAGAGAGAAAATCGTTATTGGTTCTCATGG
CCTTAGTGGACCTTGCAATTTACTGCAAACGTATTTCCGGGTTTGTTCTGCATCGATTATT
GAGGGATACCATCTGCGGGTTTAATTTATTTTAGTCAGGGATTTAGCTGGTTTGTTTTAAG
CTGGGGTTTCTAGTTTAGTAGGGTTATTAATGGATCCGGTGTGAGATATTTGTTTACTTTT
AAGGGCTGTTACGATGGTTTAGGAGCCGAGCTACTGCTTGAACTATGGCTTAGGTGTTG
TGGAAAATGATATACATCATATTTCGTCAGTTTTTTAAGAAAGAAGTGAAGAACCAAGTAT
TTTGAAGTTGTCATATTATATAATACACTCTATAACAAGTTAAAATCCATTGGATTTGGCGT
GAGCTCCTCTGGCAGTGGAAATTGATGTCATCTTTAATGACAGTTGTTTTAGTAATGAAG
CTGTTAGGGTGTTTATAGGTTTTGGAGTCAAGGAACTGGGGTACTCAGTTGGTATCGTG
CTTTTCGTTGAAGATTTTGTGTATGGCATCTGCACCTATGGTTGGGTGTGATGATTCTCG
AGACAAACATCGTTGGGTAGAGAGCAAAGTTTACATGAGAAAGGGTCATGGCAAAGGAT

CTAAAGGAAATGCTGGTTTTAATGCTCAAAATTCGACAGCACAAGTGCGGAGAGAGAAT
GATAACATGGGCAATTCGATTGCTGATAATGGAAAATCTGAAGCAGCATCGGAGGGACT
GTCGAGCCTTAGCAGGAAGCAAATAACTGTAAATCAGGATCATCCTCCGAATGAGACTA
GTAGCATGCCTGCTGTTGGAGGATTGCAGAATATAGATACTCATGTTACTTTTAAGTTGG
AGGGGTGCTCCAAGCAGGAGATATGGGAACTTCGCAAGAAATTGACTAACGAATTAGAA
CAGGTTAGGGGCACATTCAAGAAACTTGAAGCTAGAGAATTGCAGTTGAGGGGATATTC
TGTTTCGGCAGGAGTAAACACTAGTTACAGTGCCTCTCAGTTCTCTGGAAATGACATGAG
GAACAATGGCGGCAAGGAGGTCACATCTGAAGTGGCATCTGGTGGAGCGATCACGCCT
AAGCAGGCTCAACGAGAGTCTAATCCCCCTCGTCAGTTGAGCATTTCTCTCATGGAAAAC
AATCAAGCTGCAAGCGACATGGGGGAAAAGGGGAAGCGCACACCAAAGGCCAACCAGT
ATTACCGCAATTCGGAGTTTGTTTTGGGAAAAGATAAATTCCCACCTGCAGAGAGTAAGA
AATCTAAATCTACTGGTAACAAAAAGATTTCACAATCCAAAGTATTTTCTAAAGAAACAAT
GCAGGTAGGAAAAGAGTTCATGCCGCAGAAGTCTGTGAATGAGGTCTTCAAACAGTGCA
GCTTGTTGCTTACAAAGCTGATGAAACACAAGTATGGTTGGGTGTTTAACTTACCAGTTG
ATGCACAGGCGCTGGGACTTCATGATTATCATACCATAATAAAAAGGCCAATGGACCTTG
GTACTGTAAAGTCCAAGCTGGAGAAAAACCTATACAACTCACCTGCATCATTTGCCGAAG
ATGTGAAACTGACATTTTCTAATGCCATGACGTATAACCCAAAAGGGCATGAAGTTCATA
CAATGGCTGAGCAGTTGCTGCAATTATTTGAGGAGCGCTGGAAGACTATATATGAAGAA
CACCTTGATGGTAAGATGAGGTTTGGAAGTGGGCAGGGACTTGGGGCAAGCAGCAGTA
CTAAGAAACTGCCTTTCCAAGACTCGAAGAAGAATATAAAGAAATCAGAACCTGCTGGAG
GGCCATCTCCCCCAAAACCCAAATCCACAAATCACCATGCCAGTAGAACCCCAAGTGCC
AAAAAACCTAAGGCCAAAGATCCCCATAAACGTGACATGACTTATGAAGAAAAGCAGAAG
CTAAGTACAAACCTTCAAAATTTACCTCAAGAGAGATTGGAACTTATAGTTCAGATCATAA
AGAAGAGGAACCCATCTCTATGCCAGCATGATGAAGAGATAGAGGTCGACATTGACAGT
TTCGACACTGAGACTCTTTGGGAACTTGATAGATTTGTCACCAACTACAAGAAAAGCTTG
AGCAAGAATAAGAAAAAAGCGTTGCTTGCTGATCAGGCTAAAAGAGCCAGTGAACATGG
TTCAGCCAGAAATAAGCATCCGATGATTGGAAGAGAGCTACCAATGAACAATAAAAAGG
GAGAACAAGGTGAGAAGGTGGTTGAAATAGATCACATGCCACCAGTAAACCCGCCCGTT
GTTGAGGTCGAGAAGGATGGTGTTTATGCAAAAAGATCGAGCAGCTCTAGTAGCTCCAG
TAGTGATTCTGGGTCTTCTTCCAGTGACTCTGATTCTGGGAGTTCCTCAGGAAGTGAATC
CGATGCCTATGCTGCAACTTCACCTCCTGCTGGTTCGAACACATCAGCTAGGGGTTGAT
TCTGAGAATACTGTGGCAGAAGGTGAAGATGGTTTATCTTATTTGGAGTCAAATGCAGAT
GTTTGGTTTTCATTCGATGGAACCCAATTGCTTCAGAGGCGACAATGTAGGAATGCAATG
CATGTATTCTGTTTATTCACCATTGACTGTGGCATTGTTATAACCAAAAAGAGAGTAGTAT
ATTTTTAATGCAACCACATTCCTCTGGTAAGCTGCTTTCTCGATGGTCATATTATAGGATA
TGATCAGAAATCCCATCTGTATATTTTGAAAGGACATTGGTAAGAGTGTAAGAGAGACCA
CTATTGATGCTACTATAAAAAGGGATTATCTTTAGGCAGTTGCACTCTGCAGGTATGTTCT
TACCATCGTAGCCACAGGATAGGAACTACTGACTTAGTGCAAAAGACATGATAGACAGTA
AAAGAGCAGCTAAGCTCATACAATCCTATCGTGTCCTTTTTTATTGTGGCTGTATTGAAGA
AAGTAGATTTTAGCCCATTGTTCTGTACTTTGAATTATTGATTCAAAGGGTGTATCATAGG
ACTTGCTGTGACAAGTTTTGACAGCCCGTCTTATAAACTGTTTAGATGTAAAGTATATTTT
AGCCGCTGTTGTTGTAAATTTATGTTTTTCATTGCTATCAACATTTAGTTTTCAAAAAAAAA
A

SEQ ID NO: 152
GCTGCGATTGAAATTTAATCAGAAGTGAAATTGAAATTTAATCAGAACCCAAAGATGGAG
GGCCACTCGGGAGCTCTGGGATTTGGGCAAGGCTTCTCCAGGAGCTCTCAGTCTCCAA
ACCTTTCGCCCTCACCCTCGCACTCCGCCTCGGCATCAGTGACGAGCTCGGGCCAAAA
GCGCAAACGCAATGAAGTTGAGCACGCAGGGGTCGCTTCTAATTCCACGGGAATGTTCG
CAGTGCCCCCTTCCCACATATACTCCCACTTGCATCCCATGTCGATGTCGATGCCCATG
CCAATGCACAATTCGCACCCCTCCTCCCTCTCTGAGAGTCGAGACGGTGCCCTCACCTC
CAATGACGACGATGACAACCTCACCGGGGGCAACCAGTCGCAGCTGGACAGCATGAGC
GCAGGCAACACCGACGGCAGAGAGGATTTCGACGACGAAGATGACGATGATGACGATG
AAGAAGATGACGATGAAGTAGAAGGAGACGAAGAGGATCAAGATCATGACCCTGACGCC
GATGACGACTCGGACGATGGCCACGATTCCATGCGTACCTTCACCGCTGCTCGCCTGG
ACAATGGCGCTCCTAATTCCCGCAATTTAAAACCCAAGGCTGATGCTGCTGGAGTTGCA
ATTGCTCCCACTGTTAAAACTGAGCCAATTCTCGACACAGTCAAGGAAGAAAAAGTTAGC
GGTAACAACAACAACAACAGTGTCTCCGCTAACAATGCCCAAGTTGCCCCTTCTGGTTCC
GCTGTTTTACTCAGTGCTGTGAAAGAGGAGGCTAATAAGCCTACTTCTACCGATCACATC
CAGACCAGCGGAGCCTATTGTGCTAGAGAAGAATCCCTCAAGAGAGAGGAAGACGCAG

ATCGGCTTAAATTCGTGTGTTTCGGTAATGATGGAATTGATCAACACATGATCTGGCTTAT
TGGATTAAAGAATATATTTGCAAGGCAACTTCCAAACATGCCAAAAGAGTATATTGTTAGA
CTTGTAATGGACAGGAGCCACAAGTCAGTGATGATCATCAAGCAAAATCAAGTTGTTGG
GGGCATCACTTATCGCCCTTATCTTAGTCAGAAATTTGGAGAAATAGCTTTCTGTGCTATT
ACAGCGGATGAACAAGTAAAAGGCTATGGGACTAGGCTGATGAATCATTTGAAACAACAT
GCCCGTGATGTGGATGGGCTAACACACTTTCTAACATATGCTGACAATAATGCAGTTGGC
TACTTCATTAAACAGGATTTTACAAAGGAGATCAAGCTGGAGAAGGAGAGATGGCATGG
GTACATCAAGGATTACGATGGGGGAATCCTCATGGAGTGCAAAATAGACCCTAAGTTGC
CTTACACAGATCTACCAGCTATGATCCGCTGGCAGCGGCAGACAATAGATGAAAAGATA
AGGGAGCTCTCCAACTGCCATATTGTTTACTCAGGCATTGACATTCAGAAGAAAGAAGCA
GGCATTCCTAGGAAGCCCATCAAGGTTGAGGACATTCCTGGTCTAAAGGAGGCAGGCTG
GACAACAGATCAGTGGGGTCACTCGCGTTTTCGTCTACTGAATTCCCCATCCGAGGGCC
TGCCAAATCGCCAGGTCCTCCATGCCTTTATGCGGTCTCTACATAAGGCTATGGTCGAA
CATGCTGATGCTTGGCCATTCAAGGAGCCAGTAGATCCTCGTGATGTGCCTGATTACTAT
GATATCATAAAAGATCCAATGGATGTGAAGAGGATGTTTACAAATGCACGGACATATAAT
ACCCATGAAACTATTTATTATAAATGTGCAAACAGGTGATTATCTCACAGGTTGGATTCCT
ACTTTTTTAACAAACTTCAAGCTGGAACTCAGGTTAATAACAAGAGTCAGCAGTGATGGA
ATTGTCATCATAATTTCCTTTATTCAAATCAAACAAGCACTGGCTGTGCCCGCATGGTGCT
GTGCTTTATGCTTTTGCATGGTTTTCCTATAAGATGTATGAATTCGCACTGTGGTGCAATT
TTATGAATTAAACTCAAATGGAAAGGTTGCCGGCTTTTAACTATAGCTCTTGTAGATTTCC
TCAGTTGGCATGGAAGCAATTGGTGTTGCCACAGAAGAAGTTGGATTTTCCTTATATTTTT
TTACAAAAGCTGTTGTGTTTATAATTCTGCTTGTTAGTGTGAATAATGTTAATAGTTTGATC
TTCCTTGAATTTGTGGGGATACTGATATTTCTTTTTTGGGAAAGCTATGTGCTATAACAAT
AAATGCACCAATTGGTTTTTGGTTTGTCAAAAAAAAAA

SEQ ID NO: 153
GATCGCTCGAGAGTTCCTCTCTCAGCTGTTTTCCAATTCTGGACTAAACAAGGGTGCAG
GCGGCAGCCAGTTGAAACAGTGGCAGAGTGGGGGTGAACATGCAACTTAGTGGGAGCA
TGGACCCGCTGTGCAAACACACGCATAGCTGGAAGTGCTATTAACGAGTAAATACTACG
AAGCCATTCGGTCTCTGGATTCATCCGACTAATTTTGGGGTGCAACCATTTGCCAGGGG
CCCTCGATTAGACGGAGGGTGGAATGGAAGAATCGGGCAATTCACTGACCTCTGGTCC
GGATGGAAGTAAGCGAAGAGTATCTTATTTTTATGATTCAGACATCGGGAACTACTACTA
CAGTCAGGGCCACCCGATGAAGCCACATCGAATCAGAATGGCTCACAGTCTGATAGTTC
ACTATGCCCTGGATGAGAAGATGGAGGTTTGTAGACCTAATCTTCTACAGAGCAGAGAG
CTTAGAGTATTTCATGCAGATGATTACATATCTTTTCTTCAGTCGGTCACTCCTGAAACTC
AACACGAACAGCTACGACAATTGAAGAGATTTAACGTGGGGGAGGATTGCCCTGTTTTT
GATGGGCTATACAATTTCTGTCAGACCTATGCAGGGGGCTCTGTGGGAGCAGCTATCAA
ACTTAATAACAAGGAAGCAGATATTGCCATTAATTGGTCTGGTGGTCTCCACCATGCTAA
GAAGTGCGAGGCATCTGGGTTTTGCTATGTAAATGACATTGTCTTGGCTATTCTTGAACT
TCTCAAGGTCCATCAACGAGTGCTTTATATTGATATTGATATTCATCACGGGGATGGTGT
GGAGGAGGCCTTCTACAGTACTGACCGAGTCATGTCTGTGTCATTTCACAAGTTTGGAG
ACTACTTCCCAGGCACAGGCCACCTTAAAGATGTAGGATACGGGAAAGGGAAGTATTAT
TCATTAAATGTTCCTTTGAATGATGGAATTGATGACGAGAGCTACAAGAATCTCTTTAGAC
CTATAATCCAGAAGGTAATGGAAATATATCAGCCGGAAGCTGTTGTCTTGCAATGCGGAG
CAGATTCTCTGTCAGGCGATAGGCTAGGTTGTTTTAATTTATCAGTCAAAGGACATGCAG
ACTGTGTTCGATTTCTTCGGTCATTTAATGTTCCATTGGTGCTTGTTGGAGGCGGTGGTT
ACACTATTCGCAATGTTGCCCGCTGTTGGTGTTATGAGACTGCAGTTGCTGTAGGCGTC
GAACCCCAAGATAAATTGCCTTACAATGAGTACTATGAATATTTTGGGCCAGATTATACTC
TCCATGTGGCTCCTAGCAACATGGAGAATCAGAATTCAGCCAAGGAATTAGCAAAAATAA
GAAACACATTGCTGGAACAATTGAAACGAATACAACATGTTCCCAGTGTGCCATTCCAAG
AAAGGCCACCAGACACCAAGTTTCCTGAAGAGGATGAAGAGGATTATGAAAAGCGACCT
AAAGGTCATAAATGGGGTGGAGAATACTTTGGTTCGGAATCCGATGAAGAACAAAAACCT
CAGAATCGTGATATAGATATTTCGGACAAACCTGGCATAAGAAGGCAGTCACCTCCTAAT
GTTGAAGCGGCCAAAAAAATAAAAGTAGAAGAAGAAGATGGAGACATTGGAATAGTTAAT
GAGAACGATGGTGCAAAGTGGCCATTGGGTGAAGCTGGATAAAAAGTATATTATGGAGG
AAATAAGCAGGAGTTTCAGATCATTTGCACAATTCCAAGAAGATTATTTTTCTCCGGAAGC
CAAAGAGCTACCGGGAATGGGAATTTCAAATGGTCATTTCTATTTCTCTAATTTCCAGTCA
GATGGAATTGTACTTCTATCTTCATAGACTTCAAGGCTATTGTGTACATAAAGAAAGCAGA
TAATAATAAAATTTTGTCGGTTTTATTGAATGAAATTCCCCATAATGTTTTACTATTCCGTC

TGGGCTTAGAGATGTACGTTAATTGGTCATTTAAGACGACTCAGTTACATAGAAATCCAC
AAGTGGCATAGTACAAACATAAATATAAAGCTTTTTTTACCAAAAAAAAAA

SEQ ID NO: 154
CTCGAGAGTTCCTCTCTCAGCTGTTTTCCAATTCTGGACTAAACAAGGGTGCAGGCGGC
AGCCAGTTGAAACAGTGGCAGAGTGGGGGTGAACATGCAACTTAGTGGGAGCATGGAC
CCGCTGTGCAAACACACGCATAGCTGGAAGTGCTATTAACGAGTAAATACTACGAAGCC
ATTCGGTCTCTGGATTCATCCGACTAATTTTGGGGTGCAACCATTTGCCAGGGGCCCTC
GATTAGACGGTAAATAGTTGCTGAGATTTTTACTCGAATCTGCCAGTATTCAAATCTAGTC
AATATCCGTGTTGAGCTAAACAAGCGCTGAAAGTTTGCTCGAATCAGCAAGAGGGTGGA
ATGGAAGAATCGGGCAATTCACTGACCTCTGGTCCGGATGGAAGTAAGCGAAGAGTATC
TTATTTTTATGATTCAGACATCGGGAACTACTACTACAGTCAGGGCCACCCGATGAAGCC
ACATCGAATCAGAATGGCTCACAGTCTGATAGTTCACTATGCCCTGGATGAGAAGATGG
AGGTTTGTAGACCTAATCTTCTACAGAGCAGAGAGCTTAGAGTATTTCATGCAGATGATT
ACATATCTTTTCTTCAGTCGGTCACTCCTGAAACTCAACACGAACAGCTACGACAATTGA
AGAGATTTAACGTGGGGGAGGATTGCCCTGTTTTTGATGGGCTATACAATTTCTGTCAGA
CCTATGCAGGGGGCTCTGTGGGAGCAGCTATCAAACTTAATAACAAGGAAGCAGATATT
GCCATTAATTGGTCTGGTGGTCTCCACCATGCTAAGAAGTGCGAGGCATCTGGGTTTTG
CTATGTAAATGACATTGTCTTGGCTATTCTTGAACTTCTCAAGGTCCATCAACGAGTGCTT
TATATTGATATTGATATTCATCACGGGGATGGTGTGGAGGAGGCCTTCTACAGTACTGAC
CGAGTCATGTCTGTGTCATTTCACAAGTTTGGAGACTACTTCCCAGGCACAGGCCACCTT
AAAGATGTAGGATACGGGAAAGGGAAGTATTATTCATTAAATGTTCCTTTGAATGATGGA
ATTGATGACGAGAGCTACAAGAATCTCTTTAGACCTATAATCCAGAAGGTAATGGAAATA
TATCAGCCGGAAGCTGTTGTCTTGCAATGCGGAGCAGATTCTCTGTCAGGCGATAGGCT
AGGTTGTTTTAATTTATCAGTCAAAGGACATGCAGACTGTGTTCGATTTCTTCGGTCATTT
AATGTTCCATTGGTGCTTGTTGGAGGCGGTGGTTACACTATTCGCAATGTTGCCCGCTGT
TGGTGTTATGAGACTGCAGTTGCTGTAGGCGTCGAACCCCAAGATAAATTGCCTTACAAT
GAGTACTATGAATATTTTGGGCCAGATTATACTCTCCATGTGGCTCCTAGCAACATGGAG
AATCAGAATTCAGCCAAGGAATTAGCAAAAATAAGAAACACATTGCTGGAACAATTGAAA
CGAATACAACATGTTCCCAGTGTGCCATTCCAAGAAAGGCCACCAGACACCAAGTTTCCT
GAAGAGGATGAAGAGGATTATGAAAAGCGACCTAAAGGTCATAAATGGGGTGGAGAATA
CTTTGGTTCGGAATCCGATGAAGAACAAAAACCTCAGAATCGTGATATAGATATTTCGGA
CAAACCTGGCATAAGAAGGCAGTCACCTCCTAATGTTGAAGCGGCCAAAAAAATAAAAGT
AGAAGAAGAAGATGGAGACATTGGAATAGTTAATGAGAACGATGGTGCAAAGTGGCCAT
TGGGTGAAGCTGGATAAAAAGTATATTATGGAGGAAATAAGCAGGAGTTTCAGATCATTT
GCACAATTCCAAGAAGATTATTTTTCTCCGGAAGCCAAAGAGCTACCGGGAATGGGAATT
TCAAATGGTCATTTCTATTTCTCTAATTTCCAGTCAGATGGAATTGTACTTCTATCTTCATA
GACTTCAAGGCTATTGTGTACATAAAGAAAGCAGATAATAATAAAATTTTGTCGGTTTTAT
TGAATGAAAAAAAAAA

SEQ ID NO: 155
GGGGAAAACAAAGGCAGCATAAACCCTAAACCCGAAATTCCCTCTAGCCTCTGAACCTA
AACCCTAAATGTCTTCTCGACTTTGAAGCCTATAGAGTTCTCCAAAAGCTCAGTTTATCTA
TTTAGTTCATATTTTTCGTTGTGTGGTTTTACGGTTTGCGGCCAACGTTATTGTTCTCTTT
CGTGGAGCGTAATCGATCTGCAATCTTCGGGCGTCGTTCGTGCTCTCTAAGGGAGGAAG
AAGATATTTCATTCATCGTCATGGAGTTTTGGGGCGTTGAAGTGAAGCCTGGAGAGGCTT
TGACATGTGATCCTGGGGATGAAAGATACCTTCATATGTCCCAGGCTGCCATTGGGGAC
AAAGAAGGTGCTAAGGAAAATGAAAGGGTCTCCTTGTATGTCCACGTGGATGGTAAAAA
GTTTGTGCTTGGGACCCTGTCCCGTGGAAAATGTGACCAGATTGGACTGGATTTAGTGT
TTGAGAAAGAGTTCAAGCTCTCTCATACATCTCAAACTGGGAGCGTATTCGTATCTGGCT
ACACTACAGTAGATCATGAAGCCCTGGATGGATTTCCGGATGATGAAGACCTAGAGTCA
TCGGAAGATGAGGAAGAAGAATTAGCTCAGATAACGACTTTAACTGCAAAGGAGAATGG
TGGCAAAACCGGAGCCAAGCCTGTGAAGCCTGAATCTAAGTCATCTGTAACTGATAAAG
CAGCAGCTAAAGGGAAGCCAGAGGTGAAGCCACCAGTCAAAAAGCAAGAAGATGATTCT
GATTCTGATGAAGATGAAGATGAGGATGAAGATGAAGATGAAGACGATGATGATGAAGA
TGATGAGGATATGAAAGATGCGAGTGCAAGTGATGATGGGGATGAGGAAGATGACTCCG
ATGAAGAAAGTGATGATGATGAGGAGGAAGATGAGGAAACCCCAAAGCCGGCTGCTGG
GAAGAAGCGACCTATGCCAGCCTCAGACAACAAGTCACCAGCAACAGATAAGAAGGCAA
AGATCACCACACCAGCTGGAGGTCAAAAGCCAGGTGCTGACAAAGGCAAGAAAACAGAA
CACATTGCAACTCCTTATCCAAAACACGGTGCAAAGGGCCCTGCTTCAGGTGTTAAGGG

CAAAGAAACTCCACTTGGGTCAAAACAAACACCTGGTTCAAAAGTGAAAAATTCGTCTAC
ACCTGAGTCTGGCAAGAAATCTGGCCAGTTTAAATGCCAGAGTTGTTCAAGGGACTTTG
CAACAGAAGGAGCATTGAGCAGCCACAATGCTGCGAAGCATGGTGGAAAGTGAAGTATT
TCGTTTTTACATTTTCACAACTGTTTTTGCACTCTCAGTGCATCGTTGCCAGAAAATTCCT
AAACTAAGGTCAACTGTGTATCAGATCTCTAGTCAGTGTTTAGAGATTTACAGAAAGTTGT
GTACTAATTTGTATTGTAACGTCCATTTATCCAACGAGTCCTCCATTCATGCTATTTTTGT
GAGTTGCTTCTAAGGGAGGATTTCTCAAGGTAGCACAATTTTGTATTAGAGGTTAAATCT
AGTAGATGTCTGTTATGTGCGGCATTGTTGAATTGAATGATTTTATTTTTTCAAATTTGTC
GGAGGCTACTGATTTAAAAAAAAAA

SEQ ID NO: 156
ATTGATTTAAGTGGGGTTCAAGAAAATCGTATCCCCAATCCGCATCTTGGGTGCTTACAT
CAATCTGTTTATATCAACCAATCCTCGCAACAAATTGGGGAAGAAAGGCCGTTCAAATGT
GGCATGGTATATTCTCTGTGTATAATCTTCGTGTCTGGATTTTTTCCGGCCCAAGCAGCC
CAAGCAGACCATGTAGGCTTGTTGTTAAACCATGATTAATTGACTTCATGGGTGTTTTCC
AGTCTGAAATTTGGTGGAGGACTAGAACCAGTTTTTAATAGTTTGTTTTGGCAGATGAAC
TTTATCCTTCTGAAAGTTTTTGAGAACAAAAGGAATTAGCGGTTGTTTATGAAGAAGAGG
CTCCATGATGGAGACTGGTGGCAATTCTCTTCCATCCGGACCGGATGGAGTGAAGAGGA
AAGTGGCATATTTTTATGACCCAGAGGTGGGGAATTATTACTATGGACAAGGGCATCCGA
TGAAGCCTCACCGCATAAGAATGACTCATGCCCTCCTTGTTCAGTATGGTTTGCATAAAG
AAATGCAGATTCTTAAGCCCTACCCTGCGAGAGATAGAGATCTCTGCCGTTTCCATGCAG
ATGACTATGTAGCATTTTTACGAGGAATAACTCCGGAGACAATACAGGATCAAGTTAAGG
CGTTGAAACGGTTTAATGTTGGGGATGATTGCCCGGTTTTTGATGGCTTGTACCAATACT
GTCAAACTTATGCTGGGGGGTCTGTCGGAGGTGCTGTGAAACTCAATCATAAACTCTGT
GACATTGCAATTAATTGGGCTGGGGGTCTTCATCATGCTAAAAAATGTGAGGCATCCGG
ATTTTGCTATGTGAATGACATTGTGCTGGCCATTCTGGAGCTCTTGAAATATCATAAGCG
TGTTTTGTATGTTGATATTGACATTCACCATGGTGATGGAGTTGAGGAGGCTTTCTACAC
AACGGATAGGGTTATGACAGTATCATTTCACAAATTTGGGGACTATTTCCCAGGGACAGG
TGACATTAGAGACATCGGGTGCGGGAAAGGGAAATATTATGCTGTAAATGTCCCTTTAGA
TGATGGAATTGATGATGAAAGTTTTCAGTCTCTGTTTAAACCTATTATTCAACAAGTGATG
TTGGTTTACAATCCTGAGGCAATTGTTCTCCAGTGTGGAGCTGATTCTTTGTCTGGAGAC
AGGCTAGGTTGCTTTAACCTTTCTGTGAAAGGGCATGCAGAGTGTGTTCGATACATGAG
GTCATTCAATGTGCCACTTCTGATGGTCGGTGGTGGTGGGTACACAGTTCGTAATGTTG
CTCGTTGTTGGTGCTATGAGACAGGGGTTGCTGTTGGTGTGGAGATTGATGACAAAATG
CCTCAACATGAATACTACGAATATTTTGGACCAGATTACACTGTTCATGTGGCTCCAAGT
AATATGGAAAATAAGAACACGAAGCAGTATCTGGATAAAATTAGGTCCAAAATACTTGAA
AACATTAATTCATTGCCATGTGCGCCAAGTGCTCAATTCCAAGTGCAGCCTCCAGATACT
GATTTCCCTGAACTGGAAGAAGAGGATTATGATGAACGTACTAGAAGTCATAAATGGGAT
GGAGCAAGTTGTGATTCTGATTCAGAGAATGGTGATTTAAAGCATCGTAATCATGATGTT
GAGGAAAGTGCATTCCCACGGCATAATCTGGCCAATATTAGTTACAATACCAAGATTAAG
CTAGAAGGTGTTGGTACAGGTGGTCTTGACATGGCTGCTGGAACAGACACAAAAAAGAA
TGACGAGTCCTTTGAAGCCATGGACTATGAAAGTGGAGAAGAGTTACGACAGGATCATT
TTGCTTCTACAATAAATGCTTCTCAACCATGTGATCCTGCTCTTTTGACAGGAGTTCAAAA
TCAGTTGCAGTCAACAGATACTGTAAAGCCCATAGAACAAAGTGGCAATGCTCCAGGAAT
ACCACCACCTTCAGTGGCAACTGTAAGCACGGGCACTCGCCCGAGTTCAATTAGCCGGA
CATCATCTTTGAATTCAATGTCTTCAGTAAAGCAGGGTTCCATTTTGGGACCAAATCCTCC
TCAAGGGCTGAATGCATCTGGTCTACAATTTCCAGTTCCAACTTCTAATTCACCTATAAG
GCAAGGTGGGTCTTATTCTATAACTGTACAGGCCCCAGACAAGCAAGGTTTGCAAAATCA
TATGAAGGGACCTCAAAATATGCCAGGGAATTCTTGAAAATTATATATCCTTCTGCAATTT
GAAACACAGGTTACTTTGTATAATGATAATGTAATGTCATCCCCTTGGTCGCAATGTGCTT
GAGATCAGTTTAGAAGCTAATTATACACTTAGAATGCTAATTGGTGAGGTGCTAAAACTTC
TCTGTATTTTATTCTACAGTACTGTATTCGAAGATCCTGAAAATTTACTAAAACAAATGGAA
TATCAACAACCTAGGATCATTATAGCAAAAAAAAAA

SEQ ID NO: 157
AATTACCATACACCACAGCATAATCAACTGGGCGGCTTAATTCTGGAACGCGAAACGAC
CGCAGGCCGAATCGAATTTCATACAGTACATGGAAGAGGAAGCCACGAAAGCCGGGAA
GCAACGCGTAGGGTAGGTGTCGTTTTTGCCCGTTGGAAATGCCTCCGAAAGATAGAGTC
GCCTACTTCTACGACGGGGATGTGGGAAGTGTCTACTTTGGCCCAAATCATCCTATGAA
GCCACATCGACTGTGCATGACACACCATCTGGTTCTTTCTTATGAACTTCACAAGAAGAT

GGAGATCTATCGACCACACAAGGCATATCCAGTGGAACTTGCTCAGTTTCATTCAGCAGA
CTATGTGGAGTTCTTGCATCGAATAACGCCTGATACACAGCACTTGTTCACCAAGGAGCT
AGTAAAATATAATATGGGAGAAGATTGTCCTGTATTTGAGAACCTCTTTGAGTTTTGTCAA
ATTTATGCTGGTGGCACTATAGATGCTGCTCATCGACTGAACAATCAGATTTGTGATATT
GCAATCAATTGGTCTGGTGGTCTTCACCATGCTAAAAAGTGTGAGGCATCAGGTTTCTGT
TATATTAATGACTTGGTGCTGGGAATTTTGGAGTTGCTTAAACATCATGCCAGAGTTCTCT
ATGTTGATATTGATGTCCACCATGGTGATGGAGTTGAAGAAGCCTTTTATTTTACTGACA
GGGTGATGACAGTGAGCTTTCACAAGTACGGTGACATGTTTTTTCCTGGAACAGGTGAT
GTTAAGGAGGTTGGAGAAAGAGAAGGAAAATATTATGCTATTAATGTTCCTCTCAAAGAT
GGAATTGATGATGCTAGTTTCACGCGACTTTTCAAAACTATTATTACCAAGGTAGTTGACA
TATACCAGCCTGGTGCCATTGTTCTTCAGTGTGGAGCTGATTCACTTGCTGGGGATCGC
CTTGGTTGCTTTAACCTATCCATCGATGGTCATGCACAGTGTGTGAGAATTGTGAAGAAA
TTTAATCTGCCACTATTGGTTACTGGGGGAGGAGGCTACACAAAAGAAAATGTTGCTCGT
TGTTGGTCCGTGGAAACTGGAGTACTTCTTGATACAGAACTTCCTAATGAAATTCCTGAT
AATGATTACATCAAGTACTTCGCACCTGACTATTCCTTAAAGATCAACACTGCTGGGAAC
ATGGAGAATCTGAATAGCAAGACATACTTGAGCGCAATTAAAGTGCAGGTTATGGAGAAT
CTGAGGGCCATTCAACATGCACCAAGTGTACAGATGCATGAAGTGCCTCCTGACTTTTAT
ATTCCTGATATTGATGAAGATGAACTTAATCCGGATGAACGTATGGATCAGCATACTCAA
GACAGACAGATCCAGCGAGATGATGAATATTATGATGGAGACAATGACATTGATCATGAC
ATGGAAGAGGCAAGCTGATAGCCGAGAAAGATATTTGCCCCTTTGATGCAATATTAATGA
TAACGCTTTGTTGGATGTCTTTTCTTTCTACAATGACAATAAATTAGTTTTTGTAATTTATA
AAGTATTCTCGAGTGTACTTTCAGAAATTTTGAGGAGCCCTTTTCCGCCAGAAATGATTTA
TGTTGCTTTTCTGTAATAGTTCTTCCCTCCAGGGATTCTGGCTGTGTTTCTCCCCCAGTTT
TCTGAAAATCTTGGTGATAACACCAGAAGTAGAAAACTTTGCTCTATATAATTTGTGCTCG
TGTGTGTACTTGAAGATCCATCCTCACATAGTCCAATGAAGTTTTGCAGTATCATATGCC
CAATTTTTTTTGGAAAAAAAAAA

SEQ ID NO: 158
GCTCCCTTACATCGCTGGTCCGATCGTTGATAGTCACCCCACCTAGAGCCTCCCGGGTT
GTTGCATGAACTGAGATGCCTAACAATCGAAAATCGAACCCCACCATTTCAAGATCAAGA
CTCCCACATTCGCACCCAGCATAAAACGTCTCCCATCCCCAACAGATCTCGTTACCTCAG
AACAAATTAACGCCACAATCGCCATGGATTCTTCCAAGTCAGAAGAAGCAAACATTTTGC
ATGTGTTTTGGCACGAAGGAATGCTCAATCATGATTTAGGCACAGGGGTATTTGACACAT
TGGAGGATCCAGGCTTCCTGGAGGTGCTGGAAAAACATCCTGAGAATGCAGATAGAGTG
AGAAACATGCTCTCAATTCTCAGAAAGGGACCCATTGCTCCCTACACCGAATGGCACACT
GGCAGGGCTGCATATCTCTCAGAACTGTACTCCTTTCACAGACCAGATTATGTAGACATG
CTTGCCAAAACTAGTACAGCAGGTGGAAAGACCTTATGCCATGGCACACGCCTGAATCC
TGGTTCTTGGGAAGCTGCACTTCTTGCTGCTGGGACAACACTTGAAGCTATGCGCTATAT
TTTGGATGGACATGGAAAGCTATCTTATGCATTAGTGCGACCACCAGGCCACCATGCAC
AGCCAACTCAAGCTGACGGGTACTGCTTTTTGAATAATGCTGGTCTTGCAGTTGAACTAG
CTGTGGCATCTGGCTGCAAGCGAGTTGCTGTGGTTGACATCGATGTGCATTATGGGAAT
GGAACTGCAGAGGGATTCTATGAACGAGATGATGTGCTCACAATCTCTCTCCACATGAAT
CATGGTTCTTGGGGACCTTCACATCCTCAAACTGGATTTCATGATGAAGTTGGTCGAGGA
AAAGGTCTGGGGTTTAATCTAAATGTTCCTTTGCCTAATGGGACAGGAGATAAGGGATAT
GAGCATGCCATGCATGAATTGGTGGTACCAGCCATTAGCAAATTTATGCCTGAAATGATA
GTCTTGGTTATTGGTCAAGACTCCAGTGCATTTGATCCTAATGGAAGGGAGTGCTTGACA
ATGGAAGGTTATAGAAAAATTGGTCAAATAATGCGTCAACAAGCTGATCAATTTAGTGGC
GGCCGCCTTGTGGTGGTACAGGAAGGTGGCTATCATATTACTTATGCAGCATATTGTCTT
CATGCCACTCTTGAAGGTGTCCTATGTCTGCCACATCCACTTTTGTCAGATCCAATTGCT
TATTATCCAGAGCATGATATTTATAGTGAAAGAGTAACATTTATAAAAAATTACTGGCAAG
GAATCATCTCGACCACAGATAAACGTAACTGATTAGTTTACGATATGTGGTTGTGACTCT
GAGTATTTGAAGCTGTTTTACAGCTGTATATGGATAAATGTTATTAGTGTGTTGCAACAAT
ACGCTTTACCTTGTGTGTATAGTTTTATAACACTCTATGGTATCACTACCACTATGGGCCT
GTTTAGTCCAAGTTTCCTTTTTAATATTTGCCGGTATAAAGTATATGATTTTCTTAGTGTAA
AAAAAAAA

SEQ ID NO: 159
GAAACACTGATATATATACCTAAATCCCAGGTTTAAATAAAAGCTACATGTGTTCGTCGCA
TCCTTCTCTTCTCTGATTGGGGATTTCTCTGGATTTTGCTATACAGTCAGATAGCAACGAA
ACATAGCAGAGGCACGTAAGCTTTGGTTTAACCTGTAGGCATTCGTTCTACTGTGTATTT

ATCTAGTCCATAGGAACGAGACATCGGAGATTAGAGGTGAGAGGCCATGGAGGAGTCT
GGCAATGCTCTGGTATCAGGGCCTGATGGGAGTAAGAGAAGAGTTACATACTTCTACGA
TGCAGACATTGGTAATTATTACTATGGGCAAGGCCACCCAATGAAGCCACACCGCATGA
GAATGGCCCATAACTTGATTGTCCACTATGGCCTCCACCAGAGGATGGAAGTTTGCAGG
CCTCATCTAGCACAGAGCAAGGACATCCGTGCCTTTCATACTGATGATTACATACATTTC
TTGAGTAGTGTAGCACCAGACACTCAGCAGGAGCAGCTGAGGCAGTTGAAGAGGTTCAA
TGTGGGGGAGGACTGTCCTGTGTTTGATGGGCTTTTCAATTTTTGCCAGTCTTCTGCAGG
AGGCTCAATTGGGGCAGCCCTTAAACTCAATAGGAAGGATGCAGACATTGCCATCAACT
GGGCGGGGGGCCTCCACCATGCCAAGAAGTGCGAGGCTTCGGGTTTCTGCTATGTCAA
TGATATTGTGCTTGGCATCCTGGAATTGCTAAAGGTCCATCAGCGTGTACTTTATATTGAT
ATTGATATTCATCATGGTGATGGAGTTGAAGAAGCCTTCTACACTACAGACCGTGTGATG
ACTGTATCTTTTCACAAATTTGGGGATTACTTTCCAGGCACAGGTCATATTAAAGATGTAG
GATATGGGAAAGGGAAGTACTATGCTTTGAATGTTCCTTTGAATGATGGAATTGACGATG
AGAGCTACAAGCACCTTTTCAGGCCCATTATCCAGAAGGTAATGGAAGTCTATCAGCCA
GAAGCAGTTGTTCTACAATGTGGAGCAGACTCACTCTCAGGGGACAGGTTAGGGTGCTT
CAATTTGTCAGTCAAGGGACACGCAGATTGTGTTCGTTTTGTTCGGTCATTTAATATACCA
CTTATGCTGGTTGGTGGTGGTGGTTATACTATTCGTAATGTGGCTCGCTGCTGGTGCTAC
GAGACTGCTGTTGCTGTAGGTGTTGAACCCCAAGATAAATTGCCTTATAACGAGTATTAC
GAATATTTTGGGCCAGACTACACCCTCTATGTAGCCCCTAGCAATATGGAGAACCTGAAC
ACAGAAAAAGATTTGGAGAAAATGAGAAACGTGTTGCTAGAACAATTGAGTAAAATACAA
CATACACCAAGTGTACCCTTTCAGGAAAGGCCCCCTGATACAGAATTTAATGATGAGGAG
GAAGAGGACATGGAAAAACGCTCAAAATGCCGTATCTGGGATGGGGAGTACGTTGGTTC
AGAACCTGAGGAAGATGGAAAGCTTCCAAGATTTGATGCAGATACTTATGAGAGATCTGT
TCTCAAGCATGAAAACAAAAGGTTAGTGCCTGTTTCAAATGTTGAACCTCTGAAGAGAAT
AAAACAAGAGGAAGATGGAGCAGCTGTTTAAGTACTAGATGATCTTAATTGCACTGCCAT
TTCCTTCAGATGGTCACTTCTCTCATGTTCATTCAATCATTTTAATTTTCATTCTAATATGG
TTCATGCCCTAGATGAAGAAAGAAGGCCTGTGAACTTCTTTTAATTTTGATCCTGGCTTTA
AGGCTCATTTTAGAGAAGGAACTTTCATGCACTACATGGCTTAAAGATATTGATTTTGTGC
TGTGAAAGCAGAACGAAAATGAAAATGGCACTGCCACCCGGACATTTGTCCAATCTGCA
TGAGCACACTGAGGGAAAAGCTGAAGCAGAATCAGCTTTGACCAGTATTTAGTGTCTTGT
ATACAATTCTTGTTTCAGTGAATGCTGCTCCCCCTTCTCGATTACCAACCTGTACTTAAAC
ATTGTTGCTATCAGAAGAGTATTATTTTTCAATGAATGCAAGATAAATGTAAAAAAAAAA

SEQ ID NO: 160
AATTACCATACACCACAGCATAATCAACTGGGCGGCTTAATTCTGGAACGCGAAACGAC
CGCAGGCCGAATCGAATTTCATACAGTACATGGAAGAGGAAGCCACGAAAGCCGGGAA
GCAACGCGTAGGGTAGGTGTCGTTTTTGCCCGTTGGAAATGCCTCCGAAAGATAGAGTC
GCCTACTTCTACGACGGGGATGTGGGAAGTGTCTACTTTGGCCCAAATCATCCTATGAA
GCCACATCGACTGTGCATGACACACCATCTGGTTCTTTCTTATGAACTTCACAAGAAGAT
GGAGATCTATCGACCACACAAGGCATATCCAGTGGAACTTGCTCAGTTTCATTCAGCAGA
CTATGTGGAGTTCTTGCATCGAATAACGCCTGATACACAGCACTTGTTCACCAAGGAGCT
AGTAAAATATAATATGGGAGAAGATTGTCCTGTATTTGAGAACCTCTTTGAGTTTTGTCAA
ATTTATGCTGGTGGCACTATAGATGCTGCTCATCGACTGAACAATCAGATTTGTGATATT
GCAATCAATTGGTCTGGTGGTCTTCACCATGCTAAAAAGTGTGAGGCATCAGGTTTCTGT
TATATTAATGACTTGGTGCTGGGAATTTTGGAGTTGCTTAAACATCATGCCAGAGTTCTCT
ATGTTGATATTGATGTCCACCATGGTGATGGAGTTGAAGAAGCCTTTTATTTTACTGACA
GGGTGATGACAGTGAGCTTTCACAAGTACGGTGACATGTTTTTTCCTGGAACAGGTGAT
GTTAAGGAGGTTGGAGAAAGAGAAGGAAAATATTATGCTATTAATGTTCCTCTCAAAGAT
GGAATTGATGATGCTAGTTTCACGCGACTTTTCAAAACTATTATTACCAAGGTAGTTGACA
TATACCAGCCTGGTGCCATTGTTCTTCAGTGTGGAGCTGATTCACTTGCTGGGGATCGC
CTTGGTTGCTTTAACCTATCCATCGATGGTCATGCACAGTGTGTGAGAATTGTGAAGAAA
TTTAATCTGCCACTATTGGTTACTGGGGGAGGAGGCTACACAAAAGAAAATGTTGCTCGT
TGTTGGTCCGTGGAAACTGGAGTACTTCTTGATACAGAACTTCCTAATGAAATTCCTGAT
AATGATTACATCAAGTACTTCGCACCTGACTATTCCTTAAAGATCAACACTGCTGGGAAC
ATGGAGAATCTGAATAGCAAGACATACTTGAGCGCAATTAAAGTGCAGGTTATGGAGAAT
CTGAGGGCCATTCAACATGCACCAAGTGTACAGATGCATGAAGTGCCTCCTGACTTTTAT
ATTCCTGATATTGATGAAGATGAACTTAATCCGGATGAACGTATGGATCAGCATACTCAA
GACAGACAGATCCAGCGAGATGATGAATATTATGATGGAGACAATGACATTGATCATGAC
ATGGAAGAGGCAAGCTGATAGCCGAGAAAGATATTTGCCCCTTTGATGCAATATTAATGA
TAACGCTTTGTTGGATGTCTTTTCTTTCTACAATGACAATAAATTAGTTTTTGTAATTTATA

AAGTATTCTCGAGTGTACTTTCAGAAATTTTGAGGAGCCCTTTTCCGCCAGAAATGATTTA
TGTTGCTTTTCTGTAATAGTTCTTCCCTCCAGGGATTCTGGCTGTGTTTCTCCCCCAGTTT
TCTGAAAATCTTGGTGATAACACCAGAAGTAGAAAACTTTGCTCTATATAATTTGTGCTCG
TGTGTGTACTTGAAGATCCATCCTCACATAGTCCAATGAAGTTTTGCAGTATCATATGCC
CAATTTTTTTTGGAAAAAAAAAA

SEQ ID NO: 161
GGAGAGGTTCTGTCTGAGAGAAGAAGATGGATTTGAACTTGGTGAGCCATGGAGAAGAA
GAAGAAGGGGTAAGGCGTCGGAAAGTAGGAATTGTATACGACGAACGAATGTGCAAGC
ATGCTACTCCTGAGGATCAACCACATCCCGAACAACCAGATCGCATTAGGGTGATATGG
GACAAGCTCAACTCCGCCGGGGTCCTCCATAAATGTGTTATGGTGGAGGCGAAAGAAGC
ATCGGAGGAGCAATTGGCGGGGGTCCATAGTCGGAAACACATTGAGGTAATGAAAAGCA
TTGGCACTGCTAGATATAATAAGAAGAAGCGGGACAAGTTGGCAGCGTCTTACAGTTCC
ATTTATTTCAGCCAAGGCTCCTCGGAAGCCGCCCTTCTCGCTGCCGGATCCGTGGTAGA
GATATCTGAAAAAGTAGCTTCAGGGGAATTGGATGCCGGAGTTGCTATTGTTAGGCCAC
CAGGTCATCATGCAGAGGCTGACAAAGCCATGGGGTTTTGCTTGTTCAACAACATAGCT
ATTGCAGCAAAACACCTCGTCCATGAAAGGCCAGAGTTAGGTGTACAGGAAGTGTTGAT
TGTTGACTGGGATGTTCACCATGGCAATGGGACACAGCACATGTTTTGGACTGATCCAC
ATGTTTTGTATTTTTCTGTTCATAGATTTGATGCAGGAACATTCTATCCAGGAGGAGATGA
TGGGTTCTATGACAAAATTGGAGAAGGGAAAGGAGCGGGATACAATATAAATGTTCCTTG
GGAGCAAGGAAAATGTGGAGATGCAGATTACCTTGCTGTTTGGGACCATGTTTTGGTTC
CTGTTGCAAAGAGTTATGATCCAGATATGGTTCTTATTTCTGGAGGGTTTGATGCAGCAC
TTGGTGACCCATTGGGTGGATGCCGACTTACACCTTATGGGTATTCACTAATGACAAAGA
AGTTAATGGAGTTTGCAGGTGGGAAGATTGTCTTGGCCTTAGAGGGTGGCTATAATCTTA
AGTCGTTGGCAGATTCATTCCTGGCTTGTGTAGAAGCTTTGCTTAAAGATGGACCTAGTA
GAAGCTCTGTTTTGACTCATCCATTTGGATCAACATGGCGTGTAATACAGGCGGTGCGC
AAAGAACTTAGCTCATTTTGGCCAGCATTAAATGAAGAACTACAATTGCCAAGACTGCTG
AAGGATGCCTCAGAATCATTTGACAAGCTAAGTTCCAGTTCAAGTGATGAAAGCTCTGCT
TCTGAAGATGAAAAAAAAAATTGCTGAAGTGACGTCAATCATGGAAGTCTCTCCTGATCCA
TCCAGCATTCTTGCCCTAACTGCTGAAGACATTGCTCAACCACTGGCTGGGTTGAAAATT
GAAGAGGCTGGTACTGATAGCCAGAGATCATCAGATCACACTTTGTTAGATTTAACTAAT
GATGACACCCAAAAGTTGAAACAGTTCGAGGGGGAGATTTTTGTCATGATTGGTGATGAA
GAATCAGTTCCATCAGCCTCAAGTAGCAAAGATCAGAATGAGTCCACTGTAGTTTTATCA
AAAAGTAATATTAAAGCTCATAGCTGGAGATTGACTTTCTCCAGCATTTATGTTTGGTATG
CAAGTTATGGGTCAAATATGTGGAATCCAAGGTTTCTCTGTTATATTGAAGGGGGGCAGG
TTGAAGGCATGGCAAAGCGCTGCTGTGGTTCAGAGGATAAAACTCCTCCTCAAAGGATA
CAGTGGAAAGTTGTTCCTCATCGAATGTTTTTTGGGAGATCATACACAAATACATGGGGT
TCAGGAGGAGTGTCCTTTCTTGATCCAAATTGTAGTGATACGAGTGAAGCGCATGTCTGC
TTGTATAAAATAACGCTGGCACAGTTCAATGATTTGCTGCTTCAAGAGAATAATTTGAATT
GTGGGACTGAGCATCCATTAGTGGACTTATCTTCCATTGATGCAATTAGAAATGGGAATT
CTATACTTGAACTTATCAAGGATAGTTGGTATGGCACTCTCATTTACCTTGGCATGGAAG
GAGGCCTTCCAATTGTGACTTTCACGTGCTCCGTGTGTGACGTTGAAAAGTTCAAGCATG
GACAACTACCGCTTTGTCCTCCATCTTCAAGGTACGAAAATATTCTCATTCGAGGTCTCG
TCCAAGGGAAAAAACTTTCTGAAGATGATGCAACGGCCTACATCCGTGCAGCATCAACAT
CACCCCTGCTGTAAATAAAACATGAACTGATCATTTATATTCATATCTATTATTTCTTTCCC
TTGGCACGCTACCAGTTCTGCCACTCTATAATAGCAGCATAGATTCAGTAGAACCACACT
TACATGCTCATGATTAATCGATTTTAAAGAATGCCCAAGGCAATCAATTCATGTCCGTGTC
CATATGGGGTCATTCTTTCTATTCAGGTGCATGCAGTACATTGAGATGATGTGCAGGATA
TTCAACCGAGTGCAGTCATTCTTTCAAACAACTTGACAGGTGTAAACTCGTGCCGAATCG
G

SEQ ID NO: 162
TCCAGTCAATTTATAGGCATTGGCCGTTGCAGTTCCTGTACTCTCTGTTGTTTCTCTTGAA
GGAAAATCTATGGCGGACGAAGATCTGGATCTTTCCGATGTAGGGGAAGTAGAAGATGA
ACCAGGTGAGGAGATCGAATCCACTCCACCCCTCGCTGTAGGGCAGGAGAAGGAGATA
AACAGCTTGGCTTTGAAGAAAAAACTTCTGAAGGTTGGCACAAGGTGGGAAACGCCAGA
AAATGGAGACGAAGTAACAGTGCATTACACTGGAACTTTACCAGATGGTACCAAATTCGA
TTCCTCGAGAGACAGAGGAGAGCCCTTTACTTTCAAGCTTGGCCAAGGCCAAGTTATCA
AAGGATGGGACCAAGGCATTGTTACCATGAAGAAAGGGGAACGTGCACTATTCACAATA
CCTCCAGAATTGGCCTATGGTTCTTCTGGCGTGCGACCCACTATACCTCCCAATGCCAC

CCTCCAATTTGATGTGGAATTACTTTCGTGGACTAATATTGTTGATGTATGCAATGATGGA
GGCATTCTCAAGCGAATAATATCTGAAGGAGAAAAATATGAAAGACCCAAAGACCCAGAC
GAGGTGACTGTGAAATATGAAGCAAAGCTTGAGGATGGAACACTCGTGGCAAAATCCCC
AGAAGAAGGCGTGGAGTTTTATGTCAATGATGGACACTTCTGTCCAGCTATTGCGAAGG
CAGTCAAAACAATGAAGAGAGGAGAAAGTGTGATTCTCACCATAAAACCTACGTATGCCT
TTGGTGAGCGGGGTAAGGATGCTGAAGAGGGGTTTGCTGCGATTCCTCCAAATGCTACT
CTCACTACGAGTTTGGAATTAGTTTCATTTAAAGCTGTGATAGCAGTAACGGAAGATAAG
AAGGTTATCAAGAAAATCCTGAAGGAAGCTGATGGGTATGATAAACCAAGTGACGGGAC
GGTAGTGCAGATCAGGTATACTGCTAAGCTGCAAGATGGTACAATCTTTGAGAAGAAGG
GATATGAAGGCGAGGAGCCTTTCCAATTTGTAGTTGATGAAGAACAAGTCATTGCTGGTC
TTGATAAGGCAGTGGAAACCATGAAGACAGGGGAGATTGCCCTGATTACAATAGGAGCT
GAATATGGTTTTGGAAATTTTGAAACCCAGAGAGATCTGGCTGTAATTCCTCCAAATTCAA
CTCTCATCTATGAGGTGGAAATGATATCATTTACCAAGGAAAAAGAATCGTGGGATATGG
ATACAACAGAGAAAATAGAGGCCTCTAAACAGAAGAAGGAGCAAGGAAATTCTCTGTTCA
AGGTTGGAAAGTACCAGCGAGCTGCAAAGAAATATGAGAAGGCTGCAAAATATATTGAG
CACGACAGTTCCTTCAGTGCCGAGGAGAAAAAGCAATCAAAGGTTTTGAAAGTATCGTGT
AACCTGAATCATGCAGCATGCCGTCTTAAGTTGAAAGACTTCAAGGAAGCAGTTAAACTA
TGTTCAAAGGTGTTAGAGCTTGAATCACAAAATGTGAAGGCATTGTATAGAAGAGCACAA
GCATACATAGAGACAGCAGATTTGGATTTAGCTGAATTTGATATAAAGAAAGCTCTAGAG
ATAGAGCCACAAAATAGGGAAGTGCAACTGGAGTATAAAATTTTAAAGCAAAAGCAAATT
GAATATAATAAGAAAGATGCAAAGCTATATGGAAACATGTTCGCAAAGCTGAACAAATTA
GAAGCCTTTGAAGGAAAGGTATTGTCTTGAGAGCCATCGAACCTAACAGAGCAAGGGAC
AATTACTAAACGAAAGGCTTTTTTTGTGTATTTACTATTTTTATAGTTTGTAAACTGAATCA
CTGAAGCAATGGTTCGAACATAACTGTTCAATTTTTGCCAACTCTTGGGCAGATACTCTTT
TTAGGCCTTACTTTCACACCTAGTTTTGCATGGTTTATGGAAACTTATATTGCTGTTTTTCT
TTGCAACTAGGATTTTTTTTTTCTTTTTAATGAAAGCTGTTCCATGTTACAAGTTGCCAGC
GTGTCTACTGTCGTTGGCTTGCAAAGACATCTTCTTATCTTATTCCGTGTCTCCATAATCA
CCGGCTTCTTTAGTTCACCTTTGCAATTATAAGCAAGAAGCTTGACTTTGGTGAAAGTCTT
GGAAATTTTGTAGCTTCGTTGCTTCTGTATTAAATTCCAAGCAGAAGGGCACATGTTGTG
ACATCAAGTAGTAGATTGTTCTGCAGATTCTGGTTTCTATTCTTTTCTTTCGTTATACTATA
TGACAATCGAATGGACAAAGGAATTTCATCGAAAAAAAAAA

SEQ ID NO: 163
ATTTTCCATTCACTCTGAATATCAATCTCCCCCAACAAAAGCCCATGGAGATTTCACAGGT
CACCTTTACCTAGGGAAGTTCCTTCAGCTTTGGCAGTTTAACTTAAAACTACAAGGCTCC
ATGATATAACAACGTGGCAATCCTAACATGTAATTTCATTGCCAAAATCCGACGTGGATC
AACTGAACTGGACAGCCCCTGCTGTAAATAAAGCTCTCGGAGTCTTGAATCGTTTACAAA
TATTGGCAGACGAATGGTCACACCAGGCAGTTTATCGGCCTTTTACGGTCTCAATTTGTA
GGCGTTGGGCGTTGCAATACGTGTATTCTCAGTAGTCTCATTTGAAGGAAAAATCAATGG
CAGACGAAGGTCTTGAACTCTCCGATGTAGCAGAAGTCGAAGATGAACCAGGTGAGGAG
TTCGAATCCGCTCCACCCCTCGTTGTAGGGCAGGAGAAGGAGTTAAACAGCTCGGGTTT
GAAGAAAAAACTTCTGAAGGCTGGCACAAGGTGTGAAACGCCAGAAAATGGAGACGAAG
TAACAGTGCACTACACTGGGACTTTACTGGATGGTACCAAGTTCGATTCTTCGAGAGACA
GAGGAGAGCCCTTTACTTTCAACATTGGCCAAGGCCAAGTTATCAAAGGATGGGACCAA
GGCATTGTTACCATGAAGAAAAGAGAACATGCACTATTCACGATACCTCCAGAATTGGCC
TATGGCGCTTCTGGCATGCCACCTACTATACCTCCCAATGCCACCCTCCAATTTGATGTG
GAATTACTTTCTTGGACTAATATTGTTGATGTATGTAAGGATGGTGGCATTCTCAAGCGAA
TAATATCTGATGGAGAAAAATATGAAAGACCCAAAGACCCAGATGAGGTGACTGTGAAAT
ATGAAGCAAAGCTTGAGGATGGAATGCTCGTGGCAAAATCCCCAGAAGAAGGCGTGGA
GTTTTATGTCAATGATGGAAACTTTTGTCCTGCCATTGTGAAGGCAGTCAAAACAATGAA
GAAAGGAGAAAATGTGACTCTTACCATAAAACCTGCGTATGCCTTTGGTGAGCAGGGTA
AGGATGCTGAAGAGGGGTTTGCTGCAATTCCTCCAAACGCTACTATCACTATAAATCTGC
AATTAGTTTCATTTAAAGCTGTTAAAGAAGTAACAGAAGATAAGAAGGTTATCAAGAAAAT
CCTGAAGGAAGCTGATGGATATGATAAACCAAGTGATGGAACGGTAGTTCAGATCAGGT
ATACCGCTAAGCTGCAAGATGGCACAATCTTTGAGAAGAAGGGATATGCAGGCGAGGAG
CCTTTTCAATTTGTAGTAGATGAAGAACAAGTGATTGCTGGTCTTGACAAGGCAGTGGAA
ACCATGAAGACAGGGGAGGTTGCTCTGATTACAATAGGACCTGAATATGGTTTTGGAAAT
ATTGAAACCCAGAGAGATCTGGCTGTAATTCCACCATATTCAACTCTCATCTATGAGGTG
GAAATGGTATCATTTACCAAGGAAAAAGAATCATGGGATATGAATACAACAGAAAACATT
GAGGCCTCTAAACAGAAGAAAGAGCAAGGCAATTCTCTGTTTAAGGTTGGAAAGTACCT

GCGAGCTGCAAAGAAATATGATAAGGCAGCAAAATATATCGAGCACGACAATTCCTTCAG
TGCTGAGGAGAAAAAACAATCAAAGGTTTTGAAAGTATCATGTAACCTGAACCATGCAGC
ATGCTGTCTTAAACTGAAAGACTTCAAGAAAGCAGTTAAACTTTGTTCAAAGGTGTTAGA
GCTTGAATCACAAAATGTGAAGGCCTTGTATAGAAGAGCACAGGCATATATAGAGACAG
CAGATTTGGATTTAGCTGAATTTGATATCAAGAAAGCTCTAGAGATAGAGCCACAAAACA
GGGAGGTGCGACTGGAGTATCTTATTTTAAAGCAAAAGCAAATTGAATATAATAAGAAAG
ATGCAAAGCTATATGGAAACATGTTTGCAAGGCAGAACAAATTGGAAGCCATTGAGGGA
AAGGATTAGATTATCCACTCAGATAGTTGTATATCAACTAGAAACTACTGTTGGGAATGG
CCTGAGGAGACTGATTAAATGTTCCTAGCACGCAAAAAGCTGCATTTAGTTGATTAGCTT
CCTGTTTTGACTGACATTTGCGGAAGGGTGAAGGTGAACCGATGGAGGTGGAGAAGCC
AACAAAGAGTTGATGAAAACAGAAAGCCAGAGGCAACGGGTGTAGATTCTAATTAAATAG
CAGCATGTTAATGTAATACATTTAGTTTTTAGATAACTGTTAATGTGTAGTAAAGCACTAG
GAAGAGAAAGATTTGGACATGTCATTGTATTTTGACACTTGAACGGAAGCATGTTAAAGA
ACCTGTGTTATTGTTTTAAAAAAAAAA

SEQ ID NO: 164
GTCTTTTCTTGGCTGTATTTTCTGTCTTTGAAGCGGTATTTTGTACGCGAAGCTTGTTTTT
GTTGCTGGTTTTGGTGGCTTTAAGCGGGTTTTGTATGCTAAGCTTGTTTTTATTGGCGGT
TTTGGTAGTCTTTAAAGCGGTTTTGTCCAATCGCTTCTGTGCAGGTTGATCAATTTGGCG
GAAAGTCAGGCTTTTGGTGCAGCTTTTTGCAGAATCTCAGAGGTTTCGAACTCAGGATGC
CGAACCCAAAGGTTTTCTTTGATATGCAGGTCGGCGGTGCCCCAGCCGGCCGGATCGT
GATGGAGCTCTATGCGGATGTGGTGCCAAAGACGGCTGAGAACTTCCGCGCGTTGTGT
ACCGGCGAGAAAGGCACCGGCCGCTCAGGCAAGCCTCTGCACTTCAAGGGTTCGTCGT
TCCACCGTGTGATCCCAGGGTTCATGTGCCAGGGCGGTGACTTCACAAGGGGCAACGG
TACCGGTGGAGAGTCGATCTACGGCGAGAAGTTTGCCGATGAGAACTTTGTAAAGAAGC
ACACGGGGCCTGGCATCCTCTCCATGGCCAACGCCGGCCCTAACACTAACGGCTCCCA
GTTCTTCATCTGTACCGCCCAGACCTCGTGGCTGGATGGTAAGCATGTCGTATTTGGTC
AAGTTGTAGAGGGCTTGGAGGTCGTGCGCGATATCGAGAAGGTTGGATCTGGATCTGG
CAGAACTTCAAAGCCGGTTGTCATTGCCGACTCTGGACAGCTCGCTTGAATTTTTATTAT
TTACCTTCGCCTTTACGCTGCATACGTTAATAGGTTATTATTTCCTTCAACCATTACGCTG
CATAGGTTGTTAGCGTATTGTTTCCCTTTACCATTACGCTGCATGACTCCCTAGGGTTTG
TCAGCATAGGCGTTTTAAGGGTTTTTTGCATCTTTCTACTCAAGATAGTCGCTGCATAAGT
TCCTAGGGTTTGTCAGCAAAGAGGCTTCAAAGGTTTTTGTGTCTTTTCTAGTTATTATAAA
CGCTTCATAGGTTCCTAGGGTTTGTCAGCATATAGATTTCAAGGGTTTTTTTAGATCTTTG
GAGTTGAGATAAACGCTATGGCAATAACCCAGTAATGTTTGTTTTTATCATATGAAATTTT
TACATCTGGAGTTGCATTCGCAGTAAAAAAAAAA

SEQ ID NO: 165
GAAGAAAGTAAGGTTTTGTAGCGAAGAAAGTAAGGTTGATCAATTTGGCGGAAAGTCAG
GCTTTTGGTGCAGCTTTTTGCAGAATCTCAGAGGTAAGCGGTCTTTGAAAAATGAAAATA
AACAAATCGTCTGAGAAAGTAAAACCTTAAAGCCTTTCAAGGAAAAGAACTGAACTTTTTG
GCTTTGAAGAAAAATGATCTTTTTGTCTCTGAACGGCCTGATATGAAAGATCCATGTAAA
GCTCGGCAGGATTTTGTTTGTGCCTTGCAGGTTTCGAACTCAGGATGCCGAACCCAAAG
GTTTTCTTTGATATGCAGGTCGGCGGTGCCCCAGCCGGCCGGATCGTGATGGAGCTCTA
TGCGGATGTGGTGCCAAAGACGGCTGAGAACTTCCGCGCGTTGTGTACCGGCGAGAAA
GGCAACGGCCGCTCAGGCAAGCCTCTGCACTTCAAGGGTTCGTCGTTCCACCGTGTGA
TCCCAGGGTTCATGTGCCAGGGCGGTGACTTCACAAGGGGCAACGGTACCGGTGGAGA
GTCGATCTACGGCGAGAAGTTTGCCGATGAGAACTTTGTAAAGAAGCACACGGGGCCTG
GCATCCTCTCCATGGCCAACGCCGGCCCTAACACTAACGGCTCCCAGTTCTTCATCTGT
ACCGCCCAGACCTCGTGGCTGGATGGTAAGCATGTCGTATTTGGTCAAGTTGTAGAGGG
CTTGGAGGTCGTGCGCGATATCGAGAAGGTTGGATCTGGATCTGGCAGAACTTCAAAGC
CGGTTGTCATTGCCGACTCTGGACAGCTCGCTTGAATTTTTATTATTTACCTTCGCCTTTA
CGCTGCATACGTTAATAGGTTATTATTTCCTTCAACCATTACGCTGCATAGGTTGTTAGCG
TATTGTTTCCCTTTACCATTACGCTGCATGACTCCCTAGGGTTTGTCAGCATAGGCGTTTT
AAGGGTTTTTTGCATCTTTCTACTCAAGATAGTCGCTGCATAAGTTCCTAGGGTTTGTCA
GCAAAGAGGCTTCAAAGGTTTTTGTGTCTTTTCTAGTTATTATAAACGCTTCATAGGTTCC
TAGGGTTTGTCAGCATATAGATTTCAAGGGTTTTTTTAGATCTTTGGAGTTGAGATAAACG
CTATGGCAATAACCCAGTAATGTTTGTTTTTATCATATGAAATTTTTACATCTGGAGTTGC
ATTCGCAGTACTCTCCTGGTCATCGAGAATTTTTTATGTTTTTTTTAATGCGTTTAAAATAT
GCTTAAATGGGTCCGGAAAACAAAAAAAAAA

SEQ ID NO: 166
GAAGAAAGTAAGGTAAGAGGATTTTGTTAGCGAGGGGTATTCTCTGGTGGGCTTGGTAG
TCTTAAAAGCGGTGCTTTTATTCGTCTTTTCTTGGCTGTATTTTCTGTCTTTGAAGCGGTT
GATCAATTTGGCGGAAAGTCAGGCTTTTGGTGCAGCTTTTTGCAAAATCTCAGAGGTTTC
GAACTCAGGATGCCGAACCCAAAGGTTTTCTTTGATATGCAGGTCGGCGGTGCCCCAGC
CGGCCGGATCGTGATGGAGCTCTATGCGGATGTGGTGCCAAAGACGGCTGAGAACTTC
CGCGCGTTGTGTACCGGCGAGAAAGGCACCGGCCGCTCAGGCAAGCCTCTGCACTTCA
AGGGTTCGTCGTTCCACCGTGTGATCCCAGGGTTCATGTGCCAGGGCGGTGACTTCACA
AGGGGCAACGGTACCGGTGGAGAGTCGATCTACGGCGAGAAGTTTGCCGATGAGAACT
TTGTAAAGAAGCACACGGGGCCTGGCATCCTCTCCATGGCCAACGCCGGCCCTAACACT
AACGGCTCCCAGTTCTTCATCTGTACCGCCCAGACCTCGTGGCTGGATGGTAAGCATGT
CGTATTTGGTCAAGTTGTAGAGGGCTTGGAGGTCGTGCGCGATATCGAGAAGGTTGGAT
CTGGATCTGGCAGAACTTCAAAGCCGGTTGTCATTGCCGACTCTGGACAGCTCGCTTGA
ATTTTTATTATTTACCTTCGCCTTTACGCTGCATACGTTAATAGGTTATTATTTCCTTCAAC
CATTACGCTGCATAGGTTGTTAGCGTATTGTTTCCCTTTACCATTACGCTGCATGACTCC
CTAGGGTTTGTCAGCATAGGCGTTTTAAGGGTTTTTTGCATCTTTCTACTCAAGATAGTC
GCTGCATAAGTTCCTAGGGTTTGTCAGCAAAGAGGCTTCAAAGGTTTTTGTGTCTTTTCT
AGTTATTATAAACGCTTCATAGGTTCCTAGGGTTTGTCAGCATATAGATTTCAAGGGTTTT
TTTAGATCTTTGGAGTTGAGATAAACGCTATGGCAATAACCCAGTAATGTTTGTTTTTATC
ATATGAAATTTTTACATCTGGAGTTGCATTCGCAGTAAAAAAAAAA

SEQ ID NO: 167
CTCTGGTGGGCTTGGTAGTCTTAAAAGCGGTGCTTTTATTCGTCTTTTCTTGGCTGTATTT
TCTGTCTTTGAAGCGGTATTTTGTTGATCAATTTGGCGGAAAGTCAGGCTTTTGGTGCAG
CTTTTTGCAGAATCTCAGAGGTTTCGAACTCAGGATGCCGAACCCAAAGGTTTTCTTTGA
TATGCAGGTCGGCGGTGCCCCAGCCGGCCGGATCGTGATGGAGCTCTATGCGGATGTG
GTGCCAAAGACGGCTGAGAACTTCCGCGCGTTGTGTACCGGCGAGAAAGGCACCGGCC
GCTCAGGCAAGCCTCTGCACTTCAAGGGTTCGTCGTTCCACCGTGTGATCCCAGGGTTC
ATGTGCCAGGGCGGTGACTTCACAAGGGGCAACGGTACCGGTGGAGAGTCGATCTACG
GCGAGAAGTTTGCCGATGAGAACTTTGTAAAGAAGCACACGGGGCCTGGCATCCTCTCC
ATGGCCAACGCCGGCCCTAACACTAACGGCTCCCAGTTCTTCATCTGTACCGCCCAGAC
CTCGTGGCTGGATGGTAAGCATGTCGTATTTGGTCAAGTTGTAGAGGGCTTGGAGGTCG
TGCGCGATATCGAGAAGGTTGGATCTGGATCTGGCAGAACTTCAAAGCCGGTTGTCATT
GCCGACTCTGGACAGCTCGCTTGAATTTTTATTATTTACCTTCGCCTTTACGCTGCATAC
GTTAATAGGTTATTATTTCCTTCAACCATTACGCTGCATAGGTTGTTAGCGTATTGTTTCC
CTTTACCATTACGCTGCATGACTCCCTAGGGTTTGTCAGCATAGGCGTTTTAAGGGTTTT
TTGCATCTTTCTACTCAAGATAGTCGCTGCATAAGTTCCTAGGGTTTGTCAGCAAAGAGG
CTTCAAAGGTTTTTGTGTCTTTTCTAGTTATTATAAACGCTTCATAGGTTCCTAGGGTTTG
TCAGCATATAGATTTCAAGGGTTTTTTTAGATCTTTGGAGTTGAGATAAACGCTATGGCAA
TAACCCAGTAATGTTTGTTTTTATCATATGAAATTTTTACATCTGGAGTTGCATTCGCAGTA
AAAAAAAAA

SEQ ID NO: 168
GGAATCGCTTGATGCATAATATAAAATCACTCGCTTCCCCCTCTTTCTTTCGCCTAGCCG
TTTCATTGGTTGTTTGCTAAGAATTCAGAGTGTCGTACTGCGCATCTGCTTGGCTTTTGAT
TCTGTTTATCTTCGTTAACATTTCAAGTGTGTAATTTTCCTCAGCAACGTTCCAATGGCGG
ACGATTTTGAACTCCCTGAAAGTGCTGGAATGATGGAAAATGAGGACTTTGGCGATACTG
TTTTCAAGGTTGGTGAAGAGAAAGAAATTGGGAAGCAGGGGTTGAAGAAATTGTTGGTTA
AAGAAGGGGGATCGTGGGAAACACCCGAAACCGGTGATGAAGTTGAAGTTCATTACACC
GGAACACTTCTGGACGGTACAAAGTTCGATTCTAGTCGGGATAGAGGAACTCCATTCAA
ATTCAAGCTTGGTCAAGGTCAAGTAATCAAAGGATGGGATCAGGGGATTGCAACGATGA
AAAAGGGGGAAAATGCAGTCTTCACCATTCCTCCTGATCTGGCATATGGGGAATCTGGA
TCGCAACCTACAATTCCACCCAATGCGACTCTGAAATTCGATGTTGAATTGCTCTCTTGG
GCTAGTGTAAAAGATATCTGCAAGGATGGGGGCATTTTTAAGAAAATTATCAAGGAAGGG
GAGAAATGGGAGCACCCAAAAGAAGCTGATGAAGTTTTGGTTAAATATGAGGCTAGACT
GGAGGATGGAACTGTTGTATCAAAATCAGAGGAGGGTGTGGAGTTTTATGTAAAGGATG
GATATTTCTGTCCAGCTTTTGCCATAGCTGTCAAAACAATGAAGAAGGGAGAAAAAGTAT
TGTTAACAGTGAAGCCTCAGTATGGTTTTGGACACCAAGGACGCGAAGCAATTGGAAAT
GATGTTGCCCGTTCCACAAATGCAACATTGTTGGTGGATCTTGAGCTTGTGTCTTGGAAG

GTTGTCGATGAAGTTACTGATGATAAGAAAGTGCTGAAGAAAATCCTGAAGCAAGGAGAA
GGTTATGAACGACCTAATGATGGGGCAGTGGTCAAAGTGAAATACACGGGGAAGTTAGA
GGATGGCACCATATTTGAGGAAAAGGGATCTGATGAAGAACCATTCGAGTTTATGGCTG
GAGAAGAGCAAGTTGTTGATGGATTAGATAGAGCGGTTATGACAATGAAGAAAGGAGAA
GTTGCCTTGGTGAGTGTGGCAGCTGAATATGGTTACCAAACTGAGATTAAGACAGACTT
GGCAGTTGTTCCACCAAAGTCCACCTTGATTTATGAAGTGGAGCTGGTTTCATTTGTAAA
GGAGAAAGAGTCATGGGACATGAACACCGCGGAGAAGATAGAGGCAGCTGGGAAGAAG
AAAGAAGAAGGGAATGCATTATTTAAGGTCGGCAAGTATTTTAGGGCATCCAAGAAATAT
GAAAAGGCTACAAAGTATATTGAGTATGACACTTCATTCAGCGAGGAAGAGAAGAAGCA
GTCGAAGCCCTTGAAGGTTACTTGTAATCTGAACAATGCAGCTTGTAAACTCAAATTGAA
AGATTACACACAAGCAGAAAAATTATGTACAAAGGTTTTGGAGGTTGAATCGCAAAACGT
TAAGGCTTTATATAGAAGAGCACAAGCATACATTCAGACAGCAGATTTGGAGCTCGCAGA
GCTGGACATAAAGAAAGCACTCGAGATTGACCCCAACAACAGGGATGTGAAGCTCGAGT
ACAGGGCTCTTAAAGAAAAACAAAAGGAGTACAATAAGAAGGAAGCCAAATTTTATGGCA
ACATGTTTGCACGAATGAGCAAGTTGGAGGAATTAGAAAGCAGGAAATCTGGGAGCCAA
AAAGTGGAGACTGCTAATAAAGAGGAAGGGTCTGATGCCATGGCTGTAGATGGCGAGTC
TGCCTGAAGTGCTTTTGCAAGTTTAGCGATATATTTCTATGCAGTCCTTTATAGATGAGTC
TACCTTTTAGCCATTCACAATGAAGATTGTAAGTTGGGTGAACTTTTTTACCACGCTAGGT
TGATCTATTTTAAGACTCTTTTGCAAGATGCTTCTGCATTCATAGCAAAATTGTATGAGGG
ACTGGCAAGCTGAATTTGCTTGAAAGTGATGAAAGAGGATGTATTTTTGAAGTATTGCCT
TGCTCTTAGTCTAATTGACTTAGATAATGAATCTCTTGGTTTTTGTTGAAAAAAAAAA

SEQ ID NO: 169
AACAAAATAGCTGCGCGTACCACAAAGGTGACAAACGCCGGATTTCTCTTATCAGACTTG
TCAATGGCCGCCTCTCTCACTCCACTGGGTGCAGGCCTGGCTTATGCAACAATTTATGAT
CAAGCTAAAGTGAGGAAATTGGAACCCACAAAACGGTCCTTGATAGCGCTGTGCCAACA
TTCCGATTCTCAACATAGAAGATTTATTACTAGAAAATATCATGTGAACGTTCAAATTCTTA
ATCGAAGAGATGCAATCAGATTAATTGGTTTAGCAGCTGGACTATGTATTGATCTTTCTCT
GATGTATGATGCTAGAGGAGCTGGTTTACCTCCACAAGAAAATGCAAAGTTGTGTGATAC
AACGTGTGAGAAGGAACTTGAAAATGCTCCAATGATCACAACTGAATCTGGTTTACAATA
CAAGGATATCAAGATTGGGAATGGGCCAAGTCCACCTATTGGATTTCAGGTGGCAGCAA
ACTATGTAGCAATGGTACCATCTGGACAAGTTTTTGATAGTTCATTGGATAAGGGTCAAC
CTTACATCTTTCGAGTAGGTTCTGGCCAGGTGATCAAGGGTCTTGATGAAGGCCTTCTGA
GCATGAAAGTTGGTGGAAAACGTCGCTTGTACATCCCTGGCCCTCTTGCATTCCCAAAG
GGCCTTAACTCTGCTCCAGGAAGGCCGCGGGTAGCACCTAGCAGTCCTGTTATTTTTGA
CGTGAGTTTGGAGTTCATACCTGGCCTTGAAAGTGAAGAAGAGTAAACTGCTTGCATGA
CAGATTAGCAAGTCAAGTCATTGCGTAGAGTGATGGTCACCATAGACTCTTTAGTTTCAT
TCATTTGATTGATTATTCTTTTTATAAAATTTGTTTCCGTTTTCCTTGTTTGGAAGAAGGCT
TACTGTTTTGAACTTGAAAATATGTTGAAATAGGTTAAATGCACTGTTAATTCTGGACGGA
TGGAAGGATTTCCTTTCCAATCATTTCCTTGCAACAAAGAGCTGACCACGGCAAGCTTAA
GCATCTGATGAATGATGCCACATGTATATCTTTTCTTAATATAATATATATCAAGCTTTTTT
CGTAAAAAAAAAA

SEQ ID NO: 170
GTGTGAATCCATTTAAGTCTATCCAAAGGCCAAACAAATCTGAAAGAATTTTGTTTAACAC
ATTATCCATTTACGGGACAGGGCAAGCAGCAATGTCGGCCGCATCACTTTCTGCAGACA
TGGCGATCCGTGGCACTATTCTGGGGAAGACAGCATTACATGTTTTGGGACCTCAGGTT
GTATCCCAATGTCGCCAACCAGTGATGTTCAAATGTCCACCCCATACTCTCAGGAAGATG
AGATTTTCTGCTCAAGATTTGCAGTCAAAAAATTTCTATTCAGGATTTACACCATTTAAATC
AGTTTTTATTTCTACTTCAAAAAGGAGTTGGCAAGCAGGTTCTGCCAGAGCTATGTCACA
GGATGCTGCATTTCAGTCAAAGGTGACAACTAAGTGCTTTTTGGACATTGAAATTGGAGG
AGATCCTGCAGGAAGGATTGTGCTTGGACTTTTTGGTGAGGATGTACCAAAAACAGCTG
AGAATTTCCGCGCACTATGTACAGGAGAGAAAGGGTTTGGATACAAAGGATCTTCATTTC
ACCGAATCATAAAAGATTTTATGCTTCAGGGTGGAGACTTTGACAGGGGAGATGGTACA
GGAGGAAAGAGCATATATGGTCGCACTTTTGAAGACGAGAACTTCAAATTAGCCCATGTT
GGACCTGGAGTTCTAAGTATGGCTAATGCTGGCCCCAACACAAATGGCAGCCAATTTTTT
ATATGTACAGTCAAAACCCCATGGTTGGATAAACGCCACGTTGTGTTTGGACAAGTTATT
GAGGGGATGGAGATTGTTAAGAAACTCGAGTCGGAAGAAACAAATCGTACAGATCGGCC
AAAAAGGCCCTGCAGAATTGTTGATTGTGGGGAGCTTCCCTGAAGTTGCTGAAACAGAG
GCTTTAATTGTTTCAATTGAGTGATCATCTGAAAATAATTTTGTTTCGTCTGAAAAGAGGC

TTTAATTATCATAGTTTTATTCCGGCTATCTTGATCATTCACGGAAGTCCCGAGAGTCAAT
TTGCCAAGTATTCCTATTAAACTCTATTGCTCTCAGAATTGATTTAAATGCCTTTAATGGG
TGCAATGTAATAGTCTTATTATGAGTCAATATGAACTGTTAGGGTAATGCCAAAAAAAAAA
AAAAAA

SEQ ID NO: 171
GCTAGATTTCGTCGAGGGAAGTCGAGTCGGAGACAAATTCGGGTATAATATGGAGGATG
GGTGAGTCAGAATTGGGTAAAAATAGAAGGGAAGACAGAGGATCACCCCAAAACATATT
TGTCTGATTGATCGATTGATTTCTAAAAGGTCCAACGTTGTTTTCTGCTCCTCCACAACCA
AATAGTTCATCTCCCCATCTCCTGCTTTTCAGATTCACTATACATTCAAAATTCGAAATTC
GATATTCGGGACTGTGTAATCGAAGGTCTATCTATTTCAAGGTCCCCAGATCGATCTTAC
TCCGAAGGAGAATCGTTTCGATCATCTTTGCTGAAAGTGAGAGCAGAAACCTTTGAGACA
AGTCAGGGGCAATGGGGAGAATTAAGCCACAGACATTGTTGCAGCAGAGCAAAAAGAAA
AAAGTACCCGGCCGTATCAGCGTTTCCACCATTATAGTGTGCAACTTGATCATCATCTTT
CTTATGTTTTCTCTAGTGGGCATCTACAGGCAGAGGGCCAAGCGTAATCGGGCAACATC
TCGGTCTGATGGTGATGAGGAAATGGAGAACTTTGGGAGGTCCAAAATAAACAGCGTTC
CTCACCAAGCAATCGTAAATACTACAAAGGGTTTAATCACATTGGAACTTTTTGGTAAAAG
TTCTGCGCATACCGTTGAAAAGTTTGTGGAGTGGAGTGAACGTGGTTACTTCAATGGATT
ACCCTTCTATCGTGTCATTAAACACTTTGTAATTCAAGTTGGAGATCCAAAGTTTGCTGGA
AACAGGGAAGACTGGACTGTTGGTGGTCAGCTCAATGTTCAACTTGAGTTCAGTCCAAA
GCATGAAGCATTCATGCTGGGGACCTCCAAGTTGGAGGATCAGGGAGATGGATTTGAGC
TTTTCATAACAACAGCACCCATTCCGGATTTAAACGATAAGCTTAATGTCTTTGGGCGTGT
TATAAAAGGTCAAGATGTAGTTCAGGAAATAGAGGAGGTGGATACAGATGAACATTTTCA
GCCGAAATCCCCCATCATCATTAATGATGTCCGTTTAAAGGATGAGCTCTGAGACATACG
TGCAGTCCTGTTTCAAGAAATTGTGATCGATAATCAGTTTTTCTTCCAAGTATTGCTTTTC
GCCTGTTCTGTCTCGAGCATTTTATATACCCGGAGGAAACAGCGCTGTGGATACCAGGA
AGATTGAAGTGTATGTATATTTAGGCAAATACTTCGTAAGTTTTGTGATGTCTCAGTTGTT
GGCTTTTACCTGTAGAGAAGTTGAGTTCTAAGCTGGAGATTTATTTCTTTACCACCTGAAT
ATTCTCATAGTTCTAGGCCAATTATAATTCTTTAGTCCCTTCTTATGTTGGGGAATGAAAT
TTACTGTTATTACTTTACACTGAATGATCATTAATGTGCCTCAGGATTTTAGCATCTCAGTT
TGCCGCTGAACATTTGACAAGAGATTTAATGATTAAGATTTCTTTTGACATTGGGAGAGAT
TTTTTCTAACTTGCGTTGAGACAAAATGACAATGAACTGCAATCCACTTTCGACAAAAAGG
AAGAGATCAAGAATCAAGTTTGATGGGCTGACCCCATGTTAACTTATTAGTTATTATAGTG
GAGGTTCCCATTTCCAAGTCAAACAGTTTGGATACCCACTTAATGATACCAGTATAACTAT
TAATTTTTGTAAAGCCACCAAGCTTGTATAAATGATGCCTTTCTTTTCAAAAAAAAAA

SEQ ID NO: 172
CTTTCAAAGGGACGGGTTGCGAGAAATATGGCGAGGCAAAGTACTCTACTGCTCTTTTG
GAGCTTAGTATTCTTGGGTGCCATCGTATTTACTCAGGCCAAGCATGAGGAACTGGAAG
AGGTAACACACAAGGTATACTTTGATGTGGATATTGCTGGAAAACCTGCTGGTCGAGTTG
TTATTGGACTCTTTGGGAAAGCTGTACCCAAAACTGTAGAGAATTTTCGGGCGCTTTGTA
CAGGTGAGAAGGGCGTTGGAAAAAGTGGAAAGCCACTTCATTACAAGGGAAGTTTTTTC
CATAGGATCATTCCCAGCTTCATGATTCAAGGTGGTGATTTCACTCTTGGAGATGGCAGA
GGGGGAGAGTCAATCTATGGAACTAAATTTGCCGATGAGAACTTCAAACTGAAGCACAC
AGGACCAGTTTTTATTACCACAGTAACAACCGACTGGTTGGACGGCAGACATGTTGTCTT
TGGAAAAATTATTTCTGGAATGGATGTTGTATATAAAGTAGAAGCAGAAGGAAGACAGAG
TGGTCAACCAAAAAGAAAAGTAAAGATTGCAGACAGTGGAGAGCTTTCTATGGATTAGCT
AACTCTTCTAGTTGAGATCTCCATCAATTAATGGATACAAACATTGAGTTTCACTTTGGCA
AACCGACATTCCCAAATTTAATGAGCCAATATGTTTGGCTCGTCATAAAGCCTGGGATAA
TCTTAAATAATAAATATCGAGTTGGTGACTCAAAAAAAA

SEQ ID NO: 173
GAACGACTTTCAAAGGGACGGGTTGCGAGAAATATGGCGAGGCAAAGTACTCTACTGCT
CTTTTGGAGCTTAGTATTCTTGGGTGCCATCGTATTTACTCAGGCCAAGCATGAGGAACT
GGAAGAGGTAACACACAAGGTATACTTTGATGTGGATATTGCTGGAAAACCTGCTGGTC
GAGTTGTTATTGGACTCTTTGGGAAAGCTGTACCCAAAACTGTAGAGAATTTTCGGGCGC
TTTGTACAGGTGAGAAGGGCGTTGGAAAAAGTGGAAAGCCACTTCATTACAAGGGAAGT
TTTTTCCATAGGATCATTCCCAGCTTCATGATTCAAGGTGGTGATTTCACTCTTGGAGATG
GCAGAGGGGGAGAGTCAATCTATGGAACTAAATTTGCCGATGAGAACTTCAAACTGAAG
CACACAGGACCAGGATTCCTCTCCATGGCAAATGCCGGTCCTGACACGAATGGCTCTCA

GTTTTTTATTACCACAGTAACAACCGACTGGTTGGACGGCAGACATGTTGTCTTTGGAAA
AATTATTTCTGGAATGGATGTTGTATATAAAGTAGAAGCAGAAGGAAGACAGAGTGGTCA
ACCAAAAAGAAAAGTAAAGATTGCAGACAGTGGAGAGCTTTCTATGGATTAGCTAACTCT
TCTAGTTGAGATCTCCATCAATTAATGGATACAAACATTGAGTTTCACTTTGGCAAACCGA
CATTCCCAAATTTAATGAGCCAATATGTTTGGCTCGTCATAAAGCCTGGGATAATCTTAAA
TAATAAATATCGAGTTGGTGACTCAAAAAAAAAAA

SEQ ID NO: 174
GTGGCCAATGAGATTACTGAAGCTGAAGACGGATATACAACCCGGACCACATTCCCAAT
TGAATTTGTAGCAGCTCTTGTGCATGCCCACGTAGTTCTGACATCAAATTTAGGACGCAA
CATAATGTATTCAACTTCGCATTTCTTCTTCTAACGACCAGCGGTTTTCTCCCTTGAAGTG
AGGCTTCGTGTCATTTAATAACTAGGTAGGGTTTTCAGAAGTGGGGTTCATCGCATTTGC
CAGTACTTGCAATGGTACAGCGGAGAGAAGGGGGTGCGAGGTACAATATTGAAAATGGC
GGATCGTGATTATGCGAGGACAATATTGGTCAAATTGAATTGACTGGATCTCATGGTTAG
AGAGCTTAGTGGGATCGAGACGCTACAGAGACAATATACAGGAGGGGTCTCTCGTTCAT
CACTCTAGACAGTGTGTAACTGCACTTCAACAAATATATTTTAGAAAGAATCTCTGCTATT
TATGGAGATGGACGAGATTCAGGAACAGTCCCAACCCCAATCCAGTGAGAAGCAGGATA
TTTCTCAAGAATCTGACACAGGCAATGATAAAACAATTAATGCTGAAAAGATCACATCCG
AGAATGCTGAAGTTGAGGAAGATGATATGCTTCCTCCAAAAGTTAATACTGAAGTGGAAG
TTCTTCATGATAAAGTTACAAAACAAATCATCAAGGAAGGGAGTGGAAACAAACCTTCAC
GGAATTCGACATGCTTCTTGCACTACAGAGCGTGGGCTGAAAGCACTATGCACAAGTTT
CAGGACACCTGGCAAGAACAACAGCCACTCGAACTGGTTTTGGGCAGAGAGAAAAAAGA
ATTGAGTGGCTTTGCCATTGGTGTGGCTGGCATGAAAGCTGGTGAACGTGCCTTGCTTC
ATGTGGACTGGCAATTAGGTTATGGGGAAGAAGGGAACTTCTCTTTTCCAAATGTACCAC
CTAGAGCCAATCTTATATATGAAGCTGAGCTTATTGGTTTTGAGGAAGCCAAGGAAGGTA
AGGCACGTAGTGATATGACAGTAGAGGAGCGAATTGAAGCAGCAGATAGGAGACGGCA
GCAAGGTAATGAACTTTTCAAGGAGGACAAATTAGCAGAGGCTATGCAACAGTATGAAAT
GGCCCTAGCATACATGGGAGATGATTTTATGTTTCAGTTGTTTGGAAAATACAAGGATAT
GGCAAATGCTGTGAAGAACCCTTGCCATCTTAATATGGCTCAATGTTTGTTGAAGCTGAA
CCGTTATGAAGAAGCCATTGGCCAATGTAACATGGTATTAGCTGAGGATGAGAAGAATAT
CAAGGCTTTATTCAGACGTGGTAAGGCAAGAGCTACTCTAGGGCAGACTGATGATGCAA
GAGAGGACTTTCAGAAGGTTCGGAAATTTTCTCCTGAAGATAAGGCAGTGATACGAGAG
TTGCGTCTCCTTGCTGAACATGATAAGCAGGTTTATCAGAAACAGAAGGAGATGTTCAAG
GGTCTTTTTGGGCAAAAACCAGAACAGAAACCAAAGAAGTTACACTGGTTTGTTGTGTTT
TGGCAGTGGCTTCTGTCGATGATAAGAACTATTTTTAGGATGAGATCCAAAACTGACTAA
TTAATTATGGTTCAATTTGCCGATCTTGTTAATAATAAATATAACTTAATTTGTCTGTTGCC
AATAAGATCTATGCAACAATGCTATTCGACTTGGAAGTTGTGAAGTGGATCACTACTGGA
TTCCGTTACATTAGTTATTGCAAGTTGGTTATTATGTACGTTTATATCACAGAACTTTGACA
ATTGTCTGTGAACTACAGTATTAAATAGTTTTGGTGCTTGTTACTCTTAAAAAAAAAA

SEQ ID NO: 175
CTTTTTTCATTAATTTGGTTATGAGTTCGCCATCTTGTATACTATAACCGGGCAATGATGT
TTGTCACTCAGTGGTGAGGAAAGGGAGAACAATGGCAGGAGCAGGAGAAGGAACTCCC
GAGGTCACCCTCGAGACCTCCATGGGTCCTATCACAGTGGAGCTCTACCACAAGCACGC
CCCCAAGACCTGCAGGAATTTCTTGGAGCTCTCAAGAAGGGGCTATTACAACAATGTCA
AATTCCACCGCGTTATCAAGGATTTCATGGTGCAAGGCGGAGACCCAACTGGCACAGGA
AGAGGGGGGGAGTCAATTTATGGCCCAAGATTTGAAGATGAAATTACCCGAGATTTGAA
ACATACTGGAGCGGGAATATTGTCCATGGCTAATGCAGGTCCTAACACAAATGGCAGTC
AATTCTTCATTTCACTGGCACCGACACCATGGCTTGATGAAAAACACACTATTTTTGGAA
GGGTATGCAAAGGGATGGATGTTGTCAAACGGCTTGGAAATGTTCAGACTGATAAGAAT
GATCGACCTATTCATGATGTGAAAATCTTACGAACAACAGTCAAAGACTGATAAATTCCAT
ATCAAGAAGGTCAACAGCTCCTATCAGACTACGTTGGATCATCTGTTGCAGGTTACTGCA
ATTTTTAGACTTCTTTTTATTATGCTTCATCCTTTAATGCATCCTTCAGTGAAGCAACTACT
ATTGATGTTTTGGCTCTAAAAATATAGACTTCGTTTATTTAAAATTTTCCCAACATAAAATA
TGAAAGACCTAAAACAGAATCTCGTGAAATTTGCATTAGAAGATGTTACCTATATTTTGTA
TATGGTTTCAGTGGAGTTTCCTGATGTGCAGATTACACCATGAACAAATGCAATTACCCT
GTTTTATTCTATCCCGCTTTAATTAATATTGGTCATGTTTTTCTTATTCTGGCTTGGGAATT
TGGTTGGCAGATTTACAAGGTATTGTCATT

SEQ ID NO: 176

GCTCTATTATAGACCAGCTTCATCATAAGGTAAAGTTTGGCGAGCCGCGTGGACTTATTC
TTTGACCTTTTTGGCCAATAAATTTCAATCCCCGAACCATTCGGTCATCGAAATTCTCCCC
GAGATCGCTCTGTAATTCCATGCCATTTGAGCCCTTAATTTGAATTGTTTCTTTCAAGCAC
TGTATTTTAATTTTTCCCGGTGAAGTGTTAGTGTTGGTGGATTTTCCGCTGACAATCGAAT
CAGGGAATGATGGATCCGGAGTTGATGAGGCTGGCACAAGAGCAGATGAGCAAGATTT
CCCCCGATGAACTCATGAAGATGCAACGACAGATAATGGCCAATCCAGACTTAATGCGT
ATGGCATCAGAGAATATGAAGAATTTGAAGCCTGAAGATATAAGGTTTGCAGCAGAGCA
GATGAAAAATGTACGCAAGGAAGAAATGGCTGAAATAAGTGAGCGCATATCAAGGGCTT
CACCAGAAGAGATTGAAGCAATGAAAGCTCGTGCAAATCTTCAGTCTGCATATCAATTAC
AAGTAGCTCAGAACCTAAAGGATCAGGGAAATCAGCTTCATGCTCGAATGAAATACAGTG
AAGCAGCAGAGAAATATTTGCAAGCAAGGAACAATCTGACAGGAATACCCTTCTCCGAA
GCTAAGAGTTTGTTATTAGCTTCTTCCTCGAATTTAATGTCCTGCTACCTAAAGACTGGGC
AGTACGAGGAGTGTGTACAAACAGGTTCAGAGGTTCTGGCATATGATGCAATGAATGTTA
AAGCATTGTATCGCAGAGGCCAAGCCTACAAACAAATTGGAAAGCTAGAATTGGCTGTT
GCAGACTTAAGAAAAGCGGTGGAGGTTTCTCCTGAGGATGAAACAATAGCACAGGCTTT
AAGAGAAGCAAGCACAGAATTAATGGAGAAAGGGGGCACTCAAGATCAAAATGGGCCTC
GTATTGAAGAGATAATTGAGGAAGAAGCTGTTCAGCCAACTGCTGAAAAGTATCCGCAGT
CGGCTCCTATGGTGACTTCTGTTACAGAAGATGTAAGTGATGATGAGCAGGGGTCAGAA
GATCAAAATGGGTTCTCTAGAGATAGCTTTCAAGCAACTAATGCGCCCGATGGGCAGAT
GTATGCTGAGAGCTTGAGGAACTTGACTGAAAATCCTGACATGTTAAGAACTATGCAGAG
TTTGATGAAAAATGTCGACCCAGACTCCTTGGTTGCACTAAGTGGAGGTAAATTGAGCCC
TGATATGGTCAAAACTGTATCTGGTATGTTTGGTAGGATGTCACCAGAGGAGATCCAGAA
TATGATGAAAATGTCGAGCACCTTATCAAGACAAAATCCATCTACTTCATCTAGATTTGAT
GACATCACACGTGGACATTCAAACATGGATTCATCTCCACAATCTGTTTCAGTAGATAAT
GACCTTTTTGAAGAAAATCAGAATAGAGTTGGGGAGTCATCTACAAATTTGAGTTCCAGT
GCAGCCTTCTCAGGCATGCCAAACTTTTCTGCAGAAATGCAGGAGCAAGTGAGAAATCA
AATGAATGATCCAGCCACCAGACAGATGTTTACGTCAATGATACAAAATATGAGCCCTGA
AATGATGGCAAGTATGAGCGAGCAGTTTGGGGTTAAGCTTTCACCTGAGGATGCAGTTA
AGGCACAAAATGCTATGGCCTCTCTTTCTCCCAATGACTTAGACAGATTGATGAATTGGG
CAACAAGGCTTCAGACTGCAATTGACTATGCTCGAAAGATCAAGAACTGGATACTAGGAA
GGCCAGGTCTTATTTTTGCAATTTCCATGCTTCTTCTTGCTATTATTCTTCATCGGTTTGG
ATATATTGGAGACTAAGAAACGATACATCATTTGGATTTTCAAATATTGATGTAGTTGATG
CGTGCTTAAACCAATTGAGGAAGCTACTTCAGGCATGCATTTTGATCCTTTATGACCTTAT
TTCCAGCAATGCCAGCCGTATAGTTAACTGAGAAATGCTGGGAACTCTGCTGGAATTATT
TGTTAACAAGGAAGCAGTTGACTTTTGCTTGTGGATTGTACTGTGGTACATGGTATAAAT
CTATAGGCTATGTCGATTATTTTTGGCATCCTCATGCCTTTTATTAAAAACACACGGATGC
TTGAGAATGATGTTTTCCTTTGTGTTAGGTCAGTTTTTGTCCAAATCATGTGGCACCTTTG
GTTAATAACATTATCCACATCAATTTGTACATGGGATCTTCTTACGTTTATTTTTATTTTTTT
AAAAAAAAAAAAAAA

SEQ ID NO: 177
ACGGACGACTTCCATTTTTTCCACGTGAATCCAATTGTTCGCTAGCTTTTAAAGGAATTG
GTTTTACGCACATAAAAAGAAGAATGGGAGTAGAGAAGGAGATTTTGAGGCCTGGTAAT
GGCCCCAAGCCGCGTCCTGGCCAGAGCGTGACGGTCCATTGCACAGGCTACGGTAAGA
ATGAGGATCTGTCTCAGAAATTCTGGAGCACAAAAGATCCAGGACAGAAGCCTTTCACAT
TCACAATTGGGCAAGGGAGAGTTATTAAAGGATGGGATGAAGGCGTTTTGGACATGCAA
CTAGGAGAGATTTTTAAACTTCGGTGTTCACCGGATTATGGATATGGGTCAAATGGATTT
CCAGCATGGGGTATACGGCCGAATTCTGTTTTGGTTTTTGAGATTGAAGTTTTGAGTGTA
AACTAGAGTACTATATCATTATAGAATATATGATATATAAGATATAAGATATTGCCAGCAAA
CTATTTGACAGGTTATTTAATAAAGTGTGCTACATTTTCATTGTATTTTGATTAAGATTTCT
AAGGTATGATGGATGGTGTTTGTAGTAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 178
AATTAACGAAGGAATCAGGCCCTAAGCCTCAAGCTTCGGCCAACCATAGTCATAACAACA
CAACTAGGGTTTTAAGCTTTTGGCATGTTAAGTCTGGTCCACAAGTAAGAGGTTCGAATC
CGCGGAGATGCCGAATCCAAGGTGTTACCTGGACATCACTATTGGGGAGGAGCTGGAG
GGGAGGATCTTGGTGGAGCTTTACAGCGATGTGGTGCCCAAAACGGCCGAAAATTTCAG
GGCTTTGTGCACAGGAGAAAAGGGCATTGGTCCTCATACCGGCGTCCCCCTGCATTACA
AGGGGCTTCCTTTCCATAGAGTTATCAAAGGTTTTATGATACAAGGTGGAGACATCTCAG
CTCAAAATGGTACAGGGGGTGAATCAATATATGGATTAAAGTTTGATGATGAAAACTTTC

AACTGAAACATGAGCGCAGAGGAATGTTGTCAATGGCAAACTCTGGACCTAATACAAATG
GCTCTCAATTTTTCATCACAACAACAAGGACCTCCCATCTGGATGGAAAGCATGTTGTCT
TTGGGAAGGTAATTAAAGGCATGGGAGTTGTGCGCGGCATTGAGCATACTCCCACAGAA
AGTAATGACCGCCCTTCTCTAGATGTTGTAATTTCAGATTGTGGGGAGATCCCAGAAGGA
TCAGATGATGGAATAGCTAATTTCTTCAAAGATGGAGATTTGTATCCAGACTGGCCAGCT
GATCTTGATGAGAAATCTGCAGAAATTTCATGGTGGATGAATGCAGTAGATTCTGCGAAA
TGTTTTGGGAATGAAAATTACAAGAAAGGGGATTATAAGATGGCTTTAAGAAAATACAGA
AAGGCATTGCGCTATCTTGACATTTGCTGGGAGAAAGAAGAAATTGATGAAGAGAAGAG
CAATCATTTAAGGAAGACCAAGTCGCAGATATTTACAAATAGTTCTGCTTGTAAATTAAAA
TTAGGAGATTTAAAAGGTGCCTTATTGGACACAGAGTTTGCAATGCGTGATGGGGAAGAT
AATGTAAAAGCATTATTCCGTCAGGGTCAGGCATACATGGCTCTGAAAGATGTTGATTCT
GCTGTAGCAAGCTTCAAGAAAGCATTGCAGTTAGAACCTAATGATGCCGGGATTAGGAA
AGAGCTTGCAGTTGCCACGAAGATGATTAATGACAGACGTGATCAGGAACGGAGAGCCT
ATGCAAGGATGTTTCAATAGCCTATCTTATCATTTTTGTCCGCATGTAGTTTAGGATATTA
GGAAGAGTTTAGGGAAAGGACAGGCATTTGATATCTCTCATTGGACTATATCGGGATTCC
TTTTTAAAAGTACCCAGTGATATAGGAACTGAGCAAGAGGCAAACAAATTCAAATCAGGG
TTGAGCCACGATTTTACATTTTAACGAATATGATGATAACATGGTTTGTCAAGGTCTTGCA
GCAATGTCTTGCAGCATCGCCAATGTCAAGTGTGAATAATGAAGGAAAATATTTTTTCAAT
TGATTGAACATGGGCAGCATTTTGAAATTAATTCCCTTTTAAATGTGGACAGAGGCACTAT
AAGAATGCGAAATATCGTCGGAGCACGACTAATTGTTTTTTAACTTGCGCTTTTCTTTTTT
AGGCGAATGTGAAAAATTCCTTGCTGGTAGATATTGTTGAGTTATAATGGACAAAAATAG
AGCCAGTGAGCAATGGTGATTTTTCTTAATGAAAAATTGATATTATTGCC

SEQ ID NO: 179
CGAAATAGTGGAATGCGTATTGATATGGCTAAGAGTTCTCTCTGCAATTCTCAGAAACCC
TGGCAGATAATTAAGCTCAACGCTGCTACTGTTAGCGTCAGTTACCTTGGTGGTAGACAA
GAACAAGAAGCGCAAGTAGTATCTAGAGCAACAGAAGCTGTGGATTCACACGGCCGTCA
TTTGACAGTGCCTTCCAGTGTGTTCACCTCTGCATCATCTCAGATTTGAGCCCAGTCTTA
ATTACCTGAATTATAGACAGAAAATATGGGGGATGTAATAGATCTAAATGGAGATGGAGG
TGTGCTGAAAACCATAATAAGGTCAGCTAAGCCAGGTGCTATGCAGCCAACTGAAGATC
TTCCCAATGTCGATGTACATTATGAGGGCACCCTCGCAGATACTGGTGAGGTATTTGATA
CTACCCGTGAAGATAATACTCTCTTCTCATTTGAGTTGGGCAAGGGCACAGTAATCAAGG
CTTGGGATATAGCTGTTAAAACCATGAAGGTTGGTGAGGTTGCAAGAATTACTTGCAAAC
CTGAATATGCTTATGGAAGTGCTGGCTCTCCTCCAGATATTCCAGAAAATGCAACACTCA
TTTTCGAGGTGGAGTTGGTAGCATGCAAACCACGGAAAGGTTCAACTTTTGGCAGTGTTT
CAGATGAGAAGGCCAGACTAGAGGAGCTGAAGAAACAGAGAGAAATTGCTGCAGCTAG
CAAAGAAGAAGAAAAGAAACGCAGGGAGGAAGCCAAAGCAACAGCGGCTGCTCGTGTT
CAAGCAAAATTGGAGGCAAAGAAAGGGCAAGGTAGAGGGAAAGGGAAAAGCAAGGGAA
AATGATAGACTAGTTCTACAAAGCCCTAGGATGATGGACTTCATTTCTTTTGCATTAAGAT
GAAGTTATAAATATAAACTTCAATATATTATTTTTTGGTTTGCTTCCCACTTGATTAAAAAC
CAAAAAGGTCCATCTAAATTTTTGTTTTTTGGATGATATGTTTGTGCAATGTAACTGTTTAC
ACTTCTCGGTTTTTGCTTAAAAAAAAAA

SEQ ID NO: 180
AACACATTGAAAATAATATTTGTATTTTTCCACTGACATGGGACTCGGTTTAAAAATTGCC
TCTGCATCGTTTCTTCCCATATTCAACATCATGGCGACTCGTTCTCTGTGTATCCTGCTG
GTCTGCTTCATCCCTGTTCTCGCACATGTTTTATCGCTTCAAGATCCAGAGCTTGGCACA
GTCCGTGTATATTTTCAGACTACATATGGTGATATTGAATTTGGATTTTTCCCTCACGTGG
CTCCTAAAACTGTGGAACACATTTATAAACTTGTACGGCTTGGCTGCTACAATTCCAATCA
CTTTTTTCGGGTTGACAAGGGCTTTGTTGCTCAAGTGGCTGATGTTGTAGGTGGGAGAG
AAGTTCCTCTAAATTCAGAACAACGGAAGGAAGGGGAGAAAACTATTGTTGGTGAATTTA
GTGAAGTAAAACATGTTCGGGGAATTCTTTCAATGGGAAGATATAGTGATCCAGATAGTG
CATCATCCTCGTTCTCTATTCTTCTAGGAAATGCACCTCATCTAGATGGTCAGTATGCGG
TTTTTGGAAAGGTAACCAAAGGGGATGATACTCTTAAAAGATTAGAGGAAGTGCCTACTC
GCCAAGAGGGAATTTTTGTGATGCCATTGGAGCGGATCAGGATTCTTTCAACATACTACT
ATGATACTAATGAAAGAGAATCAAACTTGACCTGCGACCATGAAGTGTCAATTTTGAAAA
GGAGATTAGTAGAATCTGCTTATGAAATTGAATATCAGAGAAGAAAGTGCCTTCCATGAG
AAGTTGAACCGTCTTTTTATAGATCTTTTCAAAGATGAGTGTTGGGCTTTTCTAGCATACT
TGATTCTTCCATTTATATTCGAACATAAATTTTGTTTGTTTATTTGATTCCAATTATGACATT

AGTTCCATGTTAAATGCTAGATTTCTTATGGGGTTGGAACATTCCTCGCTGCCTTCTGGT
AATATTAGGTTATGCGTTTTTCATTCTGAAAAAAAAAA

SEQ ID NO: 181
CATAACTGTAAAGCCAGAGCAGAGAGAGAGTACAACGATGGCGAGCAAGAGGAGCCTG
AGAACGATGAATGTTTGGCCAACGCTACCGCCATTGGTCCTCCTGCTACTCTTATGCTTC
TCTTCCATGTCAAGTTCAGTGGTTGCTAAGAAGAGTGATGTCTCAGAGCTACAAATTGGA
GTGAAGCATAAACCAAAATCATGTGATATTCAAGCCCATAAAGGTGATAGAATCAAGGTG
CACTATCGGGGATCTCTTACAGATGGGACAGTTTTTGACTCAAGCTTTGAGAGGGGTGA
TCCAATAGAATTTGAACTTGGTAGTGGCCAAGTAATTAAAGGATGGGACCAAGGACTTCT
AGGAATGTGTGTTGGAGAAAAGCGCAAGTTGAGGATACCTTCCAAATTGGGTTATGGAG
CACAAGGCTCCCCTCCCAAGATTCCAGGTGGAGCAACGCTTATATTTGATACAGAACTTG
TTGCTGTTAATGGGAAAGGCATTAGTAACGATGGTGATAGCGATCTGTAAATAAGATTGC
TCCATTTTGCTTTAAAGGGCGGAGTTCATACTTTAGAAAGCGAACTTGAAGCAACTTAATT
TTAACAGGTTACATAATCTATTGTAGCATTCTCTTGAAATGGGAAATTGAGGTTGACTGTG
TACTTCTCCAGTGGACAGGAGAAAGCGATAAAATTCAAACGTTGTTGTCATTATTTCCTTC
GCTGAGTGACATGTGATACAATGAAAGGGAGTCTGAACCGTCTTTGTTGACCTTTATTAT
TACTCGATGTATATTTCTTCCCACTAAAGAGGCAAAATATTCTATTTTAAATTGAAAGACTT
TAAATTAAAAAAAAAA

SEQ ID NO: 182
CGAAACCTTCTGAAATTTCAGTTGTCATCGCGCGCTATGTCTGGGGCACCTGCAGAGCG
TCCTATCTCTTACTTTGACATCACCATCGGCGGCAAGCCCATTGGCCGCATCGTCTTCTC
ATTATATGCAGACCTAGTGCCCAAGACAGCTGAAAACTTTCGTGCATTGTGTACGGGTGA
AAAGGGTATCGGAAAGTCTGGCAAACCCCTTTGCTATGCTGGGTCTGGCTTTCACCGGG
TAATCAAGGGTTTCATGTGTCAAGGTGGTGACTTCACAGCGGGAAACGGCACTGGGGGA
GAATCGATCTACGGCGAGAAATTCGAAGACGAGGCGTTCCCCGTGAAGCACACGAAGC
CTTTCTTGCTGTCCATGGCAAATGCAGGAAAAGACACCAATGGGTCGCAGTTCTTCATCA
CTGTGAGCCAAACCCCCCACCTTGATGACAAGCACGTGGTATTCGGTGAAGTGATCAAA
GGCAAATCAATAGTCCGTGCGATTGAAAACTACCCTACTGCTTCGGGAGATGTACCTACT
TCCCCTATCATCATTTCTGCATGCGGCGTTCTCTCCCCCGATGATCCTTCTCTTGCAGCT
TCTGAGGAAACGATTGGCGACAGCTACGAGGACTACCCTGAAGACGACGATTCAGATGT
GCAGAATCCTGAAGTTGCGCTGGATATCGCACGAAAAATCCGTGAGCTTGGTAACAAGC
TCTTCAAGGAAGGGCAAATAGAGCTCGCGCTCAAGAAATACCTTAAATCGATACGGTATT
TGGATGTGCATCCTGTGTTGCCGGATGACTCTCCACCGGAGCTTAAGGACTCGTATGAC
GCACTTCTTGCGCCCCTTCTATTGAACTCGGCGCTTGCTGCTCTGCGCACGCAACCCGC
GGACGCCCAAACTGCGGTGAAGAACGCCACGCGAGCTTTGGAGCGTCTGGAGCTGAGC
GATGCAGATAAAGCCAAGGCCCTTTATCGCAGGGCGTCAGCGCATGTTATCCTCAAGCA
GGAAGATGAAGCTGAGGAGGACCTTGTTGCTGCTAGCCAGCTTTCACCTGAAGACATGG
CGATCTCCAGTAAGCTCAAGGAAGTGAAAGACGAGAAGAAGAAGAAAAGGGAAAAGGA
GAAGAAGGCATTCAAGAAAATGTTTTCGTCGTGACATACATCGTAATCATTGGAAGTATA
CCAGT

SEQ ID NO: 183
AAATGGAGGGAATGAAGTTGAAAGAATACTGCGCAAACAGAGAACAAAAATTGGTGGAC
AATGGCTTCTTCACTCCGCAGTTCTCTTTTTTCGTCATGGGCTTTGGATTCCAAGTCTGTA
TGCTCCCTATTTAACCTCAATCCAGGGAAGATGGGGCTGCCTTCCATTTCTACGCCACTG
AATTGGAGGACTTGCTGCTGTTCACACAGTTCTGAATTGTTGGAGCTAAATGAGGGGCT
CCAATCTTCTAGAAGGAAAACTGTGATGGGGTTATCCACCGTAATAGCTTTAAGCCTTGT
TTATTGTGATGAAGTAGGAGCTGTATCCACAAGCAAAAGAGCTCTAAGATCACAGAAGGT
TCCAGAGGACGAATACACAACTCTTCCAAATGGGCTCAAATACTATGATTTGAAAGTTGG
AAGCGGGACGGAAGCTGTAAAAGGTTCACGAGTTGCGGTACACTATGTAGCCAAATGGA
AAGGTATCACTTTCATGACTAGCAGGCAAGGCATGGGTATTACTGGTGGCACTCCTTATG
GATTTGATGTGGGTGCTTCTGAAAGAGGAGCTGTCTTAAAGGGATTGGACTTGGGAGTT
CAAGGGATGCGAGTTGGAGGCCAGCGCATACTTATAGTACCCCCGGAGCTTGCATATG
GTAATACAGGAATTCAAGAGATCCCACCAAATGCGACTCTTGAGTTTGATGTGGAGTTGA
TTAGCATCAAGCAAAGTCCATTTGGGTCAAGTGTAAAAATCGTTGAAGGGTAGTTGTGAA
ATTGATTAAATTGCTAATAAATATTGTTGTAATTCATTATCATCATGTTTCTTATAATTTAAG
AGCTATGAAACAACTACCTTTTGGAATGGTTTTGTTTCTTATAATTTAAGAGCTATGAAAC

AACTACCTTTTGGAATGGTTTTGTTTTTAGCATCCCAATTATATTTACAATGCTTTGAAAAA
TCAATAAAATAAGTTTCTATTTAAAAAAAAAAA

SEQ ID NO: 184
GCGAAGCAAGCAAACCAAATACCAGCAGCTCGAGGGTTTGGTTGTCGTTCCTCTTTCCC
TTTGCATTTTCAAACGATCTGGGCGCCTCCTTAACAATAACATATCAGTTTCTGCTTTCCC
TATACATTTTCGAAGTTCGGGAAATCTGTATGGCGCAAATCTGAACCTTGGCCGTGGAAA
ATCCACAGAGCTCCGGCGGAAAACAGAGGAGGCGGCCTTGAAACGCTTATAAACAGGC
TTGGATAGCCATATTGATCGGTGTTCAGGCAAGGAGGAATTCCAGATCCTGCACTTCTTA
GAAGGTTTCCGTCTTTTGTTACATTTTAATTTTCACCAGAAGGATTAATCACCTGCGAGGT
GCATTTCACCCTCTTTTGGTTGGAAACATTTATGGGCTAAATTGTGGCGTGTGAACGTTT
CATGGGAGCCATTGAGGATGAAGAACCACCCCTGAAGCGTTTAAAGGTGTCCTCACCCG
GCTTGAGAAGAGGTTTGGAGGAGGAGGCCCCTTCATTGTCTGTGGGGTCAGTGAGCAT
CTTAATGGCCAAATCATTGTCTTTGGAGGAGGGTGAGACAGTAGGTTCCAAAGGACTCA
TAAGGAGGGTGGAATTCGTAAGGATAATAACACAAGCTCTATATTCTTTGGGTTACCAGA
AAGCTGGTGCGCTTTTGGAAGAGGAGTCAGGAATACTTCTCCAGTCTTCAAATGTAGCTT
TATTTAGAAAGCAGATACTCGATGGAAAGTGGGATGAAAGTGTTGTTACACTGCGGGGT
ATTGATCAAGTGGAGGTGGAAGGAAACACACTGAAGGCTGCATCCTTTCTGATCTTGCA
GCAGAAGTTCTTTGAACTATTGGACAAAGGTAACATCCCTGAAGCTATGAAAACACTTAG
ATTAGAAATTTCTCCAATGCAATTGAACACAAAAAGAGTACATGAACTTGCAAGCTGCATT
GTGTTTCCTTCACGGTGCGAGGAGCTTGGATATTCAAAGCAGGGGAATCCAAAGTCCAG
TCAACGGATGAAAGTTTTGCAAGAAATTCAACAATTATTACCACCTTCTATAATGATCCCA
GAAAAGAGGCTGGAACGTTTAGTGGAGCAGGCCCTTAATGTGCAGCGAGAAGCATGCAT
TTTTCACAACTCCTTGGATCCGGCCCTTTCTCTATATACTGATCATCAATGTGGACGAGAT
CAAATTCCTACAACGACATTACAGGTGTTAGAGTCACATAAAAATGAAGTGTGGTTTCTA
CAGTTTTCAAACAATGGGAAGTATTTGGCTTCTGCATCAAAAGATTGTTCTGCAATAATTT
GGGAGATAACTGAAGGGGATTCTTTCTCCATGAAACATAGATTAAGTGCACACCAGAAAC
CAGTTTCATTTGTTGCTTGGAGCCCTGATGATAAACTGCTGCTTACATGCGGTATTGAGG
AGGTTGTTAAGCTTTGGAATGTTGAGACTGGTGAATGTAAACTGACCTATGATAAAGCCA
ACAGTGGATTCACTTCCTGTGGTTGGTTTCCAGATGGTGAACGGTTTATCTCTGGCGGA
GTTGATAAATGCATCTACATATGGGATCTTGAGGGAAAGGAGTTAGATTCATGGAAAGGA
CAAGGAATGCCCAAAATATCTGATCTCGCAGTCACATCTGATGGTAAAGAAATAATCAGC
ATATGTGGGGACAATGCTATTGTGATGTACAACTTGGATACAAAAACTGAAAGGTTAATT
GAGGAGGAAAGTGGAATAACTTCCCTATGTGTATCAAAGGACAGTAGATTTCTCCTCCTA
AACCTTGCAAATCAAGAGATACACCTCTGGGATATTGGAGCTCGTTCAAAGTTATTGTTA
AAGTATAAAGGTCACAGACAAGGTCGCTATGTGATAAGGTCCTGTTTTGGTGGGTCTGAT
CTTGCATTCGTTGTTAGTGGTAGTGAAGACTCACAGGTATACATTTGGCACCGAGGTAAT
GGAGAACTTTTAGCTGTTTTGCCTGGTCACTCTGGTACAGTAAACTGTGTGAGCTGGAAC
CCTGTGAACCCACATGTGTTTGCATCTGCTAGTGATGACTATACTATTCGTATATGGGGT
GTAAACAGAAACACTTTCAGGAGTAAGAATGCTAGTTCTAGTAATGGCGTCGTTCACCTT
GCAAATGGAGGGCCATAGCCAAGAGAGATAAAAGACCAATTTTGAGCAGCAGTTTAAGC
ATTGTAATAACGCTCATTATTCATTTCTGCACTTGTAAATTCTGATCTTACTTCATTTATTT
CTATTCCAAAGACCAAGCTTGTAAATTATGCTGGTTCCATATGGGGGTTAATCAGTATCC
TCGTTATTTGTGACACCAAATTATCAAATCTCACAAAATTTGTGAATGGGACTAAAAAAAA
AAAAAAAAAAAAAAAAAAA

SEQ ID NO: 185
GGCATGGTGCAGATCCTTCGCCGTCAAACACATGTACATGTAATTTTCAAGCTGGACTGA
TATTCCCCTCCTCATTTTCCATTAAGAAAACGAATGGAAATGCAGTAGAAATTGTTCAGAG
ACTCGTAGACAGAGGTTTAAGGAACATAGATAGATTTAGAAATGCCGGGAACAACTGCG
GGTGCAGGTATCGAACCCATCGAGCCTCAGTCCCTGAAGAAGCTCAGTCTCAAGTCTCT
TAAGCGCTCCTTCGATCTCTTTGCTTCCCTTCATGGCGAACCTCAACCTCCCGATCAACG
CAGTCAACGAATAAGAATTGCCTGCAAGGTGCGAGCTGAATATGAAGTCGTAAAGAACTT
GCCAACTTTGCCACAACGGGAAGTTGGCAGTTCAGTATCAAACTCAAATGTTGGAGAAA
CTCATTCATCTCTGACAACTAATCAAGCTCAAGGATTTCCAACTGACACATCAGGAGATTT
GTCGAAAGATGAAGGGAAAGAGATTACTAGCATTGCTGTTCATTTGCAGCCACAAACAG
GATTGATTGATGGAAAGGCAGGAGCAATAGCTGGAACAAGTACTGCTATTTCTTCTGTGG
GTTCTTCAGATCGGTACCAGCCAAGTGCGGCCATCATGAAAAGGCTACCCAGTAAATGG
CCACGTCCTATATGGCATCCTCCTTGGAAGAACTATCGGGTTATCAGTGGGCATTTGGG
CTGGGTGAGATCTGTTGCTTTTGATCCTGGTAATGAGTGGTTTTGTACTGGCTCTGCTGA

CCGGACAATAAAGATTTGGGAAGTAGCCACAGGAAAACTCAAGCTCACATTAACAGGGC
ATATTGAGCAGATACGAGGTTTGGCTGTAAGTTCTCGGCATCCATATTTGTTTTCTGCTG
GTGATGATAAGCAGGTCAAATGCTGGGACCTTGAGTATAACAAGGCTATTCGCTCTTACC
ATGGACATCTAAGTGGAGTTTACTGTTTGGCACTTCATCCGACATTGGATATTCTTTGTAC
TGGTGGTCGTGATTCTGTTTGCCGAGTATGGGATATTCGCACCAAAGCTCAAATATTTGC
ACTCTCTGGCCATGAGAACACTGTGTGTTCAGTTTTTACCCAAGCAATAGATCCTCAAGT
GGTGACCGGGTCTCATGATACTACAATCAAGTTATGGGACCTTGCTGCAGGAAAAACAA
TGTCTACTCTTACGTATCACAAAAAATCTGTTCGGGCAATTGCAAAACATCCTTTTGAGCA
CACTTTTGCGTCTGCATCTGCTGATAATATAAAAAAATTCAAACTTCCCAAGGGAGAGTTC
CTGCACAATATGCTATCACAGCAGAAGACAATTGTTAATGCCATGGCCATCAACGAGGAC
AATGTTCTTGTATCTGCAGGTGACAATGGAAGCTTATGGTTCTGGGATTGGAAGAGTGGT
CATAATTTTCAGCAGGCTCAAACAATTGTACAGCCTGGATCTCTGGACAGTGAAGCTGGA
ATATATGCACTCCAATATGATATAACTGGCTCTAGGCTTGTTTCTTGTGAGGCTGACAAAA
CAATTAAAATGTGGAAAGAAGACGAGACAGCCACTCCAGAGAGTCACCCAATAAATTTTA
AAGCGCCTAAAGATATCAGGCGTTTCTAAATCTTTAGTTAATTAATTGTATTTCTATTAAAA
TTGGGCATTTTAGTTGTATATGATCTATAGCCTTTTATAGAGTGCTAAATGAAGAGGTAAT
CTTTGTATTCAGCTTTGTTGTGAACTATCAATAGACGGGGATGGTCCTTTTTAGCTGCTC
CTTAAGCAGCTCAAATGATCTGTAAGAGGTAGAGACGTAGTTGTTGCTGTATTCAAGTAT
GATACATCAGCACAGTAAAGCATCTTCCTTACAAAGCTATTTTATTCAAAAAAAAAA

SEQ ID NO: 186
GCATAAGTGAGTTCATAGTCTTTTGCTCATAAAATCTATGAACAAAAACCCTTGGAACACT
TACATGGAGGGCCTTATATATTCATTTCTTTACTCATCATCGCTTCTTTTTCGTCTTTCTTC
ATCGGTGCCAGTGCTCGCGGCTTGTCTACACTGAAGTCCGATAAGGAACCTAATACATT
GCAAGCAAGATGAGGCCGATTTTGATGAAGGGCCACGAGAGGCCTCTGACGTTTCTCAA
GTATAACAGGGACGGGGACTTGCTGTTTTCTTGTGCGAAGGATCACACGCCCACCGTCT
GGTATGGACACAACGGGGAGCGCCTTGGGACCTATCGAGGACACAACGGTGCTGTCTG
GTGCTGTGATGTCTCAAGGGACTCTACACGTTTAATAACCAGCAGTGCAGATCAGACTG
CCAAGTTATGGAATGTGGAGACAGGAGCTCAACTATTTTCTTTTAACTTTGAATCCCCAG
CTAGAGCGGTGGACCTTGCCATAGGGGATAAGCTTGTTGTAATAACCACAGATCCATTTA
TGGAGTTGCCTTCAGCAATTCATATCAAACGTATTGAGAAGGATCTGTCAAAGCAGACTG
CTGATTCTGTACTTACAATAACCGGGATCAAAGGAAGGATAAACAGAGCTGTTTGGGGC
CCTCTAAATAGTACAATAATCAGTGGTGGAGAAGATTCTGTTGTTCGTATTTGGGACTCA
GAGACAGGAAAGCTGCTCAGGGAAAGTGACAAGGAAACAGGCCATCAGAAACCTATAAC
ATCTTTGTGCAAGTCTGCAGATGGATCCCACTTCCTGACAGGGTCCCTAGATAAATCAGC
CAGGCTATGGGATATCAGGACTTTAACTCTTATTAAGACGTATGTCACAGAGCGCCCTGT
AAATGCAGTTGCAATTTCTCCACTACTTGACCATGTTGTGATAGGTGGAGGTCAGGAAGC
ATCTCATGTGACTACTACTGATCGTCGTGCAGGAAAATTTGAAGCAAAATTCTTTCATAAG
ATTCTTGAAGAAGAAATTGGCGGGGTTAAGGGACACTTTGGACCTATAAATTCACTGGCA
TTTAATCCTGATGGAAGGAGCTTTGCGAGTGGTGGAGAGGACGGTTATGTAAGATTGCA
TCATTTTGATCCTGATTACTTTCACATCAAGATGTAATTTGGTACTTTGGAAAGAAGGCTT
GGTCTGTAGCGTTTCTCATTACATGATGGATTGCCAGGAATTCACAAAAAATATCGCACA
ATTGGTGTCTTGATTATGTAAGATTCTCGAGAAGTAGGAAATTTTAAGAGGTTTTAACAAA
TTTTGGGGAGGCTTGATTGATTTTTCAGACAGAATTGTTTGTATCTTGTTGGGAATTATTA
CTATTTATATACTTCAGCTGAGCTGGGTCGTGGTAAAAAAAAAAAAAAAAGGGAAAGCAAG
GTCCTACAACACCACTGCCCGACTTGGTCACCATTCATACTCCGAAAGAAGAGGAGGAA
TTCGCTGCCAAGCCAGTGCTTGGCAAGGAAACTGAGATACTTGTCTAGGTCACTTTTAGT
TCATTACATACAACATTTATATTTAAGCTTCTAATATATTCTTGACCCATCCACATGAATTC
AATTCCGGTCATATGTAGACGACTATAATGTTGTTTGTGTCCTATAACTATAGTGTTGGAT
CAACGTCGTTCAATTTTTGACTATTTTTTAACTTCGATATGATTGTTGACGTGTTAAGGGA
GGCTCTTTAGTTCTCAAGGTGTTAAAGCTTGGGGAATTGTAGTTTCTCAAGGATGTTAGT
TTATTTGGCAACGACGCTACTTGTTCATATTTGAAATTAATTAAAATTTCAAGCGTGGCAT
TTAATTTGCCAGTTTTCTCCTTTTAAAAAAAAAA

SEQ ID NO: 187
GAGGATACCGAGTTGATTACTTGCAATATTTCTTGTTTTGGGCTTCCGCGGTTCTTCAGA
ATTGAAGAGCGCTTAGGGTTTGTTATTTATTGCAGAAGTCCGATAAGGAACATAGTACAT
TGCAAGCAAGATGAGGCCGATTTTGATGAAGGGCCACGAGAGGCCTCTGACGTTTCTCA
AGTATAACAGGGACGGGGACTTGCTCTTTTCTTGTGCGAAGGATCACACGCCCACCGTC
TGGTATGGACACAACGGTGAGCGCCTTGGGACCTATCGAGGACACAACGGTGCTGTCT

GGTGCTGTGATGTCTCAAGGGACTCTACACGTTTAATAACCAGCAGTGCAGATCAGACT
GCCAAGTTATGGAATGTGGAGACAGGAAATCAACTATTTTCTTTTAACTTTGAATCCCCA
GCTAGAGCGGTGGACCTTGCCATAGGGGATAAGCTTGTTGTAATAACCACAGATCCATT
TATGGAGTTGCCTTCAGCAATTCATATCAAACGTATTGAGAAGGATCTGTCAAAGCAGAC
TGCTGATTCTGTACTTACAATAACCGGGATCAAAGGAAGGATAAACAGAGCTGTTTGGG
GCCCTCTAAATAGTACAATAATCAGTGGTGGAGAAGATTCTGTTGTTCGTATTTGGGACT
CAGAGACAGGAAAGCTGCTCAGGGAAAGTGACAAGGAAACAGGCCATCAGAAAGCTATA
ACATCTTTGTGCAAGTCTGCAGATGGATCCCACTTCCTGACAGGGTCCTTAGATAAATCA
GCCAGGCTATGGGATATCAGGACTTTAACTCTTATTAAGACGTATGTCACAGAGCGCCCT
GTAAATGCAGTTGCAATTTCTCCACTACTTGACCATGTTGTGATAGGTGGAGGTCAGGAA
GCATCTCATGTGACTACTACTGATCGTCGTGCAGGAAAATTTGAAGCAAAATTCTTTCATA
AGATTCTTGAAGAAGAAATTGGCGGTGTTAAAGGACACTTTGGACCTATAAATTCACTGG
CATTTAATCCTGATGGAAGGAGCTTTGCGAGTGGTGGAGAGGACGGTTATGTAAGATTG
CATCATTTTGATCCTGATTACTTTCACATCAAGATGTAATTTGGTACTTTGGAAAGAAGGC
TTGGTCTGTAGCATTTCTCATTACATGATGGATTGCCAGGAATTCACAAAAAATCGCACA
ATTGGTGTCTTGATTATGTAAGATTCTCGAGAAGTAGGAAATTTTAAGAGGTTTTAACAAA
TTTTGGGGAGGCTTGATTGATTTTTCAGACAGAATTGTTTGTATCTTGTTGGGAATTATTA
CTATTTATATACTTCAGCTGAGCTGGGTCGTGGTACAAACATTTTACACCCTATAACAAAA
TATAGTGTCATAAGTTTACACCAGGTAACAACTCTATATCCAAAGCTGCAGTTAAATTGTC
CCTGCCCTTCTGTGGAAGGCTTAATGAACATGTTTAAGGTTCAAAAAAAAAA

SEQ ID NO: 188
CTGCGTTATGTCATCGTCGGTCATAAACCCCAGGGTTTAAGTTGTCTTTATCACCCTCTG
CCTGGCTCGTTAGTGTAGTGAACACCATTATTTCACTGCTACCATAATATACAGAAAATCT
CCTAGAATATACAGAAAAGACCTCGAAATGGCGGAGAATAACGTCGGCGACTTCATACC
ACTTGATCGACAAGAGTACCCTTCAAAACCCGCTCCAGGCGCCGTTGACTCGAGCTTCT
GGAAATCCTTCAAGAAGAAGGAAGTTTCCCGGCAGATTGCAGGCGTTACTTGCATCAATT
TTTGCCCTGAACCTCCTCACGATTTCGCAGTTACATCGTCTACCCGGGTACATATCTATG
ACGGTAAATCTTGTGAGCTCAAGAAGACAATCACGAAGTTTAAAGATGTGGCATATTCAG
GAGTATTTCGTTCTGATGGACAGATCATTGCTGCTGGGGGTGAGACAGGGGTCATTCAG
GTCTTCAATGCAAAGTCTCAAATGGTTCTCCGGCAACTGAAGGGGCATGGCAGGCCTGT
TCGAGTTGTAAGGTATTCTCCACAGGACAAACTTCATCTACTTTCAGGTGGTGATGATAG
CATGGTCAAGTGGTGGGATATCACAACCCAGGAAGAGTTACTAAATCTAGAAGGGCATA
AGGACTATGTTCGCTGTGGTGCTGCAAGTCCATCGAGTGTTAATTTATGGGCTACAGGTT
CATATGATCACACAGTACGGCTGTGGGACTTGCGCAACTCAAAGACTGTTTTGCAATTAA
AGCATGGGAAACCACTGGAAGATGTTTTATTTTTTCCATCTGGTGGATTGCTGGCCACCG
CAGGAGGTAACGTAGTGAAGGTATGGGACATTCTTGGAGGAGGACGACCCATACACACA
ATGGAAACCCATCAAAAGACTGTCATGGCTATGTGTATTTCAAAAGTACCCAGGTCGGGC
CAGGCTTTGGGTGATGCACCTTCTCGTCTTGTTACTGCATCACTAGATGGATATATGAAA
GTTTTTGATCTTGATCATTTTAAGGTCACCCATTCTGCAAGGTATCCAGCTCCAATTCTAT
CAATGGGCATATCCTCTTTGTGTAGGACAATGGCAGTAGGAACCTCATCTGGGCTTCTTT
TCATCCGCCAGAGGAAGGGGCAAATAGAAGATAAGATACATTCAGATTCATCAGGTCTTC
AAGTAAATCCAGTTAATGATGAAAAGGATTCTGCTGTTTTGAAACCCAATCAATATCGTTA
TTATTTACGTGGTCGCAGTGAAAAACCTTCTGAAGGAGATTATGTTGTGAAAAGAATGGC
AAAAGTGTATTTCCAAGAGTATGATAAAGATTTACGTCACTTCAATCATTCAAAAGCTTTG
GTTTCTGCTCTAAAAGCTGCAGACTCCAAGGGTACAGTAGCTGTCATAGAAGAACTAGTT
GCTAGAAAGAGGCTGATACAAACTTTGTCAATTCTCAATTTGGACGAGCTCGAGTTGCTA
ATCAATTTTTTATCCAGGTTTATTCTTGTACCAAAATATTCAAGATTTCTCATTTCCCTTAC
AGATAGGGTTCTGGATGCACGTGCAGTAGATCTTGGAAAGTCAGAAAATTTAAAAAAACA
GATTGCAGACCTGAAGGGAATAGTCGTTCAGGAGCTTCGAGTGCAGCAGTCCATGCAAG
AATTGCAAGGGATTATTGAGCCACTAATTCGAGCTTCTGCTCGATAAATGATCTAAATGG
AGAGTTTTATTCATTACATGAAAGAGTATGTCACCTTTCGTGCTCCATCTATTGATTTTTAT
TTTTAAAGCCTAGAGAAAAAAGTATTTATCACAAAATGCTGAAGCATTTTTGGTTCAAGGT
TTCGTGTTCCCTTGTAATATTTCTGTTTCAGTGGGGCTAGCTGGAATGACATGTTCATCTA
TGGAAGAAGAAACAGAGGAATTTATGAGGGATCGAGAAAGAGAATTTATTTCATTATTTC
TCATCTCTTGTTTTTGATGTGCCGCTCATTCAGTTTGAGCCCTTGTAACATAATGTTCGGT
CAATGGCATGGCCATTGAAACAATTTTGGTCATGTGACGGGCATACAGAAAATTTGTATC
TAAGTATCTATACTTATCTCCAAGTATAGAAGCTTAGTGAGTCTTATTAGTTCTTAAAATAC
TAATTTTTCTGCGATATTTTGCATAAATATCTAAATTTATAAGGATTTAAAAAAAAAA

SEQ ID NO: 189
GTTCCCCAATTCCCATCAACGTTTTATAAGAATTGGTACGACTTTGCGGCCTGCGGAGAC
AAGGTGGGTCTCGCATCATGGCGTTGGAAGAAAGTTCACGAGGCAGAGCAAACAGAGA
AGAGTTCGGAGAGCATATTCATTTCTTCGATTGTTTATTCGTCTAAGACGAACAGTCCGA
GGACATGTCCGGCGGATGACTTGGTGAATTTGTTATCAAAATTTTCTTAATACTCCGAAA
GGTTCGGTCTCAGCATTTCCCTTTAACATCTTTTATATTTTATTTCAGTAGCAGACACGAA
CCAAATCAAAATACCCTTCATACGGCTTAATTCTCAGTCCCGTAATTTGGCGGTAATTTGT
TTTCCCTTTTCTGCTCGATGGTTTAATGTCGATGAATGAAGCAGTTGTGTTTCTGGGCTG
CTTTCGGGCAGTTCTCATAGGGCGGTTTTTAAGATATCCAGGGTTAGAAAACTGGTTTTC
GACAGGAAATCTGCTCACGATTTCGACCCTTAAATCGAAAGTGTACCTGGTTTCATCCAG
TTTCTTTTTTCAAAAATAGAAATTCCGTATTATTCTCCAGGTGAAGCTTACAATATCAGGTT
TCTTCTCGTCTGGACTATATCACGGCAACACTGAAATGTTTGTTAGATGGGAAGCCGAGA
AAAGTTTCGGCTATTTGATAGTTAGTTTACATTTAAAGGCAGCGGAGCAGTTTTTTTTCAC
GAAGGGTTTGATGTTTAAAGGGAGACAAGCAGTTAGTTCGAAAGGATTGGGTAACAGTA
ACAGTTTGTTTGGTAGACAGGTTCAGGCTTTATTTCGTATCGACCCAGATAGCTTCCGGG
AATTGAGAATTTGAAGAATTTAAACAGACAATTATTGCTGTCAAATTACAGGTTTCTGGCA
TCAGGGTTTCCTGCAATTAGGGTTTACTGATATCAGGTAAAGCCATGGACGTCGAAACCT
CAGGGAAACCTACAGGAAACAAAAGAACCTACACACGTCTGCCAAGGCAAGTATGCGTG
TTTTGGCAGGAAGGCAGATGCACTAGAGAATCATGCAATTTTCTGCATGTCGACGAACCT
GGATCAGTCAAAAGGGGGGGAGCAACTAATGGGTTTGCTCCAAAGAGATCATATAACGG
ATCGGATGAGAGGGACACGCTGGCCGCAGGACCTCCCGGAGGATCGAGAAGGAATATT
TCTGCTAGATGGGGAAGAGGACGAGGTGGTATATTCATATCCGACGAGAGGCAGAAGAT
TCGCAATAAAGTATGCAATTACTGGCTTGCAGGCAATTGCCAGCGTGGAGAGGAATGCA
AATATTTACACTCCTTTGTTATGGGTTCGGACGTGAAGTTCTTGACGCAGCTGTCCGGCC
ACGTTAAGGCTATTCGAGGGATTGCTTTTCCTTCTGATTCTGGTAAACTCTATTCTGGAG
GCCAAGACAAGAAGGTTATAGTGTGGGACTGCCAAACAGGGCAGGGTACAGACATCCC
ATTAAATGACGAGGTTGGATGCCTCATGAGTGAAGGCCCATGGATTTTTGTTGGCCTCC
CAAATGCGGTCAAGGCTTGGAATATTTTGACCTCAACAGAACTAAGTTTAGTTGGTCCCA
GGGGACAAGTACATGCTCTGGCTGTTGGTAATGGAATGCTTTTTGCTGGAACACATGAT
GGCAGTATATTGGCTTGGAAGTTTAGTCCTGCTTCCAATACCTTTGAGCCTGCAGCTTCC
CTTGTAGGGCATACCCAAGCTGTTGTTTCATTAGTCTCAGGAGCAGACAGACTTTACTCT
GGTTCCATGGACAAGACAATAAGAGTTTGGGATTTGGGAACTTTTCAATGTCTACAGACC
CTGAGGGATCACACATCAGTTGTAATGTCTCTTTTATGCTGGGATCAGTTTCTGTTGTCAT
GCTCTTTAGACAATACAGTGAAGGTCTGGGTTGCTACATCAAGTGGAGCCCTTGAAGTG
ACATATACTCATAATGAAGAACATGGTGTTCTTGCGCTTTGTGGAATGAACGATGAACAA
GCCAAGCCTGTCTTGCTTTGTTCTTGCAATGATAACACTGTGCGTCTTTATGACCTGCCA
TCGTTCAGCGAACGAGGTAGAATATTCTCAAGAAACGAGGTACGAACATTCCAGATTGCA
CCTGGTGGATTATTTTTCACCGGTGATGCGACAGGTGAACTTAAAGTTTGGAACTGGGC
AACCCAGAAATCCTGATGCAACTTAAAAAATTCCTGTTGCAATCAGGGCACTCTGCTTGA
AATGCAAATCTCTCTGGGAACTCCTTCACATGGCAGAAAGGTGCAGAGTCTGGAGAAGT
TGAAATTTCACGTCCTGTATACTCACTCAAGCAACTTTAGGATGAAGAGCTAAAGTATATC
AAAGATATATAAGTGGGTAATATTAACACTGAATAGCTGCAAGTGATGCATGCTTGATTTC
CGTCAGGAAAATCAGAATTTTAAA

SEQ ID NO: 190
GTCCTAATGACTGATACTGGTTTCTGGTTTACCGTCATGTTTCCCGGACCCAAGCCGCC
GGTTTTCAGAACGGCTGATGCTTTCTGAACCCAGAGCGGAGAGAGCGAGACCAGGGTC
TTTGGATTTCGATCTATGGTCGAGTCCTTCCCTTATGCGAAGGTTATGTTCCCTAAATCTG
CCAAGAAAATAATTATTTGGCATATGACATAGAGCTCGGTTTACTGATTTGCATACCGAAC
ATCTCTGGTTTTGCTTATTTTGCTGGTTCAAATCAGGGTTTCGATGCATTCTCCATAACAT
TAACACTCCTGCTTCTGTAAATTATCGAGCTATGTCGGTGCAAGAGCTTAGGGAACGTCA
TGCAGCGGCAACGGCAAAGGTTAACGCTCTCAGAGAGCGGATAAAGGCTAAGCGCCTC
CAGCTGTTGGATACAGACGTTGCAACGTATGCCAGCAGTAATGGAAGGACCCCAATCAG
CTTCAGCTTCACGGATCTGGTTTGCTGTAGAACACTTCAAGGGCACACAGGAAAAGTTTA
TTCATTGGACTGGACATCTGAGAAAAATCGCATTGTGAGTGCATCTCAAGATGGACGGTT
GATAGTTTGGAACGCTTTGACAAGCCAGAAAACACATGCAATTAAACTCCCCTGTGCCTG
GGTGATGACATGTGCATTTTCTCCTAGTGGGCAGGCAGTTGCATGTGGTGGGCTTGACA
GTGTCTGCTCCATTTTCCAGCTTAACAACCAGTTAGACAGAGATGGACATCTACCTGTTT
CCAGAATTCTTAGTGGTCATAGGAGCTATGTCTCTTCTTGTCAATATGTTCCAGATGGGG
ATACACATGTAATAACTGGCTCTGGTGATCGGACATGCATTCAGTGGGACGTAACAACA

GGACAAAGGATTGCAATCTTTGGAGGTGAATTTCCATTGGGTCACACTGCTGATGTTATG
AGTGTATCCATCAGTGCAGCAAACCCCAAGGAATTTGTTTCAGGCTCTTGTGATACAACA
ACCCGTTTATGGGATACGCGTATTGCTAGCAGGGCCATACGAACATTTCATGGACATGA
GGCAGATGTGAATACTGTGAAGTTTTTTCCTGATGGACTAAGATTTGGATCTGGCTCAGA
TGACGGTACCTGTAGGTTATTTGACATACGAACAGGGCACCAACTGCAAGTCTATCGTCA
ACCTCCCAGGGAAAATCAGTCCCCAACTGTGACTGCAATTGCATTCTCATTTTCTGGTAG
GCTTCTATTTGCAGGATATTCCAACGGGGACTGCTTTGTGTGGGACACAATATTGGAAAA
GGTAGTTCTAAACTTGGGAGAGTTACAAAATACACATAATGGCCGGATAAGCTGTCTTGG
ATTGTCAGCCGATGGGAGTGCTTTATGTACGGGGAGCTGGGACAAGAATCTAAAGATTT
GGGCATTTGGAGGCCACCGAAAAATAGTCTGATTGGTGAGAGATTATAAAAACGAAATTC
ACTTGAATGATTAGCCTTGCGAAAATGCACTCTTTATAAAGTGGGATGAGGTATGTGTTT
CCTTCCTATTGGCTAACCTGAATTCCTCCTGCAGCCCCAATTATTAAATGCTATCCTGCAA
CAGTCATGATTCGATGGGCCTAAGCTGTATTGATATTCTCTTGATCGTTGGAGAAGGCAG
ATCATTTGTAACTTTGTAATATGTATTGATATTTTATTTGTGTTTTCTCATTGCATTTGAGCT
GTGCTACACAAAAATGATCCTTGGCTCTATGATTCCCAATTCTCAGTTCAAGTACATGTGA
AGGATCTGGAGTTTACAGTCACATTTCCATATTCAATTATACATGAAAGCTTGTGGATTTG
AATGAAAAACAATGTAATTGGTTTAGTATGTTACAAGTTTATGATTACAAGGTTTAAAAAAA
AAA

SEQ ID NO: 191
CTGTCGACCCCGGTCTGGGATGATTTCTTAAACCCTCTCGACGCCCTTTCTTACGCTTAA
AACCCTTTCTTCTTTTGTTTTCATCTTCACAACCACTGGACAGAGATGTTCTGACCGCGG
ATTCACTGAGCTATTGCTGAACCAAAGATTTTTGCTTCGCGGCTACCCAGGGTTGTGTGT
CCAAATTGATATTTTCCACCACGTTTCTGGGTTTTTTTGCCCTGTGAAATAATGAAGGTGA
AGATAATATCACGTTCAACAGATGAATTTACAAGGGAACGGAGTAACGATCTTCAAAGGG
TTTTTAGAAACTTTGATCCTAACCTACATACTCAGGCAAGGGCTCAAGAATATGTCCGTG
CTCTCAATGCAGCAAAATTAGACAAGATTTTTGCAAAACCCTTTCTGGCAGCAATGAGTG
GACACATTGATGGCATTTCTGCCATGGCAAAGAGCCCACGTCATTTGAAAAGCATATTTT
CGGGTTCCGTGGATGGAGATATCCGGCTCTGGGACATTGCGGCCAGGCGCACAGTTCA
ACAATTTCCAGGTCATAGAGGAGCGGTGCGTGGTTTGACTGTGTCTACAGAAGGAGGGC
GGCTTATTTCTTGTGGAGATGATTGCACAGTTAGACTATGGGATATCCCAGTTGCTGGCA
TTGGGGAATCAAGTTATGGCTCAGAGAATGTTCAAAAGCCACTAGCAACCTATGTGGGA
AAAAATTCTTTCAGGGCTGTAGATTACCAGTGGGATAGTAATGTGTTTGCAACTGGTGGG
GCTCAAGTTGACATTTGGGATCATGACAGGTCAGAACCGACAAACAGTTTTGCTTGGGG
GTCTGATACTGTCATTTCTGTACGGTTTAATCCTGCAGAGAAAGATATATTTGCGACAACT
GCCAGTGACCGCAGTATTGTTCTATATGACCTTCGAATGGCATCACCTTTGAACAAGTTA
ATTATGCAGACAAGGAATAATGCCATTGCATGGAATCCAAGAGAACCTATGAATTTCACT
GCAGCTAATGAAGATTGCAACTGCTATAGCTATGATATGAGAAGAATGAATATTTCAACAT
GTGTGCACCAAGACCATGTTTCCGCTGTGATGGATATTGATTATTCTCCATCTGGTCGGG
AGTTTGTGACGGGATCTTATGATAGAACAGTGAGAATATTCCCCTATAATGCAGGTCACA
GCAGAGAAATCTACCATACGAAACGAATGCAAAGGGTGTTTTGTGTGAAATTTAGTGGTG
ATGCGACTTATGTTGTCTCTGGAAGTGATGATGCTAATATTCGTCTGTGGAAAGCTAAAG
CATCAGAGCAGTTAGGAGTGCTTCTCCCGAGAGAGCGCAAAAGGCACGAATATCTAGAT
GCTGTAAAGGAACGTTTTAAGCATCTTCCAGAAATTAAACGCATTGAGAGACATCGACAC
CTACCAAAGCCTATATACAAAGCAGCGCTTCTACGGCATACTGTAAATGCCGCAGCAAA
GAGGAAAGAGGAGCGTAAGAGAGCGCACAGTGCTCCTGGATCTGTGGTCACTAATCCT
CTTCGGAAAAAGCGAATTGTAGCTCAATTGGAGTGAATATCTAATTCTAGCACGCCAATG
CAGGTTATACTATGTATTGGAGTGAAAATTTAATTATAGTATATGCTAATTCAGGTTATACT
ATGTATTCAAAATATGTTATTGGGCAATCGTTATTGATTTTACCTATCGCTATCTCACTGTC
CGCCAATTTAGTGTAATTCTATGTACCAAAAAAAAAA

SEQ ID NO: 192
TCTAGGTGGGGTGTTAGGGGAATTGCTCACCGGATTTTGATTTTCTGGTCTTTTTATGTG
CGTGGCTCGTGTAGTTTGATACTGAACATTCAGGAATTGAAGCTTCTTCACAAGGATTCG
GATTTTGATTTGTTTCAGAAGGTGGGGTACAAGATTCAGTAGCGAAGTTCAGTGTTTTGC
CGAAGGGCTCTTTCAAATTAAGTAAGATGGATCACTACTACCAAGATGACTTTGATTATTT
GGTGGATGACGAGATGGTTGATTTTGCTGATGACGTCGAAGATGATGTACGCACCCGGC
GCAGGAGTGATATAGACTCAGATTCTGAAAATGATTTCGATTTAAACAATAAGTCACCTG
ACACAACAGCTCTTCAAGCCAAGAGAGGAAAAGATATTCAGGGTATTCCATGGAATCGAT
TGAACTTTACTAGGGAAAAGTATAGGGAAACTCGTTTACAGCAGTATAAGAACTACGAGA

ACCTTCCACGGCCGCGACGCAGTCGCAATCTGGACAAGGAATGTACTAATTTTGAAAGG
GGATCCAGTTTTTATGATTTTCGCCACAATACGAGATCAGTCAAGGCCACTATTGTCCAC
TTTCAGTTGAGGAATTTAGTGTGGGCAACATCAAAGCATAATGTTTACCTCATGCAGAATT
ACTCTATCATGCACTGGTCTTCACTAAAACAGAAAGGGGAAGAAGTTCTTAATGTTGCTG
GTCCGATCGTTCCATCTGTGAAACACCCCGGTTCATCACCACAAGGTCTGACAAGGGTC
CAGGTCAGTGCAATGTCTGTGAAGGACAATTTAGTTGTTGCTGGTGGTTTTCAGGGAGA
ACTTATTTGCAAGTATTTGGACAAACCGGGAGTGAGCTTCTGTACAAAAATATCTCATGAT
GAGAATGGCATCACTAATGCAGTAGAGATATACAATGATGCAAGTGGTGCAACACGACT
AATGACTGCAAATAACGACTTGGCAGTACGAGTATTTGACACTGAAAAATTTACAGTGCT
TGAACGCTTCTCTTTTCCTTGGTCTGTAAATCATACGTCTGTCAGTCCAGACGGTAAACTT
GTTGCAGTTCTTGGAGATAATGCAGATTGTTTGCTTGCAGACTGCAAGACTGGAAAGACC
GTGGGAACTCTAAGAGGTCATTTGGATTACTCTTTTGCAGCTGCCTGGCATCCAGATGGT
TATATTTTGGCTACGGGGAATCAGGACACCACCTGCAGACTTTGGGACGTTAGGAAGCT
TTCATCTTCCCTAGCTGTCTTAAAGGGTCGAATGGGTGCTATAAGATCGATACGGTTTTC
ATCAGATGGCCGCTTTATGGCCATGGCAGAGCCAGCCGATTTTGTGCATCTATATGATAC
TAGGCAGAATTATACTAAGAGCCAGGAAATTGATCTTTTTGGGGAGATTGCTGGAATTTC
ATTCAGCCCCGACACTGAAGCATTTTTTGTTGGGGTGGCTGATCGAACGTATGGAAGTC
TTCTTGAGTTCAACCGTCGACGAATGAATTACTATCTTGATTCCATCCTCTGAGTTTTGAA
AGTTAATGGGAGTGGTGTTTTCTTGAAGTGAAATGGCATCATTCTGTTGAACCAATTCTT
GTATATTAGATATGTAACATGTATGAATGTCCATAGAGCAGAGCTTTGCTCAATCAGAGG
CTTCAAAACCCAAATTCCAAGGTCTATCGAAGCCTCTTTGTTTATAAATGGTGTTGTGGTG
GATACAGCTTTCTGGTTCACGCTGTAGATTTTTTAATGTATGTAACAATTTCAGCGGATAT
AAAGTCTTCCAACTTGTAAACCAAAAAAAAAA

SEQ ID NO: 193
GGAGGATTTGGCACACGATCACCTGATCATTTTCCCAACCCTAAAATTCTTATAGCTAGG
GTTTATCCGAATAAAACCCTAAACACGGCGCGCGGTCTTCTTCTTTTATTCGGGACTTTTT
CGCAGCAGAGCATTACGATAACGAAACACGAGCAGATCGGCAGCGGCCACCCGCCGAA
GAAGAAGAAAACCCAGTAGTGCGATCAGAGTCTGAGCAGCCATGGCGGAAGCACTGGT
TCTGCGCGGCACAATGGAGGGCCACACCGATGCGGTGACGGCCATCGCGACCCCGATT
GACAACTCGGACATGATCGTCTCTTCATCCCGCGACAAGTCGATCCTCCTGTGGAACCT
AACCAAGGAGCCCGAGAAGTATGGTGTCCCACGGCGCCGCCTGACTGGCCACTCACAC
TTCGTCCAGGACGTGGTCATCTCCTCCGACGGCCAGTTTGCCCTCTCTGGCTCCTGGGA
CAGCGAGCTCCGCCTCTGGGATCTCAACACAGGCCTCACCACCAGGCGCTTCGTCGGC
CACACCAAGGACGTCCTGAGCGTTGCGTTCTCCATCGATAACCGACAGATCGTTTCCGC
CTCCCGTGACCGCACCATCAAGCTCTGGAACACTCTTGGCGAGTGCAAATACACCATCC
AGCCGGATGCAGAGGGCCACTCCAACTGGATCTCCTGCGTCCGCTTTTCCCCTTCGGC
CACCAACCCCACCATCGTATCGTGCTCATGGGACCGCACGGTAAAGGTATGGAATTTGA
CAAACTGTAAGCTCAGGAATACGCTCGTCGGACACGGAGGCTATGTGAACACGGCCGC
CGTTTCGCCCGACGGTTCGCTTTGCGCCAGTGGCGGCAAAGACGGGGTTACCATGCTC
TGGGATTTGGCGGAGGGGAAGCGGCTGTACTCGCTGGATGCCGGGGATATTATTTATG
CTCTGTGCTTTAGCCCCAACAGATACTGGCTCTGTGCTGCCACTCAGCAATGTGTGAAG
ATCTGGGATCTGGAGAGCAAGAGCATTGTGGCCGACCTTCGGCCCGATTTCATTCCCAA
CAAGAAGGCTCAGATACCTTATTGTACCAGTTTGAGTTGGAGTGCAGACGGAAGCACAT
TGTTTTCTGGTTATACAGATGGAAAAATCAGGGTGTGGGGAATTGGACATGTTTAAGCTT
TAGAGGCAATGGTAGATTATGAAGTCAACACCAGGGAGTTTGACCGTTTGGGACATTAAT
TTTGTCGAGCATGTGTTCTTTTTTCACAAAAATTAGTATAATGCATTTATGATTTCAGTTTT
GATACTATGTTAGAATTTATTTTTCTGCACGGAGGGAAGGCACTCCAAGCTTTGGAAGCA
ATATTTAATCAATGAGCATATTCTATTTCATTAAAATGGATATTTAAGGATCTATTGTATGA
AGTACAATGTTTCAGTTTTGGCATTTTAAAAAAAAAA

SEQ ID NO: 194
GCATTTTTCATTTTATATCCTTCGAATCCGTCATTCGGCTGAATGATCAGACAAACCCTCG
CAGATCCTGCTCTGTTCTGAAGCATAAACCTCCGCCAAATCCATGCGTTGAATTTCCTGA
AAAAATTCGTGGCTATTCAGTGTCGAACTGCTATCTCCACAGTCAGATCCAGTGGCGATC
AAGGTGAATTAGGGCAGAAGAACGTGAGATCGTGTACAAATGGCTGCCATAAAATCCAC
GTCTCGATCCGCGTCGGTGGCGTTTGCCCCCGACGCGCCCCTGTTGGCAGCGGGTACG
ATGGCCGGAGCCATCGACTTGTCATTCAGCTCCTTAGCCAACTTAGAGATCTTCAAGTTG
GATTTCCAGTCCGACGATCCGGAGCTCCCCGTCGTCGGAGAATGCCCAAGCAACGAAC
GACTGAATCGGCTTTCGTGGGGGAGCGCCGGTGGCTCCTTCGGAATAATTGCCGGTGG

ACTAGTTGATGGAACCATCAACATCTGGAATCCAGCTACTCTTATCAATTCGGAGGACAA
TGGAGATGCACTTATTGCACGGCTTGAACAACACACAGGACCGGTTCGTGGCCTGGAGT
TCAATACAATTTCTACAAATTTGTTGGCCTCCGGAGCTGAGGATGGGGAGCTTTGTATAT
GGGATCTTGCAAATCCTACAGCACCAACCCATTTTCCACCATTGAAGGGTGTGGGATCA
GGTGCTCAAGGGGAAATTTCTTTTTTAGCTTGGAATCGGAAGGTTCAGCATATACTGGCT
TCCACTTCATATAGTGGAACAACAGTGGTTTGGGATTTAAGACGCCAGAAACCTATCATT
AGCTTTCCAGACGCAACTCGAAGACGATGCTCTGTTTTGCAATGGAACCCTGATGCTTCA
ACACAGCTTATTGTAGCGTCAGATGATGATAATTCACCCACTTTACGGGCTTGGGACTTG
CGAAACACAATTTCTCCATACAAGGAATTTGTAGGGCATAGCAGAGGAGTCATAGCTATG
TCGTGGTGCCCAAGCGACAGTTTGTTCTTGCTAACTTGTGCAAAAGATAATCGCACTCTT
TGCTGGGACACTGGATCTGGGGAGATAGTGTGTGAATTACCAGCTGGGGCAAATTGGAA
TTTTGATGTTCAGTGGTCGCCAAAGATACCTGGAATTTTATCAACATCTTCTTTTGATGGG
AAAATTGGCATTCACAATATCGAGGCCTGCAGTCGAAATGTTTCTGGTGAGGTTGAGTTT
GGTGGTGCTATAGTACGCGGGGGACCCTCAGCTCTCTTGAAGGCCCCCAAGTGGTTAG
AGAGACCTGCTGGTGTAAGTTTTGGCTTTGGCGGAAAGCTTGCTTCATTTCGACCCAGTA
CGGTGGCTCAGGCTGCCGATCATAGACATTCTGAGGTCTTTATACACAATCTAGTTACCG
AGGACAATCTGGTTATTCGATCAACAGAATTTGAAGCTGCAATAGCAGATGGTGAAAAAG
TATCATTGCGGGCTTTGTGTGATAGAAAAGCCGAAGAGTCTCAGTCTGACGAAGAGAAA
GAAACATGGAACTTCTTAAGAGTCATGTTTGAAGATGAAGGCACGGCTAGGACAAAGCT
ACTCGAACACCTTGGATTTAAAGTGCAGTCAGAAGAGAATGGGGATTTACAGGAAACTCA
TTCCAGCAAAATTGATGACATTGGTAGTGAAATAGGAAAAACTCTTACTCTTGATGATAAA
ACAGAAGAGGATGTGTTGCCTCAGCTGAAAGGTGGACAAGATGCTGCTATCCCTCAAGA
CAATGGAGAAGATTTTTTTGATAATTTACACAGTCCCAAAGAAGAGGTCTCTCTTTCTCAT
GTAGGAAATGATTTTGTTGGAGAAAAGGATAAGGATATGGTAGTAAATGGTGCAGAGATT
GAACATGAAACTGAAGACTTGACTGAATATTCAGACTGGAATGAAGCTATTCAACATTCC
TTAGTTGTTGGTGATTATAAGGGAGCTGTTCTACAATGCCTGTCAGCAAACAGGATGGCT
GATGCTCTTATTATTGCTCACCTTGGTGGGAATTCTTTGTGGGAGAAAACTCGTGATGAA
TATTTGAAGAAAGCTAAATCTTCCTATCTCAAGGTGGTTTCTGCCATGGTCAATAATGATC
TTACAGGCCTTGTCAACAGCAGGCCATTGAAATCTTGGAAAGAAACTCTTGCTATGCTTT
GTACATATTCACAGAGGGAGGAATGGACTGTTTTATGTGATATGCTTGCTTCTAGATTGA
TAGCAGCAGGAAATGTGATGGCAGCAACACTTTGCTACATATGTGCTGGGAATATTGAAA
AGACTGTGGAAATATGGTCTCGAAGTCTTAAGTATGATTATGACGGGAGGAGCTTTGTTG
ATCATTTGCAAGATGTGATGGAAAAGACTGTTGTTCTTGCTTTAGCAACTGGACAAAAAC
GGGTTAGTCCCTCTCTGTCAAAATTGGTTGAGAATTATGCAGAATTGCTAGCAAGTCAAG
GACTTCTAACCACAGCTATGGAGTATCTAAAGCTCCTGGGCACTGAAGAATCTTCTCATG
AGCTGTCTATTTTACGAGATAGACTATACCTCTCTGGGACAGATAATAAGGTGGAGGCTT
CATCTTTTCCTTTCGAAACAAGGCAGGATCTGACAGAATCCCAGTACAATATGCATCAAA
CAGGTTTTGGGGCACCTGAAACTCAAAAAAATTATCAGGAAAACGTGCATCAGGTTTTAC
CTAGTGGGTCATACACTGACAATTATCAGCCGACTGCTAATACTCATTACATAGCAGGCT
ACCAACCGGCACCACAACAGCAACCTTCCTTCCAAAACTATTTTACTCCTGCATCTTACC
AGCCAGCACCATCTCCAAATGTTTTCTATCCATCACAAGTTTCTCAAGCTGAGCAGAGTA
ATTTTGCTCCACCTGTGAATCAGCCACCAATGAAAACATTTGTTCCATCTACACCACCAAT
CTTGAGAAATGTAGATCAATATCAGACGCCTTCTTTGAACCCTCAACTCTATCAGGGAGT
TTCTTCAGCTACAGTTGAAACACATCCCTACCAAACTGGTGCACCTGCATCTGTTTCAGT
GGGGACAACCCCAGGGCAGCCATCTGTAGTTCCTAATTTCATGGTCCCGGGACCAGTTA
CTGCACCTACAGTTACTCCCCGAGGGTTCATGCCAGTTACAACACCAACTCAACACCCG
CTAGGTAGTGCAAATCCTCCAGTTCAACCCCAAAGTCCCCAGTCCAGTCAGGTGCAGTC
TGTGACAGCTGCAACCACACCACCTCCAACAATACAAAATGTTGATACATCAAATGTGGC
AGCTGAAATTAGACCTGTAATTGGTACATTAAGAAGGCTTTACGACGAAACATCTGAAGC
TCTTGGAGGGGCAAGGGCAAATCCAGCTAAACGGCGTGAAATAGAAGATAATTCCAGAA
AAATTGGGTCTTTGTTTGCAAAATTGAACTCTGGTGATATATCTTCCAATGCTGCATCTAA
ACTGGTGCATCTCTGCCAAGCTCTGGAGTCTAGGGACTATGCAACTGCTTTTCAAATACA
GGTGGGACTGACAACAAGTGACTGGGATGAATGCAGTTTCTGGCTTGCTGCCCTTAAGC
GAATGATCAAAGTAAAACAGAATATGAGATAAATTTCTGATGGTGCTACAGTTAGATTTTT
TCCAAGTGCCATTTTTACATAATATGTGATTAGTTATTCTGTTCATTATTCTGTTTTTCATT
CAATTTGACATTGGAGTTTCAAGGCATTCCAAGGATAGCATGTACACAAGTTGAATTTTTG
CTTGAGAGTTATTGCAATTTTTGTCTTTTTGAGGACATTGTACAAAAAAAAAAAAAAAAA

SEQ ID NO: 195

GCACAGGCTATCTTATTTTGAAATCATTTTGAAATTCGTACCGTGGCACGAGGGTGTTTC
AAACCTAGCGTCCCAGGCCAGCCTCCACACGTTGTGTAGCAAAAACAACATAGGTGTAA
GCGGTGACACGATGAGCTGTCTTGTTTCATAGGACTGCAAGGGCAAAGGATATCAGCGT
TGGAATCAAATGGCGGGAGCCGGAGACTCGCAATTACAAACGAGGGCTGCAAGGGCAA
AGGATATCAGCGTCTAGTGGAATTGGAATCAAATGGCGGGAGCCGCAGACTCGCAATTA
CAAACGCTATCGGAGAGAGACAGCACCCCCAACTTCAAGAATTTGCACACCAGAGAATA
TGCTGCTCACAAGAAGAAGGTGCATTCAGTGGCATGGAATTGTACTGGTACCAAGCTTG
CTTCTGGTTCTGTTGATCAAACAGCACGTGTTTGGAACATTGAGCCACATGGTCATAGTA
AGACCAAAGATTTGGAGTTGAAGGGGCATGCTGATAGTGTTGATCAATTGTGCTGGGAC
CCAAAGCATTCAGAGTTGCTTGCTACTGCATCAGGAGACAGGACAGTGCGTCTATGGGA
TGCTCGAAGCGGAAAATGTTCTCAGCAGGTTGAACTTAGTGGGGAGAACATCAACATCA
CGTTCAAACCCGATGGGACACATATTGCTGTGGGAAACCGGGATGATGAATTGACTATA
ATCGATGTCCGCAAATTCAAGCCTCTTCATAAGCGTAAATTTAGCTATGAGGTGAATGAA
ATAGCTTGGAATACAACTGGAGAGCTTTTCTTTCTCACAACAGGAAATGGCACTGTTGAA
GTACTGTCATATCCTTCTCTTCAAGTTCTTCATACACTTGTTGCTCACACTGCGGGTTGCT
ATTGCATTGCCATTGATCCTATAGGAAGGTACTTTGCTGTTGGAAGTGCAGATGCCCTTG
TAAGCCTTTGGGATCTTTCTGAGATGTTATGTGTTCGAACATTTACCAAACTCGAGTGGC
CAGTGCGGACAATCAGCTTCAATCATGATGGGCAATACATAGCTTCTGCCAGTGAGGAT
CTTTTTATTGATATTGCTGATGTTCAGACTGGACGGACTGTACATCAAATATCTTGTCGAG
CAGCCATGAATAGTGTTGAGTGGAATCCCAAATACAATCTCTTGGCGTTTGCTGGAGATG
ATAAAAATAAGTACATGCAAGATGAAGGTGTGTTCAGAGTTTTTGGTTTTGAGACTCCATA
AATGCTCTTTAAGTTGACTTGGCTCATGTGGTTATGCTTCATCAAGGACGGCATTTTGAG
GCTAATAATTATTGTCAGATTTACACCATAAAATTACTATGGAAGTTGGATCATTATCTATG
CCATAGTGGAGTAGAACTAGATTTTGGGGAACATTGTTTTTGACTCGGACATTGAAAAAG
TGACTGCAGATTGCACATGGCCTCTTCTGTACCTAAATCTTTTCAAATGGAGTCATTGGTT
GGTTATGTCCAAGCCCTGTTAGATGGGGCACTTTAAAATCAATGCAGGCTGCAAATGG

SEQ ID NO: 196
GATGGGGGGAAGTGGGTTCGTATTCAAATTGGTGAGCTAGCAGGCCCGCCATTTGTTGA
TACATAAGCAAGGTCCGAGGGTATCAAAACTGGAGCTACAGTGTTAAAAATCACTTTTTT
AGATATTTAAAACAGTTCGGCTTATATTGGACAGCGATCAAAGTCTGCTAGAAGAAGCGT
CGATTTAACCTCAGGAGGAAAGGGGGAGGAAATCATGGCGGCAACGTCTCCCGTTGGT
GCCGGCTCGGGACGGGAGCTCGCGAATCCTCCCACTGACGGCATTTCTAACTTGAGATT
TTCCAATCACAGCGACCATTTGCTCGTCTCCTCATGGGACAGAAAGGTGCGATTATATGA
TGCAAGCGCGAATTCGTTAAAGGGTCAATTCGTACACGGAGGTCCGGTTTTAGATTGCT
GTTTTCACGACGACGCATCGGGTTTCAGTGGAAGTGCTGACAACACTGTGAGAAGGTAT
GATTTCAGCACCAGGAAGGAGGATATATTGGGAAGGCATGAAGCTCCTGTTCGTTGTGT
TGAGTATTCATATGCAGCAGGGCAAGTGATAACGGGGAGTTGGGACAAGACATTGAAAT
GTTGGGACCCTAGAGGTGCAAGTGGGCAGGAGAAGACTCTTGTTGGTACATATTCACAA
CTGGAGCGCGTGTATTCTATGTCACTAGTTGGTCACCGGTTGGTGGTAGCCACAGCTGG
CAGGCATATAAATGTTTATGATCTCCGTAACATGTCACAACCTGAACAACGAAGGGAATC
GTCCTTGAAGTATCAAACTAGATGTGTGCGTTGCTATCCAAATGGAACAGGTTTTGCCCT
TAGTTCTGTGGAAGGTCGCGTGGCTATGGAGTTTTTTGATTTATCTGAAGCTGGTCAAGC
TAAAAAATATGCCTTTAAATGTCATCGCAAGTCAGAGGCTGGCAGAGATACTGTTTATCC
AGTAAACGCAATTGCCTTTCATCCCATCTATGGTACTTTCGCCACTGGGGGCTGTGATGG
CTATGTTAATGTATGGGATGGCAACAACAAAAAGAGGTTATACCAGTATTCAAAATATCCT
ACAAGCATTGCTGCTTTATCATTTAGTAGAGATGGTCGTTTGTTGGCTGTGGCATCAAGC
TACACATTTGAGGAGGGGGAAAAACCGCATGAACCAGATGCTGTTTTTGTACGAAGTGT
AAATGAAGCTGAAGTGAAACCAAAGCCCAAAGTTTATGCTGCACCCCCATGAGGCCCGT
CTACATTAATATTCTTTCGTAGAGTATGCTTTCTAGTACATCTTCCTTCTCTTCCTGATGTA
CTTGATTGTTACATTGTTACTGTGTTGGGTATATAAGATACGAGATTCTCTGCTTTTGTTG
GCTGCTATGGTGTTTCTTTTATAATTTAATACTCTATGGATTAATCTCTTGATTCTAGAATC
TAAACTACTACCTTGCGGACATGACTGAGCATCTCTCTAACAGCTTGGGTGGGACACCT
CACTTTTTTTAGGGGTTTTGTTTGTACTAAAT

SEQ ID NO: 197
CCGGACTAATTGACTTACGCCGCTAAACTCTATACTCACATTTATTCTGTTTGAAGTCATT
TGGCTGGGAAAATCGATCATCATCAAAAGCTTGCAGCAGTTGAGTCCTCGAGAATTCATG
GCTTCAGACGACGAAGAGGGGTTTAAAAACGAAGAGGCTCCGGGTGTTGTCGATGAAG
CAGAAGTACAGGAGGGTCTGCGGGCCTGTTTTCCTTTGTCGTTTGGTAAACAAGAGAAG

AAACAGGCCCCTTTGGAATCCATCCACTCCGCCACCAAAAGGCCAGAGGATCCACGGC
CCAGACGACAGCTTGGTCCTCCAAGACCTCCACCCTCTATTTTGGCAGAACAAGAGGAT
TCTGATAGATTTGTAGGCCCTCCAAGGCCGCCCCAATTCGTTAGAGACGACAATGACGA
TGGTGAGGCGGAGATCATGATAGGACCTCCAAGACCGCCTGCTCAGTATAGCGATGAC
CATGATAACGAGGAGACCATAGGTCCTCCAAAGCCTTCATATTTAGAAAAGGGAGAGGA
GACTGACCAAATGGTTGGACCTTCCAAACGCGGAAGCGACGACGAAACCAGCGGTGAT
TCTGATGATGGAGACGATGCTGTCGATTTCAGGGTGCCTTTAAGTAATGAGATTGTGCTC
AGAGGCCATACCAAGGTTGTTTCAGCCTTGGCTATTGACCAGACAGGATCACGGGTTCT
TACAGGTAGCTATGACTATTCTGTTCGTATGTATGATTTTCAAGGTATGACTTCTCAGCTG
AAGTCATTTAGACAACTAGAACCAGCTGAAGGTCATCAAGTTCGTAGTTTAAGCTGGAGC
CCAACTTCTGATCGATTTTTGTGTGTAACTGGATCAGCCCAGGCAAAGATTTTTGATCGT
GATGGCCTCACTTTAGGAGAGTTTGTCAAAGGTGACATGTATTTGCGTGACCTTAAGAAT
ACAAAGGGGCATATATCAGGCCTGACATGTGGAGAATGGCACCCTAAGGAGAAGCAAAC
AATATTGACATGCTCAGAGGATGGTTCTCTTCGCATTTGGGATGTAAATGACTTTAATACT
CAGAAGCAGGTCATTAAGCCTAAACTTGCGAAACCTGGAAGAGTACCAGTGACTGCTTG
TGCATGGGGTCGTGACGGGAAATGCATTGCAGGTGGTGTTGGTGATGGATCAATACAG
GTGTGGAATCTTAAACCAGGATGGGGAAGCAGACCAGATTTATATGTTGCCAAGGGACA
TGATGACGATATCACAGGTCTTCAGTTTTCTGCTGATGGAAATATCCTTCTGACCAGAAG
CACTGATGAAACACTCAAGGTTTGGGATCTACGGAAGGCTATAACTCCACTGCAAGTATT
CAGAGATCTTCCAAACAACTATGCCCAAACTAATGTAGCATTTAGTCCAGATGAGAGGTT
AATATTTACAGGGACCTCTGTTGAAAGGGATGGTAATTCTGGAGGCTTATTATGTTTCTAT
GATAGGCAAACACTGGAATTGGTGTTAAGAATTGGTGTGTCCCCAGTGCACAGTGTTGT
ACGTTGTACTTGGCATCCACGGCATAATCAGGTCTTTGCTACAGTTGGGGATAAAAAAGA
AGGTGGAGCACATATACTATATGATCCAGCTTTGAGTGAGAGGGGGGCACTTGTTTGTG
TTGCAAGAGCACCCAGGAAAAAGTCCTTGGATGATTTTGAAGCCAAACCTGTTATTCACA
ACCCACATGCCCTTCCCTTGTTTCGAGATGAACCAAGTCGTAAGAGGCAACGGGAAAAA
GCACGCATGGATCCCATGAAGTCACAGAGGCCAGATTTGCCAGTCACAGGTCCTGGTTT
TGGTGGGAGGGTTGGATCAACGAAGGGCAGTTTGTTGACACAATACCTTCTCAAGGAAG
GGGGGCTGATTAAGGAGACATGGATGGAGGAAGACCCAAGAGAGGCTATACTGAAATAT
GCTGATGTTGCTGCAAAAGATCCCAAGTTTATTGCACCTGCTTATGCACAGACACAGCCG
GAAACGGTTTTTGCTGAAACTGATTCTGAGGAGGAACAGAAATGATATATAACTGATGAT
GTGGGTCATATTCATCCGGGTCTACCAATTTCCAATTTCAGGGCAACATACTATTGAATA
GAGTGCAGTTGACAAGGCAAAAATAAGAGAGCTGGTCCACATTATCAAATTTGCATTATG
ACAAGGCCAAACAAGAGATCTTGTCTATAATGTTAAATTTGTATATTATCTAACTGATTCTT
GTAATGTCTGCAATTGTATACCAAACAGGGTTGTGCTAGTTTAACATTTTAACTTAATGTA
ATCATGTAAGCTTTAGAGAGGTGGCAAGACATTTTCGGTTTATATTATTTATAATTTGTTTA
TGAAATACTTAGTTTATTATATTATCAATCCTTTCTCTATTT

SEQ ID NO: 198
GCGAGGGTGCACGATATTTCCTATCGTCCTGTTACTCGATAAAGCATAAACCTTGCAAGC
TCTGAACTAGGTTGCTTTCCATCAGACTGAGCTATTTGAGTTACAGTGACCAGGAACAGC
AACAAGTGATATCTATAGTGTTGAAATTGTGCGTGGAGTGAGGGTTTAAGGGAACAGAAA
GGACAATGAAGGAGAGGGGGCAGAGTCATGCAGGGCAACCGTCGGTCGACGAAAGGT
ACACTCAATGGAAGTCACTGGTGCCCGTTCTCTACGATTGGCTTGCTAATCACAACCTCG
TCTGGCCTTCTCTTTCCTGCCGTTGGGGTCCTCAAATGCACCAAGCAACATACAAGAACA
GTCAACGCCTGTACCTTTCGGAACAGACTGATGGCACTGTGCCAAACACATTGGTAATT
GCTACCTGTGAAGTTGTCAAACCAAGGGTTGCAGCAGCTGAACATATATCTCAGTTTAAT
GAAGAGGCACGTTCTCCTTTTGTAAAGAAGTTCAAGACGATCATACACCCTGGTGAGGTT
AACAGAATCAGAGAATTGCCACAAAACAGCAAAATTGTCGCTACACATACAGATGGCCCA
GATGTCCTCATCTGGGATGTTGATACACAACCAAACCGTCAGGCTACTCTTGGGGCAGC
AGATTCTCGTCCAGACTTGGTTTTGACTGGACATAAAGATAATGCTGAGTTTGCACTTGC
GATGAGTCCCAGTGCACCATTTGTGCTTTCTGGAGGAAAGGACAAATGTGTTTTGCTTTG
GAGTATACAAGATCATATATCTGCTGCAACTGAGCCTTCCTCTGCAAAAGCATCAAAAAC
ACCATCTTCAGCACATGGAGAGAAAGTACCAAAGATCCCTTCGATAGGTCCTAGAGGTG
TTTATAAAGGGCATAAAGATACGGTTGAAGATGTGCAATTTTGTCCTTCAAATGCACAGG
AATTCTGTAGCGTGGGTGATGACTCTGCTCTTATATTGTGGGATGCTCGAAATGGCAATG
AACCAGTCATTAAGGTTGAGAAGGCACACAATGCAGATCTTCATTGTGTTGATTGGAACC
CCCATGATGAGAATCTAATCTTGACTGGGTCTGCTGATAACTCAGTGCGCATGTTTGACC
GCCGGAATCTTACATCCAGTGGGGTTGGCTCACCTGTCCATAAATTTGAAGGTCATTCTG
CTCCTGTGCTTTGTGTGCAGTGGTGTCCAGACAAAGCATCCGTGTTTGGAAGTGCAGCA

GAGGATAGTTACTTAAATGTGTGGGACTATGAGAAGGTTGGAAAGAATGTTGGTAAGAA
GACACCGCCTGGACTTTTCTTCCAGCATGCAGGGCATAGGGATAAGGTGGTTGACTTTC
ATTGGAACAGTTTTGACCCATGGACCATTGTCAGTGTATCTGACGATGGTGAGAGTACTG
GTGGAGGAGGAACTTTACAGATTTGGCGAATGAGTGATTTGATATACAGGCCTGAAGAT
GAGGTTTTGGCAGAACTCGAAAGATTTCGAGCTCATATTCTATCTTGTCAGAACAAGTAG
AGAATGTTACGACTACCTCAGTACTTCTTTTTTTATAATTGTCTATAAACTATGCAATGTAA
ATGTTTACATTGAGGTCATGCATGAGTGTTAATTACGCTTTCACTACTGTTCACTTTTTTCT
ATAGACAGTGGAAATGGAATGATGTTGATCTGGCCATGTATAGGAAATCTAATATAAGGA
TGTCCAAAGTTGTGATTGGCATCAAGTCCAATCCAAAGTTGTGAACAATATTATATCGGTA
TTAACCACTAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 199
CTGATCTTGAGTGTCGAAGAAGAAGATTAAATGGCGAGCAGTTTCGAGGCGTGAGGGCG
TAGGCTGAAATCTTTAGCCTGAAACCCTTGGCTGAAGTGTTCGTTCGACGAAGCTCTGC
GGCCGTTAAAGCAATCCTCCGGAAGTTGTAAGATTTTGTCGGAAATCCCGGTATGCAGG
GCTTGCTGTTTGATTGAATGTTTTGGAACGGACTTGGATAGTTAATGATAAAGTAAAACTG
GCATCATGTCATCTCTTAGTAGAGAATTGGTGTTCCTTATCCTACAGTTCTTGGATGAGG
AAAAATTCAAGGAGTCGGTGCATAAGCTTGAGCAAGAGTCGGGTTTCTTTTTCAATATGA
AATATTTTGATGAGAAGGCTCAGGCTGGTGAATGGGATGAAGTTGAGCGTTATCTCTCAG
GCTTTACAAAGGTTGATGATAACCGGTATTCTATGAAAATATTCTTCGAGATCAGGAAACA
GAAGTATTTGGAGGCCTTGGACAGACAAGACCGAGCTAAAGCGGTGGATATTTTGGTGA
AAGATCTGAAGGTATTCTCAACATTTAATGAAGAACTATACAAGGAGATAACTCAGCTTCT
AACACTGGATAACTTCAGAGAAAATGAACAACTGTCCAAATACGGTGATACAAAGTCTGC
CAGAACCATTATGATGTCGGAACTTAAAAAATTGATAGAGGCCAATCCCTTATTTCGAGA
AAAGCTCATTTACCCAAACCTAAAAGCATCAAGGTTACGGACATTAATTAATCAAAGCTTG
AACTGGCAACACCAACTTTGCAAAAATCCTAGACCAAACCCAGACATAAAGACATTGTTC
ACTGATCATGCATGTGGACCTCCTAATGGAGCCCGGACACCTACGCAGCCTACGGCTTC
TCTGGGGGTGCTACCCAAAGCAACCACATTTACACCAATTGGACCCCATGGGCCTTTTC
CATCATCTTCAACCGCTACTAGTGGCTTGGCAAGTTGGATGTCAAATCCCAACATGGTGA
CATCTCCCCAAGCTCCTGTTGCTGTGGGACCTAGTGTTCCAGTTCCACCTAATCAAGCTA
CTCTTCTAAAACGTCCCAGAACCCCTCCAGGCAGTTCATCTGTGGTGGATTATCAAACTG
CTGATTCTGAGCAATTAATCAAGCGCTTGCGTCCTGTATCCCAGTCCATTGATGAGGCAA
CCTATCCAGGTCCTACTCTGCGAGTACCATGGTCAACAGATGACCTTCCAAAGACACTA
GCTCGGGCACTTAATGAACCTTATCCTGTCACAAGCATAGATTTCCACCCCTCTCAACAA
ACATTTTTACTAGTGGGCACCAAAAATGGAGAAATAACATTATGGGAGGTTGGTTCAAGA
GAAAAGCTGGCTACAAGGTCCTTTAAAATTTGGGATAATGCAAATTGCTCCAATCATTTG
GAGGCTGCATTTGTCAAAGATTCTTCAGTTTCTATCAATCGAGTATTGTGGAGCCCTGAT
GGAACATTGATAGGTATTGCTTTTACAAAGCATCTTGTTCACACATACACTTTCCAAGGAC
TTGATCTGCGACAACACTTGGAGATTGATGCACATGTTGGTGGAGTGAATGACTTAGCAT
TTTCTCATCCAAATAAGCAACTTTGCGTTGTAACTTGCGGAGATGATAAGATGATTAAGGT
CTGGGATGCTGTAACTGGTCGCAAGCTTTATAACTTTGAGGGCCATGATGCACCTGTCTA
TTCTGTTTGTCCTCATCACAAGGAAAATATCCAGTTTATATTTTCGACTGCCGTTGATGGA
AAAATTAAAGCTTGGTTGTATGATCATTTGGGATCGAGAGTAGATTATGATGCTCCTGGG
CATTCCTGTACAACGATGATGTATAGTGCTGATGGGACAAGGTTATTTTCTTGTGGCACA
AGCAAAGAAGGAGAATCCTTCCTAGTTGAATGGAATGAGAGTGAAGGAGCAATTAAGAG
AACATATTCCGGGCTTAGGAAGAAGGGTTCAGGTGTTGTGCAGTTCGATACAACGCAGA
ACCATTTTTTAGCTGTGGGGGATGAACACCTGATAAAGTTTTGGGATATGGACAGTACCA
ATATGCTTACGAGTTGTGATGCTGAGGGTGGTCTACTGAACCTGCCTCGTCTAAGATTCA
ACAAGGAAGGATCTCTTCTTGCAGTGACCACAGTAAATGGAATCAAGATTCTTGCTAATG
CAGATGGACAGAAATTACTTAAAACAATGGAGAATAGGACCTTTGATTTGCCATCAAGGG
CACACATTGATGCACCTGCAATTAAGCCATGTTGTAACTGGTCATTCAGATGCCTGGCGG
TATTCCCATTGTCGCAGCAAGTGCAACTTCGAGTCCAGCCACTGGAAGAATGGAGCGCA
TTGAGAGAACCTCCTCAGCCAATACTGTTTCTGGAATTAATGGAGTTGATCCAGCACAGA
GTTCAGAGAAACTAAGGTTATCAGATGATTTATCTGAGAAAACTAAAATATGGAAGCTGA
CTGAAATTACTGATTCTATCCAATGTCGGTGTATAACATTGCCGGAGAATGCAGCAGAGC
CTGCAAGCAAGGTTTCACGGCTTCTATACACAAATTCTGGAGTCGGATTACTAGCTCTTG
GATCTAATGCAGTGCATAAATTGTGGAAATGGAACCGCAGTGAGCAGAATCCCAGTGGC
AAGGCAACTGCAAGTGTTCATCCACAGCGCTGGCAACCAACCAGTGGTCTTCTCATGAC
CAATGATATAACTGATATAAATCCAGAAGAGGCTGTTCCATGCATAGCTCTGTCAAAGAA
TGACTCATATGTCATGTCAGCATCAGGTGGAAAAGTATCTTTATTCAACATGATGACATTT

AAGGTTATGACTACTTTCATGCCCCCACCCCCAGCATCAACGTTCTTAGCATTCCATCCA
CAGGATAACAATATCATAGCTATTGGAATGGAAGATTCAACTATCCACATATATAATGTTA
GGGTAGATGAGGTTAAAACTAAGTTGAAAGGACATCAGAAGAGAATCACAGGTCTTGCC
TTTTCCAGTACACAGAACATTCTAGTGTCTTCAGGTGCAGATGCACAGCTGTGCGTATGG
AACACAGAGACATGGGAAAAGCGTAAGTCGAAGACTATCCAGATGCCAGTTGGAAAGAC
AGTTTCTGGTGATACACGAGTTCAATTTCATTCTGACCAGCTTCATATACTTGTTGTGCAT
GAAACACAACTTGCCATATATGATGCCTATAAATTAGAACGACAATACCAGTGGGTGCCA
CAAGATGCTCTTTCAGCACCAATATTGTATGCAACATATTCATGTAATAGGCAACTTATTT
ACGCGACATTCAGTGATGGTAATATTGGTGTATATGATGCTGAAATACTTAGACCAAGAT
GCCGCATAGCCCCAACCACCTATTTAAGTTCAGGGACTAGCAGTTCTACCTCTTTACCCT
TGGTTGTTGC

SEQ ID NO: 200
GGGGACTTCAAACAAGCCGTCTTTCACTGCATGCTCCCTTTCCCGCAAAACACAAACCAA
ACACGTTAAACTCTAACCCAAACCTTCTCGCAATTCTTACCAAATCCAGTTACTCTGAGCC
AAAACCCTTGGCGGGAAAACCCCAGTTAGGAGCTTCCGGCCATGGCGAAGGACGAAGA
AGAATTCCGCGGCGAGATGGAGGAGCGCCTGGTGAACGAAGAGTACAAAATCTGGAAG
AAAAATACGCCGTTTCTTTACGATCTGGTGATAACGCACGCCCTCGAATGGCCTTCACTC
ACTGTACAGTGGCTCCCGGACCGCGAAGAGCCCCCTGGAAAAGATTATTCCGTTCAGAA
AATGATACTGGGGACTCATACTTCCGACAACGAGCCCAACTATCTGATGCTGGCCCAAG
TTCAACTCCCGCTAGAGGATGCAGAGAATGACGCCAGGCAGTATGACGACGAGCGCGG
GGAGATTGGAGGGTTCGGCTGCGCCAATGGCAAGGTACAAGTAATACAGCAAATAAATC
ACGATGGAGAGGTCAATAGAGCCCGATACATGCCACAAAATCCTTTCATTATTGCCACGA
AAACAGTTAGTGCAGAAGTCTATGTGTTTGACTACAGCAAGCATCCTTCAAAGCCTCCTC
AAGATGGTGGATGTCATCCTGATCTCAGATTGAGGGGTCATAATACAGAAGGTTATGGTT
TATCATGGAGCCCTTTTAAGCATGGCCATCTTTTAAGTGGTTCAGACGATGCACAGATCT
GTTTGTGGGACATTAATGTACCTGCCAAAAACAAAGTGCTTGAGGCCCAACAAATATTTA
AGGTGCACGAGGGTGTTGTAGAAGATGTTGCATGGCATTTAAGGCATGAGTACCTTTTT
GGGTCTGTTGGAGATGACCGCCATTTGTTGATATGGGATTTGCGTACATCTGCAACTAAT
AAACCACTGCACTCAGTAGTAGCTCATCAAGGTGAGGTTAACTGTCTTGCATTCAATCCT
TTCAATGAGTGGGTACTGGCTACAGGATCCGCAGACAGAACGGTGAAACTTTTTGATCT
GCGCAAGATATCCAGTGCTTTGCATACCTTTTCCTGTCACAAGGAAGAGGTTTTCCAAAT
AGGCTGGAGCCCCAAAAATGAAACAATATTGGCTTCTTGTAGTGCAGACAGAAGGCTTAT
GGTGTGGGACCTCAGCAGGATTGACGAATTCCAAACACCAGAGGATGCTTTAGATGGAC
CACCTGAGTTGCTGTTTATTCATGGTGGACATACTAGTAAGATATCAGATTTCTCATGGAA
TCCATGTGAGGATTGGGTTATAGCTAGTGTAGCTGAAGATAACATTCTCCAAATCTGGCA
AATGGCTGAAAATATATACCATGACGAGGAGGACGATATGCCTCCTGAAGAAGTAGTGT
AACTTTTATCTAGCTAGAAGTTGTGAAATTAAGAGGGATGTGAGGATTGGGTTATAACTA
GTGTAGCTGAAGATAACATTCTCCAAATCTGGCAAATGGCTGAAAATGTATATCATGGTG
AGGGGGATATGGCTTCCAAAGAAACATTGTAAGAGCTAGCTAGAATTTGTGAAATTAAGA
GGTGTATTCACTTTCAGAGTTTCTCAACAAATGACATGGTTCTCATTCCATTTTCTTTTATA
TAATGAGAAGCAAAACTTGGCTTAAAAAAAAAA

SEQ ID NO: 201
AAATAAGACCACCATTTGTTACCCTCTGCGAATCACCGTTTATGCTACTAGGAAGCAGAT
CGATTGGGATAAATTCATCCTCTTGGCTCGCATACAGCTTCCCTTCCAACATTCATCTCG
ATGTAAATTACAGTCACAGAACTCATCAAGCAATATGTCTCCAGGAGTGAAGCAAACGGG
CAGTCAGAAGTTCGAATCCGGCCACCAAGATGTTGTCCATGACGTCACAATGGATTACTA
CGGGAAACGCATAGCGACCTGCTCTGCAGATCGGACTATAAAGCTTTTCGGCCTGAACG
CCTCCGATACCCCAAGCCTTCTGGCCTCGCTTACGGGTCACGAAGGCCCTGTCTGGCA
GGTCGCTTGGGCTCACCCAAAGTTCGGGTCCATGCTCGCTTCCTGCTCATATGACGGAC
GCGTAATAATCTGGCGAGAGGGTCAGCAGGAGAATGAATGGTCGCAGGTTCAAGTCTTC
AAAGAGCACGAGGCCTCTGTGAATTCGATCTCATGGGCACCCAATGAATTGGGACTCTG
CCTGGCCTGTGGTTCCTCTGACGGCTCCATAACCGTCTTTACATGCCGCGAGGATGGGT
CCTGGGACAAAACGAAGATAGATCAAGCTCATCAGGTGGGCGTGACGGCAGTTTCATGG
GCGCCGGCTTCCGCTCCCGGTTCCCTTGTGGGTCAGCCCTCAGATCCCATTCAGAAGCT
CGTTTCCGGTGGGTGTGATAATACAGCTAAGGTTTGGAAATTTTACAATGGTTCTTGGAA
GCTAGATTGTTTTCCTCCTCTTCAAATGCATACAGATTGGGTTCGGGATGTTGCTTGGGC
GCCCAATTTAGGGCTTCCCAAGAGTACTATTGCTAGTTGTTCTCAGGATGGGAAGGTTGT
TATATGGACACAGGGGAAAGAAGGGGATAAATGGGAAGGGAGAATTTTGAATGATTTTA

AGATCCCCGTCTGGAGGGTCAACTGGTCCTTGACTGGAAACATCTTAGCTGTCGCTGAT
GGGAACAACAGTGTGACTCTTTGGAAGGAGGCAGTAGACGGGGATTGGAATCAGGTGA
CCACAGTACAGTGACAAATTCCTAAGAGGGAGAATTCAACATTGGATAGCTGATGAAAAA
AGACATCGAAGGACTATCAGAATGTATGAATTGCATGCAAGAGCAAGAATGAATCAGGC
ATTAAAGCGGGAATCATTTATGACTTGGCAACCTGAAAATTCTATTAAATTCACAAAAGTC
ATGACTTGGATGTTTAAAGAAAACAGTTTTAAGTAAGCTTTGGTCCTCTCTTGTGAGAATG
GTATGGTGGTACCTTTGGCTCAAGCAAAACAGTTTTATTTTTAAAGAGGGTAAGAACCAA
TCCTGGTATCAGCTAAAGCCCTTTTAAATTAAGGCTTTTATACACGATGACTCAGCTGCC
AACCGAAATCCAACTGCTTAATCAACTTTTAGCTCACCCGGAGGATCAATAAAGAGTCTC
TTCCAACCAATTACAGTTCCTCCCCTTTTGATGAAAGGTTTCCTTGTTTTTTTTGATGATCG
TTCCAAGGTAGTCCACATCCCCATTATGCTACTAGAGTGGTTCTTACTTGAGTGGGATGT
GTTTTTTTTTGTAAACAAAGTTGAGTAAGTATTTCATTAAAATCTGTGCCCAGATGCAACA
AATACCCGAGAAATCAGATAAAAAAAAAA

SEQ ID NO: 202
TTTAGCTCCACTTGAAACCTCGACCAAATATAATTTTAGTTTTAGATCTATACTTGCGCAT
CTCATTCCATCTTCTCAACAAGGATACAGGCCTCGACCAAATATAAATTTAGTTTTAGATC
TATACTGCGCATCTAATTCCATCTTCTCAATAAGGATACAAGGATACAGGCCGACGATGG
ACGGTTGAATAGATAAATTTCTATAGCTCGAAATCATCTGCTTGTTGTCGGTTGAATAGAT
AAATTTCTATAGCTCGAAATCATCTGCTTGTTTCTCCTGCGAGCTTGCCTCTTTTTGATCC
ACCTCAGCTGGTCCAGAGAAAAACACGAGATCCGTGACAAGGGAAGGGCAGCAATTGC
AAACTGGATCGAATTTTTACTGCCGAAGCTGCTTTCTTCCGGATTTGCGGATAAGGTTCT
AGGGCTTCCATCGCTTGGCAAATACGAAGTGGCTTGGTCTTCCAGTTTGTTCCTTCTCCT
CGGGCTGGCAGAAGAAGAAGGGCGACAACAAATATACAGAACTCATCGATCAATATGTC
TTCAGGAGTGAAGCAAACGGGGAGTCAGAAGTTCGAATCCGGGCACCAAGATGTTGTCC
ATGACGTCACAATGGATTACTACGGGAAAAGGATAGCGACGTGCTCCGCAGATCGGACC
ATAAAGCTTTTCGGTATGAACACCTCAGATACCCCAACCCTTCTGGCCTCGCTTACTGGT
CATGAAGGCCCTGTCTGGCAGGTTGCCTGGGCTCACCCAAAGTTCGGATCCATGCTTGC
TTCCTGCTCGTATGATCGGCGCGTAATCATCTGGCGAGAGGGTCAGCAGGAGAACGAG
TGGTCGCAGGTTCAAGTCTTCAAAGAGCACGAGGCCTCCGTGAATTCGATCTCGTGGGC
ACCCCACGAACTGGGACTCTGCCTCGCCTGCGGTTCTTCTGACGGCTCCATAACCGTCT
TCACAGGCCGCGAGGATGGATCCTGGGACAAAACGAAGATAGATCAAGCTCACCAGGT
GGGCGTGACGGCAGTTTCATGGGCGCCCGCTTCCGCTCCCGGTTCCCTTGTGGGCCAG
CCCTCAGATCCCGTTCAGAAGCTTGTGTCCGGTGGGTGTGATAACACAGCTAAGGTTTG
GAAATTTTACAATGGTTCTTGGAAGCTAGATTGTTTTCCTCCTCTGCAAATGCATACGGAT
TGGGTTCGGGATGTAGCTTGGGCACCCAATTTAGGGCTTCCCAAGAGCACTATCGCTAG
TTGTTCTCAGGATGGGAGGGTTGTTATATGGACACAGGGTAAAGAAGGGGATAAATGGG
AAGGGAAAATTTTGAATGATTTTAAGACCCCTGTCTGGAGGATCAGCTGGTCCTTGACTG
GAAACATCTTAGCTGTTGCTGATGGGAACAACAATGTGACTCTTTGGAAGGAGGCAGTG
GACGGTGAGTGGAATCAGGTGACCACAGTACAGTGACAAATGTAAGCTTCTTTTTACGAT
TTTGATTTCTTGACGTTTTAATATGGTATGGTATTAAATTTGGAAGGCCTATTCGATTGTTT
GCAAAATAATAAGTTTGTCTCGAAATTGGGTTATCCTATCCAACTTGTTGTGTCTATTGTT
CTAAAGTTTGTCTCGAAATTGGGTTATCCTATCCAACTTCTTGTGTCTA

SEQ ID NO: 203
GTCTTCTTCATGGTAGAGATAGTGTCCAACACTAAAAATAATCCAATTTTCTGGTTCCAAC
CTCAGGATATTTTACTTCAGTCTTTTGTCAGAATATTGTCTTCTCCTCGTGAAAACAGTGT
CTACCACTATAAAATCTCCAGTATTCTCTGTGCCTACTATAAATTTCCAGTCTTCTCTGTT
CTGGCTTTGTTGATATGTTGCTTAAATCCTCAAGTGTCAAAGGATTTTGCATAGTGTTTAG
CAGCATTTCCACAAGAAGAAGAAGAGGTGTTTCAATAAGATGAAGAAACGGTCACGACC
CAGCAATGGGCACCTATCGACGGCAGCGAAGAATAAAAGCCGAAAAACTGCCCCTATAA
CAAAAGACCCGTTCTTTGACTCCGCTCATAATCGAAACAAAAGTAAAGGCAAAGGCAAAA
GCAGAGGCAAAGGAGAAGAGATTTTCAGCAGTGATGAAGACGACGACGCTATTGGCAG
AGATGCTCCAGCAGAGGAAGAAGAGGAAATTGCCGAGGAAGAAAGGGAAACCGCAGAT
GAAAAGCGCCTCAGGGTGGCTAAGGCCTATCTGGATAAGATCAGGGCCATAACTAAAGC
CAATGAAGAAGATAACGAGGAGGAGGCGGGCGAAGACGAGGAAACTGAGGCAGAGCG
AAGGGGTAAAAGGGATTCTCTTGTGGCCGAGATTCTTCAGCAGGAACAGCTCGAAGAGA
GTGGGCGAGTGCAGCGCCAGCTAGCCTCCAGAGTTGTGACACCATCAAAGCTGGTAGA
GTGTCGAGTTGTTAAAAGGCACAAGCAGTCTGTCACAGCTGTGGCGCTAACAGAAGATG
ATTTGAGAGGATTTTCAGCATCAAAAGATGGCACTATTATTCACTGGGATGTTGAAACTG

GTGCAAGTGAAAAATATGAATGGCCTAGCCAAGCAGTATCTGTTTCAAGTTCGAATGAAG
TCTCCAAAACACAGAAGGGCAAGGGTTCAAAGAAACAGGGTAGCAAACATGTTTTATCAA
TGGCTGTCAGTTCTGATGGCCGTTATTTGGCAACTGGGGGTTTAGATCGTTATATTCATT
TATGGGATACTCGAACACAAAAACATATTCAGGCTTTTCGGGGTCATAGAGGAGCTGTGT
CTTGTTTAGCCTTTCGTCAAGGCACACAGCAACTTATATCAGGATCTTTTGATCGTACAAT
CAAGTTGTGGAGTGCAGAGGATAGGGCCTATATGGACACTCTTTACGGTCACCAAAGTG
AAATCCTTGCAGTTGATTGTTTGCGAAAAGAACGAGTTCTGTCTGTTGGACGCGATCACA
CTTTGAGACTATGGAAGGTCCCTGAGGAGACGCAGCTGGTCTTCAGAGGGCATGCAGC
ATCTTTAGAATGTTGCTGTTTCATCAATAATGAAGACTTCCTATCTGGCTCTGATGATGGA
AGCATTGAGCTTTGGAGTATGTTGAGAAAGAAGCCAGTTTTTATGGCAAAAAATGCACAT
GGGCATGCTATTGTAGAAAATCTTTCAGAAGATACGAGTACTAGGGAAGAGCCAGATGA
GGAAGTGACAACAAGACAATTACCTAATGGAAATAGTATTGGGAATGGTATGACAAATCA
AATGGGAATCACCCCTTCCGTAGAGTCATGGGTTGGAGCTGTCACAGTGTGCAGAGGAA
CTGATCTTGCTGCATCAGGAGCTGGTAATGGGGTGGTACGGTTATGGGCTATTGAAAAT
AGCAGTAAAAGCCTTAGAGCTCTACATGACATTCCTCTGACTGGATTTGTCAATTCACTTA
CTTTTGCACGGTCTGGACGCTTTCTTATTGCTGGAGTTGGCCAGGAACCTCGACTGGGT
AGATGGGGCCGTATTCAGGCTGCTCGCAATGGTGTGACTTTATGCCCAATTGAGCTTTC
ATGATGGATTTTGGGGGCAAATCAAGATGATGGTATCTGTTTGCATTTCCAGCACAGTTC
TGATTCTTTACTAGGGTAACTGTCACATTCAACAGAGAATCCCAGAATGGAAAGCCTCGC
ATACTTAATTATTGTTTTGATGCCTTTTCTTATAACCTGTACGATTGCCGATATATCACCAA
TTTTGCTGATTTTAATCTGAGTTTGTAGTGCGTACAAAAGGCTGCTATAAAAATATTGAAT
TTTTAGATTCCAAGATCAAAAAAAAAA

SEQ ID NO: 204
GGTGAATATGGCAAAACATACGTTAAAAAATCGAGGGGGAAGATGGATGGATGTTAAAC
ATTTGAACTTTCAGACAAACAGCCTTATAAACCCTTCGTGAGCTATATTAAGATGGATGTA
TTCTACATCCCTGAATACCCAAACCTACTCAGGCGTATCAAGCGTACGATAAACCATCTC
TCAGTCACCTACCGTCCGAGGAGAACAGCGGATTTCTATCAGTAGAATAGTTCCATTGAC
AGACGACATTTGCGACTCGGAGCACAGGGCTAGAGCAATTTCAAGGTTTTAAGCAGTTTT
CTGGAGGCTCGGATTTGGCCAGCAATACTTGCAATACTCAATTGGAGGCAAAGCTAAGG
CAGGAGCCCAGGGAGACTTGTACATGAAAAACTTCGCAGCCTGCGGTTCTGGGAGATTT
GAAGGTTGAGCAGTTTTAGCTATTAAACATTTCTTCTAAAGCAAGCTTCGACAGATTCTGT
ACAGTGATAGTTTCTGCGGCCTTTTAAACGGGGCAGAATTGAGAGATTCTAGGTCGCCG
TAAGTGCGATTTTAGGCATTCGGCTGTGCTTCCTCGTATTGCGAAAAGTTAATTAACAGA
GAGTTTCTTCAATGGCTGCTACCTTCGGCACCATTAATACTGCCACAAGTCCTCACAATC
CCAACAAATCCTTCGAGATCGTTCAGCCCCCAAATGATAGTATCTCAAGCCTATCTTTCA
GTCCAAAAGCTAACTATTTGGTGGCTACATCCTGGGATAATCAGGTTAGATGCTGGGAG
GTCCTTCAAACTGGGGCTAGCATGCCAAAAGCAGCAATGTCACATGACCAACCGGTTCT
CTGCTCGACATGGAAGGATGATGGCACTGCTGTTTTCTCTGCTGGCTGTGATAAACAAG
CTAAAATGTGGCCACTGCTGACTGGGGGTCAGCCAGTGACTGTTGCCATGCATGATGCA
CCTATTAAGGACATTGCATGGATTCCAGAAATGAACTTGCTTGCAACTGGAAGCTGGGAC
AAGACACTTAAGTATTGGGATACACGACAGTCAAATCCAGTGCATACACAGCAATTACCA
GAGCGTTGTTTTGCATTGAGTGTTCGGCATCCTCTTATGGTTGTTGGGACTGCTGATAGG
AATCTTATAATTTTTAATTTGCAGAATCCTCAAACTGAATTTAAGAGAATTTCATCTCCTCT
AAAATACCAAACAAGATGTGTTGCAGCATTTCCAGACAAGCAAGGATTTTTGGTGGGTTC
CATAGAAGGAAGAGTTGGAGTGCATCATGTAGAAGAGGCACAACAGAGTAAGAACTTCA
CTTTCAAGTGTCACAGGGATTCTAATGACATATATGCAGTTAATTCATTGAACTTCCATCC
GGTTCATCAAACCTTTGCTACAGCAGGATCTGATGGAGCATTCAATTTCTGGGATAAGGA
CAGTAAGCAGAGGCTGAAGGCAATGGCAAGGTCTAATCAACCTATTCCTTGCAGCACAT
TTAATAGTGATGGTTCTCTATATGCTTATGCGGTGAGCTATGACTGGAGTAAGGGTGCAG
AGAATCATAATCCAGCCACTGCTAAACACCATATTCTTCTCCATGTTCCACAGGAAAGTG
AGATTAAAGGGAAACCCCGAGTTACAACAAGCGGGAGGAAATAAAGTCACTTATGTCTTC
TGTTGCAATGTTGAGTGCCCTATTCTCGGTGCATGGATTTTAGCTTTATTAGTTCAGGAAT
TTCATGGAGCCCCTGTTGTTTGTGTGATGTAACTTTTTGATATATGATGTAATATACGATG
TTATATTACAAGGTATTGATATTCAATCCAATTTCATATTCGGGTTCAATGTAGTGCCTCT
CATTTTAGGGTGATAGCATGAGTTTTTTTTTAGTGTTATGACATCGACATCTACGTAGAGT
TGCATGAACAAAAAAAAAA

SEQ ID NO: 205

GATAAATATATGACATTTGCTCTTCTGTCGAGAACCTGTGGACCGGGGTAGTAGCATATA
CATATGTAGGTTGAGAGCGAGAGGATAGCGAAATCCTTACCTATAGCCACCACAAGCAG
TGGAAGTAGCACTGGCAGTGAAACTAGAAGCATAAGTAGTTACGGAAGAAGCACCACAA
GCATCAAAGCTAGAAGCAGAGTACTCCGATTTCCATTTTTCCTGGAAAGGGCCATTGAGT
GTATGGTCGTAATGGATAAGGGGACACATCAGACTAACGAGGACGAGAGCGAATCAGA
GTTCATCGACGAGGACGATGTGATTGATGAAATTTCCATAGATGAGGAAGATCTTCCAGA
TGCTGATGTGGAAGGCGAAGATGTGCAAGAAGACAACAAAAGAAGTGAGCCAGATGAGA
ACTCAAGTAGTTTGGACGATGCTATACACACTTTCGAAGGTCATGAAGATACTCTGTTTG
CTGTAGCATGCAGTCCAGTTGATGCAACATGGGTTGCATCAGGTGGTGGAGATGACAAA
GCTTTTATGTGGAGAATAGGCCATGCAACGCCATTTTTTGAGCTAAAAGGGCATACCGAT
TCTGTTGTAGCCTTGTCTTTCAGTAATGATGGGCTTTTACTTGCATCTGGTGGTTTGGAT
GGAGTAGTTCGCATCTGGGATGCTTCTACGGGAAATCTCATACATGTACTAGATGGTCCT
GGAGGGGGGATTGAGTGGGTCAGATGGCATCCAAAAGGGCATTTAGTCTTGGCAGGAT
CAGAGGATTACAGTACTTGGATGTGGAATGCTGATCTTGGAAAGTGCCTTTCAGTATATA
CTGGTCATTGCGAATCTGTCACATGTGGCGATTTTACACCGGATGGAAAGGCTATCTGTA
CTGGGTCTGCAGATGGTTCCTTGCGGGTATGGAATCCACAGACACAAGAAAGTAAACTC
ACTGTAAAAGGGTATCCATACCACACAGAGGGTCTAACATGCTTGAGTATTAGTTCAGAT
TCCACATTGGTTGTTAGTGGCTCCACAGATGGCAGTGTTCACGTGGTTAACATAAAAAAT
GGAAAGGTGGTTGCTTCCTTAGTTGGTCATTCTGGATCAATAGAGTGTGTCAGGTTTTCT
CCTAGCTTGACTTGGGTTGCAACTGGTGGGATGGATAAAAAACTAATGATTTGGGAATTA
CAAAGCTCATCACTGCGATGCACTTGTCAGCATGAGGAAGGCGTGATGAGGTTATCATG
GTCATTGTCATCCCAGCACATAATAACTTCTTCCCTTGATGGAATTGTCCGCCTTTGGGA
TAGTCGGTCAGGGGTTTGTGAAAGAGTTTTCGAGGGTCACAATGATTCAATTCAGGATAT
GGTGGTGACAGTGGATCAACGGTTTATCCTTACAGGATCAGATGACACGACTGCGAAAG
TTTTTGAAATTGGGGCATTTTGATTTTATATAATGTTTCATTCAAAATTATTGCCATGTTTG
ATTCTGAGTCACAGCAGAAGATTAGAGGAGCTGAGACTGAGGTAATAAATTGCTTTCCAC
AAGTTAACATAGGTAACTATCGACTGAAGTGAACTGGGGGGCAGAAGCTAACTATGTTG
GTTTCTTTTTCTGTAAGGTTCTTTTAACAATAATAATGTTGTGTCTACAATTTTTTTGTAATT
ATCTTTTAGCAGTTGACATTAGTATTTGAGTAAACTTTAAAAAAAAAA

SEQ ID NO: 206
CCGACGTCATTATAGAAACCCTTCATTTTTAAGTAGTTACAGAGTGATGGCTGCTATCATA
TCGTTGTTCGCAAATCGGGGCACGGTTAGGTACATGAGCATCAGTTCATTATCATAATCG
CACGCTTCCAAAGTTCATCTCTGATTTTCTTCTTCTTACACAGGAAGCAAAATCTGAGATA
ATTGTAAAGACCATACAAGCTTGAATTTGGATTTGCCACGATGCCTGTATTTAGGACTGC
ATTCAACGGCTATGCTGTCAAGTTCAGTCCATTTGTGGAAACTCGTTTGGCTGTGGCAAC
GGCCCAGAATTTTGGGATAATTGGGAATGGTCGGCAGCATGTGCTAGAACTTACACCCA
ATGGGATCGTTGAGGTCTGTGCTTTTGATTCATCTGATGGGCTTTATGATTGTACATGGT
CCGAAGCTAATGAAAATTTAGTTGTTTCAGCCAGTGGTGATGGCAGTGTTAAAATTTGGG
ATATAGCACTTCCACCTGTAGCAAATCCTATTCGAAGCTTGGAAGAGCATGCACGTGAG
GTTTACTCTGTTGATTGGAATCTTGTTAGAAAGGATTGTTTTTTAAGTGCGTCCTGGGATG
ATACTATCAGACTTTGGACGATTGACAGGCCTCAGTCCATGCGTTTGTTCAAAGAACACA
CTTATTGTATATATGCAGCAGTTTGGAATCCAAGACATGCAGATGTTTTCGCCTCTGCTTC
AGGTGATTGCACTGTGAGGATTTGGGATGTCAGAGAACCCAATGCAACCATCATAATTCC
AGCTCATGAACATGAGATTCTTTCTTGTGATTGGAACAAGTATAATGATTGTATGCTGGTG
ACGGGGTCTGTGGATAAACTAATTAAAGTATGGGATATTAGGACCTACAGGACTCCAATG
ACAGTTTTGGAAGGACACACATATGCAATCCGGAGAGTTAAATTTTCACCTCACCAGGAA
AGCCTTATTGCATCATGTTCCTACGATATGACAACATGTATGTGGGATTATAGAGCTCCA
GAGGATGCTCTTCTAGCTCGATATGATCATCATACTGAGTTTGCTGTGGGAATTGATATA
AGTGTTCTTGTGGAGGGTCTGTTGGCAAGCACTGGATGGGATGAAACTGTCTATGTTTG
GCAGCATGGAATGGATCCTCGAGCTTGTTGAGACTGGGTGTACTAGCTTGGCCTTTCCC
TCTCAGATTACTGGTTTGGTCTTTTACTTCCTCTCCGAGGATGCAGGAACTGTTTTGCAT
CCATTTTAGATAGCCATTTACATTTTACTTATTATTGGACTTGTAAAGATTTTTGTACCCTT
GTAAATTGCAAATTTAATTTACATTTCTTTTGCATTAAAAAAAAAA

SEQ ID NO: 207
GATTTAATGATCTTCCTCCTCCTTCTCATATCAGCTGATCAAAATTCATCAGAAGAGGAGA
GAAAAGAGCTATCGGCATTAAATTTTCAGACAAAGTCTTGGGATTTAGATTCGACTACTCT
TCTTCTGATACTAGCTGATCCAAATCTCAGCAGAAAAGGAGAGAAAAGAGCCGTTGGAAT
TAATTTTTTCAGACGAAGTCCTTGTGATTCTATTTTGCAGAGACTGAACAACAGTGTTACT

CAGTGCAATGGATTCCCGAAACAGGCGATCGCGCTTGAATTTGCCTCCTGGAATGTCTC
CGAGCTCCTTGCACTTAGAAACAACAGCAGGCTCTCCTGGCCTCTCGAGAGTAAATTCC
AGTCCCAGTACCCCTTCCCCCAGTCGCACAACAACGTACAGTGACCGGTTCATTCCCAG
CAGAACAGGGTCCAGGCTCAACGGATTTGCACTTATCGACAAACAGCCTCAGCCTTTGC
CTTCGCCCACTCGCTCCGCTGCCGAGGGCCGTGATGATGCATCTTCATCCTCTGCCTCT
GCATACTCCACATTGCTGAGAAACGAGCTCTTTGGGGAAGATGTTGTTGGTCCAGCCAC
TCCTGCCACTCCGGAGAAGTCCACTGGACTCTATGGCGGTTCCAGGGATTCGATCAAGT
CACCTATGAGTCCAAGCAGGAATCTTTTCAGATTTAAGAATGATCATGGAGGGAACAGTC
CCGGTTCGCCTTACTCTGCCTCTACTGTTGGGAGTGAGGGGCTTTTCTCGTCCAATGTC
GGAACTCCACCCAAGCCGGCTCGGAAGATTACTCGTTCTCCTTACAAGGTTTTGGATGC
CCCTGCACTTCAAGATGATTTTTACTTGAATCTCGTGGATTGGTCATCCAATAATGTACTT
GCAGTAGGCTTAGGCACCTGTGTCTACCTATGGAGTGCATGCACCAGTAAGGTGACAAA
ATTGTGTGATCTAGGAGTCAATGATAGTGTCTGTTCAGTTGGATGGACACCACAGGGCA
CACATCTTGCTGTGGGTACTAATATTGGAGAGGTCCAGATCTGGGATACATCTCGCTGTA
AAAAAGTTCGGACCATGGGTGGTCACTGCACTAGAGCAGGAGCACTAGCTTGGAGTTCT
TACATCTTATCATCTGGTAGCAGAGACCGGAATATTCTTCATCGTGATATTCGTGTTCAG
GATGACTTCATAAGAAAGCTTGTTGGGCATAAGTCAGAGGTCTGTGGATTGAAGTGGTCT
TATGATGATCGAGAGCTGGCATCGGGTGGAAATGATAATCAGCTCCTAGTGTGGAATCA
ACAATCGGCACAACCCTTGTTGAGATTCAATGAGCATACTGCTGCTGTTAAGGCCATTGC
ATGGTCACCCCATCAACATGGAATTCTTGCATCTGGAGGTGGGACAGCAGACCGGTGTC
TTCGCTTCTGGAACACAGCAACAGATACACGTTTGAATTGTGTAGACACTGGTAGTCAGG
TTTGTAATCTTGTATGGTGCAAAAATGTCAATGAACTAGTTAGCACGCATGGATATTCTCA
GAACCAGATAATGGTCTGGCGATATCCATCAATGTCAAAGCTGGCTACTCTAACAGGCC
ATACACTTAGGGTTCTCTATCTTGCCATCTCACCTGATGGGCAGACGATTGTCACAGGTG
CAGGGGATGAAACACTGAGATTTTGGAGTATTTTTCCATCTCCAAAGTCACAGAGCGCC
GTTCATGATTCTGGATTATGGTCTCTAGGACGAACTCATATTCGATGATATTTTTATTAGA
GGAAACATATCTTTTCCCTCCCATGAGAAAATATTTCTTGTCCAAATCCATTTTATGCTGC
TAGGTAATCCAACTGGGAACCATAAATCATAGGAATCCCTGTGAGGTGAAGGCAGGAAT
AAATTATGTCCTGAACTTATGTGTCTTCCAACAAAAAACCATAAATCGAAGCATTATTTGT
AAGATCCCCTGTGAGTTGAAGGCAAGAGTAAATTATGTCCTGAAATTATTATTCTTTGTAC
CAAAAGTTTTCAATACAGCTTGAGGTTTTACTGTCTGTTAAATATGAAGAGCGAAGCCAC
AATAGGATTTTCCCTCATTTTTGTTTGGCAACATCAAGTCACACAATCTTTCGAGAATTTA
GGCAAGCTTCATCAAAGTTGATTTGTTTGATAGTTGTGTATTCTTTGGTTTCCAAATATTTT
TTGGGTTTTCTTGTAGGCAGAAGTGAGGCTGGTTGGAAATGTGATCTCTTTAAAGGTTGT
TTCTTTTCGATGGTTAGCTATGGGTTACTGATTTCAGATTGTAATAGTTTATTGCTGAGTG
AACATCTGCATCTCATTTTTCATGCTTGTCTTTATTCAGCTATTCTGCTGCAATTGCTGAA
ATATTTCAAGCTGAAAGTTATGATTCTGGCCAAGAAGTCTACTGAAAATTTGTATTGTGGA
AATATATATATACATCAGAGTCAGTTATTTAAAAAAAAAAA

SEQ ID NO: 208
GGTTTTCAGCCCAAACCGTCAAATAATAAATCCATCGATACATCATCCATCCATCCTGATT
CGGCTTCGGCTTCGACTCCCTTTTGTACCCTTTTCCCACCCGGCCATTCATTCACATTTT
CACCCGTTTCTACGCACTGTCTAGCCGCCTTTTAATTTCCATTTAATTAAAAATCCGCTGC
AAACTCTCAGTTCCAGGTATATCTGTACGTTGGCATCTCAAGCGCCGAAAATTGTACAGT
TTGAAAGAGGGCAGGGAAGCCCTTCCCTAAATAGAAATTTATTCTCTGGTAGAGAAGTAT
AGTTTTTGGTGTGCATTTTTTTCTTCTTTTAATTCTTATTTAGAAGAGGGTAAAAGGTGGC
AGTATGGAGAAGAAGAAGGTGGTGGTGCCAATCGTCTGCCATGGACATTCAAGGCCTAT
TGTGGATTTGTTTTACAGCCCCGTCACCCCCGACGGCCTCTTCCTTATCAGCGCTAGCA
AAGATTCTAGTACAATGTTGAGGAATGGTGAAACAGGGGACTGGATCGGGACATTTGAA
GGGCATAAGGGTGCTGTGTGGAGTTGTTGTCTTGATAATCGTGCTTTACGTGCAGCATC
AGGATCAGCTGATTTTAGCGCCAAAATATGGGATGCGTTGACAGGAGATGAGTTGCACT
GTTTTGTGCACAAGCACATTGTTCGAGCATGTGCTTTTTCTGAGAGCACAAGCCTTTTAC
TCACTGGAGGACATGAGAAAATACTCCGTATTTTTGATTTGAACCGTCCTGATGCACCTC
CAAAAGAAGTTGATAATTCTCCTGGTTCAATCAGGACGGTAGCATGGCTTCACAGTGACC
AGACTATATTGAGTTCTAATTCTGATGCCGGAGGTGTGAGGCTGTGGGATTTGAGGACA
GAAAAGATTGTTCGTGTTCTGGAGACAAAATCACCAGTCACAAGTGCTGAAGTGAGTCAA
GATGGACGGTACATTACAACTGCTGATGGCAATAGTGTAAAATTCTGGGATGCTAACCAC
TTTGGAATGGTGAAAAGCTACACAATGCCATGCATGGTGGAGTCTGCTTCATTAGAACCG
ACTATGGGCAACATGTTTGTGGCTGGTGGAGAGGATATGTGGGTTCGTCTTTTTGATTTT
CATACCGGAGAAGAAATTGCTTGTAACAAGGGACATCATGGCCCGGTTCATTGTGTGCG

TTTTGCTCCTGGTGGTGAATCATATTCATCTGGATCTGAAGATGGAACAATCAGAATATG
GCAGACCTTGAATATGAATTCTGAAGAGAATGAAAGTTATGGTGTAAATGGATTGAGTGG
CAAGGTCAGAGTGGGTGTTGATGATGTGGTACAGAAGGTTGAAGGATTCCAAATTACAG
CAGATGGCCACTTGAATGACAAACCAGAGAAGCCAAATCCATAATAATGCAGTAGTGTAT
ATCCATTGCTGGGAAGCCATTGATCTGTTTGTTGAAGTTTGCTTGCCTGAGGTGTACAAT
TTTTTGTCGCTGATCAAGTGGCAAATCAAACTGCACACAGGATCCTGCTTACTGTACATA
GTAGATTATGAGGTCATATCACGTGGTTCTTTTACATCTGCAGGTGGTGCTCCTTTTTTTC
AAGCTCCTGGAGATTGGTATGGTTGAAGGATGGCAGTAAGGGTCCAGTATTAATTAGATT
ACCTTAATGATTCAGTTCTTTATGTTATCACTGTCTTGACTTTTGGATTAGATGTTTAACTA
GGAAGGTAGTAAGGTAACTAAAATGAATGTTGTAAAAAAAAGGAAGAAAATGAAAAAACA
TAAGTTTGGCCCAGATTCGGTTTATCATAAAATCTGGCTGCATATAAGGTGTCAGTTCAA
ACACAGGCAGGTCAGAATAAATAAATTTTATTTATATAGTTTTTGTTAAAAAAAAAA

SEQ ID NO: 209
CTAAGGGATACATTAAAGGCTACTAATCGCCGACAGCATTGACCAGATAGACTCTAAAAC
ATTCCTTCATTTTGTTCTGCTTCGTAAACCACCGTCCATGACTCCAGGGATTTATGGCAC
AGACGTGTTGGTGGCTATGCAAAATTTTTAGTCCGGTATGTACTGTGGTCGGATTGAGGA
AGAATTTCTGGATATTTAGGGTTTATGAATTCGTGAAGCAGAGCGAATCAACATAACCAG
GCGACTGGTATTTAGAATAATATAAAACAAGGGTAGTGTTTCAATGGAAAGGTATTCCCA
AGGCACACAGAAAAAATCAGAAATCTACACATATGAAGCTCCCTGGCAAATATATGGCAT
GAACTGGTCTGTGAGGAAAGATAAGAAATTTCGTCTTGGGATTGGTAGCTTCCTGGAGG
AGTATAACAATAGAGTAGAAATCATTGAGCTGGATGAGGAATCTGGGGAATTCAAGAGTG
ATCCAAGGCTTGCTTTTGACCATCCATATCCCACTACCAAGATAATGTTTGTTCCTGATAA
AGAATGCCAGAGGCCTGACCTTCTGGCCACTACTGGCGACTATCTGAGAATCTGGCAAG
TTTGCGAGGACCGAGTAGAGCCCAAAAGCCTCCTCAATAATAATAAGAACAGTGAATTCT
GTGCACCCTTGACTTCTTTTGACTGGAATGATGCCGATCCCAAGCGAATTGGGACATCG
AGCATCGATACCACTTGTACTATCTGGGACATAGAAAAGGAAGTGGTGGATACACAGCT
GATTGCCCATGATAAGGAGGTCTATGATATTGCCTGGGGGGGAGGTAGGGGTTTTTGCTT
CTGTATCTGCAGATGGGTCTGTGAGAGTGTTTGACTTGAGGGATAAGGAGCATTCTACTA
TTATATATGAGAGCTCACAGCCGGAGACCCCATTGCTTCGATTAGGTTGGAACAAGCAA
GATCCGAGGTTTATTGCTACTATTTTGATGGATAGTTGTAAGGTGGTTATCTTGGATATTA
GGTTTCCAACCTTGCCTGTTGCAGAGCTCCAGAGGCATCAGGCTAGTGTGAATACTATT
GCATGGGCTCCTCACAGCCCTTGCCATATTTGCACTGCAGGAGATGATTCTCAGGCACT
CATATGGGAATTATCCTCCGTTAGTCAACCACTGGTGGAGGGTGGTGGTCTAGATCCAA
TACTAGCTTACACCGCAGCTGCTGAAATTAATCAATTGCAGTGGTCTTCCATGCAGCCTG
ACTGGGTTGCAATTGCGTTTTCTAACGAAGTACAGATACTAAGGGTTTGAAATTTATTGCT
TTGTAGTTTTTCATTCAAATGTTCTAGAATTTGTCTAAGCTAGCTACTGGTGTTTAACTGAT
ATGGAAAACTTTTGCCATTCTCTTATCTGGATAGTTCCATAATGGTAAAGTACATTATGCA
AAAAAAAAA

SEQ ID NO: 210
CCAAACCAACTTTCTTCTTTTCTTTGTTCTCTACCTGTTGAAGAGGATGACCCAGGCACT
GTCAATTGAACGGATTTTGGGCTTTGTGAAATCAGAACTCAAGTCTGCTCGTCCAACAAG
TACGATATAAAAGGGTTGTTCTATTCAATTAGATTCCAGTATTGGCCAGACCATCTACTCT
CTTTCACCTACCTTTTGTTCCCATGAACATCCCTCTTTCTCATCTCCATTCTCATCATTTAG
TCTGAAAACTTCATTTTCATCACACCCACATGCTAATTTACTTTTTTGGACCATACCCACTT
CATCTTCTGTCCTATGTCATTTAAGACGTCACTCAAATGTCTTTCAAGAGCATACCGGCAT
ATTATATCCGTCCTTCTATAAGGTCTCTGGTTTGATTATTTCTTGGAGATGTCGTATATCT
GGTTTCTAGCTAGAGGAGTTTAGGGTTCCCCTGTAACTCTCTAGGTACTCGCTTTCTGTA
GGTGGGTTTGAGGTCTACACATATTTCAATCATTGCGGATTCTCCTGCTTAAGATTTTGAT
CATAATAATTTTGGTGCCAAACACTCTCAAGATCAGTTTACCGTGGCGTTCATAATAAAAG
CCTATTAGTTCACCTCTTCCGTCCAAGGGATAAGACAAACCCTAATCATATCGGGTGTCG
GCTCATCCGGAAGTATGCAATCGGAAAATAATTTGGATGAGTCTTTGCACCTAAGGGAG
GTTCAAGAGTTGCAGGGGCATACTGACACTGTGTGGGCGGTGGCATGGAATCCTGTCA
CTGGAATTGATGGAGCTCCATCTATGCTGGCCTCTTGCAGTGGTGACAAGACTGTTAGG
ATTTGGGAGAATACACACACCCTCAATTCAACATCTCCCTCGTGGGCGTGTAAGGCTGTT
CTGGAAGAAACACATACAAGAACTGTTAGATCATGTGCATGGTCACCTAATGGGAAGTTA
CTTGCAACAGCAAGTTTTGATGCAACTACTGCAATCTGGGAAAATGTTGGTGGTGAATTT
GAATGCATTGCCTCCTTAGAGGGCCATGAGAATGAAGTCAAAAGTGTTTCTTGGAGCGC
GTCTGGTATGCTACTTGCCACATGTGGTCGTGATAAGTCTGTCTGGATTTGGGACGTTCA

ACCGGGGAATGAGTTTGAGTGTGTTTCAGTACTGCAAGGCCATACACAGGATGTCAAGA
TGGTCCAGTGGCACCCTAATCGTGACATTTTGGTCTCTGCTAGTTATGATAATTCTATCAA
GGTTTGGGCAGAGGATGGAGATGGTGATGACTGGGCATGCATGCAAACATTGGGCAATT
CTGTCAGTGGTCATACATCAACAGTGTGGGCAGTATCTTTCAATTCTTCCGGAGACAGAA
TGGTCTCCTGCAGCGACGACTTAACGTTGATGGTTTGGGATACAAGCATAAACCCAGCG
GAGAGAAGTGGAAATGCTGGGCCCTGGAAACATCTTTGCACCATTTCTGGATATCATGAT
CGAACAATATTCTCAGTTCACTGGTCAAGGAGCGGTCTTATTGCTAGTGGAGCGTCCGA
TGATTGTATCAGGCTCTTTAGTGAGAGCACTGATGATTCTGTGACACCGGTTGATGGCAC
ATCGTATAAGTTGATTTTAAAAAAGGAAAAGGCTCACTCAATGGATGTGAATTCAGTCCAA
TGGCATCCCTCAGAGCCTCAGCTACTGGCTTCAGCAAGTGATGATGGTCGTATTAAGAT
ATGGGAAGTTACTCGGATAAATGGACTAGCAAACAGCCATTGACATAGTATCATTCCACA
TGACATCTAAAGGATGGGAGATTGGTATATCTTACCTGCAGGGGAAGCAAATTTCTTTAT
TGTAGCTTCTTGTTGACTCTCAGCTTCTCTGTGGGTTTGCTTTTGGGGAGTACTTTAGTC
AATAAAAGTGAAGTGAATCATGATATAAAGGGTTTAAGTAAAGAGTTGAGAATAAGCAGA
TTTTAGTGAAATACAACTGGTATGGCATTTATTGTGGCTGATTTTCCCTGTATGGCCTCTG
AAGCAGATTTCATGACTAAATTTTCATTAGGGGTTTCTGTAGTTTAACATAAATCAACATC
AAGTCTGATTTGGAATTTTTGTTTTGTTCGTTAAAATTAATTTGATTTCTCTGGTATTGTAA
AAAAAAAAAAAAAAAA

SEQ ID NO: 211
TTTTTTTTTTGAACTAGCCACGACACAGAAATCTTCCGCTTCGTGTACGTCATAAAGCGGT
TCGGAAGGCCGAACCAGACGCAGATGCAGAGTGGAAGGCGATTCCAGGATTGATTCAA
GTCACAGAGAAGAAAACCGAGGAGCAGGCGAGGCAGGTTTTTTTGAATGGTAATGGGC
GGTTTTTGATGTCAGATTTGGTGCTGTGAAAACATAAGAGCCGGCATTGTGCATTTCATA
GTTTTTGGTGGTGGGGATTGGACTGAGAAGAGTTATGCGAAACCAGCGATATAGAAGTC
CAACACTGATTTCAGAGCTTTTCTCCGGTCACCAAAACCATGAAGCGGGCTTACAAATTG
CAGGAGTTTGTTGCGCATGCTTCCAATGTCAACTGTCTCAAGATTGGGAAGAAGTCTTCC
AGAGTTCTGGTGACGGGCGGGGAAGACCACAAAGTGAATATGTGGGCTATTGGAAAAC
CGAATGCCATTCTGAGTTTATCTGGTCATTCAAGTGCTGTGGAGTCTGTGACTTTTGATT
CTGCAGAAGCTTTAGTTGTCGCTGGAGCTGCTAGTGGTACAATAAAGCTATGGGATTTG
GAAGAAGCAAAAATTGTTCGGACACTCACTGGTCATAGGTCCAATTGTATATCAGTGGAT
TTCCATCCATTTGGGGAATTTTTCGCATCTGGCTCCTTGGATACAAACCTAAAAATCTGG
GATATTAGACGTAAGGGTTGCATTCACACTTACAAGGGGCACACTCGTGGTGTTAATTCA
ATCAGATTTAGTCCAGATGGTCGTTGGGTGGTGTCAGGTGGGGAGGACAATATTGTAAA
GTTATGGGATCTAACTGCTGGAAAGCTCATGCACGACTTCAAATGCCATGAGGGTCAGA
TACAGTGCATGGATTTCCATCCTCAAGAGTTTCTTCTTGCTACAGGCTCAGCAGACAGGA
CTGTGAAATTCTGGGACCTTGAGACTTTTGAGCTTATTGGATCAGCTGGTCCTGAGACAA
CTGGAGTTCGTGCCATGATTTTCAATCCGGATGGAAGGACTCTGTTAACTGGATTGCATG
AAAGTTTGAAGGTGTTTTCCTGGGAACCTTTGAGATGCTATGATGCAGTCGATGTTGGTT
GGTCTAAATTGGCTGACCTCAACATACACGAAGGAAAGCTTCTTGGTTGTTCATACAATC
AAAGTTGTGTTGGTGTATGGGTTGTGGACATTTCGCGGGTGGGGCCATATGCTGCTGGA
AATGTATCAAGAACAAATGGCCATAATGAAGCAAAATTGGCTTCCAGTGGTCATCCATCT
GTCCAGCAATTAGATAATAACTTAAAGACCAATATGGCGAGGCTTTCCTTGTCACACAGT
ACAGAGTCAGGAATCAAGGAACCAAAGACTACCACATCGTTAACTACCACTGAAGGTCTT
TCTAGCACACCTCAACGAGCTGGAATAGCCTTTTCTTCAAAGAATCTTCCTGCAAGTTCA
GGTCCACCGTCATATGTCTCGACTCCAAAGAAAAATAGTACATCAAGAGTGCAGCCTACA
ACAAATTTTCAAACCTTAAGTAGACCTGATATAGTGCCTGTCATTGTCCCTAGAAGCAATT
CATTAAGACCGGAAACAACATCAGATGTTAAGAAAGAAATGAACAATTTTGGAAGAGTGG
TTCCATCTACAGTATCAACCAAATCAACTGATGTGATTAAATCTGGCAGCAACAGAGATG
AATCTGACAAGATAGACTCCATAAATCAGAAGCGCATGACAGGCAATGACAAAACAGAC
CTAAACATTGCTAGGGCTGAGCAACACGTTTCCTCTAGACTTGACAATACAAACACTAGT
TCTGTTGTTTGTGATGGAAATCAACCAGCGGCAAGATGGATTGGTGCAGCCAAATTCAG
AAGAAATTCACCAGTAGATCCAGTTGTAAGCCCACATGATAGAAGTCCTACTTTTCCATG
GTCTGCAACTGATGATGGAGTTACATGTCAGCCAGATCGACAAGTTACTGCACCTGAATT
ATCAAAAAGAGTGGTAGAGCCTGGTCGTGCTCGTGCTCTGGTTGCAAGTTGGGAAACAC
GAGAGAAGGCTCTCACCGCAGACACACCTGTGCTGGTCAGTGGTCGCCCCCCCACAAG
TCCTGGAGTGGACATGAACTCATTCATCCCGAGAGGAAGCCATGGGACTTCGGAAAGTG
ACTTGACAGTTAGTGATGACAACAGTGCTATAGAAGAGCTCATGCAACAGCACAATGCAT
TTACAAGCATTCTTCAAGCTCGCTTGACTAAGTTGCAGGTAATAAGGAGATTTTGGCAAA
GGAATGACTTGAAAGGTGCTATTGATGCTACGGGAAAGATGGGAGATCATTCGGTATCT

GCTGACGTTATTAGCGTACTGATTGAGAGAAGTGAAATCTTCACGCTGGATATTTGCACA
GTCATACTTCCATTGCTTACTCGGTTGCTTCAGAGCGAGACTGACAGGCACCTTACTGTC
GCTATGGAAACCCTGCTTGTGCTAGTGAAGACTTTTGGTGATGTTATCCGGGCAACTATA
TCAGCAACCCCAACAATTGGAGTTGATCTCCAAGCAGAGCAAAGGCTTGAACGCTGTAA
TCTGTGTTATGTTGAACTGGAAAACATCAAACAGATTCTTGTTCCTTTAATCAGGAGAGGT
GGAGCTGTTGCCAAGTCTGCACAAGAGTTGAGTTTAGCTCTTCAAGAAGTGTGACCCAC
TCTATTTAGGTTGACATTTTTTTTTTGTAATAAATCTTCTTTGACGGGAGGCTGTTTGTATA
TTGGGAAACACTGCATTGTTTGCAATGATAGTGGCAGTGTGGTTGAATATGCACGGCATC
AATTGCAGTGGAATTCTTATCTCAGGTCATTGGAGTGTTTAACCATTCCTATGTTGTCAGC
CACATCTTGTCAGCCACACCTTGATTCTTAGATTAAGAAGTTACTAATTTGTAGATAAATT
CTAACGAAGGTGATGATAGCATACACGTAATGAAATCACAGAGAATGTCTACGCAGTTGA
CACTGCAACTAGTCTTGCGTGAACCTGTAGTTGCGGCATGCTTCAACCTTTGTTGTAATA
TTTAAAATATTTGAGTGGCTTAAGCACTCAATTTCCTGTTTGATCAAGTCAGTTGCTAGAT
TCTTCAGTTGCTTTTCTCCGTTCATACAATCAGCAATTGGATATTGATGAATTTCATTACA
GAGTTTGATAATCTGTGAGCCATCATATATTAATTGTACATTCATAAGCA

SEQ ID NO: 212
GTTGATTCTGAACTGAAGGTATACAAAGGCGAAGAATAACAAAGTAGGAGCGCCGTCGG
TGCTCCGTCAAGCCGGCGGTCATGTCCACTCTCGAAATCGAAGCTCGCGATGTTATAAA
AATTGTGCTACAATTTTGCAAAGAGAATTCTTTGCACCAAACATTTCAAACACTGCAAAAC
GAATGCCAGGTTTCTCTTAATACTGTGGACAGCCTTGAAACATTTGTAGCAGATATCAAT
AGCGGAAGATGGGATGTGATACTGCCACAAGTTGCACAGCTTAAGCTTCCTAGAAAGAA
GCTAGAAGACTTGTATGAACAGATTGTATTGGAGATGATTGAACTCCGAGAATTAGATAC
TGCAAGAGCTATTCTAAGGCAGACTCAAGCAATGGGTTTCATGAAGCAGGAGCAACCTG
AGCGCTATCTACGCCTGGAACACCTTCTAGTTCGAACATATTTTGATCCTCGTGAGGCTT
ATCACGAGTCATCAAAAGAGAAGAGGCGGTCACAAATAGCTCAAGCTTTGGCATCCGAA
GTAACTGTTGTGCCGCCTTCTCGATTAATGGCCTTGATTGGTCAATCTTTAAAGTGGCAA
CAACATCAAGGATTACTCCCTCCAGGTACACAGTTTGACCTCTTCAGGGGCACTGCAGC
AGTGAAGGCAGATGAGGAAGAAATGTATCCAACAACACTGGCACATACAATTAAATTTGG
TAAACAAAGCCACCCAGAATGTGCTCGCTTCTCTCCAGATGGTCAATACCTTGTCTCTTG
CTCAGTTGATGGATTCATAGAGGTCTGGGACTACATCAGTGGGAAACTTAAGAAGGATCT
TCAGTATCAAGCTGATGATTCCTTTATGATGCATGATGATGCTGTCCTCTGTGTTGACTTC
AGTCGAGATTCAGAGATGCTGGCTTCTGGTTCACAGGATGGGAAAATTAAGGTATGGCG
TATACGAACAGGTCAGTGTTTAAGGCGTCTAGAGCGTGCACATTCTCAGGGTGTTACAA
GCCTTTCCTTCTCTCGTGATGGCAGTCAGCTTTTAAGCACATCATTTGACAGTACTGCAA
GAATACATGGGCTCAAATCTGGGAAAGCATTAAAGGAATTTCGTGGTCATACATCTTATG
TAAATGATGCAATATTCACAAGTGATGGGGGTCGTGTTATTACTGCTTCTAGCGATTGTA
CTGTAAAGGTTTGGGATGTGAAGACAACTGATTGTATTCAAACATTTAAACCTCCACCTC
CTCTCAAGGGAGGTGATGTATCAGTGAATTCTGTCCACCTTTTTCCAAAGAACTCAGAGC
ACATTGTAGTCTGCAACAAGGCATCATCAATCTATATCATGACACTCCAAGGGCAGGTTG
TCAAAAGTTTCTCATCTGGTAAAAGGGAAGGAGGAGATTTTGTTGCTGCATGCATTTCTC
CAAAAGGAGAGTGGATCTATTGCGTTGGGGAAGACAGAAATATCTACTGCTTCAGTCAA
CAGTCTGGGAAGCTTGAACATCTTATGAAGGCACATGACAAGGACATAATTGGTGTAACA
CCTCATCCACATAGGAATTTGTTGGTAACATATAGTGAAGATAGCACAATGAAGATATGG
AAACCCTAATATGGATTTTATACTGCTATTGGATTTGTATAGAAATAATATTTTTTTGAATT
TTGATGGTAGCGTATGGTTGAAGGAAAACTTGGATATATCATGTAAACATTTTTCCCCCAT
AAAGGAATGTATATTTTTTTTATTTACTGCACTTTATATTTCTCTGACCCTCTCTCACACAC
ATGCATTGGCACTACACAATTACACAGAAATATAAGCACTCTGCTTTTACATATGTTTAAA
AAAAAAA

SEQ ID NO: 213
GCAACATAGTACTCGATGCCGTCGTTTTGGTTCCACACTGACCAACGAGCTAAGATTGAA
GTTCCGAAGTACGATTCTTATTTCACATATTCTTGCCGTCTTCAGGTTCAGCTCGCAAAG
GGTACCGAGGAGCTCCCAGATGCAGAATTGGAGAACACTTATCTATCTATCCGACTAAC
AGAGAAGTGGTATCTCTCGCTCTAAGCTTCCATCCTGAATTATCTCCGTTTCGGCTCTCC
CACCGAACTAGCGCTACGGGCTTGTATTTTAGTAAGTAATACCGAAAAGCGGCGGAAAG
GGTTTTGAATACGTAGAGAAGAGGAAGGGAGGGATAGAGACAGAGGTGCTCGCACCAG
GAAGGGAGGGATAGAGACAGGGAGGGATAGAGACAGATACAGATGGATATTGAACTGG
AGGATCAACCCTTTGATTTGGATTTCCATCCTTCTGCTCCTATTGTTGCAGTGGCACTTAT
TACAGGCCGCCTCCAGTTGTTTCGTTATGTAGACATTTCATCAGAGCCCGAAAGATTATG

GACTGTGACTGCTCATACTGAGTCCTGTCGAGCTGCTCGATTTATTAATGCTGGTAGTTC
TGTGTTAACGGCGTCTCCAGACTGCTCAATTCTTGCCACAAATGTTGAAACTGGACAGCC
AGTAGCCAGACTTGACAATGCTCACGGGGCTGCAATAAATTGCTTAACCAATTTAACAGA
GTCAACTATTGCATCTGGAGATGAGAATGGAATCATTAAGGTTTGGGATACACGACAGAA
TTCCTGTTGCAACAAGTTTAAGGCTCATGAAGACTATATTTCGGATATGGAGTTTGTTCCT
GATACTATGCAGCTACTTGGAACAAGTGGTGATGGGACCCTTTCTGTTTGCAATCTTCGT
AAAAATAAGGTCCATGCCCGATCTGAATTTTCAGAAGATGAACTGCTTTCAGTGGCCTTG
ATGAAGAATGGGAAGAAGGTAGTATGTGGTTCACAGGAAGGAGTTCTGTTGCTTTATTCA
TGGGGTTATTTTAAGGACTGCAGCGACCGCTTTGTTGGACATCCACACTCAGTTGATGCT
TTGCTAAAGTTAGACGAAGACACAGTGTTAACTGGATCAAGTGATGGAATCATCAGGGTT
GTCTCCATTCTTCCCAACAAAATGATTGGTGTTATAGGAGAGCACAGCAGTTATCCTATT
GAGCGCCTTGCATTTTCACATGATCGGAATGTTCTTGGCAGTGCTTCACATGATCAGATC
CTAAAGCTCTGGGACATACATTATCTTCATGAAGATGATGAGCCTGAAACTAACAAGCAA
GAAGCTGTGAATGACGAAAACGTAGACATGGATTTGGATGTTGATACTGAAAAAAGACCA
AGAGGATCCAAACGAAAGAAGAGAGCAGAGAAGGGCCAAACTTCATCTCAAAAGCAATC
ATCTGATTTTTTTGCGGATATATAACATTGACAGTGTTCAAGCTTATGTATTCCTATTATTT
AAAATTTTTATTCTCACGGGTGACCTGATTTCTAAAACATTGATCCCTTGTTTTCAATTATA
TGTTTGTAATATGTAATATTTTAACGATATCTTCCTCCCAATTTTGCAATAGTGAGAGTGG
GTATTCCTTTATATATACAGGGGCCAAATGCTTTTGCAAATGCCCTTCTGATCCATCGATT
CTAAATGTATGAATTTTTCCCACCGTATACTTTTTAATGAACCGCTTTTTCCTTGAGAGGC
TATGAATGCCTGTAGAACTAATCCTTTAAGTAATTGTATATATTGAATGAGCAGTATATTTT
CTTTTTAAAAAAAAAA

SEQ ID NO: 214
GCGATACGTAGGGTCATGGCAACTAATGAAATCTACAGGAAAAGCCACCAGAAAACCAC
GTCCCTTCGCCACTAAAACTCCCATGAAAACAATAAACGCAGAATTCCCATTACTCATAA
AAACCCAACACACATCACCAAGAAATTTGTATCAGGGACGCAATTCCGAAGGGCTCCAG
GGGTTTCAGGGCCGGATCCTCTCTAAATATGGATAGGATCCAGCAAATCCCCCACACCT
GCGTAGCTCGAAAGATTAATTTGCCCCTGGGAATGTCCAAAGAGTCTCTAGCTCTGAAC
CTCCCTGCTAATTTGGCGCCCACTATGTCTCCGCCGAGTATTACTTACAGCGACAGGTTC
ATCCCCAGCAGGAAGGCCTCCAATTTCGAGGAATTCGCACTGCCGGATAAAACGTCGCC
TTCGCCGAATTCGGCTGGGGGTCAGTCCTCGTCCACCAATGGCGAAGGGCGCGATGAT
GCCTGTGCGGCTTATTCGGCTCTGCTGAGGACCGAGCTCTTTCCAGCTACCCCCGATAA
GACCGAGGGTTGCAGGAGGCCCGTGATTGGGAGCCCCAGTGGTAATGTGTTCAGGTTC
AAGTCCCAGCAGTGTAAATCCCAGAGCCCTTTTTCTTTATGTCCTGTTGGGGAAGATGGG
GATCTTAGCGAAACAGGGGCGGTTGCCAGGAAGACTACGAGAAAAATCCCTCGTTCCCC
TTTCAAGGTCTTGGATGCTCCTGCTCTTCAAGATGATTTTTATCTAAATCTTGTGGATTGG
TCATCTCATAACATACTTGCAGTGGGGTTAAGTGCATGTGTTTACCTGTGGAGCGCTAGT
AGTAGTAAGGTGACAAAACTATGCGATTTAGGTTTGGACGACAATGTCTGTTCTGTTGCG
TGGACTCAGCGGGGGACATATCTTGCTGTTGGAACAAACAATGGTGGTGTTCAGATTTG
GGATGCGGCTCATTGTAAGCAGGTTCGGACCATGGAAGGTCATTGTACCAGAGTAGGGA
CTTTAGCTTGGAATTCTCACATATTGTCATCTGGTGGTAGAGATAGAAACATTCTTCAGC
GTGACATTCGAGCTCAAGATGACTTTGTCAGTAAATTTTCTGGCCACAAGTCAGAGGTTT
GTGGGCTGAAGTGGTCCTATGACAACCGGGAACTTGCATCGGGTGGAAATGATAACCAG
CTTTTTGTATGGAACCAACAATCTCAACAACCAGTGCTGAAATATAATGAGCACACAGCT
GCTGTAAAGGCCATTGCCTGGTCTCCCCATCAGCATGGTCTTCTTGCTTCTGGAGGTGG
AACAGCAGATCGCTGCATTCGTTTCTGGAACACAGCAACCAATACTTCCTTAAATTGTGT
GGATACTGGTAGTCAGGTTTGTAACCTTGTATGGTCTAAGAATGTCAATGAGCTTGTGAG
CACCCATGGTTATTCTCAGAACCAAATTATTGTGTGGCGATACCCAACTATGTCGAAACT
TGCGACTCTAACAGGCCACACCCTCAGAGTGCTTTATCTTGCAATTTCTCCTGATGGGCA
GACTATTGTAACGGGTGCAGGTGATGAAACATTGCGCTTTTGGAATGTTTTTCCATCTTC
CAAAACACAGCAGAACACTATTCGTGACATGGGAGTTTGGTCTTCGGGTCGAACTCATAT
TAGGTAACATCAAAACTTCCTCTTCTTAGAGGAAGATGTGAAGTGGCATTGAGCAGTTGT
TGGGGAATTTGGTGGGAAACCCCATACAAGTTGACTGTATTAAGTCTCAGCTGAACAAG
GATCTGGCGTGAGCAGAAGTTGTCAACAGGTTATCTACCTGTTGTAGCTTTTAAGTAGTG
CTTTCTTAAGAGACAACCAGTTGAGGCAGTGGAACAAACTACAGGCTGTGCAGATGCTT
CCCCTTGCCTGGCTGGACAAATGTATTGGTAAAAATGCTAGATCATCTTAGAAAATAGGC
ACGCTTAGTCAAACTCATCCTTGGTGGGGATGACTTTTTCAAAGGTTTGAAAGTGTCTTG
TATCTTTGTTAGGTTCAATGAATTAATTATTTTCTCTAACACTATATTTTCTGGTATGACCG
CTCTACATTGTATATTAACCCTTGCAAATGAGATCACATTTGCAGTCCAAACAATATCCTC

TGTCCATAAGGTTGTACTTGTACTTGAACATACTTTTTGTAGAACATACTACCAAGAGAAA
TATTCAGCAAATATCATTTGAGAGTTTTCCCTGCAAGTTTGGGGATTTCTCAATCAAAAAA
AAAA

SEQ ID NO: 215
AATACGGCATACACTTACGAGCATATAGCCATTCTCGCTTCCCGGGGTGTTGGGGCTAT
ATCAAATTCTCAAGCGCAATTGACTGGCTTTCAAAATCATGAATTTGACGTTTCCCTATCG
CGATTGGAGTCGCTGAAATCTTCTGGCTCTTTGAATACCCTTAGACATTAGACGTTTAATA
CACATCGTGAAAGGATCCTAGCCACTTTAAGAATTCGCAGGGTTCCAAGGAGTTCTGGA
TATTCAACCGTGCCAATAAAAAATGGCAGGAGGCCAAGGTGAAGGGGAGGAGAAAGTA
GATAAATTATCGATGGAACTCACAGAAGATGTCATGAAAAGCATGGAGATTGGTGCTGTC
TTTAAAGATTATAATGGGAAAATCAATTCGTTGGATTTTCACAGAACGAACAATTATTTGG
TTACGGCTAGTGATGATGAGGCTATTCGGCTTTTCGACACTGCAAGTGCAACATGGCAA
AAGACCAGTTATAGTAAAAAGTATGGTGTTGACTTGATTTGCTTCACCAACCATCAAACAT
CAGTACTGTACTCTTCGAAAAATGGCTGGGATGAATCATTGCGGCATCTTTCCCTCATGG
ACAATAAATATTTGCGTTATTTCAAAGGTCACCATGATAGGGTGGTTTCCCTTTGTATGTC
ACCAAAGGGTGAGTGCTTCATGTCAGGTTCCTTGGATCGCACTGTGCTTCTTTGGGATCT
ACGAATTGATAAGTGTCAGGGTTTGATACGAGTACGAGGACGACCTGCTGTGGCATATG
ATGAACAAGGGCTTGTCTTTGCTATTTCCAATGAAGGGGGTTTAATAAAAATGTTTGATG
CTCGCTTATATGATAAGGGACCATTTGATACCTTTGTTGTGGAAGGAGACAAATCCGAGG
CATCAGGGATAAAATTTAGCAATGATGGGAAGCTAATTCTTTTGTCTACCATGGACAGCA
ACATACATGTCTTGGATGCATATCAAGGGACAACGGTGCATAGTTTCAGTGTGGAAGCT
GTTCCCAACGGTGGTGAAGCTGTTCCCAATGGTGGAACTCTGGAGGCATCTTTCAGCCC
AGATGGCAAATTTGTGATTTCAGGTTCAGGGAATGGTAATATCCATGCTTGGAGTGTCAA
CTCTGGCAAAGAAGTTGCTTGTTGGACGACTGAAGGAGTCATTCCTGCAGTTGTAAAATG
GGCCCCAAGACGCTTAATGTTTGCTAGTGGATCATCTGTGCTGTCATTATGGGTACCAGA
TTTATCAAAGTTGGCTTCTCTTACTGGTTCAAATTCTAATAGTGCTTATTGATGACATTCA
GTGGAGCCCAAGAGAGACAGATCTGACATTCCTCCAGTGACCCTAGGATGACAACAGCT
CTTAATTTCTTTCTTGGATATATATTTTTCCTTACATCTAGTGTCTTCTGTCTTCTCTGCCT
TTTGTAATTTTATATTCACTGTGCTGGGATTATCCTCTCCCCTTTTTGACCCACTGTTGTG
TGTATTTGAATTTGTACGTATAATTTTTTTTGTTTTTGTTGTTTAATATAATTAGCTCTTAAT
CACAAAAAAAAAA

SEQ ID NO: 216
GCGAGGAAGAAGAAGCGTCGAGGGTTTAGTCAAAATCCCAAAACCTTGATCCAATAGAT
GCTCATATGCACCCAGATTGTACCCCCGGAATTCGAAACCCGCTGACGCAGATGTATAT
TTGGGAAAAATTGGCCTTGGGGATGCTGATCTGCTCGAGCCTTGAAAGACTTCTGTGAG
GATCTTCATTATCGGAAATCACCCGAGGACTGCTGGAGATCAGGTGTTTCTGCAATGCA
CCGTGTGGGCAGTACAGGGAATACTTCAAATTCGTCCCGTCCTAGAAGAGAAAAGCGGT
TAACTTATGTGTTAAATGATGCAAATGATTCCAGACATTGTTCAGGGATAAACTGCTTGGT
GATATCAAAGTTATCTTTGCTTGGTGGAAATGACTATTTGTTCAGTGGTAGTCGAGATGG
AACACTCAAACGATGGGAGTTGGCTGATGATTCTGCTGTCTGCAGTGCAACATTTGAATC
TCATGTTGACTGGGTAAATGATGCGGTCCTCACAGGCGAGACACTTGTTTCTTGTTCTTC
AGACACTACTCTCAAGACCTGGCGTCCTTTCTCTGATGGTGTCTGCACCAGAACTCTTCG
TCAACATTCTGATTATGTTACATGCCTTGCAGCGGCATCAAAAAATAGCAATATTGTTGCT
TCAGGAGGTCTTGGTCGTGAAGTGTTCATATGGGATATTGAAGCAGCAATGGCTCCAGT
CTCACGGACTAGTGAAGCAATGGATGATGATACTTCGAATGGAGTTTTGAGTTCTGGGAA
TTCTGTTCTATCTACAACTGTTCGTTCTACTAATGCCACCAACAGTGCTTCTTTACACACT
TCACAATTACAAGGCTACACTCCAATTGCGGCCAAAGGCCACAAAGAATCAGTTTATGCA
TTGGCTATGAACGATGTTGGTACCTTACTCGTATCTGGAGGAACTGAGAAGGTAGTGAG
GGTGTGGGATCCAAGAAGTGGGGCAAAGCAAATGAAGCTGCGGGGGCATACTGACAAT
GTGCGAGCACTCATTTTGGATTCTACTGGCAGGTTCTGTTTATCTGGGTCTTCTGATTCT
ATCATACGGCTCTGGGATCTTGGTCAGCAGCGCTGTGTACATTCATATGCTGTGCATACA
GACTCTGTTTGGGCACTTGCAAGTACGCCAAATTTTAGTCATGTATACAGTGGTGGGAGA
GATCTTTCTTTATACCTGACAGATTTGACTACGAGAGAGAGTTTACTGCTTTGTATGGAGA
AGCATCCTCTTCTACGGTTGACATTGCAGGATGACTCAATATGGGTTGCTACAACAGATT
CTTCTTTACATAGATGGCCAGCAGAAGGACAAAATCCGCCAAAGATGTTTCAAAGGGGT
GGGTCTTTCCTGGCTGGAAACCTATCCTTTACCAGGGCAAGGGCTTGTTTGGAAGGATC
AGCGCCTGTACCTGTGAACACACAACCTTCATTTGTTATACCAGGTTCTCCAGGAATTGT
ACAACATGAAATACTGAACAACAGGCGGCATGTCTTGACGAAGGATGCTGAAGGTACTG

TAAAGCTATGGGAAATTACTCGGGGAGCAGTGCTTGACGACTATGGAAAGGTTTCCTTT
GAAGAAAAGAAAGAGGAATTATTTGAAATGGTCAGTATTCCTGCCTGGTTCACAATGGAT
ACCAGACTTGGAAGCATGTCTGTGCATTTAGATACACCCCAATGCTTTACTGCTGAAATG
TATGCAGTTGACTTAAATGTTCCAGATGCACCAGAGGAACAGAAGATAAATCTTGCACAG
GAGACTCTCCGTGGACTTTTGGCACATTGGTTGTCGAGGAGACGACAACGGTTGGCAAC
TCAAGCTTCTGCTAATGGGGATTTTCCAGCAGGGCAAGAAAATGCTCTCAGGAATCATAT
TTCATCCAGAATAGATGTCCATGATGATGCCGAAACCCATATTGCTGGAATTCTTCCAGC
ATTCGATTTCTCAACTACGTCACCACCATCAATAATTACCGAAGGTTCTCAAGGAGGCCC
ATGGCGTAAAAAGATAACAGATTTAGATGGAACAGAGGATGAAAAGGATTTCCCATGGTG
GTGTCTAGAATGTGTACTGCACGGCCGACTTTCCCCTAGAGAAAGTTTAAAGTGTAGCTT
TTATTTGCATCCATACGAGGGAACAACTGTGCAAGTGCTTACCCAGGGAAAACTCAGTG
CACCCAGAATATTGCGCATACAGAAAGTTATAAATTATGTTTTAGAGAAAATGGTCCTTGA
TAGACCACTGGATTCTAGCAACTCAGAAACAACTTTTACTCCAGGATTAAGTGGAAACCA
ATCACATGCAGCAGTAGTTGGAGATGGTTCTCTTAGATCTGGTGCACGAGTCTGGCAGC
AGAAAGCAAAGCCATTGGTGGAAATTCTTTGTAACAATCAGGTGCTCTCACCTGATATGA
GTTTGGCAACGGTGCGAACATATATCTGGAAGAAACCGGATGACTTGTATCTCTACTACA
GACTAGTGCAAAATAGATGAATCTTCGCACATCCTTTTGAGATGTGTTTAACCTTGTAATA
AAGTAGTCTGATACAATGTTTTGCTTGCTTTCTGAATTGTGTCTTTTTCAGATGGCAATCT
TCATAACTTGTTTTCTTCATCATATGCAAAATCTGTTGGATAATTAAGAGGAACAAAATGA
AAGGTGTAGTATTCATGTATTCATTTCTGGTATTGATACTTTTTCTAAGTGGTGCCATGAG
GAGGTACCTCTTTTTCATTGTGTAGACTAACCATGGAAAAAAACAGAAGAGTTCACATGT
AAAGGCTGCCTCTAACAAGGTGTGGAATAAAATTTGTCTCCAAGCAAGGAGAAAATTCGA
GGGTTATCCTCATTCCTGAGAGCATACAGCGTTATCTTTGAGACGAGTCATCAATGATAA
TATCCTCGTAAAAGGTTACAGGATCTGTTGAAGAGAGTGAATTGATTATCACTGCATGTA
AAACATCTGTCCACAGAATGTTCTTGTGTTTTAATGTAAAATTGGGCACTTCTAAGTGTTC
TAAAAGTAGCTCATATATTGATTTTGATGTAAGCAAAAAAAAAA

SEQ ID NO: 217
CCGCAACTGTCACTGATCGAAGAAAAGAAGCAGCAGCAGCAGGAGAGCAAGATCAATGA
TGAAGGGCAAAACAATCCAGATGCAAGCTGCCCATCAAAATCATGATGGCGAAACATCA
GTAGCTTGCGTTCTATGGGATTGGCATGCCAAGCATCTCATCACTGCAGGAGCCGATAA
TACCATTCTCATCCATTCTTATCCTTCATCCTCATCCTCTAAACCCATCACTCTTCGCCAT
CACAAGAATGCAGTCACTGCCCTCGCCATAAATTCCAATGTGAGAAGTCTTGCCTCTGGA
TCCGTCGATCACTCTGTTAAGCTCTACTCTTATCCAGGGGGTGAATTTCAGAGCAATGTA
ACACGGTTTACTCTGCCAATACGATCTCTGGCCTTCAACAAGTCTGGTGAGTTACTAGCT
GCTGCAGGTGACGATGAAGGCATTAAGCTAATTAGCACCATAGATAATTCCATTGCAAGG
GTGCTGAAGGGCCACAATGGCCCCGTTACTAGTATTTCTTTCGATCCGAAAAACGAGTTC
TTGGCATCTTCAGACAGTGATGGCACTGTCATATATTGGGAGCTTTCAACTGGGAAACCT
GTACATACATTGAAAAAGATTGCACCCAATACGACCTCCAATCCGACGAGTTTAAATCAA
ATCAGCTGGCGTCCTGATGGTGAGATGCTGGCTGTCCCAGGCAGGAAGAGTGAGGTTT
CAATGTATGACAGGGATACTGCAGAGAAGCTTTTCAGCCTTAAAGGCGGGCATTCAGAT
ACTATTTGTTCTCTGGCCTGGTCGCCTAATGGGAAGTATATTGCCACAGCTGGGACTGAT
CGGCAGGTCATGGTTTGGGATGCCGATAGGAGGCAGGATATTGATAAGCAACGTTTTGA
CAACCCAATTTGTTCAGTTGCCTGGAAGCCCAGTGACAATGCCTTGGCGGTGATTGATG
TACTGGGCAGATTTGGGGTATGGGAATCACCTATTGCATCTCATATGAAGTCCCCTGCA
GACGGTGCAGAGCGCTACGATAACATGGAAGATGAAGAGCCCTTGATGGCTAGATATGA
AGAGGAATTGGAAGATAGTGTATCTGGAAGCCTGAATGAGATAATAAATGACGACGATG
ACGATGATGAAATGGGTAAGATTCCAAGGAAGATCTTACAGAAGAAGCCTTCTGTTAAAG
TTGAAAAAGGCAAAGAAGAGAGCAACGCTAAGGCTTTTAAGAGTGGTCAAGACTCCTTC
AAATTAAAATCAGCAATGCAAGAGGCGTTTCAGCCTGGTGCAACGCAGCGGCAATCTGG
AAAGCGGAATTTTCTTGCTTATAACATGCTTGGAAGTGTTATCACCTTTGATAATGATGGC
TTTTCTCATATTGAGGTTGACTTCCATGATATAGGGAAAGGTTGTCGTGTGCCCTCCATG
ACTGACTATTTTGGTTTCACCATGGCCTCCCTCAGTGAAAGTGGGAGCGTTTTTGGGAGT
CCGCAGAAGGGCGAGAAGAATCCTAGTACACTCATGTATCGACCTTTCAGTAGTTGGGC
GAACAATAGTGAGTGGTCCATGCGATTTCCAATGGGAGAGGAGGTGAAAGCTGTTGCTC
TTGGTTCAGGTTGGGTTGCAGCAGTGACAAGTCTTAATTTCCTTCGGGTTTTTTCAGAGG
GCGGCTTGCAGAAATTTGTTCTTTCTATGGATGGGCCAGTAGTCACTGCAGCAGGATAT
GAGAACCTCCTTGTTGTTGTATCACATGCTTCAAATCCTCTTTTATCTGGAGATCAGGTG
CTCAGCTTCACTGTGTATGACATTTCTCAAAAAACTTGCCCTCTCTCTGGTCGGCTTCCTT
TAAGCCCTGGCTCTCATCTCACATGGCTTGGATTCAGTGAAGAAGGCTTATTAAGTTCAT

ATGATTCTGAGGGAAATTTAAGAGTGTTCACCAATGACTACAATGGCTGCTGGGTGCCAA
TTTTCAGTGCTGCAAGAGAGAGAAAGTCAGAGACTGAAAGTATCTGGATGGTGGGGCTA
AACAGTACACAGGTTTTTTTGTGTTGTGTGCAAATTGCCCGATACCTATCCACAGGTAGCC
CCAAAGCCAGTCTTGAGTGTCCTGAACTTGTCATTACCTCTAGCATGTTCTGATCTTGGA
GCAGATGACCTAGAAAATGAATACCTTAGAGGCAGCTTACTTCTTTCACAGATGCAAAAG
AAAGCTGAAGATGCAGTGGCATGTGGTCGTGAAAGTAACATGGAAGAGGATAGCATTTT
TAAAATGGAAGCTGCACTTGATCGATGCCTTTTACGTCTTATAGCAAATTGTTGCAAGGG
AGATAAATTGGTTAGAGCAACAGAGCTGGCAAGACTATTATCATTGGAAAAGTCTTTACA
GGGTGCCATAAAACTTGTCAGTGCTATGAAGCTCCCAATGCTTGCAGAACGGTTTAATAC
TATACTGGAGGAGAAGATACTTCAGGAAAATATGGAGACGATCTCTTGTAGGAGATTAAC
TTCGGAGGCACAAGACATGGACACTCCAATTTCCATTAGTGTCAAACAAGTTTCATACGG
GGCAAACCTTGGTGATAGCCCCTTTCTTCCAAATCGTCAAGTCGAACCAAAGCATTCCAC
TCCAGTGTTTTCAAAGCCTGATACAAAAATTGAAGTAGATACCTCAGAAGCTATAGCAAA
GGGCTGTGATGCACAAAATGGAAATATAAAGAGTGGTGATGCAGAGGTCCAGCCTGCAA
GCCACAATGACTCTATTCAGAAGCCAAGTAACCCCTTTGCTAAGGCATCCAATACTTCTG
CAAACCAAGCTGTACAGCGCAATGCATCTTTACTGAGTTCCATCAAGCAAATGAAGACTG
CCACAGAAAATGAAGGGAAGCGAAAAGAAAGGGCAAGGTCAGGTTCCTTGCCACAGAAA
CCAGCAAAACAAAGTAAGATTTCATGATAAAAGCATGTGCTCAATAGTGGAAACATAAGA
TACAGTCAGGCTGAAAACCCTGACAAGTAAATGCTCTATGCCATGTAGGTGACATGGAT
GTCTGATAATGATGACTAAGTGCCTCATTTATTCAATTACGACGGATTCAGTTGGCCTTTT
GTAAAAAAAAAA

SEQ ID NO: 218
AAATACCTAGTATACATAACAAGATATTATACTCTGTTCCTCGCCCTAGATTGCGATTGAA
AGCAAGCGAGACAGTCGGGGGGCGGGGTTGCTGCTGTTGAAGGCGCCTTGGGCTGTAT
AGTAACAACACAGCCACTGAAATCATTGCTGGGGCGTCGGTAAAGTTGCAGAAGCAGCA
GGACGACTAACAGGATGAAGCAGAAGCGAAAGGGCCATCAAGTTGATGACCCAAAATAT
TCTGTCCAAACCCCCCAAGAGGACGACACACCTAACGAAAGCGGCCCTGCTTCGGAAG
AAGTGGAGAGTAGTGACGAAGAAGGCGGCAATAGCAGCAACATCGAAGATGATATTATT
TATTCTTCTTCCGAAGAAGACCCCGTTGTTAGTTCTGATTACGAAGAAGACGAAGACGCA
GAAAGTGATGCAGAGGGTGTGACTGCAGAGCAAGAGTTGGAAGGGGACATCGACAATG
CCCTCCAAAATTATATGGGAACCCTTACTGTTCTTTCCAATTTTCATGGAGAGAACCTTAA
AAATGCCGAAGGTGAAGATACTTCTGGGGATGATGATGACGAGGAGGAGATGCCTAAAA
GAGCTGAAGAGTCTGATTCTCCGGAGGACGAGAATGATGAGAGGCCTAAAAGAGCTGAA
GAATCTGATTTTTCGGAGGACGAGGATGAGGAGAGGCCTAAAAGAGCTGAAGAGTCTGA
TTCTTCGGAGGACGAGGTTCCTTCCAGAAACACTGTTGGAGATGTTCCCCTCCGTTGGT
ATAAAGACGAACAACACATTGGCTATGATATTAAGGGTAAGAAGATCAAGAAGCAGCCTA
AAAAGGACCAGCTAGATTCATTCCTCGCTAGTACAGATGATTCCAGTGACTGGCGTAAG
GTATATGACGAATATAATGATGAAGAGGTTGAGCTAACAAAGGATGAAATCAAGTTTATTA
GTAGATTACGTAAAGGCACAATACCTCATGCTGATGTCAATCCATATGAGCCTTATGTTG
ATTGGTTTGACTGGAAAGATAAGGGTCATCCCTTGTCTAATGCCCCAGAACCAAAGCGG
AGATTTATCCCATCAAAATGGGAAGCTAAAAAGGTTGTGAAGCTTGTGAGAGCAATTCGG
AAAGGATGGATAACGTTTCAGAAAGCTGAAGAAAAGCCTCGCTTCTATTTGATGTGGGGA
GATGATCTCAAACCATCAGAAAAAATGGCAAATGGGCTGAGTTATATCCCGGCGCCAAA
ACCAAAGTTACCTGGGCATGAAGAATCATACAATCCTCCACCTGAGTATATTCCAACACA
GGAGGAGATCAATTCATACCAGCTTATGTATGAGGAAGACCGTCCTAAATTCATCCCTAA
AAGGTTCGACTCATTGAGAAATGTTCCAGCTTATGACAGATTCCTTTCCGAGATATTTGA
GCGATGTTTAGACCTATACTTGTGTCCAAGAACTCGGAAAAAGCGTATTAACATTGACCC
CGAATCATTGATACCTAAGCTTCCGAAGCCTAAAGATCTGCAGCCTTTTCCCTCTATTTG
CTTCCTTGAATATAAAGGCCATACAGGCGCTGTTTCATGTATATCTCCAGAGTCATCAGG
GCAGTGGCTGGCCTCTGGTTCAAAAGATGGTACTGTGCGTATTTGGGAGGTGGAAACTG
CGCGTTGTCTTAAAGTTTGGGATATTGGGAGGCCCATACAGCATATTGCATGGAATCCA
GTTTCTCAGCTTTCTATTCTGGCAGTTGCAGTGGATGAGGAAGTGCTTGTACTAAATACT
GGACTTGGAAGTGAGGATAGTCAAGAAAAAGTTGCTGAACTATTGCATGTGAAATCAAAA
CCTGTATCAGCTGATGACTTGGGTGATAACACTTCTCTGACTAAGTGGATTAAACACGAA
AAGTTTGATGGGATCAAACTTACCCATTTGAAGCCGGTGCACTTGATATCTTGGCATCAT
AAAGGAGACTACTTTGCGACTGTTGCACCCGATGGTAATACTAGGGCAGTACTTGTGCA
CCAGCTTTCTAAGCAGCAAACCCAAAACCCTTTCAAAAAGATGCAAGGGCGTGTTGTTCA
TGTTCTTTTCCACCCTAGCCGAGCAATTTTCTTTGTTGCTACAAAGACACATGTCCGGGT
TTATGACCTTGTCAAACAACAGCTTGTTAAAAGGCTTGTGACAGGTCTGCATGAGGTTTC

ATCCATGGCAGTGCATCATAAAGGAGATAATCTTCTTGTTGGGAGCAAGGAGGGAAAAG
TGTGCTGGTTTGATATGGATCTTTCAACACAACCTTACAAGACTTTAAAGAACCATTCCAA
AGATATTCATTCTGTTGCATTCCATGATTCTTATCCTTTGTTTGCATCATGCTCAGATGATT
GCAAAGCATATGTTTTTTATGGGTTGGTGTATTCTGATTTGCTTCAGAATCCATTAATTGT
CCCTTTAAAGGTGCTTCAAGGTCACCAGAGTGTAAACGGCATGGGTGTTTTAGATTGCCA
GTTTCATCCTAAGCAGCCGTGGTTGTTTACAGCAGGGGCTGATTCAGTTGTAAAGCTATA
TTGTAATTAATTTATCATCCCTTCAGCAGAATCAGTCAATATGAAGGCGACATTATTGTGC
CAAAATTGGGAGGGTAAAAGACTACCGTGTTAATAATTTTGCATATGTTCAGGGGTATTA
AAAATTCAGAGGATAAATTTCCTCACTCTCAAGTGTTAGATGGTTTTATTCGAGAAAAAAG
ACAGTCAATTCTATATCAATATTTGAAAGTGGAGGCTTGTTGGAGGTTGTCATTTATAGAA
AACAGCTCTAATATGTATCGTTATATTGAAGGGAGAGATCATTAATGCTTTTGTATCAAGC
ATAGGGTGATATTTTGACGTGATGTTCAAAATTATTGTAGGCATATGACGGGATTTTTGAC
AAGGCCTCAGGTTGTTTCTCGAGATGCTACATGAAGATTTATGTTTCAAGTTCCAAATTTT
TTTTTGGTTTAAAGTTCCGGTATTTAGATATGCCGAGAGCTTGATTGGACCGAGAGCTTG
AATATATTCCGAATCTTGATCTTAAAAAAAAAA

SEQ ID NO: 219
ACATTTCTCGCTAAACAGAGCCAAATTCACCTTCTGGCTGACGCTGCAACAGGTTGGGT
GATCCAGAGCTTCAATTTGAATATTCAAAGCATTTTTTAGCTCAAAGTTCTACCGTTTCCA
GATTAAGAAGTTAGGTGAAGAGAAATTCTCTTCCTTTTTCTTAAGTACCTTGATTTAAGAA
ACAGAATGATGTCTCTAAAAAGAGGATTCGAAGAATCCTTGGTACCAGCCAAAAGACAGA
AAACAGAGTTGTCCACTGTCACATATGGTGATGGGCCTCGACGGACATCCAGTTTGGAA
TCCCCAATCATGCTGCTGACTGGGCACCATGCAGCTATATATACAATGAAGTTCAATCCC
ACAGGAACAGTGATTGCATCTGGATCACATGAGAGGGAGATATTTTTGTGGAATGTGCAT
GGGGATTGCAAGAATTTTATGGTCTTGAAAGGTCACAAGAATGCAGTTTTGGATCTGCAC
TGGACAACCGATGGCTGCCAAATCATATCAGCAAGCCCTGATAAGACTCTTCGTGCCTG
GGATGTTGAGACAGGAAAGCAAATAAAAAAAATGGCAGAGCACTCATCCTTTGTTAATTC
TTGTTGTCCTTCACGGCGTGGGCCCCCTCTTGTAGTTAGTGGATCTGATGATGGGACTG
CAAAGCTCTGGGATTTACGTCACAGAGGAGCTATTCAAACCTTTCCAGACAAATACCAAA
TCACAGCTGTTGGGTTCTCTGATGCCGCAGACAAGATATACTCTGGTGGAATAGATAATG
AAATTAAGGTGTGGGACCTTAGAAGAGGTGAGGTCACAATGCGACTCCAAGGCCACACA
GACACAATTACAGGCATGCAGTTAAGTTCTGATGGCTCTTATCTCCTGACTAATTCGATG
GATTGTTCTCTTCGTATTTGGGATATGCGTCCATATGCCCCTCAGAATAGATGTGTGAAA
ATCTTAACAGGCCATCAACACAACTTTGAAAAAAACCTTTTGAAATGTAGTTGGTCTTCTG
ATGGAAGTAAAGTCACGGCTGGTAGTGCAGATCGTATGGTTTATATATGGGACACAACC
ACTCGACGTATATTGTACAAGCTCCCAGGCCACACTGGGTCTGTAAATGAGACTGGTTTC
CATCCTACACAGCCAATTATTGGATCATGTAGTAGTGACAAGCAGATATACTTAGGGGAG
ATTGAGCCTAATGTTGGGTATCAAGCTGTAATTTAGACATGATTCAAAGTCTAGACATTAA
TGTTTTGGAACTCTTTTTTCGAAAAAAAAAGAAAAAAAAAA

SEQ ID NO: 220
GGAGTTTTTTGCTCTGTTGGAGATTGTGAAAAGGCAGAGGAGCAGAGGCTATGGAGTTC
AGCGACACATATAAGCACACAGGGCCTTGCTGTTTCTCTCCTGATGCACGCTATTTGGC
CATAGCCGTTGACTACAGGCTGGTGATTCGAGATGTCGTTACCCTCAAGGTTGTACAGTT
GTATTCATGCATGGATAAGATCAGCAATATTGAGTGGGCTCTTGATTCAGAGTATATTCTT
TGTGGTCTATATAAGAGAGCAATGGTTCAGGCATGGTCATTATCACAACCTGAATGGACA
TGCAAGATAGATGAGGGGCCTGCAGGAATCGCCCATGCAAGATGGAGTCCCGACAGTC
GCCACATTATTACAACATCTGACTTCCAGTTGCGTCTCACAGTCTGGTCTCTTGTCAATA
CAGCCTGCATACATATACAGTGGCCTAAACATGCATCTAAAGGTGTTTCATTTACCCAGG
ATGGCAAGTTTGCAGCTATAGCTACACGGCGAGATTGCAAGGACTATGTGAATCTTCTTT
CTTGTCATACGTGGGAAGTCATGGGCACATTTACTGTTGACACTATAGATCTTGCAGATC
TTGAATGGTCACCAAATGATAGTGCAATTGTTGTTTGGGACTCCCCACTTGAGTACAAGG
TTCTTATTTACTCTCCAGATGGGCGGTGTTTATTTAAATATCAAGCTTATGACAGTTGGCT
AGGTGTGAAGACTGTTGCATGGTCTCCATGTAGCCAGTTTCTGGCAGTAGGCAGTTATG
ATCAAACACTGAGGACTTTGAATCACCTTACTTGGAAACCTTTTGCAGAATTTGTGCATGT
GAGCACTGTTCGAGGTCCTGCCAGTGCTGTTGTTTTTAAGGAAGTAGAGGAACCATGGA
ATCTTGATGTGTCTGGTCTTCACTTGAATGATGACAATGCTCATGACATCCAAGATGGCA
AGCCAGCTGAAGGCCATTCTAGGGTCCGGTACAAGGTAGTGGAATTTCCTGTCAATGTA
TCTTCACAAAAGCATCCCGTGGATAAACCAAATCCAAAGCAAGGCATTGGCTTGCTAGC
GTGGAGTCGAGATAGCCAATATTTGTTCACTCGTAATGACAATATGCCAACGGCACTTTG

GATATGGGATATTTGTCGCCTTGAGCTTGCTGCACTTCTGATACAGAAAGAGCCCATTCG
TGCAGCTGCCTGGGATCCAGTATATCCTCGTGTGGCTCTTTGTACAGGAAGCTCGCATT
TGTACATGTGGACACCCTCTGGTGCTTGTTGTGTGAATATTCCTCTGCCACAATTTGTCG
TATCAGATTTAAAATGGAACCCAGATGGAACTTCTATGCTTCTAAAGGATCGTGAGTCGT
TCTGTTGTACTTTTGTTCCAATGCTTCCTGAATTCAACGACGATGAAACTAATGAGGAATA
GTACACACAAAGCAAAAAGTGTATCCAACATCACTTGTGACAAGGCCATCGTTGTCTGCG
AATCTTCTTGGTCAACTACAAATGACCTCTGTATACTTCTCCTGGTCGTTGACACTCTAGT
GCTCACAACAGTGATTGTTTACTAGTTTCAGCTATGTCTTAAGTGTCTCCTACATTCTGCA
ACAAGTTATTTAGCTGTACCATTGTTGAAATTGTGCAAAGTTGATCCTTGGAGAGTTTATT
CCAATGTTTAAGATACAACTTGCTGCAGTGCCTTAGGAGTTCCAATTCTTCCCATTCAGC
TATATGAATAGCAACTAAGTGGTTATAGAGCTTCTCTCCATTCTTATTACTTCACTCGAGG
ACGTGATGTCTACTTCAATTTTTATAATATAATGTTGGATTTATCTATCGACCTTTTATTCA
TTCTCTTAGTCTTTGCTAGTGTAGGCATCAATTCTTATTATAAATATATTGTACTGGGGAT
CCAAGACATGGCAATATATGTCGAGATTTTCATTTTCTCAAAAAAAAAA

SEQ ID NO: 221
CTGCCACCCACATGGTAGCTTTCTTGTAAATTGCAGAAGTGTTACCATATTGTGATTAAAT
TAAATTGTGTGCTGTTTTTATAGTCATTCCCTGCATTTTTCCCTGAACGTTTCAATCCATTA
CAGGCCCTTTTTCTTGGTGTCAGTTGGTCGAATCGCGTGAGCCCTTGCTAGATTCCAATC
CCGTCCACTGCTGCGTAGTCTATTCGTCTCTCTCATTTTGCTGTTTGATATGATTATGAGG
CTTTGAAGTCTATCGCACTTAAAAAGCTAATCGCCAATTTAGCAGAAAAAAACTGACCCAT
CTTAGATAAGCCCTCCAAATTCTACTCATGAGCCCTGTTTTGCACAAAGTTTAGCATACAA
ATAGACCCCTCTGAGACAAGCCATCAAAGGTCCAATCCCTCTGGTTTCCTTTTCCTTCAA
ATTTAACCGAAAATCGGCAGATAAGGCCCCATAATCCTACCCATAAACCTGTTTTCTCTTA
ATTTTCGCATACGACTTGATTGATTTCGGATATTAGTATGGCCAAGCTCATTGAGACTCAT
TCCTGCGTCCCTTCAACGGAGAGGGGCCGTGGGATTCTCATAGCAGGAGATGCAAAGA
CCAACTCTATCATATACTGCAATGGCCGATCCGTGATCATGAGAAACCTGGACAATCCCC
TGGAAGCTTCTGTTTATGGAGAACATTCATATCCTGCAACTGTGGCTCGCTTCTCACCAA
ATGGAGAGTGGGTTGCATCAGGGGACACTTCTGGCACAGTTCGAATCTGGGGACGAGG
TTCGGATCATACTCTCAAATATGAGTACAAGGCTCTCGCTGGAAGGATTGATGATCTCGA
GTGGTCTGCCGATGGACAGCGGATCGTTGTATGTGGAGACAGTAAAGGCAAATCAATGG
TCCGGGCTTTCATGTGGGACTCTGGCACAAATGTTGGTGAATTTGATGGGCATTCAAGG
CGTGTTCTAAGTTGTTCCTTTAAACCAACGCGACCATTCCGTGTTGCCACATGTGGAGAG
GACTTCTTGGTGAACTTTTATGAAGGACCACCCTTTAGATTTAAGACATCACACAGAGAT
CATTCAAATTATGTGAACTGCGTAAGGTTCGCTCCAGATGGAAGCAAGTTTATTACTGTT
GGTTCGGATAGAAAAGGAGTGATTTTTGATGGCAAGATGGGTGAAAAGATTGGGGAGTT
GTCTAAAGAAGGTGGACACACGGGCAGTATTTATGCTGCTAGCTGGAGCCCTGACAGTA
AACAGGTACTTACTGTGTCTGCAGACAAGTCTGCTAAAATATGGGAGATTAGTGAAACTG
GTAATGGAACTGTGAAGAAAACATTGACTTTTGGAAGCCAAGGAGGAGCTGATGACATG
CTAGTTGGGTGCCTTTGGTTAAATGATTATCTGATTACTGTTTCTCTTGGTGGCATCGTCA
GCTTACTCTCTGCAGTTGATCCAGATAAACCACCAAAGACAATTTCTGGTCACATGAAAA
GTATAAATGCAATTGCATTATCTCTTCAAAGTGGACAAAGTGAGGTTTGCTCAAGCAGCT
ATGATGGTGTAATTGTTAGATGGATTCTTGGAGTTGGCTATGCTGGTCGGGTAGAGAGA
AAAGATAGTACTCAAATCAAATGCCTGGCAACAATTGAAGGAGAGCTGGTAACTTGTGGT
TTTGATAATAAGGTGAGAAGGGTACCTCTGCTATCAGAGCAACATAAAGAGTCAGAACCA
ATTGACATTGGAGCACAACCAAAGGATCTAGATGTTGCAGTTGGTTGCCCTGAGCTTACT
TTTGTTTCCACCGATGCTGGGATAATAATTATTCGTGCGTCAAAAATAGTATCAACTACTA
ATGTTGGGTATGCAGTGACTGCAGCTGCAATATCACCTGATGGAACAGAGGCTGTGGTT
GGTGGTCAGGATGGGAAATTGCGTGTGTACTCCATCAAAGGTGACACTCTTCTGGAAGA
GTCAGTCCTTGAAAGACATCGTGGTCCAATAAATGCAATCCGCTTTTCTCCTGATGGATC
CATGTTTGCATCGGGAGACCTGAACAGAGAAGCAGTTGTGTGGGACCGTATTACTAGAG
AGGTGAAACTTAAGAACATGGTCTACCACACTGCACGTATCAATTGCATTGCATGGTCTC
CAGATAGCTCTAAAGTGGCAACAGGCTCTCTTGACACCTGCATATTGATATATGAGGTGG
GAAAACCAGCATCCAGCCGAATTACAATTAAGGGGGCTCATTTGGGAGGGGTTTATGGT
CTAGCTTTTAGTGACCAGAGTACTGTTATAAGTGCAGGAGAAGATGCATGTGTCCGTGTG
TGGAGTCTCCCATAGCTCTGTAATGTCTGGATCTAATAATTTATATTTAGGGTATAGGGAT
AGATGTCTCGACATACGAGCATGGTTCTTGATGCCATCTTGTGCCATGTTTGTAGTGCTT
TTGCATGAGTTCAAATGTCTTTGTGACATATTGTCTTGAACCACCGAGGATATATCATATT
TATCTTGTAGATGGTTTCAAACGGTGCTGCTTATCTGGGGGTAATGGGGAGCATGGAGT
TTTTCAAAAAAAAAA

SEQ ID NO: 222
CTTGGGCCACCACCTTCTCTAGTAGCTGAATCTTGCAGGGGGTCGAAGTCTGTTTTTGAA
GACCAGATTGCAGCCTATTACCCAACCAAATCAGGCCTTGTTGTTGGGTTGCTACAAGTG
ACAATATCTAAGTATTGACCAGGGTGAAGAGATGCCACAACCGTCAGTTATCCTTGCAAC
AGCAGGTTATGATCACACAGTAAGGTTTTGGGAAGCCACTAGTGGCCGCTGCTATCGGA
CTCTTCAGTACCCTGATTCACAAGTCAACCATCTAGAGATAACACCTGACAAGCAGTACT
TGGCTGCAGCTGGGAATCCACATATTCGTTTATTTGAAGTCAATTCCAATAATCCTCAAC
CTGTAATTAGTTATGACTCCCACACAAACAATGTTACGGCAGTGGGATTCCAGTGTGATG
GAAAATGGATGTACTCAGGTTCGGAAGATGGTACTGTGAAGATATGGGATCTGAGAGCT
CCAGGTTTCCAGAGGGAGTATGAAAGTCGGGCTGCTGTTAATACTGTTGTCTTGCATCC
AAATCAGACAGAATTGATATCCGGAGACCAAAATGGAAATATTCGTGTGTGGGATCTCAA
TGCAAATTCTTGCAGTTGTGAACTGGTTCCAGAGGATACAGCTGTAAGATCATTGACAGT
TATGTGGGATGGAAGTCTGGTTGTAGCTGCAAATAATCATGGAACATGCTATGTTTGGAG
GTTGATGCGTGGAACACAGACAATGACGAACTTCGAGCCATTGCATAAGCTTCAAGCAC
ATAATTCATACATTCTGAAATGTCTTCTTTCACCCGAGTTCTGTGAACATCATCGGTATTT
GGCAACAACATCATCTGACCAAACTGTAAAGATATGGAATGTTGATGGCTTCACTCTGGA
GCGTACACTAACAGGACATCAACGTTGGGTTTGGGATTGCGTATTTTCTGTTGATGGAGC
TTTCCTTGTTACTGCTTCCTCAGATTCAACTGCAAGACTGTGGGACCTATCAACTGGAGA
AGCCATCAGAACATATCAAGGGCATCATAAGGCGACTGTATGCTGTGCTCTGCATGATG
GCACAGATGGTGCTTCTTGCTGATGCCTTGTTTGTATGTCCAATAGATTATAACCTATTTA
CTGTGACACTATTCTTCACACCCATGTCAATTGTCCAATGTAGCATGCCCCGCACCTTGT
ACAGAATTGTTAGAGCAAGTTGTACTTCAGGTCGTAGAGTTGCTTTTCACTTGTTATTTGT
GTAAGTGTCATTAGATTAATATGAAATTAATATGAAAGATTGTACGCTAGCTGTGATGGAC
ATTAGGAGAGAAAT

SEQ ID NO: 223
ACTTCTCAAATACCTATAGTAGAATCGGAATTGCCATGTATTTACTTCACTTCTTTAATCAT
AAATTCTCATAAACACCTGTCGCATTTTCCATTTCATCCGATTGGGCTCAATCGGAGGTC
ATTCAGGCTGATTATTTGGATTCTGAAAAGAGGGATTCCAGCCTTAAAGAGCCTTCAAGC
TCCCGAAGTTCTGCTTCAGAGAGAGGGATTCGGGAATTTCACAAGGCCAAATTCCGGAA
AATAGGTTCAGAGCAGAGGATCTGGATTTAAATTCGATTTAAAAGTCGATTCTTGTTTGAA
TTCAAACCTCCTTCGTAGATCTGAGATTTAGAGGTTTATTCTTTTGGAGCTTTCTAGTTTG
AGGGTATATGTGAGATCATATATCGGATCCAGTTGATAGTCCACTAAAGTTTCTTTTTTCC
TTCTACTGTTTAATTCATTTTGCAACATTGTGAAAATTGGTTTCAATGTTAACCAAATTTGA
AACCAAGAGCAATAGGGTAAAGGGTCTCAGTTTTCATCCAAAGAGGCCATGGATCCTGG
CAAGTCTTCACAGTGGAGTTATTCAGTTATGGGACTATCGGATGGGAACTCTGATAGACA
AATTCGATGAACATGATGGACCTGTTCGAGGAGTTCATTTTCACAAGACACAGCCACTCT
TTGTGTCTGGAGGTGATGATTACAAGATTAAGGTCTGGAACTACAAGATGCGTCAATGCC
TCTTTACATTTGTAGGGCACCTTGACTACATCCGGACAGTTCATTTTCACAATGAATATCC
TTGGATTGTTAGTGCAAGTGATGATCAGACAATAAGGCTCTGGAACTGGCAATCACGGG
TTTGCATCTCTGTCCTTACCGGTCACAATCACTATGTGATGTCTGCTTCCTTTCATCCTAA
GGAAGATCTAGTTGTTTCAGCATCGTTGGACCAAACTGTTCGTGTATGGGATATAAGTGG
CCTTAGGAAAAAAACTGTGTCTCCAGCTGATGATCTTTCGAGGCTGGCACAGATGAATAC
GGATCTTTTTGGCGGTGGTGATGTTGTAGTGAAGTATGTACTCGAGGGACATGATCGTG
GTGTGAACTGGGCTGCTTTTCATACCAGCTTGCCATTAATTGTTTCTGGTGCAGATGACC
GCCAAGTCAAGCTGTGGCGAATGAATGATACAAAGGCTTGGGAGGTTGACACTCTACGG
GGACATACAAACAACGTGTCCTGTGTGATTTTTCATGCACGACAAGACATTATTGTTTCAA
ATTCTGAAGATAAAAGTATCCGAGTGTGGGATATGTCCAAGCGAACTAGTGTTCAAACTT
TTCGTAGAGAGCATGATCGATTCTGGATTCTTGCAGCACATCCTGAGATGAATCTCCTAG
CAGCTGGGCATGACAGTGGGATGATTGTGTTCAAATTGGAGAGAGAGAGGCCTGCTTAC
GTTGTTTATGGGGGCTCGCTGTTGTATGTCAAGGATCGTTATCTGCGGACTTATGAGTTT
GCAACCCAAAAGGACAATCCACTAATACCAATTAGGAAGCCTGGTTCCATTGGTCCAAAC
CAAGGGCCAAGATCTTTGTCATACAGTCCAACAGAGAATGCTATCTTAATTTGTTCAGAT
GCTGATGGTGGTGCTTATGAACTTTATGCTGTTCCAAAGGATAGTCATGGCAGGAGTGAT
ACAGTACAGGAGGCAAAGAAAGGATTAGGAGGATCTGCTGTCTTTGTGGCTCGCAATCG
ATTTGCTGTTCTTGACAAGAATCACAATCAAGTCACTATCAAGAATTTAAAAAATGAGGTG
ACAAAAAAGTTCGATCTTCCAGTCACAGCAGATGCACTTTTCTATGCAGGGACAGGCAAC
TTGCTTTGCAGATCAGAAGATAGTGTATTTTTGTTTGATATGCAACAGAGGACGGTTTTAG
GAGAAATTCAAACCCCTAATGTCAGGTATGTAGTTTGGTCAAATGACATGGAGAATGTTG

CCTTACTGAGTAAGCATACAATCATTATTGCAAGCAAGAAGCTGTCAAGTACGTGTAGTC
TGCATGAAACTATTCGTGTTAAAAGTGGGGCTTGGGATGACAATGGCATTTTTATGTACT
CAACTTTAAATCACATAAAGTATTGTCTGCCAAATGGAGATAGTGGCATTATCAAGACGC
TGGATGTTCCAGTATACATAACAAAGGTTTCTGGAAAGTCTCTCTATTGCCTCGATAGAG
ACGGCAAGAACCGTGTTATACAGATAGATATTACCGAGTGTCTTTTCAAGCTAGCACTCA
GCAAAAAGAAATATGATTATGTTATAAACATGATTAGAAACTCTCAGCTTTGTGGTCAAGC
AATCATAGCTTACCTACAGCAGAAGGGATTTCCAGAAGTGGCTCTTCACTTTGTCAGAGA
TGAGAGAACTCGATTCAACTTGGCAGTAGAGAGTGGAAATATTGAAATAGCTGTTGCTTC
TGCAAAGGAAATTGATGAGAAGGATCACTGGTACAGGCTAGGAGTGGAGGCCCTTAGAC
AAGGTAATGCTGGAATTGTAGAGTATGCCTATCAAAGAACAAAGAATTTTGAAAGGCTTT
CTTTTCTATATCTCATAACTGGTAACCTAGACAAATTATCAAAGATGTTGAGGATTGCTGA
AATGAAAAATGATGTCATGGGCCAATTTCACAATGCACTATATTTGGGTGATATTCAAGA
GCGGATCAAGATTTTGGAGGAGTCTGGCCACCTGCACCTTGCTTATGCCACTGCATCAT
TGCATGGTCTTGCAGACATTGCTGACAGGCTTGCAGCTGATTTGGGTGGCAATATTCCA
GTTTTACCTCCAGGGAAAAAATCATCACTTCTAATGCCACCTGCCCCCATTCTGCATGGT
GGTGATTGGCCTTTGCTCAGGGTTACGAAAGGTATTTTTGAGGGTGGTTTGGAGAATTC
CACTTCTGCAGCTTATGAAGAAGAGGATGAAGAAGCTGCTGCTGACTGGGGTGAAGACA
TAGATATAGAAAACATTGAAGGGGAAAATGGTGAGGCTACAGTGTTGGATGATCAAGAA
GTTAAAGGTGGAGAGGATGATGAGGGAGGATGGGACATGGAGGATTTGGAACTTCCTC
CAGACGTAGCTGCTGCTAATGTAGGAACCAATCAGAAGACATTGTTTGTAGCCCCAACAT
TAGGTATGCCGGTAAGCCAAATTTGGATGCAAAAGTCTTCTCTCGCAGGTGAGCATGCA
GCTGCAGGCAACTTTGAAACAGCTTTACGTCTACTGACTCGCCAACTGGGCATTAAAAAT
TTCTCTCCACTGAAACCTCTATTTTTGGAACTTTATATGGGTAGTCATACATTCCTTCCTT
CCTTTGCTTCTGTACCTGCTTTTTCCTTGGCACTACAGAGAGGATGGAGTGAATCTGCTA
GTCCGAATATTAGAGGTCCCCCAGCCTTGGTCTATAGGCTTTCGGTGCTAGAGGAGAAA
CTAACGGTAGCCTACAGGGCCACGACAGAAGGTAGGTTTAGTGAAGCACTGAGGCTATT
TCTGAACATTCTGCATACAATTCCAGTTATTGTGGTAGACTCAAGAAAGGAAATTGATGA
GGTAAAAGAATTGATTGGAATTGCTAAAGAGTATGTTCTTGGTCTTCGGATGGAGGTCAA
GAGAAAAGAAATAAGGGATGATGCTGTTCGGCAGCAGGAGCTTGCCGCTTATTTTACTC
ATTGTAATCTGCAGAAAGCTCATTTGAAGCTGGCTTTGCTAAATGCAATGGGTATTTCTTA
TAGATGTAAGAACTATAATACAGCAGCAAACTTTGCTCGAAGGCTTCTGGAGACTGACCC
TTCTTCAAACCATGCAACAAAGGCTCGACAAGTTCTTCAGGTCTGTGAGAGGAACTTGCA
AGATGCGACACAGCTCAACTATGATTTCAGAAATCCTTTTGTTGTCTGTGGGGCAACTTT
TACTCCAATATACCGTGGTCAGAAAGAGGTGTCCTGTCCGTATTGCATGGCTCGTTTTGT
TCCTGACATTGCTGGGAAACTTTGTTCGATATGTGATCTTGCAATAGTGGGTTCAGATGC
ATCTGGACTATTTTGTTTTGCTACTCAAACAAGATGATCTTAGCTTTACAGAACAATGATT
TTGTGATGCCTTTTATATTCTTTGAGAGGTAAATTGTTTGCTGCCATTTCCCATCAGTATA
TGCAAGAAATGTTTCCATTTATGTCAGTTGATACCAACAGGGGCGTCAGAATAATCGTAC
AATGCATGCCAGGCTTACGTAAGCAACATGTGAGGATATTTATTTAAACATTATAAGAAG
CAGGACGCATGCATGTTTTATTTGCTGTTTCAGTTCCAGAATTGAAGCATGCACGCTGTC
TTATTAGCTGTTTCAGTTCTGGATTTTGTTTGCTGTTCTTCTTAGTAAGGCAGAAAGCTTT
ACTCACTTTGTAAGATGTGATGTAGGAATTTTGTTTACAGTGCAGGATGGCTATGCCCTC
TATTGTTCTCTATTCTGTGCATTGTGGCAAATGGTTTTGTTATACCATTTTTTTGGCCTTGT
AGGTGAATAATTGGATTTGCTTGAAAAAGTCAATCATCTTAATTCGAGGACCTAATAAAGG
TTACTACTTGAAGCTTTTCAAAAAAAAAA

SEQ ID NO: 224
CCCGAAGGGCAACTTTGACGTCACAATAGTCCTAAGTGACGGGGCCGCACAACCAGACT
GCTTATGAATACCAGGCTCCACTACCCATTGCAAGACATTTAGTCAGTTCAAAGGGAGCG
CTCTATTAAAAAAATTCCTCTTGTAAATCGTTGCAAGAGAGCTTGCCATCCAACAACAACA
ACGACTCCAGGACGAATGGATTTACTGCAGAATTACCAGGATGACAGTGAAGATTCGAA
CCCAGAACTTAGAAATCATCCACCGCTGGAAGACGCCACGGCCACTTCGGCGCCCGCA
GGAGTCGAAAATGAGACCTCTTCTTCACCAGACTCCTCCCCTTTGCGCCTTGCATTGCCA
GCAAAATCTTGCGCACCCGATGTGGACGAAACCCTAATGGCCCTCGGCGTTCCAGGCTC
CGAGAAAAAAAACAATCACAACAAGCCCATCGACCCTACGCAGCACAGCGTGACATTCA
ATCCAAGCTATGATCAGCTCTGGGCCCCGCTTTACGGCCCTGCTCATCCTTACGCCAAG
GATGGTATCGCTCAGGGCATGAGAAACCACAAGCTCGGTTTTGTGGAGGATTCGGCCAT
CGAACCGTTCATGTTCGACGAGCAGTACAATACTTTTCACAGGTATGGTTACGCTGCCGA
CCCTTCTGCTTCTTTGGGTAGTCATATTGTTGGTGATTTAGAGTCTTTGAAAAAAAACGAT
GGGGCATCTGTGTATAATTTACCTAAACGTGAGCATAAGAGGCAGAAACTCGAGAAAAAA

ATGATACAGAAGGACGAGAACGAGGAGGAGGAAAAAGAAGTTGGAGAGGAAGTTGACA
ATCCTTCCACGGAAGAGTGGCTGAAGAAGAATAGGAAAAGCCCTTGGGCTGGTAAGAAG
GAGGGTTTGCAGACTGAATTGACTGAGGAGCAGAAGAAGTATGCTCAGGAACATGCTGA
GAAGAAGGGTGACAGGGAGAAGGGTGAGAAGGTTGAAATTGTAGATAAGACTACTTTCC
ATGGCAAGGAGGAGAGGGACTACCAGGGGAGGTCTTGGATTGATCCTCCCAAGGATGC
CAAGGCAACCAATGATCATTGTTATATTCCGAAGAGGTGGGTCCATACATGGAGTGGGC
ATACAAAGGGGGTTTCCGCTATCCGATTCTTCCCCAAGTATGGTCATCTGTTGTTATCTG
CTGGCATGGATACGAAGGTGAAGATCTGGGATGTTTTTAACAGTGGGAAATGCATGAGG
ACTTACATGGGTCATTCGAAGGCGGTGCGTGACATATCATTCAGCAATGATGGTTCTAG
GTTTTTGAGTGCTGGGTATGATAGAAATATCAAATTGTGGGATACGGAGACAGGGAAGG
TGATTTCTACATTCTCCACAGGGAAAATACCATATGTTGTGAAGCTCCATCCGGACGAGG
ATAAGCAAAATGTGCTACTGGCAGGTATGAGTGATAAGAAGATCGTGCAGTGGGACATG
AATAGTGGTGAGATCACTCAAGAGTATGATCAGCATCTTGGTGCTGTGAATACCATTACT
TTTGTAGATAATAATAGGAGGTTTGTGACATCGAGCGATGACAAGTCCCTGAGGGTTTGG
GAATTCGGCATTCCTGTGGTCATTAAATACATCAGTGAGCCTCACATGCATTCAATGCCT
TCGATTTCTCTTCATCCAAACACGAATTGGCTTGCTGCACAGAGTTTGGATAATCAAATTC
TGATCTACAGTACAAGGGAGAGGTTCCAACTCAATAAGAAGAAGAGGTTTGCGGGGCAT
ATTGCAGCAGGTTATGCTTGCCAAGTTAACTTTTCACCCGATGGACGATTTGTTATGTCC
GGAGATGGCGAGGGCAGATGTTGGTTCTGGGATTGGAAAACTTGTAAAGTCTTCAGAAC
TTTAAAATGCCATGACAACGTTTGCATAGGCTGTGAGTGGCATCCTCTGGAACAGAGCAA
GGTTGCTACATGTGGCTGGGACGGAATGATTAAATATTGGGATTAAATTCAGTACAATAA
ATGCAAGATCTCTAACTCAACTTGAGTATGTATGAAGGTGCCGTATCAAAAGATTGGTAC
TTCCTTATGGACACACAAGATCGTAAGCATGGCTGAATTGACTTGAGGAAAATGAAACGA
CCTTGCTAATTGTAATGATGTAGAGGCCTCATGTGGTACCTGGTTGTCTACCAGACTATA
CTATACAGTCTGTCCTACCTTTCGAAGGAAACTTTTGAGTGACTCATGTCAATTTGTTCTT
GTATTTGACCTCCAAGAGATGTAACATTGAATTTGATTTCACCTCCATGGTTTCCAAGATG
TGAATCTTTTATTGCTTATTTGTTCACACTGGAAAAAAAAAA

SEQ ID NO: 225
GATTGCTCTAATATGAACCAGCTCAACTGGACGCTCTAAACTTACGTATGATTGAACGAA
GTGTTTGGCGCTTGTTTCGCCTTATAAATAGTGCTTCTTGTATCAGTGCGTATTGCATCA
GTGTAGTGAACACTTTGAAAGGAGTTTGGGCTTCGACTCTCTACCCTGTAAACAAGTTCA
GCTTAAACGCTGTTGCAGTAAAAGGTGTTGAAGATGGCAAGGAAGGGTTTGGGTACTGA
TCCGGCAATAGGCTCGTTAATGTCATCAAAGAAGAGGAAAGAGTATAAAGTCACCAATAG
GTTTCAAGAGGGCAAGCGGCCTCTTTATGCCATTGCTTTCAATTTCATTGATGCTCGTTA
TCATAATATTTTCGCAACAGCAGGAGGAACCCGGGTGACCATTTACCAATGCTTGGAAG
GAGGTGCTATTTCTGTTTTGCAGGCATATGTGGACGATGATAAGGATGAATCATTTTACA
CATTAAGTTGGGCATGCGATGTGAATGGTTCACCATTATTAGTTGCTGGTGGTCACAATG
GTATAATTCGAGTGCTTGATGTTGCAAATGAAAAGGTCCATAAGAGTTTCGTTGGCCATG
GAGATTCTGTGAATGAAATAAGGACCCAAGCACTGAAACCTTCCCTTATTTTATCTGCTA
GCAAAGATGAATCTGTACGACTATGGAATGTTCAAACTGGAATATGCATTTTGATATTTGC
TGGAGCTGGGGGACATCGCAATGAAGTTCTCAGTGTAGATTTTCATCCTTCAGATGTATA
TCGTATTGCAAGCTGTGGCATGGACAACACGGTCAAGATCTGGTCTATGAAAGAGTTTTG
GACCTATGTTGAGAAGTCATTTACCTGGACTGATTTGCCATCGAAATTCCCCACGAAGTA
TGTGCAATTTCCTGTTTTTATTGCTGCAGTGCATTCAAACTATGTGGATTGCACTAGATGG
CTTGGCAACTTCATCCTATCAAAGAGTGTGGACAATGAAGTAGTTTTGTGGGAGCCTTAC
AGCAAGGAGCAAAGTACTAGTGATGGAGTTGTGGACATCCTGCAGAAGTATCCAGTGCC
AGAATGTGACATTTGGTTTATCAAGTTTTCATGCGATTTCCACTATAATTCTATGGCAGTT
GGCAACAGGGAAGGAAAAGTTTATGTTTGGGAATTACAGTCAAGCCCTCCTAATTTAATT
GCCAGGTTATCACATGCACATTGCAAAAATCCTATTAGACAGACAGCTATATCTCATGAT
GGAAGTACCATTCTCTGCTGCTGTGACGATGGTAGTATGTGGCGCTGGGACGTGGTTCA
GTAAAATCAGCAATTTCCTTGCTGAATGTGCATTCGTATATGTTTCAGATATTTACACCTG
CACACTAATGACTTTGATTGTTATTAGCTTCTGGTTGATGTTGTGGAAAATCCTGGCTTTA
GCCAAGCTTTTTGTTTTGCCAGAATAGTCAACAGAGTTTTCTGTTTGCTAGAATACAGCAT
GCCCATTTTTATAAAGCTACACTTTGCGGGACATTTGAAACTAGTATGCTCCCCAGGGCC
TTGTGTTCATAAAGATGCCTTCAATGATTCAGTTGTTATTATCATCTTTAATGGAGGCAAT
GGTCTATTCACATCAGATAGTTGATGGCCACATGAGTTGTTTATACAAGTCGTTGTTTTAT
GAGAGAACCTTCTTCAGATGAATTAAAGGTATCAATATAACTATCTCAGTAATTTGAAAAA
CCTAGGATGTTGTTCTGGCAAAAAAAAAA

SEQ ID NO: 226
GCAGCTGAAAGCGCTCAAAGCGCTCGCCATGGCAATTCCACAGTGAAGAAGACCAGAG
CAGGGAGTAATGGAGTCTGGGGCAGGAGGTTCAGTTGGGGCTCGCGTGCCGTCTGCGA
AGCCAGAGATGCTGCAACAGCCACCGTATTCTAACGGTGACGATGATAATGATATGGAG
CGTGGAACGGCACCCGTCCCGTCTTCGAATCCCAATACTGTCTCGAAATGGGAGCTGGA
CAAGGATTTTCTGTGCCCTATCTGCATGCAGACGATGAAGGATGCGTTCCTCACGGCTT
GCGGCCACAGCTTCTGCTACATGTGCATCATGACGCATCTCAATAACAAGAGCAATTGC
CCCTGCTGTAGCCTCTATCTCACAAACAATCAGCTCTTTCCCAATTTCTTGCTAAATAAGC
TTTTGAAGAAGACATCTGCATGTCAAATGGCAAGCACTGCTTCACCGGTAGAGAATCTCT
GCTTGTCGCTACAACAGGGAGCAGAAGTCTCAGTAAAGGAGCTGGATTTTCTCTTGACT
CTTCTTGCTGAGAAGAAGCGGAAGATGGAACAAGAGGAAGCTGAAACTAATATGGAGAT
ACTCCTGGACTTCTTGCAACGATTGAGGCAGCAGAAACAGGCAGAGTTGAATGAGGTGC
AAGCTGATCTCCATTATATCAAGGATGACATATTAGCATTGGAGAAACGAAGACTGGAAT
TATCTAGAGCTAGAGAACGATATTCTAGGAAATTACATATGCTTTTAGATGACCCAATGGA
CACTACATTAGGTCATGCTGCAATTGATGATGGAAATAATGTTCGCACAGCTTTTGTACG
TGGTGGCCAAGGTGATGCTATTTCAGGGAAATTTCAGCAGAAAAAGGCTGAAATCAAAG
CACAAGCTAGTTCTCAAGGGATGCAAAAGAGAGCTAATTTTTGTCATTCTGATTCCCAGG
TTTTGCCTACTTTGTCAGGATTGACAATCGCAAGGAAAAGAAGAGTCCTTGCACAGTTCG
ATGATTTACAAGAGTGCTACTTGCAAAAAAGGCGACGCTGGGCCACTCAATTACGAAAAC
AATGTGATGGTGGCTTGCGAAAAGAAAGGGATGGGAACAGTATTAGTAGAGAAGGCTAT
CACGCAGGCCTTGAGGAATTCCAATCTATTCTCACAACATTTACTCGTTACAGCCGTTTG
CGGGTCATTTCAGAACTTCGACACGGGGATCTTTTCCACTCTGCAAATATTGTATCGAGC
ATTGAGTTTGATCGAGATGATGAACTATTTGCAACTGCCGGTGTTTCCAGGCGGATCAAA
GTGTTTGACTTTGCAACAGTAGTGAATGAACCTGCAGATGTGCACTGTCCTGTTGTGGAG
ATGTCCACTCGATCTAAGCTAAGTTGCTTGAGCTGGAACAAGTGTATCAAGTCTCAGATT
GCCAGCAGCGATTATGAAGGCATTGTTACTGTCTGGGATGTAAATACTCGTCAGAGTGT
CATGATGTATGAAGAACATGAAAAACGGGCATGGAGTGTGGATTTTTCACGCACAGAGC
CGACCAGGCTCATCTCAGGGAGTGATGATGGAAAGGTCAAGGTGTGGTGTACAAGGCA
AGAAACAAGTGTTCTCAATATTGACATGAAAGCAAATATTTGTTGTGTCAAGTACAATCCT
GGATCAAGCTATTATGTTGCGGTTGGTTCAGCAGATCATCACATTCATTATTATGACTTGA
GGAACCCTAGTGTTCCATTATACGAATTTAATGGTCATAGGAAAACAGTTTCCTATGTTAA
ATTCATTTCAACAAATGAACTAGCTTCTGCATCTACAGACAGTACACTGCGCTTGTGGGA
TGTGAGGGATAATTGCCTTGTACGGACATTTAAGGGGCACACTAATGAGAAGAACTTTGT
GGGGCTTACAGTCAACTCTGAATACATAGCTTGTGGCAGTGAAACAAATGGGGTGTTTG
TGTATCACAAGGCTATTTCAAAGCCTGCTGCTTGGCATCAATTTGGAAGTCCAGATTTGG
ACGACAGTGACGATGATACATCACATTTCATAAGCGCTGTTTGTTGGAAGAGTGAGAGC
CCTACAATGTTAGCTGCTAATAGCCAAGGAACAATTAAAGTTCTTGTACTTGCACCGTGA
ATTTCAAGTTTCTCATATAATTCTTCAAGGCCCTTCCTAGATATCAGAGTTGGAGAGAATT
TTGAAATTTCTTAAATTAAGATTCTTTCTGGAAATTAGCTCTCTGGATTTCTTGCTGCCTGA
GAATGTTGAGAAATGAGGTTGATGTGGATTATATGCGCAATTTTCAATCTACTCATATTTC
TATAGTGCCATATGCTTGTCGGTTGTCATTGACCTCTAATAGAATAGCCAGAGTATCTTC
CAGTGATGCGGCATGTTATTGCATGTGCTATATGCCCCTCTTAACACTAATCTGAAGGAA
TATCGGCCTCAATTTGGCAAGTTCTTTGGTCTGGGGCGAGCCAACTGGTGAAGTCTAAG
ACTATGAAGGCTAAATTTTTTAAGGGTGTCAGAACTCGAGGAAAGATTCAGGAGATGGAG
AATGCCCAGGAGTAACATTTTGTACTGATACTAAATGTATTATATCTTGGTCTTTTATCGTT
CTTGGAATACTTGCTAGCTGTCATCACACATTTTTATGCTTGAACTACGAGGAATAGTGAA
GCGTTGCCCCAACTTATGGGCTTCCCTCTTCAACTATTATTTTGTACTGATGCTAAATGTA
TTATATCTTAGTCAGTCTTTTATCGTTCTTCAATACTCGCCTGCTGTCATCAAACTTTTCTA
GGCTTGAACTACAAGGAATAGTGAAGCGTTGCTCTAACTTGTGGGCTTCCCTCTTCAACT
ATTATTGGGATTTGCTTCTAGAGCCAAATAGTTGGACACTATCAGAGCTTAGAGAATTGTT
TATATCGGTTCCCTGTAAAAGTTTGTATTTGAGGGCAGTCCGTGTCCCTTGTCAATACTTA
TGAGTTGCAGCTCATAGATTTTTTCAAGTCCCAAGATCAGCAACACGTAAATGAACATGA
AATAAATATTGTAAAATTGTGTATTGGTAAATTTGTTGAAGAAATAACAA

SEQ ID NO: 227
TTCCTTCCCGCTGAGTTGCTCGACTTCTACTAGGAAGATTTTAGGTATCCCGCAAAGTTT
AACGCATATCAGTCTCTCCTTATTCCCAATGAGTTGCTCGACTTCTATAAGACGATTTTCG
GCATCCCCCAAAGTTATGCGTTATGTTAACGCATATCAGTATCTCCTTGCTATCACAGTT
GGACGCCCAGAGGCAGCTTCCCATTAGTCGAAGCCGCCCTGGAAAAAAAACGGGTAGG
TTTTAAGGTGTTTTTTTTGTTTTTTGTATTACGAAAGTTTTTGAACTTCATATGGTTCCAGA

CCAACCTTGGATGAATTTTAATCAAATTTCAAGTGAACGCGTCTCAGCCTCTCCCCGTGG
TCACAGTTTGACGCTTAGACACAGAATCTCTGCTAGTCCAAGCCACCCCGGAAAAAGCG
GGTTCAGCTTGAAGTTCGGTGGACCTCATCAAATAAATTCGACCGTCAATGAAGTTCTGC
GAATTTCCAATTTAATTGATTTTGTTCGCCGAACGCGGTCACTTATGCTCAAAGCACCAG
TTTAACCATCAACGGCAGAGTCAAAGGCAGCATCACAACCAATTTATATTTTATACCTAGC
AGATCCAATTGATCAAGTAGCATAAGAAGCGTAGGTAGTGGAGGAATAAACAAGGGTGA
AAAAACAAGTTATCGGGACTGTGAATACGTTAGCTGACCAAGTGTATCCATTACGAGTAG
ACCCAGTATCATTCATTGGAGATTGCGGTCCAGCAGTTTCATACCCTATTTTACGGGAAA
AAGCAATAATCCCTACTTACTGGAACAGCAGCATCAGGCCAAGCCCTAGATAAGCTTTCT
GAAGTATTTGGGTAATTGAAACGGTGCGTAGCCATGGCCAACTATGTGGACTCGAAGAA
GAATTTCAAGTGTGTTCCTGCTCTGCAGCAATTTTACACAGGGGGTCCATTCAGGTTGTC
TTCGGACGGTTCATTTTTAGTGTGTGCGTGCAACGATGAAGTAAAAGTGGTAGACTTAGC
CACCGGTTCTGTGAAAAATACGTTGGAGGGAGATTCGGAGCTTATTGTTGCGCTCGCAC
TCACTCCCGATAACAAGTATTTGTTCTCAGCAAGCCGCAGTACTCAAATCAAATTCTGGG
ATTTATCCTCCGCCACCTGCAAACGAACCTGGAAGGCACATAATGGTCCTGTGGCAGAC
ATGGCCTGTGATGCTTCTGGAGGATTGCTTGCTACAGCTGGTGCAGACAGAAGTATTCT
TGTCTGGGATGTGGATGGAGGCTACTGCACACATTCATTTCGAGGGCATCAAGGAGTTG
TAACTACAGTGATATTCCATCCGGATCCCCATTGCCTTTTGCTTTTTTCTGGAAGTGATGA
TGCAACGGTTCGAATATGGGATCTTGTTGCTAAGAAGTGCATTTCTGTGCTGGAAAAGCA
TTTTTCTACTGTGACGAGTTTGGCAATCTCTGAAAATGGTTGGAATTTGCTCAGTGCTGG
TAGAGATAAGGTTGTAAATATATGGGACCTGCGTGATTATCACTGCCGGGCAACAATACC
GACATATGAACCACTAGAAGCTGTTTGTGTTCTCCCAACTGGTTCTAGGCTGGTGTCTGT
TATGAACCAAAGCCGGGCATTGCCTGAAAATCGAAAGAAAAGTGGTGCGGCTCCAGTGT
ATTTTTTGACAGTAGGGGAACGTGGCATTGTTCGTATATGGTATTCAGAAGGTGCTCTGT
GCTTATATGAGCAGAAGTCCTCGGATGCAATAATTAGCTCTGACAAAGACGAGTTGAAG
GGAGGCTTTGTATCTGCTGTTCTACTACCCCTCACTCAAGGAGTGATGTGTGTTACTGCT
GATCAGCGTTTTCTCTTTTACAATCTGGATGAGAGCGATGAAGGAAAATGTGATTTAAAG
GTCAGCAAGCGTCTTATAGGTTACAATGAGGAGATAGTGGACTTGAAATTTCTTGGGGAT
GAGGAGAAGTTTCTTGCAGTTGCAACAAATCTGGAGCAGGTGCGCATGTATGACCTTTC
ATCGATGACATGTGTATACGAGTTGTCAGGACACACAGACATAGTTCTTTGTCTAGATAC
TGTTGTTTTTTCTGGACATAGTCTGTTAGCTTCTGGGTCCAAAGATCACACAGTAAGAATT
TGGGATACAGAGAGCAAATCCTGCATATGTGTAGCAGCAGGGCACATGGGTGCTGTAG
GAGCTGTTGCCTTCTCAAAGAAAGCTAAGAACTTCTTTGTGAGTGGGAGCAGTGATCGC
ACAATCAAAGTATGGAGTTTTGCTAGCGTACTAGATTTTGGTGGTATTAGTAAGTCAATCA
AGCTTTCATCACAGGCTGCAGTTGCAGCTCATGACAAGGACATCAATTCTGTGGCTGTTG
CTCCAAATGATAGTCTTATCTGTACTGGTTCTCAGGATCGAACAGCTCGTATATGGAGAT
TACCAGATCTGGTACCAGTACTAGTGCTAAGAGGCCACAAAAGAGGAGTTTGGTGTGTG
GAATTTTCACCTGTAGATCAATGTGTGATGACAGCATCTGGTGACAAGACAATCAAAATA
TGGGCATTATCGGATGGCTCTTGCTTGAAGACATTTGAAGGTCATACTGCAAGTGTACTG
CGAGCTTCCTTCCTCACTCGTGGGACACAGTTTGTTTCTTCTGGTGCAGATGGCTTGTTG
AAATTGTGGACAATCAAGTCCAATGAATGTATTGCAACATTTGATCAGCACGAGGACAAG
ATCTGGGCAATGGCTGTGGGCAAGAAAACAGAAATGCTGGCAACAGGTGGCAGTGATTC
CCTTGTGAATTTGTGGCATGATTGCACTACCACAGATGAGGAGGAGGCTCTTCTCAAAG
AGGAGGAAGCGGCATTGAAAGACCAGGAATTACTAAATGCTCTTGCAGATACCGACTAT
GTGAAAGCAATTCAACTTGCATTTGAATTAAGAAGACCTTATAAGCTCCTTAATGTCTTCA
CTGAACTGTACAGCAAAGGACATGCCCAGGACCAGATACAGAAAGTGATACGTGAACTA
GGAAATGAAGAGCTAAGATTACTTTTAGAATATGTGCGAGAATGGAATACAAAGCCAAAA
TTTGCCCATGTTGCCCAGTTTGTATTATTTCAACTTTTCAACGTCCTTCCTCCAAAGGAGA
TCATTGAGGTACAAGGTATTAGTGAACTTTTGGAGGGTCTTATACCATATGCACAGCGAC
ATTATAGCAGGATAGATAGACTTATGAGAAGCACTTTTTTGTTGGATTACACTCTTTCATC
AATGTCAGTTCTCTCTCCCACGGAGACAGATTTGTCATCATCTAATTTATTAGCTAGAACA
GCTGATCCACTGCATGCACAGATTGATCAGTTTCACCCAACCCATTTCCCTGAGCCAAAT
TTGACTCCTATACAGTCTCTTTTAGATTCAGGCAATACAGATTCTGTGGAAGTCACAGCA
CGACGAGCCAAGAAGAAAAGAGTGTCAGGTAATGATTCAGAGAAAACAACTGTTGCTGA
AGTTAAAATTGGTGACATGGAGAATGCATTTGATGAGCCTGACGTGGCAGATCAAGGTT
CATCCAGGAAGCATAAACCTGCGAGTTCAAAGAAACGGAAGTCTATAGCCGTTGGAAAT
GCCAGTATTAAGCGTATTGCAAGTGGAAATGCAGTCACTATTGCATTGCAAGTATAAGGC
CTTGTTGATGAAGTTTGATCTACCATTGGTTTTGGAATTAATGTGAACTCTGTAGCAGGAT
TGCTGAGCAGGATTTAGAGGAACGGCTGTATGTTAACTTGAAATTTGTGTGTCGGTCAAG
GGTTGCGGGTGCAACATAGAGAAATGTGTGCCAACATTTGGTGTACCGAATGTCTTCAAT

CTGTAAGATTTTCACGGCAGTTGAACTAGTCATAGTGGAATATTATTTAAATGGTGTATTC
TAGTCACATTATTTCCAATTTTGATAGTATCTGGGCATCTGCTCATTAATTCAAAAAAAAAA

SEQ ID NO: 228
CCACATCAATTGACCTGGTTTGTCTGATTGTGCAGATAGCTTTGCTCCGGGGCAAGGAG
AACGCCAAGAGGACACCAGGGCTTCTGATACTTGGCTCGCTCGCTAACTTTTTGGAGAA
CCAGACGTCCTCATACTTTGAGGAACCTAAGCTTTTAGAATTGGAATTCATTTGGAACTA
CTGTTTTTTTTTTTTTTCGTTTTGCAGAACCAAACATCCTCATTCTTTGAGGAACCCTAGCTT
TTAGAATTAAAATTCCTTTGCGACTTTTGATTTTTTTGGTTTTGCGAGTTTAAATTTAGGCA
CATTGATTGTTTGTCATATTTTGTGTTTTGTCTGGTTATTTTACTTGAGTAGGAGTAATGGA
ATCGTCTTGCAGTTCAATGAATTCCAACAGGCATTCCACAGAGAAGCGATGTCTACGTCC
TCTGCAGAAACAGGGTGCTTCAATGAATAAACACAGCAGCGATAGATTTATTCCGGCAAG
AGGATCCATCGATTTGGACGTAGCTAGATTCATGGTCACGCAAAAGCAAAAGGACAATAA
CGATATCCATGCCCTCTCCCCTTCCCCCTCCCCTTCCAAAAAAGCATACCAGAAGGAGAT
GGCAGACACTTTATTGAAGAATGCCGGTGCTGCTGACAACAATTGCAGAATTCTTTCTTT
CAATGGAAAATCATCGACTGTTTCCCAAGGTTCACAAGAAAATGTGTTGGCTAATCTCTC
TATTTCTAGAAGAGCACGCAGATATATTCCTCAGTCTGCAGATAGAACCCTGGATGCGCC
CGACCTTCTCGATGATTATTACTTGAATCTATTGGATTGGAGTTCCACCAATGTTTTGTCC
ACAGCCCTGGGCAACACTGTATACTTGTGGGATGCTTCTAATAGCTCTATATCTGAGCTT
CTGATTGCAGATGAAGAAGAGGGTCCCGTTACCAGCGTTAGTTGGGCACCCGATGGGA
GCCAAATTGCCGTTGGGTTGAATAATTCTGTTGTACAGTTATGGGATTCGCAATCTAATA
AAAAGCTAAGAGCCTTGAAAGGTCATCATGACCGTGTTGGTGCACTTTCTTGGAATGGTC
CTATTCTCACCACAGGAGGGCTGGATGGGATTATAATCAATCATGATGTTCGCACCCGT
GACCACATTGTTCAAACATACAAAGGACATACCCAGGAGGTTTGTGGATTGAAATGGTCT
CCTTCCGGCCAACAACTTGCAAGTGGGGGAAATGACAATCTACTCTATATTTGGGACAA
GAGCATGGCATCCCATAATCCGTCCTCTCAATACTTCCACCAATTAGATGAACATTGCGC
AGCAGTCAAGGCCTTGGCATGGTGCCCCTTCCAGACCAATCTTTTGGCATCTGGTGGAG
GGACTTCAGATGGTTCTATAAAGTTTTGGAACACTCAAACAGGTGCCTGCCTGAATACCG
TCGACACCCATTCCCAGGTCTGTTCATTGTTGTGGAATAGACATGAGCGTGAGTTATTGA
GTTCCCATGGGCTAAATCAGAACCAGTTAACCTTATGGAAATACCCATCCATGGTGAAAA
TTACTGAACTTACAGGGCATACTGCTAGAGTTCTTCATATGGCACAGAGTCCTGATGGGT
ACACAGTTGCATCAGCTGCAGCAGATGAAACACTCAAGTTTTGGCAAGTTTTTGGGGCTC
CTGATGCCTCAAAGAAGACTAAGACTAAAGATACAAAAGGGGCTTTCAACATGTTTCATA
TGCACATCCGCTAAGAAGTATCTGTCAAATTTCATGGAATCAGGATGGATGTTCAACTGA
AGAGGGGCTTTCAGCAAGTTTCACGCTCACATCCGCTAAGTATTTGTCCAGATTCATGGA
ATCAGAGCGGATGTTCAACTCAAGGTAGTAATTAGATGTTTTATTTTCTGCTTGCAAAGGT
CTTTTGCAGGCGCTCTATAGTTCTGTTCTCTAGCATGAAGTGTGTATTTTATCTATTGTGG
ACCTTGATGGAATGATGTCGCAACAGCCTTGTGAGACAATGATCCATTTCTTCAGCTGCG
TGATAAGATTTAAAGTGGTTTTAGCAAGTCTTTGGAGCTTCAAAGACGGTTCCAAAAATG
ATACTTTATGGCTGTTAGCAAGCATCTGATTCACAACTATAAAAAGGCCTCCATGCTTCTA
CACATGGAATGATAATGAGGGGATGCTCAATGTTGCAGAAGATCGTTCAAGGTGCCTTTA
ACGAACATGGGATCTTGCCAAGGTGCCTTTAAGGAACATGGGATCTTGCCAGCATTCTA
AGATGTTCGAGAATGGGTTCATACAGAGAAGAGCAGCATGGTTTGTCAATTGCATTTTTT
TAATATCAAGATCTTAATGTACGCTCTTTTGTTATCCATTCTGTACTTAGCTAGTTGCACTA
CAATTTTCCAGCAAAAAAAAAA

SEQ ID NO: 229
TGAGTTTTCAAGCAACTGTTAGGCCTTATACTTATGTTGGGTGCCGGAGTCTAGTGGAGT
TATTTTGGCGAGATACAGGGGAAGAGCGCAAAAACAAGTTAGGCGGCACTGTGATGGTG
TGGTGGCTGTATGCAAGACACTAGAAGAAGAAGAAGATCGAGAGAGGAGAGCAACGGC
AACGAGAATGCTAGACGAAATTGTGGCAGACGAGGAGGAGGAATTCAACATATGGAAGA
AGAACACCCCCCTTTTATACGATGTAGTGATAACGCATGCGCTGGAATGGCCTTCCCTCA
CGGTACAGTGGCTCCCCGACCGCCACCAATCTCCCACAAAGGACTACTCTCTGCAGAAA
ATGATAGTGGGCACTCATACATCCGGGGACGAGCCCAATTATTTGATGATAGCAGAGGT
CCAGATGCCCCTGCAATACTCTGAAGACGGTAACGTTGGAGGTTTTGAATCCACCGAAG
CCAAGGTACACATAATCCAGCAGATTAATCATGAAGGAGAAGTTAATAGGGCTCAATACA
TGCCACAGAACTCATTTATTATTGCAACCAAGACAGTGAGCTCAGATGTTTATGTGTTTGA
TTACACTAAGCATTCATCAAATGCCCCTCAAGAAAGGGTTTGCAATCCTGAATTGATATTG
AAGGGTCATACTAATGAAGGATACAGTTTGTCCTGGAGCCCTCTCAAAGAAGGTCAACTA
CTAAGTGGTTCAAATGATGCGCAAATATGCTTCTGGGACATTAATGCTGCATCTGGCAGG

AAAGTGGTTGAGGCCAAGCAAATATTTAAGGTCCATGAGGGGGCAGTTGAGGATGTTTC
GTGGCATTTGAAACATGAATATCTTTTTGGATCTGTTGGAGATGATTGTCATTTGCTTATA
TGGGACACACGTACGGCTGCACCTAACAAGCCTCAACATTCAGTTGTAGCTCATGAGAG
TGAGGTTAACTCTTTGGCTTTTAATCCCTTCAATGAGTGGCTTTTAGCAACAGGATCAGC
GGACAAAACTGTTAAATTATTTGATCTCCGGAAACTCTCTTGTTCACTACATACATTTTCA
AATCATACGGAAGAAGTCTTTCAAATAGAGTGGAGTCCTATGAATGAGACAATATTAGCT
TCGTCTGGTGGTGACAGAAGACTCATGGTTTGGGACCTCAGAAGAATTGGCGACGAGCA
AACATCAGAAGATGCAGAAGATGGACCACCTGAATTGATTTTCATTCATGGAGGGCACAC
AAGCAAGATATCTGATTTCTCATGGAATCTCCATGATGATTGGTTAATTGCTAGTGTAGCA
GAAGATAATATTTTACAAATTTGGCAGATGGCAGAAAATATTTACCATGATGATGCGGATA
TTCTATGATGAAATTTCATAACTCTTACCTACTATGGTTGGTGTGCCTTGGAGAAAATTTC
CAATGCCTTCATGGGTGTGTAATCCATGATAATTACAGGAACATTGTGGATCAAATATTGT
AATTTGTCTTTAATCTTCAGGGTTCGTTACTAACAATTGAGCTCAAATCTCTATTCTGACC
AGCAATCCAGAAGTTGGATGCATTAAATATAAGGTGGATTTTAGTACCTTAAGTTAATGTA
TTTTATAAAATTATGAAGCACTCTGGACATCTGCTCTGTTCATCAAGGCAGACATCACAAG
GAAGTAGAAATTGAAGTTATGTTTCAAAAAAAAAAA

SEQ ID NO: 230
CACAAGCACAGCAGTGTGCGTTAAAGAAAGGGCGGTTTGTATTGTATAAATGGTAAATCC
GAGCGCAGAGGATGGAATGGCTACAGGTATTTATGCAGTATGCAGTGAACAGAGATGGC
AGGCTGGCCGTGATATTTTTGTGTTTATACCACATGGTCTGAGGGCGATTGATTTTGTCA
GCTGATATCTCTCAGGGCCCTGAGCCTGCCTCCCTGCCTGGCCGCCTGTGTATGAATAT
CAATGACGAAAGAAGATCATGGGGAGTCGAGGGATGAGATGGGGGAAAGAATGGTGAA
TGAGGAATACAAATTATGGAAAAAGAACACCCCATTCCTGTACGATTTGGTCATCACGCA
TGCTCTGGAGTGGCCTTCTTTAACTGTTCAATGGCTGCCTCCCAGCTGCAAGCAGCAAC
AGGACATTATCAAGGATGATGACATTGATCATCCCAACACCCAAATGGTTATTCTCGGTA
CCCATACATCAGATAACGAACCCAACTATCTAATTTTAGCTGAGGTCCAGCTCCATGATG
GAACTGAGGATGAAGATGGCGATGGCGATGTCAAACGGCCCCAGGATAAAATGAAACC
GGGAACCTCTGGAGGTGCAATGGGGAAGGTTCGAATACTTCAGCAGATAAACCATCAGA
AGGAAGTGAACCGGGCGAGATACATGCCTCAAAAACCCACAATCATTGCAACAAAGACT
GTGAATGCGGATGTTTATGTCTTCGATTACAGTAAACACCCGTCCAAGCCTCCCCAGGA
GGGACGTTGCAATCCGGAACTGCGGCTTCAAGGTCACGAGTCTGAAGGCTATGGCTTGT
CTTGGAGCCCTCTCAAGGAAGGTCATCTGTTGAGTGCTTCAGATGATGCCCAAATTTGC
CTCTGGGACATCACTGCTGCGACCAAAGCCCCCAAGGTGGTGGAAGCCAACCAAATATT
TCGGTACCACGATGGACCAGTGGAGGATGTAGCATGGCACGCAATTCATGATCATCTTT
TTGGGTCTGTAGGGGATGACCATCATTTGCTTCTATGGGATATCAGGAACGACTCTGAG
AAGCCACTTCATATTGTTGAAGCTCACCAAGCTGAGGTTAACTGTTTGGCTTTCAATCCTT
TCAATGAATGGATTGTGGCTACAGGCTCTGCAGACAGAACTGTTGCACTCCATGACATTC
GCAAGCTGGACAAAGTTCTTCATACTTGTGCACACCACATGGAAGAAGTGTTCCAAATTG
GTTGGAGTCCCCAAAATGGAGCTATATTAGCATCATGTGGATCGGATAGACGACTCATG
GTTTGGGACCTCAGTAGAATTGGCGATGAGCAAAATCCAGAGGATGCTGAGGAGGCAC
CTCCTGAACTGCTATTCATACATGGAGGACATACCAGCAAGATTTCAGACTTTTCATGGA
ATCCAGCTGAGGAATGGGTAATTGCCAGTGTTGCTGAAGACAACATTCTTCAAGTTTGGC
AGATGTCAGAGCACATTTACAATGATGACAATGATTCACCAACAGCTTAATAGGTTTGTT
GCTGTTTGGACCAGCTAATTGATTGTTTACAGGACCTCGGAAAAATGTGGAGCCTCTTTA
TGCCATATTCCATGGCACAGGGTTGAAGAAGCCTGAATATCAGCAGAAAGCGAGCTTGT
ATATTGCAGCGATGTTAATGATTCTCTTTTATAAGAGGGAAAAAGTTTATAAGCACTTGTG
AAAATTGTCTTGGACGTTTAATTTCTCTTGTGAAAGTATGGAGCTTCTCATCATTTATAGA
GTTGTGCAAAATCACCCATAATGCTATGAATTGACAGGTGACTGTAATCTTACAATTTTCT
GATGCAATTTGAAGATGGATTATGCAAGTCTTTTATCAAAAAAAAAA

SEQ ID NO: 231
TTAGTAGAAGTCAGGCTTAGCTTAATAGTAGCGGAAATAGGAGCTAAACTGTGAGACCC
GCCGGTTAGTATGCAGATACCACTTAGGGATTCCACTTACGAGTCCATTGGGGTTATTTT
GGCGGGAGACAGTCGAAGAGCACAAAATCGAAAACAAGTGGTATATTGTGGCTCTGCAA
TGGTGTGTGCAATAGAATAGAGAAGAAAAATAGCAGAAGTCAGAGGAAAGCATGGCGAT
GGCAATGGGAGACGAAAATGCGGCCGATCCAGTTGAAGAATTCAACATATGGAAGAAGA
ACACACCCTTCTTATACGATCTAGTGATAACTCATGCGCTGGAATGGCCTTCCCTCACTG
TTCAGTGGCTCCCCGACCGCCACCAATCATCCACGGCGGACTACTCCCTGCAGAAAATG
ATAGTGGGTACTCATACTTCAGAGGACGAGCCCAATTATTTGATGATAGCAGAGGTCCA

GATTCCCCTGCAAAACTCTGAAGATAATATCATTGGAGGTTTTGAATCCACCGAAGCTAA
GGTACAAATAATCCAGAAGATTAATCATGAAGGAGAAGTTAACAAGGCTCGCTACATGCC
ACAGAATTCATTTGTCATTGCAACGAAGACAGTGAGCTCAGATGTTTATGTGTTTGATTAC
AGCAAGCATCCATCAAAGGCTCCTCAAGAAAGGGTTTGCAATCCTGAATTGATATTGAAA
GGTCATTCTAATGAAGGATATGGTTTGTCATGGAGCCCTCTCAAAGAAGGTTACCTACTA
AGTGGTTCAAACGATGCACAAATATGTTTGTGGGATATAAACGCTGCATTTGGAAAGAAA
GTGCTTGAGGCCAATCAAATATTTAAGGTCCATGAGGGGGCAGTTGGGGATGTTTCATG
GCATTTGAAGCACGAATACCTTTTTGGCTCTGTTGGAGATGATTGCCATTTGCTCATATG
GGACATGCGTACAGCTGCACCTAACAAGCCTCAACAATCAGTTATAGCTCATCAGAGTG
AGGTTAACTCTTTGGCCTTTAATCCGTTCAATGAGTGGCTGTTGGCAACAGGGTCAATGG
ACAAAACTGTCAAATTATTTGATCTTCGGAAACTCAGTTGTTCGTTACATACATTTTCCAAT
CATACGGACCAAGTCTTTCAAATAGAGTGGAGTCCGATGAATGAGACAATATTAGCTTCT
TCTGGTGCTGACAGAAGACTCATGGTTTGGGACCTTGCAAGAATTGGCGAAACACCAGA
AGATGAGGAAGATGGGCCACCTGAGTTGCTTTTTGTTCATGGAGGGCACACTAGCAAGA
TATCGGATTTCTCATGGAATTTGAATGATGATCGGGTAATTGCTAGTGTAGCAGAAGATA
ACATTTTGCAAATTTGGCAGATGGCAGAAAATATTTACCATGATGATGAGGATATGCTTTG
AAGAAATTTCATAATTCCTACCTACTATCGTTGGTTTTCTGTGGAGAAAATTTCCTATCCC
TTTGTGGGTGTGTGAAAAACGAAATATAGAGGAACAATGTGGATAATGCATTGTAATTCA
TTTTTGTATTGAAGAACAATTAATCTCAAATCTTTATCTCCAATGATTCTCTATTTTCTGTC
CGGATTAAGTATAAGGCAGATTTCAGTATATAAGGTCAATGTATTTATAAAAGTATGAAAC
AGACTGGACATTTG

SEQ ID NO: 232
CAGGAAACCTTTACCACGACCAACTCCAGCCAGCTAACAATTTCCGACTGTGTACACGG
ACACAAAGTTTATTGCAAATATCTTCCCACGATCTGTAGAGATGGGATTGTTTGAGCCCT
TCAGAGCACTGGGATACATCACTGATGGCGTTCCATTCGCTGTTCAAAGGCGTGGGATT
GAGACTTTTGTTACCCTGAGCGTTGGAAAAGCTTGGCAAATTTATAATTGTGCCAAGCTG
ATTCCAGTCCTTGTAGGACCTCAAATGGATAAAAAGATAAGAGCACTTGCATGTTGGCGA
GATTTTACATTTGCTGCCACTGGGCATGACATTGCAGTGTTTAGGCGAGCTCATCAGGTG
GCTACATGGAGTGGACACAAGGCAAAAGTTACATTGCTTCTTTCTTTTGGGCAACATGTT
TTAAGTGTTGACTTAGAAGGATGCCTTTTTATATGGGCTGTTGCAGAAGTGAATCAAAAC
AAACCACCAATTGGACAAATACAACTTGGGGAGAAGTTTTCTCCAAGTTGTATCATGCAT
CCAGACACTTACCTAAACAAGGTCCTTATTGGGAGTGAAGAAGGTACCTTGCAACTTTGG
AATGTGAATACACGTAAAAAGCTTTATGAGTTTAAGGGCTGGGGGTCTTCCATTCGATGC
TGCGTCTCATCTCCAGCTCTAGATGTGGTTGGAATTGGTTGCTCAGATGGCAAAATCCAT
GTCCATAATCTACGATATGACGAAGAGATAGTGACATTTATGCATTCAACAAGAGGGGCT
GTTACTGCTTTGTCTTTCAGGACAGATGGGCAACCTCTTTTAGCAGCTGGAGGTTCCTCT
GGGGTGATAAGTATATGGAACCTTGAAAAGAAAAAGTTGCAGTCTGTAATAAAAGATGCT
CATGATTCTTCAGTATGTTCTCTTCATTTTTTTGCAAATGAACCTGTTCTGATGAGTTCTG
CAACAGATAATTCAATCAAGATGTGGATCTTTGACACAACTGATGGAGAAGCGCGACTCT
TAAAGTACAGAAGTGGCCATAGTGCACCTCCTATGTGTATAAGGTATTATGGTAAAGGGC
GGCATATTTTATCTGCTGGACAAGATCGAGCTTTCCGGATTTTCTCTGTTATACAGGATC
AACAAAGTAGAGAACTTTCACAGGGTCATGTTGGAAAACGAGCAAAGAAATTAAAAGTGA
AGGATGAAGAAATCAAACTGCCTCCAGTCATTGCTTTTGATGCAGCTGAGATTCGTGAAA
GAGACTGGTGCAATGTTGTCACCTGCCATTTGGATGATCCTTGTGCATACACATGGCGT
CTTCAGAACTTTGTAATTGGTGAACATATTTTGAAGCCATGTTTGGAGGATCCAACACCA
GTGAAGTCTTGCAGTATCAGTGCATGCGGCAATTTTGCAGTGCTTGGAACAGAAGGAGG
GTGGCTTGAGCGGTTTAATCTTCAATCGGGGATTAGTCGAGGGACATATATTGATATAGG
AGAAAAACGACAATGTGCACACAATGGAGCAGTGGTTGGTCTGGCATGTGATGCTACAA
ATACCCTTTTAATAAGTGGAGGCTATAACGGAGACATTAAGGTTTGGGATTTCAAGGGGC
GTGAGCTAAAGTTCCGATGGGAAATTGAGGTCCCATTAATTAAGATCGTATATCACCCAG
GAAATGGTATTCTAGCAACTGCAGCAGATGATATGATTCTTCGTTTATTCGATGTCACTG
CCATGCGACTTGTGCGCATATTTGTTGGGCACATGGACCGTGTCACTGATTTGTGTTTTA
GTGGAGATGGGAAGTGGTTGTTGTCATCTAGTATGGATGGGACCATAAGAGTTTGGGAT
ATAATTTCATCCAGGCAACTCAATGCAATGCACATGGATTCAGCTGTAACGGCTTTGTCA
TTGTCACCGGGTATGGATATGCTGGCAACTACTCATGTTGGTCACAATGGCATCTATCTC
TGGGCAAATCGGATGATATATTCAAAGGCTACTGACATTGAGCCCTTTATAAGTGGGAAG
CAAGTTGTGAAGGTTTCTATGCCAACTGTCTCATCCAAAAGAGAATCGGAGGAAGGGGA
TGAGAAAAGAACAATTGTTGCAGAATCAAATGTCAATAAGTCTGATGTTTCTGGTAGCTT
GATTGGGGATTCATATTCAGCTCAACTTACGCCTGAACTTGTGACGCTTGCACTGCTGCC

TAAAGCTCAGTGGCAAAGTCTTGTCAATTTAGATATTATAAAGATGCGAAATAAGCCCATT
GAACCACCCAAGAAGCCAGAGAAAGCTCCTTTCTTCTTGCCATCATTGCCTACACTCTCT
GGAGAGAGGATTTTTATCCCTAGCTCTATGAATGGGGATGGTGATCAAGATGAAACAAG
AAATGATAAAACTGTTTTTGAAGCAAGGGGTAAAAAACTGGGAGGAGAATCTTTATCATTT
ATGCAACTTCTACAATCCTGCGCAAAGATCAAAGATTTTACAACTTTCACAAATTATTTGA
AGGGTTTATCCCCCTCAGCCGTGGATATGGAATTGCGGTTGCTTCAAATTGTCGATAATG
AAAATATATCAGAAACAGAACACAGTGTAGAACTTCAAGGGATAGGAATGCTGCTGGATT
ACTTTGTAAATGAGGTGTCATGCAATAATAATTTTGAGTTTGTCCAGGCTTTGATTCGACT
ATTTTTAAAGATACACGGTGAAACCATTCGATGCCAAGTATCTTTGCAGGAGAAGGCAAG
GAAGCTACTTGAGATCCAAAGTTCTACATGGGAGAGGCTAGATACAAGCTTCCAAAATGC
GAGATGTATGATTACATTTCTTAGCAGTTCTCAGTTTTGACCAATCATTTATTTGCAGTGT
AGTTGATATGAAGGGAGAAATATGACAGTTGGTTTCAACCTGCTTGCCTATCCCTCCAAA
GAATATATAGTGTAAATTGTCTGTTGCAATAGGGATGGAGCACTAAGTGCTTACTATGGC
AAATGCTATCAATTTTTGAGGGCTTGTAATGGCATTATTTCCTTGAGTTGTTCTTGTAGGT
TGTTGAGTTACCAGTTTTGATCCTTTGATCTTTTGTTGGTTAGCTGAGGCTAATGGCGAG
GCCATGTTTTTCAGTTAATGTAAAATATGATATCTTGTGTATCATGTGACACATTTCAATTC
AGTTATTGGCTTGCAACTTCTAATTACACTCAAAAAAAAAA

SEQ ID NO: 233
CGAGGACGTATCGATCAAAAAATCTCTATCATAAGGTGCAGAATGATTGCAGCAGTGTGT
TGGGTTCCTAAGGGTGTTGCAAAGGTCCTACCAGATTCTGCTGAACCGCCTACTCAAGA
GGAGATTCAAGAGTTATTGAAGTGTAACGTTGTAGCAGAGAGTGATGATAATGAAGATAG
TGATGAAGAAAGTGAAGAGATGGATACTGAAACTGATAAAAATACCGACGCAGTGGCAA
AGGCATTAGCAGCAGCAAATGCTCTTGGGAGTCAGTCTTCAGATTTTCAAAGGCAGCAC
AAGGTTGATGATATTGCTAATGGCCTCAAAGAACTTGACATGGATCATTATGACGACGAG
GATGAAGGCATTGACATTTTTGGAAGTGGTTCTCTTGGTAATTGTTACTACCCAGCCAAT
GACATGGACCCTTATCTTGTAGAACAAGATGATGATGATGAAGATGAAATCGAAGACATG
ACAATTAAACCTTCTGACTTGATTATCTTATCAGCTCGTAATGAAGACGATGTTAGCCATC
TTGAGGTCTGGATATATGAAGAAGAAACTGAAGAGGGTGGTTCCAATATGTATGTTCATC
ATGATATCATCTTGCCAGCCTTTCCACTTTCTCTTGCTTGGCTTGATTGTAATCTGAAAGG
TGGAGAAAAAGGAAATTTTGTTGCTGTTGGAACAATGCAACCTGAAATTGAGCTCTGGGA
CCTTGATGTACTTGATGAGGTTGAACCTGCTGTAGTCCTGGGTGGTGCTGTTAAAGATGA
AGCATCTGGTAAGACGACCAAACTGAAGAAAAAGAAGAAAAATAAACAGGCTGTCAATTT
TAAAGAGGGAAGCCATACGGATGCGGTTCTTGGTTTGGCATGGAATATGGAGTACAGGA
ATGTCTTGGCTAGTGCAAGTGCTGACAAATCTGTAAAGATCTGGGATATTGTTGCTGAGA
AATGTGAGCACACAATGCAACCTCACACAGATAAGGTTCAAGCAGTAGCCTGGAATCCT
AATCAGGCAACAGTTCTTCTCAGTGGATCTTTTGATCGTTCTGTGATCATGATGGACATG
AGGGCTCCCACACACTCAGGGATACGGTGGCCAGTTCCAGCAGATGTAGAAAGCCTTG
CATGGGATCCTCATACTGATCACTCATTTATGGTTAGCGCTGAAGATGGCACTGTTCGAG
GTTTTGATATTCGTGCAGCAGCATCTACTGCAGATTTTGATGGCAAGCCAATGTTCATAC
TTCATGCACATGACAAGGCTGTCTGCGCAATTTCATACAACCCTGCTGCCCCAAGTTTAC
TCACAACAGGATCAACAGATAAGATGGTAAAGTTGTGGGATATAACAAATAATCAGCCTT
CATGCATTGCCTCAACAAATCCAAATGTTGGAGCTGTGTTTTCTGCAGCTTTTTCAAAAAA
CAGTCCGTTCTTACTTGCTACTGGAGGCTCAAAGGGAATTTTGCATGTTTGGGATACCCT
AGATAATTCTGAAGTCGCGAGAAGATTTGGAAAGTTCAGACCACAGAATTGAACTTTCAT
ATTTCTATTCAGTTTTGTTCGAAGAATAATATCGATAAAATTTATTTTCATATTCTCCAGAG
GTGGGATTGTGTATTATAATGTGTTAATACTTGGTCATTTATTTGCGGCTTTTGTTGGCAC
ATTGTTGTGCACATTGCTAAATTTCTACGAGCCTTCGGGGACAAGTGTCTAACTGTCCCA
TCAGTAACATTTCTCGAAGTTTTGCTAAAAGTTAATGTTCTCATAGGTTATTCAAAAAAAAA
A

SEQ ID NO: 234
CCAAATTCCCGCAATCTCACATTTGAAACCTGTGCTGCCCTCCATTTTATGATCTCAATTC
AACTCGAGTCTAACCTGAGCCAAACTCGATCTCTGCGCATAAAAAACTAGATTCTCCGGG
TTTCATTACCCTTTCCTCGGGCATAAATTCATTTGGGTATTTGAATATTCACCTAGGTATT
CAACATTTCTATTGGGTAGTAAAATGATTATGGATGAGAACGAGTTCTGTGATATTTTCTC
CCTGAGAAAAAGGCTATGCCTGCTGTCTTCCCAGGAGGGGGAGGAAGAAGAAGAATTA
GAGGCAATGTCACAGCTCGATGCCGGAGAATTCACCGTGACAGGCAATGAAGAGGTAG
TTGCCATTGCCGAAGATGATGTAAACACTGGAATCCTCAGTCAGGATCTTTTCAGTTCAC
AAGACTATTGTACACCATCCCAACCACAGGATTCGACCGATTTGGATAGCAAGGATAAG

GCTCCATGCCCACTATCTCCTGTGAAAAGCACGATTCAGAGAAAGAGGTGTAGACCAGA
GCTCCTAAGTAATCCACCTGATAGCATCCAGTTCTCCTTTCAACGATTGGAAAGAGTTAG
AAGTGAAGAATCTATCCAGTCCTCCTCTCAACAATTAGCAAGAGTTAGAAGTGAAGTATC
TAGTTCAGATGATTTCAAAACTCCAAAAATAACCGCATCTGGTCAAAAGAATTATGTCTCA
CAGTCGGCATTGGCTCTACGGGCTCGTGTCATGTCACCACCTTGCATTAAGAATCCATAT
TTAGATGAAAATGAAGAGTTAAATGAAAAGATTCAAAGAAGCACTCGTCGATCTCCCGCT
TGTGTAACTCCTATCCAAAGTGGAGCCTGTTTGTCCCGGTATCGTGCAGATTTTCATGAG
CTAGAGGAAATTGGTCGTGGAAATTTCAGTCGTGTTTACAAAGCTCTGAATAGACTTGAT
GGCTGCTGTTATGCAGTGAAGTGTTCACAAAGCGAGTTACGACTTGACACTGAAAGAAA
GGTGGCATTGATGGAAGTTCAATCGCTGGCAGCCCTAGGACCGCACAAAAATATTGTGG
GATATCACACTGCCTGGTTTGAGAATGATCATTTGTACATTCAAATGGAGCTTTGTGACC
ATAATCTAACAACTGCAAATGACAGAGGAATTCTAAGAACAGACACGGACTTTCTTGAAG
CAGTGTATCAAATTGCTCAGGCATTGGAGTTTATCCACGGGCGTGGAGTTGCACATCTA
GATGTCAAACCCGAGAATATTTATGTTCGTGATGGTACTTACAAGCTTGGAGACTTTGGT
CGTGCTACACTTATTAATGGAACACTCCATGTCGAAGAAGGAGATGCACGGTATATGTCC
CGTGAAATCCTGAATGATAATTATGAACATCTTGACAAGGTGGACATGTTCTCATTGGGT
GCGACCTTTTTTTGAACTTTTGATGAGAAAGCAATACCCTGGTTCTGGCAAACGGATAGAC
AGGGACACAGAAATTAAGATCCCAATTCTCCCTGGTTTTTCTATATATTTTCAAAAGCTTC
TACAGGATTTGGTCAGCAATGATCCGGGAAAAAGACCATCTGCTAAAGATGTACTCAAAA
ATCCTATATTTAATAAGGTTCGGGGAGCAAAGGAAGTCTAAGGATATGGACGAAGGATG
AATTAATTTTGAAAATTTTGACAATTCTGCAGGATGGATACTTTTGTCAAGATACATGCAA
AATTCACTTTCCTCGTGGAACGAGGATCGACTTTGGTGCTATACGGAGGACTTGCGATAA
TGCAAAACTTATAGCTGTTACTTAAGGATTTCATTGCAACTAAAAATCTTTGGCTGTAAAC
AATCAAGTAATGATAGATGTGATATGCTGGAGATTTATATTTTGGACGGAAGAGTGGAGT
TATTAGTTTGCAATATTTTTGTTAAATATGCATGGTGAACAGGCAAGATCACTGTGCTTCC
TCATAAATTGAGGCTTGCCTACGTTAATTGTTATATATGGAGAGCCATGCTAATTGTTATA
TATGGGAAGCTCTATCCCAATTGTAATTTTGTAATTTTCAATCACCTTCTTTGGTGGTATTT
CAGTTCTTGCTGTAATCACGACCTTCTGTGCTGATACAGTTTATTAATCAAATGTAATCAT
CACCTTCTGTAGTGATGTTCATGTTAAAAAAAAAA

SEQ ID NO: 235
GCTTTGCTTCCAGCACAGTGCAGAGGGGCAACGCAGCACAAAGCTGAACCGCCAAAAC
CCTAATAAGCATTTCCAAGCCACCGTCCACCGGAGGATCGGAGCGGCAAAGCAGGTCG
ATGCTGGCGCCGGCGCTGGAGATGGAGCCGGTGGAGCCGCAGTCCCTGAAGAAGCTC
AGCTTCAAGTCCCTCAAGCGCGCCCTCGACCTCTTCTCCCCCGTCCACGGCCAGATCGC
TCCTCCCGACCCCGAGAGCAAGAAGATGCGCATCAGCTACAAGCTTAATTTCGAGTACG
GTGGAGGTAGCGGGAGTGAGGACCAGGTTCCAAAGCGGAAGGAGAGCGGCGCCGCGC
AGAATCAAGGCCAACAAGCTGCCGGGGCTTCCAATGCGCTTGCTCTTCCGGGTCCAGAA
GGTTCTAAAATTCCACCAATGGAAAAGTCTCAAAATGCTTTAACTGTTGGTCCATCATTAC
GACCTCAAGGATTAAATGATGTTGGTCTACATGGAAAAGGCACTGCTATTATCTCTGCTT
CTGGATCATCAGACAGGAATTTGTCTACCTCAGCCATTATGGAAAGACTTCCAAGCAGGT
GGCCACGTCCTGTGTGGCATCCTCCATGGAAGAACTACAGGGTCATAAGTGGGCATTTG
GGATGGGTGAGATCTATTGCATTTGATCCTAGCAATCAATGGTTCTGCACTGGCTCTGCA
GATCGGACAATTAAGATATGGGATCTAGCAAGCGGAAGGCTGAAGCTAACACTGACAGG
ACATATTGAACAGATTAGGGGCCTTGCAGTGAGTAGTAAGCATACTTATATGTTCTCTGC
TGGTGATGACAAACAAGTTAAATGCTGGGATCTTGAGCAGAATAAGGTTATCCGGTCTTA
TCATGGTCATCTTAGTGGTGTTTACTGCTTGGCTCTTCATCCTACTATTGACATTTTGCTC
ACTGGAGGACGTGATTCTGTCTGTCGGGTATGGGACATCCGGAGTAAAATGCAAATTTTT
GCACTCTCTGGACATGATAATACAGTTTGTTCGGTCTTCGCTCGACCTACGGATCCACAA
GTTGTAACTGGCTCCCATGATACAACTATCAAGTTCTGGGACCTTAGACATGGAAAAACA
ATGACAACTCTCACTAACCATAAGAAATCTGTGCGAGCAATGGCCCAGCATCCTAAGGA
GAACTGTTTCGCATCTGCATCAGCTGATAACATCAAGAAATTCCAGCTTCCTCGAGGGGA
ATTTCTCCACAACATGCTCTCACAGCAGAAAACCATAATCAACACAATGGCAGTCAATGA
AGAAGGTGTGATGGCTACTGGAGGTGACAATGGGAGCTTATGGTTCTGGGACTGGAAG
AGTGGCCACAACTTCCAGCAAGCTCATACGATTGTACAGCCTGGATCCTTGGAGAGCGA
AGCTGGAATATATGCTCTCTCCTATGACTTAACTGGTTCACGGCTGGTGTCATGTGAAGC
GGACAAGACCATAAAAATGTGGAAAGAGGATGAACTTGCTACCCCAGAAACTCACCCTC
TCAATTTCAAGCCTCCCAAAGATATCAGGCGGTTCTAGTTCTTTGTAGGCATCCTATTGC
AATCTTATGTCACCGCTTGGCAACACCGTCCATCCTTTCACGCAGCCTGTTTGCTCCAAC
AGATCTTCTTGCATCTACTCTTCTTCGGAAGA

SEQ ID NO: 236
GCAGCAGAAGAGAGAGAGAGAGAGAAACGCGATCGTTCCTCAGTTTCGTCGGAGTGAA
GAAGGAGATCGGATCGGAGGCTCGCCGGCGAGGAGAGGGGAAGACATCGGACATGGA
AGAGGCCGCCAAGGAACAATCCGCCGGTTCCGGGAAGCCGAAGCTGCTCCGCTACGG
GCTGCGATCGGCCGCGAAGCCGAAGGAGGACAAGAAGGAGGAGCAGCTCCACCAGCC
GCCGCCGCCGCCGCCGCCGCAGCAGCAGGCGGCTCCGGCGCCGGCACCGGCGGCCA
CCAGGTCGTCGACGTCGGGGTCGGCGGGGGGCCGGGACCGGAGGCCGCAGCAGCAG
CACGCGGTCGACGAGAAGTACGCGCGGTGGAAGTCCCTCGTCCCAGTCCTCTACGACT
GGCTCGCCAACCACAACCTCCTCTGGCCTTCTCTCTCTTGCCGGTGGGGCCCGCAACTC
GAGCAAGCGACTTATAAGAATCGGCAGCGGCTCTACATTTCTGAGCAGACTGATGGCAG
TGTTCCAAATACTTTGGTGATAGCAAACTGTGAAGTTGTGAAACCTAGAGTTGCAGCTGC
AGAGCACGTGTCCCAGTTTAATGAGGAAGCTCGCTCTCCCTTCATAAGGAAGTACAAGA
CAATTATACATCCTGGAGAGGTTAACAGAATCAGGGAACTTCCTCAGAATCCCAATATTG
TGGCAACTCACACTGACAGCCCAGATGTTCTCATTTGGGATGTGGAATCTCAGCCTAAC
CGGCATGCTGTCTATGGAGCTACAGCTTCTCGTCCAAATCTGATTTTAACTGGACATCAA
GAGAATGCTGAATTTGCCCTTGCAATGTGTCCAGCTGAACCCTTTGTTCTCTCTGGAGGG
AAGGATAAGACGGTGGTTTTGTGGAGTATCCAAGACCATATAACAGCATCTGCAACAGAT
CAAACAACTAATAAATCTCCAGGATCTGGAGGATCCATCATTAAGAAGACTGGGGAAGGT
AATGAGGAAACTGGAAATGGCCCTTCTGTTGGACCACGAGGAATCTACTGTGGACATGA
GGATACTGTTGAAGATGTGGCTTTTTGTCCATCCACTGCACAGGAATTTTGTAGCGTTGG
TGATGATTCATGCCTTATATTGTGGGATGCACGAATTGGGACTAATCCCGTTGCTAAGGT
CGAGAAGGCACATAATGGTGACCTCCATTGTGTGGATTGGAATCCCCATGACAACAACC
TAATCTTAACCGGGTCGGCAGATAACTCTGTTAACATGTTTGATCGGCGAAATCTCACTT
CTAATGGAGTTGGTTCACCAGTCTACAAGTTTGAGGGGCATAAGGCAGCTGTTCTTTGTG
TGCAGTGGTCTCCAGACAAGCCTTCCGTCTTTGGGAGTTCTGCTGAAGACGGTCTCTTG
AACATTTGGGATTATGAGAGGGTTGATAAAAAGGTTGACAGGGCTCCAAATGCTCCTGC
GGGATTGTTTTTCCAGCATGCTGGTCACAGGGACAAAATTGTCGACTTCCACTGGAACA
CAGCTGATCCATGGACTATGGTTAGCGTATCTGATGACTGTGATACTGCTGGAGGAGGT
GGTACATTGCAGATATGGCGAATGAGCGATCTGATCTACAGGCCGGAAGAAGAGGTTTT
GGCTGAGCTGGAGAATTTCAAGGCGCATGTACTGGAATGCTCGAAGGCATGAGAGTGC
CTCGAGAACAGGCCTTCCGGGTCTCAAACACTAACTAGACAAAGCGGGTTTTTGCTGGC
TGTTACTGCTGTAAAATCTGTAGGTACTTAGCCATGGTTTAGACTCATCTGTGAGCGCCA
GGACTCCCCTCTTTACGCAGATGGTGACTGAAGCTGGTTCCGAGATCGGCATATGTAGC
TGGTAGAGGTGTGGATATTGCATAGACCGAACCTCCGCAGGTCCGCATTCTCGAGTGAG
AAACAGAGATAAATTTTAAGGGGGTTACCAAAAAAAAAA

SEQ ID NO: 237
TCTTTCACTGCATGCTCCCTTTCCCGCAAAACACAAACCAAACACGTTAAACTCTAACCC
AAACCTTCTCGCAATTCTTACCAAATCCAGTTACTCTGAGCCAAAACCCTTGGCGGGAAA
ACCCCAGTTAGGAGCTTCCGGCCATGGCGAAGGACGAAGAAGAATTCCGCGGCGAGAT
GGAGGAGCGCCTGGTGAACGAAGAGTACAAAATCTGGAAGAAAAATACGCCGTTTCTTT
ACGATCTGGTGATAACGCACGCCCTCGAATGGCCTTCACTCACTGTACAGTGGCTCCCG
GACCGCGAAGAGCCCCCTGGAAAAGATTATTCCGTTCAGAAAATGATACTGGGGACTCA
TACTTCCGACAACGAGCCCAACTATCTGATGCTGGCCCAAGTTCAACTCCCGCTAGAGG
ATGCAGAGAATGACGCCAGGCAGTATGACGACGAGCGCGGGGAGATTGGAGGGTTCG
GCTGCGCCAATGGCAAGGTACAAGTAATACAGCAAATAAATCACGATGGAGAGGTCAAT
AGAGCCCGATACATGCCACAAAATCCTTTCATTATTGCCACGAAAACAGTTAGTGCAGAA
GTCTATGTGTTTGACTACAGCAAGCATCCTTCAAAGCCTCCTCAAGATGGTGGATGTCAT
CCTGATCTCAGATTGAGGGGTCATAATACAGAAGGTTATGGTTTATCATGGAGCCCTTTT
AAGCATGGCCATCTTTTAAGTGGTTCAGACGATGCACAGATCTGTTTGTGGGACATTAAT
GTACCTGCCAAAAACAAAGTGCTTGAGGCCCAACAAATATTTAAGGTGCACGAGGGTGT
TGTAGAAGATGTTGCATGGCATTTAAGGCATGAGTACCTTTTTGGGTCTGTTGGAGATGA
CCGCCATTTGTTGATATGGGATTTGCGTACATCTGCAACTAATAAACCACTGCACTCAGT
AGTAGCTCATCAAGGTGAGGTTAACTGTCTTGCATTCAATCCTTTCAATGAGTGGGTACT
GGCTACAGGATCCGCAGACAGAACGGTGAAACTTTTTGATCTGCGCAAGATATCCAGTG
CTTTGCATACCTTTTCCTGTCACAAGGAAGAGGTTTTCCAAATAGGCTGGAGCCCCAAAA
ATGAAACAATATTGGCTTCTTGTAGTGCAGACAGAAGGCTTATGGTGTGGGACCTCAGC
AGGATTGACGAATTCCAAACACCAGAGGATGCTTTAGATGGACCACCTGAGTTGCTGTTT
ATTCATGGTGGACATACTAGTAAGATATCAGATTTCTCATGGAATCCATGTGAGGATTGG

GTTATAGCTAGTGTAGCTGAAGATAACATTCTCCAAATCTGGCAAATGGCTGAAAATATAT
ACCATGACGAGGAGGACGATATGCCTCCTGAAGAAGTAGTGTAACTTTTATCTAGCTAGA
AGTTGTGAAATTAAGAGGGATGTGAGGATTGGGTTATAACTAGTGTAGCTGAAGATAACA
TTCTCCAAATCTGGCAAATGGCTGAAAATGTATATCATGGTGAGGGGGATATGGCTTCCA
AAGAAACATTGTAAGAGCTAGCTAGAATTTGTGAAATTAAGAGGTGTATTCACTTTCAGA
GTTTCTCAACAAATGACATGGTTCTCATTCCATTTTCTTTTATATAATGAGAAGCAAAACTT
GGCTTATCATCTTTTTTTATTGAGAAAAGACACATTGAAAGAGCTTAAACCTTTAGCTTTTG
CTGTTTTCTGCTGTATTTATGTAGTTTGCAATCTGGTGAAGATTATGCCAGGTCTGCCTTT
CCTTGTAATAGTTTTTATAGACAAATTTATAGCTGAAAGTGATGCCCACGTTCTAGGTAAG
ATGGGAAACATCAGAATGTAAGGGTCTCAAATGTTGTCACTATTCAATTTTCATCTATTTC
TGGTCTGGTGGATTTTGCTGGTAAGACATTTGAAGATGTCCAACACTTGTTCTACAGTGT
ATTGTGAAATATAAGCACATAACTTGATACAAAAAAAAAAAAAAAAAA

SEQ ID NO: 238
GCCTAACTCCCCACCACACAAAAACACAAACAGGCAGAGTCCCCCGCAACTCTCCTGTC
CTGTTTCCATTAATCTCTTACGCCGTTGCGTTTTCCCTTCTGATCAGCCCTCCCATTCTGC
TCCCCTCCGTTCCTGCCCTTGGGTTGAGAAAGGCCGTTGTGGCCATTTGTCAGTGTCTG
TTATTTTTGTTTGAGCCGGCACAGCAATCTGCCAGTCTCTATCTGTAACTCACTCGGTTTT
TCTTGAAACAGCAGGAATCACCGTCACCAAGAAGACGCCATTGCAGTGTGTACTCGACT
TCTGCAAGATAGCTGGTCAGTCAAATCAAATTTACACCGACGAGCACTCGGAAATAGTTT
CTTTTCCTCCAAATTTTAATGCAAATTTTTCGACTGCGAAAAAAAAAAAATTTGGCTTGAATT
CTGGTTTTTCAGCAAGGAGGTTTTTTTCCTCTGGTTTAACAGCAGCAGCAGGATAGGGAA
ATATTTTGTGAGTACAGTTTCTGAAAAAACATTCGGACAAATTTTTGTCTGCGAGGGTTCG
AGGAATCTCGAAATTTTCACAAGCTGTTTTGGGTTTCTGGGCAAATTTTAGGGCAAATGT
TTTAAGAATGAGAAATCCAAACCGTGTGCTCGTGCATAGCAAGGAAAAAAAAATTATTCT
GTAGAGTATCAGGAAAAATCAGCAGCTCTCGCGAATTCGTCTTTTTCTCTTCCTTTTTTTT
TTGAACTTGGCATGCAGTGTGGGCTTAGAACAGTGCGTTTCAACCGTGGCTCCTTACCG
TGGTTCAAAAATGGGGAAATACATGAGAAAAGGGAAAGGCGTAGGCGAGGTAGCTGTGA
TGGAGGTCTCGCAGGGCTCGCTCGGTGTGAGAACGAGGGCTCGGACGCTTGCAGCAG
CTTCTTCACAGAAAGACCACCGGCGCCTCGGTGCCTCCAAATCTGTGACTACGAAACAC
CAGAGCTCGGCGCCTCCTGCTTCACCATGTGTGGAGTCCTCCATGCACACATGCTACCT
TGAGCTACGAAGTAGAAAGCTGGAGAAATTTTCCAGGTGTTACCACAGCGCCCATGGAG
CAACAAGCCACGGGGAGTCCAAGAGGAGCTTGAGCTTGAGCGAACCTTCGAGGCTGGC
AGTTTCAGAGGAGGCGAGAGTGGCCAGCGACAAAAGTAGTCATAGAGTCCTCCAACAAC
AGTCTTCCGTGGCTCATTCCAGGAACAATTCTGCAACGTTTTCTCACAATGCCAAGCCGG
CCAAAGCAGCCCAAAGGAAGGAGCGCCGTGACGACGACCATACGTCGGCGCGGCCAT
CCGAAGCTCCCCATGAGGACGAGGATGGCATGGAGGTAGAGGCATCATTTGGGGAGAA
CGTAATGGATCTCGACTCGAGGGAAAGAAGAACTAGAGAAACGACGCCATCGAGCTACA
CAAGGGACGTGGAGACGATGGAAACCCCGGGGTCCACGACTAGGCCTCCCTCTAACGC
CGGCAGAAGAAGATTCCAGACCGAAGGGGGGCATGGCACTCGCAACCAATTTCACGTT
CCCACCACCAACGAAATTGAAGAATTCTTCGCCGGTGCAGAGCAGCAAGAACAAAGACG
ATTTACAGATAGGTACAATTATGACCCAGTAAGTGACTCGCCTTTGCCTGGTAGGTTTGA
GTGGGTTAGATTGAGGCCATAAATATCTTTCACAAATAATTGTGGGGCTGCAGGCAGTTG
GTGTGGTTGGGAGTGCTTTACCAGCCCTTGCCATTCTTCTGTACTGCTCTGTTACCAATT
TAACTTGGGGTTTCAACGGAACTCAAAGAAAAAGGAAAACTGTAACTGCCACAAGGCCAT
GTTTGATGCCAAACTGCATCTCTCATTATTTAGGGTGCAGGCTGTATAAAATGTTGTAAAT
TGTAGTATCAATGTGTACAATATGAAGCGGCTGGCATGGCTTCTTCCCTACAATGTATTTT
GCAGGGGCATTCTTCTTTTTTATTCCCAATGAAAAAAATTCAAACAAAAAAAAAAAAAAAA
AAAAA

SEQ ID NO: 239
ACCGAATGAACCCATCCTGCGCGCGGATATCTGCATTTGATGCAAAGGAGAAGAGGGC
GAGGGTGTTAAGGAGGGCGGAGTTGCTGTAGAGGAGGGTCGGTTCGGTCATAACTCCT
GAATTTCTTGTTAGTGTTTGATGTATAAGGCTGCTCAGATGTTAAATATATTTGATTCGGT
TTGTTTTGAAGGCAAATGAAACTCTTCGGGCTTATGAAGGTTGGTTTGAATTGAGAAATTT
GGGATATATAAACAGTGGCCATGTTCTCCGGACCATGGAATTGGGGATTTCTCCATAGG
GTTTTAGTGTTTCAGGGTTTTGAGACTTGAATACCTTGAAAAGATTTGATGGGAAGGTTTT
GAGAGATAAAATCTCGAAAAGATTTTATGGGAACAGGTCTTCCGGAGATTTGAGGAATGC
AAAACATGGAGGAAAATGTGCAGAGCAGCTGGAGTCTCCACGGCAACAAAGAAATATGT
GCCCGTTACGAGATTCTAAAGAGGGTCAGCAGCGGAACTTATTTGGATGTGTACAGAGG

GCGGAGAAAGGAGGATGGTCTAATAGTAGCCCTTAAGGAGGTGCATGACTATCAAAGCT
CTTGGAGAGAGATTGAGGCGCTGCAAAGGCTTTGTGGGTGCCCCAATGTTGTGAGGCT
CTATGAGTGATTTTGGAGTTTCTGACTTCCGATCTCTATTCTGTCATTAAGTCTGCTAAGA
ATAAGGGAGAAAATGGTATTCCTGAGGCCGAGGTTAAGGCCTGGATGATTCAGATCTTG
CAAGGGTTGGCTAACTGCCATGCCAATTGGGTTATCCATCGTGACCTAAAGCCCTCTAAT
ATGCTCATTTCGGCTTATGGAATTCTCAAGCTCGCTGATTTTGGATAGGTAAGTATAGAT
GATTTAGTGGCTTGCCATCTGTTGCATCTTCTGTGTTTTTTTTCTTGTCACCTTCAATCCT
CTAGCTCTGTTTGACCCTATTTCTTGAAGTGTTCGTACATTCTGTTCTATTTCCCATTTCAT
GTTTTTAATATTTTAAAATATGCTGAAACCCTGTTTATGGTTAACTACATCTTCGTAGTAGT
TCAAAATGTTAACAAATCTCATACTTGAAAAGTTCAATCTAAAATCTTGAAAATTTTGACTG
ATTCAAAGCGCTTTGGAAACTAGATAAATGAATTGATTTTGCCTTTTACTATAACTATTCG
AGCAAATGTTGCATAGAAGCGCAAGGTTAATAGTTTCCCAAATTGAAACTGAATAGTGCT
TAAAGTTTAGTGATTCTCCGTCAGTAGGATGCCTGAACATTCAATAACTTAGTGATCTCTG
AATTCGAAAATCAATATTTGCAGGCAAGGATACTTGAAGAGCCTGAAGCGATCTATGAAG
TAGAGTATGAGCTTCCTCAAGAGGGCGATCTATGAAGTAGAGTATGAGCTTCCTCAAGA
GGATATCCTTGCTGATGCCCCAGGAGAAAGGTTGATGGATGAAGATGATAGTGTTAAGG
GAGTATGGAATGAAGGGGAGGAGGATTCATCCACTGCAGTTGAAACTAATTTTGATGATA
TGGCAGAAACTGCGAATTTGGATTTGAGCTGGAAAAATGAAGGTGATATGGTGATGCAG
GGATTCACATCTGGTGTTGGAACTCGATGGTACAGAGCTCCGGATTTTCTCTATGGAGC
AACGATCTATGGAAAAGAAATTGATTTGTGGTCGTTAGGTTGCATTCTGGGGGAGCTCTT
GATTTTAGAACCTCTCTTTTCTGGGACTTCGAACATTGATCAACTTAGCAGGCTGGTTAAA
GTTCTTAGGACTCCAACAGAAAAAAAATTGGCCTGGATGCTCCAATCTTCCTGATTATAG
GAAACTTTGTTTTCCTGGTGATGGAAGCCCTGTTGGTCTGAAGAACCATGTCCCCAATTG
CTCAGACAACATGTTTTCTATTTTGGAAAGACTTGTTTGCTATGACCCTGCAGCTAGGCT
AAATGCTAAAGAGATAGTTGAGAATAAGTATTTTGTTGAGGATCCTTATCCTGTCCTCACC
CATGAATTGAGAGTTCCCTCACCTCTAAGGGAAGAAAACAATTTTTCAGAGGATTGGGCG
AAATGGAAGGATATGGAAGTAGATTCCGACTTGGAAAACATTGATGAGTTCAATGTCGTT
CACTCAAGTGATGGTTTCTGCATTAAATTTTCATAAACTCTAATTGGTTGATCATGTTTCAT
CTTCCTCAAATTGCCTTTTATATGGATTAGTATTCTAAATTAGCTTTGGGGGCTGCATGTC
TTCAAGCATTCACCACGACAGTAAAGTAATCATTATGATTACTAATGTATTGCTTTCATGG
GGTGAGAAACATTATAGCCATGAACCTACTGTTAAATTTGTAGAATTGTCTTATACTTGCT
CAAAATAGTGACTGCATGTTTGAGAAAAAAAAAA

SEQ ID NO: 240
ACTCCAACAGAGGAAACAAAAACACATTTCCCGCCTTCGCATTTCCATTTATATCTCAATA
CCAGTTCTCCGACATTTTATCGGCCTTCAAATCTCTGATAAGAAATAATGGCTCCCGTCA
AGAGGATCGAACCAGAGAAAACGAAGGCTAACGAAGGCAAGCCGAAGCGCCGGAAAGT
AGCCTTCGCCATAGACACAGGAATTGAAGCAAATGACTGCATATCACTTCATCTTGTGAG
CACCCCAGAGGAAATGCGAGATGCAGAGGGAGTGGAGGATCAAAGCTTATCATTCAACC
CAGAATATATGCAACATTTTGTTGGGGAACATGGGAAGATCTATGGCTACAAAGGCCTGA
AAATTGATGTGTGGCTGAATGCCCTCTCATTTCATGCCTATGTGGATATCCAGTATGAGA
GCAAAGTTGAGGAAGGAAAGAGCGAGAAAGAAGCAACAGACTTGACAGATATTATGAAG
AGAATATTTGGACGTGGCTTGGTGGAAGATCGTAATGCCTTTATCCAGAGTTTTTCAAGT
AACAGTCAATCTATCGAAAGTATGATCCACAATGAAGGTGAGCGTATAGCAACACGTGAG
ATTTTGACTGATAAGGGTCTATCTGCGCAAGGAGATTCAGAGAGATTGGGAGTCTCAAAC
GAGATATTCCGCCTTGAATTGAGCGATCCACAGATACGAGAATGGCATGCACGCTTGGA
ACCGCTTGTGCTTCTTTTCGTTGAAGGAAGCCAACCTATAGAGCAAGATGACCCAAAATG
GGAAATGTATATCAGGGTCCAGAGGGAGTCATTGAGTGGAGGCAGTGCTGTGTGTAGAT
TATTGGGATTCTGTACTGTTTACCGTTTCTATCATTATCCAGACACCACCCGTCTAAGAAT
TAGCCAGATACTAGTTTTCCCTCCTTATCAAGGTAAAGGGCACGGCTTACTTCTTTTAGA
GGCTGTAAATAAAACTGCAGTTTCGAGGGATTCTTATGATGTAACAGTAGAAGAACCATC
CGAGTCCCTCCAGGAATTACGCGATTGCATGGATACAATACGCCTCCTTTCTTTTGAACC
TGTAATGCCTGCAGTTAAGTCAGCGGTGCAGAAGTTGAAGGAAGCAAACCCCTCAGACA
AAGGTGCAGCTGATCACTGTTTAGAAGGAAATGTAAACAATGAGACTGTGACTACCTCCA
GCACAAAACCCAAAAATAAAAGTGGGTGGTTTCCACCTCCAGGTCTAGTTGAAGAAGTG
AGGAAACATTTGAAGATTAGTAAGAAGCAGTTCAAACGATGCTGGGAGATTCTTCTATAT
TTGAATTTGGACCGCTCTGATTCACAATGTGAGGATAAATATCACATCTCTCTCATGGAA
CAGATAATGTCAGAACTCTTTGATAAATCGAGCGAGAAGAGTGCCAAGGGTAAACGTGT
GATTGATATAGATAACGAGTATGACAATTCAAAGACGTTTATCATGGTCCGCACAAGGAA
TCCAGGCAATGGTGAAGGTTTTTTACCCGAGGCACTGGAGGGTGGGATGGAGGTATCC

CAAGAGGATCAGTTGAAGTCATTATTTGAAGAGCGACTGGAAGAGATTGCACAAATTGCA
GAGAAAGTTCCTAGCCTCTGTAAAGCATTGCAAATGCCATAGTTCTCATTTCTATGGAGC
GTGACCTGCCTGTAGTACGAGGATTTGACCATTTGACTCTGACAAATTTTAGTAGGCTGC
AATGTTCTTTGCAGCAAAAGGTGCACGCCCTCATTCAAGTAAAAGGGTATATTTTGTCTC
ATGTTGGGGTGATAATTCTCCCTGAAAGTCTCCAAAATATATGTGCTGGTTGTAGAATAA
AAATTAGAACAAACGGCCTACTGTCATATTTCATGCTCCATGCTGCTAGCAAGATTGTTTT
TCAAGGTTCTTTTTCGGATATATAACCAGCTGAAGGGTGATAGTTTTCTTGGTACGTTTCA
CCTATTGGTCGAATTTGCTAAAATGGTCACTTCTGGTCATTGCGGATCGGTATAAAGCCG
CTATGATAGATCATATTAATCTGGGGTATATGAAGTGTGAATTATGAACTTAGCTGTACAT
TTCTCTGTTTATTTGGGATGAAGATAGATGTATGCTTAGGAGAATTTTAGACATTAAATTT
GGAAGTGTTTTTTAAAATTAATTTTGATTAGCAT

SEQ ID NO: 241
CGCACTTCTGTGCCGGCACTCCTGTCCCACTTTTGTTCGGAATTATGCCGAGGCTTCATA
CATGCGGGTGTTGCTCAGTCGTTTTTCTTTCATAAATAAAATATATGTAGCTTGCAAGGGT
TCTTTTTTATTCTAAATTTCTTGTTCAAGAAGATTTACCTCTAGCAGTTTCTTAAAATTTCC
GGTTGCCATAGTCTAGTGGGGTGAGGGTTCATTCTAGGGGATTTATTGTGTTTCACTTCC
GCATCGAAGCTTCATAGGCTGACGCATTCTGTTCATTTGATAATATCTTTTTGACCTAGCT
GAAAGCTTCTGTGCCCAGGATCTTACGTTACTGCAACCATGCCTGAGGACCGAAAAAAA
ATCCTGGAAGCATTGGCGGCCAAGAGAAAGGCCGAAGCCGAGAGCGGGGAAAAAAAGA
AAAGGCAGAAAAGTTCTTTGAATCCTGCCAAACCTGTTAGCAAACCTGTTAGTAAACCTG
TTGGTGGAATTGGTAGCAAGGGAAAGTCGACATCTGCGCCAATTTCTTCGACTAAGGCT
AAGTCTAAGCACAAGGAAGAGGTAAAGGCAAAGCGTGTAACCAAGATGGATAGGTATGA
GACTGATGAAGACGATGAGTCTGAGGAGGAAGAGGATTTAGATTCTGAATCAGATGATG
ATGAACTGTCTGATGAGGATTCTGAAGATGATATAAAATCCAAGTCTGTAAAGAAATTGC
CTCCTCAGTCCAAGGGAAAAGCTCCTGTTAAAGGAATTAGTAGTTCAAATGGGAAGGGC
AGGGATGAGAAGGGAAAGGGAATTATGAAAGACAAGGGAAAGGCGAAAGCAAAAGTGG
AGGAAAGTAGTAGTGATGCTGAAGGTGACAGTGATGATGATGGAGGTGACTTGTCTGAT
GATCCTTTGCAGGAGGTCGATCCTTCAAATATTCTTCCCTCAAAAACACGTAGACGAGCT
TCACAGCCTACGAATTATCAGTTTGCAAATATGAGTGGGGATGACGACGACGACGACGA
CAGCGACTAGATGATGAACCTGATAAAACTGTTTTGGGGCTCTCGATGGTAAGTTTTCCA
CAGTGATGAGCAATATAATATATGCTCTGAATTGTGAAATGTGGAATGCAAGATCCTCTTT
CTGGTATTAGGTGAACTAGGTAGTAATATTTGCTTTTGCAGTTAGACAGAACAGAATTCC
CTTTGGCCTTCAGTGGTCGGTATTGGTCCTATCGATATGGTATATTTTATTGCTTCACTTT
ATTTATGCTGCGCAAACTGTTGACACTCGGGATCTTTATTTTCGTTTTGACACGTCCTGTG
TTTTTACAAGTAGTTTGTTATTTGATTACATAATGCATGTTACTCTTTGGAAGGAACAGAG
CTACAGGTTCTTTTCTATAAGATATAGTTGTTGTGTAAAAAAAAAA

SEQ ID NO: 242
CTTAAACTCCATTTCTCACTCACGAGTTTACCCAAGCCTCTCTGTGCAAGCACGAATTCG
CATTTATTTCTCCATTACGTGACACAGAGGGTTAGGTGATTAATTAACTATCGGAGGATTA
ATTGAAATTCTAATTATCAGTTCGCTCCCTCCGTTTTTCTCTTCCCCCTTTCAGTCTCGTA
TGGATTAGCCACTTCGCTTGTAATAACAATGGCAGAATCATCATCATAATCATGCGGGCG
GCGAGGTGGGTTTAGGGTATAAGGTTAAACAGAGCATCCAACGACCTACCTAGACTGGT
TCACGCGCTAAGATTTTTATTGTAGTATTTTTGTTGCACATAGTGGGCATTGGGGTATTTA
TCCAAAGCCATATTGAAAAATTAAGACTTGTATCGTATGAAAAAATAAATATTCATGGCAG
ATGTACCAGAGTCTCTGCAACAAGAAAAGGATGAACAGGGAACTGATAAAAATTGCTGTG
ATGGCAAATTTCAGAAAGAGATAGACATTGATGACATGGAAGAGGAATATAATGAAAGCA
GCATAGATGATGAGGAAGAAAATTTATCTGATAATGTAGCTACCAATAATATGGGTACTAT
TCCTCAAGGGCAAGCATGCATGGCTGTTACAGTTGAAGGCATTGAGCATGCTAATTCAG
TTGGTTGTGGAAGGAATGGCAGGGAAGGCAGTGAAGAGGTCACTGCTGCTGAAGATAT
GGGGCATGTATCCATAGAAAATATAAGAGAACAAGGTCGTAATAGGAAAAGTTCGGAGC
AATTGCTTGCTTTGTATGAACAAGAAGGGCTGTTGGAAGATGATGAAGATGATGATGATG
TGGACTGGGAGCCTTTTGAAGGTGTCACTGTGCAGATGAAATGGTATTGCACCAATTGC
ACAATGGCCAATTCTGATGATTCAGTCCACTGTGATTCATGTGGAGAACATCGCAATTCA
GACATTCTCAGGCAGGGTTTCTTAGCATCACCCTATCTACCTGCAGAAAGCCCAAGCAG
TTCAGATGTTCCAGATGAGAGGCTAGAAGAAAGCAAATGTGTTATGACCACTCTCACACC
TTCCATATCTCCCATGATTGGGGTGTGTTGTTCATCATTACAGTCAGAGAGACGAACAGT
GGTTGGGTTTGATGAGCGTATGCTTCTGCATTCAGAGATTCAGATGGAGACATATCCTCA
TCCAGAGAGGCCTGATCGCCTTCGTGCAATTGCTGCAAGCCTTAGAGCTGCTGGTTTAT

TTCCAGGAAAATGTTTCTCCATACCTGCAAGAGAGGCTACTTGTGAAGAGCTGCAAACGA
TCCATTCTTTGGAGCACGTGAACGCTGTTGAATCAACAAGCTGTGGAATGTTAAGCCATT
TGTCACCAGATACGTATGCAAATGAGCATTCGTCTCTTGCGGCCAGACTTGCAGCTGGT
TTATGTGCTGACCTTGCAAAAGCAATCATGACAGGCCAAGCTCAGAATGGGTTTGCTTTG
GTAAGGCCTCCTGGTCATCATGCAGGAGTAAAAGATTCCATGGGTTTTTGTCTTCACAAC
AATGCAGCAATTGCTGTCTCAGCATCCCGGGTTGTAGGAGCAAAAAAAGTTTTAATTGTT
GATTGGGATGTACACCATGGGAATGGTACACAGGAGATCTTTGAGGCAGATCAATCGGT
GCTTTATATATCCTTACATAGACACGGAGAGGGTTTCTATCCTGGTAGTGGTGCTGTTAC
CGAGGTTGGTAGCTCCAAGGGAGAGGGATACTCGGTGAACATTCCGTGGAAATGTGGA
GGAGTTGGGGATAATGACTATATTTTTGCATTTCAACATGCCGTTCTTCCAATAGCGGAG
CAGTTTGAACCTGATTTAACAATAATATCGGCAGGCTTTGATGCTGCAAAGGGTGATCCT
CTTGGTCGCTGTGAAGTCACTCCTGATGGTTTTGCACATATGGCTCAGATGTTGAGTTGT
CTTTCAAAGGGAAAGATGCTTGTTATTTTAGAAGGAGGTTATAATCTACGTTCAATATCTG
CCTCAGCTACAGCAGTGATAAAGGTACTTCTTGGTGATAATCCTAAAGCCCTACCCATTG
ATATCCAGCCTTCCAAAGGTGGATTGCAGACACTTCTGGAAGTGTTTGAAATTCAATCAA
AGTATTGGTCTAGCTTAAAGGGGCATGACCAGAAGTTGCGCTCTCAATGGGAAGCACAG
TATGGCAGTAAAAAAAGAAAGGTGATAAGGAAAAGGCACATGCACATTGTTGGAGGTCC
AGTCTGGTGGAAATGGGGGCGTAAAAGAGTGGTGTATTATCATTGGTTTGCAAGGGTAT
CATCTAGAAAGCATTTGTGAAGTGATAGTTTCCATTTCTTCCATTTATTGTTGACAGCGCT
AAAAGATGGAATGGACAGGTTTGAAGGATGGCCTCAGCTTGCTTTTATCTGCAGATAGG
ATGTTACTTCATTTTTGTACAGATTAATTTGCGGCAAATTGGCATCATATATGACTGTCTG
GGAAATCCTGCACTAGTGGTGTAACTTGTGGCGGGAAGGAACTGCACAGACATGAGTTG
TGATTTTGATGTAATCCTAATTTTATCTATTGTCACCTTGGTTGAATTGTCAATATTTTCAT
GTCCCTTGAGGCCAATCTTTGAAATGGAGAATGTTTAATTGTAGAATACATGTGACAATAT
TAATTCTTGAAGGTTGCACATTCTTAGTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAA

SEQ ID NO: 243
TAGGAGCGAGCGAGAGTATTACCTCATTCAAGACAGGATTCAAGAGCGGAGTGTAGCAT
CGACATAAAATGGCAAGTGGAGCAGGGGCAGCAGGGGTGGTCGAATGGCACCAGAAGC
CTCCAAATCCAAAGAATCCTGTCGTCTTCTTTGACGTAACGATAGGCACCATTCCAGCCG
GACGTATTAAGATGGAGCTCTTTGCCGATATTGTTCCCAGGACAGCTGAAAACTTTCGGC
AGTTCTGCACAGGGGAATACAGGAAAGCAGGCATTCCTATTGGCTATAAAGGTTGCCAT
TTTCACAGGGTTATTAAGGACTTCATGATTCAAGCTGGTGATTTTGTGAAGGGAGATGGC
AGTGGATGCATCTCTATCTATGGTAGCAAGTTTGAAGATGAAAACTTCATTGCTAAACATA
CAGGTCCTGGCCTCCTCTCTATGGCAAATAGTGGGCCTAATACTAATGGATGCCAGTTCT
TTTTAACTTGTGCAAAATGTGATTGGCTGGACAATAAGCATGTGGTTTTTGGGAGGGTAC
TAGGAGAAGGACTCTTGGTTTTGAGGAAAATTGAAAATGTCCAAACTGGGCAACATAATC
GACCCAAGCTACCTTGTGTTATTGCTGAGTGTGGTGAAATGTAACAACCAAATGCCTGTC
AATATTGTGTCATGGGATGATTTATCTGCTCATGGATGTTAGTTCGACTAGCCTCAGTTTT
ATCTTTTGTTTCGTTCCCACAGGGAATTTTTAATGAAATGAGGAGACCATTGCTTAGATAT
TTATGGGTCTATCATTAGCAGATAAAATTCTTTATGGTCACGAACAACAGTAATGCAAGAG
CTTGAATATGTTCCAAGCTCTCCAGTTTGTTTTAATGGCATTGCGGACTTATATAAATCAG
TATACCAGTCCTTTGTTGCCTGGTTTACCAAATTTTGTT

SEQ ID NO: 244
AATATACGTGTGCCATTAAGCCTTACGGCCTCTAAAAATCGATTTAATGAACGAAGATAT
GTTGTAAGCGAGGGCCGAAGTGTTCCAACCTTATCCGTGAAGATGATGTGGGAGACCTG
CTAAGAGAAAGCCCGAGCAGACAAGGTGAGTGGACTGCTGTAAGGACCGACAATGGCA
AAATTGGTATCATCAGTATGTGCATTTTCATGTCAGCAGCGGCACCCTCACTCTCGACCT
CGTTTCCTATCCAACCGGGACCATTATAACCATTACCATAATCATTCTCATTATCATAATG
TTTGCTATTTTCCACCAATGATGATGATGCAGCAACAGTTGCAGAAGCAGAAGAGGATGA
CGACGAAGACCATTACATCCTTGTTCAAATGCAACTCTTCAAACCACACATTGCTCAAGG
GACTGAAGGAATTCATGGGCTTCAAATTCAGATTACAAGCGGCAATGCTTTCTTGCGAAA
TGAGCATCCTTGGAAGAGTGTTTGCTATATTTTTTATTGTTCATCAAGCAGCTGCCCCATT
TCCTTTTAATCATTTTGACAATTGGTTAGTCCCTCCTGCAAGTGCAGTGCTTTATTCACCT
AACACAAAGGTTCCTCGAACTGGAGAAGTTGCACTAAGAAAGTCTATTCCTGCCAATCCA
GCCATGAAAAGTATTCAGGACTTCTTGGAGGACATTTACTACTTGTTAAGATTTCCTCAAA
GAAAGCCATATGGTACAATGGAGGGAGATGTAAAAAGTGCTCTTCAGATTGCAATAAATG
AGAAGGATTCAATATTAGGTTCAGTTCCATTGGATATGAAAGAAAGGGGATTACAATTGT

ACAACTTCCTTATTGATGGTCAGGGAGGGCTGCAAGTCCTTATTGAATATATAAAAGAAA
AAGATCCAGACAAGGTCTCTGTGAACCTTTCCTCTTCATTGGATACGATTGCTCAATTGG
AACTCTTACAGGCTCCAGGTTTACCATATCTTCTACCTGAAGAATATCAGCAGTATCCAA
GGTTGAATGGAAGAGCCACAATAGAGTTTACAATGGAGAAAGGAGATAATTCAATGTTTT
CTGTGTCAAGTGGAGGTGGGCTCCAGAAAACAGCTACCATCCAGGTTGTTCTTGATGGA
TATTCGGCTCCTCTTACGGCAGGGAACTTTACAAAACTTGTGATAGATGGAGCATATAAT
GGACTTAAGCTGAAAACTACTGAGCAAGCAGTCATTTCTGACAATGAGCGTGCTGAAGC
TGGGTTTAACCTACCTATAGAGATATTGCCGGCTGGTGGGTTTGAACCATTGTATAGGAC
AACTTTGAGTGTTCAGGATGGCGAGCTACCTGTGCTCCCCTTATCTGTATATGGAGCAAT
TGCAATGGCACACAACACAATATCAGAGGACTACTCATCACCATCACAATTTTTCTTTTAT
CTATATGACAAGCGCAATGCTGGTCTAGGTGGACTTTCATTTGATGAGGGGCAATTTTCT
GTTTTTGGATACACAACAGTGGGCAAAGAGATACTTCCACAACTCAAAACAGGAGACATT
ATCAAATCGGCCAAGTTAGTGGATGGCTTTGATCACCTTGTGTTGCCATCTTCAAGTACC
TAGAATCTTTCATTCATTTCATTCAAATTCAATGCTAGAGTTGTAAACAGTGTAATAAGGA
GCAGAAGTTGTGATAGCTTTTAGGAACGATAGACTTTGTAAATTGACAAATCCATGTAATA
TTTATCCAATAGAAAAAGGAAATACATAAAAAAAAAA

SEQ ID NO: 245
GTTTTAGTTTTTGACGAGTGTCTTTCTAGGTGGGGTGTTAGGGGAATTGCTCACCGGATT
TTGATTTTCTGGTCTTTTTATGTGCGTGGCTCGTGTAGTTTGATACTGAACATTCAGGAAT
TGAAGCTTCTTCACAAGGATTCGGATTTTGATTTGTTTCAGAAGGTGGGGTACAAGATTC
AGTAGCGAAGTTCAGTGTTTTGCCGAAGGGCTCTTTCAAATTAAGTAAGATGGATCACTA
CTACCAAGATGACTTTGATTATTTGGTGGATGACGAGATGGTTGATTTTGCTGATGACGT
CGAAGATGATGTACGCACCCGGCGCAGGAGTGATATAGACTCAGATTCTGAAAATGATT
TCGATTCAAACAATAAGTCACCTGACACAACAGCTCTTCAAGCCAAGAGAGGAAAAGATA
TTCAGGGTATTCCATGGAATCGATTGAACTTTACTAGGGAAAAGTATAGGGAAACTCGTT
TACAGCAGTATAAGAACTACGAGAACCTTCCACGGCCGCGACGCAGTCGCAATCTGGAC
AAGGAATGTACTAATTTTGAAAGGGGATCCAGTTTTTATGATTTTCGCCACAATACGAGAT
CAGTCAAGGCCACTATTGTCCACTTTCAGTTGAGGAATTTAGTGTGGGCAACATCAAAGC
ATAATGTTTACCTCATGCAGAATTACTCTATCATGCACTGGTCTTCACTAAAACAGAAAGG
GGAAGAAGTTCTTAATGTTGCTGGTCCGATCATTCCATCTGTGAAACACCCCGGTTCATC
ACCACAAGGTCTGACAAGGGTCCAGGTCAGTGCAATGTCTGTGAAGGACAATTTAGTTG
TTGCTGGTGGTTTTCAGGGAGAACTTATTTGCAAGTATTTGGACAAACCGGGAGTGAGCT
TCTGTACAAAAATATCTCATGATGAGAATGGCATCACTAATGCAGTAGAGATATACAATGA
TGCAAGTGGTGCAACACGACTAATGACTGCAAATAACGACTTGGCAGTACGAGTATTTGA
CACTGAAAAATTTACAGTGCTTGAACGCTTCTCTTTTCCTTGGTCTGTAAATGTGAGTGTT
AATTCTTCGTCGAGATCACTGTGTCTTATCGGAGGGAAGTTTCTGCTGGTTTTAATTTGA
CATATCAGATGACATGAACTTATTGCAAGAAATATTACTGTCTTGTTCATCCTGTATATAG
ACGTGTATGTTAAATGTCAAATGTACATTATCCGGTTTCTGAGAGGACATTGATAACTTTG
CAGCATACGTCTGTCAGTCCAGACGGTAAACTTGTTGCAGTTCTTGGAGATAATGCAGAT
TGTTTGCTTGCAGACTGCAAGACTGGAAAGACCGTGGGAACTCTAAGAGGTCATTTGGA
TTACTCTTTTGCAGCTGCCTGGCATCCAGATGGTTATATTTTGGCTACGGGGAATCAGGA
CACCACCTGCAGACTTTGGGACGTTAGGAAGCTTTCATCTTCCCTAGCTGTCTTAAAGGG
TCGAATGGGTGCTATAAGATCGATACGGTTTTCATCAGATGGCCGCTTTATGGCCATGG
CAGAGCCAGCCGATTTTGTGCATCTATATGATACTAGGCAGAATTATACTAAGAGCCAGG
AAATTGATCTTTTTGGGGAGATTGCTGGAATTTCATTCAGCCCCGACACTGAAGCATTTTT
TGTTGGGGTGGCTGATCGAACGTATGGAAGTCTTCTTGAGTTCAACCGTCGACGAATGA
ATTACTATCTTGATTCCATCCTCTGAGTTTTGAAAGTTAATGGGAGTGGTGTTTTCTTGAA
GTGAAATGGCATCATTCTGTTGAACCAATTCTTGTATATTAGATATGTAACATGTATGAAT
GTCCATAGAGCAGAGCTTTGCTCAATCAGAGGCTTCAAAACCCAAATTCCAAGGTCTATC
GAAGCCTCTTTGTTTATAAATGGTGTTGTGGTGGATACAGCTTTCTGGTTCACGCTGTAG
ATTTTTTAATGTATGTAACAATTTCAGCGGAAAAAAAAAA

SEQ ID NO: 246
ATCTTCGGGTTTTTCCCCTGAAGAGGTCTCGGTCAAACTCGGGCTTTGAAAGAAGCGTTT
TATGGGAACTTTGTGCCTCGTTTCGTTGACGAATCTGTGCCCCTTTTCTTTGACGAATTT
GGTGAAGAGCTCTGAAGAAGAAGATGTCTTTGCTCTGCTGAGAGGTTTTCGGAACCAGG
GTTCTGCCAAAATCGGTCCGCGTAACTGTTTTGTTGGGTTGTTACCGTGTAGAGTCTGAT
TTGGTATCGTCTATGGATTCTTTCGGAATTGAAAACAAGGAAAGTTAGGGGTTTCTAAGC
CTATTTTGTCCTTGATGCCTTTAGGAAAAAGTATCGATGATTTAGCTGAATTTCTGAGTAA

TCTCAGAAAATAGTAATGGATTGCTCGGGCGACGAGGAGGAGGAGCAGTTTTTTGAGAG
TCTGGAGGAAATGTTATCACCTTCAGATTCAGGATCAGAGGCTGCAGATAATGAGACGG
GGTGCAGAAATGCCGATGCGAGGTCTAAGTACGAGATTTGGAAAAGAGCGCCCAGTAGT
ATTCAAGAAAGGCGGCAGCGGTTTCTTGTGCGGATGGGTTTGGCTAATCCCAGCGAGCT
TGGGAATCAGGTGAATTCCACATCTGCTGAATCAACTTGTTCTACGGAAACTGCAAACAT
TCCAAATGGAATTGAAAGGCTCAGAGAGAATAGTGGGGCCGTATTAAGAACTGCAGGAT
CGAGTGGCAGGAAGACGCACTGTAAAAATGTTATTAATATCGGTTTGAGAGAGGGTTCG
GTTCGGTCTAGTAGCTCAAGCAATGGTACCCCAGATGTGGGCGAGGATAATGGAGAATT
CGGGGGCACCATTTTCTCCCGTTCAGGTGGCACTTGGGAATGCATGTGCAAAATTAAGA
ACTTAGACAGTGGCAAGGAGTTTGTTGTAGACGAGTTGGGGCAAGATGGTCTATGGAAC
AAGCTTAGAGAAGTTGGAACAGATAGACAGCTAACTATGGATGAGTTTGAAAGGTCGCT
GGGGCTCTCTCCTCTTGTGCAAGAACTCATGCGCAGGGAAAGTGGGGTGGCCCAAGCA
GATTGTAACGGAGTACATCATCATGATGCAGAGATTTCAAGCAGCAAAAGGAGAAGCTG
GCTGAAAGCCCTAAAATCTGCTGCCTATTCAATGAGGCGGCCTAAAGAGGATCAAAGTA
ACTATGATTCTGAGAGAAGTGGGCGAAGATCTGGTTCCTTTGATGTTCCTTGGGGGAAG
CCTCAGTGGACTAAAGTGCGCCACTATAGGAAACGTTATAAGGAGTTTACAGCTCTTTAC
ATGGGTCAAGAGATTGAAGCCCACGAGGGCTCTATTTGGACGATGAAGTTCAGTTTGGA
TGGGCGTTATCTAGCCAGTGCTGGTCAGGATTGTGTCATTCATGTACGAGAAGTGATTG
AATCAATGAGAACGTTTGGGGCTGATACACCAGATTTATATGCTTCCAGTGCTTACTTTTC
CATGAATGGTTTGCAAGAGCTAGTCCCCTTAAGTATTGAAGACCATGCAAACAAAATGAA
AAGAGGGAAGATCATAGGCAGCAAAAAGAGTTCAAACTCAGACTGTATAGTTTTACCTAA
TAAGGTGTTCCAGCTTTCAGAAGAGCCAGTGTGTTCTTTTCATGGTCACCTTCTTGATGT
ATTTGATCTTTCATGGTCCCCATCACAGTATTTGCTTTCGTCTTCAATGGATAAAACAGTT
CGACTATGGAAACTGGGACATGAAAGTTGTCTAAAGGTGTTCTCTCATAATGATATTGTG
ACATGTATACAGTTCAATCCTGTTGATGAGAGATACTTCATAAGTGGCTCATTGGATGGA
AAAGCCCGAATTTGGAGTATTCCAGACCGTCAGGTTGTTGACTGGAGTGATCTGCGCGA
AATGGTCACTGCAGTATGTTACACTCCCGATGGACAGGGAGGACTAGTTGGATCTATCA
AGGGAAGTTGCCGTTTCTATAATACATCAGGTAACAAGCTACAGTTGGAAAATCAACTGA
ATGTGCGTAGTAAGAAGAAAAAATCTTCTGGAAAGAAGATTACTGGTTTTCAGTTTGCAC
CAGGAGGAGATTCTCAGAAAGTATTGATAACGTCTGCAGATTCACGAGTTCGAGTTTATA
ATGGTTCTGAGCTTGTTTGCAAATATAAAGGTTTTCGGAACACCTGCAGTCAAATTTCTG
CCTCATTTGCACCGAATGGGCAGCATTTTGTTTGTGCTAGTGAAGATTCACGTGTTTATA
TTTGGAATCATGAGAGCCCTCGTGGTTCTGGTGCAAGGCATGAAAAGTCCTCTTGGTCA
CATGAACACTTTCTATCACAGGGTGTTTCAGTTGCAATTCCTTGGTCTGGTATGAAGCTT
CAGCCTCCAGTTTGGAATTCACCTGAATTTATGCTTGGTCAGAGGCACAATTTGCTTTCA
TTGCAGGGAGGTAAGGATGTTGGATGTCAGAATGGTCTTCTTTCAAGAGAAGCTGGTGA
GGGGCAGGAATCAGAGACTCCTTTGCATTATATTTCTCAAGTATCACATTCATGTGGTTC
ACAGAATATGGTAGATAGGGATGGTCAAGATGATCTTTCTAGATATTCAGCTTGTATCTC
GGATTCACGTCTGTCATCTTTTATGGCTTTCCCTGAATCTCCTGGAAACCCAGATGATCT
AAATTCGAAAGTTTTTTTCTCTGATAGCTCGTCTAAAGGCTCAGCAACATGGCCTGAGGA
GAAGCTTCCTCCAACAAGAAAACAGAGCCGATCAAATAGTACTTCTTCTCACTATGACAC
TCTGAAGACTCATCTTGGTAATACCATCCAAGGCCAATCTGGAGCTTCTGCAGCTGTGG
CATGGGGATTAGTAATAGTCACTGCTGGTCACGGGGGGAGAGATCAGATCATTTCAGAAT
TATGGCTTGCCTGTTCGTCTTTGAGAAGTCAAGTATTTATAGTTAAGTGAGAGAGCAAGG
AAACTTTACCCTTTTAAATTGTCATTTCAGCATAAATACACATTTGTATTGAATGCTTTCCT
CATCTCAAATTGTATTTAATCTGCGGGCAGTATGGGAAGATTGCCATTCTTTCAGAAAAT
GAAGGTGCATCTTGATGCTTCCAAGCAGTCCTTGCATTTTTTTAGAGTATTGTGGGCTTC
CAACAATGTGGAGTTGGATAAGTTTGACACTCGATCACAATTGTCCAATCTGTTCGTTAT
CATTTTCCTTTGCTACTCAAAAAAGTGTATGGCATCTGATTGGGCCCTCTTATGCCCTGC
TGATCTGCTCTCCATTGAAGTAGAAAAAGTGGGGGACGGATCAATCAATAGCTCTACTGT
GCAACGAACCCTCTGCTCCATTCATCCAACCAGCTGAGAATGTGTGGTTATCAGTACCAC
TGACTATACACTTAAGTTATTCAGTTGATGCAGAGGTGCTCCTAAGATTGTGTCGCCCCT
ATTGCAGCATGTTCAGCAAGGGACCGTTAGCCCTTGCCTTCCCCAAGTATATAATGTGTA
TGGCAAGATGACTCGTCCAACAGCCTTGAAGATGTAGTCAAGAGTGCAAGCGATTTCCA
CTTGTTTAGTTCCCTGAAGCTGGATTTCTTCAGCCAGGCACGCTTAACTAAATTTCGTTTA
GTTCACCATGACTATTCGTTGAACTTAATGAGATTTAGTTTAACGTAGCCAGCCCGCTTCT
GCATTATCTTAGAAACGGATGTATATTGGAACAGACCAGGTGCTTTAGGTGATTAAGTGC
TGGACAAAACGAGGATCCCTTGAAATTCTTGAGAGAACTTTTTGTTGTTGTTGAGATGAA
CAACATTTGTGAGTTATCACAAATGACCATGCTCTTCGTGCCTGAGACTTTTCCCTGTAG
TATAGACTTGAGTTGTCCACGTGGTAGTGATGGTTTTGCTTTAACCCAATGCTATGTCCA

TGTAAACAGATATTTTGGACATTTTTTGATCTTACTAGTACTGCTTAAGAATTTGTAAATGT
AACTGCTGCCATGAAGATCGTTAGGGCTTTGACAGCAGCATTGTTGGTTAGCCAATTTAG
GATGTTATGATTAAATGCTAGTGTTTTCCAAGTGGTTGTAGCATATCCACTGTTATAACAT
ATTCGGGAGTTAAAAGATGAAATATAAAGGGTCCTACAGACAACCTAAGTTTCAATTCTTT
TTTTGGATTCGACATAGGACGGTGGATACAATTGTTATCCTGCAAGCAATGAAATAAATG
CGCACGAGTGCCTTTTTGTCATGCAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 247
ACGAATCAATAGTATTAAAGGGCTTTCAAATACTTGCTTTGCAAAGCCCGGGAGGCATTG
GGGACATTGTAGATTGACCTAATAGCCGATTCAAATTTATCACACGATGCCAAGCATACC
AGCAATTGGAGAGTTTACAGTCTGCGAGATCAATCGTGAGCTGCTTACGACTAAAGATGA
ATCCGATACACAAGCCAAGGATGCCTATGCAAAAATTCTGGGGCTTGTGTTCCCTCCTAT
TTCTTTCCAGATAGAAGAGGGGTTTGGCAGTGCATCTAGACAGCAGTTTGATCAAGATCT
GGATAGAGAAGACACAATTGTCACCCCAAGCACCAGTGAGGGCACAAATGCCCTGCAAG
AAGGGGGTCTTCTTTTGAAGGGTGTCTCTGTTCTAAAAAATATCCTGGCCAGTTCATTTG
GACCGATTTTTTCTCCTAATGATACAAAAGTACTAAAGAAGGTTGAGCTTCTTCAAGGAAT
CAGTTGGCATAGGCATAAGCATATTCTAGCATTTATATCGGGATCAAATCAAGTTACTGTA
CATGATTTTCAAGATCCAGAGTGGCGTGAGTCGTCTTTATTAGTTAGTGAATCTCAAAGG
GGCATTGAGGCTCTTGAATGGCGCCCAAATGGTGGCACAACACTGTCTGTTGCTTGCAG
GGGTGGGATCTGCATTTGGTCTGCATCATACCCAGGAAGTGTAGCACCTGTGAGGTCAG
GGGTTGCTTCTTTCCTGGGCACTTCTACTCGGGGAAGTAGTGTGCGATGGACATTAGTT
GACTTTCTCCAAATTCCTGGTGGCAAAGCAGTGACGGCACTATCTTGGAGTCCTACGGG
GAGATTGTTGGCCTCAGCATCACGTGAAGATTCATCATTTACAATCTGGGATGTTGCTCA
AGGAGTCGGAACTCCTCTCCGCCGAGGACTAGGAGGAATATCATTGCTGAAGTGGTCTC
CGACAGGGGATTATTTGTTCTCTGCAAAACCAAATGGGACGTTCTATCTATGGGAAACAA
ATACATGGACTTTAGAGCAATGGTCATCTTCTGGAGGCTGTGTTATTAGTGCAACTTGGG
GACCTGATGGACGTATGCTTTTTATGGCCTTCTCGGAGTCCACAACATTGGGCTCACTTC
ACTTTGCCGGAAGACCTCCATCACTCGATGCACATTTACTACCGATGGAGTTGCCAGAAA
TTGGATCCATAACAGGAGGGTTTGGCAACATTGAAAAGATGGCTTGGGATGGATGTGGA
GAGAGATTGGCTGTGTCATACACTGGTGGAGATTTGATGTACGTTGGTCTTATTGCAATC
TATGATACAAGGAGGACACCATTTATATCAGCATCATTAGTAGGATTTATTCGAGGACCA
GGAGAACAAGTCAAGCCTCTTGCATTTGCATTCCATGATAAGTTCAAGCAAGGACCCCTA
TTATCTGTGTGTTGGAGTAGTGGTCTGTGCTGTACATACCCTTTAATATTTCGAGCACACT
GATTGCTGGTCAAAACCCCTTGTAGGGTGGACTTCTGTTGTATCCAATTTTTATGGCATA
ATTAGCTAGTTTTGGTATTTGCATAGATTCCAGAATTCCATTTAGTTTTTTAGCTCGAGAA
GAAAGAGTTCTTTTCAAGAATTGTGTTGTTTTTAAAAATCGATAAGCCTTAATATGATATTT
TGAAGGAAATCACAATGCCATTGTAAAATGTTTGGCTTCCTATTCATATCGTGTCTATGTT
GGTTGCCCTTTCTTAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 248
CAGCGAAGCTGCTAATTCTACCTGTTTTAGAGCGGGGGTTTTGAAAGGGTTTTAGATTCC
ATCTGTTTTGAAGGGTTTTAGATTTGTCAGGTTTCCGGAAACGTGGATACCTGCATAATG
GAGGAAGAAAACGCAAAACACACGGAGGAGACAAGGCAGGTGCAAGTTCGATTCACAA
CGAAGTTGCAGCCTGCCCTGAGAGTTCCCACCACGTCTATTGCGATTCCTGCTCACCTC
ACCAGATATGGTCTCTCCGATATCGTCAATACTCTCCTCGGCAATGACAAACCTCAACCT
TTCGACTTTCTTGTTGAAAGTGAACTGGTTCGAACATCTCTTGAAAAGCTGCTCCTCATCA
AGGGCATATCTGCGGAAAAAATACTTAACATTGAATATATATTGGCAGTGGTCCCACCCA
AACAAGAAGAGCCATCGTTGCATGATGATTGGGTTAGTGTAGTTGATGGTTCTTATCCCA
ACTTCATATTTAGTGGGTCGTTTGATAGCATTGGAAGGATATGGAAGGGAGAAGGCTTAT
GCACACATGTTTTGGAGGGACACAGGGATGCTATCACTTCTGCTGCTTTCATCATGCCAT
CAGATTCTAGTGATAGTTTTATAAATCTAGCTACTGCCTCGAAGGATCGAACTCTACGAC
TGTGGCAGTTCAAGCCCAATGAGCACATGACAAATGGGAAGATGGTTAGACCATACAAA
CTTTTGAAAGGCCATACCTCATCTGTTCAAACTGTATCGGCATGTCCACGTAGGAATTTG
ATTTGTTCAGGTTCTTGGGATTGCTCTATAAAAATATGGCAAACTGCGGGAGAGATGGAT
ATTGAGAGCAATGCTGGTTCAGTGAAGAAGAGAAAATTAGAAGACAGCACAGAACAGAT
TATTTCTCAGATTGAAGCATCAAGGACGTTGGAAGGACATAGCCAATGTGTTTCTTCCGT
TGTCTGGCTTGAAAAAGATACGATTTATTCTGCCTCTTGGGACCATTCAGTGCGGAGTTG
GGATGTTGAAACTGGCGTTAATTCATTGACAGTGGGTTGTCGGAAAGCACTGCATTGTCT
GAGTATCGGTGGTGAAGGCTCTGCACTAATTGCTGCAGGCGGGGCTGATTCTGTGCTTC
GAATATGGGATCCGCGTATGCCAGGGACTTTTACTCCCATTCTTCAACTGTCATCGCACA

AATCCTGGATCACAGCTTGCAAGTGGCATCCAAAATCTAGGCATCATTTGATCTCAGCTT
CTCATGATGGGACATTGAAGTTGTGGGATGTCAGAAGCAAGGTCCCGCTTACCACATTG
GAGGCACATAAAGATAAGGTGTTGTGTGCAGACTGGTGGAAAGAAGATTGTGTGATATC
TGGCGGGGCTGATTCAACGTTACAGATATTTTCAAACTTAAACCTCACATGAATCGGTGG
ACATTCTCCGGTGCATGTATGGGCATATTTTTGATAGAGTTCGCGATGTTTTAAATTTATC
CAGAGTCAGTATTAGAGCAAGTCCCAATAAAAGTCTTTGG

SEQ ID NO: 249
GGCACGACTTTACCCTTGATGTGTGTTAGGGGGCACTGTTTAAAACCCTCGAGGGAGGG
CTGAGGGCTGCTACAGCATCCGCATTTGCATCTCTGCTACGACTAAGGAATCATCCTCTT
CTACAATGCCATGCCGTGATCGGTCGATTGCATTAAGTGCTGCAAGGATCAAATAGTGG
CACTGTCATGAACAGGTTGAGGTCAAAGCGCAATCATATTTTGGAATTAAGGCTAGGACA
ATCAGAACCAGAAAAAGAAGCAACATTGGCGTCTAATCGGAGCAGGGGCACGAATGCCC
CAATAGTAGTTGAGGATGATGATGATGTAGTGGTGTCATCTCCAAGGTCATTTGCTCTAG
CAAGGAGTTCAGTGTCTCAACGAAGTAGTCGCATACCAATAGTAAATGAAGAAGACTTGG
AGCTTCGACTTGGACTGGCAGTTACAGGAAGGACATCAGCAGAGCACAATCCCCGTCGC
AGGCATGGTAGAGTTCCTCCAAATAAGCCAATTGTTCTTTGTGATGATGCTGGAGAAGCT
GATCAAAGTTCTTCCAAGAAGAGGAGGACTGGACAGCAATTAAGCAGTGATGTTCAATCT
GATGAATCAAAGGAAGTTAAGCTCACATGTGCAATTTGTATATCCACGATGGAGGAGGA
GACATCAACTATATGTGGGCACATCTTCTGTAAAAAATGCATTACAAATGCTATACATCGA
TGGAAGAGATGCCCAACTTGCAGGAAAAAACTTGCAATCAACAACATCCACCGTATCTAT
ATTTCTAGCAGCACTGGTTAGCTATTCATCGTTGTACTTGTTGCCAAATTCTGCTAGTTCG
TAGGAGAAGCTATGTGTTTGTACTTCCTTATGTGTGACATAAGAGAATGTCAAAAGAGAC
ATTTCTCGCACTTGTTCTTTATCTGCACGTTAAAATGTTTTTAGAAACATCTGCCAAAGCA
CTGATGATGATATTTACAACTTGAATGATGACGAAGGCTATTGGAACAGCCATTGGGATC
TATTTAAAAAAAAAAA

SEQ ID NO: 250
AAATGCGGAGTGACAATTGAACTCTTCCCATTCCTTCAGTGCGCGGAATATCTTGGGGAA
AAAAAACCGCTGTGTTCATTTTTGTTGATCAACAATTGCTTTTTATCATCGCCACGAAGTA
CAGGGTAGGTTTAATTCGCTTTTGTTCAAGGAAGCCTGCGTGATTTTGTCGATGGTAATC
TATTCTCCTTGGGAGGTGCCCCAAATTAATGTTAATATCAACGGGAGTGGAATTTTCGTT
CCCAGCAGGCGCGAAGAGGAAATAACTAAGTTGCTCTGCAAGCACTGCATGAAATCGCA
TTAATTAAACCTTATTTTGGGTTCTTGAGTAGTGCTATCTATAAAATAATACAAAACTCGCA
ACAATAGAAGGCTCCCAGGTCACTATGGAAGAGCCACCCCCACCTGCGGTGTTACCGTC
ATCTGAGGATACCTCCATTGTTAGCTCGCATAGCTTTGTCAATGCACCCCCAACTGTCCC
TGTTGGTTTAGATGCTTCCATTCCCCAAATTTCTACGCCGGGAATTAACCAACCTGGCTT
AACCATCCCTGTACCACCCGAAGCAGCTCCTCTAACTGCGTCATTAGTTGCTGCTTCTGC
TGGCATGCCCCCTGCAGTTGTACCTTCCTTTGTCCGCCCAGCTATCGTAGCCCATCCAT
CGGTTATGCCCCCGCCCAGCATGCCTCTTGCAGCCTTGCCCATGCCAGTTGCTTCTGCT
GTACCTGTTGCTGCACCACATTTCCCCCCTTCAACCCCAAACGATAATTCTATTACTCCAT
CTATGCCCGTTCCTACTCCCATTGTAGCCTCATCTAGTGTGCCGCCTTCGGTCACAATTC
CTGGAATTGCTCCCCTGCCATTTATTGCTCCTATTCCTGTGCCTTCGTCGCGCCCTGTAG
CTCCATCGCCCTTTATGCCTCCTGCTCGTCCCTTAGGAGCTAGTGTTTCCGTTGCAATGG
ATGTAGACAATACTGATGAACAAGATCAAGATGCAGACAACAAGGGAGAAAGTCCTTCGT
CAAGTCCCGACCATCCAGAAGATCCTTCAGCAGCTGAGTATGAAATTACAGAGGAGAGC
AGAAAGGTTCGAGAGAGGCAAGAGCAGGCTATTCAGGAGCTTCTGCTGCGGAGGCGTG
CCTATGCATTGGCAGTCCCTACCAATGACTCTTCGGTGCGTGCTCGTCTTCGTCGCTTG
AACGAACCTATTACCCTGTTCGGTGAGAGGGAGATGGAGAGACGTGACCGGCTCAGAG
CGCTCATGGCGAAGCTTGATGCGGAGGGACAGCTTGAGAAGCTCATGAAGGTTCAGGA
AGAAGAAGAAGCAGCTGCCAATGTTGATGCTGAAGAGGTCCAGGAAATGGAAGGGCCT
CAGGTATATCCTTTTTATACTGAAGGCTCCCAGGAGCTGCTAAAAGCTCGGACTGAGATC
ACGAAATTCTCACTCCCTCGGGCAGTCTCGAGGTTACAGAGAGCCAGGAGAAAGAGAGA
GGATCCTGATGAGGACGAGGATGAAGAGTTGAAATGTGTGTTGCAGCAGTCAGCCCAGA
TCAATATGGATTGTAGTGAAATTGGAGATGATAGACCTCTTTCTGGATGTGCGTTTTCTTC
AGATGGAACTCTGCTAGCAACGAGTGCTTGGAGTGGTGTCACAAAATTATGGAGTGTAC
CAAACATAAATAAGGTTGCTACTTTAAAGGGACATACGGAACGAGTTACTGATGTGGCAT
TTTCTCCTACAAACTGCCATTTAGCGACTGCCTGTGCTGATCGTACTGCAATGCTATGGA
ATTCTGAGGGAGTTCTGATGAAAACATATGAAGGCCATTTGGATCGTCTTGCTCGTCTCG
CATTTCACCCATCTGGGCTATATCTGGGCACTGCTAGCTTTGACAAGACGTGGAGATTGT

GGGATGTTAACACTGGTATTGAATTGCTCTTGCAAGAAGGCCACAGCCGAAGTGTGTAT
GGGATTGCTTTTCAGTGTGATGGTTCATTAGCAGCAACCTGTGGATTAGATGGGTTGGC
ACGCATTTGGGATCTTCGTACGGGAAGAAGCATTCTTGCTCTAGAAGGCCATGTGAAAC
CAGTCCTTGGCATAGATTTCTCGCCAAATGGTTACCACTTAGCAACTGGTAGTGAGGATC
ATACCTGTCGCATTTGGGACCTAAGGAAGAGGCAATCAGTATATATCATACCTGCTCATT
CTCACCTTGTTTCACAGGTCAAATTTGAACCACAGGAGGGATATTTCTTGGTCACTGCTT
CATATGACAGTACTGCTAAGGTGTGGTCAGCACGGGACTTCAAATCCATCAAGGTATTG
GCAGGACATGAAGCAAAGGTTACGAGTGTGGATATTACAGCAGATGGGCAATATATTGC
TACTGTGTCACATGATCGTACCATAAAACTTTGGTCCAGCAAAAACAGCACTAATGATAT
GAATATTGGTTGACACCATTACAGATGGGCGATATGTTTCCAGTGTCACATGATCGTACT
AGAAAATTTTAGTCCAGCAAAACCATTAGCAACACTAATGATATGAATATTGGTTGACATC
ATGATGGGTAGTGTTGTTATGTCGTCTCTGTAGGCGATGTGAATGCAATTTGTTGCTGCC
TTCGCAGTATTCAAAACAAAACATTGAAGCAAATGCTGCTGATCCAACGTTTACAGTACTT
AATCTTCCTATAATGCAATCTGCGTTGAGAACATTCTTTTTTCTGAGATATCCATGACATG
TATATGATAAGTGGGCAGTTAGGCTTCACTTTCAGTCTTGCTTCAGCTACTCGGTTGGCA
AAAATTTTGTGAAGGGAAAGGGCAACACATCACTGGACATAATTTTAAAATATGATGCAG
AACTTTTGTTCACGAATGCCTCGTGTACTTTATACAAATTACATTTGGAATGCTAAGATGT
AAATCTACTGCTACTTGAAAAAATGCTTCTTGAACTGATATGCTCCAGG

SEQ ID NO: 251
CTTTTACTCTTGCATTTATGCCATCTGAAAACCCGAGGTGTAGTTAATGCACCTCGGCAG
AAGCCAAGATCTCCTTCGGGCTTTGCCGAGAAGTACGAGTGGATTGAAAGGAATTTGGA
GAAATAGGAGCAGGAAGTGAACGGATTTTAAAACTTACAGGATCCTGTGACATGGGCAG
GTTTTTTTGAATGGTAATGGGCGGTTTTTGATGTCAGATTTGGTGCTGTGAAAACATAAG
AGCCGGCATTGTGCATTTCATAGTTTTTGGTGGTGGGGATTGGACTGAGAAGAGTTATG
CGAAACCAGCGATATAGAAGTCCAACACTGATTTCAGAGCTTTTCTCCGGTCACCAAAAC
CATGAAGCGGGCTTACAAATTGCAGGAGTTTGTTGCGCATGCTTCCAATGTCAACTGTCT
CAAGATTGGGAAGAAGTCTTCCAGAGTTCTGGTGACGGGCGGGGAAGACCACAAAGTG
AATATGTGGGCTATTGGAAAACCGAATGCCATTCTGAGTTTATCTGGTCATTCAAGTGCT
GTGGAGTCTGTGACTTTTGATTCTGCAGAAGCTTTAGTTGTCGCTGGAGCTGCTAGTGGT
ACAATAAAGCTATGGGATTTGGAAGAAGCAAAAATTGTTCGGACACTCACTGGTCATAGG
TCCAATTGTATATCAGTGGATTTCCATCCATTTGGGGAATTTTTCGCATCTGGCTCCTTGG
ATACAAACCTAAAAATCTGGGATATTAGACGTAAGGGTTGCATTCACACTTACAAGGGGC
ACACTCGTGGTGTTAATTCAATCAGATTTAGTCCAGATGGTCGTTGGGTGGTGTCAGGTG
GGGAGGACAATATTGTAAAGTTATGGGATCTAACTGCTGGAAAGCTCATGCACGACTTCA
AATGCCATGAGGGTCAGATACAGTGCATGGATTTCCATCCTCAAGAGTTTCTTCTTGCTA
CAGGCTCAGCAGACAGGACTGTGAAATTCTGGGACCTTGAGACTTTTGAGCTTATTGGA
TCAGCTGGTCCTGAGACAACTGGAGTTCGTGCCATGATTTTCAATCCGGATGGAAGGAC
TCTGTTAACTGGATTGCATGAAAGTTTGAAGGTGTTTTCCTGGGAACCTTTGAGATGCTA
TGATGCAGTCGATGTTGGTTGGTCTAAATTGGCTGACCTCAACATACACGAAGGAAAGCT
TCTTGGTTGTTCATACAATCAAAGTTGTGTTGGTGTATGGGTTGTGGACATTTCGCGGGT
GGGGCCATATGCTGCTGGAAATGTATCAAGAACAAATGGCCATAATGAAGCAAAATTGG
CTTCCAGTGGTCATCCATCTGTCCAGCAATTAGATAATAACTTAAAGACCAATATGGCGA
GGCTTTCCTTGTCACACAGTACAGAGTCAGGAATCAAGGAACCAAAGACTACCACATCG
TTAACTACCACTGAAGGTCTTTCTAGCACACCTCAACGAGCTGGAATAGCCTTTTCTTCA
AAGAATCTTCCTGCAAGTTCAGGTCCACCGTCATATGTCTCGACTCCAAAGAAAAATAGT
ACATCAAGGGTGCAGCCTACAACAAATTTTCAAACCTTAAGTAGACCTGATATAGTGCCT
GTCATTGTCCCTAGAAGCAATTCATTAAGACCGGAAACAACATCAGATGCTAAGAAAGAA
ATGAACAATTTTGGAAGAGTGGTTCCATCTACAGTATCAACCAAATCAACTGATGTGATTA
AATCTGGCAGCAACAGAGATGAATCTGACAAGATAGACTCCATAAATCAGAAGCGCATG
ACAGGCAATGACAAAACAGACCTAAACATTGCTAGGGCTGAGCAACACGTTTCCTCTAG
ACTTGACAATACAAACACTAGTTCTGTTGTTTGTGATGGAAATCAACCAGCGGCAAGATG
GATTGGTGCAGCCAAATTCAGAAGAAATTCACCAGTAGATCCAGTTGTAAGCCCACATGA
TAGAAGTCCTACTTTTCCATGGTCTGCAACTGATGATGGAGTTACATGTCAGCCAGATCG
ACAAGTTACTGCACCTGAATTATCAAAAAGAGTGGTAGAGCCTGGTCGTGCTCGTGCTCT
GGTTGCAAGTTGGGAAACACGAGAGAAGGCTCTCACCGCAGACACACCTGTGCTGGTC
AGTGGTCGCCCCCCCACAAGTCCTGGAGTGGACATGAACTCATTCATCCCGAGAGGAA
GCCATGGGACTTCGGAAAGTGACTTGACAGTTAGTGATGACAACAGTGCTATAGAAGAG
CTCATGCAACAGCACAATGCATTTACAAGCATTCTTCAAGCTCGCTTGACTAAGTTGCAG
GTAATAAGGAGATTTTGGCAAAGGAATGACTTGAAAGGTGCTATTGATGCTACGGGAAA

GATGGGAGATCATTCGGTATCTGCTGACGTTATTAGCGTACTGATTGAGAGAAGTGAAAT
CTTCACGCTGGATATTTGCACAGTCATACTTCCATTGCTTACTCGGTTGCTTCAGAGCGA
GACTGACAGGCACCTTACTGTCGCTATGGAAACCCTGCTTGTGCTAGTGAAGACTTTTG
GTGATGTTATCCGGGCAACTATATCAGCAACCCCAACAATTGGAGTTGATCTCCAAGCAG
AGCAAAGGCTTGAACGCTGTAATCTGTGTTATGTTGAACTGGAAAACATCAAACAGATTC
TTGTTCCTTTAATCAGGAGAGGTGGAGCTGTTGCCAAGTCTGCACAAGAGTTGAGTTTAG
CTCTTCAAGAAGTGTGACCCACTCTATTTAGGTTGACATTTTTTTTTGTAATAAATCTTCTT
TGACGGGAGGCTGTTTGTATATTGGGAAACACTGCATTGTTTGCAATGATAGTGGCAGT
GTGGTTGAATATGCACGGCATCAATTGCAGTGGAATTCTTATCTCAGGTCATTGGAGTGT
TTAACCATTCCTATGTTGTCAGCCACATCTTGTCAGCCACACCTTGATTCTTAGATTAAGA
AGTTACTAATTTGTAGATAAATTCTAACGAAGGTGATGATAGCATACACGTAATGAAATCA
CAGAGAATGTCTACGCAGTTGACACTGCAACTAGTCTTGCGTGAACCTGTAGTTGCGGC
ATGCTTCAACCTTTGTTGTAATATTTAAAATATTTGAGTGGCTTAAGCACTCAATTTCCTGT
TTGATCAAGTCAGTTGCTAGATTCTTCAGTTGCTTTTCTCCGTTCATACAATCAGCAATTG
GATATTGATGAATTTCATTACAGAGTTTGATAATCTGTGAGCCATCATATATTAATTGTACA
TTCATAAGCATCAAGATACGTAATTTAAATGATTTCCTTGTGTCTGAAAAAAAAAAAAAAAA
AAACTCGAGACTAGTTCTCTCGTGGTGCGATCCTCTACTCCAAGGGCGGAAAGGCCGTT
GCAAGTAAGTTCCATTGGGAGAGTCTGGGTGATTGAAACCCGCCAGGAAGACGTACTGC
AGTATAATCTCTGCTCCGAACAGGACCTTGATCTGCGATATCAACCTCAAAACTCATACA
GCACTGAAGAAAAAAACTGTTGACTAGGAATGGAGTTGAACTCAGGTCCTGAAGTTGTTT
TTGATTGTTCTATGTCTACAAAAAAATCATGCTCTGGGGCATCTGCAAGAGAATGGCGAA
ATTTCATGTCATCCAACCCATCTGATATGTTCAAACGCACACAATCCAGGAGATTAACTG
AGCAGGCGTCTTCTCATGAAGAGAAGGAGATATTCGACACAATATTTAAAGAAGTTGAAG
AGTTAGGAGTTTCACAACTGCCCTGGAAAAAACGGAAGGCCATGGAGAACCAAAAGGTG
TTAACTTTAGGTGGAAAGCCTTCAAAGAATATCAAAACATCCATTTCCGTAGGTAAACATA
TTAGGAAGAAGCAAAAACGTACGAAGGAGGCGGCAATTAAGAATGAGGAGATTCTTCTC
GGTCGACCTGCAAAACGTGCAAAGAAAACTGAAAATATACGCAAGAGAGAGGATCGAGG
GCTGATGGCATCCGAAGGATATTTCAAATACGGGATTTTGCATGTTAAGCCTTTGGAGAA
AAATGTTGAAGAACGTGAGGAAAATGTATACAGAATGGGAAAAACAAAGAAGAAAGGCA
AATCAAAGAAACACAAGGGAAAAAAGAAGCGTTGATAGTTCTGCATTTCTGAGATGAAGC
ATCTTTCTTTTTATTGTTGGGCTTAGCTTTGGCACTATTGGAATATTTTAGCATCCAGTTTG
CCTTCAGCTATTACAAAAATATTTGAAAGGTTTTGTTTTAAATTTTTTTTTTTTTCACAATTTA
TGGATGAGGTACTCCTTATGAATATCTTCAAACTAAGAAATAACTATATATGCAAATAAAT
AAAAACGATTTGTAATTTTTCAATTTCCAAAAAAAAAA

SEQ ID NO: 252
GTGTGCAGTTTCACTACTCCCCCACACACTCTCTCTCTCTCCTTTTCCCCCAAATCAG
AAGAAGAGACGACGACTGTGTAGTAGTGAGAGAAGCGGCCGCAGACCGGGCGGTACTG
AAATTTTGATGGAGGGTTTCGGGGGAATCTCTCTCTGCTGCTGATCTCCGGGAGGGCTG
GTTCGTCTTCGGCTTCGGGAGGGGGAGGAGCGGGAGATCGTGATTTCGAAGGAGGCAG
AGATGGCCGGATCGGACGAGAACAACCCGGGCGTTGTCGGAGGTGCGCACGTTCAAGA
GGGCTTGCGGGTCGGAGCGGGGAAGATGGGGGCGGGGAATGTTCAGCAGAGACGAGC
TCTGAGCAACATCAACAGCAACATCATCGGGGCTCCTCCTTATCCATGCGCGGTCAACA
AAAGGGTCTTGTCCGAAAAAAATGTCAACTCTGAAAACGATCTCCTCAACGCTGCTCATC
GGCCAATTACTAGGCAGTTTGCTGCTCAGATGGCTTACAAGCAGCAACTTAGACCTGAG
GAGAACAAGAGGACGACCCAATCAGTCTCAAATCCCAGCAAATCTGAAGATTGTGCCAT
CTTAGATGTGGATGACGACAAGATGGCTGATGACTTTCCGGTGCCAATGTTTGTGCAAC
ACACCGAAGCAATGTTAGAAGAAATTGATCGGATGGAGGAGGTTGAGATGGAAGATGTA
GCTGAAGAACCTGTCACGGACATTGACAGCGGTGATAAAGAGAACCAGTTGGCTGTTGT
TGAGTACATTGATGACCTATACATGTTCTATCAGAAAGCCGAGGCTTCTAGTTGCGTTCC
CCCAAACTACATGGATCGGCAGCAGGATATTAATGAGCGGATGAGAGGTATACTAATTG
ACTGGCTGATTGAGGTTCATTACAAGTTTGAATTGATGGATGAGACCTTGTATCTTACGG
TCAATCTCATCGATAGATTCTTAGCTGTTCAACCTGTAGTGAAGAAAAAACTCCAGCTAGT
AGGGGTAACAGCCATGCTTCTGGCATGCAAATATGAAGAGGTCTCAGTTCCAGTAGTGG
AGGATCTCATTCTGATTTCAGACAGGGCTTATAGCAGGAAAGAAGTTCTAGAAATGGAGC
GGTTGATGGTGAATACTTTGCACTTTAACATGTCAGTGCCTACTCCTTATGTTTTCATGAG
GAGATTTCTTAAAGCCGCTCAATCTGACAAGAAGCTCGAGCTCTTGTCATTCTTCATCAT
CGAGCTTTCCCTGGTTGAATATGACATGCTGAAGTTCCCACCCTCTTTATTAGCTGCTTC
TGCAATCTACACTGCTCTGAGTACAATTACCAGAACTAAACAGTGGAGTACAACATGTGA
ATGGCACACCAGCTACTCAGAAGAACAGCTTCTGTAAGTTACTTAGTTTTTCCTTCTTCCA

CACCTTGAGATGGCTTGGTTTTTGTTGAGCTTTCTATTTCAAGCAATTTGTGATTGGGGG
GTTCTGCATTCTTTGTAGGGAATGCGCCAGATTGATGGTGACTTTCCATCAGAGGGCTG
GATCAGGGAAGCTCACTGGTGTGCACCGGAAGTATAGCACATCCAAGTTTGGTCATGCT
GCGAGAACTGAGCCTGCTAACTTTCTCTTGGACTTCCGTTTGTAGTGGGTTGCGTGTGT
CGTTGTGTACCTTGCCTAAATCAAACATAAAAAACTGACTTTTGGGCCAGGGTGGGGGG
CTTAAAAATATAGCAACATAAGTTGCCCTGCGAGTG

SEQ ID NO: 253
CCCTAAAACTCATTTCTTCCTTCCCTTCCAAGCTCTCTCTCCTCTGTTTTCTCTGCGCATG
CAAGCTCCACGAGAGGGAAAATCGGCGGCGGCCATCGTCGGAATGGGCAAGTACATGA
AGAAATCCAAGGCAATCCCCCGCGACGTCTCGCTCCTCGAGGCCTCCCCGCGCTCCCC
CTCCGCCACCGGCGTCCGGACCAGAGCCAAAACCCTAGCCTCGCGGCGCCTCCGGAG
GGCCTCCCAGCGGCGGCCGCCGCCCCCCGCTGCCGCCGCCGCCGCCGCCGCGCCGA
GCTTGGACGCCTCTCCCTGTCCGTTCTCTTACCTCCAGCTCCGGAGCCGGAGGCTCCG
GAGACCCCGTCTCGCGCCCTCCCCGGAGGCGAGGATCGACGAAGGACCCGCGGGGAG
CGGTTCTAGGGGAAGCCGCGACGCTTCGTGCTCGGCGAGGACGGCATCGTCGAGCGG
TGGAGTGGAGGGAGAAGGAGCGTGCGTCGGTCGGGGTGACAGGGGAAACGGGGGGG
AGTGTGTCCGCGATGCGGCTGTCGACGCGTCTTATGGAGAGAACGATTTGGAGATCGAA
GACAGAGACAGGAGCACAAGGGAAAGCACGCCGTGTAGCTTGATAAGGGACTCGAACG
CAAACACGCCACCAGGGTCAACCACAAGACAGCAGAGCTCATGCACTGCTCACAGAACT
CAAATGAGCATACTTAGAAGTATCCCGACCTCAGATGAGATGGAGGAGTTCTTTGCATAT
GCCGAGCAACGACAACAAAGATCATTCATTGAAAAGTACAACTTTGATATTGTCAAGGAT
CGTCCCCTCCCAGGTCGTTTTGAATGGGTGCAAGTAATTCCATGAGATTTGCATTGATGC
TGGACGAGGAGGTGTATCGCATACGGAGTCAATGTACTGGAGTAGTAAAGAGCCTGAAC
TGTGACATAACAGACGACGAGCACACTGATGAGATATTAAATTAAAGTCCCTTGACAATG
TCTAAAGAGCCGTGATCTATGAGTAGATTAGAAACCGCCTTTTTAGTTGCAAACGCCATT
GGAGTTTGTTGAGAAAGGAGAATATACCCCTTGTACATAACTTGTTTCTTCAGTATATGCT
CTTTTCTCCTAAAAAAAAAA

SEQ ID NO: 254
GCCAACTCGACAGAGAGAGAGAGAGAGAGACAACACGGTTGAAATGGACGGTCACTCC
TCGCACCTCGCCGCGCAGAACCGTTCGCGCGGCTCCCAGACCCCCTCCCCCTCCCACT
CCGCCGCCTCCGCCTCCGCCACCTCCTCCATCCACCTCAAGCGCAAGCTCTCCGCCGC
CAACGCCAGCGCCGCCTCCGCCGCCGCCGCCGCCGCCGCCGCGGCCGCAGCCGCCG
ACGACCACGCCCCACCCTTCCCTCCCTCCTCCATCTCGGCCGACACCCGCGACGGCGC
CCTCACCTCCAACGACGACCTCGAGAGCATCTCTGCCCGCGGCGGAGGGGCCGGCGA
CGACTCCGACGACGATTCCGACGATGAGGAGGAGGACGATGGCGATAACGACGGCGG
ATCCTCTCTCCGCACGTTCACCGCGGCTCGGCTCGAGAACGTGGGCCCGGCCGCGGCT
CGGAACAGGAAGATCAAGGCCGAGAGCAATGCGACGGTCAAGGTCGAGAAGGAGGACT
CCGCCAAGGACGGCGGTAATGGTGCCGGCGTGGGAGCTCTGGGGCCAGCTGCGACCT
CCGGCGCCGGTTCGGGCTCCGGAACTGTGCCGAAGGAGGATGCTGTGAAGATATTCAC
CGAGAATTTACAGGCGAGTGGGGCTTACAGTGCCAGAGAAGAGAACTTAAAAAGAGAGG
AGGAAGCAGGAAGACTCAAGTTTGAATGCCTTTCAAATGATGGTGTTGATGACCATATGG
TTTGGTTAATAGGATTAAAGAATATCTTTGCCAGGCAACTTCCTAATATGCCGAAGGAATA
CATCGTGCGTCTTGTTATGGACAGAAACCACAAGTCAGTAATGGTTATCAGACGGAATTT
AGTTGTTGGTGGTATCACTTATCGCCCATATGCCAGTCAAAAGTTTGGTGAGATAGCTTT
CTGTGCAATTAAGGCTGATGAACAAGTAAAAGGTTATGGCACAAGGCTGATGAATCACTT
AAAACAGCACGCTCGTGACGTTGATGGGCTAACTCATTTTCTGACCTATGCTGACAACAA
TGCTGTTGGTTACTTCATCAAGCAGGGTTTCACGAAAGAGATATACCTGGATAAAGATCG
ATGGCATGGGTATATTAAAGATTATGATGGGGGAATTCTTATGGAATGCAAAATTGATCC
CAAACTTCCTTATACAGATCTATCAACCATGGTTCGCCGTCAAAGGCAGGCGATTGATGA
AAAGATAAGGGAACTCTCTAATTGTCATATTGTCTACCAGGGGATTGATTTCCAGAAGAG
AGATGCAGGGGTTCCCCAAAATACCATCAAAATGGAGGATATCCCTGGCTTGAGGGAGG
CTGGGTGGACACCTGATCAATGGGGCTATTCTAGGTTTAGAGGATTGAGTGATCAAAAG
CGGTTGACTTTCTTTATTCGCCAACTTCTGAAGGTATTGAATGACCATAGTGATGCTTGG
CCATTCAAGGAACCAGTTGATGCTCGTGAGGTCCCTGATTACTATGACATAATTAAAGAC
CCTATGGATTTGAAGACAATGACCAAGAGGGTCGAATCAGAGCAATATTATGTTACGCTC
GAGATGTTCATTGCAGATGTCAAGAGGATGTTTGCTAATGCACGCACCTACAATTCCCCC
GACACTATATACTTCAAAATTGCAACAAGGTATACTTAGTCAGTCCTTGTTATGTATGTCT
TTTGTCAACCCTTCAGGGCAGATAAAAATGATTTACTGGATGTGCAGGCTGGAAGCTCAT

TTCCAGAGCAAGGTACAATCGAATCTCCAGTCTGGTGCCGGAAAAATTCAACAGTAGAG
CATTCGGTAGACTGGAGGCCCTGACCTTACTTCTCTCTATATGAATATGTGGAGCCTTGG
ATACTTACTCTGATCCATGATTGCGCTGGGGAATTAACTAGCTTCGATTGACCATGTAAC
TGAAGACTGATAGTCATATTCCCCGAAAACTGAAATTTCAGTTCCCTTAGTACAATGTAAT
TAAGATGTCTTCACTTTCTAACTCTGGTAGATGGAAATTTAACCAGATTGTGATATTGATA
TATCCGTTCACCTCAAAAAAAAAA

SEQ ID NO: 255
CCGTACCTCCGTCTTCCTTTCTATCTCCATTCGAAGCCGATCGAATAAAAACCCTAGCCA
TCCGAACTCGACTCGCCGGTGAGCTTTCTTGATTCGGTCGGGTCCGGGATGTTCAACGG
AATGATGGATCCCGAGCTCTTCAAGCTCGCGCAGGAGCAGATGAACCGCATGTCCCCC
GCCGAGCTGGCCAAGATCCAGCAACAGATGATGTCTAATCCTGAATTGATGAGAATGGC
CTCTGAGAGCATGAAGAATATGAGGCCTGAAGATTTACGACAAGCGGCAGAACAATTAA
AGCATGTTCGTCCTGAGGAGATGGCTGAGATTGGTGAGAAGATGGCTAATGCTAGCCCT
GAGGAGATTGCAGCTGTACGCGCTCGTGCGGATGCCCAGATGACTTATGAGATCAATGC
AGCTAAAATTCTCAAGAAAGAGGGGAATGAACTTCATAGCCAGGGGAGGTTTAAGGATG
CCTCGCAGAAGTATTTGCGTGCCAAGAATAACTTAAAAGGAATTCCGTCGTCTGAAGGCA
AGAATCTTTTATTGGCATGCTCCCTTAACTTGATGTCCTGCTACTTGAAAACAAGGCAGTA
CGAGGAATGCATAAAGGAAGGCTCTGAGGCTTTAGCATGTGAGGAGAAAAATCTCAAAG
CTTTCTACAGGAGGGGCCAAGCATATAGAGAATTAGGTCAATTGAAGGATGCGGTCTCT
GACTTGAGAAAGGCACATGAAATTTCTCCTGATGATGAAACAATTGCGCAGGTTCTAAGG
GATACTGAGGAAAGTTTGACCAAAGAAGGTGGTTCTGCACCAAGAGGGGTGGTCATTGA
GGAAATAACTGAAGAAGATGAGACTTTGGCCTCTGTGAACCATGAAAGCCCATCAGAATA
TTCAGAGAAGCGGCATCAGGAATCAGAGGATGCCCACAAGGGTCCAATTAATGGTGATA
TTATGGGTCAAATGACCAATTCTGAAAGCTTGAAAGCTTTGAAAGGTGACCCAGATGCAA
TCAGGTCATTCCAGAATTTTATTTCTAATGCTGATCCCACAACTTTGGCTGCAATGGGTG
CGGGAAATGCTGGAGAGGTATCCCCTGACTTGATTAAGACCGCTTCTAGTATGATCGGC
AAGATGTCAGCAGAAGAACTCCAGAAGATGATTCAGCTGGCTTCATCGTTTCCCGGGGA
AAACCCTTATGTCACAAGAAACTCAGATAGTAATTCCAATAGCTTTGGAAATGGATCAATT
CCTAATGTGTCGCCTGACATGTTAAAAACTGCGAGCGATATGATGAGTAAGATGTCACCT
GATGATCTTCAGAGGATGTTTGAAATGGCATCTTCTTCGAGAGGGAAGGACCCTTCTCTG
GATGCTAACCATGCAAGTTCAAGCTCTGGGGCGAATTTGGCCGCCAATTTGAATCATATT
TTGGGTGAAAGTGAACCGAGTTCGTCTTATCACATACCTTCAAGTTCAAGGAATATTTCAT
CTTCTCCTCTATCAAATTTCCCATCATCACCAGGTGATATGCAAGAGCAGATAAGAAATC
AAATGAAAGACCCAGCTATGAGGCAGATGTTTACATCTATGATGAAGAATATGAGTCCAG
AGATGATGGCAAACATGGGCAAACAGTTTGGACTCGAGCTTTCTCCAGAAGACGCTGCA
AAAGCCCAGGAAGCAATGTCTTCTTTATCACCAGAGATGTTGGACAAGATGATGCGCTG
GGCAGACAGGGCTCAAAGAGGAGTCGAGACAGCTAAGAAGACCAAGAACTGGTTGCTT
GGGCGGCCCGGTATGATTTTGGCCATATGCATGCTCCTCTTGGCAGTGATCCTTCACCG
GCTTGGCTTTATCGGGAGCTAGAGCGGCGGAATCCCTTTGCTCGTTCAAGCTTTATGCG
ACAAGTCGGCCAGAAAGCAGTAGCATCGGATTGGAACACGCAATGCATCTTCGTCAATC
CATCTTCAGATGAAATTGGCTCCCCTAGGCAGGGGACGTTCTTTAGTCGTCGGAAGGTC
TGTCGGATAGAATGGGTAGAACATAGCATATACCCATGAAGTGCTGGTGGTTTGTGATTT
ATGCTGACGGTTTGGACTTCCTTCCATTTAACATAAATTGTACTGCTTGTAGAGATAAAAA
TGTCCGAGCTGGAATCTCGCAAGAGATTCTGATCAAGAAATTTCATTTTTTCTTTTTAGTC
TGGTCCTATTCTATATTTATTTTACAAAAAGAAATCCATTACTTGTTCTTTAAAAAAAAAA

SEQ ID NO: 256
AAAGAGAAGCAAAAACCTGGCCTCATCGGCGCTTCTCCTCCTTCTCTGAACTCCACATCA
GGCTCGTCAGTGGCAGAATTCCCGTCTTCTTCTTGTTGACCGCCGTTTGGTTACTGTACT
CTCATCTACCGGTGGTGCCGAAGTATTGGAGGGGAGGGGATTTAGGAAGGTGATCTGG
TTATTGAATTGGGGTCGTTTGTCTGTTGAAGGGATGATAGCGGCGATATCCTGGGTCCC
TAGAGGGGCTTCGAAGGCGGTCCCCGAGGTGGCCGAACCGCCTTCCAAGGAGGAAATT
GAAGAGATTTTGAAGAGTGGAGTTGTGGAAAGAAGTGGAGATAGTGATGGTGAGGAGGA
TGATGAAAACATGGATGCAGTTGCTTCAGAAAAGGCTGATGAAGTTTCCACTGCATTATC
TGCCGCTGATGCACTCGGGAGAATTTCCAAGGTGACAAAAGCTGGATCTGGTTTTGAGG
ATATAGCTGATGGTTTGAGGAGCTTGATATGGATAATTACGATGAAGAAGATGAAGATG
TCAAGCTATTTAGCACTGGACTTGGTGACTTGTACTACCCAAGTAATGACATGGATCCCT
ACCTCAAGGATAAGGATGATGATGATGACACTGAAGAGATTGAGGACCTGTCTATTAAAC
CAATGGACTCTCTCATTGTTTGTGCACGTACAGATGATGAAGTCAATCTTCTTGAGGCGT

GTACATTCATTAGATAATGCATTTTTTTTTGCTGCACTTTCCCAGGCTGTACTGGGGTTCAC
ATAAAAACCGCTACTGATTGGTCAATTCATTGCATAAAGCAGGTCTATTTATTGGAGCCAT
CATTATCTGATGAATCAAATATGTATGTTCACCATGAAGTAGTTATTTCAGAATTTCCCCT
CTGCACAGCATGGCTTGACTGTCCAATTAAAGGTGGCGACAAAGGAAATTTTATTGCCGT
TGGCTCAATGGAGCCTGCCATTGAGATTTGGGATCTTGATATTATTGATGCTGTAGAACC
ATGTCTTGTATTAGGCGGTCAAGAAGAGTTGAAGAAGAAAAAGAAGAAAGGAAAGAAGG
CATCGATTAAGTACAAGGAGGGTAGTCACACGGATTCAGTGCTTGGCCTTGCTTGGAAC
AAGGAGTTCAGGAATATACTTGCCAGTGCAAGTGCTGACAGGCAAGTCAAAATTTGGGA
TGTTGCAGCTGGAAAATGCAATATTACCATGGAACACCACACCGACAAGGTTCAAGCAG
TTGCATGGAATCATCATGCTCCACAAGTTCTTCTTAGTGGATCTTTTGATCATTCGGTAGT
CATGAAGGATGGAAGAATACCTTCACATTCTGGATATAGATGGTCTGTGACGGCAGATGT
CGAAAGCTTGGCGTGGGATCCACATTCTGAACACTTCTTTGTGGTGTCCCTTGAAGATG
GCACCGTTAGAGGATTTGATGTACGAGCCGCTATATCTAATTCTGCCTCCCAGTCACTGC
CAAGTTTTACTCTTCATGCGCATGAAAAAGCTGTCAGCACAATCTCATACAATCCTGCAG
CACCGAATCTTCTCGCAACAGGGTCGACAGATAAGATGGTTAAACTCTGGGATCTGTCC
AATAACCAGCCTTCATGTATTGCTTCAAGAAATCCAAAAGCTGGTGCTGTCTTTTCTGTTT
CCTTCTCGGAGGATAGTCCCTTATTACTTGCTATTGGAGGTTCAAAGGGGAGACTTGAAG
TATGGGATACATCATCTGATGCAGCTGTATCCCGACGATTTGGAAAACATGGCAAGCCG
AAAACAGCAGAACCTGGTTCTTGATGCTGCCAAGTGAGAACATGTTAAAAATATCTTAGA
TCTTCTCAATGTCGGCCTTGCTTAGTTCTTAGAGCGCATTGAAATGCCTGAAGTCAGGGG
TAGTTCTGCAAGTCGAGTCATGTTCTTTTAGCAGCTTGTTATTTAGCATTTTAAGTGTCTT
CGTATTATTTTTGGGTGTTGAGAAGGAAAAAAGAAAGTGTTTTTTCTTTTTCTTGGGGAGG
TGGGGGCTTTGGCGTGAGGTTACTGCAAGTATGCAAGAGCATCACAAGGAAAAGGTCCC
AGGGAAAAGTGTTGAACATGGTAAATGAAGCTAATGCAATGGTGTCTTTTTGACTCCGAA
AAAAAAAA

SEQ ID NO: 257
AAGTGCAGGTCAGCGTCGACACCAACAAAAGCTGGCCCCTTCCCTCCTTCGTATATATC
TCGTCTCGTCTCGTCTCGTCTCTTCCGCAGAAACCCTCCCCGCAGCAGAACTCCGCCTC
GTCATACGAAAATCTCGCGTCGTTTGCAGAAAACAAGGATCGATTTGACTACCCAAAACG
CAGATAGATAGAGAGAGAGAGAGAGAGGATGAAGTTCTGCAAGAAATACCAGGAGTACA
TGCAAGGGCAAGAAGGGAAGAAGCTTCCCGGTCTTGGGTTCAAGAAGCTCAAGAAGATC
TTGAAGAGATGCAGGCGAAGAGACTCGCTTCATTCCCAGAAGGCCCTTCAAGCCGTCCA
AAATCCCCGCACTTGCCCTGCTCACTGTTCGGTGTGTGACGGAAGTTTCTTCCCCTCTCT
TCTCGAGGAGATGTCTGCTGTTTTAGGCTGTTTTAACAAGCAGGCGCAGAAGTTGCTTGA
GCTTCACCTGGCTTCGGGATTTCAGAAGTATTTGATGTGGTTCAAAGGCAAGCTTCGAG
GAAACCATGTTGCTTTAATTCAAGAAGGAAAAGATCTGGTTACTTATGCATTGATAAATGC
AATAGCCATTAGGAAGATATTGAAGAAGTATGACAAGATTCATCTCTCTACTCAAGGACA
AGCTTTCAAGTCACAAGTGCAAAGGATGCACATGGAGATCCTCCAGTCTCCATGGTTGT
GCGAGCTTATCGCCTTCCACATTAATGTACGGGAAACGAAGGCAAACTCTGGGAAGGGC
CATGCCCTTTTTGAGGGTTGCTCTCTAGTTGTCGATGATGGCAAACCATCGCTCTCTTGT
GAACTCTTCGACTCTATCAAGCTAGATATCGACTTGACCTGCTCTATATGCTTGGACACG
GTGTTTGATTCGGTTTCTCTCACTTGTGGCCACATATACTGCTATATGTGTGCTTGCTCA
GCTGCATCTGTGACGATTGTTGATGGACTGAAAGCAGCAGAACCCAAGGAAAAATGTCC
TTTATGTCGAGAGGCCAGAGTTTTTGAAGGTGCCGTACATTTGGATGAACTCAATATATT
GCTCAGTAGAAGCTGTCCGGAGTATTGGGCGGAGAGGCTTCAAACAGAGAGAGTTGAA
AGGGTTCGGCAGGCTAAGGAGCACTGGGAATCCCAGTGCCGAGCGTTCATGGGCGTGG
AATAAGGCGGTTGCCCTTTAAGGAGATTGATTTCTCTCTTGCGTGTCAGCTATTTATACA
GTCCTGTTTATAGTCAAAAGGATCCGGCAGGCGAAGAAGCACTGGGAATCCCAGTGCCG
AGCGTTTATGGGCGCGGAATAAGGCGGTTTCCCCTTTAACGAGATTGATTTCTGTCTTGC
ATGTCAGCTATTATACAGTCCTGTTTATAGTCCTGTGATGTAATAAAAAGCTGCATTGCTG
AAAACTTTCTCGCCTTGAACGCCCCTCATTGTTAAGTCTTTGCGTTCTTGGGCCTGTTGC
TCTATTCCTTTTTGAACATATAAACATGTACGTTTTCAATCAAAAAAAAA

SEQ ID NO: 258
AAGCAATGGTATCAACGCAGAGTACGCGGGAAAACCCCTCCATTTTCTTCCCTCCCCCT
CTTAAACCCTGGTTGCTTCCCGTGGTGCTCTCTCTCTCTCTCTCTAGGCAATTGGGCATG
GCGGCGGCGGCGGCGGCGTCCCTGCCGTTCAAGAAGAACTACAGGAGCTCCCAGGCG
CTGCAGCAGTTCTACGCCGGCGGTCCTTTCGCCGTCTCGTCCGACGGCTCCTTCATCGC
GTGCAACTGCGGCGACTCCATCAAGATCGTCGATTCCTCCAACGCCTCCCTCAGGCCCT

CCATCGACTGCGGCTCCGACACCATCACCGCCCTCTCCCTCAGCCCCGACGGCAAGTT
GCTCTTCTCCGCCGGCCACAGCCGCCAGATTAGGGTTTGGGACTTGTCCACCTCCACCT
GCCTGCGCTCCTGGAAGGGGCATGATGGTCCGGTGATGAGCATGGCCTGCCCCGTCTC
TGGGGGTTTGCTGGCGACGGGAGGAGCTGATAGGAAAGTCATGGTGTGGGACGTCGAC
GGTGGTTTCTGCACCCATTTCTTCAAAGGTCACGACGGCGTCGTTAGCACTGTCCTCTTC
CATCCCGACTCGAATCGCTCTCTTCTGTTCTCTGGAAGTGATGATGGAACTATACGGGTT
TGGGATCTCTTGGCAAAGAAGTGTGCTTCGACACTGAGAGGACATGATTCAACAGTCAC
TTCTCTGGCTTTTTCCGAGGATGGCTTGACATTGCTTGCCGCTGGAAGAGATAAGGTTGT
ATCTTTGTGGGACCTTCATAATTATGCCTGCAAGAAGACCATACCTATGTATGAGGTGCT
TGAATCTGTATGCGTGATACATAGTGGAACTGTTTTGGCTTCACAACTGGGGTTAGACGA
TCAGCTGAAAGTGACAAAAGAAAGTGCACAGAATATTCACTTTATTACTGTTGGTGAACG
CGGCATTTTACGGATATGGAAATCTGAAGGTTCAGTTTGCCTGTTTAAGCAAGAACATTC
TGATGTAACTGTCATCTCAGATGAGGACGACTCAAGGAGTGGCTTCACTGCTGCTGTCA
TGCTACCCTTGGATCAAGGATTGCTATGCGTGACAGCTGATCAACAGTTTTTATTCTATTA
TCCAGAGAAACATCCTGAAGGGATATTTTCGCTGACCCTATGTAGAAGACTCGTAGGCTA
CAATGAGGAAATAGTTGATATGAAGTTCCTAGGAGAGGAGGAAAATTTCCTTGCTGTTGC
TACTAATCTCGAACAGGTACGGGTTTATGAACTTGCTTCCATGTCATGTTCGTACGTCTT
GGCTGGTCATACTGAAACTGTCCTGTGCCTTGACACTTGTATTTCAAGTTCTGGAAGGAC
GCTTATTGTCACGGGAAGTAAGGACAACTCTGTTAGGTTGTGGGATTCAGAAAGCAGAC
ATTGCATTGGAGTTGGAGTAGGTCACATGGGCGCTGTAGGTGCAGTTGCTTTCTCAAGG
AAGAGACAAGATTTTTTTGTTAGTGGCAGCAGCGATCGTACACTTAAGGTTTGGAGCTTG
GATGGCATCTCAGAGGACGGTGTAGACTCAACAAATTTGAAAGCAAAAGCTGTTGTGGC
AGCTCATGATAAAGATATTAATTCTGTTGCTGTTGCACCAAATGACAGTTTAGTTTGTTCT
GGTTCTCAGGATCGCACAGCTTGTGTTTGGAGGCTTCCAGACCTGGTATCTGTAGTTGT
ACTTAAAGGGCACAAGAGGGGGATTTGGTCTGTAGAGTTTTCACCAGTTGATCAATGTGT
GCTCACTGCATCTGGCGATAAAACGGTGAAGATATGGGCCATATCTGATGGTTCTTGCTT
AAAGACCTTTGAAGGGCATGTCTCTAGTGTACTAAGAGCATCATTTCTCACCCGTGGGAC
ACAGTTTGTTTCTTGTGGTGCTGATGGCTTGGTAAAACTGTGGACTGTTAGGACAAATGA
ATGCATTGCTACATATGATCAACATAGCGATAAGGTATGGGCTTTGGCTGTTGGAAAGAA
GACCGAAATGCTTGCTACAGGTGGCAGTGATGCTGTTGTAAATCTCTGGTACGATTCAAC
TGCTTCTGACAAAGAGGATGCTTTTCGTAAAGAGGAAGAAGGTGTTTTGAAAGGTCAAGA
GTTAGAAAATGCAGTATCTGATGCTGACTACACCAAGGCAATCGAACTTGCTCTTGAACT
TCGGAGGCCTCATAAGTTATTTGAATTGTTCTCCGAACTTTGCAGGACGCGAGAAGTAG
GAGATCGTGTAGAGAGAATACTTTCTGCTCTCAGCGGCGAAGAGGTTTGTCTGCTTCTT
GAATATATCCGGGAGTGGAATGCGAAGCCAAAATTGTGCCACGTTGCCCAATCTGTGCT
ATCACAAGTCTTCAGAATTCTTTCTCCAACGGAGATAGTTGAGATTAAAGGCATCGGGGA
GCTCCTTGAAGGTCTCATTCCATATTCTCAGAGGCATTTCAGCAGGATAGACAGGCTTGT
TAGAAGCACGTATTTACTGGATTACACATTGACAGGAATGTCTGTTATTGAACCTGAAGC
AGACAGATCAGCAGTCAATGATGGGTCTCCAGATAAATCAGGCCTCGAGAAACTAGAGG
ATGGTCTTCTGGGAGAAAATGTTGGTGAGGAGAAAATCCAAAACAAGGAGGAGCTGGAA
AGTAGTGCATACAAGAAAAGAAAATTACCGAGATCTAAAGATCGTTCTAAGAAGAAATCA
AAGAATGTCGTGTACGCAGATGCAGCAGCAATCTCTTTCAGAGCTTGAATACCCGGGTC
CAGAGTTAAAAAGTGTTTATTATGACGAGCATTGCAGCGAGGACCATGTTGATGTCATTC
CCATCCTGCATTCTCAATGCCAG

SEQ ID NO: 259
AAGCAGTGGTATCAACGCAGAGTACGCGGGGACCAGGAATTCAATCATCCCGCTCCTGT
AGACCGTTGCAAATCTTCCTCTTCACGCCGCGACTGCTGGAAAAGCCCTCTCTCCGCGA
GCCAAGAAGAGGGCTTTTCTGTCCAGTCTTGCACGAGGTCGAAGATCAAGGATCACCCA
GATACGGAGAGACAGAAATAAGCAAAACCCGGAAAATGGATTCGGCTCCGAGGAGGAA
GAGCGGCGGCATCAATCTCCCGTCCGGGATGTCCGAGACCTCGCTTCGGCTTGACGGG
TTCTCGGGTTCATCATCGTCTTTCCGGGCCATCTCCAACCTGACGTCCCCGTCCAAGTC
CTCTTCCATCAGCGACAGATTCATCCCTTGCAGATCCTCTTCGAGGCTCCACACTTTCGG
GCTCGTCGAAAGGGGGTCGCCGGTTAAGGAAGGAGGGAACGAGGCGTACTCGAGGTT
GTTGAAGGCTGAGCTCTTTGGGTCTGATTTCGGTTCTCTCTCTCCTGCAGGTCAAGGGT
CGCCCATGAGCCCCAGCAAGAACATGTTGCGGTTCAAGACCGAGAGTTCGGGCCCGAA
CTCGCCGTTTTCGCCGTCGATCCTGAGGCAGGACAGCGGTTTCTCCAGTGAAGCTTCGA
CGCCTCCTAAGCCACCAAGGAAGGTGCCCAAGACGCCGCATAAGGTTCTGGATGCCCC
ATCGCTGCAAGATGACTTCTATTTGAATCTGGTTGACTGGTCATCTCAGAATACGCTTGC
GGTCGGGTTGGGCACTTGTGTTTACCTATGGAGTGCATCAAACAGCAAAGTGACAAAGC

TCTGTGATCTGGGACCTAATGATGGTGTCTGTGCAGTACAGTGGACTCGAGAGGGTTCA
TACATTTCAATTGGTACCAGCCTCGGTCAAGTTCAGATATGGGATGGAACTCAGTGTAAG
CGAGTTCGGACAATGGGTGGTCATCAGACTAGAACAGGGGTCTTGGCATGGAATTCGC
GCATATTAGCTTCTGGTAGCAGGGACCGGGTCATACTTCAGCATGATCTTCGCGTTCCA
AATGAATTCATAGGCAAGCTTGTTGGCCACAAATCTGAGGTATGTGGATTGAAATGGTCC
CATGATGACAGAGAACTCGCATCAGGTGGGAACGACAATCAGGTAATTCTCGTTTTCTG
GTAAGTCGTTTTATCTTCCATATTAAACTGAGTATGGAATATTAAACAAGCTTATTATTGGT
TGATCAGCTACTGGTGTGGAATCAGCACTCTCAGCAACCAGTTCTGAAGCTTACTGAGC
ATACAGCAGCAGTGAAAGCCATAGCATGGTCTCCCCATCAGAATGGGCTACTTGCATCA
GGAGGAGGAACTGCTGATCGATGCATACGCTTCTGGAACACCACCAATGGCCATCAGAC
GAGCAGTGTCGACACTGGGAGTCAGGTATGTAATCTGGCTTGGAGTAAGAATGTCAATG
AGTTGGTGAGCACTCATGGGTATTCACAGAATCAAATAATGGTGTGGAAATATCCATCCA
TGGCAAAGGTTGCAACTCTAACTGGCCATAGTCTTCGAGTACTTTATCTTGCCATGTCTC
CAGATGGCCAGACAATAGTGACAGGAGCTGGTGACGAGACCTTGCGGTTTTGGAATGTA
TTCCCATCGGCTAAGGCACCGGCACCGGTCAAGGATACGGGTCTCTGGTCTTTGGGGC
GAACCCACATCAGATAAAAGCTGCATGCTTGTTCAAGCTAGCTCGGTCTCTCTTGAAAAG
AAACATGTTTGATTTCATTACATGTAAAAAAATCTTGCACGGTCAACTGACATTACACTTTT
TGTAAATTGTGAGTAGAATAGGAGAAACTTTTGTACAAGATTAATACGTGTGGCATAATAA
GATTTGTAAATGATCTTTGATCAAAAAAAAAAA

SEQ ID NO: 260
CTGCCATGGCGGTACCTCGTTGAAACGCCGCAGGACGGAGAAGAAGAAGCTGCAAGAC
AACTCTCTCTCTCTCTCTCTACCTTGAGTGAGCTCCGGCTATGGAGGACGAGGCGGA
GATATACGACGGCGTGAGAGCCCAGTTCCCCCTCACCTTCGGCAAGCAGTCCAAGCCC
CAGACTTCCCTCGAATCCGTCCACAGCGCCACCCGCCGCGGCGGCCCCGCCCCCGCC
CCCGCCCCCGCCTCCTCCTCCTCTCTTCCCTCTACTACCTCCCCCTCCGCTGCCGGCGG
CGCCGGCAAGAGCAGCGGCCTCCCTTCTCTCTCCTCCTCCTCCACAGCCTGGCTCGAG
GGGCTCCGGGCCGGAAACCCTAGGGCCGGCCGCGAGGCCGGAATCGGCTCCCGCGG
TGGCGACGGTGAGGACGGCGGTCGTGCGATGATAGGTCCACCTCGCCCGCCGCCGGG
ATTTAGCGCCAATGATGATGGCGGGGGCGAGGACGACGACGACGACGGCGACGGGGT
TATGGTTGGCCCGCCGCCGCCTCCTCCTGGGAACCTTGGAGATGGTGATGACGATGAG
GAGGAGGAAGAGGCGATGATTGGGCCGCCGAGGCCGCCGGTTGTGGATTCCGACGAG
GAGGAAGAAGAGGAAGAAGAGGAGAATCGGTATCGGTTACCTCTCAGCAATGAGATCGT
GCTCAAAGGCCATAATAAGATCGTATCTGCTCTTGCGGTTGATCCCACTGGTTCTCGGGT
TCTATCCGGCAGTTATGATTACACTGTACGGATGTTCGACTTTCAAGGAATGAATTCTCG
CCTATCATCATTTAGAGATTTTGAGCCAGTTGAAGGTCATCAAGTCCGTAACTTAAGCTG
GAGTCCGACGGCAGACCGATTTCTGTGCGTAACTGGCTCTGCTCAAGCTAAGATCTATG
ATCGTGATGGGCTTACATTAGGTGAATTTGTGAAGGGAGACATGTATATCCGTGATTTGA
AAAATACTAAGGGCCATATAACTGGATTGACTTGGGGAGAGTGGCACCCTAAAACAAAG
GAGACGATTTTAACATCATCTGAGGATGGATCTCTTCGCATATGGGATGTGAATGACTTC
AAAAGTCAAAAGCAGGTTATTAAACCAAAGCTTGCAAGGCCTGGAAGAGTACCTGTGAC
AACATGCACTTGGGATCGTGAGGGCAAATGCATTGCAGGTGGTATTGGAGATGGTTCAA
TACAGATCTGGAACCTTAAACCCGGATGGGGAAGCAGGCCGGACATTCATGTCGAGCAA
GCTCATGCAGATGATATAACTGGGCTTAAGTTTTCCAGTGATGGGAAAATTTTACTGACA
AGAAGCTTTGATGATTCACTTAAGGTGTGGGATTTGCGCCTTATGAAAAATCCCCTCAAA
GTTTTTGAAGATCTTCCAAATCACTATGCTCAGACAAATATTGCATGTAGTCCTGACGAAC
AGCTATTCTTGACTGGAACTTCTGTTGAAAGGGAGTCTACAATTGGAGGCTTGTTGTGTT
TCTTTGATCGATCTAAACTTGAGCTTGTATCAAGAATTGGGATTTCTCCTACTTGTAGTGT
CGTGCAGTGTGCCTGGCACCCCAGGCTGAACCAGATCTTTGCAACTTCTGGGGATAAAA
GCCAAGGAGGAACTCACGTTCTTTATGATCCTACTCTCAGCGAGAGAGGAGCTCTTGTTT
GTGTTGCTCGTGCACCAAGGAAAAAATCTGTGGATGACTTTGAGTTAAAGCCTGTCATTC
ACAACCCCCATGCATTGCCTTTGTTCAGGGATCAGCCAAGCCGTAAGCGTCAACGTGAA
AAGATCCTCAAGGATCCATTAAAATCCCACAAACCAGAGCTTCCCATGAATGGACCTGGA
CATGGTGGAAGAGTTGGTGCAAGCAAGGGAAGTTTATTGACCCAGTATCTTCTTAAGCAA
GGAGGTATGATTAAGGAGACGTGGATGGATGAAGATCCCAGAGAAGCTATACTGAAGCA
TGCTGACGCTGCAGAAAAAAATCCAAAGTTCACCCGTGCATATGCTGAAACCCAGCCAG
ACCCTGTTTTTGCAAAATCTGATTCTGAAGATGAAGATAAATGATGTCCTATCCTGCAGTG
CTTCTTTTGGATCCTTGCTGCTGTCATTGGACATGTGAGGAAGAAAACTGGGATCGGACT
ACTGATACAGCAAGCAAATCATCTGTTGACCAGCTGGAGTGTCAACAAAGAATGGGTTTT
CTTTACTCAATTAATTTGACTGGTCTGCTGGGTTCAATTAATCGAGACATCAACTCCGAGT

TCAAAATCAGGAATTCTCTAGGAAAGTATCTCTTGCATTTTGCTGGCCTTCTTCATCATAT
GATGTGCAGTTTACATTATTATGGTTCGAGTATTATTTAGCTGCCCTATCTTAAGTCATGG
AGGCTTTACTTTGCTGAGAAATGTACCTTTAGCGATGTGATGCAATAGCAGTCTTTCATAT
TTATGTGGTGTAAGAGATGGAATCAATTGTTTCGCTTCCTCATCAGATCAATTAAGTTCAT
GCAACTAAGCTTTTCACTGTAAAAAAAAAA

SEQ ID NO: 261
MGDGSLGSGGRGNSGGGGGGGSRPEWLQQYDLIGKIGEGTYGLVFLARIKHPSTNRGKYI
AIKKFKQSKDGDGVSPTAIREIMLLREISHENVVKLVNVHINPVDMSLYLAFDYADHDLYEIIRH
HRDKVNQAINPYTVKSLLWQLLNGLNYLHSNWIIHRDLKPSNILVMGEGEEQGVVKIADFGLA
RVYQAPLKPLSDNGVVVTIWYRAPELLLGAKHYTSAVDMWAVGCIFAELLTLKPLFQGQEVK
ANPNPFQLDQLDKIFKVLGHPTQEKWPMLVNLPHWQSDVQHIQRHKYDDNALGNVVRLSSK
NATFDLLSKMLEYDPQKRITAAQALEHEYFRMEPLPGRNALVPSSPGDKVNYPTRPVDTTTD
IEGTTSLQPSQSASSGNAVPGNMPGPHVVTNRPMPRPMHMVGMQRVPASGMAGYNLNPS
GMGGGMNPSGIPMQRGVANQAQQSRRKDPGMGMGGYPPQQKQRRF

SEQ ID NO: 262
MEKYQQLAKIGEGTYGIVYKAKDKKSGELLALKKIRLEAEDEGIPSTAIREISLLKQLQHPNIVR
LYDVVHTEKKLTLVFEFLDQDLKKYLDACGDNGLEPYTVKSFLYQLLQGIAFCHEHRVLHRDL
KPQNLLINMEGELKLADFGLARAFGIPVRNYTHEVVTLWYRAPDVLMGSRKYSTQVDIWSVG
CIFAEMVNGRPLFPGSSEQDQLLRIFKTLGTPSLKTWPGMAELPDFKDNFPKYVVQSFKKIC
PKKLDKTGLDLLSRMLQYDPAKRISAEQAMGHPYFKDLKLRKPKAAGPGP

SEQ ID NO: 263
MDQYEKIEKIGEGTYGVVYKAIDRSTNKTIALKKIRLEQEDEGVPSTAIREISLLKEMQHGNIVK
LQDVVHSERRLYLVFEYLDLDLKKHMDSCPEFSKDTHTIKMFLYQILRGISYCHSHRVLHRDL
KPQNLLLDRRTNSLKLADFGLARAFGIPVRTFTHEVVTLWYRAPEILLGSRHYSTPVDVWSV
GCIFAEMVNRRPLFPGDSEIDELFKIFRIMGTPNEDSWPGVTSLPDFKSTFPKWASQDLKTVT
PTVDPAGIDLLSKMLCMDPRRRITAKVALEHEYFKDVGVIP

SEQ ID NO: 264
MVMKSKLDKYEKLEKLGEGTYGVVYKAQDKTTKEIYALKKIRLESEDEGIPSTAIREIALLKEL
QHPNVVRIHDVIHTNKKLILVFEFVDYDLKKFLHNFDKGIDPKIVKSLLYQLVRGVAHCHQQKV
LHRDLKPQNLLVSQEGILKLGDFGLARAFGIPVKNYTNEVVTLWYRAPDILLGSKNYSTSVDI
WSIGCIFVEMLNQKPLFPGSSEQDQLKKIFKIMGTPDATKWPGIAELPDWKPENFEKYPGEP
LNKVCPKMDPDGLDLLDKMLKCNPSERIAAKNAMSHPYFKDIPDNLKKLYN

SEQ ID NO: 265
MDQYEKVEKIGEGTYGVVYKAIDRLTNETIALKKIRLEQEDEGVPSTAIREISLLKEMQHGNIV
RLQDVVHSENRLYLVFEYLDLDLKKHMDSSPDFAKDPRLVKIFLYQILRGIAYCHSHRVLHRD
LKPQNLLIDRRTNALKLADFGLARAFGIPVRTFTHEVVTLWYRAPEILLGSRHYSTPVDVWSV
GCIFAEMVNQRPLFPGDSEIDELFKIFRILGTPNEDTWPGVTALPDFKSAFPKWPAKNLQDM
VPGLNSAGIDLLSKMLCLDPSKRITARSALEHEYFKDIGFVP

SEQ ID NO: 266
MEKYEKLEKVGEGTYGKVYKAKDKATGQLVALKKTRLEMDEEGVPPTALREVSLLQLLSQSL
YVVRLLSVEHVDGGSKRKPMLYLVFEYLDTDLKKFIDSHRKGPNPRPVPAATVQNFLYQLLK
GVAHCHSHGVLHRDLKPQNLLVDKEKGILKIADLGLGRAFTVPLKSYTHEVVTLWYRAPEVLL
GSAHYSIGVDMWSVGCIFAEMVRRQALFPGDSEFQQLLHIFRLLGTPTEKQWPGVTTLRDW
HVYPQWEPQNLARAVPSLGPDGVDLLSKMLKYDPAERISAKAALDHPFFDSLDKSQF

SEQ ID NO: 267
MERPATAAVSAMEAFEKLEKVGEGTYGKVYRAREKATGKIVALKKTRLHEDEEGVPPTTLRE
ISILRMLSRDPHIVRLMDVKQGQNKEGKTVLYLVFEYMETDLKKYIRGFRSSGESIPVNIVKSL
MYQLCKGVAFCHGHGVLHRDLKPHNLLMDKKTLTLKIADLGLARAFTVPIKKYTHEILTLWYR
APEVLLGATHYSTAVDMWSVGCIFAELVTKQALFPGDSELQQLLHIFRLLGTPNEKMWPGVS
SLMNWHEYPQWKPQSLSTAVPNLDKDGLDLLSQMLHYEPSRRISAKAAMEHPYFDDVNKT
CL

SEQ ID NO: 268

MGCVLGREVSSGIVTESKGRDSSEVETSKRDDSVAAKVEGEGKAEEVRTEETQKKEKVEDD
QQSREQRRRSKPSTKLGNLPKHIRGEQVAAGWPSWLSDICGEALNGWIPRRANTFEKIDKIG
QGTYSNVYKAKDLLTGKIVALKKVRFDNLEPESVRFMAREILILRHLDHPNVVKLEGLVTSRM
SCSLYLVFEYMEHDLAGLAASPAIKFTEPQVKCYMHQLLSGLEHCHNRRVLHRDIKGSNLLID
NGGVLKIGDFGLASFYDPDHKHRMTSRVVTLWYRPPELLLGANDYGVGIDLWSAGCILAELL
AGKPIMPGRTEVEQLHKIYKLCGSPSEEYWKKYKLPNATLFKPREPYRRCIRETFKDFPPSSL
PLIETLLAIDPAERGTATDALQSEFFRTEPYACEPSSLPQYPPSKEMDAKKRDDEARRLRAAS
KGQADGSKKERTRDRRVRAVPAPEANAELQHNIDRRRLISHANAKSKSEKFPPPHQDGALG
FPLGASHRFDPAVVPPDVPFTSTSFTSSKEHDQTWSGPLVDPPGAPRRKKHSAGGQRESS
KLSMGTNKGRRADSHLKAYESKSIA

SEQ ID NO: 269
MYSKSSAVDDSRESPKDRVSSSRRLSEVKTSRLDSSRRENGFRARDKVGDVSVMLIDKKVN
GSARFCDDQIEKKSDRLQKQRRERAEAAAAADHPGAGRVPKAVEGEQVAAGWPVWLSAV
AGEAIKGWLPRRADTFEKLDKIGQGTYSSVYKARDVTNNKIVALKRVRFDNLDTESVKFMAR
EIHILRMLDHPNVIKLEGLITSRMSCSLYLVFEYMEHDLTGLASRPDVKFSEPQIKCYMKQLLS
GLDHCHKHGVLHRDIKGSNLLIDNNGILKIADFGLASVFDPHQTAPLTSRVVTLWYRPPELLL
GASRYGVEVDLWSTGCILGELYTGKPILPGKTEVEQLHKIFKLCGSPSDDYWRRLHLPHAAV
FKPPQPYRRCVAEIFKELPPVALGLLETLISVDPSQRGTAAFALRSEFFTASPLPCDPSSLPKY
PPSKEIDMKLREEEARRRGAAGGKNELEKRGTKDSRTNSAYYPNAGQLQVKQCHSNANGR
SEIFGPYQEKTVSGFLVAPPKQARVSKETRKDYAEQPDRASFSGPLVPGPGFSKAGKELGH
SITVSRNTNLSTLSSLVTSRTGDNKQKSGPLVSESANQASRYSGPIREMEPARKQDRRSHVR
TNIDYRSREDGNSSTKEPALYGRGSAGNKIYVSGPLLVSSNNVDQMLKEHDRRIQEHARRA
RFDKARVGNNHPQAAVDSKLVSVHDAG

SEQ ID NO: 270
MGCIPTIISDGRRRSAAPDKRRPRPRRSSSEGEAPPHATAAGSEGGESARGAPGKERPEPA
PRFVVRSPQGWPPWLVAAVGHAIGEFVPRCADSFRKLAKIGEGTYSNVYKARDLVTGKTVA
LKKVRFDNLEAESIKFMAREILVLTRLNHPNVIKLEGPVTSRMSSGLYLAFEYMEHDLSGIAAR
QNGKFTEPQVKCFMRQLLSGLEHCHNHDVLHRDIKCSNLLIDNEGNLKIADFGLATFYDPER
KQVMTNRVVTLWYRAPELLLGATSYGIGIDLWSAGCILAELLYGKPIMPGRTEVEQLHKIFKL
CGSPSEAYWNKFKLPNANIFKPPQPYARCIAETFKDFPPSALPLLETLLSIDPDERGTATTALN
SEFFAAEPHACEPSSLPKYPPSKEMDLKLIKEKTRRDSSKRPSAIHGSRRDGIHDRAGRVIPA
PEATAENQATLHRPRAMKKANPMSRSEKFPPAHMDGVVGSSANAWLSGPASNAAPDSRR
HRSLNQNPSSSVGKASTGSSTTQETLKVAPELLQVGSSSLHPCHRMLVYGSNLTIRSK

SEQ ID NO: 271
MGCICAKQADRGPASPGSGILTGAGTGTGTRSSKIPSGLFEFEKSGVKEHGGRSGELRKLE
EKGSLSKRLRLELGFSHRYVEAEQAAAGWPSWLTAVAGDAIQGLVPLKADSFEKLEKIGQGT
YSSVFRARELANGRMVALKKVRFDNFQPESIQFMAREISILRRLDHPNIMKLEGIITSRMSNSI
YLVFEYMEHDLYGLISSPQVKFSDAQVKCYMKQLLSGIEHCHQHGVIHRDVKSSNILVNNEGI
LRIGDFGLANILNPKDRQQLTSHVVTLWYRPPELLMGSTSYGVTVDLWSVGCVFAELMFRKP
ILRGRTEVEQLHKIFKLCGSPPDGYWKMCKVPQATMFRPRHAYECTLRERCKGIATSAMKL
METFLSIEPHKRGTASSALISEYFRTVPYACDPSSLPKYPPNKEIDAKHREEARRKKARSRVR
EAEVGKRPTRIHRASQEQGFSSNIAPKEKRSYA

SEQ ID NO: 272
MAVAAPGHLNVNESPSWGSRSVDCFEKLEQIGEGTYGQVYMAKEKKTGEIVALKKIRMDNE
REGFPITAIREIKILKKLHHENVIKLKEIVTSPGPEKDEQGRPEGNKYKGGIYMVFEYMDHDLT
GLADRPGMRFSVPQIKCYMRQLLTGLHYCHINQVLHRDIKGSNLLIDNEGNLKLADFGLARSF
SNDHNANLTNRVITLWYRPPELLLGATKYGPAVDMWSVGCIFAELLHGKPIFPGKDEPEQLN
KIFELCGAPDEINWPGVSKIPWYNNFKPTRPMKRRLREVFRHFDRHALELLERMLTLDPSQRI
SAKDALDAEYFWADPLPCDPKSLPKYESSHEFQTKKKRQQQRQHEETAKRQKLQHPPQHP
RLPPVQQSGQAHAQMRPGPNQLMHGSQPPVATGPPGHHYGKPRGPSGGAGRYPSSGNP
GGGYNHPSRGGQGGSGGYNSGPYPPQGRAPPYGSSGMPGAGPRGGGGNNYGVGPSNY
PQGGGGPYGGSGAGRGSNMMGGNRNQQYGWQQ

SEQ ID NO: 273
MGCICTKGILPAHYRIKDGGLKLSKSSKRSVGSLRRDELAVSANGGGNDAADRLISSPHEVE
NEVEDRKNVDFNEKLSKSLQRRATMDVASGGHTQAQLKVGKVGGFPLGERGAQVVAGWP

SWLTAVAGEAINGWVPRRADSFEKLEKIGQGTYSSVYRARDLETNTIVALKKVRFANMDPES
VRFMAREIIIMRKLDHPNVMKLEGLITSRVSGSLYLVFEYMDHDLAGLAATPSIKLTESQIKCY
MQQLLRGLEYCHSHGVLHRDIKGSNLLVDNNGNLKIGDFGLATFFRTNQKQPLTSRVVTLWY
RPPELLLGSSDYGASVDLWSSGCILAELFAGKPIMPGRTEVEQLHKIFKLCGSPSEEYWKKS
KLPHATIFKPQQPYKRCLLETFKDFPSSALGLLDVLLAVEPECRGTASSALQNEFFTSNPLPS
DPSSLPKYPSSKEFDARLRDEEARKHKATAGKARGLESIRKGSKESKVVPTSNANADLKASI
QKRQEQSNPRSTGEKPGGTTQNNFILSGQSAKPSLNGSTQIGNANEVEALIVPDRELDSPR
GGAELRRQRSFMQRRASQLSRFSNSVAVGGDSHLDCSREKGANTQWRDEGFVARCSHPD
GGELAGKHDWSHHLLHRPISLFKKGGEHSRRDSIASYSPKKGRIHYSGPLLPSGDNLDEMLK
EHERQIQNAVRKARLDKVKTKREYADHGQTESLLCWANGR

SEQ ID NO: 274
MDPDPSPDPDPPKSWSIHTRREIIARYEILERVGSGAYSDVYRGRRLSDGLAVALKEVHDYQ
SAFREIEALQILRGSPHVVLLHEYFWREDEDAVLVLEFLRSDLAAVIADASRRPRDGGGGGA
AALRAGEVKRWMLQVLEGVDACHRNSIVHRDLKPGNLLISEEGVLKIADFGQARILLDDGNV
APDYEPESFEERSSEQADILQQPETMEADTTCPEGQEQGAITREAYLREVDEFKAKNPRHEI
DKETSIFDGDTSCLATCTTSDIGEDPFKGSYVYGAEEAGEDAQGCLTSCVGTRWFRAPELLY
GSTDYGLEVDLWSLGCIFAELLTLEPLFPGISDIDQLSRIFNVLGNLSEEVWPGCTKLPDYRTI
SFCKIENPIGLESCLPNCSSDEVSLVRRLLCYDPAARATPMELLQDKYFTEEPLPVPISALQVP
QSKNSHDEDSAGGWYDYNDMDSDSDFEDFGPLKFTPTSTGFSIQFP

SEQ ID NO: 275
MDPDPSPSPDPPKSWSIHTRREIIARYEILERVGSGAYSDVYRGRRLSDGLAVALKEVHDYQ
SAFREIEALQILRGSPHVVLLHEYFWREDEDAVLVLEFLRSDLAAVIADASRRPRGGGVAPLR
AGEGKRWMLQVLEGVDACHRNSIVHRDLKPGNLLISEEGVLKIADFGQARILLDDGNVAPDY
EPESFEERSSEQADILQQPETMEADTTCPEGQEQGAITREAYLREVDEFKAKNPRHEIDKET
SIYDGDTSCLATCTTSDIGEDPFKGSYVYGAEEAGEDAQGSLTSCVGTRWFRAPELLYGSTD
YGLEVDLWSLGCIFAELLTLEPLFPGISDIDQLSRIFNVLGNLSEEVWPGCTKLPDYRTISFCKI
ENPIGLESCLPNCSSDEVSLVRRLLCYDPAARATPMELLQDKYFTEEPLPVPISALQVPQSKN
SHDEDSAGGWYDYNDMDSDSDFEDFGPLKFTPTSTGFSIQFP

SEQ ID NO: 276
MSNQHRRSSFSSSTTSSLAKRHASSSSSSLENAGKAFAAAAVPSHLAKKRAPLGNLTNLKA
GDGNSRSSSAPSTLVANATKLAKTRKGSSTSSSIMGLSGSALPRYASTKPSGVLPSVNPSIP
RIEIAVDPMSCSMVVSPSRSDMQSVSLDESMSTCESFKSPDVEYIDNEDVSAVDSIDRKTFS
NLYISDAAAKTAVNICERDVLMEMETDEKIVNVDDNYSDPQLCATIACDIYQHLRASEAKKRP
STDFMDRVQKDITASMRAILIDWLVEVAEEYRLVPDTLYLTVNYIDRYLSGNVMNRQRLQLLG
VACMMIAAKYEEICAPQVEEFCYITDNTYFKEEVLQMESSVLNYLKFEMTAPTVKCFLRRFVR
AAQGVNEVPSLQLECMANYIAELSLLEYDMLCYAPSLVAASAIFLAKFVITPSKRPWDPTLQH
YTLYQPSDLGNCVKDLHRLCFNNHGSTLPAIREKYSQHKYKYVAKKYCPPSIPPEFFHNLVY

SEQ ID NO: 277
MNKENAVGTKSEAPTIRITRSRSKALGTSTGMLPSSRPSFKQEQKRTVRANAKRSASDENK
GTMVGNASKQHKKRTVLNDVTNIFCENSYSNCLNAAKAQTSRQGRKWSMKKDRDVHQSG
AVQIMQEDVQAQFVEESSKIKVAESMEITIPDKWAKRENSEHSISMKDTVAESSRKPQEFICG
EKSAALVQPSIVDIDSKLEDPQACTPYALDIYNYKRSTELERRPSTIYMETLQKDVTPNMRGIL
VDWLVEVSEEYKLVPDTLYLTVNLIDRSLSQKFIEKQRLQLLGVTCMLIASKYEEICPPRVEEF
CFITDNTYTSLEVLKMESRVLNLLHFQLSVPTVKTFLRRFVQAAQVSSEVPSVELEYLANYLA
ELTLVEYSFLKFLPSLMAASAVLLARWTLNQSDNPWNLTLEHYTKYKASELKAAVLALEDLQL
NTSGSTLNAIREKYRQQKVNYSLLIHSKANHEIL

SEQ ID NO: 278
MAGSDENNPGVVGGAHVQEGLRVGAGKMGAGNVQQRRALSNINSNIIGAPPYPCAVNKRV
LSEKNVNSENDLLNAAHRPITRQFAAQMAYKQQLRPEENKRTTQSVSNPSKSEDCAILDVDD
DKMADDFPVPMFVQHTEAMLEEIDRMEEVEMEDVAEEPVTDIDSGDKENQLAVVEYIDDLY
MFYQKAEASSCVPPNYMDRQQDINERMRGILIDWLIEVHYKFELMDETLYLTVNLIDRFLAVQ
PVVKKKLQLVGVTAMLLACKYEEVSVPVVEDLILISDRAYSRKEVLEMERLMVNTLHFNMSVP
TPYVFMRRFLKAAQSDKKLELLSFFIIELSLVEYDMLKFPPSLLAASAIYTALSTITRTKQWSTT
CEWHTSYSEEQLLECARLMVTFHQRAGSGKLTGVHRKYSTSKFGHAARTEPANFLLDFRL

SEQ ID NO: 279
MASRPIVPVQARGEAAIGGGAGKAAIGGGAGKQQKKNGAAEGRNRKALGDIGNLVTVRGIE
GKVQPHRPITRSFCAQLLANAQAAAAAENNKKQAVVNVNGAPSILDVPGAGKRAEPAAAAA
AAVAKAAQKKVVKPKQKAEVIDLTSDSERAIEAKKKQQHHEPTKKEGEKSSRRNMPTLTSVL
TARSKAACGMTKKPKEKVVDIDAGDAHNELAAFEYIEDIYTYYKEAENESLPRNYMSSQPEIN
EKMRAILVDWLIEIHNKFDLMPETLYLTINIIDRFLSVKAVPRRELQLLGMGALFTASKYEEIWA
PEVNDLVCIADRAYSHEQVLAMEKTILGKLEWTLTVPTHYVFLVRFIKASLGDRKLENMVYFL
AELGVMNYATLTYCPSMVAASAVYAARCTLGLTPLWNDTLKLHTGFSESQLMDCARLLVGY
HAKAKENKLQVVYKKYSSSQREGVALIPPAKALLCEGGGLSSSSSLASSS

SEQ ID NO: 280
MGLPDENNAALSKPTNLQVGGLEIGGRKFGQEIRQTRRALSVINQNLVGDRAYPCHVVNKR
GHSKRDAVCGKDQVDPVHRPLTRKFAAQTASTQQHCIEEAKKPRTAVQERNEFGDCIFVDV
EDCQPSSENQPVPMFLEIPESRLDDDMEEVEMEDIVEEEEEEPIMDIDGRDKKNPLAVVDYIE
DIYANYRRTENCSCVSANYMAQQADINEKMRSILIDWLIEVHDKFDLMHETLFLTVNLIDRFLA
RQSVVRKKLQLVGLVAMLLACKYEEVSVPVVGDLILISDKAYTRKEVLEMESLMLNSLQFNM
SVPTPYVFMRRFLKAAESDKKLEVLSFFLIELSLVEYEMVKFPPSLLAAAAIFTAQCTLYGFKQ
WTKTCEWHSNYTEDQLLECARMMVGFHQKAATGKLTGVHRKYGTSKFGYTSKCEPANFLL
GEMKNP

SEQ ID NO: 281
MGLPDENNAALSKPTNLQVGGLEIGGRKFGQEIRQTRRALSVINQNLVGDRAYPCHVVNKR
GHSKRDAVCGKDQVDPVHRPLTRKFAAQTASTQQHCIEEAKKPRTAVQERNEFGDCIFVDV
EDCQPSSENQPVPMFLEIPESRLDDDMEEVEMEDIVEEEEEEPIMDIDGRDKKNPLAVVDYIE
DIYANYRRTENCSCVSANYMAQQADINEKMRSILIDWLIEVHDKFDLMHETLFLTVNLIDRFLA
RQSVVRKKLQLVGLVAMLLACKYEEVSVPVVGDLILISDKAYTRKEVLEMEKLMLNSLQFNM
SVPTPYVFMRRFLKAAESDKKLEVLSFFLIELSLVEYEMVKFPPSLLAAAAIFTAQCTLYGFKQ
WTKTCEWHSNYTEDQLLECARMMVGFHQKAATGKLTGVHRKYGTSKFGYTSKCEAANFLL
GEMKNP

SEQ ID NO: 282
MAMVQRQGHDPSSPQEQEDGPSSFLSDDALYCEEGRFEEDDGGGGGQVDGIPLFPSQPA
DRQQDSPWADEDGEEKEEEEAELQSLFSKERGARPELAKDDGGAVAARREAVEWMLMVR
GVYGFSALTAVLAVDYLDRFLAGFRLQRDNRPWMTQLVAVACLALAAKVEETDVPLLVELQE
VGDARYVFEAKTVQRMELLVLSTLGWEMHPVTPLSFVHHVARRLGASPHHGEFTHWAFLR
RCERLLVAAVSDARSLKHLPSVLAAAAMLRVIEEVEPFRSSEYKAQLLSALHMSQEMVEDCC
RFILGIAETAGDAVTSSLDSFLKRKRRCGHLSPRSPSGVIDASFSCDDESNDSWATDPPSDP
DDNDDLNPLPKKSRSSSPSSSPSSVPDKVLDLPFMNRIFEGIVNGSPI

SEQ ID NO: 283
MEASYQPHHHGHLRQHDPSSSQQEEQVPFDALYCSEEHWGEEDEEEGLASDGLLSEERD
HRLLSPRALLDQDLLWEDEELASLFSKEEPGGMRLNLENDPSLADARREAVEWIMRVHAHY
AFSALTALLAVNYWDRFTCSFALQEDKPWMTQLSAVACLSLAAKVEETQVPLLIDFQVEDSS
PVFEAKNIQRMELLVLSSLEWKMNPVTPLSFLDYMTRRLGLTGHLCWEFLRRCENVLLSVIS
DCRFTCYLPSVIAASTMLHVINGLKPRLDVEDQTQLLGILAMGMDKIDACYKLIDDDHALRSQ
RYSHNKRKFGSVPGSPRGVMELCFSSDGSNDSWSVAASVSSSPEPHSKKSRAGEEAEDRL
LRGLEGEEDDPASADIFSFPH

SEQ ID NO: 284
MALQEEDTRRHYPTAPPFSPDGLYCEDETFGEDLADNACEYAGGGARDGLCEIKDPTLPPS
LLGQDLFWEDGELASLVSRETGTHPCWDELISDGSVALARKDAVGWILRVHGHYGFRPLTA
MLAVNYLDRFFLSRSYQRDRPWISQLVAVACLSVAAKVEETQVPILLDLQVANAKFVFESRTI
QRMELLLMSTLDWRMNSVTPISFFDHILRRFGLTTNLHRQFFWMCERLLLSVVADVRLASFL
PSVVATAAMLYVNKEIEPCICSEFLDQLLSLLKINEDRVNECYELILELSIDHPEILNYKHKRKR
GSVPSSPSGVIDTSFSCDSSNDSWGVASSVSSSLEPRFKRSRFQDQQMGLPSVNVSSMGV
LNSSY

SEQ ID NO: 285
MGQIQYSEKYFDDTYEYRHVVLPPDVAKLLPKNRLLSENEWRAIGVQQSRGWVHYAIHRPE
PHIMLFRRPLNYQQQQENQAQQNMLAK

SEQ ID NO: 286
MGSIDPPKAEQNGTAAAAVADPGQKPGAGDAMPPPPPVKHSNGTAAEPDVATKRRRMSVL
PLEVGTRVMCRWRDGKYHPVKVIERRKLNPGDPNDYEYYVHYTEFNRRLDEWVKLEQLDL
NSVETVVDEKVEDKVTGLKMTRHQKRKIDETHVEGHEELDAASLREHEEFTKVKNIATIELGR
YEIETWYFSPFPPEYNDCSKLYFCEFCLNFMKRKEQLQRHMKKCDLKHPPGDEIYRSGTLS
MFEVDGKKNKVYGQNLCYLAKLFLDHKTLYYDVDLFLFYVLCECDDRGCHMVGYFSKEKHS
EESYNLACILTLPPYQRKGYGKFLIAFSYELSKKEGKVGTPERPLSDLGLLSYKGYWTRVLLDI
LKKHKANISIKELSDMTAIKADDILNTLQSLDLIQYRKGQHVICADPKVLDRHLKAAGRGGLEV
DVSKLIWTPYREQG

SEQ ID NO: 287
MAQKHSTAPDPAAEPKKRRRVGFSGIDAGVDPNGCFKVYLVSREEEVGAPDSFCLDPVDLS
HFFEEEDGKIYGYEGLKISVWVSCVSFHSYAEIAFESKSDGGKGITDLNTALKNMFGETLVDN
KDDFLQTFSKETQFIRSTVSAGEILKHKHSDDHVNDSVSNLKVGSDVEAVRMLMGDMTAGH
LYSRLVPLVLLLVDGSSPIDVTDSSWELYLLIQKTSDQQGNFHDRLLGFAAVYRFYHYPDSSR
LRLGQILVLPLYQRKGYGRYLLEVLNNVAIADDVYDFTIEEPVDNLQHLRTCIDVQRLLSFDKV
QQAVNSTVSQLKQGKLSKKTYIPRLLPPPSVVEDARKRFKINKKQFLQCWEILVYLGLDPADK
SIQDYFSVISNRVRADILGKDSETAGKKVIEVPSDFDPEMSFVMHRAKAGGEANGIQVEDNQ
NKQEEQLQQLIDERLKDIKLIAEKVTQK

SEQ ID NO: 288
MAQKHSTAPDPAAEPKKRRRVGFSGIDAGVDPNGCFKVYLVSREEEVGAPDSFCLDPVDLS
HFFEEEDGKIYGYEGLKISVWVSCVSFHSYAEIAFESKSDGGKGITDLNTALKNMFGETLVDN
KDDFLQTFSKETQFIRSTVSAGEILKHKHSDGHVNDSVSNLKVGSDVEAVRMLMGDMTAGH
LYSRLVPLVLLLVDGSNPIDVTDSSWELYLLIQKTSDQQGNFHDRLLGFAAVYRFYHYPDSLR
LRLGQILVLPLYQRKGYGHYLLEVLNNVAIADDVYDFTIEEPVDNLQHLRTCIDVQRLLSFDKV
QQAVNSTVSQLKQGKLSKKTYIPRLLPPPSVVEDARKRFKINKKQFLQCWEILVYLGLDPADK
SIQDYFSVISNRVRADILGKDSETAGKKVIEVPSDFDPEMSFVLHRAKAGGETNGIQVEDNQN
KQEEQLQQLIDERLKDIKLIAQKVSRK

SEQ ID NO: 289
MALPMEFWGVEVKAGQPLKVNPGNAKILHLSQASLGECKSSKGNESVPLHVKFGDQKLVLG
TLSTENFPQLAFDLVFEKEFELSHNWKSGSVYFCGYKSVVHDDDDEFSDLESDSEEEDLPMI
GVENGKVAAQASAKTATASANASKVESSGKQKARIPQPMKVDEDDSDEDDDDEDEDESDE
EGVDGEADSDEEEDESDEEETPKKAEIGKKRAADSATKTPVPAKKSKLPTPQKTDGKKGGH
TATPHPAKQAGKNPANSANKSQSPKSAGQVSCKSCSKTFNSDGALQSHSKAKHGGK

SEQ ID NO: 290
MEFWGVEVKAGQPLKVNPGNAKILHLSQASLGECKSSKGNESVPLHVKFGDQKLVLGTLST
ENFPQLAFDLVFEKEFELSHNWKSGSVYFCGYKSVVHDDDDEFSDLESDSEEEDLPMIGVE
NGKVAAQASAKTATASANASKVESSGKQKASIPQPMKVDEDDSDEDDDEDDDDEDESDEG
VDGEADSDEEEDESDEEETPKKAEIGKKRAADSATKTPVPAKKSKLPTPQKTDGKKGGHTA
TPHPAKQAGKNPANSANKSQSPKSAGQVSCKSCSKTFNSDGALQSHSKAKHGGK

SEQ ID NO: 291
MEFWGVEVKSGEPLNVEPGAETVVHLSQACLGETKEKTKESVLLYVHIGVQKLVLGTLSADK
FPQIPFDLVFEKSFKLSHNWKNGSVFFSGYKTLLPCGSDADSPYSDSDTDEGLPINVTAQAD
VPAKKAPVTANANAAKPNLASAKQKVKIVESNEDGKNEGDDDEDADVSSDDDAEDDSGDE
DMVDGGDESSDEDDDDSEEGESSEEEEPKAQPSKKRPADSVLKTPASDKKSKLETPQKTD
GKKASEHVATPYPSKQAGKAIASKGQAKQQTPNSNEFSCKPCNRSFKSDQALQSHNKAKH
GGS

SEQ ID NO: 292
MDTGGNSLPSGPDGVKRKVCYFYDPEVGNYYLLQHMQVLKPVPARDRDLCRFHADDYVAF
LRSITPETQQDQLRQLKRFNVGEDCPVFDGLHSFCQTYAGGSVGGAVKLNHGLCDIAINWA
GGLHHAKKCEASGFCYVNDIVLGILELLKQHERVLYVDIDIHHGDGVEEAFYTTDRVMTVSFH
KFGDYFPGTGDIRDIGYGKGKYYSLNVPLDDGIDDESYHSLFKPIIGKVMEVFKPGAVVLQCG
ADSLSGDRLGCFNLSIKGHAECVRYMRSFNVPVLLLGGGGYTIRNVARCWCYETGVALGLE
VDDKMPQHEYYEYFGPDYTLHVAPSNMENKNSRQLLEEIRSKLLENLSKLQHAPSVPFQER

PPDTELPEADEDQEDPDERWDPDSDMDVDEDRKPLPSRVKRELIVEPEVKDQDSQKASIDH
GRGLDTTQEDNASIKVSDMNSMITDEQSVKMEQDNVNKPSEQIFPK

SEQ ID NO: 293
MDTGGNSLPSGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHM
QVLKPVPARDRDLCRFHADDYVAFLRSITPETQQDQLRQLKRFNVGEDCPVFDGLHSFCQT
YAGGSVGGAVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLGILELLKQHERVLYVDI
DIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDIGYGKGKYYSLNVPLDDGIDDESY
HSLFKPIIGKVMEVFKPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVRYMRSFNVPVLLLG
GGGYTIRNVARCWCYETGVALGLEVDDKMPQHEYYEYFGPDYTLHVAPSNMENKNSRQLL
EDIRSKLLENLSKLQHAPSVPFQERPPDTELPEADEDQEDPDERWDPDSDMDVDEDRKPLP
SRVKRELIVEPEVKDQDSQKASIDHGRGLDTTQEDNASIKVSDMNSMITDEQSVKMEQDNV
NKPSEQIFPK

SEQ ID NO: 294
MRPKDRISYFYDGDVGSVYFGPNHPMKPHRLCMTHHLVLSYELHTKMEIYRPHKAYPAELA
QFHSPDYVEFLHRITPDTQHLFPNDLAKYNLGEDCPVFENLFEFCQIYAGGTIDAARRLNNQL
CDIAINWAGGLHHAKKCEASGFCYINDLVLGILELLKYHARVLYIDIDVHHGDGVEEAFYFTDR
VMTVSFHKFGDMFFPGTGDVKEIGGKEGKFYAINVPLKDGIDDTSFTRLFKAIISKVVETYQP
GAIVLQCGADSLAGDRLGCFNLSIDGHSECVRFVKKFNLPLLVTGGGGYTKENVARCWVVE
TGVLLDTELPNEIPENEYFKYFAPDYSLKIPRGNIVLENLNSKSYLSAIKVQVLENLRNIQHAPS
VQMQEVPPDFYIPDFDEDEQNPDERMDQHTQDKQIQRDDEYYDGDNDNDHNMDDS

SEQ ID NO: 295
MTVAEDFHVNNRSKMVSQATPESRLTGGEDDNSLHNQVDELLCQELPERQVILEFEGTRPK
PYFSDHNGGENSALGVRATEDDLNSDVEAEEKQKEMTLEDMYKNDGTLYDDDEDDSDWEP
VKRQVELMRWFCTNCTMVNVEDVFLCDICGEHRDSGILRHGFYASPFMQDVGAPSVEAEV
QESREDHARSSPPSSSTVVGFDEKMLLHSEVEMKSHPHPERADRLQAIAASLATAGIFPGRC
RSLPVREITKEELQMVHSSEHVDAVEMTSHMFSSYFTPDTYANEHSARAARIAAGLCADLAS
TIISGRSKNGFALVRPPGHHAGIKHAMGFCLHNNAAVAALAAQGAGAKKVLIVDWDVHHGN
GTQEIFDGNKSVLYISLHRHEGGNFYPGTGAAHEVGTMGAEGYCVNIPWSRRGVGDNDYVF
AFHHIVLPIASAFAPDFTIISAGFDAARGDPLGCCDVTPAGYAQMTHMLSALSGGKLLVILEGG
YNLRSISSSAVAVIKVLLGDSPISEIADAVPSKAGLRTVLEVLKIQRSYWPSLESIFWELQSQW
GMFLVDNRRKQIRKRRRVLVPIWWKWGRKSVLYHLLNGHLHVKTKR

SEQ ID NO: 296
MAAAPSSPPTNRVDVFWHDGMLSHDTGRGVFDTGSDPGFLDVLEKHPENPDRVRNMVSIL
KRGPISPFISWHTATPALISQLLSFHSPEYINELVEADKNGGKVLCAGTFLNPGSWDAALLAA
GNTLSAMKYVLDGKGKIAYALVRPPGHHAQPSQADGYCFLNNAGLAVRLALDSGCKRVVVV
DIDVHYGNGTAEGFYQSSDVLTISLHMNHGSWGPSHPQSGSVDELGEDEGYGYNMNIPLPN
GTGDRGYEYAVTELVVPAVESFKPEMVVLVVGQDSSAFDPNGRQCLTMDGYRAIGRTIRGL
ADRHSGGRILIVQEGGYHVTYSAYCLHATVEGILDLPDPLLADPIAYYPEDEAFPVKVVDSIKR
YLVDKVPFLKEH

SEQ ID NO: 297
MVESSGGASLPSVGQDARKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHNLIVHYYLHRR
MEISRPFPAATTDIRRFHSEDYVTFISSVTPETVSDPAFSRQLKRFNVGEDCPVFDGIFGFCQ
ASAGGSMGAAVKLNRGDSDIALNWAGGLHHAKKSEASGFCYVNDIVLGILELLKVHKRVLYV
DIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGSGHIKDTGAGPGKNYALNVPLNDGIDD
ESFRGMFRPIIQKVMEVYQPDAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSFNVPL
MVLGGGGYTMRNVARCWCYETAVAVGVEPENDLPYNEYYEYFGPDYTLHVEPCSMENLNA
PKDLERIRNMLLEQLSRIPHAPSVPFQMTPPITQEPEEAEEDMDERPKPRIWNGEDYESDAE
EDKSQHRSSNADALHDENVEMRDSVGENSGDKTREDRSPS

SEQ ID NO: 298
MVVPSSNPHNREMAIRRRMASTFNKREDDFPSLREYNDYLEEVEEMTFNLIEGVDVPTIEAKI
AKYQEENAEQIMINRAKKAEEFAAALAASKGLPPQTDPDGALNSQAGLSVGTQGQYAPAIAG
GQPRPTGMAPQPVPLGTGLDIHGYDDEEMIKLRAERGGRAGGWSIELSKKRALEEAFGSLW
L

SEQ ID NO: 299
MAAIISCHHYHSCCSSLIASKWVGARIPTSCFGRSSTQSNNAASVRQFVTRCSSSPSSRGQ
WQPHQNGEKGRSFSLRECAISIALAVGLVTGVPSLDMSTGNAYAASPALPDLSVLISGPPIKD
PEALLRYALPINNKAIREVQKPLEDITDSLKVAGLRALDSVERNVRQASRVLKQGKNLIVSGLA
ESKKDHGVELLDKLEAGMDELQQIVEDGNRDAVAGKQRELLNYVGGVEEDMVDGFPYEVP
EEYKNMPLLKGRAAVDMKVKVKDNPNLEECVFRIVLDGYNAPVTAGNFVDLVERHFYDGME
IQRADGFVVQTGDPEGPAESFIDPSTEKPRTIPLEIMVDGEKAPVYGATLEELGLYKAQTKLP
FNAFGTMAMARDEFEDNSASSQIFWLLKESELTPSNANILDGRYAVFGYVTENQDFLADLKV
GDVIESVQVVSGLDNLANPSYKIAG

SEQ ID NO: 300
MAGEDFDIPPADEMNEDFDLPDDDDDAPVMKAGDEKEIGKQGLKKKLVKEGDAWETPDNG
DEVEVHYTGTLLDGTQFDSSRDRGTPFKFTLGQGQVIKGWDQGIKTMKKGENAIFTIPPELA
YGEAGSPPTIPPNATLQFDVELLSWTSVKDICKDGGIFKKILVEGEKWENPKDLDEVLVKYEF
QLEDGTTIARSDGVEFTVKEGHFCPAVAKAVKTMKKGEKVLLTVKPQYGFGEKGKPASGDE
GAVPPNATLQITLELVSWKTVSEVTDDKKVIKKILKEGEGYERPNEGAVVEVKLIGKLQDGTV
FVKKGHDDCEELFKFKIDEEQVVDGLDKAVMNMKKGEVALLTVAPEYAFGSSESKQDLAVV
PPSSTVYYEVELVSFVKDKESWDMNTEEKIEAAGKKKEEGNVIFKAGKYAKASKRYEKAVKY
IEYDTSFSEDEKKQAKALKVACNLNDAACKLKLKDYNQAEKLCTKVLELDSRNVKALYRRAQ
AYIELSDLDLAEFDIKKALEIDPHNRDVKLEYKVLKEKVKEFNKKDAKFYGNMFAKMSKLEPV
EKTAAKEPEPMSIDSKA

SEQ ID NO: 301
MSTVYVLEPPTKGKVVLNTTHGPLDVELWPKEAPKAVRNFVQLCLEGYYDNTIFHRIIKDFLV
QGGDPTGSGTGGESIYGDAFSDEFHSRLRFKHRGLVACANAGSPHSNGSQFFITLDRCDWL
DRKNTIFGKITGDSIYNLSGLAEVETDKSDRPLDPPPKIISVEVLWNPFEDIVPRAPVRSLVPTV
PDVQNKEPKKKAVKKLNLLSFGEEAEEEEKALVVVKQKIKSSHDVLDDPRLLKEHIPSKQVDS
YDSKTARDVQSVREALSSKKQELQKESGAEFSNSFREIADDEDDDDDDASFDARMRRQILQ
KRKELGDLPPKPKPKSRDGISARKERETSISRDKDDDDDDDQPRVEKLSLKKKGIGSEARGE
RMANADADLQLLNDAERGRQLQKQKKHRLRGREDEVLTKLETFKASVFGKPLASSAKVGDG
DGDLSDWRSVKLKFAPEPGKDRMTRNEDPNDYVVVDPLLEKGKEKFNRMQAKEKRRGRE
WAGKSLT

SEQ ID NO: 302
MASAISMHSSGLLLLQGTNGKDVTEMGKAPASSRVANMQQRKYGATCCVARGLTSRSHYA
SSLAFKQFSKTPSIKYDRMVEIKAMATDLGLQAKVTNKCFFDVEIGGEPAGRIVIGLFGDDVP
KTVENFRALCTGEKGFGYKGCSFHRIIKDFMIQGGDFTRGNGTGGKSIYGSTFEDENFALKH
VGPGVLSMANAGPSTNGSQFFICTVKTPWLDNRHVVFGQVVDGMDVVQKLESQETSRSDV
PRQPCRIVNCGELPLDG

SEQ ID NO: 303
MAASFTALSNVGSLSSPRNGSEIRRFRPSCNVAASVRPPPLKAGLSASSSSSFSGSLRLIPLS
SSPQRKSRPCSVRASAEAAAAQSKVTNKVYLDISIGNPVGKLVGRIVIGLYGDDVPQTAENFR
ALCTGEKGFGYKGSTVHRVIKDFMIQGGDFDKGNGTGGKSIYGRTFKDENFKLSHVGPGVV
SMANAGPNTNGSQFFICTVKTPWLDQRHVVFGQVLEGMDIVRLIESQETDRGDRPRKRVVV
SDCGELPVV

SEQ ID NO: 304
MAEAIDLTGDGGVMKTIVRRAKPDAVSPSETLPLVDVRYEGVLAETGEVFDSTHEDNTLFSF
EIGKGSVISAWDTALRTMKVGEVAKITCKPEYAYGSTGSPPDIPPDATLIFEVELVACKPCKGF
SVTSVTEDKARLEELKKQREIAAATKEEEKKRREEAKAAAAARVQAKLDAKKGHGKGKGKA
K

SEQ ID NO: 305
MGNPKVFFDMSIGGQPAGRIVMELYADVVPRTAENFRALCTGEKGAGRSGKPLHYKGSSFH
RVIPGFMCQGGDFTAGNGTGGESIYGSKFADENFVKKHTGPGVLSMANAGPGTNGSQFFV
CTAKTEWLDGKHVVFGQIVDGMDVVKAIEKVGSSSGRTSKPVVVADCGQLS

SEQ ID NO: 306

MPNPKVFFDMTIGGAAAGRVVMELYADTTPRTAENFRALCTGEKGVGRSKKPLHYKGSKFH
RVIPSFMCQGGDFTAGNGTGGESIYGVKFADENFIKKHTGPGILSMANAGPGTNGSQFFICT
TKTEWLDGKHVVFGKVVEGMEVVKAIEKVGSSSGRTSKPVVVADCGQLP

SEQ ID NO: 307
MAEAIDLTGDGGVMKTIVRRAKPDAVSPSETLPLVDVRYEGVLAETGEVFDSTHEDNTLFSF
EIGKGSVISAWDTALRTMKVGEVAKITCKPEYAYGSTGSPPDIPPDATLIFEVELVACKPCKGF
SVTSVTEDKARLEELKKQREIAAATKEEEKKRREEAKAAAAARVQAKLDAKKGHGKGKGKA
K

SEQ ID NO: 308
MATARSFFLCALLLLATLYLAQAKKSEDLKEVTHKVYFDVEIAGKPAGRIVMGLYGKAVPKTA
ENFRALCTGEKGTGKSGKPLHYKGSSFHRIIPSFMLQGGDFTLGDGRGGESIYGEKFADENF
KLKHTGPGLLSMANAGPDTNGSQFFITTVTTSWLDGRHVVFGKVLSGMDVVYKVEAEGRQS
GTPKSKVVIADSGELPL

SEQ ID NO: 309
MMRREISVLLQPRFVLAFLALAVLLLVFAFPFSRQRGDQVEEEPEITHRVYLDVDIDGQHLGRI
VIGLYGEVVPRTVENFRALCTGEKGKSANGKKLHYKGTPFHRIISGFMIQGGDVIYGDGKGY
ESIYGGTFADENFRIKHSHAGIISMVNSGPDSNGSQFFITTVKASWLDGEHVVFGRVIQGMDT
VYAIEGGAGTYNGKPRKKVIIADSGEIPKSKWDEER

SEQ ID NO: 310
MWATAEGGPPEVTLETSMGSFTVELYFKHAPRTSRNFIELSRRGYYDNVKFHRIIKDFIVQG
GDPTGTGRGGESIYGKKFEDEIKPELKHTGAGILSMANAGPNTNGSQFFITLAPCPSLDGKH
TIFGRVCRGMEIIKRLGSVQTDNNDRPIHDVKILRTSVKD

SEQ ID NO: 311
MSNPKVFFDILIGKMKAGRVVMELFADVTPKTAENFRALCTGEKGIGRSGKPLHYKGSTFHRI
IPNFMCQGGDFTRGNGTGGESIYGMKFADENFKIKHTGLGVLSMANAGPDTNGSQFFICTE
KTPWLDGKHVVFGKVIDGYNVVKEMESVGSDSGSTRETVAIEDCGQLSEN

SEQ ID NO: 312
MDDDFEFPASSNVENDDDDGMDMDDMGGDVPEEEDPVASPAVLKVGEEREIGKAGFKKKL
VKEGEGWETPSSGDEVEVHYTGTLLDGTKFDSSRDRGTPFKFKLGRGQVIKGWDEGIKTMK
KGENAIFTIPPELAYGESGSPPTIPPNATLQFDVELLSWSSVKDICKDGGILKKVLVEGEKWD
NPKDLDEVFVKYEASLEDGTLISKSDGVEFTVGDGYFCAALAKAVKTMKKGEKVLLTVMPQY
AFGETGRPASGDEAAVPPDASLQIMLELVSWKTVSDVTKDKKVLKKTLKEGEGYERPNDGA
AVQVRLCGKLQDGTVFVKKDDEEPFEFKIDEEQVIDGLDRAVKNMKKGEVALVTIQPEYAFG
PTESQQDLAVVPANSTVYYEVELLSFVKEKESWEMNNQEKIEAAARKKEEGNAAFKAGKYV
RASKRYEKAVRFIEYDSSFSDEEKQQAKTLKNTCNLNDAACKLKLKDFKEAEKLCTKVLEGD
GKNVKALYRRAQAYIQLVDLDLAEQDIKKALEIDPNNRDVKLEYKILKEKVREYNKRDAQFYG
NMFAKMNKLEHSRTAGMGAKHEAAPMTIDSKA

SEQ ID NO: 313
MAKPRCFMDISIGGELEGRIVGELYTDVAPKTAENFRALCTGEKGIGPHTGAPLHYKGVRFH
RVIKGFMVQGGDISAGDGTGGESIYGLKFEDENFDLKHERKGMLSMANSGPNTNGSQFFITT
TRTSHLDGKHVVFGRVVKGMGVVRSVEHVTTAAGDCPTVDVVIADCGEIPAGADDGIRNFF
KDGDTYPDWPADLDESPAELSWWMDAVDSIKAFGNGSYKKQDYKMALRKYRKALRYLDIC
WEKEGIDEVESSSLRKTKSQIFTNSSACKLKLCDLKGALLDAEFAVRDGENNAKAYFRQGQA
HMELNDIDAAAESFSKALELEPNDVGIKKELNAAKKKIFERREQEKRAYRKMFL

SEQ ID NO: 314
MTKRKNPLVFLDVSIDGDPVERIVIELFADTVPRTAENFRSLCTGEKGVGKTTGKPLHYKGSY
FHRIIKGFMAQGGDFSNGNGTGGESIYGGKFADENFKLAHDGPGLLSMANGGPNTNGSQFF
IIFKRQPHLDGKHVVFGKVMRGMEVVKKIEQVGSANGKPLQPVKIVDCGETSETGTQDAVVE
EKSKSATLKAKKKRSARDSSSESRGKRRQRKSRKERTRKRRRYSSSDSYSSESSDSDSES
YSSDTESESKSHSESSVSDSSSSDGRRRKRKSTKREKLRRQRGKDSRGEQKSARYDKKSR
HKSADSSSDSESESSSRSRSRDDKKKSSRRESARSVSKLKDAEANSPENLESPRDREIKKV
EDNSSHEEGEFSPKNDVQHNGHGTDAKFGKYDDQRPRSDGSKKSSGSMRDSPKRLANSV

PQGSPSSSPAHKASEPSSSIRARNPSRSPAPDGNSKRIRKGRGFTERFSYARRYRTPSPED
VTYRPYHYGRRNFHDRRNDRYSNYRSYSERSPHRRYRSPPRGRSPPRYQRRRSRSRSVS
RSPGGNKGRYRGRDQSRSRSRSRSRSPRRGSSPANKQLPLSERLKSRLGTRVDEHSPRR
RRSSSRSHDSSRSRSPDEVPDKHEGKAAPVSPARSRSSSPSGRGLVSYGDASPDSGIN

SEQ ID NO: 315
MSVLLVTSLGDIVVDLHADRCPLTCKNFLKLCRIKYYNGCVFHTVQKDFTAQTGDPTGTGTG
GDSVYKFLYGDQARFFMDEIHLDLKHSKTGTVAMASGGENLNASQFYFTLRDDLDYLDGKH
TVFGEVAEGLETLTRINEAYVDEKGRPYKNIRIRHTYILDDPFDDPPQLAELIPDASPEGKPKD
EVVDDVRLEDDWVPLDEQLGPAQLEEAIRAKEAHSRAVVLESIGDIPDAEIKPPDNVLFVCKL
NPVTEDEDLHTIFSRFGTVVSADVIRDFKTGDSLCYAFIEFENKDSCEQAYFKMDNALIDDRRI
KVDFSQSVAKLWSQFKRKDSQAAKGKGCFKCGAPDHMARECPGSSTRQPLSKYILKEDNA
QRGGDDSRYEMVFDEDAPESPSHGKKRRGRDDRDDRHKMSRQSVEETKFNDREGGHSV
DKHRQSERSKHREDEMSRDSKASEAGRRRIDRDFPEEERDGEKYTESHRDRDGKRGDYR
DYRKGRADVQTHGDRRGDENYRRKSAAYDDGHEGAGAARRKDSNDDHHAYRRGYGDSR
KGTRDEDDDGRGRRDDPSYRRSSGHKDSSNGGREEQKYRSGETDGKSHPERSHRGDRR
R

SEQ ID NO: 316
MRPFNGGSSIACLVLVIAAGALAESQGPHLGSARVVFQTNYGDIEFGFFPGVAPRTVDHIFKL
VRLGCYNTNHFFRVDKGFVAQVADVANGRTAPMNDEQRTEAEKTIVGEFSNVKHVRGILSM
GRYDDPDSAQSSFSILLGDAPHLDGKYAIFGRVTKGDETLKKLEQLPTRREGMFVMPTERITI
LSSYYYDTGAESCEEENSTLRRRLAASAVEVERQRMKCFP

SEQ ID NO: 317
MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRALCTGEKGTGRSGKPLHFKGSSF
HRVIPGFMCQGGDFTRGNGTGGESIYGEKFADENFVKKHTGPGILSMANAGPNTNGSQFFI
CTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGSGRTSKPVVIADSGQLA

SEQ ID NO: 318
MRFTSITSAIALFAAAASALDKPLDIKVDKAVECSRKTKAGDKIQVHYRGTLEADGSEFDASYK
RGQPLSFHVGKGQVIKGWDQGLLDMCPGEKRTLTIQPDWGYGSRGMGPIPANSVLIFETEL
VEIAGVAREEL

SEQ ID NO: 319
MGNPKVFFDMSIGGQPAGRIVMELYADVVPRTAENFRALCTGEKGAGRSGKPLHYKGSSFH
RVIPGFMCQGGDFTAGNGTGGESIYGSKFADENFVKKHTGPGVLSMANAGPGTNGSQFFV
CTAKTEWLDGKHVVFGQIVDGMDVVKAIEKVGSSSGRTSKPVVVADCGQLS

SEQ ID NO: 320
MAVATRSRWVAMSVAWILVLFGTLALIQNRLSDTGASSDPKLVHRKVGEEKKKPDDLEEVTH
KVFFDVEIGGKPAGRIVMGLFGKTVPKTVENFRALCTGEKGIGKSGKPLNYKGSQFHRIIPKF
MIQGGDFTLGDGRGGESIYGNKFSDENFKLKHTDAGRLSMTNAGPDTNGSQFFITTVTTSW
LDGRHVVFGKVLSGMDVVHKIEAEGGQSGQPKSIVVISDSGELDL

SEQ ID NO: 321
MAVTLHTNLGDIKCEIFCDEVPKAAEHNARGILSMANSGPNTNGSQFFIAYAKQPHLNGLYTI
FGRVIHGFEVLDIMEKTQTGPGDRPLAEIRLNRVTIHANPLAG

SEQ ID NO: 322
MAVATRSRWVAMSVAWILVLFGTLALIQNRLSDTGASSDPKLVHRKVGEEKKKPDDLEEVTH
KVFFDVEIGGKPAGRIVMGLFGKTVPKTVENFRALCTGEKGIGKSGKPLNYKGSQFHRIIPKF
MIQGGDFTLGDGRGGESIYGNKFSDENFKLKHTDAGRLSMANAGPDTNGSQFFITTVTTSW
LDGRHVVFGKVLSGMDVVHKIEAEGGQSGQPKSIVVISDSGELDL

SEQ ID NO: 323
MGNPKVFFDMSIGGQPAGRIVMELYADVVPRTAENFRALCTGEKGAGRSGKPLHYKGSSFH
RVIPGFMCQGGDFTAGNGTGGESIYGSKFADENFVKKHTGPGVLSMANAGPGTNGSQFFV
CTAKTEWLDGKHVVFGQIVDGMDVVKAIEKVGSSSGRTSKPVVVADCGQLS

SEQ ID NO: 324
MSPVAANAMEEAAEPEVPAPVTPSKDDADTDAAVSRFLGFCKSKLGLAEGNCVQSSTLLRK
TAHVLRSSGTVIGTGTAEEAERYWFAFVLYTVRRVGERKAEDEQNGSDETEVPLSRILKASV
LNLIDFFKEIPQFVIKAGAIVSGIYGANWDSRLEAREMQTNYVHLCILCKFYKRICGEFFILNDA
KDDMKSADSSTSDPVIMYQPFGWLLFLALRIHALSRFKDLVSSTNALVSVLAILIIHLPTRFRKF
SISDSSQLVKRSEKGVDLVGSLAYRYDTSEDEIKRTLEKANNVIAEILGITPPPASECKAENLE
NVDTDGLIYFGNLMEETSLSSILSTLEKIYEDATRNDSEFDERVFINDDDSLLVSGSLSGAAINL
TGAKRKYDSFASPAKTITRPLSPSRSPASHINGIIGGTNLRITATPVATAMTTAKWLRTFVSPL
PSKPSTDLQGFLASCDRDVTSDVIRRANIILEAIFPNSPIGERTVTGGLQNANLMDNMWAEQR
RLEALKLYYRVLEAMCRAEAQILHSNNLTSLLTNERFHRCMLACSAELVLATHKTVTMLFPAV
LERTGITAFDLSKVIESFVRHEETLPRELRRHLNTLEERLLENMVWERGSSMYNSLVVARPAL
APEINRLGLLPEPMPSLDAIALLINFSSSGLPQSPVQKHEASPGQNGDIRSPKRISTEYRSVLV
ERNFTSPVKDRLLALSNIKSKLPPPPLQSAFASPTRPHPGGGGETCAETAIHIFFSKITKLAAV
RINAMLERLQLSQQIKEGVYCLFQQILSQRTNLFFNRHIDQVILCCFYGVAKINQINLTFREIIYN
YRKQPQCKPQVFRNVFVDWSTRRNGKAGNEHVDIISFYNEIFIPSVKPLLVELGPTGATTRTN
RTSEVGNKNDAQCPGSPKISSFPTLPDMSPKKVSASHNVYVSPLRSSKMDASISHSSKSYYA
CVGESTHAYQSPSKDLVAINSRLNGNRKVRGTLNFDDVDAGLVSDSMVANSLYLQNGSSMS
SSTAKSSEKPES

SEQ ID NO: 325
MRPILMKGHERPLTFLKYNREGDLLFSCAKDHTPTVWFADNGERLGTYRGHNGAVWCCDV
SRDSMRLITGSADTTAKLWSVQNGTQLFTFNFDSPARSVDFSIGDKLAVITTDPFMELPSAIH
VKRIARDPADQASESVLVLRGHQGRIARAVWGPLNKTIISAGEDAVIRIWDSETGKLLRESDK
ETGHKKAVTSLMKSVDGSHFVTGSQDKSAKLWDIRTLTLIKTYVTERPVNAVTMSPLLDHVV
LGGGQDASAVTMTDHRAGKFEAKFFDKILQEEIGGVKGHFGPINALAFNPDGKSFSSGGED
GYVRLHHFDPDYFNIKI

SEQ ID NO: 326
MDKKRTVVPLVCHGHSRPVVDLFYSPITPDGFFLISASKDSSPMLRNGETGDWIGTFEGHKG
AVWSCCLDTNALRAASGSADFSAKLWDALSGDELHSFEHKHIVRSCAFSEDTHLLLTGGVE
KILRIFDLNRPDAPPREVDNSPGSIRTVAWLHSDQTILSSCTDIGGVRLWDVRSGKIVQTLETK
SPVTSSEVSQDGRYITTADGSTVKFWDANHFGLVKSYNMPCNIESASLEPKLGNKFIAGGED
MWVHIFDFHTGEEIGCNKGHHGPVHCVRFSPGGESYASGSEDGTIRIWQTGPANNVEGDA
NPSNGPVTGKAKVGADEVTRKVEDLQIGKEGKDWREG

SEQ ID NO: 327
MAEGLILKGTMRAHTDMVTAIAIPIDNSDMVVTSSRDKSIILWHLTKEEKVYGVPRRRLTGHS
HFVQDVVLSSDGQFALSGSWDGELRLWDLATGVSARRFVGHTKDVLSVAFSIDNRQIVSAS
RDRTIKLWNTLGECKYTIQEGEAHTDWVSCVRFSPNTLQPTIVSASWDRTIKVWNLTNCKLR
NTLAGHNGYVNTVAVSPDGSLCASGGKDGVILLWDLAEGKRLYNLEAGAIIHSLCFSPNRYW
LCAATENSIKIWDLESKSIVEDLRVDLKNEADKTDGTTTAASNKKVIYCTSLNWSADGSTLFS
GYNDGVIRVWGTGRY

SEQ ID NO: 328
MAEGLHLKGTMKAHTDMVTAIAVPIDNADMIVTSSRDKSIILWHLTKEDKVYGVPRRRLTGHS
HFVQDVVLSSDGQFALSGSWDGELRLWDLATGVSARRFVGHTKDVLSVAFSIDNRQIVSAS
RDRTIKLWNTLGECKYTIQEGEAHNDWVSCVRFSPNTLQPTIVSASWDRTVKVWNLTNCKL
RNTLQGHSGYVNTVAVSPDGSLCASGGKDGVILLWDLAEGKKLYSLEAGAIIHSLCFSPNRY
WLCAATENSIKIWDLESKSIVEDLRVDLKNEADMSDGTTGAMSSNKKVIYCTSLNWSADGST
LFSGYNDGVIRVWGIGRY

SEQ ID NO: 329
MAEGLHLKGTMKAHTDMVTAIAVPIDNADMIVTSSRDKSIILWHLTKEDKVYGVPRRRLTGHS
HFVQDVVLSSDGQFALSGSWDGELRLWDLATGVSARRFVGHTKDVLSVAFSIDNRQIVSAS
RDRTIKLWNTLGECKYTIQEGEAHNDWVSCVRFSPNTLQPTIVSASWDRTVKVWNLTNCKL
RNTLQGHSGYVNTVAVSPDGSLCASGGKDGVILLWDLAEGKKLYSLEAGAIIHSLCFSPNRY
WLCAATENSIKIWDLESKSIVEDLRVDLKNEADMSDGTTGAMSSNKKVIYCTSLNWSADGST
LFSGYNDGVIRVWGIGRY

SEQ ID NO: 330

MSGVPAPPFATTTPENGTMSSNSPAFHRDSDDDDDQGEVFLDDSDIIHEVAVDDEDLPDAD
DEADEAEEADDSLHIFTGHNGEVYSLACSPTDATLVATGAGDDKGFLWRIGHGDWAVELQG
HKDSISSLAFSLDGQLLASGSLDGVIQIWDVPSGNLKGTLDGPGGGIEWIRWHPKGHIILAGS
EDSTVWMWNADKMAYLNMFSGHGNSVTCGDFTPDGKTICTGSDDATLRIWNPKSGENIHV
VKGHPYHAEGLTSMAISSDSGLAITGAKDGSVRIVNISSGRVVSSLDAHADSVEFVGLALSSP
WAATGSLDQKLIIWDLQHSSPRATCDHEDGVTCLSWVGASRFLASGCVDGKVRVWDSLSG
DCVRTFHGHSDAIQSLSVSANEEFLVSVSIDGTARVFEIAEFH

SEQ ID NO: 331
MGTSQHQLSSCLQLLPRRRGNKNLIFRRTMASGGAAAVAPPPGYKPYRHLKTLTGHVAAVS
CVKFSNDGTLLASASLDKTLIIWSSAALSLLHRLVGHSEGVSDLAWSSDSHYICSASDDRTLRI
WSSRSPFDCLKTLRGHTDFVFCVNFNPQSSLIVSGSFDETIRIWEVKTGRCLNVIRAHSMPVT
SVHFNRDGSLIVSGSHDGSCKIWDTKNGACLKTLIDDTVPAVSFAKFSPNGKFILVATLNDTL
KLWNYATGKFLKIYTGHKNSVYCLTSTFSVTNGKYIVSGSEDRCICIWDLQGKNLIQKLEGHS
DTVISVTCHPSENKIASAGLDSDRTVRIWLQDA

SEQ ID NO: 332
MPSQKIETGHQDIVHDVAMDYYGKRVATASSDTTIKIIGVSNSSGSQHLASLSGHKGPVWQV
AWAHPKFGSILASCSYDGQVILWKEGNQNDWAQAHVFNDHKSSVNSIAWAPHELGLCLAC
GSSDGNISVFTARPDGGWDTTRIEQAHPVGVTSVSWAPSMAPGALVGSGLLDPVQKLASG
GCDNTVKVWKLYNGTWKMDCFPALQMHSDWVRDVAWAPNLGLPKSTIASASQDGTVVIWT
VAKEGEQWQGKVLKDFKTPVWRVSWSLTGNLLAVADGNNNVTLWNEAVDGEWQQVTTVE
P

SEQ ID NO: 333
MKIAGLKSVENAHDESVWAAAWVPATESRPALLLTGSLDETVKLWRPDELALERTNAGHFL
GVVSVAAHPSGVIAASASIDSFVRVFDVDTNATIATLEAPPSEVWQMQFDPKGTTLAVAGGG
SASIKLWDTATWELNATLSIPRPEQPKPSEKGNKKFVLSVAWSPDGRRLACGSMDGTISIFD
VARAKFLHHLEGHFMPVRSLVFSPVEPRLLFSASDDAHVHMYDSEGKSLVGSMSGHASWV
LSVDVSPDGAALATGSSDRTVRLWDLSMRAAVQTMSNHSDQVWGVAFRPMAGAGVRAGG
RLASVSDDKSISLYDYS

SEQ ID NO: 334
MEIDLGNLAFDVDFHPSEQLVASGLITGDLLLYRYGDGSSPEKLLEVRAHGESCRAVRFINDG
KAILTGSPDCSILATDVETGSVVARVENAHEAAVNRLVNLTESTIATGDDNGCIKVWDTRQRS
CCNTFSAHEDFISDMTFASDSMKLVVTSGDGTLSVCNLRSNKVQTRSEFSEDELLSVVIMKN
GRKVVCGTQSGTLLLYSWGFFKDCSDRFVDLSPSSVDALLKLDEDRIIAGTENGLISLIGILPN
RIIQPIAEHSDHPIERLAFSHDKKFLGSISHDQTLKLWDLNDILGSEDSPSSQAAIDDSDSDEM
DVDANPPDSSKGNKKKHSGKGNDVGNANNFFADLGD

SEQ ID NO: 335
MSQQPSVILATASYDHTIRFWEAKSGRCYRTIQYPDSQVNRLEITPHKRYLAVAGNPSIRLFD
VNSNTPQPVMSFDSHTNNVMAVGFQYDGNWMYSGSEDGTVRIWDLRARGCQREYESRGA
VNTVVLHPNQTELISGDQNGNIRVWDLTANSCSCELVPEVDTAVRSLTVMWDGSLVVAANN
NGTCYVWRLLRGSQTMTNFEPLHKLQAHNGYILKCLLSPEFCEPHRYLATASSDHTVKIWNV
EGFTLEKTLIGHQRWVWDCVFSVDGAYLITASSDTTARLWSMSTGQDIRVYQGHHKATTCC
ALHDGAEGSPG

SEQ ID NO: 336
MEDAMDMEVEVEVEAEEHSPSSSNPSGSSFRRFGLKNSIQTNFGSDYVFEITPKFDWSLMG
VSLSSNAVKLYSPTTGQYCGECRGHSDTVNGISFSGPSSPHVLHSCSSDGTIRAWDTRSFK
EVSCISAGPSQEIFSFSFGGSSDSLLSAGCKSQILFWDWRNKKQVACLEDSHVDDVTQVCFV
PHHQNKLISASVDGLICIFDTAGDINDDEHMESVINVGTSIGKVGIFGQTFEKLWCLTHIETLSV
WDWKEGTNEANFEDARKLASDSWSLDHIDYFVDCHSAEEGEGLWVIGGTNAGTLGYFPVK
YKGGAAIGSPEAVLGGGHSDVVRSVLPMSGMAGTTSKTRGIFGWTGGEDGRLCCWLSDDS
SATSRSWMSSNLVLKSSRSHHKKNRHQPY

SEQ ID NO: 337
MSQHQEYPMEYAADDYDVGEVEDDMYFHERVMGDSDTDEDEEYDHLDNKITDTSAADARR
GKDIQGIPWERLSVTREKYRRTRIEQYKNYENVPQSGESSEKDCKPTRKGGNYYEFWRNTR

SVKSTILHFQLRNLVWSTTKHDVYLMSHFSIIHWSSLTCKKTEVLDVYGHVAPREKHPGSLLE
GFTQTQVSTLAVRDKLLIAGGFQGELICKNLDRPGVSYCCRTTYDDNAITNAVEIYDYPSGAV
HFMASNNDCGVRDFDMEKFELSRHFTFPWPVNHTSLSPDGKLLVIVGDNPEGIVVDSQRGK
TIRPLQGHLDFSFASAWHPDGHIFATGNQDKTCRIWDIRNLSKSVAVLKGNLGAIRSIRFTSD
GRFMAMAEPADFVHVYDVKSGYEKEQEIDFFGEISGVSFSPDTESLFVGVWDRTYGSLLQY
NRCRNYSYLDSM

SEQ ID NO: 338
MGASSDPNPDVSDEHQKRSEIYTYEAPWHIYAMNWSVRRDKKYRLAIASLLDHPAAAAAVP
NRVEIVQLDDSTGEIRADPNLSFDHPYPATKAAFVPDKDCQRADLLATSSDFLRIWRIADDSS
RVDLRSFLNGNKNSEFCRPLTSFDWNEAEPKRIGTSSIDTTCTIWDIERETVDTQLIAHDKEV
YDIAWGGVSVFASVSADGSVRVFDLRDKEHSTIIYESSEPDTPLVRLGWNKQDPRYMATIIM
DSAKVVVLDIRYPTMPVVELQRHQASVNAIAWAPHSSCHICTAGDDSQALIWDLSSMAQPVE
GGLDPILAYTAGAEIEQLQWSSSQPDWVAIAFSLKLQ

SEQ ID NO: 339
MRGGGGGGDATGWDEDAYRESVLKEREVQTRTVFRAAFAPSPSPSPSPDAVVVASSDGS
VASYSISACLSDHRLQSLRFADAKSQNVLEAEPACFLQGHDGPAYDVKFYGEGEDSLLLSCG
DDGRIRGWMWRDITSSEAHDHSQGNSAKPVLDLVNPQSRGPWGALSPIPENNALAVDVKR
GSIYAAAGDSCAYCWDVECGKIKTVFKGHSDYLHCIAARNSSSQIITGSEDGTARIWDCRSG
KCVQVIDPDKDHKKGFFASVSCLALDASESWLVCGRGRDLSVWSISASDCIAKISTNAPAQD
VLFDDNQILLVGAEPLISRLDMNGAVLSQIHCAPQSVFSVSLHQSGVTAVGGYGGLVDVISQF
GSHLCTFRCKCI

SEQ ID NO: 340
MEAPIIDPLQGDFPEVIEEYLEHGIMKCIAFNRRGTLLAAGCTDGSCIIWDFETRGVAKELRDK
ECTAAITSVCWSKYGHRILVSASDKSLILWDVLSGEKIAHTTLQHTVLQACLHPGSSTPSICLA
CPFSSAPMIVDLNTGSTTALPVLTADVSNGATPLSRNKTSDTSVTYSPCNACFNKHGDLVYA
GTSKGEILIIDHKNVRVCAIVLVSGGAVIKNVVFSRNGQYMLTNSNDRLIRIYKNLLPPKDGLK
MLDELNESFNESDDVEKLKAIGSKCLELLHEFQDSITRVQWKAPCFSGDGEWVIGGAASRG
EHKIYIWDRAGHLVKILEGPKEALMDLAWHPVHPIIISVSLTGLVYIWAKDYTENWSAFAPDFK
ELEENEEYVEREDEFDLVPETEKVKGLDVHEDDEVDVLTVERDSVFSDSDMSQEELCFLPA
VPCLDIPEQQDKCVGSCSKLPDGNHSGSPLSVEAGQNGNASNHNSSPLEPMENSTADDTD
GVRLKRKRKPSEKGLELQAEKVKKPVKPLKSSGRLSKTNKPVIDPDSSNGVYGDDGSD

SEQ ID NO: 341
MRGVSWPEDGNNPSTSSSSQRNQQQAHAPRAVSGHAASHPSASNIFKLLVQREVSPRSKH
SSKKLWREASKCQPYPFQQSCEAVRDVRQGLISWVESASLRHLSAKYCPLVPPPRSTIAAAF
SPDGKILASTHGDHTVKLIDSQTGSCLKVLRGHRRTPWVVRFHPLYPEILASGSLDHEVRLW
DANTAECIGSRNFYRPIASIAFHARGELLAVASGHKLYIWHYNRRGETSSPTIVLRTQRSLRA
VHFHPHAAPFLLTAEVNDLDSADSAMTLATSPGYLHYPPPTVYFADAHSHERSRLADELPLM
PLPLLMWPSFTRDDGRVPLQRIDGDVGLNGQQRVDSSSSVRLWTYSTPSGQYELLLSPVES
GNSPSMPEETGNNAFSSAVEAEVSQSAMDTVEDMEVQPEERNTQFFSFSDPRFWELPLLH
GWLVGQTQAGPRSVRQSSPGDIETQSAFGEVASVSPITSGVMPVSMDPSRFGGRSGSRYR
SPGSRGVHVTGPNNDGPRDENDPQSVVSKLRSELAASLAAAASTELPCTVKLRIWPHDVKD
PCAQLDLESCRLTIPHAVLCSEMGAHFSPCGRFLAACVACVLPHLESDPGLHGQVNQDVTG
VATSPTRHPISAHQIMYELRIYSLEEATFGIVLASRPVRAAHCLTSIQFSPTSEHLLLAYGRRHS
SLLKSIVIDGENTVPIYTILEVYRVSDMELVRVLPSAEDEVNVACFHPSVGGGLIYGTKEGKLRI
LHYDSSHGLNLKSSGFLDENVPEVQTYALEC

SEQ ID NO: 342
MDSAVAIAALSLVVGAAIALLFFGNYFRKRRSEVVAMAEADLQPHPKNPSRPPPQPAAKKVH
AKSHAHGADKDKNKRHHPLDLNTLKGHGDSVTGLCFASDGRSLATACADGVVRVFKLDDAS
NKSFKFLRINLPAGGHPTAVAFGDGVSSVIVASQHLSGCSLYMYGEEKPTNLDSNKQQTKLP
MPEIKWEHHKVHEQKAILTLSGAAANYDSGDGSTIIASCSEGTDIIIWHAKTGKILGNVDTNQL
KNTMSAISPNGRFIAAAAFTADVKVWEIVYSKDGSVKGVTKVMQLKGHKSAVTWLCFTPNSE
QIVTASKDGSIRIWNINVRYHLDEDTKTLKVFPIPLQDSSGTTLHYERLSLSPDGKILAATHGS
MLQWLCIETGKVLDTAEKAHDGDITCMSWAPQSIPTGDKKVNVLATASGDKKVKLWAAPPL
PS

SEQ ID NO: 343
MEVEPKKASKTFPVKPKLKPKPRTPSGKTPESKYWSSFKTTHPLDNLSFSVPSLAFSPSPPH
LLAAAHSATVSLFSPHRTTISSFSDVVSSLSFRSDGQLLAASDLSGLIQVFDVRSRTPLRRLRS
HARPVRFVRYPVLDKLHLVSGGDDALVKYWDVAGESVVSELRGHKDYVRCGDCSPADANC
FVTGSYDHVVKLWDVRVRDGNRAATEVNHGSPVQDVIFLPSGSLVATAGGNSVKIWDLIGG
GRMVYSMESHNKTVTSICVGTMGAQQSGEEGVQLRILSVGLDGYMKVFDYSRMKVTHSMR
FPAPLLSIGFSPDSNVRAIGTSNGILYVGKRKAKENAEGGANGILGLGSVEEPRRRVLKPSFY
RYFHRGQSEKPSEGDYLVMRPKKVKLAEHDKLLKKFQHKNALISVLGGNDPEKVVAVMEEL
VARRALLKCVLNLDADELGLILTFLHKNSTVPRYSSLLLGLAKKVIDLRLEDIRASDALKGHIRN
LKRSVDEEIRIQEGLQEIQGMVSPLLRIAGRR

SEQ ID NO: 344
MQGGSSGVGYGLKYQARCISDVKADTDHTSFLTGTLSLKEENEVHLLRLSSGGTELICEGLF
SHPSEIWDLSSCPFDQRIFSTVFSTGESYGAAVWQIPELYGQLNSPQLEKIASLDAHSRKISC
VLWWPSGRHDKLVSIDEENIFLWGLDCSKKSAQVQSQESAGMLHNLSGGAWDPHDVNTVA
ATCESSIQFWDLRTMKKANSLESVHARDLDYDMRKKHLLVTSEDESGVRVWDLRMPKAPIQ
EFPGHTHWTWAVRCNPDYEGLILSAGTDSAVNLWWSSTASSDELISERLIDSPTRKLDPLLH
SYNDYEDSVYGLAWSSREPWIFASLSYDGRVVVESVKPFLSRK

SEQ ID NO: 345
MAEEEGSAELEQQLEEEFAVWKKNTPILYDLLISHALEWPSLTVHWAPLLPQPSSSAAAAAG
DPSLAAHRLVLGTHTSDGAPNFLILADALLPSSESDHCGDDAVLPKVEISQKIRVDGEVNRAR
FMPQNHNIVGAKTNGCEVYVFDCSKQAAKQHDGGFDPDLRLTGHDGEGYGLSWSPLKENY
LLSASHDKKICLWDISAAAQDKVLGAMHVFEAHEGAVGDASWHSKNDNLFGSAGDDCQLMI
WDLRTNKAQQCVKAHEKEVNSVSFNSYNDWILATASSDTTVGLFDMRKLTTPLHVFSSHEG
EVLQVEWDPNHEAVLASSSEDRRVMVWDLNRIGDEQQEGDASDGPAELLFSHGGHKAKIS
DFSWNKNEPWVISSVAEDNSVQVWQMAESICGDDDDMQAMEGYI

SEQ ID NO: 346
MGNYGEEDEDQYFDALEETASVSDRGSNSSDCCSSGSGLDENVLDSLGFEFWTKFPESVR
ARRNRFLMLTGLGIEANSVDKEDAFPPSCNEIEVYTCKVTRDDGAVQRSLDSYNCISLLQSST
SIRSNQEVESLRGDSLLSSFRGRSKESDDLTELCGMGCPESKRNAVSEFGSVSQGSIEELR
RIVASSPLVHPLLHRKLEYERELIETKQKMGAGWLRKFGSATCISGRQGDTWSDPDDLEITA
GMKMRRVRAHSSKKKYKELSSLYAAQEFLAHEGSISTMKFSMDGQYLASAGEDTVVRVWK
VTEEDRSERVNVTVDPSCLYFALNESTQLASLNTNKEHIGKAKTFQRSSDSSCVILPLKVFQIT
EKPWHEFKGHNGEVLDLSWSSKGYLLSSSTDKTVRLWRVGCDRCQRVYSHNDYVTCISFN
PVNENFFISGSIDGKVRIWNVFGGQVVAYIDCREIVSAVCYRSDGKGAIVGTMTGNCLFYSIK
DNHLQMDAQVYLHGKKKSPGKRITGFQFPPNDPGKLMITSADSVIRVLSGLDVVCKLKGPRN
SGGPMIATFTSDGKHVISASEDSNVYIWNYAGQDKTSSRVKKIWSCESFWSSNASVALPWC
GIRTVPEALAPPSRSEERRASCAENGENHHMLEEYFQKMPPYSPDCFSLSRGFFLELLPKG
SATWPEEKLSDTSPPTVSSQAISKLEYKFLKSACHSVLSSAHMWGLVIVTAGWDGRIRTYHN
YGLPVRS

SEQ ID NO: 347
MDIDFKEYRLRCELRGHEDDVRGVCVCGDGSIGTSSRDRTVRLWAPSAGERRKYEVARVLL
GHKSFVGPLAWVPPSEELPEGGIVSGGMDTLVMAWDLRNGEAQTLKGHQLQVTGIVLDGG
DIVSASVDCTLIRWKNGQLTEHWEAHKAPIQAVIRLPSGELVTGSSDTTLKLWRGKTCTQTFV
GHTDTVRGLAVMPDLGILSASHDGSIRLWAVSGECLMEMVDHTSIVYSVDSHASGLIVSGSE
DRFAKIWKDGVCFQSIEHPGCVWDVKFLEDGDIVTACSDGTIRIWTNQEDRMANSTELELFD
LELSSYKRSRKRVGGLKLEELPGLEALQVPGTSDGQTKVIREGDNGVAYAWNSTELKWDKI
GEVVDGPEDSMNRPALDGVQYDYVFDVDIGDGEPTRKLPYNRSDNPYDTADKWLLKENLPL
SYRQQIVEFILANSGQRDFNLDPSFRDPYTGSSAYVPGAPSQLAAKQARPTFKHIPKKGMLV
FDAAQFDGILKKINEFNNTLLSNQEKKNLSLTDIEISRLGAVVKILKDTSHYHSSKFADADFDLM
LKLLESWPYEMMFPVIDIFRMVILHPDGADGLLRHQEDKKDVLMESIKRATGNPSVPANFLTS
IRAVTNLFKNSAYYSWLQKHRSEMLDAFSSCSSSSNKNLQLSYATLLLNYAVLLIEKKDEEGQ
SQVLSAALELAENESLEVDARYRALVAIGSLMLDGLVKRIALDFDVEHIAKAARTSKEAKIAEV
GADIELLIKQS

SEQ ID NO: 348

MEFTEAYKQSGPCCFSPNARFIAVAVDYRLVIRDTLSLKVVQLFSCLDKISYIEWALDSEYILC
GLYKRPMIQAWSLIQPEWTCKIDEGPAGIAYARWSPDSRHILTTSDFQLRLTVWSLVNTACV
HVQWPKHASKGVSFTRDGKFAAICTRHDCKDYINLLSCHNWEIMGVFAVDTLDLADIQWSP
DDSAIVIWDSPLEYKVLVYSPDGRCLFKYQAYESGLGVKSVSWSPCGQFLAVGSYDQMLRV
LSHLTWKTFAEFTHLSNVRAPCCAAIFKEVDEPLQIDMSELSLSDDYMQGNSGDAPEGHYRV
RYDVTEVPITLPCQKPPADRPNPKQGIGLMSWSNDSQYICTRNDSMPTILWIWDMRHLELAA
ILVQKDPIRAAVWDPTGTRLVLCTGSSHLYMWTPSGAYCVSVPLSQFNITDLKWNSDGSCLL
LKDKESFCCAAAPLPPDESSDYSSDD

SEQ ID NO: 349
MATIAALDDDMVRSMSIGAVFSDFVGKLNSLDFHRKDDILVTAGEDDSVRLYDIANARLLKTT
FHKKHGTDRVCFTHHPNSLICSSTKNLDTGESLRYISMYDNRSLRYFKGHKQRVVSLCMSPI
NDSFMSGSLDHSVRMWDLRVNACQGILRLRGRPTVAYDQQGLVFAVAMEGGAIKLFDSRS
YDKGPFDAFLVGGDTSEVCDIKFSNDGKSVLLSTTNNNIYVLDAYAGDKQCGFNLEPSPSTPI
EASFSPDGQYVVSGSGDGTLHAWNISRRNEVACWNSHIGVASCLKWAPRRAMFVAASTVL
TFWIPNSEPELASAKGEAGVPPEQV

SEQ ID NO: 350
MSVAELKERHRAATETVNSLRERLKQKRVQLLDTDVAGYARTQGKTPVTFGATDLVCCRTL
QGHTGKVYSLDWTPERNRIVSVSQDGRFIVWNALTSQKTHAIRLPCAWVMTCAFAPNGQSV
ACGGLDSVCSIFNLNSPVDRDGNLPVSRMLSGHKGYVSSCQYVPDGDAHLITGSGDQTCVL
WDITTGLRTSVFGGEFQSGHTADVLSVSINGSSPRIFVSGSCDSTARMWDTRVASRAVHTY
HGHEGDVNAVKFFPDGNRFGTGSDDGTCRLFDIRTGHELQVYYQQRGIDEIPHVTSIAFSIS
GRLLIAGYSNGDCFVWDTLLAQVVLNLGSLQNSHEGRISCLGVSADGSALCTGSWDTNLKI
WAFGGIRRVT

SEQ ID NO: 351
MKKRPRGASLDQAVVDIRRREVGGLSGLSFARRLAASEGLVLRLDIYNKLKGHRGCVNTVG
FNLDGDIVISGSDDRHVKLWDWQTGKVKLSFDSGHLSNVFQAKIMPYTDDRSIVTCAADGQA
RHAQILEGGQVQTMLLAKHRGRAHKLAIDPGSPHIVYTCGEDGLVQRLDLRSNTARELFTCR
EVYGTHVEVVHLNAIAIDPRNPNLFVIGGSDEYARVYDIRNYKKWNGSHNFGRSANYFCPSHLI
GEAHVGITGLAFSGQSELLVSYNDESIYLFTQEMGLGPDPLSASTKSVDSNSSEVTSPTAVN
VDDNVTPQVYKGHRNCETVKGVGFFGPKCEYVVSGSDCGRIFIWKKKGGQLIRVMAADKHV
VNCIEPHPHIPALASSGIENDIKIWTPKAIERATLPMNVEQLKPKARGWMNRISSPRQLLLQLY
SLERWPEHGGETSSGLAAGQEELTELFFALSANGNGSPDGGGDPSGPLL

SEQ ID NO: 352
MSKRGYKLQEFVAHSSNVNCLSIGKKACRLFLTGGDDCKVNLWAIGKPNSLMSLCGHTNAV
ESVAFDSAEVLVLAGASSGVIKLWDVEEAKMVRGLTGHRSNCTAMEFHPFGEFFASGSTDT
NLKIWDIRKKGCIHTYKGHTRGISTIRFSPDGRWVVSGGNDNVVKVWDLTAGKLLHDFKFHE
NHIRSIDFHPLEFLLATGSADRTVKFWDLETFELIGSSRPEAAGVRAIAFHPDGRTLFCGLEDS
LKVYSWEPVICHDGVDMGWSTLADLCIHDGKLLGCSYYQSSVGVWVADASLIEPYGTNVKP
QQKDSGDDEIEHQESRPSAKVGTTIRSTSIMRCASPDYETKDIKNIYVDTASGNPVSSQRVG
TTNFAKVTQPLDFNDTPNLTLRRQGLVTETPDGLSGHVPSKSITQPKVVSRDSPDGKDSSRR
ESITFSRTKPGMLLRPAHSRRPSSTKYDVDRLSACAEIGVLSSAKSGSESLVDSFLNIKVAPE
DGARNGCEDNHSSVKNVSVESEKVLPLQTPKTEKCDQTVGFKEEINSVKFVNGVAVVPGRT
RTLVEKFEKREKLNSTEDQTINTPENPTLDKTPPPSLAENEEKSDRLNIVERKATRMSSHMVT
AEDRTPVTLVGSPEDQSTVMAPQRELPADESSKTPPLPVEDLEIHHGSNVSEDKATILSSQT
VSEEDSKRSTLIRNFRRRDRFKSTEGRSPVMATQRKLPTDESGKTSSLPMEDLEIKGGLNVS
EDKATSFSSRAPPREDRAHSALVRNVRKRDKFKSTNDTITVMVHQRGLSTDEASTVSVERV
ERRQLSNNVENPLNNLPPHSVPPTTTRGEPQYVGSESDSVNHEDVTELLLGNHEVFLSTLR
SRLTKLQVV

SEQ ID NO: 353
MSTFLTGTALSNPNPNKSYEVVQPPNDSVSSLSFNPKANFLVATSWDNQVRCWEIVRSGTS
LGTTPKASISHDQPVLCSTWKDDGTTVFSGGCDKQVKMWPLSGGQPMTVAMHDAPIKEIS
WIPEMNLLVTGSWDKTLRYWDTRQANPVHIQQLPERCYALTVRHPLMVVGTADRNLIIYNLQ
SPQTEFKRISSPLKYQTRCLAAFPDQQGFLVGSIEGRVGVHHLDDSQQSKNFTFKCHREGS
EIYSVNSLNFHPVHHTFATAGSDGAFNFWDKDSKQRLKAMSRCSQPIPCSTFNNDGSIFAYS
ACYDWSKGAENHNPATAKTYIFLHLPQESEVKGKPRLGTTGRK

SEQ ID NO: 354
MEVEAQQRDVNNVMCQLVDPEGTTLGPPMYLPQDVGPQQLQQMVNKLLSNEDKLPYTFYI
SDQELVVPLESYLQKNKVSVEKVLSIVYQPQAIFRIRPVNRCSATIAGHSEAVLSVAFSPDGK
QLASGSGDTTVRLWDLSTQTPMFTCKGHKNWVLSIAWSPDGKHLVSGSKAGEIQCWDPLT
GQPSGNPLVGHKKWITGISWEPVHLSSPCRRFVSSSKDGDARIWDVTLRRCVICLSGHTLAV
TCVKWGGDGVIYTGSQDCTIKVWETSQGKLIRELKGHGHWVNSLALSTEYVLRTGAFDHTG
KQYSSAEEMKQVALERYKKMKGNAPERLVSGSDDFTMFLWEPSVSKHPKTRMTGHQQLVN
HVYFSPDGQWVASASFDKSVKLWNGITGKFVAAFRGHVGPVYQISWSADSRLLLSGSKDST
LKIWDIRTKKLKRDLPGHADEVFAVDWSPDGEKVVSGGKDKVLKLWMG

SEQ ID NO: 355
MDAGSAHSSSNMKTQSRSPLQEQFLQRRNSRENLDRFIPNRSAMDFDYAHYMLTEGRKGK
ENPAVSSPSREAYRKQLAETLNMNRTRILAFKNKPPTPVELIPHELTSAQPAKPTKTRRYIPQ
TSERTLDAPDLLDDYYLNLLDWGSSNVLSIALGNTVYLWNASDGSTSELVTIDDETGPVTSVS
WAPDGRHIAVGLNNSDVQLWDSADNRLLRTLRGGHRSRVGSLAWNNHILTTGGMDGLIVN
NDVRVRSHIVDTYRGHTQEVCGLKWSASGQQLASGGNDNILHIWDRSTASSNSPTQWLHR
LEEHTAAVKALAWCPFQGNLLASGGGGGDRTIKFWNTHTGACLNSVDTGSQVCALLWNKN
ERELLSSHGFTQNQLTLWKYPSMVKIAELTGHTSRVLFMAQSPDGCTVASAAGDETLRFWN
VFGVPEVAKPAPKANPEPFAHLNRIR

SEQ ID NO: 356
MEEAIPFKNLPSREYQGHKKKVHSVAWNCTGTKLASGSVDQTARVWHIEPHGHGKVKDIEL
KGHTDSVDQLCWDPKHADLIATASGDKTVRLWDARSGKCSQQAELSGENINITYKPDGTHV
AVGNRDDELTILDVRKFKPIHKRKFNYEVNEIAWNMSGEMFFLTTGNGTVEVLAYPSLRPVD
TLMAHTAGCYCIAIDPVGRYFAVGSADSLVSLWDISEMLCVRTFTKLEWPVRTISFNHTGDYV
ASASEDLFIDISNVQTGRTVHQIPCRAAMNSVEWNPKYNLLAYAGDDKNKYQADEGVFRIFG
FESA

SEQ ID NO: 357
MGKDEEEMRGEIEERLINEEYKVWKKNTPFLYDLVITHALEWPSLTVEWLPDREEPPGKDYS
VQKLVLGTHTSENEPNYLMLAQVQLPLEDAENDARHYDDDRADVGGFGCANGKVQIIQQIN
HDGEVNRARYMPQNSFIIATKTVSAEVYVFDYSKHPSKPPLDGACSPDLRLRGHSTEGYGLS
WSKFKQGHLLSGSDDAQICLWDINATPKNKSLDAMQIFKVHEGVVEDVAWHLRHEYLFGSV
GDDQYLLIWDLRTPSVTKPVQSVVAHQSEVNCLAFNPFNEWVVATGSTDKTVKLFDLRKIST
ALHTFDAHKEEVFQVGWNPKNETILASCCLGRRLMVWDLSRIDEEQTPEDAEDGPPELLFIH
GGHTSKISDFSWNTCEDWVVASVAEDNILQIWQMAENIYHDEDDVPGEESNKGS

SEQ ID NO: 358
MMRGFSCTEDGDAPSTSSTSPPPPPPPPHRQQMQAPRASSSSSGQPTSRRSTGNVFKLLA
RREVSPRSKHSLKKFWGEASECQLCPFQQSYEAVRDVRRSLISWVEAFSLQHLSAKYCPLM
PPPRSTIAAAFSPDGKILASTHGDHTVKLIDSQTGSCLKVLRGHRRTPWVVRFHPLYPEILAS
GSLDHEVHLWDANTAECIGSRNFYRPIASIAFHAQGDLLAVASGHKLYIWHYNRSGETSSPTI
VLRTPRSLRAVHFHPHAAPFLLTAEVNDLDLTDSAMTLATSPGYLHYPPPTIYLADAHSNERS
RLEDELPLMPSPLLMWPSFTRDDGRATLPHIGGDVGLSGQQRVDSLSSGQYEFHPSPIEPS
SSTSMHEEMGTDPFSSVRESEVTQSAMNIVDNTEVQPEERSTYSFSFSDPRFWELPSVYG
WLVGQTQAAPRTAPSPGALETASALGEVASVSPVRSEFMPGGMDQPRLGGRSGSGCRSS
GSRMMRTAGLNDHPHDENYPQSVVSKLRSELEASLAAAASTELPCTVKLRVWPYDMKDPC
ALFRSESCRLTIPHAVLCSEMGAHFSPCGRFFAACVACVLPQLEADPVLHGQVDPDVTGVAT
SPTRHPVSAYQIMYELRIYSLEEATFGMVLASRSIRAAHCLTSIQFSPTSEHLLLAYGRRHNSL
LKSIVIDGENTVPIYSILEVYRVSDMELVRVLPSAEDEVNVACFHPSVGGGLVYGTKEGKLRIL
QIDSSGGLNPKSTGFLDENMAEVPTYALEC

SEQ ID NO: 359
MGEGDLPRTEAGVLRGHEGAVLAARFNGDGNYCLSCGKDRTIRLWNPHRGIHIKTYKSHGR
EVRDVHCTSDNSKLISCGGDRQIFYWDVSTGRVIRRFRGHDSEVNAVKFNDYASVVVSAGY
DRSVRAWDCRSHSTEPIQIINTFQDSVMSVCLTKTEIIGGSVDGTVRTFDIRIGREISDDLGQP
VNCISMSNDGNCILASCLDSTLRLVDRSAGELLQEYKGHTCKSYKLDCCLTNTDAHVAGGSE
DGYVFFWDLVDASVISKFRAHSSVVTSVSYHPKEDCMITASVDGTIKVWKT

SEQ ID NO: 360
MACIKGVGRSASVAMAPDGGYLATGTMAGTVDLSFSSSASLEIFGLDFQSDDRDLPLIAESP
SSERFNRLSWGKNGSGSDEFSLGLIAGGLVDGTIGLWNPLSLIRSEAGDKAIVGHLSRHKGP
VRGLEFNVIAPNLLASGADDGEICIWDLAAPREPSHFPPLRGSGSAAQGEISFLSWNSKVQHI
LASTSYNGTTVVWDLKKQKPVISFSDSVRRRCSVLQWNPDLATQLVVASDEDSSPTLRLWD
MRNIMSPVKEFAGHTRGVIAMSWCPNDSSYLVTCAKDNRTICWDTVTGEIVCELPAGSNWN
FDVHWYPKIPGVISASSFDGKIGIYNVEGCSRYGVRENEFGAATLRAPKWFKRPVGASFGFG
GKVVSFHTRSTGGPSVNSSEVFVHDIITEQTLVSRSSEFEAAIQSGDRPSLRALCEKKSQHC
ESTDDQETWGFLKVLLEDDGTARSKLLAHLGFDIPTETNDGSQEDLSQQVNALGLEDVTADK
VVQEDNNESMVFPTDNGEDFFNNLPSPRADTPVSTSADGFPTVNAAVEPSQDEVDGLEESS
DPSFDDSVQRALVVGDYKAAVALCMSANKLADALVIAHVGGASLWESTRDKYLKMSRLPYL
KVVFAMVNNDLQSLVDTRPLKFWKETLAILCSFAQGEEWAMLCNSLASKLMAAGNMLAATL
CFICAGNIDKTVEIWSRSLATEHDGMSYMDLLQDLMEKTIVLALASGQKQFSASVCKLVEKYA
EILASQGLLTTAMDYLKLLGTDDLSPELAVLRDRIAFSVEAEKGANISAFNGSQDPRGAVYGV
DQSNYGMVDTSQHYYPEAAQPQVPHTVPGSPYGENYQQPFGSSFGKGYNTPMQYQAPSQ
ASMFVPSEPPQNAQPSFVPTPVTSQPTTRSQFIPAPPLALRNPEQYQQPTLGSHLYPGSVN
PTFQPLPHAPGPVAPVPPQVSSVPGQNMPQAVAPTQMRGFMPVTNPGVVQNPGPISMQPA
TPIESAAAQPVVSPAAPPPTVQTADTSNVPAPQKPVIATLTRLYNETSEALGGSRANPAKKRE
IEDNSRKIGALFAKLNSGDISKNAADKLVQLCQALDNGDYSTALQIQVLLTTSEWDECNFWLA
TLKRMIKTRQNVRLS

SEQ ID NO: 361
MKERGKGAGRSVDERYTQWKSLVPVLYDWLANHNLVWPSLSCRWGPQLEQATYKNRQRL
YLSEQTDGSVPNTLVIANVEVVKPRVAAAEHISQFNEEARSPFVKKFKTIIHPGEVNRIRELPQ
NSKIVATHTDSPDVLIWDVETQPNRHAVLGASTSRPDLILTGHKDNAEFALAMSPTEPFVLSG
GKDRYVVLWSIQDHISTLAADPGSAKSPGSAGTNNKQSSKAAGGNDKTGDSPSIEPRGVYL
GHGDTVEDVTFCPSSAQEFCSVGDDSCLILWDARTGSSPAIKVEKAHHADLHCVDWNPHDV
NLILTGSADNTVRMFDRRNLTSGGVGSPVHTFEGHNAAVLCVQWSPDKSSVFGSSAEDGIL
NIWDHEKIGRKIETVGSKVPNSPPGLFFRHAGHRDKVVDFHWNSSDPWTIVSVSDDGESTG
GGGTLQIWRMIDLIYRPEEEVLAELDKFKSHILSCTS

SEQ ID NO: 362
MAKIAPGCEPVAGTLTPSKKREYRVTNRLQEGKRPLYAVVFNFIDSRYFNVFATVGGNRVTV
YQCLEGGVIAVLQSYIDEDKDESFYTVSWACNIDRTPFVVAGGINGIIRVIDAGNEKIHRSFVG
HGDSINEIRTQPLNPSLIVSASKDESVRLWNVHTGICILIFAGAGGHRNEVLSVDFHPSDKYRI
ASCGMDNTVKIWSMKEFWTYVEKSFTWTDLPSKFPTKYVQFPVFIAPVHSNYVDCNRWLGD
FVLSKSVDNEIVLWEPKMKEQSPGEGSVDILQKYPVPECDIWFIKFSCDFHYHSIAIGNREGKI
YVWELQSSPPVLIAKLSHPQSKSPIRQTAMSFDGSTILSCCEDGTIWRWDAITASTS

SEQ ID NO: 363
MNTAMHFGAGWRSIAEMGYTMSRLEIEPESCEDEKSLDGVGNSQGPNELPRCLDHELAHLT
NLKSRPHEHLIRDFPGRRALPVSTVKMLAGRECNYSRRGRFSSADCCHMLSRYVPVNGPSP
LDQMNSRAYVSQFSADGSLFVAGFQGSHIRIYNVDKGWKCQKNILTKSLRWTITDTSLSPDQ
RYLVYASMSPIVHIVDIGSAAMDSLANITEIHEGLDFSADSGPYSFGIFSVKFSTDGREVVAGS
SDDSIYVYDLVANKLSLRIPAHESDVNTVCFADESGHIIYSGSDDTYCKVWDRRCLSARNKPA
GVLMGHLEGITFIDSRGDGRYFISNGKDQTIKLWDIRKMGSDICRRGFRNFEWDYRWMDYP
PRARDSKHPFDLSVATYKGHSVLRTLIRCYFSPVHSTGQKYIYTGSHDSCVYIYDVVTGAQV
AALKHHKSPVRDCSWHPEYPMIVSSSWDGDIVKWEFFGNGETEIPAMKKRIRRRHLY

SEQ ID NO: 364
MEPQPQAPKKRGRKPKPKEDKKEEQLHQPPPPPPPQQQAAPAPAPAATRSSTSGSAGGR
DRRPQQQHAVDEKYARWKSLVPVLYDWLANHNLLWPSLSCRWGPQLEQATYKNRQRLYIS
EQTDGSVPNTLVIANCEVVKPRVAAAEHVSQFNEEARSPFIRKYKTIIHPGEVNRVRELPQNP
NIVATHTDSPDVLIWDVESQPNRHAVYGATASRPNLILTGHQENAEFALAMCPAEPFVLSGG
KDKTVVLWSIQDHITASATDQTTNKSPGSGGSIIKKTGEGNEETGNGPSVGPRGIYCGHEDT
VEDVAFCPSTAQEFCSVGDDSCLILWDARVGTNPVAKVEKAHNGDLHCVDWNPHDNNLILT
GSADNSVNMFDRRNLTSNGVGSPVYKFEGHKAAVLCVQWSPDKPSVFGSSAEDGLLNIWD
YERVDKKVDRAPNAPAGLFFQHAGHRDKIVDFHWNAADPWTMVSVSDDCDTAGGGGTLQI
WRMSDLIYRPEEEVLAELENFKAHVLECSKA

SEQ ID NO: 365
MGIFEPYRAVGYITTGVPFSVQRLGTETFVTVSVGKAFQVYNCAKLSLVLVGPQLPKKIRALA
SYREYTFAAYGSDIGIFKRAHQLATWSGHTAKVCLLLLFGEHILSVDVDGNAYIWAFKGMNY
NLSPVGHILLDSNFTPSCIMHPDTYLNKVILGSQEGPLQLWNISTKTKLYEFKGWNSSVSSCV
SSPALDVVAVGCADGKIHVHNIRYDEELVTFSHSMRGSVTALSFSTDGQPLLASGSSSGVVS
IWNLDKRRLQSVIRDAHDGSIISLHFFANEPVLMSSSADNSIKMWIFDTSDGDPRLLRFRSGH
SAPPLCIRFYANGRHILSAGQDRAFRLFSVVQDQQSRELSQRHVSKRAKKLKLKEEEIKLKPV
IAFDVAEIRERDWCNVVTSHMDTPQAYVWRLQNFVIGEHILRPCPNKPTPVKACMISACGNF
AILGTAGGWIERFNLQSGISRGSYIDQLEGTNSAHDGEVVGVACDATNTLMISAGYAGDIKV
WDFKGRELKSRWEIGSSLVKISYHRLNGLLATVADDFIIRLFDAVALRMVRKFEGHTDRITDL
CFSEDGKWLLSSSMDGSLRIWDIILARQVDAVFVDVSITALSLSPNMDILATTHVDQNGVFLW
VNQSMFSGDSDINLYASGKEVVTVKLPSVSSVEGSQVEESNEPTIRHSESKDVPSFRPSLEQ
IPDLVTLSLLPKSQWQSLINLDIIKVRNKPVEPPKKPEKAPFFLPSIPSLSGEILFKPSEMSDKG
DMKADEDKSKITPEVPSSRFLQLLHSCSEAKNFSPFTTYIKGLSPSTLDLELRMLQIIDDDAVD
ADADDPQDVDKRQELLSIELLMDYFIHEISCRSNFEFVQALVRLFLKIHGETIRRQSVLQNKAK
VLLETQCSVWQRVDKLFQGARCMVAFLSNSQF

SEQ ID NO: 366
MEETKVTCGSWIRRPENVNLAVLGRSPRRRGSAALEIFAFDPKSTSLSSSPLVAHVIEEIEGD
PLAIAVHPNGEDIVCFASSGSCLSFELSGQESNLKLLTKELPPLRGIGPQKCMAFSVDGSRFA
TGGVDGRLRILEWPSLRIILDEPKAHKSIRDLDFSLDSEFLATTSTDGSARIWKAEDGLPCTTL
TRRSDEKIELCRFSKDGTKPFLFCTVQRGDKAVTGVWDISTWNKIGHKRLLRKPAVVMSISL
DGKYLAQGSKDGDMCVVEVKKMEVSHWSKRLHLGTSLTSLEFCPIERVVITTSDEWGVLVT
KLNVPADWKAWQVYLLLLGLFLASLVAFYIFYENSDSFWGFPLGKDQPARPKIGSVLGDPKS
ADDQNMWGEFGPLDM

SEQ ID NO: 367
MADPVEHQHQQHQQHQLQQQRRRGWRIQGGQYLGEISALCFLHLPPPPLSLSSSPVLSLS
SGLDSESRDRPACSFRFPSAGSGSQVSLFDLASGAMVRTFYVFRGIRVHGIVLGCADFPGG
SSSSSSTLDYVIAVYGERRVKLFRLSVRLGRGAGEGSGTVLSADLELVSAAPRLSHWVMDV
RFLKENGTSEDELQRCLTVAIGCSDNSIRLWDVDKCSFVLAVSSPERCLLYSMRLWGDNLED
LQVASGTIYNEILIWKVVPNHDAPSSNELTEEGLTNSCAGNSVHECLRYEAYHICRLVGHEGS
IFRIAWSSDGSKLVSVSDDRSARIWEVHCKVQYSEDAGEVGLLFGHSARVWDCYISDNLIVT
AGEDCSCRVWGLDGQQHDVIKEHIGRGIWRCLYDPWSSLLVTGGFDSAIKVHKLDASLAEA
SAKQSNIKDLSDGTELFTTHLPNSSGHSGHMDSKSEYVRCLSFSCEDVMYIATNHGYLYHAK
LCNDGDLRWTELAQVSNEVQIICMELLPSNPYDPRIDADDWVAVGDGKGWTTVVRVVKNSD
SPKVSTSFSWAAEMDRQLLGIHWCKSLGHRFIFTADPRGALKLWRFFEVSQSSSLYPENSP
RISLIAEFKSDLGARIMCLDVAFESELLICGDLRGNLVLFPLLKDLLLDTFVVSAAKISPVNHFK
GAHGISAVSSISVAHMSFNHIELRSTGADGCICYMEYDKGLQSLNFVGMKQVKELSMIESVST
ENESTGYRTSGSYASGFASTDFIIWNLVTEAKVLQVSCGGWRRPHSYYLGDVPEMKNCFAY
VKDDIIYIRRHWIKDSKDKILPQNLRLQFHGREVHSLCFVTGDFQLRKNKQSSWIVTGCEDGT
VRLTRYTQCTDNWSSSKLLGEHVGGSAVRSICCVSNIHTTSSGTSVSDVKGIENLPKDIKGTL
MEDECNPSLLISVGAKRVLTSWLLRRRKQDGKEDDVTDLQEAENSSLPSSAGSSTFSFQWL
STDMPVKYSVPSKKSGSIKKLIGVSDTNVRCKSLLPDSEALQSKVSAVDKNEDDWRYLAVTA
FLVRHSGSRLIVCFIIVACSDATLAIRALVLPYRLWFDVALMVPLSSPVLSLQHVIIGRCQLPDE
NVQIGNVYVVISGATDGSIAFWDLTESVEAFMRRLSNIHLEKFMDCQKRPRTGRGSQGGRW
WRSLSKIACKEQPINDPVTAKAIKELNRKLTGGVACGSSSSMLDASPELDSNAANSSFEIIEV
NPFHVLNGVHQSGVNCLHVCETKHGQSSDGRFLYQLVSGGDDQALHLLKFEVLVQPPVQV
PDVPNSDIRNSILVEEFLLDEQNQKTKCTIEFISQEKIASAHNSAVKGVWTDGTWVFSTGLDQ
RVRCWISKDRGTPTELAHFIISVPEPEALDARSICWDQYQIAVAGRGMQMIEFHVPSSEIR

SEQ ID NO: 368
MPYKLSATLSNHSSDVRAVASPSDDLILSASRDSTAISWFRQSPSSFTPASVIRAGSRFVNAI
AYLPPTPRAPQGYAVVGGQDTVVNVFALGPGDKEEPEYTLVGHTDNVCALSVNSDDTIISGS
WDKTAKVWKDFALVYDLKGHQQSVWAVLAMNEKEFLTASADRTIKYWVQHKTMQTYEGHR
DAVRGLALIPDIGFASCSNDSEIRVWTMGGDVVYTLSGHTSFVYSLSVLPNGDLVSAGEDRS
VRVWRDGECSQVIVHPAISVWAVSTMPNGDIISGSSDGVVRVFSESEKRWATASELKALED
QIASQSLPSQQVGDVKKTDLPGPEALSVPGKKAGEVKMIRSGDVVEAHQWDSLASSWQKIG
EVVDAIGSGRKQLHDGKEYDYVFDVDIQEGAPPLKLPYNVSENPYTAAQRFLEQNDLPTGYL
DQVVKFIEQNTAGVKLGNDGYVDPFTGASRYQPATQSTSNTASSSYMDPFTGGSRHIAESA

PSNVPQGSHATGIIPFSKPIFFKLANVSAMQAKMFQFDEVLRNEISTATLAMRPDEVIMVNET
FTYLSKVVTSTSSARTSLGWIHIETIMQILDRWPVPQRFPVIDLGRLVTAYCMNAFSGPGDLE
KFFSCLFRTSEWTSITSGSKALTKAQETNVLLLFRTIANSLDGAPLNDMEWIKQIFRELAQTPQ
LVLNKSHRLALASVLFNFSCIGLKGPVPADVRTLHLTIILQVLRSPNDDPEVAYRTCVALGNML
YSDKTRGTPRDAQSPSPTELKSAVAAIKGGFSDPRINDVHREIMSLI

SEQ ID NO: 369
MPPQKIESGHKDTVHDLAMDYYGKRLATASSDHTINVVGVSSSGSQHLATLIGHQGPVWQIS
WAHPKFGSLLASCSYDGRVIIWREGNPNEWTQAQVFEEHKSSVNSVAWAPHELGLCLACG
SSDGNISVFTARQDGGWDTSRIDQAHPVGVTSVSWAPSTAPGALVGSGMMEPVQKLCSGG
CDNTVKVWKLYNRVWKLDCFPVLQMHTDWVRDVAWAPNLGLPKSTIASASQDGRVIIWTLA
KEGDQWQGKVLYDFRTPVWRVSWSLTGNILAVADGNNNVSLWNEAVDGEWIQVSTVEP

SEQ ID NO: 370
MSAPMLEIEARDVVKIVLQFCKENSLHQTFQTLQSECQVSLNTVDSIETFVADINSGRWDAIL
PQVAQLKLPRNTLEDLYEQIVLEMIELRELDTARAILRQTQAMGVMKQEQPERYLRLEHLLVR
TYFDPNEAYQDSTKEKRRAQIAQALAAEVTVVPPSRLMALVGQALKWQQHQGLLPPGTQFD
LFRGTAAMKQDVDDMYPTTLSHTIKFGTKSHAECARFSPDGQFLVSCSVDGFIEVWDYMSG
KLKKDLQYQADETFMMHDDPVLCVDFSRDSEMLASGSQDGKIKVWRIRTGQCLRRLERAHS
QGVTSVLFSRDGSQLLSTSFDGSARIHGLKSGKQLKEFRGHSSYVNDAIFSNDGSRVITASS
DCTVKVWDVKTSDCLQTFKPPPPLRGGDASVNSVHLFPKNADHIVVCNKTSSIYIMTLQGQV
VKSLSSGKREGGDFVAACVSPKGEWIYCVGEDRNLYCFSCQSGKLEHLMKVHEKDVIGVTH
HPHRNLVATYSEDSTMKLWKP

SEQ ID NO: 371
MDLLQSYAEDNDGDLGRHSSPEPSPPRLLPSKSAAPKVDDTTLALTVAQTNQTLARPIDPSQ
HAVAFNPTYDQLWAPICGPAHPYAKDGIAQGMRNHKLGFVEDAAIGSFLFDEQYNTFQRYG
YAADPCASTGNEYVGDLDALKQNDGISVYNIRQQEQKKYAEEYAKKKGEERGEGGREKAEV
VSDKSTFHGKEERDYQGRSWIAPPKDAKATNDHCYIPKRLVHTWSGHTKGVSAIRFFPKHG
HLILSAGMDTKVKIWDVFNSGKCMRTYMGHSKAVRDISFCNDGTKFLTAGYDKNIKYWDTET
GKVISTFSTGKIPYVVKLHPDDEKQNILLAGMSDKKIVQWDMNTGQITQEYDQHLGAVNTITF
VDDNRRFVTSSDDKSLRVWEFGIPVVIKYISEPHMHSMPSISLHPNTNWLAAQSLDNQILIYS
TRERFQLNKKKRFAGHIVAGYACQVNFSPDGRFVMSGDGEGRCWFWDWKSCKVFRTLKC
HEGVCIGCEWHPLEQSKVATCGWDGLIKYWD

SEQ ID NO: 372
MESNGNLEQTLQDGRIYRQLNSLIVAHLRDHNFPQAASAVALATMTPLNVEAPRNRLLELVA
KGLAVEKGELLRGVSHAGTNDLGGSIPASYGLVPAPWTAIDFSSLRDTKGMSKSFTKHETRH
LSDHKNVARCARFSTDGRFFATGSADTSIKLFEVSKIKQMMLPDSTDGAIRAVIRTFYDHTHP
VNDLDFHPQNTVLISAAKDHTVKFFDYSKATAKRAFRVIQDTHNVRSVAFHPSGDFLLAGTD
HPIPHLYDVNTFQCYLSANVPEFAVNAAINQVRYSSSGGMYVTASKDGTIRFWDGASANCV
RSIAGAHGAAEVTSANFTKDQRYVLSCGKDSTVKLWEVGTGRLVKQYLGATHMQLRCQAVF
NNTEEFVLSIDEPSNEIVVWDAMTAEKVARWPSNHNGPPRWIEHSPTEAAFVSCGTDRSIRF
WKETH

SEQ ID NO: 373
MSNFQGEDGEYVADDFEAEDGDEELHGRESADPESDVDEIDTPSNRFTDTTADQARRGRDI
QGIPWERLSITREKYRRTRLEQYKNYENVPQSGEKSGKDCTVTEKGNSFYEFRRNSRSVKS
TILHFQLRNLVWATSKHDVYLMSNYSVVHWSSLTGKKSEVLNLAGHVAPNEKHPGSLLEGFT
QTQVSTLAVKDRFLVAGGFQGELICKFLDRPGISFCSRTTYDDNAITNAVEIYVSPSGGIHFIA
SNNDCGVRDFDMENFELSKHFRFPWPVNHTSLSPDGKLLVIVGDDPEGILVDAKTGKTIMPL
RGHLDFSFASEWHPDGVTFATGNQDKTCRIWDIRNLSKSIAVLKGNLGAIRSIRYTSDGRYM
AIAEPADFVHVYDTKTGYKKEQEIDFFGEISGMSFSPDTESLFIGVWDRTYGSLLEYGRRRNF
SYLDCLV

SEQ ID NO: 374
MGVEEDLEDLNALAESTDAAVDGQAALASAVDSVTLQPAPPILPPVIPPPAVPVVAPVPTIPP
VLRPLAPLPIRPPVLRPPAPKRDEAGSSDSDSDHDGTAAGSTAEYEITEESRLVRERHEKAM
QDLMMKRRGAALAVPTNDKAVRARLRRLGEPMTLFGEREMERRDRLRMLMAKLDAEGQLE
KLMKAHEDEEAAASAAPEDVEEEMLQYPFYTEGSKALFNARIDIAKFSITRAALRLERARRRR

DDPDEDVDAEIDWALKKAESLSLHCSEIGDDRPLSGCSFSHDGKLLATCSMSGVAKLWDTC
RMPQVNRVLTLKGHTERATDVAFSPVQNHIATASADRTAKLWNTEGTILKTFEGHLDRLGRI
AFHPSGKYLGTTSFDKTWRLWDIESGEELLLQEGHSRSIYGIDFHRDGSLVASCGLDALARV
WDLRTGRSILALEGHVKPVLGVSFSPNGYHLATGGEDNTCRIWDLRKKKSLYTIPAHANLISE
VKFEPQEGYFLVTASYDTTAKVWSARDFKPVKTLSVHEAKITSVDITADASHIVTVSHDRTIKL
WTSNDDVKEQAMDVD

SEQ ID NO: 375
MVKAYLRYEPAAAFGVIASVESNIAYDASGKHLLAPALEKVGVWHVRQGVCTKALAPSASSA
AGPSLAVTAIASSPSSLIASGYADGSIRIWDFEKGSCETTLNGHKGAVSVLRYGKLGSLLASG
SKDNDIILWDVVGETGLYRLRGHRDQVTDLVFLDSDKKLVSSSKDKYLRVWDLETQHCMQIV
GGHHSEIWSLDTDPEERYLVTGSADPELRFYTVKNDSSDERSEADASGGVGNGDLASHNK
WDVLKQFGEIQRQSKDRVATVRFNKNGNLLACQAAGKLVEVFRVLDEAEAKRKAKRRLHRK
REKKGADVNENGDSSRGIGEGHDTMVTVADVFKLLQTIRASKKICSISFCPVAPKSSLATLAL
SLNNNLLEFHSIEADKTSKMLTIELQGHRSDVRSVTLSSDNTLLMSTSHNSVKIWNPSTGSCL
RTIDSGYGLCGLIVPQNKHALIGTKDGAIEIFDVGSGTCIEVVEAHGGSIRSIVAIPNQNGFVTG
SADHDIKFWEYGMKQKPGDNSKHLTVSNVRTLKMNDDVLVVAVSPDAQKIAVALLDCTVKV
FFMDSLKLMHSLYGHRLPVLCLDISSDGDLIVTGSADKNLMIWGLDFGDRHKSIFAHGDSIMA
VQFVGNTHYMFSVGKDRLVKYWDADKFELLLTLEGHHADIWCLAISNRGDFLVTGSHDRSIR
RWDRTEEPFFIEEEKEKRLEEMFESDLDNAFGNKYVPKEEIPEEGAVALAGKKTQETLSATD
SIIEALDIAEVELKRIAEHEEEKNNGKTAEFHPNYVMLGLSPSDFILRALSNVQTNDLEQTLLAL
PFSDALKLLSYLKDWTTYPDKVELVSRIATVLLQTHYNQLVSTPAARPLLTTLKDILHKKVKEC
KDTIGFNLAAMDHLKQLMALRSDALFQDAKVKLLEIRSQLSKRLEERTDPREAKRRKKKQKK
STNMHAWP

SEQ ID NO: 376
MGGVQAEREDKDKVSLELTEEILQSMEVGMTFRDYSGRISSMDFHRASSYLVTASDDESIRL
YDVASATCLKTINSKKYGVDLVSFTSHPMTVIYSSKNGWDESLRLLSLHDNKYLRYFKGHHD
RVVSLSLCPRNECFISGSLDRTVLLWDQRAEKCQGLLRVQGRPATAYDDPGLVFAIAFGGCV
RMFDARKYEKGPFEIFSVGGDVSDANVVKFSNDGRLMLLTTTDGHIHVLDSFRGTLLYTFNV
KPTSSKSTLEASFSPEGMFVISGSGDGSVYAWSVRGGKEVASWLSTDTEPPVIKWAPGNLM
FATGSSELSFWIPDLSKLGAYVGRK

SEQ ID NO: 377
MAAFGAAPAGNHNPNKSSEVIQPPSDSVSSLCFSPRANHLVATSWDNQVRCWELTKNGAS
VTSVPKASMSHDQPVLCSAWKDDGTTVFSGGCDKQAKMWSLMSGGQPVTVAMHDAPIKEI
AWIPEMNVLVTGSWDKTLKYWDTRQSNPVHTQQLPERCYAMTVRYPLMVVGTADRNLIVF
NLQNPQAEFKRFSSPLKYQTRCVAAFPDQQGFLVGSIEGRVGVHHLDDSQISKNFTFKCHR
DNNDIYSVNSLNFHPVHHTFATAGSDGTFNFWDKDSKQRLKAMSRCSQPIPCSTFNNDGTIY
AYSVCYDWSKGAENHNPATAKTYIFLHLPQESEVKAKPRVGTTNRK

SEQ ID NO: 378
MNCSISGEVPEEPVVSTKSGHVFERRLIERYVSDYGKCPVSGEPLTMDDVLPVKMGKIVKPR
PLQAASIPGLLSIFQNEWDSLMLSNFALEQQLHTARQELSHALYQHDAACRVIARLKKERDEA
RSLLALAERQIPMTASSDIAVNAPAMSNGRKASLDEEPGYAGKKMRPGISASIIAEITDCNLAL
SQQRKKRQIPSTLAPVEDLERYTQLSSYPLHKTGKPGITSLDICHSKDIIATGGIDTSAVLFDRS
SGQIMSTLSGHSKKVTSVNFDAQGDMVLTGSADKTVRIWQGSEDGSYNCRHILKDHTAEVQ
AITVHATNNYFATASLDNTWCFYEFSTGLCLTQVEGASGSEGYTSAAFHPDGLILGTGTSNA
DVKIWDVKTQANVTTFSGHTGAITAISFSENGYFLATAAQDGVKLWDLRKLKNFRTFSAYDK
DTGTNSVEFDHSGCYLGLAGSDIRVYQVASVKSEWNCVKTFPDLSGTGKVTCVKFGPDSKY
IAVGSMDHNLRIFGLPSEDGAMES

SEQ ID NO: 379
MAAPGVETLKKEIKELKEKIAQHRLDTDGEQPLPAAAKSKSVPEVSAALKQRRILKGHFGKIY
ALHWSADSRHLVSASQDGKLIIWNGFTTNKVHAIPLRSSWVMTCAYSPSGNLVACGGLDNL
CSVYKVPHGGNKESSSAQKTYGELAQHEGYLSCCRFIKDNEIVTSSGDSTCILWDVETKTPK
AIFNDHTGDVMSLAVFDDKGVFVSGSCDATAKLWDHRVHKQCVMTFQGHESDINSVQFFPD
GDAFGTGSDDSSCRLFDIRAYQQINKYSSDKILCGITSVAFSKTGKSLFAGYDDYNTYVWDTL
SGNQVEVLTGHENRVSCLGVSEDGKALATGSWDTLLKIWA

SEQ ID NO: 380
MGGVEDESEPASKRMKLSSRVLRGLANGSSRTEPAAGSSLDLMARPLPIEGDEEVIGSKGVI
KRVEFVRLIAKALYSLGYEKSGARLEEESGIPLQSSVVNLFMQQISDGLWDESVVTLHKIGLS
DENLVKSASFLILEQKFLELLDQEKAMDALKTLRTEITPLCIKNSRVRELSSCIISPSSCGLLNQ
NKRNSTRARSRSELLEELQKLLPPAVIIPERRLEHLVEQALVLQTDACMLHNSIDMEMSLYTD
HQCGKEHIPCRTLQILQSHNDEVWLVQFSHNGKYLASASNDRSAIIWEVDENGSVSLKHKLT
GHQKPISSVCWSPDDRQLLTCGVGETVRRWDVSSGECLRVYEKAGHGLISCAWFPDGKWI
CYGVSDRSICMCDLEGKEIECWKGQRTLSISDLEITSDGKQIISICRETAILLLDREAKYERMIE
ENQTITSFSLSKDNRYLLVNLLNQEIHLWDIKGDFRLVAKYKGLKRSRFVIRSCFGGLKQAFV
ASGSEDSQVYIWHKGSGELIEPLPGHSGAVNCVSWNPANHHMLASASDDRTIRIWGLNELN
TRHKGARPNGVHYCNGNGTS

SEQ ID NO: 381
MTQLAETYACMPSTERGRGILIAGNPKPGSNSVLYTNGRSVVILNLDNPLDISVYAEHAYPAT
VARFSPNGEWVASADSSGAVRIWGAYNDHVLKKEFKVLSGRIDDLQWSPDGLRIVASGDGK
GKSLVRAFMWDSGTNVGEFDGHSRRVLSCAFKPTRPFRIVTCGEDFLVNFYEGPPFKFKLS
RRDHSNFVNCLRFSPDGNRFISVSSDKKGIIYDGKTGEKIGELSSDGGHTGSIYAVSWSPDS
KQVITVSADKSAKIWDISEDGSGNLRKTLTSSGSGGVDDMLVGCLWQNNHLVTVSLGGTISI
YTAGDLDKAPVSFSGHMKNVSSLSVLKGDPKVILSSSYDGLIIKWIQGIGFSGRVQRKESTQI
KCLAAVDEEIVTSGYDNKVCRVSGSGDAEFIDIGCQPKDLSLALQCPEFALVSTDTGVVLLRG
AKIVSTINLGFAVTASTVAPDGTEAIIGAQDGKLRIYSISGDTLTEEAVLEKHRGAISVIHYSPDL
SMFASGDLNREAVVWDRASREVRLKNILYHTARINCLAWSPDSSTVATGSLDTCVIIYEVDKP
ASNRLTIKGAHLGGVYGLAFTDDFSVVSSGEDACIRVWKINRQ

SEQ ID NO: 382
MKVKVISRSTDEFTRERSQDLQRVFRNFDPNLRTQEKAVEYVRALNAAKLDKVFARPFVGA
MDGHVDSVSCMAKNPNYLKGIFSGSMDGDIRLWDIASRRTVCQFPGHQGPVRGLAASTDG
QILVSCGIDSTVRLWNVPVATLGESDGTHENLAKPLAVYVWKNAFWAVDHQWDGELFATAG
AQVDIWNQNRSQPISSFEWGTDTVISVRFNPGEPNVLATSGSDRSITLYDLRMSSPTRKVIM
RTKTNAISWNPMEPMNFTAANEDCNCYSYDARKLEEAKCVHKDHVSAVMDIDYSPTGREFV
TGSYDRTVRIFQYNGGHSREVYHTKRMQRVFCVKFSCDASYVISGSDDTNLRLWKAKASEQ
LGVVLPRERRKHEYHEAVKSRYKHLPEVKRIVRHRHLPKPIYKAGILRRTVNEADRRKEERR
KAHSAPGSSSAEPLRKRRIIKEIE

SEQ ID NO: 383
MVRSIKNPKKAKRKNKGSKNGDGSSSSSSIPSMPTKVWQPGVDKLEEGEELQCDPSAYNSL
HAFHIGWPCLSFDIVRDTLGLVRTEFPHQVYFVAGTQAEKPTWNSIGIFKVSNITGKRRELVP
SKPTDDADEESDSSDSDEDSDDEVGGSGTPILQLRKVGHEGCVNRIRAMNQNPHICASWG
DSGHVQIWDFSSHLNALAESEADVSQGASSVFNQAPLVKFGGHKDEGYALDWSPLVPGRL
VSGDCKNSIHLWEPTSGSTWNVDSTPFIGHAASVEDLQWSPTEENVFASCSVDGTIAIWDTR
LGKTPAASFKAHDADVNVISWNRLATCMLASGCDDGTFSIHDLRLLKEGDSVVAHFEYHKHP
VTSIEWSPHEASTLAVSSADCQLTIWDLSLEKDEEEEAEFKAKTKEQVNAPEDLPPQLLFVH
QGQKDLKELHWHAQIPGMIVSTAADGFNILMPSNIQSTLPSDGA

SEQ ID NO: 384
MERYKVIKELGDGTYGSVWKALNQQTHEIVAIKKMKRKYYIWEECINLREVKSLRKLNHPNIIK
LKEVIRENNELFFIFEYMECNLYQIMKERSTPFSETAIIKFCYQILQGLSYMHRNGYFHRDLKP
ENLLVTSDLIKIADFGLAREVLTSPPYTDYVSTRWYRAPEVLLQSPTYTTAIDMWAVGAILAEL
FTLHPLFPGESELDEIYKICGVLGTPDYETWPDGMQLAAFRNFIFPQFLPVNLSVLIPHASPEA
IDLITRLCSWDPQKRPTAEQALHHPFFRIGMSIPLSLGGHFQDNTCAAEVDTNFHSKKACKG
RGMGEKESSLECFLGLSLGLKPSLGHLGAMGSQGVGAVKQEVGSSPGCQSNPKQSLFQVL
NSRAILPLFSSSPNLNVVPVKSSLPSAYTVNSQVMWPTIAGPPAAAVTVSTLQPSILGDFKIFG
KSMGLASQYAGKEASPFS

SEQ ID NO: 385
MGEMGRGINNSSNNNNSNRPAWLQHYDLVGKIGEGTYGLVFLARSKLPNNRGLRIAIKKFK
QSKDGDGVSPTAIREIMLLREFSHENVVKLVNVHINHVDMSLYLAFDYAEHDLYEIIRHHHREKL
NHHNINQYTVKSLLWQLLNGLNYLHSNWIVHRDLKPSNILVMGEGEEHGVVKIADFGLARIYQ
APLKPLSDNGVVVTIWYRAPELLLGAKHYTSAVDMWAVGCIFAELITLKPLFQGVEVKASPNP

FQLDQLDKIFKVLGHPTIEKWPTLMNLPHWSKNLQQIQQHKYDNAGLHIGPIPAKSPAYDLLS
KMLEYDPRKRITAAQALEHEYFRIDPQPGRNALVPSQPGEKAINYPPRLVDANTDFDGTIAPQ
PSQVSSGNAPSGSIASAAVPAVRPLPQQMQLMGMQRMQNPGMAAFNLGAQASMSGLNHN
NIALQRGSSQQQAHQQVRRKEPNSGFPNTGYPPPPKSRRL

SEQ ID NO: 386
MDKYEKLEKVGEGTYGKVYKARDKMTGQLVALKKTRLEMDEEGVPPSSLREISLLQMLSQSI
YVVRLLCVEHVTKKGKPLLYLVFEYLDTDLKKFIDYRRSVNAGPLPQNVIQSFMYQLLKGVAH
CHSHGVLHRDLKPQNLLVDKSKGLLKVGDLGLGRAFTVPLKCYTHEVVTLWYRAPEVLLGST
HYSTPVDIWSVGCIFAEMVRRQPLFPGDCEIQQLLHIFTLLGTPTEEMWPGVKRLRDWHEYP
QWKPENLARAVPNLSPTGLDLISKMLQCDPAKRISAKAAMNHPYFDDLDKSQF

SEQ ID NO: 387
MDGYEKMDKVGEGTYGKVYMARDKKTGQLVALKKTRLENDGEGIPPTALREISLLQMLSQDI
YIVRLLDVKHTENKLGKPLLYLVFEYMESDLKKYIDSYRRSHTKMPPSMIKSFMYQLCRGVAY
CHSRGVMHRDLKPHNLLVDKEKGVLKIADLGLSRAFTVPVKKYTHEIVTLWYRAPEVLLGAT
HYSLPVDIWSVGCIFAEMSRMQALFTGDSEVQQLMNIFRFLGTPNEEVWPGVTKLKDWHIYP
EWKPQDISHAVPDLEPSGLDLLSQMLVYEPSKRISAKKALEHPYFDDLDKSQF

SEQ ID NO: 388
MDAYEKLEKVGEGTYGKVYKAKDKNTGQLVALKKTRLESDDEGIPPTALREISLLQMLSQDIH
IVRLLDVEHTENKNGKPLLYLVFEYMDSDLKKYIDGYRRSHTKVPPNIIKSFMYQLCQGVAYC
HSRGVMHRDLKPHNLLVDKQRGVVKIADLGLGRAFTIPIKKYTHEIVTLWYRAPEVLLGATHY
STPVDIWSVGCIFAEMVRLQALFIGDSEVQQLFKIFSFLGTPNEEIWPGVTKFRDWHIYPQWK
PQDISSAVPDLEPSGVDLLSKMLVYEPSKRISAKKALEHPYFDDLDKSQF

SEQ ID NO: 389
MDSYEKLEKVGEGTYGKVYKAKDKKTGKLVALKKTRLENDGEGIPPTALREISLLQMLSQDM
NIVRLLDVEHTENKNGKPLLYLVFEYMDSDLKKYVDGYRRSHTKMPPKIIKSFMYQLCQGVA
YCHSRGVMHRDLKPHNLLVDKQRGVLKIADLGLGRAFTVPIKKYTHEIVTLWYRAPEVLLGAT
HYSTPVDIWSVGCIFAEMSRMHALFCGDSEVQQLMSIFKFLGTPNEGVWPGVTKLKDWHIY
PEWRPQDLSRAVPDLEPSGVDLLTKMLVYEPSKRISAKKALQHPYFDDLDKSQF

SEQ ID NO: 390
MEKYEKLEKVGEGTYGKVYKGRDKRTGRLVALKKTPFHQEEGIPPTAIREISLLKSLSQCIYIV
KLLDVKASFNGKGKHVLFMVFEYADSDLKKHIDAHRQCNTKLSPRSIQSYMFQLCKGIAYCH
SHGVLHRDLKPQNILVDQKIGLLKIADLGLGRACTVPIKSYTFEVVTLWYRAPEVLLGAKRYS
MALDIWSLGCIFAELCNLQALFAGDSQIQQLINIFRLLGTPNEQLWPGVTQLSDWHEFPQWR
PQDLSKVVFNLDPNGVDLLSKMLQYDPAKRISAKEALDHPYFDSLDKSQF

SEQ ID NO: 391
MGCVCGKPSARAADYVESPAEKGASSNSRSSSMASRRLVAPAVMDQGIDAENGHEGDYRT
KLRGKQSNGADPVSLLSDDAEKQRHSRHHQHQQHHPIRPHHLRPQGEFVPNANSNPRFGN
PPRHIEGEQVAAGWPAWLTAVAGEAIKGWIPRRADSFEKLDKIGQGTYSNVYKARDLDTGKI
VALKKVRFDNLEPESVRFMAREIQVLRRLDHPNVVKLEGLVTSRMSCSLYLVFEYMDHDLAG
LAACPGIKFTEPQVKCYMQQLLRGLDHCHSRGVLHRDIKGSNLLIDNGGILKIADFGLATFFH
PDQRQPLTSRVVTLWYRPPELLLGATEYGVAVDLWSTGCILAELLAGKPIMPGRTEVEQLHKI
FKLCGSPSEDYWKKSKLPHATIFKPQQPYKRCVAETFKDFPPSALALMEVLLAIEPADRGTAT
SALKSDFFTTKPLACDPSSLPKYPPSKEFDAKIRDEEARRQRAAGGRGRDAARRPSRESRAI
PAPEANAELAISIQKRRLSSQGPSKSKSEKFNPQQEDGAVGFPIEPPRPMHIGIDAGATSRMY
SQQFGPSHSGPLSNQISSSIWGKNQKEDEIQMAPGRPSRSSKATISDFRKPGACAPQPGAD
LSHLSSLVATARSNAGIDTHKDRSGMWQHNRIDAIDGVHNNGKHEFLEVPEHPNRQDWTRF
QQPESFKGLDNYHLQDLPATHHRKDERVASKEATMNWQGYGGQGGDKIHYSGPLLPPSGN
IDEILKEHERHIQHAVRRARQDKGRPQRSNLSQNERKAFEHRSFVSGVNGNAGYSDLVNEL
PISVGSNRLKVSKTRGTEEIVELRELEREPLSSVMEKYEREHEM

SEQ ID NO: 392
MGCVCAKQSDILGEPESPKVKGSNLASSRWSVSSETKQLPQHSDSGILHHQHYYHPRDESD
EAKLKESNYGGSKRRTRQGRDPADLDMGIFVRTPSSQSEAELVAAGWPAWMAAFAGEAIH
GWIPRRAESFEKLYKIGQGTYSNVYKARDLDNGKIVALKKVRFDSLDAESVRFMAREILVLRK

LDHPNIVKLEGLVTSEVSSSLYLVFEYMEHDLAGLAACPGIKFTEPQVKCYMQQLLQGLDHC
HRHGVLHRDIKGSNLLIDNGGILKIADFGLATFFYPDQKQLLTSRVVTLWYRPPELLLGATDY
GVAVDIWSAGCILAELLAGKPILPGRTEVEQLHKIFKLCGSPSEDYWKESKLPHATIFKPQHPY
KSCIAEAFKDFSPSALALLETLLAIEPGHRGEASGALKSEFFTTEPLSCDPSSLPKYPPSKEFD
AKLRAQETRRQRDVGVRGHGSEAARRTSRLSRAGPTPNEGAELTALTQKQHSTSHATSNIG
SEKPSTKKEDYTAGLHIDPPRPVNHSYETTGVSRAYDAIRGVAYSGPLSQTHVSGSTSGKKP
KRDHVKGLSGQSSLQPSKPFIVSDSRSERIYEKSHVTDLSNHSRLAVGRNRDTTDPHKSLST
LMQQIQDGTLDGIDIGTHEYARAPVSSTKQKSAQLQRPSALKYVDNVQLQNTRVGSRQSDE
RPANKESDMVSHRQGQRIHCSGPLLHPSANIEDLLQKHEQQIQQAVRRAHHGKREALSNKS
SLPGKKPVDHRAWVSSGKGNKESPYFKGKGNKELSDLKGGPTAKVTNFRQKVM

SEQ ID NO: 393
MAVANPGQLNLQEAPSWGSRSVNCFEKLEQIGEGTYGQVYMAKEIETGEIVALKKIRMDNE
REGFPITAIREIKLLKKLQHENVIKLKEIVTSPGPEKDEQGKSDGNKYNGSIYMVFEYMDHDLT
GLAERPGMRFSVPQIKCYMKQLLIGLHYCHINQVLHRDIKGSNLLIDNNGILKLADFGLARSFC
SDQNGNLTNRVITLWYRPPELLLGSTKYGPAVDMWSVGCIFAELLYGKPILPGKNEPEQLTKI
FELCGSPDESNWPGVSKLPWYSNFKPQRQMKRRVRESFKNFDRHALDLVEKMLTLDPSQR
ISAKDALDAEYFWTDPVPCAPSSLPRYEPSHDFQTKRKRQQQRQHDEMTKRQKISQHPPQ
QHVRLPPIQNAGQGHLPLRPGPNPTMHNPPPQFPVGPSHYTGGPRGAGGQNRHPQNIRPL
HAAQGGGYNANRGYGGPPQQQGGGYPPHGMGNQGPRGGQFGGRGAGYSQGGPYGGP
VGGRGPNVGGGNRGPQFWSEQ

SEQ ID NO: 394
MQNMEDNVQSSWSLHGNKEICARYEILERVGSGTYSDVYRGRRKADGLIVALKEVHDYQSS
WREIEALQRLCGCPNVVRLYEWFWRENEDAVLVLEFLPSDLYSVIKSGKNKGENGIPEAEVK
AWMIQILQGLADCHANWVIHRDLKPSNLLISADGILKLADFGQARILEEPEAIYEVEYELPQEDI
VADAPGERLMEEDDSVKGVRNEGEEDSSTAVETNFGDMAETANLDLSWKNEGDMVMQGF
TSGVGTRWYRAPELLYGATIYGKEIDLWSLGCILGELLILEPLFSGTSDIDQLSRLVKVLGTPT
EENWPGCSNLPDYRKLCFPGDGSPVGLKNHVPSCSDSVFSILERLVCYDPAARLNAKEVLE
NKYFVEDPYPVLTHELRVPSPLREENNFSEDWAKWKDMEADSDLENIDEFNVVHSSDGFCI
KFS

SEQ ID NO: 395
MDLNQYPEDLNPELPEGTDNVDNPDNNKGSPVPSPHPPLKPLDPSERYRKGITLGQGTYGI
VYKAFDTVTNKTVAVKKIHLGKAKEGVNVTALREIKLLKELSHPNIIQLIDAYPHKQNLHIVFEF
METDLEAVIKDRNLVFSPADIKSYLQMTLKGLAVCHKKWVLHRDMKPNNLLIAADGQLKLGD
FGLARLFGSPDRKFTHQVFAVWYRAPELLFGAKQYGPAVDIWATGCIFAELLLRKPFLQGVS
DLDQIGKIFAAFGTPRQSQWPDVASLPDFVEFQFVPAPSLRSLFPMASEDALDLLSKMFTLD
PKNRITAQQALEHRYFSSVPAPTRPDLLPKPSKVDSSRPPKHASPDGPVVLSPSKARRVMLF
PNNLAGILPKQVSQSTTGGTPIEFDMPTQKLREVCPRSRITESGKKHLKRKTMDMSAALDEC
AREQEGQEGKTILDPDHQRSAKKEKHM

SEQ ID NO: 396
MAGGQENCVRITRARAACVSKASAPVIQSQVDEKKSRKRAPKRAAVDDLAANASGSQPKRR
AVLGDVTNLHAAATDCLSTAEDQVDAPNPSIKGRARNKKKEARTSTKVVKDEIHPESNPLAD
HSSNLSECQKPPAAKLAEQRSLRGVPSKAKQGGSSNSQSCSKHTDIDKDHTDPQMCTTYV
EDIYEYLRNAELKNRPSANFMETAQNDITPNMRAILVDWLVEVSEEYKLVPDTLYLTVSYIDR
YLSANPTSRHKLQLLGVSCMLIASKYEEVCPPHVEEFCYITDNTYTRDEMLSMERKILIFLNFE
MTKPTTKSFLRRFVRASQAGNKAPSLHMEFLANYLAELTLMECSFLQYLPSLIAASTVFLSRL
TLDFLTNPWNPTLAHYTGYKASQLKDCVMAIYNVQMNRKGSTLVAIREKYQQHKFKCVASLP
PPPFIAERFFEDTPN

SEQ ID NO: 397
MTGTQASNVRITRARAAKSTLNNALPPLPPAQGKPRGKRAATESNISGFSVAAEPLKRRAVL
SDVSNICKEAAAVDCLKKPKAVKVVSQNANAKGRGRGIPRNNKKITQEAEIKKETSPAICNVD
DASAGNAIGDDKQNNNVNPLKEVQDNPKELNPIAEQISVHPHCKQSVEKPNEKEIVVSDNKA
AIASLKQQSTLQSLRIPKQPKYSLKQGNPVPLANLHEDVGRSSCSDFIDIDSEYKDPQMCTAY
VTDIYANMRVVELKRRPLPNFMETTQRDINANMRSVLIDWLVEVSEEYKLVPDTLYLTVSYID
RFLSANVVNRQRLQLLGVSCMLVASKYEEICAPPVEEFCYITDNTYKKEEVLEMEISVLNRLQ
YDLTTPTTKTFLRRFIRAAQASCKVSSLHLEFMGNYLAELTLVEYDFLKYLPSLIAAAAVFVAR

MTLDPMVHPWNSTLQHYTGYKVSDMRDCICAIHDLQLNRKGCTLAAIREKYNQPKFKCVAN
LFPPPIISPQFLIDNEV

SEQ ID NO: 398
MAAPNQNALLINNNNRRPLVDIGNLVGALNAQCNISKNGARKRAFGDIGNLVEDLDAKCTISK
YWVRKRPRTNFGVNANKGASSSTQGQGIVVRGEQKAWDRIVWGNKQSCAIKMNAQHVTAT
QRGTAISISDIIDSSVQDGGIKAPSQLKARKQTVRTVTATLTARSEDSLRDVLEVPPGIDDGDR
DNPLAVVEYVEDIYHFYRKIEVRSCVPPDYMTRQLEIKDSMRGVIIDWLIEVHRTFLLMPETLY
LTVNIIDRYLSIQSVTRNELQLMGITAMFIASKYEEISPPKINDLVYITKDAYTSKQIVNMEHTILN
RLKFKLTVPTPYVFLVRFLKAAGPDKVMKNLAFFLVDLCLLHYKMIKYSPSMLAAAAVYTAQC
TLKKHPYWNKTLILHIGYSEAHLRECAHLMADLHLKAEGSNLKSVYKKYSYPIFGSVAFLSPA
KIPAGTVAAPAIDKCAHQIYLRNLR

SEQ ID NO: 399
MFPNKQTQGLVQNKKMASKAAQPKAMVPPQRVPPAANNRRALGDIGNIVADVGGKCNVTK
DGVNGKPLAQVSRPITRSFGAQLLAQAAANKGISAANNQTQVPVVIPKADVRGNKQRRTSKS
KDIPPTTVVTNESDDCVIIEQAQRIKPTCNHNVGAVGNKEKPQLLTAKPKSLTASLTSRSAVAL
RGFRFDDEMTEAEEDPLPNIDVGDRDNQLAVVEYVEDIYKFYRRTEQMSCVPDYMPRQQEI
NPKMRAVLINWLIEVHYRFGLMPETLYLTTNLIDRYLATQLVSRSNYQLVGATAMLLASKYEEI
WAPEMNDFLDILENKFERKHVLVMEKAMLNKLKFHLTVPTPYVFLVRFLKAAASDEEMENLV
FFLMELSLMQYVMIKFPPSMLAAAAVYTAQITLKKTTVWNDVLKRHTGYSEIDLKECTRLMVA
FHQSSEESKLNVVFKKYSMPEYDSVALIKPAKLPA

SEQ ID NO: 400
MAPSFDCVANAYIESCEDQEKLRQNAQILAQSGENDVDEPVSMLVQRETHYMLPEDYLQRL
RNRTLDVNVRREAVGWILKVHSFYNFGAPTAYLAVNYLDRFLSRHRMPQGVKAWMIQLMAV
ACLSLAAKMEETQVPLPSDLQREDARFIFDARTIQRMELLILSTLQWGMRSITPFSFIDYFAYR
AVQGHGHGHDATPKAVMSRAIELILSTTEEIDFMEYRPSAIAAAALLCAAEEVVPLQAVHYKR
ALSSSITDVDKDKMFGCYNLIQETIIEGGCYWTPMSLQSTEKTPVGVLDAAACLSNTPTSSYS
VKPYASVTAAKRRKLNEICSALLVSQAHPC

SEQ ID NO: 401
MAANFWTSSHCKELLDAEKVGIVHPLDKDQGLTQEDVKIIKINMSNCIRTLAQYVKLRQRVVA
TAITYCRRVYTRKSFTEYDPQLVAPTCLYLASKAEESTVQAKLVIFYMKKYSKHRYEIKDMLE
MEMKLLEALDYYLVIYHPYRPLIQFLQDAGLNDLKVTAWALVNDTYRTDLILTYPPYMIALACIY
FACIMEEKDAQAWFEELRVDMNEIKNISMEIVDYYDNYRVIPDEKMNSALNKLPHRF

SEQ ID NO: 402
MAPALSSSYECLSHLLCAEDASNVVGCWDEDESKIFCEEEEGFGIQHFPDFPVPDDDEIRVL
VRKESQYMPGKSYVQSYQNLGLDFTARQNAIGWILKVHGSYNFGPLTAYLSINYLDRFLSRN
PLPKAKVWMLQLLSVACLSLAAKMEETQVPLLLDLQAEEPDFLFEPRTIQRMELLVLSTLEWR
MLSVTPFSFVDYFLQGGGGRKPPPRAMVARANELIFNTHTVLDFLEHRPSAIAAAAVICAAEE
VLPLEAAQYKETILSCSLVDKEWVFGSYNLIQEVLIEKFSTPKKAKSASSSIPQSPVGVLDAFC
LSNNSNNTSLEASLSVNLYASVAAKRRKLNDYCNTWRMFQHSTC

SEQ ID NO: 403
MAPNCIDCAPSDLFCAEDAFGVVEWGDAETGSLYGDEDQLHYNLDICDQHDEHLWDDGEL
VAFAEKETLYVPNPVEKNSAEAKARQDAVDWILKVHAHYGFGPVTAVLSINYLDRFLSANQL
QQDKPWMTQLAAVACLSLAAKMDETEVPLLLDFQVEEAKYIFESRTIQRMELLVLSTLEWRM
SPVTPLSYIDHASRMIGLENHHCWIFTMRCKEILLNTLRDAKFLGLLPSVVAAAIMLHVIKETEL
VNPCEYENRLLSAMKVNKDMCERCIGLLIAPESSSLGSFSLGLKRKSSTINIPVPGSPDGVLD
ATFSCSSSSCGSGQSTPGSYDSNNSSILCISPAVIKKRKLNYEFCSDLHCLED

SEQ ID NO: 404
MPQIQYSEKYTDDTYEYRHVVLPPETAKLLPKNRLLNENEWRAIGVQQSRGWVHYAIHRPE
PHIMLFRRPLNYQQNQQQQAGAQSQPMGLKAQ

SEQ ID NO: 405
MDQIEYSEKYYDDTYEYRHVELPPDVARLLPKNRLLTENEWRGIGVQQSRGWVHYAIHCSE
PHIMLFRRPLNYEQNHQHPEPHIMLFRRPLNCQPNHQPQAHHPT

SEQ ID NO: 406
MDQIEYSEKYYDDTYEYRHVELPPDVARLLPKNRLLTENEWRGIGVQQSRGWVHYAIHCSE
PHIMLFRRPLNYEQNHQHPEPHIMLFRRPLNCQPNHQPQAHHPT

SEQ ID NO: 407
MPQIQYSEKYYDDTYEYRHVVLPPDVARLLPKNRLLNENEWRGIGVQQSRGWVHYAIHRPE
PHIMLFRRHLNYQQNQQQQAQQQPAQAMGLQA

SEQ ID NO: 408
MALVETEPVTLIHPEEPKKFKKKPTPGRGGVISHGLTEEEARVKAIAEIVGAMVEGCRKGEDV
DLNALKAAACRRYGLSRAPKLVEMIAALPDGERAAVLPKLKAKPVRTASGIAVVAVMSKPHR
CPHIATTGNICVYCPGGPDSDFEYSTQSYTGYEPTSMRAIRARYNPYVQTRSRIDQLKRLGH
TVDKVEFILMGGTFMSLPADYRDYFIRNLHDALSGHTSSNVEEAVCYSEHSATKCIGLTIETR
PDYCLGPHLRQMLSYGCTRLEIGVQSTYEDVARDTNRGHTVAAVADCFCLAKDAGFKVVAH
MMPDLPNVGVERDMESFREFFENPAFRADGLKIYPTLVIRGTGLYELWKTGRYRNYPPEQL
VDIIARVLALVPPWTRVYRVQRDIPMPLVTSGVEKGNLRELALARMDDLGLKCRDVRTREAGI
QDIHHKIRPEVVELVRRDYCANEGWETFLSYEDTRQDILVGLLRLRKCGHNTTCPELKGRCSI
VRELHVYGTAVPVHGRDADKLQHQGYGTLLMEQAERIAWKEHRSIKIAVISGVGTRHYYRKL
GYELEGPYMMKYLN

SEQ ID NO: 409
MLGFRDLYTSICEHLQRASGRLPIIAAATSLISTPEIAAVEKENKAPNSVDKMGMGSADESGR
FSTSNGQFMNMNNGVVKEEWKGGVPVVPSAPTTVPVITNVKLETPSSPDHDMARKRKLGFL
PLEVGTRVLCKWRDGKFHPVKIIERRKLPNGATNDYEYYVHYTEFNRRLDEWVKLEQLELDS
VETDADEKVDDKAGSLKMTRHQKRKIDETHVEGNEELDAASLREHEEFTKVKNITKIELGRYE
IETWYFSPFPSEYNNCEKLYFCEFCLNFMKRKEQLQRHMRKCDLKHPPGDEIYRSGTLSMF
EVDGKKNKVYAQNLCYLAKLFLDHKTLYYDVDLFLFYILCECDERGCHMVGYFSKEKHSEES
YNLACILTLPPYQRKGYGKFLISFSYELSKKEGKVGTPERPLSDLGLLSYRGYWTRVLLDILKK
HKSNISIKELSDMTAIKADDVLSTLQGLDLIQYRKGQHAICADPKVLDRHLKAVGRGGLEVDV
CKLIWTPYKEQ

SEQ ID NO: 410
MGSLDESTCSEEIRDEGKDSIRTKFKVESTVNNAQNGGNDNSKKKRAAGLPLEVGIRLLCKW
RDSKLHPVKIIERRKLPNGFPQDYEYYVHYTEFNRRLDEWVKLEQFELDSVETDADEKIEDK
GGSLKMTRHQKRKIDEIHVEEGQGHEDFDPASLREHEEFTKVKNIAKVELGRYEIETWYFSP
FPPEYSHCEKLFFCEFCLNFMKRKEQLQRHMRKCDLKHPPGDEIYRNGTLSMFEVDGKKNK
IYGQNLCYLAKLFLDHKTLYYDVDLFLFYVLCECDDRGCHVVGYFSKEKHSDEAYNLACILTL
PPYQRKGYGKFLIAFSYELSKKEGKVGTPERPLSDLGLLSYRGYWTRILLDILKKQRGNISIKE
LSDMTAIKVEDVISTLQVLDLIQYRKGQHVICADPKVLDRHLKAAGIAGLEVDVSKLIWTPYKE
QCG

SEQ ID NO: 411
MASAPMVGCDDSRDKHRWVESKVYMRKGHGKGSKGNAGFNAQNSTAQVRRENDNM
NSIADNGKSEAASEGLSSLSRKQITVNQDHPPNETSSMPAVGGLQNIDTHVTFKLEGCSKQE
IWELRKKLTNELEQVRGTFKKLEARELQLRGYSVSAGVNTSYSASQFSGNDMRNNGGKEVT
SEVASGGAITPKQAQRESNPPRQLSISLMENNQAASDMGEKGKRTPKANQYYRNSEFVLGK
DKFPPAESKKSKSTGNKKISQSKVFSKETMQVGKEFMPQKSVNEVFKQCSLLLTKLMKHKY
GWVFNLPVDAQALGLHDYHTIIKRPMDLGTVKSKLEKNLYNSPASFAEDVKLTFSNAMTYNP
KGHEVHTMAEQLLQLFEERWKTIYEEHLDGKMRFGSGQGLGASSSTKKLPFQDSKKNIKKS
EPAGGPSPPKPKSTNHHASRTPSAKKPKAKDPHKRDMTYEEKQKLSTNLQNLPQERLELIV
QIIKKRNPSLCQHDEEIEVDIDSFDTETLWELDRFVTNYKKSLSKNKKKALLADQAKRASEHG
SARNKHPMIGRELPMNNKKGEQGEKVVEIDHMPPVNPPVVEVEKDGVYAKRSSSSSSSSSD
SGSSSSDSDSGSSSGSESDAYAATSPPAGSNTSARG

SEQ ID NO: 412
MEGHSGALGFGQGFSRSSQSPNLSPSPSHSASASVTSSGQKRKRNEVEHAGVASNSTGM
FAVPPSHIYSHLHPMSMSMPMPMHNSHPSSLSESRDGALTSNDDDDNLTGGNQSQLDSMS
AGNTDGREDFDDEDDDDDDEEDDDEVEGDEEDQDHDPDADDDSDDGHDSMRTFTAARLD
NGAPNSRNLKPKADAAGVAIAPTVKTEPILDTVKEEKVSGNNNNNSVSANNAQVAPSGSAVL

LSAVKEEANKPTSTDHIQTSGAYCAREESLKREEDADRLKFVCFGNDGIDQHMIWLIGLKNIF
ARQLPNMPKEYIVRLVMDRSHKSVMIIKQNQVVGGITYRPYLSQKFGEIAFCAITADEQVKGY
GTRLMNHLKQHARDVDGLTHFLTYADNNAVGYFIKQDFTKEIKLEKERWHGYIKDYDGGILM
ECKIDPKLPYTDLPAMIRWQRQTIDEKIRELSNCHIVYSGIDIQKKEAGIPRKPIKVEDIPGLKEA
GWTTDQWGHSRFRLLNSPSEGLPNRQVLHAFMRSLHKAMVEHADAWPFKEPVDPRDVPD
YYDIIKDPMDVKRMFTNARTYNTHETIYYKCANR

SEQ ID NO: 413
MEESGNSLTSGPDGSKRRVSYFYDSDIGNYYYSQGHPMKPHRIRMAHSLIVHYALDEKMEV
CRPNLLQSRELRVFHADDYISFLQSVTPETQHEQLRQLKRFNVGEDCPVFDGLYNFCQTYA
GGSVGAAIKLNNKEADIAINWSGGLHHAKKCEASGFCYVNDIVLAILELLKVHQRVLYIDIDIHH
GDGVEEAFYSTDRVMSVSFHKFGDYFPGTGHLKDVGYGKGKYYSLNVPLNDGIDDESYKNL
FRPIIQKVMEIYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCVRFLRSFNVPLVLVGGGG
YTIRNVARCWCYETAVAVGVEPQDKLPYNEYYEYFGPDYTLHVAPSNMENQNSAKELAKIR
NTLLEQLKRIQHVPSVPFQERPPDTKFPEEDEEDYEKRPKGHKWGGEYFGSESDEEQKPQ
NRDIDISDKPGIRRQSPPNVEAAKKIKVEEEDGDIGIVNENDGAKWPLGEAG

SEQ ID NO: 414
MEESGNSLTSGPDGSKRRVSYFYDSDIGNYYYSQGHPMKPHRIRMAHSLIVHYALDEKMEV
CRPNLLQSRELRVFHADDYISFLQSVTPETQHEQLRQLKRFNVGEDCPVFDGLYNFCQTYA
GGSVGAAIKLNNKEADIAINWSGGLHHAKKCEASGFCYVNDIVLAILELLKVHQRVLYIDIDIHH
GDGVEEAFYSTDRVMSVSFHKFGDYFPGTGHLKDVGYGKGKYYSLNVPLNDGIDDESYKNL
FRPIIQKVMEIYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCVRFLRSFNVPLVLVGGGG
YTIRNVARCWCYETAVAVGVEPQDKLPYNEYYEYFGPDYTLHVAPSNMENQNSAKELAKIR
NTLLEQLKRIQHVPSVPFQERPPDTKFPEEDEEDYEKRPKGHKWGGEYFGSESDEEQKPQ
NRDIDISDKPGIRRQSPPNVEAAKKIKVEEEDGDIGIVNENDGAKWPLGEAG

SEQ ID NO: 415
MEFWGVEVKPGEALTCDPGDERYLHMSQAAIGDKEGAKENERVSLYVHVDGKKFVLGTLS
RGKCDQIGLDLVFEKEFKLSHTSQTGSVFVSGYTTVDHEALDGFPDDEDLESSEDEEEELAQ
ITTLTAKENGGKTGAKPVKPESKSSVTDKAAAKGKPEVKPPVKKQEDDSDSDEDEDEDEDE
DEDDDDEDDEDMKDASASDDGDEEDDSDEESDDDEEEDEETPKPAAGKKRPMPASDNKS
PATDKKAKITTPAGGQKPGADKGKKTEHIATPYPKHGAKGPASGVKGKETPLGSKQTPGSK
VKNSSTPESGKKSGQFKCQSCSRDFATEGALSSHNAAKHGGK

SEQ ID NO: 416
MMETGGNSLPSGPDGVKRKVAYFYDPEVGNYYYGQGHPMKPHRIRMTHALLVQYGLHKE
MQILKPYPARDRDLCRFHADDYVAFLRGITPETIQDQVKALKRFNVGDDCPVFDGLYQYCQT
YAGGSVGGAVKLNHKLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKYHKRVLYVDI
DIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDIGCGKGKYYAVNVPLDDGIDDESF
QSLFKPIIQQVMLVYNPEAIVLQCGADSLSGDRLGCFNLSVKGHAECVRYMRSFNVPLLMVG
GGGYTVRNVARCWCYETGVAVGVEIDDKMPQHEYYEYFGPDYTVHVAPSNMENKNTKQYL
DKIRSKILENINSLPCAPSAQFQVQPPDTDFPELEEEDYDERTRSHKWDGASCDSDSENGDL
KHRNHDVEESAFPRHNLANISYNTKIKLEGVGTGGLDMAAGTDTKKNDESFEAMDYESGEE
LRQDHFASTINASQPCDPALLTGVQNQLQSTDTVKPIEQSGNAPGIPPPSVATVSTGTRPSSI
SRTSSLNSMSSVKQGSILGPNPPQGLNASGLQFPVPTSNSPIRQGGSYSITVQAPDKQGLQ
NHMKGPQNMPGNS

SEQ ID NO: 417
MPPKDRVAYFYDGDVGSVYFGPNHPMKPHRLCMTHHLVLSYELHKKMEIYRPHKAYPVELA
QFHSADYVEFLHRITPDTQHLFTKELVKYNMGEDCPVFENLFEFCQIYAGGTIDAAHRLNNQI
CDIAINWSGGLHHAKKCEASGFCYINDLVLGILELLKHHARVLYVDIDVHHGDGVEEAFYFTD
RVMTVSFHKYGDMFFPGTGDVKEVGEREGKYYAINVPLKDGIDDASFTRLFKTIITKVVDIYQ
PGAIVLQCGADSLAGDRLGCFNLSIDGHAQCVRIVKKFNLPLLVTGGGGYTKENVARCWSVE
TGVLLDTELPNEIPDNDYIKYFAPDYSLKINTAGNMENLNSKTYLSAIKVQVMENLRAIQHAPS
VQMHEVPPDFYIPDIDEDELNPDERMDQHTQDRQIQRDDEYYDGDNDIDHDMEEAS

SEQ ID NO: 418
MDSSKSEEANILHVFWHEGMLNHDLGTGVFDTLEDPGFLEVLEKHPENADRVRNMLSILRK
GPIAPYTEWHTGRAAYLSELYSFHRPDYVDMLAKTSTAGGKTLCHGTRLNPGSWEAALLAA

GTTLEAMRYILDGHGKLSYALVRPPGHHAQPTQADGYCFLNNAGLAVELAVASGCKRVAVV
DIDVHYGNGTAEGFYERDDVLTISLHMNHGSWGPSHPQTGFHDEVGRGKGLGFNLNVPLP
NGTGDKGYEHAMHELVVPAISKFMPEMIVLVIGQDSSAFDPNGRECLTMEGYRKIGQIMRQQ
ADQFSGGRLVVVQEGGYHITYAAYCLHATLEGVLCLPHPLLSDPIAYYPEHDIYSERVTFIKNY
WQGIISTTDKRN

SEQ ID NO: 419
MEESGNALVSGPDGSKRRVTYFYDADIGNYYYGQGHPMKPHRMRMAHNLIVHYGLHQRME
VCRPHLAQSKDIRAFHTDDYIHFLSSVAPDTQQEQLRQLKRFNVGEDCPVFDGLFNFCQSSA
GGSIGAALKLNRKDADIAINWAGGLHHAKKCEASGFCYVNDIVLGILELLKVHQRVLYIDIDIHH
GDGVEEAFYTTDRVMTVSFHKFGDYFPGTGHIKDVGYGKGKYYALNVPLNDGIDDESYKHL
FRPIIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCVRFVRSFNIPLMLVGGG
GYTIRNVARCWCYETAVAVGVEPQDKLPYNEYYEYFGPDYTLYVAPSNMENLNTEKDLEKM
RNVLLEQLSKIQHTPSVPFQERPPDTEFNDEEEEDMEKRSKCRIWDGEYVGSEPEEDGKLP
RFDADTYERSVLKHENKRLVPVSNVEPLKRIKQEEDGAAV

SEQ ID NO: 420
MPPKDRVAYFYDGDVGSVYFGPNHPMKPHRLCMTHHLVLSYELHKKMEIYRPHKAYPVELA
QFHSADYVEFLHRITPDTQHLFTKELVKYNMGEDCPVFENLFEFCQIYAGGTIDAAHRLNNQI
CDIAINWSGGLHHAKKCEASGFCYINDLVLGILELLKHHARVLYVDIDVHHGDGVEEAFYFTD
RVMTVSFHKYGDMFFPGTGDVKEVGEREGKYYAINVPLKDGIDDASFTRLFKTIITKVVDIYQ
PGAIVLQCGADSLAGDRLGCFNLSIDGHAQCVRIVKKFNLPLLVTGGGGYTKENVARCWSVE
TGVLLDTELPNEIPDNDYIKYFAPDYSLKINTAGNMENLNSKTYLSAIKVQVMENLRAIQHAPS
VQMHEVPPDFYIPDIDEDELNPDERMDQHTQDRQIQRDDEYYDGDNDIDHDMEEAS

SEQ ID NO: 421
MDLNLVSHGEEEEGVRRRKVGIVYDERMCKHATPEDQPHPEQPDRIRVIWDKLNSAGVLHK
CVMVEAKEASEEQLAGVHSRKHIEVMKSIGTARYNKKKRDKLAASYSSIYFSQGSSEAALLA
AGSVVEISEKVASGELDAGVAIVRPPGHHAEADKAMGFCLFNNIAIAAKHLVHERPELGVQEV
LIVDWDVHHGNGTQHMFWTDPHVLYFSVHRFDAGTFYPGGDDGFYDKIGEGKGAGYNINV
PWEQGKCGDADYLAVWDHVLVPVAKSYDPDMVLISGGFDAALGDPLGGCRLTPYGYSLMT
KKLMEFAGGKIVLALEGGYNLKSLADSFLACVEALLKDGPSRSSVLTHPFGSTWRVIQAVRK
ELSSFWPALNEELQLPRLLKDASESFDKLSSSSSDESSASEDEKKIAEVTSIMEVSPDPSSILA
LTAEDIAQPLAGLKIEEAGTDSQRSSDHTLLDLTNDDTQKLKQFEGEIFVMIGDEESVPSASS
SKDQNESTVVLSKSNIKAHSWRLTFSSIYVWYASYGSNMWNPRFLCYIEGGQVEGMAKRCC
GSEDKTPPQRIQWKVVPHRMFFGRSYTNTWGSGGVSFLDPNCSDTSEAHVCLYKITLAQFN
DLLLQENNLNCGTEHPLVDLSSIDAIRNGNSILELIKDSWYGTLIYLGMEGGLPIVTFTCSVCD
VEKFKHGQLPLCPPSSRYENILIRGLVQGKKLSEDDATAYIRAASTSPLL

SEQ ID NO: 422
MADEDLDLSDVGEVEDEPGEEIESTPPLAVGQEKEINSLALKKKLLKVGTRWETPENGDEVT
VHYTGTLPDGTKFDSSRDRGEPFTFKLGQGQVIKGWDQGIVTMKKGERALFTIPPELAYGSS
GVRPTIPPNATLQFDVELLSWTNIVDVCNDGGILKRIISEGEKYERPKDPDEVTVKYEAKLED
GTLVAKSPEEGVEFYVNDGHFCPAIAKAVKTMKRGESVILTIKPTYAFGERGKDAEEGFAAIP
PNATLTTSLELVSFKAVIAVTEDKKVIKKILKEADGYDKPSDGTVVQIRYTAKLQDGTIFEKKGY
EGEEPFQFVVDEEQVIAGLDKAVETMKTGEIALITIGAEYGFGNFETQRDLAVIPPNSTLIYEV
EMISFTKEKESWDMDTTEKIEASKQKKEQGNSLFKVGKYQRAAKKYEKAAKYIEHDSSFSAE
EKKQSKVLKVSCNLNHAACRLKLKDFKEAVKLCSKVLELESQNVKALYRRAQAYIETADLDLA
EFDIKKALEIEPQNREVQLEYKILKQKQIEYNKKDAKLYGNMFAKLNKLEAFEGKVLS

SEQ ID NO: 423
MADEGLELSDVAEVEDEPGEEFESAPPLVVGQEKELNSSGLKKKLLKAGTRCETPENGDEV
TVHYTGTLLDGTKFDSSRDRGEPFTFNIGQGQVIKGWDQGIVTMKKREHALFTIPPELAYGA
SGMPPTIPPNATLQFDVELLSWTNIVDVCKDGGILKRIISDGEKYERPKDPDEVTVKYEAKLE
DGMLVAKSPEEGVEFYVNDGNFCPAIVKAVKTMKKGENVTLTIKPAYAFGEQGKDAEEGFA
AIPPNATITINLQLVSFKAVKEVTEDKKVIKKILKEADGYDKPSDGTVVQIRYTAKLQDGTIFEKK
GYAGEEPFQFVVDEEQVIAGLDKAVETMKTGEVALITIGPEYGFGNIETQRDLAVIPPYSTLIY
EVEMVSFTKEKESWDMNTTENIEASKQKKEQGNSLFKVGKYLRAAKKYDKAAKYIEHDNSF

SAEEKKQSKVLKVSCNLNHAACCLKLKDFKKAVKLCSKVLELESQNVKALYRRAQAYIETADL
DLAEFDIKKALEIEPQNREVRLEYLILKQKQIEYNKKDAKLYGNMFARQNKLEAIEGKD

SEQ ID NO: 424
MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRALCTGEKGTGRSGKPLHFKGSSF
HRVIPGFMCQGGDFTRGNGTGGESIYGEKFADENFVKKHTGPGILSMANAGPNTNGSQFFI
CTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGSGRTSKPVVIADSGQLA

SEQ ID NO: 425
MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRALCTGEKGNGRSGKPLHFKGSSF
HRVIPGFMCQGGDFTRGNGTGGESIYGEKFADENFVKKHTGPGILSMANAGPNTNGSQFFI
CTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGSGRTSKPVVIADSGQLA

SEQ ID NO: 426
MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRALCTGEKGTGRSGKPLHFKGSSF
HRVIPGFMCQGGDFTRGNGTGGESIYGEKFADENFVKKHTGPGILSMANAGPNTNGSQFFI
CTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGSGRTSKPVVIADSGQLA

SEQ ID NO: 427
MPNPKVFFDMQVGGAPAGRIVMELYADVVPKTAENFRALCTGEKGTGRSGKPLHFKGSSF
HRVIPGFMCQGGDFTRGNGTGGESIYGEKFADENFVKKHTGPGILSMANAGPNTNGSQFFI
CTAQTSWLDGKHVVFGQVVEGLEVVRDIEKVGSGSGRTSKPVVIADSGQLA

SEQ ID NO: 428
MADDFELPESAGMMENEDFGDTVFKVGEEKEIGKQGLKKLLVKEGGSWETPETGDEVEVH
YTGTLLDGTKFDSSRDRGTPFKFKLGQGQVIKGWDQGIATMKKGENAVFTIPPDLAYGESGS
QPTIPPNATLKFDVELLSWASVKDICKDGGIFKKIIKEGEKWEHPKEADEVLVKYEARLEDGTV
VSKSEEGVEFYVKDGYFCPAFAIAVKTMKKGEKVLLTVKPQYGFGHQGREAIGNDVARSTN
ATLLVDLELVSWKVVDEVTDDKKVLKKILKQGEGYERPNDGAVVKVKYTGKLEDGTIFEEKG
SDEEPFEFMAGEEQVVDGLDRAVMTMKKGEVALVSVAAEYGYQTEIKTDLAVVPPKSTLIYE
VELVSFVKEKESWDMNTAEKIEAAGKKKEEGNALFKVGKYFRASKKYEKATKYIEYDTSFSE
EEKKQSKPLKVTCNLNNAACKLKLKDYTQAEKLCTKVLEVESQNVKALYRRAQAYIQTADLE
LAELDIKKALEIDPNNRDVKLEYRALKEKQKEYNKKEAKFYGNMFARMSKLEELESRKSGSQ
KVETANKEEGSDAMAVDGESA

SEQ ID NO: 429
MAASLTPLGAGLAYATIYDQAKVRKLEPTKRSLIALCQHSDSQHRRFITRKYHVNVQILNRRD
AIRLIGLAAGLCIDLSLMYDARGAGLPPQENAKLCDTTCEKELENAPMITTESGLQYKDIKIGN
GPSPPIGFQVAANYVAMVPSGQVFDSSLDKGQPYIFRVGSGQVIKGLDEGLLSMKVGGKRR
LYIPGPLAFPKGLNSAPGRPRVAPSSPVIFDVSLEFIPGLESEEE

SEQ ID NO: 430
MSAASLSADMAIRGTILGKTALHVLGPQVVSQCRQPVMFKCPPHTLRKMRFSAQDLQSKNF
YSGFTPFKSVFISTSKRSWQAGSARAMSQDAAFQSKVTTKCFLDIEIGGDPAGRIVLGLFGE
DVPKTAENFRALCTGEKGFGYKGSSFHRIIKDFMLQGGDFDRGDGTGGKSIYGRTFEDENF
KLAHVGPGVLSMANAGPNTNGSQFFICTVKTPWLDKRHVVFGQVIEGMEIVKKLESEETNRT
DRPKRPCRIVDCGELP

SEQ ID NO: 431
MGRIKPQTLLQQSKKKKVPGRISVSTIIVCNLIIIFLMFSLVGIYRQRAKRNRATSRSDGDEEME
NFGRSKINSVPHQAIVNTTKGLITLELFGKSSAHTVEKFVEWSERGYFNGLPFYRVIKHFVIQV
GDPKFAGNREDWTVGGQLNVQLEFSPKHEAFMLGTSKLEDQGDGFELFITTAPIPDLNDKLN
VFGRVIKGQDVVQEIEEVDTDEHFQPKSPIIINDVRLKDEL

SEQ ID NO: 432
MARQSTLLLFWSLVFLGAIVFTQAKHEELEEVTHKVYFDVDIAGKPAGRVVIGLFGKAVPKTV
ENFRALCTGEKGVGKSGKPLHYKGSFFHRIIPSFMIQGGDFTLGDGRGGESIYGTKFADENF
KLKHTGPVFITTVTTDWLDGRHVVFGKIISGMDVVYKVEAEGRQSGQPKRKVKIADSGELSM
D

SEQ ID NO: 433
MARQSTLLLFWSLVFLGAIVFTQAKHEELEEVTHKVYFDVDIAGKPAGRVVIGLFGKAVPKTV
ENFRALCTGEKGVGKSGKPLHYKGSFFHRIIPSFMIQGGDFTLGDGRGGESIYGTKFADENF
KLKHTGPGFLSMANAGPDTNGSQFFITTVTTDWLDGRHVVFGKIISGMDVVYKVEAEGRQS
GQPKRKVKIADSGELSMD

SEQ ID NO: 434
MEMDEIQEQSQPQSSEKQDISQESDTGNDKTINAEKITSENAEVEEDDMLPPKVNTEVEVLH
DKVTKQIIKEGSGNKPSRNSTCFLHYRAWAESTMHKFQDTWQEQQPLELVLGREKKELSGF
AIGVAGMKAGERALLHVDWQLGYGEEGNFSFPNVPPRANLIYEAELIGFEEAKEGKARSDMT
VEERIEAADRRRQQGNELFKEDKLAEAMQQYEMALAYMGDDFMFQLFGKYKDMANAVKNP
CHLNMAQCLLKLNRYEEAIGQCNMVLAEDEKNIKALFRRGKARATLGQTDDAREDFQKVRK
FSPEDKAVIRELRLLAEHDKQVYQKQKEMFKGLFGQKPEQKPKKLHWFVVFWQWLLSMIRTI
FRMRSKTD

SEQ ID NO: 435
MAGAGEGTPEVTLETSMGPITVELYHKHAPKTCRNFLELSRRGYYNNVKFHRVIKDFMVQG
GDPTGTGRGGESIYGPRFEDEITRDLKHTGAGILSMANAGPNTNGSQFFISLAPTPWLDEKH
TIFGRVCKGMDVVKRLGNVQTDKNDRPIHDVKILRTTVKD

SEQ ID NO: 436
MMDPELMRLAQEQMSKISPDELMKMQRQIMANPDLMRMASENMKNLKPEDIRFAAEQMKN
VRKEEMAEISERISRASPEEIEAMKARANLQSAYQLQVAQNLKDQGNQLHARMKYSEAAEK
YLQARNNLTGIPFSEAKSLLLASSSNLMSCYLKTGQYEECVQTGSEVLAYDAMNVKALYRRG
QAYKQIGKLELAVADLRKAVEVSPEDETIAQALREASTELMEKGGTQDQNGPRIEEIIEEEAV
QPTAEKYPQSAPMVTSVTEDVSDDEQGSEDQNGFSRDSFQATNAPDGQMYAESLRNLTEN
PDMLRTMQSLMKNVDPDSLVALSGGKLSPDMVKTVSGMFGRMSPEEIQNMMKMSSTLSRQ
NPSTSSRFDDITRGHSNMDSSPQSVSVDNDLFEENQNRVGESSTNLSSSAAFSGMPNFSAE
MQEQVRNQMNDPATRQMFTSMIQNMSPEMMASMSEQFGVKLSPEDAVKAQNAMASLSPN
DLDRLMNWATRLQTAIDYARKIKNWILGRPGLIFAISMLLLAIILHRFGYIGD

SEQ ID NO: 437
MGVEKEILRPGNGPKPRPGQSVTVHCTGYGKNEDLSQKFWSTKDPGQKPFTFTIGQGRVIK
GWDEGVLDMQLGEIFKLRCSPDYGYGSNGFPAWGIRPNSVLVFEIEVLSVN

SEQ ID NO: 438
MPNPRCYLDITIGEELEGRILVELYSDVVPKTAENFRALCTGEKGIGPHTGVPLHYKGLPFHR
VIKGFMIQGGDISAQNGTGGESIYGLKFDDENFQLKHERRGMLSMANSGPNTNGSQFFITTT
RTSHLDGKHVVFGKVIKGMGVVRGIEHTPTESNDRPSLDVVISDCGEIPEGSDDGIANFFKD
GDLYPDWPADLDEKSAEISWWMNAVDSAKCFGNENYKKGDYKMALRKYRKALRYLDICWE
KEEIDEEKSNHLRKTKSQIFTNSSACKLKLGDLKGALLDTEFAMRDGEDNVKALFRQGQAYM
ALKDVDSAVASFKKALQLEPNDAGIRKELAVATKMINDRRDQERRAYARMFQ

SEQ ID NO: 439
MGDVIDLNGDGGVLKTIIRSAKPGAMQPTEDLPNVDVHYEGTLADTGEVFDTTREDNTLFSF
ELGKGTVIKAWDIAVKTMKVGEVARITCKPEYAYGSAGSPPDIPENATLIFEVELVACKPRKG
STFGSVSDEKARLEELKKQREIAAASKEEEKKRREEAKATAAARVQAKLEAKKGQGRGKGK
SKGK

SEQ ID NO: 440
MGLGLKIASASFLPIFNIMATRSLCILLVCFIPVLAHVLSLQDPELGTVRVYFQTTYGDIEFGFFP
HVAPKTVEHIYKLVRLGCYNSNHFFRVDKGFVAQVADVVGGREVPLNSEQRKEGEKTIVGEF
SEVKHVRGILSMGRYSDPDSASSSFSILLGNAPHLDGQYAVFGKVTKGDDTLKRLEEVPTRQ
EGIFVMPLERIRILSTYYYDTNERESNLTCDHEVSILKRRLVESAYEIEYQRRKCLP

SEQ ID NO: 441
MASKRSLRTMNVWPTLPPLVLLLLLCFSSMSSSVVAKKSDVSELQIGVKHKPKSCDIQAHKG
DRIKVHYRGSLTDGTVFDSSFERGDPIEFELGSGQVIKGWDQGLLGMCVGEKRKLRIPSKLG
YGAQGSPPKIPGGATLIFDTELVAVNGKGISNDGDSDL

SEQ ID NO: 442
MSGAPAERPISYFDITIGGKPIGRIVFSLYADLVPKTAENFRALCTGEKGIGKSGKPLCYAGSG
FHRVIKGFMCQGGDFTAGNGTGGESIYGEKFEDEAFPVKHTKPFLLSMANAGKDTNGSQFFI
TVSQTPHLDDKHVVFGEVIKGKSIVRAIENYPTASGDVPTSPIIISACGVLSPDDPSLAASEETI
GDSYEDYPEDDDSDVQNPEVALDIARKIRELGNKLFKEGQIELALKKYLKSIRYLDVHPVLPD
DSPPELKDSYDALLAPLLLNSALAALRTQPADAQTAVKNATRALERLELSDADKAKALYRRAS
AHVILKQEDEAEEDLVAASQLSPEDMAISSKLKEVKDEKKKKREKEKKAFKKMFSS

SEQ ID NO: 443
MASSLRSSLFSSWALDSKSVCSLFNLNPGKMGLPSISTPLNWRTCCCSHSSELLELNEGLQS
SRRKTVMGLSTVIALSLVYCDEVGAVSTSKRALRSQKVPEDEYTTLPNGLKYYDLKVGSGTE
AVKGSRVAVHYVAKWKGITFMTSRQGMGITGGTPYGFDVGASERGAVLKGLDLGVQGMRV
GGQRILIVPPELAYGNTGIQEIPPNATLEFDVELISIKQSPFGSSVKIVEG

SEQ ID NO: 444
MGAIEDEEPPLKRLKVSSPGLRRGLEEEAPSLSVGSVSILMAKSLSLEEGETVGSKGLIRRVE
FVRIITQALYSLGYQKAGALLEEESGILLQSSNVALFRKQILDGKWDESVVTLRGIDQVEVEGN
TLKAASFLILQQKFFELLDKGNIPEAMKTLRLEISPMQLNTKRVHELASCIVFPSRCEELGYSK
QGNPKSSQRMKVLQEIQQLLPPSIMIPEKRLERLVEQALNVQREACIFHNSLDPALSLYTDHQ
CGRDQIPTTTLQVLESHKNEVWFLQFSNNGKYLASASKDCSAIIWEITEGDSFSMKHRLSAH
QKPVSFVAWSPDDKLLLTCGIEEVVKLWNVETGECKLTYDKANSGFTSCGWFPDGERFISG
GVDKCIYIWDLEGKELDSWKGQGMPKISDLAVTSDGKEIISICGDNAIVMYNLDTKTERLIEEE
SGITSLCVSKDSRFLLLNLANQEIHLWDIGARSKLLLKYKGHRQGRYVIRSCFGGSDLAFVVS
GSEDSQVYIWHRGNGELLAVLPGHSGTVNCVSWNPVNPHVFASASDDYTIRIWGVNRNTFR
SKNASSSNGVVHLANGGP

SEQ ID NO: 445
MPGTTAGAGIEPIEPQSLKKLSLKSLKRSFDLFASLHGEPQPPDQRSQRIRIACKVRAEYEVV
KNLPTLPQREVGSSVSNSNVGETHSSLTTNQAQGFPTDTSGDLSKDEGKEITSIAVHLQPQT
GLIDGKAGAIAGTSTAISSVGSSDRYQPSAAIMKRLPSKWPRPIWHPPWKNYRVISGHLGWV
RSVAFDPGNEWFCTGSADRTIKIWEVATGKLKLTLTGHIEQIRGLAVSSRHPYLFSAGDDKQ
VKCWDLEYNKAIRSYHGHLSGVYCLALHPTLDILCTGGRDSVCRVWDIRTKAQIFALSGHEN
TVCSVFTQAIDPQVVTGSHDTTIKLWDLAAGKTMSTLTYHKKSVRAIAKHPFEHTFASASADN
IKKFKLPKGEFLHNMLSQQKTIVNAMAINEDNVLVSAGDNGSLWFWDWKSGHNFQQAQTIV
QPGSLDSEAGIYALQYDITGSRLVSCEADKTIKMWKEDETATPESHPINFKAPKDIRRF

SEQ ID NO: 446
MRPILMKGHERPLTFLKYNRDGDLLFSCAKDHTPTVWYGHNGERLGTYRGHNGAVWCCDV
SRDSTRLITSSADQTAKLWNVETGAQLFSFNFESPARAVDLAIGDKLVVITTDPFMELPSAIHI
KRIEKDLSKQTADSVLTITGIKGRINRAVWGPLNSTIISGGEDSVVRIWDSETGKLLRESDKET
GHQKPITSLCKSADGSHFLTGSLDKSARLWDIRTLTLIKTYVTERPVNAVAISPLLDHVVIGGG
QEASHVTTTDRRAGKFEAKFFHKILEEEIGGVKGHFGPINSLAFNPDGRSFASGGEDGYVRL
HHFDPDYFHIKM

SEQ ID NO: 447
MRPILMKGHERPLTFLKYNRDGDLLFSCAKDHTPTVWYGHNGERLGTYRGHNGAVWCCDV
SRDSTRLITSSADQTAKLWNVETGNQLFSFNFESPARAVDLAIGDKLVVITTDPFMELPSAIHI
KRIEKDLSKQTADSVLTITGIKGRINRAVWGPLNSTIISGGEDSVVRIWDSETGKLLRESDKET
GHQKAITSLCKSADGSHFLTGSLDKSARLWDIRTLTLIKTYVTERPVNAVAISPLLDHVVIGGG
QEASHVTTTDRRAGKFEAKFFHKILEEEIGGVKGHFGPINSLAFNPDGRSFASGGEDGYVRL
HHFDPDYFHIKM

SEQ ID NO: 448
MAENNVGDFIPLDRQEYPSKPAPGAVDSSFWKSFKKKEVSRQIAGVTCINFCPEPPHDFAVT
SSTRVHIYDGKSCELKKTITKFKDVAYSGVFRSDGQIIAAGGETGVIQVFNAKSQMVLRQLKG
HGRPVRVVRYSPQDKLHLLSGGDDSMVKWWDITTQEELLNLEGHKDYVRCGAASPSSVNL
WATGSYDHTVRLWDLRNSKTVLQLKHGKPLEDVLFFPSGGLLATAGGNVVKVWDILGGGRP
IHTMETHQKTVMAMCISKVPRSGQALGDAPSRLVTASLDGYMKVFDLDHFKVTHSARYPAPI
LSMGISSLCRTMAVGTSSGLLFIRQRKGQIEDKIHSDSSGLQVNPVNDEKDSAVLKPNQYRY
YLRGRSEKPSEGDYVVKRMAKVYFQEYDKDLRHFNHSKALVSALKAADSKGTVAVIEELVAR

KRLIQTLSILNLDELELLINFLSRFILVPKYSRFLISLTDRVLDARAVDLGKSENLKKQIADLKGIV
VQELRVQQSMQELQGIIEPLIRASAR

SEQ ID NO: 449
MDVETSGKPTGNKRTYTRLPRQVCVFWQEGRCTRESCNFLHVDEPGSVKRGGATNGFAP
KRSYNGSDERDTLAAGPPGGSRRNISARWGRGRGGIFISDERQKIRNKVCNYWLAGNCQR
GEECKYLHSFVMGSDVKFLTQLSGHVKAIRGIAFPSDSGKLYSGGQDKKVIVWDCQTGQGT
DIPLNDEVGCLMSEGPWIFVGLPNAVKAWNILTSTELSLVGPRGQVHALAVGNGMLFAGTHD
GSILAWKFSPASNTFEPAASLVGHTQAVVSLVSGADRLYSGSMDKTIRVWDLGTFQCLQTLR
DHTSVVMSLLCWDQFLLSCSLDNTVKVWVATSSGALEVTYTHNEEHGVLALCGMNDEQAK
PVLLCSCNDNTVRLYDLPSFSERGRIFSRNEVRTFQIAPGGLFFTGDATGELKVVWNWATQKS

SEQ ID NO: 450
MSVQELRERHAAATAKVNALRERIKAKRLQLLDTDVATYASSNGRTPISFSFTDLVCCRTLQ
GHTGKVYSLDWTSEKNRIVSASQDGRLIVWNALTSQKTHAIKLPCAWVMTCAFSPSGQAVA
CGGLDSVCSIFQLNNQLDRDGHLPVSRILSGHRSYVSSCQYVPDGDTHVITGSGDRTCIQW
DVTTGQRIAIFGGEFPLGHTADVMSVSISAANPKEFVSGSCDTTTRLWDTRIASRAIRTFHGH
EADVNTVKFFPDGLRFGSGSDDGTCRLFDIRTGHQLQVYRQPPRENQSPTVTAIAFSFSGRL
LFAGYSNGDCFVWDTILEKVVLNLGELQNTHNGRISCLGLSADGSALCTGSWDKNLKIWAFG
GHRKIV

SEQ ID NO: 451
MKVKIISRSTDEFTRERSNDLQRVFRNFDPNLHTQARAQEYVRALNAAKLDKIFAKPFLAAMS
GHIDGISAMAKSPRHLKSIFSGSVDGDIRLWDIAARRTVQQFPGHRGAVRGLTVSTEGGRLIS
CGDDCTVRLWDIPVAGIGESSYGSENVQKPLATYVGKNSFRAVDYQWDSNVFATGGAQVDI
WDHDRSEPTNSFAWGSDTVISVRFNPAEKDIFATTASDRSIVLYDLRMASPLNKLIMQTRNN
AIAWNPREPMNFTAANEDCNCYSYDMRRMNISTCVHQDHVSAVMDIDYSPSGREFVTGSY
DRTVRIFPYNAGHSREIYHTKRMQRVFCVKFSGDATYVVSGSDDANIRLWKAKASEQLGVLL
PRERKRHEYLDAVKERFKHLPEIKRIERHRHLPKPIYKAALLRHTVNAAAKRKEERKRAHSAP
GSVVTNPLRKKRIVAQLE

SEQ ID NO: 452
MDHYYQDDFDYLVDDEMVDFADDVEDDVRTRRRSDIDSDSENDFDLNNKSPDTTALQAKR
GKDIQGIPWNRLNFTREKYRETRLQQYKNYENLPRPRRSRNLDKECTNFERGSSFYDFRHN
TRSVKATIVHFQLRNLVWATSKHNVYLMQNYSIMHWSSLKQKGEEVLNVAGPIVPSVKHPGS
SPQGLTRVQVSAMSVKDNLVVAGGFQGELICKYLDKPGVSFCTKISHDENGITNAVEIYNDA
SGATRLMTANNDLAVRVFDTEKFTVLERFSFPWSVNHTSVSPDGKLVAVLGDNADCLLADC
KTGKTVGTLRGHLDYSFAAAWHPDGYILATGNQDTTCRLWDVRKLSSSLAVLKGRMGAIRSI
RFSSDGRFMAMAEPADFVHLYDTRQNYTKSQEIDLFGEIAGISFSPDTEAFFVGVADRTYGS
LLEFNRRRMNYYLDSIL

SEQ ID NO: 453
MAEALVLRGTMEGHTDAVTAIATPIDNSDMIVSSSRDKSILLWNLTKEPEKYGVPRRRLTGHS
HFVQDVVISSDGQFALSGSWDSELRLWDLNTGLTTRRFVGHTKDVLSVAFSIDNRQIVSASR
DRTIKLWNTLGECKYTIQPDAEGHSNWISCVRFSPSATNPTIVSCSWDRTVKVWNLTNCKLR
NTLVGHGGYVNTAAVSPDGSLCASGGKDGVTMLWDLAEGKRLYSLDAGDIIYALCFSPNRY
WLCAATQQCVKIWDLESKSIVADLRPDFIPNKKAQIPYCTSLSWSADGSTLFSGYTDGKIRV
WGIGHV

SEQ ID NO: 454
MAAIKSTSRSASVAFAPDAPLLAAGTMAGAIDLSFSSLANLEIFKLDFQSDDPELPVVGECPS
NERLNRLSWGSAGGSFGIIAGGLVDGTINIWNPATLINSEDNGDALIARLEQHTGPVRGLEFN
TISTNLLASGAEDGELCIWDLANPTAPTHFPPLKGVGSGAQGEISFLAWNRKVQHILASTSYS
GTTVVWDLRRQKPIISFPDATRRRCSVLQWNPDASTQLIVASDDDNSPTLRAWDLRNTISPY
KEFVGHSRGVIAMSWCPSDSLFLLTCAKDNRTLCWDTGSGEIVCELPAGANWNFDVQWSP
KIPGILSTSSFDGKIGIHNIEACSRNVSGEVEFGGAIVRGGPSALLKAPKWLERPAGVSFGFG
GKLASFRPSTVAQAADHRHSEVFIHNLVTEDNLVIRSTEFEAAIADGEKVSLRALCDRKAEES
QSDEEKETWNFLRVMFEDEGTARTKLLEHLGFKVQSEENGDLQETHSSKIDDIGSEIGKTLTL
DDKTEEDVLPQLKGGQDAAIPQDNGEDFFDNLHSPKEEVSLSHVGNDFVGEKDKDMVVNG
AEIEHETEDLTEYSDWNEAIQHSLVVGDYKGAVLQCLSANRMADALIIAHLGGNSLWEKTRD

EYLKKAKSSYLKVVSAMVNNDLTGLVNSRPLKSWKETLAMLCTYSQREEWTVLCDMLASRLI
AAGNVMAATLCYICAGNIEKTVEIWSRSLKYDYDGRSFVDHLQDVMEKTVVLALATGQKRVS
PSLSKLVENYAELLASQGLLTTAMEYLKLLGTEESSHELSILRDRLYLSGTDNKVEASSFPFET
RQDLTESQYNMHQTGFGAPETQKNYQENVHQVLPSGSYTDNYQPTANTHYIAGYQPAPQQ
QPSFQNYFTPASYQPAPSPNVFYPSQVSQAEQSNFAPPVNQPPMKTFVPSTPPILRNVDQY
QTPSLNPQLYQGVSSATVETHPYQTGAPASVSVGTTPGQPSVVPNFMVPGPVTAPTVTPRG
FMPVTTPTQHPLGSANPPVQPQSPQSSQVQSVTAATTPPPTIQNVDTSNVAAEIRPVIGTLR
RLYDETSEALGGARANPAKRREIEDNSRKIGSLFAKLNSGDISSNAASKLVHLCQALESRDYA
TAFQIQVGLTTSDWDECSFWLAALKRMIKVKQNMR

SEQ ID NO: 455
MAGAADSQLQTLSERDSTPNFKNLHTREYAAHKKKVHSVAWNCTGTKLASGSVDQTARVW
NIEPHGHSKTKDLELKGHADSVDQLCWDPKHSELLATASGDRTVRLWDARSGKCSQQVEL
SGENINITFKPDGTHIAVGNRDDELTIIDVRKFKPLHKRKFSYEVNEIAWNTTGELFFLTTGNG
TVEVLSYPSLQVLHTLVAHTAGCYCIAIDPIGRYFAVGSADALVSLWDLSEMLCVRTFTKLEW
PVRTISFNHDGQYIASASEDLFIDIADVQTGRTVHQISCRAAMNSVEWNPKYNLLAFAGDDKN
KYMQDEGVFRVFGFETP

SEQ ID NO: 456
MAATSPVGAGSGRELANPPTDGISNLRFSNHSDHLLVSSWDRKVRLYDASANSLKGQFVHG
GPVLDCCFHDDASGFSGSADNTVRRYDFSTRKEDILGRHEAPVRCVEYSYAAGQVITGSWD
KTLKCWDPRGASGQEKTLVGTYSQLERVYSMSLVGHRLVVATAGRHINVYDLRNMSQPEQ
RRESSLKYQTRCVRCYPNGTGFALSSVEGRVAMEFFDLSEAGQAKKYAFKCHRKSEAGRD
TVYPVNAIAFHPIYGTFATGGCDGYVNVWDGNNKKRLYQYSKYPTSIAALSFSRDGRLLAVA
SSYTFEEGEKPHEPDAVFVRSVNEAEVKPKPKVYAAPP

SEQ ID NO: 457
MASDDEEGFKNEEAPGVVDEAEVQEGLRACFPLSFGKQEKKQAPLESIHSATKRPEDPRPR
RQLGPPRPPPSILAEQEDSDRFVGPPRPPQFVRDDNDDGEAEIMIGPPRPPAQYSDDHDNE
ETIGPPKPSYLEKGEETDQMVGPSKRGSDDETSGDSDDGDDAVDFRVPLSNEIVLRGHTKV
VSALAIDQTGSRVLTGSYDYSVRMYDFQGMTSQLKSFRQLEPAEGHQVRSLSWSPTSDRFL
CVTGSAQAKIFDRDGLTLGEFVKGDMYLRDLKNTKGHISGLTCGEWHPKEKQTILTCSEDGS
LRIWDVNDFNTQKQVIKPKLAKPGRVPVTACAWGRDGKCIAGGVGDGSIQVWNLKPGWGS
RPDLYVAKGHDDDITGLQFSADGNILLTRSTDETLKVWDLRKAITPLQVFRDLPNNYAQTNVA
FSPDERLIFTGTSVERDGNSGGLLCFYDRQTLELVLRIGVSPVHSVVRCTWHPRHNQVFATV
GDKKEGGAHILYDPALSERGALVCVARAPRKKSLDDFEAKPVIHNPHALPLFRDEPSRKRQR
EKARMDPMKSQRPDLPVTGPGFGGRVGSTKGSLLTQYLLKEGGLIKETWMEEDPREAILKY
ADVAAKDPKFIAPAYAQTQPETVFAETDSEEEQK

SEQ ID NO: 458
MKERGQSHAGQPSVDERYTQWKSLVPVLYDWLANHNLVWPSLSCRWGPQMHQATYKNS
QRLYLSEQTDGTVPNTLVIATCEVVKPRVAAAEHISQFNEEARSPFVKKFKTIIHPGEVNRIRE
LPQNSKIVATHTDGPDVLIWDVDTQPNRQATLGAADSRPDLVLTGHKDNAEFALAMSPSAPF
VLSGGKDKCVLLWSIQDHISAATEPSSAKASKTPSSAHGEKVPKIPSIGPRGVYKGHKDTVED
VQFCPSNAQEFCSVGDDSALILWDARNGNEPVIKVEKAHNADLHCVDWNPHDENLILTGSA
DNSVRMFDRRNLTSSGVGSPVHKFEGHSAPVLCVQWCPDKASVFGSAAEDSYLNVWDYE
KVGKNVGKKTPPGLFFQHAGHRDKVVDFHWNSFDPWTIVSVSDDGESTGGGGTLQIWRMS
DLIYRPEDEVLAELERFRAHILSCQNK

SEQ ID NO: 459
MSSLSRELVFLILQFLDEEKFKESVHKLEQESGFFFNMKYFDEKAQAGEWDEVERYLSGFTK
VDDNRYSMKIFFEIRKQKYLEALDRQDRAKAVDILVKDLKVFSTFNEELYKEITQLLTLDNFRE
NEQLSKYGDTKSARTIMMSELKKLIEANPLFREKLIYPNLKASRLRTLINQSLNWQHQLCKNP
RPNPDIKTLFTDHACGPPNGARTPTQPTASLGVLPKATTFTPIGPHGPFPSSSTATSGLASW
MSNPNMVTSPQAPVAVGPSVPVPPNQATLLKRPRTPPGSSSVVDYQTADSEQLIKRLRPVS
QSIDEATYPGPTLRVPWSTDDLPKTLARALNEPYPVTSIDFHPSQQTFLLVGTKNGEITLWEV
GSREKLATRSFKIWDNANCSNHLEAAFVKDSSVSINRVLWSPDGTLIGIAFTKHLVHTYTFQG
LDLRQHLEIDAHVGGVNDLAFSHPNKQLCVVTCGDDKMIKVWDAVTGRKLYNFEGHDAPVY
SVCPHHKENIQFIFSTAVDGKIKAWLYDHLGSRVDYDAPGHSCTTMMYSADGTRLFSCGTSK
EGESFLVEWNESEGAIKRTYSGLRKKGSGVVQFDTTQNHFLAVGDEHLIKFWDMDSTNMLT

SCDAEGGLLNLPRLRFNKEGSLLAVTTVNGIKILANADGQKLLKTMENRTFDLPSRAHIDAAS
ATSSPATGRMERIERTSSANTVSGINGVDPAQSSEKLRLSDDLSEKTKIWKLTEITDSIQCRCI
TLPENAAEPASKVSRLLYTNSGVGLLALGSNAVHKLWKWNRSEQNPSGKATASVHPQRWQ
PTSGLLMTNDITDINPEEAVPCIALSKNDSYVMSASGGKVSLFNMMTFKVMTTFMPPPPAST
FLAFHPQDNNIIAIGMEDSTIHIYNVRVDEVKTKLKGHQKRITGLAFSSTQNILVSSGADAQLCV
WNTETWEKRKSKTIQMPVGKTVSGDTRVQFHSDQLHILVVHETQLAIYDAYKLERQYQWVP
QDALSAPILYATYSCNRQLIYATFSDGNIGVYDAEILRPRCRIAPTTYLSSGTSSSTSLPLVVAA
HPHEPNQFAIGLSDGAVQVLEPSESEGKWGVSPPPENGVVPAVVAGPSTSNQGSEQAPR

SEQ ID NO: 460
MAKDEEEFRGEMEERLVNEEYKIWKKNTPFLYDLVITHALEWPSLTVQWLPDREEPPGKDY
SVQKMILGTHTSDNEPNYLMLAQVQLPLEDAENDARQYDDERGEIGGFGCANGKVQVIQQI
NHDGEVNRARYMPQNPFIIATKTVSAEVYVFDYSKHPSKPPQDGGCHPDLRLRGHNTEGYG
LSWSPFKHGHLLSGSDDAQICLWDINVPAKNKVLEAQQIFKVHEGVVEDVAWHLRHEYLFG
SVGDDRHLLIWDLRTSATNKPLHSVVAHQGEVNCLAFNPFNEWVLATGSADRTVKLFDLRKI
SSALHTFSCHKEEVFQIGWSPKNETILASCSADRRLMVWDLSRIDEFQTPEDALDGPPELLFI
HGGHTSKISDFSWNPCEDWVIASVAEDNILQIWQMAENIYHDEEDDMPPEEVV

SEQ ID NO: 461
MSPGVKQTGSQKFESGHQDVVHDVTMDYYGKRIATCSADRTIKLFGLNASDTPSLLASLTG
HEGPVWQVAWAHPKFGSMLASCSYDGRVIIWREGQQENEWSQVQVFKEHEASVNSISWA
PNELGLCLACGSSDGSITVFTCREDGSWDKTKIDQAHQVGVTAVSWAPASAPGSLVGQPSD
PIQKLVSGGCDNTAKVWKFYNGSWKLDCFPPLQMHTDWVRDVAWAPNLGLPKSTIASCSQ
DGKVVIWTQGKEGDKWEGRILNDFKIPVWRVNWSLTGNILAVADGNNSVTLWKEAVDGDW
NQVTTVQ

SEQ ID NO: 462
MSSGVKQTGSQKFESGHQDVVHDVTMDYYGKRIATCSADRTIKLFGMNTSDTPTLLASLTG
HEGPVWQVAWAHPKFGSMLASCSYDRRVIIWREGQQENEWSQVQVFKEHEASVNSISWA
PHELGLCLACGSSDGSITVFTGREDGSWDKTKIDQAHQVGVTAVSWAPASAPGSLVGQPSD
PVQKLVSGGCDNTAKVWKFYNGSWKLDCFPPLQMHTDWVRDVAWAPNLGLPKSTIASCSQ
DGRVVIWTQGKEGDKWEGKILNDFKTPVWRISWSLTGNILAVADGNNNVTLWKEAVDGEW
NQVTTVQ

SEQ ID NO: 463
MKKRSRPSNGHLSTAAKNKSRKTAPITKDPFFDSAHNRNKSKGKGKSRGKGEEIFSSDEDD
DAIGRDAPAEEEEEIAEEERETADEKRLRVAKAYLDKIRAITKANEEDNEEEAGEDEETEAER
RGKRDSLVAEILQQEQLEESGRVQRQLASRVVTPSKLVECRVVKRHKQSVTAVALTEDDLR
GFSASKDGTIIHWDVETGASEKYEWPSQAVSVSSSNEVSKTQKGKGSKKQGSKHVLSMAV
SSDGRYLATGGLDRYIHLWDTRTQKHIQAFRGHRGAVSCLAFRQGTQQLISGSFDRTIKLWS
AEDRAYMDTLYGHQSEILAVDCLRKERVLSVGRDHTLRLWKVPEETQLVFRGHAASLECCC
FINNEDFLSGSDDGSIELWSMLRKKPVFMAKNAHGHAIVENLSEDTSTREEPDEEVTTRQLP
NGNSIGNGMTNQMGITPSVESWVGAVTVCRGTDLAASGAGNGVVRLWAIENSSKSLRALH
DIPLTGFVNSLTFARSGRFLIAGVGQEPRLGRWGRIQAARNGVTLCPIELS

SEQ ID NO: 464
MAATFGTINTATSPHNPNKSFEIVQPPNDSISSLSFSPKANYLVATSWDNQVRCWEVLQTGA
SMPKAAMSHDQPVLCSTWKDDGTAVFSAGCDKQAKMWPLLTGGQPVTVAMHDAPIKDIAW
IPEMNLLATGSWDKTLKYWDTRQSNPVHTQQLPERCFALSVRHPLMVVGTADRNLIIFNLQN
PQTEFKRISSPLKYQTRCVAAFPDKQGFLVGSIEGRVGVHHVEEAQQSKNFTFKCHRDSNDI
YAVNSLNFHPVHQTFATAGSDGAFNFWDKDSKQRLKAMARSNQPIPCSTFNSDGSLYAYAV
SYDWSKGAENHNPATAKHHILLHVPQESEIKGKPRVTTSGRK

SEQ ID NO: 465
MVVMDKGTHQTNEDESESEFIDEDDVIDEISIDEEDLPDADVEGEDVQEDNKRSEPDENSSS
LDDAIHTFEGHEDTLFAVACSPVDATWVASGGGDDKAFMWRIGHATPFFELKGHTDSVVAL
SFSNDGLLLASGGLDGVVRIWDASTGNLIHVLDGPGGGIEWVRWHPKGHLVLAGSEDYSTW
MWNADLGKCLSVYTGHCESVTCGDFTPDGKAICTGSADGSLRVWNPQTQESKLTVKGYPY
HTEGLTCLSISSDSTLVVSGSTDGSVHVVNIKNGKVVASLVGHSGSIECVRFSPSLTWVATG

GMDKKLMIWELQSSSLRCTCQHEEGVMRLSWSLSSQHIITSSLDGIVRLWDSRSGVCERVF
EGHNDSIQDMVVTVDQRFILTGSDDTTAKVFEIGAF

SEQ ID NO: 466
MPVFRTAFNGYAVKFSPFVETRLAVATAQNFGIIGNGRQHVLELTPNGIVEVCAFDSSDGLY
DCTWSEANENLVVSASGDGSVKIWDIALPPVANPIRSLEEHAREVYSVDWNLVRKDCFLSAS
WDDTIRLWTIDRPQSMRLFKEHTYCIYAAVWNPRHADVFASASGDCTVRIWDVREPNATIIIP
AHEHEILSCDWNKYNDCMLVTGSVDKLIKVWDIRTYRTPMTVLEGHTYAIRRVKFSPHQESLI
ASCSYDMTTCMWDYRAPEDALLARYDHHTEFAVGIDISVLVEGLLASTGWDETVYVWQHG
MDPRAC

SEQ ID NO: 467
MDSRNRRSRLNLPPGMSPSSLHLETTAGSPGLSRVNSSPSTPSPSRTTTYSDRFIPSRTGS
RLNGFALIDKQPQPLPSPTRSAAEGRDDASSSSASAYSTLLRNELFGEDVVGPATPATPEKS
TGLYGGSRDSIKSPMSPSRNLFRFKNDHGGNSPGSPYSASTVGSEGLFSSNVGTPPKPARK
ITRSPYKVLDAPALQDDFYLNLVDWSSNNVLAVGLGTCVYLWSACTSKVTKLCDLGVNDSVC
SVGWTPQGTHLAVGTNIGEVQIWDTSRCKKVRTMGGHCTRAGALAWSSYILSSGSRDRNIL
HRDIRVQDDFIRKLVGHKSEVCGLKWSYDDRELASGGNDNQLLVWNQQSAQPLLRFNEHT
AAVKAIAWSPHQHGILASGGGTADRCLRFWNTATDTRLNCVDTGSQVCNLVWCKNVNELV
STHGYSQNQIMVWRYPSMSKLATLTGHTLRVLYLAISPDGQTIVTGAGDETLRFWSIFPSPKS
QSAVHDSGLWSLGRTHIR

SEQ ID NO: 468
MEKKKVVVPIVCHGHSRPIVDLFYSPVTPDGLFLISASKDSSTMLRNGETGDWIGTFEGHKG
AVWSCCLDNRALRAASGSADFSAKIWDALTGDELHCFVHKHIVRACAFSESTSLLLTGGHEK
ILRIFDLNRPDAPPKEVDNSPGSIRTVAWLHSDQTILSSNSDAGGVRLWDLRTEKIVRVLETK
SPVTSAEVSQDGRYITTADGNSVKFWDANHFGMVKSYTMPCMVESASLEPTMGNMFVAGG
EDMWVRLFDFHTGEEIACNKGHHGPVHCVRFAPGGESYSSGSEDGTIRIWQTLNMNSEEN
ESYGVNGLSGKVRVGVDDVVQKVEGFQITADGHLNDKPEKPNP

SEQ ID NO: 469
MERYSQGTQKKSEIYTYEAPWQIYGMNWSVRKDKKFRLGIGSFLEEYNNRVEIIELDEESGE
FKSDPRLAFDHPYPTTKIMFVPDKECQRPDLLATTGDYLRIWQVCEDRVEPKSLLNNNKNSE
FCAPLTSFDWNDADPKRIGTSSIDTTCTIWDIEKEVVDTQLIAHDKEVYDIAWGEVGVFASVS
ADGSVRVFDLRDKEHSTIIYESSQPETPLLRLGWNKQDPRFIATILMDSCKVVILDIRFPTLPVA
ELQRHQASVNTIAWAPHSPCHICTAGDDSQALIWELSSVSQPLVEGGGLDPILAYTAAAEINQ
LQWSSMQPDWVAIAFSNEVQILRV

SEQ ID NO: 470
MQSENNLDESLHLREVQELQGHTDTVWAVAWNPVTGIDGAPSMLASCSGDKTVRIWENTH
TLNSTSPSWACKAVLEETHTRTVRSCAWSPNGKLLATASFDATTAIWENVGGEFECIASLEG
HENEVKSVSWSASGMLLATCGRDKSVWIWDVQPGNEFECVSVLQGHTQDVKMVQWHPNR
DILVSASYDNSIKVWAEDGDGDDWACMQTLGNSVSGHTSTVWAVSFNSSGDRMVSCSDDL
TLMVWDTSINPAERSGNAGPWKHLCTISGYHDRTIFSVHWSRSGLIASGASDDCIRLFSEST
DDSVTPVDGTSYKLILKKEKAHSMDVNSVQWHPSEPQLLASASDDGRIKIWEVTRINGLANS
H

SEQ ID NO: 471
MKRAYKLQEFVAHASNVNCLKIGKKSSRVLVTGGEDHKVNMWAIGKPNAILSLSGHSSAVES
VTFDSAEALVVAGAASGTIKLWDLEEAKIVRTLTGHRSNCISVDFHPFGEFFASGSLDTNLKI
WDIRRKGCIHTYKGHTRGVNSIRFSPDGRWVVSGGEDNIVKLWDLTAGKLMHDFKCHEGQI
QCMDFHPQEFLLATGSADRTVKFWDLETFELIGSAGPETTGVRAMIFNPDGRTLLTGLHESL
KVFSWEPLRCYDAVDVGWSKLADLNIHEGKLLGCSYNQSCVGVWVVDISRVGPYAAGNVS
RTNGHNEAKLASSGHPSVQQLDNNLKTNMARLSLSHSTESGIKEPKTTTSLTTTEGLSSTPQ
RAGIAFSSKNLPASSGPPSYVSTPKKNSTSRVQPTTNFQTLSRPDIVPVIVPRSNSLRPETTS
DVKKEMNNFGRVVPSTVSTKSTDVIKSGSNRDESDKIDSINQKRMTGNDKTDLNIARAEQHV
SSRLDNTNTSSVVCDGNQPAARWIGAAKFRRNSPVDPVVSPHDRSPTFPWSATDDGVTCQ
PDRQVTAPELSKRVVEPGRARALVASWETREKALTADTPVLVSGRPPTSPGVDMNSFIPRG
SHGTSESDLTVSDDNSAIEELMQQHNAFTSILQARLTKLQVIRRFWQRNDLKGAIDATGKMG

DHSVSADVISVLIERSEIFTLDICTVILPLLTRLLQSETDRHLTVAMETLLVLVKTFGDVIRATISA
TPTIGVDLQAEQRLERCNLCYVELENIKQILVPLIRRGGAVAKSAQELSLALQEV

SEQ ID NO: 472
MSTLEIEARDVIKIVLQFCKENSLHQTFQTLQNECQVSLNTVDSLETFVADINSGRWDVILPQV
AQLKLPRKKLEDLYEQIVLEMIELRELDTARAILRQTQAMGFMKQEQPERYLRLEHLLVRTYF
DPREAYHESSKEKRRSQIAQALASEVTVVPPSRLMALIGQSLKWQQHQGLLPPGTQFDLFR
GTAAVKADEEEMYPTTLAHTIKFGKQSHPECARFSPDGQYLVSCSVDGFIEVWDYISGKLKK
DLQYQADDSFMMHDDAVLCVDFSRDSEMLASGSQDGKIKVWRIRTGQCLRRLERAHSQGV
TSLSFSRDGSQLLSTSFDSTARIHGLKSGKALKEFRGHTSYVNDAIFTSDGGRVITASSDCTV
KVWDVKTTDCIQTFKPPPPLKGGDVSVNSVHLFPKNSEHIVVCNKASSIYIMTLQGQVVKSFS
SGKREGGDFVAACISPKGEWIYCVGEDRNIYCFSQQSGKLEHLMKAHDKDIIGVTPHPHRNL
LVTYSEDSTMKIWKP

SEQ ID NO: 473
MDIELEDQPFDLDFHPSAPIVAVALITGRLQLFRYVDISSEPERLWTVTAHTESCRAARFINAG
SSVLTASPDCSILATNVETGQPVARLDNAHGAAINCLTNLTESTIASGDENGIIKVWDTRQNS
CCNKFKAHEDYISDMEFVPDTMQLLGTSGDGTLSVCNLRKNKVHARSEFSEDELLSVALMK
NGKKVVCGSQEGVLLLYSWGYFKDCSDRFVGHPHSVDALLKLDEDTVLTGSSDGIIRVVSIL
PNKMIGVIGEHSSYPIERLAFSHDRNVLGSASHDQILKLWDIHYLHEDDEPETNKQEAVNDEN
VDMDLDVDTEKRPRGSKRKKRAEKGQTSSQKQSSDFFADI

SEQ ID NO: 474
MDRIQQIPHTCVARKINLPLGMSKESLALNLPANLAPTMSPPSITYSDRFIPSRKASNFEEFAL
PDKTSPSPNSAGGQSSSTNGEGRDDACAAYSALLRTELFPATPDKTEGCRRPVIGSPSGNV
FRFKSQQCKSQSPFSLCPVGEDGDLSETGAVARKTTRKIPRSPFKVLDAPALQDDFYLNLVD
WSSHNILAVGLSACVYLWSASSSKVTKLCDLGLDDNVCSVAWTQRGTYLAVGTNNGGVQI
WDAAHCKQVRTMEGHCTRVGTLAWNSHILSSGGRDRNILQRDIRAQDDFVSKFSGHKSEV
CGLKWSYDNRELASGGNDNQLFVWNQQSQQPVLKYNEHTAAVKAIAWSPHQHGLLASGG
GTADRCIRFWNTATNTSLNCVDTGSQVCNLVWSKNVNELVSTHGYSQNQIIVWRYPTMSKL
ATLTGHTLRVLYLAISPDGQTIVTGAGDETLRFWNVFPSSKTQQNTIRDMGVWSSGRTHIR

SEQ ID NO: 475
MAGGQGEGEEKVDKLSMELTEDVMKSMEIGAVFKDYNGKINSLDFHRTNNYLVTASDDEAI
RLFDTASATWQKTSYSKKYGVDLICFTNHQTSVLYSSKNGWDESLRHLSLMDNKYLRYFKG
HHDRVVSLCMSPKGECFMSGSLDRTVLLWDLRIDKCQGLIRVRGRPAVAYDEQGLVFAISN
EGGLIKMFDARLYDKGPFDTFVVEGDKSEASGIKFSNDGKLILLSTMDSNIHVLDAYQGTTVH
SFSVEAVPNGGEAVPNGGTLEASFSPDGKFVISGSGNGNIHAWSVNSGKEVACWTTEGVIP
AVVKWAPRRLMFASGSSVLSLWVPDLSKLASLTGSNSNSAY

SEQ ID NO: 476
MHRVGSTGNTSNSSRPRREKRLTYVLNDANDSRHCSGINCLVISKLSLLGGNDYLFSGSRD
GTLKRWELADDSAVCSATFESHVDWVNDAVLTGETLVSCSSDTTLKTWRPFSDGVCTRTLR
QHSDYVTCLAAASKNSNIVASGGLGREVFIWDIEAAMAPVSRTSEAMDDDTSNGVLSSGNS
VLSTTVRSTNATNSASLHTSQLQGYTPIAAKGHKESVYALAMNDVGTLLVSGGTEKVVRVW
DPRSGAKQMKLRGHTDNVRALILDSTGRFCLSGSSDSIIRLWDLGQQRCVHSYAVHTDSVW
ALASTPNFSHVYSGGRDLSLYLTDLTTRESLLLCMEKHPLLRLTLQDDSIWVATTDSSLHRW
PAEGQNPPKMFQRGGSFLAGNLSFTRARACLEGSAPVPVNTQPSFVIPGSPGIVQHEILNNR
RHVLTKDAEGTVKLWEITRGAVLDDYGKVSFEEKKEELFEMVSIPAWFTMDTRLGSMSVHLD
TPQCFTAEMYAVDLNVPDAPEEQKINLAQETLRGLLAHWLSRRRQRLATQASANGDFPAGQ
ENALRNHISSRIDVHDDAETHIAGILPAFDFSTTSPPSIITEGSQGGPWRKKITDLDGTEDEKD
FPWWCLECVLHGRLSPRESLKCSFYLHPYEGTTVQVLTQGKLSAPRILRIQKVINYVLEKMVL
DRPLDSSNSETTFTPGLSGNQSHAAVVGDGSLRSGARVWQQKAKPLVEILCNNQVLSPDM
SLATVRTYIWKKPDDLYLYYRLVQNR

SEQ ID NO: 477
MMKGKTIQMQAAHQNHDGETSVACVLWDWHAKHLITAGADNTILIHSYPSSSSSKPITLRHH
KNAVTALAINSNVRSLASGSVDHSVKLYSYPGGEFQSNVTRFTLPIRSLAFNKSGELLAAAGD
DEGIKLISTIDNSIARVLKGHNGPVTSISFDPKNEFLASSDSDGTVIYWELSTGKPVHTLKKIAP
NTTSNPTSLNQISWRPDGEMLAVPGRKSEVSMYDRDTAEKLFSLKGGHSDTICSLAWSPNG

KYIATAGTDRQVMVWDADRRQDIDKQRFDNPICSVAWKPSDNALAVIDVLGRFGVWESPIAS
HMKSPADGAERYDNMEDEEPLMARYEEELEDSVSGSLNEIINDDDDDDEMGKIPRKILQKKP
SVKVEKGKEESNAKAFKSGQDSFKLKSAMQEAFQPGATQRQSGKRNFLAYNMLGSVITFDN
DGFSHIEVDFHDIGKGCRVPSMTDYFGFTMASLSESGSVFGSPQKGEKNPSTLMYRPFSSW
ANNSEWSMRFPMGEEVKAVALGSGWVAAVTSLNFLRVFSEGGLQKFVLSMDGPVVTAAGY
ENLLVVVSHASNPLLSGDQVLSFTVYDISQKTCPLSGRLPLSPGSHLTWLGFSEEGLLSSYD
SEGNLRVFTNDYNGCWVPIFSAARERKSETESIWMVGLNSTQVFCVVCKLPDTYPQVAPKP
VLSVLNLSLPLACSDLGADDLENEYLRGSLLLSQMQKKAEDAVACGRESNMEEDSIFKMEAA
LDRCLLRLIANCCKGDKLVRATELARLLSLEKSLQGAIKLVSAMKLPMLAERFNTILEEKILQEN
METISCRRLTSEAQDMDTPISISVKQVSYGANLGDSPFLPNRQVEPKHSTPVFSKPDTKIEVD
TSEAIAKGCDAQNGNIKSGDAEVQPASHNDSIQKPSNPFAKASNTSANQAVQRNASLLSSIK
QMKTATENEGKRKERARSGSLPQKPAKQSKIS

SEQ ID NO: 478
MKQKRKGHQVDDPKYSVQTPQEDDTPNESGPASEEVESSDEEGGNSSNIEDDIIYSSSEED
PVVSSDYEEDEDAESDAEGVTAEQELEGDIDNALQNYMGTLTVLSNFHGENLKNAEGEDTS
GDDDDEEEMPKRAEESDSPEDENDERPKRAEESDFSEDEDEERPKRAEESDSSEDEVPSR
NTVGDVPLRWYKDEQHIGYDIKGKKIKKQPKKDQLDSFLASTDDSSDWRKVYDEYNDEEVE
LTKDEIKFISRLRKGTIPHADVNPYEPYVDWFDWKDKGHPLSNAPEPKRRFIPSKWEAKKVV
KLVRAIRKGWITFQKAEEKPRFYLMWGDDLKPSEKMANGLSYIPAPKPKLPGHEESYNPPPE
YIPTQEEINSYQLMYEEDRPKFIPKRFDSLRNVPAYDRFLSEIFERCLDLYLCPRTRKKRINIDP
ESLIPKLPKPKDLQPFPSICFLEYKGHTGAVSCISPESSGQWLASGSKDGTVRIWEVETARCL
KVWDIGRPIQHIAWNPVSQLSILAVAVDEEVLVLNTGLGSEDSQEKVAELLHVKSKPVSADDL
GDNTSLTKWIKHEKFDGIKLTHLKPVHLISWHHKGDYFATVAPDGNTRAVLVHQLSKQQTQN
PFKKMQGRVVHVLFHPSRAIFFVATKTHVRVYDLVKQQLVKRLVTGLHEVSSMAVHHKGDN
LLVGSKEGKVCWFDMDLSTQPYKTLKNHSKDIHSVAFHDSYPLFASCSDDCKAYVFYGLVY
SDLLQNPLIVPLKVLQGHQSVNGMGVLDCQFHPKQPWLFTAGADSVVKLYCN

SEQ ID NO: 479
MMSLKRGFEESLVPAKRQKTELSTVTYGDGPRRTSSLESPIMLLTGHHAAIYTMKFNPTGTVI
ASGSHEREIFLWNVHGDCKNFMVLKGHKNAVLDLHWTTDGCQIISASPDKTLRAWDVETGK
QIKKMAEHSSFVNSCCPSRRGPPLVVSGSDDGTAKLWDLRHRGAIQTFPDKYQITAVGFSD
AADKIYSGGIDNEIKVWDLRRGEVTMRLQGHTDTITGMQLSSDGSYLLTNSMDCSLRIWDMR
PYAPQNRCVKILTGHQHNFEKNLLKCSWSSDGSKVTAGSADRMVYIWDTTTRRILYKLPGHT
GSVNETGFHPTQPIIGSCSSDKQIYLGEIEPNVGYQAVI

SEQ ID NO: 480
MEFSDTYKHTGPCCFSPDARYLAIAVDYRLVIRDVVTLKVVQLYSCMDKISNIEWALDSEYILC
GLYKRAMVQAWSLSQPEWTCKIDEGPAGIAHARWSPDSRHIITTSDFQLRLTVWSLVNTACI
HIQWPKHASKGVSFTQDGKFAAIATRRDCKDYVNLLSCHTWEVMGTFTVDTIDLADLEWSP
NDSAIVVWDSPLEYKVLIYSPDGRCLFKYQAYDSWLGVKTVAWSPCSQFLAVGSYDQTLRT
LNHLTWKPFAEFVHVSTVRGPASAVVFKEVEEPWNLDVSGLHLNDDNAHDIQDGKPAEGHS
RVRYKVVEFPVNVSSQKHPVDKPNPKQGIGLLAWSRDSQYLFTRNDNMPTALWIWDICRLE
LAALLIQKEPIRAAAWDPVYPRVALCTGSSHLYMWTPSGACCVNIPLPQFVVSDLKWNPDGT
SMLLKDRESFCCTFVPMLPEFNDDETNEE

SEQ ID NO: 481
MAKLIETHSCVPSTERGRGILIAGDAKTNSIIYCNGRSVIMRNLDNPLEASVYGEHSYPATVAR
FSPNGEWVASGDTSGTVRIWGRGSDHTLKYEYKALAGRIDDLEWSADGQRIVVCGDSKGK
SMVRAFMWDSGTNVGEFDGHSRRVLSCSFKPTRPFRVATCGEDFLVNFYEGPPFRFKTSH
RDHSNYVNCVRFAPDGSKFITVGSDRKGVIFDGKMGEKIGELSKEGGHTGSIYAASWSPDS
KQVLTVSADKSAKIWEISETGNGTVKKTLTFGSQGGADDMLVGCLWLNDYLITVSLGGIVSLL
SAVDPDKPPKTISGHMKSINAIALSLQSGQSEVCSSSYDGVIVRWILGVGYAGRVERKDSTQI
KCLATIEGELVTCGFDNKVRRVPLLSEQHKESEPIDIGAQPKDLDVAVGCPELTFVSTDAGIIII
RASKIVSTTNVGYAVTAAAISPDGTEAVVGGQDGKLRVYSIKGDTLLEESVLERHRGPINAIRF
SPDGSMFASGDLNREAVVWDRITREVKLKNMVYHTARINCIAWSPDSSKVATGSLDTCILIYE
VGKPASSRITIKGAHLGGVYGLAFSDQSTVISAGEDACVRVWSLP

SEQ ID NO: 482

MPQPSVILATAGYDHTVRFWEATSGRCYRTLQYPDSQVNHLEITPDKQYLAAAGNPHIRLFE
VNSNNPQPVISYDSHTNNVTAVGFQCDGKWMYSGSEDGTVKIWDLRAPGFQREYESRAAV
NTVVLHPNQTELISGDQNGNIRVWDLNANSCSCELVPEDTAVRSLTVMWDGSLVVAANNHG
TCYVWRLMRGTQTMTNFEPLHKLQAHNSYILKCLLSPEFCEHHRYLATTSSDQTVKIWNVD
GFTLERTLTGHQRWVWDCVFSVDGAFLVTASSDSTARLWDLSTGEAIRTYQGHHKATVCCA
LHDGTDGASC

SEQ ID NO: 483
MLTKFETKSNRVKGLSFHPKRPWILASLHSGVIQLWDYRMGTLIDKFDEHDGPVRGVHFHKT
QPLFVSGGDDYKIKVWNYKMRQCLFTFVGHLDYIRTVHFHNEYPWIVSASDDQTIRLWNWQ
SRVCISVLTGHNHYVMSASFHPKEDLVVSASLDQTVRVWDISGLRKKTVSPADDLSRLAQM
NTDLFGGGDVVVKYVLEGHDRGVNWAAFHTSLPLIVSGADDRQVKLWRMNDTKAWEVDTL
RGHTNNVSCVIFHARQDIIVSNSEDKSIRVWDMSKRTSVQTFRREHDRFWILAAHPEMNLLA
AGHDSGMIVFKLERERPAYVVYGGSLLYVKDRYLRTYEFATQKDNPLIPIRKPGSIGPNQGPR
SLSYSPTENAILICSDADGGAYELYAVPKDSHGRSDTVQEAKKGLGGSAVFVARNRFAVLDK
NHNQVTIKNLKNEVTKKFDLPVTADALFYAGTGNLLCRSEDSVFLFDMQQRTVLGEIQTPNV
RYVVWSNDMENVALLSKHTIIIASKKLSSTCSLHETIRVKSGAWDDNGIFMYSTLNHIKYCLPN
GDSGIIKTLDVPVYITKVSGKSLYCLDRDGKNRVIQIDITECLFKLALSKKKYDYVINMIRNSQL
CGQAIIAYLQQKGFPEVALHFVRDERTRFNLAVESGNIEIAVASAKEIDEKDHWYRLGVEALR
QGNAGIVEYAYQRTKNFERLSFLYLITGNLDKLSKMLRIAEMKNDVMGQFHNALYLGDIQERI
KILEESGHLHLAYATASLHGLADIADRLAADLGGNIPVLPPGKKSSLLMPPAPILHGGDWPLLR
VTKGIFEGGLENSTSAAYEEEDEEAAADWGEDIDIENIEGENGEATVLDDQEVKGGEDDEGG
WDMEDLELPPDVAAANVGTNQKTLFVAPTLGMPVSQIWMQKSSLAGEHAAAGNFETALRLL
TRQLGIKNFSPLKPLFLELYMGSHTFLPSFASVPAFSLALQRGWSESASPNIRGPPALVYRLS
VLEEKLTVAYRATTEGRFSEALRLFLNILHTIPVIVVDSRKEIDEVKELIGIAKEYVLGLRMEVKR
KEIRDDAVRQQELAAYFTHCNLQKAHLKLALLNAMGISYRCKNYNTAANFARRLLETDPSSN
HATKARQVLQVCERNLQDATQLNYDFRNPFVVCGATFTPIYRGQKEVSCPYCMARFVPDIA
GKLCSICDLAIVGSDASGLFCFATQTR

SEQ ID NO: 484
MDLLQNYQDDSEDSNPELRNHPPLEDATATSAPAGVENETSSSPDSSPLRLALPAKSCAPD
VDETLMALGVPGSEKKNNHNKPIDPTQHSVTFNPSYDQLWAPLYGPAHPYAKDGIAQGMRN
HKLGFVEDSAIEPFMFDEQYNTFHRYGYAADPSASLGSHIVGDLESLKKNDGASVYNLPKRE
HKRQKLEKKMIQKDENEEEEKEVGEEVDNPSTEEWLKKNRKSPWAGKKEGLQTELTEEQK
KYAQEHAEKKGDREKGEKVEIVDKTTFHGKEERDYQGRSWIDPPKDAKATNDHCYIPKRWV
HTWSGHTKGVSAIRFFPKYGHLLLSAGMDTKVKIWDVFNSGKCMRTYMGHSKAVRDISFSN
DGSRFLSAGYDRNIKLWDTETGKVISTFSTGKIPYVVKLHPDEDKQNVLLAGMSDKKIVQWD
MNSGEITQEYDQHLGAVNTITFVDNNRRFVTSSDDKSLRVWEFGIPVVIKYISEPHMHSMPSI
SLHPNTNWLAAQSLDNQILIYSTRERFQLNKKKRFAGHIAAGYACQVNFSPDGRFVMSGDG
EGRCWFWDWKTCKVFRTLKCHDNVCIGCEWHPLEQSKVATCGWDGMIKYWD

SEQ ID NO: 485
MARKGLGTDPAIGSLMSSKKRKEYKVTNRFQEGKRPLYAIAFNFIDARYHNIFATAGGTRVTI
YQCLEGGAISVLQAYVDDDKDESFYTLSWACDVNGSPLLVAGGHNGIIRVLDVANEKVHKSF
VGHGDSVNEIRTQALKPSLILSASKDESVRLWNVQTGICILIFAGAGGHRNEVLSVDFHPSDV
YRIASCGMDNTVKIWSMKEFWTYVEKSFTWTDLPSKFPTKYVQFPVFIAAVHSNYVDCTRW
LGNFILSKSVDNEVVLWEPYSKEQSTSDGVVDILQKYPVPECDIWFIKFSCDFHYNSMAVGN
REGKVYVWELQSSPPNLIARLSHAHCKNPIRQTAISHDGSTILCCCDDGSMWRWDVVQ

SEQ ID NO: 486
MESGAGGSVGARVPSAKPEMLQQPPYSNGDDDNDMERGTAPVPSSNPNTVSKWELDKDF
LCPICMQTMKDAFLTACGHSFCYMCIMTHLNNKSNCPCCSLYLTNNQLFPNFLLNKLLKKTS
ACQMASTASPVENLCLSLQQGAEVSVKELDFLLTLLAEKKRKMEQEEAETNMEILLDFLQRL
RQQKQAELNEVQADLHYIKDDILALEKRRLELSRARERYSRKLHMLLDDPMDTTLGHAAIDD
GNNVRTAFVRGGQGDAISGKFQQKKAEIKAQASSQGMQKRANFCHSDSQVLPTLSGLTIAR
KRRVLAQFDDLQECYLQKRRRWATQLRKQCDGGLRKERDGNSISREGYHAGLEEFQSILTT
FTRYSRLRVISELRHGDLFHSANIVSSIEFDRDDELFATAGVSRRIKVFDFATVVNEPADVHCP
VVEMSTRSKLSCLSWNKCIKSQIASSDYEGIVTVWDVNTRQSVMMYEEHEKRAWSVDFSRT
EPTRLISGSDDGKVKVWCTRQETSVLNIDMKANICCVKYNPGSSYYVAVGSADHHIHYYDLR
NPSVPLYEFNGHRKTVSYVKFISTNELASASTDSTLRLWDVRDNCLVRTFKGHTNEKNFVGL

TVNSEYIACGSETNGVFVYHKAISKPAAWHQFGSPDLDDSDDDTSHFISAVCWKSESPTMLA
ANSQGTIKVLVLAP

SEQ ID NO: 487
MANYVDSKKNFKCVPALQQFYTGGPFRLSSDGSFLVCACNDEVKVVDLATGSVKNTLEGDS
ELIVALALTPDNKYLFSASRSTQIKFWDLSSATCKRTWKAHNGPVADMACDASGGLLATAGA
DRSILVWDVDGGYCTHSFRGHQGVVTTVIFHPDPHCLLLFSGSDDATVRIWDLVAKKCISVL
EKHFSTVTSLAISENGWNLLSAGRDKVVNIWDLRDYHCRATIPTYEPLEAVCVLPTGSRLVSV
MNQSRALPENRKKSGAAPVYFLTVGERGIVRIWYSEGALCLYEQKSSDAIISSDKDELKGGF
VSAVLLPLTQGVMCVTADQRFLFYNLDESDEGKCDLKVSKRLIGYNEEIVDLKFLGDEEKFLA
VATNLEQVRMYDLSSMTCVYELSGHTDIVLCLDTVVFSGHSLLASGSKDHTVRIWDTESKSCI
CVAAGHMGAVGAVAFSKKAKNFFVSGSSDRTIKVWSFASVLDFGGISKSIKLSSQAAVAAHD
KDINSVAVAPNDSLICTGSQDRTARIWRLPDLVPVLVLRGHKRGVWCVEFSPVDQCVMTAS
GDKTIKIWALSDGSCLKTFEGHTASVLRASFLTRGTQFVSSGADGLLKLWTIKSNECIATFDQ
HEDKIWAMAVGKKTEMLATGGSDSLVNLWHDCTTTDEEEALLKEEEAALKDQELLNALADT
DYVKAIQLAFELRRPYKLLNVFTELYSKGHAQDQIQKVIRELGNEELRLLLEYVREWNTKPKF
AHVAQFVLFQLFNVLPPKEIIEVQGISELLEGLIPYAQRHYSRIDRLMRSTFLLDYTLSSMSVLS
PTETDLSSSNLLARTADPLHAQIDQFHPTHFPEPNLTPIQSLLDSGNTDSVEVTARRAKKKRV
SGNDSEKTTVAEVKIGDMENAFDEPDVADQGSSRKHKPASSKKRKSIAVGNASIKRIASGNA
VTIALQV

SEQ ID NO: 488
MESSCSSMNSNRHSTEKRCLRPLQKQGASMNKHSSDRFIPARGSIDLDVARFMVTQKQKD
NNDIHALSPSPSPSKKAYQKEMADTLLKNAGAADNNCRILSFNGKSSTVSQGSQENVLANLS
ISRRARRYIPQSADRTLDAPDLLDDYYLNLLDWSSTNVLSTALGNTVYLWDASNSSISELLIAD
EEEGPVTSVSWAPDGSQIAVGLNNSVVQLWDSQSNKKLRALKGHHDRVGALSWNGPILTT
GGLDGIIINHDVRTRDHIVQTYKGHTQEVCGLKWSPSGQQLASGGNDNLLYIWDKSMASHN
PSSQYFHQLDEHCAAVKALAWCPFQTNLLASGGGTSDGSIKFWNTQTGACLNTVDTHSQV
CSLLWNRHERELLSSHGLNQNQLTLWKYPSMVKITELTGHTARVLHMAQSPDGYTVASAAA
DETLKFWQVFGAPDASKKTKTKDTKGAFNMFHMHIR

SEQ ID NO: 489
MLDEIVADEEEEFNIWKKNTPLLYDVVITHALEWPSLTVQWLPDRHQSPTKDYSLQKMIVGT
HTSGDEPNYLMIAEVQMPLQYSEDGNVGGFESTEAKVHIIQQINHEGEVNRAQYMPQNSFII
ATKTVSSDVYVFDYTKHSSNAPQERVCNPELILKGHTNEGYSLSWSPLKEGQLLSGSNDAQI
CFWDINAASGRKVVEAKQIFKVHEGAVEDVSWHLKHEYLFGSVGDDCHLLIWDTRTAAPNK
PQHSVVAHESEVNSLAFNPFNEWLLATGSADKTVKLFDLRKLSCSLHTFSNHTEEVFQIEWS
PMNETILASSGGDRRLMVWDLRRIGDEQTSEDAEDGPPELIFIHGGHTSKISDFSWNLHDDW
LIASVAEDNILQIWQMAENIYHDDADIL

SEQ ID NO: 490
MTKEDHGESRDEMGERMVNEEYKLWKKNTPFLYDLVITHALEWPSLTVQWLPPSCKQQQD
IIKDDDIDHPNTQMVILGTHTSDNEPNYLILAEVQLHDGTEDEDGDGDVKRPQDKMKPGTSG
GAMGKVRILQQINHQKEVNRARYMPQKPTIIATKTVNADVYVFDYSKHPSKPPQEGRCNPEL
RLQGHESEGYGLSWSPLKEGHLLSASDDAQICLWDITAATKAPKVVEANQIFRYHDGPVEDV
AWHAIHDHLFGSVGDDHHLLLWDIRNDSEKPLHIVEAHQAEVNCLAFNPFNEWIVATGSADR
TVALHDIRKLDKVLHTCAHHMEEVFQIGWSPQNGAILASCGSDRRLMVWDLSRIGDEQNPE
DAEEAPPELLFIHGGHTSKISDFSWNPAEEWVIASVAEDNILQVWQMSEHIYNDDNDSPTA

SEQ ID NO: 491
MAMAMGDENAADPVEEFNIWKKNTPFLYDLVITHALEWPSLTVQWLPDRHQSSTADYSLQK
MIVGTHTSEDEPNYLMIAEVQIPLQNSEDNIIGGFESTEAKVQIIQKINHEGEVNKARYMPQNS
FVIATKTVSSDVYVFDYSKHPSKAPQERVCNPELILKGHSNEGYGLSWSPLKEGYLLSGSND
AQICLWDINAAFGKKVLEANQIFKVHEGAVGDVSWHLKHEYLFGSVGDDCHLLIWDMRTAAP
NKPQQSVIAHQSEVNSLAFNPFNEWLLATGSMDKTVKLFDLRKLSCSLHTFSNHTDQVFQIE
WSPMNETILASSGADRRLMVWDLARIGETPEDEEDGPPELLFVHGGHTSKISDFSWNLNDD
RVIASVAEDNILQIWQMAENIYHDDEDML

SEQ ID NO: 492

MGLFEPFRALGYITDGVPFAVQRRGIETFVTLSVGKAWQIYNCAKLIPVLVGPQMDKKIRALA
CWRDFTFAATGHDIAVFRRAHQVATWSGHKAKVTLLLSFGQHVLSVDLEGCLFIWAVAEVN
QNKPPIGQIQLGEKFSPSCIMHPDTYLNKVLIGSEEGTLQLWNVNTRKKLYEFKGWGSSIRC
CVSSPALDVVGIGCSDGKIHVHNLRYDEEIVTFMHSTRGAVTALSFRTDGQPLLAAGGSSGVI
SIWNLEKKKLQSVIKDAHDSSVCSLHFFANEPVLMSSATDNSIKMWIFDTTDGEARLLKYRSG
HSAPPMCIRYYGKGRHILSAGQDRAFRIFSVIQDQQSRELSQGHVGKRAKKLKVKDEEIKLPP
VIAFDAAEIRERDWCNVVTCHLDDPCAYTWRLQNFVIGEHILKPCLEDPTPVKSCSISACGNF
AVLGTEGGWLERFNLQSGISRGTYIDIGEKRQCAHNGAVVGLACDATNTLLISGGYNGDIKV
WDFKGRELKFRWEIEVPLIKIVYHPGNGILATAADDMILRLFDVTAMRLVRIFVGHMDRVTDL
CFSGDGKWLLSSSMDGTIRVWDIISSRQLNAMHMDSAVTALSLSPGMDMLATTHVGHNGIY
LWANRMIYSKATDIEPFISGKQVVKVSMPTVSSKRESEEGDEKRTIVAESNVNKSDVSGSLIG
DSYSAQLTPELVTLALLPKAQWQSLVNLDIIKMRNKPIEPPKKPEKAPFFLPSLPTLSGERIFIP
SSMNGDGDQDETRNDKTVFEARGKKLGGESLSFMQLLQSCAKIKDFTTFTNYLKGLSPSAV
DMELRLLQIVDNENISETEHSVELQGIGMLLDYFVNEVSCNNNFEFVQALIRLFLKIHGETIRC
QVSLQEKARKLLEIQSSTWERLDTSFQNARCMITFLSSSQF

SEQ ID NO: 493
MIAAVCWVPKGVAKVLPDSAEPPTQEEIQELLKCNVVAESDDNEDSDEESEEMDTETDKNT
DAVAKALAAANALGSQSSDFQRQHKVDDIANGLKELDMDHYDDEDEGIDIFGSGSLGNCYY
PANDMDPYLVEQDDDDEDEIEDMTIKPSDLIILSARNEDDVSHLEVWIYEEETEEGGSNMYV
HHDIILPAFPLSLAWLDCNLKGGEKGNFVAVGTMQPEIELWDLDVLDEVEPAVVLGGAVKDE
ASGKTTKLKKKKKNKQAVNFKEGSHTDAVLGLAWNMEYRNVLASASADKSVKIWDIVAEKC
EHTMQPHTDKVQAVAWNPNQATVLLSGSFDRSVIMMDMRAPTHSGIRWPVPADVESLAWD
PHTDHSFMVSAEDGTVRGFDIRAAASTADFDGKPMFILHAHDKAVCAISYNPAAPSLLTTGST
DKMVKLWDITNNQPSCIASTNPNVGAVFSAAFSKNSPFLLATGGSKGILHVWDTLDNSEVAR
RFGKFRPQN

SEQ ID NO: 494
MIMDENEFCDIFSLRKRLCLLSSQEGEEEEELEAMSQLDAGEFTVTGNEEVVAIAEDDVNTGI
LSQDLFSSQDYCTPSQPQDSTDLDSKDKAPCPLSPVKSTIQRKRCRPELLSNPPDSIQFSFQ
RLERVRSEESIQSSSQQLARVRSEVSSSDDFKTPKITASGQKNYVSQSALALRARVMSPPCI
KNPYLDENEELNEKIQRSTRRSPACVTPIQSGACLSRYRADFHELEEIGRGNFSRVYKALNRL
DGCCYAVKCSQSELRLDTERKVALMEVQSLAALGPHKNIVGYHTAWFENDHLYIQMELCDH
NLTTANDRGILRTDTDFLEAVYQIAQALEFIHGRGVAHLDVKPENIYVRDGTYKLGDFGRATLI
NGTLHVEEGDARYMSREILNDNYEHLDKVDMFSLGATFFELLMRKQYPGSGKRIDRDTEIKIP
ILPGFSIYFQKLLQDLVSNDPGKRPSAKDVLKNPIFNKVRGAKEV

SEQ ID NO: 495
MLAPALEMEPVEPQSLKKLSFKSLKRALDLFSPVHGQIAPPDPESKKMRISYKLNFEYGGGS
GSEDQVPKRKESGAAQNQGQQAAGASNALALPGPEGSKIPPMEKSQNALTVGPSLRPQGL
NDVGLHGKGTAIISASGSSDRNLSTSAIMERLPSRWPRPVWHPPWKNYRVISGHLGWVRSI
AFDPSNQWFCTGSADRTIKIWDLASGRLKLTLTGHIEQIRGLAVSSKHTYMFSAGDDKQVKC
WDLEQNKVIRSYHGHLSGVYCLALHPTIDILLTGGRDSVCRVWDIRSKMQIFALSGHDNTVC
SVFARPTDPQVVTGSHDTTIKFWDLRHGKTMTTLTNHKKSVRAMAQHPKENCFASASADNI
KKFQLPRGEFLHNMLSQQKTIINTMAVNEEGVMATGGDNGSLWFWDWKSGHNFQQAHTIV
QPGSLESEAGIYALSYDLTGSRLVSCEADKTIKMWKEDELATPETHPLNFKPPKDIRRF

SEQ ID NO: 496
MEEAAKEQSAGSGKPKLLRYGLRSAAKPKEDKKEEQLHQPPPPPPPQQQAAPAPAPAATR
SSTSGSAGGRDRRPQQQHAVDEKYARWKSLVPVLYDWLANHNLLWPSLSCRWGPQLEQA
TYKNRQRLYISEQTDGSVPNTLVIANCEVVKPRVAAAEHVSQFNEEARSPFIRKYKTIIHPGEV
NRIRELPQNPNIVATHTDSPDVLIWDVESQPNRHAVYGATASRPNLILTGHQENAEFALAMC
PAEPFVLSGGKDKTVVLWSIQDHITASATDQTTNKSPGSGGSIIKKTGEGNEETGNGPSVGP
RGIYCGHEDTVEDVAFCPSTAQEFCSVGDDSCLILWDARIGTNPVAKVEKAHNGDLHCVDW
NPHDNNLILTGSADNSVNMFDRRNLTSNGVGSPVYKFEGHKAAVLCVQWSPDKPSVFGSS
AEDGLLNIWDYERVDKKVDRAPNAPAGLFFQHAGHRDKIVDFHWNTADPWTMVSVSDDCD
TAGGGTLQIWRMSDLIYRPEEEVLAELENFKAHVLECSKA

SEQ ID NO: 497

MAKDEEEFRGEMEERLVNEEYKIWKKNTPFLYDLVITHALEWPSLTVQWLPDREEPPGKDY
SVQKMILGTHTSDNEPNYLMLAQVQLPLEDAENDARQYDDERGEIGGFGCANGKVQVIQQI
NHDGEVNRARYMPQNPFIIATKTVSAEVYVFDYSKHPSKPPQDGGCHPDLRLRGHNTEGYG
LSWSPFKHGHLLSGSDDAQICLWDINVPAKNKVLEAQQIFKVHEGVVEDVAWHLRHEYLFG
SVGDDRHLLIWDLRTSATNKPLHSVVAHQGEVNCLAFNPFNEWVLATGSADRTVKLFDLRKI
SSALHTFSCHKEEVFQIGWSPKNETILASCSADRRLMVWDLSRIDEFQTPEDALDGPPELLFI
HGGHTSKISDFSWNPCEDWVIASVAEDNILQIWQMAENIYHDEEDDMPPEEVV

SEQ ID NO: 498
MGKYMRKGKGVGEVAVMEVSQGSLGVRTRARTLAAASSQKDHRRLGASKSVTTKHQSSA
PPASPCVESSMHTCYLELRSRKLEKFSRCYHSAHGATSHGESKRSLSLSEPSRLAVSEEAR
VASDKSSHRVLQQQSSVAHSRNNSATFSHNAKPAKAAQRKERRDDDHTSARPSEAPHEDE
DGMEVEASFGENVMDLDSRERRTRETTPSSYTRDVETMETPGSTTRPPSNAGRRRFQTEG
GHGTRNQFHVPTTNEIEEFFAGAEQQEQRRFTDRYNYDPVSDSPLPGRFEWVRLRP

SEQ ID NO: 499
MQNMEENVQSSWSLHGNKEICARYEILKRVSSGTYLDVYRGRRKEDGLIVALKEVHDYQSS
WREIEALQRLCGCPNVVRLYEVILEFLTSDLYSVIKSAKNKGENGIPEAEVKAWMIQILQGLAN
CHANWVIHRDLKPSNMLISAYGILKLADFGSMSFLKRAIYEVEYELPQEDILADAPGERLMDE
DDSVKGVWNEGEEDSSTAVETNFDDMAETANLDLSWKNEGDMVMQGFTSGVGTRWYRA
PDFLYGATIYGKEIDLWSLGCILGELLILEPLFSGTSNIDQLSRLVKVLGLQQKKNWPGCSNLP
DYRKLCFPGDGSPVGLKNHVPNCSDNMFSILERLVCYDPAARLNAKEIVENKYFVEDPYPVL
THELRVPSPLREENNFSEDWAKWKDMEVDSDLENIDEFNVVHSSDGFCIKFS

SEQ ID NO: 500
MAPVKRIEPEKTKANEGKPKRRKVAFAIDTGIEANDCISLHLVSTPEEMRDAEGVEDQSLSFN
PEYMQHFVGEHGKIYGYKGLKIDVWLNALSFHAYVDIQYESKVEEGKSEKEATDLTDIMKRIF
GRGLVEDRNAFIQSFSSNSQSIESMIHNEGERIATREILTDKGLSAQGDSERLGVSNEIFRLEL
SDPQIREWHARLEPLVLLFVEGSQPIEQDDPKWEMYIRVQRESLSGGSAVCRLLGFCTVYRF
YHYPDTTRLRISQILVFPPYQGKGHGLLLLEAVNKTAVSRDSYDVTVEEPSESLQELRDCMD
TIRLLSFEPVMPAVKSAVQKLKEANPSDKGAADHCLEGNVNNETVTTSSTKPKNKSGWFPP
PGLVEEVRKHLKISKKQFKRCWEILLYLNLDRSDSQCEDKYHISLMEQIMSELFDKSSEKSAK
GKRVIDIDNEYDNSKTFIMVRTRNPGNGEGFLPEALEGGMEVSQEDQLKSLFEERLEEIAQIA
EKVPSLCKALQMP

SEQ ID NO: 501
MPEDRKKILEALAAKRKAEAESGEKKKRQKSSLNPAKPVSKPVSKPVGGIGSKGKSTSAPIS
STKAKSKHKEEVKAKRVTKMDRYETDEDDESEEEEDLDSESDDDELSDEDSEDDIKSKSVK
KLPPQSKGKAPVKGISSSNGKGRDEKGKGIMKDKGKAKAKVEESSSDAEGDSDDDGGDLS
DDPLQEVDPSNILPSKTRRRASQPTNYQFANMSGDDDDDDDSD

SEQ ID NO: 502
MADVPESLQQEKDEQGTDKNCCDGKFQKEIDIDDMEEEYNESSIDDEEENLSDNVATNNMG
TIPQGQACMAVTVEGIEHANSVGCGRNGREGSEEVTAAEDMGHVSIENIREQGRNRKSSEQ
LLALYEQEGLLEDDEDDDDVDWEPFEGVTVQMKWYCTNCTMANSDDSVHCDSCGEHRNS
DILRQGFLASPYLPAESPSSSDVPDERLEESKCVMTTLTPSISPMIGVCCSSLQSERRTVVGF
DERMLLHSEIQMETYPHPERPDRLRAIAASLRAAGLFPGKCFSIPAREATCEELQTIHSLEHV
NAVESTSCGMLSHLSPDTYANEHSSLAARLAAGLCADLAKAIMTGQAQNGFALVRPPGHHA
GVKDSMGFCLHNNAAIAVSASRVVGAKKVLIVDWDVHHGNGTQEIFEADQSVLYISLHRHGE
GFYPGSGAVTEVGSSKGEGYSVNIPWKCGGVGDNDYIFAFQHAVLPIAEQFEPDLTIISAGFD
AAKGDPLGRCEVTPDGFAHMAQMLSCLSKGKMLVILEGGYNLRSISASATAVIKVLLGDNPK
ALPIDIQPSKGGLQTLLEVFEIQSKYWSSLKGHDQKLRSQWEAQYGSKKRKVIRKRHMHIVG
GPVWWKWGRKRVVYYHWFARVSSRKHL

SEQ ID NO: 503
MASGAGAAGVVEWHQKPPNPKNPVVFFDVTIGTIPAGRIKMELFADIVPRTAENFRQFCTGE
YRKAGIPIGYKGCHFHRVIKDFMIQAGDFVKGDGSGCISIYGSKFEDENFIAKHTGPGLLSMA
NSGPNTNGCQFFLTCAKCDWLDNKHVVFGRVLGEGLLVLRKIENVQTGQHNRPKLPCVIAE
CGEM

SEQ ID NO: 504
MAKLVSSVCAFSCQQRHPHSRPRFLSNRDHYNHYHNHSHYHNVCYFPPMMMMQQQLQK
QKRMTTKTITSLFKCNSSNHTLLKGLKEFMGFKFRLQAAMLSCEMSILGRVFAIFFIVHQAAAP
FPFNHFDNWLVPPASAVLYSPNTKVPRTGEVALRKSIPANPAMKSIQDFLEDIYYLLRFPQRK
PYGTMEGDVKSALQIAINEKDSILGSVPLDMKERGLQLYNFLIDGQGGLQVLIEYIKEKDPDKV
SVNLSSSLDTIAQLELLQAPGLPYLLPEEYQQYPRLNGRATIEFTMEKGDNSMFSVSSGGGL
QKTATIQVVLDGYSAPLTAGNFTKLVIDGAYNGLKLKTTEQAVISDNERAEAGFNLPIEILPAG
GFEPLYRTTLSVQDGELPVLPLSVYGAIAMAHNTISEDYSSPSQFFFYLYDKRNAGLGGLSFD
EGQFSVFGYTTVGKEILPQLKTGDIIKSAKLVDGFDHLVLPSSST

SEQ ID NO: 505
MDHYYQDDFDYLVDDEMVDFADDVEDDVRTRRRSDIDSDSENDFDSNNKSPDTTALQAKR
GKDIQGIPWNRLNFTREKYRETRLQQYKNYENLPRPRRSRNLDKECTNFERGSSFYDFRHN
TRSVKATIVHFQLRNLVWATSKHNVYLMQNYSIMHWSSLKQKGEEVLNVAGPIIPSVKHPGS
SPQGLTRVQVSAMSVKDNLVVAGGFQGELICKYLDKPGVSFCTKISHDENGITNAVEIYNDA
SGATRLMTANNDLAVRVFDTEKFTVLERFSFPWSVNHTSVSPDGKLVAVLGDNADCLLADC
KTGKTVGTLRGHLDYSFAAAWHPDGYILATGNQDTTCRLWDVRKLSSSLAVLKGRMGAIRSI
RFSSDGRFMAMAEPADFVHLYDTRQNYTKSQEIDLFGEIAGISFSPDTEAFFVGVADRTYGS
LLEFNRRRMNYYLDSIL

SEQ ID NO: 506
MDCSGDEEEEQFFESLEEMLSPSDSGSEAADNETGCRNADARSKYEIWKRAPSSIQERRQ
RFLVRMGLANPSELGNQVNSTSAESTCSTETANIPNGIERLRENSGAVLRTAGSSGRKTHCK
NVINIGLREGSVRSSSSSNGTPDVGEDNGEFGGTIFSRSGGTWECMCKIKNLDSGKEFVVD
ELGQDGLWNKLREVGTDRQLTMDEFERSLGLSPLVQELMRRESGVAQADCNGVHHHDAEI
SSSKRRSWLKALKSAAYSMRRPKEDQSNYDSERSGRRSGSFDVPWGKPQWTKVRHYRKR
YKEFTALYMGQEIEAHEGSIWTMKFSLDGRYLASAGQDCVIHVREVIESMRTFGADTPDLYA
SSAYFSMNGLQELVPLSIEDHANKMKRGKIIGSKKSSNSDCIVLPNKVFQLSEEPVCSFHGHL
LDVFDLSWSPSQYLLSSSMDKTVRLWKLGHESCLKVFSHNDIVTCIQFNPVDERYFISGSLD
GKARIWSIPDRQVVDWSDLREMVTAVCYTPDGQGGLVGSIKGSCRFYNTSGNKLQLENQLN
VRSKKKKSSGKKITGFQFAPGGDSQKVLITSADSRVRVYNGSELVCKYKGFRNTCSQISASF
APNGQHFVCASEDSRVYIWNHESPRGSGARHEKSSWSHEHFLSQGVSVAIPWSGMKLQPP
VWNSPEFMLGQRHNLLSLQGGKDVGCQNGLLSREAGEGQESETPLHYISQVSHSCGSQN
MVDRDGQDDLSRYSACISDSRLSSFMAFPESPGNPDDLNSKVFFSDSSSKGSATWPEEKLP
PTRKQSRSNSTSSHYDTLKTHLGNTIQGQSGASAAVAWGLVIVTAGHGGEIRSFQNYGLPV
RL

SEQ ID NO: 507
MPSIPAIGEFTVCEINRELLTTKDESDTQAKDAYAKILGLVFPPISFQIEEGFGSASRQQFDQD
LDREDTIVTPSTSEGTNALQEGGLLLKGVSVLKNILASSFGPIFSPNDTKVLKKVELLQGISWH
RHKHILAFISGSNQVTVHDFQDPEWRESSLLVSESQRGIEALEWRPNGGTTLSVACRGGICI
WSASYPGSVAPVRSGVASFLGTSTRGSSVRWTLVDFLQIPGGKAVTALSWSPTGRLLASAS
REDSSFTIWDVAQGVGTPLRRGLGGISLLKWSPTGDYLFSAKPNGTFYLWETNTWTLEQWS
SSGGCVISATWGPDGRMLFMAFSESTTLGSLHFAGRPPSLDAHLLPMELPEIGSITGGFGNI
EKMAWDGCGERLAVSYTGGDLMYVGLIAIYDTRRTPFISASLVGFIRGPGEQVKPLAFAFHD
KFKQGPLLSVCWSSGLCCTYPLIFRAH

SEQ ID NO: 508
MEEENAKHTEETRQVQVRFTTKLQPALRVPTTSIAIPAHLTRYGLSDIVNTLLGNDKPQPFDF
LVESELVRTSLEKLLLIKGISAEKILNIEYILAVVPPKQEEPSLHDDWVSVVDGSYPNFIFSGSF
DSIGRIWKGEGLCTHVLEGHRDAITSAAFIMPSDSSDSFINLATASKDRTLRLWQFKPNEHMT
NGKMVRPYKLLKGHTSSVQTVSACPRRNLICSGSWDCSIKIWQTAGEMDIESNAGSVKKRK
LEDSTEQIISQIEASRTLEGHSQCVSSVVWLEKDTIYSASWDHSVRSWDVETGVNSLTVGCR
KALHCLSIGGEGSALIAAGGADSVLRIWDPRMPGTFTPILQLSSHKSWITACKWHPKSRHHLI
SASHDGTLKLWDVRSKVPLTTLEAHKDKVLCADWWKEDCVISGGADSTLQIFSNLNLT

SEQ ID NO: 509
MNRLRSKRNHILELRLGQSEPEKEATLASNRSRGTNAPIVVEDDDDVVVSSPRSFALARSSV
SQRSSRIPIVNEEDLELRLGLAVTGRTSAEHNPRRRHGRVPPNKPIVLCDDAGEADQSSSKK

RRTGQQLSSDVQSDESKEVKLTCAICISTMEEETSTICGHIFCKKCITNAIHRWKRCPTCRKKL
AINNIHRIYISSSTG

SEQ ID NO: 510
MEEPPPPAVLPSSEDTSIVSSHSFVNAPPTVPVGLDASIPQISTPGINQPGLTIPVPPEAAPLT
ASLVAASAGMPPAVVPSFVRPAIVAHPSVMPPPSMPLAALPMPVASAVPVAAPHFPPSTPND
NSITPSMPVPTPIVASSSVPPSVTIPGIAPLPFIAPIPVPSSRPVAPSPFMPPARPLGASVSVAM
DVDNTDEQDQDADNKGESPSSSPDHPEDPSAAEYEITEESRKVRERQEQAIQELLLRRRAY
ALAVPTNDSSVRARLRRLNEPITLFGEREMERRDRLRALMAKLDAEGQLEKLMKVQEEEEAA
ANVDAEEVQEMEGPQVYPFYTEGSQELLKARTEITKFSLPRAVSRLQRARRKREDPDEDED
EELKCVLQQSAQINMDCSEIGDDRPLSGCAFSSDGTLLATSAWSGVTKLWSVPNINKVATLK
GHTERVTDVAFSPTNCHLATACADRTAMLWNSEGVLMKTYEGHLDRLARLAFHPSGLYLGT
ASFDKTWRLWDVNTGIELLLQEGHSRSVYGIAFQCDGSLAATCGLDGLARIWDLRTGRSILA
LEGHVKPVLGIDFSPNGYHLATGSEDHTCRIWDLRKRQSVYIIPAHSHLVSQVKFEPQEGYFL
VTASYDSTAKVWSARDFKSIKVLAGHEAKVTSVDITADGQYIATVSHDRTIKLWSSKNSTNDM
NIG

SEQ ID NO: 511
MKRAYKLQEFVAHASNVNCLKIGKKSSRVLVTGGEDHKVNMWAIGKPNAILSLSGHSSAVES
VTFDSAEALVVAGAASGTIKLWDLEEAKIVRTLTGHRSNCISVDFHPFGEFFASGSLDTNLKI
WDIRRKGCIHTYKGHTRGVNSIRFSPDGRWVVSGGEDNIVKLWDLTAGKLMHDFKCHEGQI
QCMDFHPQEFLLATGSADRTVKFWDLETFELIGSAGPETTGVRAMIFNPDGRTLLTGLHESL
KVFSWEPLRCYDAVDVGWSKLADLNIHEGKLLGCSYNQSCVGVWVVDISRVGPYAAGNVS
RTNGHNEAKLASSGHPSVQQLDNNLKTNMARLSLSHSTESGIKEPKTTTSLTTTEGLSSTPQ
RAGIAFSSKNLPASSGPPSYVSTPKKNSTSRVQPTTNFQTLSRPDIVPVIVPRSNSLRPETTS
DAKKEMNNFGRVVPSTVSTKSTDVIKSGSNRDESDKIDSINQKRMTGNDKTDLNIARAEQHV
SSRLDNTNTSSVVCDGNQPAARWIGAAKFRRNSPVDPVVSPHDRSPTFPWSATDDGVTCQ
PDRQVTAPELSKRVVEPGRARALVASWETREKALTADTPVLVSGRPPTSPGVDMNSFIPRG
SHGTSESDLTVSDDNSAIEELMQQHNAFTSILQARLTKLQVIRRFWQRNDLKGAIDATGKMG
DHSVSADVISVLIERSEIFTLDICTVILPLLTRLLQSETDRHLTVAMETLLVLVKTFGDVIRATISA
TPTIGVDLQAEQRLERCNLCYVELENIKQILVPLIRRGGAVAKSAQELSLALQEV

SEQ ID NO: 512
MAGSDENNPGVVGGAHVQEGLRVGAGKMGAGNVQQRRALSNINSNIIGAPPYPCAVNKRV
LSEKNVNSENDLLNAAHRPITRQFAAQMAYKQQLRPEENKRTTQSVSNPSKSEDCAILDVDD
DKMADDFPVPMFVQHTEAMLEEIDRMEEVEMEDVAEEPVTDIDSGDKENQLAVVEYIDDLY
MFYQKAEASSCVPPNYMDRQQDINERMRGILIDWLIEVHYKFELMDETLYLTVNLIDRFLAVQ
PVVKKKLQLVGVTAMLLACKYEEVSVPVVEDLILISDRAYSRKEVLEMERLMVNTLHFNMSVP
TPYVFMRRFLKAAQSDKKLELLSFFIIELSLVEYDMLKFPPSLLAASAIYTALSTITRTKQWSTT
CEWHTSYSEEQLLECARLMVTFHQRAGSGKLTGVHRKYSTSKFGHAARTEPANFLLDFRL

SEQ ID NO: 513
MQAPREGKSAAAIVGMGKYMKKSKAIPRDVSLLEASPRSPSATGVRTRAKTLASRRLRRAS
QRRPPPPAAAAAAAAPSLDASPCPFSYLQLRSRRLRRPRLAPSPEARIDEGPAGSGSRGSR
DASCSARTASSSGGVEGEGACVGRGDRGNGGECVRDAAVDASYGENDLEIEDRDRSTRE
STPCSLIRDSNANTPPGSTTRQQSSCTAHRTQMSILRSIPTSDEMEEFFAYAEQRQQRSFIEK
YNFDIVKDRPLPGRFEWVQVIP

SEQ ID NO: 514
MDGHSSHLAAQNRSRGSQTPSPSHSAASASATSSIHLKRKLSAANASAASAAAAAAAAAAA
ADDHAPPFPPSSISADTRDGALTSNDDLESISARGGGAGDDSDDDSDDEEEDDGDNDGGS
SLRTFTAARLENVGPAAARNRKIKAESNATVKVEKEDSAKDGGNGAGVGALGPAATSGAGS
GSGTVPKEDAVKIFTENLQASGAYSAREENLKREEEAGRLKFECLSNDGVDDHMVWLIGLK
NIFARQLPNMPKEYIVRLVMDRNHKSVMVIRRNLVVGGITYRPYASQKFGEIAFCAIKADEQV
KGYGTRLMNHLKQHARDVDGLTHFLTYADNNAVGYFIKQGFTKEIYLDKDRWHGYIKDYDG
GILMECKIDPKLPYTDLSTMVRRQRQAIDEKIRELSNCHIVYQGIDFQKRDAGVPQNTIKMEDI
PGLREAGWTPDQWGYSRFRGLSDQKRLTFFIRQLLKVLNDHSDAWPFKEPVDAREVPDYY
DIIKDPMDLKTMTKRVESEQYYVTLEMFIADVKRMFANARTYNSPDTIYFKIATRLEAHFQSKV
QSNLQSGAGKIQQ

SEQ ID NO: 515
MFNGMMDPELFKLAQEQMNRMSPAELAKIQQQMMSNPELMRMASESMKNMRPEDLRQAA
EQLKHVRPEEMAEIGEKMANASPEEIAAVRARADAQMTYEINAAKILKKEGNELHSQGRFKD
ASQKYLRAKNNLKGIPSSEGKNLLLACSLNLMSCYLKTRQYEECIKEGSEALACEEKNLKAFY
RRGQAYRELGQLKDAVSDLRKAHEISPDDETIAQVLRDTEESLTKEGGSAPRGVVIEEITEED
ETLASVNHESPSEYSEKRHQESEDAHKGPINGDIMGQMTNSESLKALKGDPDAIRSFQNFIS
NADPTTLAAMGAGNAGEVSPDLIKTASSMIGKMSAEELQKMIQLASSFPGENPYVTRNSDSN
SNSFGNGSIPNVSPDMLKTASDMMSKMSPDDLQRMFEMASSSRGKDPSLDANHASSSSGA
NLAANLNHILGESEPSSSYHIPSSSRNISSSPLSNFPSSPGDMQEQIRNQMKDPAMRQMFTS
MMKNMSPEMMANMGKQFGLELSPEDAAKAQEAMSSLSPEMLDKMMRWADRAQRGVETA
KKTKNWLLGRPGMILAICMLLLAVILHRLGFIGS

SEQ ID NO: 516
MIAAISWVPRGASKAVPEVAEPPSKEEIEEILKSGVVERSGDSDGEEDDENMDAVASEKADE
VSTALSAADALGRISKVTKAGSGFEDIADGLRELDMDNYDEEDEDVKLFSTGLGDLYYPSND
MDPYLKDKDDDDDTEEIEDLSIKPMDSLIVCARTDDEVNLLEVYLLEPSLSDESNMYVHHEVV
ISEFPLCTAWLDCPIKGGDKGNFIAVGSMEPAIEIWDLDIIDAVEPCLVLGGQEELKKKKKKGK
KASIKYKEGSHTDSVLGLAWNKEFRNILASASADRQVKIWDVAAGKCNITMEHHTDKVQAVA
WNHHAPQVLLSGSFDHSVVMKDGRIPSHSGYRWSVTADVESLAWDPHSEHFFVVSLEDGT
VRGFDVRAAISNSASQSLPSFTLHAHEKAVSTISYNPAAPNLLATGSTDKMVKLWDLSNNQP
SCIASRNPKAGAVFSVSFSEDSPLLLAIGGSKGRLEVWDTSSDAAVSRRFGKHGKPKTAEPG
S

SEQ ID NO: 517
MKFCKKYQEYMQGQEGKKLPGLGFKKLKKILKRCRRRDSLHSQKALQAVQNPRTCPAHCS
VCDGSFFPSLLEEMSAVLGCFNKQAQKLLELHLASGFQKYLMWFKGKLRGNHVALIQEGKD
LVTYALINAIAIRKILKKYDKIHLSTQGQAFKSQVQRMHMEILQSPWLCELIAFHINVRETKANS
GKGHALFEGCSLVVDDGKPSLSCELFDSIKLDIDLTCSICLDTVFDSVSLTCGHIYCYMCACS
AASVTIVDGLKAAEPKEKCPLCREARVFEGAVHLDELNILLSRSCPEYWAERLQTERVERVR
QAKEHWESQCRAFMGVE

SEQ ID NO: 518
MVSTQSTRENPSIFFPPPLKPWLLPVVLSLSLSRQLGMAAAAAASLPFKKNYRSSQALQQFY
AGGPFAVSSDGSFIACNCGDSIKIVDSSNASLRPSIDCGSDTITALSLSPDGKLLFSAGHSRQI
RVWDLSTSTCLRSWKGHDGPVMSMACPVSGGLLATGGADRKVMVWDVDGGFCTHFFKG
HDGVVSTVLFHPDSNRSLLFSGSDDGTIRVWDLLAKKCASTLRGHDSTVTSLAFSEDGLTLL
AAGRDKVVSLWDLHNYACKKTIPMYEVLESVCVIHSGTVLASQLGLDDQLKVTKESAQNIHFI
TVGERGILRIWKSEGSVCLFKQEHSDVTVISDEDDSRSGFTAAVMLPLDQGLLCVTADQQFL
FYYPEKHPEGIFSLTLCRRLVGYNEEIVDMKFLGEEENFLAVATNLEQVRVYELASMSCSYVL
AGHTETVLCLDTCISSSGRTLIVTGSKDNSVRLWDSESRHCIGVGVGHMGAVGAVAFSRKR
QDFFVSGSSDRTLKVWSLDGISEDGVDSTNLKAKAVVAAHDKDINSVAVAPNDSLVCSGSQ
DRTACVWRLPDLVSVVVLKGHKRGIWSVEFSPVDQCVLTASGDKTVKIWAISDGSCLKTFEG
HVSSVLRASFLTRGTQFVSCGADGLVKLWTVRTNECIATYDQHSDKVWALAVGKKTEMLAT
GGSDAVVNLWYDSTASDKEDAFRKEEEGVLKGQELENAVSDADYTKAIELALELRRPHKLFE
LFSELCRTREVGDRVERILSALSGEEVCLLLEYIREWNAKPKLCHVAQSVLSQVFRILSPTEIV
EIKGIGELLEGLIPYSQRHFSRIDRLVRSTYLLDYTLTGMSVIEPEADRSAVNDGSPDKSGLEK
LEDGLLGENVGEEKIQNKEELESSAYKKRKLPRSKDRSKKKSKNVVYADAAAISFRA

SEQ ID NO: 519
MDSAPRRKSGGINLPSGMSETSLRLDGFSGSSSSFRAISNLTSPSKSSSISDRFIPCRSSSRL
HTFGLVERGSPVKEGGNEAYSRLLKAELFGSDFGSLSPAGQGSPMSPSKNMLRFKTESSGP
NSPFSPSILRQDSGFSSEASTPPKPPRKVPKTPHKVLDAPSLQDDFYLNLVDWSSQNTLAVG
LGTCVYLWSASNSKVTKLCDLGPNDGVCAVQWTREGSYISIGTSLGQVQIWDGTQCKRVRT
MGGHQTRTGVLAWNSRILASGSRDRVILQHDLRVPNEFIGKLVGHKSEVCGLKWSHDDREL
ASGGNDNQLLVWNQHSQQPVLKLTEHTAAVKAIAWSPHQNGLLASGGGTADRCIRFWNTT
NGHQTSSVDTGSQVCNLAWSKNVNELVSTHGYSQNQIMVWKYPSMAKVATLTGHSLRVLY
LAMSPDGQTIVTGAGDETLRFWNVFPSAKAPAPVKDTGLWSLGRTHIR

SEQ ID NO: 520

MEDEAEIYDGVRAQFPLTFGKQSKPQTSLESVHSATRRGGPAPAPAPASSSSLPSTTSPSAA
GGAGKSSGLPSLSSSSTAWLEGLRAGNPRAGREAGIGSRGGDGEDGGRAMIGPPRPPPGF
SANDDGGGEDDDDDGDGVMVGPPPPPPGNLGDGDDDEEEEEAMIGPPRPPVVDSDEEEE
EEEEENRYRLPLSNEIVLKGHNKIVSALAVDPTGSRVLSGSYDYTVRMFDFQGMNSRLSSFR
DFEPVEGHQVRNLSWSPTADRFLCVTGSAQAKIYDRDGLTLGEFVKGDMYIRDLKNTKGHIT
GLTWGEWHPKTKETILTSSEDGSLRIWDVNDFKSQKQVIKPKLARPGRVPVTTCTWDREGK
CIAGGIGDGSIQIWNLKPGWGSRPDIHVEQAHADDITGLKFSSDGKILLTRSFDDSLKVWDLR
LMKNPLKVFEDLPNHYAQTNIACSPDEQLFLTGTSVERESTIGGLLCFFDRSKLELVSRIGISP
TCSVVQCAWHPRLNQIFATSGDKSQGGTHVLYDPTLSERGALVCVARAPRKKSVDDFELKP
VIHNPHALPLFRDQPSRKRQREKILKDPLKSHKPELPMNGPGHGGRVGASKGSLLTQYLLKQ
GGMIKETWMDEDPREAILKHADAAEKNPKFTRAYAETQPDPVFAKSDSEDEDK

SEQ ID NO: 521
GATTTTAAGTAACTCAATTAGCAGTTCCAACATTAAACCATTATTATTACCCCTTTTATC

SEQ ID NO: 522
CTCAAAAAGTACTTGGATGCGTGCGGTGACAACGGACTCGAACCGTACACTGTCAAATC
T

SEQ ID NO: 523
TTGTCAAGTTGCAGGACGTAGTGCACAGTGAGAGGCGTCTATATCTAGTTTTTGAGTACT

SEQ ID NO: 524
GAAGAAATTATATAACTAGATACAAGGTTAGCTAGGTATATAATAGCGGTACAAGTCTTT

SEQ ID NO: 525
GGACAAATCAAGTAGAACTTCTCTCGGCAGCATCAGTTTTTCTAATCCATGCCTTGTTGC

SEQ ID NO: 526
CTCAGTTCTGATAATGCCTCGGATATATGGCCGAGTGTTCGCTGGACGGCCTCTTATGTT

SEQ ID NO: 527
GGAGATTCTGAACTGCAACAGCTCCTACACATTTTCAGACTGTTGGGTACTCCAAATGAA

SEQ ID NO: 528
GACTGGTAAAATCGTTGCACTAAAAAAGGTCCGGTTTGACAACTTGGAACCTGAAAGCGT

SEQ ID NO: 529
AAACACCAATCTATCAACACTGTCGAGTTTAGTCACTAGTAGAACCGGAGATAACAAACA

SEQ ID NO: 530
CTATGATCCTGAGCGCAAGCAAGTTATGACCAATAGAGTCGTTACACTATGGTACCGAG
C

SEQ ID NO: 531
TGTTGTGAAGGTAGTTATAGCCATCGATTAGACAGTGATTAAAGTAGTACCCGTGCCAAT

SEQ ID NO: 532
CCACATACAAGAGTTGTTACGCTACACATCCTATACCATCAAAGGAACGTTGGAATGCCA

SEQ ID NO: 533
TATGATCGACACAAGCATTTTGTGTTGGAGCCTCAGCTAATTGTATGTCATCGAGTACTT

SEQ ID NO: 534
AAAATTTTTGCTACGGATAATGTTGTGAGGCGAGGCAGTCGAAATTACGGAGGTTGACTT

SEQ ID NO: 535
ATGCAGGGATCAAATTTGTGAGTACTACGTAAAATTTTGCTACGGAGGCGAGGCAGTCG
A

SEQ ID NO: 536
GAAGAATACAGGCTCGTACCTGATACACTGTACCTGACTGTTAACTACATAGATCGGTAT

SEQ ID NO: 537
TCCACCCTAAATGCGATACGTGAAAAGTATAGACAACAGAAGGTAAACTATTCATTACTG

SEQ ID NO: 538
AGGCTTCTAGTTGCGTTCCCCCAAACTACATGGATCGGCAGCAGGATATTAATGAGCGG
A

SEQ ID NO: 539
GAGAAAAATGACAGATTGATATCGATGATGATGACTGTCGTGTCATCAGTAGTGTGCTTT

SEQ ID NO: 540
TTTCCAATTGTAGTTCGTCTTTTATTGTAACAATAAATTGATAGATACTGATTCGAAATA

SEQ ID NO: 541
ACATTTATGCTAACTATAGGAGAACGGAGAATTGTAGCTGCGTCTCTGCTAACTACATGG

SEQ ID NO: 542
TTCTGGCTTAAAGGCTATTCTTTGTGCACAATGACCTGAGGGAGGTCTCGACAGACCACT

SEQ ID NO: 543
TTCATCCGGGTCCTGGTTATCATACTCTTATATATGTTGGGGAATAACGGTTCATATGTT

SEQ ID NO: 544
GGGTGTGCTTAATAGTTCTTATTAGTCTTAGCTTATTATCTTTGATTGGACATGCTATAA

SEQ ID NO: 545
CTTGCTAAGTAGACATGTTATATTTCTAATGCTTTGAGAACAATATTACAGTATAATTAG

SEQ ID NO: 546
AATCATCGACTAGACCGATGGTCAAAGTGGTAATCATGTAATTAAACGCGTTTGTCATTG

SEQ ID NO: 547
ATGGAAAAATCTATGGATATGAAGGATTGAAGATATCCGTCTGGGTAAGCTGTGTATCAT

SEQ ID NO: 548
TTATGATTTGAGAAAACCCTTGCAGGCTGCGATTTGCGGATCATGACAGCATAGTTTTGC

SEQ ID NO: 549
GTTTTGTTGTGAGGGCTTGGTAGGTTTTCATTATATTGTAATGTCGACGACAGAGATTTT

SEQ ID NO: 550
CCAATTAATGTTACTGCTCAAGCTGACGTACCTGCGAAAAAAGCACCAGTGACTGCTAAT

SEQ ID NO: 551
TGATGTCAAAACGTAGCTCTTTTTTGTGTGAGCTATCCTGCTAAATTAAACCTCAGCAAA

SEQ ID NO: 552
ACATGAGTATTATGAATACTTCGGTCCTGACTATACACTTCATGTTGCTCCGAGTAACAT

SEQ ID NO: 553
GAATTGGCGATCACAATCTACTGTAGTCAATACTCAAGTGGGAGGTGTAAATAGATTCCA

SEQ ID NO: 554
GATCATGTGTAATCAGTATATCAGGTTAGAAACAGTACTCTTGAGCTTAGCGGGCACTGT

SEQ ID NO: 555
TCCTGTGAAGGTGGTCGACTCAATCAAAAGGTACCTTGTAGATAAGGTACCTTTTCTCAA

SEQ ID NO: 556
GCATTTTATACGACGGATAGAGTCATGACCGTATCTTTCCATAAGTTTGGGGACTTCTTC

SEQ ID NO: 557
CCTCGTTTCTTTGCGGTTCGGACGCATCATGGATGTATCTCCAAAGAGTAATCTGTCGAT

SEQ ID NO: 558
AATTCAGATCTATTAGTGAAAGTTGGCATGAGTCTCAATCTTAGGGGAATACAGTACGGA

SEQ ID NO: 559
TGATATGAGTATCATAACTCGGATGGTGACAACTTTGTACTACGGTCGGCACCGGTAGAT

SEQ ID NO: 560
CATATACAATCTTAGTGGATTAGCTGAGGTCGAAACTGACAAGAGTGATCGCCCGTTGG
A

SEQ ID NO: 561
CATGGCTAACGCTGGCCCTAGCACTAATGGGAGCCAATTTTTCATATGCACTGTAAAGAC

SEQ ID NO: 562
AACAAAGTCTACCTTGACATTAGCATCGGTAACCCTGTCGGGAAACTAGTCGGAAGAATT

SEQ ID NO: 563
TGTGCTTGGATATACTGTATAAGCATTCTATATTATGCTTGTTGGCTTCGTTTTGAGGGA

SEQ ID NO: 564
TTAACGTCGACCGCTTCTCTGCCCCTTGAATTTTCCCGAGAAAACCAGGAACCTGCCAAA

SEQ ID NO: 565
TGTTGAATACGATGTATTATAATGTTGGTGTCTTGGTGAAATACAGAATTATGCTTGCGT

SEQ ID NO: 566
ATCGCTGTGGCTGATCTCGTCGCTCCGGCTTTTCATAAAAATCATGGCTGAGGCAATCG
A

SEQ ID NO: 567
CTCGCAACCCTATATCTCGCTCAGGCGAAGAAGTCTGAGGATTTGAAAGAGGTGACTCA
C

SEQ ID NO: 568
TGTTTTTGGGTACACGCAGTTAGGATAACTAGCATGAAAGCCCGATCCCGCATATACAG
G

SEQ ID NO: 569
GAGGACTAGCCGGAACTTCATCGAACTCTCTCGGAGGGGTTACTACGATAACGTCAAGT
T

SEQ ID NO: 570
GATGGCTAGCACTGTGTAGAAAGGTGAATTTAAAGTACTTGTCTACACTGCTTATTAAAT

SEQ ID NO: 571
TGAGACTGTCTTGGCGTGTATTTTGGAATAAACTATTATCACGTTTTGTTAAATATAATA

SEQ ID NO: 572
TTACAAAATGGCTCTCAGAAAGTATCGAAAGGCCCTGCGCTATCTGGATATCTGCTGGG
A

SEQ ID NO: 573
AATTTTATGTTTGCTACTGCTTAGTGCTTAATGGACTTGCGTAGGTATTCAAATTACAGA

SEQ ID NO: 574
TGGAACCGTGGTATCGGCTGACGTTATCCGTGATTTTAAGACTGGAGATAGTTTATGCTA

SEQ ID NO: 575
CTTTGATGTATCCTCAGTGTACTGCTTTTAGCTATGTATAGATCGAGTCAACTCATTGAA

SEQ ID NO: 576
TTTTTATTATTTACCTTCGCCTTTACGCTGCATACGTTAATAGGTTATTATTTCCTTCAA

SEQ ID NO: 577
ATTTGTCCATGACAATCGTAGTCGAAGACACGATACGCTCTTAGATGGTACGGAAATCTG

SEQ ID NO: 578
TGAATAGAGATAACTTTTCTGAGTGTGAATTGGATATTACGTTGCAAATAGCCGAATGAA

SEQ ID NO: 579
GCTTTAGGTTAGGGATCCCTGTAAGCTGATGATAGATATTGGAGATGGTACTTGTAAGAT

SEQ ID NO: 580
TGTTGTGTTTGGAAAGGTGCTGTCTGGGATGGATGTTGTCCACAAGATTGAGGCTGAAG
G

SEQ ID NO: 581
GGAAAGCGGGGAATGAGCATGTGGATATTATCTCTTTCTACAATGAAATATTCATTCCTT

SEQ ID NO: 582
CATCAGGACGTTGACTCTAATTAAGACATATGTGACAGAGCGCCCTGTTAATGCGGTTAC

SEQ ID NO: 583
CTTTAGGTTTGATCTGTCTGTTTTGTCTATCCTGCGAGTTTCGAGCATGTGCGTGTGTGA

SEQ ID NO: 584
CAGCCCCAATAGATACTGGCTCTGTGCCGCTACTGAGAACAGTATTAAAATCTGGGACC
T

SEQ ID NO: 585
AAGAATGAAGCTGATATGAGTGATGGAACTACGGGGGCCATGAGCTCAAATAAGAAGGT
C

SEQ ID NO: 586
TGACTACAATTAGCACCTCACCATTATCGAACTGTATAATTGTGCTTGCCTGCTATTATT

SEQ ID NO: 587
TTGAAGCGGAAATATATATTTATGCTACTACATAAGTAATGTACTACTTGACAAGATGAG

SEQ ID NO: 588
TACTCGATGTGGTATAGAATTTATCCAATGTACTCCTAAATGTAGATACATCGTGTATTG

SEQ ID NO: 589
GCTTCGTCTGATACCACTATCAAGATAATAGGCGTGAGCAATAGCTCTGGATCACAGCAC

SEQ ID NO: 590
GGTCGGCTTGCTAGTGTATCTGATGACAAGAGCATATCACTCTATGATTACTCATGAAGG

SEQ ID NO: 591
GAAAGGAGAAAAGCATGGAGATCGATCTCGGAAACCTCGCATTCGACGTCGATTTTCAT
C

SEQ ID NO: 592

GATTCAGTACCCGGATTCGCAAGTCAACCGGTTGGAGATAACTCCACATAAGCGGTACC
T

SEQ ID NO: 593
TTCCATGTATCAAGCCGCATCAATGTTTGTCGCTGCAATTAACATGTGTGCAGTCGATCC

SEQ ID NO: 594
TTCAGCGCATTGTGTAAATGTAGATAGGTGATATATTTCTCGTTGCAATGTAGGGTAAGA

SEQ ID NO: 595
TCCAATAATCACATTTACCATCAACAGGCATCAGCAACATACTGTTGTAGTGTAATTAAT

SEQ ID NO: 596
GGGCATTCTGACTACCTGCACTGTATAGCTGCACGGAACTCTTCTAGTCAGATTATAACA

SEQ ID NO: 597
AATCGTCTGGTAGATTGTCAAAAACTAATAAACCTGTGATTGATCCGGATTCTAGTAATG

SEQ ID NO: 598
AGTTGAGGATTCTCCACTATGACAGCTCTCATGGCTTGAATCTAAAGTCATCTGGTTTTC

SEQ ID NO: 599
GAACAATCATTCTGTAGAACACTAGAGTCTATATGCTTGACTGTATCGGTTAATTAATTC

SEQ ID NO: 600
AGATAGCGATAGAGTTATACTGCATGTACTGAGGTAAATGTTTTGATTACTCCACCCAAT

SEQ ID NO: 601
AAGAATTGTTAGGAGGTGTATACTTTCTGTAACTGTATTCAATGAGCATACACCTGACGG

SEQ ID NO: 602
CAACTCATATAATGACTGGATTCTGGCAACCGCGTCTTCAGACACAACAGTTGGACTATT

SEQ ID NO: 603
AGTGTAAAAGGATGCCCCTAATAGATTATATGCCAAGTGTAGTATATATAATAGTGCTTT

SEQ ID NO: 604
AAGAATCTACAGTTGTCTTATGCTACTCTATTACTCAATTATGCTGTGCTATTGATTGAG

SEQ ID NO: 605
TCTGAATACATACTTTGTGGTCTCTATAAAAGACCAATGATACAGGCATGGTCATTAATT

SEQ ID NO: 606
TAAATCTTCTCATGTGCCTGGCGTAAATTTTGCAGTTATTACTAGACCAAGATAGTTTCA

SEQ ID NO: 607
ACATGGATTCGATCAATCGCCACATGACAACTAAAACAAGCGGTTCACGTGATTGTAATT

SEQ ID NO: 608
AGATGAGTATGCTCGGGTGTATGATATTCGCAATTACAAGTGGAATGGATCGCATAATTT

SEQ ID NO: 609
TCTTTGATTCTGTTGTATGGTGTATCTTATTGTATCTTCTATCTGCCCCCCATGTAATTC

SEQ ID NO: 610
TTCGTTGTGTAGTACTGGGAGTTACTACTTGTATGTATGTAAATCATGTGGCGTCTGTCC

SEQ ID NO: 611
GGAGATGTGTAATATGTCTGAGCGGTCACACTCTAGCTGTTACATGCGTAAAGTGGGGA
G

SEQ ID NO: 612
CCACCGTTGCGTAACTCGAATAGCCGGATTTTCGTTTTCGTTTTTATTTCCCCGTTAATT

SEQ ID NO: 613
TGAGATGCTCTGTGTGAGGACTTTTACGAAACTTGAATGGCCCGTAAGGACAATAAGCTT

SEQ ID NO: 614
TGGGTTGTTGCGACGGGTTCTACAGATAAGACTGTTAAGTTATTTGATCTACGCAAGATC

SEQ ID NO: 615
GCAGAGGTGCCTACATATGCTTTAGAATGCTAGTAGCTTGGAAGTGCAACACGCTCGTG
A

SEQ ID NO: 616
AGTAAAGTTTAACGACTATGCATCTGTCGTAGTATCAGCCGGCTATGATCGTTCAGTGCG

SEQ ID NO: 617
CGTTAGGATAGTCTTTAAAGGAGTTGGTGATTATTGATTTCCACCCAATATATGTAGCGT

SEQ ID NO: 618
GAGCAAGCTACTTACAAAAATCGACAGCGTCTTTACCTATCTGAACAGACAGATGGCAGT

SEQ ID NO: 619
TCCTTCCGACAAGTACCGTATTGCAAGTTGTGGTATGGACAATACGGTTAAAATCTGGTC

SEQ ID NO: 620
TTTCACTCGATGACGGTTGGCCGGATAAATAATCGCTTATATAGTCCTAATAAGTTCCAT

SEQ ID NO: 621
ATATGTAGGTGGTAGAGGTGTGGATATTGCATAGACCGAACCTCCGCAGGTCCGCATTC
T

SEQ ID NO: 622
CCATTGAACTACTTATGGATTACTTTATACATGAAATATCATGCCGGAGTAATTTTGAGT

SEQ ID NO: 623
AGCATTAGAGACCTGGATTTTAGTCTAGATTCAGAGTTTTTGGCTACGACATCTACTGAT

SEQ ID NO: 624
AAAGGTTTATCCCTCATTGGATTTGATATATAAACTGAGAGTGTTTTGCCCCCCATTAAA

SEQ ID NO: 625
GTACAGCGTGTATTTCTTGTTACGATACTTGAGGGGTTAGAGGCACCTACGAATTAGGAA

SEQ ID NO: 626
ATATCCTTATGAATGAAGTTTGGATGATAAGTGGCGCCAGACTTTCTACTCACCCTTTTT

SEQ ID NO: 627
TGATCACATCGTTGTTTGCAATAAGACGTCATCAATTTATATCATGACTCTACAGGGACA

SEQ ID NO: 628
TTTTCCCAGTGTACTGCGAGAGTGATGCTACATAAGTTTACTCTTGTGTCTAACTTTTCC

SEQ ID NO: 629
AGATTCTACAGATGGCGCTATACGAGCTGTTATACGGACATTTTATGACCATACACATCC

SEQ ID NO: 630
TGCTACGGGAAACCAGGACAAAACTTGTAGGATTTGGGACATACGAAACTTATCTAAGTC

SEQ ID NO: 631
CAAGTCATATAGTTACAGTGTCGCATGACAGAACAATTAAGCTCTGGACTAGTAACGACG

SEQ ID NO: 632
TGCCACATCGTAACCATCATAGCACTTATCATCTAATTATGGTGAAAGGGAGTTATATAT

SEQ ID NO: 633
GTTTATACTTATAAACAACAGAGAGACAACTGTACAGGTGTTGTAAACACTCCCAGTGTG

SEQ ID NO: 634
CTGTGTTTTAGCCCGAGGGCCAATCACTTAGTTGCTACTTCGTGGGATAATCAGGTACG
G

SEQ ID NO: 635
GCAAAGTAGAGTTTAAGTTTCGTTGTGCTTGGACCGGAAAACTCACATGCTTAGAGTTTA

SEQ ID NO: 636
AAGATTTGGGCATAACTTGTATGAACTTTTTCTGTTGTCGACACTGTAATTACACGAGCT

SEQ ID NO: 637
AAACAGATGCATGTATGCTTCATAACTCTATAGATATGGAAATGTCACTGTACACTGATC

SEQ ID NO: 638
TTATTGGTGCACAGGACGGAAAATTGCGCATATATTCTATTTCAGGTGATACATTAACAG

SEQ ID NO: 639
AGGCACAGACACTTGCCTAAACCAATATACAAGGCAGGTATTCTAAGGCGCACCGTGAA
T

SEQ ID NO: 640
CATGCGAAGGTTTCTGGGAATTTTCAGTAGAAAATTCGGTCGTGGCGGCCATCCTCGAT
A

SEQ ID NO: 641
TTAAGCTGATAGCTTTAGTTCCTACGTGGAATGTATAAATGCACCATTGTCCATAAGGCA

SEQ ID NO: 642
GGATGCTCTGGTTACATGACTACTCCTTAGGGAATCAGTCAGACATTTTAAATAACTTCC

SEQ ID NO: 643
TCATTAAGCGGTACTGGCAGAGGACATGTCTATTTATACAAGCAAATGGTCCTATTGGCT

SEQ ID NO: 644
ATGTTGGTCAGACCTCAAATATTGTACTCCCCACACTAGGGAGCATTTACGGTGAATATA

SEQ ID NO: 645
TCCTCTCGACCCTTAGAGTCCTCTGCGAATCTTGTTGTTAGTTACTGTGTACGCTGTAAC

SEQ ID NO: 646
AAGCATGTTTTGAATTTATGGTGGTGGCATGTGGATATTTGAACTTGGTTGAGAAAAATT

SEQ ID NO: 647
CATTCCTATTGAAGGGTCAACCTTTAATTTTGGCTAGCAGGACTGTATAGGATTATATGC

SEQ ID NO: 648
TTATTGTATTTTAGATTCTTGATGGCCATCTAAACTTCTGGCTGCTTGGTGCAACATTGA

SEQ ID NO: 649
ATAGCTAATGATTCCATGCTATCCATGGTATCTACTTCACGATAATAAAGGTCTTAGTCC

SEQ ID NO: 650
CACCTAATAGGCCTGAGTATTGCTCACCACTATGCTGATATGGGGAGCAATAACGTTAGT

SEQ ID NO: 651
TTTCTTTTCACTTTGTACTAATGATCATTGTGACCACAAAATCTTTATACACAATACAGA

SEQ ID NO: 652
CTTGTCACTATCCTCATATTGATATCACCTCGTGTATGTTGTGGGGTGGCAAAATTACTT

SEQ ID NO: 653
TATTTTAACTCAGCGACTTACCAGCCTAGTAAGCAATGGGGAGCTTGCATGTATTAGTTT

SEQ ID NO: 654
ATTCGTCCTGGTCCTTTAGGACATGTACTTATGTCCATGCAAGTGCTTCTTGCCTAAGCT

SEQ ID NO: 655
TTCTAGGCGATATATATCGCCGTAACTTTGGATGTGTTAAGAATATAGGGGATCATTAGC

SEQ ID NO: 656
AGTTGCAGAGTGTGTAGCAACTGATGAGCATAGTTGTTATGTTTCTCAACTCAGTTGCAC

SEQ ID NO: 657
AAGAAACTCATACACTGGACAGGCCAACCTTCCAAATATGTGTTTAGAAAACCTTTGTCT

SEQ ID NO: 658
AAGGGGTGCTATCCATATCTAGAATCTACCATGCTCAATGAGGTATCTTCATTAGTATAC

SEQ ID NO: 659
ATCTAATGCTAGTTTATTGATTTCTATGATCCAAGACCTCGTCATAGATCAAGTGCCTAG

SEQ ID NO: 660
TTGTTATTAAATACCATTCAATATGCTTATGATTCATGAATGCTTAAGAGATTCTGCTGC

SEQ ID NO: 661
GCTTCTAAACTGTAGAAGCCTGTTATCTTTAGACTCGTGGTTATGTGAACTACTTTTACA

SEQ ID NO: 662
GGCTGTGGGGATTCGAGCCTGATGGTTATGCACTGTGGCCAGCAAGATGTTGAAGTTTT
A

SEQ ID NO: 663
GCCTGATGGTTATGCACTGTAAGTGATCTGATTTGATTAACTATTTTATCAATTAATTTT

SEQ ID NO: 664
ATGGTCATTATCCGAGATAGTGCGCTTTGTCATGGGAAAATGACTATTGAATGTGAGTTT

SEQ ID NO: 665
TTTTCTGGTGCATCCTTAACACAGCTTGGTTACATGGTGAATTACAGTATTTGAAGGAGT

SEQ ID NO: 666
AGATTTAATGCCACTTAGGTGATCGGTGACCCACTTGTACATATAGATGTTGGCGATGTT

SEQ ID NO: 667
AAGAAATTCATCAATTCTTTGAAATTATTGTTCCCTTTTGATGCGGCCCCTTTCTGGAGG

SEQ ID NO: 668
TAAAGTATATTTTAGCCGCTGTTGTTGTAAATTTATGTTTTTCATTGCTATCAACATTTA

SEQ ID NO: 669
GGTTTTCCTATAAGATGTATGAATTCGCACTGTGGTGCAATTTTATGAATTAAACTCAAA

SEQ ID NO: 670
TTTACTATTCCGTCTGGGCTTAGAGATGTACGTTAATTGGTCATTTAAGACGACTCAGTT

SEQ ID NO: 671
TCAAATCTAGTCAATATCCGTGTTGAGCTAAACAAGCGCTGAAAGTTTGCTCGAATCAGC

SEQ ID NO: 672
AGAAAGTTGTGTACTAATTTGTATTGTAACGTCCATTTATCCAACGAGTCCTCCATTCAT

SEQ ID NO: 673
CAGTACTGTATTCGAAGATCCTGAAAATTTACTAAAACAAATGGAATATCAACAACCTAG

SEQ ID NO: 674
TTGCTCTATATAATTTGTGCTCGTGTGTGTACTTGAAGATCCATCCTCACATAGTCCAAT

SEQ ID NO: 675
GTGTGTATAGTTTTATAACACTCTATGGTATCACTACCACTATGGGCCTGTTTAGTCCAA

SEQ ID NO: 676
GAAGCAGAATCAGCTTTGACCAGTATTTAGTGTCTTGTATACAATTCTTGTTTCAGTGAA

SEQ ID NO: 677
AAATCAAGATTAAAATCCGAAACCAAGGCTAACCAGCAAACTGTGAGGTGTACATTGTTG

SEQ ID NO: 678
TTCCAAGCAGAAGGGCACATGTTGTGACATCAAGTAGTAGATTGTTCTGCAGATTCTGGT

SEQ ID NO: 679
GTTAATGTAATACATTTAGTTTTTAGATAACTGTTAATGTGTAGTAAAGCACTAGGAAGA

SEQ ID NO: 680
GAGGCTTCAAAGGTTTTTGTGTCTTTTCTAGTTATTATAAACGCTTCATAGGTTCCTAGG

SEQ ID NO: 681
GAAGATTGTAAGTTGGGTGAACTTTTTTACCACGCTAGGTTGATCTATTTTAAGACTCTT

SEQ ID NO: 682
AAAATAGCTGCGCGTACCACAAAGGTGACAAACGCCGGATTTCTCTTATCAGACTTGTCA

SEQ ID NO: 683
TTTAATTATCATAGTTTTATTCCGGCTATCTTGATCATTCACGGAAGTCCCGAGAGTCAA

SEQ ID NO: 684
GTGGAGTGAACGTGGTTACTTCAATGGATTACCCTTCTATCGTGTCATTAAACACTTTGT

SEQ ID NO: 685
GCTAACTCTTCTAGTTGAGATCTCCATCAATTAATGGATACAAACATTGAGTTTCACTTT

SEQ ID NO: 686
GGATCACTACTGGATTCCGTTACATTAGTTATTGCAAGTTGGTTATTATGTACGTTTATA

SEQ ID NO: 687
ATGAACAAATGCAATTACCCTGTTTTATTCTATCCCGCTTTAATTAATATTGGTCATGTT

SEQ ID NO: 688
TTTGCTTGTGGATTGTACTGTGGTACATGGTATAAATCTATAGGCTATGTCGATTATTTT

SEQ ID NO: 689
ATATAAGATATAAGATATTGCCAGCAAACTATTTGACAGGTTATTTAATAAAGTGTGCTA

SEQ ID NO: 690
TTTTAAATGTGGACAGAGGCACTATAAGAATGCGAAATATCGTCGGAGCACGACTAATTG

SEQ ID NO: 691
ATAGACTAGTTCTACAAAGCCCTAGGATGATGGACTTCATTTCTTTTGCATTAAGATGAA

SEQ ID NO: 692
GATTTCTTATGGGGTTGGAACATTCCTCGCTGCCTTCTGGTAATATTAGGTTATGCGTTT

SEQ ID NO: 693
AATTGAGGTTGACTGTGTACTTCTCCAGTGGACAGGAGAAAGCGATAAAATTCAAACGTT

SEQ ID NO: 694
AAGGAAGGGCAAATAGAGCTCGCGCTCAAGAAATACCTTAAATCGATACGGTATTTGGAT

SEQ ID NO: 695
TAATTTAAGAGCTATGAAACAACTACCTTTTGGAATGGTTTTGTTTTTAGCATCCCAATT

SEQ ID NO: 696
TTGTAAATTATGCTGGTTCCATATGGGGGTTAATCAGTATCCTGGTTATTTGTGACACCA

SEQ ID NO: 697
GTTGTGAACTATCAATAGACGGGGATGGTCCTTTTTAGCTGCTCCTTAAGCAGCTCAAAT

SEQ ID NO: 698
TCAATTCCGGTCATATGTAGACGACTATAATGTTGTTTGTGTCCTATAACTATAGTGTTG

SEQ ID NO: 699
CATTTTACACCCTATAACAAAATATAGTGTCATAAGTTTACACCAGGTAACAACTCTATA

SEQ ID NO: 700
ATGGAGAGTTTTATTCATTACATGAAAGAGTATGTCACCTTTCGTGCTCCATCTATTGAT

SEQ ID NO: 701
TTTCACGTCCTGTATACTCACTCAAGCAACTTTAGGATGAAGAGCTAAAGTATATCAAAG

SEQ ID NO: 702
AATGCACTCTTTATAAAGTGGGATGAGGTATGTGTTTCCTTCCTATTGGCTAACCTGAAT

SEQ ID NO: 703
ATTGGGCAATCGTTATTGATTTTACCTATCGCTATCTCACTGTCCGCCAATTTAGTGTAA

SEQ ID NO: 704
TTTCAGCGGATATAAAGTCTTCCAACTTGTAAACCGGTGCTGTGAAGATTAAAAGTCCTT

SEQ ID NO: 705
GCTTTAGAGGCAATGGTAGATTATGAAGTCAACACCAGGGAGTTTGACCGTTTGGGACA
T

SEQ ID NO: 706
CATTCAATTTGACATTGGAGTTTCAAGGCATTCCAAGGATAGCATGTACACAAGTTGAAT

SEQ ID NO: 707
CATAAAATTACTATGGAAGTTGGATCATTATCTATGCCATAGTGGAGTAGAACTAGATTT

SEQ ID NO: 708
CTCTTGATTCTAGAATCTAAACTACTACCTTGCGGACATGACTGAGCATCTCTCTAACAG

SEQ ID NO: 709

CAGGGTTGTGCTAGTTTAACATTTTAACTTAATGTAATCATGTAAGCTTTAGAGAGGTGG

SEQ ID NO: 710
GTAAATGTTTACATTGAGGTCATGCATGAGTGTTAATTACGCTTTCACTACTGTTCACTT

SEQ ID NO: 711
AATTAAAGCTTGGTTGTATGATCATTTGGGATCGAGAGTAGATTATGATGCTCCTGGGCA

SEQ ID NO: 712
TTATCTAGCTAGAAGTTGTGAAATTAAGAGGGATGTGAGGATTGGGTTATAACTAGTGTA

SEQ ID NO: 713
AATGAATCAGGCATTAAAGCGGGAATCATTTATGACTTGGCAACCTGAAAATTCTATTAA

SEQ ID NO: 714
TTCTTGACGTTTTAATATGGTATGGTATTAAATTTGGAAGGCCTATTCGATTGTTTGCAA

SEQ ID NO: 715
TTCTTATAACCTGTACGATTGCCGATATATCACCAATTTTGCTGATTTTAATCTGAGTTT

SEQ ID NO: 716
CAATTTCATATTCGGGTTCAATGTAGTGCCTCTCATTTTAGGGTGATAGCATGAGTTTTT

SEQ ID NO: 717
TCCACAAGTTAACATAGGTAACTATCGACTGAAGTGAACTGGGGGGCAGAAGCTAACTA
T

SEQ ID NO: 718
TTTAGATAGCCATTTACATTTTACTTATTATTGGACTTGTAAAGATTTTTGTACCCTTGT

SEQ ID NO: 719
TTGCTGAAATATTTCAAGCTGAAAGTTATGATTCTGGCCAAGAAGTCTACTGAAAATTTG

SEQ ID NO: 720
AAACATAAGTTTGGCCCAGATTCGGTTTATCATAAAATCTGGCTGCATATAAGGTGTCAG

SEQ ID NO: 721
ATGTTCTAGAATTTGTCTAAGCTAGCTACTGGTGTTTAACTGATATGGAAAACTTTTGCC

SEQ ID NO: 722
TTTGGGGAGTACTTTAGTCAATAAAAGTGAAGTGAATCATGATATAAAGGGTTTAAGTAA

SEQ ID NO: 723
AGAAGTTACTAATTTGTAGATAAATTCTAACGAAGGTGATGATAGCATACACGTAATGAA

SEQ ID NO: 724
GAATTTTGATGGTAGCGTATGGTTGAAGGAAAACTTGGATATATCATGTAAACATTTTTC

SEQ ID NO: 725
TTAATGAACCGCTTTTTCCTTGAGAGGCTATGAATGCCTGTAGAACTAATCCTTTAAGTA

SEQ ID NO: 726
TTTCTCTAACACTATATTTTCTGGTATGACCGCTCTACATTGTATATTAACCCTTGCAAA

SEQ ID NO: 727
TATATTCACTGTGCTGGGATTATCCTCTCCCCTTTTTGACCCACTGTTGTGTGTATTTGA

SEQ ID NO: 728
GAGCATACAGCGTTATCTTTGAGACGAGTCATCAATGATAATATCCTCGTAAAAGGTTAC

SEQ ID NO: 729
TTTATTCAATTACGACGGATTCAGTTGGCCTTTTGTAACATTCAAGTATCCATCTATCAC

SEQ ID NO: 730
ATGTTCAGGGGTATTAAAAATTCAGAGGATAAATTTCCTCACTCTCAAGTGTTAGATGGT

SEQ ID NO: 731
CAAAGTCTAGACGTTAATGTTTTGGAACTCTTTTTTCGAATTTGTGCCTATTGAATCACT

SEQ ID NO: 732
TATAAATATATTGTACTGGGGATCCAAGACATGGCAATATATGTCGAGATTTTCATTTTC

SEQ ID NO: 733
CTTTTGCATGAGTTCAAATGTCTTTGTGACATATTGTCTTGAACCACCGAGGATATATCA

SEQ ID NO: 734
GTTTGTATGTCCAATAGATTATAACCTATTTACTGTGACACTATTCTTCACACCCATGTC

SEQ ID NO: 735
AGATCTAGTTGTTTCAGCATCGTTGGACCAAACTGTTCGTGTATGGGATATAAGTGGCCT

SEQ ID NO: 736
TGCCGTATCAAAAGATTGGTACTTCCTTATGGACACACAAGATCGTAAGCATGGCTGAAT

SEQ ID NO: 737
TTGATGGCCACATGAGTTGTTTATACAAGTCGTTGTTTTATGAGAGAACCTTCTTCAGAT

SEQ ID NO: 738
ATTTCTATAGTGCCATATGCTTGTCGGTTGTCATTGACCTCTAATAGAATAGCCAGAGTA

SEQ ID NO: 739
TTCACGGCAGTTGAACTAGTCATAGTGGAATATTATTTAAATGGTGTATTCTAGTCACAT

SEQ ID NO: 740
TGCAGGCGCTCTATAGTTCTGTTCTCTAGCATGAAGTGTGTATTTTATCTATTGTGGACC

SEQ ID NO: 741
TGTCTTTAATCTTCAGGGTTCGTTACTAACAATTGAGCTCAAATCTCTATTCTGACCAGC

SEQ ID NO: 742
CATTTATAGAGTTGTGCAAAATCACCCATAATGCTATGAATTGACAGGTGACTGTAATCT

SEQ ID NO: 743
GGAGAAAATTTCCTATCCCTTTGTGGGTGTGTGAAAAACGAAATATAGAGGAACAATGTG

SEQ ID NO: 744
ACCAATCATTTATTTGCAGTGTAGTTGATATGAAGGGAGAAATATGACAGTTGGTTTCAA

SEQ ID NO: 745
AAGTTAATGTTCTCATAGGTTATTCATTGGAGTTGTCTCGTATGTACGCTGTGCCGTAGT

SEQ ID NO: 746
CTCATAAATTGAGGCTTGCCTACGTTAATTGTTATATATGGAGAGCCATGCTAATTGTTA

SEQ ID NO: 747
GCAGATCATGTAATTGTATCTCAAATTATAGTATCCGTATTCTGTACAAATGCTCCGGAA

SEQ ID NO: 748
TCTTTACGCAGATGGTGACTGAAGCTGGTTCCGAGATCGGCATATGTAGCTGGTAGAGG
T

SEQ ID NO: 749
TTCACATTGAGGGTTGCCGTCGGTATTCGCCGATGATATCCTGTTTTACGCGCAACAGTT

SEQ ID NO: 750
TCATTATTTAGGGTGCAGGCTGTATAAAATGTTGTAAATTGTAGTATCAATGTGTACAAT

SEQ ID NO: 751
GCATTCACCACGACAGTAAAGTAATCATTATGATTACTAATGTATTGCTTTCATGGGGTG

SEQ ID NO: 752
AAAGGGTATATTTTGTCTCATGTTGGGGTGATAATTCTCCCTGAAAGTCTCCAAAATATA

SEQ ID NO: 753
AAATTTCCGGTTGCCATAGTCTAGTGGGGTGAGGGTTCATTCTAGGGGATTTATTGTGTT

SEQ ID NO: 754
GCAGTGATAAAGGTACTTCTTGGTGATAATCCTAAAGCCTTACCCATGGATATCCAGCCT

SEQ ID NO: 755
TTCTTTAACAAGGTAAAAATCCCCCCCTTGGCATGTAGCTCAATTAGTTGTAATGGAACT

SEQ ID NO: 756
AGTTGTAAACAGTGTAATAAGGAGCAGAAGTTGTGATAGCTTTTAGGAACGATAGACTTT

SEQ ID NO: 757
TGAACCAATTCTTGTATATTAGATATGTAACATGTATGAATGTCCATAGAGCAGAGCTTT

SEQ ID NO: 758
AGCCAGGCACGCTTAACTAAATTTCGTTTAGTTCACCATGACTATTCGTTGAACTTAATG

SEQ ID NO: 759
CAAAACCCCTTGTAGGGTGGACTTCTGTTGTATCCAATTTTTATGGCATAATTAGCTAGT

SEQ ID NO: 760
AATTTGGTGATTATTCCTTACCATATCGTACTGTACAGATACGGTAAGGTCGAAATATAT

SEQ ID NO: 761
CATGCCGTGATCGGTCGATTGCATTAAGTGCTGCAAGGATCAAATAGTGGCACTGTCAT
G

SEQ ID NO: 762
CAAACATAAATAAGGTTGCTACTTTAAAGGGACATACGGAACGAGTTACTGATGTGGCAT

SEQ ID NO: 763
ATTTATGGATGAGGTACTCCTTATGAATATCTTCAAACTAAGAAATAACTATATATGCAA

SEQ ID NO: 764
CTTGGTTTTTGTTGAGCTTTCTATTTCAAGCAATTTGTGATTGGGGGGTTCTGCATTCTT

SEQ ID NO: 765
ATGTCTAAAGAGCCGTGATCTATGAGTAGATTAGAAACCGCCTTTTTAGTTGCAAACGCC

SEQ ID NO: 766
TTGCAACAAGGTATACTTAGTCAGTCCTTGTTATGTATGTCTTTTGTCAACCCTTCAGGG

SEQ ID NO: 767
GGCGGAATCCCTTTGTTCTTTCGAGCTTTACGTGACAAGTCGGCCAGAAAGCAGTAGCA
T

SEQ ID NO: 768
TTGATGTACGAGCCGCTATATCTAATTCTGCCTCCCAGTCACTGCCAAGTTTTACTCTTC

SEQ ID NO: 769
GTCTTGCATGTCAGCTATTATACAGTCCTGTTTATAGTCCTGTGATGTAATAAAAAGCTG

SEQ ID NO: 770
AAGTAGGAGATCGTGTAGAGAGAATACTTTCTGCTCTCAGCGGCGAAGAGGTTTGTCTG
C

SEQ ID NO: 771
AATTGTGAGTAGAATAGGAGAAACTTTTGTACAAGATTAATACGTGTGGCATAATAAGAT

SEQ ID NO: 772
TGATGTGCAGTTTACATTATTATGGTTCGAGTATTATTTAGCTGCCCTATCTTAAGTCAT