5-25-2010

Organic Cation Transporter Preferentially Expressed in Hematopoietic Cells and Leukemias and Uses Thereof

Jeffrey A. Moscow
University of Kentucky, jmoscow@uky.edu

Xin Lu
University of Kentucky, xin.lu@uky.edu

Craig Jordan
University of Kentucky, jordan.craig@uky.edu

Click here to let us know how access to this document benefits you.

Follow this and additional works at: https://uknowledge.uky.edu/markey_patents

Part of the Oncology Commons

Recommended Citation
Moscow, Jeffrey A.; Lu, Xin; and Jordan, Craig, "Organic Cation Transporter Preferentially Expressed in Hematopoietic Cells and Leukemias and Uses Thereof" (2010). Markey Cancer Center Faculty Patents. 1.
https://uknowledge.uky.edu/markey_patents/1

This Patent is brought to you for free and open access by the Cancer at UKnowledge. It has been accepted for inclusion in Markey Cancer Center Faculty Patents by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
ORGANIC CATION TRANSPORTER PREFERENTIALLY EXPRESSED IN HEMATOPOIETIC CELLS AND LEUKEMIAS AND USES THEREOF

Inventors: Jeffrey Moscow, Lexington, KY (US); Xin Lu, Shanghai (CN); Craig Jordan, Rochester, NY (US)

Assignee: University of Kentucky Research Foundation, Lexington, KY (US)

Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 451 days.

App. No.: 11/521,487
Filed: Sep. 15, 2006

Prior Publication Data

Div. application No. 10/849,551, filed on May 20, 2004, now abandoned.

Provisional application No. 60/471,709, filed on May 20, 2003.

Int. Cl.
C12Q 1/00 (2006.01)
G01N 33/53 (2006.01)
C07K 14/435 (2006.01)

U.S. Cl. 435/4, 530/350

Field of Classification Search None

See application file for complete search history.

References Cited

U.S. PATENT DOCUMENTS

FOREIGN PATENT DOCUMENTS
WO 02/46415 A2* 6/2002

OTHER PUBLICATIONS


Hillier et al., GenBank AA03971, May 9, 1997.


* cited by examiner

Primary Examiner—Bridget E Bunner
(74) Attorney, Agent, or Firm—McDermott Will & Emery LLP

ABSTRACT

A novel organic cation transporter (OCT) gene, OCT 6, and use thereof is described. The OCT6 gene is preferentially expressed in human hematopoietic tissues, including CD34+ cells and leukemia cells. Its narrow tissue distribution, substrate specificity, and close homology to other cell membrane transporters make OCT6 an attractive target for the treatment of myeloid diseases.

7 Claims, 10 Drawing Sheets
<table>
<thead>
<tr>
<th>SEQ ID NO.</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>OCT6</td>
<td>MGRHFEQTYDVGAVLQNLSDGTYLQNFLSFGMFVTPH</td>
</tr>
<tr>
<td>4</td>
<td>OCT1</td>
<td>MRSYDQVFAFLGGEWPGFQLFIFFLSSASIPNGFGNGMSVVFLAGTPE</td>
</tr>
<tr>
<td>5</td>
<td>OCT2</td>
<td>MALQFVQQVLAELAGDFGRFQGLQHLLCANTCLNLSPFPYFFFAHVFMLSDEP</td>
</tr>
<tr>
<td>6</td>
<td>OCTN1</td>
<td>MRSYDQVFAFLGGEWPGFQLFIFFLSSASIPNGFGNGMSVVFLAGTPE</td>
</tr>
<tr>
<td>7</td>
<td>OCTN2</td>
<td>MRSYDQVFAFLGGEWPGFQLFIFFLSSASIPNGFGNGMSVVFLAGTPE</td>
</tr>
<tr>
<td>8</td>
<td>OCT3</td>
<td>MALQFVQQVLAELAGDFGRFQGLQHLLCANTCLNLSPFPYFFFAHVFMLSDEP</td>
</tr>
<tr>
<td>9</td>
<td>OAT4</td>
<td>MALQFVQQVLAELAGDFGRFQGLQHLLCANTCLNLSPFPYFFFAHVFMLSDEP</td>
</tr>
<tr>
<td>10</td>
<td>OAT3</td>
<td>MALQFVQQVLAELAGDFGRFQGLQHLLCANTCLNLSPFPYFFFAHVFMLSDEP</td>
</tr>
<tr>
<td>11</td>
<td>OAT1</td>
<td>MALQFVQQVLAELAGDFGRFQGLQHLLCANTCLNLSPFPYFFFAHVFMLSDEP</td>
</tr>
<tr>
<td>12</td>
<td>OAT1</td>
<td>MALQFVQQVLAELAGDFGRFQGLQHLLCANTCLNLSPFPYFFFAHVFMLSDEP</td>
</tr>
</tbody>
</table>

**FIG. 2A**
FIG. 2D
FIG. 4
FIG. 5

OCT6 RNA levels (relative to MOLT4)
ORGANIC CATION TRANSPORTER PREFERENTIALLY EXPRESSED IN HEMATOPOIETIC CELLS AND LEUKEMIAS AND USES THEREOF

FIELD OF THE INVENTION

The invention relates to a gene encoding an organic cation transporter, OCT6, and its use as a target for the treatment of hematological malignancies, and in particular, leukemia. The invention further relates to screening methods for identifying agonists and antagonists/binding partners of OCT6 transport activity.

BACKGROUND OF THE INVENTION

The lipid bilayer of the cellular membrane insulates the intracellular milieu from exposure to hydrophilic compounds. Unlike lipophilic compounds that can diffuse through cellular membranes, water-soluble compounds usually require specific transport mechanisms to gain access to the intracellular space. The regulation of the traffic of polar compounds in both directions across the cellular membrane is a complex process involving several large families of transport proteins.

Most often in cancer research, drug transport is thought of as a mechanism of cellular drug resistance, as drug efflux pumps such as the products of the MDR1 and MRP genes have been shown to be mechanisms of resistance to lipid-soluble anticancer drugs. However, drug transport is a two-way street, and mechanisms also exist for pumping drugs into cells. For polar, water-soluble anticancer agents, drug uptake, and not drug efflux, is the critical determinant of cellular drug accumulation.

Most cancer chemotherapy employs drugs that are lipid-soluble that can easily penetrate the cell membrane of cancer cells. One advantage of using lipid-soluble drugs is that they easily gain intracellular access to different types of cancer cells, so many cancer cells appear to be initially sensitive to these drugs. The disadvantage is that cancer cells learn to increase the activity of drug efflux pumps in the cell membrane to pump lipid-soluble drugs out of the cell, resulting in drug resistance.

In contrast, potential water-soluble anticancer drugs may not survive the preclinical screening process since there is a great deal of variability in the expression of drug transport genes in different types of cancer cells. Variability in transport gene expression may result in variability in accumulation of polar, water-soluble drugs. One approach to more effectively utilize water-soluble anticancer drugs is to identify which of the dozens of transport genes are actually expressed in tumors.

The importance of carrier-mediated anticancer drug uptake is exemplified in reduced folate carrier (RFC) mediated uptake of methotrexate (MTX). Methotrexate (MTX), a reduced folate analogue, is scavenged and retained in cells by mechanisms designed to secure folates from the environment. The major mechanism of MTX uptake at pharmacologic concentrations is the reduced folate carrier (RFC), an OAT transporter with a Km for MTX between approximately 0.8-26 µM. Decreased RFC activity has been observed in several in vitro models of transport-mediated MTX resistance (Biochem. Pharmacol. 11: 1233-1234, 1960). Once rodent and human genes encoding proteins with RFC activity were isolated, the molecular explanations for decreased RFC activity emerged. RFC1 transfection into the transport-deficient MTX^R ZR75 cell line resulted in a 20-fold increase in 6-hour MTX uptake and a concomitant 250-fold increase in sensitivity to MTX relative to control cell clones, showing that the RFC1 gene reconstitutes RFC activity and has a significant impact on MTX cytotoxicity (Moscow, et al., Cancer Res. 55: 3790-3794, 1995).

In different cell lines, MTX transport deficiency has been ascribed either to mutations in the RFC gene or in decreased expression of the RFC gene product. Several studies have demonstrated that RFC1 gene expression is an important determinant of sensitivity to MTX. In vitro studies, we have found that RFC1 RNA levels correlate with MTX sensitivity in a panel of non-selected cell lines, including breast cancer cell lines (Moscow et al., Int J Cancer. 72: 184-190, 1997).

A plethora of genes with the ability to transport MTX out of the cell have been reported, including MRP1, MRP2, MRP3, MRP4, the organic anion transporters hOAT2 and hOAT3, and the mitoxantrone-resistance protein (BCRP/MXR). However, despite the multitude of MTX export genes, clinical studies have shown a relationship between the expression of RFC1, the mechanism of MTX uptake, and prognosis in Acute Lymphoid Leukemia (ALL) and osteosarcoma. As a result, RFC1 expression and MTX uptake are now implicated as determinants of clinical sensitivity in several types of tumors. Thus, the role of RFC1 in mediating sensitivity of its cytotoxic drug substrates has become a prototype that illustrates the potential role of transporters, like OAT and OCT genes, in determination of anticancer drug selectivity and toxicity.

However, there is a need to identify additional channels, or transporters, that are found in specific cancers, to enable the targeting of different cancers with anticancer agents that are substrates for those transporters.

SUMMARY OF THE INVENTION

The present invention is directed towards a membrane protein that functions to transport hydrophilic substances across cellular membranes. The protein, OCT6, is a new member of the organic cation transporter (OCT) family (SLC22 gene family). Tissue distribution of this protein is distinct from other OCT protein family members; being detected in leukemia, leukemia blast cells and CD34+ cells.

In one aspect, the present invention provides a novel target for hematological malignancies such as leukemia, an OCT6 transporter.

In another aspect of the present invention there is a method for screening potential substrates that selectively bind the OCT6 transporter. The method involves contacting a cell which overexpresses an OCT6 transporter gene with a test compound and determining whether the test compound is a substrate for the OCT6 transporter.

In another aspect, there is a method for screening potential anti-cancer agents in a cell overexpressing an OCT6 transporter gene. The method comprises determining viability of a cell which expresses OCT6 transporter gene incubated in the presence and absence of a test compound and identifying the test compound as a potential anti-cancer agent if there is cellular influx of the test compound and cell death.
In another aspect of the invention, a test kit is provided for screening candidate drugs for hematologic malignancies comprising a mammalian cell line or cells which overexpress OCT6, a control substrate and a detectable substance.

In still another aspect of the invention, there are immunogenic compositions for treating hematological malignancies. In a preferred embodiment, immunogenic compositions for treating leukemia comprise a substrate that binds selectively to a leukemia cell expressing the OCT6 transporter gene. In another preferred embodiment of the invention, the substrate comprises an antibody that selectively binds to the OCT6 transporter protein. Preferably, the OCT6 transporter protein allows cellular uptake of the substrate which then causes cell death. In one embodiment the substrate is cytotoxic and in another preferred embodiment the substrate is coupled with a cytotoxic agent.

In still another aspect, the present invention provides a method for impairing a leukemia cell comprising contacting the cell with a cytotoxic OCT6 transporter protein. In one embodiment the substrate is a cytotoxic agent and in another embodiment the substrate is coupled to a cytotoxic agent.

In yet another aspect, the present invention provides a method for treating hematological malignancies comprising administering to a subject in need thereof an immunogenic composition comprising a substrate that binds selectively to a cell expressing the OCT6 transporter gene. In a preferred embodiment the OCT6 transporter protein allows cellular uptake of the substrate which then causes cell death. In another preferred embodiment the substrate is cytotoxic. In another preferred embodiment, the substrate is coupled with a cytotoxic agent.

**BRIEF DESCRIPTION OF THE DRAWINGS**

**FIG. 1.** A shows the predicted hydrophobic profile of OCT6.

**FIG. 1.** B. is a dendrogram showing phylogenetic relationship between OCT6 (SEQ ID NO:2) and other OCT and OAT proteins, including, OCTN1 (SEQ ID NO:4), OCT3 (SEQ ID NO:5), OCTN2 (SEQ ID NO:6), OCT2 (SEQ ID NO:7), OCT1 (SEQ ID NO:8), OAT5 (SEQ ID NO:9), OAT4 (SEQ ID NO:10), OAT3 (SEQ ID NO:11), and OAT1 (SEQ ID NO:12).

**FIG. 2A-F.** is the CLUSTALW alignment of OCT6 and other OCT and OAT proteins. The bottom row represents areas of consensus.

**FIG. 3.** shows the normal tissue distribution of OCT6 RNA determined by RT-PCR using a cDNA panel. Only 1000x (highest) cDNA concentration is shown. Panel A. 1, salivary gland; 2, thyroid; 3, adrenal; 4, pancreas; 5, ovary; 6, uterus; 7, prostate; 8, skin; 9, peripheral blood leukocytes; 10, bone marrow; 11, fetal brain; 12, fetal liver. Panel B. 1, brain; 2, heart; 3, kidney; 4, spleen; 5, liver; 6, colon; 7, lung; 8, small intestine; 9, muscle; 10, stomach, 11, testis; 12, placenta.

**FIG. 4.** shows quantitative RT-PCR for the transporter gene OCT6 performed with RNA extracted from peripheral blood leukocytes, CD34+ cells and additional hematopoietic cell lines. Fresh discarded buffy coats that were twice sorted by FACS using CD14 (monocytes), CD15 (granulocytes), CD3 (T-cells) and CD20 (B-cells). Purities of 99% or better were obtained. For peripheral WBC and sorted subsets, the average±SD represent pooled results from samples from 2 individuals performed in triplicate or quadruplicate. For CD34-selected mobilized peripheral blood (MPB), the results from each of 3 individuals are shown. For CD34-selected bone marrow (CD34+BM), the results are from one individual. OCT6 levels were normalized to the expression of actin RNA, as a control for equivalence of mRNA template. The units, in log scale, are arbitrary and based on a standard curve of OCT6 RT-PCR in serially diluted HL60 RNA. Unity is defined as the level of OCT6 RNA found in MOLT4 cells. **FIG. 5.** shows quantitative RT-PCR for the gene OCT6 using RNA extracted from leukemic blasts obtained from patients at the time of diagnosis. OCT6 levels were normalized to the expression of actin RNA, as a control for equivalence of mRNA template. The OCT6 RNA levels in placenta, liver, kidney and MOLT-4 cell line were determined concurrently and shown for comparison. The units, in log scale, are arbitrary and based on a standard curve of OCT6 RT-PCR in serially diluted HL60 RNA. Unity is defined as the level of OCT6 RNA found in MOLT4 cells.

**DETAILED DESCRIPTION OF THE INVENTION**

The present invention is based on the discovery and isolation of a new member of the SLC22 gene family (the OCT family of proteins) that is unusual for its distinct pattern of tissue distribution. Rather than the typical high levels of expression in liver, kidney or placenta, high levels of RNA for this transporter were found in some leukemia cell lines, in CD34+ cells, and in circulating leukemia blast cells.

All patents, patent applications and literature cited in this description are incorporated herein by reference in their entirety. In the case of inconsistencies, the present disclosure, including definitions, will prevail.

**OCT Family**

Two families of proteins involved in maintaining homeostasis of charged organic compounds are the organic anion transporters (OATs) which carry the SLC21 designation and the organic cation transporters (OCTs), which carry the SLC22 designation (See Table 1). OATs and OCTs each have characteristic patterns of tissue expression, with predominant expression in a tissue involved in the transport of xenobiotics, i.e., liver, kidney or placenta.

<table>
<thead>
<tr>
<th>Gene Family</th>
<th>Gene Name</th>
<th>Locus Link</th>
<th>Alternative Names</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLC21</td>
<td>SLC21A1</td>
<td>6577</td>
<td></td>
</tr>
<tr>
<td>SLC21A2</td>
<td>6578</td>
<td>PGT</td>
<td></td>
</tr>
<tr>
<td>SLC21A3</td>
<td>6579</td>
<td>OATP-A</td>
<td></td>
</tr>
<tr>
<td>SLC21A4</td>
<td>28237</td>
<td>OAT-K1</td>
<td></td>
</tr>
<tr>
<td>SLC21A5</td>
<td>28236</td>
<td>OATP2</td>
<td></td>
</tr>
<tr>
<td>SLC21A6</td>
<td>10599</td>
<td>LST-1</td>
<td></td>
</tr>
<tr>
<td>SLC21A7</td>
<td>28235</td>
<td>OATP3, OATP-3</td>
<td></td>
</tr>
<tr>
<td>SLC21A8</td>
<td>28234</td>
<td>LST2, OATP8, SLC21A8, OATP-8</td>
<td></td>
</tr>
<tr>
<td>SLC21A9</td>
<td>11309</td>
<td>OATP-B</td>
<td></td>
</tr>
<tr>
<td>SLC21A10</td>
<td>28233</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC21A11</td>
<td>28232</td>
<td>OATP-D</td>
<td></td>
</tr>
<tr>
<td>SLC21A12</td>
<td>28331</td>
<td>LOC51737, OATP-E, POAT</td>
<td></td>
</tr>
<tr>
<td>SLC21A13</td>
<td>28230</td>
<td>OATP5, OATP-5</td>
<td></td>
</tr>
<tr>
<td>SLC21A14</td>
<td>53019</td>
<td>OATP-3</td>
<td></td>
</tr>
<tr>
<td>SLC22</td>
<td>SLC22A1</td>
<td>6580</td>
<td>OCT1</td>
</tr>
<tr>
<td>SLC22A2</td>
<td>6582</td>
<td>OCT2</td>
<td></td>
</tr>
<tr>
<td>SLC22A3</td>
<td>6581</td>
<td>OCT3</td>
<td></td>
</tr>
<tr>
<td>SLC22A4</td>
<td>6583</td>
<td>OCTN1</td>
<td></td>
</tr>
<tr>
<td>SLC22A5</td>
<td>6584</td>
<td>OCTN2, CDSP, SCD</td>
<td></td>
</tr>
<tr>
<td>SLC22A6</td>
<td>9356</td>
<td>NK1, OAT1, OAT-1</td>
<td></td>
</tr>
<tr>
<td>SLC22A7</td>
<td>10864</td>
<td>NLT, OAT2, OCT-2</td>
<td></td>
</tr>
<tr>
<td>SLC22A8</td>
<td>9376</td>
<td>OAT3, OAT-3</td>
<td></td>
</tr>
<tr>
<td>SLC22A9</td>
<td>6584</td>
<td>OAT4, OAT-4</td>
<td></td>
</tr>
</tbody>
</table>

The OAT and OCT carriers result in increased cellular accumulation of their respective substrates, despite the fact
that they are carriers that mediate facilitative diffusion. For carriers, the degree of intracellular accumulation may not exceed the extracellular concentration. However, the presence of the carrier allows uptake in comparison to no uptake in the absence of the carrier, and drugs that bind an intracellular target or which are chemically modified in the cells, e.g., by phosphorylation or polyglutamylation, may be eliminated from the substrate pool and not available for transport back across the cellular membrane.

The first five members of the SLC22 family of transporters, OCT1, OCT2, OCT3, OCTN1, and OCTN2, have been characterized as organic cation transporters. The uptake of many cations, such as tetraethylammonium (TEA), N-1-methyl-3-isopropylammonium (MIP), choline, p-cresolamine, amantadine, and morphine are mediated by these polypeptide transporters. In general, these transporters are potential-dependent, but independent of sodium and proton gradients. These genes are all characterized by the presence of 11 or 12 transmembrane domains, as predicted by hydrophobicity analysis, and all have a large hydrophilic loop between transmembrane domain (TMD) 1 and TMD2.

OCT substrates are shown below in Table 2. Tetraethylammonium (TEA) is the classic substrate for OCT transporters. In addition, OCT1, OCT2 and OCT3 transport 1-methyl-4-phenylpyridinium (MPP). Compared to OCT2, OCT1 has a higher affinity for some cations (for example maprotiline and phentolamine), a similar affinity for others (for example, dehydrogenase and quinidine), and a lower affinity for corticosterone (See Koepsell et al., Ann. Rev. Physiol. 60: 243–266, 1998). OCT3 is an electroneutral transporter for TEA and guanidine. Other physiologic substrates for OCT transporters include dopamine, histamine, epinephrine and norepinephrine, acetylcholine, and 5-hydroxytryptamine (Burckhardt et al., Am J Physiol Renal Physiol. 278: F853–66, 2000), suggesting an important role for these transporters in the central nervous system, in addition to their role in hepatic and renal clearance. Interestingly, despite its cationic nature, recent studies have identified cimetidine as a selective inhibitor, but not a substrate for several organic cation transporters, including OCT1, OCT2, OCT3, hOCTN1, and hOCTN2.

TABLE 2

<table>
<thead>
<tr>
<th>OCT Substrates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Name</td>
</tr>
<tr>
<td>OCT1</td>
</tr>
<tr>
<td>OCT1</td>
</tr>
<tr>
<td>OCT2</td>
</tr>
<tr>
<td>OCT2</td>
</tr>
<tr>
<td>OCT2</td>
</tr>
<tr>
<td>OCT2</td>
</tr>
<tr>
<td>OCT2</td>
</tr>
<tr>
<td>OCT2</td>
</tr>
<tr>
<td>OCT2</td>
</tr>
<tr>
<td>OCT2</td>
</tr>
<tr>
<td>OCT3</td>
</tr>
<tr>
<td>OCT3</td>
</tr>
<tr>
<td>OCTN1</td>
</tr>
<tr>
<td>OCTN2</td>
</tr>
<tr>
<td>OCTN2</td>
</tr>
<tr>
<td>OCTN2</td>
</tr>
<tr>
<td>OCTN2</td>
</tr>
<tr>
<td>OCTN2</td>
</tr>
<tr>
<td>OCTN2</td>
</tr>
<tr>
<td>OCTN2</td>
</tr>
</tbody>
</table>

OCT1 and OCT2 are predominantly expressed in the kidney and liver. These transporters are located on the basolateral surface of renal tubules and, therefore, play a role in the removal of organic cations from the blood. OCT3 is most abundantly expressed in placenta. In addition, other tissue-specific roles have been implicated for these transporters. As noted above, OCTs may play a role in transport of endogenous neuroleptic substrates, and OCT3 has been implicated in the disposal of cationic neurotoxins and neurotransmitters in the brain (Wu et al., J Biol Chem. 273: 32776-86, 1998). Dhillon et al. (Clin Pharmacol Ther. 65: 205, 1999) used RT-PCR followed by functional transport studies (TEA) to identify OCT1 expression in a human mammary epithelial cell line (MCF12A). Further, the OCT1 gene has been shown to be up regulated in lactating mammary epithelial cells. The OCTN1 gene, cloned from a cDNA, shows sequence similarity to organic cation transporter genes, which is highly expressed in kidney as well as trachea, bone marrow and fetal liver. Recombinant OCTN1 expressed in mammalian cells exhibited saturable uptake of TEA that was pH sensitive. Several others suggest that OCTN1 is a renal proton/organic cation antipporter functioning at the epithelial apical membrane. The uptake of pyrillamine, quinidine, verapamil and L-carnitine were increased by expression of OCTN1 in Xenopus oocytes.

Another OCT protein family member, OCTN2, cloned from a human placental trophoblast cell line, is expressed widely in human tissues including kidney, placenta and heart. OCTN2 is more closely related to OCTN1 than to OCT1, OCT2 and OCT3 (Biochem Biophys Res Commun. 246: 589-95, 1998). Transfection of OCTN2 has demonstrated its role in the transport of TEA and carnitine. OCTN2-mediated transport of TEA is sodium independent, whereas transport of carnitine is sodium-dependent. The role of sodium in OCTN2-mediated carnitine transport not only involves the electrogenic gradient, but the presence of sodium also alters the affinity of OCTN2 for carnitine. Germline mutations of OCTN2 result in primary carnitine deficiency, a syndrome of progressive cardiomypathy and skeletal myopathy. The symptoms associated with this syndrome are thought to result not only from generalized carnitine deficiency from decreased renal carnitine reabsorption, but also from inability of cardiac and skeletal myocytes, which ordinarily express OCTN2, to accumulate carnitine. This syndrome demonstrates that tissue-specific OCT-mediated transport is essential for accumulation of required cations in specific tissues. The present invention identifies a new transport protein in the OCT family, OCT6, preferentially expressed in leukemia cell lines, leukemia blast cells and CD34+ cells. The cell surface localization and the transporter function of the OCT6 gene product suggest its usefulness as a target in the diagnosis and treatment of hematologic malignancies.

As used herein, the term “antibody” refers to an immunoglobulin molecule with a specific amino acid sequence evoked in by an antigen, and characterized by reacting specifically with the antigen in some demonstrable way. As used herein, the term “carrier” refers to a diluent, adjuvant, excipient, or vehicle with which the compositions of the present invention are administered. As used herein, “compound” refers to any agent, chemical, substance, or substrate, whether organic or inorganic, or any protein including antibodies, peptides, polypeptides, peptides, and the like. As used herein, the term cytotoxin or cytotoxic agent includes any specific substance, which may or may not be antibody, that inhibits or prevents the functions of cells, causes destruction of cells, or both.
As used herein, the term "derivative" refers to something produced by modification of something pre-existing; for example, a substance or chemical compound that may be produced from another substance or compound of similar structure in one or more steps.

As used herein, the term "fragment" refers to a part of a larger entity, said larger entity comprising by non-limiting example, an antibody, compound or substance.

As used herein, the term "leukemia blast" or "leukemic blast" refers to lymphoblasts, the abnormal immature white blood cells associated with leukemia.

As used herein, the term "monoclonal antibody" is not limited to antibodies produced through hybridoma technology. The term "monoclonal antibody" refers to an antibody that is derived from a single clone, including any eukaryotic, prokaryotic, or phage clone, and not the method by which it is produced.

As used herein, the term "pharmaceutically acceptable carrier" refers to a carrier that may be administered to a subject, together with one or more liver protecting agents and one or more mushroom powder or extract of the present invention, and which does not destroy the pharmacological activity thereof and is nontoxic when administered in doses sufficient to deliver a therapeutic amount of the compound.

As used herein, the term "substrate" refers to a substance, compound, agent, antibody or derivatives and/or fragment thereof, acted upon by the OCT6 transporter protein (e.g., a substance that is taken across the cellular membrane by action of the OCT6 transporter protein).

OCT6 (SEQ ID NO:1) was first identified as a potential OCT gene by assembling and sequencing ESTs as described in Example 1 (amino acid sequence of OCT6 is SEQ ID NO:2). The gene sequence proved to be identical to the recently submitted cDNA ORB1 (GenBank AF268892) submitted by M. Okabe and T. Miki, incorporated herein in its entirety. It is also contained within the submitted BAC clone CTA-331P3 (SEQ ID NO: 3) (GenBank AAC002464) located at chromosome 6q21, incorporated herein in its entirety. The gene has a predicted protein structure typical of transport proteins with two groups of six transmembrane domains separated by a hydrophilic region (FIG. 1A). CLUSTALW alignment produced a dendrogram showing the phylogenetic relationship between OCT6 and other OAT and OCT proteins (FIG. 1B). This dendrogram suggests that the distinction between OAT and OCT genes, based on functional studies, obscures the common origin of both families of transporters. The actual CLUSTALW alignment of these genes is shown in FIG. 2 and demonstrates multiple regions of conservation among all of these genes.

Next, according to the methods described in Example 3, quantitative RT-PCR analysis of the expression of OCT6 was performed, along with the expression of other OCT genes, in 50 cell lines. The results are shown in Table 3. The two highest expressing cell lines for OCT6 in this panel were two leukemia cell lines, HL60, a human promyelocytic leukemia cell line, and MOLT4, a human acute lymphoblastic leukemia (T-cell) cell line. There was only a low level of expression detected in most of the other cell lines.

<table>
<thead>
<tr>
<th>No.</th>
<th>Cell Line</th>
<th>source</th>
<th>OCT1</th>
<th>OCT2</th>
<th>OCT3</th>
<th>OCTN 2</th>
<th>OCT6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CCRF-CEM</td>
<td>Leukemia</td>
<td>0.7</td>
<td>0.7</td>
<td>0.2</td>
<td>0.1</td>
<td>5.7</td>
</tr>
<tr>
<td>2</td>
<td>HL-60</td>
<td>Leukemia</td>
<td>0.5</td>
<td>1.3</td>
<td>0.0</td>
<td>0.4</td>
<td>7.6</td>
</tr>
<tr>
<td>3</td>
<td>K-562</td>
<td>Leukemia</td>
<td>1.4</td>
<td>1.2</td>
<td>0.2</td>
<td>1.4</td>
<td>5.2</td>
</tr>
<tr>
<td>4</td>
<td>MOLT-4</td>
<td>Leukemia</td>
<td>0.1</td>
<td>1.1</td>
<td>0.5</td>
<td>0.6</td>
<td>46.8</td>
</tr>
<tr>
<td>5</td>
<td>RPMI-8226</td>
<td>Leukemia</td>
<td>2.8</td>
<td>2.0</td>
<td>0.1</td>
<td>3.7</td>
<td>6.02</td>
</tr>
<tr>
<td>6</td>
<td>SR</td>
<td>Leukemia</td>
<td>1.9</td>
<td>1.1</td>
<td>0.0</td>
<td>0.9</td>
<td>2.6</td>
</tr>
<tr>
<td>7</td>
<td>AS49-A5CC</td>
<td>Lung cancer</td>
<td>1.7</td>
<td>1.2</td>
<td>161</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>8</td>
<td>HOP-62</td>
<td>Lung cancer</td>
<td>0.8</td>
<td>4.8</td>
<td>0.6</td>
<td>2.4</td>
<td>4.1</td>
</tr>
<tr>
<td>9</td>
<td>NCI-H226</td>
<td>Lung cancer</td>
<td>4.8</td>
<td>0.5</td>
<td>0.1</td>
<td>21.1</td>
<td>4.8</td>
</tr>
<tr>
<td>10</td>
<td>NCI-H23</td>
<td>Lung cancer</td>
<td>0.5</td>
<td>0.7</td>
<td>0.0</td>
<td>0.3</td>
<td>5.2</td>
</tr>
<tr>
<td>11</td>
<td>NCI-H460</td>
<td>Lung cancer</td>
<td>0.7</td>
<td>1.0</td>
<td>0.0</td>
<td>1.7</td>
<td>1.8</td>
</tr>
<tr>
<td>12</td>
<td>COLO205</td>
<td>Colon Ca.</td>
<td>4.9</td>
<td>5.3</td>
<td>30.9</td>
<td>2.2</td>
<td>3.6</td>
</tr>
<tr>
<td>13</td>
<td>HCC-2998</td>
<td>Colon Ca.</td>
<td>1.5</td>
<td>1.0</td>
<td>0.0</td>
<td>2.6</td>
<td>5.4</td>
</tr>
<tr>
<td>14</td>
<td>HCT-116</td>
<td>Colon Ca.</td>
<td>1.7</td>
<td>2.1</td>
<td>0.1</td>
<td>2.8</td>
<td>9.7</td>
</tr>
<tr>
<td>15</td>
<td>HCT-15</td>
<td>Colon Ca.</td>
<td>0.9</td>
<td>1.7</td>
<td>0.1</td>
<td>3.5</td>
<td>4.2</td>
</tr>
<tr>
<td>16</td>
<td>HT-29</td>
<td>Colon Ca.</td>
<td>1.9</td>
<td>1.2</td>
<td>18.1</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>17</td>
<td>KM-12</td>
<td>Colon Ca.</td>
<td>0.6</td>
<td>1.0</td>
<td>12.2</td>
<td>0.7</td>
<td>2.1</td>
</tr>
<tr>
<td>18</td>
<td>SW-620</td>
<td>Colon Ca.</td>
<td>1.0</td>
<td>2.6</td>
<td>40.4</td>
<td>3.7</td>
<td>1.9</td>
</tr>
<tr>
<td>19</td>
<td>SF-268</td>
<td>CNS Tumor</td>
<td>0.4</td>
<td>0.8</td>
<td>0.0</td>
<td>0.9</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>SF-295</td>
<td>CNS Tumor</td>
<td>0.5</td>
<td>1.2</td>
<td>0.2</td>
<td>1.1</td>
<td>2.5</td>
</tr>
<tr>
<td>21</td>
<td>SF-539</td>
<td>CNS Tumor</td>
<td>0.5</td>
<td>0.6</td>
<td>2.3</td>
<td>0.2</td>
<td>5.3</td>
</tr>
<tr>
<td>22</td>
<td>SNB-75</td>
<td>CNS Tumor</td>
<td>0.8</td>
<td>1.8</td>
<td>0.0</td>
<td>0.6</td>
<td>2.3</td>
</tr>
<tr>
<td>23</td>
<td>U251</td>
<td>CNS Tumor</td>
<td>0.8</td>
<td>0.9</td>
<td>0.0</td>
<td>0.6</td>
<td>7.4</td>
</tr>
<tr>
<td>24</td>
<td>LOCMV1</td>
<td>Melanoma</td>
<td>2.9</td>
<td>2.1</td>
<td>0.1</td>
<td>0.4</td>
<td>3.6</td>
</tr>
<tr>
<td>25</td>
<td>MALME-3M</td>
<td>Melanoma</td>
<td>1.5</td>
<td>1.5</td>
<td>0.0</td>
<td>2.3</td>
<td>3</td>
</tr>
<tr>
<td>26</td>
<td>M14</td>
<td>Melanoma</td>
<td>1.9</td>
<td>1.4</td>
<td>0.0</td>
<td>1.9</td>
<td>4.7</td>
</tr>
<tr>
<td>27</td>
<td>SK-MEL-2</td>
<td>Melanoma</td>
<td>2.1</td>
<td>1.9</td>
<td>0.0</td>
<td>2.2</td>
<td>3.9</td>
</tr>
<tr>
<td>28</td>
<td>SK-MEL-5</td>
<td>Melanoma</td>
<td>2.6</td>
<td>1.5</td>
<td>0.0</td>
<td>1.9</td>
<td>2.7</td>
</tr>
<tr>
<td>29</td>
<td>UACC-257</td>
<td>Melanoma</td>
<td>3.2</td>
<td>3.6</td>
<td>0.0</td>
<td>1.1</td>
<td>5.4</td>
</tr>
<tr>
<td>30</td>
<td>IGROV1</td>
<td>Ovarian Ca.</td>
<td>4.9</td>
<td>5.0</td>
<td>17.9</td>
<td>1.8</td>
<td>2.5</td>
</tr>
<tr>
<td>31</td>
<td>OVCAR-3</td>
<td>Ovarian Ca.</td>
<td>1.4</td>
<td>0.1</td>
<td>0.0</td>
<td>2.2</td>
<td>1.4</td>
</tr>
<tr>
<td>32</td>
<td>OVCAR-4</td>
<td>Ovarian Ca.</td>
<td>2.6</td>
<td>1.4</td>
<td>0.0</td>
<td>8.9</td>
<td>3.4</td>
</tr>
<tr>
<td>33</td>
<td>OVCAR-5</td>
<td>Ovarian Ca.</td>
<td>3.5</td>
<td>2.7</td>
<td>105</td>
<td>10.0</td>
<td>4.8</td>
</tr>
<tr>
<td>34</td>
<td>OVCAR-8</td>
<td>Ovarian Ca.</td>
<td>1.1</td>
<td>1.0</td>
<td>0.0</td>
<td>0.8</td>
<td>1.6</td>
</tr>
<tr>
<td>35</td>
<td>SK-OV-3</td>
<td>Ovarian Ca.</td>
<td>3.9</td>
<td>1995</td>
<td>92</td>
<td>8.5</td>
<td>9.8</td>
</tr>
<tr>
<td>36</td>
<td>A498</td>
<td>Renal Ca.</td>
<td>2.2</td>
<td>13.4</td>
<td>180</td>
<td>4.7</td>
<td>1.3</td>
</tr>
<tr>
<td>37</td>
<td>ACHN</td>
<td>Renal Ca.</td>
<td>1.1</td>
<td>1.1</td>
<td>0.7</td>
<td>1.2</td>
<td>1.1</td>
</tr>
<tr>
<td>38</td>
<td>CAKI-1</td>
<td>Renal Ca.</td>
<td>3.5</td>
<td>2.5</td>
<td>4.8</td>
<td>1.8</td>
<td>2.8</td>
</tr>
</tbody>
</table>
TABLE 3-continued

<table>
<thead>
<tr>
<th>No.</th>
<th>Cell Line</th>
<th>source</th>
<th>OCT1</th>
<th>OCT2</th>
<th>OCT3</th>
<th>OCTN 2</th>
<th>OCT6</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>RXF-393</td>
<td>Renal Ca.</td>
<td>1.7</td>
<td>1.2</td>
<td>3.0</td>
<td>0.6</td>
<td>1.2</td>
</tr>
<tr>
<td>40</td>
<td>TX-10</td>
<td>Renal Ca.</td>
<td>3.6</td>
<td>5.0</td>
<td>16.8</td>
<td>2.5</td>
<td>8</td>
</tr>
<tr>
<td>41</td>
<td>DU-31</td>
<td>Renal Ca.</td>
<td>4.4</td>
<td>1.6</td>
<td>31.2</td>
<td>1.2</td>
<td>2.3</td>
</tr>
<tr>
<td>42</td>
<td>PC-3</td>
<td>Prostate Ca.</td>
<td>2.1</td>
<td>0.8</td>
<td>9.6</td>
<td>3.3</td>
<td>4.7</td>
</tr>
<tr>
<td>43</td>
<td>DU-145</td>
<td>Prostate Ca.</td>
<td>1.1</td>
<td>1.1</td>
<td>3.4</td>
<td>1.6</td>
<td>3</td>
</tr>
<tr>
<td>44</td>
<td>MCF-7</td>
<td>Breast Ca.</td>
<td>0.8</td>
<td>1.8</td>
<td>0.0</td>
<td>10.4</td>
<td>3.5</td>
</tr>
<tr>
<td>45</td>
<td>NCI/ADR-RES</td>
<td>Breast Ca.</td>
<td>1.4</td>
<td>1.3</td>
<td>1.1</td>
<td>2.0</td>
<td>2.1</td>
</tr>
<tr>
<td>46</td>
<td>MDA-MB-231</td>
<td>Breast Ca.</td>
<td>1.2</td>
<td>0.4</td>
<td>3.9</td>
<td>4.8</td>
<td>1.8</td>
</tr>
<tr>
<td>47</td>
<td>HS578T</td>
<td>Breast Ca.</td>
<td>1.0</td>
<td>1.5</td>
<td>0.0</td>
<td>1.2</td>
<td>8.3</td>
</tr>
<tr>
<td>48</td>
<td>MDA-MB-435</td>
<td>Breast Ca.</td>
<td>1.9</td>
<td>0.6</td>
<td>0.1</td>
<td>0.7</td>
<td>2.7</td>
</tr>
<tr>
<td>49</td>
<td>BT-549</td>
<td>Breast Ca.</td>
<td>1.2</td>
<td>0.8</td>
<td>0.1</td>
<td>0.3</td>
<td>2.6</td>
</tr>
<tr>
<td>50</td>
<td>T-47D</td>
<td>Breast Ca.</td>
<td>0.7</td>
<td>1.1</td>
<td>0.1</td>
<td>4.2</td>
<td>8.7</td>
</tr>
</tbody>
</table>

OCT6 is unique among the known members of OCT and OAT genes because of its pattern of tissue distribution. The pattern of expression of the OCT6 gene in the 50 cell lines suggested that its expression might be restricted to hematopoietic tissues. The restricted pattern of expression observed for OCT6 also suggests that therapies using OCT6-specific substrates are unlikely to have widespread toxicity to normal tissues. Therefore, we examined OCT6 expression in a cDNA panel representing a wide cross-section of normal tissues according to the methods of Example 4 (FIG. 3). This study revealed that OCT6 RNA levels are highest in testis and fetal liver, with lower but detectable levels in peripheral blood leukocytes and bone marrow. Since fetal hematopoiesis occurs in the liver, it is possible that the fetal liver sample may have included both hepatocytes and hematopoietic cells. OCT6 RNA levels were also barely detectable in pancreatic and adrenal tissue. Unlike other OCT genes, expression was not detectable in liver, kidney or placenta.

To determine whether OCT6 RNA expression in hematopoietic cells was lineage-specific, leukocytes were sorted from discarded buffy coat specimens by flow cytometry, and purified subpopulations were examined for OCT6 RNA expression according to the methods described in Example 5. OCT6 expression was also examined in a population of CD34+ cells. As can be seen in FIG. 4, the expression of OCT6 was highly enriched in CD34+ cells in comparison to the other cell populations. Also, significant levels of OCT6 expression (relative to MOLT4) were found in other hematopoietic cell lines: U937, a human histiocytic lymphoma cell line; THP-1, a human acute monocytic leukemia cell line; KG-1, a human erythroleukemia cell line; and MV-4-11, a human biphenotypic (B-cell and myelomonocytic) leukemia cell line.

The high levels of OCT6 RNA in some leukemia cell lines and CD34+ cells also raised the question as to whether this gene was highly expressed in actual leukemias. To address this issue, the RNA levels of OCT6 in 25 samples of peripheral leukemia cells were measured according to the methods set out in Example 6. The FAB classification of these samples are shown in Table 4. These results are shown in FIG. 8, and demonstrate that the majority of specimens contained RNA levels for OCT6 that exceeded the level found in MOLT4 cell line, the second highest expressing cell line among those examined, and exceed by orders of magnitude the levels found in placenta, kidney and liver.

Due to the OCT6 protein’s location on the cellular membrane and its function as an intracellular transporter, the OCT6 transporter protein has been identified as a therapeutic target. Basic principles of cellular pharmacology suggest that increase in intracellular accumulation will lead to increased intracellular effect. For anticancer drugs, this principle has been studied extensively in the context of lipophilic drugs, which require no specific mechanism for cellular uptake, and export pumps such as the product of the multidrug resistance gene, MDR1, whose overexpression of MDR1 leads to increased cellular resistance by decreasing intracellular concentrations of drug (Moscovici, J. A., Schneider, E. S., Ivy, S. P., and Cowan, K. H. Multidrug resistance. In: H. M. Pinedo, D. L. Longo, and B. A. Chabner (eds.), Cancer chemotherapy and biological response modifiers. Annual 17, New York: Elsevier, 1997). The same principle applies to charged, hydrophilic drugs of the present invention, except that the determinants of sensitivity depend on uptake as opposed to efflux. As such, cells expressing an OCT6 transporter are likely to be highly sensitive to cytotoxic OCT6 substrates.
Drug Screening

Accordingly, the present invention provides methods for screening potential substrates of, and potential therapeutic agents against hematological malignancies like leukemia that overexpress, the OCT6 transporter. In particular, potential therapeutic agents are screened for the ability to be a substrate recognized by an OCT6 transporter protein. Preferably, potential substrates are screened for the ability to confer cytotoxic effects on a cell overexpressing OCT6 transporter protein. More preferably, agents are screened for the ability to preferentially cause cellular uptake into, and cell death of, cells overexpressing the OCT6 transporter. Most preferably, the agents are screened for the ability to cause cell death of cancer cells such as leukemia overexpressing the OCT6 transporter as compared to normal cells.

A method for screening potential substrates of the OCT6 transporter protein comprises providing a cell or cell line which expresses OCT6 and a test compound, incubating the test compound and cell line and analyzing the cell or cell line to determine if there was a cellular influx of the test compound. Analysis of the cell line to determine whether cellular uptake of the test compound occurred can be accomplished by any means known in the art. For example, a test compound can be tagged with a detectable label prior to contact with a cell and then observed under microscopy or by other means for its location. Non-limiting examples of labels include green fluorescent protein, alkaline phosphatase, horseradish peroxidase, rease, f3-galactosidase, CAT, luciferase, an immunogenic tag peptide sequence, an extrinsically activatable enzyme, an extrinsically activatable toxin, an extrinsically activatable fluor, an extrinsically activatable quenching agent, a radioactive element or an antibody.

A method for screening candidate anti-cancer agents comprises determining the viability of a mammalian cell which expresses OCT6 incubated in the presence and absence of a test compound and identifying the test compound as a potential anti-leukemia agent if there is a cellular uptake of the test compound and cell death. Analysis of cell viability can be accomplished by any means known in the art.

It is well known in the art that viability of a cell can be determined by contacting the cell with a dye and viewing it under a microscope. Viable cells can be observed to have an intact membrane and do not stain, whereas dying or dead cells having "leaky" membranes do stain. Incorporation of the dye by the cell indicates the death of the cell. The most common dye used in the art for determining viability is trypan blue. Viability of cells can also be determined by detecting DNA synthesis. Cells can be cultured in cell medium with labeled nucleotides (e.g., [3H]thymidine). The uptake or incorporation of the labeled nucleotides indicates DNA synthesis and cell viability. In addition, colonies formed by cells cultured in medium indicate cell growth and is another means to test viability of the cells.

Identification and/or observation of cells undergoing apoptosis can be another method of determining cell viability. Apoptosis is a specific mode of cell death recognized by a characteristic pattern of morphological, biochemical, and molecular changes. Cells going through apoptosis appear shrunken, and rounded; they also can be observed to become detached from culture dish. Thermophological changes involve a characteristic pattern of condensation of chromatin and cytoplasm which can be readily identified by microscopy. When stained with a DNA-binding dye, such as H33258, apoptotic cells display classic condensed and punctate nuclei instead of homogeneous and round nuclei.

The hallmark of apoptosis is the endonucleolysis, a molecular change in which nuclear DNA is initially degraded at the linker sections of nucleosomes to give rise to fragments equivalent to single and multiple nucleosomes. When these DNA fragments are subjected to gel electrophoresis, they reveal a series of DNA bands which are positioned approximately equally distant from each other on the gel. The size difference between the two bands next to each other is about the length of one nucleosome (i.e., 20 base pairs). This characteristic display of the DNA bands is called a DNA ladder and it indicates apoptosis of the cell. Apoptotic cells can be identified by flow cytometric methods based on measurement of cellular DNA content, increased sensitivity of DNA to denaturation, or altered light scattering properties. These methods are well known in the art and are within the contemplation of the invention.

Abnormal DNA breaks are also characteristic of apoptosis and can be detected by any means known in the art. In one embodiment, DNA breaks are labeled with biotinylated dUTP (b-dUTP). Cells are fixed and incubated in the presence of biotinylated dUTP with either exogenous terminal transferase (terminal DNA transferase assay, TdT assay) or DNA polymerase (nick translation assay; NT assay). The biotinylated dUTP is incorporated into the chromosome at the places where abnormal DNA breaks are repaired, and are detected with fluorescein conjugated to avidin under fluorescence microscopy.

Kits

The present invention provides kits that can be used in the above screening methods. In one embodiment, a kit comprises a substantially isolated polypeptide comprising an OCT6 epitope which is specifically immunoactive with only test compound(s) that are substrates of the OCT6 transporter protein. Binding of a test compound to the OCT6 epitope is indicative that the test compound is an OCT6 substrate. In another embodiment, a kit comprises a cell line that overexpresses an OCT6 transporter protein. Binding and/or cellular uptake of a test compound via the OCT6 protein is indicative that the test compound is a OCT6 substrate. Preferably, the kits of the present invention further comprise a control compound or antibody which does not react with the OCT6 transporter protein. In another specific embodiment, the kits of the present invention contain a means for detecting the binding of a test compound to an OCT6 epitope and/or cellular uptake of a test compound. For example, the test compound may be conjugated to a detectable substrate such as a fluorescent compound, an enzymatic substrate, a radioactive compound or a luminescent compound, or a second antibody which recognizes the first antibody may be conjugated to a detectable substrate.

The detectable substance may be coupled or conjugated either directly to the test compound (or fragment thereof) or indirectly, through an intermediate (such as, for example, a linker known in the art) using techniques known in the art. See, for example, U.S. Pat. No. 4,744,900 for metal ions which can be conjugated to antibodies for use as diagnostics according to the present invention. Further non-limiting examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, radioactive materials, postiron emitting metals using various positron emission tomographies, nonradioactive paramagnetic metal ions, immunogenic tag peptide sequences, extrinsically activatable toxins, extrinsically activatable quenching agents, or antibodies.

Non-limiting examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/bi-
Examples of suitable radioactive material include $^{125}$I, $^{131}$I, $^{111}$In or $^{99m}$Tc.

**Immunogenic Compositions**

The present invention also provides immunogenic compositions for the treatment of hematological malignancies. Non-limiting exemplary hematological malignancies include, but are not limited to, Hodgkin’s disease, leukemia such as, acute lymphoid (lymphocytic or lymphoblastic) leukemia (ALL), acute myeloid (myelogenous or myeloblastic) leukemia (AML), acute lymphoid leukemia, biphenotypic (ALL, biphentypic), acute undifferentiated leukemia (AUL), chronic myeloid (myelogenous or granulocytic) leukemia (CML), erythroleukemia, granulocytic leukemia, lymphoma, monocytic leukemia, myeloma, myelomonocytic leukemia, myelodysplastic syndromes, non-Hodgkin lymphoma, progranulocytic leukemia.

According to the invention immunogenic compositions for the treatment of hematological malignancies comprise a substrate recognized by an OCT6 transporter protein. Preferably, the substrate is a compound that binds selectively or specifically to a OCT6 transporter protein. In a preferred embodiment, the compound binds selectively to the OCT6 transporter protein encoded by a nucleotide sequence of SEQ ID NO.1. The compound may be a cytoxic or coupled or conjugated with a cytotoxic agent. Preferably the cytotoxic agent is a chemotherapeutic agent.

The present invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically effective amount of a compound, and a pharmaceutically acceptable carrier.

Cell surface proteins like the OCT6 transporter can be utilized in antibody-based targeting strategies. In still another aspect of the invention, antibodies can be developed by known methods in the art against the external epitope of OCT6 transporter protein. In a preferred embodiment, antibodies are substrates of the OCT6 protein. The antibodies may be polyclonal antibodies or monoclonal antibodies.

Polyclonal antibodies to an antigen-of-interest can be produced by various procedures well known in the art. For example, a polypeptide of the invention can be administered to various host animals including, but not limited to, rabbits, mice, rats, etc. to induce the production of sera containing polyclonal antibodies specific for the antigen. Various adjuvants may be used to increase the immunological response, depending on the host species, and include but are not limited to, Freund’s incomplete and complete, mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronics polyols, polyoxyxans, liposomes, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guérin) and corynebacterium parvum. Such adjuvants are also well known in the art.

Monoclonal antibodies can be prepared using a wide variety of techniques known in the art including the use of hybridoma, recombinant, and phage display technologies, or a combination thereof. For example, monoclonal antibodies can be produced using hybridoma techniques including those known in the art and taught, for example, in Harlow et al., Antibodies: A Laboratory Manual. (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, et al., in: Immunologic Methods, 3d ed. (W. H. Freeman and Co., 1985); and in Antibodies: A Laboratory Manual, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, et al., in: Monoclonal Antibodies and T-Cell Hybridomas 563-681 (Elsevier, N.Y., 1981) (said references incorporated by reference in their entireties).

The present invention further encompasses antibodies or fragments thereof conjugated to a diagnostic or therapeutic agent. The antibodies can be used diagnostically to, for example, monitor the development or progression of a tumor as part of a clinical testing procedure to, e.g., determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, radioactive materials, positron emitting metals using various positron emission tomographies, and nonradioactive paramagnetic metal ions. The detectable substance may be coupled or conjugated either directly to the antibody (or fragment thereof) or indirectly, through an intermediate, such as, for example, a linker known in the art, using techniques known in the art. See, for example, U.S. Pat. No. 4,741,900 for metal ions which can be conjugated to antibodies for use as diagnostics according to the present invention.) Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin; and examples of suitable radioactive material include $^{125}$I, $^{131}$I, $^{111}$In or $^{99m}$Tc.
Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is to be administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

The compounds of the invention can be formulated as neutral or salt forms. Pharmacologically acceptable salts include those formed with anions such as those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., and those formed with cations such as those derived from sodium, potassium, ammonium, calcium, ferrous hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc.

The amount of the compound of the invention which will be effective in the treatment, inhibition and prevention of hematological malignancies can be determined by standard clinical techniques. In addition, in vitro assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each patient's circumstances. Effective doses may be extrapolated from dose-response curves derived from in vitro or animal model test systems.

Various other delivery systems are known and can be used to administer a composition of the invention, e.g., encapsulation in liposomes, microparticles, microcapsules, recombinant cells capable of expressing the compound, receptor-mediated endocytosis (See, e.g., Wu and Wu, J. Biol. Chem. 262:4429-4432 (1987)), construction of a nucleic acid as part of a retroviral or other vector, etc. Methods of introduction include but are not limited to intracutaneous, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, oral routes. The compounds or compositions may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (e.g., oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents. Administration can be systemic or local. In addition, it may be desirable to introduce the pharmaceutical compounds or compositions of the invention into the central nervous system by any suitable route, including intraventricular and intrathecal injection; intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir, such as an Omnaya reservoir. Pulmonary administration can also be employed, e.g., by use of an inhaler or nebulizer, and formulation with an aerosolizing agent.

In a specific embodiment, it may be desirable to administer the pharmaceutical compounds or compositions of the invention locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, e.g., in conjunction with a wound dressing after surgery, by injection, by means of a catheter, by means of a suppository, or by means of an implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as stialastic membranes, or fibers. Preferably, when administering a protein, including an antibody, of the invention, care must be taken to use materials to which the protein does not absorb.

In another embodiment, the compound or composition can be delivered in a vesicle, in particular a liposome (See Langer, Science 240:1527-1533 (1990); Treat et al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein and Friedler (eds.), Liss, New York, pp. 353-365 (1989); Lopez-Berestein, ibid., pp. 317-327; see generally ibid.)

In yet another embodiment, the compound or composition can be delivered in a controlled release system. In one...

In a specific embodiment where the compound of the invention is a nucleic acid encoding a protein, the nucleic acid can be administered in vivo to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by use of a retroviral vector (see U.S. Pat. No. 4,980,286), or by direct injection, or by use of microparticle bombardment (e.g., a gene gun; Biologic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox-like peptide which is known to enter the nucleus (see, e.g., Joliot et al., Proc. Natl. Acad. Sci. USA 88:1864-1868 (1991)), etc. Alternatively, a nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

EXAMPLES

The following examples are presented for the illustrative purposes and it is to be understood that the present invention is not limited to those precise embodiments, and that various changes and modifications can be effected therein by one skilled in the art without departing from the scope and spirit of the invention as defined by the appended claims.

Example 1

OCT6 Nucleotide Sequence Identification and Analysis

OCT6 was first identified as a potential OCT gene by assembling and sequencing ESTs. BLAST searches of human ESTs in GenBank data base identified AI040384 (654 bp), AA033971 (714 bp) and H70190 (474 bp) sequences from three fetal liver IMAGE clones, 1656502, 429904 and 212935 respectively. IMAGE clone 1656502 (3', insert 1337 bp) ended the predicted stop codon, while IMAGE clone 429904 (5', insert 996 bp) and IMAGE clone 212935 (5', insert 966 bp) aligned with the 5'-coding region. All clones were obtained from the IMAGE Consortium through the American Type Culture Collection (Manassas, Va.). Each clone was sequenced in both directions. The sequences were determined using ABI Prism™ 377 DNA sequencer (Perkin-Elmer). Our assemblage proved to be identical to the recently submitted cDNA OKB1 (AF268892) submitted by M. Okabe and T. Abe. We have dubbed this gene OCT6 as OCTN1 and OCTN2 may be considered as OCT4 and OCT5 respectively.

The OCT6 gene (SEQ ID NO:1) is also contained within BAC clone CTA-331P3 (SEQ ID NO:3) (GenBank AC002464) located at chromosome 6q21. It is divided into 6 exons that span 42 kb on the human genome, from nucleotide 79,570 to nucleotide 120490 on CTA-331P3.

The gene has a predicted protein structure typical of transport proteins with 2 groups of 6 transmembrane domains separated by a hydrophilic region (Fig. 1A). The large hydrophilic region between TM1 and TM2 is typical of OCT and OAT genes and is presumed to be located on the outside surface of the cell membrane. The OCT6 protein contains four potential sites for N-glycosylation and phosphorylation, which will be described below in Methods. Of interest, the protein sequence also contains a 22 amino acid leucine zipper motif, starting at amino acid 146, suggesting that there may be a physical interaction between OCT5 and ion channels or other membrane-associated proteins.

CLUSTALW alignment produced a dendrogram showing the phylogenetic relationship between OCT5 and other OCT and OAT proteins (Fig. 1B). This dendrogram suggests that the distinction between OAT and OCT genes, based on functional studies, obscures the common origin of both families of transporters. The actual CLUSTALW alignment of these genes is shown in Fig. 2 and demonstrates multiple regions of conservation among all of these genes.

The hydropathy profile analysis, multiple sequence alignments of amino acid sequences using CLUSTALW and the phylogenetic tree were all produced with MacVector software.

Example 2

Molecular Cloning of OCT6

BLAST searches of human ESTs in GenBank data base identified AI040384 (654 bp), AA033971 (714 bp) and H70190 (474 bp) sequences from three fetal liver IMAGE clones, 1656502, 429904 and 212935 respectively. IMAGE clone 1656502 (3', insert 1337 bp) ended the predicted stop codon, whereas IMAGE clone 429904 (5', insert 996 bp) and IMAGE clone 212935 (5', insert 966 bp) aligned with the 5'-coding region. All clones were obtained from the IMAGE Consortium through the American Type Culture Collection (Manassas, Va.). Each clone was sequenced in both directions. The sequences were determined using ABI Prism™ 377 DNA sequencer (Perkin-Elmer).

Example 3

Quantitative RT-PCR of OCT6 RNA Levels in Cancer Cell Lines

Total RNA isolated from 50 cell lines used in the NCI drug screening program was provided by the Developmental Therapeutics Program, NCI. Quantitative RT-PCR for detecting OAT-X transporter gene expression was performed by using a Roche LightCycler, which uses real time fluorescence detection for quantitative measurement of PCR products. A gene-specific primer pair was designed with Oligo 4.0 software and purchased from Integrated DNA Technologies, Inc. (Corvallis, Iowa) (F: 5'-GGCATTACATGCCCAAGACCAG-3') and (F: 5'-TGGGAACCTCAAGCAGCATTGGAG-3') (SEQ ID NO:13) and (F: 5'-TGGGAACCTCAAGCAGCATTGGAG-3') (SEQ ID NO:14). The specificity of the PCR reaction was confirmed by directly determining the DNA sequence of the PCR product. First, cDNA was synthesized from total RNA using SuperScript First-Strand Synthesis System (GIBCO/BRL) in a 20 μl volume following the instructions supplied by the manufacturer. The cDNA treated with RNase H for 20 minutes at 37°C and stored at −20°C. Then, 2 μl of cDNA reaction was amplified in a standard PCR reaction condition, using 0.3 μM primer concentration, with the addition of SYBR Green I Dye. After 30 seconds denaturation at 95°C, the amplification reaction proceeded through 45-50 cycles of 95°C denaturation for 0 second, 62-65°C annealing
for 10 seconds and a 72°C extension for 40 seconds, with
slopes of 20°C/s, 20°C/s and 2°C/s, respectively.

Quantification was performed using the LightCycler analysis
software. The log-linear portion of the standard amplification
curve was identified, and the ‘crossing point’, a threshold of
relative fluorescence, was determined as the best fit through
the log-linear region above the background fluorescence
(noise) band. The quantification of PCR product then was
derived by plotting fluorescence data in the log linear region
of each sample to determine a calculated number of cycles
needed to reach the fluorescence crossing point. The calculated
number of cycles required to reach the crossing point is
proportional to the amount of target RNA in the sample. The
relative amount of product was described in arbitrary units by
interpolation of the data using a standard curve of a series of
dilutions of a standard cell line RNA. The quantitative mea-
surement of each gene in each cell line was normalized to the
relative amount of actin RNA in each cell line, as a control for
equivalent cDNA loading in each sample. The results rep-
sent the average of 3 independent determinations performed
in duplicate.

A melting curve analysis was performed with positive con-
tral RNA prior analysis of the cell lines to enhance sensitivity
and the specificity of the data. Amplified products usually
melt quickly at a temperature characteristic for the products.
The fluorescence signal was acquired at temperatures just
below the Tm of the specific PCR product and above the Tm
of the primer dimers. All specific PCR products displayed a
single, sharply melting curve with a narrow peak. In addition,
PCR products were confirmed for specificity and correct size
by visualization of the LightCycler products on a 1% agarose
gel.

Example 4

Tissue Distribution

First strand cDNAs derived from 24 adult and fetal tissues
(RAPID-SCAN gene expression panel, OriGene Technolo-
gies, Rockville, Md.). The PCR primers used in this study
were the same as used in the quantitative RT-PCR studies. The
PCR reaction samples were denatured at 94°C for 50 sec-
onds, annealed and extended at 64°C for 30 sec for 35 cycles.
The PCR products were then visualized on 1% agarose gels.

Example 5

Cell Sorting

All human specimens were obtained in accordance with
institutional IRB guidelines. Leukocytes from fresh discar-
deduffy coats were isolated after RBC lysis with ammox-
nium chloride and labeled with lineage specific antibodies
(CD14, monocytes; CD15, granulocytes; CD3, T-cells; and
CD20, B-cells), and isolated using a FACS Vantage flow
cytometer. Each population was sorted twice to ensure puri-
ties of at least 99%. CD34 cells were obtained from discarded
aliquots of G-CSF-mobilized peripheral blood stem cell col-
celations from cancer patients. For each sample, the PCR
results represent the pooled average of cells from 2 individu-
als performed in triplicate or quadruplicate.

Example 6

OCT6 RNA Levels in Leukemic Blasts

Total RNA was extracted from leukemia specimens using
QIAGEN RNeasy midi kit. 150 ng of total RNA were used as
a template for the first strand cDNA synthesis with the Oligo
(dT) primer using the super script system (GIBCO BRL)
according to the manufacturer’s protocol. Quantitative real-
time RT-PCR was performed using an iCycler thermal cycler
with methods similar to those described above for the Roche
LightCycler. The results represent the average of 3 indepen-
dent determination performed in duplicate.

Although illustrative embodiments of the present invention
have been described in detail, it is to be understood that the
present invention is not limited to those precise embodiments,
and that various changes and modifications can be effected
therein by one skilled in the art without departing from the
scope and spirit of the invention as defined by the appended
claims.
<210> SEQ ID NO 2
<211> LENGTH: 578
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (264)...(264)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (268)...(269)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (274)...(275)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (410)...(410)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<400> SEQUENCE: 2

Met Gly Ser Arg His Phe Glu Gly Ile Tyr Asp His Val Gly His Phe
1   5      10     15
Gly Arg Phe Glu Arg Val Leu Tyr Phe Ile Cys Ala Phe Glu Asn Ile
20  25      30
Ser Cys Gly Ile His Tyr Leu Ala Ser Val Phe Met Gly Val Thr Pro
35   40   45
His His Val Cys Arg Pro Pro Gly Asn Val Ser Gin Val Val Phe His
50  55  60
Asn His Ser Ser Trp Ser Leu Glu Asp Thr Gly Ala Leu Leu Ser Ser
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly</td>
<td>Gin</td>
<td>Lys</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
<td>Val</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td>Glu</td>
<td>Gin</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Gin</td>
<td>Asn</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Glu</td>
<td>Ile</td>
<td>Trp</td>
</tr>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
</tr>
<tr>
<td>Glu</td>
<td>Leu</td>
<td>Ser</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>Cys</td>
<td>Ser</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>Asn</td>
<td>Lys</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>Glu</td>
<td>Thr</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Ser</td>
<td>Leu</td>
</tr>
<tr>
<td>100</td>
<td>105</td>
<td></td>
<td>110</td>
</tr>
<tr>
<td>Gly</td>
<td>Tyr</td>
<td>Glu</td>
<td>Tyr</td>
</tr>
<tr>
<td></td>
<td>Gly</td>
<td>Ser</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td>Phe</td>
<td>Pro</td>
<td>Cys</td>
</tr>
<tr>
<td></td>
<td>Asp</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>120</td>
<td></td>
<td>125</td>
</tr>
<tr>
<td>Tyr</td>
<td>Ile</td>
<td>Tyr</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
<td>Trp</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Thr</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
<td>Glu</td>
<td>Trp</td>
</tr>
<tr>
<td>130</td>
<td>135</td>
<td></td>
<td>140</td>
</tr>
<tr>
<td>Asn</td>
<td>Leu</td>
<td>Val</td>
<td>Cys</td>
</tr>
<tr>
<td></td>
<td>Arg</td>
<td>Lys</td>
<td>Trp</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>Met</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Gin</td>
<td>Pro</td>
<td>Leu</td>
</tr>
<tr>
<td>145</td>
<td>150</td>
<td></td>
<td>155</td>
</tr>
<tr>
<td>Phe</td>
<td>Met</td>
<td>Phe</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Leu</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Thr</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td>Trp</td>
<td>Phe</td>
<td>Ser</td>
</tr>
<tr>
<td>165</td>
<td>170</td>
<td></td>
<td>175</td>
</tr>
<tr>
<td>Asp</td>
<td>Arg</td>
<td>Leu</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Leu</td>
<td>Trp</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Met</td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>185</td>
<td></td>
<td>190</td>
</tr>
<tr>
<td>Phe</td>
<td>Leu</td>
<td>Phe</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>Val</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
<td>Phe</td>
<td></td>
</tr>
<tr>
<td>195</td>
<td>200</td>
<td></td>
<td>205</td>
</tr>
<tr>
<td>Met</td>
<td>Ala</td>
<td>Ala</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Ala</td>
<td>Met</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>215</td>
<td></td>
<td>220</td>
</tr>
<tr>
<td>Val</td>
<td>Gly</td>
<td>Phe</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td>Met</td>
<td>Gly</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td>Met</td>
<td>Lys</td>
<td>Ser</td>
</tr>
<tr>
<td>225</td>
<td>230</td>
<td></td>
<td>235</td>
</tr>
<tr>
<td>Trp</td>
<td>Ala</td>
<td>Ser</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td>His</td>
<td>Ser</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Gly</td>
<td>Thr</td>
</tr>
<tr>
<td>245</td>
<td>250</td>
<td></td>
<td>255</td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td>Leu</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Xaa</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>Xaa</td>
<td>Xaa</td>
<td>Tyr</td>
</tr>
<tr>
<td>260</td>
<td>265</td>
<td></td>
<td>270</td>
</tr>
<tr>
<td>Ile</td>
<td>Xaa</td>
<td>Xaa</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
<td>Val</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>Pro</td>
<td>Phe</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>Cys</td>
<td>Trp</td>
<td>Val</td>
</tr>
<tr>
<td>275</td>
<td>280</td>
<td></td>
<td>285</td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td>Glu</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>Trp</td>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Arg</td>
<td>Tyr</td>
<td>Glu</td>
</tr>
<tr>
<td>290</td>
<td>295</td>
<td></td>
<td>300</td>
</tr>
<tr>
<td>Ala</td>
<td>Gin</td>
<td>Lys</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td>Met</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Asn</td>
<td>Arg</td>
<td>Ala</td>
</tr>
<tr>
<td>305</td>
<td>310</td>
<td></td>
<td>315</td>
</tr>
<tr>
<td>Cys</td>
<td>Lys</td>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Leu</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Gly</td>
<td>Pro</td>
<td>Val</td>
</tr>
<tr>
<td>325</td>
<td>330</td>
<td></td>
<td>335</td>
</tr>
<tr>
<td>Asn</td>
<td>Ser</td>
<td>Pro</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>Gin</td>
<td>Lys</td>
<td>His</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Tyr</td>
<td>Leu</td>
</tr>
<tr>
<td>340</td>
<td>345</td>
<td></td>
<td>350</td>
</tr>
<tr>
<td>Asn</td>
<td>Trp</td>
<td>Ser</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>Arg</td>
<td>Thr</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Thr</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td>Trp</td>
<td>Phe</td>
</tr>
<tr>
<td>355</td>
<td>360</td>
<td></td>
<td>365</td>
</tr>
<tr>
<td>Thr</td>
<td>Gly</td>
<td>Ser</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Ser</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Asn</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>370</td>
<td>375</td>
<td></td>
<td>380</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td>Asn</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>Asn</td>
<td>Leu</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Val</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>385</td>
<td>390</td>
<td></td>
<td>395</td>
</tr>
<tr>
<td>Pro</td>
<td>Ala</td>
<td>Tyr</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>Cys</td>
<td>Ala</td>
<td>Xaa</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Gly</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>405</td>
<td>410</td>
<td></td>
<td>415</td>
</tr>
<tr>
<td>Thr</td>
<td>Val</td>
<td>Leu</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Phe</td>
<td>Cys</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Ala</td>
<td>Cys</td>
</tr>
<tr>
<td>420</td>
<td>425</td>
<td></td>
<td>430</td>
</tr>
<tr>
<td>Val</td>
<td>Met</td>
<td>Val</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>Lys</td>
<td>His</td>
<td>Tyr</td>
</tr>
<tr>
<td></td>
<td>Gly</td>
<td>Val</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>435</td>
<td>440</td>
<td></td>
<td>445</td>
</tr>
<tr>
<td>Met</td>
<td>Val</td>
<td>Gly</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>Gly</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Gly</td>
<td>Leu</td>
<td>Ile</td>
</tr>
<tr>
<td>450</td>
<td>455</td>
<td></td>
<td>460</td>
</tr>
<tr>
<td>Tyr</td>
<td>Thr</td>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>Pro</td>
<td>Thr</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Leu</td>
<td>Ala</td>
</tr>
<tr>
<td>465</td>
<td>470</td>
<td></td>
<td>475</td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td>Ser</td>
<td>Met</td>
</tr>
<tr>
<td></td>
<td>Arg</td>
<td>Leu</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Ala</td>
<td>Pro</td>
</tr>
<tr>
<td>485</td>
<td>490</td>
<td></td>
<td>495</td>
</tr>
</tbody>
</table>
Val Asp Leu Ser Ser Ile Trp Ile Phe Ile Pro Glu Leu Phe Val Gly
  500  505  510
Thr Met Ala Leu Leu Ser Gly Val Leu Thr Leu Lys Leu Pro Glu Thr
  515  520  525
Leu Gly Lys Arg Leu Ala Thr Thr Glu Glu Ala Ala Lys Leu Glu
  530  535  540
Ser Glu Arg Glu Ser Ser Lys Leu Leu Leu Thr Thr Asn Arg
  545  550  555  560
Ser Gly Leu Leu Gly Thr Glu Ala Ile Thr Pro Arg Ser Gly Leu
  565  570  575

Gly Glu

<210> SEQ ID NO 3
<211> LENGTH: 123805
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 3

aagcttgccc aaccatgggc ccaagggcga cattggtgct aagatgcct tgaatgcagc  60
ccaccacaa tttgtaaact tttcttaaag cttgagatct tttttcaaat ttttttttaa 120
gctcatacg tatacggtat tttttatgtg gtgcaagaca atttgtctcc  180
ttccaatgtg gcocaggaag gccaaaaagat tgtcaacccc tttgagttct ctaggcaact 240
gggccccagt gaaatttgta tcaagggagga tagtagagtc ecgtagtagc octgaggaga 300
tgcccaccag acctcataag atacacacag ttccccaggag gcacccacact gttggcaaca 360
tcaagaaatc gttaggattt aggaaagca tcaaacccccc ggtgtggtgg agctgtgtgt 420
ttaaggaagtt cccttttttt catttactgc cccatcgaag actgagacat aatcgtggca 480
caggatgtcg ctctcaagtta gottgctggcc gctctcctct ccatactctaa atcacaacct 540
cttctgttg gttagagtaa aatcctcttg aaccacctac cttgctgtgc taggtcaatctc 600
tagacaatct taagttcgact ccattaagt ttcctttaag aaatgtttgtt aaatatcctt 660
tcttctcagct gmtgatctccgt tgtatggct gcaagacacac tttgagttct ccaacattcca 720
gttcccaat ctttttacta tgtgaactat ctcacacattt aagctctgtact aaatcctt 780
tctctgtaa aagctgcaat tcataaccag aagttacagtg tacatacaca 840
gaggtgaact ttccccccctt atgtaaacac tttgagttga ccagagttgaatttctttctgg 900
ccccctccg ttccctggct caggtgtctcg gcccagctctgtttgcccaaat gaaatgtcaga 960
ggttagggctc taagagacag cattctggat gttgctgtact tctctctgacc tctgtgaccc 1020
ggacaagcta gggaatcacc cttttttatat cttgtgactt acgtgctgactt ttttctgg 1080
caggtttccg ttttttctctg tgtctctcatt ctctaatgacta acgtgctattg 1140
gttggtagt accttgcagat gaagggccct attggctcagat gggagttgtgc caaacagctaa 1200
gcagatgtgtt accctgttttc aataactcttc ggtttggagag ggtattttttttt taataaactgg 1260
gttgctgctg tgtgttatga attaactaaag gggagttgctc tgtctttttttt caattcactaa 1320
gcataagac cttctataact cttccttcttg gttctgtgct gatgtgagat 1380
gaggagaaa ggggacagc cttccttcttg gttctgtgct gatgtgagat 1440
tgtgcctctt ctctctttttt ctttcctttcttc acccagtttc gttctctttt ccatacttttt 1500
aacccctgtg tggggcacag cttccttcttg gttctctttt ctttcctttt ccatacttttt 1560
aacacactgt gaaaggtta tattggttc cattgtggtt gttgtggcatt ttgctctccttcttggtggctc 1620
aggtgcaagt gagctgaaca tacattaaaaa atacaacccc ttaagagctg actggttaag 1680
acttaagccc agtatctttcc aagagatgagt gtctaggggc acacctccaga tcttgaacctg 1740
cctgagtgt caagcaacgg tcgcacagt ttgatttcct ttctctataaa atgatacgtt 1800
gttctgatgg atctctggag tcctcctggga gttctcttgag gttcattgcct aaacccctac 1860
tctctctcttc acctcttctt ggctgctgaga ttcctgatca ttcctctaatg ttggtggcac 1920
tgacattgta gaaggcataca gaggaatgg aagcttcctac attataacg cagaatattga 1980
ggcaagagcc caaactatatt gcaaaaacca tggcagatgt tggataggaag aagacacgtga 2040
gacacagcca gcacacagagc cttcattaacc tctgacccca aagagctcttg aacctgaagtt 2100
cocaaagctc ctttcttccc cccaggcact cactgtctttc aaagcacttc caatctcaag 2160
ttgggagatt ttgcggcgct cagggctcrc cgcgacactca ggcaacaccc cttcctctat 2220
ttcgattttct cctctgcaac atggagagta tacaattgga tggatttttta aagcagacgtc 2280
agagctaaaca tcctcaacct cagggactcc aggaaagcga caactaaacat aaggtctctga 2340
aaccaatgaa aaaaaacagcc aagcttctgcgc accaaacagtt ttcacaccat ggacgccagc 2400
ttgcaagagga aaccaagtgg agtgatccct tctgtgcccc aagctgctgag cttggaagac 2460
ttcggcctggc atgagagcag gacatgaaat acaggtggag ccaacagcata cccctctgtgta 2520
gaaagctgag ggagggctcc cagggacact tgtgcaacat ttcacactaag gaggagataac 2580
ttgaggagag aacagtctcgc acggagagag cctgctgctc tctgagatgt catcttggg 2640
ttcgacgca aagtcggcag agcacactca tctgagaggg aagaataggg ccagaaaaacat 2700
aatccacacc aagattagca ggcaacacct cctcacgtgc ctcacacagc gcacataagcc 2760
cctgctgccca ccaacacacag tggaaactct ttcacattcat cgggcatctga tggacacact 2820
tcgagtattc ttcagataat gcacccaatc tcaagctctc tctggtctcc 2880
ttcgacacg gatttcaaacag aacccctgct gcacccaggtgt gttccgtcctc gcagctccgc 2940
catgcacggat gggtggggtc ttgcaccaag tcaacagcttg ttcgccccagg agttcagacag 3000
cagctgcccc aaactgttaga gattctgtat cttttttttttttttttt ttttttttttttttttt 3060
aaggcagccgc aaggggtctgc aacocctgtaa ccagctccact ttgcaggggc gagggtgggtc 3120
gattctctg aagcggaggtgt ttgcagcaca ccacagagca aagcgaggat cctgcttcga 3180
tcataaaacc aaaaattagtc gggttgtgtgc gcagacccga ttaacctcag ctacggcagca 3240
ggtgcgcccc ggagacactggg ttcgaacctag gaggcagaggg ttcgagctga tggagactcg 3300
gcagtctgtg ccaagctgctgg caaacaagag gaaacactcat cttaaaaaaa aaaaaaacaac 3360
accaaaaaaa aacaactgtat gagaagacct aacaagttttt aacaagttttt aacctgcactt 3420
agacaacacc tcaaacttttt tttttgttct tttttgttct ttttaaagat ttcacggcaca 3480
ggcattgggt ctccacacta taattccccggc aatgggggag acggagaggg gttgcacactt 3540
tgaggttaagg aatggagac ccagacgccc tcaacacctg ccgaaactctg tgggtttctaa 3600
aatataaaaa actgccccaggt ttggggtgggg gggtctgtct gcgtccaggtt actcggaggg 3660
ctgagagcag aagatctgtt gaaacgtggga ggaagaggtt aactgcagct ggattacgctg 3720
cacggacatc cagctcgggaa cacaagctga gactgcgctt caaaaaata aaaaagaaaaa 3780
gatattccgc cactccagca caaaaatcct aacgagacga aacatgaggg aaatttataac 3840
aagacacccc ataatccatat tgcgtcaaatc cagtgtatat gagaaaatcct taaaatatcct 3900
aacaggaaaa aaattttgtg gaaacacgca gacaagatg aagctgctgca gcagacttga 3960
ggaagagttg aagttaaaa actctagtac ctagatctct cagctagtt ggtggaggag 6420
ggaaccccttg aagaaagagg ggaggcatgg gggagcctgg gcctcaggag aagctcggga 6480
ggtgctccct ctccttggtct ctcctcttta atctctcatc cgcgccttgag 6540
gctcctccct ctcctcttctcc tctccttcct tcctccctct tctcctctctg 6600
cattttaaac tctctgatttc tctctctctc tctcctctct ctctctctct 6660
cctcctgaac tgcctgaccc tgcctgcttc gctcctctct cctcctctct 6720
cgcggtgcgc ccggcggtgc ccggcggtgc ccggcggtgc ccggcggtgc 6780
ggaagatgtga aggtctttttc tagtggctcc cacaatgccc tagggactct cgatgtgctt 6840
ttggtaggt gcacatgcca tctctgaacc aatgactatg cccgatcttc acgtggtgcct 6900
cacagggccc aacggtgggt tgcctggtga tgggatgaac cctcaccag ccccagggaa 6960
gtgaagactt ggcgcagcaga gggcctcttg atgaagcttg agctctggttc tgggaagatgt 7020
ggggtatcct ttggggctat catctccat cagcagtcga aactctcttg cctatggaag 7080
getatgacca aagctgaagat ttcgctttta aatgtgagaa tctgcagttg catacttaag 7140
aaggggagaa aaataaaat aaaaatttaa ggtgtttaaa atagatcaac tcacaaagct 7200
tgtccaccag acacaaacct gaaaggataa catattaacac tgcaggggtt atttttttct 7260
aagaaaatat ccattttaac acaagggatc cttaagggaa gcacacagt 7320
tcttcttttcc caacaccaaa tccaacatttg aatgacagga gtaagagctt 7380
tccacagttg aatccacagc ccctgggaatt gggcaggagc aggagccctg gaggtcagga 7440
gttggagcc ggcctggaaca acaactgtagg aacccttttc tacatataat aacaaattta 7500
actggtggtt tgtgggaagt cctctgtacc tctctcaaac ggcgggtgtg gcgaagataa 7560
cattgaaac gcggttacct gaggcagagt gggcccagctgc gttgcaagct acctcaccgcc 7620
tgaggagcag aagctggtgg cctcctcctc aaaaataaaa aataaattct gaaaaaaataa 7680
ataaataaat aataaaataa aataaaatat aataaattact attataaca acttagaactt 7740
caacgcggac atctctgctg aggcttcttg ggttgtgtgta tttttatata cttacaacac 7800
tccacaaagc aaccttattta cttcttactt ccacaaaggg aaaaaaaaaa tggagagat 7860
tgtataacc ccaacagcttg acacacttaac tgtcggcacc cagctgtgta acggccagtt 7920
gtatgtacgc ggattgcctt ccctgcattag aacccaatag atctccatac tggttgggctt 7980
tctcttgtttt cagccagctg acgtggttaa gggcaggaga aagcactagag aacgtctattg 8040
caagcttttt tttgactactc acgcagmtgc gcggcaagaa tagcactaag tggaaaaaggg 8100
atccaaagaa agaatcaaaa cgttactgta ctttttattta aacaaatattt actactattg 8160
tctctctctct ttggatggag ggcggccact cttgatttttc cccagaaattt atctctctttt 8220
ctcgctggct cagcggccgg ggatggtactt cttctactag ttcgcttcgc tctctctctct 8280
taacgctag gggtttctag ctctctctct cagcctagct taagcctag 8340
ggtcctttt ccttctgttc cctcctcttg cctctctctct cctctctctct 8400
gcatactt ggcggcccaact tttgaatcctt gggcaggagc aagcctctgtaa aagcctctgtaa 8460
ggaagatgtag tgttggtggtg ggttcctaat aaactgtcttg cactaaagggccagcagctg 8520
pgatgctttc tcctgagcagc agagcagtaa aggagtatttt tggagatctttt tggagtat 8580
gtggccctct tggaggtgga taggcccttg cccagaggttg aagcctctgtaa aagcctctgtaa 8640
agagaagaga gtcgcctgtga ctttttatttt cttctctctct cttctctctct 8700
agtttgtggga attggacagt tttctgctca gtcagtatct cttatatttt tattgttatt 8760
 tttaacagaaa gaaaagaggt ctttttagaat aattggaaac tcacactctg catgagggaca 8820
gataagggg gaacgctggaagg attgtgtcag agaaagagcga gcatgtctgtt ttgacgtag 8880
caaagaatat gtggacatgagg ggtaggaggg ccagggccacc tcagggagca ttgaagagttc 8940
agcagtacag ttcagaggctt attotccaag taagtggaaggt taatagggcagt ttgggtactag 9000
atgtatacag aagaggttct cttgagttt gtaagtcctg aggctttgg tttgtttgaga 9060
gcagagtct ctctcagttg cccaggcttg aggctggtggtag taagatctca gttatagtgg 9120
atctctgcgt ccttggtcata agcgccacct cttccccccc cagcctcccg aagtagctgg 9180
gtcaaggttg cgatacaacc aagttgtgctt attttgatttc ttatatggtttttt 9240
agaaaaaggg gttttgtagct gttgctctgga gctgtcctggc tctaaggtact 9300
acaccaacccag atgtggttggaag ttataggtgt gtgacacttgt ggtggtgcag 9360
aactgttatttt gggagggacaa tttataaaca atataaggaa taagacagagacaagaggggtg 9420
gcataagggca agccctgtgaag atgcgtactata tagatgtcctcgcggcgtgcactg gaaagggcag 9480
cgtcagcata aagatgcaaga atggaaagaag ctgcagggag gagaagctac tatagtagtg 9540
gaagtattct gcgtaggtgct atacaggaag acaagatgtgt gtcagacttg gtagctcata 9600
gggggtgtgt cggcgtagca gattgggagc aagcgtggagg aagccaacagt gtgaagagaa 9660
gggcatcatt gcggcacctttt ggacactcgct ggtagccgtgg aatggttaggtg 9720
gggatgcgtta gaagacaggtt ggaaatacaga atgtgtgattc agagacacgg ccgacctctca 9780
gtcagcaatc cagacagagaga agggaggagc ggaggaaggg gctagagaaaa tgtagagga 9840
aggggggtcg aagaaatagtt gagaagggaga agagagagag ttcgaggaag taagggaggg 9900
gtttaagaaa aaggggtcgg ccaggtgtgtg ggtgctaatgc tcaacactttt gggagccaaac 9960
ggccacaggg gcatactgga ggcagcaggt ccaagacacc actggagcaag atagtcaagac 10020
ccacatttcca caaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aatgtagctgg ggcaggttgtg 10080
ggcacctgta gtcocagcta ctcaagggagc agggagcttc ggtgtctggtt agcggcaggg 10140
gtcagctgct gctgaagggca tctatcaact gcagctcgag cggggtgacca taacagagcag 10200
ttgctctagg aaaaaaaaaa aagaggtgtgg gttggttgag taatgtagttt tagtaaaagg 10260
cctgagaggt gtgggtagtt gcagctgccaact cttggtccttt taccccatac cctgttacg 10320
aagagtccct tcaattccaa acattaaaaa aggtggtgag aactcataag gaaacataaca 10380
ttccccattt attaaatattt tttttatttt tttttttttt ttttttttttt tatgctattt 10440
agaaaaactt taacttattta aaaaaatcata ttggtcttacat tgttgtattt aatttttttt 10500
acataataaa ttgggttttaa aagtagtttcag tagtctcaaaa acagtaacttt gatacatcatac 10560
ccacacttgg agaattggta tagotaagt ggaggaggaga agcatggaagta actacagttta 10620
acggtgtaaca gcgaagaaaag aacctgttttt tttttttttttt tagtagggaa atggtagggtt 10680
tttttttttt tggcagcaggt tccacacagc gtcgcctacat cttttttttt cggctccact 10740
gtagagctcc acacaccacct acacagggaa acactctaac atggagaaa acgggtaag 10800
gtagactag caggtgtggg caggtgtgtgg gtaggtgattc ggcocacaaaag aggtacggcc 10860
acgtctacag ttcggagcaact gttgagttgac cttatattttt taagaggact tttggcagag 10920
taatccataa ggtctctggag aatgagatctc atcggtattcct cagttggttgc ctgaaatccaa 10980
tgcgagaggt gttttttttttt tgcagacagtc gcagagagag agagagagag agagagagag 11040
gtagagataa aggcaagaggt ggtagcagagtc cagccacacagc ccagagggag gtaggggtct 11100
ccaaagcctg gaaaaagcag gaaagttctc cccccatcgc ctttgagaga agcactggcag 11160
agctgacgcc tggttctcag accctctgcc tctcgaacca gttgagagat gattttctgt 11220
ggtctgcaag cacattctg tctgtgttga ttgttaaag cagcctcagg aatotaacag 11280
aggtggggag ctcggccaca ggtcgaggtc ctgagctagtg gggtgtgtgc tcaggtctct 11340
cattttccac cccccacccag gggcgcacac aatcttttat tttggtttta ttcaccttgag 11400
gottaatctc acaagcagcct tatttctctca tttgtgataac ttttttctctta 11460
gtctgtcctc gattgtagttt tagaaaaacc acagatatttt ttttttttttttaaaat 11520
tccacacgtt ttacccagca agtatagaaag tcaacaaaata ttctctact gatactcttc 11580
ttttttctc tcaattttta agctggcaca ttggcgccag cggggcaccag ggcctgcaag 11640
gggagctcct cctggacgta ccaectctcg cctgacccac cgacagaggtg acggggggcg 11700
tgtgccggag cagggccgacg tgtggcaaaag gcagccccgg caagggagacc tccctctgtg 11760
aatgtctgtg cattttagct gctgargggg attgctccca cccctctcaa aggtataggg 11820
tgctgacctct cagcaccacag gcagactgta ctcgccaggcc gccgccccct 11880
gagacacctgg ctccctctctc ttcctctctga ctcdcctctgg gggagatag tgctgctct 11940
ggtgtgatct cttcccccaag cccagagctg ggtctctattt tttgcagacc aagggagcgg 12000
gggatctttgg ttcaggggga getccatcag tagctgccgcc cctctgtgtcc cttttgttt 12060
tctcctttgag gttcagcaac cccacctgag cgcgtcctgc cttctctgtg 12120
cggtgtgtcg ccagcagccg gccggcaggggg gcggcttcgg gcagcccttc ggaactcag 12180
caggggtcct gcggcaccac acgtcacttc cccccaccag cagcaccacgc ctcacggtgct 12240
ggtgatcctc cttttctctt ccacccctcc oggggccccg aggcccacagc tccctctgg 12300
ccttctcact cgttcggacct cttccctctgt ttctcgagct ggtgcgctcag agagtcctta 12360
tctggacacct cttgtggttgg tattctgtcgg ggcgctcccct cttttgctcg 12420
ccggtgttgg cccccatcct tccggccagaca cttggcagctt cttggccacc accccagacc 12480
ttcctgttttt ccgtctgccctt aatacccttc gcgttctgca tgaaccctcct acctcgtgca 12540
cagtgccttg ctggttctct gcacccacgg acaccctctg cggcctctgc tctgtctctg 12600
tccacccagc tggccctccc ttcgctatct cccacacactt ccctttaaccg gaaactactc 12660	tctccctcc caagacactc tcacacctgc cttcccccgc agtccttcct ttcctgctca 12720
tttccttcttc tgctgtaagct catgtctttct ggaactctcttt ccacccgcac cttgtgctcg 12780
tacatgtcct cgtattgatg taccgttcct tccatctttt cttcttaaccag tttgtttcgg 12840
aaaaatccca tctgctccac tgtgcttagaa ctttaaatag tggtaaaccac 12900
aggtggttgctg ctctcccaggtc ggtgtctctg cttgcgctcc cccacgcggt 12960
tttcctacag cccgctctgc ctgctgctgtg cctgggcccc cttggcccaa atggccccca agcgaacaca 13020
agaagttcttt tctctcctata aacaaaccag acaggttggc gccttgccagt gcaccggggag 13080
agctggagct acaggggccag agcttgagtc gcgtggccca atccacaggg cccatctgccc 13140
tctgctgtgct cggctggctg tgtgctaaag ccccccacgg gcagagcagag tggtaaagctg 13200
caggggtcgct atccctgctc acctctctgt ggtggacctgc gtcggctgtc 13260
tgctgtagttt cctctctctgc ttcctgggtc atctgctctg ttcctccctc cctcgccca 13320
gggtggagcc cccctccccc gcccaccacc ccatccctg tctcaacccct tccaaatacg 13380
gttttacta aactcccctg agttcactta ttgccacatt cttctccgcccc cggcagccct 13440
ccagcacaag cacgacacatt gagaccgaata gaaagctcag aacaagagct 15900
gtgtgccact accactctga tcaacaaggc tcacccaaaa aaaaaaaga gaaaccttaa 15960
tccattaatg tagtggcctag tagtggctata gccatagca gaagagaaaa acctagcctt 16020
tacctttcct cagttacaaaa agttttctca agaaggtatta aagatttaaata taaaagcct 16080
cagtttataa aatactctaga aaaaagaaaac cggatctccc tcttctaatgc tgaacctgcg 16140
aaaaatatt tggctcaagcc cttaaagcaca attgcaacaa aacaaacaaat ttggaacaagt 16200
ggagcttaatg atatgaagga gctttgctcac acgaagaaag aaccatacaca gatgannea 16260
gacagcattcag aagagtggttt gcacaaatttt ctcaacactat gcaagtcacaa aaggtctaat 16320
atccacagaa ttaataagagata ttaaattcaca aagacaaaaac caataactct cattaaaaat 16380
gggtccaaggg gacaagcaga taacttctcaaa aagagacagt ccaacaggc gacaagataa 16440
tgaaaaaggt tcatctctaa cgagttcctag atcagacgca acaaacaaca 16500
ccactcaca cagctcagaa gatgtcctct gaaagcattaa aacaacacag atgctgtgga 16560
ggtgggtrag aaasaggagt gctttacatag ttgttgtagag aatgttaactt aggctcgcca 16620
cggctggaag aggtgtagag attttgcataa gccatctaaa ccgagctcagg accggcctaa 16680
gcgtacccctcatg tatttctctt ctcaccccag gaaacagcat cattatcaca aaaaagacaa 16740
tgcccttgcga ttgcttatcag acgtatcttc acagtacgaa gacagtgagaa acataatgag 16800
tgccctcacg cagttgcagg gataaagaa gttgctgcga tatacgcgta ggtatactat 16860
aactgctataa aacaagatca aacagattct ttcagaccaata tgggttgagg ctggagccgg 16920
tactcactaa tgaatattcct atggacacgaa aaaaacattta cccgtgctct cacattataa 16980
tgggagctta aacaagccag acatagtca taaactctgag gaaacagcag aactgtgact 17040
dactaggggg gaggcgaagg agcaggattt gggtgaaag aactcactatt aggcaactattg 17100
cctcacaactt aggctcaattg gcggcctctga aacaactaggc atgatggcccc ctgtattccta 17160
aaataaaagt tattattttttt aaaaaagat attataata cacattataa tatattatat tattgaagatg 17220
atatatatatt tggggctcag cacagtataac tattagcatca taggaagaca cagcagctatt 17280
tataaagatttgccacaa gggaggttaa gaaagagcaggtc atgggagtctg tccatctataa 17340
genccatttt ggtctgctagg cttgggtttgg tggagagagaa atccggggtcct gtaagcttgg 17400
aatggaataata ggtatgcccc aagctcaccag acagtttaga ttcacgtgta caaatatacaaa 17460	tggtgctct ctcaccaaaaa aatttacacct tcatctactc ctctctttgaca atgatgacggag 17520
ntagcggcttg ctcacctagaa gcctctggag ggactttagca gggagaaatg ccgctcttcc 17580	tagttcagg ggtgctgcttg cctagtacctg tcgctacgctg gacacgttccc tctgcaagccc 17640
cagcgctgctg atctctctct gcagatcctg tccctctgacct ctaactccta 17700
aatcagacatt ggacagatccct cctgtaggaag gttggcagaaaa gattgggtac 17760
cggcctttgg gtcggctcttc tggccaccgc tgggctaac gttcgctgggg tgggctttcc 17820
agcatccag acgaggattgt ttagagacggc tctgcagctct gacacgggtaag tctggctttcc 17880
ccccctctgag aagcctcgag ccagcaccctc ttcagagggc gctgtcctca ggttctctcttctt 17940
gaatggttata ctcagacctg aggggagata ggtggctctgt gctacaggata ccctctacttt 18000
tccttaggcgtccttacagcgagtttctctt ccacctactc cccctcttact 18060
aatatttgtgt tttatattctttt tttgtagaaaa ttaaaggcat ttggactttc cagaggttag 18120
tcatttacca atataaaaaacctt gaccccaatttataaaaaa aatatgacccct cttctttttct 18180
aaacactg cagaacac ggaacaca aacaacat gaaacactt ccaacttat 18240
tttactaag tgaagattc tatattcttc ttcagttcct ggctctctgt ctgatttttc 18300
cgtattgta acctgctgcc ataaaagagt tctcaaccct aatgatattg tgtuatttag 18360
cctacagct gctagggtcag gcccctagtg ccataggctca atacattct 18420
gcattttgca tcaaggcaag cactataata ttctcatttt tctgaagcgct ccaattttg 18480
aatgctagtt ctgtgctcgg gatcacaggg ggtgctcagtt ttttcctata tcaaaaaaag 18540
cacagacaa aatatctttc ttcaggtgtgc atcaacccca cctaattttt tctottgttt 18600
cctcctttct gattctccta acaagtggggc tcaactatttt ttcocacgat agacaaga 18660
tttacagaaaa aagggctggag atctcagaga ataatgttgctg aataagaaga ggggtgggagat 18720
tttctactga acactagctg tcaagaaattg gcaaatccct gcctgctgtag ttttctttgc 18780
cagggaaaa aacacgtotcga aggctgaggg cagagggggaa tctctggcgt ttgagagga 18840
atttgaacct gggaacagag gttctgttacg agataagagct gcttttttctg ttcactcttc 18900
ataatgtaao ctgtaacattac atactgctga ctgagttcgtt gggagagagtt 19960
ctactaaagaa aacatttttgt ttoacaagat agcctccaa ttcagcaagct cctgaatgaa 19200
tttgagaggg gaagacacatt ctgcaagtct tccctacttt ttcactttttt tccacttttctt 19260
aatccacttt gttcctgcag tttacacagac gctgacacat cttctgtttgc 19140
acctactgga gttaatagatta atctctacttt ctactctacttt tcatttttttctttaaatgttt 19220
tctgggtata aagccagcag aagaagcttaa gcatagcaccc ccctgtgttgg ttctctttttt 19260
tgtcataca taattgttcat tgtacagactc ttctttttttt catttaaaat ctaataa 19320
tttgctagc tcacaaacacac agtttctctc agttttctgta cttggaatca 19380
cgtagcattt aasaagctgtc ttttttttaaacc tactacattt actaagctttgac 19440
tccggagaaga caatactagat attatgtaat atatatatat tatataaaac ttcttgctt 19500
gctgtagat gttaggcctt ctcagaaagtt cttctagacat agcagatgtttag cagccacaaag 19560
cgagaaaaac tgcagctgtgg ctctcgacaat cttgaagagtc tttgagactc ctgagaaaac 19620
atgtcataca acctctgaag cagacaaccc cttgctttcc tttgcttttc atgaaatgc 19680
ctctactattt gcatcttggc gcctgctgacat tcataggct gcctacccctgcggtg 19740
acctctatga atgcagcaca gctgttttatc tcatctattc caggacttaa gcacagactg 19800
aaagaagact gagaggagat atatattaat ttgctctttt ttgctctctt caattttgct 19860
tttccttcct ttgctctctt ctaaaagcttt ctaaaaccaca ttctctccaa cttacacgct 19920
tttctttaaa tatttgaaat tttttagtct tcaaaagttg gtaattaaa ttctcccca 19980
aatattggct gctgtttagta gtagagatgt tgaattcaca acocctgtttg atccaacagc 20040
actagaaagag gtttttcccac ctcataataa gataggagat atccacccag gagaacaaatt 20100
gctttgcctt tcttcttttc cttttctttt cttttctttt ctttttttttt tttttttttttt 20160
cacacgctgaa caggaacttt cttgagaatg gcctgctcct aaggtgcttt cttactctca 20220
gaactattttt caaaggtttat tttgctctttt cttaaacaca ttttctgctg gttaaacttttt 20280
tttgccccca tcaacattat ttcttctccc ccctcctttat cccattttat ccagcattttt 20340
gccocatctt ttttttaaa tctgaagccc ctttctcactt ccttttttttt ttttttttttt 20400
ttggttttct tttgagagtt cctctctggtaa taagtggatg tttttttttttt ttttttttttt 20460
cctattttt catttttttt tttgagagag ggctcctgtt gtacgcttgg cttttttttttt 20520
atggtgcagct cctttgtcag ttgcaacttcc cttttctttt tttttttttttttt 20580
cagctctctg atgtgtcaag attacaggtg ccacaccaac ctgcgtgactt tgtatatatt 20640
tagtagagat gggttcccag cagcctggcc aggcggtagt tagaatcctt acetccaaatg 20700
atacaacgtg ctacagcttcc ctgatgctag ggtactcagg agacaccaac ctggtccctgc 20760
cattatctcc ccaatcctttt ttataaagta gtttgctgac atcagcaaatacctattt 20820
ttcgctttact actgtaagaa tattgccca catttcgtaa gttggctgt cgtatataa 20880
gtctgctgct aataattttcttctaatgatc tgtcattttt cttctgaagct acctcttctg 20940
gtttagccat ctgcaactttat attataactc atctctagactt gggggtgagt ttaaaagact 21000
ttatctctgca tgttatattgg tgtgctctctg taactgceecca tgaagttcttt ctggtccctg 21060
gccgatag agtccaatatt tcaaagacagg agaactgcag aagccaaaga gtaattttcg 21120
tgacagagtg cgtgaatagga aaagtttaat acacatagag caggtctaaat ggaagctgta 21180
agctttatct ctagcctcctg gaaactcaag ctctccagc gggggtgagt gggggtgagt 21240
gaaaggggag tggcaggttg tggggtcctag gataaattc caaggggatcc agggttcctg 21300
cctgctgctga gtcgctgccct gatggttggcg ccacagccga cggggccagcg tttctctgtcc 21360
tggcgtggac ccagctgacat atcagagtag cagctgtcaga aatattttc aacacaaatc 21420
ttattagattcct ttcaccacaa catttaaaag ctagaaagact tgggctgccct tgggtgatata 21480
actctaaagc cataatttggg ttataccgtt gtaatcagag aggcaagttgg 21540
gtcacaccag aaggggactgc tttggttcag gaggaggtct ttcgctttgct tgaattttcg 21600
tataaaacta atccctccaa aaggttagtc atcagatgccc caggaagggag ggggacgagg 21660
ttggttcgttc aacccgatag ggtcagctgt ggctggactgc ttcctttctctgctcaac 21720
tctcgtgtagt cattcttttaa ccagctgtgcttt cttactctctc tattctctctttctctc 21780
atctcgctctg gaaactttct tttctattct ttttaaatat tgtgctttccct cttattttttt 21840
ctctctctctct ctgcccaaccccttattattt ccagctgtcaccatatctag 21900
ccttggatct tttactcctg cttctctcttt ctgtattttc cctgcgtattt ccagctgtcaccatatctag 21960
ttaattatctcctt ctttcaggtct tttctttata tctccatata tgggtattctg 22020
ttttttataac ctttcataaat ttattcttttctttcttttctttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttctttt
-continued

tgatataaca aggtagagat gggtgaagaa ctgatacaca gagtaagtgct aggatctctt 22980

taaagggcct gaagacctgta ctgaggtggtt aggaaacagc atgaaacagag gcctagagag 23040

ggaataagta ggtggtctac tctggtcctt gttattacac aagggctctg cctgtctagt 23100

atctcttgct tccctaaatt ggtagacttt tggataaca atatctcttt 23160

tgagatcttt ctcatacttc cttttaaaat atatggtgct cccaaattaat cctactagtc 23220

rttacagag tctgctttac atggactattt cttatatact tcttattactt cttcatgct 23280

tcattgtgag tagggctgtgc aatacatcaaa agcaacagca ctctcacagag atgacacctg 23340

ctctttactt ttgcgtgcaaa agttgctgctt tttctctttt attgtccact tattggagttc 23400

gggttggaga gctctgcttt ctttctttta acaactaaaa acacgcttcag tgggcaataa 23460

agcacagaa gttctctcttt gttagcagcag agtgctctgt gcctacaccc caactctgccc 23520

tggcagactt cttcagcact tttttttttt ctttttttct aaaaaaataa tattotgaaca 23580

agggactaag ctgaaactgc tttttttttt gttctcttctg cttttttctg aaaaaaataa 23640

attgtaagac acaatccttt ctaggactttt cttttttttt cttttttttt 23700

agatagctaa gctgtcctata gctagatattc tttttttttt ttgctgtgcag ggggacatgg 23760

tcccaactta caagcctaca gtttctcttt cctgcgccttct ggtgtgccc 23820

ctggtatttt ttagaatcct cttctcttcct ctaggtttcc cccggagactc aggggtgttg 23880

cttttctttt cttcctccttt gtttggtttt aatccagact gttttgcttt 23940

agataagact gttgatattc ttgctctcttc aagaaacagc atgacccagt tggccacagt 24000

attcacaagag gacactcttg ggtggaattt taaagctact aacgtgttac ctgcctgacc 24060

tgatgctgat ggtccagattt gatctactac ctgctctattc tggacacttg 24120

aacactgtag gttcatttcct cttcattgg aactcagaa ataggtcttt cctaactctt 24180

catctgagaa cgtgctcaag ggacactgac atgtgctact atacactatg gtcgagaaac 24240

gcaacactgag cagggctctgc ggcacatagc cacaaggctg aaacactccc ctagctgca 24300

aaaggaacac gccaaaatttt ttcttcagct cttgtgctgt gcctgctgca tagaaacttc 24360

tataacctaa gacaccatga aacaccaaca gttacattc attccagact cgggagaaa 24420

ttctcttcct ctaactgttt tcattatttt ttctttttttt ttttttttttt ttctttttttt 24480

atatactct aagagcagaa aggggctttt tattcttttc tagaatattc gagatccaaa 24540

tacagactgac gaaactcagc ccccttttttt tttttttttt ttttttttttt ttttttttttt 24600

ccccttacaac agtttggaaat atttttggttt ttttttttttt ttttttttttt ttttttttttt 24660

tgctccatgca ggtcgatagtg cagtttcacg aagcaacctt ggcacatagc 24720

ccagggcttg cttcagaact ccctccctaa cagtttttgc agaatcactt 24780

tgacacacaa ccacactaagtt atttttgttaa ttttttaaga ttggtgggatc cgcacatttg 24840

tgacacacaa ccacactaagtt atttttgttaa ttttttaaga ttggtgggatc cgcacatttg 24900

tctccactaag cctccagttg gatggtgtcc actccgccgc caggtcttcac ccctccagttg 24960

tgtgtaattt cctcggttcc ggtggatactg gctgctgctgc ttttttttttt ttttttttttt 25020

agctggttga ccctctccagc cagtttttttt ttttttttttt ttttttttttt ttttttttttt 25080

atttgacact gtagacttattt ctgctctttt ttttttttttt ttttttttttt ttttttttttt 25140

ccaggtggac ccgtgctgagc ctttctcagg aaacatcagc ccctccagttg 25200

aatgataat aacacttccac aatccatggc taacttttttttttt ttttttttttt 25260

gttgagagtt ttaagaaaat cgtagagcgtt aagcataagag ggtcataagag 25320
gttgccata taatactttgt ccaactttaccttttgcct cgcctaatgt cgccttttac 30120
gctgctacta aagcacttac ccgagctggg aagaaata aatttttag gatctcagtt 30180
ccoagttcg gggagcggt cccacatcgt tgggaacctt aagggcatttt ctacatcgtt 30240
gggaggaaga gagaagagag gaaaggg cttttat cactataccttttttttac 30300
atgctgttta tctctctaatg gaaagatgtgg cctagatgtg gactacatcagtt 30360
tctctttggag gttgagttgcc ccctagtgaat cagcagcgtg ttcagttgtgtaa ccac 30420
atatttgctg acacacgag aacacaccata ctaaagtcttt tgcctaaaga caatggcccc 30480
atatattatc tatttcacta tcaactgcca ctctggygtt ggaagatgtc ttcgctccta 30540
gctctctag gattcggtt tagatcagagt tacaagccag cccacacact ctaagtatttt tttttattt 30600
ctatttagaga tgggtgcttc ccaattttggtc ctacttggtc tggacattag cgggaagaggt gatgggtggc 30660
agcaccacc gctctgggtctt ccagttgtgct ctcagggcagt gggagcttgg gatgggttaac 30720
cacctttggctt ataggctcgc tggctcactg atctatgtat atcctttattct gataggcacc 30780
actgtttcct cactgttacag ttttttctttt ttcctcttac gctatctgat aagtttctcct 30840
actgtttcct ttttttctttttt ttaccttttgg cttggttaatt atgatattc cttctttttct 30900
ataaatatta ggaatgaact ctattagcacc gcacagcgct cccctttttctt gccccttggtg 30960
ccagggccaa gagctttctgc cttgagatttt cagtttctca aacccctgatc aaccccttgctt 31020
gtcctcatag gtttcagggg cagctgtttga gttgtgcttt tggagcttc ttcctttcctt 31080
ctctctcatg atgagttcacc aaggtgcatc cccctctcgc cggctttagtt ctctctcagcg 31140
ggtctgacct cttgagttg cccctttttctt tgcctttttttt tcctctttttttt ttcctttttttt 31200
ttggtctgctt cctttttttctt gatgggtctt gcctctttttttt ttaccttttctttt ccacagc 31260
tgctctctcttg gtttttttctt gttttttttttt tccatctttttttttt ttcctttttttttt 31320
actgtttctc ctctctctct ctcctttttttt cttggttaatt atgatattc cttctttttct 31380
tagatattttt atctttttttttt atactttttttttt atctttttttttt atctttttttttt 31440
ccagggccaa gagctttctgc cttgagatttt cagtttctca aacccctgatc aaccccttgctt 31500
accctatatat cctattatat gttttttttttt ttcctttttttt ttcctttttttt ttcctttttttt 31560
ttggtctgctt cctttttttctt gatgggtctt ctctctctct ctctctctctct ctctctctctc 31620
cctttttttt gatgggtctt ctctctctctct ctctctctctct ctctctctctct ctctctctctc 31680
tagatattttt atctttttttttt atactttttttttt atctttttttttt atctttttttttt 31740
tgctctctctct ctcctttttttt ttcctttttttt ttcctttttttt ttcctttttttt ttcctttttttt 31800
tgctctctctct ctcctttttttt ttcctttttttt ttcctttttttt ttcctttttttt ttcctttttttt 31860
atattattttt atactttttttttt atactttttttttt atactttttttttt atactttttttttt 31920
tgctctctctct ctcctttttttt ttcctttttttt ttcctttttttt ttcctttttttt ttcctttttttt 31980
atattattttt atactttttttttt atactttttttttt atactttttttttt atactttttttttt 32040
atattattttt atactttttttttt atactttttttttt atactttttttttt atactttttttttt 32100
atattattttt atactttttttttt atactttttttttt atactttttttttt atactttttttttt 32160
atattattttt atactttttttttt atactttttttttt atactttttttttt atactttttttttt 32220
atattattttt atactttttttttt atactttttttttt atactttttttttt atactttttttttt 32280
atattattttt atactttttttttt atactttttttttt atactttttttttt atactttttttttt 32340
atattattttt atactttttttttt atactttttttttt atactttttttttt atactttttttttt 32400
taggctggtgcc ttatgttaagg agaccaagag ttgtctttce atagcaatgg gaataaaaaa
34860
tacgtgacca taacctggcagg agaggggtta aatgtagagc tgggaacctg tgaagttct
34920
tgactttttt tattttttaag gogaatagg aagttagagta cagctgagac cgaatgact
34980
agtgtgatgaa cgcgggtcagc taagggtgta atgtaaagtt tttctctatga cgaagagcag
35040
agtgtgcacgg tgtgtgcaata gaagaaccttg aatggcgcaag tgaagacgct atagctcct
35100
ggcaagggcact gcgcagtaaag tgcctagca aaccacaaac caattgacga taagggccagc
35160
tggggggccgc ggtggccgggg agggccccct tactgacccaa ctctctcttaa cccgggctgta
35220
gaggggcaacct taggtgtgtag cgtgtgagggg tcggacagcg ccacccctctct
35280
tgagaattg aatcctatga gatatatatttt cctgtaagaa taaatctgtat tttttataatcc
35340
gaaataactg gggggaaaaaa gatgtatatc tcaggtatct gcgtgagacga ggtagaagag
35400
agagagagag tcggagctagc ctaggaatga atgttgcgga aaattggaggg ccaatattttt
35460
acaggtgctgg acacaacagcg gcctcctctct gcagatctccttt tactgtgcct
35520
atgccgatc agtgccgccca cactgctgta cccggtgaccct gcgggccctcg ctacaggtcc
35580
ccttcctgac ctcgaggtcc cttgggtcctct ggtggccgcttc cagcaagtttctcc
35640
tgaagagagcc gcggactcgg tccttgcctg ttcgaatccaa gggcgcacgg ccagagcacag
35660
tgatatccgg acgggacttc ttctttctctg gcaaaatggt gagaagttgg ttttttattaat
35700
agagataaaag aaatctcata ccaaatattttt aagtatcatgg aagtocatccct
35760
gcctccctcc tctttttatct ttcgaatctaa atccagagac ggtccttatta aaaaagccccc
35820
aattataaaa gactgotgaa caagctgtgc gcggccggtggt cttgctcttt gacagccggag
35880
tataaaaaa gacagttaag ggaattatcgc cagctctcctcgttctcctgc gttatgttccc
35940
tgcctcaccc tccatttttt actcacctgc gctacagctga tttcattagc tttttttttt
36000
agtgtcctcc cacaagctca aagttggggc ggctttttctt gaaacgcctgg atgttctctc
36060
cattggcaca ttattatcctca aaggttagctt acacagcggc tactgcagctg tatttttctt
36120
atgtgttctt tataatatgg cttctttataag tttttttatg aaatttgcaagttttttcccc
36180
aagtttttctttttttttt tttttttgtt aatgtttattc tttttatttc cagctgagac
36240
agagagagag tcggagctagc ctaggaatga atgttgcgga aaattggaggg ccaatattttt
36300
ittttattaa gacctcttcc tcaatagttt ggcagttgct tataaatagtt atttttttttt
36360
tggttcggcc tttcataaat tcttttttctt ccatcttttt ctttatttt attttatatatttttt
36420
tatataacaa gtcgtaaaaa ttgatgattt ctaagtttttt cttgcttttg gaagaaaaagt
36480
agtgtggtt cagcagctaa ttcccttttt ttgtaaatgg gataaatctt gaagttcttt
36540
tttctctctac ttcgattcctt atctacgattctg aagctttgat cgcgttattt ctgtacttctt
36600
gcatatatt attatgtccaa gtagttccaa tacaacaaac ataattaatgttttttttct
36660
aagaaatcaco aggtagttgg acctcagatat gtcgtgtcgg aacaaagaat gttatatgtgag
36720
tatataaatgt tataatgttta taagagagac cagcagctatc ataataacaa atatagtgtat
36780
agtttttttt gatgttgat ttttttttta ttttcttttt ttttttttttt ctgtttttttttt
36840
ctgaggtttgg cgtttttgcct tttttttggt ttggcagtttt tctcttttct
36900
tctccttct gctttttttttt tttttttttt ttttttttttt ttttttttttt ttttttttttt
36960
tatattttttt ttctgagagc ggttcttttctg ttttttttttt gttgattttg accggacctaa
37020
tcaatgctca cttggggcttt taatcttttt gcgggattct atcgagcttct ctcagcttctt
37080
tgagagtcag gcgggctcag atggtagctg gcgggcttctt ctcagcttctt 37140
acagtcgc gaacttcga ctatggatg gattacatgt gttgccatg tgggcagcgt ggctcgaac ctctgggtgc acagagttgc 37200
caccaacgct ccctcagata tgtctgagtt acagggatta tgtatgtcc tgtgctcatt 37260
atatatatata attgcaacact gtaattgaca tctagctgtat agttttagaa 37320
ttgtatccca tgtctgagct tgttggtcct aacgctttat ccagctttct tgttagaggtg 37380
agggcagcag attgctgatc gtcagagctgt cgagaagctga ctgacgaactc tgtggttaacc 37440
cctgtcctt ctataataaca aaaaaactac gttgtgagtt gcacagggcct tgtatctcag 37500
ctacttggga ggtatagcc aagagagctg tgtggctctg gaggagagg tgtgggctag 37560
cagagacggt accacgtcgc tgcagctgg gtagggagtt gagagctcog ttaaaaaa 37620
aagaaaaa gaagaaaaa aataagcgcg cattgtgaaat gtagatattta 37680
atatattct tatttcttat tttattact tattttattata aatgagatcct 37740
cccaagttgt tttataaccta agtctagacaa acttggtgct tctcttgctg gagaagaggg 37800
tgagaggttc ctgcagagagct gattttagtg ccacagccag aagagagcaca ctaatagcg 37860
aaaaaggggct atgggcagat aagtctttaaa cttcttcggta actttcttggc tttcattctg 37920
cttgtccgg atagagatttta aaaaaattttt ggaaggagtttc 38040
agctgtttaa ctctctgttg tggctaccca gtataagcgc aacatattggc tattggaaga 38120
cattagata tattttgctt ctctctgttg acagagttgc aagagttttt tattttgctt 38180
ccttggattac ccaatcagatg tggctccagg acctctcagc ctcaccccac ccctctcctc 38220
tatcaccatg ccctccactgc tcaagggcag tccacacgtg aacctgcaca 38280
tagagacact cagcttgctg tacaaggggg tttctatcct ccagagactg atcttctcctg 38340
tctctctcttt tgtgaaaaaa atctcctgtat aggtgagctc tgtggccctca gacctgtctg 38400
gtctcaggt ggcgctctct ttttttacct gtaaacacata aatattttta cgggaggtcc 38460
aatctcata aagagatgtc atcatataaat ccaagagctct tattagagaaaa gattattaca 38520
acaatagacaa aacagttggt cagaaaaca ctaatctaca agtctttggac tggaccccc 38580
tacttcaata taattcata tttgctgctgt cctcaggatt tttgatgacag aacgtcgagct 38640
acaggtagaa atctggggct atggaagatt agtattaata taaagatgtc aaccatatcg 38700
tcagcttggct ccatactctcg tgagagcatt atattttgctt aaaggttaaa tctgtgaaaat 38760
agctttgagg aacaggttag tggcactttc tgtggaggtg aatttttaaa ttagccttctt 38820
cccttctggtg cccttctggtg cttcagttggt ctttcttcttgcttt 38880
atgctcctgt acctgctcttg tggaggaagct gtagggtgctct tttcagttctt 38940
acagtctgca tggctgcttg gggaggaacgt cttacctcc gccagctttc tcttgtcttc 39000
ccagaaagtc ccctctgggga ctatgggtgtg acacattttt aacattttaa attatcagttg 39060
tttgtacctt cggagtttt cgtaggaacaa ctaaaatttt gctccatact ccctcttggtt 39120
gcgcctgggc tggctccattt tggccggttt ccctctcctc aggagagttg cttcttttcttt 39180
atatatatat tagcctgtct ccctctctct cttccttttgcc ccctgatcap cccctcctttttt 39240
gcaggctttg ggtccctctct tttctctctt cttccggaatt cccagatct 39300
ccctaagaata tgaatagaat actggttacac ccctctctct cttataaaa 39360
atttagattc cctctctttaa gagaattaaat atattttatat ggataatctt 39420
ggggcagcag agcttttttc aagtattcct gaggagagag gctctttttct ccctctcagccc 39480
agctgcggct cccctctggat cctttttaaa tgtttaagtt tgttttgctag 39540
-continued

cgggctcaca gttattggcc cgcttgggcc tcacaaatgt tggggattac aggtgtgac 41940
caccacacct ggcccaagaga tagcctttta aataccaaaaa caataatcag tgaagtatgt 42000
gaatagagga tctctactct atttgtagctg aaaaaagtag aaaaaatgataa gagaataata 42060
gaaagttcctg ttacaacaac tttcaccct actgagaaat acagggagaga atgtaagcaca 42120
ttggaggttgt tcgtcttggc gtgagaagac tcacatataat catctcaacact gtctttttttt 42180
taatgctatct ctccttaaag cccttgaagg aagttttttttt tttttttttg gacaaatctaa 42240
ggtgtcctctc aagatttaaa gcttatatttt atataaaaaac agaataatttt atgtagaca 42300
gttgctgaata acataagga gccctggagca aacaggtata aatactctgaat atgcatatac 42360
gacatcgttttt atgatatata ctggtgttttt gcaaatatgt aaggtgtcat ttcaaatcata 42420
gaagagaaaa gttgattact tagttaatgg aagtgagaaa aacccgatct ttcctgagaaa 42480
taaggtgaatt tcatattttta cacttttataa ccaatcatc tcagatgagg tcctattttt 42540
aaracagagca agttaaagga agttaagaaag aatatgaatg gtttctctca tgaatagggg 42600
aagaatatct ttgagcataat atgaacacca ggaagaaag aaaaagaaat atgataattc 42660
ggacatataa atataaaatat aatcataaat gcaaaaaaag accataagga aagtcggaaaat 42720
tcaaaacaca gtgcgggagg aagtttataaa ttctctcataat atataaagt ttatatctcc 42780
ltatataaaa cactctcaaa aacagtgac aatacacaaca aacacagaa aaaaaaaga 42840
aaraaaaaag agatgtgacac tgtgaaagaaa aaccaataata gggtctttaa agatattggg 42900
gatgctcaat tcatactata agaagatttg agaataaaata aatcataaatag tagtacatt 42960
ltaaatgaaca aatgctttagaatt ccttcgaagt ttgctcataat gttgggtgtag tgaagaagtg 43020
agaaaaataaa aacagatctaa aacacatgaat atggaggtga taatggccaa aagttatatga 43080
agagatataa tgcataatcc atcaagattatt aatatataaa tatattttttctc cccagacact 43140
cacctgcgggg ataattgacca aacatagcct tcaacagtgc acatacgacag tattattatt 43200
gcaaaaaatatt cgtctgtagc ttttattttga attttgaagaaactgtaa aacccaaatgt 43260
tcaaaatag aagtgttggct acaatttgggt ttgtaactgt tttttttttt aatataatgga 43320
aaaattttgt gcatactacttg agaaagaaag aatatagcgt gtaacgtctc cttaaaaga 43380
gtagagtgag ttcgcaagcc aattgtgataa acaaaaaaag gaaaaaaatatatatttttt 43440
ltttcatttt ttaattcata aatttttataa aacgtatattc accactttta gtaggcggg 43500
getgcaataa caaaataacaa tagtttaagt gtgtgtaaaca aaaaaatatt atcttttcaac 43560
agttgttattg cttgctgacat cccacacgg ggccagcact ggtgtgggttctt cttgagggg 43620
gagagagtat acaagactttt caagtcattt ttttttttttttt tattaaaaa tttttttttt 43680
lttttatttttt tttttttttt cttttttttttt ctttttttttttt ttttttttttttt 43740
getgcttttttttttttt tttttttttttttt cttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
ttgagggaga cacaattcag tccacagtac caagttattt gatctatgta gttggtcttg 44340
ggccccaggga agaagagcag atgtaggaga atagaatggg aaggggttatt ttcattgtt 44400
tgctttcttg accttcttcag ttcatacaac gagggtccaa caatcattt 44460
caatcttgtc ecctctgtgct actgcacca aagttctttgta gttcacttag aagatcaact 44520
gaggtcagaa ggttgctttcc gataaaacct tcctgctgac ttaactggac acaggaagggc 44590
agctagagag agaggagagag agagaaccttg ggtagatgta tcggagttga aaggttgaggt 44640
atctacaggt gcattatctg ctgagagtgg ttcttttggt taatgagcaca gcttgtgggct 44700
tgccccagcag caacaggttt cttatccatt gttcagcgatt cttccctcgg tgggacactc 44760
cacaaccttg gttgtagttg tagaatcttc aagggcataatt ccttagaatt ttaaagttc 44820
aggcatagg ataacattag tagggagagc cacatctcctcg toctattctc ctcctattat 44880
acttagggtt actgtgcagat tccataagg ttaagagctca gccccacaaac accgccctattc 44940
cctcagatgc ttggttaaag tccttgccagcc ccctaatcctttt tttttcgtaa gatggaggtt 45000
cctctctctg cccctgctcg tcgagcttgttg cacgatatcc gcccggccct cccctcatct 45060
ccggggttca agcagggttg ctgctctcagc ccctctagag tcgagctgtg ccgccagtcg 45120
ccacattgcc cagcattatt tgtctatcttt agatagacca ggggttccacca aatataccca 45180
gggtggtcct gaaactctcg cctcatctgc cggcctcccagc atggcggtg 45240
tgagcctcg gcgcagcgcct catcctctcg actctactata gattttgtg tatctagcctaa 45300
cctcatcatt ttgacttatt gccccacagca actcaagagaa atccttctattct 45360	tagtctttac gatgctttttaat aagggcaca agtgacggagcc caggggaggag gttacagag gg 45420
gttgagccag cggaggtctcc aagggcatag gttcttgctc tgtgaggttg gacggccgaag gacggctctcc 45480
cctcctcagcgtg gttggtgtca tggctcacc gcgcggaggag ctctccccaaa ccctgtgtta 45540
aggggttttt agtgggtttgct ctcatcagggc ataggtgattt aacactcgggg ccaccttgtga 45600
tccgtctcag cccctgtccct cttttttcctctgcttattg ggtggattga gttatccactc 45660
atccataaag cccttcgttc ctctggagaa cggccttatct caaagctca gccagactgct 45720
cctcagaggtt ccaatattttt gctataacac agttcagttt gaaaggggttt tgtcttagac 45780
agttaaaagc gttccacttt tccctatcacc agagggaaac ctccttggatt ttagtgcttctg 45840
ttttggttatttgac cccgggagcag cagctgctatt tatttttttta cttctcagttctg 45900	ttctagagaa gggggctgctc tctctacccag aagggaggttt ctgttacatt ctgtatagaggg 45960
agacattgacag ttgattcaaat cagagtaatgg aagaatgtatt ttaattagagc atgaagggaa 46020
ttaaaagaaagtg atcccttcatg aggcagacac gcacaaggggtt cctgggtcattt 46080
ttttggttatttg aataacctattgattt cttctggttatt gttatcctgtt 46140
agggagtcaat accttcgctgct aatgtccctt tccctctgtgct ctgaccctaat agggcaacctg 46200
cctcttcgctctt ctagcatttt aataactgttt tttgggttcc tttttatgct 46260
gctataatct tattattcgt tataaatga cagttgagggac atcagagtctt catctcacctgatcagcttc 46320
acatcctggg ttggtggttgg atgtactttg tccctcaagac ttttttttac gcacaggtctt 46380	ttgatgctct ttggatccttg cctgagctc cttattttttct attggatcagga 46440
cctggagggatg cggggttcagc ggtctctcac ttatttttaacct cttcattttat ttttatttaa 46500
gttcagttg gttctgatgct tctgatacacta gttcatttttg aatgggtacc ttagcttacata 46560
aaccagcaca caccagcacaa caccacgaca agtattttat cttgagatgcttg 46620
ccttgcctct acacacacac acacacacac acacacacac acacacacac acacacacac atgatctgag 49080
gaccttgggt tttttcttca aagacagggt cttactatgt taccaggttt gaggattgac 49140
ttcgagctg aaggagctt cctatctcag ccttacaagct atagacatgt 49200
gccacacac gcattacact gcgttttgca ttcgtttaac agagtacgtg cagacagtctg 49260
tcttacagg tctttcttcct aagagctcgt tctgtgctaa tgasattcata aaccctttac 49320
tccacttcca cccctaggaa aaccctggaa ggcagagggg aacagggggg aggatttttc 49380
ttaactttgca cgtttaagag ccaagatttca agagttcctat gatccacacac tctaatattc 49440
cattgagct gttatgcaca aatggccacag aaggggttaac ttattaaaac ttccaatatta 49500
tttctcaagt ttctttaggg ttcgaaacctc aagatctgag tcagggctca 49560
gtctcttgc ttcacagggg cagcatctgc accaggggga tgaatgtctg 49620
ccccctcaat acgtgtaagag aacccaccca aaaaaaagtt aaggtttcag gaggcttca 49680
ggctttaact cccctaacaa aaaaataact cgttttaaat gccacaccc tcataactat 49740
ccaccagacc attacgcttc acacacacag cttgtggagga gatccttccaa accatagcgc 49800
caggcaagga ggcgtgttgtc catgggcttc taaacctcaac tttagttggc cttacctctg 49860
ttcatatcgc ttttctgat cc aaggtttatc tgcctgcttc ggggtatgaa 49920
aaccacaaaa ataagcttcca tgcctctcaag gggctgcaac agggcgagat ccattgatcg 49980
cttcagacgc ttcgtgtcag ctcctcttgctg gcaacagaga tggagagcc acagccaaac 50040
gtctcgggga cagagcagat ttcttgagaa accgcaacctg ctaacctagaa acagggagg 50100
atgggtatt aagttgtaaa aaggggatag cagcttgcctg ttcacgggaa caataattca 50160
tcttataatg ggctacatgg ggtgtcagt gctttgtccttg tgtgacagctg gggacagaa 50220
aactttgggt agggggcctt ttcacccctt tttgtagtaat attgtgtgctg aectgtcaga 50280
ccagctcagt cagggggagtc ctacacccctgt ggcgtgaaag gaaatgaaaga cccacacac 50340
gaatacagaca ggggtcagat ggaaacacag cgggtctcaca gcttctcaag cttgaaaac 50400
agacaacaga tttaacacaca tatttattaa cagcccaacca gctctgacag cttgttctct 50460
agatatattaa ttactaataa atgctccctta tgtgaacaaaa agggattgag ccaattaaaa 50520
taagagttg gactgtttga tggctgacgg agccgtctgt tacaagcagct acgcttcagt 50580
cctgttgttg aggttgctga ataatctttta atgtctgcgg ctttgggtcgc caggtgtctc 50640
ttgtgcctct tccacacacc ccaacactcc tccgggtgag tttttagggcc atattaaga 50700
tgttacagtt tgtcgaatat tttgttttag gcagagtgcg gggccagttc atggccagat 50760
tttgggggc cctgactaaca ctagctcctcttt cttcctgtct ctaacacaaca taacatatca 50820
gggcagcttgg ctcacgctcg ccagagctcgt gatctcacaag tggacaactt 50880
acacacacac acacaacaacact gtaaagacac aattacatcgt gaaactcagc acgctcacaag 50940
ttgtttttata aatttgcaac tcttcgaaac acacacagtt cttgacctaa ggcttacgctg 51000
gccctgctg ctcttatatt ttatatatatt ctcttaccac aagcctgtcgc ctcttctgtc 51060
gttggctcata cccttctgag tctgggtgcc actatactgc catgcttcac gataattaga 51120
acetttgcttc ttactttttat cttatcactc atctgagccgt tttgtcttca gaaggtgcag 51180
atgagttgctg catgtcgccag aagatctgagcg tgtccatttg agcattacac ccccttaggg 51240
gacacacact taagtcctctc ataggaattc tttgcaacac cctctctctctc tttgcaacatg 51300
caaatcctt aaccaaatat tttaacctttt cttactcagag ttaatcattt cttacataaaa 51360
ctcaaatgta gotataatat ttacctgttg atctgggggc tgtattctaa caggggaactg 56160
ccagcctctt aaacacccat ctggccctgc ttctgatatt ctccgcctga tagaactaag 56220
agttggggct caagggcttg tctgaccaag cactggggcg ctagaatttt ggccggctgc 56280
cctggaaaa gagaactgct gaggggtttt ttcctccaaa tataaaggg ggggtgcaga 56340
agggtagggg tgaacactc tccotctttgg cacctttagct ttagctggta aataaactatg 56400
cctgtaacc ttctctctctt tctggcta atctttacat tttctccccc ttcatcctcg 56460
ttgtaaaaaag ttccaaagtt gaaaccaac caaccccaac atgctccatt gttaactcgtg 56520
cgytcttcgg cttctcctct taccctaccac ggtggattgt gttaagactc tgggttacc 56580
tccagcttag tttccatcctt ccaagtctct tctggtgccc tttgaaactag atttggaccc 56640
ccagcagttt gggcacatgg tggagcagc cggcgcttgag agacaccaac acatggtgca 56700
tagaggaatt aaagcaaaac acacctttat gatggcgtg aagtgggaat caggggtcc 56760
acagcctcct aagctgagac ccgaagacag cagatattaac acatatttac taacacaaaa 56820
ccagtcttta ccgatttttc ttagaaatcc aataattaca acagatctcc atctttggaaa 56880
cgaagggtac ggtgtaattt aagagttgg ttggtcctag taaccgcagc agggactgct 56940
ccttaaggcct cagctggcttc atgctatttt ttgggtctta aagatgaacct taacccgattt 57000
tcgccctcgg tggcggcagcc tgtctcttgc cttcaattcc gtaaactcac acacctccag 57060
tggtgggcctt aggccctaag tggagctgtg cagtggtttc ttttacggcc 57120
gttttgggctc cagttttacc ggaggccctc tccaaattgt ccattttctg 57180
atgtaagttc ttcctagatt cgccagacat cggagtctgc tttttccttg 57240
ccotatctgg tctttatgctc aagagctatt ttgtaaagga ttgaaagttc cagacaaata 57300
atgggctcag ccctataatttttttctt gtttatagtt cgggttcct cccgattta 57360
tgaagaggg gattagcact cctotgtgaa tttgaattaa attgtagaat acacccagtt 57420
ctcaagcggg aacaacagaa gcacacatcct atcaactgct cttcctattttt tattgggccc 57480
tctggccctgg cttggaggcc gtggagcagc cggaggtttt cttcgggtgta ttccagttcg 57540
atcataaccc gcacaagcgct cttccactcc cggccttctt aaggaatata cacgaagcagc 57600
agagcagccg aacctttttct gttggttttt cctctgtggag gtggttttt tttgcattttt 57660
tggcagagtgt tctcgtgggc cagcagctag cttctctgga ggtggtttgt gtacttctcc 57720
agotacacca cagacagtga aacacacaca ccaaacacaa cacaagagct cacaaaccc 57780
tttcacacctt tttctcttgag cggcagaaac catctttctg cttactttttg cttattttta 57840
catcaacatgy ttttaccatt caagctttct aggcsnnattttt cttttcccattt cttttttcac 57900
tgaaagagtc tattagtaagg tttcaggaag gttttgctaat ttctctttttc 57960
tttttttttt ttttagttag gggaaaaca ttccttccac tggacacctt ctttaacctc 58020
tggttttttt tggagacca cttccttttaa actgctccag ttgagcggct cagggagtgt 58080
gcagcgtaaa tactctcagg aacagccttg ctgctccccga ttgagggga acaacagcata 58140	tggagggag ggtggagggag ggggaagga ggacgaagga aagacagaga 58200
ctgaaaaggt ccgacacca cttctccccg gtcaggggtt tggaaacctc tctcccctcc 58260
taccgtaatt tgggctgtgg caggggttgg gtttgcattg cttcctcatgg ccttttctct 58320	tgagatggag cccctttttc cttaaactat atccatcttt ctcacccgatt gttggactct 58380
agagtctcgc cagatctcttt tgcacatgctt tttcttcccc cttcactttc ttcgtccacc 58440
cccattata gccacacatttt ctcagagcac ctccttggga aaaaaacctta gtaagctccc 58500
-continued

gggatctcgg cccAACACA ccACCTTCGA TAACAGAAAA AGCCAGTCG ATGGATGGG 60900
ccggggtgtt cccggggagac aagggctgggt tcgaatctgt ctgtctgact ctagagcttg 60960
ggctgacacta aaaaagactc ccctttcttc taagagctgc atgggtatc ac 10200
gtgacagcag tgggagacat tagagatcgt aagagagtaa atgggatata gatggctag 10800
ggcttggcag aagggagctgt gcagatgtag tatttttttt ctaaaaggt tatacttttaa 11140
acccatttctt gccctgtagc gcggagagag gtccaggtac acagctgtctt attaaatgaa 11200
ggacaataat ttataaatat ccctctctta atccctcttc tagttcttga tagoaacttg 11260
ccccctcgag cccgggagag tagaaatgaa aagcttgaag cagcagagat gagaagaggt 11320
ccgagggat cagggctgc acataaatagg ctttaagata ctattttttttt 13800
ccctttcttc taagagctgc aaggggttgag atgcagctag atccagaggt 14440
ctttctgcc ctctttcttt ttttttaata taatattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
agttggaatt tattttggtg tcattcagaga ccaacacata caagttgcact taactgtcact
ctgaaagatg aatagagatat atccagtgag gtattggaggg atttccaggc tgacaggtcg
aacatgacaag agggc tgtgatggct cggacacgaa tcattggaac tgaaggagcc
ctgacctttc agagcagaggg aagttgatagg gaagctgact ggaggtgata ccctgagagg
yagcaggaat ctagctcctg gcttagctca aagctatgat gaggagatt gaacatcttgc
aagagccagg gcacctaggc gggaggtgaga cagggctgac caagccactg aacatcttag
agatctatct cattctcgag agacggccgg ccacactgtt ctagcttcat aagccgtacga gaagaggggt
"gctgtgaggag cctgtagtgg tgtgatgaga actggaacca aattcctccta caacagctggc
"tttctcattc catctatgatt tactatcctcc cctgctctcat taactatgat ttcaccttgc
"agaggacctt tctccagaaaa gatatattgct acaaatatact ctttggtctct ttctggagat
"tccacaaat caattaatga attacaattt cttgagtggaa ggtgaaaca cattggtgc
"aacagttgctt cccagagaaa ttttagaacc cttgtgcttc aatcctcacc aacagttgctt cttggtgc
"ctccaatctc caaatcctatt ccaagctcact tcaacaccc cttgcctgaa cttggtgc
"atacctgcgt ggaacagagac tccaaacact cttctcctct cgtgtcctgc tccttgtaac
"cttatgac gcctgcgcaaat aagtagtttc cctttgctgt gcagactgt gttttatcact
"gttagtagtg aactgctcttt gcaaaaaata ttgacagtggag aagaaacctg cattagctggc
"gtcattggtg tttagctcwt ccaagctaatt ccctttggtt aacggtaaaa caagagagct
"aacagcagct totaaataca gocaaaaacta attttcctct cttcgcaagaa tgaagaagct
"ctttgtaga cctcctgaag tccacagatt tttgtattag ggaaggggcct gaacctgctg
"aataagttag cgtattttaa tgtataacag caatgcttcc catgctgctg ttcttgtaac
"ttcactagc ctcagaggtc agtgcgaggg ccaagatct tgaaccctcc
"caaatcctatgtgcttgaa gcacatacata tcattgtaac aacactaaat gatgaattg
"gatatttttta agacttttttt attttggccttta ccactgact caccctggacc cccctgacttc
"taccaggaag caacagaaaa acacatcctatt cccgcttgaa aacataacact agcaagctgaa
"gacactctcg acacatctaatt tttcaacttc acaaatccta gacagacaaa ttttagttagc
"ccccgctgc caaacgtgcct ctaaaaaccc cctgcctgttag tttggtggag aaggtgatatt
"gagaattgtt tttatctatt ttgtccttttt gttgataaatt aaatatatttt ttaaatatttt
"tttttaaaat cctctgtgat caacctgaggt cagggatttttg agacccagctt ggccaanctg
"gtgaaacccc atctctattaa aaaaaactaa attaggtggg tgtgtgtggt tgtgctgtga
"aacaccattatctccagagga aaccactttg tacctgggag ccaacaggttg
"ctgtgcacag atccacacttt cattccctcaacctcgccatcctgagctaa acagcagcagatg
"gaaacaaaa caaacaacaa caaacaaca aaccaccat attgtccttc gcacacacag atccaaatgcctg
"aattagtttc cttcctctag acacataatt tttagctaac aagtttaact gataaatgtat
"actagtctca aaccagattta tagaggcttcc acaataccttt tgaacacata ccaagtcagat
"actcttacaa atccacactc aagagatcct caccatttt tttagcttcc aacagctcagat
"gtatataattt tttgataaat aacaccatttt atttgttttc gttgtcataa taacatgttga
"attttaatag attatatag tcataacagac acaatttttta taattaatttttaa taacattgagc
taataagacta aatcagcccc ccacacagtttccctggg aatcagctctttc acacagagttt
"atcactgtca taagaaaaac cgggttctgt tttttcctca taactctattt attggaaac
"tgagattaa tttcctttcaacctgacac caaataacatttt ttttgaagaagcttaatcag
""
agacacata ggtaaacctg ctgcacatgt gcatagcattt gttaactgtc atggctgtgg $8040$
tggagaagc cgaagtggaa caacccaggg ttacgctcat ggccactcttg gttcaggcgtg $8100$
gtttttagttt gttcagatcg tattactagc aggtctttat tagcctttaca cttgccagtc $8160$
ttggcgccac cttcagactc atctggtaaa ttgaatgagc taaccagctg ggaaatgcagc $8220$
tcagccgctc ctcgatcttt ttggccgatt cttcattcag aatggacttg ttctggttca $8290$
aacaccttcg acaatattctg ttatcttctg tagagacttg agtcctcttt gccctgagg $8340$
ccgttcctca attcccaggg tgtcaggtgtt cctgctcttt ggtgccttgg aagtgtctga $8400$
ttacagctgt gagccacata gcccagctaa tttaacttgt tagacacactg tagaattgat $8460$
gtctctagtt cattagcttt tttgagcact gggcaagcag aacgccataaa tgctgctcag $8520$
tttgggtcttt atttctgctgg attcgacccg ggcgaactgc tctcattcag aaggggacagc $8580$
agccatcttg gcgcaagcag cacgctgccc cgacctttga aagcaacttg agggaggacc $8640$
actgcagccc aagagtttga gtcgacgttg ggcacatag ctagcccccg tttgaaaattc $8700$
tcgcacaaga aagacagaaacc aagggggcgg ggaacccag tagtcagaaag ttgccaacag $8760$
cggcgggtgg tggfagtggca cgctgttatc ccccgctcttg gmagagcttg aagttgggcc $8820$
atccagagtt cagaggttgg gggacacgct cagccgctgt attaaattccc gtctcctacta $8880$
aagatacaca aaaaaattcag tggggtgtggt ggtattgagc ttgactccaa gctacttggg $8940$
agagctgaggc aagagactgt ctcgaacgca ggagcaggag ttcgctgag aagcaagattg $9000$
cgcattgca ttcagacgttt ggtgacaggg ttcggctcctt gtttccctaa aagagccataa $9060$
axaaaaaaaa aaaaaaagaatta aaaaaagaattttggcctgggctc aacacacactc $9120$
ggggggcaaga gtagatttctt cggagagagaa ctggtctgactt acltttttagg gacacagggc $9180$
cacagcctc ggagacaggg accttgagaa aatgagcagg ggtatatttg gaaacaggcc $9240$
cctcagatgg tgtcagacaca ttcaggtgct gttatcagtg aggggtgcaaa ttagaccaac $9300$
gcgcacagtct gatgctagttt ggacacaggg tgtgggctcgt gtcgatagtg ggagaaggtg $9360$
tggagaagct gttgagggcg ctcgactgag aacgacagcg tggatattccc ccagcccagc $9420$
ttcagctggg ctcctcctgg gtagaagtcg ggcggctgtag ttcggagcgg ttcgttggcc $9480$
caaggggagg acgggtgcttg gcacagcccta tggagacgg tcacagacgct ttacagctat $9540$
aacatagtgca gttgccactcc tcctcccttc ccacacccac caacctccct tccacactttg $9600$
tgtaggatt ttaggtcctgg atgcttccgg tccttcctta gtaacaacttta taatcctgtg $9660$
catacttggc aattttataa tagactcttc gtttaccctt tggccatcctt ttaaaactcaat $9720$
tgcgggttgtt gttgactttg ctctttgtaa tttctgtgaga aacatgtctg agcttattgttgc $9780$
gggtcggggg accctttgaa tagaaaggg aacgctggat gagggttggg cttccgggcc $9840$
ccttcggacaa cttgagccac gaaactttac ttcagggctca gggacacttaa gtcaacccaa $9900$
tggacaaacta caacccctta attgacagta atttctctctaat ctcagcccc $9960$
tggcagttc cctcccccttc tttccacacc cccctccttc aagtccgacc gatatttacctc $10020$
cctctagtgt cttcagatgt ctttttttaa tcttcccaag gcaagttgtt tgtgggtttgg $10080$
cctctctctta ggaagctgcc aaggtggcct ggtgggtgta ggcagccacctg $10140$
cagctctttct cccttcctgttgg tttggaacact tgcggcagttt aagcttgctactt $10200$
tattattccc tttgtgctga caattttcctctt cttgccacactt attcaactactttattttg $10260$
gcaccatcatta tatttctaaa tggatgtgctt ctgagacagg gtaatgatata $10320$
-continued

cagaatgcct atattaataa tgaattcacat atatatataa ttagtggaa gaaacaaaaa 70380
aacaaggtagc atgtggaagg gctgaaaaat gatgaggact gataattttag gaggagcaga 70440
gaaacaatagc cgtgactgaag tgaagaaatgt agccagttgct ctctcagaga ccagaattt 70500
cagagcaaga aatacgacgg caaaatctca gagactaaggc aagtggtgggt tgtactgtgt 70560
tgctgtctttt gaggatgacg aaggtgacca gttgtggcttg aagtgaagcg gatgtggaa 70620
aagagtgtagc ggaggtggttg caagagaagca gcgcagagtct tggcggacct ttggacat 70680
ctttgtagctt gatagggagc acgttagcag aaggaggtgg acctattctt ctatattcct 70740
gctatgtgtgg tgtaggagggg aagcaagggag agcagctgagg aagacagagag tagtacagaa 70800
tggggtctttt tataaaacga aagacaacca gaaattttta ataggaaaagg ataaggtgag 70860
tgaagcaaaa agggagtagc taaagacac caagaatgttt ggcttggcaca aactcaagggg 70920
agggatgtca tttacagaga tggaggaatgc tgtgaaggaa caggaagttgag gtggcagcag 70980
tgttgaatgtgg ctggtagggat gtggtggagca gatggtgagga agtggtggagt gtttacagt 71040
ctataggtgt ggagtggattgt cagtggtgac acgtgaacct ggggatattc aatgatgtgt 71100
tggaaagcccc ttaagatgtcc ctgtaaagaa caagattggta agttggttttag ggtgttcaca 71160
acccactaccc tctccacattatta taatattaaga cgcgcccttag acaatatactg aatgtagat 71220
catcagacctg tttcaataata accttttatct aagaacagac tagtgggtctca gttttggtgc 71280
catagatattg aatgctgcacc atctttgaaga ctatgcgttt acctaggttag atcgtaaaaa 71340
gaaagagaga gaaatcagag gctgggggca cccacacttt aaaaaatggg aagagagagag 71400
ggacactaga gaactgagata taacaagcag caaggggacca ggcacatatca gaggttggyg 71460	tgccatggga gattggaagat gtttcaaggaa ggacaaacca atattattgt gccaaaacc 71520
gtggtagctg aatactgtgg ggcaaaaaaa aatgagttggtctgcttaga taccctcagtt 71580
ggttatatttt atttttcattt gtttttgagct attttttcata atttttttattg ttagaacaac 71640
tgacaaatag tgaatattgc tttgcattgata aaaaacttta atgtgctcagta caatgcaatt 71700
aatgcacattgg tggagatgct gtagaagcct aaaaaaccag ttttgcatt 71760
atatcatcgt gattccacac ccaagaagata tttttttgaaga tataagggtaattgagcagaa 71820	ttttttaatt gttcacaatc atacagtgt ggattcaca aactagatgtt ttagccacac 71880	tttacacccg cgttgtaggtt agaccttttt gttcaggggc tacaaccctg acacagcatgt 71940
atatargagc aatgggtgttg gcaagttggaa caagagtgttg aatatttttg cattaaaaaca 72000	atatataaatatatcagacg aacagagttg aacagataggtt aatagtggtt aatctgacct 72060	tatgtaggtgc acttactgtg aatggagctt acagctgttg aagctgtcgctttt 72120
gtggtagcga cattgagattgt gaagttggcttg gctggttttac gataatgtttt 72180
ataaacactg tacacttcgg ctacttttaaaatatatattac gttttttttct ttccttattt 72240
ataaatcaac cttctagatcg tatacattctt ctttttttttat tttttttttttt 72300
cttgtgact ttgttctaatatatcttgcc ttaatacaca aacaaagtgt acagctgtgt 72360
aaaaatattttt gttctttttata ttctctttctt attaactattt ttaatataaaaa aaaaaaattttttttttttttttttttt 72420	tttactttttt aacctttttttt gtaataaagttt aagagagaaa cacattatat agcactttggcc 72480
acacaggaaga aggtacataa atatctgtgg ccctacccct acactcttgct ccaacctggaa 72540
ttgcttcaag ggcaataataa cttctgggggct tgtctgtttgctatgacaacatcttgacattt 72600
tggagagcgtt ctctagggag cttctttttacag tgggtttctaat tttttttttttttt 72660
ttgagaagcac gtttttttttc tggggtttcag cggagttttgc aagagtttgct tttggctcaca 72720
taacctctcc tttggtgttt tgaagaacta tggctctcttg aactgtgtgct atggcccccct 77520
ccttcctccc taggcaccgc cttgacaca actactctct eactacccct caccttcctcc 77580
coaacctgta tcctcctag aacctcacc tgaagaacct aggcttttctt aagcttctttt 77640
agtggtcagg aacactctca atttataaac cctctccaag tgttttatatta ctttaagggt 77700
tagtctctcc aaggtatatc atgttgcgga aagaaagttc tagatataggt tgtgctatac 77760
aaaaataagt caaatctttaa tgtgttttatt tgtgctcttg tggacaatag ggagagggag 77820
catcttggga aattgacctt cctctcttctt aatatattat caagtaaag tttgctgcttg 77880
ttactctctct gcttttgaagac tgtetacagg gatgctctct tgtctttttg gcaagtttaa 77940
atgtctcaaca cagcctcctga agatctctcgcc gctttcggga gcatcagac caactcttgta 78000
tgctctctgc tttgaagtct ctgcctcttgcc ccttaaactt tttgaccttct tttgagaaaaac 78060
gatggagttc ctctgtgattt aaggttgagg tggagaagtt cagagagggt gcattaagcta 78120
tggaaagtcttg aagcaaaaaa ttaaagagag aatcgtctcc tttcctctag gaactcaagg 78180
atttctgccc ttatatgcttt gattacccg ccacacgtgg ggccagtgac tgccagggctg 78240
acctagagag gcacactctca ctgctggatat agtgcggaaac ttgctttttt cttgcacgcg 78300
tggtgcaccc caattgacct ttcctctctctta aatgtcttcctt aacctctaca gttccagaag 78360
aaggtctcacc cttggtatgg ggagcaaccc cggaaaccccc cgcctactcc aataagttggc 78420
tgggaatgcc atgtttccgct tatttttatg cagatctccg gaggccgaag cggccggattc 78480
aagaagctcg gaatcagaaa tctctctcctcc ttaaagcttg aaacccctgg tcataaaaaa 78540
aataccacat tcatcagctctgctt gatgtgccaag gcatcctccct gcaactctggg 78600
aagagataag tttgcaacgg ggagattggag attcagttgag gcagagatcg caccactgca 78660
cctcagctcttg ggcgcagagag gagacactcct tgttcaaaaa aaaaaaaaaaaa aaaaaaaa 78720
aacaacaaaaa aaaatcactcataa cgggtttcctcc cccggcttcct tgaatctctca gtaggaggt 78780
cctctaatatt aatgaaattt tatttttggta tagaagaggt gttctactac 78840
gttgocagaag cttgcttaca actctggtcg ctcagaacct ctcactgcct gcgcctcaca 78900
agaactgtcgg tattactgct gagacccacac ggcggcttta ttttacttttttattttt 78960
gagggaggttc caacctgggt gtaggtcttg gtagttcatg gcacgcatctt gcgcctccgc 79020
aaoctcgcctc cccaggtttgt aagagtctct cttgct-catg cctcgcagct aagtcgaggt 79080
acagggactg cgcacagagc cccgggttaat tttgatatatttt aagttgcagct gggttttccc 79140
atgtgggtga ggcactgttg gaatcctgca cctaggtttga ctcggccggcg gcggccctcc 79200
aagaagttcg ggatctctgg gttgacctcg ctcgggctcc cttctgcaga tttttttcaca 79260
aagttctcgg gtagatttctt gggatcattaatttttaa gtagttgcttg cgttggctca 79320
tggctgaagc tatataattc tgcggctcca cggagccggag tagggatgttgc atttccaccc 79380
ctgggtcttt cttcagctgaca atttttttca gggtctgggg cttccccgca gctgtgcgggc 79440
cccgacagaca ttgctggcagag gccccggttg cggccgggcttg ggtggcggctg aagccgcgc 79500
ctggggggag ggaggcagcg cgcggcggcg cggcgccgacg cgcggccggc cgggagaggg 79560
cggaggcgag cccggggggc gggagatgccc ccgccccggc cgcggagcag cgcggagcag 79620
gaggagacag atacacggtt cgcctggtgc cgcctggtgc ggccggcaggg cggagagcag 79680
agccggggggag gcggaggttcgg cgtgctggcg cgtgctggcg cggagcagctgt cccggccagc 79740
ctggcggccc cccgagagcg cgcacccgtgg ggcggccgca cttgaggggg aaggggtggc 79800
-continued

caccataact tacagcctga tgagttgctc tacagtgac acacotgtgt taccccaacc 84600
caaatctact cctctctatgg ctttcctccaa ggcotactcc gtcgctgtcc cccagacgta 84660
tctgcggct catacactac atacactaagtt tctactaatg ggtggtaaag gagcaacctcc 84720
ctgacttgctct tttggctcact acatgttgct gtagacccac ccaatttattg tggctacaag 84780
gttgctgtata ccttcaccgtt acatagttact tatggaacac aatatttcc gcctgcatca 84840
acctgtgctca cagcagccc tgtggtgcgtc caggttgccgc ttaatgtgcgt gcctctctcc 84900
geaatgtgtc tttgggtctag atgtatatttt cggcctaccag cagagatact ctgctagaca 84960
tatagtgctag tttgttattgt cttgctgaca acatcagcag atatatattt acacagctag 85020
atgctagcagc ctcgctaccaca tttgctatct atggctatge acctgctgccc tagaactacaa 85080
acgggtccagt cagtttaatgg tgtactaactcg tggaggtcgt gcctaaaccaaac 85140
aggtgcatct acacagtttgc aggtgtgcgt gaaaaattaa cattaatattt ataattcatc 85200
hgacgctagc tttgtatattt gttgtctgtc tagcagcagct ccttcggtttt ggcgtggtgc 85260
tactcgtcatctcgctgtctactggtatgt gtcggctacag aggcttgccga 85320
tgctgctccc tgtgctgcaag tttgctaaacct aatgctccac ctgctgacact 85380
tggaaccagtt cttctgctacat gttgatttgag atgtggctgc gaaattaacttg 85440
taccagaaaa tagtttaaccg agaataactag atacagctgg gcgtattttg catcctatcc 85500
agcatctcct aagacgctttc taataggtgttg aggggcttttc accaccacccg aagcactcctg 85560
aatcagacaa gctaaagaggg ggacagaaaaa gaccaataacct agttgagtag gctggccagg 85620
gttgtgtgtgt ttggtggtgaa ctgctgctggtc gcgacaagaag aagggagcagc caaacgagaa 85680
ggagagggga aagagagagga agaatagctc ttggtgatgt tgtgctgatcg tgggggtcacc 85740
tggtgccgaa agaaagacaa atggagcggag aagttgaagaat caaatagcag cagctggatg 85800
gactctgaata ctggagctc gaagttgaggt gcggatggttg gcaatgggaa gccagtgggga 85860
agtaagaggg ggagggctca gttctctgcaa cttcgagctttt caggggttgga gctggtccag 85920
ctagagacta tttggggaacc aaccccaaat gctcagagta aacctctagg aagtttaacac 85980
tgctgcgtgt cgctgcggcc tgtggtgtgtg tgtggatcgt gcataatcttc gggatggygg 86040
tcataaaatataataatatag gcacacactac ctactagattg atgtttctattct ctagctataa 86100
tctcataagtt gccggtgtgtc ctgctgctact caccacacct agttttagct 86160
aggaacgctat tttctagctt cttgtgcttca gttgctgtag cgggtgtgat cttgctcacc 86220
aagaacaccac ctagttatact ctggtgtgggct gtagctctag aacacgacta ctcctccattc 86280
taaacccaaactgctacctgctcgctgtgagtatg cagtgattgct tccttcacac 86340
tactctcaact aataacttattt cctgctaccgc atgcatctgat gcattgggtag tttttagtct 86400
ctgtagctctg ctttgctatccgttggtgttg ggtgagcagc gccagacgaag 86460
cttggaattgg tttttagttg gttccacagct tttgctgcttc tttgctgctgacc 86520
cagacacatat caacacgtctttttttagt gctgctgcgtgc cggtaaacc 86580
atttttaaa catactaaactgttgtaatg cggagcttta accaagtaattt gtttttaaggg 86640
ctgtagctgtcttggtggag ttttttagtttg gttcctgccct gctcggcacc 86700
tttgtagtatt tttgctggtgccgctttgcgtgc tttgctgctgacc 86760
tcatactttctggtgatag tttatatatttttttctctc tttgctggtgctgctgacc 86820
gcgagctgcc ctgctgctgctc cttttctttccgttgcagcctgctcgtgctgctgctg 86880
ttcagctctg cagctgctcagttggtgcgg ctgagcctgcttcgagctgcgcctgcgtg 86940
ttttaaatct ttctactctc tcatgaaagt acagtaagct ttctcaggg ctgatgatgc
89340
tgcataact cactcagatag aatgcagatt acctatatct cactatatct aatcctatct
89400
ctatcactc actatatct cactatatct cactatatct aatcctatct aacatgatag
g89460
actcttagt gttgctggct gctgtaagct gctgtaagct gctgtaagct acagcacaat
89520
gcagcttct tttgcatata atgattcttt ttttctctggg taataactcga gtagtaagct
89580
tgtcggagca aatgcagatg ctaatttttag tttctgagga aatcctatct aatcctatct
89640
taggaggttg ctaactctct cttctcagttc aagctagttata aatgttcctc ttttcttcga
89700
gttcactaa tttctggttct tttgctcttt ttaactttag ctaactctct cttactgagta
89760
tgtagtactc cttgctggttt aatgtccatt tttctcagttc aatgttcctc ttttcttcga
89820
tttctggttt cttgctggttt gttgctggtt ctttctggttt ctttctggttt ctttctggttt
89880
tctcactcactc ctatctggtt ctaactctct ctttctggttt ctttctggttt ctttctggttt
90000
tgttctctctct ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90060
tttttctggtt ttttgctcttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90120
ttttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90180
attgtagcct ttttagcctt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90240
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90300
cttccttcct ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90360
tgttctctctct ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90420
tctcactcactc ctatctggtt ctaactctct ctttctggttt ctttctggttt ctttctggttt
90480
ttttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90540
attgtagcct ttttagcctt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90600
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90660
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90720
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90780
attgtagcct ttttagcctt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90840
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90900
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
90960
attgtagcct ttttagcctt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91020
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91080
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91140
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91200
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91260
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91320
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91380
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91440
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91500
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91560
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91620
ttttctcctc ctttctggttt ctttctggttt ctttctggttt ctttctggttt ctttctggttt
91680
tgagggtgct aaaaactttt gtctttttgtt atagctggag ctttttatct gattctctct t94080
catctgaagg agtggatagt agtgggtct cttgtcaca caacatctca ctttttaaact 94140
cceataacce cacgctcttg cggggagct ggtggagag atgtgtaaaa catgagggga 94200
tctatccccc cctggtctct tggcgcagta aagggctcct acaagatctg atgytttttaa 94260
aaatgggagt ttctcgcaca aagcttctct tttgaccttg gcatcatacg ttaagatgtg 94320
tctgtttctt cttggtctct gcctgatttg tcggccctcc cccaggttgg gsgaatgttaa 94380
gtcaataaaa cttccttttt tttgtaaaatg caaccttgcct gccgagtgc gctagtttca 94440
tgaaattgga caatacagca gctgtgccct cacttttttg aatggtccaa cattggatag 94500
aaactctttaa agttgcattt ctttttttca tgaaggggtg ggtgtgtgat tataaaaggt 94560
atgatgtgct ggtgcttttt ctggggtgct tcaagggccc aagggctcct cttgggtctct 94620	
tgtagtggaa ttcatttttt ggtggggcct tcggcatacg tggctttttt agttgatag 94680
aggggttgtg agtgctacca cttttcttttg tagggctagg agtgctggag tcgctgggaa 94740
cctatcctct acacctgaccc cttgcttttt ttcagccagct tcagccccgg ctgggctggc 94800
ttcagttcct agcagcagct ggtagccttg agggtaagaa gagttgacat ccctctcttg 94860
cctacaggg caacctggaga ctttctctctt gggggggggt ggtggcttct gggggagtgg 94920
ggcagctag cgcggctaca gcctctcact cttggcggag aagttttttt agttttaagt 94980
actatgccc tttcagacca cttcataacc tcaaggtatt tttttatt tttttatttt ctttttggc 95040
ctggccactc ctgaggtctg tcgaagagct tcagtttggt ctacacccaa aatgctctcg 95100
gtgcagagcc tttttttcttc ggagcagcag cctgtcctgg gctagttcct gttcctttgt 95160
gcagggact tcgatgttct tgcggtgagg agtggaggtgg aagctttcctg ttcgaagagc 95220
caggggctac ccctctcttc tccggggtggt tggagagcct gtaaagactt ctgaggaagg 95280					ttttcttaat tcgctgtgaa cccgctttcc gcagggagaa cttctgtcgct acctggcaca 95340
tggataggt gggggttggg ccaatttttta atgtccctct cccaggttga cttccagact actactaacg 95400
ccectctccag agatggtcgc tgcgtctgca tttctttttat cccagaggg ccttcttggg 95460
gcttgtctcc cctaatggag aacgtcatct gcgggagtaggg gtttgctcct cggggagtct 95520
cccagggcct ctgggtctgc ccctcttttg cccagaggg aagggctctcc aagttttttg 95580
tgctgggcac gggtgatctc gtctgctcctg ctcagctgcc cggggaagtt ggtgctgagc 95640
cagagacgcttc agatggtcgct tggagctcag ttgtttttgt cccagaggg cctttcttgg 95700
cagacacact ctggggaggt tgggtctctt gcctgctcctt cccagaggg cggggaagtt 95760
tgccagcaca tggccagaggg tggagctcag tgggtctcctt cccagaggg cctgctcttt 95820
cctgtttgtta aagggggaag tggagctcag tgggtctcctt cccagaggg cctgctcttt 95880
tcgagtcgc ccaggggca cctcctgctc cttctttctt tttctctcct tccacccctttc 95940
agctgtttcc ttttggtgct cttggcagcag tttctctcct tttctctcct tccacccctttc 96000
tttataatct aatataaaa aggtgataaa cttgatccac ataagaaag attttatggta 96060
atcttcacaa ttttgtaaaa ttagaggtct cttgagccca cccaggtgtaga aacatgctg 96120
ggcagggagt gcctgcagac actccctcaggt gcagggagtc actcagatgc aggggggtttg 96180
gggaggtgtgc aagggagaaa cgggagaggt gtctgctgag tgaatgttgtt ccattttgacc 96240
ttcaggttatt ccagagcacg ctatactcct cttcctactt taattgttttt ataggatagt 96300
tgctggatgc cttctactta tgcgttttac attcactaatc ttataagcct ttgttcaggg 96360
tttgctatgc tttataattt caggggtttg aatatttttg ggtatoccac aatttaatatg 96420
ttcatactg ttgtgtcttg ggccataattg tcctattgct gcgtattaattt acttcgaagaa 96480
tatgtcaact gttggtcttg ggtctataat ttgggactct tattttttttta aaaaatatatat 96540
taataattttt aatatttagat ttatgggaata gtaacacaga tagtagcagag 96600
tcctcatattg accttctaacc aagttcctct ctttgtaaca tctttcatca cctattttgt 96660
cagaactcaac aacaagtctat taattatgtta cctatttaata ctaataatac tttattgtaa 96720
tcttacctgtt ttcctcttcct gtcttctctct ctgtaggggt cccctccacag agacaccacac 96780
tgatttttact ttgcttctagc agtcttccag cttctcttta tttttctgta cctttgcagat 96840
cattctgtaga attgcctcagg taaccttgtag atgaacccca aagtggtgctttttttaaaa 96900
tttaaacttaa gaaatatttg ttttggaaaa gacacacaca gataatgaagt gccccttctca 96960
tcgagatcgtg aggaggatac acgacatcaca ctagacgctga ctggggatgt aatgcttgt 97020
cactctggagt ctcttagtgt tcgaagttcct cttggttcctcc cactgttaaaag atataatatttt 97080
tctttttctc tctgcttttgt ctgtagagaag aagttaactaa gctatgccaac ccccttagaa 97140
gtaatttgcc ggaatattcag ttgctctccct cttgaagggcattgcctaca gatagtaaat 97200
gatattctaa gtataacagatt gtttctttct cttttcttttt ttttttttca cttctattttt 97260
ttataggtata tggttagtatt tatttggcag ctcttccccag aaaaagagaa caaatggataa 97320
acacacacac acacacacac acacacacac acatataagaga ctttatatatag gatgttgttt 97380
acatgtgttt aagcctgtag aagttcctctg acgctttact tgcagccttg gataactaggt 97440
aagctatgcg tgatattgca tccacaagctc ggcctgggtct gtaagctttctc atacatcttc 97500
acacacacac acacacacac acacacacac acatataagaga ctttatatatag gatgttgttt 97560
caacoagagaca atataatttgcc ttctctctca cttttttggt tttactcaac ctoataagta 97620
tggacgttagc cccctcaacat cttggtgtgg ttgatcttac ttagctact gattcatacg 97680
tcaatctctt cccagacacag tcaacacacag ccacacacac ctaatattttt gcagctacag 97740
tggggcttccc tttaacacagt caagttcaaca caaaaattata caaacaacagt atgaattcaat 97800
gtatatatttt tattgtattt ggcattatac atcgctatttttttt tttgttgtgc caaatgttgtt 97860
cagctctctca ctgggtctag gggcctgcct tcaaggctct cttgttttttt ttctctttttctg 97920
gtggagacag gcagctgttag aaatataacac acaagcacaac gacacacacac aaaaaaagag 97980
tggggcttgg gggccacacag cccacacagctt gagggtgttgc aatgtgtctggt aaatttttgc 98040
tgtgacttgtg atttattgga tacaagacat gggggcaggg ttgaaggtggt gacgctcttc 98100
caatggtgctg caacgctcgct gcgtgtctcg cagggggcct cccctctctgt 98160
gggccacacag gggggcaggg agagagagag gcagacagac cagcagcttttttttttttttta 98220
atactgacag ccctatttatttt ttttcataat tcgggtctct cttgatctg gggcagacca 98280
agtgtatctgg atgcggagat ccagagcagcc aggagggtgac cacaattttttt ccacgatc 98340
agggagatgg taggcttctct gaaaaactagc aaggagcagt gccagagacct aacctttaaaa 98400
caagagatgg tggagatagag ctctctctat tccactcttca ccctcctttg ggcacccgtc 98460
ttttgctcgtg aagtagggcg tcctttctctg ggtagctgcg ggtcacagctgatctggt 98520
cggattgtaa cgggacctcc ctaacacactg gggcacttctt ctagatccgg gacgctcttc 98580
ctggaacctcc gggtggtctt ccggacagaa cagagctttgt tctttgctct cttttttctca 98640
tgctctcttct cttggtctgc tggatggtgg ctgctttcttt gatgttatggt ttgtagaagct 98700
aggacattag taatagaaaa ataaagagata acttatcaacactatag ggtgataactt 98760
atatataact atgtctatga ttctatatata tctagcataaa ctctttgttat tttatatatt 98820
ttatatatg gaacagctctg tgcctgctct cttgctccttg cactctgggt cttgtgcaccc 98880
cagcactggtc caacggaact tttttcaggt tcagtcattg gctgttcctcc cacagctgcctt 99440
tcctttttgt ttgtaaatgat ttctctattct tagttgccgg ttagatattc taggctattt 99000
tgtagttacct ctttttagt gtaataatctg tctgctatctt ctctttggtg 99060
aatgtatatc aagaactacag atctctggaac tgggtggtct gcgtgcctctt tggagtgcatc 99120
tgctcttgaa ctctctgctg agatgtgactg aatgataattata gttatatata ctaacccaca 99180
gatataataa ttgctataaat tattctataa tttatccatt ctgtctatag ttaagaatat 99240
cagcatggag ccctcttccaa ttgtgttattg ttctttttcct gcctatcacg cgataataat 99300
aatggtcctt tctctcttgg catgacccctgc agatcctctc attcatatag tgcttctcag 99360
aatctctctt gtctcttttct otaacggtgcct cttggtgta tggaggtccac cocctctatc 99420
gtctgacagc ccctcttccaa ttgctgacag cttgcttccct atatacactct taactggagt 99480
ctgaggacgt caggggctct ctgggtctca ggcctagaaa tagcttgccac atgtgcaggtg 99540
cagaagtgcgt agatagtgtgca gctctcaactg cttaggggatat acggagtttg tccttctttt 99600
agtttgtggt atgatagtctc tggctgtctgc aagactggttc ctttggtgcata tgggtcaca 99660
tatgacgaca acagagtggaa acacacgctcg tgacacgagc ggaacactgtc gctgacccga 99720
aatggtcttt gatcctgtatt ccagctccct gcttgcttgct cagcctggtc gcattgtcttg 99780
actttttgtc acattttttcag cagggtagaat caaactatat gcaggtttgt actgctctag 99840
agttttttttt cctctcttttttt tttggggtta cacccagaaat tattttttaa actggttctt 99900
attatatctt ctatctattt atcactttcttt cttgctatact cttactttttatttatgtaa 99960
gatagacttgc ggctgtagcact tagttataaaa tagtctgtgtcct atgcatgtttt 100020
aatgactggc cccagcagaaa ttgcttact catcgtgttt ctgaactttc ttctttgtag 100080
tgggtgattc aaataataat cactacttcg aagattttcg tgcacaaatg atacagaaaaa 100140
saagacttct gaccataaaact ctcactataa gatacctttt actatatttttgttattttttt 100200
ttcttttttt actgcttttt gttatattgat tagctctttcag tcaactctctt attcttgtaaaa 100260
aatattgtggc ttgataatac agcctttttttt gctggtgtaga acaaataaaa 100320
acagactggtc taactatcaca aacatcagaaatttatccttcactgcgtggagaa 100380
gcagcagata aggggacacttg agttgctggg cttctgctgttt tttgcaaatga 100440
gagtccttttc ttggtggagtgc taatattttc ctgaactgtgca gggagttttg 100500
gtagagggc cggagaatgct ttctgtgctac ccctctgtactc ttctcaagttgt ggagttggg 100560
tctactctgc atgctttattt cactcttctaa gggactccctt cttataagttg tggctacggtc 100620
gattagtttt caacacatgtt ttgctttggact gatactattct taacaagctata ccttggtcact 100680
atattttttt taatattattc ataatatatca atataaatagc aagaatgtttt 100740
ttctctgtgcat acagctagca taccacgca gacagatggc cttattttgtt ctcataagttt 100800
atgagttattc ctagtactag cttgagctgtg tttggtagtgc tggaccttact gacaatattctgtaa 100860
atctctcagcctc tggctgtata cccgct cacagcttttc tttttttttt ctgacgggttgc 100920
aatgagttttcctacgtactag cgggtattgct tagctcttgttggcagtttcttgagttgggtg 101040
ctgcaactgtc gcgtcttggtttaacactaa ttgcttttttttctgctagttc 101100
aatctctgtgtcct gatggtgttcttattata aaaaagggcataaataaatgacccagaggt 101160
-continued

```
aaaggggag gccatatggg aaaaaagttt ggcaagtcct caaaaatgta cacataggt 108300
tacccatac gcccagaaatt ctaaccttgg tatatacaca agagaataca aacaactaatg 108360
tcaacacaata ttcgtgctcag gagctggtcag acaagctgta cttcaatatcag 108420
gaaatgagc acatacaagc tgaataagata aataatggtt tatccgtata atgggctgac 108480
tccagagagaa aaaaagaaatg aagttgtaaat aatcggcata acgcagagaag cagctggaaa 108540
tcttacaaa gggaaaagaa gtcagacaca aaagggcaca taaaagtttga cttcattatatt 108600
ggagaaatgct cagaaacaggg aatactaatgag acagagagag tagatgaaattg ttggtcagag 108660
actggaggaga cgggagaaaag gggggtggtcag gttcttggtgc acagagttttct ttcggggagt 108720
gattaagtaga taacctgaaatt aagaagattg cagttggtcga caaacctgca aataactcaaa 108780
aaacacacaa attgctgact ttaaaggtt gaaattttag gcatttgaat tatatacaca 108840
ctttaaaaoct tcgtcagata tattttgatt ttaaaacaaag agatgctcttc ttcacattaag 108900
aaaaagatagc gcagagatac aacatttttac aagagagaaaa ggtagcaaga cctctgatttt 108960
tagatcctcag ccaggggagcct tggggaagac aagcggggu 109020
atcttttccc cagagcagcct ccttttcctttt gtcctttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
aattctgtt tcagcttgc aaatacacc gcgtgcttct ctaacccccaa ggaacaaag 110700
ccacaccctt agaatcactc cgtgaattgt ggctttctaat ttctgtgact ttttcctgttg 110760
atgctcctgg ggaagtattg tgtggtgaa tcttggtggat ggagatacga ggcagatcctc 110820
tttactcaac accttttatt agctgcttac agatgctaat getaagtaaa ccaagccccag 110880
cgtctcagt cagctgatga gaaagggcag cacggtgtta tcagctttctt aaaaataaat 110940
aaatgattga cagaagggcct tataacatcag tgggtcttgga gctttcaccag ggtcctggag 111000
gttcaaggg tttctcactg tggaggtggt gttacagagag aatattcacc ccaatagctc 111060
agtggagatca agggccctga gatgagaag aaaaagcggca agtcttgagag aatagagagg 111120
aggagagacg cattgagatgg cgctagaggg ccaagggcgt tagtgggagct tgtgtagat 111180
tgtcaggtg tgtcgtctact cctcataaggc aatggtcact gacacaaaaagtt acctcggagaa 111240
tttgtcagtt tattagggcg tggatctcag gcagagctgaa tgtttaattg cattgtgaag 111300
agatggggctg agtgggtaacc cagatgttct tgtgagttgt tgttatagga tttctcctat 111360
ccataaacc gcagagttcg tgctattcct cagcagaaaa taacacttgc 111420
agggcaaatc attacgcttt aatttaatct gcgaatattat agtattgaga agtacttacca 111480
atctttggtta atgtcaacag acgctttcacc aaaaacgact ttgttgatgt gtagtttagaa 111540
gtggcatgga ttaggtgttg gctattaaacag aagttgcaacatgcgacagagaa gtgtggtgca 111600
cacaaatact acagatcatt cgaatccagct tgtcctcttt aacaccccaca ccaagctcct 111660
catagtcctg cccttctctga aagggagaat attctgttag gcggcgccttt tttctgacctgt 111720
ccactgcgcag gggagagatt ggtgattgcctg cttttcttaa aagttttttaa aacactttgg 111780
agcagctggag aaacagggga caaaagtctgcg aatctgctgc acaacactttt gcacggtggg 111840
acatggtcag gagaaggtgt atttttctta gaactacaacat tgcgaaaggta taacacata 111900
atagcggcct gcagagagtt gatcataaat taataatgat tggagagagtatta 111960
gggctgctctaa atataatatta caaaaagttt ttttatgatatt ttttcttctgag aatccaaat 112020
tgcaccaatctt gcaccccctc gctctctctc ccaaaaccc acatgctgac 112080
taatcccttc attttatatg cgcacatcag cattctaat cacatatcatt 112140
aatatatatt aagatagctc atatcttcga aaaaagttt caagctctctt tttctctctt 112200
ttgtaaagtt cattataacaa aaaaatttac accctaaacattttgctgta acagcttgag 112260
tgatataacg tgcagccatcct tctgtggtcag tcggcttcct ccctaaacat 112320
ttttaaccct gcagacccct aactcctgtaa actgctagtc gctggcgc 112380
cttcagcttc gacaacccca aatccttctc atctgctctcct cctgctggttc 112440
cctctgtgaa aggaggtact gatgtagttg tttttctgag cagttgcctt ttcttctcag 112500
attaatactcg tgtgaatgctt cgcctatcct aaccgatata ggggacgac 112560
ttgggggtgt gttctccttg cttctgctgt gccaaatattctt ggggctggcct ttccttcggc 112620
acatcctgg acagccccaa tttctcctgg ctttttttttt ggtcctgtct ggggacttctt 112680
gggctcgtgt ctgtattttt tttgagaggt cctgtcttgg gctttttttcag ggtgctgtgc 112740
attatatctctc tctttctctct aaccgatata ggggacgac 112800
cgatcggtag gtgttctgtt cttctttctggt tttttgctggcgccttttaa aactatatctt 112860
caggggttttt ttgttactttttttttcagccagc ttttctgttag 112920
ggacagcc ttcgctcatat cttcttttctt gttctgtccct ttctgctgtt ggagagagtc 112980
-continued

taggagttct cccataaact cctttcagaa tattgtattgc gaaatattt ttatccagttc 113040
tgcaggttgc ctctttcctc cattgtagct gctttgtgat gcagcgttct aattcttcatg 113100
aagtctaatg ttcctctcct gcctgtcaca aaataaatat cttattactc 113160
ccatagtta gcaaaatatt tttttgctcc attatagcagt ttgattttaa caataaagac 113220
cctctgtgata cagatggtact gatgatggca ttttattgga atgagaaat acaaccacag 113280
tataattgaa cattattgag acceaaagga caactgcata cattttatac ataataactc 113340
tctcagatg ttcgctatttg ggtggtgagg gtaggggaaag atgaaagggac atataagaaa 113400
aaaacatga cgggggtcag ttccttatttt tttcattgtg tgggggttc gaagaacgaatt 113460
cctttgtcat tttttttttt tctataattgc atttgtcatttag ggtcagactc cctcaatagt 113520
tataataagttv aataactcaat atcaattttt tttctactcat gcaggtcata 113580
tgctgttggct gcctttgctt gttttgggttt gttgtttgtgaa atcaagggtc 113640
tctctgtcatgt tgcatttct tttttgtcag ggaaccccag tcctcgtggct gcagggct 113700
ttattgtctcg gctggtgtcttt ttaagcagttg atccttctca agctgtgttct ccctttttat 113760
ttggtctcat tgcggttgcc agagagacact tttttttccttc tctagaggg agatagttag 113820
gacaacaca aataatgtag catcaagcggc aagtgggaca ggggagctgc ctgtaaatctg 113880
tcagataacttt tctataagtg cctttttgctta tttttgtcatg tttacttcatg 113940
aagcacaca aataatgtag catcaagcggc aagtgggaca ggggagctgc ctgtaaatctg 114000
ttggtctcaact tgcggttgcc aagagtaggga ttttatttttt cttttgctttg 114060
ntattgaggt gaatgggcat ttttttttcttg cccttctctt ttccttttttt tcctttttat 114120
atattgaggt gaatgggcat ttttttttcttg cccttctctt ttccttttttt tcctttttat 114180
ttggatttc aatttttttt tattggcact ttaaaccctg acaaatctgt caaccocatt 114240
ttattgagtt ttttttttcttg cccttctctt ttccttttttt tcctttttat 114300
tgctatgaggt aataatattt cccaaagcag aaatatttttat cgaatatctc ttaaaccctg 114360
tgcctacaag tcaagcaaac ttttttttcttg cccttctctt ttccttttttt tcctttttat 114420
tgaggtctcttgc gagaatgtac ctgtctttt gacctttattt ccctttttat 114480
atattgaggt gaatgggaggt aataatattt cccaaagcag aaatatttttat cgaatatctc ttaaaccctg 114540
acacagggct gtatagttgttg tggagtgtcactg ttttttttcttg cccttctctt ttccttttttt tcctttttat 114600
acacagggct gtatagttgttg tggagtgtcactg ttttttttcttg cccttctctt ttccttttttt tcctttttat 114660
ctcttttattt ataatgtatgt ttggcctttc ttttttttcttg cccttctctt ttccttttttt tcctttttat 114720
atattgaggt atatatattttt cccaaagcag aaatatttttat cgaatatctc ttaaaccctg 114780
atattgaggt atatatattttt cccaaagcag aaatatttttat cgaatatctc ttaaaccctg 114840
acacagggct gtatagttgttg tggagtgtcactg ttttttttcttg cccttctctt ttccttttttt tcctttttat 114900
acacagggct gtatagttgttg tggagtgtcactg ttttttttcttg cccttctctt ttccttttttt tcctttttat 114960
acacagggct gtatagttgttg tggagtgtcactg ttttttttcttg cccttctctt ttccttttttt tcctttttat 115020
acacagggct gtatagttgttg tggagtgtcactg ttttttttcttg cccttctctt ttccttttttt tcctttttat 115080
acacagggct gtatagttgttg tggagtgtcactg ttttttttcttg cccttctctt ttccttttttt tcctttttat 115140
acacagggct gtatagttgttg tggagtgtcactg ttttttttcttg cccttctctt ttccttttttt tcctttttat 115200
acacagggct gtatagttgttg tggagtgtcactg ttttttttcttg cccttctctt ttccttttttt tcctttttat 115260
acacagggct gtatagttgttg tggagtgtcactg ttttttttcttg cccttctctt ttccttttttt tcctttttat 115320
acacagggct gtatagttgttg tggagtgtcactg ttttttttcttg cccttctctt ttccttttttt tcctttttat 115380
-continued

gcccagagac acctgtccca cagactctcc tctgtaaatg ttcctatctc 122520

gcgctcctc cttaaatctc ccaagggcct tttctcaac cactgtctta tcaagtggagc 122580
tttcacttct tcaaggggtta gatgtggataa ttctcttaggt cacttataac 123640
cctctcttccc tggctcacaata taacaacacac aatgtattgt tttttttctt ggaatcctat 124700
ttaaaaataac aatctgtaatg ctatgttggttc cttccgcttt agaggataggt tatttgcctc 125760
gtgctcagtc tggcagggcc aacgttggcc cttcccactg tctagggatt aacacetgccc 126820
cagacatccg tttcttctgt cttcctcaggt cagggggtg ccagacatcag ctggctctgtc 127890
tagcagaga tagatataaaa ccccttctagc gaatttcccc cttgcaacacac cttgcctctgc 128940
gaaatgctgac ggccagggcag cctgtgaggcc agagggtgta agtaaattc ggttggattc 129990
agagatgcatt ggttctcttcg aattttaaaatg tttctatttt ctgcgctttc aagcttgctgg 130950
agagatgcatt ggttctcttcg aattttaaaatg tttctatttt ctgcgctttc aagcttgctgg 131910
gggtataga atagagaggtt ataactccc tggaaatgct gctattatcc cactgtggctgc 132870
agagatgcatt ggttctcttcg aattttaaaatg tttctatttt ctgcgctttc aagcttgctgg 133830
tgacgctctc tggagctacat atatatat caatctctgct cctttttctca ggttgtggaa 134790
tccagcgtacc cacaaggtg tcgagcagc gtgggtgtctc cttgagccttc gctgagccccc 135850
agaactctc catcctctct ctggctcact ctctgcaaccct ggaaatcactg cagaggttatt 136810
tatggagag ccggtctctcaat ttacaggcgg gtaagaagcc tagcagtgtag ccctctcttt 137770
cagagaggggtg ccggtctctcaat ttacaggcgg gtaagaagcc tagcagtgtag ccctctcttt 138830
agaactctc catcctctct ctggctcact ctctgcaaccct ggaaatcactg cagaggttatt 139890
tatggagag ccggtctctcaat ttacaggcgg gtaagaagcc tagcagtgtag ccctctcttt 139890
<210> SEQ ID NO 4
<211> LENGTH: 551
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

Met Arg Asp Tyr Asp Glu Val Ile Ala Phe Leu Gly Glu Trp Gly Pro
1      5       10      15
Phe Gln Arg Leu Ile Phe Phe Leu Leu Ser Ala Ser Ile Ile Pro Ann
20     25       30
Gly Phe Ann Gly Met Ser Val Val Phe Leu Ala Gly Thr Pro Glu His
35     40       45
Arg Cys Arg Val Pro Asp Ala Ala Ann Leu Ser Ser Ala Ser Trp Arg Ann
50     55       60
Ann Ser Val Pro Leu Arg Leu Arg Asp Gly Arg Glu Val Pro His Ser
70     75       80
Cys Ser Arg Tyr Arg Leu Ala Thr Ile Ala Asn Phe Ser Ala Leu Gly
85     90      95
Leu Glu Pro Gly Arg Asp Val Asp Leu Gly Gin Leu Glu Gin Glu Ser
100    105     110
Cys Leu Asp Gly Trp Glu Phe Ser Gin Asp Val Tyr Leu Ser Thr Val
115    120     125
Val Thr Glu Trp Ann Leu Val Cys Glu Asp Ann Trp Lys Val Pro Leu
<table>
<thead>
<tr>
<th>130</th>
<th>135</th>
<th>140</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr Ser Leu Phe Gly Val Leu Leu Gly Ser Phe Val Ser 145</td>
<td>150</td>
<td>155</td>
</tr>
<tr>
<td>Gly Cln Leu Ser Asp Arg Phe Gly Arg Lys Asn Val Leu Phe Ala Thr 165</td>
<td>170</td>
<td>175</td>
</tr>
<tr>
<td>Met Ala Val Glu Thr Gly Phe Ser Phe Leu Gln Ile Phe Ser Ile Ser 180</td>
<td>185</td>
<td>190</td>
</tr>
<tr>
<td>Trp Glu Met Phe Thr Val Leu Phe Val Ile Val Gly Met Gly Gln Ile 195</td>
<td>200</td>
<td>205</td>
</tr>
<tr>
<td>Ser Asn Tyr Val Val Ala Phe Ile Leu Gly Thr Glu Ile Leu Gly Lys 210</td>
<td>215</td>
<td>220</td>
</tr>
<tr>
<td>Ser Val Arg Ile Ile Phe Ser Thr Leu Gly Val Cys Thr Phe Ala 225</td>
<td>230</td>
<td>235</td>
</tr>
<tr>
<td>Val Gly Tyr Met Leu Leu Pro Leu Phe Ala Tyr Phe Ile Arg Asp Trp 245</td>
<td>250</td>
<td>255</td>
</tr>
<tr>
<td>Arg Met Leu Leu Ala Leu Thr Val Pro Gly Val Leu Cys Val Pro 260</td>
<td>265</td>
<td>270</td>
</tr>
<tr>
<td>Leu Trp Trp Phe Ala Leu Glu Ser Arg Pro Arg Trp Leu Ile Ser Gin Arg 275</td>
<td>280</td>
<td>285</td>
</tr>
<tr>
<td>Arg Phe Arg Glu Ala Glu Asp Ile Ile Gin Lys Ala Ala Lys Met Asn 290</td>
<td>295</td>
<td>300</td>
</tr>
<tr>
<td>Asn Ile Ala Val Pro Ala Val Ile Phe Asp Ser Val Glu Leu Asn 305</td>
<td>310</td>
<td>315</td>
</tr>
<tr>
<td>Pro Leu Lys Gln Gln Lys Ala Phe Leu Asp Leu Phe Arg Thr Arg 325</td>
<td>330</td>
<td>335</td>
</tr>
<tr>
<td>Asn Ile Ala Ile Met Thr Ile Met Ser Leu Leu Leu Trp Met Leu Thr 340</td>
<td>345</td>
<td>350</td>
</tr>
<tr>
<td>Ser Val Gly Tyr Phe Ala Leu Ser Leu Asp Ala Pro Asn Leu His Gly 355</td>
<td>360</td>
<td>365</td>
</tr>
<tr>
<td>Asp Ala Tyr Leu Asn Cys Phe Leu Ser Ala Leu Ile Glu Ile Pro Ala 370</td>
<td>375</td>
<td>380</td>
</tr>
<tr>
<td>Tyr Ile Thr Ala Trp Leu Leu Arg Thr Leu Pro Arg Arg Tyr Ile 385</td>
<td>390</td>
<td>395</td>
</tr>
<tr>
<td>Ile Ala Ala Val Leu Phe Trp Gly Gly Val Leu Leu Phe Ile Gln 405</td>
<td>410</td>
<td>415</td>
</tr>
<tr>
<td>Leu Val Pro Val Asp Tyr Phe Leu Ser Ile Gly Leu Val Met Leu 420</td>
<td>425</td>
<td>430</td>
</tr>
<tr>
<td>Gly Lys Phe Gly Ile Thr Ser Ala Phe Ser Met Leu Tyr Val Phe Thr 435</td>
<td>440</td>
<td>445</td>
</tr>
<tr>
<td>Ala Glu Leu Tyr Pro Thr Leu Val Arg Asn Met Ala Val Gly Val Thr 450</td>
<td>455</td>
<td>460</td>
</tr>
<tr>
<td>Ser Thr Ala Ser Arg Val Gly Ser Ile Ala Pro Tyr Phe Val Tyr 465</td>
<td>470</td>
<td>475</td>
</tr>
<tr>
<td>Leu Gly Ala Tyr Asn Met Leu Pro Tyr Ile Val Met Gly Ser Leu 485</td>
<td>490</td>
<td>495</td>
</tr>
<tr>
<td>Thr Val Leu Ile Gly Ile Leu Thr Leu Phe Phe Pro Glu Ser Leu Gly 500</td>
<td>505</td>
<td>510</td>
</tr>
<tr>
<td>Met Thr Leu Pro Glu Thr Leu Glu Gln Met Gin Lys Val Lys Trp Phe 515</td>
<td>520</td>
<td>525</td>
</tr>
<tr>
<td>Arg Ser Gly Lys Thr Arg Asp Ser Met Glu Thr Glu Asn Pro 530</td>
<td>535</td>
<td>540</td>
</tr>
<tr>
<td>Lys Val Leu Ile Thr Ala Phe 545</td>
<td>550</td>
<td></td>
</tr>
</tbody>
</table>
<210> SEQ ID NO 5
<211> LENGTH: 551
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 5

Met Ala Gln Phe Val Gln Val Leu Ala Glu Ile Gly Aep Phe Gly Arg
  1      5      10     15
Phe Gln Ile Gln Leu Leu Ile Leu Cys Val Leu Asn Phe Leu Ser
    20    25     30
Pro Phe Tyr Phe Ala His Val Phe Met Val Leu Asp Glu Pro His
    35    40     45
His Cys Ala Val Ala Trp Val Lys Asn His Thr Phe Asn Leu Ser Ala
    50    55     60
Ala Glu Gln Leu Val Leu Ser Val Pro Leu Asp Thr Ala Gly His Pro
    65    70     75     80
Glu Pro Cys Leu Met Phe Arg Pro Pro Pro Ala Asn Ala Ser Leu Gln
    85    90     95
Asp Ile Leu Ser His Arg Phe Asn Glu Thr Gln Pro Cys Asp Met Gly
   100   105    110
Trp Glu Tyr Pro Glu Asn Arg Leu Pro Ser Leu Lys Asn Gln Phe Asn
   115   120    125
Leu Val Cys Asp Arg Lys His Leu Lys Asp Thr Thr Gln Ser Val Phe
   130   135
Met Gly Gly Leu Val Gly Thr Leu Met Phe Gly Pro Leu Cys Asp
   145   150    155    160
Arg Ile Gly Arg Lys Ala Thr Ile Leu Ala Glu Leu Leu Phe Thr
   165   170    175
Leu Ile Gly Leu Ala Thr Ala Phe Val Pro Ser Phe Glu Leu Tyr Met
   180   185    190
Ala Leu Arg Phe Ala Val Ala Thr Val Ala Gly Leu Ser Phe Ser
   195   200    205
Asn Val Thr Leu Leu Thr Glu Trp Val Gly Pro Ser Thr Arg Thr Gln
   210   215    220
Ala Val Val Leu Ala Gln Cys Asn Phe Ser Leu Gly Gln Met Val Leu
   225   230    235    240
Ala Gly Leu Ala Tyr Gly Phe Asn Phe Arg Leu Leu Gln Ile Thr
   245   250    255
Gly Thr Ala Pro Gly Leu Leu Leu Phe Phe Tyr Phe Thr Ala Leu Pro
   260   265
Glu Ser Ala Arg Trp Leu Leu Thr Arg Gly Arg Met Asp Glu Ala Ala
   270   275    280    285
Gln Leu Ile Gln Lys Ala Ala Ser Val Asn Arg Arg Lys Leu Ser Pro
   290   295    300
Glu Leu Met Asn Gln Leu Val Pro Glu Lys Thr Gly Pro Ser Gly Asn
   305   310    315    320
Ala Leu Asp Leu Phe Arg His Pro Gln Leu Arg Lys Val Thr Leu Ile
   325   330    335
Ile Phe Cys Val Trp Phe Val Asp Ser Leu Gly Tyr Tyr Gly Leu Ser
   340   345
Leu Gln Val Gly Asp Phe Gly Leu Asp Val Tyr Leu Thr Gln Leu Ile
   355   360    365
Phe Gly Ala Val Glu Val Pro Ala Arg Cys Ser Ser Ile Phe Met Met
Gln Arg Phe Gly Arg Lys Trp Ser Gln Leu Gly Thr Leu Val Leu Gly
370 375 380
Gly Leu Met Cys Ile Ile Ile Phe Ile Pro Ala Asp Leu Pro Val
385 390 395 400
Val Val Thr Met Leu Ala Val Val Gly Lys Met Ala Thr Ala Ala Ala
405 410 415
Phe Thr Ile Ser Tyr Val Tyr Ser Ala Glu Leu Phe Pro Thr Ile Leu
425 430
Arg Gln Thr Gly Met Gly Leu Val Gly Ile Phe Ser Arg Ile Gly Gly
440 445
Ile Leu Thr Pro Leu Val Ile Leu Leu Gly Glu Tyr His Ala Ala Leu
450 455 460
Pro Met Leu Ile Tyr Gly Ser Leu Pro Ile Val Ala Gly Leu Cys
465 470 475 480
Thr Leu Leu Pro Glu Thr His Gly Gln Gly Leu Lys Asp Thr Leu Gln
485 490 495 500
Asp Leu Glu Leu Gly Pro His Pro Arg Ser Pro Lys Ser Val Pro Ser
505 510
Glu Lys Glu Thr Glu Ala Lys Gly Arg Thr Ser Ser Pro Gly Val Ala
520 525
Phe Val Ser Ser Thr Tyr Phe
530 535 540
545 550

<210> SEQ ID NO 6
<211> LENGTH: 557
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 6
Met Arg Asp Tyr Asp Glu Val Thr Ala Phe Leu Gly Glu Trp Gly Pro
1 5 10 15
Phe Gln Arg Leu Ile Phe Phe Leu Leu Ser Ala Ser Ile Ile Pro Asn
20 25 30
Gly Phe Thr Gly Leu Ser Ser Val Phe Leu Ile Ala Thr Pro Glu His
35 40 45
Arg Cys Arg Val Pro Asp Ala Ala Asn Leu Ser Ser Ala Thr Pro Asn
50 55 60
His Thr Val Pro Leu Arg Leu Arg Asp Gly Arg Glu Val Pro His Ser
65 70 75 80
Cys Arg Arg Tyr Arg Leu Ala Thr Ile Ala Asn Phe Ser Ala Leu Gly
85 90 95
Leu Glu Pro Gly Arg Asp Val Asp Leu Gly Gin Leu Glu Gin Glu Ser
100 105 110
Cys Leu Asp Gly Trp Gly Phe Ser Gln Asp Val Tyr Leu Ser Thr Ile
115 120 125
Val Thr Glu Trp Asn Leu Val Cys Glu Asp Asp Thr Lys Ala Pro Leu
130 135 140
Thr Ile Ser Leu Phe Phe Val Gly Val Leu Gly Ser Phe Ile Ser
145 150 155 160
Gly Gin Leu Ser Arg Arg Phe Gly Arg Lys Asn Val Leu Phe Val Thr
165 170 175
Met Gly Met Gln Thr Gly Phe Ser Phe Leu Gln Ile Phe Ser Lys Asn
180 185 190
<210> SEQ ID NO 7
<211> LENGTH: 555
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 7

Met Pro Thr Thr Val Asp Asp Val Leu Glu His Gly Gly Gly Phe His
1  5   10  15
<table>
<thead>
<tr>
<th>Residue</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td>Phe Gln Lys Gln Met Phe Phe Leu Leu Ala Leu Leu Ser Ala Thr</td>
</tr>
<tr>
<td></td>
<td>20 25 30</td>
</tr>
<tr>
<td>Gly</td>
<td>Trp Ser Pro Ile Tyr Val Gly Ile Val Phe Leu Gly Phe Thr Pro Asp</td>
</tr>
<tr>
<td></td>
<td>35 40 45</td>
</tr>
<tr>
<td>Ser</td>
<td>His Arg Cys Arg Ser Pro Gly Val Ala Glu Leu Ser Leu Arg Cys Gly</td>
</tr>
<tr>
<td></td>
<td>50 55 60</td>
</tr>
<tr>
<td>Ala</td>
<td>Trp Ser Pro Ala Glu Glu Leu Asn Tyr Thr Val Pro Gly Pro Gly Pro Pro</td>
</tr>
<tr>
<td></td>
<td>65 70 75 80</td>
</tr>
<tr>
<td>Leu</td>
<td>Ala Gly Glu Ala Ser Pro Arg Glu Cys Arg Arg Tyr Gly Val Asp Trp</td>
</tr>
<tr>
<td></td>
<td>85 90 95</td>
</tr>
<tr>
<td>Ala</td>
<td>Asn Gln Ser Thr Phe Asp Cys Val Asp Pro Leu Ala Ala Ser Leu Asp Thr</td>
</tr>
<tr>
<td></td>
<td>100 105 110</td>
</tr>
<tr>
<td>Thr</td>
<td>Asn Arg Ser Arg Leu Pro Leu Gly Pro Cys Arg Asp Gly Trp Val Tyr</td>
</tr>
<tr>
<td></td>
<td>115 120 125</td>
</tr>
<tr>
<td>Gly</td>
<td>Glu Thr Pro Gly Ser Ile Val Thr Glu Phe Asn Leu Val Cys Ala</td>
</tr>
<tr>
<td></td>
<td>130 135 140</td>
</tr>
<tr>
<td>Ser</td>
<td>Asn Ser Trp Met Leu Asp Leu Phe Gln Ser Ser Val Asn Val Gly Phe</td>
</tr>
<tr>
<td></td>
<td>145 150 155 160</td>
</tr>
<tr>
<td>Gly</td>
<td>Phe Ile Gly Ser Met Ser Ile Gly Tyr Ile Ala Asp Arg Phe Gly Arg</td>
</tr>
<tr>
<td></td>
<td>165 170 175</td>
</tr>
<tr>
<td>Thr</td>
<td>Lys Leu Cys Leu Leu Thr Thr Val Leu Ile Asn Ala Ala Ala Gly Val</td>
</tr>
<tr>
<td></td>
<td>180 185 190</td>
</tr>
<tr>
<td>Val</td>
<td>Leu Met Ala Ile Ser Pro Thr Tyr Thr Trp Met Leu Ile Phe Arg Leu</td>
</tr>
<tr>
<td></td>
<td>195 200 205</td>
</tr>
<tr>
<td>Ser</td>
<td>Ile Gln Gly Leu Val Ser Lys Ala Gly Trp Leu Ile Gly Tyr Ile Leu</td>
</tr>
<tr>
<td></td>
<td>210 215 220</td>
</tr>
<tr>
<td>Asp</td>
<td>Ile Thr Glu Phe Val Gly Arg Arg Tyr Arg Thr Val Gly Ile Phe</td>
</tr>
<tr>
<td></td>
<td>225 230 235 240</td>
</tr>
<tr>
<td>Gly</td>
<td>Tyr Gln Val Ala Tyr Thr Val Gly Leu Leu Val Leu Ala Gly Val Ala</td>
</tr>
<tr>
<td></td>
<td>245 250 255</td>
</tr>
<tr>
<td>Ala</td>
<td>Tyr Ala Leu Pro His Trp Arg Trp Leu Gln Phe Thr Val Ala Leu Pro</td>
</tr>
<tr>
<td></td>
<td>260 265 270</td>
</tr>
<tr>
<td>Asp</td>
<td>Asn Phe Phe Phe Leu Leu Tyr Tyr Trp Cys Ile Pro Glu Ser Pro Arg</td>
</tr>
<tr>
<td></td>
<td>275 280 285</td>
</tr>
<tr>
<td>Gly</td>
<td>Trp Leu Ile Ser Gln Asn Lys Asn Ala Glu Ala Met Arg Ile Ile Lys</td>
</tr>
<tr>
<td></td>
<td>290 295 300</td>
</tr>
<tr>
<td>Ser</td>
<td>His Ile Ala Lys Lys Asn Gly Lys Ser Leu Pro Ala Ser Leu Gln Arg</td>
</tr>
<tr>
<td></td>
<td>305 310 315 320</td>
</tr>
<tr>
<td>Ala</td>
<td>Leu Arg Leu Glu Glu Thr Gly Lys Leu Asn Pro Ser Phe Leu</td>
</tr>
<tr>
<td></td>
<td>325 330 335</td>
</tr>
<tr>
<td>Thr</td>
<td>Asp Leu Val Arg Thr Pro Gln Ile Arg Lys His Thr Met Ile Leu Met</td>
</tr>
<tr>
<td></td>
<td>340 345 350</td>
</tr>
<tr>
<td>Val</td>
<td>Tyr Asn Trp Phe Thr Ser Val Leu Tyr Gln Gly Leu Ile Met His</td>
</tr>
<tr>
<td></td>
<td>355 360 365</td>
</tr>
<tr>
<td>Ser</td>
<td>Met Gly Leu Ala Gly Asp Asn Ile Tyr Leu Asp Phe Phe Tyr Ser Ala</td>
</tr>
<tr>
<td></td>
<td>370 375 380</td>
</tr>
<tr>
<td>Ala</td>
<td>Leu Val Glu Phe Pro Ala Ala Phe Met Ile Ile Leu Thr Ile Asp Arg</td>
</tr>
<tr>
<td></td>
<td>385 390 395 400</td>
</tr>
<tr>
<td>Ala</td>
<td>Ile Gly Arg Arg Tyr Pro Trp Ala Ala Ser Asn Met Val Ala Gly Ala</td>
</tr>
<tr>
<td></td>
<td>405 410 415</td>
</tr>
<tr>
<td>Asp</td>
<td>Ala Cys Leu Ala Ser Val Phe Ile Pro Gly Asp Leu Gln Trp Leu Lys</td>
</tr>
<tr>
<td></td>
<td>420 425 430</td>
</tr>
</tbody>
</table>
-continued

Ile Ile Ile Ser Cys Leu Gly Arg Met Gly Ile Thr Met Ala Tyr Glu
   435 440 445
Ile Val Cys Leu Val Asn Ala Glu Tyr Pro Thr Phe Ile Arg Asn
   450 455 460
Leu Gly Val His Ile Cys Ser Ser Met Cys Asp Ile Gly Gly Ile Ile
   465 470 475 480
Thr Pro Phe Leu Val Tyr Arg Leu Thr Asn Ile Trp Leu Glu Leu Pro
   485 490 495
Leu Met Val Phe Gly Val Leu Gly Leu Val Ala Gly Gly Leu Val Leu
   500 505 510
Leu Leu Pro Glu Thr Lys Gly Ala Leu Pro Glu Thr Ile Glu Glu
   515 520 525
Ala Glu Asn Met Glu Arg Pro Arg Gly Asn Lys Gly Lys Met Ile Tyr
   530 535 540
Leu Gln Val Gln Gly Leu Asp Ile Pro Leu Asn
   545 550 555

<210> SEQ ID NO 8
<211> LENGTH: 554
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 8

Met Pro Thr Val Asp Asp Ile Leu Glu Gln Val Gly Glu Ser Gly Trp
   1   5   10  15
Phe Gln Lys Gln Ala Phe Leu Ile Leu Cys Leu Leu Ser Ala Ala Phe
   20  25   30
Ala Pro Ile Cys Val Gly Ile Val Phe Leu Gly Phe Thr Pro Asp His
   35  40   45
His Cys Gln Ser Pro Gly Val Ala Glu Leu Ser Gln Arg Cys Gly Trp
   50  55   60
Ser Pro Ala Glu Glu Leu Asn Tyr Thr Val Pro Gly Leu Gly Pro Ala
   65  70   75  80
Gly Glu Ala Phe Leu Gly Gln Cys Arg Arg Tyr Glu Val Asp Trp Asn
   85  90   95
Gln Ser Ala Leu Ser Cys Val Asp Pro Leu Ala Ser Leu Ala Thr Asn
  100 105  110
Arg Ser His Leu Pro Leu Gly Pro Cys Gln Asp Gly Trp Val Tyr Asp
  115 120  125
Thr Pro Gly Ser Ser Ile Val Thr Glu Phe Asn Leu Val Cys Ala Asp
  130 135  140
Ser Trp Lys Leu Asp Leu Phe Gln Ser Cys Leu Asn Ala Gly Phe Leu
  145 150  155  160
Phe Gly Ser Leu Gly Val Gly Tyr Phe Ala Asp Arg Phe Gly Arg Lys
  165 170  175
Leu Cys Leu Leu Gly Thr Val Leu Val Asn Ala Val Ser Gly Val Leu
  180 185  190
Met Ala Phe Ser Pro Asn Tyr Met Ser Met Leu Phe Arg Leu Leu
  195 200  205
Gln Gly Leu Val Ser Lys Gly Asn Trp Met Ala Gly Tyr Thr Leu Ile
  210 215  220
Thr Glu Phe Val Gly Ser Gly Ser Arg Arg Thr Val Ala Ile Met Tyr
  225 230  235  240
Gln Met Ala Phe Thr Val Gly Leu Val Ala Leu Thr Gly Leu Ala Tyr
  245 250  255
| Ala | Leu | Pro | His | Trp | Arg | Trp | Leu | Gln | Leu | Ala | Val | Ser | Leu | Pro | Thr | 260 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Phe | Leu | Phe | Leu | Tyr | Tyr | Trp | Cys | Val | Pro | Glu | Ser | Pro | Arg | Trp |     |     |     |     |
| 275 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Leu | Ser | Gln | Lys | Arg | Asn | Thr | Glu | Ala | Ile | Lys | Ile | Met | Asp | His |     |     |     |     |
| 290 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Ile | Ala | Gln | Lys | Asn | Gly | Lys | Leu | Pro | Pro | Ala | Asp | Leu | Lys | Met | Leu |     |     |     |     |
| 305 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Ser | Leu | Glu | Glu | Asp | Val | Thr | Glu | Leu | Ser | Pro | Ser | Phe | Ala | Asp |     |     |     |     |     |     |     |     |     |
| 325 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Leu | Phe | Arg | Thr | Pro | Arg | Leu | Arg | Lys | Arg | Thr | Phe | Ile | Leu | Met | Tyr |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 340 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Leu | Trp | Phe | Thr | Asp | Ser | Val | Leu | Tyr | Gln | Gly | Leu | Ile | Leu | His | Met |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 355 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Gly | Ala | Thr | Ser | Gly | Asn | Leu | Tyr | Leu | Asp | Phe | Leu | Tyr | Ser | Ala | Leu |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 370 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Val | Glu | Ile | Pro | Gly | Ala | Phe | Leu | Ile | Thr | Ile | Asp | Arg | Val |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 385 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Gly | Arg | Ile | Tyr | Pro | Met | Ala | Met | Ser | Asn | Leu | Ala | Glu | Ala | Ala |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 405 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Cys | Leu | Val | Met | Ile | Phe | Ser | Pro | Asp | Leu | His | Trp | Leu | Asn | Ile |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 420 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Ile | Ile | Met | Cys | Val | Gly | Arg | Met | Gly | Ile | Thr | Ile | Ala | Ile | Gln | Met |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 435 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Ile | Cys | Leu | Val | Asn | Ala | Glu | Leu | Tyr | Pro | Thr | Phe | Val | Arg | Asn | Leu |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 450 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Gly | Val | Met | Val | Cys | Ser | Ser | Leu | Cys | Asp | Ser | Leu | Gly | Ile | Gly | Ile | Thr |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 465 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Pro | Phe | Ile | Val | Phe | Arg | Leu | Arg | Glu | Val | Trp | Gln | Ala | Leu | Pro | Leu |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 485 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Ile | Leu | Phe | Ala | Val | Leu | Gly | Leu | Ala | Ala | Gla | Val | Thr | Leu | Leu |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 500 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Leu | Pro | Glu | Thr | Lys | Gly | Val | Ala | Leu | Pro | Glu | Thr | Met | Lys | Asp | Ala |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 515 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Glu | Asn | Leu | Gly | Arg | Lys | Ala | Lys | Pro | Lys | Glu | Asn | Thr | Ile | Tyr | Leu |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 530 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|
| Lys | Val | Glu | Thr | Ser | Glu | Pro | Ser | Gly | Thr |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 545 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |-----|

<210> SEQ ID NO 9
<211> LENGTH: 539
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

Met Ala Phe Glu Leu Leu Ser Gln Val Gly Leu Gly Arg Phe 1 5 10 15
Gln Met Leu His Leu Val Phe Ile Leu Pro Ser Leu Met Leu Ile 20 25 30
Pro His Ile Leu Leu Glu Asn Phe Ala Ala Ile Pro Gly His Arg 35 40 45
Cys Trp Val His Met Leu Asp Asn Thr Gly Ser Gly Asn Glu Thr 50 55 60
Gly Ile Leu Ser Glu Asp Ala Leu Leu Arg Ile Ser Ile Pro Leu Asp
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Ser Asn Leu Arg Pro Glu Lys</td>
<td>Cys Arg Phe Phe Val Val His Pro Glu Trp</td>
<td>95</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Gln Leu Leu His Leu Asn Gly Ile His Ser Thr Ser Glu Ala Asp Thr</td>
<td>100</td>
<td>105</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>Glu Pro Cys Val Asp Gly Trp Val Tyr Asp Glu Ser Tyr Phe Pro Ser</td>
<td>115</td>
<td>120</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>Thr Ile Val Thr Lys Trp Asp Leu Val Cys Asp Tyr Glu Ser Leu Lys</td>
<td>130</td>
<td>135</td>
<td>140</td>
<td></td>
</tr>
<tr>
<td>Ser Val Val Gln Phe Leu Leu Leu Thr Gly Met Leu Val Gly Gly Ile</td>
<td>145</td>
<td>150</td>
<td>155</td>
<td>160</td>
</tr>
<tr>
<td>Ile His His Gly Val Ser Asp Arg Phe Gly Arg Arg Phe Ile Leu Arg</td>
<td>165</td>
<td>170</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>Trp Cys Leu Leu Gln Leu Ala Ile Thr Asp Thr Cys Ala Ala Phe Ala</td>
<td>180</td>
<td>185</td>
<td>190</td>
<td></td>
</tr>
<tr>
<td>Pro Thr Phe Pro Val Tyr Cys Val Leu Arg Phe Leu Ala Gly Phe Ser</td>
<td>195</td>
<td>200</td>
<td>205</td>
<td></td>
</tr>
<tr>
<td>Ser Met Ile Ile Ile Ser Asn Asn Ser Leu Pro Ile Thr Glu Trp Ile</td>
<td>210</td>
<td>215</td>
<td>220</td>
<td></td>
</tr>
<tr>
<td>Arg Pro Asn Ser Lys Ala Leu Val Val Ile Leu Ser Ser Gly Ala Leu</td>
<td>225</td>
<td>230</td>
<td>235</td>
<td>240</td>
</tr>
<tr>
<td>Ser Ile Gly Gln Ile Ile Leu Gly Gly Leu Ala Tyr Val Phe Arg Asp</td>
<td>245</td>
<td>250</td>
<td>255</td>
<td></td>
</tr>
<tr>
<td>Trp Gln Thr Leu His Val Val Ala Ser Val Pro Phe Leu Gly Leu Leu</td>
<td>260</td>
<td>265</td>
<td>270</td>
<td></td>
</tr>
<tr>
<td>Leu Leu Gln Arg Trp Leu Val Glu Ser Ala Arg Trp Leu Ile Ile Thr</td>
<td>275</td>
<td>280</td>
<td>285</td>
<td></td>
</tr>
<tr>
<td>Asn Lys Leu Asp Glu Gly Leu Lys Ala Leu Arg Lys Val Ala Arg Thr</td>
<td>290</td>
<td>295</td>
<td>300</td>
<td></td>
</tr>
<tr>
<td>Asn Gly Ile Lys Asn Ala Glu Thr Leu Asn Ile Gly Val Val Arg</td>
<td>305</td>
<td>310</td>
<td>315</td>
<td>320</td>
</tr>
<tr>
<td>Ser Thr Met Gln Glu Glu Leu Asp Ala Ala Gin Thr Lys Thr Thr Val</td>
<td>325</td>
<td>330</td>
<td>335</td>
<td></td>
</tr>
<tr>
<td>Cys Asp Leu Phe Arg Asn Pro Ser Met Arg Lys Arg Ile Cys Ile Leu</td>
<td>340</td>
<td>345</td>
<td>350</td>
<td></td>
</tr>
<tr>
<td>Val Phe Leu Arg Phe Ala Asn Thr Ile Pro Phe Tyr Gly Thr Met Val</td>
<td>355</td>
<td>360</td>
<td>365</td>
<td></td>
</tr>
<tr>
<td>Asn Leu Gln His Val Gly Ser Asn Ile Phe Leu Leu Gin Val Leu Tyr</td>
<td>370</td>
<td>375</td>
<td>380</td>
<td></td>
</tr>
<tr>
<td>Gly Ala Val Ala Leu Ile Val Arg Cys Leu Ala Leu Leu Thr Leu Asn</td>
<td>385</td>
<td>390</td>
<td>395</td>
<td>400</td>
</tr>
<tr>
<td>His Met Gly Arg Arg Ile Ser Gin Ile Leu Phe Met Phe Leu Val Gly</td>
<td>405</td>
<td>410</td>
<td>415</td>
<td></td>
</tr>
<tr>
<td>Leu Ser Ile Leu Ala Asn Thr Phe Val Pro Lys Glu Met Gin Thr Leu</td>
<td>420</td>
<td>425</td>
<td>430</td>
<td></td>
</tr>
<tr>
<td>Arg Val Ala Leu Ala Cys Leu Gly Ile Gly Cys Ser Ala Ala Thr Phe</td>
<td>435</td>
<td>440</td>
<td>445</td>
<td></td>
</tr>
<tr>
<td>Ser Ser Val Ala Val His Phe Ile Glu Leu Ile Pro Thr Val Leu Arg</td>
<td>450</td>
<td>455</td>
<td>460</td>
<td></td>
</tr>
<tr>
<td>Ala Arg Ala Ser Gly Ile Asp Leu Thr Ala Ser Arg Ile Gly Ala Ala</td>
<td>465</td>
<td>470</td>
<td>475</td>
<td>480</td>
</tr>
<tr>
<td>Leu Pro Leu Leu Met Thr Leu Thr Val Phe Phe Thr Thr Leu Pro Thr</td>
<td>485</td>
<td>490</td>
<td>495</td>
<td></td>
</tr>
</tbody>
</table>
Ile Ile Tyr Gly Ile Phe Pro Ile Ile Gly Gly Leu Ile Val Phe Leu
500 505 510

Leu Pro Glu Thr Lys Asn Leu Pro Leu Pro Asp Thr Ile Lys Asp Val
515 520 525

Glu Asn Gln Lys Lys Asn Leu Lys Glu Lys Ala
530 535

<210> SEQ ID NO 10
<211> LENGTH: 550
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 10

Met Ala Phe Ser Lys Leu Leu Glu Gln Ala Gly Gly Val Gly Leu Phe
1  5  10  15

Gln Thr Leu Gln Val Leu Thr Phe Ile Leu Pro Cys Leu Met Ile Pro
20 25 30

Ser Gln Met Leu Leu Glu Asn Phe Ser Ala Ala Ile Pro Gly His Arg
35 40 45

Cys Thr Thr His Met Leu Asp Asn Gly Ser Ala Val Ser Thr Asn Met
50 55 60

Thr Pro Lys Ala Leu Leu Thr Ile Ser Ile Pro Gly Pro Gly Asn Gln
65 70 75 80

Gly Pro His Gln Cys Arg Arg Phe Arg Gln Pro Gln Trp Gln Leu Leu
85 90 95

Asp Pro Asn Ala Thr Ala Thr Ser Trp Ser Glu Ala Asp Thr Glu Pro
100 105 110

Cys Val Asp Gly Trp Val Tyr Asp Arg Ser Val Phe Thr Ser Thr Ile
115 120 125

Val Ala Lys Trp Asp Leu Val Cys Ser Ser Glu Gly Leu Lys Pro Leu
130 135 140

Ser Gln Ser Ile Phe Met Ser Gly Ile Leu Val Gly Ser Phe Ile Trp
145 150 155 160

Gly Leu Leu Ser Tyr Arg Phe Gly Arg Lys Pro Met Leu Ser Trp Cys
165 170 175

Cys Leu Gln Leu Ala Val Ala Gly Thr Ser Thr Ile Phe Ala Pro Thr
180 185 190

Phe Val Ile Tyr Cys Gly Leu Arg Phe Val Ala Ala Phe Gly Met Ala
195 200 205

Gly Ile Phe Leu Ser Ser Leu Thr Leu Met Val Glu Trp Thr Thr Thr
210 215 220

Ser Arg Arg Ala Val Thr Met Thr Val Gly Cys Ala Phe Ser Ala
225 230 235 240

Gly Gln Ala Ala Leu Gly Gly Leu Ala Phe Ala Leu Arg Asp Thr Arg
245 250 255

Thr Leu Gln Leu Ala Ser Val Pro Phe Phe Ala Ile Ser Leu Ile
260 265 270

Ser Thr Thr Leu Pro Glu Ser Ala Arg Trp Leu Ile Ille Lys Gly Lys
275 280 285

Pro Asp Gln Ala Leu Gln Glu Leu Arg Lys Ala Arg Ile Asn Gly
290 295 300

His Lys Glu Ala Lys Arg Leu Thr Ile Glu Val Leu Met Ser Ser Val
305 310 315 320

Lys Glu Glu Val Ala Ser Ala Lys Gly Glu Pro Arg Ser Val Leu Asp Leu
Phe Cys Val Pro Val Leu Arg Trp Arg Ser Cys Ala Met Leu Val Val 
340 345 350
Asn Phe Ser Leu Leu Ile Ser Tyr Tyr Gly Leu Val Phe Asp Leu Gln 
355 360 365
Ser Leu Gly Arg Asp Ile Phe Leu Leu Gln Ala Leu Phe Gly Ala Val 
370 375 380
Arg Phe Leu Gly Arg Ala Thr Thr Ala Leu Leu Leu Ser Phe Leu Gly 
395 390 395 400
Arg Arg Thr Ile Gln Ala Gly Ser Gln Ala Met Ala Gly Leu Ala Ile 
405 410 415
Leu Ala Asn Met Leu Val Pro Gln Asp Leu Gln Thr Leu Arg Val Val 
420 425 430
Phe Ala Val Leu Gly Lys Gly Cys Phe Gly Ile Ser Leu Thr Cys Leu 
435 440 445
Thr Ile Tyr Lys Ala Glu Leu Phe Pro Thr Pro Val Arg Met Thr Ala 
450 455 460
Asp Gly Ile Leu His Thr Val Gly Arg Leu Gly Ala Met Met Gly Pro 
465 470 475 480
Leu Ile Leu Met Ser Arg Gln Ala Leu Pro Leu Leu Pro Pro Leu Leu 
485 490 495
Tyr Gly Val Ile Ser Ile Ala Ser Ser Leu Val Leu Phe Phe Leu 
500 505 510
Pro Glu Thr Gln Gly Leu Pro Leu Pro Asp Thr Ile Gln Asp Leu Glu 
515 520 525
Ser Gln Lys Ser Thr Ala Ala Gln Gly Asn Arg Gln Glu Ala Val Thr 
530 535 540
Val Glu Ser Thr Ser Leu 
545 550

<210> SEQ ID NO 11
<211> LENGTH: 542
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11
Met Thr Phe Ser Glu Ile Leu Asp Arg Val Gly Ser Met Gly His Phe 
2 3 4 5 6 7 8 9 10 11
Gln Phe Leu His Val Ala Ile Leu Gly Leu Pro Ile Leu Asn Met Ala 
12 13 14 15 16 17 18 19 20 21
Asn His Asn Leu Leu Gln Ile Phe Thr Ala Ala Thr Pro Val His His 
22 23 24 25 26 27 28 29 30 31
Cys Arg Pro Pro His Asn Ala Ser Thr Gly Pro Trp Val Leu Pro Met 
32 33 34 35 36 37 38 39 40 41
Gly Pro Asn Gly Lys Pro Glu Arg Cys Leu Arg Phe Val His Pro Pro 
42 43 44 45 46 47 48 49 50 51
Asn Ala Ser Leu Pro Asn Asp Thr Glu Arg Ala Met Gln Pro Cys Leu 
52 53 54 55 56 57 58 59 60 61
Asp Gly Trp Val Tyr Asn Ser Thr Lys Asp Ser Ile Val Thr Glu Trp 
62 63 64 65 66 67 68 69 70 71
Asp Leu Val Cys Asn Ser Asn Lys Leu Lys Met Ala Gln Ser Ile 
72 73 74 75 76 77 78 79 80 81
Phe Met Ala Gly Ile Leu Ile Gly Gly Leu Val Leu Gly Asp Leu Ser 
82 83 84 85 86 87 88 89 90 91

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp Arg Phe Gly Arg Arg Arg Ile Leu Thr Cys Ser Tyr Leu Leu Leu</td>
<td>145</td>
</tr>
<tr>
<td>Ala Ala Ser Gly Ser Gly Ala Phe Ser Pro Thr Phe Pro Ile Tyr</td>
<td>164</td>
</tr>
<tr>
<td>Met Val Phe Arg Phe Leu Cys Gly Phe Gly Ile Ser Gly Ile Thr Leu</td>
<td>180</td>
</tr>
<tr>
<td>Ser Thr Val Ile Leu Asn Val Glu Trp Val Pro Thr Arg Met Arg Ala</td>
<td>195</td>
</tr>
<tr>
<td>Ile Met Ser Thr Ala Leu Gly Tyr Cys Tyr Thr Phe Gly Gln Phe Ile</td>
<td>210</td>
</tr>
<tr>
<td>Leu Pro Gly Leu Ala Tyr Ala Ile Pro Gln Trp Arg Trp Leu Gln Leu</td>
<td>225</td>
</tr>
<tr>
<td>Thr Val Ser Ile Pro Phe Phe Val Phe Leu Ser Ser Trp Trp Thr</td>
<td>245</td>
</tr>
<tr>
<td>Pro Glu Ser Ile Arg Trp Leu Val Leu Ser Gly Lys Ser Ser Glu Ala</td>
<td>260</td>
</tr>
<tr>
<td>Leu Lys Ile Leu Arg Arg Val Ala Val Phe Asn Gly Lys Lys Glu Glu</td>
<td>275</td>
</tr>
<tr>
<td>Gly Glu Arg Leu Ser Leu Glu Glu Leu Lys Leu Arg Leu Gln Lys Glu</td>
<td>290</td>
</tr>
<tr>
<td>Ile Ser Leu Ala Lys Ala Lys Tyr Thr Ala Ser Arg Leu Phe Arg Ile</td>
<td>305</td>
</tr>
<tr>
<td>Pro Met Leu Arg Arg Met Thr Phe Cys Leu Ser Leu Ala Trp Phe Ala</td>
<td>325</td>
</tr>
<tr>
<td>Thr Gly Phe Ala Tyr Tyr Ser Leu Ala Met Gly Val Glu Glu Phe Gly</td>
<td>340</td>
</tr>
<tr>
<td>Val Asn Leu Tyr Ile Leu Gln Ile Ile Phe Gly Gly Val Asp Val Pro</td>
<td>355</td>
</tr>
<tr>
<td>Ala Lys Phe Ile Thr Ile Leu Ser Leu Ser Tyr Leu Gly Arg His Thr</td>
<td>370</td>
</tr>
<tr>
<td>Thr Gln Ala Ala Leu Leu Ala Gly Gly Ala Ile Leu Ala Leu</td>
<td>385</td>
</tr>
<tr>
<td>Thr Phe Val Pro Leu Asp Leu Gln Thr Val Arg Thr Val Leu Ala Val</td>
<td>405</td>
</tr>
<tr>
<td>Phe Gly Lys Gly Cys Leu Ser Ser Ser Phe Ser Cys Leu Phe Leu Tyr</td>
<td>420</td>
</tr>
<tr>
<td>Thr Ser Glu Leu Tyr Pro Thr Val Ile Arg Glu Thr Gly Met Gly Val</td>
<td>435</td>
</tr>
<tr>
<td>Ser Asn Leu Trp Thr Arg Val Gly Ser Met Val Ser Pro Leu Val Lys</td>
<td>450</td>
</tr>
<tr>
<td>Ile Thr Gly Glu Val Gln Pro Phe Ile Pro Asn Ile Ile Tyr Gly Ile</td>
<td>465</td>
</tr>
<tr>
<td>Thr Ala Leu Gly Gly Ser Ala Ala Leu Phe Leu Pro Glu Thr Leu</td>
<td>485</td>
</tr>
<tr>
<td>Asn Gln Pro Leu Pro Glu Thr Ile Glu Asp Leu Glu Asn Trp Ser Leu</td>
<td>500</td>
</tr>
<tr>
<td>Arg Ala Lys Lys Gly Pro Lys Gln Gly Pro Gly Val Glu Lys Ala Ser Gln</td>
<td>515</td>
</tr>
<tr>
<td>Arg Ile Pro Leu Gln Pro His Gly Pro Gly Leu Gly Ser Ser</td>
<td>530</td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 12
<211> LENGTH: 550
<212> TYPE: PRT
<table>
<thead>
<tr>
<th>Residue</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Met Ala Phe Asn Asp Leu Leu Glu Gln Val Gly Gly Val Gly Arg Phe</td>
</tr>
<tr>
<td>5</td>
<td>Gln Gln Ile Gln Val Thr Leu Val Val Leu Pro Leu Leu Leu Met Ala</td>
</tr>
<tr>
<td>10</td>
<td>Ser His Asn Thr Leu Gln Asn Phe Thr Ala Ala Ile Pro Thr His</td>
</tr>
<tr>
<td>15</td>
<td>Cys Arg Pro Pro Ala Asp Ala Asn Leu Ser Lys Asn Gly Gly Leu Glu</td>
</tr>
<tr>
<td>20</td>
<td>Val Trp Leu Pro Arg Asp Arg Glu Gly Glu Pro Gly Ser Cys Leu Arg</td>
</tr>
<tr>
<td>25</td>
<td>Phe Thr Ser Pro Gln Trp Gly Leu Pro Phe Leu Asn Gly Thr Glu Ala</td>
</tr>
<tr>
<td>30</td>
<td>Asn Gly Thr Gly Ala Thr Glu Pro Cys Thr Asp Gly Trp Ile Tyr Asp</td>
</tr>
<tr>
<td>35</td>
<td>Asn Ser Thr Phe Pro Ser Thr Ile Val Thr Glu Trp Asp Leu Val Cys</td>
</tr>
<tr>
<td>40</td>
<td>Ser His Arg Ala Leu Arg Gln Leu Ala Gln Ser Leu Tyr Met Val Gly</td>
</tr>
<tr>
<td>45</td>
<td>Val Leu Leu Gly Ala Met Val Phe Gly Tyr Leu Ala Asp Arg Leu Gly</td>
</tr>
<tr>
<td>50</td>
<td>Arg Arg Lys Val Leu Ile Leu Asn Tyr Leu Gln Thr Ala Val Ser Gly</td>
</tr>
<tr>
<td>55</td>
<td>Thr Cys Ala Ala Phe Ala Pro Asn Phe Pro Ile Tyr Cys Ala Phe Arg</td>
</tr>
<tr>
<td>60</td>
<td>Leu Leu Ser Gly Met Ala Leu Ala Gly Ile Ser Leu Asn Cys Met Thr</td>
</tr>
<tr>
<td>65</td>
<td>Leu Asn Val Glu Trp Met Pro Ile His Thr Arg Ala Cys Val Gly Thr</td>
</tr>
<tr>
<td>70</td>
<td>Leu Ile Gly Tyr Val Tyr Ser Leu Gly Glu Phe Leu Leu Ala Gly Val</td>
</tr>
<tr>
<td>75</td>
<td>Ala Tyr Ala Val Pro His Trp Arg His Leu Gin Leu Leu Val Ser Ala</td>
</tr>
<tr>
<td>80</td>
<td>Pro Phe Phe Ala Phe Ile Tyr Ser Trp Phe Ile Glu Ser Ala</td>
</tr>
<tr>
<td>85</td>
<td>Arg Trp His Ser Ser Ser Gly Arg Leu Asp Leu Thr Leu Arg Ala Leu</td>
</tr>
<tr>
<td>90</td>
<td>Gln Gin Val Ala Arg Ile Asn Gly Lys Arg Glu Glu Gly Ala Lys Leu</td>
</tr>
<tr>
<td>95</td>
<td>Ser Met Gin Val Leu Arg Ala Ser Leu Gin Lys Glu Leu Thr Met Gly</td>
</tr>
<tr>
<td>100</td>
<td>Lys Gin Ala Ser Ala Met Gin Leu Leu Arg Cys Pro Thr Leu Arg</td>
</tr>
<tr>
<td>105</td>
<td>His Leu Phe Leu Cys Leu Ser Met Leu Trp Phe Ala Thr Ser Phe Ala</td>
</tr>
<tr>
<td>110</td>
<td>Tyr Tyr Gly Leu Val Met Asp Leu Gin Gly Phe Gly Val Ser Ile Tyr</td>
</tr>
<tr>
<td>115</td>
<td>Leu Ile Gin Val Ile Phe Gly Ala Val Asp Leu Pro Ala Lys Leu Val</td>
</tr>
<tr>
<td>120</td>
<td>Gly Phe Leu Val Ile Asn Ser Leu Gly Arg Arg Pro Ala Gin Met Ala</td>
</tr>
</tbody>
</table>
What is claimed is:

1. A method of screening candidate substrates of the organic cation transporter 6 (OCT6) comprising:
   a. providing a test agent;
   b. providing mammalian cells or a mammalian cell line which express OCT6;
   c. incubating the test agent with the cells or cell line; and
   d. determining whether the test agent is a substrate for OCT6, wherein the mammalian cells or mammalian cell line provided in step b are leukemia cells or a leukemia cell line, respectively.

2. The method of claim 1 wherein the test agent is coupled to a detectable substance.
3. The method of claim 2 wherein the detectable substance is selected from the group consisting of extrinsically activatable enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, radioactive materials, positron emitting metals using various positron emission tomographies, nonradioactive paramagnetic metal ions, immunogenic tag peptide sequences, extrinsically activatable toxins, extrinsically activatable quenching agents, and antibodies.
4. The method of claim 1 wherein the step of determining whether the test agent is a substrate for OCT6 comprises analyzing whether the test agent is located intracellularly.
5. The method of claim 1, wherein step (d) comprises determining the viability of the cells or cell line.
6. The method of claim 5, wherein the viability of the cells or cell line is determined by applying a dye to the cells or cell line, wherein incorporation of the dye by the cells is indicative of death of the cells or cell line.
7. The method of claim 6, wherein the dye is trypan blue.

* * * * *