

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2002247939 B2**

(54) Title
Treatment of CNS disorders using D-amino acid oxidase and D-aspartate oxidase antagonists

(51) International Patent Classification(s)
C12Q 1/26 (2006.01) **A61K 31/60** (2006.01)
A61K 31/00 (2006.01) **A61K 31/7024** (2006.01)
A61K 31/04 (2006.01) **A61K 33/02** (2006.01)
A61K 31/145 (2006.01) **A61P 25/00** (2006.01)
A61K 31/155 (2006.01) **A61P 25/16** (2006.01)
A61K 31/17 (2006.01) **A61P 25/18** (2006.01)
A61K 31/175 (2006.01) **A61P 25/22** (2006.01)
A61K 31/19 (2006.01) **A61P 25/24** (2006.01)
A61K 31/192 (2006.01) **A61P 25/28** (2006.01)
A61K 31/194 (2006.01) **C07D 209/12** (2006.01)
A61K 31/195 (2006.01) **C07D 209/18** (2006.01)
A61K 31/196 (2006.01) **C07D 209/42** (2006.01)
A61K 31/198 (2006.01) **C07D 213/79** (2006.01)
A61K 31/22 (2006.01) **C07D 213/80** (2006.01)
A61K 31/221 (2006.01) **C07D 233/84** (2006.01)
A61K 31/223 (2006.01) **C07D 233/88** (2006.01)
A61K 31/245 (2006.01) **C07D 239/40** (2006.01)
A61K 31/255 (2006.01) **C07D 257/04** (2006.01)
A61K 31/351 (2006.01) **C07D 309/36** (2006.01)
A61K 31/404 (2006.01) **C07H 11/04** (2006.01)
A61K 31/405 (2006.01) **C07J 7/00** (2006.01)
A61K 31/41 (2006.01) **C12N 9/06** (2006.01)
A61K 31/4164 (2006.01) **C12N 15/09** (2006.01)
A61K 31/426 (2006.01) **C12Q 1/68** (2006.01)
A61K 31/4402 (2006.01) **G01N 33/15** (2006.01)
A61K 31/4425 (2006.01) **G01N 33/50** (2006.01)
A61K 31/455 (2006.01) **G01N 33/53** (2006.01)
A61K 31/505 (2006.01) **G01N 33/566** (2006.01)
A61K 31/57 (2006.01) **G01N 33/68** (2006.01)

(21) Application No: **2002247939** (22) Date of Filing: **2002.01.15**

(87) WIPO No: **WO02/066672**

(30) Priority Data

(31) Number	(32) Date	(33) Country
60/333,881	2001.11.19	US
60/305,445	2001.07.13	US
60/261,883	2001.01.16	US
60/345,211	2001.10.22	US

(43) Publication Date: **2002.09.04**

(43) Publication Journal Date: **2003.02.27**

(44) Accepted Journal Date: **2007.01.25**

(71) Applicant(s)
Serono Genetics Institute S.A.

(72) Inventor(s)
Chumakov, Ilya;Cohen, Daniel

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
29 August 2002 (29.08.2002)

PCT

(10) International Publication Number
WO 02/066672 A2

- (51) International Patent Classification⁷: **C12Q 1/00**
- (21) International Application Number: PCT/IB02/01262
- (22) International Filing Date: 15 January 2002 (15.01.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/261,883 16 January 2001 (16.01.2001) US
60/305,445 13 July 2001 (13.07.2001) US
60/345,211 22 October 2001 (22.10.2001) US
60/333,881 19 November 2001 (19.11.2001) US
- (71) Applicant (for all designated States except US): **GENSET**
[FR/FR]; Intellectual Property Department, 24, rue Royale,
F-75008 Paris (FR).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **COHEN, Daniel**
[FR/FR]; 1, boulevard Richard Wallace, F-92220
Neuilly-sur-Seine (FR). **CHUMAKOV, Ilya** [FR/FR];
196, rue des Chevreuilleilles, F-77000 Vaux-le-Penil (FR).
- (74) Common Representative: **GENSET**; Intellectual Prop-
erty Department, 24, rue Royale, F-75008 Paris (FR).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GI, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

Published:

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

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



WO 02/066672 A2

(54) Title: TREATMENT OF CNS DISORDERS USING D-AMINO ACID OXIDASE AND D-ASPARTATE OXIDASE ANTAG-
ONISTS

(57) Abstract: Compounds that are antagonists of D-amino acid oxidase and D-aspartate oxidase, methods of treating CNS disorders
including bipolar disorder, psychosis and schizophrenia using the compounds, and pharmaceutically acceptable compositions that
contain the antagonists are disclosed.

Treatment of CNS Disorders Using D-Amino Acid Oxidase and D-Aspartate Oxidase Antagonists**FIELD OF THE INVENTION**

This invention provides means to identify compounds useful in the treatment of CNS-related disorders such as schizophrenia, bipolar disorder, depression and other mood disorders, means to determine the predisposition of individuals to said disorders, as well as means for the disease diagnosis and prognosis of said disorders. More specifically, this invention relates to means of treating said disorders using antagonists of D-amino acid oxidase (DAO) and D-aspartate oxidase (DDO).

BACKGROUND

Advances in the technological armamentarium available to basic and clinical investigators have enabled increasingly sophisticated studies of brain and nervous system function in health and disease. Numerous hypotheses both neurobiological and pharmacological have been advanced with respect to the neurochemical and genetic mechanisms involved in central nervous system (CNS) disorders, including psychiatric disorders and neurodegenerative diseases. However, CNS disorders have complex and poorly understood etiologies, as well as symptoms that are overlapping, poorly characterized, and difficult to measure. As a result future treatment regimes and drug development efforts will be required to be more sophisticated and focused on multigenic causes, and will need new assays to segment disease populations, and provide more accurate diagnostic and prognostic information on patients suffering from CNS disorders.

Neurological Basis of CNS Disorders

Neurotransmitters serve as signal transmitters throughout the body. Diseases that affect neurotransmission can therefore have serious consequences. For example, for over 30 years the leading theory to explain the biological basis of many psychiatric disorders such as depression has been the monoamine hypothesis. This theory proposes that depression is partially due to a deficiency in one of the three main biogenic monoamines, namely dopamine, norepinephrine and/or serotonin.

In addition to the monoamine hypothesis, numerous arguments tend to show the value in taking into account the overall function of the brain and no longer only considering a single neuronal system. In this context, the value of dual specific actions on the central aminergic systems including second and third messenger systems has now emerged.

Endocrine Basis of CNS Disorders

It is furthermore apparent that the main monoamine systems, namely dopamine, norepinephrine and serotonin, do not completely explain the pathophysiology of many CNS disorders. In particular, it is clear that CNS disorders may have an endocrine component; the hypothalamic-pituitary-adrenal

(HPA) axis, including the effects of corticotrophin-releasing factor and glucocorticoids, plays an important role in the pathophysiology of CNS disorders.

In the hypothalamus-pituitary-adrenal (HPA) axis, the hypothalamus lies at the top of the hierarchy regulating hormone secretion. It manufactures and releases peptides (small chains of amino acids) that act on the pituitary, at the base of the brain, stimulating or inhibiting the pituitary's release of various hormones into the blood. These hormones, among them growth hormone, thyroid-stimulating hormone and adrenocorticotrophic hormone (ACTH), control the release of other hormones from target glands. In addition to functioning outside the nervous system, the hormones released in response to pituitary hormones also feed back to the pituitary and hypothalamus. There they deliver inhibitory signals that serve to limit excess hormone biosynthesis.

CNS Disorders

Neurotransmitter and hormonal abnormalities are implicated in disorders of movement (e.g. Parkinson's disease, Huntington's disease, motor neuron disease, etc.), disorders of mood (e.g. unipolar depression, bipolar disorder, anxiety, etc.) and diseases involving the intellect (e.g. Alzheimer's disease, Lewy body dementia, schizophrenia, etc.). In addition, these systems have been implicated in many other disorders, such as coma, head injury, cerebral infarction, epilepsy, alcoholism and the mental retardation states of metabolic origin seen particularly in childhood.

CNS disorders can encompass a wide range of disorders, and a correspondingly wide range of genetic factors. Examples of CNS disorders include neurodegenerative disorders, psychotic disorders, mood disorders, autism, substance dependence and alcoholism, mental retardation, and other psychiatric diseases including cognitive, anxiety, eating, impulse-control, and personality disorders. Disorders can be defined using the Diagnosis and Statistical Manual of Mental Disorders fourth edition (DSM-IV) classification.

Even when considering just a small subset of CNS disorders, it is evident from the lack of adequate treatment for and understanding of the molecular basis of the psychotic disorders schizophrenia and bipolar disorder that new targets for therapeutic invention and improved methods of treatment are needed. For both schizophrenia and bipolar disorder, all the known molecules used for the treatment have side effects and act only against the symptoms of the disease. There is a strong need for new molecules without associated side effects and directed against targets which are involved in the causal mechanisms of schizophrenia and bipolar disorder. Therefore, tools facilitating the discovery and characterization of these targets are necessary and useful.

The aggregation of schizophrenia and bipolar disorder in families, the evidence from twin and adoption studies, and the lack of variation in incidence worldwide, indicate that schizophrenia and bipolar disorder are primarily genetic conditions, although environmental risk factors are also involved at some level as necessary, sufficient, or interactive causes. For example, schizophrenia occurs in 1% of

the general population. But, if there is one grandparent with schizophrenia, the risk of getting the illness increases to about 3%; one parent with Schizophrenia, to about 10%. When both parents have schizophrenia, the risk rises to approximately 40%.

Identification of schizophrenia susceptibility gene on chromosome 13q31-q33.

The identification of genes involved in a particular trait such as a specific central nervous system disorder, like schizophrenia, can be carried out through two main strategies currently used for genetic mapping: linkage analysis and association studies. Linkage analysis requires the study of families with multiple affected individuals and is now useful in the detection of mono- or oligogenic inherited traits. Conversely, association studies examine the frequency of marker alleles in unrelated trait (T+) individuals compared with trait negative (T-) controls, and are generally employed in the detection of polygenic inheritance.

Genetic link or “linkage” is based on an analysis of which of two neighboring sequences on a chromosome contains the least recombinations by crossing-over during meiosis. To do this, chromosomal markers, like microsatellite markers, have been localized with precision on the genome. Genetic link analysis calculates the probabilities of recombinations on the target gene with the chromosomal markers used, according to the genealogical tree, the transmission of the disease, and the transmission of the markers. Thus, if a particular allele of a given marker is transmitted with the disease more often than chance would have it (recombination level between 0 and 0.5), it is possible to deduce that the target gene in question is found in the neighborhood of the marker. Using this technique, it has been possible to localize several genes demonstrating a genetic predisposition of familial cancers. In order to be able to be included in a genetic link study, the families affected by a hereditary form of the disease must satisfy the “informativeness” criteria: several affected subjects (and whose constitutional DNA is available) per generation, and at best having a large number of siblings.

Results of previous linkage studies supported the hypothesis that chromosome 13 was likely to harbor a schizophrenia susceptibility locus on 13q32 (Blouin JL et al., 1998, Nature Genetics, 20:70-73; Lin MW et al., 1997, Hum. Genet., 99(3):417-420). These observations suggesting the presence of a schizophrenia locus on the chromosome 13q32 locus had been obtained by carrying out linkage studies. Linkage analysis had been successfully applied to map simple genetic traits that show clear Mendelian inheritance patterns and which have a high penetrance, but this method suffers from a variety of drawbacks. First, linkage analysis is limited by its reliance on the choice of a genetic model suitable for each studied trait. Furthermore, the resolution attainable using linkage analysis is limited, and complementary studies are required to refine the analysis of the typical 20 Mb regions initially identified through this method. In addition, linkage analysis has proven difficult when applied to complex genetic traits, such as those due to the combined action of multiple genes and/or environmental factors. In such cases, too great an effort and cost are needed to recruit the adequate number of affected families

required for applying linkage analysis to these situations. Finally, linkage analysis cannot be applied to the study of traits for which no large informative families are available.

Novel schizophrenia gene: g34872 (sbg1).

More recently, instead of using linkage studies, a novel schizophrenia and bipolar disorder related gene referred to as the g34872 gene located on the chromosome 13q31-q33 locus was identified using an alternative method of conducting association studies. This alternative method involved generating biallelic markers (primarily single nucleotide polymorphisms (SNPs)) in the region of interest, identifying markers in linkage disequilibrium with schizophrenia, and conducting association studies in unrelated schizophrenia and bipolar disorder case and control populations.

In summary, a BAC contig covering the candidate genomic region was constructed using 27 public sequence-tagged site (STS) markers localised on chromosome 13 in the region of 13q31-q33 to screen a 7 genome equivalent proprietary BAC library. From these materials, new STSs were generated allowing construction of a dense physical map of the region. In total, 275 STSs allowed identification of 255 BACs that were all sized and mapped by in situ chromosomal hybridisation for verification. New biallelic markers were generated by partial sequencing of insert ends from subclones of some of the BAC inserts localized to the human chromosome 13q31-q33 region. In a first phase of the analysis, a first set of 34 biallelic markers on 9 different BACs across the chromosome 13q31-q33 candidate locus were analysed in schizophrenic cases and controls, thereby identifying a subregion showing an association with schizophrenia. Following this first analysis, further biallelic markers were generated as described above in order to provide a very high density map of the target region. A minimal set of 35 BACs was identified and fully sequenced which resulted in several contigs including a contig of over 900kb comprising sequences of the target region.

These biallelic markers were used in association studies in order to refine a particular subregion of interest, which contained a candidate schizophrenia gene, g34872. The biallelic markers were genotyped in several studies carried out in different populations to confirm the association with the subregion. Association studies were first performed on two different screening samples of schizophrenia cases and controls from a French Canadian population comprising 139 cases and 141 controls, and 215 cases and 241 controls, respectively, as well on bipolar disorder cases and controls from an Argentinian population. The results obtained after several studies using this population indicated a genomic region of about 150kb showing a significant association with schizophrenia. This association was then confirmed in separate studies using cases and controls from a U.S. schizophrenia population, as well as in further samples from the Argentinian and French Canadian populations.

The approximately 150kb genomic region associated with schizophrenia was found to contain the candidate gene g34872. In addition to characterizing the intron-exon structure of the g34872 gene, a range of mRNA splicing variants including tissue specific mRNA splicing variants were identified, and

the existence of the mRNA was demonstrated. Subsequently, a peptide fragment derived from the g34872 polypeptide product, the amino acid sequence of which is shown in SEQ ID No 5, demonstrated a decrease in locomotor movement frequency, and an increase in stereotypy when injected intraperitoneally in mice. Further discussion of the identification of the g34872 gene is provided in co-pending U.S. patent application ser. no. 09/539,333 titled "Schizophrenia associated genes, proteins and biallelic markers" and co-pending International Patent Application No. PCT/IB00/00435, both filed 30 March 2000 and which disclosures are hereby incorporated by reference in their entireties.

g34872 interacting proteins and schizophrenia.

There is a strong need to identify genes involved in schizophrenia and bipolar disorder. There is also a need to identify genes involved in the g34872 pathway and genes whose products functionally interact with the g34872 gene products. These genes may provide new intervention points in the treatment of schizophrenia or bipolar disorder and allow further study and characterization of the g34872 gene and related biological pathway. The knowledge of these genes and the related biological pathways involved in schizophrenia will allow researchers to understand the etiology of schizophrenia and bipolar disorder and will lead to drugs and medications which are directed against the cause of the diseases. There is also a great need for new methods for detecting a susceptibility to schizophrenia and bipolar disorder, as well as for preventing or following up the development of the disease. Diagnostic tools could also prove extremely useful. Indeed, early identification of subjects at risk of developing schizophrenia would enable early and/or prophylactic treatment to be administered. Moreover, accurate assessments of the eventual efficacy of a medicament as well as the patient's eventual tolerance to it may enable clinicians to enhance the benefit/risk ratio of schizophrenia and bipolar disorder treatment regimes.

The present invention thus relates to any gene encoding for proteins which interact with g34872 polypeptides, herein referred to as g34872 binding partners. By yeast 2-hybrid technology, the inventors have cloned several g34872 binding partners. The inventors demonstrate that D-amino acid oxidase is included in the group of said g34872 binding partners. Knowledge of g34872 binding partner permits the development of medicaments for the treatment of CNS disease mediated by genes selected from the group comprising g34872, D-amino acid oxidase and any other g34872 binding partners. Furthermore, knowledge of g34872 binding partners provides a means for the detection of g34872, g34872-binding partners, g34872-binding partners complexes or interactions between g34872 and its binding partners.

g34872 interacting proteins and schizophrenia: D-amino acid oxidase.

D-Amino acid oxidase (DAO) was one of the first enzymes to be described and the second flavoprotein to be discovered in the mid 1930s. DAO converts D-amino acids into the corresponding .alpha.-keto acids. It does this by catalyzing the dehydrogenation of D-amino acids to their imino

counterparts and a reduced flavin-product complex. The reduced flavin is then (re)oxidized by dioxygen to yield FADox and H₂O₂, whereas the imino acid spontaneously hydrolyzes to the keto acid and NH₄⁺. Although DAO is present in most organisms and mammalian tissues, its physiological role in vertebrates has been unclear. DAO oxidizes: D-Met, D-Pro, D-Phe, D-Tyr, D-Ile, D-Leu, D-Ala and D-Val. D-Ser, D-Arg, D-His, D-norleucine and D-Trp are oxidized at a low rate. D-Ornithine, cis-4-hydroxy-D-proline, D-Thr, D-Trp-methyl ester, N-acetyl-D-Ala and D-Lys are oxidized at a very low rate. D-Asp, D-Glu and their derivatives, Gly and all the L-amino acids are not oxidized (or are at a rate which is undetectable). D-Aspartate oxidase (DDO) oxidizes only D-Asp, D-Glu and their following derivatives: D-Asn, D-Gln, D-Asp-dimethyl-ester and N-methyl-D-Asp.

CNS disorders are a type of neurological disorder. CNS disorders can be drug induced; can be attributed to genetic predisposition, infection or trauma; or can be of unknown etiology. CNS disorders comprise neuropsychiatric disorders, neurological diseases and mental illnesses; and include neurodegenerative diseases, behavioral disorders, cognitive disorders and cognitive affective disorders. There are several CNS disorders whose clinical manifestations have been attributed to CNS dysfunction (i.e., disorders resulting from inappropriate levels of neurotransmitter release, inappropriate properties of neurotransmitter receptors, and/or inappropriate interaction between neurotransmitters and neurotransmitter receptors). Several CNS disorders can be attributed to a cholinergic deficiency, a dopaminergic deficiency, an adrenergic deficiency and/or a serotonergic deficiency. CNS disorders of relatively common occurrence include presenile dementia (early onset Alzheimer's disease), senile dementia (dementia of the Alzheimer's type), Parkinsonism including Parkinson's disease, Huntington's chorea, tardive dyskinesia, hyperkinesia, mania, attention deficit disorder, anxiety, dyslexia, schizophrenia, psychosis, bipolar disorder, depression and Tourette's syndrome.

Neurotransmitter and hormonal abnormalities are implicated in disorders of movement (e.g. Parkinson's disease, Huntington's disease, motor neuron disease, etc.), disorders of mood (e.g. unipolar depression, bipolar disorder, anxiety, etc.) and diseases involving the intellect (e.g. Alzheimer's disease, Lewy body dementia, schizophrenia, etc.). In addition, neurotransmitter and hormonal abnormalities have been implicated in a wide range of disorders, such as coma, head injury, cerebral infarction, epilepsy, alcoholism and the mental retardation states of metabolic origin seen particularly in childhood.

Schizophrenia

In developed countries schizophrenia occurs in approximately one per cent of the adult population at some point during their lives. There are an estimated 45 million people with schizophrenia in the world, with more than 33 million of them in the developing countries. Moreover, schizophrenia accounts for a fourth of all mental health costs and takes up one in three psychiatric hospital beds. Most schizophrenia patients are never able to work. The cost of schizophrenia to society is enormous. In the United States, for example, the direct cost of treatment of schizophrenia has been estimated to be close

to 0.5% of the gross national product. Standardized mortality ratios (SMRs) for schizophrenic patients are estimated to be two to four times higher than the general population and their life expectancy overall is 20 % shorter than for the general population.

The most common cause of death among schizophrenic patients is suicide (in 10% of patients) which represents a 20 times higher risk than for the general population. Deaths from heart disease and from diseases of the respiratory and digestive system are also increased among schizophrenic patients.

Schizophrenia comprises a group of psychoses with either 'positive' or 'negative' symptoms. Positive symptoms consist of hallucinations, delusions and disorders of thought; negative symptoms include emotional flattening, lack of volition and a decrease in motor activity.

A number of biochemical abnormalities have been identified and, in consequence, several neurotransmitter based hypotheses have been advanced over recent years; the most popular one has been "the dopamine hypothesis," one variant of which states that there is over-activity of the mesolimbic dopamine pathways at the level of the D₂ receptor. However, researchers have been unable to consistently find an association between various receptors of the dopaminergic system and schizophrenia.

Bipolar Disorder

Bipolar disorders are relatively common disorders, occurring in about 1.3% of the population, and have been reported to constitute about half of the mood disorders seen in psychiatric clinics with severe and potentially disabling effects. Bipolar disorders have been found to vary with gender depending of the type of disorder; for example, bipolar disorder I is found equally among men and women, while bipolar disorder II is reportedly more common in women. The age of onset of bipolar disorders is typically in the teenage years and diagnosis is typically made in the patient's early twenties. Bipolar disorders also occur among the elderly, generally as a result of a neurological disorder or other medical conditions. In addition to the severe effects on patients' social development, suicide completion rates among bipolar patients are reported to be about 15%.

Bipolar disorders are characterized by phases of excitement and often depression; the excitement phases, referred to as mania or hypomania, and depressive phases can alternate or occur in various admixtures, and can occur to different degrees of severity and over varying duration. Since bipolar disorders can exist in different forms and display different symptoms, the classification of bipolar disorder has been the subject of extensive studies resulting in the definition of bipolar disorder subtypes and widening of the overall concept to include patients previously thought to be suffering from different disorders. Bipolar disorders often share certain clinical signs, symptoms, treatments and neurobiological features with psychotic illnesses in general and therefore present a challenge to the psychiatrist to make an accurate diagnosis. Furthermore, because the course of bipolar disorders and

various mood and psychotic disorders can differ greatly, it is critical to characterize the illness as early as possible in order to offer means to manage the illness over a long term.

The costs of bipolar disorders to society are enormous. The mania associated with the disease impairs performance and causes psychosis, and often results in hospitalization. This disease places a heavy burden on the patient's family and relatives, both in terms of the direct and indirect costs involved and the social stigma associated with the illness, sometimes over generations. Such stigma often leads to isolation and neglect. Furthermore, the earlier the onset, the more severe are the effects of interrupted education and social development.

The DSM-IV classification of bipolar disorder distinguishes among four types of disorders based on the degree and duration of mania or hypomania as well as two types of disorders which are evident typically with medical conditions or their treatments, or to substance abuse. Mania is recognized by elevated, expansive or irritable mood as well as by distractability, impulsive behavior, increased activity, grandiosity, elation, racing thoughts, and pressured speech. Of the four types of bipolar disorder characterized by the particular degree and duration of mania, DSM-IV includes:

- bipolar disorder I, including patients displaying mania for at least one week;
- bipolar disorder II, including patients displaying hypomania for at least 4 days, characterized by milder symptoms of excitement than mania, who have not previously displayed mania, and have previously suffered from episodes of major depression;
- bipolar disorder not otherwise specified (NOS), including patients otherwise displaying features of bipolar disorder II but not meeting the 4 day duration for the excitement phase, or who display hypomania without an episode of major depression; and
- cyclothymia, including patients who show numerous manic and depressive symptoms that do not meet the criteria for hypomania or major depression, but which are displayed for over two years without a symptom-free interval of more than two months.

The remaining two types of bipolar disorder as classified in DSM-VI are disorders evident or caused by various medical disorder and their treatments, and disorders involving or related to substance abuse. Medical disorders which can cause bipolar disorders typically include endocrine disorders and cerebrovascular injuries, and medical treatments causing bipolar disorder are known to include glucocorticoids and the abuse of stimulants. The disorder associated with the use or abuse of a substance is referred to as "substance induced mood disorder with manic or mixed features".

Diagnosis of bipolar disorder can be very challenging. One particularly troublesome difficulty is that some patients exhibit mixed states, simultaneously manic and dysphoric or depressive, but do not fall into the DSM-IV classification because not all required criteria for mania and major depression are met daily for at least one week. Other difficulties include classification of patients in the DSM-IV groups based on duration of phase since patients often cycle between excited and depressive episodes at

different rates. In particular, it is reported that the use of antidepressants may alter the course of the disease for the worse by causing "rapid-cycling". Also making diagnosis more difficult is the fact that bipolar patients, particularly at what is known as Stage III mania, share symptoms of disorganized thinking and behavior with bipolar disorder patients. Furthermore, psychiatrists must distinguish between agitated depression and mixed mania; it is common that patients with major depression (14 days or more) exhibit agitation, resulting in bipolar-like features. A yet further complicating factor is that bipolar patients have an exceptionally high rate of substance, particularly alcohol abuse. While the prevalence of mania in alcoholic patients is low, it is well known that substance abusers can show excited symptoms. Difficulties therefore result for the diagnosis of bipolar patients with substance abuse.

Depression

Depression is a serious medical illness that affects 340 million people worldwide. In contrast to the normal emotional experiences of sadness, loss, or passing mood states, clinical depression is persistent and can interfere significantly with an individual's ability to function. As a result, depression is the leading cause of disability throughout the world with an estimated cost of \$53 billion each year in the United States alone.

Symptoms of depression include depressed mood, diminished interest or pleasure in activities, change in appetite or weight, insomnia or hypersomnia, psycho-motor agitation or retardation, fatigue or loss of energy, feelings of worthlessness or excessive guilt, anxiety, inability to concentrate or act decisively, and recurrent thoughts of death or suicide. A diagnosis of unipolar major depression (or major depressive disorder) is made if a person has five or more of these symptoms and impairment in usual functioning nearly every day during the same two-week period. The onset of depression generally begins in late adolescence or early adult life; however, recent evidence suggests depression may be occurring earlier in life in people born in the past thirty years.

The World Health Organization predicts that by the year 2020 depression will be the greatest burden of ill-health to people in the developing world, and that by then depression will be the second largest cause of death and disability. Beyond the almost unbearable misery it causes, the big risk in major depression is suicide. Within five years of suffering a major depression, an estimated 25% of sufferers try to kill themselves. In addition, depression is a frequent and serious complication of heart attack, stroke, diabetes, and cancer. According to one recent study that covered a 13-year period, individuals with a history of major depression were four times as likely to suffer a heart attack compared to people without such a history.

Depression may also be a feature in up to 50% of patients with CNS disorders such as Parkinson's disease and Alzheimer's disease.

Low levels of the dopamine metabolite HVA are found in the CSF in patients with depression. In addition, dopamine agonists produce a therapeutic response in depression.

Presently, antidepressants are designed to address many of the symptoms of depression by increasing neurotransmitter concentration in aminergic synapses. Distinct pharmacologic mechanisms allow the antidepressants to be separated into seven different classes. The two classical mechanisms are those of tricyclic antidepressants (TCAs) and monoamine oxidase inhibitors (MAOIs). The most widely prescribed agents are the serotonin selective reuptake inhibitors (SSRIs). Three other classes of antidepressants, like the SSRIs, increase serotonergic neurotransmission, but they also have additional actions, namely dual serotonin and norepinephrine reuptake inhibition; serotonin-2 antagonism/reuptake inhibition; and α_2 antagonism plus serotonin-2 and -3 antagonism. The selective norepinephrine and dopamine reuptake inhibitors define a novel class of antidepressant that has no direct actions on the serotonin system.

For CNS disorders such as schizophrenia, bipolar disorder, depression and other mood disorders, all the known molecules used for treatment have side effects and act only against the symptoms of the disease. There is a strong need for new molecules without associated side effects or reduced side effects which are directed against targets that are involved in the causal mechanisms of such CNS disorders. It would be desirable to provide a useful method for the prevention and treatment of such CNS disorders by administering a DAO antagonist compound to a patient susceptible to or suffering from such a disorder. Alternatively, it would be desirable to provide a useful method for the prevention and treatment of such CNS disorders by administering a DDO antagonist compound to a patient susceptible to or suffering from such a disorder.

For CNS disorders such as Parkinson's Disease, Alzheimer's Disease, and other neurodegenerative disorders there are limited numbers of pharmaceutical compositions available for treatment and the known molecules used for treatment have side effects and act only against the symptoms of the disease. There is a strong need for new molecules without associated side effects or reduced side effects which are directed against targets that are involved in the causal mechanisms of such CNS disorders. It would be desirable to provide a useful method for the prevention and treatment of such CNS disorders by administering a DAO activator compound to a patient susceptible to or suffering from such a disorder. Alternatively, it would be desirable to provide a useful method for the prevention and treatment of such CNS disorders by administering a DDO activator compound to a patient susceptible to or suffering from such a disorder.

The pharmaceutical compositions of the present invention are useful for the prevention and treatment of such CNS disorders.

Treatment

As there are currently no cures for CNS disorders such as schizophrenia, bipolar disorder, depression and other mood disorders, the objective of treatment is to reduce the severity of the symptoms, if possible to the point of remission. Due to the similarities in symptoms, schizophrenia, depression and bipolar disorder are often treated with some of the same medicaments. Both diseases are often treated with antipsychotics and neuroleptics.

For schizophrenia, for example, antipsychotic medications are the most common and most valuable treatments. There are four main classes of antipsychotic drugs which are commonly prescribed for schizophrenia. The first, neuroleptics, exemplified by chlorpromazine (Thorazine), has revolutionized the treatment of schizophrenic patients by reducing positive (psychotic) symptoms and preventing their recurrence. Patients receiving chlorpromazine have been able to leave mental hospitals and live in community programs or their own homes. But these drugs are far from ideal. Some 20% to 30% of patients do not respond to them at all, and others eventually relapse. These drugs were named neuroleptics because they produce serious neurological side effects, including rigidity and tremors in the arms and legs, muscle spasms, abnormal body movements, and akathisia (restless pacing and fidgeting). These side effects are so troublesome that many patients simply refuse to take the drugs. Besides, neuroleptics do not improve the so-called negative symptoms of schizophrenia and the side effects may even exacerbate these symptoms. Thus, despite the clear beneficial effects of neuroleptics, even some patients who have a good short-term response will ultimately deteriorate in overall functioning.

The well known deficiencies in the standard neuroleptics have stimulated a search for new treatments and have led to a new class of drugs termed atypical neuroleptics. The first atypical neuroleptic, Clozapine, is effective for about one third of patients who do not respond to standard neuroleptics. It seems to reduce negative as well as positive symptoms, or at least exacerbates negative symptoms less than standard neuroleptics do. Moreover, it has beneficial effects on overall functioning and may reduce the chance of suicide in schizophrenic patients. It does not produce the troubling neurological symptoms of the standard neuroleptics, or raise blood levels of the hormone prolactin, excess of which may cause menstrual irregularities and infertility in women, impotence or breast enlargement in men. Many patients who cannot tolerate standard neuroleptics have been able to take clozapine. However, clozapine has serious limitations. It was originally withdrawn from the market because it can cause agranulocytosis, a potentially lethal inability to produce white blood cells. Agranulocytosis remains a threat that requires careful monitoring and periodic blood tests. Clozapine can also cause seizures and other disturbing side effects (e.g., drowsiness, lowered blood pressure, drooling, bed-wetting, and weight gain). Thus it is usually taken only by patients who do not respond to other drugs.

Researchers have developed a third class of antipsychotic drugs that have the virtues of clozapine without its defects. One of these drugs is risperidone (Risperdal). Early studies suggest that it is as effective as standard neuroleptic drugs for positive symptoms and may be somewhat more effective for negative symptoms. It produces more neurological side effects than clozapine but fewer than standard neuroleptics. However, it raises prolactin levels. Risperidone is now prescribed for a broad range of psychotic patients, and many clinicians seem to use it before clozapine for patients who do not respond to standard drugs, because they regard it as safer. Another new drug is Olanzapine (Zyprexa) which is at least as effective as standard drugs for positive symptoms and more effective for negative symptoms. It has few neurological side effects at ordinary clinical doses, and it does not significantly raise prolactin levels. Although it does not produce most of clozapine's most troubling side effects, including agranulocytosis, some patients taking olanzapine may become sedated or dizzy, develop dry mouth, or gain weight. In rare cases, liver function tests become transiently abnormal.

Outcome studies in schizophrenia are usually based on hospital treatment studies and may not be representative of the population of schizophrenia patients. At the extremes of outcome, 20 % of patients seem to recover completely after one episode of psychosis, whereas 14-19% of patients develop a chronic unremitting psychosis and never fully recover. In general, clinical outcome at five years seems to follow the rule of thirds: with about 35 % of patients in the poor outcome category; 36 % in the good outcome category, and the remainder with intermediate outcome. Prognosis in schizophrenia does not seem to worsen after five years.

Whatever the reasons, there is increasing evidence that leaving schizophrenia untreated for long periods early in course of the illness may negatively affect the outcome. However, the use of drugs is often delayed for patients experiencing a first episode of the illness. The patients may not realize that they are ill, or they may be afraid to seek help; family members sometimes hope the problem will simply disappear or cannot persuade the patient to seek treatment; clinicians may hesitate to prescribe antipsychotic medications when the diagnosis is uncertain because of potential side effects. Indeed, at the first manifestation of the disease, schizophrenia is difficult to distinguish from bipolar manic-depressive disorders, severe depression, drug-related disorders, and stress-related disorders. Since the optimum treatments differ among these diseases, the long term prognosis of the disorder also differs the beginning of the treatment.

For both CNS disorders such as schizophrenia, bipolar disorder, depression and other mood disorder, known molecules used for the treatment have side effects and act only against the symptoms of the disease. There is a strong need for new molecules without associated side effects and directed against targets which are involved in the causal mechanisms of such CNS disorders. Therefore, tools facilitating the discovery and characterization of these targets are necessary and useful.

The aggregation of schizophrenia and bipolar disorder in families, the evidence from twin and adoption studies, and the lack of variation in incidence worldwide, indicate that schizophrenia, depression, and bipolar disorder are primarily genetic conditions, although environmental risk factors are also involved at some level as necessary, sufficient, or interactive causes. For example, schizophrenia occurs in 1% of the general population. But, if there is one grandparent with schizophrenia, the risk of getting the illness increases to about 3%; one parent with Schizophrenia, to about 10%. When both parents have schizophrenia, the risk rises to approximately 40%.

Consequently, there is a strong need to identify genes involved in such CNS disorders. The knowledge of these genes will allow researchers to understand the etiology of schizophrenia, depression, bipolar disorder and other mood disorders and could lead to drugs and medications which are directed against the cause of the diseases, not just against their symptoms.

There is also a great need for new methods for detecting a susceptibility to such CNS disorders as schizophrenia, depression and bipolar disorder, as well as for preventing or following up the development of the disease. Diagnostic tools could also prove extremely useful. Indeed, early identification of subjects at risk of developing such CNS disorders would enable early and/or prophylactic treatment to be administered. Moreover, accurate assessments of the eventual efficacy of a medicament as well as the patient's eventual tolerance to it may enable clinicians to enhance the benefit/risk ratio of treatment regimes for CNS disorders such as those for schizophrenia, depression, bipolar disorder or other mood disorders.

SUMMARY OF THE INVENTION

The present invention stems from an identification of novel polymorphisms including biallelic markers located on human chromosome 13q31-q33 locus, an identification and characterization of novel schizophrenia-related genes located on

- 5 human chromosome 13q31-q33 locus, and from an identification of genetic associations between alleles of biallelic markers located on human chromosome 13q31-q33 locus and disease, as confirmed and characterized in a panel of human subjects. The novel polymorphisms and the schizophrenia-associated gene sequences has been filed in US Patent No. 6,476,208 and International Patent Application WO
10 2000/058510, which disclosures are hereby incorporated by reference in their entireties.

- CNS disorders which can be treated in accordance with the present invention include presenile dementia (early onset Alzheimer's disease), senile dementia (dementia of the Alzheimer's type), Parkinsonism including Parkinson's disease, Huntington's chorea, tardive dyskinesia, hyperkinesia, mania, attention deficit hyperactivity disorder
15 (ADHD), attention deficit disorder (ADD), anxiety disorders, dyslexia, psychotic disorders, schizophrenia, bipolar disorder, major depressive episodes, manic episodes, hypomanic episodes, depression, autistic disorders, substance abuse, excessive aggression, tic disorders and Tourette's syndrome. Preferred disorders of the present invention include schizophrenia, depression and bipolar disorder. Further preferred
20 embodiments of schizophrenia and schizophreniform disorders include: schizophrenia (catatonic), schizophrenia (disorganized), schizophrenia (paranoid), schizophrenia (undifferential), schizophrenia (residual), schizophreniform disorder, brief reactive psychosis, schizoaffective disorder, induced psychotic disorder, schizotypal personality disorder, schizoid personality disorder, paranoid personality disorder and delusional
25 (paranoid) disorder.

- The present invention pertains to methods for providing treatment of CNS disorders to a subject susceptible to such a disorder, and for providing treatment to a subject suffering from a CNS disorder. In particular, the method comprises administering to a patient an amount of a DAO or DDO antagonist or inhibitor
30 compound effective for providing some degree of reversal or amelioration of the progression of the CNS disorder, reversal or amelioration of the symptoms of the CNS disorder, and reversal or amelioration of the reoccurrence of the CNS disorder.

- The present invention further pertains to methods for providing prevention of CNS disorders to a subject susceptible to such a disorder, and for providing treatment
35 to a subject suffering from a CNS disorder. In particular, the method comprises administering to a patient an amount of a DAO or DDO antagonist compound effective

for providing some degree of prevention of the progression of the CNS disorder (i.e., provide protective effects), prevention of the symptoms of the CNS disorder, and prevention of the reoccurrence of the CNS disorder.

The present invention further pertains to the genomic sequence of DAO, novel
5 exons discovered in the DAO gene, novel polymorphic biallelic markers (SNPs) discovered in the DAO gene, methods of detecting persons susceptible to a CNS disorder, novel methods of antagonizing, inhibiting or reducing the activity of DAO, novel methods of agonizing, promoting, increasing the activity of DAO, and a novel composition which affects DAO activity. The present invention further pertains to
10 nucleic acid molecules comprising the genomic sequences of a novel human gene encoding g34872 (sbgl) proteins, proteins encoded thereby, as well as antibodies thereto, as described in US Patent No. 6,476,208 and International Patent Application WO 2000/058510, which disclosures are hereby incorporated by reference in their entireties. The invention also deals with the cDNA sequences encoding the g34872,
15 DAO and DDO proteins, and variants thereof. Oligonucleotide probes or primers hybridizing specifically with a g34872, DAO, and DDO genomic or cDNA sequence are also part of the present invention, as well as DNA amplification and detection methods using said primers and probes.

A further object of the invention consists of recombinant vectors comprising any
20 of the nucleic acid sequences described above, and in particular of recombinant vectors comprising a g34872, DDO, and DAO regulatory sequence or a sequence encoding a g34872, DDO, and DAO protein, as well as of cell hosts and transgenic non human animals comprising said nucleic acid sequences or recombinant vectors.

The invention also concerns to biallelic markers of the g34872, DAO and DDO
25 gene and the use thereof. Included are probes and primers for use in genotyping biallelic markers of the invention.

An embodiment of the invention encompasses any polynucleotide of the invention attached to a solid support polynucleotide may comprise a sequence disclosed in the present specification; optionally, said polynucleotide may comprise, consist of,
30 or consist essentially of any polynucleotide described in the present specification; optionally, said determining may be performed in a hybridization assay, sequencing assay, microsequencing assay, or an enzyme-based mismatch detection assay; optionally, said polynucleotide may be attached to a solid support, array, or addressable array; optionally, said polynucleotide may be labeled.

35 Finally, the invention is directed to drug screening assays and methods for the screening of substances for the treatment of schizophrenia, bipolar disorder or a related

CNS disorder based on the role of g34872, DAO, or DDO nucleotides and polynucleotides in disease. One object of the invention deals with animal models of schizophrenia, including mouse, primate, non-human primate bipolar disorder or related CNS disorder based on the role of g34872, DAO, or DDO in disease. The

- 5 invention is also directed to methods for the screening of substances or molecules that inhibit the expression of g34872, DAO, or DDO, as well as with methods for the screening of substances or molecules that interact with a g34872, DAO, or DDO polypeptide, or that modulate the activity of a g34872, DAO, or DDO polypeptide.

As noted above, certain aspects of the present invention stem from the identification of genetic associations between schizophrenia and bipolar disorder and alleles of biallelic markers of g34872 gene and the DAO gene. The invention provides appropriate tools for establishing further genetic associations between alleles of biallelic markers in the g34872 and DAO locus and either side effects or benefit resulting from the administration of agents acting on CNS disorders or symptoms such as schizophrenia, depression or bipolar disorder, or schizophrenia or bipolar disorder symptoms, including agents like chlorpromazine, clozapine, risperidone, olanzapine, sertindole, quetiapine and ziprasidone.

The invention provides appropriate tools for establishing further genetic associations between alleles of biallelic markers of DAO and g34872 with a trait. Methods and products are provided for the molecular detection of a genetic susceptibility in humans to schizophrenia and bipolar disorder. They can be used for diagnosis, staging, prognosis and monitoring of this disease, which processes can be further included within treatment approaches. The invention also provides for the efficient design and evaluation of suitable therapeutic solutions including individualized strategies for optimizing drug usage, and screening of potential new medicament candidates.

A preferred embodiment of the invention includes a method of treating a central nervous system disorder in a patient in need thereof, the method comprising administering said patient an effective amount of a composition or compound comprising a DAO antagonist or inhibitor or a DDO antagonist or inhibitor.

Further preferred is a method of treating psychosis, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a DAO antagonist or inhibitor or a DDO inhibitor or antagonist.

Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a DAO antagonist or inhibitor or a DDO inhibitor or antagonist.

Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a DAO antagonist or inhibitor or a DDO inhibitor or antagonist.

A preferred embodiment of the invention includes a method of treating a central nervous system disorder in a patient in need thereof, the method comprising administering said patient an effective amount of a composition or compound comprising a DAO antagonist or inhibitor and a DDO antagonist or inhibitor.

Further preferred is a method of treating psychosis, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a DAO antagonist or inhibitor and a DDO inhibitor or antagonist.

Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a DAO antagonist or inhibitor and a DDO inhibitor or antagonist.

Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a DAO antagonist or inhibitor and a DDO inhibitor or antagonist.

A preferred embodiment of the invention includes a method of treating a central nervous system disorder in a patient in need thereof, the method comprising administering said patient an effective amount of a composition or compound comprising a g34872 antagonist or inhibitor.

Further preferred is a method of treating psychosis, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a g34872 inhibitor or antagonist.

Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a g34872 inhibitor or antagonist.

Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a g34872 inhibitor or antagonist.

A preferred embodiment of the invention includes a method of treating a central nervous system disorder in a patient in need thereof, the method comprising administering said patient an effective amount of a composition or compound comprising a DAO antagonist or inhibitor or a DDO antagonist or inhibitor in combination with a g34872 antagonist or inhibitor composition or compound.

Further preferred is a method of treating psychosis, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a DAO antagonist or inhibitor or a DDO antagonist or inhibitor in combination with a g34872 antagonist or inhibitor composition or compound.

Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a DAO antagonist or inhibitor or a DDO antagonist or inhibitor in combination with a g34872 antagonist or inhibitor composition or compound.

Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a DAO antagonist or inhibitor or a DDO antagonist or inhibitor in combination with a g34872 antagonist or inhibitor composition or compound.

A preferred embodiment of the invention includes a method of treating a central nervous system disorder in a patient in need thereof, the method comprising administering said patient an effective amount of a composition or compound comprising a combination of a DAO antagonist or inhibitor, a DDO antagonist or inhibitor, and a g34872 antagonist or inhibitor composition or compound.

Further preferred is a method of treating psychosis, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a combination of a DAO antagonist or inhibitor, a DDO antagonist or inhibitor, and a g34872 antagonist or inhibitor composition or compound.

Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a combination of a DAO antagonist or inhibitor, a DDO antagonist or inhibitor, and a g34872 antagonist or inhibitor composition or compound.

Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising a combination of a DAO antagonist or inhibitor, a DDO antagonist or inhibitor, and a g34872 antagonist or inhibitor composition or compound.

A preferred embodiment of the invention includes a method of treating a central nervous system disorder in a patient in need thereof, the method comprising administering said patient an effective amount of a composition or compound comprising at least one of the following: a DAO antagonist or inhibitor, a DDO antagonist or inhibitor, or a g34872 antagonist or inhibitor composition or compound.

Further preferred is a method of treating psychosis, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising at least one of the following: a DAO antagonist or inhibitor, a DDO antagonist or inhibitor, or a g34872 antagonist or inhibitor composition or compound.

Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising at least one of the following: a DAO antagonist or inhibitor, a DDO antagonist or inhibitor, or a g34872 antagonist or inhibitor composition or compound.

Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition or compound comprising at least one of the following: a DAO antagonist or inhibitor, a DDO antagonist or inhibitor, or a g34872 antagonist or inhibitor composition or compound.

It should be appreciated that compositions or compounds known in the art to be used in methods of treating a central nervous system disorder, a psychosis, a schizophrenic disorder, or a bipolar disorder

in a patient in need thereof, which are known to or inherently act to inhibit or antagonize DAO, DDO or g34872 are preferentially excluded from the present invention.

A further preferred embodiment of the invention relates to methods to inhibit DAO activity. Furthermore, the invention relates to a method to treat schizophrenia by inhibiting DAO activity. Further preferred is a method to treat schizophrenia by inhibiting DAO activity using a composition comprising a ketimine to inhibit DAO activity.

Another preferred embodiment is directed to a method to inhibit DDO activity. Furthermore, the invention relates to a method to treat schizophrenia by inhibiting DDO activity. Further preferred is a method to treat schizophrenia by inhibiting DDO activity using a composition comprising a ketimine to inhibit DDO activity.

Another preferred embodiment of the invention relates to methods of inhibiting the interaction between DAO and g34872.

Another preferred embodiment of the invention relates to a method of inhibiting the interaction between g34872 and DDO.

Another embodiment of the invention relates to any polypeptide fragment of a DAO polypeptide of SEQ ID NOs:7, 8, 9, 10, or 18 which antagonizes the interaction between said DAO polypeptide and a g34872 polypeptide of SEQ ID NO:14, or fragment thereof. Further preferred is a fragment of a DAO polypeptide comprising amino acids 23-347 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 227-347 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 31-347 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 51-347 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 66-347 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 101-347 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 126-347 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 146-347 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 175-347 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 180-347 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 1-189 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 1-205 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 31-189 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 31-205 of SEQ ID NO:7. Further preferred is a fragment of a DAO polypeptide comprising amino acids 84-205 of SEQ ID NO:7.

A further preferred embodiment of the invention relates to compositions which bind to a DAO polypeptide or fragment thereof. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 23-347 of SEQ ID NO:7. Further preferred are compositions which

bind to a fragment of a DAO polypeptide comprising amino acids 227-347 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 31-347 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 51-347 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 66-347 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 101-347 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 126-347 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 146-347 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 175-347 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 180-347 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 1-189 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 1-205 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 31-189 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 31-205 of SEQ ID NO:7. Further preferred are compositions which bind to a fragment of a DAO polypeptide comprising amino acids 84-205 of SEQ ID NO:7.

A further preferred embodiment is directed to a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide, or a fragment thereof. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 23-347 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 227-347 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 31-347 of SEQ ID NO:7. Further preferred a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 51-347 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering

therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 66-347 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 101-347 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 126-347 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 146-347 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 175-347 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 180-347 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 1-189 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 1-205 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 31-189 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 31-205 of SEQ ID NO:7. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 84-205 of SEQ ID NO:7.

A further preferred embodiment is directed to a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide, or a fragment thereof.

Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 23-347 of SEQ ID NO:7.

Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 227-347 of SEQ ID NO:7.

Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 31-347 of SEQ ID NO:7.

Further preferred a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 51-347 of SEQ ID NO:7. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 66-347 of SEQ ID NO:7. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 101-347 of SEQ ID NO:7. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 126-347 of SEQ ID NO:7. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 146-347 of SEQ ID NO:7. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 175-347 of SEQ ID NO:7. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 180-347 of SEQ ID NO:7. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 1-189 of SEQ ID NO:7. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide

comprising amino acids 1-205 of SEQ ID NO:7. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 31-189 of SEQ ID NO:7. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 31-205 of SEQ ID NO:7. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a DAO polypeptide comprising amino acids 84-205 of SEQ ID NO:7.

A further preferred embodiment of the invention relates to compositions which bind to a g34872 polypeptide of SEQ ID NO:14, or fragment thereof. Further preferred are compositions which bind to a g34872 polypeptide comprising amino acids 65-153 of SEQ ID NO:14, or fragment thereof. Further preferred are compositions which bind to a polypeptide of SEQ ID NO:16 or fragment thereof.

A further preferred embodiment is directed to a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a g34872 polypeptide of SEQ ID NO:14, or fragment thereof. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a g34872 polypeptide comprising amino acids 65-153 of SEQ ID NO:14, or fragment thereof. Further preferred is a method of treating schizophrenia, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a polypeptide of SEQ ID NO:16 or fragment thereof.

A further preferred embodiment is directed to a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a g34872 polypeptide of SEQ ID NO:14, or fragment thereof. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a g34872 polypeptide comprising amino acids 65-153 of SEQ ID NO:14, or fragment thereof. Further preferred is a method of treating bipolar disorder, the method comprising administering to a patient suffering therefrom a therapeutically effective amount of a composition comprising a composition which binds to a polypeptide of SEQ ID NO:16 or fragment thereof.

A further preferred embodiment of the invention relates to any polypeptide fragment of a g34872 polypeptide of SEQ ID NO:14 which antagonizes the interaction between said g34872 polypeptide or fragment thereof and a DAO polypeptide or fragment thereof. Further preferred is any fragment of g34872 which antagonizes the increase in DAO activity by a g34872 polypeptide. Further preferred is a fragment of a g34872 polypeptide comprising the amino acids of SEQ ID NO:16.

A further preferred embodiment of the invention relates to compositions which antagonize the interaction between a g34872 polypeptide of SEQ ID NO:14, or a fragment thereof, and a DAO polypeptide of SEQ ID NOs:7-10 or 18, or a fragment thereof.

A further preferred embodiment of the invention relates to compositions which antagonize the interaction between a g34872 polypeptide of SEQ ID NO:14, or a fragment thereof, and a DDO polypeptide of SEQ ID NOs:21 or 22, or a fragment thereof.

Another embodiment of the invention relates to methods of increasing the activity of DAO with a g34872 polypeptide or fragment thereof. Furthermore, the invention relates to methods of increasing the activity of DDO with a g34872 polypeptide or fragment thereof.

A further embodiment of the invention relates to methods of inhibiting the glycosylation of DAO.

A further embodiment of the invention relates to methods of enhancing the multimerization of DAO.

A further embodiment of the invention relates to methods of inhibiting translation of DAO.

A further embodiment of the invention relates to differential identification of DAO variants.

A preferred embodiment of the invention is directed to a composition or a compound which reduces, inhibits or antagonizes DAO activity. Further preferred, the composition or compound is a competitive inhibitor or antagonist of DAO activity. Further preferred, the composition or compound is a noncompetitive inhibitor or antagonist of DAO activity. Further preferred, the composition or compound is an uncompetitive inhibitor or antagonist of DAO activity. Further preferred, the composition or compound is an allosteric inhibitor or antagonist of DAO activity. Further preferred, the composition or compound is a reversible inhibitor or antagonist of DAO activity. Further preferred, the composition or compound is an irreversible inhibitor or antagonist of DAO activity.

A further embodiment is directed to a composition or compound which reduces, inhibits or antagonizes DDO activity. Further preferred, the composition or compound is a competitive inhibitor or antagonist of DDO activity. Further preferred, the composition or compound is a noncompetitive inhibitor or antagonist of DDO activity. Further preferred, the composition or compound is an uncompetitive inhibitor or antagonist of DDO activity. Further preferred, the composition or compound is an allosteric inhibitor or antagonist of DDO activity. Further preferred, the composition or compound is a reversible inhibitor or antagonist of DDO activity. Further preferred, the composition or compound

is an irreversible inhibitor or antagonist of DDO activity. Further preferred are compositions or compounds which reduce, inhibit or antagonize the activity of DAO and DDO.

Further preferred is a method of treating a CNS disorder with a composition or compound which reduces, inhibits or antagonizes DAO activity. Further preferred is a method of treating a CNS disorder with a composition or compound which reduces, inhibits or antagonizes DDO activity. Further preferred is a method of treating a CNS disorder with a composition or compound which reduces, inhibits or antagonizes both DAO and DDO activity. Further preferred is a method of treating a CNS disorder with a composition or compound which reduces, inhibits or antagonizes either DAO and DDO activity. Further preferred is a method of treating a CNS disorder with a first composition or compound which reduces, inhibits or antagonizes DAO in combination with a second composition which reduces, inhibits or antagonizes DDO activity. Further preferred is a method of treating a CNS disorder with a composition or compound which reduces, inhibits or antagonizes DAO in combination with another composition. Further preferred is a method of treating a CNS disorder with a composition or compound which reduces, inhibits or antagonizes DAO in combination with another composition routinely used in the treatment of said CNS disorder. Further preferred is a method of treating a CNS disorder with a composition or compound which reduces, inhibits or antagonizes DAO in combination with another composition unrelated to the treatment of said CNS disorder. Further preferred is a method of treating a CNS disorder with a composition or compound which reduces, inhibits or antagonizes DDO in combination with another composition. Further preferred is a method of treating a CNS disorder with a composition or compound which reduces, inhibits or antagonizes DDO in combination with another composition routinely used in the treatment of said CNS disorder. Further preferred is a method of treating a CNS disorder with a composition or compound which reduces, inhibits or antagonizes DDO in combination with another composition unrelated to the treatment of said CNS disorder.

Preferred compositions or compounds of the invention which reduce, inhibit or antagonize DAO or DDO activity are selected from, but not limited to, the list comprising:

- i. **IRI**, 2-oxo-3-pentynoate;
- ii. **CMI**, Aminoguanidine (Guanyldiazine; Carbamimidic hydrazide; Pimagedine; GER 11; Hydrazinecarboximidamide) or hydrochloride salt (Guanyldiazine hydrochloride), bicarbonate salt, nitrate salt, sulfate (2:1) salt, sulfate (1:1) salt, and hemisulfate salt thereof;
- iii. **FI**, benzoic acid;
- iv. **FI**, sodium benzoate;
- v. **FI**, 2-aminobenzoate;
- vi. **FI**, 3-aminobenzoate;
- vii. **FI**, 4-aminobenzoate (p-aminobenzoate, PABA, Vitamin Bx, Vitamin H1);

- viii. **CMI**, Methylglyoxal bis(guanyldihydrazone) (also known as: Methyl GAG; Mitoguazone; 1,1'-((Methylethanediyldiene)dinitrilo)diguandine; Hydrazinecarboximidamide, 2,2'-(1-methyl-1,2-ethanediyldiene)bis-; Pyruvaldehyde bis(amidinohydrazone); Megag; Mitoguazona [INN-Spanish] ; Guanidine, 1,1'-((methylethanediyldiene)dinitrilo)di-; 1,1'-((Methylethanediyldiene)dinitrilo)diguandine);
- ix. **CMI**, Methylglyoxal bis(guanyldihydrazone), dihydrochloride;
- x. **CMI**, phenylglyoxal bis(guanyldihydrazone) (PhGBG);
- xi. **CMI**, glyoxal bis(guanyldihydrazone) (GBG; Guanidine, 1,1'-(ethanediyldenedinitrilo)di-(8CI); Hydrazinecarboximidamide, 2,2'-(1,2-ethanediyldiene)bis- (9CI));
- xii. **CMI**, indole-propionic (IPA, 3-(3-Indolyl)propanoic acid);
- xiii. **CMI**, 3-indole-acetic acid (Heteroauxin, IAA);
- xiv. **CMI**, Indole-3-acetic acid Sodium salt;
- xv. **CMI**, Indole-3-acetone;
- xvi. **CMI**, Indole-3-acetamide;
- xvii. **CMI**, Indole-3-acetyl-L-aspartic acid;
- xviii. **CMI**, Indole-3-acetyl-L-alanine;
- xix. **CMI**, Indole-3-acetyl-glycine;
- xx. **CMI**, Indole-3-acetaldehyde Sodium Bisulfite Addition compound;
- xxi. **CMI**, Indole-3-carboxylic acid;
- xxii. **CMI**, Indole-3-pyruvic acid (3-(3-Indolyl)-2-oxopropanoic acid);
- xxiii. **FI**, salicylic acid (2-Hydroxybenzoic acid);
- xxiv. **FI**, salicylic acid Sodium Salt;
- xxv. **FI**, Salicylic acid Potassium Salt;
- xxvi. **IRI**, Dansyl chloride (5-(Dimethylamino)naphthalene-1-sulfonyl chloride);
- xxvii. **IRI**, Dansyl fluoride (5-(Dimethylamino)naphthalene-1-sulfonyl fluoride);
- xxviii. **CMI**, dansyl glycine;
- xxix. **CMI**, Alanine tetrazole;
- xxx. **FI**, benzoic tetrazole;
- xxxi. **CMI**, tetrazole;
- xxxii. **CMI**, Riboflavin 5'-pyrophosphate (RPP, 5-Phospho-alpha-D-ribosyl diphosphate, P-Rib-PP, P-RPP);
- xxxiii. **IRI**, DL-propargylglycine (DL-PG, 2-Amino-4-pentynoic acid);
- xxxiv. **IRI**, L-C-Propargylglycine;
- xxxv. **IRI**, N-Acetyl-DL-propargylglycine;
- xxxvi. **FII**, ($\pm$)-Sodium 3-hydroxybutyrate;

- xxxvii. **FI**, Trigonelline Hydrochloride (1-Methylpyridinium-3-carboxylate);
- xxxviii. **FI**, N-methylnicotinate;
- xxxix. **FI**, Methyl 6-methylnicotinate;
 - xl. **FI**, Ethyl 2-methylnicotinate;
 - xli. **CMI**, Kojic acid (2-Hydroxymethyl-5-hydroxy-gamma-pyrone, 5-Hydroxy-2-hydroxymethyl-4-pyranone);
 - xl.ii. **CMI**, derivatives of kojic acid, such as: 6-(PYRROLIDINOMETHYL)-KOJIC ACID HYDROCHLORIDE, 6-(MORPHOLINOMETHYL)-KOJIC ACID, 6-(DIETHYLAMINOMETHYL)-KOJIC ACID Hydrochloride;
 - xl.iii. **IRI**, O-(2,4-dinitrophenyl)hydroxylamine;
 - xl.iv. **CMI**, 2,4-DINITROPHENYL GLYCINE;
 - xl.v. **CMI**, Hydroxylamine Hydrochloride;
 - xl.vi. **IRI**, Methyl-p-nitrobenzenesulfonate (Methyl 4-nitrobenzenesulfonate);
 - xl.vii. **FIV**, Aminoethylcysteine-ketimine (AECK, Thialysine ketimine, 2H-1,4-Thiazine-5,6-dihydro-3-carboxylic acid, S-Aminoethyl-L-cysteine ketimine, 2H-1,4-Thiazine-3-carboxylic acid, 5,6-dihydro-);
 - xl.viii. **FIV**, 1,4-thiazine derivatives;
 - xl.ix. **CMI**, 4-Phenyl-1,4-sulfonazan (Tetrahydro-4-phenyl-4H-1,4-thiazine 1-oxide, 4H-1,4-Thiazine, tetrahydro-4-phenyl-, 1-oxide);
 - 1. **CMI**, Phenothiazine (Thiodiphenylamine, 10H-Phenothiazine, AFI-Tiazin, Agrazine, Antiverm, Dibenzo-1,4-thiazine);
 - li. **CMI**, 3,4-Dihydro-2H-1,4-thiazine-3,5-dicarboxylic acid (3,4-Dhtca, CAS#86360-62-5);
 - lii. **CMI**, Nifurtimox (Nifurtimox [BAN:INN], 1-((5-Nitrofurfurylidene)amino)-2-methyltetrahydro-1,4-thiazine-4,4-dioxide, 3-Methyl-4-(5'-nitrofurylidene-amino)-tetrahydro-4H-1,4-thiazine-1,1-dioxide, BAY 2502, 4-((5-Nitrofurfurylidene)amino)-3-methylthiomorpholine 1,1-dioxide, etc);
 - lii.iii. **FIV**, 3-(1-Pyrrolidinylmethyl)-4-(5,6-dichloro-1-indancarbonyl)-tetrahydro-1,4-thiazine hydrochloride (R 84760; R 84761; **Thiomorpholine**, 4-((5,6-dichloro-2,3-dihydro-1H-inden-1-yl)carbonyl)-3-(1-pyrrolidinylmethyl)-, monohydrochloride, (R-(R*,S*))-);
 - liv. **FIV**, ketimine reduced forms;
 - lv. **CMI**, cystathionine;
 - lvi. **FIII**, cystathionine ketimine;
 - lvii. **FIV**, lanthionine ketimine;
 - lviii. **FIV**, thiomorpholine-2-carboxylic acid;
 - lix. **CMI**, thiomorpholine-2,6-dicarboxylic acid;

- lx. **FIV**, TMDA (1,4-Thiomorpholine-3,5-dicarboxylic acid);
- lxi. **IRI**, 1-chloro-1-nitroethane;
- lxii. **FI**, anthranilate;
- lxiii. **FI**, Ethyl 2-aminobenzoate (ethyl anthranilate);
- lxiv. **FI**, Methyl 2-aminobenzoate (Methyl anthranilate);
- lxv. **FI**, picolinate;
- lxvi. **FI**, Ethyl picolinate (2-(Ethoxycarbonyl)pyridine, Ethyl 2-pyridinecarboxylate, ;
- lxvii. **CMI**, L-Leucine methyl ester, hydrochloride;
- lxviii. **CMI**, L-leucine ([[(S)-(+)-leucine]]);
- lxix. **IRI**, Fluorodinitrobenzene (1-Fluoro-2,4-dinitrobenzene, 2,4-DNFB, Benzene, 1-fluoro-2,4-dinitro-, VAN, etc);
- lxx. **IRI**, Dinitrochlorobenzene (1-Chloro-2,4-dinitrobenzene, 1,3-Dinitro-4-chlorobenzene, etc);
- lxxi. **IRI**, 1,2-cyclohexanedione;
- lxxii. **IRI**, Allylglycine (D-Allylglycine, 4-Pentenoic acid, 2-amino-);
- lxxiii. **CMI**, 2-amino-2,4-pentadienoate;
- lxxiv. **CMI**, 2-hydroxy-2,4-pentadienoate;
- lxxv. **CMI**, 2-amino-4-keto-2-pentenoate;
- lxxvi. **FII**, 2-hydroxybutyrate;
- lxxvii. **FII**, Sodium 2-hydroxybutyrate;
- lxxviii. **IRI**, N-chloro-D-leucine;
- lxxix. **CMI**, N-Acetyl-D-leucine;
- lxxx. **CMI**, D-Leu (D-2-Amino-4-methylpentanoic acid);
- lxxxi. **IRI**, D-propargylglycine; 2-Amino-4-pentynoic acid; D,L-Propargylglycine; L-2-Amino-4-pentynoic acid ;
- lxxxii. **CMI**, Progesterone (4-Pregnene-3,20-dione);
- lxxxiii. **CMI**, FAD (Flavin adenine dinucleotide , 1H-Purin-6-amine, flavin dinucleotide, Adenosine 5'-(trihydrogen pyrophosphate), 5'-5'-ester with riboflavin, etc);
- lxxxiv. **CMI**, 6-OH-FAD;
- lxxxv. **IRI**, Phenylglyoxal (2,2-Dihydroxyacetophenone);
- lxxxvi. **IRI**, Phenylglyoxal Monohydrate (2,2-Dihydroxyacetophenone monohydrate);
- lxxxvii. **FIII**, Cyclothionine (Perhydro-1,4-thiazepine-3,5-dicarboxylic acid, 1,4-Hexahydrothiazepine-3,5-dicarboxylic acid, 1,4-Thiazepine-3,5-dicarboxylic acid, hexahydro-);
- lxxxviii. **CMI**, alpha-alpha'-iminodipropionic (Alanopine; 2,2'-Iminodipropionic acid; L-Alanine, N-(1-carboxyethyl)-);

- lxxxix. **CMI**, Meso-Diaminosuccinic acid (3-Aminoaspartic acid ; Diaminosuccinic acid; CAS RN: 921-52-8); meso-2,3-Diaminosuccinic acid (CAS RN: 23220-52-2);
- xc. **CMI**, Thiosemicarbazide (thiocarbamoyl hydrazide);
- xc. **CMI**, Thiourea (Sulfourea; Thiocarbamide);
- xcii. **CMI**, Methylthiouracil (4(6)-Methyl-2-thiouracil, 4-Hydroxy-2-mercapto 6-methylpyrimidine);
- xciii. **CMI**, Sulphathiazole (N1-(2-Thiazolyl)sulfanilamide, 4-Amino-N-2-thiazolylbenzenesulfonamide);
- xciv. **CMI**, Sulfathiazole Sodium Salt (4-Amino-N-2-thiazolylbenzenesulfonamide sodium salt);
- xcv. **CMI**, Thiocyanate;
- xcvi. **FI**, 3-METHYLBENZYL THIOCYANATE;
- xcvii. **CMI**, methimazole (2-Mercapto-1-methylimidazole, 1-Methylimidazole-2-thiol);
- xcviii. **FII**, Dicarboxylic hydroxyacids;
- xcix. **FII**, 1,3-Acetonedicarboxylic acid (3-Oxoglutaric acid);
 - c. **CMI**, D-tartaric acid ([(2S,3S)-(-)-tartaric acid, unnatural tartaric acid]);
 - ci. **CMI**, L-tartaric acid ([(2R,3R)-(+)-tartaric acid, natural tartaric acid]);
 - cii. **CMI**, DL-tartaric acid;
 - ciii. potassium tartrate;
- civ. **FII**, D-malic acid; [(R)-(+)-malic acid, (R)-(+)-hydroxysuccinic acid];
- cv. **FII**, L-malic acid; [(S)-(-)-malic acid, (S)-(-)-hydroxysuccinic acid];
- cvi. **FII**, DL-Malic acid (DL-hydroxysuccinic acid);
- cvi. **FII**, Alpha-keto acids that are analogues of the amino acids alanine, leucine, phenylalanine, phenylglycine, tyrosine, serine, aspartate, etc and salts and derivatives thereof;
- cvi. **FII**, pyruvic acid (2-Oxopropionic acid, alpha-Ketopropionic acid);
- cix. **FII**, sodium pyruvate;
- cx. **FII**, Pyruvic acid methyl ester (methyl pyruvate);
- cx. **FI**, Phenylpyruvic acid;
- cxii. **FII**, Calcium phenylpyruvate (calcium pyruvate);
- cxiii. **FI**, Phenylpyruvic acid Sodium salt (Sodium phenylpyruvate);
- cxiv. **FII**, 4-hydroxyphenyl pyruvic acid;
- cxv. **FII**, sodium alpha-ketoisovaleric acid (3-Methyl-2-oxobutyric acid Sodium salt, 3-Methyl-2-oxobutanoic acid sodium salt, a-Ketoisovaleric acid Sodium salt; Ketovaline Sodium salt);
- cxvi. **FI**, benzoylformic acid (a-Oxophenylacetic acid, Phenylglyoxylic acid);
- cxvii. **FII**, 4-methylthio-2-oxopentanoic acid;
- cxviii. **FII**, 4-Methyl-2-oxopentanoic acid (4-Methyl-2-oxovaleric acid; alpha-Ketoisocaproic acid; ;

- cxix. **FII**, 4-methylthio-2-oxybutanoic acid;
- cxx. **FII**, 2-oxybutanoic acid (hydroxybutyrate; 2-Hydroxybutyric acid; alpha-Hydroxy-n-butyric acid;
- cxxi. **FII**, DL-alpha-Hydroxybutyric acid Sodium Salt (sodium ($\pm$)-2-Hydroxybutyrate);
- cxxii. **FII**, Indole-3-pyruvic acid (alpha-Keto analogue of tryptophan);
- cxxiii. The reaction product between cysteamine and bromopyruvate;
- cxxiv. **CMI**, cysteamine (2-Aminoethanethiol; 2-Mercaptoethylamine);
- cxxv. **CMI**, pantetheine;
- cxxvi. **CMI**, S-adenosylmethionine;
- cxxvii. **IRI**, Ethyl bromopyruvate;
- cxxviii. **IRI**, Methyl bromopyruvate;
- cxxix. **IRI**, Bromopyruvate; and
- cxix. **CMI**, 5-S-Cysteinyl dopamine,

wherein IRI indicates Irreversible Inhibitor compositions; CMI indicates Competitive Inhibitor compositions not included in Formula I-IV compositions; FI indicates Formula I compositions as described herein; FII indicates Formula II compositions as described herein; FIII indicates Formula III compositions as described herein; and FIV indicates Formula IV compositions as described herein. It should be appreciated that Formula I-IV compositions are competitive, noncompetitive, uncompetitive or allosteric inhibitors of DAO or DDO.

Preferred compositions to be used in methods of the invention to reduce, inhibit, or antagonize DAO or DDO catalytic activity in vitro or in vivo are selected from the above list of compositions "i" through and including "cxix"; more preferred are compositions selected from irreversible inhibitor compositions, Formula I compositions, Formula II compositions, Formula III compositions and Formula IV compositions; even more preferred are compositions selected from Formula I compositions, Formula II compositions, Formula III compositions and Formula IV; most preferred are compositions selected from Formula I and Formula IV. Further preferred compositions to be used in methods of the invention to reduce, inhibit, or antagonize DAO or DDO catalytic activity in vitro or in vivo are selected from the group comprising benzoate, aminoethylcysteine ketimine (AECK), and derivatives thereof.

In a further preferred embodiment, preferred compositions or compounds to be used in methods of the invention of treating a CNS disorder are selected from the above list of compositions "i" through and including "cxix"; more preferred are compositions selected from irreversible inhibitor compositions, Formula I compositions, Formula II compositions, Formula III compositions and Formula IV compositions; even more preferred are compositions selected from Formula I compositions, Formula II compositions, Formula III compositions and Formula IV; most preferred are compositions selected from Formula I and Formula IV. Further preferred compositions to be used in methods of the invention

of treating a CNS disorder are selected from the group comprising benzoate, aminoethylcysteine ketimine (AECK), and derivatives thereof.

A highly preferred compound or composition of the invention to reduce, inhibit or antagonize DAO or DDO activity is selected from the list comprising, but not limited to: Aminoethylcysteine-ketimine (AECK, Thialysine ketimine, 2H-1,4-Thiazine-5,6-dihydro-3-carboxylic acid, S-Aminoethyl-L-cysteine ketimine, 2H-1,4-Thiazine-3-carboxylic acid, 5,6-dihydro-); aminoethylcysteine (thialysine); cysteamine; pantetheine; cystathionine and S-adenosylmethionine.

Another preferred embodiment of the invention is directed to a compound or composition which reduces, inhibits or antagonizes the oxidation or degradation of an amino acid or derivative thereof.

Another preferred embodiment of the invention is directed to a compound or composition which reduces, inhibits or antagonizes the oxidation or degradation an L-amino acid or derivative thereof.

Another preferred embodiment of the invention is directed to a compound or composition which reduces, inhibits or antagonizes the oxidation or degradation of an D-amino acid or derivative thereof.

Another preferred embodiment of the invention is directed to a compound or composition which reduces, inhibits or antagonizes the oxidation or degradation of glycine or derivative thereof. A further preferred embodiment of the invention is directed to a compound or composition which reduces, inhibits or antagonizes the oxidation or degradation of at least one D-amino acid selected from the list

comprising: D-Met, D-Pro, D-Phe, D-Tyr, D-Ile, D-Leu, D-Ala, D-Val, D-Ser, D-Arg, D-His, D-norleucine, D-Trp, D-Ornithine, cis-4-hydroxy-D-proline, D-Thr, D-Trp-methyl ester, N-acetyl-D-Ala, D-Lys, D-Asp, D-Glu, D-Asn, D-Gln, D-Asp-dimethyl-ester and N-methyl-D-Asp. Further preferred is a composition which reduces, inhibits, or antagonizes the oxidation or degradation of D-serine. Further preferred is a composition or compound which reduces, inhibits or antagonizes the oxidation or

degradation of D-Ser, N-methyl-D-Asp, D-Asp or Gly. A preferred compound or composition of the invention which reduces, inhibits or antagonizes the oxidation or degradation of an amino acid, or derivative thereof, is selected from the list including, but not limited to comprising: Aminoethylcysteine-ketimine (AECK, Thialysine ketimine, 2H-1,4-Thiazine-5,6-dihydro-3-carboxylic acid, S-Aminoethyl-L-cysteine ketimine, 2H-1,4-Thiazine-3-carboxylic acid, 5,6-dihydro-); aminoethylcysteine (thialysine); cysteamine; pantetheine; cystathionine and S-adenosylmethionine. A preferred compound or

composition of the invention which reduces, inhibits or antagonizes the oxidation or degradation of D-Met, D-Pro, D-Phe, D-Tyr, D-Ile, D-Leu, D-Ala, D-Val, D-Ser, D-Arg, D-His, D-norleucine, D-Trp, D-Ornithine, cis-4-hydroxy-D-proline, D-Thr, D-Trp-methyl ester, N-acetyl-D-Ala, D-Lys, D-Asp, D-Glu, D-Asn, D-Gln, D-Asp-dimethyl-ester, N-methyl-D-Asp or Gly is selected from the list including, but not limited to comprising: Aminoethylcysteine-ketimine (AECK, Thialysine ketimine, 2H-1,4-Thiazine-5,6-dihydro-3-carboxylic acid, S-Aminoethyl-L-cysteine ketimine, 2H-1,4-Thiazine-3-carboxylic acid, 5,6-dihydro-); aminoethylcysteine (thialysine); cysteamine; pantetheine; cystathionine and S-

adenosylmethionine. A preferred compound or composition of the invention which reduces, inhibits or antagonizes the oxidation or degradation of D-Ser is selected from the list including, but not limited to comprising: Aminoethylcysteine-ketimine (AECK, Thialysine ketimine, 2H-1,4-Thiazine-5,6-dihydro-3-carboxylic acid, S-Aminoethyl-L-cysteine ketimine, 2H-1,4-Thiazine-3-carboxylic acid, 5,6-dihydro-); aminoethylcysteine (thialysine); cysteamine; pantetheine; cystathionine and S-adenosylmethionine.

Another embodiment of the invention is directed to a composition which reduces, inhibits or antagonizes the oxidation of Reduced-Flavin Adenine Dinucleotide (Re-FAD). Another embodiment of the invention is directed to a composition which reduces, inhibits or antagonizes the reduction of Oxidized-Flavin Adenine Dinucleotide (Ox-FAD). A further embodiment is directed to a composition which reduces, inhibits or antagonizes the activity of flavokinase. A further embodiment is directed to a composition which reduces, inhibits or antagonizes the activity of FAD pyrophosphorylase. A further embodiment is directed to a composition which binds to or interacts with Re-FAD or Ox-FAD. A further embodiment is directed to a composition which binds to or interacts with flavokinase or FAD pyrophosphorylase.

A further preferred embodiment is directed to a composition or compound which increases, agonizes or promotes the activity of cystathionine beta-synthase. A preferred composition which increases, agonizes or promotes the activity of cystathionine beta-synthase comprises S-adenosylmethionine or homocysteine. Another preferred composition which increases, agonizes or promotes the activity of cystathionine beta-synthase is pyridoxine or derivative thereof.

A further preferred embodiment of the invention is directed to a method of screening for a composition which binds to or interacts with DAO, DDO, Re-FAD, Ox-FAD, flavokinase, FAD pyrophosphorylase, cystathionine beta synthase, L-amino acid oxidase, or glutamine transaminase. A further preferred embodiment of the invention is directed to a method of screening for a composition which reduces, inhibits or antagonizes the activity of DAO, DDO, flavokinase, FAD pyrophosphorylase, L-amino acid oxidase, or glutamine transaminase. A further preferred embodiment of the invention is directed to a method of screening for a composition which promote, increase, or agonize the activity of cystathionine beta synthase, L-amino acid oxidase, or glutamine transaminase.

Thus, in one aspect is provided a method of identifying a candidate molecule for the treatment of disease or for increasing the levels of or decreasing the degradation of amino acids, said method comprising : a) contacting a DAO, DDO, flavokinase, FAD pyrophosphorylase, L-amino acid oxidase, or glutamine transaminase polypeptide or a biologically active fragment thereof with a test compound; and b) determining whether said compound selectively binds to said polypeptide; wherein a determination that said compound selectively binds to said polypeptide indicates that said compound is a candidate molecule for the treatment of disease or for increasing the levels of or decreasing the degradation of amino acids.

Also provided is a method of identifying a candidate molecule for the treatment of disease or for increasing the levels of or decreasing the degradation of amino acids, said method comprising: a) contacting a DAO, DDO, flavokinase, FAD pyrophosphorylase, L-amino acid oxidase, or glutamine transaminase polypeptide or a biologically active fragment thereof polypeptide with a test compound; and b) determining whether said compound selectively inhibits the activity of said polypeptide; wherein a determination that said compound selectively inhibits the activity of said polypeptide indicates that said compound is a candidate molecule for the treatment of disease or for increasing the levels of or decreasing the degradation of amino acids.

Also described is a method of identifying a candidate molecule for the treatment of disease or for increasing the levels of or decreasing the degradation of amino acids, said method comprising : a) providing a cell comprising a DAO, DDO, flavokinase, FAD pyrophosphorylase, L-amino acid oxidase, or glutamine transaminase polypeptide or a biologically active fragment thereof; b) contacting said cell with a test compound; and c) determining whether said compound selectively inhibits DAO, DDO, flavokinase, FAD pyrophosphorylase, L-amino acid oxidase, or glutamine transaminase activity; wherein a determination that said compound selectively inhibits the activity of said polypeptide indicates that said compound is a candidate molecule for the treatment of disease or for increasing the levels of or decreasing the degradation of amino acids.

Further provided is a method of identifying a candidate molecule for the treatment of disease or for increasing the levels of or decreasing the degradation of amino acids, said method comprising: a) contacting a cystathionine beta synthase, L-amino acid oxidase, or glutamine transaminase polypeptide or a biologically active fragment thereof polypeptide with a test compound; and b) determining whether said compound selectively increases the activity of said polypeptide; wherein a determination that said compound selectively increases the activity of said polypeptide indicates that said compound is a candidate molecule for the treatment of disease or for increasing the levels of or decreasing the degradation of amino acids.

Another embodiment is method of identifying a candidate molecule for the treatment of disease or for increasing the levels of or decreasing the degradation of amino acids, said method comprising : a) providing a cell comprising a cystathionine beta synthase, L-amino acid oxidase, or glutamine transaminase polypeptide or a biologically active fragment thereof; b) contacting said cell with a test compound; and c) determining whether said compound selectively increases cystathionine beta synthase, L-amino acid oxidase, or glutamine transaminase activity; wherein a determination that said compound selectively increases the activity of said polypeptide indicates that said compound is a candidate molecule for the treatment of disease or for increasing the levels of or decreasing the degradation of amino acids.

A further preferred embodiment of the invention is directed to a method of antagonizing, reducing or inhibiting DAO activity in vitro. Further preferred is a method of antagonizing, reducing or inhibiting DAO activity in vivo. Further preferred is a method of antagonizing, reducing or inhibiting DAO activity in vitro or in vivo comprising the step of contacting DAO with a composition which reduces, inhibits or antagonizes the activity of DAO. A preferred activity of DAO to be inhibited is the oxidation of a substrate, preferably the substrate is a D-Amino Acid, preferably the D-amino acid is D-Ser, D-Asp, or N-methyl-D-Asp.

A further preferred embodiment of the invention is directed to a method of antagonizing, reducing or inhibiting DDO activity in vitro. Further preferred is a method of antagonizing, reducing or inhibiting DDO activity in vivo. Further preferred is a method of antagonizing, reducing or inhibiting DDO activity in vitro or in vivo comprising the step of contacting DDO with a composition which reduces, inhibits or antagonizes the activity of DDO. A preferred activity of DDO to be inhibited is the oxidation of a substrate, preferably the substrate is a D-Amino Acid, preferably the D-amino acid is D-Asp, D-Glu, D-Asn, D-Gln, D-Asp-dimethyl-ester or N-methyl-D-Asp.

Another embodiment of the invention is directed to compositions which increase the levels of at least one D-amino acid in vitro. Further preferred are compositions which increase the levels of at least one D-amino acid in vivo, preferably in tissues of mammals, further preferably in tissues of mice, rats, dogs, cows, pigs, apes, monkeys or humans. Still further preferred are compositions which increase levels of at least one D-amino acid in tissues of the central nervous system, preferably the brain or spinal cord. Still further preferred are compositions which increase levels of at least one D-amino acid in tissues of the brain, preferably the hippocampus, amygdala, substantia nigra, cerebellum, corpus callosum, caudate nucleus, cerebral cortex, thalamus, or pituitary gland. Other preferred tissues in which compositions of the invention increase levels of at least one D-amino acid include, but are not limited to the kidney, liver, adipose, muscle, and testis.

A preferred embodiment of the invention is directed to a use of a polypeptide of SEQ ID NO:15, or a fragment thereof, in a method to increase DAO activity. Further preferred is a use of a polypeptide of SEQ ID NO:15, or a fragment thereof, in a method to increase DDO activity. Further preferred is a use of a polypeptide of SEQ ID NO:15, or a fragment thereof, in a method to decrease serine racemase activity.

A preferred embodiment of the invention is directed to a use of a polypeptide of SEQ ID NO:15, or a fragment thereof, in a method of increasing production of compounds or compositions which are the product of a reaction involving DAO as a catalyst.

A preferred embodiment of the invention is directed to a method of screening for compositions or compounds that bind to g34872 polypeptides (SEQ ID NO:15) or g34872 polynucleotides (SEQ ID NO:14), or fragments thereof. Further preferred is a method of contacting g34872 polypeptides, or

fragments thereof, with DAO thereby increasing DAO activity above a basal level. Further preferred is a method of reducing, inhibiting, antagonizing or blocking the interaction of DAO and g34872. Further preferred is a method of treating a CNS disorder by blocking the interaction of g34872 and DAO. Further preferred is a method of treating a CNS disorder with a compound or composition which reduces, blocks, inhibits or antagonizes the interaction between g34872 and DAO.

The preferred DAO polypeptides of the invention include polypeptides of SEQ ID NO:7-10 and 19, and fragments thereof as well as polynucleotides that encode the same. The preferred DDO polypeptides of the invention include polypeptides of SEQ ID NO:22 and 23, and fragments thereof, as well as polynucleotides that encode the same. Preferred DAO polynucleotides of the invention include SEQ ID NO:2-6, and 18, and fragments thereof, as well as polypeptides encoded by the same. Preferred DDO polynucleotides of the invention include SEQ ID NO:20 and 21, and fragments thereof, as well as polypeptides encoded by the same.

Preferred biallelic markers of DAO are described in SEQ ID NO:1, as well as represented by 47-mers of marker 24-1443-126 (SEQ ID NO:24), marker 24-1457-52 (SEQ ID NO:26), and marker 24-1461-256 (SEQ ID NO:29).

Another embodiment of the invention is directed at compositions which differentially bind to polypeptides of SEQ ID NO:7. Another embodiment of the invention is directed at compositions which differentially bind to polypeptides of SEQ ID NO:8. Another embodiment of the invention is directed at compositions which differentially bind to polypeptides of SEQ ID NO:9. Another embodiment of the invention is directed at compositions which differentially bind to polypeptides of SEQ ID NO:10. Further preferred are compositions which bind to polypeptides of SEQ ID NO:10 but not to polypeptides of SEQ ID NO:7, 8, or 9. Further preferred are compositions which bind to polypeptides of SEQ ID NO:9 but not to polypeptides of SEQ ID NO:7, 8, or 10. Further preferred are compositions which bind to polypeptides of SEQ ID NO:8 but not to polypeptides of SEQ ID NO:7, 9, or 10. Further preferred are compositions which bind to polypeptides of SEQ ID NO:7 but not to polypeptides of SEQ ID NO:8, 9, or 10. Further preferred are compositions which bind to polypeptides of SEQ ID NO:8, 9, or 10 but not to polypeptides of SEQ ID NO:7.

Another embodiment of the invention is directed to a composition which differentially binds to a monomeric polypeptide comprising SEQ ID NO:7, 8, 9, 10, or 15, or a polypeptide fragment thereof. Further preferred is a composition which binds to a monomeric polypeptide of SEQ ID NO:7, or a fragment thereof, but not to a homo- or hetero- multimeric form comprising at least a monomer of a polypeptide of SEQ ID NO:7, or a fragment thereof. Further preferred is a composition which binds to a monomeric polypeptide of SEQ ID NO:8, or a fragment thereof, but not to a homo- or hetero- multimeric form comprising at least a monomer of a polypeptide of SEQ ID NO:8, or a fragment thereof. Further preferred is a composition which binds to a monomeric polypeptide of SEQ ID NO:9,

or a fragment thereof, but not to a homo- or hetero- multimeric form comprising at least a monomer of a polypeptide of SEQ ID NO:9, or a fragment thereof. Further preferred is a composition which binds to a monomeric polypeptide of SEQ ID NO:10, or a fragment thereof, but not to a homo- or hetero- multimeric form comprising at least a monomer of a polypeptide of SEQ ID NO:10, or a fragment thereof. Further preferred is a composition which binds to a monomeric polypeptide of SEQ ID NO:15, or a fragment thereof, but not to a homo- or hetero- multimeric form comprising at least a monomer of a polypeptide of SEQ ID NO:15, or a fragment thereof.

Another embodiment of the invention is directed to a composition which binds to a multimeric polypeptide comprising at least one polypeptide of SEQ ID NO:7, 8, 9, 10, or 15, or a fragment thereof. Further preferred is a composition which binds to a homo- or hetero- multimeric form comprising at least one monomer of a polypeptide of SEQ ID NO:7, or a fragment thereof, but does not bind to a monomeric polypeptide of SEQ ID NO:7, or a fragment thereof. Another embodiment of the invention is directed to a composition which binds to a homo- or hetero- multimeric form comprising at least one monomer of a polypeptide of SEQ ID NO:8, or a fragment thereof, but does not bind to a monomeric polypeptide of SEQ ID NO:8, or a fragment thereof. Another embodiment of the invention is directed to a composition which binds to a homo- or hetero- multimeric form comprising at least one monomer of a polypeptide of SEQ ID NO:9, or a fragment thereof, but does not bind to a monomeric polypeptide of SEQ ID NO:9, or a fragment thereof. Another embodiment of the invention is directed to a composition which binds to a homo- or hetero- multimeric form comprising at least one monomer of a polypeptide of SEQ ID NO:10, or a fragment thereof, but does not bind to a monomeric polypeptide of SEQ ID NO:10, or a fragment thereof. Another embodiment of the invention is directed to a composition which binds to a homo- or hetero- multimeric form comprising at least one monomer of a polypeptide of SEQ ID NO:15, or a fragment thereof, but does not bind to a monomeric polypeptide of SEQ ID NO:15, or a fragment thereof.

Another embodiment of the invention is directed at compositions which differentially bind to polynucleotides of SEQ ID NO:2. Another embodiment of the invention is directed at compositions which differentially bind to polynucleotides of SEQ ID NO:3. Another embodiment of the invention is directed at compositions which differentially bind to polynucleotides of SEQ ID NO:4. Another embodiment of the invention is directed at compositions which differentially bind to polynucleotides of SEQ ID NO:5. Another embodiment of the invention is directed at compositions which differentially bind to polynucleotides of SEQ ID NO:6. Further preferred are compositions which bind to polynucleotides of SEQ ID NO:6 but not to polynucleotides of SEQ ID NO:2, 3, 4, or 5. Further preferred are compositions which bind to polynucleotides of SEQ ID NO:5 but not to polynucleotides of SEQ ID NO:2, 3, 4, or 6. Further preferred are compositions which bind to polynucleotides of SEQ ID NO:4 but not to polynucleotides of SEQ ID NO:2, 3, 5, or 6. Further preferred are compositions which

bind to polynucleotides of SEQ ID NO:3 but not to polynucleotides of SEQ ID NO:2, 4, 5, or 6. Further preferred are compositions which bind to polynucleotides of SEQ ID NO:2 but not to polynucleotides of SEQ ID NO:3, 4, 5, or 6. Further preferred are compositions which bind to polynucleotides of SEQ ID NO:3, 4, 5, or 6 but not to polynucleotides of SEQ ID NO:2.

A further preferred embodiment of the invention is directed to a genomic sequence comprising polynucleotides of SEQ ID NO:1. Further preferred are methods to genotype regions of the polynucleotides of SEQ ID NO:1.

An embodiment of the invention is directed to a purified or isolated nucleic acid comprising the sequence of SEQ ID NO:1 or complement thereof. Further preferred is a purified or isolated nucleic acid comprising at least 10 consecutive nucleotides of the sequence of SEQ ID NO:1 or complement thereof. Still further preferred is a nucleic acid comprises at least 15 consecutive nucleotides of the sequence of SEQ ID NO:1 or complement thereof.

An another embodiment of the invention is directed to a purified or isolated nucleic acid comprising at least 10 consecutive nucleotides of the sequence of SEQ ID NO:1, or complement thereof, of one or more exons. Further preferred is a purified or isolated nucleic acid of SEQ ID NO:1, or complement thereof, comprising the sequence of at least 10 consecutive nucleotides from nucleotides 40389 to 40670 of SEQ ID NO:1, or complement thereof. Also preferred is a purified or isolated nucleic acid of SEQ ID NO:1, or complement thereof, comprising the sequence of at least 10 consecutive nucleotides from nucleotides 42666 to 42778 of SEQ ID NO:1, or complement thereof. Also preferred is a purified or isolated nucleic acid of SEQ ID NO:1, or complement thereof, comprising the sequence of at least 10 consecutive nucleotides from nucleotides 43416 to 43519 of SEQ ID NO:1, or complement thereof. Also preferred is a purified or isolated nucleic acid of SEQ ID NO:1, or complement thereof, comprising the sequence of at least 10 consecutive nucleotides from nucleotides 61159 to 61402 of SEQ ID NO:1, or complement thereof. Also preferred is a purified or isolated nucleic acid of SEQ ID NO:1, or complement thereof, comprising the sequence of at least 10 consecutive nucleotides from nucleotides 64050 to 64711 of SEQ ID NO:1, or complement thereof. Also preferred is a purified or isolated nucleic acid of SEQ ID NO:1, or complement thereof, comprising the sequence of at least 10 consecutive nucleotides from nucleotides 68126 to 68261 of SEQ ID NO:1, or complement thereof. Also preferred is a purified or isolated nucleic acid of SEQ ID NO:1, or complement thereof, comprising the sequence of at least 10 consecutive nucleotides from nucleotides 84906 to 85541 of SEQ ID NO:1, or complement thereof.

A further preferred embodiment of the invention is directed to a purified or isolated nucleic acid comprising the sequence of SEQ ID NO:2 or complement thereof. A still further preferred embodiment of the invention directed to a purified or isolated nucleic acid comprising the sequence of SEQ ID NO:3 or complement thereof. Another further preferred embodiment of the invention directed to a purified or

isolated nucleic acid comprising the sequence of SEQ ID NO:4 or complement thereof. Another further preferred embodiment of the invention directed to a purified or isolated nucleic acid comprising the sequence of SEQ ID NO:5 or complement thereof. Another further preferred embodiment of the invention directed to a purified or isolated nucleic acid comprising the sequence of SEQ ID NO:6 or complement thereof. Another further preferred embodiment of the invention directed to a purified or isolated nucleic acid comprising the sequence of SEQ ID NO:14 or complement thereof. Another further preferred embodiment of the invention directed to a purified or isolated nucleic acid comprising the sequence of SEQ ID NO:16 or complement thereof. Another further preferred embodiment of the invention directed to a purified or isolated nucleic acid comprising the sequence of any one of the sequences of SEQ ID NO:18, 20, or 21, or complement thereof.

Another embodiment of the invention is directed to a purified or isolated nucleic acid comprising at least 10 consecutive nucleotides of at least one of the sequences of SEQ ID NO:2-6, or complement thereof. Further preferred is a purified or isolated nucleic acid comprising at least 15 consecutive nucleotides of at least one of the sequences of SEQ ID NO:2-6, or complement thereof.

Another embodiment of the invention is directed to a purified or isolated nucleic acid comprising at least 10 consecutive nucleotides of the sequence of SEQ ID NO:14, or complement thereof. Further preferred is a purified or isolated nucleic acid comprising at least 15 consecutive nucleotides of the sequence of SEQ ID NO:14, or complement thereof.

Another embodiment of the invention is directed to a purified or isolated nucleic acid encoding the polypeptide of SEQ ID NO:7. Further preferred is a purified or isolated nucleic acid encoding at least 10 consecutive amino acids of the polypeptide of SEQ ID NO:7. Still further preferred is a purified or isolated nucleic acid, wherein said nucleic acid encodes at least 15 consecutive amino acids of the polypeptide of SEQ ID NO:7.

Another embodiment of the invention is directed to a purified or isolated nucleic acid encoding the polypeptide of SEQ ID NO:8. Further preferred is a purified or isolated nucleic acid encoding at least 10 consecutive amino acids of the polypeptide of SEQ ID NO:8. Still further preferred is a purified or isolated nucleic acid, wherein said nucleic acid encodes at least 15 consecutive amino acids of the polypeptide of SEQ ID NO:8.

Another embodiment of the invention is directed to a purified or isolated nucleic acid encoding the polypeptide of SEQ ID NO:9. Further preferred is a purified or isolated nucleic acid encoding at least 10 consecutive amino acids of the polypeptide of SEQ ID NO:9. Still further preferred is a purified or isolated nucleic acid, wherein said nucleic acid encodes at least 15 consecutive amino acids of the polypeptide of SEQ ID NO:9.

Another embodiment of the invention is directed to a purified or isolated nucleic acid encoding the polypeptide of SEQ ID NO:10. Further preferred is a purified or isolated nucleic acid encoding at

least 10 consecutive amino acids of the polypeptide of SEQ ID NO:10. Still further preferred is a purified or isolated nucleic acid, wherein said nucleic acid encodes at least 15 consecutive amino acids of the polypeptide of SEQ ID NO:10.

Another embodiment of the invention is directed to a purified or isolated
5 nucleic acid encoding the polypeptide of SEQ ID NO:15. Further preferred is a purified or isolated nucleic acid encoding at least 10 consecutive amino acids of the polypeptide of SEQ ID NO:15. Still further preferred is a purified or isolated nucleic acid, wherein said nucleic acid encodes at least 15 consecutive amino acids of the polypeptide of SEQ ID NO:15.

10 Another embodiment of the invention is directed to a purified or isolated nucleic acid encoding the polypeptide of SEQ ID NO:17. Further preferred is a purified or isolated nucleic acid encoding at least 10 consecutive amino acids of the polypeptide of SEQ ID NO:17. Still further preferred is a purified or isolated nucleic acid, wherein said nucleic acid encodes at least 15 consecutive amino acids of the
15 polypeptide of SEQ ID NO:17.

Another embodiment of the invention is directed to a purified or isolated nucleic acid encoding the polypeptide of SEQ ID NO:19. Further preferred is a purified or isolated nucleic acid encoding at least 10 consecutive amino acids of the polypeptide of SEQ ID NO:19. Still further preferred is a purified or isolated nucleic
20 acid, wherein said nucleic acid encodes at least 15 consecutive amino acids of the polypeptide of SEQ ID NO:19.

Another embodiment of the invention is directed to a purified or isolated nucleic acid encoding the polypeptide of SEQ ID NO:22. Further preferred is a purified or isolated nucleic acid encoding at least 10 consecutive amino acids of the
25 polypeptide of SEQ ID NO:22. Still further preferred is a purified or isolated nucleic acid, wherein said nucleic acid encodes at least 15 consecutive amino acids of the polypeptide of SEQ ID NO:22.

Another embodiment of the invention is directed to a purified or isolated nucleic acid encoding the polypeptide of SEQ ID NO:23. Further preferred is a
30 purified or isolated nucleic acid encoding at least 10 consecutive amino acids of the polypeptide of SEQ ID NO:23. Still further preferred is a purified or isolated nucleic acid, wherein said nucleic acid encodes at least 15 consecutive amino acids of the polypeptide of SEQ ID NO:23.

A further preferred embodiment of the invention is directed at the biallelic
35 markers.

In another aspect, the present invention provides an isolated or purified nucleic acid encoding a DAO polypeptide selected from the group consisting of:

- a) a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence selected from the group of sequences consisting of SEQ ID NOs:8 to 10; and
- b) a nucleic acid molecule comprising a nucleic acid sequence selected from the group of sequences consisting of SEQ ID NOs:1 and 3 to 6, or a sequence complementary thereto.

In a further aspect, the present invention provides a purified or isolated DAO polypeptide selected from the group consisting of:

- a) a polypeptide comprising an amino acid sequence selected from the group of sequences consisting of SEQ ID NOs:8 to 10;
- b) a polypeptide encoded by a nucleic acid molecule comprising a nucleic acid sequence selected from the group of sequences consisting of SEQ ID NOs:1 and 3 to 6.

In yet another aspect, the present invention provides a method of identifying a candidate molecule for the treatment of schizophrenia, depression or bipolar disorder, said method comprising:

- a) contacting a DAO or DDO or a biologically active fragment thereof polypeptide with a test compound; and
- b) determining whether said compound selectively inhibits the activity of said polypeptide;

wherein a test compound that selectively inhibits the activity of said polypeptide is identified as a candidate molecule for the treatment of schizophrenia, depression or bipolar disorder.

In another aspect, the present invention provides a method of identifying a candidate molecule for the treatment of schizophrenia, depression or bipolar disorder, said method comprising:

- a) contacting a DAO or DDO or a biologically active fragment thereof polypeptide with a test compound; and
- b) determining whether said compound selectively binds said polypeptide;

wherein a test compound that selectively binds said polypeptide is identified as a candidate molecule for the treatment of schizophrenia, depression or bipolar disorder.

In a further aspect, the present invention provides a method of screening for antagonists of a DAO polypeptide, comprising the steps of:

- a) contacting a test compound with a DAO polypeptide of the invention;

- b) detecting the level of DAO activity; and
- c) comparing the activity to the activity of a control test without the test compound, whereby a decrease in the level of the DAO activity over the control indicates that the test compound is an antagonist of DAO.

5 In yet a further aspect, the present invention provides a method of screening for compounds that inhibit the expression of the DAO mRNA or polypeptide, comprising the steps of:

- a) incubating cells expressing a DAO polypeptide of the invention, (i) in the presence and (ii) in the absence, of a test compound; and
- 10 b) detecting the level of the DAO mRNA or polypeptide in the cells.

In another aspect, the present invention provides a method of detecting an association with schizophrenia, depression or bipolar disorder, comprising determining the identity of a nucleotide at a DAO-related biallelic marker, or the complement thereof in a biological sample.

- 15 The present invention further provides a method of detecting an association between a DAO-related biallelic marker and schizophrenia, depression or bipolar disorder, comprising determining the identity of a polymorphic nucleotide at a DAO-related biallelic marker, or the complement thereof in a biological sample, wherein the polymorphic nucleotide is associated with schizophrenia, depression or bipolar
- 20 disorder.

Preferably, said determining is performed by a hybridization assay, a sequencing assay, a microsequencing assay, or an enzyme-based mismatch detection assay.

In a further aspect, the present invention provides a method of estimating the frequency of an allele of a biallelic marker in a population comprising;

- 25 a) genotyping individuals from said population for said biallelic marker according to the method of the invention; and
- b) determining the proportional representation of said biallelic marker in said population.

In another aspect, the present invention provides a method of detecting an

30 association between a genotype and a trait, comprising the steps of:

- a) determining the frequency of at least one biallelic marker in trait positive population according to the method of the invention;
- b) determining the frequency of at least one biallelic marker in a control population according to the method of the invention; and
- 35 c) determining whether a statistically significant association exists between said genotype and said trait.

In another aspect, the present invention provides a method of determining whether an individual is at risk of schizophrenia, depression or bipolar disorder, comprising:

- a) genotyping at least one biallelic marker according to the method of the invention;
- 5 and
- b) correlating the result of step a) with a risk of developing schizophrenia, depression or bipolar disorder.

Preferably, said biallelic marker is selected from the group of markers described in SEQ ID NOs:1, 24, 26, and 29, and the complements thereof.

- 10 In a further aspect, the present invention provides a method of diagnosing, or detecting a predisposition to or susceptibility to schizophrenia, depression or bipolar disorder in a subject, comprising

- a) obtaining a nucleic acid sample from said subject; and
- b) determining the identity of a nucleotide at a DAO-related polymorphism, or the
- 15 complement thereof in said biological sample.

The present invention further provides a method of diagnosing, or detecting a predisposition to or susceptibility to schizophrenia, depression or bipolar disorder in a subject, comprising

- a) obtaining a nucleic acid sample from said subject;
- 20 b) determining the identity of a nucleotide at a DAO-related polymorphism, or the complement thereof in said biological sample; and
- c) determining whether the subject has a DAO polymorphism associated with schizophrenia, depression or bipolar disorder.

- 25 In another aspect, the present invention provides for the use of a compound capable of inhibiting the conversion of a D-amino acid into the corresponding α -keto acid for the manufacture of a medicament for the treatment of schizophrenia or bipolar disorder.

Preferably, said compound inhibits the activity of a DAO or DDO polypeptide.

- 30 In another aspect, the present invention provides for the use of an antagonist of DAO or DDO in the manufacture of a medicament for the treatment of psychosis, schizophrenia, bipolar disorder, depression or a mood disorder.

- 35 In another aspect, the present invention provides a method of treating psychosis, schizophrenia, bipolar disorder, depression or a mood disorder in a patient in need thereof, comprising administering said patient an effective amount of a composition or compound comprising a DAO antagonist or a DDO antagonist or inhibitor.

Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the
5 field relevant to the present invention as it existed before the priority date of each claim of this application.

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of
10 any other element, integer or step, or group of elements, integers or steps.

Brief description of the Drawings

Figure 1 demonstrates the activity of yeast expressed recombinant g34782 and DAO polypeptides.

15 Figure 2 demonstrates the activity of bacterial expressed recombinant g34872 and DAO polypeptides.

Figure 3 demonstrates the in vitro activation of purified DAO by g34872 using D-serine as a substrate.

Figure 4 demonstrates the dose dependent affect of g34872 on DAO activity.

Figure 5 demonstrates the kinetics of the interaction between g34872 and DAO.

Figure 6 is a table demonstrating the results of a DAO biallelic marker association analysis between French Canadian schizophrenia cases and controls.

BRIEF DESCRIPTION OF THE SEQUENCES PROVIDED IN THE SEQUENCE LISTING

- SEQ ID NO:1 genomic sequence of DAO;
- SEQ ID NO:2 DAO cDNA;
- SEQ ID NO:3 novel cDNA with Exons U 2 3 4 5 6 7 8 9 10 11Long;
- SEQ ID NO:4 novel cDNA with Exons B C Ulong V 2 3 4 5 6 7 9 10 11Long;
- SEQ ID NO:5 novel cDNA with Exons U 2 4 5 6 7 8 9 10 11Long;
- SEQ ID NO:6 novel cDNA with Exons B 2 3 7 8 9 10 11;
- SEQ ID NO:7 polypeptide of DAO from cDNA of SEQ ID NO:2 and 3;
- SEQ ID NO:8 polypeptide of DAO from cDNA of SEQ ID NO:4;
- SEQ ID NO:9 polypeptide of DAO from cDNA of SEQ ID NO:5;
- SEQ ID NO:10 polypeptide of DAO from cDNA of SEQ ID NO:6;
- SEQ ID NO:11-12 polynucleotides comprising g34872 biallelic markers 99/16105-152 and 99/5919-215;
- SEQ ID NO:13 polynucleotides of g34872, including polymorphisms;
- SEQ ID NO:14 polypeptides of g34872, wherein the amino acid at position 10 is tyrosine or serine, the amino acid at position 30 is lysine or arginine, the amino acid at position 50 is glutamate or a premature stop, the amino acid at position 60 is arginine or glycine, and the amino acid at position 115 is aspartate or alanine;
- SEQ ID NO:15 g34872 polynucleotide encoding polypeptide of SEQ ID NO:16 used in 2-Hybrid experiments;
- SEQ ID NO:17 polynucleotide of DAO encoding polypeptide of SEQ ID NO:18;
- SEQ ID NOs:19 and 20 polynucleotides of DDO encoding polypeptides of SEQ ID NOs:21 and 22, respectively; and
- SEQ ID NOs:23-26 polynucleotides comprising DAO biallelic markers 24-1443/126, 24-1457/52, 27-93/181, and 24-1461/256, respectively, noting polymorphic base at position 24.

The g34872 genomic sequence and biallelic markers are described in SEQ ID NO:1 of US Patent Application No:09/539,333 and International Patent Application No:PCT/IB00/00435, which disclosures are hereby incorporated by reference in their entireties.

In accordance with the regulations relating to Sequence Listings, the following codes have been used in the Sequence Listing to indicate the locations of biallelic markers within the sequences and to identify each of the alleles present at the polymorphic base. The code "r" in the sequences indicates that one allele of the polymorphic base is a guanine, while the other allele is an adenine. The code "y" in the sequences indicates that one allele of the polymorphic base is a thymine, while the other allele is a cytosine. The code "m" in the sequences indicates that one allele of the polymorphic base is an adenine, while the other allele is a cytosine. The code "k" in the sequences indicates that one allele of the polymorphic base is a guanine, while the other allele is a thymine. The code "s" in the sequences indicates that one allele of the polymorphic base is a guanine, while the other allele is a cytosine. The code "w" in the sequences indicates that one allele of the polymorphic base is an adenine, while the other allele is a thymine.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to methods for providing prevention of a CNS disorder to a subject susceptible to such a disorder, and for providing treatment to a subject suffering from a CNS disorder. In particular, the method comprises administering to a patient an amount of a DAO or DDO antagonist compound effective for providing some degree of prevention or amelioration of the progression of the CNS disorder (i.e., provide protective effects), amelioration of the symptoms of the CNS disorder, and amelioration of the reoccurrence of the CNS disorder.

CNS disorders which can be treated in accordance with the present invention include presenile dementia (early onset Alzheimer's disease), senile dementia (dementia of the Alzheimer's type), Parkinsonism including Parkinson's disease, Huntington's chorea, tardive dyskinesia, hyperkinesia, mania, attention deficit hyperactivity disorder (ADHD), attention deficit disorder (ADD), anxiety disorders, dyslexia, psychotic disorders, schizophrenia, bipolar disorder, major depressive episodes, manic episodes, hypomanic episodes, depression, autistic disorders, substance abuse, excessive aggression, tic disorders and Tourette's syndrome. Preferred disorders of the present invention include schizophrenia and bipolar disorder. Further preferred embodiments of schizophrenia and schizophreniform disorders include: schizophrenia (catatonic), schizophrenia (disorganized), schizophrenia (paranoid), schizophrenia (undifferentiated), schizophrenia (residual), schizophreniform disorder, brief reactive psychosis, schizoaffective disorder, induced psychotic disorder, schizotypal

personality disorder, schizoid personality disorder, paranoid personality disorder and delusional (paranoid) disorder.

The identification of genes involved in a particular trait such as a specific central nervous system disorder, like schizophrenia, can be carried out through two main strategies currently used for genetic mapping: linkage analysis and association studies. Linkage analysis requires the study of families with multiple affected individuals and is now useful in the detection of mono- or oligogenic inherited traits. Conversely, association studies examine the frequency of marker alleles in unrelated trait (T+) individuals compared with trait negative (T-) controls, and are generally employed in the detection of polygenic inheritance.

In the present application, additional biallelic markers located in the DAO gene associated with schizophrenia are disclosed. The identification of these biallelic markers in association with schizophrenia has allowed for the further definition of the chromosomal region suspected of containing a genetic determinant involved in a predisposition to develop schizophrenia and has resulted in the identification of novel gene sequences disclosed herein which are associated with a predisposition to develop schizophrenia. Furthermore, biallelic markers in the g34872 gene, previously described, as well as in the DAO gene presently described can be used alone or in combination to determine individuals at risk for developing a CNS disorder. Moreover, biallelic markers in the g34872 gene, previously described, as well as in the DAO gene presently described can be used alone or in combination to determine individuals who will benefit from the treatment described by the present invention. Additionally, the sequence information provides a resource for the further identification of new genes and markers in those regions. Additionally, the sequences comprising the the schizophrenia-associated genes are useful, for example, for the isolation of other genes in putative gene families, the identification of homologs from other species, treatment of disease and as probes and primers for diagnostic or screening assays as described herein. Furthermore, the identified polymorphisms are used in the design of assays for the reliable detection of genetic susceptibility to schizophrenia and bipolar disorder. They are also used in the design of drug screening protocols to provide an accurate and efficient evaluation of the therapeutic and side-effect potential of new or already existing medicament or treatment regime.

Definitions

The term "treat" or "treating" means to ameliorate, alleviate symptoms, eliminate the causation of the symptoms either on a temporary or permanent basis, or to prevent or slow the appearance of symptoms of the named disorder or condition.

The dose of the compound is that amount effective to prevent occurrence of the symptoms of the disorder or to treat some symptoms of the disorder from which the patient suffers. By "effective amount", "therapeutically effective amount" "therapeutic amount" or "effective dose" is meant that

amount sufficient to elicit the desired pharmacological or therapeutic effects, thus resulting in effective prevention or treatment of the disorder. Prevention of the disorder is manifested by delaying the onset of the symptoms of the disorder to a medically significant extent. Treatment of the disorder is manifested by a decrease in the symptoms associated with the disorder or an amelioration of the reoccurrence of the symptoms of the disorder. A therapeutically effective amount of a compound of the present invention can be easily determined by one skilled in the art by administering a quantity of a compound to an individual and observing the result. In addition, those skilled in the art are familiar with identifying individuals having a CNS disorder readily able to identify individuals who suffer from the CNS disorder.

The terms "antagonist" and "inhibitor" are considered to be synonymous and can be used interchangeably throughout the disclosure. The "antagonist" compounds of the invention may be administered together with a typical or atypical anti-CNS disorder drug, such as an antipsychotic drug. Typical antipsychotics include: haloperidol, fluphenazine, perphenazine, chlorpromazine, molindone, pimozide, trifluoperazine and thioridazine, thiadiazole, oxadiazole and others. Atypical antipsychotics include: clozapine, risperidone, olanzapine, sertindole, M100907, ziprasidone, seroquel, zotepine, amisulpride, iloperidone, phenelzine and others. Typical antidepressant and anti-anxiety agents include: heterocyclic antidepressants (TCAs, tetracyclics, and the like), SSRIs, mixed serotonin and norepinephrine reuptake inhibitors, dopamine reuptake inhibitors and MAOIs. The antagonists may also be used to treat individuals for whom the above drugs are contraindicated. The present invention also provides a method for the treatment or prevention of schizophrenia, bipolar disorder, or other CNS disorders without concomitant therapy with other antipsychotic, antidepressant, anti-anxiety, or other drugs, in a patient who is non-responsive. The antipsychotic, antidepressant, anti-anxiety, or other drugs may be administered at a subtherapeutic doses, i.e., at a lower dose than the dosage that is typically used for treatments with the above drugs alone. Drugs used for the treatment of schizophrenia, bipolar disorder, depression, and other CNS disorders, that are either recognized as a DAO or DDO inhibitor or that inherently act as an inhibitor of DAO or DDO are specifically excluded from the definition of DAO or DDO "antagonist" and may be specifically excluded from the present invention. Further, any molecule, compound or drug disclosed herein may be specifically excluded from the invention.

"Alkyl" means a branched or unranked saturated hydrocarbon chain containing 1 to 8 carbon atoms, such as methyl, ethyl, propyl, iso-propyl, butyl, iso-butyl, tert-butyl, n-pentyl, n-hexyl, and the like, unless otherwise indicated.

"Alkoxy" means the group --OR wherein R is alkyl as herein defined. Preferably, R is a branched or unbranched saturated hydrocarbon chain containing 1 to 3 carbon atoms.

"Halo" means fluoro, chloro, bromo, or iodo, unless otherwise indicated.

"Phenyl" includes all possible isomeric phenyl radicals, optionally monosubstituted or multi-substituted with substituents selected from the group consisting of alkyl, alkoxy, hydroxy, halo, and haloalkyl.

Preferred heteroaryl rings include pyrrole, furan, thiophene, pyridine, pyrimidine, pyridazine, pyrazine, triazole, tetrazole, pyrazole, imidazole, isothiazole, thiazole, isoxazole and oxazole. Preferred "heteroaryl fused to phenyl" rings indole, isoindole, benzofuran, benzothiophene, quinoline, isoquinoline, quinoxaline, quinazoline, benzotriazole, indazole, benzimidazole, benzothiazole, benzisoxazole, and benzoxazole. It is assumed that "heteroaryl fused to phenyl" rings are included when using the term heteroaryl rings. The term "saturated or partially unsaturated heterocycloalkyl ring" means a saturated or partially unsaturated (but not aromatic, or fully saturated) heterocycle having 5-7 ring atoms, and containing 1-3 heteroatoms selected from N, O, or S. Preferred saturated or partially unsaturated heterocycloalkyl rings include piperidine, piperazine, morpholine, tetrahydropyran, thiomorpholine, or pyrrolidine.

The term "pharmaceutically acceptable salt" refers to salts of the subject compounds which possess the desired pharmacological activity and which are neither biologically nor otherwise undesirable. The salts can be formed with inorganic acids such as acetate, adipate, alginate, aspartate, benzoate, benzenesulfonate, bisulfate, butyrate, citrate, camphorate, camphorsulfonate, cyclopentanepropionate, digluconate, dodecylsulfate, ethanesulfonate, fumarate, glucoheptanoate, glycerophosphate, hemisulfate, heptanoate, hexanoate, hydrochloride, hydrobromide, hydroiodide, 2-hydroxyethanesulfonate, lactate, maleate, methanesulfonate, 2-naphthalenesulfonate, nicotinate, oxalate, thiocyanate, tosylate and undecanoate. Base salts include ammonium salts, alkali metal salts such as sodium and potassium salts, alkaline earth metal salts such as calcium and magnesium salts, salt with organic bases such as dicyclohexylamine salts, N-methyl-D-glucamine, and salts with amino acids such as arginine, lysine, and so forth. Also, the basic nitrogen-containing groups can be quarternized with such agents as lower alkyl halides, such as methyl, ethyl, propyl, and butyl chloride, bromides and iodides; dialkyl sulfates like dimethyl, diethyl, dibutyl and diamyl sulfates, long chain halides such as decyl, lauryl, myristyl and stearyl chlorides, bromides and iodides, aralkyl halides like benzyl and phenethyl bromides and others. Water or oil-soluble or dispersible products are thereby obtained. Furthermore, pharmaceutical and pharmaceutically acceptable compositions are described *infra*.

The compounds of this invention possess asymmetric centers and thus can be produced as mixtures of stereoisomers or as individual stereoisomers. The individual stereoisomers may be obtained by using an optically active starting material, by resolving a racemic or non-racemic mixture of an intermediate at some appropriate stage of the synthesis, or by resolution of the compound of formula (I). It is understood that the individual stereoisomers as well as mixtures (racemic and non-racemic) of

stereoisomers are encompassed by the scope of the present invention. The compounds of this invention possess at least one asymmetric centers and thus can be produced as mixtures of stereoisomers or as individual R- and S-stereoisomers. The individual enantiomers may be obtained by resolving a racemic or non-racemic mixture of an intermediate at some appropriate stage of the synthesis. It is understood that the individual R- and S- stereoisomers as well as mixtures of stereoisomers are encompassed by this invention.

"Isomers" are different compounds that have the same molecular formula.

"Stereoisomers" are isomers that differ only in the way the atoms are arranged in space.

"Enantiomers" are a pair of stereoisomers that are non-superimposable mirror images of each other.

"Diastereoisomers" are stereoisomers which are not mirror images of each other.

"Racemic mixture" means a mixture containing equal parts of individual enantiomers. "Non-racemic mixture" is a mixture containing unequal parts of individual enantiomers or stereoisomers.

"Substituted Alkyls" include carboxyalkyls such as acetyl, aminoalkyls, dialkylaminoalkyls, hydroxyalkyls and mercaptoalkyls, alkylsilyl.

The present invention relates to compounds of Formulae I-VI including , but not limited to the specific examples presented herein. Further, any of these compounds may take the form of a pharmaceutically acceptable salt.

It should be appreciated that the compounds of the invention described herein can be synthesized by an artisan skilled in the art of organic chemistry.

The term "psychotic condition" as used herein means pathologic psychological conditions which are psychoses or may be associated with psychotic features. Such conditions include, but are not limited to the psychotic disorders which have been characterized in the DSM-IV-R, Diagnostic and Statistical Manual of Mental Disorders, Revised 4th Ed. (1994), including schizophrenia and acute mania. The DSM-IV-R was prepared by the Task Force on Nomenclature and Statistics of the American Association, and provides clear descriptions of diagnostic categories. The skilled artisan will recognize that there are alternative nomenclatures, nosologies, and classification systems for pathologic psychological conditions and that these systems evolve with medical scientific progress.

The term "schizophrenia" encompasses, or alternatively may be specifically limited to, Schizophrenia, Schizophreniform Disorder, Schizoaffective Disorder, Delusional Disorder, Brief Psychotic Disorder, Psychotic Disorder Due to a General Medical Condition, Psychotic Disorder Not Otherwise Specified, or described elsewhere herein. The symptoms of these disorders are in large part as defined in the Diagnostic and Statistical Manual of Mental Disorder, fourth edition (DSMIV). The sections of the DSMIV that relate to these disorders are hereby incorporated by reference.

The term "bipolar disorder" as used herein refers to a condition characterized as a Bipolar Disorder, in the DSM-IV-R. Diagnostic and Statistical Manual of Mental Disorders, Revised, 3rd Ed. (1994) as category 296.xx. To further clarify, Applicants contemplate the treatment of both bipolar disorder I and bipolar disorder II as described in the DSM-IV-R. The term further includes cyclothymic disorder. Cyclothymic disorder refers to an alternation of depressive symptoms and hypomanic symptoms. The skilled artisan will recognize that there are alternative nomenclatures, nosologies, and classification systems for pathologic psychological conditions and that these systems evolve with medical scientific progress.

As used herein, the term "non-responsive" in relation to major depressive disorder means patients who have not had a reasonable clinical response (e.g. a 50% reduction in Hamilton Depression Scale (HAM-D) from a patient's baseline score after treatment with one or more clinical courses of conventional antidepressants).

A "major depressive episode" is defined as at least two weeks of depressed mood or loss of interest, which may be accompanied by other symptoms of depression. The symptoms must persist for most of the day (i.e. for at least two thirds of the patients' waking hours), nearly every day (i.e. for at least ten out of fourteen days) for at least two consecutive weeks. A "depressed mood" is often described by the patient as feeling sad, hopeless, helpless or worthless. The patient may also appear sad to an observer, for example, through facial expression, posture, voice and tearfulness. In children and adolescents, the mood may be irritable. A "loss of interest" is often described by the patient as feeling less interested in hobbies or not feeling any enjoyment in activities that were previously considered to be pleasurable.

A major depressive episode may be accompanied by other symptoms of depression including significant weight loss when not dieting or weight gain (e.g. a change of more than 5% body weight in one month), or decrease or increase in appetite; insomnia or hypersomnia; psychomotor agitation or retardation; fatigue or loss of energy; feelings of worthlessness or excessive or inappropriate guilt; diminished ability to think or concentrate; or indecisiveness; and recurrent thoughts of death, recurrent suicidal ideation with or without a specific plan, or a suicide attempt.

A "manic episode" is defined by a distinct period during which there is an abnormally and persistently elevated, expansive, or irritable mood. This period of abnormal mood must last at least 1 week (or less if hospitalization is required). The mood disturbance must be accompanied by at least three additional symptoms from a list that includes inflated self-esteem or grandiosity, decreased need for sleep, pressure of speech, flight of ideas, distractibility, increased involvement in goal-directed activities or psychomotor agitation, and excessive involvement in pleasurable activities with a high potential of painful consequences. If the mood is irritable (rather than elevated or expansive), at least four of the above symptoms must be present. The disturbance must be sufficiently severe to cause

marked impairment in social or occupational functioning or to require hospitalization, or it is characterized by the presence of psychotic features.

A "hypomanic episode" is less severe than a manic episode. The symptoms of a hypomanic episode are generally the same as those which define a manic episode, except that delusions and hallucinations are not present and the episode is not severe enough to cause marked impairment of social and occupational functioning or to require hospitalisation of the individual.

The term "autistic disorder" as used herein means a condition characterized as an Autistic Disorder in the DSM-IV-R as category 299.xx, including 299.00, 299.80, and 299.10, preferably 299.00.

The term "anxiety disorder" includes, but is not limited to, obsessive-compulsive disorder, psychoactive substance anxiety disorder, post-traumatic stress disorder, generalized anxiety disorder, anxiety disorder NOS, and organic anxiety disorder.

The term "substance abuse" as used herein means the undesired physical and/or psychological dependence on a drug. The term refers to dependence on a substance such as cocaine, psychedelic agents, marijuana, amphetamines, hallucinogen, phencyclidine, benzodiazepines, alcohol and nicotine.

The term "attention deficit hyperactivity disorder and "ADHD" as used herein mean a condition or disorder characterized by a persistent pattern of inattention, hyperactivity, impulsivity, or any combination thereof.

The term "excessive aggression" as used herein refers to a condition characterized by aggression that is so excessive that it interferes with the individual's daily functions, relationships, and may threaten the safety of the individual, for example in a situation in which violent suicide is contemplated. The excessive aggression which may be treated using the method claimed herein is independent of a psychotic condition and not directly related to the consumption of a drug or other substance.

A tic is a sudden, rapid recurrent, nonrhythmic, stereotyped motor movement or vocalization, experienced as irresistible but suppressible for varying lengths of time. Common simple motor tics include eye blinking, neck jerking, shoulder shrugging, facial grimacing, and coughing. Common simple vocal tics include throat clearing, grunting, sniffing, snorting, and barking. Common complex motor tics include facial gestures, grooming behaviors, jumping, touching, stamping, and smelling an object. Common complex vocal tics include repeating words or phrases out of context, coprolalia (use of socially unacceptable words, frequently obscene) palilalia (repeating one's own sounds or words), and echolalia (repeating the last heard sound, word or phrase). The term "tic disorder" as used herein means includes tic disorders featuring one or more motor tics and one or more tic and more vocal tics, and vocal tics. Examples include Transient Tic Disorder, Tourette's Disorder, Chronic Vocal Tic Disorder, and Tic Disorder not otherwise specified as described by DSM-IV-R.

The terms "comprising", "consisting of", or consisting essentially of" have distinct meaning and each term may be substituted for another herein to change the scope of the invention.

As used interchangeably herein, the term "oligonucleotides", and "polynucleotides" include RNA, DNA, or RNA/DNA hybrid sequences of more than one nucleotide in either single chain or duplex form. The term "nucleotide" as used herein as an adjective to describe molecules comprising RNA, DNA, or RNA/DNA hybrid sequences of any length in single-stranded or duplex form. The term "nucleotide" is also used herein as a noun to refer to individual nucleotides or varieties of nucleotides, meaning a molecule, or individual unit in a larger nucleic acid molecule, comprising a purine or pyrimidine, a ribose or deoxyribose sugar moiety, and a phosphate group, or phosphodiester linkage in the case of nucleotides within an oligonucleotide or polynucleotide. Although the term "nucleotide" is also used herein to encompass "modified nucleotides" which comprise at least one modifications (a) an alternative linking group, (b) an analogous form of purine, (c) an analogous form of pyrimidine, or (d) an analogous sugar, for examples of analogous linking groups, purine, pyrimidines, and sugars see for example PCT publication No. WO 95/04064, the disclosure of which is incorporated herein by reference. However, the polynucleotides of the invention are preferably comprised of greater than 50% conventional deoxyribose nucleotides, and most preferably greater than 90% conventional deoxyribose nucleotides. The polynucleotide sequences of the invention may be prepared by any known method, including synthetic, recombinant, *ex vivo* generation, or a combination thereof, as well as utilizing any purification methods known in the art.

The term "purified" is used herein to describe a polynucleotide or polynucleotide vector of the invention which has been separated from other compounds including, but not limited to other nucleic acids, carbohydrates, lipids and proteins (such as the enzymes used in the synthesis of the polynucleotide), or the separation of covalently closed polynucleotides from linear polynucleotides. A polynucleotide is substantially pure when at least about 50 %, preferably 60 to 75% of a sample exhibits a single polynucleotide sequence and conformation (linear versus covalently close). A substantially pure polynucleotide typically comprises about 50 %, preferably 60 to 90% weight/weight of a nucleic acid sample, more usually about 95%, and preferably is over about 99% pure. Polynucleotide purity or homogeneity may be indicated by a number of means well known in the art, such as agarose or polyacrylamide gel electrophoresis of a sample, followed by visualizing a single polynucleotide band upon staining the gel. For certain purposes higher resolution of can be provided by using HPLC or other means well known in the art. A polypeptide is substantially pure when at least about 50%, preferably 60 to 75% of a sample exhibits a single polypeptide sequence. A substantially pure polypeptide typically comprises about 50%, preferably 60 to 90% weight/weight of a protein sample, more usually about 95%, and preferably is over about 99% pure. Polypeptide purity or homogeneity is indicated by a number of means well known in the art, such as polyacrylamide gel electrophoresis of a sample, followed by visualizing a single polypeptide band upon staining the gel. For certain purposes higher resolution can be provided by using HPLC or other means well known in the art. The term purified may

also is used herein to describe a chemical composition of the invention which have been separated from other compounds.

The term "isolated" requires that the material be removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or DNA or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotide could be part of a vector and/or such polynucleotide or polypeptide could be part of a composition, and still be isolated in that the vector or composition is not part of its natural environment.

The term "primer" denotes a specific oligonucleotide sequence which is complementary to a target nucleotide sequence and used to hybridize to the target nucleotide sequence. A primer serves as an initiation point for nucleotide polymerization catalyzed by either DNA polymerase, RNA polymerase or reverse transcriptase.

The term "probe" denotes a defined nucleic acid segment (or nucleotide analog segment, e.g., polynucleotide as defined herein) which can be used to identify a specific polynucleotide sequence present in samples, said nucleic acid segment comprising a nucleotide sequence complementary of the specific polynucleotide sequence to be identified.

The terms "trait" and "phenotype" are used interchangeably herein and refer to any clinically distinguishable, detectable or otherwise measurable property of an organism such as symptoms of, or susceptibility to a disease for example. Typically the terms "trait" or "phenotype" are used herein to refer to symptoms of, or susceptibility to schizophrenia or bipolar disorder; or to refer to an individual's response to an agent acting on schizophrenia or bipolar disorder; or to refer to symptoms of, or susceptibility to side effects to an agent acting on schizophrenia or bipolar disorder.

The term "allele" is used herein to refer to variants of a nucleotide sequence. A biallelic polymorphism has two forms. Typically the first identified allele is designated as the original allele whereas other alleles are designated as alternative alleles. Diploid organisms may be homozygous or heterozygous for an allelic form.

The term "heterozygosity rate" is used herein to refer to the incidence of individuals in a population, which are heterozygous at a particular allele. In a biallelic system the heterozygosity rate is on average equal to $2P_a(1-P_a)$, where P_a is the frequency of the least common allele. In order to be useful in genetic studies a genetic marker should have an adequate level of heterozygosity to allow a reasonable probability that a randomly selected person will be heterozygous.

The term "genotype" as used herein refers the identity of the alleles present in an individual or a sample. In the context of the present invention a genotype preferably refers to the description of the

biallelic marker alleles present in an individual or a sample. The term “genotyping” a sample or an individual for a biallelic marker involves determining the specific allele or the specific nucleotide(s) carried by an individual at a biallelic marker.

The term “mutation” as used herein refers to a difference in DNA sequence between or among different genomes or individuals which has a frequency below 1%.

The term “haplotype” refers to a combination of alleles present in an individual or a sample on a single chromosome. In the context of the present invention a haplotype preferably refers to a combination of biallelic marker alleles found in a given individual and which may be associated with a phenotype.

The term “polymorphism” as used herein refers to the occurrence of two or more alternative genomic sequences or alleles between or among different genomes or individuals. “Polymorphic” refers to the condition in which two or more variants of a specific genomic sequence can be found in a population. A “polymorphic site” is the locus at which the variation occurs. A polymorphism may comprise a substitution, deletion or insertion of one or more nucleotides. A single nucleotide polymorphism is a single base pair change. Typically a single nucleotide polymorphism is the replacement of one nucleotide by another nucleotide at the polymorphic site. Deletion of a single nucleotide or insertion of a single nucleotide, also give rise to single nucleotide polymorphisms. In the context of the present invention “single nucleotide polymorphism” preferably refers to a single nucleotide substitution. Typically, between different genomes or between different individuals, the polymorphic site may be occupied by two different nucleotides.

The terms “biallelic polymorphism” and “biallelic marker” are used interchangeably herein to refer to a polymorphism having two alleles at a fairly high frequency in the population, preferably a single nucleotide polymorphism. A “biallelic marker allele” refers to the nucleotide variants present at a biallelic marker site. Typically the frequency of the less common allele of the biallelic markers of the present invention has been validated to be greater than 1%, preferably the frequency is greater than 10%, more preferably the frequency is at least 20% (i.e. heterozygosity rate of at least 0.32), even more preferably the frequency is at least 30% (i.e. heterozygosity rate of at least 0.42). A biallelic marker wherein the frequency of the less common allele is 30% or more is termed a “high quality biallelic marker.” All of the genotyping, haplotyping, association, and interaction study methods of the invention may optionally be performed solely with high quality biallelic markers.

The location of nucleotides in a polynucleotide with respect to the center of the polynucleotide are described herein in the following manner. When a polynucleotide has an odd number of nucleotides, the nucleotide at an equal distance from the 3' and 5' ends of the polynucleotide is considered to be “at the center” of the polynucleotide, and any nucleotide immediately adjacent to the nucleotide at the center, or the nucleotide at the center itself is considered to be “within 1 nucleotide of the center.” With

an odd number of nucleotides in a polynucleotide any of the five nucleotides positions in the middle of the polynucleotide would be considered to be within 2 nucleotides of the center, and so on. When a polynucleotide has an even number of nucleotides, there would be a bond and not a nucleotide at the center of the polynucleotide. Thus, either of the two central nucleotides would be considered to be "within 1 nucleotide of the center" and any of the four nucleotides in the middle of the polynucleotide would be considered to be "within 2 nucleotides of the center", and so on. For polymorphisms which involve the substitution, insertion or deletion of 1 or more nucleotides, the polymorphism, allele or biallelic marker is "at the center" of a polynucleotide if the difference between the distance from the substituted, inserted, or deleted polynucleotides of the polymorphism and the 3' end of the polynucleotide, and the distance from the substituted, inserted, or deleted polynucleotides of the polymorphism and the 5' end of the polynucleotide is zero or one nucleotide. If this difference is 0 to 3, then the polymorphism is considered to be "within 1 nucleotide of the center." If the difference is 0 to 5, the polymorphism is considered to be "within 2 nucleotides of the center." If the difference is 0 to 7, the polymorphism is considered to be "within 3 nucleotides of the center," and so on. For polymorphisms which involve the substitution, insertion or deletion of 1 or more nucleotides, the polymorphism, allele or biallelic marker is "at the center" of a polynucleotide if the difference between the distance from the substituted, inserted, or deleted polynucleotides of the polymorphism and the 3' end of the polynucleotide, and the distance from the substituted, inserted, or deleted polynucleotides of the polymorphism and the 5' end of the polynucleotide is zero or one nucleotide. If this difference is 0 to 3, then the polymorphism is considered to be "within 1 nucleotide of the center." If the difference is 0 to 5, the polymorphism is considered to be "within 2 nucleotides of the center." If the difference is 0 to 7, the polymorphism is considered to be "within 3 nucleotides of the center," and so on.

The term "upstream" is used herein to refer to a location which, is toward the 5' end of the polynucleotide from a specific reference point.

The terms "base paired" and "Watson & Crick base paired" are used interchangeably herein to refer to nucleotides which can be hydrogen bonded to one another by virtue of their sequence identities in a manner like that found in double-helical DNA with thymine or uracil residues linked to adenine residues by two hydrogen bonds and cytosine and guanine residues linked by three hydrogen bonds (See Stryer, L., *Biochemistry*, 4th edition, 1995).

The terms "complementary" or "complement thereof" are used herein to refer to the sequences of polynucleotides which is capable of forming Watson & Crick base pairing with another specified polynucleotide throughout the entirety of the complementary region. This term is applied to pairs of polynucleotides based solely upon their sequences and not any particular set of conditions under which the two polynucleotides would actually bind.

The terms “DAO gene”, when used herein, encompasses genomic, mRNA and cDNA sequences encoding any D-amino acid oxidase proteins of the invention, including the untranslated regulatory regions of the genomic DNA.

The terms “g34872 gene”, when used herein, encompasses genomic, mRNA and cDNA sequences encoding any g34872 protein, including the untranslated regulatory regions of the genomic DNA.

The terms “DDO gene”, when used herein, encompasses genomic, mRNA and cDNA sequences encoding any D-aspartate oxidase protein, including the untranslated regulatory regions of the genomic DNA.

As used herein the term “13q31-q33-related biallelic marker” relates to a set of biallelic markers residing in the human chromosome 13q31-q33 region. The term 13q31-q33-related biallelic marker encompasses all of the biallelic markers disclosed in Table 6b of US Patent application 09/539,333 and international application PCT/IB00/00435, which disclosures are incorporated by reference in their entireties, and any biallelic markers in linkage disequilibrium therewith, as well as any biallelic markers disclosed in Table 6c (of same US Patent application 09/539,333 and international application PCT/IB00/00435) and any biallelic markers in linkage disequilibrium therewith. The preferred chromosome 13q31-q33-related biallelic marker alleles of the present invention include each one the alleles described in Tables 6b (of same US Patent application 09/539,333 and international application PCT/IB00/00435) individually or in groups consisting of all the possible combinations of the alleles listed.

As used herein the term “Region D-related biallelic marker” relates to a set of biallelic markers in linkage disequilibrium with the subregion of the chromosome 13q31-q33 region referred to herein as Region D. The term Region D-related biallelic marker encompasses the biallelic markers A1 to A242, A249 to A251, A257 to A263, A269 to A270, A278, A285 to A299, A303 to A307, A324, A330, A334 to A335, A346 to A357 and A361 to A489 disclosed in Table 6b and any biallelic markers in linkage disequilibrium with markers A1 to A242, A249 to A251, A257 to A263, A269 to A270, A278, A285 to A299, A303 to A307, A324, A330, A334 to A335, A346 to A357 and A361 to A489, of US Patent application 09/539,333 and international application PCT/IB00/00435, which disclosures are incorporated by reference in their entireties.

As used herein the term “sbg1-related biallelic marker” relates to a set of biallelic markers in linkage disequilibrium with the sbg1 gene or an sbg1 nucleotide sequence. The term sbg1-related biallelic marker encompasses the biallelic markers A85 to A219 disclosed in Table 6b and any biallelic markers in linkage disequilibrium therewith, of US Patent application 09/539,333 and international application PCT/IB00/00435, which disclosures are incorporated by reference in their entireties.

As used herein the term “g34665-related biallelic marker” relates to a set of biallelic markers in linkage disequilibrium with the g34665 gene or an sbg1 nucleotide sequence. The term g34665-related

biallelic marker encompasses the biallelic markers A230 to A236 disclosed in Table 6b and any biallelic markers in linkage disequilibrium therewith.

The term “polypeptide” refers to a polymer of amino acids without regard to the length of the polymer; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide. This term also does not specify or exclude post-expression modifications of polypeptides, for example, polypeptides which include the covalent attachment of glycosyl groups, acetyl groups, phosphate groups, lipid groups and the like are expressly encompassed by the term polypeptide. Also included within the definition are polypeptides which contain one or more analogs of an amino acid (including, for example, non-naturally occurring amino acids, amino acids which only occur naturally in an unrelated biological system, modified amino acids from mammalian systems etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

As used herein, the term “non-human animal” refers to any non-human vertebrate, birds and more usually mammals, preferably primates, farm animals such as swine, goats, sheep, donkeys, and horses, rabbits or rodents, more preferably rats or mice. As used herein, the term “animal” is used to refer to any vertebrate, preferably a mammal. Both the terms “animal” and “mammal” expressly embrace human subjects unless preceded with the term “non-human”.

As used herein, the term “antibody” refers to a polypeptide or group of polypeptides which are comprised of at least one binding domain, where an antibody binding domain is formed from the folding of variable domains of an antibody molecule to form three-dimensional binding spaces with an internal surface shape and charge distribution complementary to the features of an antigenic determinant of an antigen., which allows an immunological reaction with the antigen. Antibodies include recombinant proteins comprising the binding domains, as well as fragments, including Fab, Fab', F(ab)₂, and F(ab')₂ fragments.

As used herein, an “antigenic determinant” is the portion of an antigen molecule, in this case an sbg1 polypeptide, that determines the specificity of the antigen-antibody reaction. An “epitope” refers to an antigenic determinant of a polypeptide. An epitope can comprise as few as 3 amino acids in a spatial conformation which is unique to the epitope. Generally an epitope comprises at least 6 such amino acids, and more usually at least 8-10 such amino acids. Methods for determining the amino acids which make up an epitope include x-ray crystallography, 2-dimensional nuclear magnetic resonance, and epitope mapping e.g. the Pepscan method described by Geysen et al. 1984; PCT Publication No. WO 84/03564; and PCT Publication No. WO 84/03506.

A complete description of “Variants and Fragments”, “Identity Between Nucleic Acids Or Polypeptides”, “Stringent Hybridization Conditions”, “DNA Constructs that Enables Directing

Temporal and Spatial Expression of *sbg1* Nucleic Acid Sequences in Recombinant Cell Hosts and in Transgenic Animals" are fully detailed in co-pending U.S. patent application ser. no. 09/539,333 titled "Schizophrenia associated genes, proteins and biallelic markers" and co-pending International Patent Application No. PCT/IB00/00435, both filed 30 March 2000 and which disclosures are hereby incorporated by reference in their entireties.

Genomic Sequences of g34872 and DAO polynucleotides

Particularly preferred g34872 nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising, consisting essentially of, or consisting of a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of nucleotide positions 213818 to 243685 of US Patent Application 09/539,333 SEQ ID No:1, or the complements thereof (US Patent application 09/539,333 and international application PCT/IB00/00435, which disclosures are incorporated by reference in their entireties).

DAO polynucleotides of the invention are described in SEQ ID NO:1 of the present invention. Particularly preferred nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising, consisting essentially of, or consisting of a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of nucleotide positions 6000-86600 of SEQ ID No:1. Nucleic acids of the invention encompass DAO nucleic acid from any source, including primate, non-human primate, mammalian and human DAO nucleic acids.

Further preferred nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 1 or the complements thereof, wherein said contiguous span comprises a DAO related biallelic marker. Optionally, said biallelic marker is selected from the group comprising 24-1443/126, 24-1457/52, or 24-1461/256. Preferably, said biallelic marker is 24-1461/256.

It should be noted that nucleic acid fragments of any size and sequence may also be comprised by the polynucleotides described in this section.

Thus, the invention embodies purified, isolated, or recombinant polynucleotides comprising a nucleotide sequence selected from the group consisting of the exons of the *DAO* gene (SEQ ID NO:1), or a sequence complementary thereto. Preferred are purified, isolated, or recombinant polynucleotides comprising at least one exon of the *DAO* gene, or a complementary sequence thereto or a fragment or a variant thereof. Also encompassed by the invention are purified, isolated, or recombinant nucleic acids comprising a combination of at least two exons of the *DAO* gene selected from the group consisting of exons Z, A, B, C, Ulong, U, V, Z, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 11long, wherein the polynucleotides are arranged within the nucleic acid in the same relative order as in SEQ ID NO:1.

Particularly preferred nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100 or 200 nucleotides of SEQ ID No 1. or the complements thereof.

Another object of the invention consists of a purified, isolated, or recombinant nucleic acid that hybridizes with an DAO nucleotide sequence of SEQ ID NO:1, or a complementary sequence thereto or a variant thereof, under the stringent hybridization conditions as defined above.

The present invention further embodies purified, isolated, or recombinant polynucleotides comprising a nucleotide sequence selected from the group consisting of the introns of the *DAO* gene (SEQ ID NO:1), or a sequence complementary thereto.

In other embodiments, the present invention encompasses the DAO gene as well as DAO genomic sequences consisting of, consisting essentially of, or comprising the sequence of nucleotide positions of SEQ ID No 1, a sequence complementary thereto, as well as fragments and variants thereof.

The invention also encompasses a purified, isolated, or recombinant polynucleotide comprising a nucleotide sequence of DAO having at least 70, 75, 80, 85, 90, or 95% nucleotide identity with SEQ ID NO:1 or a complementary sequence thereto or a fragment thereof.

These nucleic acids, as well as their fragments and variants, may be used as oligonucleotide primers or probes in order to detect the presence of a copy of a gene comprising an g34782, DAO or DDO nucleic acid sequence in a test sample, or alternatively in order to amplify a target nucleotide sequence within an g347982, DAO or DDO nucleic acid sequence or adjoining region.

Additional preferred nucleic acids of the invention include isolated, purified, or recombinant DAO polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100 or 200 nucleotides of SEQ ID NO:1, or the complements thereof, wherein said contiguous span comprises at least one biallelic marker. Optionally, said contiguous span comprises an DAO-related biallelic marker. It should be noted that nucleic acid fragments of any size and sequence may also be comprised by the polynucleotides described in this section. Either the original or the alternative allele may be present at said biallelic marker.

Yet further nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 or 500 nucleotides, to the extent that said span is consistent with the nucleotide position range, of SEQ ID NO:1, wherein said contiguous span comprises at least 1, 2, 3, 5, or 10 of the following nucleotide positions of SEQ ID No 1: 215820 to 215941, 216661 to 217009, 230409 to 290721, 231272 to 231411, 234202 to 234321, 240528 to 240567, 240528 to 240827 and 240528 to 240996, or the complements thereof, as well as polynucleotides having at least 70, 75, 80, 85, 90, or 95% nucleotide identity with said span, and polynucleotides capable of hybridizing with said span.

The present invention also comprises a purified or isolated nucleic acid encoding an DAO protein having the amino acid sequence of any one of SEQ ID NOs:7-10 or a peptide fragment or variant thereof.

While this section is entitled "Genomic Sequences of *sbg1*," it should be noted that nucleic acid fragments of any size and sequence may also be comprised by the polynucleotides described in this section, flanking the genomic sequences *sbg1* on either side or between two or more such genomic sequences.

***DAO* cDNA Sequences**

The expression of the *DAO* gene has been shown to lead to the production of several mRNA species. Several cDNA sequences corresponding to these mRNA are set forth in SEQ ID NOs:2-6. The invention encompasses a purified, isolated, or recombinant nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NOs:2-6, complementary sequences thereto, splice variants thereof, as well as allelic variants, and fragments thereof. Moreover, preferred polynucleotides of the invention include purified, isolated, or recombinant *DAO* cDNAs consisting of, consisting essentially of, or comprising a nucleotide sequence selected from the group consisting of SEQ ID NOs:2-6. Particularly preferred nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 8, 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 75, 80, 100, 200 or 500 nucleotides, to the extent that the length of said contiguous span is consistent with the length of the SEQ ID NOs:2-6, or the complements thereof.

It should be noted that nucleic acid fragments of any size and sequence may also be comprised by the polynucleotides described in this section.

The invention also pertains to a purified or isolated nucleic acid comprising a polynucleotide having at least 70, 80, 85, 90 or 95% nucleotide identity with a polynucleotide selected from the group consisting of SEQ ID NOs:2-6, advantageously 99% nucleotide identity, preferably 99.5% nucleotide identity and most preferably 99.8% nucleotide identity with a polynucleotide selected from the group consisting of SEQ ID NOs:2-6, or a sequence complementary thereto or a biologically active fragment thereof.

Another object of the invention relates to purified, isolated or recombinant nucleic acids comprising a polynucleotide that hybridizes, under the stringent hybridization conditions defined herein, with a polynucleotide selected from the group consisting of SEQ ID NOs:2-6, or a sequence complementary thereto or a variant thereof or a biologically active fragment thereof. The *DAO* cDNA forms of SEQ ID NOs:2-6 are further described in the sequence listing.

Primers used to isolate the particular *DAO* cDNAs or for genotyping are listed in SEQ ID NO:1.

Biallelic markers for *DAO*, and genotyping primers thereof, are listed in SEQ ID NOs:1, 24, 26, and 29.

Polynucleotides of g34872 are listed in SEQ ID NO:14 and 16. g34872 biallelic marker 99-16105-152

of SEQ ID NO:12 and g34872 biallelic marker 99-5919-215 of SEQ ID NO:13 are listed and primers to make are described therein. cDNA of g34872 is listed in SEQ ID NO:14 and polynucleotides used in 2-hybrid experiments are listed in SEQ ID NO:16.

The present inventors have also identified novel exons and variations in cDNA sequence as obtained from various tissues and these are listed as Exons 11 long, Z, A, B, C, and UL of SEQ ID NO:1, and in polynucleotides of SEQ ID NOs:2-6. Novel forms of DAO polypeptides are listed in SEQ ID NO:8-10.

These variants represent rare and novel forms of DAO which are preferably used to screen for compositions to use in methods of treating CNS disorders.

It should be noted that nucleic acid fragments of any size and sequence may also be comprised by the polynucleotides described in this section, flanking the genomic sequences of g34872, DAO and DDO on either side or between two or more such genomic sequences.

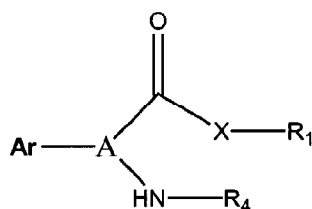
DAO and DDO Antagonists

The term "antagonist" as used herein refers to the inhibition of enzymatic reaction whereby DAO or DDO converts a D-amino acid substrate into the corresponding α -keto acid. The antagonists may be specified as either competitive, non-competitive, uncompetitive, allosteric, or irreversible inhibitors of DAO or DDO enzymatic activity. The term "activity" or "enzymatic activity" of DAO or DDO refers to the enzymatic reaction above. Antagonists may be specified in terms of the degree of inhibition of DAO or DDO activity. Preferred antagonists reduce DAO or DDO activity by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 95%. Inhibitory effect may also be specified as an inhibition constant or K_i (M) values. Preferred antagonists have a K_i (M) with a numeric value less than 5×10^{-2} , 1×10^{-2} , 5×10^{-3} , 1×10^{-3} , 5×10^{-4} , 1×10^{-4} , 5×10^{-5} , 1×10^{-6} , 5×10^{-7} , 1×10^{-7} . It is noted that there is an inverse relationship between the K_i (M) numeric value and the inhibitory effect, i.e., as the K_i (M) value decreases, the inhibitory effect increases. Antagonists may also be specified in terms of their specificity for DAO or DDO. Therefore, included in the present invention are antagonists that inhibit DAO or DDO activity but do not inhibit other human flavoproteins (p-Hydroxybenzoate hydroxylase, cholesterol oxidase and glucose oxidase) or has a K_i (M) numeric value for other human flavoproteins greater than 1×10^{-2} , 5×10^{-2} . It should be appreciated from the definition that the generic terms "antagonist" and "inhibitor" can be used interchangeably to indicate any composition which inhibits DAO or DDO activity as defined above. In addition, specific types of antagonists or inhibitors can be set forth independently as described in the specification, for example a competitive inhibitor.

Over 200 inhibitors of DAO and DDO have been studied to date. DAO and DDO antagonists may be selected from the compositions presented *supra*, or other antagonists known in the art, or made using the methods described herein, or known in the art. Alternatively, DAO and DDO antagonists can

be purchased from commercial suppliers. A non-limiting list of compounds useful in accordance with the invention is provided in Table I. DAO and DDO antagonists are further comprise the families of compositions selected from the groups comprising: Competitive Inhibitor compositions, Irreversible Inhibitor compositions, Formula I, Formula II, Formula III, Formula IV, Formula V, and Formula VI compositions, and subgroups thereof, as presented herein. Further preferred representative compositions of the Formulae I-VI, and subgroups thereof, include, but are not limited to the detailed description *infra*.

Formula I compositions, or pharmaceutically acceptable salts thereof, are represented by the structure comprising:

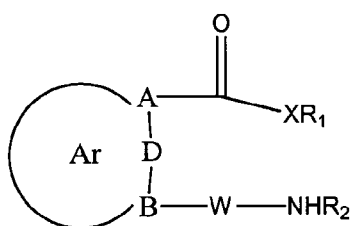


wherein:

- a) A is alkyl such as methyl, ethyl, propyl or butyl; branched chain alkyl such as isobutyl, isopropyl, isopentyl or cycloalkyl such as cyclopropyl, cyclopentyl or cyclohexyl. Such groups may themselves be substituted with C₁-C₆ alkyl, halo, hydroxyl or amino;
- b) X is O or N;
- c) Ar is an aromatic mono-, bi- or tricyclic fused heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to five position(s) with hydrogen, halogen, hydroxyl, -CN, COR₂, --CONR₂ R₃, --S(O)_n R₂, --OPO(OR₂)OR₃, --PO(OR₃)R₃, --OC(O)NR₂ R₃, --COOR₂, --CONR₂ R₃, --SO₃H, --NR₂ R₃, --NR₂ COR₃, --NR₃ COOR₃, --SO₂ NR₂ R₃, --N(R₂)SO₂ R₃, --NR₂ CONR₂ R₂, --SO₂ NHCOR₂, --CONHSO₂ R₂, --SO₂ NHCN, --OR₁, C₁-C₆ straight or branched chain alkyl or alkenyl, or C₁-C₆ branched or straight chain alkyl or alkenyl which is substituted with one or more, halogen, hydroxyl, amino, carboxy, carboxamide, nitrile, nitro, alkoxy, trifluoromethyl, sulfur, sulfonate, phosphonate, phosphate, Ar¹, N₃ or a combination thereof and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and a combination thereof;
- d) R₄ is H, alkyl, Ar¹, O, substituted alkyl;
- e) R¹ is (C₁ - C₆) alkyl, Ar¹, (C₁ - C₄) alkoxycarbonylmethyl, substituted alkyl;

- f) R_2 and R_3 are each, independently, hydrogen, C_1 - C_6 straight or branched chain alkyl or alkenyl, or C_1 - C_6 branched or straight chain alkyl or alkenyl which is substituted with one or more, halogen, hydroxyl, amino, carboxy, carboxamide, nitrile, nitro, alkoxy, trifluoromethyl, sulfur, sulfonate, phosphonate, phosphate, Ar^1 , or N_3 ; and
- g) Ar^1 is a mono-, bi- or tricyclic, carbo- or heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to three position(s) with halo, hydroxyl, nitro, trifluoromethyl, C_1 - C_6 straight or branched chain alkyl or alkenyl, C_1 - C_4 alkoxy, C_1 - C_4 alkenyloxy, phenoxy, benzyloxy, amino, or a combination thereof; wherein the individual ring sizes are 3-7 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and a combination thereof.

Further preferred Formula I compositions, or pharmaceutically acceptable salts thereof, are Formula Ia compositions, or pharmaceutically acceptable salts thereof, comprising the structure:

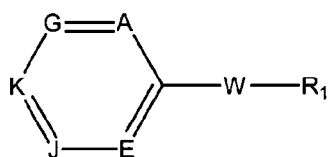


wherein:

- A and B consist of C or N and D may contain 0-2 members consisting of C or N;
- W is C_1 - C_4 alkyl such as $(CH_2)_n$, branched chain alkyl;
- n is 0-4. Further, when $n = 0$ it is assumed that $-NHR_2$ is covalently bound to B;
- X is O or N;
- R_2 is H, alkyl, Ar^1 , or O substituted alkyl;
- R^1 is $(C_1 - C_6)$ alkyl, Ar^1 , $(C_1 - C_4)$ alkoxycarbonylmethyl, or substituted alkyl;
- Ar is an aromatic mono-, bi- or tricyclic fused heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to six position(s) with halo, hydroxyl, nitro, trifluoromethyl, C_1 - C_6 straight or branched chain alkyl or alkenyl, C_1 - C_4 alkoxy, C_1 - C_4 alkenyloxy, phenoxy, benzyloxy, amino, C_3 - C_6 cycloalkyl or a combination thereof; wherein the individual ring sizes are 5-6 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and a combination thereof; and
- Ar^1 is a mono-, bi- or tricyclic, carbo- or heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to three position(s) with halo, hydroxyl, nitro, trifluoromethyl, C_1 - C_6 straight or branched chain alkyl or alkenyl, C_1 - C_4 alkoxy, C_1 - C_4

alkenyloxy, phenoxy, benzyloxy, amino, or a combination thereof; wherein the individual ring sizes are 3-7 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and a combination thereof.

Further preferred Formula Ia compositions, or pharmaceutically acceptable salts thereof, are Formula Ib compositions, or pharmaceutically acceptable salts thereof, comprising the structure:



wherein:

- a) A, G, K, J, E are members of a six membered carbo or heterocyclic aromatic ring, wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of C, N and a combination thereof;
- b) A, G, K, J, E may each independently be unsubstituted or substituted with hydrogen, halogen, hydroxyl, -CN, COR₂, --CONR₂ R₃, --S(O)_n R₂, --OPO(OR₂)OR₃, --PO(OR₃)R₃, --OC(O)NR₂ R₃, --COOR₂, --CONR₂ R₃, --SO₃H, --NR₂ R₃, --NR₂ COR₃, --NR₃ COOR₃, --SO₂ NR₂ R₃, --N(R₂)SO₂ R₃, --NR₂ CONR₂ R₂, --SO₂ NHCOR₂, --CONHSO₂ R₂, --SO₂ NHCN, --OR₁, C₁-C₆ straight or branched chain alkyl or alkenyl, or C₁-C₆ branched or straight chain alkyl or alkenyl which is substituted with one or more halogen, hydroxyl, amino, carboxy, carboxamide, nitrile, nitro, alkoxy, trifluoromethyl, sulfur, sulfonate, phosphonate, phosphate, Ar¹, or N₃;
- c) R₁ is CN, COR₂, --CONR₂ R₃, --S(O)_n R₂, --OPO(OR₂)OR₃, --PO(OR₃)R₃, --OC(O)NR₂ R₃, --COOR₂, --CONR₂ R₃, --SO₃H, --NR₂ R₃, --NR₂ COR₃, --NR₃ COOR₃, --SO₂ NR₂ R₃, --N(R₂)SO₂ R₃, --NR₂ CONR₂ R₂, --SO₂ NHCOR₂, --CONHSO₂ R₂, --SO₂ NHCN, SCN, COCO₂H, C₁-C₆ straight or branched chain alkyl or alkenyl, or C₁-C₆ branched or straight chain alkyl or alkenyl which is substituted with one or more halogen, hydroxyl, amino, carboxy, carboxamide, nitrile, nitro, alkoxy, trifluoromethyl, sulfur, sulfonate, phosphonate, phosphate, Ar¹, or N₃;
- d) W is N, (CH₂)_x, or -NCH₂;
- e) x=0-4;
- f) n=0-2;
- g) R₂ and R₃ are each, independently, hydrogen, C₁-C₆ straight or branched chain alkyl or alkenyl, or C₁-C₆ branched or straight chain alkyl or alkenyl which is substituted with one or more halogen, hydroxyl, amino, carboxy, carboxamide, nitrile, nitro, alkoxy, trifluoromethyl, sulfur, sulfonate, phosphonate, phosphate, Ar¹, or N₃; and
- h) Ar¹ is a mono-, bi- or tricyclic, carbo- or heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to three position(s) with halo, hydroxyl, nitro, trifluoromethyl, C₁-C₆ straight or branched chain alkyl or alkenyl, C₁-C₄ alkoxy, C₁-C₄

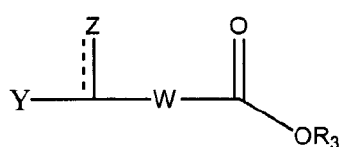
alkenyloxy, phenoxy, benzyloxy, amino, or a combination thereof; wherein the individual ring sizes are 5-6 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and a combination thereof.

Specific examples of Formulae I, Ia, and Ib compositions, or pharmaceutically acceptable salts thereof, include, but not limited to, the list comprising:

- a) Benzoic acid;
- b) 2-Aminobenzoate;
- c) 3-Aminobenzoate;
- d) 4-Aminobenzoate;
- e) Salicylic acid;
- f) N-Methylnicotinate;
- g) Methyl-6-methylnicotinate;
- h) Ethyl-2-methylnicotinate;
- i) Anthranilate;
- j) Ethyl-2-aminobenzoate;
- k) Methyl-2-aminobenzoate;
- l) Picolinate;
- m) Ethyl-2-pyridinecarboxylate;
- n) 3-Methylbenzyl thiocyanate;
- o) Phenyl pyruvic acid;
- p) Phenylglyoxilic acid;
- q) 1-Methyl pyridinium-3-carboxylate;
- r) Befloxatone; (5*R*)-5-(Methoxymethyl)-3-[4-[(3*R*)-4,4,4-trifluoro-3-hydroxybutoxy]phenyl]-2-oxazolidinone;
- s) Bupropion; 1-(3-Chlorophenyl)-2-[(1,1-dimethylethyl)amino]-1-propanone;
- t) Cotinine; 1-Methyl-5-(3-pyridinyl)-2-pyrrolidinone;
- u) Duloxetine; (γ*S*)-*N*-Methyl-γ-(1-naphthalenyloxy)-2-thiophenepropanamine;
- v) Fenpentadiol; 2-(4-Chlorophenyl)-4-methyl-2,4-pentanediol;
- w) Fluvoxamine; (*E*)-5-Methoxy-1-[4-(trifluoromethyl)phenyl]-1-pentanone *O*-(2-aminoethyl)oxime;
- x) Iproclozide; 4-(Chlorophenoxy)acetic acid 2-(1-methylethyl)hydrazide;
- y) Iproniazid; 4-Pyridinecarboxylic acid 2-(1-methylethyl)hydrazide;
- z) Levophacetoperane; α-Phenyl-2-piperidinemethanol acetate;

- aa) Rolipram; 4-[3-(Cyclopentyloxy)-4-methoxyphenyl]-2-pyrrolidinone;
 bb) Tranlycypromine; (1*R*,2*S*)-*rel*-2-Phenylcyclopropanamine; and
 cc) Milnacipran; (1*R*,2*S*)-*rel*-2-(Aminomethyl)-*N,N*-diethyl-1-phenylcyclopropanecarboxamide.

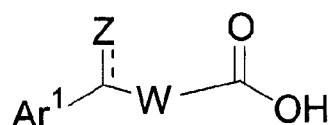
Formula II compositions, or pharmaceutically acceptable salts thereof, are represented by the structure comprising:



wherein:

- a) $W = (CH_2)_n$;
 b) $n = 0-5$;
 c) Z is O or hydroxyl;
 d) $Y = H, Ar^1, R_4(CH_2)_x, R_1S(CH_2)_x, R_1SO(CH_2)_x, R_1SO_2(CH_2)_x, R_1SO_3(CH_2)_x, HNR_1SO_2(CH_2)_x, R_1R_2N(CH_2)_x, R_1O(CH_2)_x, CF_3$, or OH;
 e) $x = 0-6$;
 f) R_1, R_2 and R_3 are each independently hydrogen, C_1-C_6 straight or branched chain alkyl or C_1-C_6 branched or straight chain alkyl substituted with one or more halogen, hydroxyl, amino, carboxy, carboxamide, nitrile, nitro, alkoxy, trifluoromethyl, sulfur, sulfonate, phosphonate, phosphate, or Ar^1 ;
 g) R_4 is halogen, CN, N_3 , C_1-C_6 straight or branched chain alkyl or C_1-C_6 branched or straight chain alkyl substituted with one or more halogen, hydroxyl, nitro, alkoxy, trifluoromethyl, sulfonate, phosphonate, phosphate, Ar^1 , $-COR_1$, $-COOR_1$, $-CONR_1R_2$, CN, $-NR_1$, $-NR_1R_2$, $-SR_1$, $-SO_2NHCN$, or N_3 ; and
 h) Ar^1 is a mono-, bi- or tricyclic, carbo- or heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to three position(s) with halo, hydroxyl, nitro, trifluoromethyl, C_1-C_6 straight or branched chain alkyl or alkenyl, C_1-C_4 alkoxy, C_1-C_4 alkenyloxy, phenoxy, benzyloxy, amino, or a combination thereof; wherein the individual ring sizes are 5-6 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and a combination thereof.

Further preferred Formula II compositions, or pharmaceutically acceptable salts thereof, are Formula IIa compositions, or pharmaceutically acceptable salts thereof, comprising the structure:



wherein:

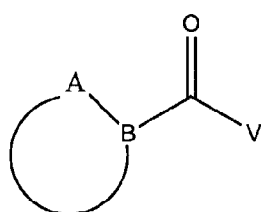
- a) Y is Ar¹;
- b) Z is a carbonyl or hydroxyl;
- c) W is (CH₂)_n wherein (n= 0,1, or 2) and R₃ = H; and
- d) Ar¹ is a mono-, bi- or tricyclic, carbo- or heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to three position(s) with halo, hydroxyl, nitro, trifluoromethyl, C₁-C₆ straight or branched chain alkyl or alkenyl, C₁-C₄ alkoxy, C₁-C₄ alkenyloxy, phenoxy, benzyloxy, amino, or a combination thereof; wherein the individual ring sizes are 5-6 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and a combination thereof.

Specific examples of Formula II and IIa compositions, or pharmaceutically acceptable salts thereof, include, but are not limited to, the list comprising:

- a) 2-Oxo-propionic acid;
- b) 5-Guanidiono-2-oxo-pentanoic acid;
- c) 2-Oxo-succinamic acid;
- d) 2-Oxo-succinic acid;
- e) 3-Mercapto-2-oxo-propionic acid;
- f) 3-(1*H*-Imidazol-4-yl)-2-oxo-propionic acid;
- g) 3-Methyl-2-oxo-pentanoic acid;
- h) Oxo-acetic acid;
- i) 4-Carbamoyl-2-oxo butyric acid;
- j) 2-Oxo-pentanedioic acid;
- k) 4-Methyl-2-oxo-pentanoic acid;
- l) 6-Amino-2-oxo-hexanoic acid;
- m) 4-Methylsulfanyl-2-oxo-butyric acid;
- n) 2-Oxo-3-phenyl propionic acid;
- o) 3-Hydroxy-2-oxo-propionic acid;
- p) 3-Hydroxy-2-oxo-butyric acid;
- q) 3-(1*H*-Indol-3-yl)-2-oxo-propionic acid;

- r) 3-(4-Hydroxy-phenyl)-2-oxo-propionic acid;
- s) 3-methyl-2-oxo-butyric acid;
- t) 2-Hydroxy butyric acid;
- u) 3-Hydroxy butyric acid; and
- v) 3-Oxoglutaric acid.

Formula III compositions, or pharmaceutically acceptable salts thereof, are represented by the structure comprising:

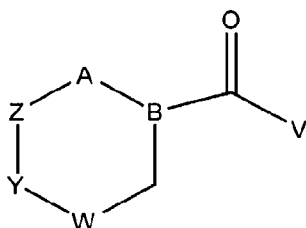


wherein:

- a) A and B taken together, form a 5-8 membered saturated or partially unsaturated heterocyclic ring containing at least one additional O, S, SO, SO₂, NH, or NR¹ heteroatom in any chemically stable oxidation state;
- b) V is O, OR₁, NR₂, NR₁R₂, CHR₁R₂, CH₂R₃, CHR₃R₄, or CH₂N₃;
- c) R₁ and R₂ are independently hydrogen, C₁-C₆ straight or branched chain alkyl or C₁-C₆ branched or straight chain alkyl substituted with one or more halogen, hydroxyl, amino, carboxy, carboxamide, nitro, alkoxy, trifluoromethyl, sulfur, sulfonate, phosphonate, or Ar¹;
- d) R₃ and R₄ are either halogen, C₁-C₆ straight or branched chain alkyl or C₁-C₆ branched or straight chain alkyl substituted with one or more hydroxyl, amino, carboxy, carboxamide, nitro, alkoxy, trifluoromethyl, sulfur, sulfonate, phosphonate, Ar¹, --OC(O)R₁, --COOR₁, --CONR₁R₂, CN, NR₁, NR₁R₂, SR₁, SO₂NHCN, or N₃; and
- e) Ar¹ is a mono-, bi- or tricyclic, carbo- or heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to three position(s) with halo, hydroxyl, nitro, trifluoromethyl, C₁-C₆ straight or branched chain alkyl or alkenyl, C₁-C₄ alkoxy, C₁-C₄ alkenyloxy, phenoxy, benzyloxy, amino, or a combination thereof; wherein the individual ring sizes are 5-6 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and a combination thereof.

Specific examples of Formula III compositions include, but are not limited to, Cystathionine ketimine and Cyclothionine.

Further preferred Formula III compositions, or pharmaceutically acceptable salts thereof, are Formula IV compositions, or pharmaceutically acceptable salts thereof, represented by the structure comprising :

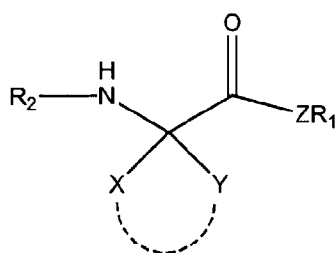


wherein:

- a) W-Y-Z-A-B comprise a six membered saturated or partially saturated carbocyclic or heterocyclic ring, wherein the heterocyclic ring contains heteroatom(s) selected from the group consisting of O, N, S, and any combination thereof;
- b) B is either C, CH or N;
- c) A, W, Y, Z are each independently CH₂, CHR₃, CR₃R₄, O, S, SO, SO₂, NH, NR₁, NR₁R₂, or C=O;
- d) V is O, OR₁, NR₂, NR₁R₂, CHR₁R₂, CH₂R₃, CHR₃R₃, or CH₂N₃;
- e) R₁ and R₂ are independently hydrogen, C₁-C₆ straight or branched chain alkyl or C₁-C₆ branched or straight chain alkyl substituted with one or more, halogen, hydroxyl, amino, carboxy, carboxamide, nitrile, nitro, alkoxy, trifluoromethyl, sulfur, sulfonate, phosphonate, phosphate, or Ar¹;
- f) R₃ and R₄ are each independently halogen, --OC(O)R₁, --COOR₁, --CONR₁R₂, CN, --NR₁, --NR₁R₂, --SR₁, --SO₂NHCN, N₃, C₁-C₆ straight or branched chain alkyl or C₁-C₆ branched or straight chain alkyl substituted with one or more halogen, hydroxyl, nitro, alkoxy, trifluoromethyl, sulfonate, phosphonate, Ar¹, --OC(O)R₁, --COOR₁, --CONR₁R₂, CN, --NR₁, --NR₁R₂, --SR₁, --SO₂NHCN, or N₃; and
- g) Ar¹ is a mono-, bi- or tricyclic, carbo- or heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to three position(s) with halo, hydroxyl, nitro, trifluoromethyl, C₁-C₆ straight or branched chain alkyl or alkenyl, C₁-C₄ alkoxy, C₁-C₄ alkenyloxy, phenoxy, benzyloxy, amino, or a combination thereof; wherein the individual ring sizes are 5-6 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and any combination thereof.

Specific examples of Formula IV compositions, or pharmaceutically acceptable salts thereof, include, but are not limited to, Aminoethylcysteine-ketimine (2H-1,4-thiazine-5,6-dihydro-3-carboxylic acid), Thiomorpholine-2-carboxylic acid, Lanthionine ketimine, and 1,4-Thiomorpholine-3,5-dicarboxylic acid.

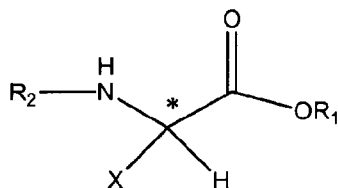
Formula V compositions, or pharmaceutically acceptable salts thereof, are represented by the structure comprising:



wherein:

- a) Z is O or NH;
- b) R¹ is (C₁-C₆) alkyl, Ar¹, or (C₁-C₄) alkoxy carbonylmethyl;
- c) X, Y, independently of one another, are H, Ar¹, (C₁-C₆) alkyl (which can be interrupted or substituted by heteroatoms, such as N, P, O, S or Si, it being possible for the heteroatoms themselves to be substituted by (C₁-C₃) alkyl once or several times), (C₂-C₆) alkenyl, (C₁-C₆) haloalkyl, or halogen. When X and Y are each carbon they may be covalently joined to form a saturated or partially unsaturated carbocyclic compound of 3-8 members consisting independently of C, N, O, and S, further wherein ring members may themselves be unsubstituted or substituted with halo, hydroxyl, carboxy, nitro, trifluoromethyl, C₁-C₆ straight or branched chain alkyl or alkenyl, C₁-C₄ alkoxy, C₁-C₄ alkenyloxy, phenoxy, benzyloxy, amino, substituted alkyl, Ar¹, or a combination thereof;
- d) R₂ is H, alkyl, Ar¹, or O substituted alkyl; and
- e) Ar¹ is a mono-, bi- or tricyclic, carbo- or heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to three position(s) with halo, hydroxyl, nitro, trifluoromethyl, C₁-C₆ straight or branched chain alkyl or alkenyl, C₁-C₄ alkoxy, C₁-C₄ alkenyloxy, phenoxy, benzyloxy, amino, or a combination thereof; wherein the individual ring sizes are 3-7 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and any combination thereof.

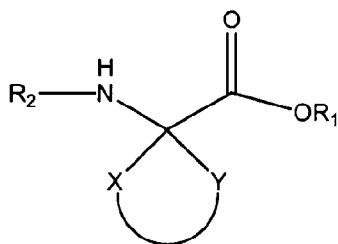
Further preferred Formula V compositions, or pharmaceutically acceptable salts thereof, are Formula Va compositions, or pharmaceutically acceptable salts thereof, comprising the structure:



wherein:

- a) * = asymmetric center and
- b) $R^1 = (C_1 - C_6)$ alkyl, Ar^1 , $(C_1 - C_4)$ alkoxy carbonylmethyl and
- c) X is H, $(C_1 - C_6)$ alkyl (which can be interrupted or substituted by heteroatoms, such as N, P, O, S or Si, it being possible for the heteroatoms themselves to be substituted by $(C_1 - C_3)$ alkyl once or several times), $(C_2 - C_6)$ alkenyl, $(C_1 - C_6)$ haloalkyl, halogen, or Ar^1 ;
- d) R_2 is H, alkyl, Ar^1 , or O substituted alkyl;
- e) Ar^1 is a mono-, bi- or tricyclic, carbo- or heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to three position(s) with halo, hydroxyl, nitro, trifluoromethyl, $C_1 - C_6$ straight or branched chain alkyl or alkenyl, $C_1 - C_4$ alkoxy, $C_1 - C_4$ alkenyloxy, phenoxy, benzyloxy, amino, or a combination thereof; wherein the individual ring sizes are 3-7 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and any combination thereof.

Further preferred Formula V compositions, or pharmaceutically acceptable salts thereof, include Formula Vb compositions, or pharmaceutically acceptable salts thereof, comprising the structure:



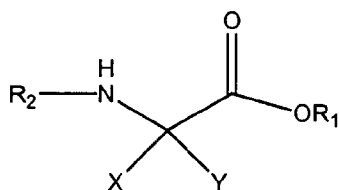
wherein:

- a) X and Y are each carbon;

- b) X and Y are connected by a saturated or partially saturated ring of 3-8 carbons and such a ring may itself be substituted in one to five position(s) with halo, hydroxyl, carboxy, amino, nitro, cyano, trifluoromethyl, C₁-C₆ straight or branched chain alkyl or alkenyl, C₁-C₄ alkoxy, C₁-C₄ alkenyloxy, or substituted alkyl groups;
- c) R¹ is (C₁-C₆) alkyl, Ar¹, or (C₁-C₄) alkoxycarbonylmethyl;
- d) R₂ is H, alkyl, Ar¹, or O substituted alkyl; and
- e) Ar¹ is a mono-, bi- or tricyclic, carbo- or heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to three position(s) with halo, hydroxyl, nitro, trifluoromethyl, C₁-C₆ straight or branched chain alkyl or alkenyl, C₁-C₄ alkoxy, C₁-C₄ alkenyloxy, phenoxy, benzyloxy, amino, or a combination thereof; wherein the individual ring sizes are 3-7 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and any combination thereof.

Further preferred Formula Vb compositions include rings joining X and Y which comprise 3-6 members.

Further preferred Formula V compositions, or pharmaceutically acceptable salts thereof, are Formula Vc compositions, or pharmaceutically acceptable salts thereof, comprising the structure:

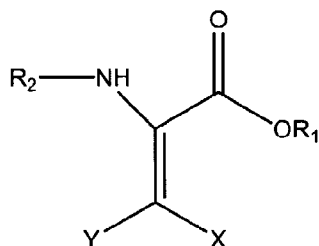


wherein:

- a) X, Y, independently of one another, are H, Ar¹, (C₁-C₆) alkyl (which can be interrupted or substituted by heteroatoms, such as N, P, O, S or Si, it being possible for the heteroatoms themselves to be substituted by (C₁-C₃) alkyl once or several times), (C₂-C₆) alkenyl, (C₁-C₆) haloalkyl, or halogen such as naphthyl or phenyl;
- b) R₂ is H, alkyl, Ar¹, or O substituted alkyl; and
- c) Ar¹ is a mono-, bi- or tricyclic, carbo- or heterocyclic ring, wherein the ring is either unsubstituted or substituted in one to three position(s) with halo, hydroxyl, nitro, trifluoromethyl, C₁-C₆ straight or branched chain alkyl or alkenyl, C₁-C₄ alkoxy, C₁-C₄ alkenyloxy, phenoxy, benzyloxy, amino, or a combination thereof; wherein the individual ring

sizes are 3-7 members; and wherein the heterocyclic ring contains 1-6 heteroatom(s) selected from the group consisting of O, N, S, and any combination thereof.

Formula VI compositions, or pharmaceutically acceptable salts thereof, are represented by the structure comprising:



wherein:

- a) R¹ is (C₁-C₆) alkyl, Ar¹, or (C₁-C₄) alkoxy carbonylmethyl;
- b) R₂ is H, alkyl, Ar¹, or O substituted alkyl;
- c) Y is H, Ar¹, (C₁-C₆) alkyl (which can be interrupted or substituted by heteroatoms, such as N, P, O, S or Si, it being possible for the heteroatoms themselves to be substituted by (C₁-C₃) alkyl once or several times), (C₂-C₆) alkenyl, (C₁-C₆) haloalkyl, or halogen; and
- d) X is alkyl or phenyl.

In further preferred Formula VI compositions, Y is Ar¹. In still further preferred Formula VI compositions, Y is phenyl, naphthyl, 3-formylindole, imidazole, or pyrazole.

Additional compounds of the present invention include but are not limited to: (irreversible or suicide inhibitors): 2-Oxo-pentynoate, Dansyl chloride, Dansyl Fluoride, Propargylglycine, *N*-Acetyl-propargylglycine, *O*-(2,4-dinitrophenyl)hydroxylamine, 1-Chloro-2-nitroethane, 1,2-cyclohexadione, Allylglycine, *N*-chloro-D-leucine, Phenylglyoxal, Ethyl bromopyruvate, Methyl bromopyruvate, and Bromopyruvate. Further compounds include, but are not limited to: Methylglyoxal bis(guanylhydrazone), Hydrazinecarboximidamide, Pyruvaldehyde bis(amidinohydrazone), 3-(3-Indolyl)propanoic acid, 3-indole-acetic acid, Indole-3-acetone, Indole-3-acetamide, Indole-3-acetyl-L-aspartic acid, Indole-3-acetyl-L-alanine, Indole-3-acetyl-glycine, Indole-3-carboxylic acid, Indole-3-pyruvic acid, dansyl glycine, Alanine tetrazole, tetrazole, Riboflavin 5'-pyrophosphate, 5-Hydroxy-2-hydroxymethyl-4-pyranone, Hydroxylamine Hydrochloride, Tetrahydro-4-phenyl-4H-1,4-thiazine 1-oxide, Phenothiazine, 3,4-Dihydro-2H-1,4-thiazine-3,5-dicarboxylic acid, Nifurtimox (1-((5-Nitrofurfurylidene)amino)-2-methyltetrahydro-1,4-thiazine-4,4-dioxide), 2-amino-2,4-pentadienoate, 2-amino-4-keto-2-pentenoate, *N*-Acetyl-D-leucine, D-Leu (D-2-Amino-4-methylpentanoic acid,

Progesterone (4-Pregnene-3,20-dione, FAD (Flavin adenine dinucleotide), 6-OH-FAD (Flavin adenine dinucleotide), *N*-(1-carboxyethyl)-L-Alanine, 3-Aminoaspartic acid, thiocarbamoyl hydrazide, 5-S-Cysteinyl dopamine, phosphatidyl serine, 4-Hydroxy-2-mercapto 6-methylpyrimidine, 4-Amino-*N*-2-thiazolylbenzenesulfonamide, Thiocyanate, 2-Mercapto-1-methylimidazole, tartaric acid, 2-Aminoethanethiol, S-adenosylmethionine, and Thiourea.

The non-limiting listing of antagonists listed herein, or in Table I, may be altered or derivatized utilizing methods known in the art to produce one or more of the following compounds: a) a prodrug; b) a compound with greater enzymatic activity; c) a compound with more specificity for DAO or DDO; d) a compound with lower toxicity; or e) a compound lacking unwanted side effects. Methods for measuring DAO or DDO activity are well known in the art and may be performed using methods disclosed herein or disclosed in a reference cited herein. All of the references cited below for the exemplary DAO or DDO antagonists are incorporated by reference herein in their entireties.

Exemplary DAO or DDO Antagonists

- a) 2-oxo-3-pentynoate; (*Biochemistry* 1999 May 4; 38 (18):5822-8).
- b) Aminoguanidine; (*J Neurochem* 1998 Mar;70(3):1323-6).
- c) Benzoate (benzoic acid) and salts thereof (e.g., sodium benzoate); (*Neurosci Lett* 1996 Nov 8;218(3):145-8).
- d) *o*-, *m*-, and *p*-aminobenzoate (*J. Biochem. (Tokyo)* 1976 Nov, 80(5): 1101-1108)
- e) Methylglyoxal bis(guanylhydrazone) (MGBG); phenylglyoxal bis(guanylhydrazone) (PhGBG); glyoxal bis(guanylhydrazone) (GBG); (*Anticancer Drug Des* 1996 Oct;11(7):493-508).
- f) Alpha-keto acids that are analogs of the amino acids alanine, valine, leucine, phenylalanine, phenylglycine, tyrosine, tryptophan; serine, aspartate; etc. (e.g., pyruvic acid, alpha-ketoisovaleric acid, 4-methylthio-2-oxopentanoic acid, 4-methylthio-2-oxobutanoic acid, phenylpyruvic acid, indol-3-pyruvic acid, benzoylformic acid, 4-hydroxyphenyl pyruvic acid, and salts and derivatives thereof), indole-propionic, 3-indole-acetic acid, salicylic acid, and salts and derivatives thereof (*Enzyme Microb Technol* 1996 Apr;18(5):379-82).
- g) Dansyl chloride; (*FEBS Lett* 1995 Apr 24;363(3):307-10).
- h) Alanine tetrazole and benzoic tetrazole; (*Res Commun Chem Pathol Pharmacol* 1994 Feb;83(2):209-22).
- i) Riboflavin 5'-pyrophosphate (RPP); (*Anal Biochem* 1992 May 1;202(2):348-55).
- j) D-propargylglycine (D-PG); (*J Biochem (Tokyo)* 1991 Jan;109(1):171-7).
- k) D,L-beta-hydroxybutyrate; (*J Histochem Cytochem* 1991 Jan;39(1):81-6).
- l) Trigonelline, i.e., N-methylnicotinate; (*J Biochem (Tokyo)* 1990 May;107(5):726-31).
- m) Kojic acid and salts thereof; (*J Biol Chem* 1989 Feb 15;264(5):2509-17).
- n) O-(2,4-dinitrophenyl)hydroxylamine; (*Biochemistry* 1987 Mar 24;26(6):1717-22).

- o) Benzoate; (D'Silva, C., Willams, C.H., Jr., & Massey, V. (1986) *Biochemistry* 25, 5602-5608).
- p) Methyl-p-nitrobenzenesulfonate; (*J Biol Chem* 1984 May 10;259(9):5585-90).
- q) Aminoethylcysteine-ketimine (2H-1,4-thiazine-5,6-dihydro-3-carboxylic acid); 1,4-thiazine derivatives, ketimine reduced forms (thiomorpholine-2-carboxylic acid and thiomorpholine-2,6-dicarboxylic acid); (*Biochim Biophys Acta* 1983 Oct 17;748(1):40-7).
- r) The reaction product between cysteamine and bromopyruvate; (*J Appl Biochem* 1983 Aug-Oct;5(4-5):320-9).
- s) 1-chloro-1-nitroethane; (*J Biol Chem* 1983 Jan 25;258(2):1136-41)
- t) Benzoate; anthranilate; picolinate; L-leucine; (*J Biol Chem* 1982 Sep 10;257(17):9958-62).
- u) Fluorodinitrobenzene; (Nishino, T., Massey, V., and Williams, C.H., Jr. (1980) *J Biol Chem* 255, 3610-3616).
- v) 1,2-cyclohexanedione; (*Eur J Biochem* 1981 Oct;119(3):553-7).
- w) Allylglycine; 2-amino-2,4-pentadienoate; 2-hydroxy-2,4-pentadienoate; (*Biochemistry* 1978 Dec 26;17(26):5620-6).
- x) 2-amino-4-keto-2-pentenoate; (*Biochemistry* 1978 Dec 26;17(26):5613-9).
- y) D,L-2-hydroxybutyrate; (*J Cell Biol* 1978 Apr;77(1):59-71).
- z) P-aminobenzoate; (*J Biochem (Tokyo)* 1976 Nov;80(5):1073-83).
- aa) N-chloro-D-leucine; (*J Biol Chem* 1976 Oct 10;251(19):6150-3).
- bb) D-propargylglycine; (*Biochemistry* 1976 Jul 13;15(14):3070-6).
- cc) D-2-amino-4-pentynoic acid (D-propargylglycine); (*J Biochem (Tokyo)* 1975 Jul; 78(1):57-63).
- dd) Progesterone (*Biochim Biophys Acta* 1978 Jan 12;522(1):43-8).
- ee) Long chain, medium chain and short chain free fatty acids (*Biochem Int* 1990 Dec;22(5):837-42).
- ff) 6-OH-FAD (*Biochim Biophys Acta* 1999 Apr 12;1431(1):212-22).
- gg) Phenylglyoxal, L-tartrate (*Eur J Biochem* 1992 Apr 1;205(1):127-32).
- hh) Cyclothionine, TMDA, alpha-alpha'-iminodipropionic (*Physiol Chem Phys Med NMR* 1986;18(1):71-4).
- ii) Inhibitors disclosed in *J Histochem Cytochem* 1990 Sep;38(9):1377-81.
- jj) Meso-Diaminosuccinic aci(*Eur J Biochem* 1981 Jul;117(3):635-8).
- kk) Thiosemicarbazide, thiourea, methylthiouracyl, sulphathiazole, thiocyanate, and methimazole (*Endocrinol Exp* 1976;10(4):243-51).
- ll) Dicarboxylic hydroxyacids (*Enzymologia* 1967 Dec 31;33(6):325-30).
- mm) Malic and tartaric acid (*Boll Soc Ital Biol Sper* 1966 Oct 31;42(20):1455-7).

It should be appreciated that DAO and DDO antagonists of the invention include the compounds listed above and throughout the specification, as well as the salts and derivatives thereof these compounds.

Methods of Screening for Compounds Modulating DAO or DDO Expression and/or Activity

Methods that can be used for testing antagonistic compounds for their ability to inhibit or decrease the activity of a DAO or DDO polypeptide or inhibit or decrease the expression of a DAO or DDO gene product (mRNA or polypeptide) are well known in the art. Suitable DAO and DDO polypeptides useful for methods of screening include both recombinant DAO and DDO or DAO and DDO polypeptides purified from tissue (e.g., hog kidneys). Preferred DAO and DDO polypeptides, and polynucleotides useful to make said polypeptides, are the human DAO and DDO sequences of Figures 1 and 2. Preferred antagonists of the present invention are antagonists of the polypeptides of Figures 1 and 2. Further preferred antagonists of the present invention inhibit the oxidative deamination of D-amino acids. Further preferred antagonists of the present invention inhibit the oxidative deamination of D-Serine or D-Aspartate. The assays described herein and known in the art for measuring DAO or DDO enzymatic activity can be performed either *in vitro* or *in vivo*.

Antagonists according to the present invention include naturally occurring and synthetic compounds and small molecules. Antagonists of the present invention may either block binding of DAO or DDO to either its cofactor, FAD, or substrate, or block enzymatic activity, e.g., oxidative deamination of D-amino acids. Whether any candidate antagonist of the present invention can enhance or inhibit DAO or DDO activity is determined using well known methods in the art for measuring DAO or DDO activity. One method for screening involves contacting a sample comprising a DAO or DDO polypeptide with a test compound and assaying DAO or DDO activity in the presence of a substrate. The level of DAO or DDO activity is compared to a sample that does not contain the test compound, whereby a decreased DAO or DDO level of activity over the standard indicates that the candidate compound is an antagonist of DAO or DDO. DAO or DDO activity can be measured as an isolated or purified enzyme or in a biological sample comprising cells or tissue expressing DAO or DDO.

Alternatively, one of skill in the art can identify compounds that inhibit expression of a DAO or DDO gene product (mRNA or polypeptide). Cells expressing DAO or DDO (e.g., liver, kidney, or brain cells) are incubated in the presence and absence of the test compound. By measuring the expression level of a DAO or DDO gene product in the presence and absence of the test compound or the level of DAO or DDO activity in the presence and absence of the test compound, compounds can be identified that suppress expression of a DAO or DDO gene product. Alternatively, constructs comprising a DAO or DDO regulatory sequence operably linked to a reporter gene (e.g. luciferase, chloramphenicol acetyl transferase, LacZ, green fluorescent polypeptide, beta galactosidase ,etc.) can be introduced into host

cells and the effect of the test compounds on expression of the reporter gene detected. Cells suitable for use in the foregoing assays include, but are not limited to, cells having the same origin as tissues or cell lines in which the polypeptide is known to be expressed (e.g., kidney, liver and brain). The quantification of the expression of a DAO or DDO polypeptide may be realized either at the mRNA level (using for example Northern blots, RT-PCR, preferably quantitative RT-PCR with primers and probes specific for the DAO or DDO mRNA of interest) or at the polypeptide level (by measuring DAO or DDO enzymatic activity or by using polyclonal or monoclonal antibodies in immunoassays such as ELISA or RIA assays, Western blots, immunochemistry).

In other aspects, an assay is a cell-based assay in which a cell which expresses a DAO or DDO protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to inhibit, activate, or increase DAO or DDO activity determined. Determining the ability of the test compound to inhibit, activate, or increase DAO or DDO activity can be accomplished by monitoring the bioactivity of the DAO or DDO protein or biologically active portion thereof. Preferably, amino acid oxidation is monitored. The cell, for example, can be of mammalian origin, bacterial origin or a yeast cell. For example, in some embodiments, the cell can be a mammalian cell, bacterial cell or yeast cell.

The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is used with peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K. S. (1997) *Anticancer Drug Des.* 12:145, the disclosure of which is incorporated herein by reference in its entirety).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt et al. (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann et al. (1994). *J. Med. Chem.* 37:2678; Cho et al. (1993) *Science* 261:1303; Carrell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carrell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop et al. (1994) *J. Med. Chem.* 37:1233, the disclosures of which are incorporated herein by reference in their entireties.

Libraries of compounds may be presented in solution (e.g., Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner U.S. Pat. No. 5,223,409), spores (Ladner U.S. Pat. No. '409), plasmids (Cull et al. (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devin (1990) *Science* 249:404-406); (Cwirla et al. (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382);

(Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*), the disclosures of which are incorporated herein by reference in their entireties.

Determining the ability of the test compound to inhibit DAO or DDO activity can also be accomplished, for example, by coupling the DAO or DDO protein or biologically active portion thereof with a radioisotope or enzymatic label such that binding of the DAO or DDO protein or biologically active portion thereof to its cognate target molecule can be determined by detecting the labeled DAO or DDO protein or biologically active portion thereof in a complex. Preferably, a DAO or DDO 'target molecule' is a molecule with which a DAO or DDO protein binds or interacts in nature, such that DAO or DDO -mediated function is achieved. In one example, a DAO target molecule is a g34872 polypeptide. For example, compounds (e.g., DAO or DDO protein or biologically active portion thereof) can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. The labeled molecule is placed in contact with its cognate molecule and the extent of complex formation is measured. For example, the extent of complex formation may be measured by immuno precipitating the complex or by performing gel electrophoresis.

It is also within the scope of this invention to determine the ability of a compound (e.g., DAO or DDO protein or biologically active portion thereof) to interact with its cognate target molecule without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with its cognate target molecule without the labeling of either the compound or the target molecule. McConnell, H. M. et al. (1992) *Science* 257:1906-1912, the disclosure of which is incorporated herein by reference in its entirety. A microphysiometer such as a cytosensor is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between compound and receptor.

In a preferred embodiment, the assay comprises contacting a cell which expresses or which is responsive to a DAO or DDO protein or biologically active portion thereof, with a target molecule to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to inhibit or increase the activity of the DAO or DDO protein or biologically active portion thereof, wherein determining the ability of the test compound to inhibit or increase the activity of the DAO or DDO protein or biologically active portion thereof, comprises determining the ability of the test compound to inhibit or increase a biological activity of the DAO or DDO expressing cell (e.g.,

determining the ability of the test compound to inhibit or increase transduction, protein:protein interactions, substrate binding).

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a DAO or DDO target molecule (i.e. a molecule with which DAO or DDO interacts) with a test compound and determining the ability of the test compound to modulate (e.g. stimulate or inhibit respectively) the activity of the DAO or DDO target molecule. Determining the ability of the test compound to modulate the activity of a DAO or DDO target molecule can be accomplished, for example, by determining the ability of the DAO or DDO protein to bind to or interact with the DAO or DDO target molecule.

Determining the ability of the DAO or DDO protein to bind to or interact with a DAO or DDO target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the DAO or DDO protein to bind to or interact with a DAO or DDO target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by contacting the target molecule with the DAO or DDO protein or a fragment thereof and measuring induction of a cellular second messenger of the target (i.e. intracellular Ca^{2+} , diacylglycerol, IP₃, etc.), detecting catalytic/enzymatic activity of the target on appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, e.g., luciferase), or detecting a target-regulated cellular response, for example, signal transduction or protein:protein interactions.

In other preferred embodiments, an assay of the present invention is a cell-free assay in which a DAO or DDO protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the DAO or DDO protein or biologically active portion thereof is determined. Binding of the test compound to the protein can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the DAO or DDO protein or biologically active portion thereof with a known compound which binds DAO or DDO (e.g., a DAO or DDO target molecule) to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a DAO or DDO protein, wherein determining the ability of the test compound to interact with a DAO or DDO protein comprises determining the ability of the test compound to preferentially bind to DAO or DDO or biologically active portion thereof as compared to the known compound.

In another embodiment, the assay is a cell-free assay in which a DAO or DDO protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (e.g., inhibit the activity of the DAO or DDO or activate the activity of the DAO or DDO) the activity of the protein or biologically active portion thereof is determined. Determining the ability of the test compound to modulate the activity of a protein can be accomplished, for example, by

determining the ability of the protein to bind to a target molecule by one of the methods described above for determining direct binding. This can also be accomplished for example using a technology such as real-time Biomolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo et al. (1995) *Curr. Opin. Struct. Biol.* 5:699-705, the disclosures of which are incorporated herein by reference in their entireties. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (e.g., BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In an alternative embodiment, determining the ability of the test compound to modulate the activity of a DAO or DDO protein can be accomplished by determining the ability of the DAO or DDO protein to further modulate the activity of a downstream effector a DAO or DDO target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined as previously described.

In yet another embodiment, the cell-free assay involves contacting a DAO or DDO protein or biologically active portion thereof with a known compound which binds the DAO or DDO protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the DAO or DDO protein, wherein determining the ability of the test compound to interact with the DAO or DDO protein comprises determining the ability of the DAO or DDO protein to preferentially bind to or modulate the activity of a DAO or DDO target molecule.

The cell-free assays of the present invention are amenable to use of both soluble and/or membrane-bound forms of isolated proteins (e.g. DAO or DDO proteins or biologically active portions thereof or molecules to which DAO or DDO targets bind). In the case of cell-free assays in which a membrane-bound form an isolated protein is used it may be desirable to utilize a solubilizing agent such that the membrane-bound form of the isolated protein is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, TritonTM X-100, TritonTM X-114, ThesitTM, Isotridecypoly(ethylene glycol ether)_n, 3-(3-cholamidopropyl)dimethylamminio-1-propane sulfonate (CHAPS), 3-(3-cholamidopropyl)dimethylamminio-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either a DAO or DDO protein or a target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a DAO or DDO protein, or interaction of a DAO or DDO protein with a target molecule in the presence and absence of a candidate compound, can be

accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/ DAO or DDO fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or DAO or DDO protein, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of DAO or DDO binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a DAO or DDO protein or a DAO or DDO target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated DAO or DDO protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with DAO or DDO protein or target molecules but which do not interfere with binding of the DAO or DDO protein to its target molecule can be derivatized to the wells of the plate, and unbound target or DAO or DDO protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the DAO or DDO protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the DAO or DDO protein or target molecule.

In yet another aspect of the invention, the proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Pat. No. 5,283,317; Zervos et al. (1993) *Cell* 72:223-232; Madura et al. (1993) *J. Biol. Chem.* 268:12046-12054; Bartel et al. (1993) *Biotechniques* 14:920-924; Iwabuchi et al. (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300, the disclosures of which are incorporated herein by reference in their entireties), to identify other proteins which bind to or interact with the DAO or DDO proteins, and/or are involved in the activity of the DAO or DDO proteins.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a DAO or DDO protein or a fragment thereof is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the

other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a DAO or DDO-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the DAO or DDO protein.

This invention further pertains to novel agents identified by the above-described screening assays and to processes for producing such agents by use of these assays. Accordingly, in one embodiment, the present invention includes a compound or agent obtainable by a method comprising the steps of any one of the aforementioned screening assays (e.g., cell-based assays or cell-free assays). For example, in one embodiment, the invention includes a compound or agent obtainable by a method comprising contacting a cell which expresses a DAO or DDO target molecule with a test compound and determining the ability of the test compound to bind to, or modulate the activity of, the DAO or DDO target molecule. In another embodiment, the invention includes a compound or agent obtainable by a method comprising contacting a cell which expresses a DAO or DDO target molecule with a DAO or DDO protein or biologically-active portion thereof, to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with, or modulate the activity of, the DAO or DDO target molecule. In another embodiment, the invention includes a compound or agent obtainable by a method comprising contacting a DAO or DDO protein or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to inhibit the activity of, the DAO or DDO protein or biologically active portion thereof. In yet another embodiment, the present invention included a compound or agent obtainable by a method comprising contacting a DAO or DDO protein or biologically active portion thereof with a known compound which binds the DAO or DDO protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with, or modulate the activity of the DAO or DDO protein.

Antagonist compounds that inhibit DAO or DDO activity or inhibit expression of a DAO or DDO gene product can also be identified using *in vivo* screens. In these assays, the test compound is administered (e.g. IV, IP, IM, orally, or otherwise), to the animal, for example, at a variety of dose levels. The effect of the test compound on DAO or DDO activity or gene product expression is determined by comparing the levels of DAO or DDO activity or gene product expression, respectively, in the tissues of test and control animals that express DAO or DDO. Suitable test animals include, but

not limited to, rodents (e.g., mice and rats) and primates. Humanized non-human animals, such as humanized mice, can also be used as test animals, that is, animals in which the endogenous polypeptide is ablated (knocked out) and the homologous human polypeptide added back by standard transgenic approaches. Such animals express only the human form of a polypeptide.

In vivo assays also include animal models for CNS disorders. These models include, but are not limited to: conditioned avoidance behavior in rats model; gerbil foot-tapping model; ferret emesis model; separation-induced vocalization model; behavioral activity assessment of mice and rats in the omnitech digiscan animal activity monitors; blockade of amphetamine-stimulated locomotion in rats model; prepulse inhibition (PPI) of acoustic startle in rats model; inhibition of apomorphine-induced climbing behaviour model; and the DOI-induced head twitches and scratches model as described herein and known in the art.

Other antagonists of the present invention include antisense and triple helix tools to inhibit expression of a DAO or DDO gene product. In antisense approaches, nucleic acid sequences complementary to a DAO or DDO mRNA or genomic sequence are hybridized to the DAO or DDO mRNA or genomic DNA intracellularly, thereby blocking the expression of the DAO or DDO polypeptide encoded by the mRNA. The antisense nucleic acid molecules to be used in DAO or DDO therapy may be either DNA or RNA sequences. Preferred methods using antisense polynucleotide according to the present invention are the procedures described by Sczakiel *et al.*(1995), which disclosure is hereby incorporated by reference in its entirety. Other preferred antisense polynucleotides according to the present invention are sequences complementary to either a sequence of DAO or DDO mRNAs comprising the translation initiation codon ATG or a sequence of DAO or DDO.

It is preferable that the antisense polynucleotides comprise sequences complementary to a DAO or DDO initiation codon (ATG) or genomic DNA containing a splicing donor or acceptor site. It is also preferable that the antisense polynucleotides of the invention have a 3' polyadenylation signal that has been replaced with a self-cleaving ribozyme sequence, such that RNA polymerase II transcripts are produced without poly(A) at their 3' ends, these antisense polynucleotides being incapable of export from the nucleus, such as described by Liu *et al.*(1994), which disclosure is hereby incorporated by reference in its entirety. The DAO or DDO antisense polynucleotides may also comprise, within the ribozyme cassette, a histone stem-loop structure to stabilize cleaved transcripts against 3'-5' exonucleolytic degradation, such as the structure described by Eckner *et al.*(1991), which disclosure is hereby incorporated by reference in its entirety.

The antisense nucleic acids should have a length and melting temperature sufficient to permit formation of an intracellular duplex having sufficient stability to inhibit the expression of the DAO or DDO mRNA in the duplex. Strategies for designing antisense nucleic acids suitable for use in DAO or

DDO therapy are disclosed in Green *et al.*, (1986) and Izant and Weintraub, (1984), the disclosures of which are incorporated herein by reference.

In some strategies, antisense molecules are obtained by reversing the orientation of the DAO or DDO coding region with respect to a promoter so as to transcribe the opposite strand from that which is normally transcribed in the cell. Another approach involves transcription of DAO or DDO antisense nucleic acids *in vivo* by operably linking DNA containing the antisense sequence to a promoter in a suitable expression vector.

Alternatively, oligonucleotides which are complementary to the strand normally transcribed in the cell may be synthesized *in vitro*. Thus, the antisense nucleic acids are complementary to the corresponding mRNA and are capable of hybridizing to the mRNA to create a duplex. The antisense sequences may also contain modified sugar phosphate backbones to increase stability and make them less sensitive to RNase activity. Examples of modifications suitable for use in antisense strategies include 2' O-methyl RNA oligonucleotides and polypeptide-nucleic acid (PNA) oligonucleotides. Further examples are described by Rossi *et al.*, (1991), which disclosure is hereby incorporated by reference in its entirety.

Various types of antisense oligonucleotides complementary to the sequence of the DAO or DDO cDNA or genomic DNA may be used. For example, stable and semi-stable antisense oligonucleotides described in International Application No. PCT WO94/23026, hereby incorporated by reference, can be used. In these molecules, the 3' end or both the 3' and 5' ends are engaged in intramolecular hydrogen bonding between complementary base pairs. These molecules are better able to withstand exonuclease attacks and exhibit increased stability compared to conventional antisense oligonucleotides.

In yet another method of using antisense technology to inhibit expression of a DAO or DDO polypeptide, the covalently cross-linked antisense oligonucleotides described in International Application No. WO 96/31523, hereby incorporated by reference, is used. These double- or single-stranded oligonucleotides comprise one or more, respectively, inter- or intra-oligonucleotide covalent cross-linkages, wherein the linkage consists of an amide bond between a primary amine group of one strand and a carboxyl group of the other strand or of the same strand, respectively, the primary amine group being directly substituted in the 2' position of the strand nucleotide monosaccharide ring, and the carboxyl group being carried by an aliphatic spacer group substituted on a nucleotide or nucleotide analog of the other strand or the same strand, respectively.

The antisense oligodeoxynucleotides and oligonucleotides disclosed in International Application No. WO 92/18522, incorporated by reference, may also be used. These molecules are stable to degradation and contain at least one transcription control recognition sequence which binds to control polypeptides and are effective as decoys therefor. These molecules may contain "hairpin" structures,

"dumbbell" structures, "modified dumbbell" structures, "cross-linked" decoy structures and "loop" structures.

Further, the cyclic double-stranded oligonucleotides described in European Patent Application No. 0 572 287 A2, hereby incorporated by reference may be used. These ligated oligonucleotide "dumbbells" contain the binding site for a transcription factor and inhibit expression of the DAO or DDO under control of the transcription factor by sequestering the factor.

Use of the closed antisense oligonucleotides disclosed in International Application No. WO 92/19732, hereby incorporated by reference, is also an alternative. Because these molecules have no free ends, they are more resistant to degradation by exonucleases than are conventional oligonucleotides. These oligonucleotides may be multifunctional, interacting with several regions which are not adjacent to the target mRNA.

The appropriate level of antisense nucleic acids required to inhibit DAO or DDO expression may be determined using *in vitro* expression analysis. The antisense molecule may be introduced into the cells by diffusion, injection, infection or transfection using procedures known in the art. For example, the antisense nucleic acids can be introduced into the body as a bare or naked oligonucleotide, oligonucleotide encapsulated in lipid, oligonucleotide sequence encapsulated by viral polypeptide, or as an oligonucleotide operably linked to a promoter contained in an expression vector. The expression vector may be any of a variety of expression vectors known in the art, including retroviral or viral vectors, vectors capable of extrachromosomal replication, or integrating vectors. The vectors may be DNA or RNA.

The antisense molecules are introduced onto cell samples at a number of different concentrations preferably between $1 \times 10^{-10} \text{M}$ to $1 \times 10^{-4} \text{M}$. Once the minimum concentration that can adequately control DAO or DDO expression is identified, the optimized dose is translated into a dosage suitable for use *in vivo*. For example, an inhibiting concentration in culture of 1×10^{-7} translates into a dose of approximately 0.6 mg/kg bodyweight. Levels of oligonucleotide approaching 100 mg/kg bodyweight or higher may be possible after testing the toxicity of the oligonucleotide in laboratory animals. It is additionally contemplated that cells from the vertebrate are removed, treated with the antisense oligonucleotide, and reintroduced into the vertebrate.

In a preferred application of this invention, the polypeptide encoded by the DAO or DDO is first identified or the enzymatic activity measured, so that the effectiveness of antisense inhibition on translation can be monitored using techniques that include but are not limited to antibody-mediated tests such as RIAs and ELISA, functional assays, or radiolabeling, and assays to measure DAO or DDO activity.

An alternative to the antisense technology that is used according to the present invention to inhibit expression of a DAO or DDO gene product comprises using ribozymes that will bind to a DAO

or DDO target sequence via their complementary polynucleotide tail and that will cleave the corresponding DAO or DDO RNA by hydrolyzing its target site (namely "hammerhead ribozymes"). Briefly, the simplified cycle of a hammerhead ribozyme comprises (1) sequence specific binding to the target DAO or DDO RNA via complementary antisense sequences; (2) site-specific hydrolysis of the cleavable motif of the target DAO or DDO strand; and (3) release of cleavage products, which gives rise to another catalytic cycle. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, *Nature* 20334:585-591 (1988). Indeed, the use of long-chain antisense polynucleotide (at least 30 bases long) or ribozymes with long antisense arms are advantageous. A preferred delivery system for antisense ribozyme is achieved by covalently linking these antisense ribozymes to lipophilic groups or to use liposomes as a convenient vector. Preferred antisense ribozymes according to the present invention are prepared as described by Rossi *et al.*, (1991) and Sczakiel *et al.* (1995), the specific preparation procedures being referred to in said articles being herein incorporated by reference.

The DAO or DDO genomic DNA may also be used to inhibit the expression of the DAO or DDO based on intracellular triple helix formation. Triple helix oligonucleotides are used to inhibit transcription from a genome. They are particularly useful for studying alterations in cell DAO or DDO activity. The DAO or DDO cDNAs or genomic DNA or a fragment of those sequences, can be used to inhibit DAO or DDO expression in individuals having a CNS disorder associated with expression of a particular DAO or DDO. Similarly, a portion of the DAO or DDO genomic DNA can be used to study the effect of inhibiting DAO or DDO transcription within a cell. Traditionally, homopurine sequences are considered the most useful for triple helix strategies. However, homopyrimidine sequences may also be used to inhibit DAO or DDO expression. Such homopyrimidine oligonucleotides bind to the major groove at homopurine:homopyrimidine sequences.

To carry out DAO or DDO therapy strategies using the triple helix approach, the sequences of the DAO or DDO genomic DNA are first scanned to identify 10-mer to 20-mer homopyrimidine or homopurine stretches which could be used in triple-helix based strategies for inhibiting DAO or DDO expression. Following identification of candidate homopyrimidine or homopurine stretches, their efficiency in inhibiting DAO or DDO expression is assessed by introducing varying amounts of oligonucleotides containing the candidate sequences into tissue culture cells which express the DAO or DDO. Treated cells are monitored for altered DAO or DDO enzymatic activity or reduced DAO or DDO expression as described above.

The oligonucleotides which are effective in inhibiting DAO or DDO expression in tissue culture cells may then be introduced *in vivo* using the techniques and at a dosage calculated based on the *in vitro* results, as described for antisense polynucleotides.

In some embodiments, the natural (beta) anomers of the oligonucleotide units can be replaced with alpha anomers to render the oligonucleotide more resistant to nucleases. Further, an intercalating agent such as ethidium bromide, or the like, can be attached to the 3' end of the alpha oligonucleotide to stabilize the triple helix. For information on oligonucleotides suitable for triple helix formation see Griffin *et al.* (Science 245:967-71, 1989), which is hereby incorporated by this reference.

Pharmaceutical and Physiologically Acceptable Compositions and Administration Thereof

The compounds and compositions for use in the invention can be prepared utilizing readily available starting materials and employing common synthetic methodologies well-known to those skilled in the art. Alternatively, compounds useful in the practice of the invention can be purchased from commercial vendors, such as Sigma Chemical Company (St. Louis, MO).

The relative activity, potency and specificity of a DAO or DDO antagonist can be determined by a pharmacological study in animals according to the method of Nyberg *et al.* [Psychopharmacology 119, 345-348 (1995)], described herein, or known in the art. The test provides an estimate of relative activity, potency and, through a measure of specificity, an estimate of therapeutic index. Other animal studies which may be used include, but are not limited to, studies involving conditioned avoidance, apomorphine induced climbing, blockade of 5-hydroxy-tryptophan-induced head twitching and other animal models disclosed herein or known in the art. Although the differential metabolism among patient populations can be determined by a clinical study in humans, less expensive and time-consuming substitutes are provided by the methods of Kerr *et al.* [Biochem. Pharmacol. 47, 1969-1979 (1994)] and Karam *et al.* [Drug Metab. Dispos. 24, 1081-1087 (1996)]. Similarly, the potential for drug-drug interactions may be assessed clinically according to the methods of Leach *et al.* [Epilepsia 37, 1100-1106 (1996)] or in vitro according to the methods of Kerr *et al.* [op. cit.] and Turner and Renton [Can. J. Physiol. Pharmacol. 67, 582-586 (1989)]. In addition, the relative activity, potency and specificity of a DAO or DDO antagonist may be tested using various in vitro assays.

The effective dose can vary, depending upon factors such as the condition of the patient, the severity of the symptoms of the disorder, and the manner in which the pharmaceutical composition is administered. For human patients, the effective dose of typical compounds generally requires administering the compound in an amount of at least about 1, often at least about 10, and frequently at least about 25 mg/24 hr./patient. For human patients, the effective dose of typical compounds requires administering the compound which generally does not exceed about 500, often does not exceed about 400, and frequently does not exceed about 300 mg/24 hr./patient. In addition, administration of the effective dose is such that the concentration of the compound within the plasma of the patient normally does not exceed 500 ng/ml, and frequently does not exceed 100 ng/ml.

The compounds and compositions of the present invention can be administered to a patient at dosage levels in the range of about 0.1 to about 1,000 mg per day. For a normal human adult having a body weight of about 70 kilograms, it is estimated that a dosage in the range of about 0.01 to about 100 mg per kilogram of body weight per day is sufficient. The specific dosage used, however, can vary. For example, the dosage can depend on a numbers of factors including the requirements of the patient, the severity of the condition being treated, and the pharmacological activity of the compound being used. The determination of optimum dosages for a particular patient is well-known to those skilled in the art. The amount of active ingredient that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host treated and the particular mode of administration. It will be understood, however, that the specific dose level for any particular patient will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, sex, diet, time of administration, route of administration, and rate of excretion, drug combination and the severity of the particular disease undergoing therapy.

In some embodiments, various combinations of DAO or DDO antagonists can be used in the practice of the invention. Thus, compositions containing more than one DAO antagonist can be used to in therapeutic methodologies according to the invention. Alternatively, compositions containing more than one DDO antagonist can be used in the disclosed methodologies. In yet another embodiment, combinations of at least one DAO antagonist and at least one DDO antagonist can be used in treatment methodologies disclosed herein.

Preferred compounds useful according to the method of the present invention have the ability to pass across the blood-brain barrier of the patient. As such, such compounds have the ability to enter the central nervous system of the patient. The log P values of typical compounds useful in carrying out the present invention generally are greater than 0, often are greater than about 1, and frequently are greater than about 1.5. The log P values of such typical compounds generally are less than about 4, often are less than about 3.5, and frequently are less than about 3. Log P values provide a measure of the ability of a compound to pass across a diffusion barrier, such as a biological membrane. See, Hansch, et al., J. Med. Chem., Vol. 11, p. 1 (1968). Alternatively, the compositions of the present invention can bypass the blood brain barrier through the use of compositions and methods known in the art for bypassing the blood brain barrier (e.g., U.S. Patent Nos. 5,686,416; 5,994,392, incorporated by reference in their entireties) or can be injected directly into the brain. Suitable areas for injection include the cerebral cortex, cerebellum, midbrain, brainstem, hypothalamus, spinal cord and ventricular tissue, and areas of the PNS including the carotid body and the adrenal medulla. The compositions can be administered as a bolus or through the use of other methods such as an osmotic pump.

The compounds of the present invention can be administered to a patient alone or as part of a composition that contains other components such as excipients, diluents, and carriers, all of which are

well-known in the art. The compositions can be administered to humans and animals either orally, rectally, parenterally (intravenous, by intramuscularly or subcutaneously), intracisternally, intravaginally, intraperitoneally, intravesically, locally (powders, ointments or drops), or as a buccal or nasal spray, or inhaled.

Compositions suitable for parenteral injection can comprise physiologically acceptable sterile aqueous or nonaqueous solutions, dispersions, suspensions or emulsions, and sterile powders for reconstitution into sterile injectable solutions or dispersions. Examples of suitable aqueous and nonaqueous carriers, diluents, solvents or vehicles include water, ethanol, polyols (propyleneglycol, polyethyleneglycol, glycerol, and the like), suitable mixtures thereof, vegetable oils (such as olive oil) and injectable organic esters such as ethyl oleate. Proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersions and by the use of surfactants.

These compositions can also contain adjuvants such as preserving, wetting, emulsifying, and dispensing agents. Prevention of the action of microorganisms can be ensured by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, and the like. It may also be desirable to include isotonic agents, for example sugars, sodium chloride, and the like. Prolonged absorption of the injectable pharmaceutical form can be brought about by the use of agents delaying absorption, for example, aluminum monostearate and gelatin.

Solid dosage forms for oral administration include capsules, tablets, pills, powders, and granules. In such solid dosage forms, the active compound is admixed with at least one customary inert excipient (or carrier) such as sodium citrate or dicalcium phosphate or (a) fillers or extenders, as for example, starches, lactose, sucrose, glucose, mannitol, and silicic acid; (b) binders, as for example, carboxymethylcellulose, alginates, gelatin, polyvinylpyrrolidone, sucrose and acacia; (c) humectants, as for example, glycerol; (d) disintegrating agents, as for example, agar-agar, calcium carbonate, potato or tapioca starch, alginic acid, certain complex silicates and sodium carbonate; (e) solution retarders, as for example paraffin; (f) absorption accelerators, as for example, quaternary ammonium compounds; (g) wetting agents, as for example, cetyl alcohol and glycerol monostearate; (h) adsorbents, as for example, kaolin and bentonite; and (i) lubricants, as for example, talc, calcium stearate, magnesium stearate, solid polyethylene glycols, sodium lauryl sulfate, or mixtures thereof. In the case of capsules, tablets, and pills, the dosage forms may also comprise buffering agents.

Solid compositions of a similar type may also be employed as fillers in soft and hard-filled gelatin capsules using such excipients as lactose or milk sugar as well as high molecular weight polyethylene glycols, and the like. Solid dosage forms such as tablets, dragees, capsules, pills, and granules can be prepared with coatings and shells, such as enteric coatings and others well-known in the art. They may contain opacifying agents and can also be of such composition that they release the active

compound or compounds in a certain part of the intestinal tract in a delayed manner. Examples of embedding compositions which can be used are polymeric substances and waxes. The active compounds can also be in micro-encapsulated form, if appropriate, with one or more of the above-mentioned excipients.

Liquid dosage forms for oral administration include pharmaceutically acceptable emulsions, solutions, suspensions, syrups, and elixirs. In addition to the active compounds, the liquid dosage forms may contain inert diluents commonly used in the art, such as water or other solvents, solubilizing agents and emulsifiers, as for example, ethyl alcohol, isopropyl alcohol, ethyl carbonate, ethyl acetate, benzyl alcohol, benzyl benzoate, propylene glycol, 1,3-butylene glycol, dimethylformamide, oils, in particular, cottonseed oil, groundnut oil, corn germ oil, olive oil, castor oil and sesame oil, glycerol, tetrahydrofurfuryl alcohol, polyethylene glycols and fatty acid esters of sorbitan or mixtures of these substances, and the like. Besides such inert diluents, the composition can also include adjuvants, such as wetting agents, emulsifying and suspending agents, sweetening, flavoring, and perfuming agents.

Suspensions, in addition to the active compounds, may contain suspending agents, as for example, ethoxylated isostearyl alcohols, polyoxyethylene sorbitol and sorbitan esters, microcrystalline cellulose, aluminum metahydroxide, bentonite, agar-agar and tragacanth, or mixtures of these substances, and the like.

Compositions for rectal administrations are preferably suppositories which can be prepared by mixing the compounds of the present invention with suitable nonirritating excipients or carriers such as cocoa butter, polyethylene glycol or a suppository wax, which are solid at ordinary temperatures but liquid at body temperature and therefore, melt in the rectum or vaginal cavity and release the active component.

Dosage forms for topical administration of a compound of this invention include ointments, powders, sprays, and inhalants. The active component is admixed under sterile conditions with a physiologically acceptable carrier and any preservative, buffers, or propellants as may be required. Ophthalmic formulations, eye ointments, powders, and solutions are also contemplated as being within the scope of this invention.

In addition, the compounds of the present invention can exist in unsolvated as well as solvated forms with pharmaceutically acceptable solvents such as water, ethanol, and the like. In general, the solvated forms are considered equivalent to the unsolvated forms for the purposes of the present invention.

Animal Models***Conditioned Avoidance Behavior in Rats***

The conditioned avoidance model is a standard behavioural test predictive of antipsychotic activity. One of the major pharmacological properties of currently employed clinical antipsychotic drugs in animals is their ability to block conditioned avoidance responding. See e.g., Cook, L. and Davidson, A. B.: Behavioral pharmacology: Animal models involving aversive control of behavior. In Psychopharmacology, A Generation of Progress, ed by M. A. Lipton, A. Dimascio and K. Killam, pp. 563-567, Raven Press, New York, 1978; Davidson, A. B. and Weidley, E. Differential Effects of Neuroleptic and other Psychotropic Agents on Acquisition of Avoidance in Rats, 18 Life Sci. 1279-1284 (1976), incorporated by reference herein in their entireties. There is a high correlation between their activity and potency on a conditioned avoidance test and their clinical efficacy and potencies as antipsychotic drugs. See e.g., Creese, I., Burt, D. R. and Snyder, S. H.: Dopamine receptor binding predicts clinical and pharmacological properties of antischizophrenic drugs. Science (Washington D.C.) 192:481-483, 1976, incorporated by reference herein in its entirety.

In a conditioned avoidance test, animals learn to respond during a conditioned stimulus in order to avoid mild shock presentation. A response during the conditioned stimulus is termed an avoidance response, a response during shock is termed an escape response; a response failure is when the animal fails to respond during either the conditioned stimulus or the shock presentation and is indicative of motor impairment. Animals rapidly learn to avoid 99% of the time. Antipsychotic drugs decrease the percentage of avoidance responses without interfering with the ability of the animal to respond since the animals do emit escape responses. The percentage of response failures is considered a measure of motor impairment.

Rats are required to press a response lever in an experimental chamber in order to avoid or escape foot-shock. Each experimental session consists of 50 trials. During each trial, the chamber is illuminated and a tone presented for a maximum of 10 sec. A response during the tone immediately terminates the tone and the houselight, ending the trial. In the absence of a response during the tone alone, tone+foot-shock (2.0 mA) is presented for a maximum of 10 sec. A response during shock presentation immediately terminates the shock, the tone and the houselight, ending the trial.

For drug screening, an appropriate dose, e.g., 3.0 mg/kg, is administered in an appropriate manner, e.g., i.p. or s.c., for an appropriate time, 30 min, before the start of the experimental session. The treated group may receive only a single dose of the DAO or DDO antagonist or alternatively, may be treated daily (e.g., sid, bid, or tid) for at least 1 day, 3 days, 1 week, 2 weeks, 1 month, or 2 months, prior to the experiment. A drug is considered active if it reduces the % avoidance responding to at least 50% without producing greater than 50% response failures. For active drugs, a dose-response curve is subsequently determined.

Gerbil Foot-Tapping

Male or female Mongolian gerbils (35-70 g) are anaesthetised by inhalation of an isoflurane/oxygen mixture to permit exposure of the jugular vein in order to permit administration of test or control compounds or vehicle in an injection volume of 5 ml/kg i.v. Alternatively, test compounds may be administered orally or by subcutaneous or intraperitoneal routes. The treated group may receive only a single dose of the test compound or may be treated daily (e.g., sid, bid, or tid) for at least 1 day, 3 days, 1 week, 2 weeks, 1 month, or 2 months, prior to the assay. A skin incision is then made in the midline of the scalp to expose the skull. An anxiogenic agent (e.g. pentagastrin) and/or a control agent (saline, DAO or DDO antagonist, D-Ser, D-Asp, etc.) is infused directly into the cerebral ventricles (e.g. 3 pmol in 5 μ l i.c.v., depending on test substance) by vertical insertion of a cuffed 27 gauge needle to a depth of 4.5 mm below bregma. The scalp incision is closed and the animal allowed to recover from anaesthesia in a clear perspex observation box (25 cm.times.20 cm.times.20cm). The duration and/or intensity of hind foot tapping is then recorded continuously for approximately 5 minutes. Alternatively, the ability of test compounds to inhibit foot tapping evoked by aversive stimulation, such as foot shock or single housing, may be studied using a similar method of quantification. Preferred antagonists of the present invention are able to inhibit induced foot-tapping in the gerbil.

Ferret Emesis

Individually housed male ferrets (1.0-2.5 kg) are dosed orally by gavage with test or control compounds or vehicle. Ten minutes later they are fed with approximately 100 g of tinned cat food. The treated group may receive only a single dose of the test compound or may be treated daily (e.g., sid, bid, or tid) for at least 1 day, 3 days, 1 week, 2 weeks, 1 month, or 2 months, prior to the experiment. At 60 minutes following oral dosing, cisplatin (10 mg/kg) is given i.v. via a jugular vein catheter inserted under a brief period of halothane anaesthesia. The catheter is then removed, the jugular vein ligated and the skin incision closed. The ferrets recover rapidly from the anaesthetic and are mobile within 10-20 minutes. The animals are observed continuously during recovery from the anaesthetic and for 4 hours following the cisplatin injection, after which time the animals are killed humanely. The numbers of retches and vomits occurring during the 4 hours after cisplatin administration are recorded by trained observers.

Separation-Induced Vocalisation

Male and female guinea-pigs pups are housed in family groups with their mothers and littermates throughout the study. Experiments are commenced after weaning when the pups are 2 weeks old. Before entering an experiment, the pups are screened to ensure that a vigorous vocalisation

response is reproducibly elicited following maternal separation. The pups are placed individually in an observation cage (55 cm.times.39 cm.times.19cm) in a room physically isolated from the home cage for 15 minutes and the duration of vocalisation during this baseline period is recorded. Only animals which vocalise for longer than 5 minutes are employed for drug challenge studies (approximately 50% of available pups may fail to reach this criterion). The treated group may receive only a single dose of the test compound or may be treated daily (e.g., sid, bid, or tid) for at least 1 day, 3 days, 1 week, 2 weeks, 1 month, or 2 months, prior to the experiment. On test days each pup receives an oral dose or an s.c. or i.p. injection of test compound or vehicle and is then immediately returned to the home cage with its mother and siblings for 30 to 60 minutes (or for up to 4 hours following an oral dose, dependent upon the oral pharmacokinetics of the test compound) before social isolation for 15 minutes as described above. The duration of vocalisation on drug treatment days is expressed as a percentage of the pre-treatment baseline value for each animal. The same subjects are retested once weekly for up to 6 weeks. Between 6 and 8 animals receive each test compound at each dose tested. Preferred antagonists of the present invention are effective in the attenuation of separation-induced vocalisations by guinea-pig pups as hereinafter defined.

Behavioral Activity Assessment Of Mice And Rats In The Omnitech Digiscan Animal Activity Monitors

The purpose of this test is to evaluate compounds for antipsychotic-like central nervous system (CNS) effects and a variety of other behavioral effects generally associated with CNS activity. This test has the capacity to determine drug effects on many aspects of locomotor activity in rodents, including horizontal activity (beam breaks), total distance traveled (in cm), number of movements, movement time (in sec), rest time (in sec), vertical activity (beam breaks), number of vertical movements, vertical time (in sec), stereotypy counts, number of stereotypic episodes, stereotypy time (in sec), margin and center time (in sec), clockwise and counterclockwise revolutions, and time (in sec) spent in each corner of the activity monitor. Generally, however, drug effects on behavior are assessed using total distance traveled (in cm) as the most accurate measure of locomotor activity.

Male CD-1 albino mice weighing 20 to 40g (Charles River Laboratories) or male Sprague-Dawley rats weighing 150 to 300g (Harlan Laboratories) are used for these studies. The treated group may receive only a single dose of the test compound before the experiment or may be treated daily (e.g., sid, bid, or tid) for at least 1 day, 3 days, 1 week, 2 weeks, 1 month, or 2 months, prior to the experiment.

The Omnitech Digiscan animal activity monitor consists of a 16".times.16".times.12" plexiglas cubicle enclosed inside 2 sets of 16 infrared photobeam sensors spaced 1 inch apart on all four sides of the bottom of the cubicle. An additional set of photobeam sensors are placed directly above the lower

photobeam sensors, which measure vertical activity. Interruption of any beam should generate a flash of the LED indicator located in the center of the monitor mainframe. A diagnostic test of each of the 24 monitors is generally performed prior to the start of an experiment, in which all the photobeams are checked for any interruption. Each activity monitor can be divided into four 8" square quadrants using a plexiglas insert that fits inside the plexiglas cubicle, of which 2 quadrants (front left and rear right) can be used for activity testing. Generally, this divided arrangement is utilized for mouse activity studies (2 mice per divided monitor) as opposed to rat studies (1 rat per undivided monitor). Up to 999 data samples can be taken for up to 999 minutes duration. Generally, 6 data samples of 10-minute duration each are collected for mice (1-hour test), or 6 samples of 5-minute duration for rats (30-minute test).

Once the animal is placed in the activity chamber, the chambers are individually activated to begin collecting data. Activity levels are generally monitored with the overhead lights turned off, as the dark-stimulation tends to produce less variation in the data. The following types of data (with brief definitions) are collected during each experiment:

Variable 1: Horizontal activity--total number of beam interruptions that occurred in the horizontal sensor.

Variable 2: Total distance (in cm) traveled--a more accurate indicator of ambulatory activity as it takes into account any diagonal movement.

Variable 3: Number of movements--number of discrete movements separated by at least 1 second.

Variable 4: Movement time (sec)--amount of time in ambulation.

Variable 5: Rest time (sec)--difference between sample time and time spent moving.

Variable 6: Vertical activity--total number of beam interruptions that occurred in the vertical sensor as the animal rears up.

Variable 7: Number of vertical movements--each time the animal rears up and interrupts the vertical sensor (separated by at least 1 second).

Variables 8, 9, 10, and 11: Time spent in corners (left and right front, left and right rear)--time spent by the animal in close proximity to two adjoining walls of the cage.

Variable 12: Vertical time (sec)--time spent interrupting the vertical beams during rearing.

Variable 13: Stereotypy counts--number of beam breaks that occur during a period of repeated interruption (stereotypy) of the same beam (or set of beams).

Variable 14: Number of stereotypy--number of times the monitor observes stereotypic behavior, separated by at least 1 second.

Variable 15: Stereotypy time (sec)--total amount of time that stereotypic behavior is exhibited.

Variable 16: Clockwise revolutions--number of times the animal circles with at least a 2" diameter (will not pick up tighter rotating movements).

Variable 17: Anticlockwise revolutions--number of times the animal circles with at least a 2" diameter.

Variable 18: Margin time (sec)--time spent by the animal in close proximity (within 1 cm) to the walls of the plexiglas cage.

Variable 19: Center time (sec)--time spent by the animal away from the walls of the cage.

Data can be expressed as either actual counts, time (in sec), centimeters traveled, or percent inhibition of activity relative to vehicle-treated control animals tested concurrently. Significant changes in activity (i.e., cm traveled), relative to controls, are determined by t-test or analysis of variance and Newman-Keul's multiple-range test. Stimulation of activity levels is indicated by negative values. The dose which could be expected to decrease activity levels by 50% (ED.sub.50) and the 95% confidence limits (CL) around that value are estimated by regression analysis using at least three data points which fall on the linear portion of the dose-effect curve.

Blockade Of Amphetamine-Stimulated Locomotion In Rats

The blockade of amphetamine-stimulated locomotion procedure is a modification of the Locomotor Activity Protocol in the Omnitech Digiscan Activity Monitors described above. The blockade of amphetamine-stimulated locomotion procedure uses the central nervous system stimulant d-amphetamine to assess antipsychotic activity of dopaminergic agents.

Male Sprague-Dawley rats (Harlan Labs) are used for these studies. The treated group may receive only a single dose of the test compound before the experiment or may be treated daily (e.g., sid, bid, or tid) for at least 1 day, 3 days, 1 week, 2 weeks, 1 month, or 2 months, prior to the experiment. For the IP studies, amphetamine is given 20 minutes prior to the drug, after which a 30 minute locomotor activity test is conducted. For the oral study, drug is dosed 30 minutes prior to the test, while amphetamine is given 15 minutes prior to the test, which allows time for oral absorption. Locomotor activity (centimeters travelled per 30 minute test) is measured in 16".times.16" open chambers. Amphetamine generally produces a 2- to 3-fold increase in locomotion over saline controls. Drug effects are reported as percent reversal of amphetamine-stimulated locomotion. Significant changes in amphetamine-stimulated locomotion, relative to amphetamine treated controls, are determined by t-test. The dose which would reverse amphetamine-stimulated locomotion by 50% (ED.sub.50) and the 95% confidence limits are estimated by regression analysis.

Protocol For The Prepulse Inhibition Of Acoustic Startle Model In Rats

Prepulse inhibition (PPI) of acoustic startle is a form of sensorimotor gating which occurs when a weak stimulus precedes a startling stimulus, resulting in diminution of the startle response amplitude. Schizophrenic patients exhibit reduced prepulse inhibition of acoustic startle compared to control subjects, consistent with a loss of sensorimotor gating. Thus, an animal model utilizing this phenomenon is quite useful in the study of known and potential antipsychotic agents. In rats, for example, PPI can be blocked with direct dopamine agonists (DA) such as apomorphine, or the indirect DA agonist amphetamine, and this effect can be antagonized with dopamine antagonists such as haloperidol.

Male Sprague-Dawley rats from Harlan Labs (180-280 g) are housed in groups of five rats per cage and maintained on a 12-hour light/dark cycle with free access to food pellets and water. The treated group may receive only a single dose of the test compound before the experiment or may be treated daily (e.g., sid, bid, or tid) for at least 1 day, 3 days, 1 week, 2 weeks, 1 month, or 2 months, prior to the experiment.

Startle chambers (SR-LAB, San Diego Instruments) consisting of a Plexiglas cylinder resting on a Plexiglas frame within a ventilated sound-attenuating enclosure are used. Acoustic stimuli are presented via a loudspeaker mounted above the rat. A piezoelectric device is mounted below the Plexiglas frame, which detects and transduces the motion occurring inside the cylinder during the 100 msec after the onset of the startling stimulus. The average responses during the 100 msec record window (100.times.1 msec readings) are recorded by microcomputer and interface assembly (San Diego Instruments). Each of the chambers are calibrated to one another to ensure consistent levels of loudspeaker performance over a wide range of decibel (dB) levels (67 to 125 dB). Sound levels are assessed with a dB meter (e.g., Radio Shack). Each stabilimeter (which houses the piezoelectric device) is adjusted to produce equal response sensitivity to a constantly vibrating calibrator.

Animals treated with the test compound may receive only a single dose of the test compound before the experiment or may be treated daily (e.g., sid, bid, or tid) for at least 1 day, 3 days, 1 week, 2 weeks, 1 month, or 2 months, prior to the experiment. Prior to the experiment, each animal is pretreated with saline, test, or control compounds (e.g., apomorphine, haloperidol, clozapine, etc.).

Each test session begins with a 5-minute test acclimation period of 70 dB white noise. The test session lasts a total of 30 minutes; several sequential tests are done to obtain an adequate number of rats per treatment group. The first and last trials are 120 dB pulse-alone trials presented 7 to 23 seconds apart, during which time the rats habituate rapidly to the noise bursts. These data are not included in the PPI calculation. The middle trials consists of 120 dB pulse-alone trials and trials of each of the following five trial types in pseudorandom order: (1) no stim, (2) 72 dB prepulse 100 msec prior to 120 dB startle, (3) 74 dB prepulse 100 msec prior to 120 dB startle, (4) 78 dB prepulse 100 msec prior to 120

dB startle, and (5) 86 dB prepulse 100 msec prior to 120 dB startle. The prepulses (2, 4, 8, and 16 dB over 70 dB background noise) are of 20-msec duration, while the startle stimuli were 40-msec duration. When the prepulse is paired with the 120 dB pulse, no obvious acoustic difference can be detected by the human ear as compared to the 120 dB pulse alone. Prepulse inhibition of the acoustic startle reflex is expressed as the percent inhibition of the 120 dB startle amplitude produced when a 2 to 16 dB (over background) prepulse precedes the startling stimulus.

Inhibition Of Apomorphine-Induced Climbing Behaviour

In Animal Pharmacology Studies, the antipsychotic activity of the test compounds can be tested by the inhibition of apomorphine-induced climbing behaviour (P. Protais et al: "Psychopharmacology", 50, 1-6, 1976). Male Swiss mice weighing 22-24g are used. Animals treated with the test compound may receive only a single dose of the test compound before the experiment or may be treated daily (e.g., sid, bid, or tid) for at least 1 day, 3 days, 1 week, 2 weeks, 1 month, or 2 months, prior to the experiment. Animals are administered orally with test drug or 0.25% agar at time 0. After 60 minutes, apomorphine is subcutaneously injected at a dose of 1 mg/kg, and after further 70 minutes the animal's behaviour is assessed. Two additional assessments are performed at 10 min intervals. For assessment, each animal is placed on the bottom of a small upright box (11.times.7.5.times.4.5 cm). The walls of the box are made of translucent methacrylate except one of the lateral surfaces (7.5 cm wide) which is a 3mm wire mesh. The position of the animal is scored for 2 minutes according to the following criteria: 0=four paws on the floor; 1=three paws on the floor; 2=two paws on the floor; 3=one paw on the floor; and 4=four paws holding the wire mesh. If an animal keeps several positions within the 2 min observation, the seconds elapsed in each position will be recorded. Finally, mean scoring is calculated. Under these experimental conditions, the effective dose 50% (ED.sub.50) values are calculated.

Inhibition Of DOI-Induced Head Twitches And Scratches

The antipsychotic activity of the test compounds can also be tested by the inhibition test of 1-(2,5-dimethoxy-4-iodophenyl)-2-aminopropane (DOI)-induced head twitches and scratches (M. Oka et al: "J. Pharm. Exp. Ther.", 264(1), 158-165, 1993). Male N.M.R.I. mice weighing 22-26g are used. After the animals are weighed, they are individually placed in transparent cages two hours prior to experiment. Animals treated with the test compound may receive only a single dose of the test compound before the experiment or may be treated daily (e.g., sid, bid, or tid) for at least 1 day, 3 days, 1 week, 2 weeks, 1 month, or 2 months, prior to the experiment. Test compound is given p.o. at time 0. At time 60 min DOI at the dose of 3 mg/kg i.p. dissolved in saline is administered. The number of head twitches and scratches were assessed as well as the presence or absence of escape attempts. The effective dose 50% (ED.sub.50) values obtained under the above experimental conditions are calculated.

Human Clinical Trials

The activity of a DAO or DDO antagonist for treating or alleviating schizophrenia, bipolar disorder, or another CNS disorder of the present invention can be demonstrated by human clinical trials. For example, a study can be designed as a double-blind, parallel, placebo-controlled multicenter trial. Subjects are randomized into four groups, placebo and three increasing dosages of test compound, e.g., 25, 50, and 75 mg. The dosages are administered in a manner disclosed herein or practiced by the skilled practitioner, e.g., orally with food. Subjects are observed at four visits to provide baseline measurements. Further visits, e.g., 5-33, are served as the treatment phase for the study.

During the visits, subjects are observed for signs of psychotic behavior or bipolar behavior such as agitation, mood swings, tremor, delirium, social withdrawal, and concentration abilities. Treatment groups are compared with respect to the number and percent of subjects who ever had the symptom during the double-blind portion of the study (visits 5 through 33), at a severity that was worse than during the baseline visits (1 through 4).

DAO, DDO, and Biallelic Markers Thereof In Methods Of Genetic Diagnostics

The DAO and DDO genomic and cDNA sequences, and the biallelic markers of the present invention can also be used to develop diagnostics tests capable of identifying individuals who express a detectable trait as the result of a specific genotype or individuals whose genotype places them at risk of developing a detectable trait at a subsequent time. The trait analyzed using the present diagnostics may be used to diagnose any detectable trait, including predisposition to schizophrenia or bipolar disorder, age of onset of detectable symptoms, a beneficial response to or side effects related to treatment against schizophrenia or bipolar disorder. Such a diagnosis can be useful in the monitoring, prognosis and/or prophylactic or curative therapy for schizophrenia or bipolar disorder.

The diagnostic techniques of the present invention may employ a variety of methodologies to determine whether a test subject has a genotype associated with an increased risk of developing a detectable trait or whether the individual suffers from a detectable trait as a result of a particular mutation, including methods which enable the analysis of individual chromosomes for haplotyping, such as family studies, single sperm DNA analysis or somatic hybrids.

The diagnostic techniques concern the detection of specific alleles present within the human DAO or DDO genes, preferably within a DAO or DDO exon or coding sequence. More particularly, the invention concerns the detection of a nucleic acid comprising at least one of the nucleotide sequences of SEQ ID Nos. 1 to 6 or a fragment thereof or a complementary sequence thereto including the polymorphic base.

These methods involve obtaining a nucleic acid sample from the individual and, determining, whether the nucleic acid sample contains at least one allele or at least one biallelic marker haplotype, indicative of a risk of developing the trait or indicative that the individual expresses the trait as a result of possessing a particular the human DAO or DDO-related polymorphism or mutation (trait-causing allele).

Preferably, in such diagnostic methods, a nucleic acid sample is obtained from the individual and this sample is genotyped using methods well known in the art, or as described for example in PCT/IB00/00435 incorporated herein by reference. The diagnostics may be based on a single biallelic marker or a on group of biallelic markers.

In each of these methods, a nucleic acid sample is obtained from the test subject and the biallelic marker pattern of one or more of a biallelic marker of the invention is determined.

In one embodiment, a PCR amplification is conducted on the nucleic acid sample to amplify regions in which polymorphisms associated with a detectable phenotype have been identified. The amplification products are sequenced to determine whether the individual possesses one or more human DAO or DDO polymorphisms associated with a detectable phenotype. Alternatively, the nucleic acid sample is subjected to microsequencing reactions to determine whether the individual possesses one or more DAO or DDO-related polymorphisms associated with a detectable phenotype resulting from a mutation or a polymorphism in the DAO or DDO genomic sequence. In another embodiment, the nucleic acid sample is contacted with one or more allele specific oligonucleotide probes which, specifically hybridize to one or more human chromosome DAO or DDO-related alleles associated with a detectable phenotype. In another embodiment, the nucleic acid sample is contacted with a second oligonucleotide capable of producing an amplification product when used with the allele specific oligonucleotide in an amplification reaction. The presence of an amplification product in the amplification reaction indicates that the individual possesses one or more DAO or DDO-related alleles associated with a detectable phenotype. In a preferred embodiment, the detectable trait is schizophrenia or bipolar disorder. Diagnostic kits comprise any of the polynucleotides of the present invention.

These diagnostic methods are extremely valuable as they can, in certain circumstances, be used to initiate preventive treatments or to allow an individual carrying a significant haplotype to foresee warning signs such as minor symptoms.

Diagnostics, which analyze and predict response to a drug or side effects to a drug, may be used to determine whether an individual should be treated with a particular drug. For example, if the diagnostic indicates a likelihood that an individual will respond positively to treatment with a particular drug, the drug may be administered to the individual. Conversely, if the diagnostic indicates that an individual is likely to respond negatively to treatment with a particular drug, an alternative course of

treatment may be prescribed. A negative response may be defined as either the absence of an efficacious response or the presence of toxic side effects.

Clinical drug trials represent another application for the markers of the present invention. One or more markers indicative of response to an agent acting against schizophrenia or to side effects to an agent acting against schizophrenia may be identified using the methods described above. Thereafter, potential participants in clinical trials of such an agent may be screened to identify those individuals most likely to respond favorably to the drug and exclude those likely to experience side effects. In that way, the effectiveness of drug treatment may be measured in individuals who respond positively to the drug, without lowering the measurement as a result of the inclusion of individuals who are unlikely to respond positively in the study and without risking undesirable safety problems.

DAO and DDO in the Prevention and Treatment of Disease

In large part because of the risk of suicide, the detection of susceptibility to schizophrenia, bipolar disorder as well as other psychiatric disease in individuals is very important. Consequently, the invention concerns a method for the treatment of schizophrenia or bipolar disorder, or a related disorder comprising the following steps:

- selecting an individual whose DNA comprises alleles of a DAO or DDO-related biallelic marker or of a group of DAO or DDO-related markers associated with schizophrenia or bipolar disorder;
- following up said individual for the appearance (and optionally the development) of the symptoms related to schizophrenia or bipolar disorder; and
- administering a treatment acting against schizophrenia or bipolar disorder or against symptoms thereof to said individual at an appropriate stage of the disease.

Another embodiment of the present invention comprises a method for the treatment of schizophrenia or bipolar disorder comprising the following steps:

- selecting an individual whose DNA comprises alleles of a DAO or DDO-related biallelic marker or of a group of DAO or DDO-related markers associated with schizophrenia or bipolar disorder;
- administering a preventive treatment of schizophrenia or bipolar disorder to said individual.

In a further embodiment, the present invention concerns a method for the treatment of schizophrenia or bipolar disorder comprising the following steps:

- selecting an individual whose DNA comprises alleles of a DAO or DDO-related biallelic marker or of a group of DAO or DDO-related markers associated with schizophrenia or bipolar disorder;
- administering a preventive treatment of schizophrenia or bipolar disorder to said individual;
- following up said individual for the appearance and the development of schizophrenia or bipolar disorder symptoms; and optionally
- administering a treatment acting against schizophrenia or bipolar disorder or against symptoms thereof to said individual at the appropriate stage of the disease.

For use in the determination of the course of treatment of an individual suffering from disease, the present invention also concerns a method for the treatment of schizophrenia or bipolar disorder comprising the following steps:

- selecting an individual suffering from schizophrenia or bipolar disorder whose DNA comprises alleles of a DAO or DDO-related biallelic marker or of a group of DAO or DDO-related markers, associated with the gravity of schizophrenia or bipolar disorder or of the symptoms thereof; and
- administering a treatment acting against schizophrenia or bipolar disorder or symptoms thereof to said individual.

The invention also concerns a method for the treatment of schizophrenia or bipolar disorder in a selected population of individuals. The method comprises:

- selecting an individual suffering from schizophrenia or bipolar disorder and whose DNA comprises alleles of a DAO or DDO-related biallelic marker or of a group of DAO or DDO-related markers associated with a positive response to treatment with an effective amount of a medicament acting against schizophrenia or bipolar disorder or symptoms thereof,
- and/or whose DNA does not comprise alleles of a DAO or DDO-related biallelic marker or of a group of DAO or DDO-related markers associated with a negative response to treatment with said medicament; and
- administering at suitable intervals an effective amount of said medicament to said selected individual.

In the context of the present invention, a "positive response" to a medicament can be defined as comprising a reduction of the symptoms related to the disease. In the context of the present invention, a "negative response" to a medicament can be defined as comprising either a lack of positive response to the medicament which does not lead to a symptom reduction or which leads to a side-effect observed following administration of the medicament.

The invention also relates to a method of determining whether a subject is likely to respond positively to treatment with a medicament. The method comprises identifying a first population of individuals who respond positively to said medicament and a second population of individuals who respond negatively to said medicament. One or more biallelic markers is identified in the first population which is associated with a positive response to said medicament or one or more biallelic markers is identified in the second population which is associated with a negative response to said medicament. The biallelic markers may be identified using the techniques described herein.

A DNA sample is then obtained from the subject to be tested. The DNA sample is analyzed to determine whether it comprises alleles of one or more biallelic markers associated with a positive response to treatment with the medicament and/or alleles of one or more biallelic markers associated with a negative response to treatment with the medicament.

In some embodiments, the medicament may be administered to the subject in a clinical trial if the DNA sample contains alleles of one or more biallelic markers associated with a positive response to treatment with the medicament and/or if the DNA sample lacks alleles of one or more biallelic markers associated with a negative response to treatment with the medicament. In preferred embodiments, the medicament is a drug acting against schizophrenia or bipolar disorder.

Using the method of the present invention, the evaluation of drug efficacy may be conducted in a population of individuals likely to respond favorably to the medicament.

Another aspect of the invention is a method of using a medicament comprising obtaining a DNA sample from a subject, determining whether the DNA sample contains alleles of one or more biallelic markers associated with a positive response to the medicament and/or whether the DNA sample contains alleles of one or more biallelic markers associated with a negative response to the medicament, and administering the medicament to the subject if the DNA sample contains alleles of one or more biallelic markers associated with a positive response to the medicament and/or if the DNA sample lacks alleles of one or more biallelic markers associated with a negative response to the medicament.

The invention also concerns a method for the clinical testing of a medicament, preferably a medicament acting against schizophrenia or or bipolar disorder or symptoms thereof. The method comprises the following steps:

- administering a medicament, preferably a medicament susceptible of acting against schizophrenia or or bipolar disorder or symptoms thereof to a heterogeneous population of individuals,
- identifying a first population of individuals who respond positively to said medicament and a second population of individuals who respond negatively to said medicament,
- identifying biallelic markers in said first population which are associated with a positive response to said medicament,
- selecting individuals whose DNA comprises biallelic markers associated with a positive response to said medicament, and
- administering said medicament to said individuals.

In any of the methods for the prevention, diagnosis and treatment of schizophrenia and bipolar disorder, including methods of using a medicament, clinical testing of a medicament, determining whether a subject is likely to respond positively to treatment with a medicament.

Such methods are deemed to be extremely useful to increase the benefit/risk ratio resulting from the administration of medicaments which may cause undesirable side effects and/or be inefficacious to a portion of the patient population to which it is normally administered.

Once an individual has been diagnosed as suffering from schizophrenia or bipolar disorder, selection tests are carried out to determine whether the DNA of this individual comprises alleles of a

biallelic marker or of a group of biallelic markers associated with a positive response to treatment or with a negative response to treatment which may include either side effects or unresponsiveness.

The selection of the patient to be treated using the method of the present invention can be carried out through the detection methods described above. The individuals which are to be selected are preferably those whose DNA does not comprise alleles of a biallelic marker or of a group of biallelic markers associated with a negative response to treatment. The knowledge of an individual's genetic predisposition to unresponsiveness or side effects to particular medicaments allows the clinician to direct treatment toward appropriate drugs against schizophrenia or bipolar disorder or symptoms thereof.

Once the patient's genetic predispositions have been determined, the clinician can select appropriate treatment for which negative response, particularly side effects, has not been reported or has been reported only marginally for the patient.

The biallelic markers of the invention have demonstrated an association with schizophrenia and bipolar disorders. However, the present invention also comprises any of the prevention, diagnostic, prognosis and treatment methods described herein using the biallelic markers of the invention in methods of preventing, diagnosing, managing and treating related disorders, particularly related CNS disorders.

EXAMPLES

Construction of the plasmids for the protein expression in bacteria and in yeast

Expression of the recombinant proteins without tag was carried out with pET11a vector (Stratagen). The coding sequences with appropriated sites (NdeI in 5' and HindIII in 3') were obtained by PCR (TaqPlusPrecision System, Stratagen) with the primers corresponding to the ORF limits.

The generated PCR products were purified (Qiaquick PCR, Qiagen), digested with NdeI and HindIII, gel purified (Microspin, PolyLabo), and ligated into a vector open with the same enzymes. The constructs were transfected into the DH10B bacterial host (Gibco BRL), plasmid DNAs were extracted and sequenced to select proper coding sequences.

The plasmids for expression of human DAAO and g34872 in yeast were constructed with pESC-LEU shuttle vector (Stratagen).

Expression and purification of the recombinant g34872 protein without tag

The plasmids were then transfected into the BL21(DE3) CodonPlus RIL bacterial host (Stratagen), the bacteria were allowed to grow in 0.8 liter of LB media until an A600 of 0.7 was achieved. Expression of fusion proteins was induced by the addition of 1 mM isopropyl-1-thio--D-galactopyranoside and further cultured for 3 h. Bacterial pellets were prepared and immediately frozen (-80°C), then thawed in the water bath at 30°C ; AEBSF was added at 2mM. Bacterial cells were suspended in 25 ml of BugBuster

extraction agent (Novagen) supplemented with protease inhibitor mixture (SetIII, Calbiochem) and with 10 mM EGTA. The suspension was incubated 30 min at room temperature, then benzonase was added (Novagen) and incubation was continued for 15 min. The lysate was centrifuged at $10,000 \times g$ at 4°C for 30 min. Bacterial proteins were fractionated from the supernatant by salt precipitation. The protein pellet corresponding to 35 - 55% of ammonium sulfate saturation was dissolved in 2 ml of 50 mM TrisHCl buffer pH8/50 mM NaCl with 10 mM DTT, the solution was clarified by centrifugation and applied on Ultragel AcA44 (Pharmacia) column (1.6x 65cm) equilibrated with 20 mM TrisCl buffer pH8/50 mM NaCl buffer. Eluted proteins were analyzed by electrophoresis, the fractions containing MN2R protein were pooled and concentrated by ultrafiltration (10K cut, Biomax-15, Sigma). The proteins were then applied on DEAE-Macroprep (Bio-Rad) column (1x2 cm) equilibrated with 20 mM TrisCl pH8 and eluted with linear salt gradient (from 0 to 1M NaCl, 20 column volumes). The fractions containing MN2R protein were pooled, concentrated by ultrafiltration (10K cut) and applied on Superdex 75 (Pharmacia) column (1x27 cm) equilibrated with 20 mM TrisCl pH8 buffer. The fractions from single major pique were pooled, concentrated to 5 mg /ml and saved at 4°C . Yield of the purified electrophoretically homogenous protein was typically 5 mg per liter of bacterial culture. Denaturing electrophoresis of the proteins in 10% NuPage custom gels was done according the manufacturer recommendations (NuPage by Novagen), MES/SDS running buffer was used. Molecular weight markers See-Blue were from Invitrogen. Proteins were visualized after staining with Coomassie Brilliant Blue G colloidal solution (Sigma).

Purification of the natural DAAO from pig kidney

The crude preparation of pig kidney DAAO was purchased from Sigma. Proteins were dissolved in 50 mM TrisCl pH8 (1g in 10 ml); the solution was clarified by centrifugation and applied on Sephadex G-50 medium column (2.6 x 40 cm) equilibrated with 10 mM TrisCl pH8/ 100 mM NaCl. The desalted proteins were then concentrated 3 fold by ultrafiltration (30K cut, Biomax-15, Sigma) equilibrated with 10 mM TrisCl pH8/100 mM NaCl/ 1 mM DTT/ 10 mM ATP and applied on DEAE-Sepharose column (1.6 x 7cm) in the same buffer without ATP. The column was washed with two column volumes of 10 mM TrisCl pH8/100 mM NaCl, followed by one volume of 10 mM TrisCl pH8/125 mM NaCl and then proteins were eluted with 10 mM TrisCl pH8/150 mM NaCl buffer. The fractions were assayed for DAAO enzyme activity, pooled and concentrated by ultrafiltration. The proteins were then applied on Ultragel AcA44 column (1.6 x 65 cm) equilibrated with 10 mM TrisCl pH8/100 mM NaCl/ 1 mM DTT and eluted with the same buffer. The fractions containing electrophoretically pure DAAO were concentrated by ultrafiltration and kept at 4°C .

Expression and purification of the recombinant human DAAO

The plasmid was transfected into the BL21(DE3) CodonPlus RIL bacterial host (Stratagen), the bacteria were allowed to grow in 3 liters of LB media until an A600 of 0.7 was achieved. Expression of fusion proteins was induced by the addition of 1 mM isopropyl-1-thio-beta-D-galactopyranoside and further cultured for 5 h. Bacterial pellets were extracted with BugBuster extraction agent (Novagen) in presence of 2mM AEBSF, benzonase was routinely used. The lysate was adjusted to pH8 with 50 mM TrisCl and centrifuged at $10,000 \times g$ at 4 °C for 20 min. The proteins were precipitated from the extract with ammonium sulfate (from 30 to 50% saturation), collected by centrifugation at $10,000 \times g$ at 4 °C for 60 min and dissolved in 50 mM TrisCl pH8 (10 ml); the solution was clarified by centrifugation and applied on Sephadex G-75 column (2.6 x 40 cm) equilibrated with 10 mM TrisCl pH8/ 100 mM NaCl. The following steps of the purification were almost identical to those described for pig kidney DAAO. The only exception was the elution of the protein from DEAE-resin : it was done with 10 mM TrisCl pH8/300 mM NaCl buffer. The yield of the purified electrophoretically homogenous DAAO protein was 0.7 mg per liter of bacterial culture.

Expression and extraction of g34872 and human DAAO proteins in yeast *S. cerevisiae*

Yeast *S.cerevisiae* YPH499 and FY1679-18B strains were grown on YPD rich medium. The plasmids were transfected in yeast cells by standard lithium acetate method ; the transformants were selected on YNG synthetic medium, grown at 30°C in 1 liter of synthetic medium lacking leucine with 2% raffinose as a carbon source up to culture density 1 u A600/ ml . The cells were collected by centrifugation at 20°C, the medium was replaced by YNGal (with 2% D-galactose) and the incubation were continued for 20 h. The cells were pelleted, washed with ice cold water, resuspended in 20 mM TrisCl buffer pH8/ 2 mM AEBSF and vortexed 8 times for 1 min with a glass beads (Sigma) to extract the proteins. The lysate was centrifuged at $10,000 \times g$ at 4 °C for 30 min, the supernatant (S1) was collected and kept at 4°C. The pellet was resuspended in 20 mM TrisCl buffer pH8/ 2 mg/ml saponine/ 0.3% sarkosyl and vortexed 3 times for 1 min. The pellet extract was clarified by centrifugation (S2) and immediately frozen at -20°C. The protein concentration was detected by Bradford reagent (Bio-Rad), the expression was confirmed by Western blot procedure with rabbit anti-g34872-his6 serum (dilution 1/5000) and by DAAO enzyme activity detection with D-serine a substrate.

DAAO enzyme activity detection

The assay mixture was typically composed of D-serine (Aldrich) 200 mM, FAD (Sigma) at 0.1 mM, sodium phosphate buffer pH8 at 75 mM, HR-peroxidase (Sigma) 1U/ml. The mixture was air-saturated just before use. o-Dianizidine (Sigma) was added in the mixture. In the typical assay 5 µl of the enzyme (DAAO and mixes) was added to 25 µl of the assay mixture, the incubation was stopped with 50 µl of

20% H₂SO₄. The activity was observed as absorbance of the peroxidase-oxidized o-Dianizidine at 540nm. The reactions containing high protein concentrations were centrifuged 15 min at 14000rpm before absorbance measuring.

The control experiences were done to establish that g34872 protein do not influence peroxidase enzyme activity. The assay of peroxidase was done in the conditions identical to those for DAAO assay ; hydrogen peroxide (Gibco) was used as a substrate, and no effect of g34872 on HRP activity was confirmed.

Example 1:

Yeast cells were transformed with the plasmids constructed in pESC-Leu expression vector. One plasmid can express hDAAO, second one - C-terminal tagged g34872, third - is the vector without insertion (control). After the induction of the expression (2% galactose in the medium) these cells were incubated for 24h, than the extracts were maid and combined as follow: different volumes of DAAO extract were mixed with either g34872-cHis6 or with the vector extracts. The same volumes of DAAO extracts were also mixed with BSA (external control). After 30min of the pre-incubation the combined extracts were used for DAAO activity measuring. All yeast extracts and BSA solution had the same total protein concentration. DAAO activity was determined with D-serine at 37°C. See Figure 1 for demonstration that g34872 activates DAO.

Example 2:

Purified recombinant human DAAO was added in the E.coli extract containing expressed g34872 and in BSA solution. The concentrations of DAAO were 0.5 and 0.3 mg/ml. Total protein concentrations of the bacterial extract and of BSA were 12mg/ml. After 30 min pre-incubation the mixes were used for DAAO activity measuring at 37°C. D-serine was used as a substrate. (figure 2)

Example 3:

In vitro activation of purified DAAO in the presence of g34872 protein : the effect of activation depends on g34872 concentration.

Purified DAAO and g34872 were mixed and incubated 50min before activity essay, T°amb. Total protein concentration was the same in all the mixes. D-Serine was used as the substrate for DAAO, the pH of the reaction was 8,0. Proteins used: purified porcine DAAO, concentrations in the mixes were always 50 ng/μl purified recombinant g34872 concentrations in the mixes were from 0 to 450 ng/μl

bovine serum albumin (BSA) concentrations in the mixes were from 0 to 450 ng/μl . The range of the concentrations of g34872 protein can be considered as physiological as corresponding to the data found for lumenal Golgi proteins. See figure 3.

Example 4:**Estimation of the limits of g34872 concentrations necessary for DAAO activation in vitro**

Pig kidney DAAO was mixed with g34872 in PBS and incubated 50 min at 20°C. DAAO concentration was 50 ng/μl in all the mixes. The enzymatic activity of DAAO was measured at 20°C with 200 mM D-serine, pH was 8.0. Pig kidney DAAO was mixed with g34872 in PBS and incubated 1h at 20°C. DAAO concentration was 50 ng/μl in all the mixes. The enzymatic activity of DAAO was measured at 20°C with 200 mM D-serine, pH was 8.0. See figure 4.

Example 5:**DAAO Kinetics in the presence of g34872 protein : g34872 is an allosteric activator of DAAO**

Pig kidney DAAO was mixed with g34872 in PBS and incubated 30 min at 20°C. DAAO concentration was 200 ng/μl and g34872 concentration was 2 μg/μl in the protein mixture. The control mixture (without g34872) was composed of 200 ng/μl DAAO and 2 μg/μl BSA. The enzymatic activity of DAAO was measured at 20°C with D-serine, the substrate concentration used were from 0 to 100 mM, other components of the mixture and pH were standard.

V_{max} observed for g34872&DAAO mix corresponds to K_m= 4 mM, V_{max} observed for DAAO&BSA mix corresponds to K_m= 4mM. This result shows no change in DAAO affinity for its substrate (D-serine) and suggests that g34872 interacts with DAAO in the site other than the active site of the enzyme. See figure 5.

Example 6: Biallelic Markers of the invention

Validated polymorphisms (occurring at a frequency of >5% in the general population) have been discovered in the DAO gene (SEQ ID NO:1). These polymorphisms, also referred to as Biallelic markers, are represented by SEQ ID NOs:23-26 and by numbers 24/1443-126, 24/1457-52, 27/93-181, and 24/1461-256, respectively, wherein the polymorphic base is located at position 24. Polynucleotides comprising amplicons and microsequencing primers for detecting each DAO biallelic marker of the invention are described in SEQ ID NO:1. As shown in Figure 6, Marker 27-93/181 (SEQ ID NO:25) and 24-1461/256 (SEQ ID NO:26) have been determined to be significantly associated with schizophrenia $p=0.0066$ and 0.0111 , respectively. Markers of the invention can be further used to determine if an individual is at risk for schizophrenia, as demonstrated in Figure 6, as well as other related CNS disorders, preferably depression and bipolar disorder.

Example 7: Syntheses of compounds or compositions of the invention.Compound Preparation:

The DAO and DDO antagonist compositions and compounds of the invention can be prepared by a variety of methods which are well known to one of skill in the art. General schemes include but are not limited to those described *infra*.

Preparation of Compounds of Formula I, Ia, Ib

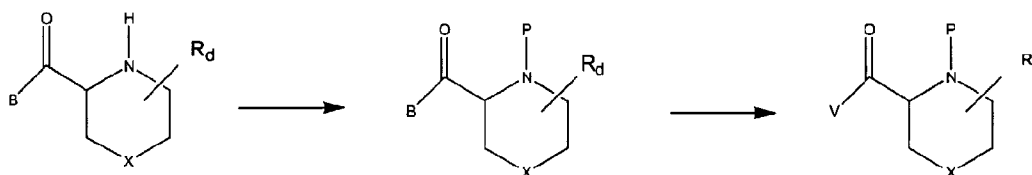
A vast number of the compounds of Formulae I, Ia, and Ib are commercially available or readily synthesized via common methods known to the skilled artisan from commercially available compounds. Specifically, substituents can be introduced into aromatic rings such as phenyl, naphthyl or substituted naphthyl or phenyl by way of electrophilic substitution reactions such as Friedel Crafts alkylations, acylations, and nitration in concentrated nitric acid. Transforming aromatic groups into organometallic salts such as Grignard reagents or introduction of substituents via aryl diazonium compounds are also common methods of aromatic ring modification. Example of these manipulations and other relevant transformations are discussed in standard texts such as March, Advanced Organic Chemistry(Wiley), Carey and Sundberg, Advanced Organic Chemistry(Vol 2.) and Keeting, Heterocyclic Chemistry (all 17 volumes).

Preparation of Compounds of Formula II.

Compounds of Formula II are commercially available or readily synthesized by the skilled artisan utilizing known synthetic techniques.

Preparation of Compounds of Formula IV Substituted at Position Z

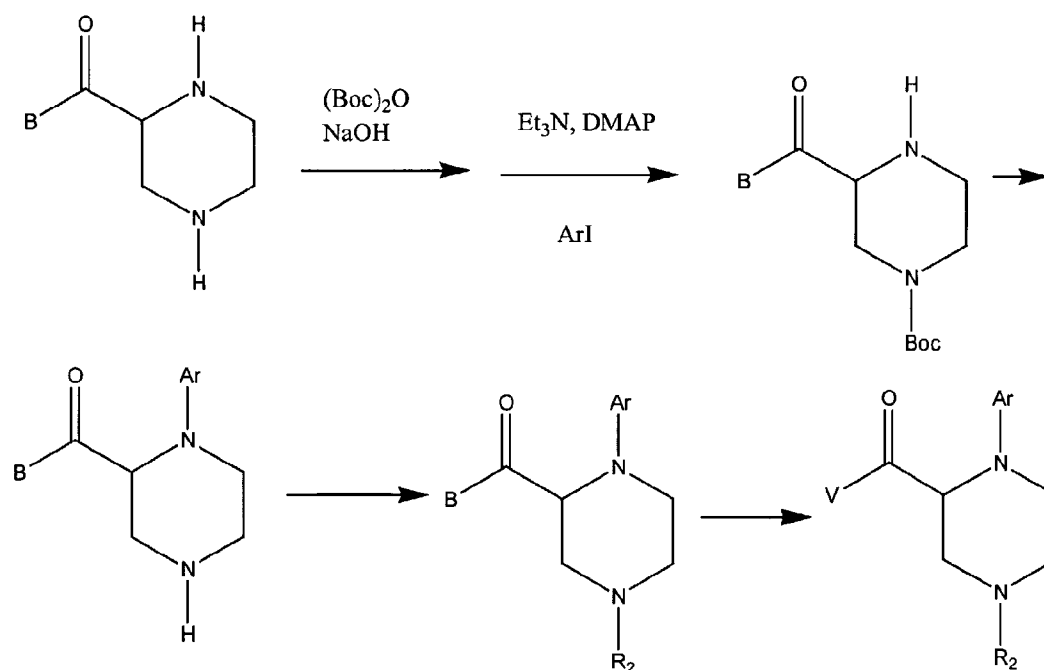
For the manipulation of R₁ it is understood that the skilled artisan may choose to prepare R₁ before, after or concurrent with the preparation of the heterocyclic ring. For compounds in which A is nitrogen, a preferred method of making the compounds is.



SCHEME 4a

Where R_d is a derivatizable group or can be manipulated or substituted, such compounds are known and can be prepared by known methods. (P) is a protecting group such as aryl and (B) is a suitable blocking group. For clarity, groups at position (Y) of formula IV are not shown.

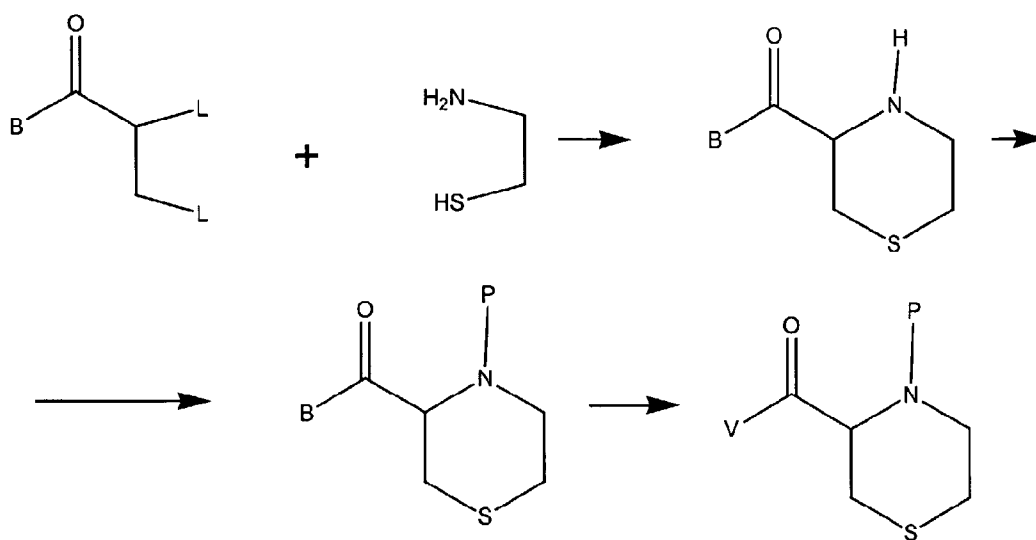
For preparation and elaboration of the heterocyclic ring it is understood that the skilled artisan may choose to prepare R_1 before, after or concurrent with the preparation of the heterocyclic ring. For clarity, the substituents at Z and Y are not shown. For compounds in which X is nitrogen, a preferred method of manipulating R_2 is shown. In the schemes below, L is any acceptable leaving group, and B is a blocking group as above. Boc is an example of a preferred, and art recognized blocking group. The skilled artisan will recognize that the choice of blocking group is within the skill of the artisan working in organic chemistry.



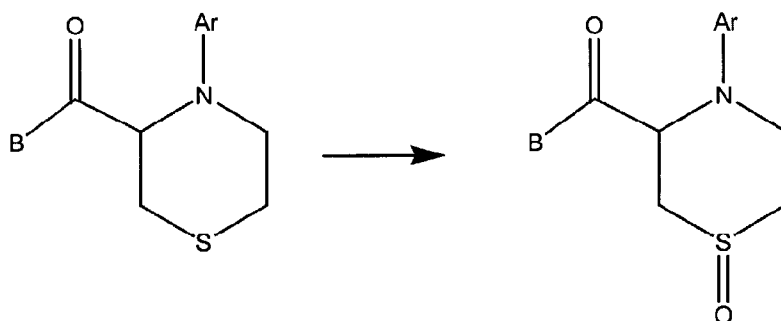
SCHEME 4b

For compounds containing a sulfur in the heterocyclic ring the preferred methods of ring formation are shown. For the preparation and elaboration of the heterocyclic ring it is understood that the skilled artisan may choose to prepare R_1 before, after or concurrent with the preparation of the heterocyclic ring. For clarity groups at position Z and Y are not shown.

106

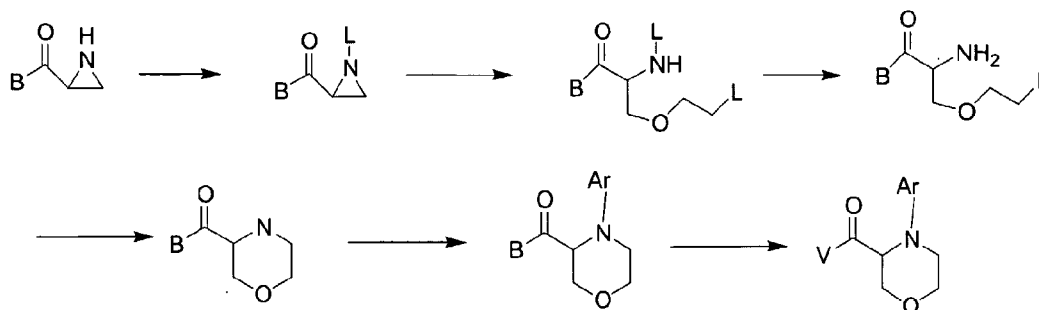
**SCHEME 4c**

Where X is sulfur, further elaboration of the heterocyclic ring can be accomplished after the ring has been formed. For example, oxidation of the ring sulfur atom using known methods can provide the corresponding sulfoxides and sulfones as shown.

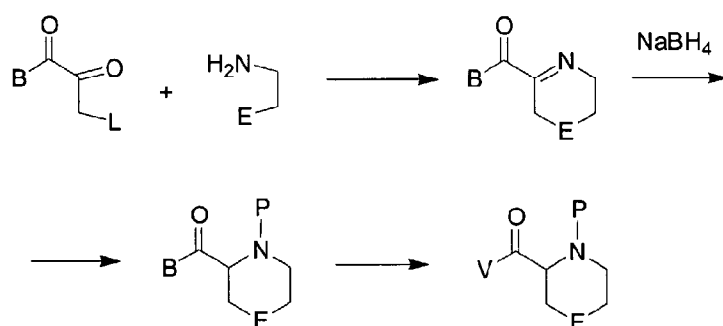
**SCHEME 4d**

For compounds containing an oxygen in the heterocyclic ring, the preferred methods of ring formation are shown. A bifunctional moiety, for example a halo hydroxy species is reacted with an aziridine below. The halo moiety serves as a leaving group, useful in ring closure reactions. Upon formation of the ring, elaboration of the invention proceeds as described above.

107

**SCHEME 4e**

Another acceptable strategy for making the heterocyclic ring of the invention, having E as sulfur, nitrogen or oxygen includes the following scheme. This is a preferred route by which to also prepare compounds in which A is nitrogen and A-B is unsaturated.

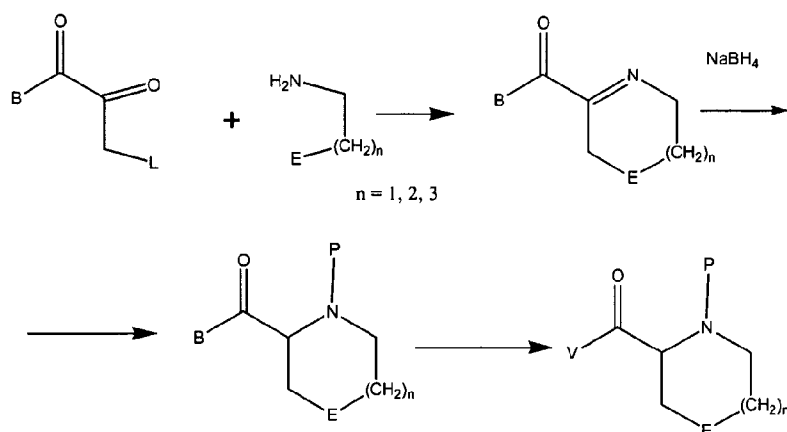
**SCHEME 4f**

Preparation of Other Preferred Compounds of Formula IV and Formula III

Of course the skilled artisan will recognize that scheme I can be applied to a substituent at (Y) for all of the described groups. Where Z is a ketal or thioether the compounds of the invention may be prepared from a compound having a carbonyl in the ring. Such compounds are prepared by known method, and many of such compounds are known or commercially available. Thus the skilled artisan will appreciate that a hydroxy, amino, imino, alkoxy or other group may be manipulated into a carbonyl compound.

The skilled artisan will also recognize that the above synthetic routes for compounds of formula IV can be applied to compounds of formula III in which the ring size is seven and eight members in size.

Symbols B, L, P and V are defined as described above. The following example is exemplary but not limiting.

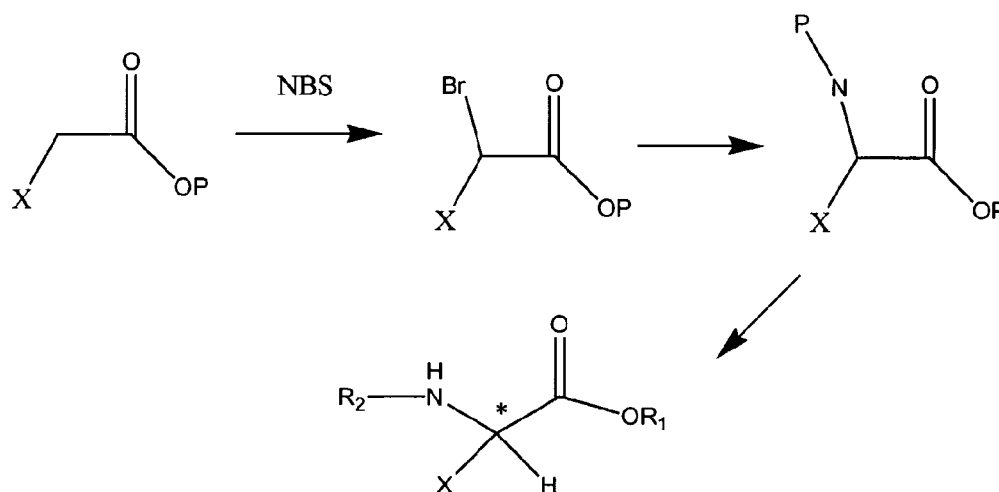


SCHEME 4g

Synthesis of Compounds of Formula Va

Compounds of Formula Va can be synthesized by a variety of methods. The best known route, which can be used for different alpha amino acids is the Strecker synthesis route. In that method a suitable aldehyde is treated with ammonia and HCN, so that an alpha-amino nitrile is formed, which is subsequently subject to a hydrolysis reaction.

Another acceptable strategy for the synthesis of compounds of formula Va is through the following scheme:

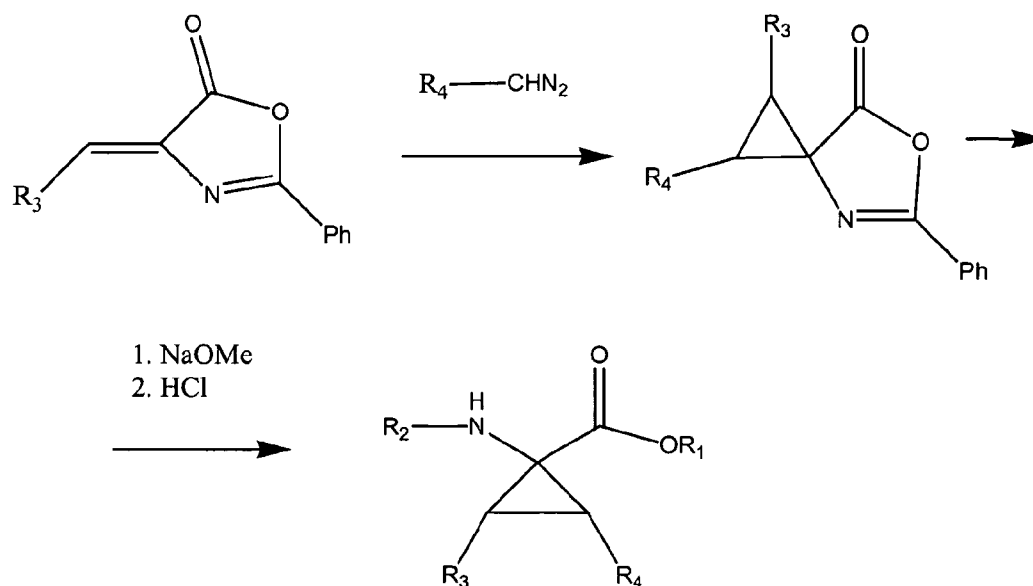


SCHEME 5a

in which P is a protecting group such as tertiary butyl which may be be the same as R_1 . X is a group as described above. The protected compound is brominated using a halogenating reagent such as PBr_3 , NBS or CBr_4 followed by halogen displacement using NH_3 or protected amine derivatives such as potassium phthalimide. Incorporation of R_1 and R_2 can be readily accomplished by the skilled artisan.

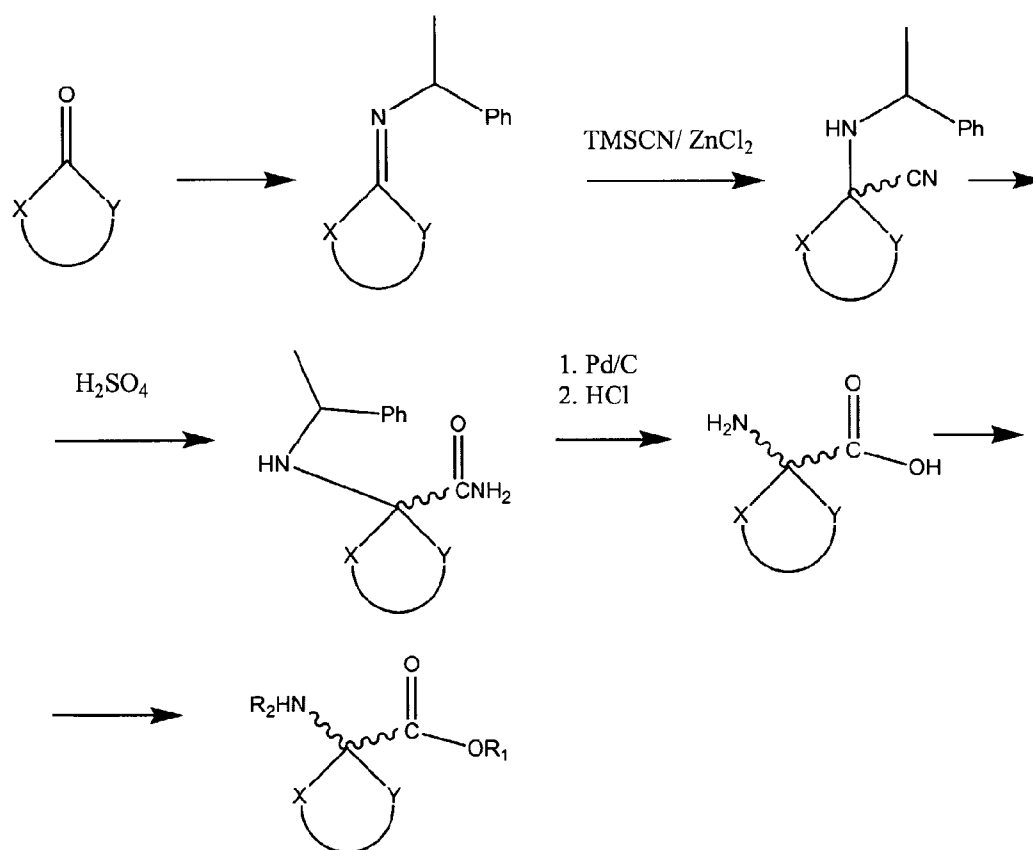
Synthesis of Compounds of Formula Vb in Which X and Y Comprise a Cyclopropane Ring

Among the various routes for the construction of α -amino acids, 1,3-dipolar cycloaddition of diazoalkanes with α,β -dehydroamino acid derivatives has been widely utilized. Hence the scheme below demonstrates that R_3 substituents of dehydroamino acids which are preferably alkyl or Ar^1 can be protected as the imino esters, where Ar^1 is as defined above. The skilled artisan will recognize that such compounds can be reacted with diazo substituted compounds which are preferably alkyl or Ar^1 to produce the resulting protected cyclopropane derivative. Reaction of such compounds with basic alcoholic solutions such as sodium methoxide followed by acid hydrolysis can provide the corresponding R_3 , R_4 substituted cyclopropane amino acids. Further derivativization of R_1 and R_2 can be readily accomplished by known methods.

**SCHEME 5b**

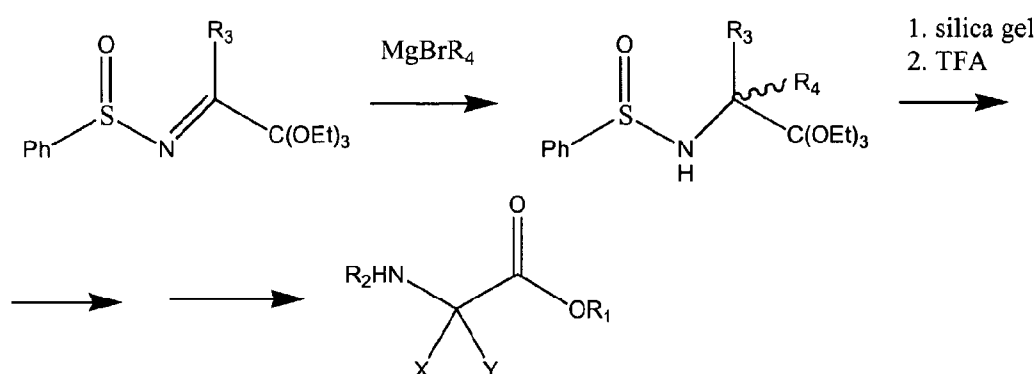
Synthesis of Compounds of Formula Vb in Which X and Y Comprise Rings of 5-8 Members

Substituted carbocyclic or heteroatom containing rings of preferably 5, 6, 7, 8 members can be transformed into amino acid derivatives consistent with the compounds represented by formula Vb. One of several well established routes is the conversion of a cyclic ketone containing compound to the corresponding amino acid derivative. Such cyclic keto compounds are abundant in the literature and are readily synthesized by the skilled artisan. The starting compound may be protected or unprotected. Trimethylsilyl cyanide addition to an imine derivative of the starting ketone provides cyano addition products. Hydrolysis and reductive cleavage of the protected amine generates the amino acid. Further derivativization of R_1 and R_2 can be readily accomplished by known methods.

**SCHEME 5c**Synthesis of Compounds of Formula Vc

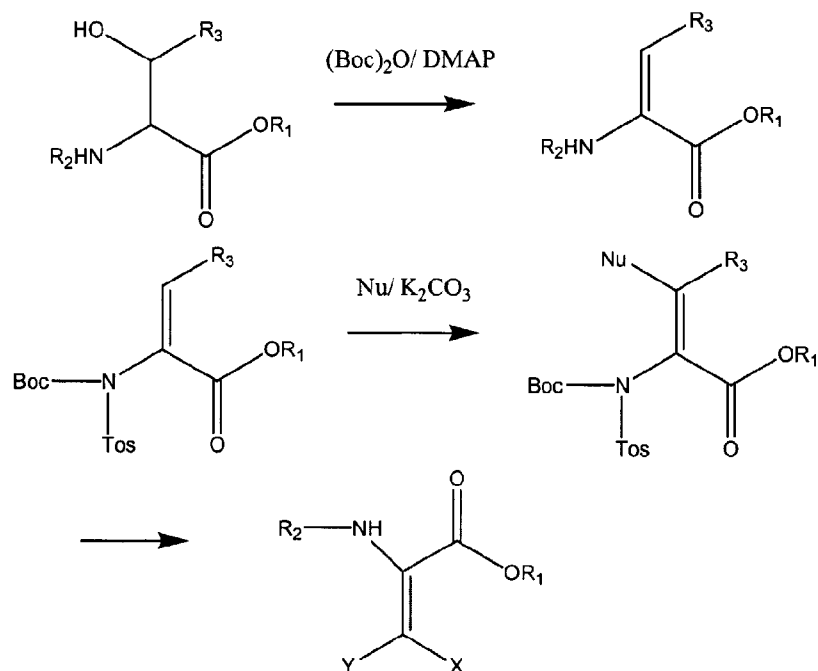
Compounds of formula Vc can be synthesized from sulfenimine derivatives of compounds substituted with R_3 where R_3 is preferably alkyl or aryl. There are several routes to the preparation of substituted sulfenimines that can be readily synthesized by the skilled artisan.

Addition of R_4 in the form of an organometallic reagent such as alkyl magnesium bromide followed by treatment with trifluoroacetic acid provides the corresponding disubstituted amino acid which can be further derivatized at R_1 and R_2 by known methods.



SCHEME 5d

Synthesis of Compounds of Formula VI



SCHEME 6

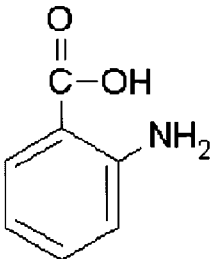
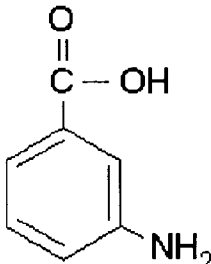
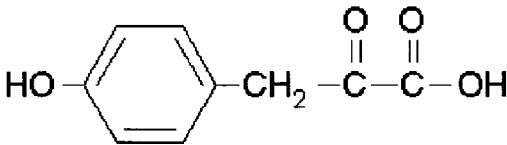
Mono or disubstituted dehydroamino derivatives can be synthesized from a substituted amino alcohol. Such amino alcohols are readily synthesized by one skilled in the art by methods similar to the procedures described earlier. Dehydration of the monosubstituted amino alcohol by (Boc)₂O /DMAP provides the dehydroamino derivative. Addition of nucleophiles (Nu) in the presence of base generates the disubstituted dehydroamino derivative

These steps may be varied to increase yield of desired product. The skilled artisan will also recognize the judicious choice of reactants, solvents and temperatures is an important component in successful synthesis. While the determination of optimal conditions, etc. is routine, it will be understood that to make a variety of compounds can be generated in a similar fashion, using the guidance of the schemes above.

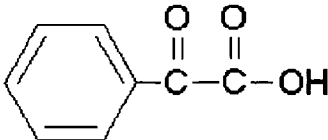
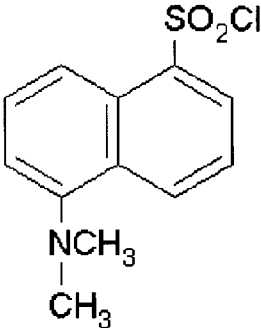
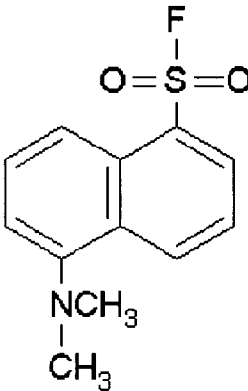
It is recognized that the skilled artisan in the art of organic chemistry can readily carry out standard manipulations of organic compounds without further direction; that is, it is well within the scope and practice of the skilled artisan to carry out such manipulations. These include but are not limited to , reduction of carbonyl compounds to their corresponding alcohols, oxidations of hydroxyls and the like, acylations, aromatic substitutions, both electrophilic and nucleophilic, etherifications, esterifications and saponifications and the like. Example of these manipulations are discussed in standard texts such as March, Advanced Organic Chemistry(Wiley), Carey and Sundberg, Advanced Organic Chemistry(Vovl 2.) and Keeting, Heterocyclic Chemistry (all 17 volumes).

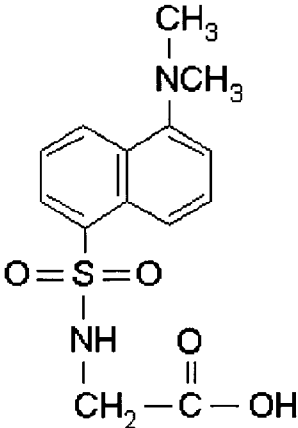
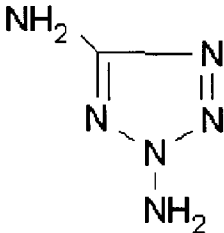
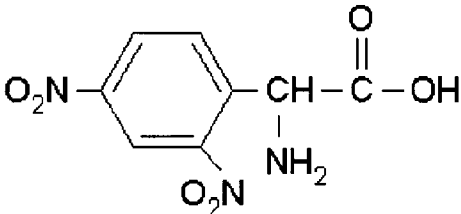
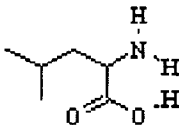
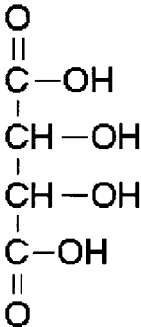
The skilled artisan will readily appreciate that certain reactions are best carried out when other functionality is masked or protected in the molecule, thus avoiding any undesirable side reactions and/or increasing the yield of the reactions. These reactions are found in the literature and are also well within the scope of the artisan. Examples of many of these manipulations can be found in T. Greene, Protecting Groups in Organic Synthesis. Of course, amino acids used as starting materials with reactive side chains are preferably blocked to prevent undesired side reactions.

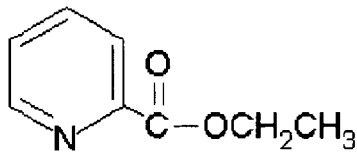
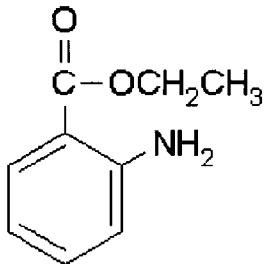
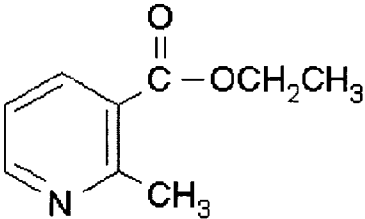
Table I

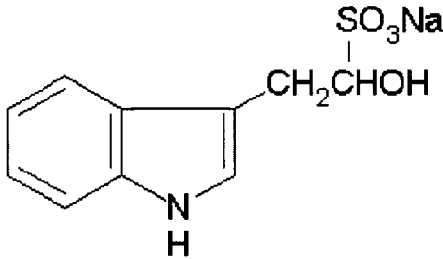
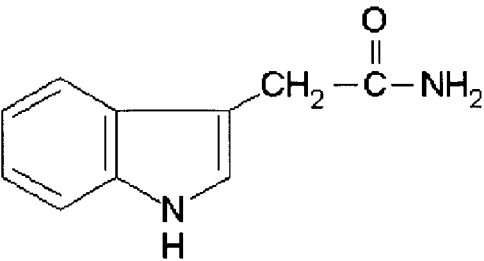
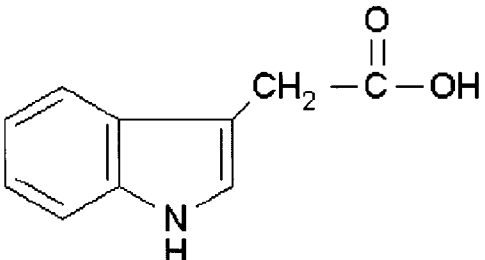
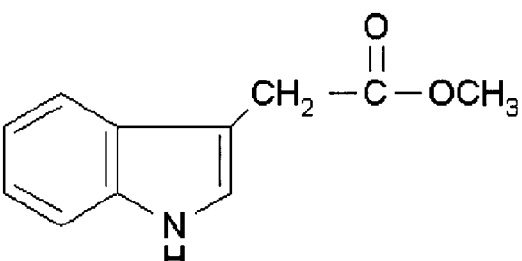
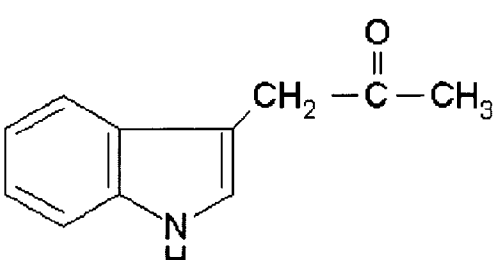
NAME	STRUCTURE
2- aminobenzoate	
2-hydroxybutyrate	$\text{CH}_3 \text{ CH}_2 \underset{\text{OH}}{\underset{ }{\text{CH}}} - \overset{\text{O}}{\underset{ }{\text{C}}} - \text{ONa}$
2-hydroxybutyrate	$\text{CH}_3 \text{ CH}_2 \underset{\text{OH}}{\underset{ }{\text{CH}}} - \overset{\text{O}}{\underset{ }{\text{C}}} - \text{ONa}$
3-aminobenzoate	
3-hydroxybutyrate	$\text{CH}_3 \underset{\text{OH}}{\underset{ }{\text{CH}}} \text{ CH}_2 - \overset{\text{O}}{\underset{ }{\text{C}}} - \text{ONa}$
4-hydroxyphenylpyruvate	
Acetone dicarboxylate	$\text{HO} - \overset{\text{O}}{\underset{ }{\text{C}}} - \text{CH}_2 - \overset{\text{O}}{\underset{ }{\text{C}}} - \text{CH}_2 - \overset{\text{O}}{\underset{ }{\text{C}}} - \text{OH}$

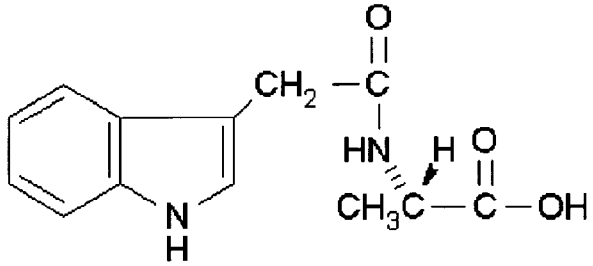
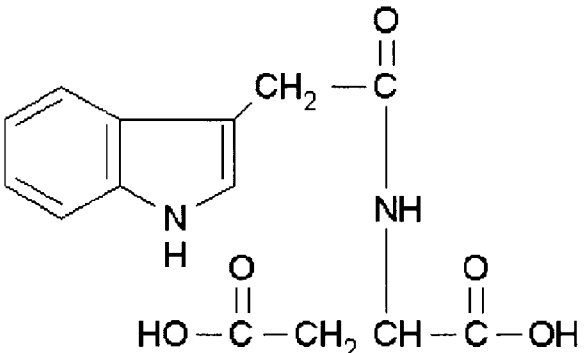
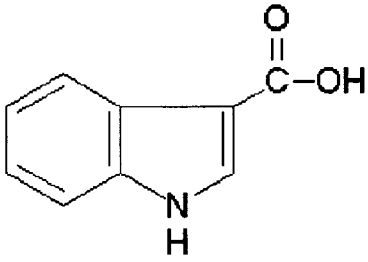
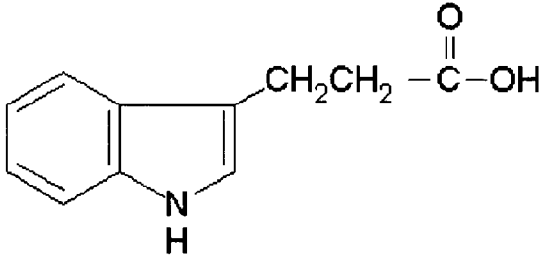
Aminoguanidine bicarbonate	$\begin{array}{c} \text{NH} \\ \\ \text{H}_2\text{NNH}-\text{C}-\text{NH}_2 \end{array} \cdot \text{H}_2\text{CO}_3$
Aminoguanidine HCL	$\begin{array}{c} \text{NH} \\ \\ \text{H}_2\text{NNH}-\text{C}-\text{NH}_2 \end{array} \cdot \text{HCl}$
Aminoguanidine hemisulfate	$\begin{array}{c} \text{NH} \\ \\ \text{H}_2\text{NNH}-\text{C}-\text{NH}_2 \end{array} \cdot \frac{1}{2}\text{H}_2\text{SO}_4$
Aminoguanidine nitrate	$\begin{array}{c} \text{NH} \\ \\ \text{H}_2\text{NNH}-\text{C}-\text{NH}_2 \end{array} \cdot \text{HNO}_3$
Aminoguanidine sulfate	$\begin{array}{c} \text{NH} \\ \\ \text{NH}_2\text{NH}-\text{C}-\text{NH}_2 \end{array} \cdot \text{H}_2\text{SO}_4 \cdot \text{H}_2\text{O}$
Benzoic acid copper salt	$\left[\begin{array}{c} \text{O} \\ \\ \text{C}-\text{O}^- \\ \\ \text{C}_6\text{H}_5 \end{array} \right]_2 \text{Cu}^{2+} \cdot 2\text{H}_2\text{O}$
Benzoic acid	$\begin{array}{c} \text{O} \\ \\ \text{C}-\text{OH} \\ \\ \text{C}_6\text{H}_5 \end{array}$

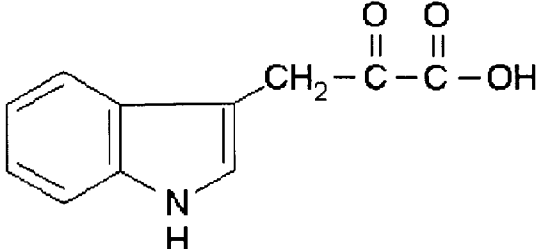
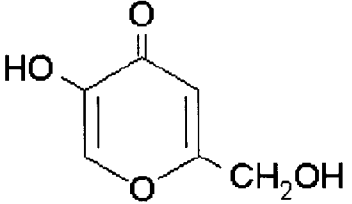
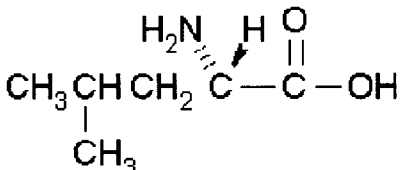
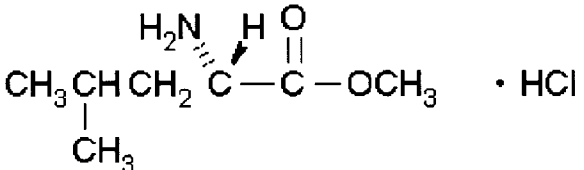
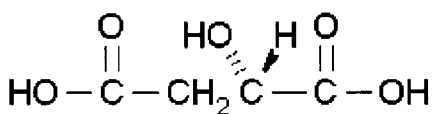
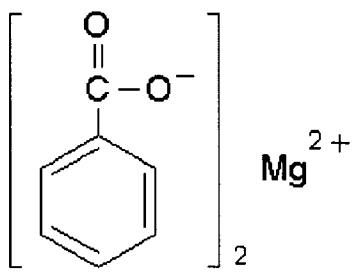
Benzoylformic acid	
cysteamine	$\text{H}_2\text{NCH}_2\text{CH}_2\text{SH}$
dansylchloride	
dansylfluoride	

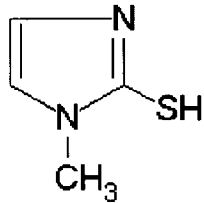
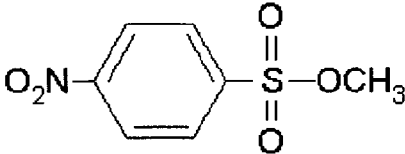
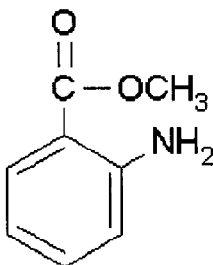
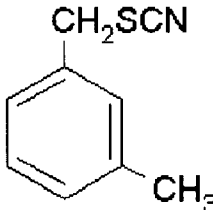
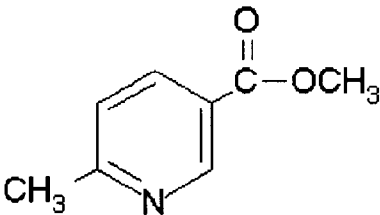
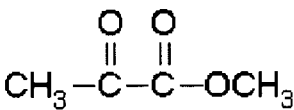
dansylglycine	 <chem>CN(C)c1ccc2cc(ccc2c1)S(=O)(=O)NC(C(=O)O)</chem>
diaminetetrazole	 <chem>Nc1nn[nH]n1N</chem>
dinitrophenylglycine	 <chem>O=[N+]([O-])c1cccc(c1C(N)C(=O)O)[N+](=O)[O-]</chem>
d-leucine	 <chem>CC(C)C[C@@H](C(=O)[O-])[NH3+]</chem>
DL-tartarate	 <chem>O=C(O)[C@H](O)[C@@H](O)C(=O)O</chem>

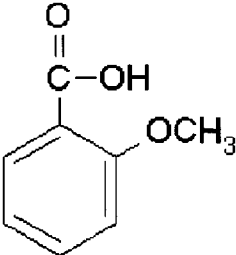
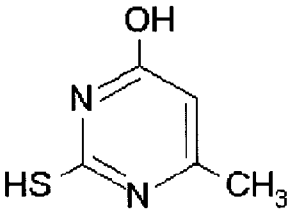
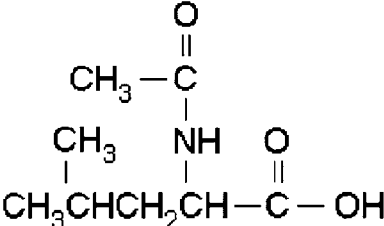
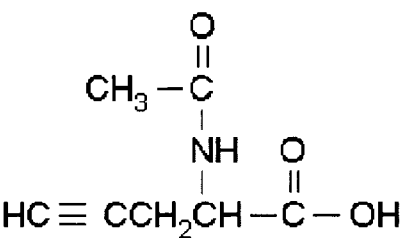
d-malicacid	$\text{HO}-\overset{\overset{\text{O}}{\parallel}}{\text{C}}-\text{CH}_2-\overset{\overset{\text{HO}}{\mid}}{\underset{\mid}{\text{C}}}-\overset{\overset{\text{H}}{\mid}}{\underset{\mid}{\text{C}}}-\overset{\overset{\text{O}}{\parallel}}{\text{C}}-\text{OH}$
D-tartarate	$\begin{array}{c} \text{O} \\ \parallel \\ \text{C}-\text{OH} \\ \mid \\ \text{CH}-\text{OH} \\ \mid \\ \text{CH}-\text{OH} \\ \mid \\ \text{C}-\text{OH} \\ \parallel \\ \text{O} \end{array}$
Ethyl-2-picolinate	
ethylanthranilate	
ethylbromopyruvate	$\text{BrCH}_2-\overset{\overset{\text{O}}{\parallel}}{\text{C}}-\overset{\overset{\text{O}}{\parallel}}{\text{C}}-\text{OCH}_2\text{CH}_3$
ethylmethylnicotinate	
HydroxylamineHCL	$\text{H}_2\text{NOH} \cdot \text{HCl}$

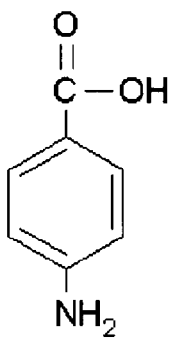
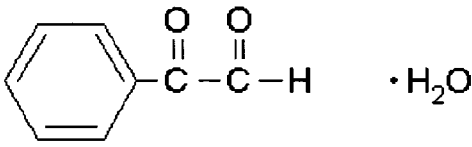
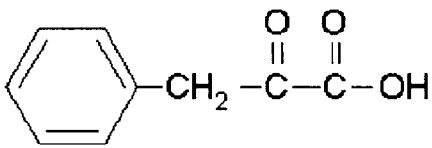
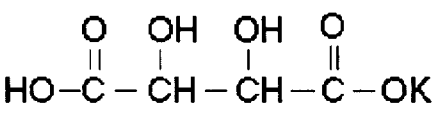
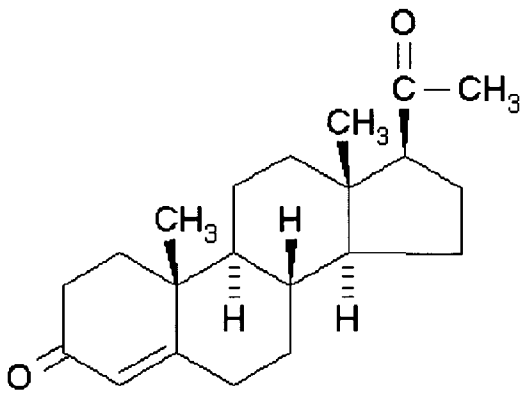
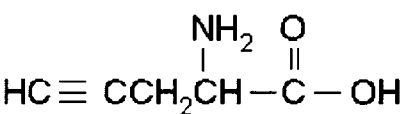
Indole-3-acetaldehyde sodium bisulfite	 <chem>NC1=CC=C2C(=C1)C(=CC=C2)CC(S(=O)(=O)O)O</chem>
Indole-3-acetamide	 <chem>NC(=O)CC1=CC=C2C(=C1)C(=CC=C2)N</chem>
Indole-3-acetate	 <chem>OC(=O)CC1=CC=C2C(=C1)C(=CC=C2)N</chem>
Indole-3-aceticmethylester	 <chem>COC(=O)CC1=CC=C2C(=C1)C(=CC=C2)N</chem>
Indole-3-acetone	 <chem>CC(=O)CC1=CC=C2C(=C1)C(=CC=C2)N</chem>

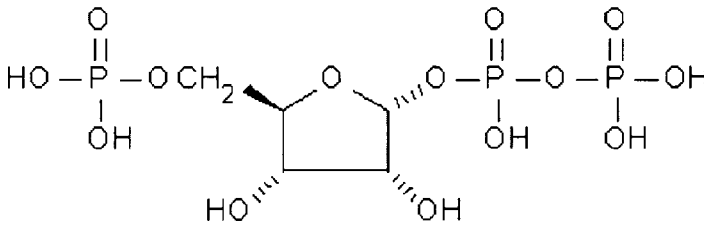
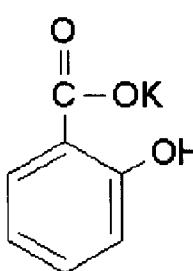
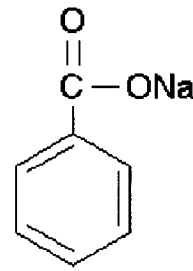
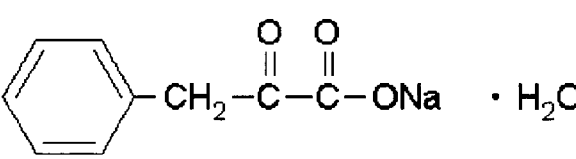
Indole-3-acetylalanine	 The structure shows an indole ring with a hydrogen atom on the nitrogen. At the 3-position of the indole ring, there is a CH_2 group connected to a carbonyl group ($\text{C}=\text{O}$). This carbonyl group is further connected to a chiral center (an alpha-carbon). This alpha-carbon is bonded to a hydrogen atom (H), a methyl group (CH_3), and an amino group (NH). The amino group is connected to another carbonyl group ($\text{C}=\text{O}$), which is part of a carboxylic acid group ($\text{C}=\text{O}$ and OH).
Indole-3-acetyl ASP	 The structure shows an indole ring with a hydrogen atom on the nitrogen. At the 3-position of the indole ring, there is a CH_2 group connected to a carbonyl group ($\text{C}=\text{O}$). This carbonyl group is connected to an amide group (NH). The amide group is connected to a chiral center (an alpha-carbon). This alpha-carbon is bonded to a hydrogen atom (H), a carboxylic acid group ($\text{C}=\text{O}$ and OH), and a CH_2 group. This CH_2 group is connected to another carbonyl group ($\text{C}=\text{O}$), which is part of a carboxylic acid group ($\text{C}=\text{O}$ and OH).
Indole-3-carboxylate	 The structure shows an indole ring with a hydrogen atom on the nitrogen. At the 3-position of the indole ring, there is a carboxylic acid group ($\text{C}=\text{O}$ and OH).
Indole-3-propionate	 The structure shows an indole ring with a hydrogen atom on the nitrogen. At the 3-position of the indole ring, there is a CH_2CH_2 group connected to a carboxylic acid group ($\text{C}=\text{O}$ and OH).

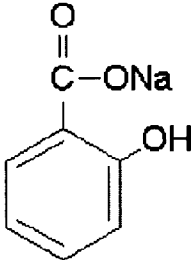
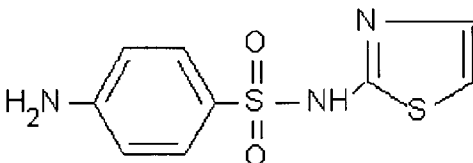
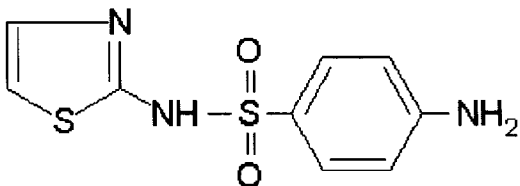
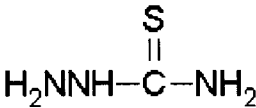
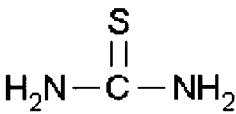
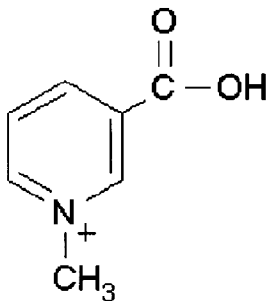
Indole-3-pyruvate	
Kiojic acid	
L-leucine	
L-leumethylester HCL	
L-malic acid	
Magnesium benzoate	

methimazole	
Methyl-4-nitrobenzenesulfonate	
methylanthranilate	
methylbenzthiocyanate	
methylmethylnicotinate	
methylpyruvate	

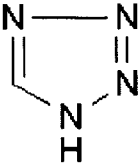
methylsalicylate	
methyluracil	
MGAG diHCL	$\text{H}_2\text{N}-\overset{\text{NH}}{\underset{\parallel}{\text{C}}}-\text{NH}-\text{N}=\overset{\text{CH}_3}{\underset{ }{\text{C}}}-\text{CH}=\text{N}-\text{NH}-\overset{\text{NH}}{\underset{\parallel}{\text{C}}}-\text{NH}_2 \quad \begin{array}{l} \bullet 2\text{HCl} \\ \bullet x\text{H}_2\text{O} \end{array}$
N-acetyl-D-leucine	
N-acetyl-D-L-propargylglycine	
Nitroethane	$\text{CH}_3\text{CH}_2\text{NO}_2$

p-aminobenzoate	
Phenylglyoxal monohydrate	
phenylpyruvate	
Potassium tartrate	
progesterone	
proparglyglycine	

pyruvate	$\text{CH}_3-\overset{\overset{\text{O}}{\parallel}}{\text{C}}-\overset{\overset{\text{O}}{\parallel}}{\text{C}}-\text{OH}$
RPP	
salicylate	
Sodium benzoate	
Sodium alpha-ketoisovalerate	$\text{CH}_3-\overset{\overset{\text{CH}_3}{\mid}}{\text{CH}}-\overset{\overset{\text{O}}{\parallel}}{\text{C}}-\overset{\overset{\text{O}}{\parallel}}{\text{C}}-\text{ONa}$
Sodium phenylpyruvate	
Sodium pyruvate	$\text{CH}_3-\overset{\overset{\text{O}}{\parallel}}{\text{C}}-\overset{\overset{\text{O}}{\parallel}}{\text{C}}-\text{ONa}$

Sodium salicylate	
Sodium sulfathiazole	 (Na⁺)
Sulfathiazole	
thiosemicarbazide	
thiourea	
Trigonelline	 Cl⁻

126

Unsubstituted tetrazole	 <chem>c1nn[nH]n1</chem>
-------------------------	---

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. An isolated or purified nucleic acid encoding a DAO polypeptide selected from the group consisting of:
 - 5 a) a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence selected from the group of sequences consisting of SEQ ID NOs:8 to 10; and
 - b) a nucleic acid molecule comprising a nucleic acid sequence selected from the group of sequences consisting of SEQ ID NOs:1 and 3 to 6, or a sequence
10 complementary thereto.
2. A purified or isolated DAO polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence selected from the group of sequences consisting of SEQ ID NOs:8 to 10;
 - 15 b) a polypeptide encoded by a nucleic acid molecule comprising a nucleic acid sequence selected from the group of sequences consisting of SEQ ID NOs:1 and 3 to 6.
3. A method of identifying a candidate molecule for the treatment of schizophrenia,
20 depression or bipolar disorder, said method comprising:
 - a) contacting a DAO or DDO or a biologically active fragment thereof polypeptide with a test compound; and
 - b) determining whether said compound selectively inhibits the activity of said polypeptide;
 - 25 wherein a test compound that selectively inhibits the activity of said polypeptide is identified as a candidate molecule for the treatment of schizophrenia, depression or bipolar disorder.
4. A method of identifying a candidate molecule for the treatment of schizophrenia,
30 depression or bipolar disorder, said method comprising:
 - a) contacting a DAO or DDO or a biologically active fragment thereof polypeptide with a test compound; and
 - b) determining whether said compound selectively binds said polypeptide;
 - wherein a test compound that selectively binds said polypeptide is identified as a
35 candidate molecule for the treatment of schizophrenia, depression or bipolar disorder.

5. A method of screening for antagonists of a DAO polypeptide, comprising the steps of:
- a) contacting a test compound with a DAO polypeptide of claim 2;
 - b) detecting the level of DAO activity; and
- 5 c) comparing the activity to the activity of a control test without the test compound, whereby a decrease in the level of the DAO activity over the control indicates that the test compound is an antagonist of DAO.
6. A method of screening for compounds that inhibit the expression of the DAO mRNA or polypeptide, comprising the steps of:
- a) incubating cells expressing a DAO polypeptide of claim 2, (i) in the presence and (ii) in the absence, of a test compound; and
 - b) detecting the level of the DAO mRNA or polypeptide in the cells.
- 15 7. A method of detecting an association between a DAO-related biallelic marker and schizophrenia, depression or bipolar disorder, comprising determining the identity of a polymorphic nucleotide at a DAO-related biallelic marker, or the complement thereof in a biological sample, wherein the polymorphic nucleotide is associated with schizophrenia, depression or bipolar disorder.
- 20 8. A method according to claim 7, wherein said determining is performed by a hybridization assay, a sequencing assay, a microsequencing assay, or an enzyme-based mismatch detection assay.
- 25 9. A method of estimating the frequency of an allele of a biallelic marker in a population comprising;
- a) genotyping individuals from said population for said biallelic marker according to the method of claim 7; and
 - b) determining the proportional representation of said biallelic marker in said
- 30 population.
10. A method of detecting an association between a genotype and a trait, comprising the steps of:
- a) determining the frequency of at least one biallelic marker in trait positive
- 35 population according to the method of claim 7;

- b) determining the frequency of at least one biallelic marker in a control population according to the method of claim 7; and
- c) determining whether a statistically significant association exists between said genotype and said trait.

5

11. A method of determining whether an individual is at risk of schizophrenia, depression or bipolar disorder, comprising:

- a) genotyping at least one biallelic marker according to the method of claim 7; and
- b) correlating the result of step a) with a risk of developing schizophrenia, depression or bipolar disorder.

10

12. A method according to any one of claims 7 to 11, wherein said biallelic marker is selected from the group of markers described in SEQ ID NOs:1, 24, 26, and 29, and the complements thereof.

15

13. A method of diagnosing, or detecting a predisposition to or susceptibility to schizophrenia, depression or bipolar disorder in a subject, comprising

- a) obtaining a nucleic acid sample from said subject;
- b) determining the identity of a nucleotide at a DAO-related polymorphism, or the complement thereof in said biological sample; and
- c) determining whether the subject has a DAO polymorphism associated with schizophrenia, depression or bipolar disorder.

20

14. Use of an antagonist of DAO or DDO in the manufacture of a medicament for the treatment of psychosis, schizophrenia, bipolar disorder, depression or a mood disorder.

25

15. A method of treating psychosis, schizophrenia, bipolar disorder, depression or a mood disorder in a patient in need thereof, comprising administering to said patient an effective amount of a composition or compound comprising a DAO antagonist or a DDO antagonist or inhibitor.

30

Dated this fifth day of January 2007

Serono Genetics Institute S.A.

Patent Attorneys for the Applicant:

F B RICE & CO

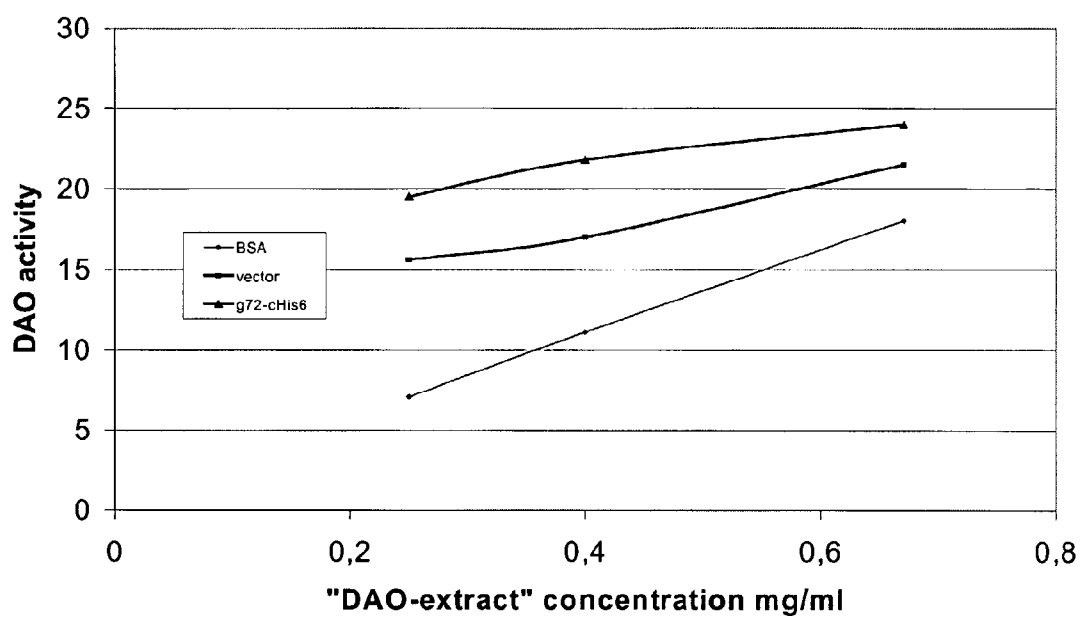
Figure 1**Recombinant human DAO activity in the mixes
of yeast *S. cerevisiae* extracts**

Figure 2

human DAO activity in the presence of E.coli g72 extract

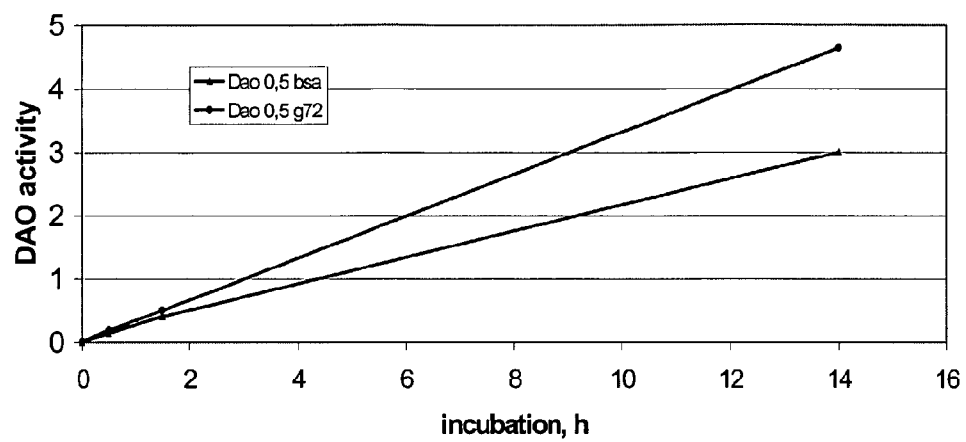


Figure 3

DAO activity in the presence of the different concentrations
of g72

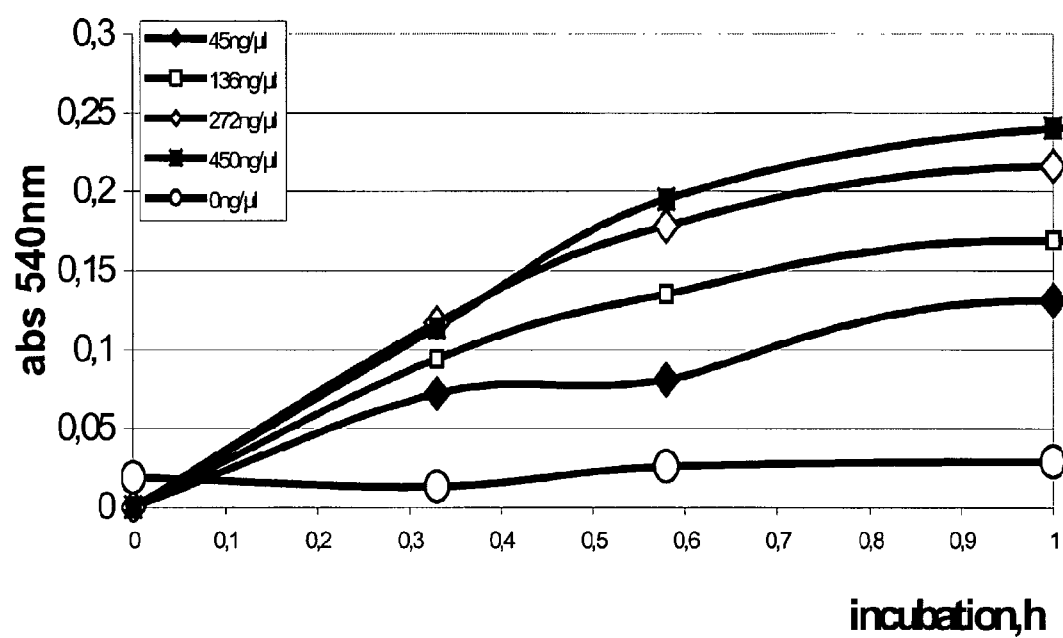
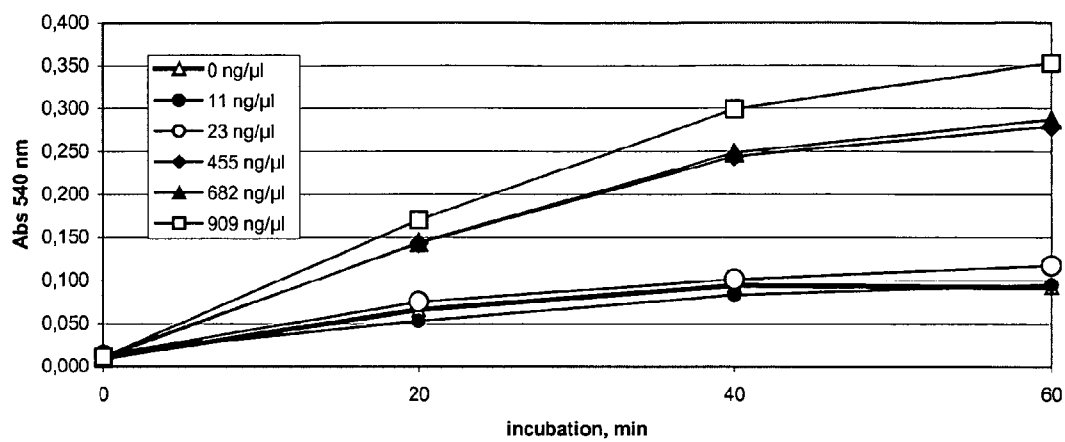


Figure 4**DAO activity in the presence of the different concentrations of g72**

5/6

DAO kinetics

Figure 5

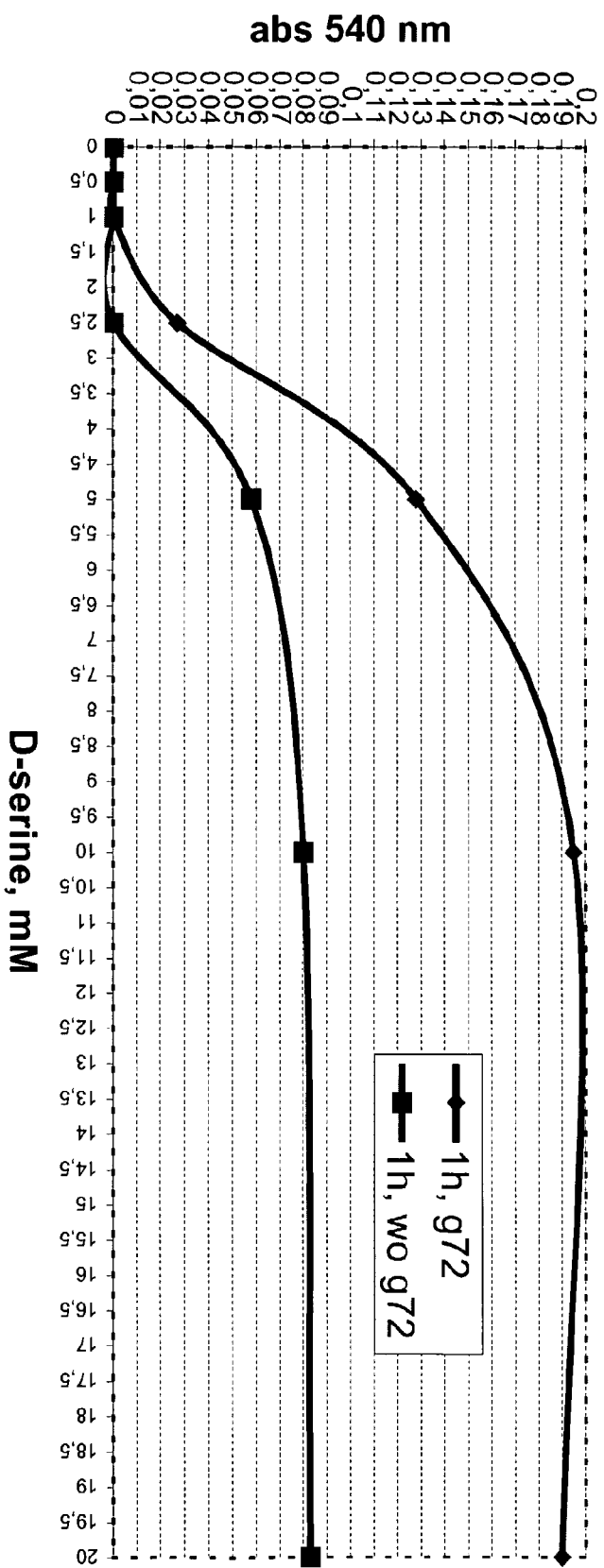


Figure 6:

FIGURE 6: ASSOCIATION RESULTS DAAO - PJ 27 Algene sample (213 cases, 241 controls)

Mks	Location	Position	ALLELIC TEST				GENOTYPIC TEST	
			Chosen Allele	Allelic freq Diff.	Chi Sq.	p.value	Chi sq.	p.value
24-1443/126	5' of gene	132905	T	0.004	0.0171	8.96E-01	0.0206	9.90E-01
24-1457/52	intron B	94290	C	0.029	0.6981	4.03E-01	1.0478	5.92E-01
27-93/181	intron 3	65365	T	0.079	7.3783	6.60E-03	8.9084	1.16E-02
24-1461/256	intron 10	55058	G	0.077	6.4495	1.11E-02	6.5728	3.74E-02
(*) : exact test								

<110> GENSET

<120> Treatment of CNS Disorders Using D-Amino Acid Oxidase and D-Aspartate Oxidase Antagonists

<130> 101.WO1

<150> 60/261,883

<151> 2001-01-16

<150> 60/305,445

<151> 2001-07-13

<150> 60/

<151> 2001-10-22

<150> 60/333,881

<151> 2001-11-19

<160> 26

<170> Patent.pm

<210> 1

<211> 86592

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> 38388..40388

<223> 5'regulatory region

<220>

<221> exon

<222> 40389..40670

<223> exon Z

<220>

<221> exon

<222> 42666..42778

<223> exon A

<220>

<221> exon

<222> 43416..43519

<223> exon B

<220>

<221> exon

<222> 61159..61402

<223> exon C

<220>

<221> exon

<222> 64050..64711

<223> exon Ulong

<220>

<221> exon

<222> 64578..64711

<223> exon U

<220>
<221> exon
<222> 68126..68261
<223> exon V

<220>
<221> exon
<222> 69488..69690
<223> exon 2

<220>
<221> exon
<222> 71942..72056
<223> exon 3

<220>
<221> exon
<222> 73962..74038
<223> exon 4

<220>
<221> exon
<222> 74701..74766
<223> exon 5

<220>
<221> exon
<222> 77478..77532
<223> exon 6

<220>
<221> exon
<222> 78762..78866
<223> exon 7

<220>
<221> exon
<222> 81507..81589
<223> exon 8

<220>
<221> exon
<222> 83181..83298
<223> exon 9

<220>
<221> exon
<222> 83879..83977
<223> exon 10

<220>
<221> exon
<222> 84906..85423
<223> exon 11

<220>
<221> exon
<222> 84906..85541
<223> exon 11long

<220>
<221> misc_feature

<222> 85542..86592
<223> 3'regulatory region

<220>
<221> allele
<222> 6199
<223> 24-1443-126 : polymorphic base C or T

<220>
<221> allele
<222> 44814
<223> 24-1457-52 : polymorphic base A or C

<220>
<221> allele
<222> 72600
<223> 27-93-181 : polymorphic base A or G

<220>
<221> allele
<222> 84046
<223> 24-1461-256 : polymorphic base A or G

<220>
<221> primer_bind
<222> 6074..6094
<223> 24-1443.pu

<220>
<221> primer_bind
<222> 6559..6579
<223> 24-1443.rp complement

<220>
<221> primer_bind
<222> 44763..44782
<223> 24-1457.pu

<220>
<221> primer_bind
<222> 45257..45277
<223> 24-1457.rp complement

<220>
<221> primer_bind
<222> 72359..72378
<223> 27-93.rp

<220>
<221> primer_bind
<222> 72763..72780
<223> 27-93.pu complement

<220>
<221> primer_bind
<222> 83791..83811
<223> 24-1461.pu

<220>
<221> primer_bind
<222> 84307..84324
<223> 24-1461.rp complement

```
<220>
<221> primer_bind
<222> 6180..6198
<223> 24-1443-126.mis

<220>
<221> primer_bind
<222> 6200..6218
<223> 24-1443-126.mis complement

<220>
<221> primer_bind
<222> 44795..44813
<223> 24-1457-52.mis

<220>
<221> primer_bind
<222> 44815..44833
<223> 24-1457-52.mis complement

<220>
<221> primer_bind
<222> 72581..72599
<223> 27-93-181.mis

<220>
<221> primer_bind
<222> 72601..72619
<223> 27-93-181.mis complement

<220>
<221> primer_bind
<222> 84027..84045
<223> 24-1461-256.mis

<220>
<221> primer_bind
<222> 84047..84065
<223> 24-1461-256.mis complement

<220>
<221> misc_binding
<222> 6187..6211
<223> 24-1443-126.probe

<220>
<221> misc_binding
<222> 44802..44826
<223> 24-1457-52.probe

<220>
<221> misc_binding
<222> 72588..72612
<223> 27-93-181.probe

<220>
<221> misc_binding
<222> 84034..84058
<223> 24-1461-256.probe

<400> 1
attattggaa caggccacac ttgcgagggg agtccctgcc tcagaaagat tcagaaaagc 60
tagacagtca ctggaagaac aattacaacc gcaagacggt caaacactaa acaccgtat 120
```

gcctcagaac	cgtacagata	atggccaaat	agatggggct	ctgggcattt	ctgagagcac	180
ctgcctgggtg	gcaccccatc	ctaattggacc	atgcccctcca	gtctccaagt	ggctcttcag	240
agctcacatc	cgaacacctc	ctatgctaca	ggttcttcta	gccccagggt	cccaaccacc	300
ccaaggccac	agaggccagc	cccaactcca	tcttctacat	gtgtcacagg	aaactttctc	360
atagtgtat	ttattatgta	ctgcgggggt	gggggcatg	tcataaaaaga	aatgtcctcc	420
cttttttatt	catctccttc	taacaagcat	caaagtctca	gtcgctagca	tgtgacttac	480
agaagctctc	atgggaacaa	gacaagacca	tactgttacc	gtgacactca	cggcctccct	540
gactggtttc	tgctgttgat	tctgcctcaa	atgctcctca	aatgcacctt	gctgtccgc	600
ctccacccta	gagctcgct	gactgccac	ttgcccgtta	agagtgggt	taggtctcac	660
tcctgccaga	aaggctctgc	cagggtgctct	caacagtcac	cccctcctgt	ggctctcaca	720
aaccccagca	cctctcggtc	actctctccc	tcctatctgg	ttgtgactgt	cttccatgct	780
cacttagaag	ctctctgagg	ccaagaactg	tgtgtactgt	tgcttctttg	tttacctggg	840
cctagcccat	tgccctatac	acaggagaat	gcaaataaat	catatgctta	atgaatgagt	900
cgatgaatga	atgatgaata	aagggaatct	aacttagttt	taacaaatcc	aggtttttga	960
atgatctcac	agccattcat	ttatcttggt	atgtcagggg	agtgactcca	ccctcatttc	1020
acacgtactct	tggggctcaat	gctctaactt	acttggcctc	cagttagtgg	gaaattacaa	1080
gctacacttc	aagcctctga	ctaggacctg	ccatgaagta	cttgggaatc	agtggagtat	1140
cactgtgggg	tgaggtgtct	gaggcgaggc	ccaccaatct	ccatacttct	ccccggggcc	1200
ctctgctga	gagggctctc	ctgcttcctt	tggcagactc	tggtttggtc	ttctgggttc	1260
ggcgttggtg	tcacctcctt	caggaagcat	ttctggctaa	gggtgccccac	tctatagcag	1320
ctgggtgtaa	acctctctaa	gcaaacagca	taactttctg	tcctctcaat	tgactctgag	1380
ttctgagagc	acagcctgga	gctggcacgg	tgcttggcac	agagagctga	aatggcacac	1440
cctagtgttc	ccagtggctc	gactccccag	gctctccatc	aggacgcagc	cctctcccac	1500
ctctgatgga	tatgggacca	tggaatgctt	tgtccagcag	caactcttgc	ctccctcaca	1560
gaagggaaca	cctagcccat	cagactcacc	tttccctact	ggaaaagtcc	actcccagca	1620
agatatcttc	ctcgggtgtc	tggcgccgc	tgctgtacac	caccaccatg	taccggaccc	1680
ggctccgcca	ggcgtctctc	aggcgactg	cctggaacag	ggcagacatg	ctctcactaa	1740
cctgcctttg	gaggtggtgc	ctccctccca	tctccaatgc	aagatcaaca	ctttcagtgt	1800
tctaccttct	cctctgggag	ttaaaaatga	agagaaaatt	cttggctggg	catgggtggt	1860
caggctgtga	attccagcac	tttgggaggc	caaggtaggc	agatcacttg	aggtcaggag	1920
ttcgagacca	gcctggccaa	catggcaaaa	ccccatctct	tactaaaaat	acaaaaatta	1980
gctgggcatg	gtggcggtgc	cctgtaatcc	cagctactca	ggggactgag	gcacgagaat	2040
ctcttgaacc	cggaagcgg	aggttgagc	gagctgagat	catgccacca	catgccagcc	2100
agagcgacag	agtgagattc	tgtctcaaac	aacaacaaca	caacaaaaca	caaagcggaa	2160
gttcttgaca	gcaggaaacca	ggcctcgttt	ctctctgtag	caccagggac	gccgcctggc	2220
tcagaggaat	accacaaaat	gcaagaaatc	agtgaacaca	tgaaatccaa	agaaagtctg	2280
tatttagctt	atttaactgc	cgtggagacc	gtttctatcc	ctcctccgc	ccctctgggg	2340
aactgaggag	tcaacctggc	tttggcttta	gtgcacaatt	tgagaatttg	ttgtaacctt	2400
aaagcttttc	cccttatcat	tcacgaatgg	ttccccacca	ggtttcacaa	ttaaaaatta	2460
aaacttgctg	gctgggcacg	gtggcttaca	cctataactc	caggactttg	ggaggcagag	2520
gcaggagaat	catttgaggc	caggagtcca	agaccagcct	gggcaaaaata	gcaagacccc	2580
atatccacaa	aattttttta	aaaataaggc	aggggtgtac	acacttgtag	tcccagctat	2640
ccaggaggct	gaggtgggaa	gattgcttga	acccaggagt	ttgaggctgt	agtgaactaa	2700
gatctagcca	gcctactcca	gcctgggcaa	cagagcaaga	ccctcatctc	acaaaaatta	2760
aaaaaaaaatt	ttttaacttg	acattctcac	tgcttcttac	cagcttgatt	ctgtcttcgc	2820
aacgcagaag	gttgatcatc	acctgaagat	gttgaggcag	atcacctgtt	ggaccaataa	2880
agaaagcttt	aaaaggtctc	ttacctactc	tctaggaaaa	aaaacctctg	aaaggctgac	2940
tttgagggtc	tggaaaaaga	ttgagaagtt	aaaatttgct	tacctacacc	acaggagaat	3000
caccacaaaa	acttcaagtc	tgaatttctc	ttacaccact	ctgaatactg	tgcgacgtgg	3060
atgggtgaca	tggagcttac	tgtcatgttg	ttaaaagtgt	ctcttatttc	ctgaaatata	3120
tacagtatag	gtttccaaat	acaaaatgtg	aaaaatacac	gcaagcctag	agaaaaatgt	3180
tatttcattc	aagccaatgt	tactcggcag	gttgggtgct	ctagaaacga	cagctgtggc	3240
tggaaagtaag	gcatttgcta	agagttaatc	attagagaaa	aaggacagag	catcacgttt	3300
cctcttcaaa	caacttcttc	ttctatacac	agtctcgcac	tgtcaccacg	gctggagtgc	3360
agtgttacga	tctcagctca	ctgcaacctc	cgcctcctgg	attcaacaga	ttcttctgct	3420
tcagcctcct	aagtagctgg	gattacaggt	gcccataccc	agaccgggt	aattttttgt	3480
tttatagtag	agatggagtt	tcaccatgtt	ggccaggctg	gtctcgaact	cctgacctca	3540
agtgatctgc	acgctcagc	ctcccaaagt	actgggatta	caggcataag	ctaccccacc	3600
caggcccccac	ttcaaaactc	tgcaatttcc	gctggaggca	gacattattt	ccataaccgg	3660
ggggggcgggg	ggaaatgttt	aagtgactct	acagatagca	gctgtatgct	ggttgcccag	3720
agaaataatt	tgaatagaaa	ccaatctgtc	attttctctt	ttcttgctaa	aaattatgta	3780
ctcttttttc	ttcactatgt	aaaacaggca	gtaaccaggg	acggcttctg	aacttctctg	3840
agctgccccca	gggttcaggga	gggtgtctctg	gagtgacagt	aggaaagtct	cttactggcc	3900

atgagtctcg	cgcgaagcag	agaccctgtc	agaagaagcg	cacactttca	cggaggggaa	3960
agttgtaagg	gaggtgcata	attagtaagt	agcaggtgtg	actccaaggt	tgcttttttt	4020
ctctagctta	cacatttttc	tttatatctg	caaggatttc	tttctgaaga	aagggtcatc	4080
tgtagagatg	ctaataatcag	cctggtgtgg	tggctcacac	ctgtaatccc	agtgccttgg	4140
gaggccgagg	caggagactc	acttgaggcc	aggcattcaa	gaccagtctg	ggcaacatgg	4200
caagacccca	tctctacaga	aaagtaaaaa	attagctggg	ctttctgggt	cacatctgta	4260
gtcccagcta	cttgggaggg	caaggcagga	ggatcgctgg	agcccaggag	tttgagatca	4320
ccctgggcaa	cacgataaga	ccctgtctct	acaaaggaaa	aaaaattact	ctatacatca	4380
caattacaac	cccaaaagga	tcaataatgc	ttcacacactc	aaatgctcca	aaaggaaaca	4440
ttgtgtttgt	tccttttgca	aaagcatctt	tttcatttta	agggagaagg	acagatgatg	4500
tccaaattgc	acttctctgc	tcagagagga	attgggtcat	tagaaatttg	tgctcttagc	4560
caggagggta	gatctcatgt	taagcgttct	ttctttttct	ttttttttca	atagagacag	4620
ggttttgcca	tgtgtcccag	gctggtctcg	aactcctgga	ctcaaatgat	cctctcacct	4680
cagcctccca	aagtgtctgg	attataggca	tgagccacca	aaccagggcc	aattaagcat	4740
tctttccaca	ataagtaaaa	tttaaaaaag	aaaagaacca	tgccctctct	atctgtcctc	4800
tcagttata	caattccaca	gtgtataaca	ccctgtgttg	accctgcttc	ctatgtgag	4860
cgatttggag	ataagggttc	acattaaaga	aagccataga	cctcccagc	cccttctcc	4920
acccgtcatg	tcaccaatgc	aacacaacga	caacgaccat	gagctgggtc	ttcacctgcc	4980
tgggccctcc	caccatctac	ccgagtcaca	gaactgcatt	ggggaaagca	aaaacaaacc	5040
cctgtctgat	aaatgcctaa	atgaaaggga	cattttccac	acagataaac	ttctttcagt	5100
gggattgttt	gctgagatat	ggaactgtcg	acagacagaa	atccaaaccc	cagtctgaca	5160
tccacacaca	aaaaaatcag	agaatataag	ccctagaaag	ggtctcaaat	tgactggact	5220
ggctgaaaca	aactgaacta	cttttccaag	gacagaatta	accctcaatt	gtactcagct	5280
ctgcacagtg	gttactgggg	ggcctctggg	acattcagga	gacttgatgg	taattctagg	5340
gaaaaaaagg	aactaacgta	agtctagtct	gcgtctgtcc	caaggtcatt	tacagaccaa	5400
ctgtggacag	ctggcgcccc	ctctgccttc	cgacctcacc	gtccactcca	gacctcaggg	5460
cacaagagtc	agccagctgg	tggcttgcat	cctacccttc	tagtctttgg	attagaggaa	5520
ggaggtatct	gacacttagt	gagcagagct	tgagcatttg	ctttgtcata	tgtgttacaa	5580
ttaaaacatg	aacaacagct	acattttctaa	gagggcagaa	taattagcaa	attcaagaac	5640
gaagaatctg	gctgagtatg	gcacctcaaa	ccataaatcc	caatgctttg	agaggctgag	5700
gtggggagat	ggccttgagg	caggagttag	agaccagcct	gagcaacata	gtgagttaga	5760
cctcatctac	acacacacac	acacacacac	acacaactag	ctgggtgtgg	tgacacgcat	5820
ctatggtccc	agctactcag	gaagctaagg	ctggacaatc	acttgagccc	aggaggtcga	5880
ggctgcagtg	agctatgatc	aggccactgc	acgccagcct	gggcaagaga	atgagatccg	5940
tctctaaaaa	aacttttcat	ataattaaaa	aaaaaaaaag	aatgaagggt	ctgtttatag	6000
ctgtattgta	ctagaagtca	togtaataac	aatgatagtt	acccatata	atatacagca	6060
ctactactag	gtaggatgtg	tacacgcata	actctaaatt	tccatattgt	ctgaggcacc	6120
agtatttgat	ggccttggtg	aagactagga	aactcaggct	tagaagtcca	cctgtaccgg	6180
cttagtaagt	tggagaacya	ggatcagaag	acaggctctgc	ctggcttcaa	aacaaatact	6240
atttccacaa	accacactgc	ctccttgtag	aggacagtta	ttttctttgc	ttaaaacaga	6300
cctaaatatt	atcaacatca	gtatgtgaaa	atactgactg	agccttggtg	tttgctataa	6360
attgcatggg	gtagaattct	aacctgagca	ctcagatcta	aaatgaagct	gaatgacttg	6420
aggttaaaaa	aacaaaatgt	tcacaagaaa	actggccaca	atagctgggt	ggtttcacct	6480
gctgctgttc	tgaagggtaa	aggccttctc	agctcacaga	cattcaatta	tgcatgcctc	6540
ctcaagaaa	tgccctgaga	tgctgtccac	ctacgacaaa	gatccactta	catgcaagca	6600
ctttttcttc	tttctttctt	tttgagatag	ggctcctttc	ttttgtcacc	caggctggag	6660
tgagtgggcg	caatcggtgg	tcactgagca	acacagtggg	caacatagtg	agacctcatt	6720
tacatacaca	cccacaaaaa	actagctggc	tgtggtgaca	catcagcctc	gacctcctgg	6780
gctcaagcaa	tcctcccacc	tcagccccc	accttgcctg	gattacaggc	atgcgccacc	6840
acgcccagct	aattttttgt	ttttttgcag	agatagggtt	tcactgtgtt	gttcgggtta	6900
gtctggaact	cctgggttca	agcgagatct	gcccaccttg	gcctcccaaa	tcctggaatt	6960
acaggcaaga	gccaccgtgc	ctggccataa	tggtgttttg	ttgttattgt	ttttaagaaa	7020
cagagctctct	ctctgtccac	caggctggag	tgcagtggcg	tgatccctagc	ttgctgcagc	7080
ctcaaaactcc	tgggctcaag	cgatcctccc	aactcagcct	cccaaagcac	tgggattaca	7140
ggtgtgagat	accatgcagg	gccacgcaag	catttcttga	attcctcttt	ctaactgcct	7200
tcagctctga	gtcaagtctc	ctaagaaaac	cagtcttact	acttagtagg	cacttcttat	7260
ttaaactcag	tttgatctct	acctatttac	ttctgtctac	ttcctaaaaa	caaactatta	7320
cagaatcaag	acttctctact	acagtgtcta	tctcagagtt	ggagccaaag	gcccttcaag	7380
aaattctcca	aatgagtggt	tttcaaatgc	ttggagaaat	ccatcccaag	attaggtata	7440
cagcactcca	gatggttatt	ttcaagtggg	cgacactctg	ctataattca	ttttggtgca	7500
tttggttaaaa	agtcaggctg	taacttacag	cctgcaatta	actgataaac	tacagagagg	7560
aaatctttgc	atcccagcag	gatgctgctg	accttactcc	tgacgcagac	agacatgaca	7620
taaaaggttg	gaaaatgtgc	gtggtctgct	caagagagag	catctgagcc	tctgcctgca	7680

ctgggtcactg	caaacctgcg	tccactatgt	ctaaggcctt	caaactcagc	aacatcacca	7740
acaatggaag	tttcctctgc	tgtccagaaa	agaagctcca	atgtaagagt	atcaacttag	7800
agccctcacc	tgcattgctt	tgggggtgct	gaagactccg	ctggccttga	gggctgcttc	7860
cctgttgtaa	gaagagggct	gcgcctttca	ccatgaaaaa	gctctcactt	aagctgggaa	7920
ggataagacc	agagcacagt	tagaccggaa	ttcagacagg	aaaatggaca	aagaattact	7980
gcaggggaaa	aagcttttagc	gtggacaaat	ggcatgtaaa	atgcaaatag	gatgaaactg	8040
cttttataat	aattccacgt	agtacttttc	tcaaaccctg	cttttgctaa	aagcttgctg	8100
ctggagaatt	ttcgtgacaa	aataatgctt	ctgtgacaac	acccaaagt	ctacataggc	8160
tctccagggc	ccctttctgc	agaatactgg	acaggatctc	cactgtcata	taacattttc	8220
ttctttcttt	tttttttgag	acgcagtttc	actctgtcat	ccaggctgga	gtacagtggg	8280
gtgatctcaa	ctcactgcaa	cctctgcctc	ctgggttcaa	gcgattctca	tgtctcagcc	8340
tcccagaag	ctgagattac	aggcatgtgc	caccatggcc	agctaattat	tgtattatta	8400
gtagagacat	ggttttacca	tgttgccag	gctggtctca	aactcctgac	ctccagcaat	8460
ctgcctgctt	aggcctcctg	gagtgtgtg	attacaggcg	tgaccacgcc	cagccataac	8520
atttttctaag	aaaagagaac	aactccctga	ttaggagagg	gcagtctact	ttgtgaattc	8580
tctgtctctt	gattgtgatc	tctgttctta	actctctggc	ttttaacaac	tccattgttt	8640
cttgggtgact	tccttctgat	gaatacaagg	atgaaattac	actttcacta	gttgtttgca	8700
ttttaagaaa	agtggggagg	ggccgggtgg	ctcaagcctg	taatcccagc	attttgggag	8760
gccaaggcag	tggatcactt	gaggtcagga	gttcgagacc	agtctggcca	acatggtgaa	8820
acctgtctc	tactaaaaat	gcaaaaatta	gccaggcgtg	gtggcacatg	cctgtaatcc	8880
cagctactca	gaaggtctgag	gcacaagaat	cgcttacttg	agccccagg	acggaggttg	8940
gagttagacc	agatcgacc	acgcaccact	gcactccagc	ctgggcgaca	gagcaagact	9000
gtgtctcaaa	aaaaaaaaaa	gaaagaaaga	aagaaagaaa	aaagtgggtg	gatactgact	9060
tgtgatttaa	cttagtcaag	gttgtcctgt	ccactattct	tgaggaaaac	ctcaagtggg	9120
cccaatgaat	ttctcagcag	aatgaatctt	tggcctttgt	tatttttagct	agcaataaca	9180
tttataacta	cctataactt	taaaaattac	aattaaaaaa	tgtttatttg	ggaggctggg	9240
gtggaaggat	catttgagcc	caggagtctg	agaccagcct	gggcaacgtt	gtgagacccc	9300
gtcgtacatc	aaaagttttt	atttttaatt	tcactttcat	gacttggcta	tcaagtctgg	9360
cttttgcaaa	aattaagaca	taagaaaaga	atgcttcagc	tatgaattac	tatcaattgt	9420
tcaaaaatac	catcaactct	caaaaattat	cataaaatac	accaaataa	ttaacaacgg	9480
ctttgcccga	ggtgggggag	gaggagggaat	agattatctt	ctagtatttt	ccaaatgttc	9540
tatatataac	atataattaac	ctttaaaaca	tctacttttg	tttgattctc	aaaataatat	9600
aaaacactac	tataataatt	aaaaagaaca	ttctaattct	aataatttca	taaaaggagg	9660
tcacagttca	aattgtaggc	aactataaaa	atctcgtctc	tgaacaacca	atgaacatat	9720
acatgatattg	aaggaaaaaat	ccctaagaaa	aagcagtctt	ctaattaaag	agaaccttga	9780
aattaagtaa	atcaattcct	gacagaaaga	cgaagatgtt	ttctgtaata	caagaaagca	9840
agatcacctt	tgcctcagac	atctaattgt	agtgtttaa	cgttcgaatt	ctggaataaa	9900
aaactcagca	agtcctaaag	tatgactctg	ggtgccaaga	aaatgccaca	ggaactagca	9960
tttccaatca	gcagctcctg	agatcaggaa	gactgttatg	ttctatgata	taaagtccac	10020
aataaaatct	gttagttttt	ctggttaaat	gctcatgcta	aaaatagtga	ctgctcaaat	10080
attaagtaag	aagacttagt	tttgcttctc	tgttcagtc	tctgaattcc	aggcaattgg	10140
ttttcgatat	cttgtgacac	caatacttga	catctaacag	cattttgtcc	actactgcag	10200
atgcactgcc	gagtcatcct	ttccacccct	tcacaggcat	atatttgtgc	tgcaagggtc	10260
aagtgttgag	gagctcagga	ttataaataa	cgaagaaac	gagaagcagc	ctttctttgc	10320
tgtctcacc	tactcatag	gaagtataaa	cgactttagc	atccatctgg	ccgatctcat	10380
ttcacaggct	gcagaatcac	ctaacccttt	ccacctgcaa	agcttgtcac	tctctccttc	10440
cttagaatct	cacagctgag	tatgttttca	gaactgttct	tagacacaga	tcatttacta	10500
tttattctca	tcaaaatctg	aaacagctat	gcgagagggt	ccaaactcat	gaaacctaaa	10560
acaaccatca	gttcatcgaa	gcagctggga	aaatcttttc	gagacaacat	caactgcttt	10620
tgttcatgag	attaaaaaaa	aaaaattcat	actgaccaga	aacccaagca	cgctggaaac	10680
agccaaccat	taacgatgac	ctttgccttg	gaaaccatga	gcaaaaattc	cccttggttt	10740
cccttatatt	tcctttggaa	aaaaaaagga	acaatgcaac	agactaggct	ggtttctact	10800
tgtgatcact	tacaaggcca	gctgttcttc	ctccatgttc	ctacactgat	aagaatcagg	10860
gactcctgct	ctacgcagta	agtccaggat	gcattgattg	gggcccctgga	acactctgcc	10920
tctgttcccc	cacgacaatc	aagtaacagg	catttactgt	aaaaagcaag	actggaagct	10980
gcaggggaagc	ccaagtagca	gcgcattatc	ccgaagctgt	gagatcacc	tgctctctgc	11040
aaatacagtc	aggagatata	gccagaggaa	accgaacgac	atgactctcc	gggtgggggg	11100
tgggggtggga	ggccgcagag	catggtcagt	cacaggattt	atgaaaacaa	gatgcagaaa	11160
gtctctgtga	cccggcttcc	tggcttctct	tctgagctca	ctctgggccc	agagcctcat	11220
gcgcccctctg	cgtgctgac	ctgaatactg	tatctgacga	ctgcagcttc	tgatgccag	11280
aggcacaggc	tcccgaattca	tcagaccctc	aaagtgtccc	actgggggaag	tccatgaaga	11340
aatccacatt	ggtgatggca	cgctcacttt	accagggtgc	tggggccagg	aagcccaaac	11400
ccacaagcca	tccatcccag	ccaccagaaa	gtcactcttc	tcacaaaaga	tctgagtgtc	11460

ctaaaaggag	tgactaaagt	tacaaaaggt	cagacgcaga	cagacaaaac	ggaaatgtct	11520
tcctccaccg	ctgtaagaaa	aatcttgatg	agggataaaa	aaaaaaaaag	ccgctgcctt	11580
ctctaccgcg	caactggaat	gtttttatct	ccaccacaca	gatctgttct	cggacactga	11640
ttactgccat	tcgggaagct	tcataagatt	aaagttttct	caaagcattg	aagacagaca	11700
aaaaacctca	atcaatgctc	ctcaaaaaac	cccaggcccc	caaaatataa	acagccagtg	11760
tcattccagaa	accaagccat	ggcaggaaaac	cagtaatcag	gggtggtcata	cgtactaatt	11820
tgagctggaa	acctctggac	agcagaagca	gtgggttggc	tgaaggaaga	tgacagaagtc	11880
ggtaaaaataa	aagaggttcg	tggtctcagt	gtcacatctc	ctaacgctcc	ctacaactgc	11940
cctccgagct	ctggccatct	gtcccttatg	gagatcagga	aaagccagga	ggctgccgag	12000
tgcttccacg	agggtctggg	agccaactcc	tcctcagagt	cctaccgaa	aagcaaatgg	12060
ctcttgtaga	actcttgctc	tcctctgata	ttttggctga	aaaaggccct	tgtcccagca	12120
catcctgatg	aaagagggcc	attcagcaaa	acagctgagg	ttcctctaatt	cactgcactc	12180
ctacgggctt	ttctgtaggc	cggagaaaaca	agcacggggg	tgtgcattcg	acattgtgag	12240
ggcaacaac	tggccccaag	gaaccaaccc	caagcaacaa	gaccccttc	cgattcaaat	12300
caacattctg	aaggtgactc	ctttctttca	aatcagcatc	catttaccga	acgggtgacgg	12360
tgacgtgggc	agctgccgca	gttagttatt	ctgcgtactc	aaagcacggg	tacatcctga	12420
aaattcttca	gtcattgctaa	cagctatctg	aggggacacg	ccaggtagag	gggaccacat	12480
gcacacctat	gaggagctct	gggatacga	cgggtgccca	ggcaggtcag	gctgcaaagg	12540
tcctaaaggt	tggaggtgtg	atcccaaac	ctccaggcac	aagccagcca	agagctgtgt	12600
ttttagcggt	tccttcagtg	agagaaataa	gttcaggatg	tgaataacca	tgacgcagga	12660
gagaatggaa	taagtaccct	aagaaagggg	ctcggtcagg	gtttacaaga	gggaggaggg	12720
agcattttaac	tggtagcttc	tggaacaatt	cctgaaggaa	gcagcactga	gtaggggctt	12780
ctctccctc	ggctcacaa	tgaccaagcg	atcctcccta	cggattaagt	gaaacacaca	12840
ttaccatgat	tctggttttg	caggtgagga	aacccagct	tgectaggag	cacatatctc	12900
tacaagatgg	ggctggactc	acatctatct	gccccacgce	cacctgctta	acctctgtta	12960
agcagctgtt	ctactcatcc	agaatgaaa	tcagagccat	tatgctgcgg	tcacatccgc	13020
tcattgcctgc	ccaggtgcct	aatggcaaa	ccactaaggc	actgagaagt	cagaatgtgg	13080
atcacatctt	ccgtccttct	tcccagtggt	tgaatgcata	atgctgtggg	aagagagaga	13140
aggaaccatt	caagcaaa	gaactccagg	aagacgagac	tgtgccgggg	ttcttccatc	13200
tgcccaagta	gaatcagaa	gggcaggggg	cccacagcct	tatcctaccc	accactgcgg	13260
tcattagttg	ggcaacaggc	acatcctttg	gccttcttgc	actgcataga	ggctaaggag	13320
ttctgtaaac	cacacagcca	cgctgaccaa	gaagtgcgtt	tcaaggtaag	tttctcatca	13380
acaggactat	tatttactga	ggatctccca	tgtggccaag	gctgtaggag	gtacttagct	13440
acgccacgtc	attgaactct	ggcagttctg	cagggttaagg	tattttctcc	atatgacaaa	13500
cgaagggaagc	ccgtcaacaa	ttccaaaata	gaatcaccag	ggatagcatg	gacaacgccc	13560
atggtgactg	ccgcgcttta	agggttaaga	aaagtaaaaa	ctgggggtga	tgactcattc	13620
ctgtaatccc	agcacttttg	gaggctgaga	tgggtggatc	atttgaggtc	aggagttcga	13680
gactagcatg	cactacatgg	caaaaccctg	tcttacttaa	aaatacaaaa	attagccagg	13740
tgtggtgggtg	catgcctgta	atcccagcta	ctcaggaggc	tgaggcagag	aatcacttga	13800
accaggaggg	cggaggttgc	agtgcagcaa	gatgcaccca	ctgtactcca	gcctgggcga	13860
caaagtgaga	cactgtcttg	ggcaggggcg	gtggggaaca	aaagtaaaaa	caatggctctg	13920
ggaattcata	ttctcggggt	ccaatttaca	ttctaccata	tatactctga	tttaaccctta	13980
gaattaaccc	ctagaattcc	ttacagggtt	ctgttcattc	atccaaacag	gcaaacattt	14040
gcagagcatg	gagcacaggg	taagccaagc	cagcccaagc	tctgataagg	gcaaagacag	14100
ccatcctctt	ttaggaatgg	gtatatgtgc	tgggtctctg	gggtgtctgcc	ctgtgcata	14160
gaaacagcat	ttcttgaaga	acaaaaatag	taggtataga	aacatcacag	tatggaatat	14220
ccaaacaccc	ctgaattcca	actctgggtc	tacattgaaa	caacctatca	aactcctaaa	14280
acacattcat	gccaggttcc	agcctcagca	gagtctaatt	cggaaggctc	gtgatgagtc	14340
ctgggcatct	acttttttaa	aaagtccag	ggagctgggc	atggtagctc	atgcctgtaa	14400
tcacagcact	ttgggagggtc	aaagtggggg	aatcagttga	ccctggaggt	tcaagattaa	14460
cctgggcaac	gtaacaagat	cccatctcta	caaaaaata	aaaataaaat	tagctaggct	14520
tggtggtgtg	tgctgttagt	cccagctgct	caggaggctg	agggtggaga	atcacttgag	14580
cctggtgagg	tcaaggctac	agtgcagctg	gaccacacca	ctgcactcca	acctgggaga	14640
cagatcttgt	ctccagaaag	ttccaggggg	tgttcttgat	gcacagccaa	gttttaaaaa	14700
cctcagaatc	aaataacatc	atggccaggc	atggtggctc	acgcctgtaa	tcctggcact	14760
ttgggaggcc	aagtggggtg	gatcacttga	ggtcaggagt	tcaagaccag	cctggaaaac	14820
atggtgaaac	ccagtcttta	ctaaaaatac	aaaaattagc	tgagcgtggg	gacgcacact	14880
tgtagtccca	gcttcttggg	aagctgaggc	acgagaatca	cttgtaccct	ggaggtcgag	14940
gtgcactga	tgggagattg	tgatcctgga	gtcccactg	cactccagcc	tggatgagag	15000
tgagactgtc	tcaaaaacaa	acacacaaac	aaacaacatc	agaagacaca	gagaaaaacag	15060
tcttctccat	gggcttcata	aagatacctc	tcacataggt	acacgtcgat	gttttctgct	15120
ggtaaaaggt	aacaccaaca	aaaaggcatg	gtgctctcag	aaggtgggtg	atgtgattag	15180
gtgcaataaa	gggagggtcat	gctagggtca	aaaacaaaat	aatactctct	ttggaagcag	15240

taaaacagat	gctagttcttc	tactacacac	tttcagagac	ctgaatgttc	ttctggccct	15300
ctaagggaga	cgctgcatca	tgacaatacg	aaatgatgac	agtgaagca	aaaacagatc	15360
agacctgtgc	tgtgtgaaac	agacatgggg	tctcgctatg	ttgcccaggc	tgggtctccaa	15420
ctcctgagct	caagcgattg	ttccgccttg	gcctcccaaa	gtgctggggg	gacagctgtg	15480
agccaccgag	accaacctca	gatcagacct	ttgacaaact	ctgctgtgga	caaagcattc	15540
tgggtgaatgt	caactcatct	gatcttcaca	aaaccgtgtg	gaagaccaga	caggcattat	15600
tacactaatt	tatgcctaag	gaaacaggga	gttaaatagt	acaaatttag	gattttctgat	15660
gctgtatctc	gaaaaaaaaa	tagagaatat	gagcctgaag	aagaggccct	gtaaaggggtc	15720
ccagattgat	gggacaggct	gagacaaacg	gaatcacttt	tccttggaata	gaactaaccc	15780
tcaatggtac	cccactctgc	atgggtgatta	ctgaggggac	tgtcaattgt	ccagcgaact	15840
tgatggtaat	tctaggagaa	aaaggaacta	atgtaatgct	gtcagcatag	aaagatgggt	15900
gccaacgagc	attccaaaaa	ggaggctctg	ttaattcggg	ttcgatcaac	aagtatttgc	15960
tgagtgtcta	ttgtgtccgg	tcagtgtctaa	ggcctgagaa	tttagaagtg	aaacagaccc	16020
ggtttccacc	catgccacag	accactccac	acctggtctg	gagtgcact	ggagggccag	16080
gcaggcacag	gacagtaact	tcgatataag	gcagcaagtt	ccacggtgga	aggaggtgga	16140
agggtcacag	gcaggtacac	acagggttcc	agggagccct	ccctggaaga	aatgaacct	16200
gcaggccct	gaaggatcag	taaacagaga	ggcataaggg	gcaggagagt	aagatgatta	16260
tgctacatgt	accttattgt	gaacccagga	ggatttggcc	tctgtcataa	aaggccccc	16320
tgtgggttca	taaacctcaa	tttacaatt	gtgctttata	tatcagttcc	ttataagttt	16380
ggttagcgta	aattggtttc	ttagaacttg	atcatccctg	agtgaactca	caaatttcaag	16440
tttcagaatg	tgcaaaccta	agaaacaaac	ctcatgcttg	tggttgagac	atcgactgt	16500
caacatcaca	aattctcagc	acctgaatgc	ctggtatact	atcaacatat	attgttttaa	16560
atatgtaaat	aatagctttc	tagttataga	gagttgttcc	ctacattttt	ccacttaatt	16620
tttacaatcc	cattccctcg	atgaaacaac	cccagcctgg	gcaacatggc	aaaatcctgc	16680
ctctacaaaa	aatacaaaaa	ttagctgggt	gtgggtggcg	gcacctgcag	ttccgggtat	16740
ttgggaggct	gaagtgggag	aatcacttga	gcccgggagg	cagaggttgc	agtgcagcaa	16800
gattgtaccg	ctgcactcca	gcctgggaga	cagagggaga	ctctgtcca	cccacccccg	16860
cctcccaaaa	gaaaagaaaa	gaaaagaaaag	aaaatgaaac	caccaagact	gggagaagat	16920
aaatgacttg	tctgtggtea	tctggctaatt	aagaggtaga	atggggctga	aaaagttcgg	16980
tgctcttctt	gaagaatcca	taggtcagaa	agcagacca	tctgacctgc	agcaatagca	17040
gcaacgtgga	aagctaatac	actgacctca	aaaccactct	cagtgcaggct	ctggatggat	17100
tcagaacccc	aggcctagca	aagtgaagtt	gataaagatg	taaaggagat	cgaaaattca	17160
ccatttgagg	agagattagc	taaagactgc	aggctggatg	gaaaattctt	ttccatggttc	17220
ttccacagggt	tcttccctca	tttggaactc	gtgttttaaaa	gtcacaaaaga	ccctgagttg	17280
ggccaagggtc	tcgttcttct	tactgtggg	ccttgacagt	caacatggca	gggcctcgtt	17340
ccaaatgtca	ctcttcagag	cctaagaaaa	caagtaactt	tagggacaca	cctgtcaacc	17400
ggagctccca	aattgtaccc	ccctaaacac	ataatgctga	gcatagaaaa	attccagctc	17460
tgacagcggt	tatacttagg	gaaaggggtc	acagacaagg	aatgctggca	gggctcatta	17520
caaatatctt	tgctgctgga	acatgtattg	tttggtctaga	aggcgtaggc	ttctctcaga	17580
gagaaggaat	gtccaaaagt	atttcagaca	gtaagagaca	ttctctgagc	cagctacaca	17640
gctctccttc	aaaccaacgg	gtagcgcaa	gcagctgaac	tgaccagcga	gctcgcaaaa	17700
gcaagctttt	tttttttttt	ctccctaaat	aagacagcaa	gtgatgtgtc	ttggcttggt	17760
ttagcaaat	ttaagatagt	tcctgatga	ccccagagc	cctcaggccc	catggaagct	17820
ggagctaag	catcttcttc	caagcatcat	ctgctctacc	aggatctaag	ccccttcacg	17880
agggcagaag	gtataaaggc	tgactgtgc	gggaaatgct	atggcagcaa	agacagccaa	17940
acacgccaga	aataacaggc	acatgaaggga	aatgtttctg	agacagctca	aaaattccga	18000
gaagagatta	ttccgactgt	cccaggttct	cagccctgtc	tatggtatgc	agccccatac	18060
cacagtcatt	tgtcaccgag	tcctaacttt	gtcagaggcc	cctcctttca	ggtctctcag	18120
gcaccaccca	gttctggccc	tcctcaccoc	cgtgagccag	gcgacatcca	agcagcccca	18180
cgggtgcacc	ggctctgtgc	tgcatctctc	gaatgtccct	gaaggccagg	gctggttgat	18240
tctctaccca	ctctctgtct	agtatgggag	ccactggcca	gatgtgttga	ctgaacactt	18300
aagatgcagt	aaagtgtgacc	aagaaactgg	gtttgtcatt	ttatttcatt	ttagttaatt	18360
taaatttaag	tttaatttagc	tacatgaggc	tatcagctgt	ggtataggac	agcagagctc	18420
cgggaagcttt	tggcctgggtg	agaagaatca	ggacaagcgc	ctccctggcc	tctcgccac	18480
tctgcacagc	cgctaaccat	tgctctcatg	acattctttc	ccagccccag	aacttttagc	18540
catgtgacat	catctattga	ttagagtcca	aacttcttgt	gctaactctc	tatgggttgc	18600
cacaatttagc	cattgtatgt	cgttaaccta	aatttcatte	atctgctatg	tcctgacctt	18660
aggggcttag	aatatagtta	gaaaacagta	tttcagaata	aaaaaccatt	cttgtattac	18720
ctctcgcact	attccccctg	ttctccatgc	ttcgcattct	ctgttctatc	cccagctata	18780
gcactgtccc	cataaagcct	actgtggttc	tcggctcatg	gtgtccttcc	ttccatctgc	18840
ctcccgacgt	catgcctgtc	ttccagtgtt	tatgccttct	ccaggaagct	ttctcttggtg	18900
gccctcgctg	tgagctatag	ctcctccctt	tcaacatctt	ctagcacctc	ctcttactgt	18960
gcagggtgagg	acactgaggc	ttaagggtta	agtcacttgc	tcaaggtcac	atacagtcgg	19020

tccctctgtat	ccgcggggttc	ctcatcagtg	gattcaacca	actacagaca	gaaaatacag	19080
tatttgaggg	atgctgaact	ctttgaatta	gtgggttctg	cgggtgctta	agcatccatg	19140
gattttgtta	tccctcgcaa	aggcggggg	cctgaaacca	atcccccttg	atactgaagg	19200
aaagaccacc	cttagtgata	ggaacctagg	aaccceaagt	ccctcatttc	caaactcgtg	19260
tccctgaccc	acttatttac	taactagtgg	tgaagccatc	ttcctgccag	tatattttta	19320
cttcacaatg	ggatgtgagg	gccaggatgc	acatgctttt	taaatctccc	tctgtgcttg	19380
acatacagta	gattgaaagt	aagtgtctgt	agatacactg	gccaaagctg	gctcttctct	19440
gaagtgcagta	ttccaggagt	aactcaccct	ggatcatctct	gtgccctggg	cacactgggc	19500
actccccaca	cacagggtga	acctggcaaa	taagactcac	agcatcatgc	cacgtgcgag	19560
ttaaagccac	ctggagggtca	ggtcagggtct	tccctgacaac	tgagtgtctc	aaataacaca	19620
acagcagcta	agttccccac	atcaccttga	gtgtctggag	agctaggcct	atgacttctc	19680
tgtctcagga	tccctctcag	tgcccagaaa	acagtggaca	tcaataaatg	taacaccaat	19740
aacatcttcg	ttgagcgcta	tgctaagcac	atcaggatag	ttaaactcatt	tattccccag	19800
tgtccatctc	tcagtgtttt	atacatacgg	gaactgaggc	tcaattagcc	gagcgtgggt	19860
tcgtgtcctc	gtaatcccag	ctattgggag	gcacaagaat	cactcgaacc	caggagatgg	19920
aggttgcagt	gagccgagat	tgtgccacag	cactgcaaca	gagtaagact	ccgtcttaaa	19980
aaaaaaaaaa	aacagaaaaa	aaaacaaaac	aacactgagg	ctcagggagg	ttaagtcacc	20040
tgcccaagtt	catgagacca	aggagccggg	aagcaggaag	gggaaggcag	gagtgttaact	20100
ctgaaacctc	tgctcttagg	cactggcttt	cagctgaact	gatacctctg	gaaaacagtc	20160
tcaaaaaagt	ccacttctcc	tcccaacaat	tcagacctaa	aaaccatttg	gcgggggaagg	20220
gcagggcaag	cttctgagtt	ggggaggggg	tgtgggatcc	caagctgagg	tgtctgtttg	20280
caagcagggg	gcaaagggca	tctgtgcagg	gagggggctg	caagggagac	agagactgct	20340
cacaggcaag	gaatgaaata	ttaaacatta	atgttaatat	taataatttat	aattaatata	20400
tttatgatata	atagcatata	tacatattat	attaattaat	tataactata	ttaatataat	20460
taattataac	tatattaata	taatcaatta	taactatatt	aatataatcg	attataacta	20520
tattaatata	atcgattata	actatattaa	tataatcgat	tataactata	ttaatataat	20580
cgattatatt	aatattaata	taatcgatta	tattaatatt	aatataatcg	attatattaa	20640
tattaatata	atcaatatta	acaaatatat	actatataat	ataaataata	cctaagttta	20700
tataatatgc	ataatgttaa	tatttattaa	tatttcaggg	acaatgggag	tcattgaatat	20760
ggagagacaa	aactagaatg	aaccccaagg	tgctgcatca	gaattgaagg	taccagctctg	20820
aactcatagt	tttcaaccta	ttgaaataaa	tatagatgca	cgtgtgtgtg	tatgcacgta	20880
catacaaaatg	ttccctaatt	ctgcccattg	agaggcctgt	ggtttagcaac	accccaacag	20940
caataagcag	acctagcttg	gctcctaatt	ttcattttcc	actaaaagga	accagagccc	21000
cttggtataaa	ggactgattc	cacagtgagg	tagggagcat	ctgttgccag	aaagcaagaa	21060
agcacttaaa	gaatgatgtg	gacatgtcaa	agggacacag	aagccagcct	ggatgagatc	21120
ccactggccc	taactgtcca	caaggacaat	ttgagcaagg	atgtcaacaa	tttaagagca	21180
gattataatga	taactgaataa	aacagaaaaa	tacaagaagt	tgaaacggac	attgatggca	21240
gacaggatat	taacataaatt	ttaaaagtac	tctccaagga	atgcttctga	atgatgaagg	21300
ggaaaaagaat	aactgtacag	tgaaaaagcc	tggtaaaacc	caccttagtg	accaaagtga	21360
atgtcaccat	agtgggacaa	aaggaaatca	agtgccacct	tatgggattc	aacgaggacg	21420
cagcatccct	tgggtgatgt	tccagccaaa	tacacgtgcc	cgggtggaatc	acacaagaac	21480
atcagacaca	ctcacactga	gggacactct	gcaaactgac	agtactgggc	acaaacatgt	21540
ccaggtcatg	gtcgaccgca	gtggctcatg	cctgtaatcc	cagcattttg	ggaggctgag	21600
gtgggaggat	caactgaggt	caggggttcg	agaccagcct	ggccaacatg	gcaacacctt	21660
attctctact	aaaaatacaa	aaattagccg	agcgtgtgtg	agcatgcccg	taattccagc	21720
tacttggggc	gctaaggcac	aagaatcgct	tgaaccgggg	aggtggaggt	tgacgcgagc	21780
tgagatatca	ccgctgcact	ccagcttggg	cgacagagtg	agtttccaac	tcaaaaaata	21840
aaaaaataaa	ataaaatcca	ggccacaaga	gtcaaagaaa	gactgaggaa	ggttccagac	21900
tgacaggagag	ccaagagaca	ggataactag	atgcaatggg	cagtcctgaa	ttggatcttt	21960
tgttatgaag	gacaacgctg	ggacatatgg	tgactcttga	atgggggttag	aggactagac	22020
ggtgggaatg	catcagagtc	agtgtcccgc	gtggatggct	gtgttgcggt	tctgtgggag	22080
aatgccctgg	tctgtattcc	aagggtaatg	gagtagcagg	ttgacaaatt	actttcaaat	22140
ggttcaaaaa	agaaagttct	tttactgtga	cttgcaattc	ttatgtaagc	tggaatttat	22200
ctcaaaatta	acgagaattt	tttatcgacg	tagtatttta	catattttatg	gaaaacatgt	22260
aagtatttgt	tacatgcata	aactgtgtaa	tgaccaagtc	agagtatctg	gggtatccat	22320
gaccttgagt	attaatcatt	tgtatgtgtt	gggagcatta	caagttttcg	agttaccaat	22380
tttttttttt	ttcctttgag	acagggctct	actctgtcgc	ccaggctgga	gtgcagtggtg	22440
acgaccacgg	ctcacgcagc	acagcctcca	cctcccaggc	tcaagcgatc	cttccacctc	22500
aaccacccaa	gtagctggga	ctacaggtgt	gtgctgccac	ccccagctaa	ttttttaatt	22560
tttttgtaga	gacaggggtct	cactatgctg	ccagggtctg	tactgaactc	ctaggctcaa	22620
gagatcctcc	cacctcggtc	tcccaaagtg	ctgggatcat	aggcatgagc	caccataccc	22680
agccaaattt	tttaaagtta	ttttttaaat	ctccacttaa	ttcgattttg	gtaaaacacg	22740
acctgtaatt	tttctttatc	ggtaggtaat	aaaagcttca	gatgatattta	ctgatcactg	22800

gtatgggcat	atttcatgac	tttgcccttt	catctcttgc	atagttttac	cctcaccaag	22860
caagaccttc	cctgcctcag	cactgtttgc	cctcttcctg	ttttccagaa	cagaagtggc	22920
cctgtttcgt	gcccagagca	gaagagaacg	atgaagagct	ctgctctccc	aggtcttcc	22980
ggctgtgtg	tgtccaggtt	ttgagggcct	ctcacataca	cggctctgga	ccacgtaaga	23040
tctaatttta	gcattttcct	gctcggagac	cacaatgttt	ggaacagcag	gggtgacct	23100
gcccgtgcag	gcctcctatt	gtgaaggcca	cgcgaagcca	ggataccgca	gcccgtcagg	23160
atgtgactca	gcacctgtc	tcagtgtctg	ggcggccagc	agctctggca	ccaagtgtcg	23220
ctgtgacct	cacctcttaa	gaccacaaat	acccagggtg	attgggtggg	taggcataga	23280
gcatacacctc	tccctgttaa	gacaacttgc	ttgtccatcc	attatgctgg	gcttccttgt	23340
gaacaccaca	ggtatctatc	aggaagagtt	cttccgagga	actgatctgc	tggtattttc	23400
aggacaccaa	gaatcaagag	attggctctg	tttctctctt	tgctttgact	accaggaaac	23460
tcaaagtcag	atctgtggcc	aaattcttgt	aaccatacca	atgctatgtc	atgtattaca	23520
tgtacaaacc	ttccctctac	ttcatcttat	tttctctctg	tttcttctgt	tcccgatttt	23580
ctcactaatg	ttacattcta	ttgttctcta	tgaatgttgt	aagggtgttc	aaatcctttt	23640
tggagggcac	tactgtagat	acaacacaca	ttaccctga	gggataagga	ctctttttga	23700
ctccacacag	aatccctggc	atttggcaaa	gaacccatat	ttaggcacta	aatacacatg	23760
ggctgaatgg	aaaaagccaa	tagctaagta	aaaaccacct	ccattaccat	attgtttcac	23820
aagaggttct	tttcccttcc	atctcatgag	gtggggcctg	gtcaggagtc	cccagggcct	23880
gggaattagg	ttccttaggg	agccttcttg	ctgtaggggc	agccaacagg	tcagtggcct	23940
tgactccaga	cctaaagagc	cactcctaga	ctcccagctg	caacagacac	agcgtggcac	24000
gggtgggcct	ggccactggg	gaagtgacaa	gtgatttcca	gatgctgcag	ccagcctggc	24060
tctttccaga	ccacactgaa	ggcccttccc	tgtgggaatt	ctgatggggc	ccagatttgg	24120
ggaaaacacgc	ctcagggact	cttggcaagt	gcgtggccagg	cctggaccag	gaatgacttc	24180
tgtgggcaca	gggagagacc	aggcatttcc	taacacagga	ccttgaacag	ccttctctga	24240
aacaaagtct	ttctaaaaat	agcttcaaaa	gtaaccattc	aagaaaagaa	agaaaaaaaa	24300
aactgtaaaa	gtaaaggcac	tcaagaatga	tatttcccag	ataaaagcct	ggcacagggt	24360
tcagaaggaa	ttgcaggaaa	acaggtcaag	gctgggtttt	tcctcttagg	tgctcatttg	24420
ttaacattgg	tctttggagg	ggaacaagt	cggcagggaag	ggctggcact	gaaaatgatg	24480
gccactgggt	ataggccagg	gccagacact	gtacacagaa	caagactctc	tgagggcctc	24540
aggagggccc	tgagaggagg	aaggcaggtg	gtggggccag	ggtcagacat	gcaagtgcag	24600
taagtggcaa	ggccgatgcc	ccatccagaa	gctccgctct	gaccacacgc	aggtcttccc	24660
ggcatgtcct	catttatgcg	gcagtctctt	gtatctcact	gcaattctgc	ccccacactg	24720
caggctggcc	agcgtggcct	cctcataagc	acatcaccct	gcaccccgac	actgactaca	24780
cccacaaagc	aggagccccc	gcaccctcca	gcccacatgc	tcagttcgct	ttgaaaatgg	24840
ctcctctcgg	gggctgggcg	cagtggctca	tgcctgtaat	cccagcactt	tgaggaggccg	24900
aggtggttgg	attgcttgtg	gtcaggaatt	caagaccagc	ctggccaaca	tggtgaaacc	24960
ccatctctac	taaaaatata	aaaaattcgc	caggcatggg	ggtacaagcc	tatagtccta	25020
gctacccaag	aggctgaggc	aggagaatca	cttgaaccga	ggaggcagag	gttgctagtt	25080
gccgagatcg	tgccactgca	ctctagcctg	ggtgacggag	caagactgtc	tcaaaaaaaa	25140
gaaaaaaaaa	aagaaaatgg	ctcctctgga	ttttgattaa	tcctattttg	attaatcctg	25200
gtttctcatt	ttcagccttc	cttgaagcag	catgacccat	ctggatgtcc	tcctcatctc	25260
aggaattttc	taataagctg	tctaaatcca	gagatccgac	cacagaacaa	tgaatgccaa	25320
agatgagttc	taaagatgcg	agtactttct	ttctaaacgg	acgctgcttt	gtgtatggct	25380
ctgtcctcgg	ggcagacgc	ggcaggctaa	gcccctggga	ggaggaggtg	agtcacagca	25440
gaggttcact	tcctctcagt	agcccgctg	gttttctcca	ctgcagggtc	agacatagc	25500
cctgaccacg	ctagaccccc	ataagcgcat	gacctgtctc	tcaccgtggg	aataaaactc	25560
gtgatagtca	gttacaataa	cacagcaaat	gatgagcagc	acaatataaa	cacagatcta	25620
gattggtggg	tctgaggact	cattcttaaa	tttggaggcc	atcaccta	cttgtctttt	25680
cactttacat	agcaggagac	agggacccag	agaagtgaag	aggcgttgcc	ttaggttgca	25740
cagcagatga	cgcctctcaa	gatggaccct	aggttgtctg	actccgtctc	acagctttgc	25800
cccatttatc	atgaagatga	acgctggtaa	cactgctacc	tacgagctga	gcttgcacgc	25860
acattcctgg	tgtgtacatg	catgctgca	cgtcacgca	atgtgctaag	tgcacaggaa	25920
ggagaccaga	gcccgtaggc	gttcttttga	agtctaagta	ctgggtgttc	gaaagttaaa	25980
tgaacacctac	tagactctga	gcaaaattcg	ttttacgtta	accttaatga	aaagttaaat	26040
taagtcttga	cagaattaac	tcttcacgtc	tctgtctcca	tttgtcccca	ttctagaatg	26100
agttttctaa	ttaaaaaaaa	tatatagggc	cgggtgcagt	ggctcacgcc	tgtaatccca	26160
gcactttggg	aggccgaggc	gagtggatca	cctgaggctca	ggagttcgag	accaacctgg	26220
ccaacatggg	gaaaccccg	ctctactaaa	aatacgaaaa	attagctagg	ggtgggtggc	26280
catgctctga	atccacagct	ctcgggaggc	tgaggcagga	gaatcactgg	aacctggag	26340
gcagagggtg	cagtgaacca	agatcgtgcc	actgcactcc	agcctggtga	cagagcaagt	26400
actcccaccc	ccccccacaa	aaaaaaagta	tatatgtgtg	tgtgtgtata	tatatatata	26460
tagctaggca	cagtggctca	tgcctggaat	cccagcactt	tgaggaggccg	atgtgggcag	26520
atcacttgag	tccaggaggt	caagatcagc	ctgggcaaca	cagtgaagcc	ctgtctctac	26580

caaaaataca	aggtggtgtg	cacctgtggt	cccagctact	tgggaggctg	aggtgggagg	26640
accaattgag	cccaggagggt	cggggctgca	gtgagctgta	atcatgccac	tgtactccag	26700
tctgggcaac	agagcaagac	tctgtctcaa	aaagaagaaa	agagagagag	agggaaaaaa	26760
aattgaaggc	aaattctgat	tttcaaatca	aacgtttcaa	caaactgcag	aaataaaacc	26820
cgagttaaac	caaaaggaac	agccaaacag	cacaatgacc	ccaatgttta	aatatgcccc	26880
aatgtttaaa	agtgggagtc	aatgggaggc	cactacctac	aaggccacag	gggttagggc	26940
aggactcagg	tcctgaatc	acagcagcct	gcattcaaac	cctggctcag	gcctcccacc	27000
agcctcgtgg	aactggtttc	ctaaaaatgag	gagagtcctc	actttgcagg	cttgtgacaa	27060
caagatgaca	gcaagtgcga	aagttccaag	cccagagcct	gcagcctgca	gaagctggcc	27120
tcattaccac	ccggatgttc	tcgggctgct	agcacatgaa	ggggatacgt	gacaatccct	27180
gctttaagta	cagctcaggg	agttgacggg	acctgcccga	gcacatagtg	atgccgctaa	27240
tggctcacca	ggaagaatgg	actgcaaagc	ctggttcttc	tgataaactc	cattctgtct	27300
cccagtgtgg	gttctgatgc	atagggagga	ggaaaagaca	gtgcttggtg	tttggggtga	27360
agagcacagg	ttttggagtc	aatgagacat	ggagtatgag	ggtctcagct	ctaccgttta	27420
ctactaaata	aaaacaggcc	actgacctct	ctgggggtta	gtcttctcct	ccagggaaatg	27480
ggaattcaaa	tgtctttaca	gggttttcac	aaagattaac	tgaaataatg	cacacaaggc	27540
aatcacagag	tggagtattg	gtgctccctt	ttctctcctc	catccctgct	ttattttttc	27600
gcctgggcac	ttaccaaacac	acgattattg	cgcttggtta	ttttattttac	tgtcttgtct	27660
cctcaacaga	atgtcagctt	ccagagcagg	aatttttatt	ttgtttgttg	ctatatccct	27720
agccctaaa	acagggtctt	gcacacagta	ggagctcaaa	aaatatttgt	tgaatgaata	27780
gctcacaagc	agacagatga	ggacagaggg	gtcttgagac	tgatctaaca	gcaccgatat	27840
tactaaactg	caacggaggc	aacggtgagg	agaatttctc	tgtcctttgt	ttcctgaaag	27900
tccaagacca	cttttagttg	ctcaacagga	aacaatactc	aacttacaag	acctctaggg	27960
cctatccagg	gcaaaactgg	cactgtgagg	caggaggtca	ggcagccctg	tccttagggt	28020
ggctcacggg	ctagtgggca	gggccagctt	cttcataatg	gctcagaggg	gccccgtgct	28080
tggtttaata	ctctgttggg	gccatcttga	aattcttaat	aatgtttgtt	gttgttgttt	28140
gtttgttggg	ttgagacaga	gtctcactct	gtcgccagg	gtggaatgca	gtgggtgtgat	28200
ctcagctcac	tgtaacctcc	acctcccggg	ttccagtgat	tctcctgcct	cagcctccca	28260
agtagctggg	attacaggca	cgcgccacca	tactcggcta	atttttgtat	tttttagcaga	28320
gacaggggtt	caccatgttg	cccaggatgg	tctcaaaact	ctgacctcaa	gtgatccgcc	28380
cgctcgggca	tcccaagtg	ctaggattac	aggcttgagc	cactgagccc	agcctcttaa	28440
taatgttttt	aaaagggggt	ctcccatgtt	cattttgcac	tgggcttcac	aaattacgca	28500
gccagtcctg	cattacagga	aatatttctg	tacctaaagta	catatactac	aaagcaagta	28560
ccaacacca	aggaacacat	aaggagagaa	aaacgcctgt	gagaagaaaa	aggaagacac	28620
gaatcattec	caacagaagc	tgttaccatg	aaggaaagta	gggcaggggc	atgtgttgaa	28680
tgtctactat	gggagaaggg	gttcgcatca	tgagcacatt	taattctgac	aaccacccca	28740
caagctgtgt	actatactgg	ccatttgaaa	ctaaggcctg	cccgagatca	tataatagcc	28800
taggaggtga	caaaagacag	acacaggagc	caaaccatg	cccatccctc	cctaagccca	28860
aaatcataga	aaaaaaaaaa	taagaatcaa	catgggcggt	tatttttaag	gccagcatgt	28920
tcaaggtggg	ggcaaatcca	agagacacta	agcctcagag	catgaacaag	catgtgggtg	28980
ctgagtggag	gggaccagtg	tttaccaggg	tgatgtcaga	ctctgcaagg	ctcgctcccc	29040
gtgtttcttg	tctcttccca	tgagcaccag	gcacccctta	ccatcccca	actaggcaca	29100
tctgtaacgc	tgaatggaag	cctacttgtt	tacatgtgtt	ctatgttaga	ctgggggcat	29160
ccctagaaca	cacacagatt	gactggtggg	cagaattctg	ctaggtgcat	gcacccgagt	29220
gagcctttct	ctttgaatgt	gggagggacc	agagaacaag	atgggagagc	tgttccctta	29280
attaggtctg	gctgcacatt	aaaggcggtg	agacagtcac	tccagtata	acaatctgtc	29340
ataagacct	acagaagcag	actctcctgt	tggccttgaa	gaagcaagca	ccacgaattc	29400
tccacagctg	caagaaaatg	aattcaggcc	aggcgtgggtg	gtcacgcct	gtaatccag	29460
cactttggga	ggctgaggcg	ggtggatcac	ctgaagtccg	gagtttgaga	ccagcctgac	29520
caatacggtg	aaaccccatc	tctactaaaa	atacaaaaat	ttgcagggtg	cacctacagt	29580
cccagctact	cgggaggctg	agacaggaca	aaaatttgca	gggtacacct	acagtcacag	29640
ctactcggga	ggctgagaca	ggacaaaaat	ttgcagggtg	cacctacagt	cccagctact	29700
cgggaggctg	agacaggaga	attacttgaa	cccaggaggc	agaggctgca	gtgagccgag	29760
atcgccacc	tgcactccag	cctgggcaac	agagcagaaa	aaaaaaaaaa	aaagtaaaaa	29820
aaaaaagaaa	atgaattcag	ccaagaacca	cgtgagctta	aaagaggacc	ctgggggttca	29880
gacaagacct	cagccccggc	cagcaagcct	tgtgagtcc	cgaacagaga	acccagctat	29940
accgtgtcca	gattcctgac	ccatggaagc	tgtgagataa	taaacatggc	ctgggtgcgg	30000
tggctggcca	ggcatgatgg	ctcatgcctg	taatcccagc	actttgggag	gccaaaggcg	30060
gcagatcacc	tgagttcagg	tgctcgagac	caacctgccc	gacatgatga	aacctgtctc	30120
ctactaaaaa	tacaaaattg	gccaggtgtg	gtggatatgc	cctgtaatcc	cagatacttg	30180
ggaggtgag	gcaggagaat	cgcttgaacc	cgggaggcgg	aggttgcagc	gagctgagat	30240
tgcgccattg	cactctagcc	tgagcaacgt	gagcgaaact	ccatttcaa	tttaaacaaa	30300
ataaacatat	attgtttaac	tgttaaagta	gtggtaactt	gtcatgcagc	aggcaatgac	30360

tgatacagta	acctatgcac	acatccatct	ccagtcacga	cacagaactt	ggatgcacgg	30420
ggtgcatgac	acctcttggc	aggacttaac	tggacagaca	agcaacaaag	acaataaagc	30480
ccaggctaag	atggactgcc	aagggcaggg	aggaacccca	gagtggtggac	aggtgcaaaa	30540
gtaggggtgt	tcaatgaaga	ggggaagcat	ggtctgcagg	gcaatgacat	gccaaccccc	30600
atccactctg	acactgtagg	ggagggggtg	aaggcaaaac	cacacttcaa	aaggctgtag	30660
ggagaatggg	gtccctgggg	gacttccaag	tggagaccaa	aaggggaagg	gagtgcgagg	30720
agaaaggcag	aggagtcagg	gagttcacag	tttaccactg	aaaccaaata	aaacagaaga	30780
gacaaaatcc	tgcagctcgc	tctggcccaa	acctttgcta	gggcaggcaa	tcacaaatga	30840
gcaaatata	ataattctaa	tgaccacgtt	cccgcaattg	acttggaat	gctggattaa	30900
aaaaaaaaaa	cttactcct	gatccacacc	ctggggacaa	tattatctcc	ccagtgtcct	30960
acctagccca	caactactta	tgtctcatgc	cagactgagc	cagctcccgg	gatggcaagg	31020
gagccaggag	ctgctgccag	cagggccatc	tgctcaccaa	ttcccacagt	ctgaacggca	31080
cagcttccaa	agagggacta	cgagcgccca	gcagcagcct	gcacatgcag	aaggcagggg	31140
agagcgaggg	aaatggatct	atactgctct	gctgcaatca	tctgcatgct	gggtgtgaga	31200
tgatcagttc	ttgagacact	tcccagaagg	cttcagaaa	tactgtctga	gttacaacac	31260
tgcttctccc	aagtctgtat	tcttatttgc	atcttatagg	aatgtagccg	ggtaaggag	31320
gaaggctgct	tcaagtcaaa	gggcatccat	ggtggcgccc	ctctcaggcc	tggaccacgc	31380
acctgcagga	gtcgccccc	ttaattctcc	tctgccgtga	actaacactg	cacatcagca	31440
atactttgtg	aagaccgagc	acagcaacca	agcccactgt	ggacctgatt	ctaggcaagg	31500
aacttttttt	ttttttgaga	caggtctctg	ctgtcaccca	ggatggagtg	cagcggcaca	31560
atctcagctc	actgcagctc	cgacctcctg	ggttcaagtg	attctcctac	ctcaacctac	31620
ttgagtagct	gggactacag	gcgtgcgcca	ccatgccccag	ctaatttttc	tatttttttt	31680
gcggagatgg	ggtcttgcca	tggtgtgtag	gctggtctca	aactacgggg	ctcaaagcaa	31740
tccactcacc	ttggcttccc	aaagtattgg	gattacaggc	gtgagccact	ggggctggcc	31800
tagacaagga	acatcacaca	actactccac	aacctgaaa	ggtcacagtg	tcactcctgt	31860
tttatagggtg	gaacaattga	gaccacacaga	gctgtaagaa	cacacaaagg	taaaggaaact	31920
cagccaccag	cacaggagcc	agatgccaaa	cttaggtctg	cctgactcag	aacccactg	31980
ccttccctct	acacctggct	gtttctccta	catgtctgga	atttactgca	gggtcaaagg	32040
ttcatccatt	taaaactgttc	acttttatec	acttacttat	tttgagacag	agtctcgctc	32100
tgttggcccag	gctggagtg	agtgcgcaa	actcggtcca	ctgcaacctc	ccgggttcaa	32160
acaattctcc	tgctcagcc	tcccagtagc	ctgggattac	aggagcgcac	caccaggtccg	32220
gctgattttt	gtatttttag	tagaaagggg	gtttcaccat	gttggccagg	ctgggtctcga	32280
actccggagc	tccagtgatc	cgtcccgcct	tggccttcca	aagtgtctgg	attagagatg	32340
tgagccaccg	tgccagcca	tttaaactgt	ttaaatgcta	cacaaaggca	gagaaatgag	32400
gccgtcacta	agggatttga	gagcagttag	ggatacaaca	agggcacaca	gacctgcatt	32460
gtaaggcggg	tgtggcacct	gtcacagata	gggatgccag	gggctccctg	ctttctctga	32520
agagaggga	atcacaata	tctggggcag	gcgcactttt	agctggtcat	gaggactaca	32580
gccaggtgaa	aaggaaactgg	cctagggaac	gtggttgacg	ggggagcagg	gagtggtccc	32640
aatggactgg	aaaaggcaca	tgcgagaggg	gaggttgga	aggccacca	cgccggtgac	32700
gctggagctc	agaaaagact	ccgaggacca	gaagggaaga	gcatcaaggg	accagggggt	32760
gatatgccaa	ggtagagagg	atgggtctga	ggtgttcttc	tgtgacggga	cagaagagat	32820
gtgaggacct	gaagaggcgc	caccaggga	cttgagaaaa	agaaggcagg	tggctcggaa	32880
gaccgatcc	atgtgacct	gcaatttatt	ggatatggat	gaatcagagc	tgactttttc	32940
gatgacccaa	aagatgaaac	tattaattaa	ggctagacgg	aggggagaaa	agagagagaa	33000
ataccctgct	acctccagtt	tttctcccta	cagcacctcc	tagagatgag	gtgatggcag	33060
ctcaccttca	ggatccattg	gaaacaaaga	gaaatctctc	cctaattctc	ctgaccccaa	33120
acagaaagca	accaactatt	ccataatttt	cttctctagg	agagattcag	agagaagagg	33180
cctgaaaatg	caaatataca	cacggatgtg	aatgagtttc	agtcaaagct	taaaccagga	33240
cacaactttt	cttgtgtagc	gagggaaggc	aggaggcagg	atgggtgcct	gtgcctaaaa	33300
aggtgagtg	gacgtcaagg	tggctgaaga	gggttttaaa	agggcaccta	gggggacaag	33360
cccagagccc	agcatcccac	cctaaatgag	aacacaggtc	tccaactcca	gcccagggtc	33420
tccctgtggt	cacaatgggt	ttggggacct	gctgacagt	gcacggaagg	actctcggtg	33480
gtggtcagaa	tgaccacaaca	tcccaggagg	cacctgccac	cagttggcat	gagtccttgg	33540
tgctggccct	ggcgtgctg	ctaataccac	ccagtggact	taggcatgct	ccctcacctg	33600
tatgtccaag	acagctaatt	caacagtact	actaacctgg	tcccagaaa	ggcggcagag	33660
taggatcaac	ttggcattag	aggtctcgct	tcaaatacag	gatttcccaa	ttccaatctt	33720
gggcctcagc	caccaaccgg	gaaaaccccc	ctccaagggc	tgttttgaga	atatggaaag	33780
ctgccaaagc	tttatttccc	atccctttgt	aaggtccccc	gcagctccca	actagaagag	33840
aaagggacct	tttattagca	agaaaaaggg	caggtgcagt	gactgtcatc	acgcctgtaa	33900
tcccagcact	ttgggaggct	gaggtggggg	gactccttaa	gccccagagt	ttgagaccag	33960
cctcaacaac	acagtgtgat	ctcatctctt	caaaacacat	ttaaaaaat	tagccaggtg	34020
tggtggcgca	cgctgtagt	cccagctact	tgggaggtg	aggtgggaag	atagcttggc	34080
cccaggagtt	caaggctgca	gtgagctatg	atcacaccat	tgcactccgg	cctggacaac	34140

aaagaccctg	tctctaaaaa	ataaaaaataa	aaagtttttaa	ttttaaaaaa	ggttaaatac	34200
taatgagaaa	ggtccacata	caaatttttca	tgctactgct	atattataagt	tcttattaac	34260
tgaaaaacact	atgctataag	ctattaaagg	taactaaaaa	ataaaaaatag	ctctttgccc	34320
caaagacagt	ctaagtgaca	gagctcagta	ctcactcatc	aacaacagcc	ctgagagaga	34380
caagagatgg	aagagattcc	aaccccaaga	aaagggctgga	ggggagccag	gggaggaggc	34440
atgggggagg	ggacctccaa	gagggagcta	ggggagggca	gatggggagg	agagcccaga	34500
agggagctgg	gggaggggaa	catgggggag	gggagcatgg	gggaggggag	ctccaagggg	34560
gagcatgggg	gaggggagat	ggggagggga	gctctaggag	ggagctgggg	gaggggagat	34620
ggggagggga	gctccaggaa	ggagctgggg	gaggggagca	ttggggaggg	gagatggggg	34680
aggggagctc	caagagtggg	ctgggggagg	ggagatgggg	aggggagctc	caggaggagg	34740
ctgggggagg	ggagatgggg	aggggagctc	caggaggagg	ctgggggaaa	agggcttggg	34800
gagggtagct	caatgacggg	acgtggggat	ggggagctaa	ggaggagatc	tgggagaggg	34860
gagcttgggg	aggggagatg	ggggagggga	gatgggggag	gaaagctggg	gaaggggatc	34920
tgggagaggg	aagcttgggg	agaggatcta	gggaggggag	ctgggggttg	ggagatgggg	34980
aggagaaatg	ggagaggagc	ttggggaggg	ggatctaggg	aggaaatctg	gggcaaggga	35040
gctgagggga	gggaacttgg	ggaggggatt	tagggagggg	agctggggg	aggggagctc	35100
agagagggga	cttcagggag	gggagatggg	agaggggatt	ggggagggat	ggttagggat	35160
gggatctagg	gaggggattt	gggggagggg	agcttggggg	ggggatctgg	gggaggggag	35220
tggggaagag	agatggggag	tcgggggagg	ggaacctgga	gggagggatc	tggggaaggg	35280
gattttgggg	aggagaacag	gtggagagag	gagctggtgg	ggagggcagt	tgggggcagg	35340
catctggggg	agatttgggg	ggaagggaag	ctgggcgccc	acaggagccg	ctgtgagggtg	35400
ggcaagcccc	tcttccagtt	cctcctcgac	agtcagctc	cagactcca	ctccaccctc	35460
ccctgcttcc	accagacag	tctgatctgc	aactcgccc	atgactgcc	ccattgggaa	35520
tccagctgct	tctagcctgg	gaacctgac	gtgggcctcg	acctgaccaa	tcaaaaaccc	35580
caggggtgatg	gagcaaatgt	gtcctgtatc	ttgagcataa	cattaaaagt	gaggaccag	35640
cagaagtccc	ccagcgagg	cccagaaata	aggaatctct	ttgattcttg	caggctagt	35700
tttccctacc	cacataatct	ttagaaatca	tgtgtgccgt	aataaaaagt	agtatttccc	35760
ctccctcac	tcaagcacac	agaaacatcg	gagaaaagct	gagcatattt	ctaccagttc	35820
tgcatatgat	tttgaccaga	acacctgct	gtcggtaatg	aatggttgac	cccaatttct	35880
gaacacatat	ttccttttcc	aattaatttt	ccttcccctc	atgagataaa	acagactatt	35940
tttttttaaa	gaacaatatt	cctgaaaatt	tatttacttt	ttttaaaact	atgaggtcag	36000
agtttaagac	tggctccttg	gtatgaagga	atacatgata	ttaatataac	aaagggctga	36060
atcttccata	aatcaacaaa	acacccaaac	aaaggcagaa	cttaattttt	ggcaaagaaa	36120
aaacaaaaat	gtttttggtg	tccattagtg	aatacatcag	ctgaggactg	ccatcttgga	36180
atcttttaaa	tgagcagagc	taaagatttc	tcataagcac	aattaaagca	ccctgaattg	36240
atacctttag	ggggttgagt	atctgtttca	aatcagcaaa	gtgcttaccg	caaaagggaac	36300
accttcccaa	aagcaagatg	aaaaagtggg	ggcagagtgt	catgattatt	actttttttt	36360
ttaagcagaa	gaatagtctg	caagaaaata	cataaaaatg	ctcaagttag	gccgggcaca	36420
gtagctcatg	actgtaatcc	cagtacttca	ggaggccaag	cagggaagatg	tcttgaggcc	36480
aggagtccaa	gaccagcctg	ggcaacacag	caagaccttg	tctctattag	aaaataataa	36540
gttaccaaaa	aatgctcaaa	atggtaatgt	aaaggtggta	gaataatggc	aattattttc	36600
cttcttttcc	aagcgttata	taatattatt	aaagtggcta	gacatatgta	tggattttta	36660
agcacttcag	ttttatgtgt	tttaggtata	atttctaaag	cactaaaaaa	ttggcatatt	36720
cttttttttt	tttttttaag	acggagtttt	gctctgtcgc	caggctggag	tgcagtggcg	36780
caatcttggc	tcaactctgc	tcccgggttc	aagcgaattc	ccgcctcag	cccccaagt	36840
agctgggact	acaagcacac	actaccacgc	ccggttaatt	ttttctgttt	ttttagtaga	36900
gacagggttt	caccatgttg	gccaggatgg	tctcgatctc	ctgacctgtg	gatccaccgc	36960
cctcagcctc	ccaaagtgtc	ggaattacag	gcgtgaccca	ccgcaccag	cccaaatgtg	37020
gcataattct	tttgaacgtt	ttccctttgg	gagaggaaaca	agagcattcc	ttacctgctt	37080
gggagaaaga	cttaggaaca	agaattgaaa	gtctgcttac	ctgaggttta	atthtcgato	37140
ttcttcgctg	ccagcctcca	actacagaga	aagaaagaga	atatcacacc	acaggcacca	37200
ctgtcaacac	gcctcgggag	gcagtctcac	atthttctacc	ccggtactgg	aaaaagataa	37260
agatatccag	gaaacctagc	tacttctaaa	cagccgtgcc	ctttcctcac	caatcccgtt	37320
ctgtcccttg	gagtcatttc	cgtgggggaa	ttttcagggt	tccaaatgtt	gacccacatt	37380
cctgccgcag	tccaggggat	ggagtcctgt	tagctcaaca	tttcctatct	ggtgttgtta	37440
cccagcacgg	tcttttagcc	ctcagcctc	aactttccga	ggttggtctg	gaccttatcc	37500
tgtttttctc	ttttaagggg	aggggggtcat	gttttaagag	aatccacttc	ctccgcagag	37560
ccaggcaata	acagctgagt	gatgaacacc	atthttcaaaa	aaccaacca	ggcaagactt	37620
gcacagtggg	aggtggccag	gaatcaggcc	gtctgtttgt	gggtcttgaa	agctcttgat	37680
ggttctcgaa	aagacttaaa	catttgatac	gaaacatcct	aggctatcgg	tttatttata	37740
taaatgcaag	aaagagatat	ttaatatttt	ctgaaatcta	aaaggccacg	agtttgggct	37800
ccagaagtac	ctatgactta	tttttattht	ttttctttca	gagagcaaac	tgaaaaataag	37860
aaggaaacac	atacacaccc	cccaaacac	tccgcaccgc	tgggacttgg	catgtttttt	37920

atgttgcaca	gaggcgccca	ttgaatggga	aagagaaacc	tggaaagctg	tgatggctgg	37980
gagagatgca	gggtctgatcg	aggacagaaa	tgaggcagga	gccaaagggcg	aaggaaaaag	38040
ggccagagaga	taatgtaggg	aggggcctgg	gcagcaaggg	acacccacca	ggaggtggca	38100
acttcaacca	agaatgagta	caccagcccg	gcgcagtggc	tcacacctgg	aatcccagca	38160
ctgcaggagg	ccgaggtggg	cggatcacct	gaggtcagga	gttcgagacc	agcctagcca	38220
acaaggtgac	acgctgtctc	tactaaaaat	acaaaaatta	gccaggcacg	gtgacatgta	38280
cctgtaatcc	cagctacctg	ggaggtctgag	gcaggagaat	cacttgaacc	caggaggcgg	38340
aggttgcagt	gagccgagat	tgcaccactg	cactccagcc	tggtgaaaga	gcaatacttc	38400
gtttcaaaaa	aaaaaaaaaa	gcgtacacca	gagggcctgg	gagtccttac	atcatattaa	38460
gatgaagtac	atacaagatt	ctgcagaggg	acctacccca	cactgaagga	gaagtggaaa	38520
gagcagggaa	ggcttcaatg	accacaaaac	aagtcacaag	agcaacaaat	tgacaaagag	38580
tatgttgggg	tctaaccacg	tggttctcaa	tgcgagcaaa	gttctctcaa	caggagacat	38640
tggaagatgt	ctggaggcat	ttttctggag	gtcactactg	gcacctagt	ggtagaggac	38700
aggatcccac	aacacacagg	atggtccccc	cacaagagag	aatgttctgg	tcccaagtat	38760
caacagtggg	ttgagaaact	ctggtccaat	ccaaaaaagt	gcctggagat	ctgcccatga	38820
aaattagtca	ctttgaatgt	ttctcagaaa	taaccaatgtt	atgatccatt	cctgaaaatt	38880
attgatttat	ctattcttgt	gctctgcctg	ttcacacaaa	ggaactgaga	taagattcac	38940
aggaaaatga	cataagtga	ataatcaaag	actgggaaaa	aaagaaaatt	aaaagagaaa	39000
agagagcctt	aggactgagg	gctgtgatcc	cccgtttccc	acgccggcag	caggcctggc	39060
tgtgtcagga	aagcactgcc	ctaagtgtct	gactcattat	gaagtgtcaa	tttgaagagt	39120
gatgacgtga	cttgggggta	cggacttcac	aatcatttaa	ctctcggtca	ctctctgagg	39180
ttctcagatg	aaaggccatc	tcaggtcagt	tattccaggg	aaactacatc	tgccaaaggaa	39240
cacatgaaag	aggttaattca	gtccttttag	atgagccagg	gccacacac	aggaagcaac	39300
tcaagcgagg	gcggaccagg	gcagaaccgg	cctggcctag	gtctcctgac	cccatacaca	39360
cttgctgtct	ccatcccacc	ttgcttctca	cctcaacaca	tctgaacgag	ggccttgctt	39420
tcgggaaaca	tcccagcgca	ttcaaagcca	agcaatgaat	gctgcagctt	tgctatgatc	39480
aaataaaagc	tgcttgagtt	ttactttatg	tttatcaggg	tatttgccac	ttggtaaaaa	39540
taatgcttta	tataataatt	gaaaatgtat	ctagtagaca	caacacaaat	gtccaacaaa	39600
atgtggtaca	tccatacaat	ggagtattat	acagccatga	aaaggaatga	agtactgcca	39660
catgctacaa	ctttgaaaac	gtgatgtctga	gtcaaagaag	gtcaaagaag	ccagacacaa	39720
aaaggccacac	aggggtgtgat	tccatttata	taaaatgtcc	agaataagca	aatccataga	39780
tacagaaaagt	agattagtgg	ttgcctaagg	ttagggagaa	tggggcaggg	ggaggtgca	39840
ggtaggggct	aacgggtaca	gggttcctat	tttggggagg	atgaaaatgt	tctggaatta	39900
gatggtggtg	gttgacacat	ctcatgtata	tactaaaaac	cactgaattg	tacactttaa	39960
aatggcaaat	ttatggtatg	taactaaaaa	ataataagac	cttaaaatgc	gtaagacaag	40020
aacagattag	gttgcaagta	actctaggaa	catgggggtt	tgaatcagaa	atctgggctg	40080
acaggttcag	ctgaaagcca	acctctcccc	ttacctcaca	taaaactttg	tgatgagaac	40140
ctgagattaa	gtgcataaat	tgccacagac	cagtgaaccg	aacagaaaaa	agccccctct	40200
cgctgtggga	aggaaaggcg	ctcctgcaac	ctaacttctc	agtagcaggc	tattgatcgc	40260
cagtgttctt	ttgcctctaa	tcagcgtgta	gagggggatt	actagaacct	tctgtgtata	40320
gataactcat	gaatggcctc	tcctctccaa	ggagggggct	gtgaagggtc	aacttcccag	40380
ccactctgaa	aatgtccctg	ccaatcccag	caaaaacagg	ctgaagaact	acctaccag	40440
gagacagggc	tgtcaagcca	aatgcaaaac	ttatttctct	gtctctctca	gacacacaaa	40500
ctcccccgct	tgtatcagtc	aacttcccc	actccctccc	acaaagaaag	gggtggaaga	40560
gccagagctg	tggtgcgga	acttctggg	cctggaccg	cagggcgct	cctccagctt	40620
tctctaaaca	cagctaaggg	tctgcaggcg	gacactcagc	cttggttatag	gtaagagttt	40680
agaccagagg	ccttgacggg	ttcttcaaga	gatggtgggc	aagattgcgc	gaccagaggg	40740
tcatccctgc	agctacagag	ggctgacctg	ctcagaggcc	caaggcccca	gcctaggaca	40800
agccaggcca	acctgcagg	ctaagagggc	aacagtgcct	tcaatcaacc	ccagaggaaa	40860
aagtggccag	gcaaacggac	ctgggcccaca	cacagaccca	caaaaaacgcg	cacagtgcc	40920
ggacacgcaa	cccaggaatg	cacctatgca	atcaccagga	atgggtcaca	gccacacaga	40980
aagatagatg	cacataaaca	cacaggcctg	agtgatgtta	cagaaaggaa	aagccagact	41040
aaggctgcac	gcacagacgt	gaaacacagc	cacacagagc	ccacagcacg	ctcggtcacc	41100
gtcacacagt	gacacgggca	cgcttacaga	cagaactcca	gaggcggcag	gcggggaaac	41160
aatctcacac	gtttgtaggg	gcactcccag	atgcctgtct	cacgctggca	cagtccccgg	41220
tacggcaggt	cagcaacagt	cacatctcac	atcgacagac	caggcataca	ggcaaaaggc	41280
ctagaactac	cccgccaca	ggtctcagaa	ccagcggtct	acgcagtca	ccaatcaagg	41340
gtcccagttg	cacatccagt	caccctgga	ccctggtcac	actgcagagt	cactcacaaa	41400
tggtgagctcc	gcacagcgca	cagtcctccc	cagacagagg	tcaaccacaag	atgggggtca	41460
cactgaaat	cacagtcccc	acacaatcac	gaggtcacat	ctgcacacac	agtccttgca	41520
cagtcaccct	taggggtcac	aacgcacaca	gtctccgct	aacaggggtc	accccaagat	41580
gggggtaacc	ccctgatgtg	ggtcacagcg	cacacacagt	atcctgcaga	cactcccaga	41640
gggtgcgcac	tgcatcacaca	gtcctgcaa	agtcgcccc	cgataggggt	cacaccgcac	41700

acaaagtc	ccgacagctc	cccaagacag	gggtcacatcg	cacacagtc	tcgcacggtc	41760
accccggtcc	ggctgcccgg	ctctgttcct	acggcggggc	cccaggagagc	ccgcgcagcc	41820
gccccctgc	cccgcaagcg	cggccccagc	tccggcgggc	tcggcgcggc	gtccggcggc	41880
ccaggccggg	cgcggcgagc	ccggggctca	cctcgctgtt	gctggccgag	gaggaggcgg	41940
cgtcgggcgt	gggcgagcgc	tgcagggtca	ccaggggccat	ggctgcggcg	cgggtgcgag	42000
gcgccacaga	cgtctcgagc	tagagccgcc	accgccaccg	ccgcccgggc	cgggccccgg	42060
gcctcctgga	gccgcgcgcg	ggcggccggg	ccgagccggg	ccgggcccgc	ccctccccct	42120
cggcgctcgc	accgcccccg	ccccagctc	ccgcctcccc	cgcggcgcg	cgcaggcctc	42180
agtgcgcgga	gtgggcgggg	aagcgggcag	ggcgggacga	ggaggcgcg	gtgcgcgggg	42240
gcctgaggg	ctgcccaggg	cctcggtctg	tcgatcacgt	ccctcgcgcg	cccgacacac	42300
gcgccccgc	ccgcgcgcgc	cgtatcagg	cctgggactc	gggggcgcgc	gcgcgcgcgc	42360
gagcccgtag	gcccagggg	ccctgcccgc	tgtctgcct	gggaaactg	aggcccgcg	42420
accgtgcaga	caggactgta	cagcgaccag	gaaataaaa	acgtcctggg	gccggcgcg	42480
gtggctcacg	cctgtaatcc	cagcactttg	ggaggctgag	gcgggcggat	tacgaggtca	42540
agagatcgag	accatcctgg	ccaacatggt	gaaacccgt	ctctactaaa	aagacaaaa	42600
ttagctgggc	gcagtgggtg	cgcctgttag	tcccagctac	tcgggaggct	gaggcaagag	42660
aatcgcttga	atctgggagg	cggagggtgc	aatgagctga	gatcgcgcca	ctgcactcca	42720
gcctgggcga	cagagcgaga	ctcgtctca	aaaaaacaaa	aaacaaaaaa	caaaaacagt	42780
aagcaaaata	gattcgccctg	attttgcaga	ggttaatcaa	gttattaggc	acgtttttta	42840
aaaagtat	tgctaattct	tttcaatgaa	ttctttctgg	gtgttctgaa	accagccaa	42900
ctccttgagg	gtcagggaag	gcttcccaga	agagctttat	tctgaggctt	gggcttgagc	42960
ataagcagga	ttaacaggtg	aaagaacaga	gagacagctc	tccaagcagg	ggggatcagc	43020
gtgcccgtga	gcaggaagaa	gtttgtcaac	cggaggccag	cactcaggga	agggaagagg	43080
ggaggaatgg	ctggagtctc	catcctctct	ggaaagatcg	ctccggctgc	tgcggtgatg	43140
agggaccacg	gggcagaggg	ctgagggaga	ccaggggagg	ggctgctgct	gttgtcccgg	43200
ggagaggtga	ccagttatgg	ggatggagag	gggaacatgg	aataagatac	caagaaggca	43260
attctggctt	gacttagtag	taggaaactt	ttcttttagc	caaaatctca	tctcccggct	43320
cccaccccca	acctctgcat	gttgacacaag	cactcgcaaa	cgcagtggtc	ccagcctgcc	43380
ccgcagctta	gcaaatttgt	cttactgccc	aacaggaaac	ccacgcagcc	tcctggattc	43440
ttcccctgcc	tcctctctgt	cctggggctg	tgacctctct	catgttattc	acagggtctc	43500
agcacgat	atctcaagg	tgattctagt	ggggggcact	gtagcttcta	cggagcgttt	43560
ctaagagggg	atttgtggga	atgtttgtgg	ttgtcttgc	gatggagggg	gagagctcct	43620
ggcatttaga	gtgcaagagc	cttggatgct	aaatgtcttc	caatgcactg	gacagctctc	43680
ccaacaagaa	ttgtccatt	cccacaaaat	gtttcctggg	tgaaaaaccc	atttatagta	43740
atttgaagcc	agaacctaac	tccatttcat	gcacaaacac	tagtcttctc	tccttctctc	43800
cttcttctct	tccttctctc	ctgccttctc	tccttctctc	ctctcttctc	ctcacttttt	43860
ttctgaaaca	gggtctcact	cccgtcaccc	aggctgaagt	gcaatgtcac	aatcatagct	43920
cactgcagcc	tcctctctcc	aggctcaaat	catctcctcg	cttcagctctc	ctgagtacaa	43980
cgggtacaca	ccaccacacc	cagctccttt	aaaaaaaaag	tttaactatg	ttgcccaggc	44040
aatcctcctg	cttccgcctt	ccaaagtgtc	gggattacag	acagaagcca	ccatggctag	44100
cctgggtat	tttactgaat	tttcagaaag	gtgactatgt	tgaaaccctg	tctctcctaa	44160
aaatacaaaa	aattagccag	gcattggtgg	gggcacctat	aatctcagct	actcaggagg	44220
ctgaggcagg	agaatcactt	gaacccggga	ggcagagggt	gcagcaatct	gagatcgtgc	44280
cactgcactc	cagcctgtgt	gacacagcaa	gcagagaga	aagagagaag	ggaaggagg	44340
ggaggggagg	ggagaggagg	ggagaggagg	ggaggggagg	ggaggggagg	ggaggggagg	44400
ggagaggagg	ggaggggagg	ggaggggagg	ggaggggagg	ggaggggagg	ggaggggagg	44460
ggaggggagg	ggaggggagg	ggaggggagg	ggaggggagg	ggaggggagg	ggaggggagg	44520
gaaggggaaa	tacactttgt	tttgcttgag	agttttgtca	agagttgttc	atccatcctt	44580
agggaaaagg	aggtaatgga	tggcaacgcc	tctgtcaata	ttagagcatc	ccacacaagg	44640
tgcccaaca	tgtagctgca	ctctaggtag	acagacagtc	ataggtaact	aaatgtcaaa	44700
tataagggaa	aattgtggac	aaaattcagt	tgagttagaga	atattttatt	tctcaaatcc	44760
aagcacattg	attattggca	ggcccatgct	tctgagatgc	ccctgtgtcc	tctmagggag	44820
tagtggtctga	gcatttccac	attgtaatgc	atgttgtttc	attatgattt	atttttcttt	44880
tatgtctctc	ttacattagt	tttcaaattt	gagagtttga	gaatccccctg	gagaaaaatac	44940
agattgctag	acccacctc	ccagagtttc	gaattcacaa	ggtttgctgt	agggctggaa	45000
aatttgcacg	tctaacaaat	tcacaggcaa	tgctgatgct	tctgtctggg	gacgacagtc	45060
tgagaactac	tgctatata	aatgcaatgg	cctcttcacc	aagaaattcc	tacctagatc	45120
tgatcctgg	accgctcgt	ggccccaat	cctaattccc	ctgctctggc	cggcctgctt	45180
ttccactcac	cccactttt	ttggaggcag	ttccacccc	ttctcacttc	ctcttagagc	45240
tgagagccct	tttcttccc	acaactaact	cttgctagaa	atcacctcca	aaaagctttc	45300
cctgcccctt	aagcagtgtc	atttccagga	tctcgtagcc	ctcacccctac	ccttaaacac	45360
acagcaagtg	tcagtctgcc	ttatcataat	gggtccatct	ctctgtcttg	tcccattacc	45420
gtagagccag	gaacggtccc	taagaaaagg	ctcaggaatc	aggctgggac	cagcgtgagg	45480

gtgcaaaatg	taagagggtg	cccccaaaaa	ctcaatgatt	aagataaata	gtatttttaat	45540
gcaatatattt	agaaaaatcaa	aattaatgcc	aaatccatga	tgaataaaaat	attttttaaaa	45600
tttgctttttt	ttttttttttt	ttaattgaga	cagagtcttg	ctctgttgcc	caggctggag	45660
tgcagtgtgg	cacaatctct	gcctcttggg	ttcaagcagt	tctcctgcct	cagcctcccg	45720
agtagctggg	attacagacc	cccaccacca	tgaccggcta	atttttgtat	tttttagtaga	45780
gatgggggtt	caccatgttg	gccaggtctg	tctcaaatcc	ctgaaatcag	tgatctgcct	45840
gcctcggcct	cccaaaatgc	tgggattaca	ggtgtgagcc	actgcacctg	gtcaaaaat	45900
ttacaaaaat	tttttaagag	ccaagggtct	attctgtcac	ccaggactgg	gtgtagtgg	45960
gcaatcctag	ctcacttcag	ccttgaactc	tgggtctcaag	ccatcctcct	gcctctgcct	46020
ccggagtacc	tgagactaca	ggtgtacacc	accacgcctg	gctgacttta	tttttgccag	46080
aaactgggtg	ttgctatgtt	gccaggtctg	gtttcaaact	cctggaggca	ctcaatcccc	46140
cgaccttggc	ctcccaaagc	tttgggatta	ccggcatgag	ccaccacacc	tggccaaagt	46200
atcaaatattt	taagtaaaat	tggcatcagt	attgtgtcac	tgattcttcc	acttacttca	46260
gacttcagt	tagctcagca	aagcactttt	attgatcctg	tctttatttg	attcttttac	46320
acttttggcc	attctaaagc	cttttgtgaa	aatggcctgt	ggttcagctg	ggcatgggtg	46380
cgtgcacctg	taatcccagc	tactcgggag	gctgtggcag	gagaatcgcc	tgaaaccagg	46440
aggtggaggc	tgcagtgggc	tgagatcgtg	ccacttttga	cactctgtct	caaaaaaaaa	46500
aaaaaaaaaa	aaaaaggaag	cctgtcggct	tgactccagt	agcctctgat	gggggtggag	46560
ggacaagggg	aagtgaagc	tcccaggcct	cagtcaaggc	aggtcccaag	aagccctgag	46620
catggaggag	gggaacaatc	cagtagaggc	agctctgaag	ttttctccca	tgcattagag	46680
ccctttccaa	tcagtatcat	gatttttcat	catataatag	tttatttaat	catctttgac	46740
ctcctccttg	tagtcccagc	tcacttttgt	aaacagtga	aaacagtga	ttattgagct	46800
atttgcctct	tgctaaggca	caatgcaag	tgctttgtga	gtgtgtgggg	gacatgattt	46860
attaacatgt	gactgtcccc	ccacttatac	tccaagatca	cctcctccag	gaagccttcc	46920
ttgccccgtg	gctgggttag	gcaccccttc	tctgtgtctc	tacagccctc	gtgcattagt	46980
gacaatggca	ttgtggatct	gccctaggcc	catttctggg	ttgggacact	ttaggtacat	47040
tcattcttgt	caccctgtga	ttctcatttc	atgggtgagg	aaattgatgc	acagagtgg	47100
taaggcactg	gccccaaagt	atgtaactaa	ggagtgggtga	acctgggttc	cccagtgttt	47160
tctgctttag	aactcaggca	aagacagggt	cttcaggac	agcctcagaa	agtgttgggtg	47220
caaatagggt	tggtgcaaaa	gtaatgtcgg	tttttgtcat	tttttttttt	tttaattgggtg	47280
caaaagtaat	tgccggttttg	tcattaatga	ccaactatta	taagtaatag	ttcccttttt	47340
tttttttttt	gagatggaat	cttgctctgt	tggccaggct	ggagtgcagt	ggcttgatct	47400
tggctctctg	caaactccgc	ttcctgggtt	caagtgatcc	tctgcctca	gcctcccaag	47460
tagctgggat	tacagggtgc	caccccatg	cccagcta	ttttgtattt	ttagtagaaa	47520
cgggggtttca	ccatgttggc	caggctgggtc	ccgaactcct	gacctcaagt	gatccaccca	47580
cctcggcctc	cccaaaagtgc	tgggattaca	ggtgtgagcc	actgcacctg	gccagtagt	47640
tgccgttttag	agcaataaac	ttgtaatttc	ttcattccaa	aatgatattc	47700	
agaggtaata	aagctctgat	aggctgaata	atggcctgca	aagatgtcca	tattccaaat	47760
ccctagaatc	cctgcctatg	ttaccttgca	tgctaagagg	gttttacaga	tgtgattaaa	47820
ctcaggatgt	ttagatgggg	aaattttcct	ggaggaggcc	caagagggtc	taatgtaatc	47880
acaagggtcc	ttataagagg	gaggtgagaa	ggtcagagtc	agtagtaaga	gatgtgacaa	47940
cggaaactgag	ggattagagt	gaaggaagag	gccacaatcc	aaggaatgca	ggcagttgct	48000
aaaagtggaa	aaacacccaa	aaatgaattc	tcttttcaga	gcctccagaa	agaatggagc	48060
cctgctgata	ttcttttctt	ttcttttttg	agttagggtc	tgctcagag	agctgtcacc	48120
caggctggag	tgcagtggca	tcacatagc	tcacagcagc	ctcgacctcc	agggtccaag	48180
ggattctccc	acctcagcct	cctgagtagc	tgcgactaca	gacacacacc	actatgcccg	48240
gttgactttt	tttaattatt	attatacttt	aagttctggg	gtacatgtgc	agaatgtgca	48300
ggcttggttac	atagggtatac	acgtgccatg	gtggtttgct	gcacccatca	accgtcatc	48360
tgcattagat	atctctccta	atgttatccc	tcccctggcc	caccaccccc	tgactggccc	48420
cgggtgtgtga	tgttccccat	gcccggttga	tttttaaggg	ttttgtttgt	ttgtttgttt	48480
tttttagagac	gagggtctca	gctgggtgca	gtggctcatg	cctgtaattc	cagcactttg	48540
ggaggtaagg	cgggcagatt	gcttcagccc	aggagttcaa	gaccagcctg	ggcaacatgg	48600
cgaaacccaa	aaatgcaaaa	aattaactgg	gcatgggtggc	acatgcctga	ggctgaggtg	48660
ggagtatcgt	ctgagcctgg	gagatcaagg	ctgcagttag	ccatgatcat	gccactgtgc	48720
tccagcctgg	ttgatgggg	gagaccctgt	gtctaaaaaa	taaaagaaat	gaaggtcttg	48780
ctgtgtttcc	taggctgttc	ttgaactcct	aggctcaagc	aatcctcctg	cctcagccac	48840
cccagttgct	tggattacag	gcacaagcca	ccatgtccaa	tcttggaac	gtcttgattt	48900
tagacttctg	atctctacaa	ttgcaagaga	ataaatttat	gttgttttta	gccacgaaat	48960
ctctgggaat	ttgttacagc	agccatacga	aatgtaata	aaactcaacc	tccatttggg	49020
ctttaaaaaa	catatcatta	taatgccatt	accaggtata	ttccagggtg	ttcccaagcg	49080
ttgtgtcatt	ttctcattca	ctcaactcat	ccaataaact	atgtttgttg	ctctcctggg	49140
cactagtcta	ggaatctggg	ttccatcagt	gaacaaaatg	gaatcactgc	ccttgaagag	49200
cattcaatca	agtgggaaat	atagtaaaaa	tatatatata	tgcaaatatg	tttaaaatca	49260

tatgtggttaa	atatattgca	tttaaatgaa	ttaatatggcc	gggcacggtg	gctcatgcct	49320
gtaatcccag	cactttggga	ggccgaggcc	agtggatcac	ttgaggccag	gagttcgaga	49380
ccagcctggc	caacatggcg	aaaccccgtc	tctactaaaa	gtacaaaaat	tagccagttg	49440
tgggtggtgg	tgctgtaat	cccaggtact	cgaggaggctg	aggcacaaaa	atcgcttgaa	49500
ctgagggggt	gcgaggttg	cagtgagccg	agatcatgcc	actgcactcc	agcctgggtg	49560
acagagtgag	actgtctcaa	aataataata	ataataatta	atataatgaa	ttaatattgg	49620
taagggtcct	tagaacaaga	taggcactga	tatgtgtcaa	ataaatgaaa	tatgatgtcc	49680
aatcatgaaa	aagcttggga	gaaaaacaaa	gcaggctaag	ggcagagtaa	tggaggaggc	49740
cacttagaca	aatggtcagg	gaagcttctg	ggtgaggtga	tatttgagca	gaggaatcac	49800
catgacagca	ccaccaggga	ggtgtagaaa	ccctgggata	tgcttggttc	attcaaaactg	49860
gcctccccac	taaggaaactg	tgaggtactt	tttctgagac	ccattttctt	tctgtctgtg	49920
tcacccaggc	tggagcgag	tggcgagatc	tcggctcact	gcaacctcct	ccccccaggc	49980
tcaagtgata	ctcccacctc	agcctcctga	gtagctagga	ttacagggtg	gtgccaccat	50040
acccagctaa	tttttgtatt	tttagtagag	tcggcggttc	accatgttgg	ccaggccagg	50100
ctgccacctt	ggcttcctac	agtgtctgga	ttacagggtg	gagccttcag	acccagccga	50160
gacccactgt	ctttctctgt	aaaattgata	tgaaagtgat	agtgtctggc	cgggcactagt	50220
ggctcacgcc	tgtaatccca	gcactttggg	aggccaaggt	gggcagataa	cctgaggtca	50280
ggagtccaag	accagcctgt	ccaagacggt	gaaacctgt	ctctactgaa	aatacaaaaa	50340
ttagccagg	gtggtggtg	gtgcctataa	tctcagctac	tcaggaggct	gaggcaggag	50400
aatcgcttga	acccaggaag	cagaggttac	agtgagtcga	ggtcccgcga	cttccactcca	50460
gcctggacaa	caaagcaaga	ctccatctca	aaaaaaaaaa	aaaaaagaaa	agaaaagaaa	50520
agaaagtgg	agtgtgacc	tcagagcttg	gttgtgtcaa	ttgaacagca	tactatgcag	50580
gaaaggcaga	gcgtgctgtc	ctatttacta	atagctaccta	aggtatttgg	ttgaatttgt	50640
tccccacaaa	attcactagt	ccctgtgaat	gggaccttat	ttggaaatga	ggtctttgca	50700
gctgatcaag	ttaagatgag	gtcattaggg	cggggccccta	ttcgcatatg	actgtgtccg	50760
tatgaaaagg	gggaaatttg	ctgggcgagg	tagctcatgc	ctataatccc	agcacttttg	50820
gagaccaagg	cggttggtgc	acctgaggtc	aggagttcga	gaccagcctg	acaaacatgg	50880
agaaacctg	tctctattac	aaatacaaaa	ttagccaggc	gtggtggtgc	atgtctgtaa	50940
tcccagctac	ttcggaggct	gaggcaggag	aatcaettga	acccgggggg	tggagggttc	51000
agtgaactga	gattgcgcca	ttgcactcca	gcctgggcaa	caagagcgaa	actgcatctc	51060
aaaaaataaa	caaacaaaaca	aataaataaa	taaataataa	aaggggaaat	ttggaccacg	51120
agccaagggg	aaaatgcttc	ctgaagggtg	tagtttgtgt	gccacaagcc	aaagagcacc	51180
cgagatggc	agcaaacac	cagagctagg	agtgagaagt	gaggagcaga	tttgctgtgg	51240
cttctgaaga	aaccagcaac	tcgatttcag	agttccaggc	tccagaactg	agagagtaaa	51300
tgctgtggt	ttaagcctcc	cagtttgtgg	cactttgtta	cagcagccac	aggaaaggaa	51360
cgcatctaac	atgatcattt	catcagctgc	agaaaatgag	gctcagagca	aggctagggt	51420
ttgaaccac	gccaactaga	cccagacca	catacatggg	ttgttgacct	ttctagcctg	51480
ggaggtgaca	tttggatgt	tccactattt	cgagggaacg	gtgttcagag	gactcaaacg	51540
ttctgccacc	tgggccaggg	tgtccagtgt	tagatatgga	agtcagggtat	ctggggcttc	51600
tgacaaagca	tctctctggg	tgggtcaata	ggcaccagca	gccaggcagt	tgagggtatct	51660
ctgcccctgt	gcggagtgg	ctgaagcctc	ccttcctcta	cccatccctt	catttaacct	51720
gcttgccaga	tcaagtggt	ccctggcctc	tggcagggg	gtattcacct	gaatttgctt	51780
tttattcact	atgatcaca	gcaacacact	gacccttgct	gggcctcaga	atctcaactc	51840
ttggtgggg	gcgtggccc	atgctggtaa	tcccagcact	gggaggccaa	ggtgggtgaa	51900
tcacttagag	ccagagtta	gagaccagcc	tggccaacat	ggcaaaaacc	tgtctctact	51960
aaaaatacaa	aaattagcca	ggcatggtgg	catgcacctg	tagtcccagc	tactcaggag	52020
gctgatgcac	agaatcact	tgaatccagg	agatggagg	tgcaatgagc	caagatcaca	52080
ccactgcact	ccagcctggg	tgacggagt	agactctgtc	tcaaaaaaca	aacaaaagaa	52140
tctcaactct	taatatggaa	tgattaatgt	cctataagaa	agcatttggt	agatcaattg	52200
ggtatatgca	tgtgatatcc	ctcctagata	gctctcagtg	gtcctcatct	cctgatattc	52260
atgcctgtg	agctctcttc	acacaataca	tagaactgac	ctgtgtaacc	actaggatat	52320
cacagatact	acagcatgtg	gcttctgagg	ctaggggtca	taaaagacac	tgaagctgct	52380
gtcttgctgt	ctcttggtat	tctcattctg	ggggaatcca	gctaccatgt	catggggaca	52440
tttaggacat	ttaagcagcc	aaagagagag	ggccgcagtg	caagaaactg	aggtctctctg	52500
ccaatgacca	gcactaacct	attgtcatgt	gaatgcacca	ccttgaaaat	ggatcctcca	52560
gccccagtc	ggcctttatt	tattttattt	cttattaatt	gagaccgggt	ctcattctgt	52620
ctcccaggct	ggagtacagt	ggcaccatct	tggctcgtg	taacctctgc	ctcctgggtt	52680
caagcgattc	tcattcttca	gcctcccagg	tagctgggat	tacaggcggtg	cgctaccatg	52740
cccagctagt	tttttgtat	tttttagtag	gacagggttg	cgccatgttg	cccaggctgg	52800
tctcaaacctc	ctggcctcca	gtgatctgcc	tatctcggac	ccccaaagt	ctgagattac	52860
aggcaagagc	cattgtgcca	ggccccccat	tcaagccttc	agatgagatc	acagccatgg	52920
ccaacatctg	gagtgaacc	tcatgagaca	ctctgagcca	gagctgcccc	gctgagctgc	52980
ttccagattc	caggcccaaa	gaaaatgtat	gagataataa	atgtttattg	ttttaagctg	53040

ctaaatttta	aggtaacttg	ttatgcagca	atagataact	tttatatgct	gccataaaaa	53100
tattataaaa	ccatgcacta	gtacagaaag	atttttataa	aatattaagt	ggaagaaaaag	53160
aaaagcaggc	caccaaacag	cgtaggacag	tagaccccat	ttttgaaaga	aaaatgtgaa	53220
gagttaaaaa	actctacca	aaggggaaaa	aaaagagggc	atcaatggag	agatggagaa	53280
gctttgtttt	tggatgggaa	gactcagtat	ggtagagtta	acaacctctc	gaaattaaat	53340
taaatggaaa	tgctattgaa	atcccaactt	gattcttttg	agtgtaggga	tctttacaac	53400
agataactgg	acaagctaac	atttattggg	tatatatgtg	cgttgcatca	cgtgacagtc	53460
actgtttcat	cttaattcca	ccataggaga	aagtcctctc	ttatttaatt	ttctgagag	53520
taaaagtactg	ctattacctg	ttccccctcc	cattttactt	aggaggtttc	aagaggggac	53580
ttgtctgaga	tcctggaaac	cgtggagggtg	agatgacatc	aagcatgttt	gatattttaca	53640
tggtgtccct	tggtccctct	gccacatggc	ctccccactg	tgccctggtt	tccttaagta	53700
ccagcccaag	gacacatgga	taggaaaggt	ggagctgggg	caccagccca	gtctgcctga	53760
ctccagagtc	cctggtctta	atcactaaac	caccocagaa	aagtaaccgt	gggagaagag	53820
acctgcaaac	taggaaaaag	aagattaaag	ggaaggaaac	tggtctgcta	gatattaaaa	53880
catatgacaa	agctgtagaa	attaaaaacag	aatggggccg	ggtgcagcag	cttctgcctg	53940
taatcccagc	actttgggag	gccaaaggta	gtggatcacc	tgaggtcagg	agttctagat	54000
cagtgtgacc	aataatggtaa	aacctgtctc	ctactaaaag	tataaaaaatt	agctgggcat	54060
agtgggtgtg	acctgtagtc	ccagctactc	tgaggtctga	gccaggagaa	ttacttgaac	54120
ctgggaggca	gaggttgacg	tgagccaagc	tcacactact	acactccagc	ctgggctaca	54180
gagcgagact	ccagttcaaaa	aaaaaaaaaa	aagaaaagaa	aaaaaagaag	aagaaaaaaa	54240
aaaaccgggc	gcagtggctc	atgcctgtaa	tcacagcact	ttgggaggcc	gaggtgggtg	54300
gatcacctga	ggtcaggaat	tcaagaccag	cctggccaac	atggtgaaac	cctgtctcta	54360
ctaaaaaacac	aaaatcagca	gggtgtggtg	ctgcctgcgt	ataatcccag	ctacttggga	54420
gactgaggca	agagaatccc	ttgaacctgg	aaggcagagg	ttgcagtga	ccaagactgc	54480
gccactgcac	tcagcctggg	gcaacaagag	caaaattcca	tctcaaaaac	aaaacaaaac	54540
aaaaacaaaa	acaaaaacaa	acatagaatg	ggtgtctggc	tagaaaagcca	caaacagatg	54600
aatggagcac	aacattaaat	ccaggataaa	actcaaacac	acaggaaaat	ctggtgaacg	54660
ataataggag	gtaggtatct	gtgaagcgtg	gagtaagagt	gagttactca	accaacagtt	54720
ctgtgtccac	tggttagcaa	gtaaatgcgg	atccctaccc	cactgttagt	ctccaaaatt	54780
aatcccaacc	gagtcagtta	aatgtttaaa	aaaccttcaa	aagttgccag	gcattggtgg	54840
tcacgcctgt	aatcccagca	ctttgggagg	ccgaggcggg	tggtatcacct	gaggtcagga	54900
gttctagatc	agtgtgacca	atatagcaaa	acccacctc	tactaaaaac	acaaaaatta	54960
gctgggcatg	gtcaggggag	cctatcgtcc	cagctactca	ggaggctgag	ccaggagatt	55020
tactagaacc	caggaggcag	agggttcagt	gggccaaagat	cacaccaccc	acactccagc	55080
ctgggcaaca	gagttagact	cgttctcaga	aaaaaaaaaa	aaaaaaacct	tcaaaagtta	55140
tagaaagtct	gtgtgagtaa	tttttaatat	gaagcaaggg	agagagatga	agcagggttt	55200
ctaaacatga	ctatagccat	ataaaagtat	gttataaagc	tgggcgtggt	ggctcacgcc	55260
tgtaatccca	gcactttggg	aggctgaggc	gggtggatca	cctgaggtca	ggagtttag	55320
accagcctga	ccaacatgga	gaaacccctg	ctctactaaa	aatacaaaaa	ctagctgggt	55380
atggtggcgc	atgtctgtaa	tcacagctcc	tcaggaggct	gaggcaggag	aattgtctga	55440
agtcgggagg	tggaggttgc	agttagccga	gatcgacca	ttgcactcca	gcctgggcaa	55500
caagagcgaa	actccgactc	aaaaaaaaaa	aatgtttata	aaccacacac	cactataaag	55560
aaatgataat	gcaaaaatca	taaaggacga	aaaaaaaaag	gaaatagatt	taactacaaa	55620
aaagtttttg	ttttgtcttt	gtttttttag	agtttagagtc	ttgttctttt	tcacagctg	55680
gtacaatcat	agctcactgc	caccttgaac	ctctggctc	aagcaatcct	cctgcctctg	55740
aaactgcgtt	tgcaaaaatt	ataactgaga	aaacgatgac	agtgaagag	atctgacctta	55800
actgactcca	tcttgcttct	aacctccaag	ctgtccgtgt	tcattcctgg	gtgtaggcca	55860
aactaacttt	gggaggaatt	tagtttatag	tttaactttg	tcaaagttaa	actaatgagt	55920
taatagccca	ttttccaaaa	caaaccctct	tcctgcctgg	ggactagact	gcctttgcag	55980
gactaacaac	ttattatagc	taccagatta	gaaattatgg	tttaggagtc	atgcagctga	56040
agcctacaag	attctgaatc	tcacaaattg	ctcctggaga	taacatcacc	attgtaaaaa	56100
ctaagatcag	atcttgacat	atcttgacag	ctctgcactc	gatggatcag	ctggcactac	56160
ccaaatggat	aaacaggctc	atctgatctg	tggtccccc	ccagaaaactg	accagcata	56220
agaggaccgc	ttcaactcct	ataactttgt	ctccaacctg	aacaatcaac	actccctac	56280
tttctgaccc	cctaccaccc	aaattaccct	taaaaaacct	agccaggcgc	agtggctcat	56340
gcctgtaatc	ccagcacttt	gggaggctaa	ggcaggcgga	tcacctgagg	tcagggttcg	56400
agaccaacca	tgccaacat	agtgaacccc	catctctact	aaaaatacaa	aattagccag	56460
gtgtggtagt	gtgcgcctgt	aatcccagct	actcaggagg	ctgaggcagg	agaatgcct	56520
gaacccggga	cacagagggtg	gcagtgagcc	aagatcactc	cactgcactc	cagcctgtat	56580
gacaagagca	aaactcggtc	tcaaaaaaaa	aacaaaaaaa	aaaaccacac	gaaaaaaacc	56640
ctgaaccatg	atcctaaact	ctttcactat	tgagttccc	ctgacttgat	acattggctc	56700
tgtctaggca	gcgggcaagg	ataaccatt	gggcagttgc	acctcagcct	cctgagtagc	56760
tgggattaca	aatgcaagcc	acagctaaaa	aaattttcaa	acctttgtag	gacagacaaa	56820

ttggggaaaa	catttgcaac	aaagggatga	tacacataca	tataaagagt	tctttcaagg	56880
ctgggacacag	tggctcacgc	ctgtaatccc	agcactttgg	gaggccgagg	caggcagatc	56940
acgaggtcag	gagttcaaga	ccagcctggc	caatatggtg	aaaccccatc	tgtactaaaa	57000
atacaaaaat	tagccgggtg	tgggtggcatg	cgctgtaat	cccagttact	caggaggctg	57060
aggcaggaga	attgcttgaa	cccgggaagc	agagattgca	gtgagccgag	atcgccaccac	57120
tgcactccag	cctgggtgac	agagtggagac	tccatctcaa	aaaaaaaaaa	aaaaaaaaag	57180
agttatttca	aattaataag	aaaaataaacc	aacacaattc	aatagaaaaa	tggggaaaaa	57240
agaataggca	ctttacaaag	aaataaatac	aaagcccagt	ggacatgaaa	ctttgccatc	57300
tccttagcag	ggcctggagt	aagatgggga	gaagggaagg	tgcaaaaatt	aaggaggctg	57360
tcactctcag	ggccatgtaa	gtacaaagtg	ggcatatgag	ggtaagtgcc	tccttaaatg	57420
tgtaaattgc	tagagccctg	ctggatggca	gtctggcaaa	atggatcaat	atttttaaatg	57480
tacaaaccct	ggcacaaatga	ttccattttt	aggaactgac	cttatggaaa	cgatcaggca	57540
agtggtccaa	gaaacacatc	taggatgttt	ttaatgtcga	caaattagaa	atgacaggta	57600
aattcaaccc	tacggactga	ctttaaaaat	tgttacatct	ggctgggcat	ggtggctcac	57660
gcctgtaatc	ccagcatttt	gggagaccaa	catgggagga	tcgcttgagc	ccaggagttc	57720
aagcactgct	tgggcaacat	agggagaccc	cgctgcata	aaaaaaaaaa	aaagtataaa	57780
ttagccaggt	tgggtgtgca	tgccgtgagt	tctaactact	caggaggctg	aggaggagg	57840
atcacttgag	ccctagaggt	caagactaca	gtgaactgtg	attgcccac	tgtactccag	57900
cctgggcaat	agagtgggc	cctgtctcaa	aaaagaaaaa	aaaatgttac	atccaggtag	57960
attggcatoc	tgtgtaaaaa	ggatgccatc	ctgtagtccc	agctgcttgg	gaggctgagg	58020
caggagaatc	gtttgagccc	aggaattcga	ggcttcagt	agctatgttc	acaccactgc	58080
acttcagcct	aggcaacaga	gcaagacttt	gtcaataaat	taaaagaaaa	ataaaaaagta	58140
cgctactggt	ctatagtgg	cgtggaaaga	ggctcagggt	atgccatatt	ggtacaagt	58200
aacagaggaa	ccaacatag	tcacatgata	ctatttttgc	cacctgcccg	tgtttatgtt	58260
cacatttga	aatatttgc	caaagtaatg	gtccctattt	cctggtggtg	ggattaattg	58320
caggggattc	ttactttctt	ctttatgcct	gctgcataaa	tacttgaaat	ccttaaatat	58380
tgcttaatac	ttgaaaaagt	gattaaagct	aattttgtct	gagaaaagaga	gtgggagtta	58440
acctgttatt	ctgtaacttc	ctggccccac	cagggttgac	tcctgcagag	cattctccag	58500
gtaaatgttt	ttgccttggc	ctgactgtat	ttcagaacta	ccaggaggtc	gttttgttta	58560
tcaaccaccc	agtggggtca	aaaagaccct	taactcttac	aattccagcc	aaataaacag	58620
aagtgtcttt	cgaaagtcta	gggcctccca	ttactaggat	cagtgtgttt	aggacttccag	58680
ggtagtggaa	agggccttgg	ttccacagag	ctgtctcagg	gcacttaaat	ttccctaagt	58740
gtaaaatgga	cagcttcaac	cgtatcagtg	ttctctacct	ttctcttttc	ttttcttttg	58800
agacagggtc	ttgctctggt	accaggtctg	gagtgcgaatg	gcaagatctc	agctcagtc	58860
cgctcaaac	accaggtcta	atcaatcctt	ctacctcagc	ctcccaagta	actgggacta	58920
caggcctgtg	ccaccatgct	tggctaattt	tttgtagaga	tggggtttca	tcagtgtgcc	58980
caggctggta	tcaaacgcct	gggtcaaga	gatcctcctg	ccccagcctc	tcaaagtgtc	59040
gcgattacag	cgctgagcca	ctgtgcctgg	cttttctcta	aactcactct	cctttttaat	59100
aaagataaaa	ttcttacacc	cttctagtgt	ggtacctttc	tccttattcc	aatagccgag	59160
aagatactgt	ggaactttac	tttctgtaga	ttatatcag	aaaacaatag	ttgtcccca	59220
agctcatttt	ccaaaattaa	ataataatc	taagtatgct	tgtttgtaca	cagtacagga	59280
ctttctgaag	ccacaggcca	cctccagtc	tggctactga	tgccctgggt	ccttctctgg	59340
ctctcaatta	aaagctatag	tgtagtgaat	gagtaccca	gttctgggac	acaacctggg	59400
tgaggtgcg	caggtaaaat	acaggcggt	ctgggggagg	tggccacgc	ctgtaatcct	59460
agcactttgg	gaggccaag	tgggaggatc	aggagttcaa	gaccagcctg	gccaacatgg	59520
caaacatgt	ctatactaaa	aataaaaaaa	ttagcctggt	gcagtggcac	atgcctataa	59580
tgccagctac	ttgggaggt	gaggcacaa	aatcacttga	accaggagg	cggagtgtgc	59640
agtgcgcaa	gaccacgcca	ctgcactcca	gcctgggcaa	cagagcgaga	ccctatctca	59700
aaaaaaaaaa	aaatatagat	acacacacac	acacacacac	acacacacac	acacacacac	59760
acacacacat	atggtgtcct	ggaatctatt	tcctagatct	ggcaacccta	acctagtcca	59820
catttgggac	tctgtctcca	ggcagtgtga	ctataagcac	agtctgtctt	tccttttttc	59880
tttgtctcac	cctctttctt	cttctttcct	tcctccctcc	ttgcctgcct	gctttctcct	59940
tccttcattt	ttcttctctc	cttctctccc	ctccactccc	tcctccttcc	ttcctttatt	60000
ccttacttcc	tcctctctct	tcctctctct	tttcttccct	aatttgtgtca	agtgcacaa	60060
tccttaattt	aaatatgcag	cttgatgaat	ttttacatat	gcataaactc	ctgcaaccac	60120
taccagatt	aaggagcacg	tttccagcat	cccaggaaat	tttctcatgc	ctcttgctgg	60180
tcagtatctc	ccccagaggt	aaccactctt	ctcacagcct	gttattgtca	attaattttg	60240
tatgttcttg	aatttcataa	aagtggaggt	atgcaatatg	agctcttaag	tgtctggctg	60300
cttcttctta	acctaatgac	tgagattcat	tcaggttgct	atatataaca	gtattttccc	60360
ttttcattgc	tgtataatat	tcatttgtgt	gaattttttt	ttggaggggg	gagttttgtt	60420
tcctgaaaac	accacaattt	gtttatccat	cctctgtctc	atagatatatt	ggttgtttcc	60480
agtttggggt	gtaaattcaa	aataaaatcc	taagggtcca	ctaaatgaac	accttctctg	60540
gcaaagggaa	ccccagaaaa	actttaaaaa	ctttgtttcc	agccatgatg	agacaggagg	60600

tcaggcacac	cacattacac	tcccttcctt	ccttttgttg	tttagatata	agaaaagatc	60660
agcatcaatg	ctaaaaataga	gggctgagta	tggtgactca	cacctgtaat	cccagtcocc	60720
tgaggagactg	aggaaggcag	atcacttgag	gccagaagtt	cgagaccagc	ctgggcaaca	60780
tggtgaaact	ctgtctctac	aaaataaaat	aaaataaaat	aaaataatta	gccaggcacg	60840
gtggtgcgtg	tcctgtggtc	ccagctactg	gggaggctga	ggtgagagga	tcgcttgagc	60900
ccagggaagca	gaggctgcag	tgagtcata	tctttccact	gcactccagc	atgggtaata	60960
gagtgcagct	ctgtctcaaa	aaaaaaaaaa	aaagagagag	agattataag	actgacagaa	61020
cagacttttt	gtggcaataa	gataccaaat	tataaacaca	gcctaaggcc	atgtcaggca	61080
agggttaagt	cagggtcccc	tactcttaag	gaataaaacta	tgttctaatt	atgttacaag	61140
atctttcttt	ttctctagca	gcgaaacaag	cactggcctc	agaagaagca	atattaaaac	61200
agttacaact	catctagcac	acagacaccc	aactgacacc	ctgttccctc	agtcataaca	61260
acaactacag	ctttgattga	acaagagact	gagtttggtg	actttctcct	aataaaaaaga	61320
tcactgacta	tggtgctgtt	ctgggtgggt	tacgaaaccc	caacctcatg	tgccctgcat	61380
tcctgaaaag	acattttgat	gtgtagggtt	taattgtaat	acattgattg	attgattgat	61440
caattgattg	attgagatag	ggtcttactc	tggtgcccag	gctggagtgc	agtggcacga	61500
tcacaactca	ctgcaacctc	tgccctcctg	gctcaagcaa	tcctcccacc	tcagcctccc	61560
aagtagctgg	gactacaggt	gcacgcaact	gcgcccggct	actttttgta	ttttttgtag	61620
agacaggggt	ttcgccatgt	tgcccaagct	ggtctcaaac	tcctgggctc	aagcgatcca	61680
cccaccttgg	actccaaaag	tgctagtatt	ataggcatga	gccaccatgg	ctggcctaata	61740
tgtaatacat	ttaaatgtta	agtctccacc	ccaaagtga	catgggtgtg	atgttacatg	61800
cacattttgt	catacacatg	tggtggggcc	accttcataa	atattcatag	cttctcctgt	61860
aactgtctgg	atataatcatt	cagccaaccc	cttcagcaca	aagctcctaa	cccaaccctc	61920
cctccttcaa	agtgcccgtc	tctgttcttg	gtaggaggca	tacttcccag	gccatggact	61980
ggtcaccttg	tggtctataa	ccccctataa	gaaataagat	ttcttctcct	ctctgaattt	62040
acacattttg	gatttttttt	tttttttttt	ttagtttaaca	ggggctatga	acattcttac	62100
agaagccttt	tgattgatgt	gtgttttcat	ttatcttggg	tatatatata	ggcgtgggca	62160
tgatagatat	taggatagcc	atctttaact	tcagtggatg	ctggagcaag	tttctgaatt	62220
tcaactctga	agtggggatg	ataataacag	cacctgcctt	acagggctgt	ttcgagattc	62280
aaagagaaaa	tctgggtgaag	gcaggggtgcg	gtggctcagc	cctataatcc	caccactttg	62340
ggaggccaag	gtgggcagat	cacctgaggt	caggagttca	agaccagcct	ggccaacatg	62400
gtgaaaacct	gtctctacta	aaaatagaaa	aacaatgagc	cagggtgaggt	ggatatgtgc	62460
tgtaaaacca	gccactcggg	agtctgaggc	aggagaattg	cttgaatctg	ggaggcagat	62520
gttgacagtga	gttgagatgg	caccactgca	ctccagcctg	ggcgacagag	tgagactctg	62580
tctcagaaaa	aaataaaaaa	gaaaaaaaga	aaatccaggt	atctagaatt	ggtacaccgc	62640
aatttcaaaa	acgtaaatta	ttgctgtgat	ggcagtgggg	agcatgaaga	tattggacta	62700
acttttatga	atgttcaagt	gtccccatga	tgaattaaac	acacagggaa	ctttataaag	62760
gccatatggt	atataagtga	tacatgacta	ttgtattaaa	attcaaaacta	gttagatata	62820
aagtaaaaag	tggttttcac	cctatccatt	ttttattatt	gaagaaaaaa	aaatatgtca	62880
tagcgtgggtg	gcttatgcct	gtaatcccaa	ccctttggga	ggtcgggggtg	ggatgattgc	62940
ttgaggccag	gagtttgaga	ccagcttggg	caaaatagca	agaccctgtc	tttcaaaaaa	63000
gtaagtaatt	tggttgggtg	ttatggcatg	catctgtagt	cctggctagg	ctgaagcaga	63060
aggattgctt	gagcgcagga	gttcaaggcg	ccactgcact	ctggcctggg	tgacagagtg	63120
agatcctctc	tctctctctc	tctctttttt	tttttttttt	ttttgttttt	tgagactggg	63180
tctcactctg	tcacccaggc	tagagtgcag	tggtctgact	ttggttccct	gcaagctccg	63240
cctccaggtg	caagtgtatt	tcttgccctc	gcctcccgag	tagctgagat	tacggacatg	63300
tgccaccaag	gcccgtcaat	ttttgtattt	ttagttaga	taggggttca	ccaacatgtt	63360
agccaggctg	gtctcaaacg	cctaacctca	agtgatccat	ccacctcggc	ctcccaaagt	63420
gctgcgatta	caggcaagag	ccactgcgcc	tgccctgacc	ctgtctgtta	tcttttcttt	63480
ttcttttttt	ttgttttctt	tttttttttt	agacagagta	tcgctctgta	gcccaggctg	63540
gagtgtgcag	tggtgccatc	ttggtcactc	gctacctcca	cccaccaggt	tcaagcaatt	63600
ctcctgcctc	agcctcctgt	gtagccagga	ttacaggcac	accccaccac	tcctggctga	63660
ttttttgtaa	tttttagtaga	gacggggttt	gcctatgttg	gccaggctgg	tctcgaaactc	63720
ctgacctcag	gtgatccacc	cacctaggcc	tcccaaagtg	tcagaattac	agggtgaggg	63780
cactgtgccc	agccgaccct	cttttaaaaa	aggaaaaaat	actatgcagt	gagtatattg	63840
catgcatttt	cttatttcat	cttcgtcttt	ttatttgatg	atactaaagg	cagggtgttag	63900
aggctggatt	gctaaaagctg	acccaaagaa	tgccctccctc	agggtgtggt	ggctccctctc	63960
tctcaggcct	cagtcttccc	atctgtacag	tgaggtgcct	gcagatctct	gggctctaaa	64020
aatcacagct	ccatgtttat	ccctggcaga	ggaagggcct	ggagtccctgc	tgcttgctgc	64080
tctgggatac	gggagcaaaag	agccacgcac	ccctatggcc	cacacaggcg	tcacctccag	64140
tctctccttg	gcctcatctc	ccagagctcc	tggaatggca	tcgggctggc	ccagggaagg	64200
cctgtcctgt	gcctctcctt	tcccctcagg	ggctgccagg	ctgaccaccc	ccaccgagg	64260
ccaggccctac	agtgcctccat	ggaacgtcct	gacctcctcc	cagggtggca	gcaggaagaa	64320
ggaagaaagg	ggatcctctc	cagctggcca	gagagacaga	ccttcttgtg	ctcatcaacc	64380

ctccaagaat	gcctgccttc	cctccttccc	ccaaggcctg	tccacagggg	cttgagatca	64440
gccagaaaag	tcaggcaact	tttcagggac	tgggagcgag	gtctcccggc	cgggcctggg	64500
tccagtctct	gtgggcagtg	cagtgcgag	ccccaccct	caagccgtgc	cctgtccata	64560
gtccagact	ttgacctgc	actccagtc	gggctggcgg	acagagggct	ggaaacaaga	64620
cgctccagaa	tcaggagctt	cccctcagga	aatagcatcc	tgtgtccccg	cactgcagtt	64680
gtctggtctc	tccagcagtt	tggtagcttc	ggtagtggtg	agatgcacct	ttgagctggg	64740
gacaggggtt	gggagagggg	agaggcaaag	gatttcatgt	cctcccaatg	tcaaagacag	64800
ggctcaacat	tacagcctaa	ggcaggtgac	aggaaaggag	agatccagcc	tctcaaacat	64860
ccagcagaga	gaccataggt	aagtgatttt	tccctcccca	agcctcagtt	tcttcacctg	64920
gaacatgggg	atcataactc	ccctcttaca	gcgtgagtct	gagtgttaaa	agaggtgggtg	64980
catgtaaagt	gcttagagca	gatctaggca	catagcaagt	actcaaatgg	tagttattat	65040
tatttttggt	gggggagttg	gtaggctggt	tctcaactt	ttatagcttc	tgttccattt	65100
caaggataaa	ctctgcaaat	aacttcatga	gaagtgcgg	tgtggtgcaa	ccagggagaa	65160
ctaattatgt	tcattcaaat	gcctcatctc	tggcttactg	atTTTTTTTT	ttaaaaagaa	65220
gtctttcata	ttctttgcta	tgggcacata	gcaatcaaag	gcacagctg	tctcagattg	65280
ccttctaggg	gacaagggag	gtcctaggca	gataaatgca	agactgaaag	acaagcagaa	65340
agcatcaagt	ggcaactgca	tgccaaactgc	ctaaatat	ttttggagca	gtgcagaaag	65400
cggcgataga	actgggtcta	ggctccgaatg	ctgtcccata	ctgactgcgt	aaccttgggt	65460
gggtgacttc	tcctccctaa	acctcagtc	cagcctccag	aatgagggcg	gtaaccttcc	65520
ctacttccta	gagcagttga	gaggattgag	aggattatgt	cggtagctga	tctacaggtg	65580
tctggcaagt	ggcagagacc	aaaatacatt	ggttcccttc	ctgtccaca	cttacacaga	65640
cattctaact	acacacacac	acacacacac	acacacacac	acacacacaa	atataataat	65700
ccagctggtt	tgcattctct	gggatacata	ctccaagctt	gctgggttga	agtaatgatg	65760
taaaacagag	gagaacggca	acactaataa	aaacatcagc	aacaacacga	aaatgtccaa	65820
ccgaataact	gagctgggtg	cgtttaagtc	caaaagctca	ttacctacac	gcataaatga	65880
ttttacctaa	ggctggatct	gccacatctg	acaatctgtc	tctggcttgt	catgaggacc	65940
tcatgcattt	atTTTgtatt	ttaaaacaca	cacacacaca	cacacacaca	cacacacacg	66000
ttgtataat	cagtgtcaac	tttgactcat	atcttgaatt	tttttaaaaa	aagataattg	66060
acttaggact	cacacttttt	tcctttttaa	tttttttttt	tttttttttt	ttgacagagt	66120
ttcactcttg	tcacttgggc	tggagtgcaa	tggcatgatt	tctgcccact	gcaattccca	66180
cctcccaggt	tcaagggatt	ctcctgcctc	ggcctcccg	gtagctggaa	tttcaggcgt	66240
gcaccacat	gccaaagctaa	tttttttgta	tttttgtaga	gacagggttt	caccatattg	66300
gccaggtctg	tcttgaactc	ccgacctcaa	gtgatctgcc	agcctcgacc	tcccaaagtg	66360
ctggaattaa	agacgtgagc	cactgtgccc	ggcctttttg	atTTTccatt	ctatttcctac	66420
caacactcta	aaaattccta	caggcatttt	atTTTatttt	atTTTatctt	atTTTatat	66480
atattttatg	tttgaatgc	aggactctga	agcttcagct	gttcttattt	accggcttga	66540
ttctcagatt	tttcaaaaca	tgtgatttac	tggcaagcat	ggcatttaag	cacctaggct	66600
tatgagtcag	gttggcctgg	gctgtgcctc	tcaccacctg	ggtagtcagg	agctgatatt	66660
ccagttagga	gacaataagg	caaggagctt	tgtcagctct	cataaaagtt	tatagatgag	66720
gtcgggcatg	gtggctcagc	cctgtaatcc	tagcactttg	agagtctgag	gccagcaaat	66780
cacctgaggt	cagaagtgtg	agaccagcct	ggccaaacatg	gtgaaacctt	gtctctacta	66840
aaaatacaaa	aattagccag	gcattgtggt	gcattgctgt	aatcccagct	actcaggagg	66900
ctgaggcagg	agaatcacct	gaaccgggga	ggcagagtct	gcagtgcagc	aagattgtac	66960
cattgcactc	cagcctgggc	gacaagagtg	aaaattcctg	ctcgaaaaaa	taaagtttat	67020
agatgaggaa	cattaggttc	gattaggttc	aaccaactca	tcctgggttg	cctgggactc	67080
tgtatgcactg	acttttagtc	tgaagtctgc	catcctggga	ggacctcag	ccctgggcaa	67140
gctggggagg	ttggtcacc	tcactcagtc	aagttgagca	acttgcccag	ggttacatgg	67200
ctggtgtgtg	cccaagtcag	gctgcgaacc	tgggtctgtc	tgactctcag	cctgggccat	67260
actgtctctt	agattcttca	tggagaatta	ggaaaaatac	agaaagccct	ttattctctc	67320
gccttctcat	tgttaacata	taaaaaatgg	caagcggggc	ggtagcagtg	cacacacctg	67380
aaagcccagc	gctttgggag	gctgaggggg	gaggattgct	tgagcctagg	aattggaggt	67440
ggcagtgcag	tatgattgtg	ccactgcact	ccagcctggg	tgacagagtg	agacctgtgc	67500
tcttaaaaaa	aagaaaaaga	gtggtcagct	ctccggaaat	tatgcagaca	gtcaaaaagc	67560
ccagagaggg	gaattaaact	agccaaggtc	gcacagcaag	gcagaagtga	agccaggtct	67620
gactctgcct	ttctcttctc	ctcttttttt	ttttgaggca	gaatttcgct	ctgttgccca	67680
gactggaatg	cagtgtgtgc	aactcgactc	gctgcacact	ctgctgccc	ggttcaagcg	67740
attctcctgc	ctcagcctcc	cgagtgcctg	ggattacagg	cgctgccc	cgcgctggc	67800
taatttttgt	agttttcagt	agagatgggg	tttcaccatc	ttggccagac	tggtcttgaa	67860
gtcctgcact	cgtgatccac	ccgctcggc	ctcccgaggt	attgggatta	caggcgtaag	67920
ccactgcagc	tggctcctcc	tctctctctt	tgttctctga	atgtctttgt	tctatgtgat	67980
ttttcaaaat	gctaggagac	aggaaggagg	ctgctgtgtg	ttgaggccct	actctgtgcc	68040
aggcgtggta	ccaagaactt	ttgttaaact	tcttatttaa	tccttaaaat	gacctgtgta	68100
gattgggatt	aacctgtgtt	tgcagatgaa	gagcttgtgt	ctccagaggc	aaagtatggg	68160

ggaagaggga	agagagaaga	ccaaggggtcc	ctgagagggg	ctgtcaccta	agccccagta	68220
tccaagctcg	ggctcgaagc	tggaaaggaga	attgcctaga	ggaacgatac	ctttctgttt	68280
gttggttcta	tctccaactt	ggcttctgaa	accccaacag	agtccagttc	ttgtgggctg	68340
gagccgtttt	ccctccttta	taaaactagg	ccatattaag	aatgtcccg	tgtccagggc	68400
cacaggcccg	agttgccagg	agctgaggtc	tgcgggagga	gagttgtgag	tgaagaggag	68460
ggaagttga	atttggtctc	tctgggcaca	aataattctc	ttgttctgcc	tcagcaggag	68520
cctgcagaat	atttccctgc	tgtgcgggct	taagtagctt	caaggttaa	agctggtagg	68580
ctttctaacc	ttctcagggc	ccaatcagcc	ctgtgcccc	aggcaggtgg	agttctgtgc	68640
tgaagacca	agttctgagg	ccagacactg	cgtctgtcat	gctcatagct	gcatttgcta	68700
gctgccagcc	tggcacatgg	taggtgtgca	ttaagcgtgt	gttgagttaa	ctcaaattga	68760
aattaagtca	cagctgtacc	atttaactgg	ctgtgtgact	tcaggtaagt	cacatcacct	68820
ctctgaacca	cagtttctct	ctctgtaaga	cgggactgat	aacagcagcc	cctacctcat	68880
gagagtgttg	ggagacttgg	atgaatggat	gcttgtgaag	cacttagtgc	cggggccagc	68940
tggctcacag	taggtgctcc	acaaatgtca	gtatattact	tcttttgcac	caggcagctt	69000
gttaaatttg	tacgttttgg	catctgttgc	aatttcccat	ccatcccttc	aagcataggt	69060
tattagaggt	tgaagcatct	tgcccaaagt	taacggccag	taggtggcag	agctaggtcc	69120
tgaagccaga	gcccacgca	ctaaccaccg	gcctaccag	cctacagttg	gtcgtgcctt	69180
ctgtcgggtc	ttttctattc	ccagcccaga	aactgggtgt	ctggggacgc	tccccagaga	69240
aagtgtcatc	attcaccagc	cgtgtgactg	tggccaagtc	tcggtcactt	ctccatacct	69300
cagtgtttcc	atttgcaaaa	cgggaacaat	gatattcctt	cctcctaggg	gtcatcggga	69360
aggtcaata	taaaaagggc	ttggtgggtg	ctggcacctt	ctaagccttc	agtggtgggt	69420
ggcaatggcg	ctaaggatga	tggagatgat	ggtgatgatg	gtgtgcctca	accttctctt	69480
cccacaggct	gctgcaatgc	gtgtgggtgg	gattggagca	ggagtcatcg	ggctgtccac	69540
cgcctctgc	atccatgagc	gctaccactc	agtcctgcag	ccactggaca	taaaggtcta	69600
cgcggaccgc	ttcacccac	tcaccaccac	cgcgtgggt	gcccgcctct	ggcagcccta	69660
cctttctgac	cccaacaacc	cacaggaggc	gtgagtgagg	gtcacatagg	gtagcctggg	69720
gtgcccattg	acctaatgtc	gcagagggag	tcagggttcc	catcaccaag	agcaagcccc	69780
ttgtggaagc	tactgatcta	gcataaaata	aagaaaatgc	caggcgtggg	ggttcacgcc	69840
tttaatccta	gcactttggg	aggtcagggg	gggaggatca	cttgaggcca	ggagttccag	69900
atcagcctgg	gcaacgtggg	gaaaccccat	ctctaccaaa	aatacaaaaa	attagccggg	69960
catggtggcg	cacactgta	atcccagcta	ctcgggaggc	tgaggcagga	aaaccatttg	70020
agcctaggag	gtgaagggtg	cagtgaagtg	agattccggc	actgcactcg	tgacagagtg	70080
agactctgtt	tcaaaaagaa	aaaaataaag	aaaagattca	taaataattaa	gccccttgct	70140
ctgtgccaga	tactaggagg	ctttgtctcg	tcttccttaa	actgggtgcc	tgtcaatacc	70200
acatgattgg	tgaatctgga	aaacttctct	tgttttaatt	tatacatttt	tatttatttt	70260
ttgagattgt	gtttcactct	tgtcgcccag	actggagtgc	aatggcgtga	tcctggctca	70320
ctgcacctgc	tgcctcccag	gttcaagcga	ttctctcgcc	tcagcttccc	aagtaactgg	70380
gattacaggc	atctgccacc	acgcctgggt	aatttttgta	tttttagtag	agatggagtt	70440
tcagtgtggc	cagactggtc	tcgaactcct	gacctcaagt	gatctgcccc	ccttgacctc	70500
ccaaagtgtc	gggattacag	gcattgagca	tcattgcctg	ccaaatttta	tctttttaaa	70560
tagagatagg	gtctactat	gttgcccggg	ctggcttga	actcctgggc	tcaagtgatc	70620
tgcctctcct	ggcctcccaa	aagtgtcggg	attacaggca	tgagccatca	tgccttgcca	70680
aattttatct	ttttaaatag	agatagggtc	tcactatgtt	gcccgggctg	gtcttgaact	70740
ctcgggtcct	agtgtactgc	cctccttgcc	ctcccaaagt	gctgggatta	cagggtgtag	70800
cccttgccac	cagctgaatc	tagaaaaact	ctaagtgggt	gaacatctaa	gtgggtggat	70860
ggatgcacag	atttatcaaa	taaattgcaa	aggtcattat	ggtagttag	aaactgccag	70920
atggttcagc	aaatggaaca	cccaatgaat	agcagctcaa	acagattaaa	aaaaaatttt	70980
taagaggcat	cctgtcacc	aggctgaagt	gcagtacat	gatcatagct	cattgcagcc	71040
ttgacctcct	gggtcacaag	gatcctccca	cctcagcctc	ccaagtagcg	aggacacaca	71100
tgcattgctat	catgcctgga	taattttctt	tattttttgt	agagccaggg	tcttctctatg	71160
ttaccaggcc	ttgtctcaaa	ctcctgacct	caagtgacct	tcctgcctca	gcctctgaa	71220
gagctgggat	tataggcatg	agccactgca	cccagccaga	attttaattc	acacagctgt	71280
aaaaaaccaa	tgattctgat	cagtgggcag	tgattcgggg	cccaggctcc	ttccatctaa	71340
tggctctgcc	gtttttccac	gtgcttttga	ggtcacctca	atgtcaccat	tcacatgggc	71400
tggtgactgc	agaaggatca	tgcaggacca	cacgtgcaga	gtcttttagag	tccttgggcc	71460
agaaacaagt	cacacagtca	catttagctg	cagaggggtc	taggaaatgt	aggtgagctg	71520
tgtgcccagg	gggaggagga	aaagtgttag	gagcggcaag	tcgatgtctg	ccaccagtgt	71580
ttacaaggag	gggtgcttgc	agccagactg	aacagtgtgg	ctcataatcc	ccaaagccag	71640
gtcaaggact	tcctgcaaac	tcacagcca	tgtaattcca	tgttgagggt	gcactccata	71700
tggttatgat	ggggcatcct	tcattccctc	tcttctttat	tctattaatg	gggaaatatt	71760
ggaaaattta	ggaggagaa	gacccaaggc	atttggggag	ttgcaggagt	gaacgtgggtg	71820
gatttctggg	ttttggacac	acccaagct	cctgatcatg	ccacagcccc	atgccagctg	71880
acctaagggt	ttttgcccag	ctcagggcag	tgggtgatcg	aactcttcat	gaccttcca	71940

gggactggag	ccaacagacc	tttgactatc	tccctagcca	tgtccattct	cccaacgctg	72000
aaaacctggg	cctgttccca	atctcgggct	acaacctctt	ccatgaagcc	attccgggtg	72060
gtgaacagtt	cttgaccatg	agggatgagc	acccagggct	ggggtagtga	gggtgggtgc	72120
agcagagcct	taatcacaga	tgagggcggg	gtgctttgag	tctcgtaggc	aacagactcc	72180
tgggttcaaa	acagggtttg	tttaaatctt	atctttgctt	tgaaaattat	ttttgtttta	72240
cattttgctg	ttaaattggc	agaggacaaa	gaatcttctg	atgcccaggg	gaaactagcc	72300
tttgattagc	atggctaaaa	tacaaacatg	ttctgcagtg	acgggcactt	ggtgctgaag	72360
ccaaaagggt	tcaagtgcc	ctgaaggctc	caaggctttt	tatcaagaag	gaataaaata	72420
ctcatcaaa	caaaaactgc	caaaacattt	attatgtgcc	aggtccagtc	ctaactaatt	72480
tacagctagc	gactaattta	attctcttta	taaccgggag	gtaagggtcg	tccttatcct	72540
cacttaataa	atgagaaaac	ggaggctcca	agaaatggag	taacttgccc	aaggccacar	72600
agctcgccag	tggcagagct	gggatttgaa	cccaggccat	ctgtgactcc	atgggtgtca	72660
gtgtgctaac	aggaacagca	cagccctggg	acggtttgct	caggctcctt	ggagaggggtg	72720
gtctggcgct	gtgcccagag	ccccgtgcca	gctctcaagg	ttcattcaac	ctttggcact	72780
gtgctaagg	ctttatccac	attatctggt	accttccatg	ggaccaagag	tatttttttt	72840
tttttgagac	aggtgtccac	tgtattgccc	aggtggaggt	gcattggcat	gatctcggtc	72900
cactgcaacc	tctgcttcc	gggttcaagg	cattctcctg	tctcagcctc	ctgagtaatt	72960
gggattacag	gtgcgcaact	ccacgcctgg	ctaatttttg	tatttttaag	agatgggggt	73020
tcactatgac	ggccgggctg	gtctcgaact	cctgacctca	agtgatctgc	ctgccttggc	73080
ctcccaaggt	gctgggatta	caggcgtgaa	ccactgcacc	cggccaagag	tgattattaa	73140
ctccatgata	cagacaagga	aactgagctc	cagagaattc	aagtagcaag	tgatgaggct	73200
ggggtctctg	acactatgct	ctgctgtctg	acactatgct	ctggtgcttt	ctctcatccc	73260
cggggactct	cactgtttct	gctttctctc	ccctatttct	gacttttccc	ctataactca	73320
ccctcggctc	tactcttacc	cttaccataa	ataggggtta	agaacatgaa	ctctggaact	73380
aagctgtatg	ggttaaaatc	tcaacaccac	catttattag	ctgtgtaatc	ttagacaagt	73440
tatttaattct	ttctaagcct	caattggctc	atctgtaaac	tggggaaaga	atagcatcca	73500
cccaatggc	ttcttgtgaa	gattaaatgg	accagtataa	gaaaatgctt	ggaacagtgc	73560
cttatatgca	cttagcatta	cataagctc	tgtcattatc	atcttttttt	ttttttgaga	73620
tgaagtctcg	ctctgtggcc	caggctggag	tgcagcggca	caatttcggc	ttactgcaac	73680
ctccagctcc	cagattcaag	caattctctc	gctccaatct	cctgagtatc	taggattaca	73740
ggcatgcacc	accatatctt	gctaattttt	gtattattat	ttagtataaa	cagggtttca	73800
ccatgttggc	cagactggtc	tggagctcct	gacctcaggt	gatccaccca	tctcagcttc	73860
ccaaagtgtc	gggattacag	gtgtgagcca	cctcgcctgg	cccattatca	ttattattga	73920
cttccatccc	accagtgcc	ccctttgtcc	ttcctcttca	ggacccttcc	tgggaaggaca	73980
cagttctggg	atttcggaag	ctgaccccca	gagagctgga	tatgttccca	gattacgggt	74040
gagtttattg	tcacaggcaa	aggggactgg	ggcctgacga	gttagcagac	ctgtccagaa	74100
ggcagcagag	ccagatttcc	gctctacccc	aggccctggg	accctggctc	taaggtttct	74160
cctggctcct	ggctcagctc	cttcagagag	gctaccactc	caaacctggc	cttgggctgg	74220
gaaggtaggg	ggatgaaat	cacagatctc	aagcccagaa	gctccatata	accatattgt	74280
ttttagatg	aagatactga	gtttcagaga	ggctaagtga	cttcctaagg	tcacacagcc	74340
aagtggccaa	actgggatcc	caaccagctc	gtatgacccc	acaccctctc	tttcttttct	74400
ctacagcctg	atgctctctc	ggtcttctcc	tcaccccaac	ccacaccaca	cctgaatccc	74460
ctctacgaat	gcacattcaa	tctccacttg	cattttccaa	tgtcagatat	ggccttttct	74520
gatagaaaaa	attgagctga	tttcccttgc	aaaccacgtc	cccccttgaa	cttcacgtag	74580
tggctcctgg	actacccttt	ggggccacag	aacaacattg	ctcccacctc	catttcacag	74640
ccttcaataa	tagctctgat	ttttaccttt	atttccacct	tttgcttact	gtgactctag	74700
ctatggctgg	ttccacacaa	gcctaattct	ggagggaag	aactatctac	agtggctgac	74760
tgaagggtga	gattttaagc	ttcactttga	gggaggtacc	tcccagagac	caagttgtag	74820
tgggaagatg	ttcgtgggct	tcctcagca	tggactaacc	cccagggttg	aagaataccc	74880
ttaggcctgg	tgtgggagct	atccttggtc	ctgatcaccg	ctgggcacag	aggcaatgga	74940
tccctgagct	agctgagcat	cagaaccacc	tgggcagctg	tttacacatg	atgtccatta	75000
acaacctctt	tcaaatccct	aatgtttgtg	taatagtttt	agatgtattc	ttttaagggt	75060
tccagataca	tttacatcat	ctgcaaaaac	taagctgctc	tttattttta	tctccctctc	75120
tctttttttt	tcttctctaa	ctgctttggc	caataacctc	agaacaatgt	tactcataca	75180
gatgatagtg	gatgtctttg	acttgcttct	catgttaata	ggaatcttgc	agtgttctaa	75240
attagcaaac	actcacttga	gagatacatt	ggtatttata	ttcacataca	ttcatattaa	75300
gggagattcc	ataccttttt	tgtgtgtgtg	agatggagtc	tgcctctgtc	acccaggtcg	75360
gagtgcagtg	gtgcgactct	ggctcactgc	aagctctgcc	tccctgggttc	atgtcattct	75420
cctgcctcag	cctcctgagt	agctcgcgact	acaggtgccc	gccaccacca	cacctgcta	75480
atctttttgta	tttttagtag	agatgggttt	tcacctgtt	agccaggatg	gtctcaatct	75540
cctgacctcg	taatctgccc	gcctcggcct	cccaaatg	tgggattaca	ggtgtcagcc	75600
accacgcccc	gcctgattcc	atacattttt	tatatattac	tgttttttaa	agatttttag	75660
gccaggcatg	gtggttcata	cctgtaatcc	tagcactttg	agaggccgag	gtgggcagat	75720

cacttgagcc	caggagttca	agaccagcct	gggcaacatg	gcaaaaccct	gtctctacag	75780
aaaaattcaa	aaatcagcca	ggtataatgg	tgcattgctg	tagtcacagc	tacttaggag	75840
gctgaggtgg	gaggatggct	ttatcccggg	aaggagaggc	tgcagtgagc	tgtgatcatg	75900
ccactgcact	ccagcctggg	gggacagggc	gagaccctgt	ctcaaaaaaa	aaaaaaaaa	75960
ttaaaaaat	atggaatata	tagtggcctt	tatcagatga	cctcagaaga	ttttttttta	76020
atgtagattt	taggacccca	ctctatacct	gctgaattag	aacttctggg	ataagggtca	76080
taaatttgct	ttttttcatt	tttttgagac	aaaatcctac	tttgtcacc	aggctggagt	76140
gggatgtagt	ggtatgaaca	caactcacag	cagcctcaac	ttcctgggct	caaagtatcc	76200
tcccacctca	gcctccaaag	tagctgggac	cacatgcatt	tgcacacatg	cctatctaat	76260
ttttaaatat	ttttgtagag	ataggggtct	actatgttgc	ccaggctggg	ctcaaaccct	76320
tgggctcaag	caatcttcct	gcctcagcct	cccaaagtgc	tgggattaca	ggcgtgagca	76380
aacaggccta	gcaaaaattt	gcattttaag	aagcttcctg	gcgattctaa	ttatcagcca	76440
tgtttgggaa	tattgtact	aagacatggc	tatttctcct	aacctgggga	cacatgaccc	76500
ttgtccagtc	ttttccagga	aaaacatgcc	ctcaagatgt	ttttctatct	tgaggaaatg	76560
atggaaatga	agggtatgct	tcacctcctt	tttggcctat	ttcctgttct	ttcctgttct	76620
ttggatgttt	ctagtgtatt	tctttcttct	ttcttttttt	tttttttttt	tttgagacag	76680
agtcttgctc	tgtcaccacg	gctggagtgc	agtggcgcaa	tctcggtctc	ctgcaagctc	76740
cgctcctggg	gttcatgcct	ttctcctgcc	tcagcctccc	gagtagcttg	gaatacaggc	76800
gcctgccacc	acgtctggct	aatttttttt	ttgtattgtt	agtagagacg	gggtttcacc	76860
gtgttagcca	ggatgggtct	aatctcctga	cctcgtgatc	cgeccacctc	ggcctcccaa	76920
agtactggga	ttacaggcgt	gagccaccgc	gcctcgctgt	ttctagtgtg	tttctaactc	76980
tgatagatgt	tttctctatg	ggatgtttta	aaggagggtg	gatgtcctca	gccacctctc	77040
ctctctcatgc	ccggtctctg	acaaaaggga	atttggcact	ggtacaactc	tccccttctc	77100
tactctgaat	ctcattgcct	ttgctgttac	aaagcaatgt	gggtggtcata	ggaagtgtct	77160
ggggctaaga	ggcctgggtt	tgagttccaa	ctccatcatt	gactcactct	atggccttca	77220
gcaaggccct	tccccactc	catctgcccc	acaaggggct	tggaccatct	ctggtttctc	77280
aaaggagatt	ttgtggacca	ccagtccagt	aggtgctcat	gagctgattt	gatgacacag	77340
ccatcttctc	aagcagcatc	ctgtgcaact	aacgtccgca	gaaggttggt	tggggaaaagg	77400
tcctctgtgc	acccttcttg	gtgggatggg	ggcagatagc	tgaatactgg	gctttttgat	77460
gtgtttgato	atcccagggt	aactgagagg	ggagtgaagt	tcttccagcg	gaaagtggag	77520
tcttttgagg	aggtgagttg	cagggctgat	gcggtggatg	gggcagggaa	gaagtgggga	77580
ggcctctgct	tcttgctgct	gagtcggggg	ctcccttctc	aggctcctag	ggtccccaca	77640
ggcctgcctc	agcacccttg	cccagaagc	actcaggtat	tctgaaggga	ggaagtctct	77700
gccttcatgt	tggtagtggg	aacaaaggaa	cactgggato	atggtggcca	ttaggagctg	77760
atttatatct	gagactcaat	gagttttggg	tctagagagc	tggccgcatt	ttctcagtgt	77820
cagctgcact	ccaaggctag	aacttggttg	cttcttagcc	ctaccgacat	ctgtgttggt	77880
ctttctgcaa	atccagggcc	ctcagctgac	ttcactctaa	agaagcacca	ccaccttaaa	77940
taatgacagg	aaaagccacc	atctccaggc	accagcaaaa	agagctttac	tgtatggctt	78000
cattcaatcc	cagcatctaa	aacctgctt	ggcacaagga	aggcgctccg	tacatgtagc	78060
tactagtgtc	atgtcatgaa	gactaacctg	ctctggtcag	gccctgatgg	acaccgaaga	78120
tacatggctg	acccaatgca	gtcctcatte	tcagtcatte	actcaggaac	aatagtagcg	78180
tcttgcaatg	tgtgtgtccc	ttaaacttact	cgtggtgaga	gtcactgggg	ctgggttggt	78240
gagcttaggg	gctcacgatg	cgtgcttgag	atgagatcat	ctcatctgta	gacagagctg	78300
gggttccaag	gtgtctctg	caaatgtctt	ggcagagtga	aaggcaagag	aataaagta	78360
aaaggagtc	gaaggagaaa	gagaactctc	tctgcttctt	ttctgacttc	ttttgggagg	78420
ttccaggga	atttccccca	tccaaagaac	tgtttttaca	ccacttttat	attcagagtt	78480
gtgcaggagc	ctcataacag	cctatgaaca	gccatgggca	gcctcatttt	acaggggcag	78540
ctgagaatta	aggaggtaac	cagacatttt	caaggtcaca	cgtcagataa	atggcagtat	78600
gaaaatttga	agccaggccc	ctctgattec	tatttgagac	ctctccccac	tgttcatcag	78660
ggagttagaca	gatttgagggt	agaagaaagg	ggaagagaga	acagggggata	ccagggtctt	78720
ccccaccttt	catccccac	tacctggttg	gttgctacca	ggtggcaaga	gaaggcgag	78780
acgtgattgt	caactgcact	ggggtatggg	ctggggcgct	acaacgagac	cccctgctgc	78840
agccaggccg	ggggcagatc	atgaagggtg	gtgtgagggt	gagaccctta	ccttttggtt	78900
ataggaagat	cattctgcat	gcttatttca	tcctcaaga	tcatggacaa	atcagggaaca	78960
tctgttagag	gaaccccccg	gactgcaggg	aattgacatg	taaaaaaac	aaacctgtcc	79020
cacccccatt	gctctcttct	aggatttctt	cttgatcggt	aagcatgcat	gtatgcgctt	79080
gtacctatgt	gggagcagca	tatgcctgta	ttgcaataaa	aatagcaaac	attagagtgt	79140
ttaccaagcg	cgagatacag	tcctaagcac	tttattgtgt	ttattattat	tatttaattat	79200
taatttgtgt	ttatttatta	tcattgttat	tattattttt	gagacagggt	atcacctcat	79260
tgccagggtt	agagtgcagt	atcttgatca	tggctcactg	tagccttgac	ctcccaggct	79320
cccaccttag	cctactgagt	agctgagact	acaagcgcat	gccaccacca	tgctcagcca	79380
atttttttat	tttttgtaga	gaaaggattt	caccatattg	ctcaggctgg	tctcaaacct	79440
ctgggctcaa	gtgatcccc	caccttggtc	tgtcaaaagt	ctgggattac	aggcgtgagc	79500

caccacgctc	agcctattgt	gttaattaat	ttagtgatgg	ccacagccct	tcgagctggg	79560
tactaccata	tcgttattgt	catcttacag	atgaagaaat	tgaggcacag	aggagttaag	79620
taacaggcac	aagttcacac	ggtagtacgc	agtgaattg	ggattggaat	ccaggcaacc	79680
tggcttaaga	gcctgtgcgt	gcaagcattg	ttccatgcct	cctcttgctg	tgtgtgtgca	79740
tatgagggtg	tgtgtgtgtg	catgtatgtg	tgtgtgtgta	tgtaagggtg	tgtgtgcata	79800
tgtgtgtgtg	catgtgtgtg	gggtgtgtgtg	catgtgtgtg	gggtgtgtgtg	catgtgtgtg	79860
aggggtgtgtg	catatgtgat	gggtgtgtgca	catatgtgtg	gggtgtgtgtg	catgtgtgtg	79920
aggggtgtgtg	catatgtgtg	atgggtgtgtg	tgcacgtatg	tgggggtgat	tgtgcatgta	79980
tgtgagggtg	tatgtgcata	tgtgtgatgg	tgtgctgtgca	tgcacacccat	gtgaggggtg	80040
tgtgtgtgtg	catgtgtgtg	aagggtgtgtg	cgcattgtgtg	agtgtatgta	cgtgtgtgat	80100
gggtgtgtgtg	tgtgagggtg	tgtatgcatg	tgtgtgaggg	tatgtgtgtg	catgtgtgtg	80160
aggggtgtgtg	catgtatgta	aggggtgtgtg	tacatgcatg	tgtgtgaggg	gtatattgtgt	80220
ggatgcatgt	gagggtgtgt	gtgtgcatgt	gtgtgagggg	gtgtgtctga	gggtgtgtgt	80280
gtgcatgtgc	acctgtgagt	gttcataggt	gtgcaggtgt	gtgtgcttct	gtgtgtaggg	80340
gtgcgtgtgtg	tggttccctaa	tgtgggtgtaac	tggggtgtaac	aaccaaataa	gtgactgaag	80400
cataagctctc	aaatcatcga	ggtttatgga	gccagcttga	gggcgcaccc	aggaaaaacg	80460
cgagtcacag	atgcacctgt	gactcctttt	tccaaagagg	ttctcaggag	atttagtctt	80520
tatacatctt	ctttaaaaaa	aaaaaagtga	gagaaggggtg	tagcagcgag	agaatgattg	80580
catacttgtg	aaacttttagt	tagtgcccag	taaactctaca	ttttacataa	gatgaagggt	80640
tgggccaggc	gtgggtgactc	acacctgtaa	tcccagcact	ttgggaggct	gaggcagggtg	80700
gatcacgagg	tcaggagttc	gagaccagcc	tggccaacgt	ggtgaaaccc	catctctact	80760
aaaaatacaa	aaaattagct	gggtgtgtgtg	gcggtgtgcct	gtaatcccag	ctactcggga	80820
ggctgacgca	ggagaatcgc	ttgaacccgg	gaggcagagg	ctgcagtgtg	cggagactac	80880
accactgcac	tccagcctgg	caacagagcg	aggctgtctc	aaaaaaaaaa	aaaaaaaaaa	80940
aaattgaagg	tttgaaggaa	aaaggaatgg	aggaagttct	gtatctggga	agataagctt	81000
gtcattgatg	ttatcagtgt	ggagtctgtt	gaaagggctg	gtttctgctt	accctctagg	81060
gaagaaagcc	taacttttgt	caggctcattg	agggagggga	tataatgaga	cgtgtcggac	81120
ctcccttccc	ccgcagctg	tgaactcagc	tccaaggttt	ctctggggct	cctggggcca	81180
agaggggggtc	tgttcagtcg	gttggggact	tagaatttta	tttttatttc	tcatgtgtat	81240
gcatttccat	gtgtgtactg	gtgcttttct	tcggacatgt	gggtgaggag	aaacaatgct	81300
tcaggggagca	gggggtggctg	ccaattaggg	cagctcttcc	tgcaagaggc	aagcagtcag	81360
gtgcagactt	gggccatagt	gtcatgagag	gtcttataag	gaatcagcct	ggccactctt	81420
gtcaggacat	ctggccacag	aggggagcaa	gggcagccac	attgactcac	ctccgctgat	81480
gagactttcc	tgccctgaat	caacaggtgg	acgccccttg	gatgaagcac	ttcattctca	81540
cccatgaccc	agagagaggc	atctacaatt	ccccgtacat	catcccaggg	taaaattgga	81600
ctgttctcgg	gcagaagagt	gggtcccttc	atgcccctct	catgaccctg	ctgcctcccc	81660
caagttcctt	actccctgca	gttgttccct	ttcaattgtt	ttatgtactt	agctattttt	81720
tattattatt	ttttgagaca	gagtttctct	cttattgccc	aggctggagt	gtaattgggtg	81780
gatcttggct	cactgcaacc	tctgcctccc	aggttcaagc	aattatcctg	cctcagcctc	81840
ccaagtagct	gagattacag	gtgcccacca	ccacatccag	ctaatttttt	gtatttttag	81900
tagagacagg	gtttcaccat	gttggccagg	ctggctctga	actcctgact	cagggtgatcc	81960
acctaccctt	gcctcccaaa	gtgctgggat	tacaggcgtg	agccaccgtg	cctggccccc	82020
tttcaatggt	tttagtgagt	ttgagctact	gaatccctgg	gaaggcagac	tcagcctcga	82080
ctgaggtcta	ccgtgaacat	tcttttggat	gacatagtgt	gtgatgctgg	agacaaaggc	82140
agtggatgta	atgtggtgac	actaaaagtg	gtatgtaggt	ggctcacgcc	tgtaatccca	82200
gcactttgctg	aggccaatgt	gggaggatct	cttgagccca	ggagttcaag	accagcttgg	82260
gcaacatggc	aagaccccg	ctctacaaaa	atacaaaaat	tagccggggc	tgatgggtgta	82320
tgccataggt	cccagctatt	cgagaggctg	agatggggag	attgcttgaa	cctgggaggt	82380
tgaggatgca	gtgagccatg	ttcacaccac	tgtactccag	cttggggccac	agagcgagac	82440
cccatctcaa	aaaaaaaaaa	aagtgggtgtg	aatggcaata	atgggagtgg	gaatgggaat	82500
gggtgattggg	gctgatgggtg	atgataatgt	taacgggtgga	gatgacaatg	tcactgaaac	82560
cagtgggtggt	gttcatggga	tgacaatatt	gttgatagcg	gaatgggtgg	attagggata	82620
atattgtatt	gatgggggag	acagcgttca	tgggggtggt	gattagcgta	agagttgtag	82680
agtgggtgatg	ttaatggagg	tgggtctgggtg	ctgatgagga	gatcaatggt	gatgaagggtg	82740
tgattgggag	tggggatgggt	agctgggtgct	gatgggaaatg	acactatcaa	tgatgttaat	82800
actgtagcag	agctgacagt	ctcaaaggca	atgttaataa	catggttgca	ccaaccatgt	82860
tatctcaatg	gcgatgttac	tgggtgtcgtg	gagatgacaa	tatcaatggc	aatgtttagtg	82920
gtgggtgggtg	aatgatgaat	gcagttgggtg	gtgatgacct	attaatgata	gtagcaaaga	82980
caatgttgtt	gatggagatg	acaacattga	tgggaagtgg	gatgggaagag	ttcgttgggtg	83040
gtgttgatgg	tgatgacagt	ggcaattgag	gtagtgtatg	tgggtgggtgt	agcagagggtg	83100
acaaggttga	tggtaatgac	ctttattcat	ctcagagcct	tcatttttct	tcactccttga	83160
ccctcctcat	ttgtatctag	gacccagaca	gttactcttg	gaggcatctt	ccagttggga	83220
aactggagtg	aactaaacaa	tatccaggac	cacaacacca	tttgggaagg	ctgctgcaga	83280

ctggagccca	cactgaaggt	aaggtagggg	ggagtagcag	tgccctaacc	caaggtcgtg	83340
ggagcttggg	aatgaggaca	cttcaggagc	ggaagatgcc	accgctggga	taactgggca	83400
aattaattcc	agcaagggat	gtggaacata	acagaatttg	ataatgtaca	gggaagttct	83460
tgctatgggc	taatgaatcc	tgtctggcca	tggtcgagag	cccttggttt	tcacatttgt	83520
ctgcgagtga	tgatgacagt	agtgatgggtg	atgaggatga	gttggtagctg	atgggtgagga	83580
aaatgctgag	aatggtaata	gtgatgggtga	taaggtgggtg	acagttgtta	aaattatggt	83640
ggtggctgat	ggtgagggta	gtggttgatg	atggaattgg	tggaaagggtg	gaagcagtaa	83700
tggtaatgat	gttggtagct	gataaagatg	gtgttgggtg	tagtgggtgat	tgataaagat	83760
gactgtgatt	atattagtgg	tggtgggtgat	gagattctaa	aagctaactc	cctactacct	83820
aaaaatggca	gcaggaaaaa	aaaatccaga	aatgagtgat	cagcactttt	ctttccagaa	83880
tgcaagaatt	attggtgaac	gaactggctt	ccggccagta	cgccccaga	ttcggctaga	83940
aagagaacag	cttcgcactg	gacctcaaaa	cacagaggta	tgctcccatg	gcaaggaaag	84000
taatgccctc	ttccactcct	cagatggctc	tggtcatctt	agggarcagt	catgtctgat	84060
ctcaagttcc	acacaggtc	catagcaggc	aggggcagtg	gtggctaata	ttccctctc	84120
tataaatggg	gaaactgagg	ctcaatgatg	gttaaggacc	tgctcaaggt	tacatagagg	84180
ggcagtgggtg	atgttaaatgg	aggtgggtgct	gagatgatca	atgttgataa	tggtgtgatt	84240
gggagtgggg	atggtagctg	gtgctgatgg	aaatgacact	atcaagtatg	ttagtaccac	84300
agcagagggtg	acgatctcaa	aggcagtgtt	aacatggctg	cactaactgt	ctcattggca	84360
atattaatcg	tgtggcagag	atgacagtat	caatggcagt	gttaatgatg	gtgggtgaaat	84420
ggtgaatggg	gttggttttc	taaagtctgt	ggtcaataaa	caggaaaatg	tgtacttact	84480
ggatgtgtac	ttcgtgtcag	acacagcagc	aagtccatta	catgaatgac	cttattaaat	84540
ctcctctgga	gctctttggg	atagggacag	ttctccctat	gcttcggatg	aggaaaactgg	84600
ggtgaattaa	gaggtgaagt	cacttgccca	agtcagacca	ctgggtggaag	gcagggtctgg	84660
gatgtgattt	gaatttgact	ccaaggctat	ttccagatat	ccattttgtg	gctgccccat	84720
catctcttgc	aactgttcca	gggggtcccc	accattccac	ccgggtgcca	agagaagctc	84780
aggtggcatc	tggtcttgcc	caggactctt	cggtgagctc	ctgagtcttc	cagggcagaa	84840
gagcttcatc	tattctttcc	actgtccctc	tcggacctgg	ccacctctc	tcttgccctc	84900
cctaggatcat	ccacaactat	ggccatggag	gctacgggct	caccatccac	tggggatgtg	84960
ccctggaggc	agccaagctc	tttgggagaa	tcctggaaga	aaagaaattg	tcacaaatgc	85020
caccatccca	ctcctgaaga	ctccagtgc	tgctgcctcc	ccccacaaga	actccctctc	85080
ccctcagcc	aatgaaatcaa	tgtgtctcct	cataagccat	tgcttctccc	tcacttcttt	85140
cctcaaagaa	gcatgagggtg	agagaaagcc	acaaagtcag	tgcttgagga	agggttcagc	85200
ccaacatggg	gcccctctca	tcactgaaat	ccctctacct	tctctgggtc	tggtattata	85260
aagaacagct	gaggctgtca	ttccatgagt	cttcagaaga	aaggacagct	cagaaaatca	85320
aagaggccaa	ctgcccagag	ccacagaaaa	tgaggataaa	ttgaggctaa	gtaacctgat	85380
tacaagttgt	actaacatat	taaaggttct	gaaaagtcct	gcagcaaaaga	caactatctg	85440
atgttgttta	accagtgct	tgctaaacct	atctggctat	ggaactcttt	tgcccagagc	85500
accatgaat	gcatgacac	aaatctgaga	aaatgctgga	acagattttg	ttgtatctgt	85560
tgtgtttgtt	gtaggagggt	atacatataa	ctgggggtgtg	gagggggcag	agaggtgagg	85620
cactgaacta	gtaacacatg	gtgtttgttc	cacatctaga	attccaaatg	gcatcagcta	85680
ttcaccgagt	ggccccatga	gcaccacgta	acctttgagg	agggggccact	ggagggatca	85740
ttccacaagg	aacccttcca	tagagaactg	tttttagtcca	ttttctgttg	cttataacag	85800
aatatctgaa	actggagatt	tttttttttt	tttttttgaga	caggatctca	ctctgtcacc	85860
caggctgggtg	tgcatgggca	tgattttggc	tcactgcaac	ctccgcctcc	caggctcaaa	85920
tgatcctccc	tcctcagcca	cccagtagc	tgggactaca	ggcgcttgct	accatgccca	85980
gctaattttg	tgtgtgtgtg	tgtgtgtgtg	tgtgttttgt	agagagtgtt	ttgtagagac	86040
tggttttggc	catgttgtcc	aggctggcgt	tgaactcctg	ggatcaagtg	atcctcctgc	86100
ctcagcctcc	aaagtgggtg	gattataggc	ataagccacc	acgcctggcg	gaaactgtgg	86160
aattaataga	gaaaagggaat	ttatttatta	ccgttataga	gtctgagaag	tccaaggttg	86220
aggggccaca	tctggtgaga	gccttctctc	tggtgggtgc	agaggtgggg	actctctgca	86280
gagtcacag	gaggcttagg	gcatcacgtg	gtgagggggc	tgattgtgct	aatgtgctag	86340
ctcagctctg	tccttgtct	tagaaagcca	ccattttcct	tcccaagatg	accattaat	86400
ccattaaacct	aataacccat	taattgataa	atggattaat	ccatttatga	gagcagcgct	86460
cttaggatcc	aatcacctct	taaaggcgcc	acctctccag	accaccacta	aggtgggtgga	86520
ctaaggacta	agtctcaacg	tgagtttttg	cagggacgtt	taagcaatag	caagaactaa	86580
actcaccaag	ca					86592

<210> 2

<211> 1573

<212> DNA

<213> Homo sapiens

<220>

<221> 5'UTR
<222> 1..143

<220>
<221> CDS
<222> 144..1187

<220>
<221> 3'UTR
<222> 1188..1573

<220>
<221> polyA_signal
<222> 1549..1554

<400> 2
 tgcactccag tccgggctgg cggacagagg gctggaaaca agacgctcca gaatcaggag 60
 cttccctca ggaaatagca tcctgtgtcc ccgcaactgca gttgtctggt ctctccagca 120
 gtttggtact tccggctgct gca atg cgt gtg gtg att gga gca gga gtc 173
 Met Arg Val Val Val Ile Gly Ala Gly Val
 1 5 10
 atc ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc 221
 Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
 15 20 25
 ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc 269
 Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
 30 35 40
 acc acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac 317
 Thr Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp
 45 50 55
 ccc aac aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat 365
 Pro Asn Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr
 60 65 70
 ctc ctg agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc 413
 Leu Leu Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe
 75 80 85 90
 cta atc tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc 461
 Leu Ile Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser
 95 100 105
 tgg aag gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg 509
 Trp Lys Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu
 110 115 120
 gat atg ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att 557
 Asp Met Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile
 125 130 135
 ctg gag gga aag aac tat cta cag tgg ctg act gaa agg tta act gag 605
 Leu Glu Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu
 140 145 150
 agg gga gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg 653
 Arg Gly Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val
 155 160 165 170
 gca aga gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct 701
 Ala Arg Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala
 175 180 185
 ggg gcg cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc 749
 Gly Ala Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile
 190 195 200
 atg aag gtg gac gcc cct tgg atg aag cac ttc att ctc acc cat gac 797
 Met Lys Val Asp Ala Pro Trp Met Lys His Phe Ile Leu Thr His Asp
 205 210 215
 cca gag aga ggc atc tac aat tcc ccg tac atc atc cca ggg acc cag 845
 Pro Glu Arg Gly Ile Tyr Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln
 220 225 230

29

```

aca gtt act ctt gga ggc atc ttc cag ttg gga aac tgg agt gaa cta      893
Thr Val Thr Leu Gly Gly Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu
235                240                245                250
aac aat atc cag gac cac aac acc att tgg gaa ggc tgc tgc aga ctg      941
Asn Asn Ile Gln Asp His Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu
                255                260                265
gag ccc aca ctg aag aat gca aga att att ggt gaa cga act ggc ttc      989
Glu Pro Thr Leu Lys Asn Ala Arg Ile Ile Gly Glu Arg Thr Gly Phe
                270                275                280
cgg cca gta cgc ccc cag att cgg cta gaa aga gaa cag ctt cgc act      1037
Arg Pro Val Arg Pro Gln Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr
                285                290                295
gga cct tca aac aca gag gtc atc cac aac tat ggc cat gga ggc tac      1085
Gly Pro Ser Asn Thr Glu Val Ile His Asn Tyr Gly His Gly Gly Tyr
300                305                310
ggg ctc acc atc cac tgg gga tgt gcc ctg gag gca gcc aag ctc ttt      1133
Gly Leu Thr Ile His Trp Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe
315                320                325                330
ggg aga atc ctg gaa gaa aag aaa ttg tcc aga atg cca cca tcc cac      1181
Gly Arg Ile Leu Glu Glu Lys Lys Leu Ser Arg Met Pro Pro Ser His
                335                340                345
ctc tga agactccagt gactgctgcc tccccccaca agaactccct tctcccctca      1237
Leu *
gccaatgaat caatgtgctc cttcataagc cattgcttct ccctcacttc tttcctcaaa      1297
gaagcatgag gtgagagaaa gccacaaagt cagtgcctgg agaagggttc agcccaacat      1357
ggggccctc tcactactga aatccctcta ctttctctgg gtctggcatt ataaagaaca      1417
gctgaggctg tcattccatg agtcttcaga agaaaggaca gctcagaaaa tcaaagaggc      1477
caactgcccc gagccacaga aaatggagga taattgaggc taagtaacct gattacaagt      1537
tgtactaaca tattaagggt tctgaaaagt cctgca                                1573

<210> 3
<211> 1691
<212> DNA
<213> Homo sapiens

<220>
<221> 5'UTR
<222> 1..143

<220>
<221> CDS
<222> 144..1187

<220>
<221> 3'UTR
<222> 1188..1691

<400> 3
tgcactccag tccgggctgg cggacagagg gctggaaca agacgctcca gaatcaggag      60
cttcccctca ggaaatagca tctgtgtcc ccgactgca gttgtctggt ctctccagca      120
gtttggtact tccggctgct gca atg cgt gtg gtg att gga gca gga gtc      173
                Met Arg Val Val Val Ile Gly Ala Gly Val
                1                5                10
atc ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc      221
Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
                15                20                25
ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc      269
Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
                30                35                40
acc acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac      317
Thr Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp
                45                50                55
ccc aac aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat      365

```

Pro	Asn	Asn	Pro	Gln	Glu	Ala	Asp	Trp	Ser	Gln	Gln	Thr	Phe	Asp	Tyr	
60						65					70					
ctc	ctg	agc	cat	gtc	cat	tct	ccc	aac	gct	gaa	aac	ctg	ggc	ctg	ttc	413
Leu	Leu	Ser	His	Val	His	Ser	Pro	Asn	Ala	Glu	Asn	Leu	Gly	Leu	Phe	
75					80					85					90	
cta	atc	tcg	ggc	tac	aac	ctc	ttc	cat	gaa	gcc	att	ccg	gac	cct	tcc	461
Leu	Ile	Ser	Gly	Tyr	Asn	Leu	Phe	His	Glu	Ala	Ile	Pro	Asp	Pro	Ser	
				95					100					105		
tgg	aag	gac	aca	gtt	ctg	gga	ttt	cgg	aag	ctg	acc	ccc	aga	gag	ctg	509
Trp	Lys	Asp	Thr	Val	Leu	Gly	Phe	Arg	Lys	Leu	Thr	Pro	Arg	Glu	Leu	
			110					115					120			
gat	atg	ttc	cca	gat	tac	ggc	tat	ggc	tgg	ttc	cac	aca	agc	cta	att	557
Asp	Met		Pro	Asp	Tyr	Gly	Tyr	Gly	Trp	Phe	His	Thr	Ser	Leu	Ile	
		125					130					135				
ctg	gag	gga	aag	aac	tat	cta	cag	tgg	ctg	act	gaa	agg	tta	act	gag	605
Leu	Glu	Gly	Lys	Asn	Tyr	Leu	Gln	Trp	Leu	Thr	Glu	Arg	Leu	Thr	Glu	
	140					145				150						
agg	gga	gtg	aag	ttc	ttc	cag	cgg	aaa	gtg	gag	tct	ttt	gag	gag	gtg	653
Arg	Gly	Val	Lys	Phe	Phe	Gln	Arg	Lys	Val	Glu	Ser	Phe	Glu	Glu	Val	
155				160					165					170		
gca	aga	gaa	ggc	gca	gac	gtg	att	gtc	aac	tgc	act	ggg	gta	tgg	gct	701
Ala	Arg	Glu	Gly	Ala	Asp	Val	Ile	Val	Asn	Cys	Thr	Gly	Val	Trp	Ala	
			175					180					185			
ggg	gcg	cta	caa	cga	gac	ccc	ctg	ctg	cag	cca	ggc	cgg	ggg	cag	atc	749
Gly	Ala	Leu	Gln	Arg	Asp	Pro	Leu	Leu	Gln	Pro	Gly	Arg	Gly	Gln	Ile	
			190				195					200				
atg	aag	gtg	gac	gcc	cct	tgg	atg	aag	cac	ttc	att	ctc	acc	cat	gac	797
Met	Lys	Val	Asp	Ala	Pro	Trp	Met	Lys	His	Phe	Ile	Leu	Thr	His	Asp	
		205				210						215				
cca	gag	aga	ggc	atc	tac	aat	tcc	ccg	tac	atc	atc	cca	ggg	acc	cag	845
Pro	Glu	Arg	Gly	Ile	Tyr	Asn	Ser	Pro	Tyr	Ile	Ile	Pro	Gly	Thr	Gln	
	220					225				230						
aca	gtt	act	ctt	gga	ggc	atc	ttc	cag	ttg	gga	aac	tgg	agt	gaa	cta	893
Thr	Val	Thr	Leu	Gly	Gly	Ile	Phe	Gln	Leu	Gly	Asn	Trp	Ser	Glu	Leu	
235				240					245					250		
aac	aat	atc	cag	gac	cac	aac	acc	att	tgg	gaa	ggc	tgc	tgc	aga	ctg	941
Asn	Asn	Ile	Gln	Asp	His	Asn	Thr	Ile	Trp	Glu	Gly	Cys	Cys	Arg	Leu	
			255					260				265				
gag	ccc	aca	ctg	aag	aat	gca	aga	att	att	ggt	gaa	cga	act	ggc	ttc	989
Glu	Pro	Thr	Leu	Lys	Asn	Ala	Arg	Ile	Ile	Gly	Glu	Arg				

<210> 4
 <211> 2620
 <212> DNA
 <213> Homo sapiens

<220>
 <221> 5'UTR
 <222> 1..1155

<220>
 <221> CDS
 <222> 1156..1818

<220>
 <221> 3'UTR
 <222> 1819..2620

<400> 4
 gaaacccacg cagcctcctg gattcttccc cgtccctccc tctgtcctgg ggctgtgacc 60
 tcttccatgt tattcacagg gtctcagcac gattcatctc aaagcagcga aacaagcact 120
 ggctcagaa gaagcaatat taaaacagtt acaactcatc tagcgcacag acacccaact 180
 gacaccctgt tcctccagtc ataacaacaa ctacagcttt gattgaacaa gagactgagt 240
 ttgtaacttt tctcctaata aaaagatcac tgactatgga ctgcttctgg tggggttacg 300
 aaaccgcaac ctcatgtgcc tgcatttctt gaaaagacat tttgatgtag gaagggcctg 360
 gagtctgtct gcttgcgtct ctgggatacg ggagcaaaga gccacgcac ctcatggccc 420
 acacaggcgt cacctccagt ctctccttgg cctcatctcc ccagcgtcct ggaatggcat 480
 cgggctggcc cagggaagccc ctgtcctgtg cctctccttt cccctcaggg gctgccaggg 540
 tgaccacccc caccgcaggc caggcctaca gtgccccatg gaacgtcctg accctccccc 600
 aggttggcag caggaagaag gaagaaaggg gatcctctcc agctggccag agagacagac 660
 cttcttctgtc tcataacccc tccaagaatg cctgcctccc ctccctcccc caaggcctgt 720
 ccacaggggc ttgagatcag ccagaaaagt caggcaactt ttcagggact gggagcgagg 780
 tctcccgccc gggcctgggt ccagtctctg tgggcagtgc agtgccgagc cccaccctc 840
 aagccgtgcc ctgtccatag ctccagactt tgacctgca ctccagtcg ggctggcgga 900
 cagagggctg gaaacaagac gctccagaat caggagcttc ccctcaggaa atagcatcct 960
 gtgtccccgc actgcagttg tctggtctct ccagcagttt ggtacttccg atgaagagct 1020
 tgtgtctcca cggcaaaagt atgggggaag agggaaagaa gaagaccaag ggtccctgag 1080
 aggggctgtc ccctaagccc cagtatccaa gctcgggctc gaagctggaa ggagaattgc 1140
 ctagaggctg ctgca atg cgt gtg gtg gtg att gga gca gga gtc atc ggg 1191
 Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly
 1 5 10
 ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc ctg cag 1239
 Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln
 15 20 25
 cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc acc acc 1287
 Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr
 30 35 40
 acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac ccc aac 1335
 Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn
 45 50 55 60
 aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat ctc ctg 1383
 Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu
 65 70 75
 agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc cta atc 1431
 Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile
 80 85 90
 tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc tgg aag 1479
 Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys
 95 100 105
 gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg gat atg 1527
 Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met
 110 115 120
 ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att ctg gag 1575

32

```

Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu
125          130          135          140
gga aag aac tat cta cag tgg ctg act gaa agg tta act gag agg gga      1623
Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly
          145          150          155
gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg gca aga      1671
Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg
          160          165          170
gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct ggg gcg      1719
Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala
          175          180          185
cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc atg aag      1767
Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys
          190          195          200
gac cca gac agt tac tct tgg agg cat ctt cca gtt ggg aaa ctg gag      1815
Asp Pro Asp Ser Tyr Ser Trp Arg His Leu Pro Val Gly Lys Leu Glu
          205          210          215          220
tga actaaacaat atccaggacc acaacacccat ttgggaaggc tgctgcagac      1868
*
tgagagcccac actgaagaat gcaagaatta ttggtgaacg aactggcttc cggccagtac      1928
gccccagat tgggctagaa agagaacagc ttgcactgg accttcaaac acagaggtea      1988
tccacaacta tggccatgga ggctacgggc tcaccatcca ctggggatgt gccctggagg      2048
cagccaagct ctttgggaga atcctggaag aaaagaaatt gtccagaatg ccaccatccc      2108
acctctgaag actccagtga ctgctgcctc cccccacaag aactcccttc tccctcagc      2168
caatgaatca atgtgtcctt tcataagcca ttgcttcttc ctacttctt tctcaaaga      2228
agcatgaggt gagagaaagc cacaaagtca gtgcctggag aagggttcag cccaacatgg      2288
ggccctctc atcactgaaa tccctctacc ttctctgggt ctggcattat aaagaacagc      2348
tgaggctgtc attccatgag tcttcagaag aaaggacagc tcagaaaatc aaagaggcca      2408
actgcccaga gccacagaaa atggaggata attgaggcta agtaacctga ttacaagttg      2468
tactaacata ttaaagggtc tgaaaagtcc tgcagcaaag acaactatct gatgttgttt      2528
aaccagtgct ttgctaaacc tatctggcta tggaactctt ttgcccagag caccatgaa      2588
tgccatgaca caaatctgag aaaatgctgg aa      2620

<210> 5
<211> 1576
<212> DNA
<213> Homo sapiens

<220>
<221> 5'UTR
<222> 1..143

<220>
<221> CDS
<222> 144..380

<220>
<221> 3'UTR
<222> 381..1576

<400> 5
tgactccag tccgggctgg cggacagagg gctggaaca agacgctcca gaatcaggag      60
cttccccctca ggaaatagca tcctgtgtcc ccgactgca gttgtctggt ctctccagca      120
gtttgggtact tccggctgct gca atg cgt gtg gtg att gga gca gga gtc      173
          Met Arg Val Val Val Ile Gly Ala Gly Val
          1          5          10
atc ggg ctg tcc acc gcc ctg tgc atc cat gag cgc tac cac tca gtc      221
Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
          15          20          25
ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctg      269
Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
          30          35          40
acc acc acc gac gtg gct gcc ggc ctg tgg cag ccc tac ctt tct gac      317

```



```

      35      40      45
ggc ctc tgg cag ccc tac ctt tct gac ccc aac aac cca cag gag gcg      308
Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu Ala
50      55      60      65
gac tgg agc caa cag acc ttt gac tat ctc ctg agc cat gtc cat tct      356
Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His Ser
      70      75      80
ccc aac gct gaa aac ctg ggc ctg ttc cta atc tcg ggc tac aac ctc      404
Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn Leu
      85      90      95
ttc cat gaa gcc att ccg gtg gca aga gaa ggc gca gac gtg att gtc      452
Phe His Glu Ala Ile Pro Val Ala Arg Glu Gly Ala Asp Val Ile Val
      100      105      110
aac tgc act ggg gta tgg gct ggg gcg cta caa cga gac ccc ctg ctg      500
Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp Pro Leu Leu
      115      120      125
cag cca ggc cgg ggg cag atc atg aag gtg gac gcc cct tgg atg aag      548
Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro Trp Met Lys
      130      135      140      145
cac ttc att ctc acc cat gac cca gag aga ggc atc tac aat tcc ccg      596
His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr Asn Ser Pro
      150      155      160
tac atc atc cca ggg acc cag aca gtt act ctt gga ggc atc ttc cag      644
Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly Ile Phe Gln
      165      170      175
ttg gga aac tgg agt gaa cta aac aat atc cag gac cac aac acc att      692
Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His Asn Thr Ile
      180      185      190
tgg gaa ggc tgc tgc aga ctg gag ccc aca ctg aag aat gca aga att      740
Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn Ala Arg Ile
      195      200      205
att ggt gaa cga act ggc ttc cgg cca gta cgc ccc cag att cgg cta      788
Ile Gly Glu Arg Thr Gly Phe Arg Pro Val Arg Pro Gln Ile Arg Leu
      210      215      220      225
gaa aga gaa cag ctt cgc act gga cct tca aac aca gag gtc atc cac      836
Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu Val Ile His
      230      235      240
aac tat ggc cat gga ggc tac ggg ctc acc atc cac tgg gga tgt gcc      884
Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp Gly Cys Ala
      245      250      255
ctg gag gca gcc aag ctc ttt ggg aga atc ctg gaa gaa aag aaa ttg      932
Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu Lys Lys Leu
      260      265      270
tcc aga atg cca cca tcc cac ctc tga agactccagt gactgctgcc      979
Ser Arg Met Pro Pro Ser His Leu *
      275      280
tccccccaca agaactccct tctccctca gccaatgaat caatgtgctc cttcataagc      1039
cattgcttct ccctcacttc tttcctcaaa gaagcatgag gtgagagaaa gccacaaagt      1099
cagtgccttg agaagggttc agcccaacat ggggccctc tcatcactga aatccctcta      1159
ccttctcttg gtctggcatt ataaagaaca gctgaggtg tcatccatg agtcttcaga      1219
agaaaggaca gctcagaaaa tcaaagaggc caactgccca gagccacaga aaatggagga      1279
taattgaggc taagtaacct gattacaagt tgtactaaca tattaaggt tctgaaaagt      1339
cctgca      1345

```

<210> 7

<211> 347

<212> PRT

<213> Homo sapiens

<400> 7

```

Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
1      5      10      15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile

```

			20					25					30			
Lys	Val	Tyr	Ala	Asp	Arg	Phe	Thr	Pro	Leu	Thr	Thr	Thr	Asp	Val	Ala	
		35					40					45				
Ala	Gly	Leu	Trp	Gln	Pro	Tyr	Leu	Ser	Asp	Pro	Asn	Asn	Pro	Gln	Glu	
	50					55					60					
Ala	Asp	Trp	Ser	Gln	Gln	Thr	Phe	Asp	Tyr	Leu	Leu	Ser	His	Val	His	
65					70					75					80	
Ser	Pro	Asn	Ala	Glu	Asn	Leu	Gly	Leu	Phe	Leu	Ile	Ser	Gly	Tyr	Asn	
				85					90					95		
Leu	Phe	His	Glu	Ala	Ile	Pro	Asp	Pro	Ser	Trp	Lys	Asp	Thr	Val	Leu	
			100					105					110			
Gly	Phe	Arg	Lys	Leu	Thr	Pro	Arg	Glu	Leu	Asp	Met	Phe	Pro	Asp	Tyr	
		115					120					125				
Gly	Tyr	Gly	Trp	Phe	His	Thr	Ser	Leu	Ile	Leu	Glu	Gly	Lys	Asn	Tyr	
	130					135					140					
Leu	Gln	Trp	Leu	Thr	Glu	Arg	Leu	Thr	Glu	Arg	Gly	Val	Lys	Phe	Phe	
145					150				155						160	
Gln	Arg	Lys	Val	Glu	Ser	Phe	Glu	Glu	Val	Ala	Arg	Glu	Gly	Ala	Asp	
				165					170					175		
Val	Ile	Val	Asn	Cys	Thr	Gly	Val	Trp	Ala	Gly	Ala	Leu	Gln	Arg	Asp	
			180					185					190			
Pro	Leu	Leu	Gln	Pro	Gly	Arg	Gly	Gln	Ile	Met	Lys	Val	Asp	Ala	Pro	
		195					200					205				
Trp	Met	Lys	His	Phe	Ile	Leu	Thr	His	Asp	Pro	Glu	Arg	Gly	Ile	Tyr	
	210					215					220					
Asn	Ser	Pro	Tyr	Ile	Ile	Pro	Gly	Thr	Gln	Thr	Val	Thr	Leu	Gly	Gly	
225					230					235					240	
Ile	Phe	Gln	Leu	Gly	Asn	Trp	Ser	Glu	Leu	Asn	Asn	Ile	Gln	Asp	His	
				245					250					255		
Asn	Thr	Ile	Trp	Glu	Gly	Cys	Cys	Arg	Leu	Glu	Pro	Thr	Leu	Lys	Asn	
			260					265					270			
Ala	Arg	Ile	Ile	Gly	Glu	Arg	Thr	Gly	Phe	Arg	Pro	Val	Arg	Pro	Gln	
		275					280					285				
Ile	Arg	Leu	Glu	Arg	Glu	Gln	Leu	Arg	Thr	Gly	Pro	Ser	Asn	Thr	Glu	
	290					295					300					
Val	Ile	His	Asn	Tyr	Gly	His	Gly	Gly	Tyr	Gly	Leu	Thr	Ile	His	Trp	
305					310					315					320	
Gly	Cys	Ala	Leu	Glu	Ala	Ala	Lys	Leu	Phe	Gly	Arg	Ile	Leu	Glu	Glu	
				325					330					335		
Lys	Lys	Leu	Ser	Arg	Met	Pro	Pro	Ser	His	Leu						
			340					345								

```
<210> 8
<211> 220
<212> PRT
<213> Homo sapiens
```

<400>	8															
Met	Arg	Val	Val	Val	Ile	Gly	Ala	Gly	Val	Ile	Gly	Leu	Ser	Thr	Ala	
1				5					10					15		
Leu	Cys	Ile	His	Glu	Arg	Tyr	His	Ser	Val	Leu	Gln	Pro	Leu	Asp	Ile	
			20					25					30			
Lys	Val	Tyr	Ala	Asp	Arg	Phe	Thr	Pro	Leu	Thr	Thr	Thr	Asp	Val	Ala	
			35				40					45				
Ala	Gly	Leu	Trp	Gln	Pro	Tyr	Leu	Ser	Asp	Pro	Asn	Asn	Pro	Gln	Glu	
	50					55					60					
Ala	Asp	Trp	Ser	Gln	Gln	Thr	Phe	Asp	Tyr	Leu	Leu	Ser	His	Val	His	
65				70					75					80		
Ser	Pro	Asn	Ala	Glu	Asn	Leu	Gly	Leu	Phe	Leu	Ile	Ser	Gly	Tyr	Asn	
				85					90					95		
Leu	Phe	His	Glu	Ala	Ile	Pro	Asp	Pro	Ser	Trp	Lys	Asp	Thr	Val	Leu	
			100					105					110			
Gly	Phe	Arg	Lys	Leu	Thr	Pro	Arg	Glu	Leu	Asp	Met	Phe	Pro	Asp	Tyr	

		115					120					125				
Gly	Tyr	Gly	Trp	Phe	His	Thr	Ser	Leu	Ile	Leu	Glu	Gly	Lys	Asn	Tyr	
		130					135					140				
Leu	Gln	Trp	Leu	Thr	Glu	Arg	Leu	Thr	Glu	Arg	Gly	Val	Lys	Phe	Phe	
145					150					155					160	
Gln	Arg	Lys	Val	Glu	Ser	Phe	Glu	Glu	Val	Ala	Arg	Glu	Gly	Ala	Asp	
				165					170					175		
Val	Ile	Val	Asn	Cys	Thr	Gly	Val	Trp	Ala	Gly	Ala	Leu	Gln	Arg	Asp	
			180					185					190			
Pro	Leu	Leu	Gln	Pro	Gly	Arg	Gly	Gln	Ile	Met	Lys	Asp	Pro	Asp	Ser	
		195					200					205				
Tyr	Ser	Trp	Arg	His	Leu	Pro	Val	Gly	Lys	Leu	Glu					
		210					215				220					

```
<210> 9
<211> 78
<212> PRT
<213> Homo sapiens
```

```

<400> 9
Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
1          5          10          15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
          20          25          30
Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
          35          40          45
Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
          50          55          60
Ala Thr Leu Pro Gly Arg Thr Gln Phe Trp Asp Phe Gly Ser
65          70          75

```

```
<210> 10
<211> 281
<212> PRT
<213> Homo sapiens
```

<400>	10															
Met	Arg	Val	Val	Val	Ile	Gly	Ala	Gly	Val	Ile	Gly	Leu	Ser	Thr	Ala	
1				5					10					15		
Leu	Cys	Ile	His	Glu	Arg	Tyr	His	Ser	Val	Leu	Gln	Pro	Leu	Asp	Ile	
			20					25					30			
Lys	Val	Tyr	Ala	Asp	Arg	Phe	Thr	Pro	Leu	Thr	Thr	Thr	Asp	Val	Ala	
		35					40					45				
Ala	Gly	Leu	Trp	Gln	Pro	Tyr	Leu	Ser	Asp	Pro	Asn	Asn	Pro	Gln	Glu	
	50					55					60					
Ala	Asp	Trp	Ser	Gln	Gln	Thr	Phe	Asp	Tyr	Leu	Leu	Ser	His	Val	His	
65					70					75					80	
Ser	Pro	Asn	Ala	Glu	Asn	Leu	Gly	Leu	Phe	Leu	Ile	Ser	Gly	Tyr	Asn	
				85					90					95		
Leu	Phe	His	Glu	Ala	Ile	Pro	Val	Ala	Arg	Glu	Gly	Ala	Asp	Val	Ile	
			100					105					110			
Val	Asn	Cys	Thr	Gly	Val	Trp	Ala	Gly	Ala	Leu	Gln	Arg	Asp	Pro	Leu	
		115					120					125				
Leu	Gln	Pro	Gly	Arg	Gly	Gln	Ile	Met	Lys	Val	Asp	Ala	Pro	Trp	Met	
	130					135					140					
Lys	His	Phe	Ile	Leu	Thr	His	Asp	Pro	Glu	Arg	Gly	Ile	Tyr	Asn	Ser	
145					150					155					160	
Pro	Tyr	Ile	Ile	Pro	Gly	Thr	Gln	Thr	Val	Thr	Leu	Gly	Gly	Ile	Phe	
			165						170					175		
Gln	Leu	Gly	Asn	Trp	Ser	Glu	Leu	Asn	Asn	Ile	Gln	Asp	His	Asn	Thr	
			180					185					190			
Ile	Trp	Glu	Gly	Cys	Cys	Arg	Leu	Glu	Pro	Thr	Leu	Lys	Asn	Ala	Arg	
		195					200					205				

37

```

Ile Ile Gly Glu Arg Thr Gly Phe Arg Pro Val Arg Pro Gln Ile Arg
    210                      215                      220
Leu Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu Val Ile
225                      230                      235                      240
His Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp Gly Cys
                      245                      250                      255
Ala Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu Lys Lys
                      260                      265                      270
Leu Ser Arg Met Pro Pro Ser His Leu
    275                      280

```

<210> 11
 <211> 456
 <212> DNA
 <213> Homo sapiens

<220>
 <221> allele
 <222> 152
 <223> 99-16105-152 : polymorphic base A or G

<220>
 <221> misc_binding
 <222> 140..164
 <223> 99-16105-152.probe

<220>
 <221> primer_bind
 <222> 133..151
 <223> 99-16105-152.mis

<220>
 <221> primer_bind
 <222> 153..171
 <223> 99-16105-152.mis complement

<220>
 <221> primer_bind
 <222> 1..19
 <223> 99-16105.pu

<220>
 <221> primer_bind
 <222> 437..456
 <223> 99-16105.rp complement

```

<400> 11
cgctttgttg tattctttgt tatttatcca ttttgccaaa ttatctgcaa gtagaaatat    60
cgaaataaga agctcttttag caatttactt tggatattgg ttttcttttg aaggacagtt    120
attaaaatag cttgtaggat tactcathtt crtttttctt cttttttaa ataaagcaat    180
gtcatcactt ttttccctgt attatatctt tcctcaataa ttgatatgct acattaaagg    240
aacacaaaat ggtcttaatt atgcaataat gatcaaggca aagagtgttt cctgggaact    300
aatgggttgc tgagaggagg tgatggcttg aggtccagct ggttattaag ccgcaggaaa    360
tgctgcaggc caagatttgt attatttctc tgagatgaaa atgaacccaa aaaaaggcaa    420
aatgggtttt tctccactaa tgggtaaaat gaactc                                456

```

<210> 12
 <211> 463
 <212> DNA
 <213> Homo sapiens

<220>
 <221> allele

```

<222> 215
<223> 99-5919-215 : polymorphic base A or G

<220>
<221> misc_binding
<222> 203..227
<223> 99-5919-215.probe

<220>
<221> primer_bind
<222> 196..214
<223> 99-5919-215.mis

<220>
<221> primer_bind
<222> 216..234
<223> 99-5919-215.mis complement

<220>
<221> primer_bind
<222> 11..29
<223> 99-5919.pu

<220>
<221> primer_bind
<222> 445..465
<223> 99-5919.rp complement

<400> 12
tttctcttga ctacagcaat gcagatttca attctgccat tgaattccca gacatattcg      60
tcatcccccatt tttcatcccc caccaccctg ccattttctt cgtgttaact tgttttcctg      120
actcacagaa atcacctttt cctgtataca tttttaggat gtcagacttt attctaataga      180
tttctcctag ttgcccccca aaattgtatt ctacrgtgtg attttaaagc tgaattttca      240
agatgatatt tcatatctat attttcacaa gcttttcttc tatgaatgtt attgtcagct      300
gtcagggtgt gagatgggtac ttgatactac attctttcca agctgttgcc tgaatcggtt      360
taagacaaaag tcattactag gctgtaaaact gttgctctgc aaaattgagc agcacgtatt      420
taaccactca tacttcttag ctctccaaca ctttgagter ata                          463

<210> 13
<211> 742
<212> DNA
<213> Homo sapiens

<220>
<221> 5'UTR
<222> 1..46

<220>
<221> CDS
<222> 47..508

<220>
<221> 3'UTR
<222> 509..742

<220>
<221> polyA_signal
<222> 718..723

<220>
<221> allele
<222> 21
<223> 8-135-112 : polymorphic base C or T

```

```
<220>
<221> allele
<222> 75
<223> 8-135-166 : polymorphic base A or C
```

```
<220>
<221> allele
<222> 135
<223> 99-16038-118 : polymorphic base A or G
```

```
<220>
<221> allele
<222> 194
<223> 8-137-152 : polymorphic base G or T
```

```
<220>  
<221> allele  
<222> 224  
<223> 8-137-182 : polymorphic base A or G
```

```
<220>  
<221> allele  
<222> 390  
<223> 8-130-220 : polymorphic base A or C
```

```
<220>
<221> allele
<222> 406
<223> 8-130-236 : polymorphic base A or G
```

```
<220>
<221> allele
<222> 578
<223> 8-131-199 : polymorphic base A or C
```

```
<220>
<221> allele
<222> 641
<223> 8-132-97 : polymorphic base C or T
```

```
<220>
<221> allele
<222> 708
<223> 8-132-164 : polymorphic base C or T
```

```
<220>
<221> allele
<222> 723
<223> 8-132-179 : polymorphic base A or T
```

<400> 13																												
tc	at	ct	ct	gc	tt	ca	ca	at	gc	yg	at	ga	tt	ta	gct	ggg	gag	ga	cc	ca	aa	at	g	ct	g	aa		55
																						Met	Leu	Glu				
																						1						
aa	g	ct	g	at	g	g	t	g	ct	ga	t	m	t	ct	ca	g	ct	tt	tt	ca	aa	aga	tcc	aga	tat	aca		103
Lys	Leu	Met	Gly	Ala	Asp	Xaa	Leu	Gln	Leu	Phe	Arg	Ser	Arg	Tyr	Thr													
	5					10					15																	
tt	g	g	t	aaa	at	c	tac	tt	c	ata	g	g	t	ttt	caa	arg	agc	att	ctt	ctg	agc						151	
Leu	Gly	Lys	Ile	Tyr	Phe	Ile	Gly	Phe	Gln	Xaa	Ser	Ile	Leu	Leu	Ser													
20					25					30					35													
aaa	tct	gaa	aac	tct	cta	aac	tct	att	gca	aag	gag	aca	gaa	kaa	gga												199	
Lys	Ser	Glu	Asn	Ser	Leu	Asn	Ser	Ile	Ala	Lys	Glu	Thr	Glu	Xaa	Gly													

40

40	45	50	
aga gag acg gta aca agg aaa gaa rga tgg aag aga agg cat gag gac			247
Arg Glu Thr Val Thr Arg Lys Glu Xaa Trp Lys Arg Arg His Glu Asp			
55	60	65	
ggc tat ttg gaa atg gca cag agg cat tta cag aga tca tta tgt cct			295
Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser Leu Cys Pro			
70	75	80	
tgg gtc tct tac ctt cct cag ccc tat gca gag ctt gaa gaa gta agc			343
Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu Glu Val Ser			
85	90	95	
agc cat gtt gga aaa gtc ttc atg gca aga aac tat gag ttc ctt gmc			391
Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu Phe Leu Xaa			
100	105	110	115
tat gag gcc tct aar gac cgc agg cag cct cta gaa cga atg tgg acc			439
Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg Met Trp Thr			
120	125	130	
tgc aac tac aac cag caa aaa gac cag tca tgc aac cac aag gaa ata			487
Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His Lys Glu Ile			
135	140	145	
act tct acc aaa gct gaa tga gtttggagc agattcttcc cagccaatcc			538
Thr Ser Thr Lys Ala Glu *			
150			
ttctgatgac aatgtagtct ggccaacatc ttacttggam tctgacggac tctgtgtctg			598
ggaccagct gataacacgt ggtgatggga ttgtatttgc aaytctctgg tcagtaagtg			658
ataaaatgcc atttctatgc acccacctgg cctgtgtgac tgggagaaty tctcttttta			718
ttaawtgtgc ttcaagtttt aaca			742

<210> 14

<211> 153

<212> PRT

<213> Homo sapiens

<400> 14

Met Leu Glu Lys Leu Met Gly Ala Asp Xaa Leu Gln Leu Phe Arg Ser	
1	5
Arg Tyr Thr Leu Gly Lys Ile Tyr Phe Ile Gly Phe Gln Xaa Ser Ile	10
	15
20	25
Leu Leu Ser Lys Ser Glu Asn Ser Leu Asn Ser Ile Ala Lys Glu Thr	30
35	40
Glu Xaa Gly Arg Glu Thr Val Thr Arg Lys Glu Xaa Trp Lys Arg Arg	45
50	55
His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser	60
65	70
Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu	75
80	85
Glu Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu	90
95	100
Phe Leu Xaa Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg	105
110	115
Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His	120
125	130
Lys Glu Ile Thr Ser Thr Lys Ala Glu	135
140	145
	150

<210> 15

<211> 476

<212> DNA

<213> Homo sapiens

<400> 15

cat gag gac ggc tat ttg gaa atg gca cag agg cat tta cag aga tca	
His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser	48
1	5
	10
	15

41

```

tta tgt cct tgg gtc tct tac ctt cct cag ccc tat gca gag ctt gaa      96
Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu
      20      25      30
gaa gta agc agc cat gtt gga aaa gtc ttc atg gca aga aac tat gag      144
Glu Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu
      35      40      45
ttc ctt gcc tat gag gcc tct aag gac cgc agg cag cct cta gaa cga      192
Phe Leu Ala Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg
      50      55      60
atg tgg acc tgc aac tac aac cag caa aaa gac cag tca tgc aac cac      240
Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His
      65      70      75      80
aag gaa ata act tct acc aaa gct gaa tga gtttggaagc agattcttcc      290
Lys Glu Ile Thr Ser Thr Lys Ala Glu *
      85      90
cagccaatcc ttctgatgac aatgtagtct ggccaacatc ttcactggac tctgacggac      350
tctgtgtctg ggaccagct gataaacagt ggtgatggga ttgtatttgc aactctctgg      410
tcagtaagtg ataaaaatgcc atttctatgc acccacctgg cctgtgtgac tgggagaatc      470
tctctt      476

```

<210> 16

<211> 89

<212> PRT

<213> Homo sapiens

<400> 16

```

His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser
1      5      10      15
Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu
      20      25      30
Glu Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu
      35      40      45
Phe Leu Ala Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg
      50      55      60
Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His
      65      70      75      80
Lys Glu Ile Thr Ser Thr Lys Ala Glu
      85

```

<210> 17

<211> 1633

<212> DNA

<213> Homo sapiens

<400> 17

```

ttgggggtcca ttgcaaccog aggcgagact agagttccca agcgagaagg gaagaggcag      60
tgggtgcacg tggaaggcgg acagagggct ggaaacaaga cgctccagaa tcaggagctt      120
cccctcagga aatagcatcc tgtgtccccg cactgcagtt gtctggtctc tccagcagtt      180
tggtacttcc ggctgctgca atg cgt gtg gtg gtg att gga gca gga gtc atc      233
      Met Arg Val Val Val Ile Gly Ala Gly Val Ile
      1      5      10
ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc ctg      281
Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val Leu
      15      20      25
cag cca ctg cac ata aag gtc tac gcg gac cgc ttc acc cca ctc acc      329
Gln Pro Leu His Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr
      30      35      40
acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac ccc      377
Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro
      45      50      55
aac aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat ctc      425
Asn Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu
      60      65      70      75

```

42

```

ctg agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc cta      473
Leu Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu
      80                      85                      90
atc tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc tgg      521
Ile Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp
      95                      100                      105
aag gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg gat      569
Lys Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp
      110                      115                      120
atg ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att ctg      617
Met Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu
      125                      130                      135
gag gga aag aac tat cta cag tgg ctg act gaa agg tta act gag agg      665
Glu Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg
      140                      145                      150                      155
gga gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg gca      713
Gly Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala
      160                      165                      170
aga gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct ggg      761
Arg Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly
      175                      180                      185
gcg cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc atg      809
Ala Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met
      190                      195                      200
aag gtg gac gcc cct tgg atg aag cac ttc att ctc acc cat gac cca      857
Lys Val Asp Ala Pro Trp Met Lys His Phe Ile Leu Thr His Asp Pro
      205                      210                      215
gag aga ggc atc tac aat tcc ccg tac atc atc cca ggg acc cag aca      905
Glu Arg Gly Ile Tyr Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln Thr
      220                      225                      230                      235
gtt act ctt gga ggc atc ttc cag ttg gga aac tgg agt gaa cta aac      953
Val Thr Leu Gly Gly Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu Asn
      240                      245                      250
aat atc cag gac cac aac acc att tgg gaa ggc tgc tgc aga ctg gag      1001
Asn Ile Gln Asp His Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu Glu
      255                      260                      265
ccc aca ctg aag aat gca aga att att ggt gaa gca act ggc ttc cgg      1049
Pro Thr Leu Lys Asn Ala Arg Ile Ile Gly Glu Ala Thr Gly Phe Arg
      270                      275                      280
cca gta cgc ccc cag att cgg cta gaa aga gaa cag ctt cgc act gga      1097
Pro Val Arg Pro Gln Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr Gly
      285                      290                      295
cct tca aac aca gag gtc atc cac aac tat ggc cat gga ggc tac ggg      1145
Pro Ser Asn Thr Glu Val Ile His Asn Tyr Gly His Gly Gly Tyr Gly
      300                      305                      310                      315
ctc acc atc cac tgg gga tgt gcc ctg gag gca gcc aag ctc ttt ggg      1193
Leu Thr Ile His Trp Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe Gly
      320                      325                      330
aga atc ctg gaa gaa aag aaa ttg tcc aga atg cca cca tcc cac ctc      1241
Arg Ile Leu Glu Glu Lys Lys Leu Ser Arg Met Pro Pro Ser His Leu
      335                      340                      345
tga agactccagt gactgctgcc tccccccaca agaactccct tctccctca      1294
*
gccaatgaat caatgtgctc cttcataagc cattgcttct cctcacttc tttcctcaaa      1354
gaagcatgag gtgagagaaa gccacraagt cagtgcctgg agaagggttc agcccaacat      1414
ggggccctc tcatcactga aatccctcta ccttctctgg gtctggcatt ataaagaaca      1474
gctgaggetg tcattccatg agtcttcaga agaaaggaca gctcagaaag tcaaagaggc      1534
caactgccc gagccacaga aaatggagga taattgaggc taagtaacct gattacaagt      1594
tgtactaaca tattaagggt tctgaaaagt cctgcaaaa      1633

```

<210> 18

<211> 347

<212> PRT

<213> Homo sapiens

<400> 18

```

Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
1          5          10          15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
          20          25          30
Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
          35          40          45
Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
          50          55          60
Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His
65          70          75          80
Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn
          85          90          95
Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys Asp Thr Val Leu
          100          105          110
Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met Phe Pro Asp Tyr
          115          120          125
Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu Gly Lys Asn Tyr
          130          135          140
Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly Val Lys Phe Phe
145          150          155          160
Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg Glu Gly Ala Asp
          165          170          175
Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp
          180          185          190
Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro
          195          200          205
Trp Met Lys His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr
          210          215          220
Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly
225          230          235          240
Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His
          245          250          255
Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn
          260          265          270
Ala Arg Ile Ile Gly Glu Ala Thr Gly Phe Arg Pro Val Arg Pro Gln
          275          280          285
Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu
          290          295          300
Val Ile His Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp
305          310          315          320
Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu
          325          330          335
Lys Lys Leu Ser Arg Met Pro Pro Ser His Leu
          340          345

```

<210> 19

<211> 1200

<212> DNA

<213> Homo sapiens

<400> 19

```

atg gac aca gca cgg att gca gtt gtc ggg gca ggt gtg gtg ggg ctc      48
Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu
1          5          10          15
tcc acg gct gtg tgc atc tcc aaa ctg gtg ccc cga tgc tcc gtt acc      96
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr
          20          25          30
atc att tca gac aag ttt act cca gat acc acc agt gat gtg gca gcc      144
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala
          35          40          45

```

44

gga atg ctt att cct cac act tat cca gat aca ccc att cac acg cag	192
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln	
50 55 60	
aag cag tgg ttc aga gaa acc ttt aat cac ctc ttt gca att gcc aat	240
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn	
65 70 75 80	
tct gca gaa gct gga gat gct ggt gtt cat ttg gta tca ggt tgg cag	288
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Trp Gln	
85 90 95	
ata ttt cag agc act ccg act gaa gaa gtg cca ttc tgg gct gac gtg	336
Ile Phe Gln Ser Thr Pro Thr Glu Glu Val Pro Phe Trp Ala Asp Val	
100 105 110	
gtt ctg gga ttt cga aag atg act gag gct gag ctg aag aaa ttc ccc	384
Val Leu Gly Phe Arg Lys Met Thr Glu Ala Glu Leu Lys Lys Phe Pro	
115 120 125	
cag tat gtg ttt ggt cag gct ttt aca acc ctg aaa tgt gaa tgc cct	432
Gln Tyr Val Phe Gly Gln Ala Phe Thr Thr Leu Lys Cys Glu Cys Pro	
130 135 140	
gcc tac ctc ccg tgg ttg gag aaa agg ata aag gga agt gga ggc tgg	480
Ala Tyr Leu Pro Trp Leu Glu Lys Arg Ile Lys Gly Ser Gly Gly Trp	
145 150 155 160	
aca ctc act cgg cga ata gaa gac ctg tgg gaa ctt cat ccg tcc ttt	528
Thr Leu Thr Arg Ile Glu Asp Leu Trp Glu Leu His Pro Ser Phe	
165 170 175	
gac atc gtg gtc aac tgt tca ggc ctt gga agc aga cag ctt gca gga	576
Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser Arg Gln Leu Ala Gly	
180 185 190	
gac tca aag att ttc cct gta agg ggc caa gtc ctc caa gtt cag gct	624
Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val Leu Gln Val Gln Ala	
195 200 205	
ccc tgg gtg gag cat ttt atc cga gat ggc agt ggg ctg aca tat att	672
Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser Gly Leu Thr Tyr Ile	
210 215 220	
tat cct ggt aca tcc cat gta acc cta ggt gga act agg caa aaa ggg	720
Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly Thr Arg Gln Lys Gly	
225 230 235 240	
gac tgg aat ctg tcc ccg gat gca gaa aat agc aga gag att ctt tcc	768
Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser Arg Glu Ile Leu Ser	
245 250 255	
cga tgc tgt gct ctg gag ccc tcc ctc cac gga gcc tgc aac atc agg	816
Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly Ala Cys Asn Ile Arg	
260 265 270	
gag aag gtg ggc ttg agg ccc tac agg cca ggc gtg cga ctg cag aca	864
Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly Val Arg Leu Gln Thr	
275 280 285	
gag ctc ctt gcg cga gat gga cag agg ctg cct gta gtc cac cac tat	912
Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro Val Val His His Tyr	
290 295 300	
ggc cat ggg agt ggg ggc atc tca gtg cac tgg ggc act gct ctg gag	960
Gly His Gly Ser Gly Gly Ile Ser Val His Trp Gly Thr Ala Leu Glu	
305 310 315 320	
gcc gcc agg ctg gtg agc gag tgt gtc cat gcc ctc agg acc ccc att	1008
Ala Ala Arg Leu Val Ser Glu Cys Val His Ala Leu Arg Thr Pro Ile	
325 330 335	
ccc aag tca aac ctg tag atgacataaaa atgacagcaa agagactgag	1056
Pro Lys Ser Asn Leu *	
340	
agactgttga tcaaagcaca gaacagggttc aaataaacttt tccactgcat gaaagttaa	1116
ttagacatttt ctttgttttc aacattagaa gtggtgtaac atgtaagctg agcacggtag	1176
catgcctata gtcccagcta cttg	1200

<210> 20

<211> 2056

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> 1025

<223> n=a, g, c or t

<400> 20

atg gac aca gca cgg att gca gtt gtc ggg gca ggt gtg gtg ggg ctc	48
Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu	
1 5 10 15	
tcc acg gct gtg tgc atc tcc aaa ctg gtg ccc cga tgc tcc gtt acc	96
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr	
20 25 30	
atc att tca gac aag ttt act cca gat acc acc agt gat gtg gca gcc	144
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala	
35 40 45	
gga atg ctt att cct cac act tat cca gat aca ccc att cac acg cag	192
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln	
50 55 60	
aag cag tgg ttc aga gaa acc ttt aat cac ctc ttt gca att gcc aat	240
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn	
65 70 75 80	
tct gca gaa gct gga gat gct ggt gtt cat ttg gta tca ggg ata aag	288
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Ile Lys	
85 90 95	
gga agt gga ggc tgg aca ctc act cgg cga ata gaa gac ctg tgg gaa	336
Gly Ser Gly Gly Trp Thr Leu Thr Arg Ile Glu Asp Leu Trp Glu	
100 105 110	
ctt cat ccg tcc ttt gac atc gtg gtc aac tgt tca ggc ctt gga agc	384
Leu His Pro Ser Phe Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser	
115 120 125	
aga cag ctt gca gga gac tca aag att ttc cct gta agg ggc caa gtc	432
Arg Gln Leu Ala Gly Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val	
130 135 140	
ctc caa gtt cag gct ccc tgg gtg gag cat ttt atc cga gat ggc agt	480
Leu Gln Val Gln Ala Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser	
145 150 155 160	
ggg ctg aca tat att tat cct ggt aca tcc cat gta acc cta ggt gga	528
Gly Leu Thr Tyr Ile Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly	
165 170 175	
act agg caa aaa ggg gac tgg aat ctg tcc ccg gat gca gaa aat agc	576
Thr Arg Gln Lys Gly Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser	
180 185 190	
aga gag att ctt tcc cga tgc tgt gct ctg gag ccc tcc ctc cac gga	624
Arg Glu Ile Leu Ser Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly	
195 200 205	
gcc tgc aac atc agg gag aag gtg ggc ttg agg ccc tac agg cca ggc	672
Ala Cys Asn Ile Arg Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly	
210 215 220	
gtg cga ctg cag aca gag ctc ctt gcg cga gat gga cag agg ctg cct	720
Val Arg Leu Gln Thr Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro	
225 230 235 240	
gta gtc cac cac tat ggc cat ggg agt ggg ggc atc tca gtg cac tgg	768
Val Val His His Tyr Gly His Gly Ser Gly Gly Ile Ser Val His Trp	
245 250 255	
ggc act gct ctg gag gcc gcc agg ctg gtg agc gag tgt gtc cat gcc	816
Gly Thr Ala Leu Glu Ala Ala Arg Leu Val Ser Glu Cys Val His Ala	
260 265 270	
ctc agg acc ccc att ccc aag tca aac ctg tag atgacataaa atgacagcaa	869
Leu Arg Thr Pro Ile Pro Lys Ser Asn Leu *	
275 280	

```

agagactgag agactgttga tcaaagcaca gaacagggttc aaataacttt tccactgcat 929
gaaagtttaa ttagacattt ctttgttttc aacattagaa gtggtgtaac atgtaagctg 989
agcacggtag catgcctata gtcccagcta cttg>nm_00 4032atggac acagcacgga 1041
ttgcagttgt cggggcagggt gtggtggggc tctccacggc tgtgtgcatc tccaaactgg 1101
tgccccgatg ctccgttacc atcatttcag acaagtttac tccagatacc accagtgatg 1161
tggcagccgg aatgcattatt cctcacactt atccagatac acccattcac acgcagaagc 1221
agtggttcag agaaaccttt aatcacctct ttgcaattgc caattctgca gaagctggag 1281
atgtctggtg tcatttggtg tcagggataa aggggaagtgg aggctggaca ctactcggc 1341
gaatagaaga cctgtgggaa cttcatccgt cctttgacat cgtggtcaac tgttcaggcc 1401
ttggaagcag acagcttgca ggagactcaa agattttccc tgtaaggggc caagtcctcc 1461
aagttcaggc tccctgggtg gagcatttta tccgagatgg cagtgggctg acatatattt 1521
atcctggtac atcccatgta accctaggtg gaactaggca aaaaggggac tggaatctgt 1581
ccccggatgc agaaaatagc agagagattc tttcccgatg ctgtgctctg gagccctccc 1641
tccacggagc ctgcaacatc agggagaagg tgggcttgag gccctacagg ccaggcgtgc 1701
gactgcagac agagctcctt gcgcgagatg gacagaggct gcctgtagtc caccactatg 1761
gccatgggag tgggggcacg tcagtgcact ggggcactgc tctggaggcc gccaggctgg 1821
tgagcgagtg tgtccatgcc ctcaggaccc ccattcccaa gtcaaacctg tagatgacat 1881
aaaatgacag caaagagact gagagactgt tgatcaaagc acagaacagg ttcaaataac 1941
ttttccactg catgaaagtt taattagaca tttctttgtt ttcaacatta gaagtgggtg 2001
aacatgtaag ctgagcacgg tagcatgcct atagtcacag ctacttg 2048

```

<210> 21

<211> 341

<212> PRT

<213> Homo sapiens

<400> 21

```

Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu
1          5          10          15
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr
20          25          30
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala
35          40          45
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln
50          55          60
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn
65          70          75          80
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Trp Gln
85          90          95
Ile Phe Gln Ser Thr Pro Thr Glu Glu Val Pro Phe Trp Ala Asp Val
100          105          110
Val Leu Gly Phe Arg Lys Met Thr Glu Ala Glu Leu Lys Lys Phe Pro
115          120          125
Gln Tyr Val Phe Gly Gln Ala Phe Thr Thr Leu Lys Cys Glu Cys Pro
130          135          140
Ala Tyr Leu Pro Trp Leu Glu Lys Arg Ile Lys Gly Ser Gly Gly Trp
145          150          155          160
Thr Leu Thr Arg Arg Ile Glu Asp Leu Trp Glu Leu His Pro Ser Phe
165          170          175
Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser Arg Gln Leu Ala Gly
180          185          190
Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val Leu Gln Val Gln Ala
195          200          205
Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser Gly Leu Thr Tyr Ile
210          215          220
Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly Thr Arg Gln Lys Gly
225          230          235          240
Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser Arg Glu Ile Leu Ser
245          250          255
Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly Ala Cys Asn Ile Arg
260          265          270
Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly Val Arg Leu Gln Thr
275          280          285

```

47

Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro Val Val His His Tyr
 290 295 300
 Gly His Gly Ser Gly Gly Ile Ser Val His Trp Gly Thr Ala Leu Glu
 305 310 315 320
 Ala Ala Arg Leu Val Ser Glu Cys Val His Ala Leu Arg Thr Pro Ile
 325 330 335
 Pro Lys Ser Asn Leu
 340

<210> 22

<211> 282

<212> PRT

<213> Homo sapiens

<400> 22

Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu
 1 5 10 15
 Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr
 20 25 30
 Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala
 35 40 45
 Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln
 50 55 60
 Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn
 65 70 75 80
 Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Ile Lys
 85 90 95
 Gly Ser Gly Gly Trp Thr Leu Thr Arg Arg Ile Glu Asp Leu Trp Glu
 100 105 110
 Leu His Pro Ser Phe Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser
 115 120 125
 Arg Gln Leu Ala Gly Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val
 130 135 140
 Leu Gln Val Gln Ala Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser
 145 150 155 160
 Gly Leu Thr Tyr Ile Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly
 165 170 175
 Thr Arg Gln Lys Gly Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser
 180 185 190
 Arg Glu Ile Leu Ser Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly
 195 200 205
 Ala Cys Asn Ile Arg Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly
 210 215 220
 Val Arg Leu Gln Thr Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro
 225 230 235 240
 Val Val His His Tyr Gly His Gly Ser Gly Gly Ile Ser Val His Trp
 245 250 255
 Gly Thr Ala Leu Glu Ala Ala Arg Leu Val Ser Glu Cys Val His Ala
 260 265 270
 Leu Arg Thr Pro Ile Pro Lys Ser Asn Leu
 275 280

<210> 23

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> oligonucleotide 24-1443-126

<400> 23

tacggcttag taagttggag aacyaggatc agaagacagg tctgcct

<220>
<221> allele
<222> 24
<223> polymorphic base C or T

<210> 24
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide 24-1457-52

<220>
<221> allele
<222> 24
<223> polymorphic base C or A

<400> 24
tctgagatgc ccctgtgtcc tctmagggag tagtggtga gcatttc

47

<210> 25
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide 27-93-181

<220>
<221> allele
<222> 24
<223> polymorphic base C or T

<400> 25
cccagctctg ccactggcga gctytgtggc cttgggcaag ttactcc

47

<210> 26
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide 24-1461-256

<220>
<221> allele
<222> 24
<223> polymorphic base A or G

<400> 26
gatggctctg gcattttcag ggarcagtca tgtctgatct caagttc

47

SEQUENCE LISTING

<110> GENSET

<120> TREATMENT OF CNS DISORDERS USING D-AMINO ACID OXIDASE AND
D-ASPARTATE OXIDASE ANTAGONISTS

<130> 101.WO1

<150> 60/261,883
<151> 2001-01-16

<150> 60/305,445
<151> 2001-07-13

<150> 60/
<151> 2001-10-22

<150> 60/333,881
<151> 2001-11-19

<160> 26

<170> Patent.pm

<210> 1
<211> 86592
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 38388..40388
<223> 5'regulatory region

<220>
<221> exon
<222> 40389..40670
<223> exon Z

<220>
<221> exon
<222> 42666..42778
<223> exon A

<220>
<221> exon
<222> 43416..43519
<223> exon B

<220>
<221> exon
<222> 61159..61402
<223> exon C

<220>
<221> exon
<222> 64050..64711
<223> exon Ulong

2002247939 08 Oct 2004

<220>
<221> exon
<222> 64578..64711
<223> exon U

<220>
<221> exon
<222> 68126..68261
<223> exon V

<220>
<221> exon
<222> 69488..69690
<223> exon 2

<220>
<221> exon
<222> 71942..72056
<223> exon 3

<220>
<221> exon
<222> 73962..74038
<223> exon 4

<220>
<221> exon
<222> 74701..74766
<223> exon 5

<220>
<221> exon
<222> 77478..77532
<223> exon 6

<220>
<221> exon
<222> 78762..78866
<223> exon 7

<220>
<221> exon
<222> 81507..81589
<223> exon 8

<220>
<221> exon
<222> 83181..83298

<223> exon 9

<220>
<221> exon
<222> 83879..83977
<223> exon 10

<220>
<221> exon
<222> 84906..85423
<223> exon 11

2002247939 08 Oct 2004

```

<220>
<221> exon
<222> 84906..85541
<223> exon 11long

<220>
<221> misc_feature
<222> 85542..86592
<223> 3'regulatory region

<220>
<221> allele
<222> 6199
<223> 24-1443-126 : polymorphic base C or T

<220>
<221> allele
<222> 44814
<223> 24-1457-52 : polymorphic base A or C

<220>
<221> allele
<222> 72600
<223> 27-93-181 : polymorphic base A or G

<220>
<221> allele
<222> 84046
<223> 24-1461-256 : polymorphic base A or G

<220>
<221> primer_bind
<222> 6074..6094
<223> 24-1443.pu

<220>
<221> primer_bind
<222> 6559..6579
<223> 24-1443.rp complement

<220>
<221> primer_bind
<222> 44763..44782
<223> 24-1457.pu

<220>
<221> primer_bind
<222> 45257..45277
<223> 24-1457.rp complement

<220>
<221> primer_bind
<222> 72359..72378
<223> 27-93.rp

<220>
<221> primer_bind
<222> 72763..72780
<223> 27-93.pu complement

<220>

```

2002247939 08 Oct 2004

```
<221> primer_bind
<222> 83791..83811
<223> 24-1461.pu

<220>
<221> primer_bind
<222> 84307..84324
<223> 24-1461.rp complement

<220>
<221> primer_bind
<222> 6180..6198
<223> 24-1443-126.mis

<220>
<221> primer_bind
<222> 6200..6218
<223> 24-1443-126.mis complement

<220>
<221> primer_bind
<222> 44795..44813
<223> 24-1457-52.mis

<220>
<221> primer_bind
<222> 44815..44833
<223> 24-1457-52.mis complement

<220>
<221> primer_bind
<222> 72581..72599
<223> 27-93-181.mis

<220>
<221> primer_bind
<222> 72601..72619
<223> 27-93-181.mis complement

<220>
<221> primer_bind
<222> 84027..84045
<223> 24-1461-256.mis

<220>
<221> primer_bind
<222> 84047..84065
<223> 24-1461-256.mis complement

<220>
<221> misc_binding
<222> 6187..6211
<223> 24-1443-126.probe

<220>
<221> misc_binding
<222> 44802..44826
<223> 24-1457-52.probe

<220>
<221> misc_binding
```

<222> 72588..72612
<223> 27-93-181.probe

<220>
<221> misc binding
<222> 84034..84058
<223> 24-1461-256.probe

<400> 1

attattggaa	caggccacac	ttgcgaggga	agtccctgcc	tcagaaagat	tcagaaaagc	60
tagacagtca	ctggaagaac	aattacaacc	gcaagacggt	caaacactaa	acaccgctat	120
gcctcagaac	cgtacagata	atggccaaat	agatggggct	ctgggcattt	ctgagagcac	180
ctgcctggtg	gcaccccatc	ctaattggacc	atgccctcca	gtctccaagt	ggctcttcag	240
agctcacatc	cgaacacctc	ctatgctaca	ggttcttcta	gccccagggt	cccaaccacc	300
ccaaggccac	agaggccagc	cccaactcca	tcttctacat	gtgtcacagg	aaactttctc	360
atagtgtctat	ttattatgta	ctgcgggggg	ggggggccatg	tcataaaaga	aatgtcctcc	420
cttttttatt	catctccttc	taacaagcat	caaagtctca	gtcgctagca	tgtgacttac	480
agaagctctc	atgggaacaa	gacaagacca	tactgttacc	gtgacactca	cggcctccct	540
gactggtttc	tgctgttgat	tctgcctcaa	atgctcctca	aatgcacctt	gctgctccgc	600
ctccacccta	gagctcgctt	gactgcccac	ttgcccgtaa	agagtcgggt	taggcttcac	660
tcttgccaga	aaggctctgc	cagggtgctct	caacagtcac	cccctcctgt	ggctctcaca	720
aaccccagca	cctctcggtc	actctctccc	tcttatctgg	ttgtgactgt	cttccatgct	780
cacttagaag	ctctctgagg	ccaagaactg	tgtgtactgt	tgcttctttg	tttacctggg	840
cctagcccat	tgctctatac	acaggagaat	gcaaataaat	catatgctta	atgaatgagt	900
cgatgaatga	atgatgaata	aagggaatct	aatctagtgt	taacaaatcc	aggttttgca	960
atgatctcac	aggcattcat	ttatcttggtg	atgtcagggg	agtgactcca	ccctcatttc	1020
acacgcattc	tggggtcaat	gctctaactt	acttggcctc	cagttagtgg	gaaattacaa	1080
gctacacttc	aagcctctga	ctaggacctg	ccatgaagta	cttgggaatc	agtggagtat	1140
cactgtgggg	tgaggtgtct	gaggcgaggc	ccaccaatct	ccatacttct	ccccggggcc	1200
ctctgcctga	gagggtctcc	ctgcttccct	tggcagactc	tggtttggcc	ttctgggttc	1260
ggcggtgttg	tcacctcctt	caggaagcat	ttctggctaa	ggtgccccac	tctatagcag	1320
ctgggtgtaa	acctctctaa	gcaaacagca	taactttctg	tcctctcaat	tgactctgag	1380
ttctgagagc	acagcctgga	gctggcacgg	tgcttggcac	agagagctga	aatggcacac	1440
cctagtgttc	ccagtggctc	gactccccag	gctctccatc	aggacgcagc	cctctcccac	1500
ctctgatgga	tatgggacca	tggaatgctt	tgtccagcag	caactcttgc	ctccctcaca	1560
gaagggaaca	cctagcccat	cagactcacc	tttcttactt	ggaaaagtcc	actcccagca	1620
agatattctc	ctcggtgtcc	tggcgcccg	tgctgtacac	caccaccatg	taccggacct	1680
ggtccgcccc	ggcgctctcc	aggcgactg	cctggaacag	ggcagacatg	ctctcactaa	1740
cctgcctttg	gaggtggtgc	ctccctccca	tctccaatgc	aagatcaaca	ctttcagtgt	1800
tctaccttcc	cctctgggag	ttaaaaatga	agagaaaatt	cttggctggg	catgggtggt	1860
caggcctgta	atcccagcac	tttgggaggc	caaggtaggc	agatcacttg	aggtcaggag	1920
ttcgagacca	gcctggccaa	catggcaaaa	ccccatctct	tactaaaaat	acaaaaatta	1980
gctgggcatg	gtggcgtg	cctgtaatcc	cagctactca	ggggactgag	gcacgagaat	2040
ctcttgaacc	cggaagcg	aggttgcatg	gagctgagat	catgccacca	catgccagcc	2100
agagcgacag	agtgaattc	tgtctcaaac	aacaacaaca	caacaaaaca	caaagcggaa	2160
gttcttgaca	gcaggaacca	ggcctcggtt	ctctctgtag	caccagggac	gccgcctggc	2220
tcagaggaat	cacccaaaat	gcaagaaatc	agtgaacaca	tgaaatccaa	agaaagtctg	2280
tatttagctt	atttaactgc	cgtggagacc	tgtttcatcc	ctcctcccgc	ccctctgggg	2340
aactgaggag	tcaacctggc	tttggcttta	gtgcacaatt	tgagaatttg	ttgtaacctt	2400
aaagcttttc	cccttatcat	tcacgaatgg	ttccccacca	ggtttcacaa	ttaaaaatta	2460
aaacttgctg	gctgggcacg	gtggcttaca	cctataactc	caggactttg	ggaggcagag	2520
gcaggagaat	catttgaggc	caggagtcca	agaccagcct	gggcaaaata	gcaagacccc	2580
atatccacaa	aattttttta	aaaataaggc	aggggtgtac	acacttgtag	tcccagctat	2640
ccaggaggct	gaggtgggaa	gattgcttga	acccaggagt	ttgaggctgt	agtgaactaa	2700
gatcatgcca	ctgcactcca	gcctgggcaa	cagagcaaga	ccctcatctc	acaaaaatta	2760
aaaaaaaaatt	ttttaacttg	acattctcac	tgcttcttac	cagcttgatt	ctgtcttcgc	2820
aacgcagaag	gttgatcatc	acctgaagat	gttgaggcag	atcacctgtt	ggaccaataa	2880
agaaagcttt	aaaaggctct	ttacctactc	tctaggaaaa	aaaacctctg	aaaggctgac	2940
tttgagggct	tggaaaaaga	ttgagaagtt	aaaatttgtc	tacctacacc	acaggagaat	3000
caccacaaaa	acttcaagtc	tgaattttct	ttacaccact	ctgaatactg	tgcgacgtgg	3060

atggggtgaca	tggagcttac	tgtcatgttg	ttaaaagtgtg	ctcttatttc	ctgaaataca	3120
tacagtatag	gtttccaaat	acaaaatgtg	aaaaatacag	gcaagcctag	agaaaaatgt	3180
tatttcattc	aagccaatgt	tactcggcag	gttgggggtgc	ctagaaacga	cagctgtggc	3240
tggaagtaag	gcatttgcta	agagttaatc	attagagaaa	aaggacagag	catcacgttt	3300
cctcttcaaa	caacttcttc	ttctatacag	agtctcgcac	tgtcacccag	gctggagtgc	3360
agtggtagca	tctcagctca	ctgcaacctc	cgccctcctgg	attcaacaga	ttcttctgct	3420
tcagcctcct	aagtagctgg	gattacaggt	gcccatacacc	agaccgggct	aattttttgta	3480
tttatagtag	agatggagtt	tcaccatggt	ggccaggctg	gtctcgaact	cctgacctca	3540
agtgatctgc	acgcctcagc	ctcccaaagt	gctgggatta	caggcataag	ctaccccacc	3600
caggccccac	ttcaaacttc	tgcatttttc	actggaggca	gacattattt	ccataaccgg	3660
ggggggcgggg	ggaaatgttt	aagtgactct	acagatagca	gctgtatgct	ggttgcccag	3720
agaaataatt	tgaatagaaa	ccaatctgtc	attttctctt	ttcttgctaa	aaattatgta	3780
ctcttttttc	ttcactatgt	aaaacaggca	gtaaccaggg	acggcttctg	aacttctctg	3840
agctgccccca	gggttcagga	ggtgttcctg	gagtgcagtg	aggaaagtct	cttactggcc	3900
atgagtctcg	cgcgaaagcag	agaccctgtc	agaagaagcg	cacactttca	cggaggggaa	3960
agttgtaagg	gaggtgcata	attagtaagt	agcagggtgtg	actccaaggt	tgcttttttt	4020
ctctagctta	cacatttttc	tttatatctg	caaggatttc	tttctgaaga	aagggtcatc	4080
tgtagagtag	ctaatacag	cctggtgtgg	tggctcacac	ctgtaatccc	agtgttttgg	4140
gaggccgag	caagtagctc	acttgaggcc	aggcattcaa	gaccagtctg	ggcaacatgg	4200
caagacccca	tctctacaga	aaagtaaaaa	atttagctggg	ctttctggtt	cacatctgta	4260
gtcccagcta	cttgggaggc	caaggcagga	ggatcgctgg	agcccaggag	tttgagatca	4320
ccctgggcaa	cacgataaga	ccctgtctct	acaaaggaaa	aaaaattact	ctatacatca	4380
caattacaac	cccaaaagga	tcaataatgc	ttacacactc	aaatgctcca	aaaggaaaca	4440
ttgtgtttgt	tccttttgca	aaagcatctt	tttcatttta	agggagaagg	acagatgatg	4500
tccaaattgc	acttctgtc	tcagagagga	attgggtcat	tagaaatttg	tgctcttagc	4560
caggagggtta	gatctcatgt	taagcgttct	ttctttttct	ttttttttca	atagagacag	4620
ggttttgcca	tggtgcccag	gctggtctcg	aactcctgga	ctcaaataat	cctctcacct	4680
cagcctccca	aagtgtctggg	attataggca	tgagccacca	aaccagggcc	aattaagcat	4740
tctttccaca	ataagtaaaa	tttaaaaaag	aaaagaacca	tgccctctt	atctgtcctc	4800
tccagttata	caattccaca	gtgtataaca	ccctgtgttg	accctgcttc	ctatgatgag	4860
cgatttgag	ataagggttc	acattaaaga	aagccataga	cctccccagc	cccttctctc	4920
acccgtcatg	tcaccaatgc	aacacaacga	caacgaccat	gagctgggtc	ttcacctgcc	4980
tgggcccctcc	caccatctac	ccgagtcaca	gaactgcatt	ggggaaagca	aaaacaaacc	5040
cctgtctgat	aaatgcctaa	atgaaaggga	cattttccac	acagataaac	ttctttcagt	5100
gggattgttt	gctgagatat	ggaactgtg	acagacagaa	atccaaacc	cagtctgaca	5160
tccacacata	aaaaaatcat	agaatataag	ccctagaaa	ggtctcaaat	tgactggact	5220
ggctgaaaca	aactgaacta	cttttccaa	gacagaatta	accctcaatt	gtactcagct	5280
ctgcacagtg	gttactgggg	ggcctctggt	acattcagga	gacttgatgg	taattctagg	5340
gaaaaaaagg	aactaacgta	agtctagtct	gcgtctgtcc	caaggtcatt	tacagaccaa	5400
ctgtggacag	ctggcgcccc	ctctgccttc	cgacctcatc	gtccactcca	gacctcaggg	5460
cacaagagtc	agccagctgg	tggcttgcac	cctacccttc	tagtctttgg	attagaggaa	5520
ggagggtatct	gacacttagt	gagcagagct	tgagcatttg	ctttgtcata	tgtgttacaa	5580
ttaaaacatg	aacaacagct	acatttctaa	gagggcagaa	taattagcaa	attcaagaac	5640
gaagaatctg	gctgagtatg	gcacctcaaa	cctataatcc	caatgctttg	agaggctgag	5700
gtggggagat	ggcttgaggc	caggagtgg	agaccagcct	gagcaacata	gtgagtgaga	5760
cctcatctac	acacacacac	acacacacac	acacaactag	ctgggtgtgg	tgacacgcat	5820
ctatggtccc	agctactcag	gaagctaagg	ctggacaatc	acttgagccc	aggaggtcga	5880
ggctgcagtg	agctatgatc	aggccactgc	acgccagcct	gggcaagaga	atgagatccg	5940
tctctaaaaa	aacttttcat	ataattaaaa	aaaaaaaaag	aatgaagggt	ctgtttatag	6000
ctgtattgta	ctagaagtca	tcgtaataac	aatgatagtt	acccatatat	atatacagca	6060
cctactacag	gtaggatagt	tacacgcata	actctaaatt	tccatattgt	ctgaggcacc	6120
agtatttgat	gccattgta	aagactagga	aactgaggct	tagaagtcga	cctgttacgg	6180
cttagtaagt	tggagaaacya	ggatcagaag	acaggctgtc	ctggcttcaa	aacaaaact	6240
atttccacaa	accacactgc	ctcctgttac	agagactgta	ttttctttgc	ttaaaacaga	6300
cctaaatatt	atcaacatca	gtatgtgaaa	atactgactg	agccttggtg	tttgctataa	6360
attgcatggt	gtagaattct	aacctgagca	ctcagatcta	aaatgaagct	gaatgacttg	6420
aggttaaaca	aacaaaatgt	tcacaagaaa	actggccaca	atagctggtt	ggtttcacct	6480
gctgctgttc	tgaaaggtaa	aggccttctc	agctcacaga	cattcaatta	tgactgcct	6540
ctccaagaaa	tgccctgaga	tgctgtccac	ctacgacaaa	gatccactta	catgcaagca	6600
ctttttcctc	tttctttctt	tttgagatag	ggtccttttc	ttttgtcacc	caggctggag	6660
tgcagtggcg	caatcggtgc	tactgagca	acacagttag	caacatagtg	agacctcatt	6720

tacatacaca	cccacaaaaa	actagctggc	tgtggtgaca	catcagcctc	gacctcctgg	6780
gctcaagcaa	tcctcccacc	tcagccccc	accttgctgg	gattacaggc	atgcgccacc	6840
acgcccagct	aatttttgtg	ttttttgcag	agataggggt	tcactgtgtt	gttcgggtta	6900
gtctggaact	cctgggttca	agcgagatct	gccacacctg	gcctcccaaa	tcctggaatt	6960
acaggcaaga	gccaccgtgc	ctggccataa	gtgtgttttg	ttgttattgt	ttttaagaaa	7020
cagagtctct	ctctgtcacc	caggctggag	tgacgtggcg	tgatcctagc	ttgctgcagc	7080
ctcaaaactcc	tgggctcaag	cgatcctccc	aactcagcct	cccaaagcac	tgggattaca	7140
ggtgtgagat	accatgcagg	gccacgcaag	catttcttga	attcctcttt	ctaactgcct	7200
tcagctctga	gtcaagtctc	ctaagaaaac	cagtcttact	acttagtagg	cacttcttat	7260
ttaaactcag	tttgatcctc	accctattac	ttctgtctac	ttcctaaaaa	caaactatta	7320
cagaatcaag	acttcctact	acagtgtcta	tctcagagtt	ggagccaaag	gcccttcaag	7380
aaattctcca	aatgagtggt	tttcaaatgc	ttggagaaat	ccatcccaag	attaggtata	7440
cagcactcca	gatgggttatt	ttcaagtggg	cgacatctgg	ctataattca	ttttggtgca	7500
tttggttaaaa	agtcaggctg	taacttacag	cctgcaatta	actgataaac	tacagagagg	7560
aaatcctttgc	atcccagcag	gatgctgctg	accttactcc	tgacgcagac	agacatgaca	7620
taaaagggttg	gaaaatgtgc	gtggtctgct	caagagagag	catctgagcc	tctgcctgca	7680
ctggctcactg	caaacctgcg	tccactatgt	ctaaggcctt	caaactcagc	aacatcacca	7740
acaatggaag	tttctctctg	tgtccagaaa	agaagctcca	atgtaagagt	atcaacttag	7800
agccctcacc	tgcatgcttg	tgggggtgct	gaagactccg	ctggccttga	gggctgcttc	7860
cctgttgtaa	gaagaggggt	gcgcctttca	ccatgaaaaa	gctctcactt	aagctgggaa	7920
ggataagacc	agagcacagt	tagaccggaa	ttcagacagg	aaaatggaca	aagaattact	7980
gcaggggaaa	aagcttttagc	gtggacaaat	ggcatgtaaa	atgcaaatag	gatgaaactg	8040
cttttataat	aattccacgt	agtacttttc	tcaaaccttg	cttttgctaa	aagcttgctg	8100
ctggagaatt	ttcgtgacaa	aataatgctt	ctgtgacaac	acccaaagtt	ctacataggc	8160
tctccagggc	ccctttctgc	agaatactgg	acagggatct	cactgtcata	taacattttc	8220
ttctttcttt	tttttttgag	acgcagtttc	actctgtcat	ccaggctgga	gtacagtggg	8280
gtgatctcaa	ctcactgcaa	cctctgcctc	ctgggttcaa	gcgattctca	tgtctcagcc	8340
tccccagaag	ctgagattac	aggcatgtgc	caccatggcc	agctaattat	tgtattatta	8400
gtagagacat	ggttttacca	tgttggccag	gctgggtctca	aactcctgac	ctccagcaat	8460
ctgcctgctt	aggcctcctg	gagtgtgtgt	attacaggcg	tgaccacgcc	cagccataac	8520
attttctaag	aaaagagaac	aactccctga	ttaggagagg	gcagtctact	ttgtgaattc	8580
tcatgctctt	gctgttgatc	tctgcttcta	actctctggc	ttttaacaac	tccattgttt	8640
cttggtgact	tcccttgatg	gaatacaagg	atgaaattac	actttcacta	gttgtttgca	8700
ttttaagaaa	agtggggagg	ggccgggtgg	ctcaagcctg	taatcccagc	attttgggag	8760
gccaaaggcag	tggatcactt	gaggtcagga	gttcgagacc	agtctggcca	acatggtgaa	8820
accctgtctc	tactaaaaat	gcaaaaaata	gccaggcgtg	gtggcacatg	cctgtaatcc	8880
cagctactca	gaaggctgag	gcacaagaat	cgcttacttg	agccccaggg	acgggagttg	8940
gagtgcagcca	agatcgcacc	acgcaccact	gcactccagc	ctgggcgaca	gagcaagact	9000
gtgtctcaaa	aaaaaaaaaa	gaaagaaaga	aagaaagaaa	aaagtgggtg	gatactgact	9060
tgtgatttaa	cttagtcaag	gttgtcctgt	ccactattct	tgaggaaaac	ctcaagttgg	9120
cccaatgaat	ttctcagcag	aatgaatctt	tggcctttgt	tatttttagct	agcaataaca	9180
tttataacta	cctataactt	taaaaattac	aattaaaaaa	tgtttatttg	ggaggctggg	9240
gtggaaggat	catttgagcc	caggagtctg	agaccagcct	gggcaacggt	gtgagacccc	9300
gtcgtacatc	aaaagttttt	atttttaatt	tcactttcat	gacttggtta	tcaagtctgg	9360
cttttgcaaa	aattaagaca	taagaaaaga	atgcttcagc	tatgaattac	tatcaattgt	9420
tcaaaaaatac	catcaactct	caaaattatg	cataaaatac	accaaataa	ttaacaacgg	9480
ctttgcggga	ggtgggggag	gaggaggaat	agattatctt	ctagtatttt	ccaaatgttc	9540
tatattaaac	atatattaac	ctttaaaaca	tctacttttg	tttgattctc	aaaataatat	9600
aaaacactac	tatataattt	aaaaagaaca	ttctaactct	aataatttca	taaaaggagg	9660
tcacagttca	aattgtaggc	aactataaaa	atttcgctct	tgaacaacca	atgaacatat	9720
acatgatttg	aaggaaaaat	ccctaagaaa	aagcagctct	ctaatttaaag	agaaccttga	9780
aaattaagtaa	atcaattcct	gacagaaaag	cgaagatggt	ttctgtaata	caagaaagca	9840
agatcacctt	tgccccagac	atctaattgt	agtagttaaa	cgttcgaatt	ctggaataaa	9900
aaactcagca	aagtctaaag	tatgactctg	ggtgccaaga	aaatgccaca	ggaactagca	9960
tttccaatca	gcagctcctg	agatcaggaa	gactgttatg	ttctatgata	taaagtccac	10020
aataaaaatct	gttagttttt	ctggttaaat	gctcatgcta	aaaatagtga	ctgctcaaat	10080
attaagtaag	aagacttagt	tttgccttct	tgttcagtcc	tctgaattcc	aggcaattgg	10140
ttttcgatat	cttgtgacac	caatacttga	catctaacag	cattttgtcc	actactgcag	10200
atgcactgcc	gagtcacctc	ttccaccctc	tcacaggcat	atatttgtgc	tgcaagggtc	10260
aagtgttgag	gagctcagga	ttataaataa	cgaaagaaac	gagaagcagc	ctttctttgc	10320
tgtctcacc	tcactcatag	gaagtaaaaa	gctcttttagc	atccatctgg	ccgatctcat	10380

ttcacaggct	gcagaatcac	ctaacccttt	ccacctgcaa	agcttgtcac	tctctccttc	10440
cttagaatct	cacagctgag	tatgttttca	gaactgttct	tagacacaga	tcattttacta	10500
ttttattctca	tcaaaatctg	aaacagctat	gcgagagggt	ccaaactcat	gaaacctaaa	10560
acaaccatca	gttcatcgaa	gcagctggga	aaatcttttc	gagacaacat	caactgcttt	10620
tgttcatgag	attaaaaaaa	aaaaattcat	actgaccaga	aaccaagca	cgctggaaac	10680
agccaacat	taacgatgac	ctttgccttg	gaaacatga	gcaaaaattc	cccttgggtt	10740
cccttatatt	tcctttggaa	aaaaaaagga	acaatgcaac	agactaggct	ggtttcactc	10800
tgtgatcact	tacaaggcca	gctgttcttc	ctccatgttc	ctacactgat	aagaatcagg	10860
gactcctgct	ctacgcatga	agtcaggatg	gcattgattg	gggccctgga	acactctgcc	10920
tctgttcccc	cacgacaatc	aagtaacagg	catttactgt	aaaaagcaag	actggaagct	10980
gcagggaagc	ccaagtagca	gcgcattatc	ccgaagctgt	gagatcaccc	tgcgtcctgc	11040
aaatacagtc	aggagataca	gccagaggaa	accgcacgac	atgactctcc	gggtgggggg	11100
tgggggtggga	ggccgcagag	catggtcagt	cacaggattt	atgaaaacaa	gatgcagaaa	11160
gtctctgtga	cccggtcttc	tggcttctct	tctgagctca	ctctgggccc	agagcctcat	11220
gcgccctctg	cgtggctgac	ctgaatactg	tatctgacga	ctgcagcttc	tgatgccag	11280
aggcacaggc	tcccgatcca	tcagaccctc	aaagtgtccc	actggggaag	tccatgaaga	11340
aatccacatt	ggtgatggca	cgctcacttt	accagggtgc	tggggccagg	aagcccaaac	11400
ccacaagcca	tccatcccag	ccaccagaa	gtcactcttc	tcacaaaaga	tctgagtgtc	11460
ctaaaaggag	tgactaaaag	tacaaaaggt	cagacgca	cagacaaaac	ggaaatgtct	11520
tcctccaccg	ctgtaagaaa	aatcttgatg	aggataaaa	aaaaaaaaag	ccgctgccct	11580
ctctaccgcg	caactggaat	gtttttatct	ccaccacaca	gatctgttct	cggacactga	11640
ttactgccat	tcgggaagct	tcataagatt	aaagtttctc	caaagcattg	aagacagaca	11700
aaaaacctca	atcaatgctc	ctcaaaaaac	cccaggcccc	caaaatataa	acagccagt	11760
tcattccagaa	accaagccat	ggcaggaaac	cagtaatcag	ggtggtcata	cgtactaatt	11820
tgagctggaa	acctctggac	agcagaagca	gtgggttggc	tgaaggaaga	tgcagaagtc	11880
ggtaaaataa	aagaggttcg	tggtgcagt	gtcacatct	ctaacgctcc	ctacaactgc	11940
cctccgagct	ctggccatct	gctccctatg	gagatcagga	aaagccagga	ggctgccgag	12000
tgcttccacg	agggtctggg	agccaaactc	tcctcagagt	cctaccgaa	aagcaaattg	12060
ctcttgtgga	actcttgtct	tcctctgata	ttttggctga	aaaaggccct	tgtcccagca	12120
catcctgatg	aaagagggcc	attcagcaaa	acagctgagg	ttcctcta	cactgcactc	12180
ctacgggctt	ttctgtaggc	cggagaaaca	agcaccgggg	tgtgcattcg	acattgtgag	12240
ggcaaacaac	tggccccaag	gaaccaaccc	caagcaacaa	gaccccttc	cgattcaaat	12300
caacattctg	aaggatgact	ctttctttca	aatcagcatc	catttacc	acggtgacgg	12360
tgacgtgggc	agctgccgca	gttagttatt	ctgcgtactc	aaagcacggt	tacatcctga	12420
aaattcttca	gtcatgctaa	cagctatctg	aggggacacg	ccaggtagag	gggaccacat	12480
gcacacctat	gagatgctct	gggatacga	cggtgcccaa	ggcaggtcag	gctgcaaagg	12540
tcctaaaggt	tggaggtgtg	atcccaaacc	ctccaggcac	aagccagcca	agagctgtgt	12600
ttttagcggt	tccttcagtg	agagaaataa	gttcaggatg	tgaataacca	tgacgcagga	12660
gagaatggaa	taagtaccct	aagaaagggg	ctcggctagg	gtttacaaga	gggaggagg	12720
agcatttaac	tgggtgacttc	tggaaacaatt	cctgaaggaa	gcagcactga	gtaggggctt	12780
ctcttccttc	ggctcacaag	tgaccaagcg	atcctcccta	cggattaagt	gaaacacaca	12840
ttaccatgat	tctggttttg	cagggtgagga	aacccagct	tgccataggag	cacatatctc	12900
tacaagatgg	ggctggactc	acatctatct	gccccacgcc	cacctgtta	acccctgtta	12960
agcagctgtt	ctactcatcc	agaatgaaaa	tcagagccat	tatgctgcgg	tcacatccgc	13020
tcatgcctgc	ccagggtgct	aatggcaaa	ccactaaggc	actgagaagt	cagaatgtgg	13080
atcacatctt	ccgtccttct	tcccagtggt	tgaatgcatc	atgctgtgga	aagagagaga	13140
aggaaccatt	caagcaaaca	gaactccagg	aagacgagac	tgtgccgggg	ttcttccatc	13200
tgcccaagta	gaaatcagaa	gggcagggga	cccacagcct	tatcctaccc	accactgccg	13260
tcatagttag	gggacaggac	acatcctttg	gcccttctgc	actgcataga	ggctaaggag	13320
ttctgtaaac	cacacagcca	cgctgaccaa	gaagtcgctt	tcaaggtaa	tttctcatca	13380
acaggactat	tatttactga	ggatctccca	tgtggccaag	gctgtaggag	gtacttagct	13440
acgccacgtc	attgaactct	ggcagttctg	cagggtaaag	tattttctcc	atatgacaaa	13500
cgaaggaagc	ccgtcaacaa	ttccaaaata	gaatcaccag	ggatagcatg	gacaacgccc	13560
atgggtgactg	ccgcgcttta	aggtttaaga	aaagtataaa	ctgggggtga	tgactcattc	13620
ctgtaatccc	agcacttttg	gaggctgaga	tgggtggatc	atttgagggtc	aggagtctga	13680
gactagcatg	gtcaacatgg	caaaaccttg	tctctactaa	aaatacaaaa	attagccagg	13740
tgtgggtggg	catgctgtga	atcccagcta	ctcaggaggc	tgaggcagag	aatcacttga	13800
accaggagg	cggaggttgc	agtgaagcaa	gatcgacca	ctgtactcca	gcctgggcga	13860
caaagtgaga	cactgtcttg	ggcagggggc	gtggggaaca	aaagtaaaaa	caatgggtctg	13920
ggaattcata	tttctgggtt	ccaatttaca	ttctaccata	tatactctga	ttaaccctta	13980
gaattaaccc	ctagaattcc	ttacagggtt	ctgttcattc	atccaaacag	gcaaacattt	14040

gcagagcatg	gagcacaggg	taagccaagc	cagcccaagc	tctgataagg	gcaaagacag	14100
ccatcctctt	taaggaatgg	gtatatgtgc	tggtgatctg	ggtgtctgcc	ctgctgcata	14160
gaaacagcat	ttcttgaaga	acaaaaatag	taggtataga	aacatcacag	tatggaatat	14220
ccaaacaccc	ctgaattcca	actctgggtca	tacattgaaa	caacctatca	aactcctaaa	14280
acacattcat	gcccagggtcc	agcctcagca	gagtctaatt	cggaagggtct	gtgatgagtc	14340
ctgggcatct	acttttttaa	aaagttccag	ggagctgggc	atggtagctc	atgcctgtaa	14400
tcacagcact	ttgggaggtc	aaagtgggag	aatcagttga	cccctggagt	tcaagattaa	14460
cctgggcaac	gtaacaagat	cccattctcta	caaaaaaata	aaaataaaat	tagctaggct	14520
tggtgggtgtg	tgctgtagt	cccagctgct	caggaggctg	aggtgggaga	atcacttgag	14580
cctggtgagg	tcaaggctac	agtgaagctgt	gaccacacca	ctgcactcca	acctgggaga	14640
cagatcttgt	ctccagaaag	ttccaggggg	tgcttctgat	gcacagccaa	gttttaaaaa	14700
cctcagaatc	aaataacatc	atggccaggc	atgggtggctc	acgcctgtaa	tcctggcact	14760
ttgggaggcc	aaggtgggtg	gatcacttga	ggtcaggagt	tcaagaccag	cctggaaaac	14820
atggtgaaac	ccagtcttta	ctaaaaatac	aaaaattagc	tgagcgtggt	gacgcacact	14880
tgtagtccca	gcttcttggg	aagctgaggc	acgagaatca	cttgtaccct	ggaggtcgag	14940
gctgcactga	gtggagattg	tgatcctgga	gtccccactg	cactccagcc	tggtatgagag	15000
tgagactgtc	tcaaaaacaa	acacacaaac	aaacaacatc	agaagacaca	gagaaaacag	15060
tcttctccat	gggcttcata	aagatacctc	tcacataggt	acacgtcgat	gttttctgct	15120
gttaaaagggt	aacacccaata	aaaaggcctc	gtgctctcag	aaggtgggtg	atgtgattag	15180
gtgcaataaaa	gggaggtcat	gctaggggtca	aaaacaaaat	aatactctct	ttggaagcag	15240
taaaacagat	gctagtcttc	tactacacac	tttcagagac	ctgaatgttc	ttctggccct	15300
ctaagggaga	cgctgcatca	tgacaatacg	aaatgatgac	agtgaagca	aaaacagatc	15360
agacctgtgc	tgtgtgaaac	agacatgggg	tctcgctatg	ttgccaggc	tggtctccaa	15420
ctcctgagct	caagcgattg	ttccgccttg	gcctcccaaa	gtgctgggggt	gacagctgtg	15480
agccaccgag	accaacctca	gatcagacct	ttgacaaact	ctgctgtgga	caaagcattc	15540
tggtgaatgt	caactcatct	gatcttcaca	aaaccgtgtg	gaagaccaga	caggcattat	15600
tacactaatt	tatgcctaag	gaaacaggga	gttaaatagt	acaaatttag	gatttctgat	15660
gctgtatctc	gaaaaaaaag	tagagaatat	gagcctgaag	aagaggccct	gtaaagggtc	15720
ccagattgat	gggacaggct	gagacaaacg	gaatcacttt	tccttgata	gaactaacc	15780
tcaatggtac	cccactctgc	atggtgatta	ctgaggggac	tgtcaattgt	ccagcgaact	15840
tgatggtaat	tctaggagaa	aaaggaacta	atgtaatgct	gtcagcatag	aaagatgggt	15900
gccaacgagc	attccaaaaa	ggaggctctg	ttaattcgggt	ttcgatcaac	aagtatttgc	15960
tgagtgtcta	ttgtgtccgg	tcagtgtctaa	ggcctgagaa	tttagaagtg	aaacagacc	16020
ggtttccacc	catgccacag	accactccac	acctggtctg	gagtgcact	ggagggccag	16080
gcaggcacag	gacagtaact	tcgatataag	gcagcaagtt	ccacggtgga	aggaggtgga	16140
aggtgcacag	gcacgtaac	taaacgggtc	agggaaggct	ccctggaaga	aatgaagcct	16200
gcgaggccct	gaaggatcag	taaacagaga	ggcataaggg	gcaggagagt	aagatgatta	16260
tgctacatgt	accttattgt	gaaccagga	ggatttggcc	tctgtcataa	aaggccccc	16320
tgtgggttca	taaacctcaa	tttacaatt	gtgctttata	tatcagttcc	ttataagttt	16380
ggttagcgta	aattggtttc	ttagaacttg	atcatccctg	agtgaactca	caaattcaag	16440
tttcagaatg	tgcaaacctc	agaaacaaac	ctcatgcttg	tggttgagac	atcgactgt	16500
caacatcaca	aattctcagc	acctgaatgc	ctggtatact	atcaacatat	attgttttaa	16560
atatgtaaat	aatagctttc	tagttataga	gagtttgctc	ctacattttt	ccacttaatt	16620
tttacaatcc	cattcccctg	atgaaacaac	cccagcctgg	gcaacatggc	aaaatcctgc	16680
ctctacaaaa	aatacaaaaa	ttagctgggt	gtgggtggcg	gcacctgcag	tcccgggtat	16740
ttgggaggct	gaagtgggag	aatcacttga	gcccgggagg	cagaggttgc	agtgaagcaa	16800
gattgtaccg	ctgcactcca	gcctgggaga	cagagggaga	ctctgtccca	cccacccccg	16860
cctcccaaaa	gaaaagaaaa	gaaagaaaaag	aaaatgaaac	caccaagact	gggagaagat	16920
aaatgacttg	tctgtggtca	tctggctaata	aagaggtaga	atggggctga	aaaagttcgg	16980
tgctcttcc	gaagaatcca	taggtcagaa	agcagcacca	tctgacctgc	agcaatagca	17040
gcaacgtgga	aagctaatac	actgacctca	aaaccactct	cagtgaggct	ctggatggat	17100
tcagaacccc	aggcctagca	aagtgaagtt	gataaagatg	taaaggagat	cgaaaattca	17160
ccatttgagg	agagattagc	taaagactgc	aggtcggatg	gaaaattctt	tccatggttc	17220
tcccacagggt	tcttccctca	tttggaactc	gtgttttaaaa	gtcacaaaaga	ccctgagttg	17280
ggccaagggtc	tcgttcttct	tcactgtggg	ccttgcaagt	caacatggca	gggcctcggt	17340
ccaaatgtca	ctcttcagag	cctaagaaaa	caagtaactt	tagggacaca	cctgtcaacc	17400
ggagctccca	aattgtaccc	ccctaaacac	ataatgctga	gcatagaaaa	attccagctc	17460
tgagagcggt	tatacttagg	gaaaggggtc	acagacaagg	aatgctggca	gggctcatta	17520
caaatatctt	tgctgctgga	acatgtattg	tttggttaga	aggcgtaggc	ttctctcaga	17580
gagaagggaat	gtccaaaagt	atttcagaca	gtaagagaca	ttctctgagc	cagctacaca	17640
gctctccttc	aaaccaacgg	gtagcggcaa	gcagctgaac	tgaccagcga	gctcgcaaaa	17700

gcaagctttt	ttttttttt	ctccctaagt	aagacagcaa	gtgatgtgtc	ttggcttggt	17760
ttagcaaat	ttaagatagt	tccctgatga	ccccaagagc	cctcaggccc	catggaagct	17820
ggagctaagt	catcttcctc	caagcatcat	ctgctctacc	aggatctaag	ccccttcacg	17880
agggcagaag	gtataaaggc	tgactgtgac	gggaaatgct	atggcagcaa	agacagccaa	17940
acacgccaga	aataacaggc	acatgaagga	aatgtttctg	agacagctca	aaaattccga	18000
gaagagatta	tcccgaactg	cccaggttct	cagccctgtc	tatggtatgc	agccccatac	18060
cacagtcatt	tgtcaccgag	tcctaacttt	gtcagaggcc	cctcctttca	ggtctctcag	18120
gcaccaccca	gttctggccc	tcctcacccc	cgtgagccag	gcgacatcca	agcagcccca	18180
cgggtgcacc	ggctctgtgc	tgcattctct	gaatgtccct	gaaggccagg	gctgttgtat	18240
tctctaccca	ctctctgtct	agtatgggag	ccactggcca	gatgtgttga	ctgaacactt	18300
aagatgcagt	aagtgtgacc	aagaaactgg	gtttgtcatt	ttatttcatt	ttagttaatt	18360
taaattttaag	tttaattagc	tacatgaggc	tatcagctgt	ggtataggac	agcagagctc	18420
cgggaagctt	tggcctgggt	agaagaatca	ggacaagcgc	ctccctggcc	tctcgccac	18480
tctgcacagc	cgctaaccat	tgctctcatg	acattctttc	ccagccccag	aacttttagc	18540
catgtgacat	catctattga	ttagagtcca	aacttcttgt	gctaactctc	tatgggttgc	18600
cacaattagc	cattgtatgt	cgtaaaccta	aatttcattc	atctgctatg	tcctgacctt	18660
aggggcttag	aatatagtta	gaaaacagta	tttcagaata	aaaaaccatt	cttgtattac	18720
ctctcgact	attccccctg	ttctccatgc	ttcgacttct	ctgttctatc	cccagctata	18780
gcactgctcc	cataaagcct	actcggttgc	tcggctcatg	gtgtccttcc	tcccactctg	18840
ctcccgacgt	catgacgtgc	ttccagtgtt	tatgccttct	ccaggaagct	ttctcttgtg	18900
gccctcgctg	tgagctatag	ctcctccctt	tcaacatctt	ctagcacctc	ctcttactgt	18960
gcaggtgagg	acactgaggc	ttaagggtta	agtcaacttg	tcaaggtcac	atacagctcg	19020
tcctctgtat	ccgcgggttc	ctcatcagtg	gattcaacca	actacagaca	gaaaatacag	19080
tatttgaggg	atgctgaact	ctttgaatta	gtgggttctg	cgggtgctta	agcatccatg	19140
gattttgtta	tcctcggcaa	aggcgggggt	cctgaaacca	atcccccttg	atactgaagg	19200
aaagaccacc	cttagtgata	ggaacctagg	aacccaagtt	ccctcatttc	caaactcgtg	19260
tccctgaccc	acttattttac	taactagtgg	tgaagccatc	ttcctgccag	tatatattta	19320
cttcacaatg	ggatgtgagg	gccaggatgc	acatgctttt	taaatctccc	tctgtgcttg	19380
acatacagta	gattgaaagt	aagtgtctgg	agatacactg	gccaagctgt	gctcttctct	19440
gaagtacagta	ttccaggagt	aactcaccct	ggtcatctct	gtgccctggg	cacactgggc	19500
actccccaca	cacaggttga	acctggcaaa	taagactcac	agcatcatgc	cacgtgagag	19560
ttaaagccac	ctggaggtca	ggtcaggtct	tcctgacaac	tgagtgtctc	aaataacaca	19620
acagcagcta	agttccccac	atcaccttga	gtgtctggag	agctaggcct	atgacttctc	19680
tgtctcagga	tcctctctcag	tgcccagaaa	acagtggaca	tcaataaatg	taacaccaat	19740
aactccttcg	ttgagcgctt	tgctaagcac	atcaggtatg	ttaactcatt	tattccccag	19800
tgtccatctc	tcagtgcttt	atacatagcg	gaactgagtc	tcaattagcc	gagcgtgggt	19860
tcgtgctcct	gtaatcccag	ctattgggag	gcacaagaat	cactcgaacc	caggagatgg	19920
aggttgacgt	gagccgagat	tgtgccacag	cactgcaaca	gagtaagact	ccgtcttaaa	19980
aaaacaaaaa	aacagaaaac	aaaacaaaca	aacactgagg	ctcaggaggg	ttaagtcacc	20040
tgcccaagtt	catgagacca	aggagccggg	aagcaggaag	gggaaggcag	gagtgttaact	20100
ctgaaacctc	tgtctttagg	cactggcttt	cagctgaact	gatacctctg	gaaaacagtc	20160
tcaaaaaagt	ccacttctcc	tcccaacaat	tcagacctaa	aaaccatttg	gcgggggaagg	20220
gcagggcaag	cttctgagtt	ggggaggggg	tgtgggatcc	caagctgagg	tgtctgttgg	20280
caagcagggg	gcaaagggca	tctgtgcagg	gagggggctg	caagggagac	agagactgct	20340
cacaggcaag	gaatgaaata	ttaaaccatta	atgttaatat	taatatttat	aattaatata	20400
tttatgatat	atagcatata	tacatattat	attaattaat	tataactata	ttaatatata	20460
taattataac	tatattaata	taatcaatta	taactatatt	aatataatcg	attataacta	20520
tattaatata	atcgattata	actatattaa	tataatcgat	tataactata	ttaatatata	20580
cgattatatt	aatattaata	taatcgatta	tattaatatt	aatataatcg	attatattaa	20640
tattaatata	atcaatatta	acaaatata	actatataat	ataaataata	cctaagttta	20700
tataatattg	ataatgttaa	tatttattaa	tatttcaggg	acaatgggag	tcatgaatat	20760
ggagagacaa	aactagaaatg	aacccaagg	tgctgcatca	gaattgaagg	taccagtctg	20820
aactcatagt	tttcaacctt	ttgaaataaa	tatagatgca	cgtgtgtgtg	tatgcacgta	20880
catacaaatg	ttccctaatt	ctgcccattg	agaggcctgt	ggttagcaac	acccaacag	20940
caataagcag	acctagcttg	gctcctaagt	ttcatttttc	actaaaagga	accagagccc	21000
cttgagataa	ggactgattc	cacaggtggg	tagggagcat	ctgttgccag	aaagcaagaa	21060
agcacttaaa	gaatgatgtg	gacatgtcaa	agggacacag	aagccagcct	ggatgagatc	21120
ccactggccc	taactgtcca	caaggacaat	ttgagcaagg	atgtcaacaa	tttaagagca	21180
gattataaac	cactgaataa	aacagaaaaa	tacaaagaat	tgaaacggac	attgatggca	21240
gacaggatat	taacataatt	ttaaagtatc	tctccaagga	atgcttctga	atgatgaagg	21300
ggaaaagaat	aactgtacag	tggaaaagcc	tggtaaaacc	caccttagtg	accaaagtga	21360

atgtcaccat	agtgggacaa	aaggaaatca	agtgccacct	tatgggattc	aacgaggacg	21420
cagcatccct	tgggtgatgt	tccagccaaa	tacacgtgcc	cggtggaatc	acacaagaac	21480
atcagacaca	ctcacactga	gggacactct	gcaaactgac	agtactgggc	acaaacatgt	21540
ccaggtcatg	gtcgaccgca	gtggctcatg	cctgtaatcc	cagcattttg	ggaggctgag	21600
gtgggcgcat	cacttgagggt	caggggttcg	agaccagcct	ggccaacatg	gcaacaccct	21660
attctctact	aaaaatacaa	aaattagccg	agcgtgggtg	agcatgcccg	taatcccagc	21720
tacttggggc	gctaaggcac	aagaatcgct	tgaacccggg	aggtggagggt	tgcagcgagc	21780
tgagatatca	ccgctgcact	ccagcttggg	cgacagagtg	agtttccaac	tcaaaaaata	21840
aaaaaataaa	ataaaatcca	ggccacaaga	gtcaaagaaa	gactgaggaa	ggttccagac	21900
tgaggagag	ccaagagaca	ggataactag	atgcaatggg	cagtcctgaa	ttggatcttt	21960
tgttatgaag	gacaacgctg	ggacatatgg	tgactcttga	atgggggttag	aggactagac	22020
ggtgggaatg	catcagagtc	agtgtcccgc	gtggatggct	gtgttgcggt	tctgtgggag	22080
aatgccctgg	tctgtattcc	aagggtaatg	gagtagcagg	ttgacaaatt	actttcaaat	22140
ggttcaaaaa	agaaagttct	tttactgtga	cttgcaattc	ttatgtaagc	tggaaattat	22200
ctcaaaatta	acgagaatct	tttatcgacg	tagtatttta	catatttatg	gaaaacatgt	22260
aagtatttgt	tacatgcata	aactgtgtaa	tgaccaagtc	agagtatctg	gggtatccat	22320
gaccttgagt	attaatcatt	tgtatgtgtt	gggagcatta	caagttttcg	agttaccaat	22380
tttttttttt	ttcctttgag	acagggtctt	actctgtcgc	ccaggctgga	gtgcagtggg	22440
acgaccacgg	ctcacgcagc	acagcctcca	cctcccaggc	tcaagcgatc	cttcacacct	22500
aaccacccaa	gtagctggga	ctacagggtg	gtgctgccac	ccccagctaa	ttttttaatt	22560
tttttgtaga	gacagggtct	cactatgctg	ccagggtctg	tactgaactc	ctaggctcaa	22620
gagatcctcc	cacctcggtc	tcccaaagtg	ctgggatcat	aggcatgagc	caccataccc	22680
agccaaatct	tttaaagtta	ttttttaaat	ctccacttaa	ttcgattttg	gtaaaacacg	22740
acctgtaatt	tttctttatc	ggtaggtaat	aaaagcttca	gatgatttta	ctgatcactg	22800
gtatgggcat	atttcatgac	tttgcccttt	catctcttgc	atagttttac	cctcaccaag	22860
caagaccttc	cctgcctcag	cactgtttgc	cctcttcgtg	ttttccagaa	cagaagtggc	22920
cctgtttcgt	gcccagagca	gaagagaacg	atgaagagct	ctgctctccc	aggtcttcc	22980
ggtctgtgtg	tgtccagggt	ttgagggcct	ctcacataca	cggctctgga	ccacgtaaga	23040
tctaatttta	gcattttcct	gctcggagac	cacaatgttt	ggaacagcag	gggctgacct	23100
gcccgtgcag	gcctcctatt	gtgaagggca	cgcaagccca	ggataaccgca	gccctgcagg	23160
atgtgactca	gcacccgtgc	tcagtgtctg	ggcggccagc	agctctggca	ccaagtgtctg	23220
ctgctgacct	cacctcttaa	gaccacaaat	acccagggtg	attggtggga	taggcatgca	23280
gcacagctc	tcctgtttaa	gacaacttgc	ttgtccatcc	attatgctgg	gcttccttgt	23340
gaacaccaca	ggtatctatc	aggaagagtt	cttccgagga	actgatctgc	tggtattttc	23400
aggacaccaa	gaatcaagag	attggtcttg	tttctctctt	tgctttgact	accaggaaac	23460
tcaaagtcag	atctgtggcc	aaattctggt	aacctatacca	atgctatgtc	atgtattaca	23520
tgtacaaacc	ttccccttac	ttcatcttat	tttcttctgc	tttcttcgtg	tcccgaatttt	23580
ctcactaatg	ttacattcta	ttgttctcta	tgaatgttgt	aagggtgtttc	aaatcctttt	23640
tggagggcac	tactgtagat	acaacacaca	ttacccttga	gggataagga	ctcttttttga	23700
ctccacacag	aatccctggc	atttggcaaa	gaaccatata	ttaggcacta	aatacacatg	23760
ggctgaatgg	aaaaagccaa	tagctaagta	aaaaccacct	ccattaccat	attgtttcac	23820
aagagggtct	tttcccttcc	atctcatgag	gtggggcctg	gtcaggagtc	cccagggcct	23880
gggaattagg	ttccttaggg	agccttcttg	ctgtaggggc	agccaacagg	tcagtggcct	23940
tgactccaga	cctaaagagc	cactcctaga	ctcccagctg	caacagacac	agcgtggcac	24000
gggtgggcct	ggccactggg	gaagtgacaa	gtgatttcca	gatgctgcag	ccagcctggc	24060
tctttccaga	ccacactgaa	ggccccttcc	tgtgggaatt	ctgatggggc	ccagatttgg	24120
ggaaacacgc	ctcgaggact	cttggaaggt	gcgtgccagg	cctggaccag	gaatgacttc	24180
tgtgggcaca	gggagagacc	aggcatttcc	taacacagga	ccttgaacag	ccttctctga	24240
aacaaagtct	ttctaaaaat	agcttcaaaa	gtaacatttc	aagaaaagaa	agaaaaaaaa	24300
aactgtaaaa	gtaaaggcac	tcaagaatga	tatttcccag	ataaaagcct	ggcacagggt	24360
tcagaggaac	ttgcaggaaa	acagggtcaag	gctgggtttt	tcctcttagg	tgtcacttgg	24420
ttaacattgg	tctttggagg	ggaacaagtg	cggcaggaag	ggctggcact	gaaaaatgat	24480
gccactgggt	ataggccagg	gccagacact	gtacacagaa	caagactctc	tggaggcctc	24540
aggagggccc	tgagaggagg	aaggcagggt	gtggggccag	ggtcagacat	gcaagtgagc	24600
taagtggcaa	ggccgatgcc	ccatccagaa	gccccgctct	gaccacacgc	aggctctccc	24660
ggcatgtcct	catttatgcg	gcagtctctt	gtatctcact	gcaattctgc	ccccacactg	24720
caggctggcc	agcgtggcct	cctcataagc	acatcaccct	gcaccccgac	actgactaca	24780
cccacaaagc	aggagcccc	gcaccctcca	gcccattcgc	tcagtctcgt	ttgaaaatgg	24840
ctcctctcgg	gggctgggcg	cagtggctca	tgctgtgaat	cccagcactt	tgggaggccg	24900
agggtggttg	attgcttgtg	gtcaggaatt	caagaccagc	ctggccaaca	tgggtgaaacc	24960
ccatctctac	taaaaaatata	aaaaattcgc	caggcatggt	ggtacaagcc	tatagtcccta	25020

gctacccaag	aggctgaggc	aggagaaatca	cttgaaccca	ggaggcagag	gttgcagtta	25080
gccgagatcg	tgccactgca	ctctagcctg	ggtgacggag	caagactgtc	tcaaaaaaaa	25140
gaaaaaaaaa	aagaaaatgg	ctcctctgga	ttttgattaa	tcctattttg	attaatcctg	25200
gtttctcatt	ttcagccttc	cttgaagcag	cttagcccat	ctggatgtcc	tcctcatctc	25260
aggaattttc	taataagctg	tctaaatcca	gagatccgac	cacagaacaa	tgaatgccaa	25320
agatgagttc	taaagatgcg	agtactttct	ttctaaacgg	acgctgcttt	gtgtatggct	25380
ctgctcctgg	gggcagacgc	ggcaggctaa	gccctgcgga	ggaggaggtg	agtcccagca	25440
gagggtcact	tcctctcagt	agcccggctg	gttttctcca	ctgcagggtc	agaccatagc	25500
cctgaccag	ctagaccccc	ataagcgcac	gaccttgctc	tcaccgtggg	aataaaactc	25560
gtgatagtca	gttacaataa	cacagcaaata	gatgagcagc	acaatataaa	cacagatcta	25620
gattgggtgg	tctgaggact	cattctttaa	tttggaggcc	atcacctaata	cttgtctttt	25680
cactttacat	agcaggagac	agggacccag	agaagtgaag	aggcgttgcc	ttaggttgca	25740
cagcagatga	cgctctcaa	gatggaccct	aggttgtctg	actccgtctc	acagctttgc	25800
cccatttatc	atgaagatga	acgctggtaa	cactgctacc	tacgagctga	gcttgacgc	25860
acattcctgg	tgtgtacatg	catgctgca	cgctcacgca	atgtgctaag	tgcacaggaa	25920
ggagaccaga	gccctgaggc	gttcttttga	agtctaagta	ctggtgtttc	gaaagtttaa	25980
tgaacacctac	tagactctga	gcaaaattcg	ttttacgtta	accttaataa	aaagtttaata	26040
taagttctga	cagaattaac	tcttcacgtc	tctgtcctca	tttgtcccca	ttctagaatg	26100
agttttctaa	ttaaaaaaa	tatatagggc	cggtgtcagt	ggctcacgcc	tgtaatccca	26160
gcactttggg	aggccgaggc	gactggatga	cctgagctca	ggagtccgag	accaacctgg	26220
ccaacatggt	gaaaccccg	ctctactaaa	aatacgaata	attagctagg	ggtgggtggc	26280
catgctgtga	atcccagcta	ctcgggaggc	tgaggcagga	gaatcactgg	aacctggag	26340
gcagaggttg	cagttagcca	agatcgtgcc	actgcactcc	agcctggtga	cagagcaagt	26400
actcccatcc	ccccccacaa	aaaaaaagta	tatatgtgtg	tgtgtgtata	tatatatata	26460
tagctaggca	cagtggctca	tgcttggaat	cccagcactt	tgaggaggcc	atgtgggcag	26520
atcacttgag	tccaggagtt	caagatcagc	ctgggcaaca	cagttagacc	ctgtctctac	26580
caaaaatata	aggtgggtgtg	cacctgtggt	cccagctact	tgaggaggctg	aggtgggagg	26640
accaattgag	cccaggaggt	cggggctgca	gtgagctgta	atcatgccac	tgtactccag	26700
tctgggcaac	agagcaagac	tctgtctcaa	aaagaagaaa	agagagagag	agggaaaaaa	26760
aattgaaggc	aaattctgat	tttcaaatca	aacgttccaa	caaactgcag	aaataaaacc	26820
cgagttaaac	caaaaggaac	agccaaacag	cacaatgacc	ccaatgttta	aatatgcccc	26880
aatgtttaaa	agtgggagtc	aatgggaggc	cactacctac	aaggccacag	gggttagggc	26940
aggactcagg	tcctgaatc	acagcagcct	gcattcaaac	cctggctcag	gcctcccacc	27000
agcctcgtgg	aactggtttc	ctaaaatgag	gagagtccct	actttgcagg	cttgtgacaa	27060
caagatgaca	gcaagtgcac	aagttccaag	cccagagcct	gcagcctgca	gaagctggcc	27120
tcattaccac	ccggatgttc	tccgggctg	acagatgaa	gggatacgt	gacaatccct	27180
gctttaagta	cagctcaggg	agttgacggg	acctgcccaa	gcacatagt	atgccgctaa	27240
tggtcacca	ggaagaatgg	actgcaaagc	ctggttcttc	tgataaaactc	cattctgtct	27300
cccagtgtgg	gttctgatgc	atagggagga	ggaaaagaca	gtgcttggt	tttgggggtga	27360
agagcacagg	ttttggagtc	aatgagacat	ggagtatgag	ggtctcagct	ctaccgttta	27420
ctactaaata	aaaacaggcc	actgacctct	ctgggggtta	gtcttctcct	ccagggaatg	27480
ggaattcaaa	tgtccttaca	gggttttcac	aaagattaac	tgaataaatg	cacacaaggc	27540
aatcacagag	tggagtatgg	gtgctccctt	ttctctctc	catccctgct	ttattttttc	27600
gcctgggcac	ttaccaacac	acgattattg	cgcttggtta	ttttatttac	tgtcttgtct	27660
cctcaacaga	atgtcagctt	ccagagcagg	aatttttatt	ttgtttgttg	ctatattccc	27720
agccccataa	acagggcttg	gcacacagta	ggagctcaaa	aaatatttgt	tgaatgaata	27780
gctcacaagc	agacagatga	ggacagaggg	gtcttgagac	tgatctaaca	gcaccgatata	27840
tactaaactg	caacggaggc	aacggtggga	agaatttctc	tgtcctttgt	ttcctgaaag	27900
tccaagacca	cttttagttg	ctcaacagga	aacaatactc	aacttacaag	acctctaggg	27960
cctatccagg	gcaaactggg	cactgtgagg	caggaggtca	ggcagccctg	tccttagggg	28020
ggctcacggg	ctagtgggca	ggccagctt	cttcatatgt	gctcagaggg	gccccgtgct	28080
tggtttaata	ctctgttggt	gccatcttga	aattcttaata	aatgtttgtt	gttgtgtgtt	28140
gtttgttggt	ttgagacaga	gtctcactct	gtgcgccagg	gtggaatgca	gtgggtgtgat	28200
ctcagctcac	tgtaacctcc	acctcccggg	ttccagtgat	tctcctgcct	cagctcccca	28260
agtagctggg	attacaggca	cgcgccacca	tactcggcta	atttttgtat	ttttagcaga	28320
gacaggggtt	caccatgttg	cccaggatgg	tctcaaaactc	ctgacctcaa	gtgatccgcc	28380
cgctcgggcc	tcccaaagtg	ctaggattac	aggcgtgagc	cactgagccc	agcctcttaa	28440
taatgttttt	aaaaggggct	ctcccatgtt	cattttgcac	tgggcttcac	aaattacgca	28500
gccagtcctg	cattacagga	aatatttctg	tacctaagta	catatactac	aaagcaagta	28560
ccaaacacca	aggaacact	aaggagagaa	aaacgcctgt	gagaagaaaa	aggaagacac	28620
gaatcattcc	caacagaagc	tgttaccatg	aagggaagta	gggcaggggc	atttgttgaa	28680

tgtctactat	gggagaaggg	gttcgcatca	tgagcacatt	taattctgac	aaccacccta	28740
caagctgtgt	actatactgg	ccatttgaaa	ctaaggcctg	cccagatca	tataatagcc	28800
taggaggtga	caaaggacag	acacaggagc	caaaccctatg	cccattccctc	cctaagtcca	28860
aatcataga	aaaaaaaaaa	taagaatcaa	catgggcggt	tatttttaag	gccagcatgt	28920
tcaaggtggg	ggcaaatcca	agagacacta	agcctcagag	catgaacaag	catgtgggtg	28980
ctgagtggag	gggaccagtg	tttaccaggg	tgatgtcaga	ctctgcaagg	ctcgctcccc	29040
gtgtttctgg	tctcttccca	tgagcaccag	gcacccttta	ccatcccca	actaggcaca	29100
tctgtaacgc	tgaatggaag	cctacttggt	tacatgtggt	ctatgttaga	ctgggggcat	29160
ccctagaaca	cacacagatt	gactggtggg	cagaattctg	ctaggtgcat	gcacccgagt	29220
gagcctttct	ctttgaatgt	gggagggacc	agagaacaag	atgggagagc	tgttccctta	29280
attaggctgt	gctgcacatt	aaaggcggtg	agacagtcat	tccagtata	acaatctgtc	29340
ataagaccct	acagaagcag	actctcctgt	tggccttgaa	gaagcaagca	ccacgaattc	29400
tccacagctg	caagaaaatg	aattcaggcc	aggcgtggtg	gctcacgcct	gtaatcccag	29460
cactttggga	ggctgaggcg	ggtggatcac	ctgaagtcag	gagtttgaga	ccagcctgac	29520
caatacgggtg	aaaccccatc	tctactaaaa	atacaaaaat	ttgcagggtg	cacctacagt	29580
cccagctact	cgaggaggctg	agacaggaca	aaaatttgca	gggtacacct	acagtcccag	29640
ctactcggga	ggctgagaca	ggacaaaaat	ttgcagggtg	cacctacagt	cccagctact	29700
cgggagcctg	agacaggaga	attacttgaa	cccaggaggc	agaggctgca	gtgagccgag	29760
ctgcgaccac	tgacactccag	cctgggaacc	agagcagaaa	aaaaaaaaaa	aaagtaaaaa	29820
aaaaaagaaa	atgaattcag	ccaagaacca	cgtgagctta	aaagaggacc	ctgggggttca	29880
gacaagacct	cagccccggc	cagcaagcct	tgtgagttcc	cgaacagaga	acccagctat	29940
accgtgtcca	gattcctgac	ccatggaagc	tgtgagataa	taaacaatggc	ctgggtgctg	30000
tggctggcca	ggcatgatgg	ctcatgcctg	taatcccagc	actttgggag	gccaaggcgg	30060
gcagatcacc	tgagttcagg	tgctcgagac	caacctgccc	gacatgatga	aacctgtct	30120
ctactaaaaa	tacaaaattg	gccagggtgtg	gtggtatgcg	cctgtaatcc	cagatacttg	30180
ggaggctgag	gcaggagaat	cgcttgaacc	cgggaggcgg	aggttgagc	gagctgagat	30240
tgcgccattg	cactctagcc	tgagcaacgt	gagcgaaact	ccatttcaaa	tttaaacaaa	30300
ataaacatat	attgtttaac	tgttaagtta	gtggttaact	gtcatgcagc	aggcaatgac	30360
tgatacacgta	acctatgcac	acatccatct	ccagtacgga	cacagaactt	ggatgcacgg	30420
ggtgcatgac	acctcttggc	aggacttaac	tggacagaca	agcaacaaa	acaataaagc	30480
ccaggctaag	atggactgcc	aagggcaggg	aggaacccca	gagtgtggac	aggtgcaaaa	30540
gtaggggtgt	tcaatgaaga	ggggaagcat	ggtctgcagg	gcaatgacat	gccaaccccc	30600
atccactctg	acactgtagg	ggagggggtg	aaggcaaaac	cacacttcaa	aaggctgtag	30660
ggagaatggg	gtccctgggg	gacttccaag	tggagaccaa	aaggggaagg	gagtgcggag	30720
agaaaagcag	aggagtcagg	gagttcacag	tttaccctg	aaaccaaata	aaacagaaga	30780
gacaaaatcc	tgacatcgcg	tgtggcccaa	accttactga	ggcaggaa	tcacaaatga	30840
gcaaaattata	ataattctaa	tgaccacgtt	cccgaattg	acttggaat	gctggattaa	30900
aaaaaaaaaa	cttcactcct	gatccacacc	ctggggacaa	tattatctcc	ccagtgtcct	30960
acctagccca	caactactta	tgtctcatgc	cagactgagc	cagctcccgg	gatggcaagg	31020
gagccaggag	ctgctgccag	cagggccatc	tgctcaccaa	ttcccacagt	ctgaacggca	31080
cagcttccaa	agagggacta	cgagcggcca	gcagcagcct	gcacatgcag	aaggcagggg	31140
agagcgaggg	aaatggatct	atactgctct	gctgcaatca	tctgcatgct	gggtgtgaga	31200
tgatcagttc	ttgagacact	tcccagaagg	ccttcagaaa	tactgtctga	gttacaacac	31260
tgcttcctcc	aagtctgtat	tcttatttgc	atcttatagg	aatgtagccg	ggtaaaggag	31320
gaaggctgct	tcaagtcaaa	gggcatccat	ggtgggcgcc	ctctcaggcc	tggaccagc	31380
acctgcagga	gtcggccctt	ttaattctcc	tctgccgtga	actaacactg	cacatcagca	31440
atactttgtg	aagaccgagc	acagcaacca	agcccactgt	ggacctgatt	ctaggcaagg	31500
aacttttttt	ttttttgaga	cagggctctg	ctgtcaccca	ggatggagtg	cagcggcaca	31560
atctcagctc	actgcagtct	cgacctcctg	ggttcaagtg	attctcctac	ctcaacctac	31620
ttgagtagct	gggactacag	gcgtgcgcca	ccatgcccg	ctaatttttc	tatttttttt	31680
gcggagatgg	ggtcttgcca	tggtgtgtag	gctgggtctca	aactacgggg	ctcaaagcaa	31740
tccactcacc	ttggcttccc	aaagtattgg	gattacaggc	gtgagccact	ggggctggcc	31800
tagacaagga	acatcacaca	actactccac	aaccttgaaa	ggtcacagtg	tcacccctgt	31860
tttatagggtg	gaacaattga	gaccacaga	gctgtaagaa	cacacaaagg	taaaggaaact	31920
cagccaccag	cacaggagcc	agatgccaaa	cttaggtctg	cctgactcag	aacccactg	31980
ccttccctct	acacctggct	gtttctccta	catgtctgga	atttactgca	gggtcaaagg	32040
ttcatccatt	taaactgttc	acttttatca	acttacttat	tttgagacag	agtctcgctc	32100
tggtgcccag	gctggagtgc	agtgcgcaa	actcggtca	ctgcaacctc	ccgggttcaa	32160
acaattctcc	tgctcagcc	tcccagtag	ctgggattac	aggagcgac	caccaggccg	32220
gctgattttt	gtatttttag	tagaaagggg	gtttcaccat	gttgccagg	ctggtctcga	32280
actccggagc	tccagtgatc	cgtcccgct	tggccttcca	aagtgtggtg	attagagatg	32340

tgagccaccg	tgcccagcca	tttaaaactgt	ttaaatgcta	cacaaaggca	gagaaatgag	32400
gccgtcacta	agggatttga	gagcagttag	ggatacaaca	agggcacaca	gacctgcatt	32460
gtaaggcggg	tgtggcacct	gtcacagata	gggatgccag	gggctccctg	ctttctctga	32520
agagagggaa	atcacaaata	tctggggcag	gcgcactttt	agctggtcac	gaggactaca	32580
gccaggtgaa	aaggaactgg	cctaggggaa	gtgtgtgacg	ggggagcagg	gagtagtccc	32640
aatggactgg	aaaaggcaca	tgcgagaggg	gaggggtggaa	agggcaccaa	cgccgggtgac	32700
gctggagctc	agaaaagact	ccgaggacca	gaaggaagaa	gcacaaaggg	accagggggt	32760
gatatgccaa	ggtagagagg	atgggtctga	ggtgttcctc	tgtgacggga	cagaagagat	32820
gtgaggacct	gaagaggcgc	caccagggaa	cttgagaaaa	agaaggcagg	tggctcggaa	32880
gacccgatcc	atgtgaccct	gcaattttatt	ggatatggat	gaatcagagc	tgactttttc	32940
gatgacccaa	aagatgaaac	tattaattaa	ggctagacgg	aggggagaaa	agagagagaa	33000
ataccctgct	acctccagtt	tttctcccta	cagcacctcc	tagagatgag	gtgatggcag	33060
ctcaccttca	ggatccattg	gaaacaaaga	gaaatctctc	cctaattctc	ctgaccccaa	33120
acagaaagca	accaactatt	ccataatttt	cttctctagg	agagattcag	agagaagagg	33180
cctgaaaatg	caaattaaca	cacggatgtg	aatgagtttc	agtcaaagct	taaaccagga	33240
cacaactttt	cttgtgtagc	gaggaaggct	aggaggcagg	atggtgtcct	gtgcctaaaa	33300
aggtgagtgt	gacgtcaagg	tggctgaaga	ggggttttaa	agggcaccta	gggggacaa	33360
cccagtgccc	agcatcccac	cctaaatgag	aacacaggtc	tccaactcca	gcccagggtc	33420
tccctgtggg	cacaatgggt	ttggggacct	gctgacagtc	gcacggaagg	actctcgggt	33480
gtggtcagaa	tgaccaacaa	tcccaggagg	cacgtgccac	cagttggcat	gagtccttgg	33540
tgctggccct	ggcgctgctg	ctaattccacc	ccagtggact	taggcatgct	ccctcacctg	33600
tatgtccaag	acagctaatt	caacagtact	actaacctgg	tcccagaaa	ggcggcagag	33660
taggatcaac	ttggcattag	aggtctcgct	tcaaatacag	gatttcccaa	ttccaatctt	33720
gggcctcagc	caccaaccgg	gaaaaccccc	ctccaagggc	tgttttgaga	atatggaaa	33780
ctgccaaagc	tttattttccc	atccctttgt	aaggtccctt	gcagctccca	actagaagag	33840
aaagggacct	tttatttagca	agaaaagggc	caggtgcagt	gactgtcatc	acgcctgtaa	33900
tcccagcact	ttggggaggct	gaggtgggag	gatcacttaa	gccaagagt	ttgagaccag	33960
cctcaacaac	acagtgtgat	ctcatctctt	caaaacacat	ttaaaaaaat	tagccaggtg	34020
tggtggcgca	cgctgtagt	cccagctact	tgggaggctg	aggtgggaag	atagcttggc	34080
cccaggagtt	caaggctgca	gtgagctatg	atcacaccat	tgactccgg	cctggacaac	34140
aaagaccctg	tctctaaaaa	ataaaaaata	aaagttttta	ttttaaaaaa	ggttaaatac	34200
taatgagaaa	ggtccacata	caaatttttca	tgtcactgct	atttataagt	tcttattaac	34260
tgaaaacact	atgctataag	ctatttaaagg	taactaaaaa	ataaaaaatag	ctctttgccc	34320
caaagacagt	ctaagtgaca	gagctcagta	ctcactcatc	aacaacagcc	ctgagagaga	34380
caagagatgg	aagagattcc	aaccccaga	aagggctgga	ggggagccag	gggaggaggc	34440
atggggggag	ggacctccaa	gagggagcta	ggggagggca	gatggggagg	agagcccaga	34500
agggagctgg	gggaggggaa	catgggggag	gggagcatgg	gggaggggag	ctccaagggg	34560
gagcatgggg	gagggggagat	ggggagggga	gctctaggag	ggagctgggg	gaggggagat	34620
ggggagggga	gctccaggaa	ggagctgggg	gaggggagca	ttggggaggg	gagatggggg	34680
aggggagctc	caagagttag	ctggggggagg	ggagatgggg	aggggagctc	caggagggag	34740
ctggggggagg	ggagatgggg	aggggagctc	caggagggag	ctgggggaaa	agggcttggg	34800
gagggtagct	caatgacggg	acgtgggggat	ggggagctaa	ggaggagatc	tggggagagg	34860
gagcttgggg	aggggagatg	ggggagggga	gatgggggag	gaaagctggg	gaaggggatc	34920
tgggagaggg	aagcttgggg	agaggatcta	gggaggggag	ctgggggtgg	ggagatgggg	34980
aggagaaaatg	ggagaggagc	ttggggaggg	ggatctaggg	aggaaatctg	gggcaaggga	35040
gctgaggggag	gggaacttgg	ggaggggatt	tagggagggg	agctggggga	aggggagctc	35100
agagaggggga	cttcaggggag	gggagatggg	agaggggatt	ggggagggat	ggttagggat	35160
gggatctagg	gaggggattt	gggggagggga	agcttgggga	ggggatctgg	gggaggggag	35220
tggggaagag	agatggggag	tcgggggagg	ggaacctgga	gggagggatc	tggggaaggg	35280
gattttgggg	aggagaacag	gtggagagag	gagctggtgg	ggagggcagt	tgggggcagg	35340
catctggggg	agattttggg	ggaaggggag	ctgggcgccc	acaggagccg	ctgtgaggtg	35400
ggcaagcccc	tctttcagtt	cctcctcgac	agtcagctc	cagacttcca	ctccaccctt	35460
ccctgcttcc	acctagacag	tctgatctgc	aactcgccc	atgactgccc	ccattgggaa	35520
tccagctgct	tctagcctgg	gaacctgtac	gtgggcccctg	acctgaccaa	tcaaaaaacc	35580
caggggtgatg	gagcaaatgt	gtcctgtatc	ttgagcataa	cattaaaaagt	gaggaccag	35640
cagaagtccc	ccagcgagga	cccagaaata	aggaatctct	ttgattcttg	caggctagtg	35700
tttccctacc	cacataatct	ttagaaatca	tgtgtgccgt	aataaaaagt	agtatttccc	35760
ctcccttcac	tcaagcacac	agaaacatcg	gagaaaaagct	gagcatattt	ctaccagttc	35820
tgcatatgag	tttgaccaga	acaccctgct	gtcggtaaat	aatggttgac	cccaatttct	35880
gaacacatat	ttccttttcc	aattaatttt	ccttcccctc	atgagataaa	acagactatt	35940
tttttttaaa	gaacaatatt	cctgaaaatt	tatttacttt	ttttaaaact	atgaggtcag	36000

agtttaagac	tggctccttg	gtatgaagga	atacatgata	ttaatataac	aaagggtcga	36060
atctttccata	aatcaacaaa	acacccaaac	aaaggcgagaa	cttaattttt	ggcaaagaaa	36120
aaacaaaaat	gttttttggtg	tccattagtg	aatacatcag	ctgaggactg	ccatcttgga	36180
atctttttaa	tgagcagagc	taaagatttc	tcataagcac	aattaaagca	ccctgaattg	36240
ataccttttag	ggggttgagt	atctgtttca	aatcagcaaa	gtgcttaccg	caaaaggaac	36300
accttaccaa	aagcaagatg	aaaaagttag	ggcagagtgt	catgattatt	actttttttt	36360
ttaagcagaa	gaatagtctg	caagaaaata	cataaaaatg	ctcaagttag	gccgggcaca	36420
gtagctcatg	actgtaatcc	cagtacttca	ggaggccaag	caggaagatg	tcttgaggcc	36480
aggagtccaa	gaccagcctg	ggcaacacag	caagaccttg	tctctattag	aaaataataa	36540
gttaccaaaa	aatgctcaaa	atggtaatgt	aagggtggta	gaataatggc	aattattttc	36600
cttctttttcc	aagcgttata	taatattatt	aaagtggcta	gacatatgta	tggattttta	36660
agcacttcag	ttttatgtgt	tttaggtata	atttctaaag	cactaaaaaa	ttggcatatt	36720
cttttttttt	tttttttaag	acggagtttt	gctctgtcgc	caggctggag	tgcaagtggcg	36780
caatcttggc	tcactctgcc	tcccgggttc	aagcgattcc	cccgcctcag	cccccaagt	36840
agctgggact	acaagcacac	actaccacgc	ccggttaatt	ttttctgttt	ttttagtaga	36900
gacagggttc	caccatgttg	gccaggatgg	tctcgatctc	ctgacctgtg	gatccacccg	36960
cctcagcctc	ccaaagtgtc	ggaattacag	gcgtgaccca	ccgcacccag	cccaaaattg	37020
gcatattctt	tttgaacgtt	ttcccttttg	gagaggaaca	agagcattcc	ttacctgctt	37080
gggagaaaaga	cttaggaaca	agaattgaaa	gtctgcttac	ctgaggttta	attttcgatc	37140
ttcttcgctg	ccagcctcca	actacagaga	aagaaagaga	atatcacacc	acaggacca	37200
ctgtcaacac	gcctcgggcg	gcagtctcac	attttctacc	ccggtactgg	aaaaagataa	37260
agatatccag	gaaacctagc	tacttctaaa	cagccgtgcc	ctttcctcac	caatcccgtt	37320
ctgtcccttg	gagtcatttc	cgtgggggaa	ttttcaggtt	tccaaatgtt	gacccacatt	37380
cctgccgcag	tccaggggat	ggagtctgtt	tagctcaaca	tttctatctt	ggtgttggtta	37440
cccagcacgg	tcttttagcc	ctcagccctc	aactttccga	ggttgttctg	gaccttatcc	37500
tgtttttctc	ttttaagggg	agggggtcat	gtttaaagag	aatccacttc	ctccgcagag	37560
ccaggcaata	acagctgagt	gatgaacacc	attttcaaaa	aaccaacca	ggcaagactt	37620
gcacagtgga	aggtggccag	gaatcaggcc	gtctgtttgt	gggtcttgaa	agctcttgat	37680
ggttctcgaa	aagacttaaa	catttgatac	gaaacatcct	aggctatcgg	tttatttata	37740
taaatgcaag	aaagagatat	ttaatatatt	ctgaaatcta	aaaggccacg	agtttgggct	37800
ccagaagtac	ctatgactta	tttttatatt	ttttctttca	gagagcaaac	tgaaaataag	37860
aaggaaaacac	atacacaccc	cccaaacaac	tccgcaccgc	tgggacttgg	catgtttttt	37920
atgttgacac	gaggcgccca	ttgaatggga	aagagaaacc	tggaaagctg	tgatggctgg	37980
gagagatgca	gggctgatcg	aggacagaaa	tgaggcagga	gccaagggcg	aaggaaaaag	38040
ggtccagaga	taatgtaggg	aggggcctgg	gcagcaaggg	acaccaccca	ggaggtggca	38100
acttcaacca	agaatgagta	caccagcccg	gcgcagtggc	tcacacctgg	aatcccagca	38160
ctgcaggagg	ccgaggtggg	cggatcacct	gaggtcagga	gttcgagacc	agcctagcca	38220
acaaggtgac	acgctgtctc	tactaaaaat	acaaaaatta	gccaggcacg	gtgacatgta	38280
cctgtaatcc	cagctacctg	ggaggctgag	gcaggagaat	cacttgaacc	caggaggcgg	38340
aggttgacgt	gagccgagat	tgcaccactg	cactccagcc	tggtgaaaga	gcaatacttc	38400
gtttcaaaaa	aaaaaaaaaa	gcgtacacca	gagggcctgg	gagtcacctac	atcatattaa	38460
gatgaagtac	atacaagatt	ctgcagaggc	acctacccca	cactgaagga	gaagtggaaa	38520
gagcagggaa	ggcttcaatg	accacaaaac	aagtcacaag	agcaacaaat	tgacaaagag	38580
tatgttgggg	tctaaccagg	tggttctcaa	tgcgagagca	gttctctcaa	caggagacat	38640
tggaaagatgt	ctggaggcat	ttttctggag	gtcactactg	gcacctagtg	ggtagaggac	38700
aggatcccac	aacacacagg	atggtcccc	cacaagagag	aatgttctgg	tcccaagtat	38760
caacagtggg	ttgagaaact	ctggtccaat	ccaaaaaagt	gcctggagat	ctgcccataga	38820
aaattagtc	ctttgaatgt	ttctcagaaa	taacaatgtt	atgatccatt	cctgaaaatt	38880
attgatttat	ctattcttgt	gctctgcctg	ttcacacaaa	ggaactgaga	taagattcac	38940
aggaaaatga	cataagttag	ataatcaaa	actgggaaaa	aaagaaaatt	aaaagagaaa	39000
agagagcctt	aggactgagg	gctgtgatcc	cccgtttccc	acgccggcag	caggcctggc	39060
tgtgtcagga	aagcactggc	ctaagtgtct	gactcattat	gaagttgcaa	tttgaagagt	39120
gatgacgtga	cttgggggta	cggacttcac	aatcatttaa	ctctcggtca	ctctctgagg	39180
ttctcagatg	aaaggccatc	tcaggtcagt	tattccaggg	aaactacatc	tgccaaggaa	39240
cacatgaaa	aggttaattca	gtccttttag	atgagccagg	gccacacac	aggaagcaac	39300
tcaagcgagg	gcggaaccag	gcagaaccgg	cctggcctag	gtctcctgac	cccatacaca	39360
cttgcgtgtc	ccatcccacc	ttgcttctca	cctcaacaca	tctgaacgag	ggccttgctt	39420
tcgggaaaca	tcccagcgca	ttcaaagcca	agcaatgaat	gctgcagctt	tgctatgatc	39480
aaataaaaagc	tgctgtagtt	ttactttatg	tttatcaggg	tcattggcac	ttggtaaaaa	39540
taatgcttta	tataataatt	gaaaaatgat	ctagtagaca	caacacaaat	gtccaacaaa	39600
atgtggtaca	tccatacaat	ggagtattat	acagccatga	aaaggaaatga	agtactgcca	39660

catgctacaa	tatgcatgaa	ctttgaaaaac	gtgatgctga	gtcaaagaag	ccagacacaa	39720
aaggccacac	aggggtgtgat	tccattttata	taaaatgtcc	agaataagca	aatccataga	39780
tacagaaagt	agattagtgg	ttgcctaagg	ttagggagaa	tggggcaggg	ggaggctgca	39840
ggtgagggct	aacgggtaca	gggttcctat	tttggggagg	atgaaaatgt	tctggaatta	39900
gatggtggtg	gttgcacaat	ctcatgtata	tactaaaaac	cactgaattg	tacactttaa	39960
aatggcaa	ttatggtatg	taactaaaaa	ataataagac	cttaaaatgc	gtaagacaag	40020
aacagattag	gttgcaagta	actctaggaa	catggggttt	tgaatcagaa	atctgggctg	40080
acaggttcag	cctgaagcca	acctctcccc	ttacctcaca	taaacttttg	tgatgagaac	40140
ctgagattaa	gtgcataaat	tgccagacag	cagtgaaccg	aacagaaaac	agccccctct	40200
cgctgtggga	aggaaaggcg	ctcctgcaac	ctaacttctc	agtagcaggc	tattgatcgc	40260
cagtgttctt	ttgcctctaa	tcagcgtgta	gagggggatt	actagaacct	tctgtgtata	40320
gataactcat	gaatggcctc	tcctctccaa	ggagggggct	gtgaaggttc	aacttcccag	40380
ccactctgaa	aatgtcccctg	ccaatcccag	caaaacaagg	ctgaagaact	accctaccag	40440
gagacagggc	tgtcaagcca	aatgcaaaca	ttattctctt	gtctctctca	gacacacaaa	40500
cctccccctg	gttatcagtc	aacttcccc	actccctccc	acaaagaaaag	gggctgaaga	40560
gccagatgc	tggctgcgga	acttccctggg	cctgggaccg	cagggccgct	cctccagtct	40620
tctctaaaca	cagctaaggg	tctgcaggcg	gacactcagc	cttggttatag	gtaagagttt	40680
agaccagtg	ccttgacggg	ttcttcaaga	gatggtgggc	aagattgcgc	gaccagaggg	40740
tcattccctg	agctacagag	ggctgacctg	ctcagaggcc	caaggcccca	gcctaggaca	40800
agccagggca	accctgcag	ctaagaggcg	aacagtggcc	tcaatcaacc	ccagaggaaa	40860
aagtggccag	gcaaacggac	ctgggcaca	cacagacca	caaaaacgcg	cacagtcca	40920
ggacacgcaa	cccaggaatg	cacctatgca	atcaccaga	atgggtcaca	gccacacaga	40980
aagatagatg	cacataaaca	cacaggcctg	agtgatgtta	cagaaaggaa	aagccagact	41040
aaggctgcac	gcacagacgt	gaaacacagc	cacacagagc	ccacagcacg	ctcggtcacc	41100
gtcacacagt	gacacgggca	cgctacaga	cagaactcca	gaggcggcag	gcggggaaac	41160
aatctcacac	gtttgtaggg	gcactcccag	atgcctgtct	cacgctggca	cagtccccgg	41220
tacggcaggt	cagcaacagt	cacatctcac	atcgcacagc	caggcataca	ggcaaagggc	41280
ctagaactac	cccggccaca	ggtctcagaa	ccagcggtc	acgcagtcac	ccaatcaagg	41340
gtcccagttg	cacatccagt	cacccctgga	ccctggtcac	actgcagagt	cactcacaaa	41400
tgggagttcc	gacagacgca	cagtcctccc	cagacagagg	tcaacccaag	atgggggtca	41460
cacctgaaat	cacagtcccc	acacaatcac	gaggtcacat	ctgcacacac	agtctctgca	41520
cagtcaccc	taggggtcac	aacgcacaca	gtctccgcct	aacaggggtc	acccaagat	41580
gggggtaaac	ccctgatgtg	ggtcacagcg	cacacacagt	atcctgcaga	cactcccaga	41640
ggggctgcac	tgcatacaca	gtccctgcaa	agtcgcccc	cgataggggt	cacaccgcac	41700
acaaagtccc	cgcacagctc	cccaagacag	ggtcacatcg	cacacagtc	tcgcacggtc	41760
accccgttcc	ggctgcccgg	ctctgttcc	atcgccagggc	cccagaggag	ccgcagggc	41820
gccccctgc	cccgcacgcg	cgccccagc	tccggcgggc	tccggcgggc	gtccggcggc	41880
ccaggccggg	cgcggcgagc	ccggggctca	cctcgctgtt	gctggccgag	gaggaggcgg	41940
cgctgggctg	gggcgagcgc	tgagggtca	ccagggccat	ggctgcggcg	cggtgcgagg	42000
gcgccacaga	cgtctcgagc	tagagccgcc	accgccaccg	ccgcccgggc	cgggcccggg	42060
gectcctgga	gcccgcgcgc	ggcgccggg	ccgagccggg	ccgggcccgc	ccctccccct	42120
cgcgctcgcc	accgcccccg	ccccagctc	ccgctccc	cgccggcgcg	cgcaggcctc	42180
agtgcgcgga	gtgggcgggg	aagcgggcag	ggcgggacga	ggaggcgcg	gtgcgcgggg	42240
gccctgaggg	ctgcccagag	cctcggtg	tcgatcacgt	ccctcgcgcg	cccacacac	42300
gcgccccgc	ccgcgcgcgc	cgctatcagg	cctgggactc	gggggcgcgc	gcgcccgcgc	42360
gagcccgtac	gccccagggg	ccctgcccgc	tgctctgcct	ggggaaactg	aggcccggcg	42420
accgtgcaga	caggactgta	cagcgaccag	gaaataaaaag	acgtcctggg	gccgggcgcg	42480
gtggctcacg	cctgtaatcc	cagcactttg	ggaggctgag	gcgggcggat	tacgaggtca	42540
agagatcgag	accatcctgg	ccaacatggt	gaaaccccg	ctctactaaa	aagacaaaaa	42600
ttagctgggc	gcagtgggtg	gcgctgtag	tcccagctac	tcgggaggct	gaggcaagag	42660
aatcggttga	atctgggtgc	cggaggttgc	aatgagctga	gatcgcgcca	ctgcaactca	42720
gcctgggcga	cagagcgaga	ctcgttctca	aaaaaacaaa	aaacaaaaaa	caaaaacagt	42780
aagcaaaata	gattcgccctg	attttgcaga	ggttaatcaa	gttattagc	acgtttttta	42840
aaaagtattt	tgctaattctt	tttcaatgaa	ttctttctgg	gtgttctgaa	acccagccaa	42900
ctccttggag	gtcagggaa	gcttcccaga	agagctttat	tctgaggctt	gggcttgagc	42960
ataagcagga	ttaacaggtg	aaagaacaga	gagacagctc	tccaagcagg	ggggatcagc	43020
gtgccctgaa	gcaggaagaa	gtttgtcaac	cggaggccag	cactcaggga	aggggaagagg	43080
ggaggaatgg	ctggagtctc	catcctctct	ggaaagatcg	ctccggtgc	tgctgtgatg	43140
agggaccacg	gggcagaggg	ctgagggaga	ccagggagga	ggctgctgct	gttgtcccgg	43200
ggagaggtga	ccagttatgg	ggatggagag	gggaacatgg	aataagatac	caagaaggca	43260
attctggctt	gacttagtag	taggaaactt	ttcttttagc	caaaatctca	tctcccggct	43320

cccacccccca	acctctgcat	gttgacaaag	cactcgcaaa	cgcagtggtc	ccagcctgcc	43380
ccgcagctta	gcaaatttgt	cttactgccc	aacaggaaac	ccacgcagcc	tcctggatcc	43440
ttccccgtcc	ctccctctgt	cctggggctg	tgacctcctc	catgttattc	acagggctctc	43500
agcacgattc	atctcaaagg	tgattctagt	ggggggcact	gtagcttcta	cggagcgttt	43560
ctaagagggg	atgtgtggga	atgtttgtgg	ttgtcttgct	gatggagggg	gagagctcct	43620
ggcattttaga	gtgcaagagc	cttgatgct	aaatgtcttc	caatgcactg	gacagtctcc	43680
ccaacaagaa	ttgctccatt	cccacaaaat	gtttcctggg	tgaaaaaccc	atztatagta	43740
atgtgaagcc	agaacctaac	tccatttcat	gcatcaacac	tagtcttcct	tccttccttc	43800
cttccttcct	tccttccttc	ctgecttcct	tccttccttc	ctctctttct	ctcacttttt	43860
ttctgaaaca	gggtctcact	cccgtcacc	aggctgaagt	gcaatgtcac	aatcatagct	43920
cactgcagcc	tccatctccc	aggctcaa	catcctcctg	cttcagtctc	ctgagtacaa	43980
cgggtacaca	ccaccacacc	cagctccttt	aaaaaaaaag	tttaactatg	ttgccaggc	44040
aatcctcctg	cttcgcgcctt	ccaaagtgtc	gggattacag	acagaagcca	ccatggctag	44100
cctggtat	tttactgaat	tttcagaaa	gtgactatgt	tgaaaccctg	tctctcctaa	44160
aaatacaaaa	aattagccag	gcatggtggc	gggcacctat	aatctcagct	actcaggagg	44220
ctgaggcagg	agaatcactt	gaaccgggga	ggcagagggt	gcagcaatct	gagatcgtgc	44280
cactgcactc	cagcctgtgt	gacacagcaa	gacagagaga	aagagagaag	ggaagggagg	44340
ggaggggagg	ggagaggagg	ggagaggagg	ggagaggagg	ggaggggaga	ggaggggagg	44400
ggaggggagg	ggaggggagg	ggaggggagg	ggaggggagg	ggaggggagg	ggaggggagg	44460
gaagggaaaa	tacactttgt	tttgcttgag	agttttgtca	agagttgttc	atccatcctt	44520
agggaaaaag	aggtaatgga	tggcaacgcc	tctgctaata	ttagagcatc	ccacacaagg	44580
tgccccacaac	tgtagctgca	ctctaggtag	acagacagtc	ataggtactt	aaatgtcaaa	44640
tataagggaa	aattgtggac	aaaattcagt	tgagtagaga	atattttatt	tctcaaattcc	44700
aagcacattg	attattggca	ggcccatgct	tctgagatgc	ccctgtgtcc	tctmagggag	44760
tagtggctga	gcatttccac	attgtaatgc	atgttgtttc	attatgattt	atttttcttt	44820
tatgtctctc	ttacattagt	tttcaaattt	gagagtgtga	gaatcccctg	gagaaaatac	44880
agattgctag	acccacctc	ccagagtttc	gaattcacaa	ggtttgctgt	agggctggaa	44940
aatttgcacg	tctaacaaat	tcacaggcaa	tgctgatgct	tctgtctggg	gacgacagtc	45000
tgagaactac	tgctataaca	aatgcaatgg	cctcttcacc	aagaaattcc	tacctagatc	45060
tgatcctggt	acccgtccgt	ggcccccaat	cctaattcccc	ctgctctggc	cggcctgctt	45120
ttccactcac	cccaactttt	ttggaggcag	tctccacccc	ttctcacttc	ctcttagagc	45180
tgagagccct	tttcttcccc	acaactaact	cttgctagaa	atcacctcca	aaaagctttc	45240
cctgcccctt	aagcagtgtc	atttccagga	tctcgtagcc	ctcaccttac	ccttaaacac	45300
acagaaagct	tcagctgtcc	ttatcataat	gggtccatct	ctctgtcttg	tcccattacc	45360
gtagagccag	gaagcgtccc	taagaaaag	ctcaatgatt	aggctgggag	cagcgtgagg	45420
gtgcaaaatg	taaggggtg	cccccaaaa	ctcaatgatt	aagataaata	gtattttaat	45480
gcaatatttt	agaaaatcaa	aattaatgcc	aaatccatga	tgaataaaat	atttttaaaa	45540
tttgcttttt	tttttttttt	ttaattgaga	cagagtcttg	ctctgttgcc	caggctggag	45600
tgagtggtgg	cacaatctct	gcctcttggt	ttcaagcagt	tctcctgcct	cagcctcccg	45660
agtagctggg	attacagacc	cccaccacca	tgaccgggcta	atttttgtat	ttttagtaga	45720
gatgggggtt	caccatgttg	gccaggctgg	tctcaaattc	ctgaaatcag	tgatctgcct	45780
gcctcggcct	cccaaaatgc	tgggattaca	gggtgagacc	actgcacctg	gtcaaaaatat	45840
ttacaaaaat	tttttaagag	ccaaggtctc	attctgtcac	ccaggactgg	gtgtagtggg	45900
gcaatcctag	ctcacttcag	ccttgaactc	tgggtcgaag	ccatcctcct	gcctctgcct	45960
ccggagtacc	tgagactaca	ggtgtacacc	accacgcctg	gctgacttta	tttttgccag	46020
aaactgggtg	ttgctatggt	gcccaggctg	gtttcaaact	cctggaggca	ctcaatcccc	46080
cgaccttggt	ctcccaaaag	tttgggatta	ccggcatgag	ccaccacacc	tggccaaaag	46140
atcaaaatttt	taagtaaaat	tggcatcagt	attgtgtcac	tgattcttcc	acttacttca	46200
gacttcagtg	tagctcagca	aagcactttt	attgatcctg	tctttatttg	attctttttac	46260
aactttggcc	attctaaagc	cttttgtaga	aatggcctgt	ggttcagctg	ggcatgggtg	46320
cgtgcacctg	taatcccagc	tactcgggag	gctgtggcag	gagaatcgcc	tgaaaccagg	46380
aggtggaggc	tgagtgggc	tgagatcgtg	ccacttttga	cactctgtct	caaaaaaaaa	46440
aaaaaaaaaaa	aaaaaggaag	cctgtcggct	tgactccagt	agcctctgat	gggtggaggt	46500
ggacaagggg	aagtgaagc	tcccaggcct	cagtcagggc	aggtcccaag	aagccctgag	46560
catggaggag	gggaacaatc	cagtagaggc	agctctgaag	ttttctccca	tgcattagag	46620
ccctttccaa	tcagtatcat	gatttttcat	catataatag	tttatttaat	catctttgac	46680
ctcctccttg	tagtcccagc	tcacttttgt	aactaataaa	aaacagttag	ttattgagct	46740
atgtgctctc	tgctaaggca	caatgcaaag	tgctttgtga	gtgtgtgggg	gacatgattt	46800
attaacatgt	gactgtcccc	ccacttatac	tccaagatca	cctcctccag	gaagccttcc	46860
ttgccccgtg	gctgggttag	gcacccttcc	tctgtgctcc	tacagcccct	gtgcattagt	46920
						46980

gacaatggca	ttgtggatct	gccctaggcc	catttctggg	ttgggacact	ttaggtacat	47040
tcattcttgt	caccctgtga	ttctcatttc	atgggtgagg	aaattgatgc	acagagtggg	47100
taaggcactg	gccccaaagt	atgtaactaa	ggagtgggtga	acctggttca	cccatgtttt	47160
tctgctttag	aactcaggca	aagacaggtt	cttccaggac	agcctcagaa	agtgttgggtg	47220
caaattaggt	tgggtgaaaa	gtaattgctg	tttttgtcat	tttttttttt	tttaatgggtg	47280
caaaagtaat	tgcggttttg	tcattaatga	ccaactatta	taagtaatag	ttcccttttt	47340
tttttttttt	gagatggaat	cttgctctgt	tgcccaggct	ggagtgcagt	ggcttgatct	47400
tggctctctg	caaactccgc	ttcctgggtt	caagtgatcc	tcctgcctca	gcctcccaag	47460
tagctgggat	tacaggtgcc	cacccccatg	cccagctaata	ttttgtattt	ttagtagaaa	47520
cgggggttca	ccatgttggc	caggctggtc	ccgaactcct	gacctcaagt	gatccaccca	47580
cctcggcctc	cccaaagtgc	tgggattaca	ggtgtgagcc	actgcacctg	gccagtagtt	47640
tgcctgttaa	agcaaataac	ttgtaatttc	tccttaatta	ttcattccaa	aatgatattc	47700
agaggtaata	aagctctgat	aggctgaata	atggcctgca	aagatgtcca	tattccaaat	47760
ccctagaatc	cctgcctatg	ttaccttgca	tgctaagagg	gttttacaga	tgtgattaaa	47820
ctcaggatgt	ttagatgggg	aaattttcct	ggaggaggcc	caagagggtcc	taatgtaatc	47880
acaaggggtcc	ttataagagg	gaggtgagaa	ggtcagagtc	agtagtaaga	gatgtgacaa	47940
cggaaactgag	ggattagagt	gaaggaagag	gccacaatcc	aaggaatgca	ggcagttgct	48000
aaaagtggaa	aaacaccaa	aaatgaattc	tcctttcaga	gcctccagaa	agaatggagc	48060
cctgctgata	tccttttctt	ttcttttttg	agcttagggc	ttgctcacag	agctgtcacc	48120
caggctggag	tgcagtggca	tcctcatagc	tcacaggagc	ctcgacctcc	agggtctcaag	48180
ggattctccc	acctcagcct	cctgagtagc	tgcgactaca	gacacacacc	actatgcccg	48240
gttgactttt	tttaattatt	attatacttt	aagtctctggg	gtacatgtgc	agaatgtgca	48300
ggcttggttac	ataggtatac	acgtgccatg	gtggtttgct	gcacccatca	acccgtcatc	48360
tgcattagat	atttctccta	atgttatccc	tcctctggcc	caccaccccc	tgactggccc	48420
cgggtgtgtga	tgttccccat	gcccgggtga	tttttaaggg	ttttgtttgt	ttgtttgttt	48480
ttttagagac	gagggtctca	gctgggtgca	gtggctcatg	cctgtaattc	cagcactttg	48540
ggaggtaagg	cgggcagatt	gcttcagccc	aggagttcaa	gaccagcctg	ggcaacatgg	48600
cgaaaccaaaa	aaatgcaaaa	aattaactgg	gcatggtggc	acatgcctga	ggctgaggtg	48660
ggagtatcgt	ctgagcctgg	gagatcaagg	ctgcagttag	ccatgatcat	gccactgtgc	48720
tccagcctgg	ttgatggggt	gagaccctgt	gtctaaaaaa	taaaagaaat	gaaggtcttg	48780
ctgtgtttcc	taggctgttc	ttgaactcct	aggctcaagc	aatcctcctg	cctcagccac	48840
cccagttgct	tggattacag	gcacaagcca	ccatgtccaa	tcctggcaac	gtcttgattt	48900
tagactttctg	atctctacaa	ttgcaagaga	ataaatttat	gttgttttta	gccacgaaat	48960
ctctgggaat	ttgttacagc	agccatacga	aatgaatata	aaactcaacc	tccatttggg	49020
ctttaaaaaa	catatcatta	taatgccatt	cccagataata	ttccaggtgc	ttcccaagcg	49080
ttgtgtcatt	ttctcattca	ctcaactcat	ccaataaata	atgtttgttg	ctctctaggg	49140
cactagtcta	ggaatctggg	ttccatcagt	gaacaaaatg	gaatcactgc	ccttgaagag	49200
cattcaatca	agtgggaaat	atagtaaaaa	tatatatata	tgcaaatatg	tttaaaatca	49260
tatgtggtaa	atatattgca	tttaaatgaa	ttaataggcc	gggcacgggtg	gctcatgcct	49320
gtaatcccag	cactttggga	ggccgaggcc	agtggatcac	ttgaggccag	gagttcgaga	49380
ccagcctggc	caacatggcg	aaaccccgtc	tctactaaaa	gtacaaaaat	tagccagttg	49440
tgggtggtggg	tgctgttaat	cccaggtact	cgggaggctg	aggcacaaaa	atcgcttgaa	49500
ctgaggggggt	gcggaggttg	cagttagccg	agatcatgcc	actgcactcc	agcctgggtg	49560
acagagttag	actgtctcaa	aataataata	ataataatta	attaaatgaa	ttaatattgg	49620
taagggtcct	tagaacaaga	taggcactga	tatgtgtcaa	ataaatgaaa	tatgatgtcc	49680
aatcatgaaa	aagcttggga	gaaaaacaaa	gcaggctaag	ggcagagtaa	tggaggaggc	49740
cacttagaca	aatggtcagg	gaagcttctg	ggtgaggtga	tatttgagca	gaggaatcac	49800
catgacagca	ccaccaggga	ggtgtagaaa	ccctgggatc	tgcttggttc	attcaaaactg	49860
gcctccccac	taaggaaact	tgaggtactt	tttctgagac	ccattttctt	tctgtctgtg	49920
tcacccaggc	tggagcgag	tggcgcgatc	tcggctcact	gcaacctcct	ccccccaggc	49980
tcaagtgatc	ctcccacctc	agcctcctga	gtagctagga	ttacaggtgt	gtgccaccat	50040
accagctaa	tttttgtatt	tttagtagag	tcggcgtttc	accatgtttg	ccaggccagg	50100
ctgccacctt	ggcttcctac	agtgtctggga	ttacaggtgt	gagccttcag	accagccga	50160
gacccactgt	ctttctctgt	aaaattgata	tgaaagtgat	agtgtctggc	cgggcatagt	50220
ggctcacgcc	tgtaatccca	gcactttggg	aggccaagggt	gggcagataa	cctgaggtca	50280
ggagttcaag	accagcctgt	ccaagacgggt	gaaacctgt	ctctactgaa	aatacaaaaa	50340
ttagccaggt	gtggtggtgg	gtgcctataa	tctcagctac	tcaggagggt	gaggcaggag	50400
aatcgcttga	accaggaag	cagaggttac	agttagtcga	ggtcccgcga	cttcaactca	50460
gctggacaa	caaagcaaga	ctccatctca	aaaaaaaaaa	aaaaaaagaa	agaaaaagaa	50520
agaaaagtgg	agtgtctgacc	tcagagcttg	gttgtgtcaa	ttgaacagca	tactatgcag	50580
gaaaggcaga	gcgtgtctgc	ctatttacta	atagtaccta	aggtattggg	ttgaattgtg	50640

tccccacaaa	attcactagt	ccctgtgaat	gggaccttat	ttggaaatga	ggtcttttgca	50700
gctgatcaag	ttaagatgag	gtcatttaggg	cggggcccta	ttcgcatatg	actgtgtccg	50760
tatgaaaagg	gggaaatttg	ctggggcgcg	tagctcatgc	ctataatccc	agcacttttg	50820
gagaccaagg	cgggtggatc	acctgaggtc	aggagttcga	gaccagcctg	acaaacatgg	50880
agaaaccctg	tctctattac	aaatacaaaa	ttagccaggc	gtgggtgggtg	atgtctgtaa	50940
tcccagctac	ttcggaggct	gaggcaggag	aatcacttga	acccgggggg	tggaggttgc	51000
agtgaactga	gattgcgcca	ttgcactcca	gcctgggcaa	caagagcgaa	actgcatctc	51060
aaaaaataaa	caaacaaaca	aataaataaa	taaataataa	aaggggaaat	ttggaccag	51120
agccaagggg	aaaatgcttc	ctgaagggtt	tagttgtgct	gccacaagcc	aaagagcacc	51180
cgagatggtc	agcaaaccac	cagagctagg	agtgagaagt	gaggagcaga	tttgcgtggc	51240
cttctgaaga	aaccagcaac	tcgatttcag	agttccaggc	tccagaactg	agagagtaaa	51300
tgcctgtggt	ttaagcctcc	cagtttgtgg	cactttgtta	cagcagccac	aggaaaggaa	51360
cgcactaacc	atgatcattt	catcagctgc	agaaaatgag	gctcagagca	aggctagggt	51420
ttgaaccag	gccaaactaga	ccccagacca	catacatggg	ttgttggcct	ttctagcctg	51480
ggaggtgaca	gtttggatgt	tccactattt	gcagggaacg	gtgttcagag	gactcaaagc	51540
ttctgccacc	tgggccaggg	tgtccagtg	tagatatgga	agtcagggtat	ctggggcttc	51600
tgacaaagca	tctctctggg	tgggtcaata	ggcaccagca	gccaggcagt	tgagggatct	51660
ctgcccctgt	gcggagttgt	ctgaagcctc	ccttccctga	cccattccctt	catttaacct	51720
gctggccaga	tcaagtgatt	cctggcctc	tgccaggggt	gtattcacct	gaatttgctt	51780
ttttattcact	atgactcaca	gcaacacact	gacccttgct	gggcctcaga	atctcaactc	51840
ttggtggggt	gcggtggccc	atgctggtaa	tccagcact	gggaggccaa	ggtgggtgaa	51900
tcacttgagg	ccacgagtta	gagaccagcc	tggccaacat	ggcaaaaacc	tgtctctact	51960
aaaaatacaa	aaattagcca	ggcatgggtg	catgcacctg	tagtcccagc	tactcaggag	52020
gctgatgcac	agaatcact	tgaatccagg	agatggaggt	tgcaatgagc	caagatcaca	52080
ccactgcact	ccagcctggg	tgacggagtg	agactctgtc	tcaaaaaaca	aacaaaagaa	52140
tctcaactct	taatatggaa	tgattaatgt	cctataagaa	agcattgggt	agatcaattg	52200
ggtatatgca	tgtgatatcc	ctcctagata	gctctcagtg	gtcctcatct	cctgatattc	52260
atgccctgtg	gagtctcttc	acacaatata	tagaactgac	ctgtgtaacc	actaggatat	52320
cacagatact	acagcatgtg	gcttctgagg	ctaggggtca	taaaagacac	tgaagctgct	52380
gtcttgtctg	ctcttggtt	tctcattctg	ggggaatcca	gctaccatgt	catggggaca	52440
tttaggacat	ttaagcagcc	aaagagagag	ggccgcatgg	caagaaactg	aggtctcctg	52500
ccaatgacca	gcactaacct	attgtcatgt	gaatgcacca	ccttgaaaat	ggatcctcca	52560
gccccagtc	ggcctttatt	tattttattt	cttattaatt	gagaccgggt	ctcattctgt	52620
ctcccaggct	ggagtacagt	ggcaccatct	tggctcgctg	taacctctgc	ctcctgggtt	52680
caagcgattg	tcattcttca	gcctcccag	tagctgggtt	tacaggcgtg	cgctaccatg	52740
cccagtagtc	ttttttgtat	tttttagtga	gacaggggtt	cgccatgttg	cccagggttg	52800
tctcaaactc	ctggcctcca	gtgatctgcc	tatctcgagc	ccccaaagtg	ctgagattac	52860
aggcaagagc	cattgtgcc	ggcccccat	tcaagccttc	agatgagatc	acagccatgg	52920
ccaacatctg	gagtgcgaac	tcatgagaca	ctctgagcca	gagctgcca	gctgagctgc	52980
ttccagattc	caggcccaaa	gaaaatgtat	gagataataa	atgtttattg	ttttaagctg	53040
ctaaatttta	aggtaacttg	ttatgcagca	atagataact	tttatatgct	gccataaaaa	53100
tattataaaa	ccatgacta	gtacagaaag	atttttataa	aatattaagt	ggaagaaaag	53160
aaaagcaggc	caccaaacag	cgtaggacag	tagaccccat	ttttgaaaga	aaaatgtgaa	53220
gagttaaaaa	actctaccaa	aaggggaaaa	aaaagagggc	atcaatggag	agatggagaa	53280
gctttgtttt	tggatgggaa	gactcagtat	ggtagagtta	acaacctctc	gaaattaaat	53340
taaatggaaa	tgctattgaa	atcccaactt	gattcttttg	agtgtaggga	tctttacaac	53400
agataactgg	acaagctaac	atttattggg	tatatatgtg	cgttgcatca	cgtgacagtc	53460
actgtttcat	cttaattcca	ccataggaga	aagtccctct	ttattttaatt	tttctgagag	53520
taaagtactg	ctattacctg	ttcccccttc	cattttactt	aggagggttc	aagaggggac	53580
ttgtctgaga	tccctgaaac	cgtggagggtg	agatgacatc	aagcatgttt	gatattttaca	53640
tgtgtgcctt	tgggctctct	gccacatggc	ctccccactg	tgccttggtt	tccctaagta	53700
ccagcccaag	gacacattga	taggaaagg	ggagctgggg	caccagccca	gtctgcctga	53760
ctccagagtc	cctggtctta	atcactaaac	caccccgaaa	aagtaaccgt	gggagaagag	53820
acctgcaaac	taggaaaaag	aagattaaag	ggaaggaaatc	tgttctgcta	gatattaaaa	53880
catatgacaa	agctgtagaa	attaaaacag	aatggggccg	ggtgcagcag	cttatgcctg	53940
taatcccagc	actttgggag	gccaagggtga	gtggatcacc	tgaggtcagg	agttctagat	54000
cagtgtgacc	aatatggtaa	aaccctgtct	ctactaaaag	tataaaaaatt	agctgggcat	54060
agtgggtgtg	acctgtagtc	ccagctactc	tgcaggctga	gccaggagaa	ttacttgaac	54120
ctgggaggca	gaggttgacg	tgagccaagc	tcacactact	acactccagc	ctgggctaca	54180
gagcgagact	ccagttcaaa	aaaaaaaaaa	aagaaagaaa	aaaaaagaag	aagaaaaaaa	54240
aaaaccgggc	gcagtggtct	atgcctgtaa	tcccagcact	ttgggaggcc	gaggtgggtg	54300

gatcacctga	ggtcaggaat	tcaagaccag	cctggccaac	atggtgaaac	cctgtctcta	54360
ctaaaaacac	aaaatcagca	gggtgtggtg	ctgcatgcgt	ataatcccag	ctacttggga	54420
gactgaggca	agagaatccc	ttgaacctgg	aaggcagagg	ttgcagtga	ccaagactgc	54480
gccactgcac	tccagcctgg	gcaacaagag	caaaattcca	tctcaaaaac	aaaacaaaac	54540
aaaaacaaaa	acaaaaacaa	acatagaatg	gggtgctggcc	tagaaagcca	caaacagatg	54600
aatggagcac	aacattaagt	ccaggaataa	actcaaacac	acaggaaagt	ctggtgaacg	54660
ataataggag	gtaggtatct	gtgaagcgtg	gagtaagagt	gagttactca	accaacagtt	54720
ctgctgccac	tggctagcaa	gtaaatgcgg	atccctaccc	cactgctagt	ctccaaaatt	54780
aattccaaat	gagtcagtta	aatgtttaaa	aaaccttcaa	aagttgccag	gcatgggtggc	54840
tcacgcctgt	aatcccagca	ctttggggagg	ccgaggcggg	tggatcacct	gaggtcagga	54900
gttctagatc	agtgtgacca	atatagcaaa	accccacctc	tactaaaaac	acaaaaatta	54960
gctgggcatg	gtcagaggcg	cctatcgtcc	cagctactca	ggaggctgag	ccaggagatt	55020
tactagaacc	caggaggcag	aggttgcagt	gggccaagat	cacaccaccc	acactccagc	55080
ctgggcaaca	gagtgaact	cgttctcaga	aaaaaaaaaa	aaaaaaacct	tcaaaagtta	55140
tagaaagtct	gtgtgagtaa	tttttaatat	gaagcaaggg	agagagatga	agcagggttt	55200
ctaaatccga	ctatagccat	ataaaagtat	gttataaagc	tgggcgtggt	ggctcacgcc	55260
tgtaatccca	gcactttggg	aggctgaggc	gggtggatca	cctgaggtca	ggagtttgag	55320
accagcctga	ccaacatgga	gaaaccccg	ctctactaaa	aatacaaaaa	ctagctgggt	55380
atggtggcgc	atgtctgtaa	tcccagctcc	tcaggaggct	gaggcaggag	aattgcttga	55440
agtcgggagg	tggaggttgc	agtgaagccga	gatcgcacca	ttgcactcca	gcctgggcaa	55500
caagagcgaa	actccgactc	aaaaaaaaaa	aatgttataa	aaccacacac	cactataaag	55560
aaatgataat	gcaaaaatca	taaaggacga	aaaaaaaaag	gaaatagatt	taactacaaa	55620
aaagtttttg	ttttgctttt	gttttttttag	agttagagtc	ttgttctttt	tcccaggctg	55680
gtacaatcat	agctcactgc	caccttgaac	tcttgggctc	aagcaatcct	cctgcctctg	55740
aaactgcgtt	tgcaaaaatt	ataactgaga	aaacgatgac	agtgaagag	atctgacct	55800
actgactcca	tcttgcttct	aacctccaag	ctgtccgtgt	tcattcctgg	gtgtaggcca	55860
aactaacttt	gggaggaatt	tagtttatag	tttaactttg	tcaaagttta	actaagatgt	55920
taatagccca	ttttccaaaa	caaaccctt	tcctgcctgg	ggactagact	gcctttgcag	55980
gactaacaaa	ttattatagc	taccagatta	gaaattatgg	tttaggagtc	atgcagctga	56040
agcctacaag	attctgaatc	tcccaaattg	ctcctggaga	taacatcacc	attgtaaaac	56100
ctaagatcag	tgcttgacat	attttgcaga	cctcgcactc	gatggatcag	ctggcactac	56160
ccaaatggat	aaacaggctc	atctgatctg	tggtccccac	ccagaaactg	acccagcata	56220
agaggaccgc	ttcaactcct	ataactttgt	ctccaacctg	aacaatcaac	actcccctac	56280
ttctgaccc	cctaccacc	aaattaccct	taaaaacctt	agccaggcgc	agtggctcat	56340
gcctgtaatc	ccagcacttt	gggaggctaa	ggcaggcgga	tcacctgagg	tcagggttcg	56400
agaccaacca	tggccaacat	agtgaaaacc	catctctact	aaaaatacaa	aattagccag	56460
gtgtggtagt	gtgcgcctgt	aatcccagct	actcaggagg	ctgaggcagg	agaatcgct	56520
gaacccggga	cacagagggtg	gcagtgaagc	aagatcactc	cactgcactc	cagcctgtat	56580
gacaagagca	aaactcggtc	tcaaaaacaa	aacaaaacaa	aaaaccaca	gaaaaaaacc	56640
ctgaaccatg	atcctaaact	ctttcactat	tgcatgtccc	ctgacttgat	acattggctc	56700
tgtctaggca	gcgggcaagg	ataaccatt	gggcagttgc	acctcagcct	cctgagtagc	56760
tgggattaca	aatgcaagcc	acagctaaaa	aaattttcaa	acctttgtag	gacagacaaa	56820
ttggggaaaa	catttgcaac	aaagggatga	tacacataca	tataaagagt	tctttcaagg	56880
ctgggcacag	tggctcacgc	ctgtaatccc	agcactttgg	gaggccgagg	caggcagatc	56940
acgagggtcag	gagttcaaga	ccagcctggc	caatatgggtg	aaaccccatc	tgtactaaaa	57000
atacaaaaaat	tagccgggtg	tgggtggcatg	cgctgtaat	cccagttact	caggaggctg	57060
aggcaggaga	attgcttgaa	cccgggaagc	agagattgca	gtgagccgag	atcgaccac	57120
tgactccag	cctgggtgac	agagtgaagc	tccatctcaa	aaaaaaaaaa	aaaaaaaaag	57180
agttatttca	aattaataag	aaaaataacc	aacacaattc	aatagaaaaa	tggggaaaaa	57240
agaataggca	ctttacaaa	aaataaatac	aaagccagct	ggacatgaaa	ctttgccatc	57300
tccttagcag	ggcttgagg	aagatgggga	gaaggaaaga	tgcaaaat	aaggaggctg	57360
tcactctcag	ggccatgtaa	gtacaaaagt	ggcatatgag	ggtaagtgcc	tccttaaatg	57420
tgtaaatgca	tagagccctg	ctggatggca	gtctggcaaa	atggatcaat	attttaaatg	57480
tacaaaccct	ggcacaatga	ttccattttt	aggaactgac	cttatggaaa	cgatcaggca	57540
agtgtgccaa	gaaacacatc	taggatgttt	ttaatgtcga	caaattagaa	atgacaggta	57600
aattcaaccc	tacggactga	ctttaaaaat	tgttacatct	ggctgggcat	ggtggctcac	57660
gcctgtaatc	ccagcatttt	gggagaccaa	catgggagga	tcgcttgagc	ccaggagttc	57720
aagaccagtc	tgggcaacat	agggagaccc	cgctcgctaca	aaaaaaaaaa	aaagtaaaaa	57780
ttagccaggt	tgggtggtgca	tgccgttagt	tctaactact	caggaggctg	aggaggagg	57840
atcacttgag	ccctagaggt	caagactaca	gtgaactgtg	attgcgccac	tgtactccag	57900
cctgggcaat	agagtgaggc	cctgtctcaa	aaaagaaaaa	aaaatgttac	atccaggtag	57960

attggcatcc	tgtgtaaaaa	ggatgccatc	ctgtagtccc	agctgcttgg	gaggctgagg	58020
caggagaatc	gtttgagccc	aggaattcga	ggcttcagtg	agctatgttc	acaccactgc	58080
acttcagcct	aggcaacaga	gcaagacttt	gtcaataaat	taaaagaaaa	ataaaaagta	58140
cgtcactgtt	ctatagtggg	cgtggaaaga	ggctcagggg	atgccatatt	ggtacaagtg	58200
aacagaggaa	ccaacatatg	tcacatgata	ctatTTTTTgc	cacctgcccg	tgTTTTatgtt	58260
cacatttgga	aatatttggc	caaagtaatg	gtccctattt	cctgggtggg	ggattaattg	58320
caggggattc	ttactttctt	ctttatgcct	gctgcataaa	tacttgaaat	ccttaaatac	58380
tgcttaatac	ttgaaaaagt	gattaaagct	aattttgtct	gagaaagaga	gtgggagtta	58440
acctgttatt	ctgtaacttc	ctggccccac	caggggttgac	tcctgcagag	cattctccag	58500
gtaaattgtt	ttgccctggc	ctgactgtat	ttcagaacta	ccaggagggtc	gttttgttta	58560
tcaaccaccc	agtgggggtca	aaaagaccct	taacttctac	aattccagcc	aaataaacag	58620
aagttgcttt	cgaaggtcta	gggcctccca	ttactaggat	cagtgaagttt	aggacttcag	58680
ggtagtgga	agggccttgg	tcccacagag	ctgtctcagg	gcacttaaat	ttccctaagt	58740
gtaaaatgga	cagcttcaac	cgtatcagtg	tttctcacct	ttctcttttc	ttttcttttg	58800
agacaggggtc	ttgctctgtt	acccaggctg	gagtgaatg	gcaagatctc	agctcagtg	58860
cgctcaacc	acccaggcta	atcaatcctt	ctacctcagc	ctcccaagta	actgggacta	58920
caggcctgtg	ccaccatgct	tggctaattt	tttgtagaga	tggggtttca	tcatgttgcc	58980
caggctggta	tcaaacgcct	gggctcaaga	gatcctcctg	ccccagcctc	tcaaagtgt	59040
gcgattacag	gcgtgagcca	ctgtgcctgg	ctttttctta	aactcactct	cctttttatg	59100
aaagataaaa	ttcttacacc	cttcctagt	ggtagctttt	tccttattcc	aatagccgag	59160
aagatactgt	ggaactttac	tttctgtaga	ttatatcacg	aaaacaatag	ttgtccccc	59220
agctcatttt	ccaaaattaa	ataataattc	taagtatgct	tgtttgtaca	cagtacagga	59280
ctttctgaag	ccacaggcca	cctccagtc	tggctactga	tgcttgggg	ccttctctgg	59340
ctctcaatta	aaagctatag	tgtagtga	gagtacccca	gttctgggac	acaacctggg	59400
tgaggggtgc	caggtaaaat	acagggcggt	ctgggggagg	tggcccacgc	ctgtaatcct	59460
agcacttttg	gaggccaaag	tgggaggatc	aggagttcaa	gaccagcctg	gccaacatgg	59520
caaacaatgt	ctatactaaa	aataaaaaaa	ttagcctggg	gcagtggcac	atgcctataa	59580
tgccagctac	ttgggagggt	gaggcacaag	aatcacttga	accagggagg	cggagtttgc	59640
agtgaagcaa	gaccacgcca	ctgcactcca	gcctgggcaa	cagagcgaga	ccctatctca	59700
aaaaaaaaaa	aaatatagat	acacacacac	acacacacac	acacacacac	acacacacac	59760
acacacacat	atggtgtcct	ggaatctatt	tcctagatct	ggcaacccta	acctagtcca	59820
catttgggcc	tctgtctcca	ggcagtggtga	ctataagcac	agtctgtctt	tccttttttc	59880
tttgtctcac	cctctttctt	cttctttcct	tcttccctcc	ttgctgtcct	gctttctcct	59940
tctttcattt	ttcttctctc	cttctctccc	ctccactccc	tcctccttcc	ttcctttatt	60000
ccttacttcc	tctctctctt	tctctctctc	tttcttccct	aattgtgtca	agtgcacaa	60060
tcttaatttt	aaatatgcag	cttgatgaat	ttttacatat	gcataaactc	ctgcaaccac	60120
taccagatt	aaggagcacg	tttccagcat	cccaggaaat	tttctcatgc	ctcttgcctg	60180
tcagtatctc	ccccagagg	aaccactctt	ctcacagcct	gttattgtca	attaattttg	60240
tatgttcttg	aatttcataa	aagtggaagt	atgcaatatg	agctcttaag	tgtctggctg	60300
cttcttctta	acctaattgac	tgagattcat	tcagggtgtc	atatataaca	gtattttccc	60360
ttttcattgc	tgtataatat	tccattgtgt	gaattttttt	ttggaggggg	gagttttgtt	60420
tcctgaaaac	accacaattt	gtttatccat	cctctgtctc	atagatatatt	ggttgtttcc	60480
agtttggggt	gtaaattcaa	aataaaatcc	taagggtcca	ctaaatgaac	acccttcttg	60540
gcaaagggaa	ccccagaaaa	actttaaaaa	ctttgtttcc	agccatgatg	agacaggagg	60600
tcaggcacac	cacattacac	tcccttccct	ccttttgtgg	tttagatata	agaaaagatc	60660
agcatcaatg	ctaaaataga	gggctgagta	tgggtgactca	cacctgtaat	cccagtcccc	60720
tgggagactg	aggaaggcag	atcacttgag	gccagaagtt	cgagaccagc	ctgggcaaca	60780
tggtgaaact	ctgtctctac	aaaataaaat	aaaataaaat	aaaataatta	gccaggcacg	60840
gtgggtgcgtg	tcctgtggtc	ccagctactg	gggaggctga	ggtgagagga	tcgcttgagc	60900
ccaggaagca	gaggctgcag	tgagtcatga	tctttccact	gcactccagc	atgggtaata	60960
gagtgaagct	ctgtctcaaa	aaaaaaaaaa	aaagagagag	agattataag	actgacagaa	61020
cagacttttt	gtggcaataa	gataccaaat	tataaacaca	gcctaaggcc	atgtcaggca	61080
aggggttaagt	caggtgcccc	tactcttaag	gaataaacta	tgttctaatt	atgttacaag	61140
atttttcttt	ttctctagca	gcgaaacaag	cactggcctc	agaagaagca	atattaaaac	61200
agttacaact	catctagcac	acagacaccc	aactgacacc	ctgttccctc	agtcataaca	61260
acaactacag	ctttgattga	acaagagact	gagtttggtg	actttctcct	aataaaaaag	61320
tcactgacta	tggactgctt	ctgggtgggg	tacgaaaccg	caacctcatg	tgcttgcatt	61380
tcctgaaaag	acattttgat	gtgtaggttc	taattgtaat	acattgattg	attgattgat	61440
caattgattg	attgagatag	ggctttactc	tggtgcccag	gctggagtg	agtggcacga	61500
tcacaactca	ctgcaacctc	tgccctcctg	gctcaagcaa	tcctccacc	tcagcctccc	61560
aagtagctgg	gactacaggt	gcacgcaact	gcgcccggct	actttttgta	ttttttgtag	61620

agacaggggt	ttcgccatgt	tgcccaagct	ggtctcaaac	tcctgggctc	aagcgatcca	61680
cccaccttgg	actccaaaag	tgctagatatt	ataggcatga	gccaccatgg	ctggcctaata	61740
tgtaatacat	ttaaatgtta	agtctccacc	ccaaagtga	catgggttgt	atgttacatg	61800
cacatttgtt	catacacatg	tggtggggcc	accttcataa	atattcatag	cttctctgt	61860
aacctgctgg	atatatcatt	cagccaaccc	cttcagcaca	aagctcctaa	cccaacccct	61920
cctccttcaa	agtgcccgctc	tctgttcttg	gtaggaggca	tacttcccag	gccatggact	61980
ggtcaccttg	tgggctataa	ccccttataa	gaaataagat	ttcttctcct	ctctgaattt	62040
acacatttgt	gatttttttt	tttttttttt	ttagttaaca	ggggctatga	acattcttac	62100
agaagccttt	tgattgatgt	gtgttttcat	ttatcttggg	tatatatata	ggcgtgggca	62160
tgatagatat	taggatagcc	atctttaact	tcagtggatg	ctggagcaag	tttctgaatt	62220
tcaactctga	agtggggatg	ataataacag	cacctgcctt	acagggctgt	ttcgagattc	62280
aaagagaaaa	tctgggtaag	gcagggtgcg	gtggctcacg	cctataatcc	caccactttg	62340
ggaggccaag	gtgggcagat	cacctgaggt	caggagtcca	agaccagcct	ggccaacatg	62400
gtgaaaccct	gtctctacta	aaaatagaaa	aacaatgagc	caggtgaggt	ggtatgtgcc	62460
tgtaaaccce	gccactcggg	agtctgaggc	aggagaattg	cttgaatctg	ggaggcagat	62520
gttgcagtga	gttgagatgg	caccactgca	ctccagcctg	ggcgacagag	tgagactctg	62580
tctcagaaaa	aaataaaaaa	gaaaaaaaga	aaatccaggt	atttagaatt	ggtacaccgc	62640
aatttcaaaa	acgtaaatta	ttgctgtgat	ggcagtgggg	agcatgaaga	tattggacta	62700
acttttatga	atgttcaagt	tgatccatga	tgaattaaac	acacagggaa	ctttataaag	62760
gccatatgtt	atataagtga	tacatgacta	ttgtattaaa	attcaaacta	gttagatata	62820
aagtaaaaaa	tggttttcac	cctatccatt	ttttattatt	gaagaaaaaa	aaatatgtca	62880
tagcgtgggtg	gcttatgcct	gtaatcccaa	ccctttggga	ggtcgggggtg	ggatgattgc	62940
ttgaggccag	gagtttgaga	ccagcttggg	caaaatagca	agaccctgtc	tttcaaaaaa	63000
gtaagtaatt	tggtctgggtg	ttatggcatg	catctgtagt	cctggctagg	ctgaagcaga	63060
aggattgctt	gagcgcagga	gttcaaggcg	ccactgcact	ctggcctggg	tgacagagtg	63120
agatcctctc	tctctctctc	tctctttttt	tttttttttt	ttttgttttt	tgagactggg	63180
tctcactctg	tcaccagggc	tagagtgcag	tggtttgatc	ttggttcaact	gcaagctccg	63240
cctcccagtt	caagtgattt	tcttgccctc	gcctcccag	tagctgagat	tacggacatg	63300
tgccaccacg	gccggctaata	ttttgtattt	ttagtagaga	tagggtttca	ccaacatgtt	63360
agccaggctg	gtctcaaacg	cctaacctca	agtgatccat	ccacctcggc	ctcccaaagt	63420
gctgcgatta	caggcaagag	ccactgcgcc	tggcctgacc	ctgtctgtta	tcttttcttt	63480
ttcttttttt	ttgttttctt	tttttttttt	agacagagta	tcgctctgta	gcccaggctg	63540
gagtgtgcag	tggtgccatc	ttggctcact	gctacctcca	cccaccaggt	tcaagcaatt	63600
ctcctgcctc	agcctcctgt	gtagccagga	ttacaggcac	accccaccac	tcctggctga	63660
ttttttgtaa	tttttagtaga	gacgggggtt	cgccatgttg	gccaggctgg	tctcgaactc	63720
cttgacctcag	gttagtcacc	gacccatggc	tcccaaagtg	tcagaattac	aggtgttattg	63780
cactgtgccc	agccgaccct	cttttaaaaa	aggaaaaaat	actatgcagt	gagtattttg	63840
catgcatttt	cttatttcat	cttcgtcttt	ttatttgatg	atactaaagg	caggtgttag	63900
aggctggatt	gctaaagctg	acccaaagaa	tgccctccctc	agggctgggt	ggtccctctc	63960
tctcaggcct	cagtcttccc	atctgtacag	tgagggtgct	gcagatctct	gggctctaaa	64020
aatcacagct	ccatgtttat	ccctggcaga	ggaaggccct	ggagtccctgc	tgcttgctgc	64080
tctgggatac	gggagcaaa	agccacgcat	cctcatggcc	cacacaggcg	tcacctccag	64140
tctctccttg	gcctcatctc	cccagcgtcc	tggaatggca	tcgggctggc	ccaggggagcc	64200
cctgtcctgt	gcctctcctt	tcccctcagg	ggctgccagg	ctgaccaccc	ccaccgcagg	64260
ccaggcctac	agtgccccat	ggaacgtcct	gaccctcccc	caggggtggca	gcaggaagaa	64320
ggaagaaagg	ggatcctctc	cagctggcca	gagagacaga	ccttcttgtg	ctcatcaacc	64380
ctccaagaat	gcctgccttc	cctccttccc	ccaaggcctg	tccacagggg	cttgagatca	64440
gccagaaaag	tcaggcaact	tttcagggac	tgggagcgag	gtctcccggc	cgggcctggg	64500
tccagtctct	gtgggcagtg	cagtgccgag	ccccaccctt	caagccgtgc	cctgtccata	64560
gctccagact	ttgacctgc	actccagtcc	gggctggcgg	acagagggct	ggaaacaaga	64620
cgctccagaa	tcagagcctt	ccctcagga	aatagcatcc	tgtgtccccg	cactgcagtt	64680
gtctgggtctc	tccagcagtt	tggtacttcc	ggtgagtggc	agatgcacct	ttgagctggg	64740
gacaggggtt	gggagagggg	agaggcaaa	gatttcatgt	cctcccaatg	tcaaagacag	64800
ggctcaacat	tacagcctaa	ggcaggtgac	aggaaaggag	agatccagcc	tctcaaacat	64860
ccagcagaga	gaccataggt	aagtgatttt	tccctcccca	agcctcagtt	tcttcacctg	64920
gaacatgggg	atcataactc	ccctcttaca	gcgtgagtct	gagtgttaaa	agaggtgggtg	64980
catgtaaagt	gcttagagca	gatctaggca	catagcaagt	actcaaattg	tagttattat	65040
tatttttggg	gggggagttg	gtaggctggg	tctcaaactt	ttatagcttc	tgttccattt	65100
caaggataaaa	ctctgcaaat	aacttcatga	gaagtagccg	tgtgggtgcaa	ccaggggagaa	65160
ctaattatgt	tcattcaaat	gcctcatctc	tggcttactg	attttttttt	ttaaaaagaa	65220
gtctttcata	ttcttttgcta	tgggcacata	gcaatcaaag	gcatcagctg	tctcagattg	65280

ccttctaggg	gacaagggag	gtcctaggca	gataaatgca	agactgaaag	acaagcagaa	65340
agcatcaagt	ggcaactgca	tgccaactgc	ctaaatatatt	ttttggagca	gtgcagaaa	65400
cgccgataga	actgggtcta	ggtccgaatg	ctgtcccata	ctgactgcgt	aaccttgggt	65460
gggtgacttc	tcctccctaa	acctcagtc	cagcctccag	aatgagggcg	gtaaccttcc	65520
ctacttccta	gagcagttga	gaggattgag	aggattatgt	cggtactgca	tctacaggtg	65580
tctggcaagt	ggcagagacc	aaaatacatt	ggttcccttc	ctgctccaca	cttacacaga	65640
catttctaate	acacacacac	acacacacac	acacacacac	acacacacaa	atataataat	65700
cccagctgtt	tgcatcttct	gggatacata	ctccaagctt	gctgggttga	agtaatgatg	65760
taaaacagag	gagaacggca	acactaataa	aaacatcagc	aacaacacga	aaatgtccaa	65820
ccgaataaact	gagctgggtg	cgtttaagtc	caaaagctca	ttacctacac	gcatgaatga	65880
ttttaccta	ggctggatct	gccacatctg	acaatctgtc	tctggcttgt	catgaggacc	65940
tcatgcattt	atthttgtatt	ttaaaacaca	cacacacaca	cacacacaca	cacacacacg	66000
ttgctataat	cagtgtcaac	tttgactcat	atcttgaatt	tttttaaaaa	aagataattg	66060
acttaggact	cacacttttt	tcctttttaa	tttttttttt	tttttttttt	ttgacagagt	66120
ttcactcttg	tcacctgggc	tggagtgcaa	tggcatgatt	tctgcccact	gcaatctcca	66180
cctcccaggt	tcaagggatt	ctcctgcctc	ggcctcccga	gtagctggaa	tttcaggcgt	66240
gcaccacat	gccaagctaa	tttttttgta	tttttgtaga	gacaggggtt	caccatattg	66300
gccaggctgg	tcttgaactc	ccgacctcaa	gtgatctgcc	agcctcgacc	tcccaaagtg	66360
ctggaatttaa	agacgtgagc	cactgtgccc	ggcctttttg	atthttccatt	ctatttctac	66420
caacactcta	aaaatttccta	caggtctttt	atthttattht	atthttatctt	atthttatatt	66480
atatttttatg	tttgaaatgc	aggactctga	agcttcagct	gttcctattht	accggcctga	66540
ttctcagatt	tttcaaacca	tgtgattttac	tggcaagcat	ggcattttaag	cacctaggct	66600
tatgagtcag	gctggcctgg	gctctgcctc	tcaccacctg	ggtgtccagg	agctgatatt	66660
ccagttagga	gacaataagg	caaggagctt	tgtcagctct	cataaaaagt	tatagatgag	66720
gtcgggcatg	gtggctcacg	cctgtaatcc	tagcactttg	agagtctgag	gccagcaaat	66780
cacctgaggt	cagaagtttg	agaccagcct	ggccaacatg	gtgaaacctt	gtctctacta	66840
aaaatacaaaa	aatttagccag	gcatgtttgt	gcatgcctgt	aatcccagct	actcaggagg	66900
ctgaggcagg	agaatcacct	gaaccgggga	ggcagagtct	gcagttagcc	aagattgtac	66960
cattgcactc	cagcctgggc	gacaagagtg	aaaatttcctg	ctcgaaaaaa	taaagtttat	67020
agatgaggaa	actgagggttc	gattaggatt	aaccaactca	tcctgggttg	cctgggactc	67080
tgatgactg	acttttagtc	tgaaagtctg	catcctggga	ggacctcag	ccctgggcaa	67140
gctggggagg	ttggtcaccc	tcactcagtc	aagttgagca	acttgcccag	ggttacatgg	67200
ctgggtgtgtg	cccaagtcag	gctgcgaacc	tgggtctgtc	tgactctcag	cctggggccat	67260
actgtctctt	agattcttca	tggagaatta	ggaaaaatac	agaaagccct	ttattcctct	67320
gccttctcat	tgttaacata	taaaaatggt	caagcgggcg	ggtgcagtgg	cacacacctg	67380
aaagcccagc	gctttgggag	gctgaggggg	gaggattgct	tgagcctagg	aattggagg	67440
ggcagttagc	tatgattgtg	ccactgcact	ccagcctggg	tgacagagtg	agacctgtgc	67500
tcttaaaaaa	aagaaaaaga	gtggtcagct	ctccggaaat	tatgcagaca	gtcaaaaagc	67560
ccagagaggg	gaattaactt	agccaaggct	gcacagcaag	gcagaagtga	agccaggctc	67620
gactctgcct	ttctcttctc	ctcttttttt	ttttgaggca	gaatttcgct	ctgttgccca	67680
gactggaatg	cagtgggtgcg	aactcgactc	gctgcaacct	ctgctgcccc	ggttcaagcg	67740
attctcctgc	ctcagcctcc	cgagtagctg	ggattacagg	cgcctgccac	cgcgctggc	67800
taatttttgt	agttttcagt	agagatgggg	tttcaccatc	ttggccagac	tggctttgaa	67860
gtcctgacct	cgtgatccac	ccgcctcggc	ctcccagggt	attgggatta	caggcgtaag	67920
ccactgcagc	tggtcctccc	tctctccttt	tgttcctgca	atgtctttgt	tctatgtgat	67980
ttttcaaaat	gctaggagac	aggaaggagg	ctgctgtgtg	ttgagggcct	actctgtgcc	68040
aggcgtggta	ccaagaactt	ttgctaaact	tcttatttaa	tccttaaaat	gacctgtga	68100
gattgggatt	aaccctgttt	tgcagatgaa	gagcttgtgt	ctccagaggc	aaagtatggg	68160
ggaagaggga	agagagaaga	ccaagggctc	ctgagagggg	ctgtcaccta	agccccagta	68220
tccaagctcg	ggctcgaagc	tggaaaggaga	attgcctaga	ggaacgatac	ctttctgttt	68280
gttggttcta	tctccaactt	ggcttctgaa	accccaacag	agtcagttc	ttgtgggctg	68340
gagccgtttt	ccctccttta	taaaactagg	ccatatttaag	aatgtcccg	tgtccagggc	68400
cacaggccccg	agttgccagg	agctgagggtc	tgcgggagga	gagttgtgag	tgaagaggag	68460
ggaaagtga	atthggctct	tctgggcaca	aataattctc	ttgttctgcc	tcagcaggag	68520
cctgcagaat	atthccctgc	tgtgcgggct	taagtagctt	caagggttaa	agctggtagg	68580
ccttctaatac	ttctcagggc	ccaatcagcc	ctgtgcccc	aggcagggtg	agttctgtgc	68640
tggaaagacca	agttctgagg	ccagacactg	cgtctgtcat	gctcatagct	gcatttgcta	68700
gctgccagcc	tggcacatgg	taggtgtgca	ttaagcgtgt	gttgagtta	ctcaaattga	68760
aattaagtca	cagctgtacc	atttaactgg	ctgtgtgact	tcaggtaagt	cacatcacct	68820
ctctgaacca	cagtttcttc	ctctgtaaga	cgggactgat	aacagcagcc	cctacctcat	68880
gagagtgttg	ggagacttgg	atgaatggat	gcttgtgaag	cacttagtgc	cggggccagc	68940

2002247939 08 Oct 2004

tggtcacag	taggtgctcc	acaaatgtca	gtatattact	tcttttgcag	caggcagctt	69000
gttaaatttg	ttacgttttg	catcttggtc	aatttcccat	ccatcccctc	aagcataggt	69060
tattagaggt	tgaagcatct	tgcccaaagt	taacggccag	taggtggcag	agctgagtc	69120
tgaagccaga	gcccacgcga	ctaaccaccg	gcctaccag	cctacagttg	gtcgtgccct	69180
ctgctgggtc	ttttctatct	ccagcccaga	aactgggtgt	ctggggacgc	tccccagaga	69240
aagttgcatc	attcaccagc	cgtgtgactg	tgcccaagtc	tccgtcactt	ctccatacct	69300
cagtgtttcc	at ttgcaaaa	cggaacaat	gatattcctt	cctcctaggg	gtcatcgga	69360
aggtcaaata	taaaaagggc	ttggtggtgt	ctggcacctt	ctaagccttc	agtggatggt	69420
ggcaatggcg	ctaaggatga	tggagatgat	ggtgatgatg	gtgtgcctca	acccttcctt	69480
cccacaggct	gctgcaatgc	gtgtggtggt	gattggagca	ggagtcatcg	ggctgtccac	69540
cgccctctgc	atccatgagc	gctaccactc	agtcctgcag	ccactggaca	taaaggtcta	69600
cgcggaaccg	ttcaccaccac	tcaccaccac	cgacgtggct	gccggcctct	ggcagcccta	69660
cctttctgac	cccaacaacc	cacaggaggc	gtgagtgagg	gtcacatagg	gtagcctggg	69720
gtgccatgg	acctaagtct	gcagagggag	tcagggttcc	catcaccaag	agcaagcccc	69780
ttgtggaagc	tactgatcta	gcataaaata	aagaaaatgc	caggcgtggt	ggttcacgcc	69840
tttaatccta	gcactttggg	aggtcgaggt	gggaggatca	cttgaggcca	ggagttccag	69900
atcagcctgg	gcaacgtggt	gaaaccccat	ctctacaaa	aatacaaaaa	attagccggg	69960
catggtggcg	cacacctgta	atcccagcta	ctcgggaggc	tgaggcagga	aaaccatttg	70020
agcctaggag	gtgaaggtgg	cagtgaagctg	agattccgcc	actgcactcg	tgacagatgt	70080
agactctgtt	tcaaaaagaa	aaaaataaag	aaaagattca	taaatattaa	gccccttgct	70140
ctgtgccaga	tactaggagg	ccttgctctcg	tcttccttaa	actgggtgcc	tgtcaataacc	70200
acatgattgg	tgaatctgga	aaacttcctc	tgttttaatt	tatacatttt	tatttatttt	70260
ttgagattgt	gtttcactct	tgtcgccag	actggagtgc	aatggcgtga	tcctggctca	70320
ctgcatcctc	tgccctccag	gttcaagcga	ttctcctgcc	tcagcttccc	aagtaactgg	70380
gattacaggc	atctgccacc	acgcctggct	aatttttgta	tttttagtag	agatggagtt	70440
tcattgttgg	cagactggct	tcgaactcct	gacctcaagt	gatctgcca	ccttgacctc	70500
ccaaagtgt	gggattacag	gcatgagcca	tcattgcctt	ccaaatttta	tctttttaaa	70560
tagagatagg	gtctcactat	gctgcccggg	ctggtcctga	actcctgggc	tcaagtgatc	70620
tgccctcctt	ggcctcccaa	aagtgtctga	attacaggca	tgagccatca	tgccctgcca	70680
aattttatct	ttttaaatag	agatagggtc	tcactatggt	gcccgggctg	gtcttgaact	70740
cctgggtctc	agtgatctgc	cctccttggc	ctcccaaagt	gctgggatta	cagggtgtgag	70800
cccttgacc	cagctgaatc	tagaaaactt	ctaagtgggt	gaacatctaa	gtgggtggat	70860
ggatgcacag	at ttatcaaa	taaattgcaa	aggtcattat	ggtagttag	aaactgccag	70920
atggttcagc	aaatggaaca	ccaatgaat	agcagctcaa	acagattaaa	aaaaaatttt	70980
taagaggcat	cctgtcaccc	aggctgaagt	gcagtacat	gatcatagct	cattgcagcc	71040
ttgacctcct	gggtcaagt	gatcctccca	cctcagcctc	ccaagtacgc	aggacacaca	71100
tgcatgctat	catgcctgga	taattttctt	tattttttgt	agagccaggg	tcttcctatg	71160
ttaccaggc	ttgtctcaaa	ctcctgacct	caagtgacct	tcctgcctca	gcctcctgaa	71220
gagctgggat	tataggcatg	agccactgca	cccagccaga	at tttaattc	acacagctgt	71280
aaaaaaccaa	tgattctgat	cagtgggcag	tgattcgggg	cccaggctcc	ttccatctaa	71340
tggtcttgcc	gtttttccac	gtgcttttga	ggtcacctca	atgtcaccat	tcacatgggc	71400
tggtgactgc	agaaggatca	tgcaggacca	cacgtgcaga	gtcttttagag	tccttgggcc	71460
agaaacaagt	cacacagtca	catttagctg	cgagagggtc	taggaaatgt	aggtgagctg	71520
tgtgcccagg	gggaggagga	aaagtgttag	gagcggcaag	tcgatgtctg	ccaccagtgt	71580
ttacaaggag	gggtgcttgc	agccagactg	aacagtgtgg	ctcataatcc	ccaaagccag	71640
gtcaaggact	tcactgaaac	tcactagcca	tgtaatccca	tgctggaggt	gcactccata	71700
tggttatgat	ggggcatcct	tcattccctc	tcttctttat	tctattaatg	gggaaatatt	71760
ggaaaattta	ggaggagaa	gacccaagc	atttggggag	ttgcaggagt	gaacgtggtg	71820
gatttctggg	ttttggacac	accccaagct	cctgatcatg	ccacagcccc	atgccagctg	71880
acctaagggt	ttttgccag	ctcagggcag	tggtgatcg	aactcttcat	gaccttcca	71940
gggactggag	ccaacagacc	tttgactatc	tcctgagcca	tgtccattct	cccaacgctg	72000
aaaacctggg	cctgttccta	atctcgggct	acaacctctt	ccatgaagcc	attccggtgg	72060
gtgaacagtt	cttgaccatg	agggatgagc	accagggtct	ggggtagtga	gggtgggtgc	72120
agcagagcct	taatcacaga	tgagggcggg	gtgctttgag	tctcgtaggc	aacagactcc	72180
tggtttcaaa	acaggtttgg	tttaaatctt	at ttttgctt	tgaaaattat	ttttgtttta	72240
cattttgctg	ttaaattggc	agaggacaaa	gaatcttctg	atgccaggg	gaaactagcc	72300
tttgattagc	atggctaaaa	tacaaacatg	ttctgcagtg	acgggcactt	ggtgctgaag	72360
ccaaaagggt	tcaagtgcc	ctgaaggctc	caaggctttt	tatcaagaag	gaataaaata	72420
ctcatcaaag	caaaaactgc	caaagcattt	attatgtgcc	aggtccagtc	ctaactaatt	72480
tacagctagc	gactaattta	attctcttta	taaccgggag	gtaagggtg	tccttatcct	72540
cacttaataa	atgagaaaac	ggaggctcca	agaaatggag	taacttgccc	aaggccacar	72600

agctcgccag	tggcagagct	gggatttgaa	cccaggccat	ctgtgactcc	atggtgtcca	72660
gtgtgctaac	aggaacagca	cagccctggg	acggtttgct	caggctcctt	ggagaggggtg	72720
gtctggcgct	gtgccagag	ccccgtgcc	gctctcaagg	ttcattcaac	ctttggcact	72780
gtgctaagg	ctttatccac	attatctgtt	acctttcatg	ggaccaagag	tatttttttt	72840
tttttgagac	aggggtctcac	tgtattgccc	agggctggagt	gcattggcat	gatctcggct	72900
cactgcaacc	tctgcttcc	gggttcaagc	cattctcctg	tctcagcctc	ctgagtaatt	72960
gggattacag	gtgcgcacta	ccacgcctgg	ctaatttttg	tatttttaag	agatgggggtt	73020
tcactatgac	ggccgggctg	gtctcgaact	cctgacctca	agtgatctgc	ctgccttggc	73080
ctcccaaagt	gctgggatta	caggcgtgaa	ccactgcacc	cggccaagag	tgattattaa	73140
ctccatgata	cagacaagga	aactgagtct	cagagaattc	aagtagcaag	tgatgaggct	73200
gggggtctctg	acactatgct	ctgctgtctg	acactatgct	ctgttgcttt	ctctcatccc	73260
cggggactct	cactgtttct	gctttctctc	ccctatttct	gacttttccc	ctataactca	73320
ccctcgggtct	tactcttacc	cttaccataa	ataggggtta	agaacatgaa	ctctggaact	73380
aagctgtatg	ggtaaaaatc	tcaacaccac	cattttattag	ctgtgtaatc	ttagacaagt	73440
tatttaaatct	ttctaagcct	caattgggtcc	atctgtaaac	tggggaaaga	atagcatcca	73500
ccccaatggc	ttcttgtgaa	gattaaatgg	accagtataa	gaaaatgctt	ggaacagtgc	73560
cttatatgca	cttagcatta	cataagtctc	tgtcattatc	attttttttt	ttttttgaga	73620
tgaagtctcg	ctctgtggcc	caggctggag	tgacgcggca	caatttcggc	ttactgcaac	73680
ctccagctcc	cagattcaag	caattctcct	gcctcaatct	cctgagtatc	taggattaca	73740
ggcatgcacc	accatactct	gctaattttt	gtattattat	ttagtataaa	cagggtttca	73800
ccatgtttggc	cagactggct	tggagctcct	gacctcaggt	gatccacca	tctcagcttc	73860
ccaaagtgtct	gggattacag	gtgtgagcca	cctcgcctgg	cccattatca	ttattattga	73920
cttccatccc	accagtgcc	ccctttgtcc	ttcctcttca	ggacccttcc	tgggaaggaca	73980
cagttctggg	atttcggaag	ctgaccccca	gagagctgga	tatgttccca	gattacgggt	74040
gagtttattg	tcacaggcaa	aggggactgg	ggcctgacga	gttagcagac	ctgtccagaa	74100
ggcagcagag	ggtagaggca	ccagatttcc	tgtcctaccc	aggccctgg	accctggctct	74160
cctggctcctt	ggtccagctc	cttcagagag	gctacccact	caaacctggc	cttgggctgg	74220
gaaggtaggg	ggtatgaaat	cacagatctc	aagcccagaa	gctccatatc	accatattgt	74280
tttgtagatg	aagatactga	gtttcagaga	ggctaagtga	cttcctaagg	tcacacagcc	74340
aagtggccaa	actgggattc	caaccagtct	gtatgacccc	acacccctcc	tttcttttct	74400
ctacagcctg	atgcctctct	ggtcttctcc	tcaccccacc	ccacaccaca	cctgaatccc	74460
ctctacgaat	gcacattcaa	tctccacttg	cattttccaa	tgtcagatat	ggccttttct	74520
gatagaaaaa	ttttccttgc	attgagctca	aaaccacgtc	cccccttgaa	cttcacgtag	74580
tggctctggc	actacccttt	gggccacag	aacaacattg	ctcccacctc	catttcacag	74640
ccttcaaata	tagctctgat	ttttaccttt	atttccacct	tttgcttact	gtgactctag	74700
ctatggctgg	ttccacacaa	gcctaattct	ggagggaaag	aactatctac	agtggctgac	74760
tgaaggtga	gattttaagc	ttcactttga	gggaggtacc	tcccagagac	caagttgtag	74820
tgggaagatgg	ttcgtgggct	tccctcagca	tggactaacc	cccaggtttg	aagaataccc	74880
ttaggcctgg	tgtgggagct	atccttgggtc	ctgatcaccg	ctgggcacag	aggcaatgga	74940
tctgagcct	agctgagcat	cagaaccacc	tgggcagctg	tttacacatg	atgtccatta	75000
acaacctctt	tcaaatccct	aatgtttgtg	taatagtttt	agatgtattc	ttttaagggtt	75060
tccagataca	tttacatcat	ctgcaaaaaca	taagctgcct	tttattttta	tctccctctc	75120
tctttttttt	tcttctctaa	ctgctttggc	caatacctct	agaacaatgt	tactcataca	75180
gatgatagtg	gatgtctttg	acttgttctt	catgttaata	ggaatcttgc	agtgttctaa	75240
attagcaaac	actcacttga	gagatacatt	ggtattttata	ttcacataca	ttcatattaa	75300
gggagattcc	ataccttttt	tgtgtgtgtg	agatggagtc	tcgctctgtc	accaggctg	75360
gagtgacgtg	gtgcgatctt	ggctcactgc	aagctctgcc	tcctgggttc	atgtcattct	75420
cctgcctcag	cctcctgagt	agctgcgact	acagggtgct	gccaccacca	cacctggcta	75480
atttttttga	tttttagtag	agatgggttt	tcaccatggt	agccaggatg	gtctcaatct	75540
cctgacctcg	taatctgccc	gcctcggcct	cccaaaatgc	tgggattaca	ggtgtcagcc	75600
accacgcccc	gcctgattcc	atacattttt	tatatattac	tgttttttaa	agatttttag	75660
gccaggcatg	gtggttcata	cctgtaatcc	tagcactttg	agaggccgag	gtgggcagat	75720
cacttgagcc	caggagttca	agaccagcct	gggcaacatg	gcaaaacctc	gtctctacag	75780
aaaaattcaa	aaatcagcca	ggtataatgg	tggctgcctg	tagtcacagc	tacttaggag	75840
gctgaggtgg	gaggatggct	ttatcccggg	aaggagaggc	tgcatgtagc	tgtgatcatg	75900
ccactgcact	ccagcctggg	gggacagggc	gagaccctgt	ctcaaaaaaa	aaaaaaaaga	75960
ttaaaaaaat	atggaatata	tagtggcttt	tatcagatga	cctcagaaga	ttttttttta	76020
atgtagattt	taggacccca	ctctatacct	gctgaattag	aacttctggg	ataaggttca	76080
taaatttgct	ttttttcatt	tttttgagac	aaaatcttac	tttgtcaccc	aggctggagt	76140
gggatgtagt	ggtatgaaca	caactcacag	cagcctcaac	ttcctgggct	caaatgatcc	76200
tcccacctca	gcctccaaag	tagctgggac	cacatgcatg	tgccacaatg	cctatctaatt	76260

ttttaaatat	ttttgtagag	ataggggtctc	actatgttgc	ccaggctggg	ctcaaacc	76320
tgggctcaag	caatcttcct	gcctcagcct	cccaaagtgc	tgggattaca	ggcgtgagca	76380
aacaggccta	gcaaaaat	gcattttaag	aagcttcctg	gcgattctaa	ttatcagcca	76440
tgtttgggaa	tcattgtact	aagacatggc	tatttctcct	aacctgggga	cacatgaccc	76500
ttgtccagtc	ttttccagga	aaaacatgcc	ctcaagatgt	ttttctatct	tgaggaaatg	76560
atggaaatga	gatagtcca	agggtatgct	tcaccttctt	tttggcttat	ttcctgttct	76620
ttggatgttt	ctagtgtatt	tctttctttc	ttcttttttt	tttttttttt	tttgagacag	76680
agtcttgctc	tgtcaccag	gctggagtgc	agtggcgcaa	tctcggtcca	ctgcaagctc	76740
cgcctcctgg	gttcatgcct	ttctcctgcc	tcagcctccc	gagtagcttg	gaatacaggc	76800
gcctgccacc	acgtctggct	aatttttttt	ttgtattgtt	agtagagacg	gggtttcacc	76860
gtgttagcca	ggatggctct	aatctcctga	cctcgtgatc	cgccacctc	ggcctcccaa	76920
agtactggga	ttacaggcgt	gagccaccgc	gcctcgtgtt	ttctagtgtg	tttctaatac	76980
tgatagatgt	ttttcctatg	ggatgtttta	aaggaggggt	gatgtcctca	gccacacctc	77040
ctcctcatgc	ccggcttctg	acaaagggga	atttggcact	ggtacaactc	tccccttctc	77100
tactctgaat	ctcattgcct	ttgctgttac	aaagcaatgt	ggtggtcata	ggaagtgtct	77160
ggggctaaga	ggcctgggtt	tgagtcca	ctccatcatt	gactcactct	atggccttca	77220
gcaaggccct	tccccactc	catctgcca	acaaggggct	tggaccatct	ctggtttctc	77280
aaaggagatt	ttgtggacca	ccagtccagt	agggtctcat	gagctgattt	gatgacacag	77340
ccatcttctc	aagcagcatc	ctgtgcaact	aacgtccgca	gaaggttggt	tggggaaagg	77400
tccctgtgcc	acccttcttg	gtgggtaggg	ggcagatagc	tgaatactgg	gctttttgat	77460
gtgtttgatc	atcccaggtt	aactgagagg	ggagtgaagt	tcttcagcgg	gaaagtggag	77520
tcttttgagg	agggtgagtt	cagggctgat	gcggtggatg	gggcagggaa	gaagtaggga	77580
ggcctctgct	tcttgctgct	gagtcggggg	ctcccttctc	aggctcctag	ggtccccaca	77640
ggcctgcctc	agcacccttg	ccccagaagc	actcaggtat	tctgaaggga	ggaagtctct	77700
gccttcatgt	tggtagtggg	aacaaaggaa	caactgggatc	atggtggcca	ttaggagctg	77760
atttatatct	gagactcaat	gagttttggg	tctagagagc	tggccgcatt	ttctcagtgt	77820
cagctgcact	ccaaggtcag	aacttggttg	cttcctagcc	ctaccgacat	ctgtgttggt	77880
ctttctgcaa	agtccaggcc	ctcagctgac	tcacctctaa	agaagcacca	ccaccaataa	77940
taatgacagg	aaaagccacc	atctccaggc	accagcaaaa	agagctttac	tgtatggctt	78000
cattcaatcc	cagcatctaa	aacctgtgct	ggcacaagga	aggcgctccg	tacatgtagc	78060
tactagtgtc	atgtcatgaa	gactaacctg	ctctggtcag	gccctgatgg	acaccgaaga	78120
tacatggtcg	acccaatgca	gtcctcattc	tcagtcattc	actcaggaac	aatagtagcg	78180
tcttgcaatg	tgtgtgtccc	ttaaacttact	cgtggtgaga	gtcactgggg	ctgggttggg	78240
gagcttaggg	gtcacgatg	cgtgcttgag	atgagatcat	ctcatctgta	gacagagctg	78300
gggttccaac	gtgtcttctg	caaatgtctt	ggcagagtag	aaggcaagag	aataaaagta	78360
aaaggagtca	gaaggagaaa	gagaactctc	tctgcttctc	ttctgacttc	ttttgggagg	78420
ttccagggaag	atttccccc	tccaaagaac	tgttttacaa	ccacttttat	attcagagtt	78480
gtgcaggagc	ctcataacag	cctatgaaca	gccatgggca	gcctcatttt	acaggggcag	78540
ctgagaatta	aggaggtaac	cagacatttt	caaggtcaca	cgtcagataa	atggcagtat	78600
gaaaatttga	agccaggccc	ctctgattcc	tcattgagac	ctctccccac	tgttcatcag	78660
ggagtagaca	gattgagggt	agaagaaagg	ggaagagaga	acaggggata	ccagggtctt	78720
ccccaccttt	catccccac	taccctgttg	gttgctacca	ggtggcaaga	gaaggcgag	78780
acgtgattgt	caactgcact	ggggtatggg	ctggggcgct	acaacgagac	cccctgctgc	78840
agccaaggccg	ggggcagatc	atgaagggtg	gtgtgagggt	gagaccctta	ccttttgta	78900
ataggaagat	cattctgcat	gcttatttca	tccctcaaga	tcatggacaa	atcaggaaca	78960
tctgttagag	gaaccccccg	gactgcaggg	aattgacatg	taaaaaaaac	aaacctgtcc	79020
cacccccatt	gctctctttc	aggatttcct	cttgatcggt	aagcatgcat	gtatgcgctt	79080
gtacctatgt	gggagcagca	tatgcctgta	ttgcaataaa	aatagcaaac	attagagtgt	79140
ttaccaagcg	cgagatacag	tcctaagcac	tttattgtgt	ttattattat	tattaattat	79200
taattgtgtt	attattatta	tcattgttat	tattattttt	gagacagggt	atcactccat	79260
tgcccagggt	agagtgcagt	atcttgatca	tggctcactg	tagccttgac	ctccagggt	79320
cccaccttag	cctactgagt	agctgagact	acaagcgcat	gccaccacca	tgctcagcca	79380
attttttttat	tttttgtaga	gaaaggattt	caccatattg	ctcaggctgg	tctcaaactc	79440
ctgggctcaa	gtgatcccc	cacctgggcc	tgtcaaagtg	ctgggattac	aggcgtagc	79500
caccacgctc	agcctattgt	gttaattaat	ttagtgtatg	ccacagccct	tcgagctggg	79560
tactaccata	tcgttattgt	catcttacag	atgaagaaat	tgaggcacag	aggagttaag	79620
taacaggcac	aagttcacac	ggtagtacgc	agtgcatttg	ggattggaat	ccaggcaacc	79680
tggcttaaga	gcctgtgcgt	gcaagcattg	ttccatgcct	cctcttgctg	tgtgtgtgca	79740
tatgagggta	tgtgtgtgtg	catgtatgtg	tgtgtgtgta	tgtaagggta	tgtgtgcata	79800
tgtgtgtgtg	catgtgtgag	ggtgtgtgtg	catgtgtgag	ggtgtgtgtg	catgtgtgtg	79860
aggggtgtgtg	catatgtgat	ggtgtgtgca	catatgtgag	ggtgtgtgtg	catgtgtgtg	79920

aggggtgtgtg	catatgtgtg	atgggtgtgtg	tgcacgtatg	tgggggtgat	tgtgcatgta	79980
tgtgaggggtg	tatgtgcata	tgtgtgatgg	tgtgctgtga	tgcacaccat	gtgaggggta	80040
tgtgtgtgtg	catgtgtgtg	aaggtgtgtg	cgcatgtgag	agtgtatgta	cgtgtgtgat	80100
gggtgtgtgtg	tgtgagggta	tgtatgcata	tgtgtgaggg	tatgtgtgtg	catgtgtgtg	80160
aggggtgtgtg	catgtatgta	aggggtgtgtg	tacatgcatg	tgtgtgaggg	gtatatgtgt	80220
ggatgcatgt	gaggggtgtgt	gtgtgcatgt	gtgtgaggg	gtgtgtctga	gggtgtgtgt	80280
gtgcatgtgc	acctgtgagt	gttcataggt	gtgcaggtgt	gtgtgcttct	gtgtgtaggg	80340
gtgctgtgtg	gtgttcctaa	tgtgggctga	tgggtgtaac	aaccaaata	gtgactgaag	80400
cataagtctc	aaatcatcga	ggtttatgga	gccagcttga	gggcgcaccc	aggaaaaacg	80460
cgagtcacag	atgcacctgt	gactcctttt	tccaaagagg	ttctcaggag	atttagtctt	80520
tatacatttt	ctttaaaaaa	aaaaaagtga	gagaaggggtg	tagcagcgag	agaatgattg	80580
catacttgtg	aaacttttagt	tagtgcccag	taaatctaca	ttttacataa	gatgaagggt	80640
tggggccaggc	gtggtgactc	acacctgtaa	tcccagcact	ttgggagggt	gaggcagggtg	80700
gatcacgagg	tcaggagttc	gagaccagcc	tggccaacgt	gggtgaaaccc	catctctact	80760
aaaaatacaa	aaaattagct	gggtgtgtgtg	gcgggtgcct	gtaatcccag	ctactcggga	80820
ggctgacgca	ggagaatcgc	ttgaacccgg	gaggcagagg	ctgcagttag	ccgagactac	80880
accactgcac	tccagcctgg	caacagagcg	aggctgtctc	aaaaaaaaaa	aaaaaaaaaa	80940
aaattgaagg	tttgaaggaa	aaaggaatgg	aggaagtctc	gtatctggga	agataagctt	81000
gtcattgaag	ttatcagttg	ggagtctgtt	gaaagggctg	gtttctgctt	acccttagg	81060
gaagaaagcc	taacttttgt	caggtcattg	agggagggga	tataatgaga	cgtgtcggac	81120
ctcccttccc	cccgcagctg	tgaactcagc	tccaaggttt	ctctggggct	cctggggcca	81180
agagggggctc	tgttcagtcg	gttggggact	tagaatttta	tttttatttc	tcattgtgat	81240
gcattttacat	gtgtgtactg	gtgcttttct	tccgacatgt	gggtgaggag	aaacaatgct	81300
tcagggagca	gggggtggctg	ccaattaggg	cagctcttcc	tgcaagaggc	aagcagtcag	81360
gtgcagactt	gggccatagt	gtcatgagag	gtcttataag	gaatcagcct	ggccactctt	81420
gtcaggacat	ctggccacag	aggggagcaa	gggcagccac	attgactcac	ctccgctgat	81480
gagactttcc	tgccctgaat	caacaggtgg	acgccccttg	gatgaagcac	ttcatttctca	81540
cccatgaccc	agagagaggc	atctacaatt	ccccgtacat	catcccaggg	taaaattgga	81600
ctgttctcgg	gcagaagagt	ggcccccttc	atgcccctct	catgaccctg	ctgcctcccc	81660
caagctcctt	actccctgca	gttgttccct	ttcaatgttt	ttatgtactt	agctattttt	81720
tattattatt	ttttgagaca	gagtttcaact	cttattgccc	aggctggagt	gtaatggtgc	81780
gatcttggct	cactgcaacc	tctgcctccc	aggttcaagc	aattatcctg	cctcagcctc	81840
ccaagtagct	gagattacag	gtgcccacca	ccacatccag	ctaatttttt	gtatttttag	81900
tagagacagg	gtttcaccat	gttggccagg	ctggtcttga	actcctgact	caggtgatcc	81960
acctaccctt	gcttcccaaa	gtgctgggat	tacaggcgtg	agccaccgtg	cctggccccc	82020
tttcaactgt	ttgagctagt	ttgagctagt	gaatccctgg	gaaggcagac	tcagcctcga	82080
ctgaggtcta	ccgtgaacat	tcttttggat	gacaatagt	gtgatgctgg	agacaaaggc	82140
agtggatgta	atgtgggtgac	actaaaagt	gtatgtaggt	ggctcacgcc	tgtaatccca	82200
gcactttgcg	aggccaatgt	gggaggattt	cttgagccca	ggagtccaag	accagcttgg	82260
gcaacatggc	aagaccccg	ctctacaaaa	atacaaaaat	tagccggggc	tgatggtgta	82320
tgccataggt	cccagctatt	cgagaggctg	agatgggagg	attgcttgaa	cctgggagggt	82380
tgaggatgca	gtgagccatg	ttcacaccac	tgtactccag	cttggggccac	agagcgagac	82440
cccatctcaa	aaaaaaaaaa	aagtgggtgtg	aatggcaata	atgggagtgg	gaatgggaat	82500
gggtgattggg	gctgatgggtg	atgataatgt	taacgggtgga	gatgacaatg	tcactgaaac	82560
cagtgggtgg	gttcatggga	tgacaatatt	gttgatagcg	gaatgggtgg	attagggata	82620
atattgtatt	gatggggaag	acagcgttca	tgggggtgg	gattagcgta	agagttgtag	82680
agtgggtgatg	ttaatggagg	tggctctggtg	ctgatgagga	gatcaatgtt	gatgaagggtg	82740
tgattggggag	tggggatgg	agctgggtgct	gatggaaatg	acactatcaa	tgatgttaat	82800
actgtagcag	agctgacagt	ctcaaaggca	atgttaataa	catggttgca	ccaaccatgt	82860
tatctcaatg	gcgatgttac	tgggtgtcgtg	gagatgacaa	tatcaatggc	aatgttagtg	82920
gtgggtgggtg	aatgatgaat	gcagttgggtg	gtgatgacct	attaatgata	gtagcaaaga	82980
caatgttgtt	gatggagatg	acaacattga	tggagtggt	gatggaagag	ttcgttgttg	83040
gtgttgatgg	tgatgcagat	ggcaattgag	gtagtgtgg	tgggtgtgtt	agcagaggtg	83100
acaaggttga	tggtaatgac	ctttattcat	ctcagagcct	tcatttttct	tcacatttga	83160
ccctcctcat	ttgtatctag	gacccagaca	gttactcttg	gaggcatctt	ccagttggga	83220
aactggagtg	aactaaacaa	tatccaggac	cacaacacca	tttgggaagg	ctgctgcaga	83280
ctggagccca	cactgaagg	aaggtaggga	ggagttagcag	tgcctaaac	caaggtcgtg	83340
ggagcttgg	aatgaggaca	cttcaggacg	ggaagatgcc	accgctggga	taactgggca	83400
aattaattcc	agcaagggat	gtggaacata	acagaatttg	ataatgtaca	gggaagttct	83460
tgctatgggc	taatgaatcc	tgtctggcca	tggctgagag	cccttggttt	tcacatttgt	83520
ctgcgagtga	tgatgacagt	agtgatgggtg	atgaggatga	gttggtactg	atggtgagga	83580

aaatgctgag	aatggtaata	gtgatggtga	taagggtggtg	acagttgtta	aaattatggt	83640
ggtggctgat	ggtgagggtg	gtggttgatg	atggaattgg	tggaaagggtg	gaagcagtaa	83700
tggtaatgat	gttggtagct	gataaagatg	gtgttggtgg	tagtggtgat	tgataaagat	83760
gactgtgatt	atattagtgg	tgggtggtgat	gagattctaa	aagctaactc	cctactacct	83820
aaaaatggca	gcaggaaaaa	aaaatccaga	aatgagtgat	cagcactttt	ctttccagaa	83880
tgcaagaatt	attggtgaac	gaactggctt	ccggccagta	cgcccccaga	ttcggctaga	83940
aagagaacag	cttcgcactg	gaccttcaaa	cacagaggta	tgctcccatg	gcaaggaaag	84000
taatgccctc	ttccactcct	cagatggctc	tggcattttc	agggarcagt	catgtctgat	84060
ctcaagttcc	acacaggctc	catagcaggc	aggggcagtg	gtggctaata	tcccctcctc	84120
tataaatggg	gaaactgagg	ctcaatgatg	gttaaggacc	tgctcaaggt	tacatagagg	84180
ggcagtggtg	atgttaatgg	aggtggtgct	gatgagatca	atgttgataa	tgggtgtgact	84240
gggagtgggg	atggtagctg	gtgctgatgg	aaatgacact	atcaagtatg	ttagtaccac	84300
agcagagggtg	acgatctcaa	aggcagtgtt	aacatggctg	cactaactgt	ctcattggca	84360
atattaatcg	tgtggcagag	atgacagtat	caatggcagt	gttaatgatg	gtggtgaaat	84420
ggtgaatggg	gttggttttc	taaagtctgt	ggtcaaataa	caggaaaatg	tgtacttact	84480
ggatgtgtac	ttcgtgtcag	acacagcagc	aagtccatta	catgaatgac	cttattaaat	84540
ctcctctgga	gctctttggg	atagggacag	ttctccctat	gcttcggatg	aggaaactgg	84600
ggtgaattaa	gaggtgaagt	acttgccca	agtcagacca	ctggtggaag	gcagggcttg	84660
gatgtgattt	gaatttgact	ccaaggctat	ttccagatat	ccattttgtg	gctgccccat	84720
catctcttgc	aactgttcca	gggggtcccc	accattccac	cccggtgcca	agagaagctc	84780
aggtggcatc	tggctttgcc	caggactctt	cgggaggctc	ctgagtcttc	cagggcagaa	84840
gagcttcac	tattctttcc	actgtccctc	tcggacctgg	ccacctcttc	tcttgccctc	84900
cctaggtcat	ccacaactat	ggccatggag	gctacgggct	caccatccac	tggggatgtg	84960
ccttgagggc	agccaagctc	tttgggagaa	tcctggaaga	aaagaaattg	tccagaatgc	85020
caccatccca	cctctgaaga	ctccagtgc	tgctgcctcc	ccccacaaga	actcccttct	85080
cccctcagcc	aatgaatcaa	tgtgtctcct	cataagccat	tgcttctccc	tcacttcttt	85140
cctcaaagaa	gcatgagggtg	agagaaagcc	acaaagtcag	tgcttgaga	agggttcagc	85200
ccaacatggg	gcccctctca	tcactgaaat	ccctctacct	tctctgggtc	tggcattata	85260
aagaacagct	gaggctgtca	ttccatgagt	cttcagaaga	aaggacagct	cagaaaatca	85320
aagaggccaa	ctgcccagag	ccacagaaaa	tggaggataa	ttgaggctaa	gtaacctgat	85380
tacaagttgt	actaacatat	taaaggttct	gaaaagtcct	gcagcaaaga	caactatctg	85440
atgttgttta	acccagtgc	tgctaaacct	atctggctat	ggaactcttt	tgcccagagc	85500
acccatgaat	gccatgacac	aaatctgaga	aaatgctgga	acagattttg	ttgtatctgt	85560
tgtgtttgtt	gtaggaggtt	atacatacaa	ctgggggtgtg	gagggggcag	agaggtgagg	85620
cactgaacta	gtaacacatg	gtgtttgttc	cacatctaga	attccaaatg	gcacgagcta	85680
ttcaccaggt	ggcccccata	gcaccacgta	acctttgagg	agggggccact	ggatggagta	85740
tcccacaagg	aaccccttca	tagagaactg	ttttagtcca	ttttctgttg	cttataacag	85800
aatatctgaa	actggagatt	tttttttttt	ttttttgaga	caggatctca	ctctgtcacc	85860
caggctggtg	tgacgtggca	tgattttggc	tcactgcaac	ctccgcctcc	caggctcaaa	85920
tgatcctccc	tcctcagcca	cccagtagtc	tgggactaca	ggcgcttgct	accatgcccc	85980
gctaattttg	tgtgtgtgtg	tgtgtgtgtg	tgtgttttgt	agagagtgtt	ttgtagagac	86040
tgggtttggc	catgttgtcc	aggctggcgt	tgaactcctg	ggatcaagtg	atcctcctgc	86100
ctcagcctcc	aaagtgggtg	gattatagtc	ataagccacc	acgcctggcg	gaaactgtgg	86160
aattaataga	gaaaaggaat	ttatttatta	ccgttataga	gtctgagaag	tccaaggttg	86220
agggggccaca	tctggtgaga	gccttctctc	tggctggtgc	agaggtgggg	actctctgca	86280
gagtcaccag	gaggcttagg	gcatcacgtg	gtgagggggc	tgattgtgct	aatgtgctag	86340
ctcagctctg	tcccttgtct	tagaaagcca	ccattttcct	tcccaagatg	accattaat	86400
ccattaacct	aataacccat	taattgataa	atggattaat	ccatttatga	gagcagcgct	86460
cttaggatcc	aatcacctct	taaaggcgcc	acctctccag	accaccacta	aggtggtgga	86520
ctaaggacta	agtctcaacg	tgagttttgg	cagggacgtt	taagcaatag	caagaactaa	86580
actcaccaag	ca					86592

<210> 2

<211> 1573

<212> DNA

<213> Homo sapiens

<220>

<221> 5'UTR

<222> 1..143

2002247939 08 Oct 2004

<220>
 <221> CDS
 <222> 144..1187

<220>
 <221> 3'UTR
 <222> 1188..1573

<220>
 <221> polyA_signal
 <222> 1549..1554

<400> 2

```

tgcactccag tccgggctgg cggacagagg gctggaaaca agacgctcca gaatcaggag      60
cttccctca ggaaatagca tcctgtgtcc ccgactgca gttgtctggt ctctccagca      120
gtttgttact tccggctgct gca atg cgt gtg gtg gtg att gga gca gga gtc      173
               Met Arg Val Val Val Ile Gly Ala Gly Val
                   1             5             10
atc ggg ctg tcc acc gcc ctg tgc atc cat gag cgc tac cac tca gtc      221
Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
                   15             20             25
ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctg      269
Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
                   30             35             40
acc acc acc gac gtg gct gcc ggc ctg tgg cag ccc tac ctt tct gac      317
Thr Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp
                   45             50             55
ccc aac aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat      365
Pro Asn Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr
                   60             65             70
ctc ctg agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc      413
Leu Leu Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe
                   75             80             85             90
cta atc tcg ggc tac aac ctg ttc cat gaa gcc att ccg gac cct tcc      461
Leu Ile Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser
                   95             100             105
tgg aag gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg      509
Trp Lys Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu
                   110             115             120
gat atg ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att      557
Asp Met Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile
                   125             130             135
ctg gag gga aag aac tat cta cag tgg ctg act gaa agg tta act gag      605
Leu Glu Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu
                   140             145             150
agg gga gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg      653
Arg Gly Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val
                   155             160             165             170
gca aga gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct      701
Ala Arg Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala
                   175             180             185
ggg gcg cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc      749
Gly Ala Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile
                   190             195             200
atg aag gtg gac gcc cct tgg atg aag cac ttc att ctc acc cat gac      797
Met Lys Val Asp Ala Pro Trp Met Lys His Phe Ile Leu Thr His Asp
                   205             210             215
cca gag aga ggc atc tac aat tcc ccg tac atc atc cca ggg acc cag      845
Pro Glu Arg Gly Ile Tyr Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln
                   220             225             230
aca gtt act ctt gga ggc atc ttc cag ttg gga aac tgg agt gaa cta      893

```

2002247939 08 Oct 2004

```

Thr Val Thr Leu Gly Gly Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu
235                240                245                250
aac aat atc cag gac cac aac acc att tgg gaa ggc tgc tgc aga ctg      941
Asn Asn Ile Gln Asp His Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu
                255                260                265
gag ccc aca ctg aag aat gca aga att att ggt gaa cga act ggc ttc      989
Glu Pro Thr Leu Lys Asn Ala Arg Ile Ile Gly Glu Arg Thr Gly Phe
                270                275                280
cgg cca gta cgc ccc cag att cgg cta gaa aga gaa cag ctt cgc act      1037
Arg Pro Val Arg Pro Gln Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr
                285                290                295
gga cct tca aac aca gag gtc atc cac aac tat ggc cat gga ggc tac      1085
Gly Pro Ser Asn Thr Glu Val Ile His Asn Tyr Gly His Gly Gly Tyr
                300                305                310
ggg ctc acc atc cac tgg gga tgt gcc ctg gag gca gcc aag ctc ttt      1133
Gly Leu Thr Ile His Trp Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe
                315                320                325                330
ggg aga atc ctg gaa gaa aag aaa ttg tcc aga atg cca cca tcc cac      1181
Gly Arg Ile Leu Glu Glu Lys Lys Leu Ser Arg Met Pro Pro Ser His
                335                340                345
ctc tga agactccagt gactgctgcc tccccccaca agaactccct tctccctca      1237
Leu *
gccaatgaat caatgtgctc cttcataagc cattgcttct ccctcacttc tttcctcaaa      1297
gaagcatgag gtgagagaaa gccacaaagt cagtgcctgg agaagggttc agcccaacat      1357
ggggccctc tcatcactga aatccctcta ctttctctgg gtctggcatt ataaagaaca      1417
gctgaggctg tcattccatg agtcttcaga agaaaggaca gctcagaaaa tcaaagaggc      1477
caactgcca gagccacaga aaatggagga taattgaggc taagtaacct gattacaagt      1537
tgtactaaca tattaaggt tctgaaaagt cctgca                                1573

```

<210> 3

<211> 1691

<212> DNA

<213> Homo sapiens

<220>

<221> 5'UTR

<222> 1..143

<220>

<221> CDS

<222> 144..1187

<220>

<221> 3'UTR

<222> 1188..1691

<400> 3

```

tgcactccag tccgggctgg cggacagagg gctggaaaca agacgctcca gaatcaggag      60
cttccctca ggaaatagca tcctgtgtcc ccgcactgca gttgtctggt ctctccagca      120
gtttggtact tccggctgct gca atg cgt gtg gtg gtg att gga gca gga gtc      173
                Met Arg Val Val Val Ile Gly Ala Gly Val
                1                5                10
atc ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc      221
Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
                15                20                25
ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc      269
Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
                30                35                40
acc acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac      317
Thr Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp
                45                50                55

```

2002247939 08 Oct 2004

ccc aac aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat	365
Pro Asn Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr	
60 65 70	
ctc ctg agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc	413
Leu Leu Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe	
75 80 85 90	
cta atc tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc	461
Leu Ile Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser	
95 100 105	
tgg aag gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg	509
Trp Lys Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu	
110 115 120	
gat atg ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att	557
Asp Met Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile	
125 130 135	
ctg gag gga aag aac tat cta cag tgg ctg act gaa agg tta act gag	605
Leu Glu Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu	
140 145 150	
agg gga gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg	653
Arg Gly Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val	
155 160 165 170	
gca aga gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct	701
Ala Arg Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala	
175 180 185	
ggg gcg cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc	749
Gly Ala Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile	
190 195 200	
atg aag gtg gac gcc cct tgg atg aag cac ttc att ctc acc cat gac	797
Met Lys Val Asp Ala Pro Trp Met Lys His Phe Ile Leu Thr His Asp	
205 210 215	
cca gag aga ggc atc tac aat tcc ccg tac atc atc cca ggg acc cag	845
Pro Glu Arg Gly Ile Tyr Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln	
220 225 230	
aca gtt act ctt gga ggc atc ttc cag ttg gga aac tgg agt gaa cta	893
Thr Val Thr Leu Gly Gly Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu	
235 240 245 250	
aac aat atc cag gac cac aac acc att tgg gaa ggc tgc tgc aga ctg	941
Asn Asn Ile Gln Asp His Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu	
255 260 265	
gag ccc aca ctg aag aat gca aga att att ggt gaa cga act ggc ttc	989
Glu Pro Thr Leu Lys Asn Ala Arg Ile Ile Gly Glu Arg Thr Gly Phe	
270 275 280	
cgg cca gta cgc ccc cag att cgg cta gaa aga gaa cag ctt cgc act	1037
Arg Pro Val Arg Pro Gln Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr	
285 290 295	
gga cct tca aac aca gag gtc atc cac aac tat ggc cat gga ggc tac	1085
Gly Pro Ser Asn Thr Glu Val Ile His Asn Tyr Gly His Gly Gly Tyr	
300 305 310	
ggg ctc acc atc cac tgg gga tgt gcc ctg gag gca gcc aag ctc ttt	1133
Gly Leu Thr Ile His Trp Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe	
315 320 325 330	
ggg aga atc ctg gaa gaa aag aaa ttg tcc aga atg cca cca tcc cac	1181
Gly Arg Ile Leu Glu Lys Lys Leu Ser Arg Met Pro Pro Ser His	
335 340 345	
ctc tga agactccagt gactgctgcc tccccccaca agaactccct tctccctca	1237
Leu *	
gccaatgaat caatgtgctc cttcataagc cattgcttct ccctcacttc tttcctcaaa	1297
gaagcatgag gtgagagaaa gccacaaagt cagtgcctgg agaagggttc agcccaacat	1357
ggggccctc tcatcactga aatccctcta ccttctctgg gtctggcatt ataaagaaca	1417
gctgaggctg tcattccatg agtcttcaga agaaaggaca gctcagaaaa tcaaagaggc	1477
caactgcca gagccacaga aaatggagga taattgaggc taagtaacct gattacaagt	1537

2002247939 08 Oct 2004

tgtactaaca	tattaaaggt	tctgaaaagt	cctgcagcaa	agacaactat	ctgatgttgt	1597
ttaacccagt	gcttgctaaa	cctatctggc	tatggaactc	ttttgcccag	agcaccatg	1657
aatgccatga	cacaaatctg	agaaaatgct	ggaa			1691

<210> 4
 <211> 2620
 <212> DNA
 <213> Homo sapiens

<220>
 <221> 5'UTR
 <222> 1..1155

<220>
 <221> CDS
 <222> 1156..1818

<220>
 <221> 3'UTR
 <222> 1819..2620

<400> 4

gaaacccacg	cagcctcctg	gattcttccc	cgccctccc	tctgtcctgg	ggctgtgacc	60
tcctccatgt	tattcacagg	gtctcagcac	gattcatctc	aaagcagcga	aacaagcact	120
ggcctcagaa	gaagcaatat	taaaacagtt	acaactcatc	tagcgcacag	acacccaact	180
gacaccctgt	tcctccagtc	ataacaacaa	ctacagcttt	gattgaacaa	gagactgagt	240
ttggtaactt	tctcctaata	aaaagatcac	tgactatgga	ctgcttctgg	tggggttacg	300
aaaccgcaac	ctcatgtgcc	tgcatttcct	gaaaagacat	tttgaatgag	gaagggcctg	360
gagtcctgct	gcttgctgct	ctgggatacg	ggagcaaaga	gccacgcac	ctcatggccc	420
acacaggcgt	cacctccagt	ctctccttgg	cctcatctcc	ccagcgtcct	ggaatggcat	480
cgggctggcc	cagggagccc	ctgtcctgtg	cctctccttt	cccctcaggg	gctgccaggc	540
tgaccacccc	caccgcaggc	caggcctaca	gtgccccatg	gaacgtcctg	accctcccc	600
agggtggcag	caggaagaag	gaagaaagg	gatcctctcc	agctggccag	agagacagac	660
cttcttgtgc	tcataacccc	tccaagaatg	cctgcccccc	ctccttcccc	caaggcctgt	720
ccacaggggc	ttgagatcag	ccagaaaagt	caggcaactt	ttcagggact	gggagcgagg	780
tctccgggct	gggcttgggt	ccagtcctct	tgggcagtgc	agtgccgagc	cccacccttc	840
aagccgtgcc	ctgtccatag	ctccagactt	tgacctgca	ctccagtcgg	ggctggcgga	900
cagagggtcg	gaaacaagac	gctccagaat	caggagcttc	ccctcaggaa	atagcatcct	960
gtgtccccgc	actgcagttg	tctggtctct	ccagcagttt	ggtaacttccg	atgaagagct	1020
tgtgtctcca	gaggcaaagt	atgggggaag	aggggaagaga	gaagaccaag	ggtccctgag	1080
aggggctgtc	ccctaagccc	cagtatccaa	gctcgggctc	gaagctggaa	ggagaattgc	1140
ctagaggctg	ctgca atg	cgt gtg gtg	gtg att gga	gca gga gtc	atc ggg	1191
	Met Arg Val Val	Val Ile Gly	Ala Gly Val	Ile Gly		
	1	5	10			
ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc ctg cag	1239					
Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln						
15	20	25				
cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc acc acc	1287					
Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr						
30	35	40				
acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac ccc aac	1335					
Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn						
45	50	55	60			
aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat ctc ctg	1383					
Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu						
65	70	75				
agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc cta atc	1431					
Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile						
80	85	90				
tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc tgg aag	1479					
Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys						

2002247939 08 Oct 2004

```

          95              100              105
gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg gat atg      1527
Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met
          110              115              120
ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att ctg gag      1575
Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu
          125              130              135              140
gga aag aac tat cta cag tgg ctg act gaa agg tta act gag agg gga      1623
Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly
          145              150              155
gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg gca aga      1671
Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg
          160              165              170
gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct ggg gcg      1719
Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala
          175              180              185
cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc atg aag      1767
Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys
          190              195              200
gac cca gac agt tac tct tgg agg cat ctt cca gtt ggg aaa ctg gag      1815
Asp Pro Asp Ser Tyr Ser Trp Arg His Leu Pro Val Gly Lys Leu Glu
          205              210              215              220
tga actaaacaat atccaggacc acaacaccat ttgggaaggc tgctgcagac      1868
*
tgagagcccac actgaagaat gcaagaatta ttggtgaacg aactggcttc cggccagtac      1928
gccccagat tcggctagaa agagaacagc ttcgcactgg accttcaaac acagagggtca      1988
tccacaacta tggccatgga ggctacgggc tcaccatcca ctggggatgt gccctggagg      2048
cagccaagct ctttgggaga atcctggaag aaaagaaatt gtccagaatg ccaccatccc      2108
acctctgaag actccagtga ctgctgcctc cccccacaag aactcccttc tcccctcagc      2168
caatgaatca atgtgctcct tcataagcca ttgcttctcc ctcaacttctt tcctcaaaga      2228
agcatgaggt gagagaaaagc cacaaaagtca gtgcctggag aagggttcag cccaacatgg      2288
ggccccctctc atcactgaaa tccctctacc ttctctgggt ctggcattat aaagaacagc      2348
tgaggctgtc attccatgag tcttcagaag aaaggacagc tcagaaaatc aaagaggcca      2408
actgcccaga gccacagaaa atggaggata attgaggcta agtaacctga ttacaagttg      2468
tactaacata ttaaaggttc tgaaaagtcc tgcagcaaag acaactatct gatgttgttt      2528
aaccacgtgc ttgctaaaacc tatctggcta tggaactctt ttgcccagag caccatgaa      2588
tgccatgaca caaatctgag aaaatgctgg aa                                2620

```

<210> 5

<211> 1576

<212> DNA

<213> Homo sapiens

<220>

<221> 5'UTR

<222> 1..143

<220>

<221> CDS

<222> 144..380

<220>

<221> 3'UTR

<222> 381..1576

<400> 5

```

tgcaactccag tccgggctgg cggacagagg gctggaaaaca agacgctcca gaatcaggag      60
cttccccctca ggaaatagca tcctgtgtcc ccgcactgca gttgtctggt ctctccagca      120
gtttggtact tccggctgct gca atg cgt gtg gtg gtg att gga gca gga gtc      173
Met Arg Val Val Val Ile Gly Ala Gly Val

```

1

5

10

2002247939 08 Oct 2004

```

atc ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc      221
Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
              15                      20                      25
ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc      269
Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
              30                      35                      40
acc acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac      317
Thr Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp
              45                      50                      55
ccc aac aac cca cag gag gcg acc ctt cct gga agg aca cag ttc tgg      365
Pro Asn Asn Pro Gln Glu Ala Thr Leu Pro Gly Arg Thr Gln Phe Trp
              60                      65                      70
gat ttc gga agc tga cccccagaga gctggatatg ttcccagatt acggctatgg      420
Asp Phe Gly Ser *
75
ctggttccac acaagcctaa ttctggaggg aaagaactat ctacagtggc tgactgaaag      480
gttaactgag aggggagtg agttcttcca gcggaaagtg gagtcttttg aggaggtggc      540
aagagaaggc gcagacgtga ttgtcaactg cactggggta tgggctgggg cgctacaacg      600
agaccccctg ctgcagccag gccgggggca gatcatgaag gtggacgccc cttggatgaa      660
gcacttcatt ctcacccatg acccagagag aggcattctac aattccccgt acatcatccc      720
agggacccag acagttactc ttggaggcat cttccagttg ggaaactgga gtgaactaaa      780
caatatccag gaccacaaca ccatttggaagg aggctgctgc agactggagc ccacactgaa      840
gaatgcaaga attattggtg aacgaactgg cttccggcca gtacgcccc agattcggct      900
agaaagagaa cagcttcgca ctggaccttc aaacacagag gtcattccaca actatggcca      960
tgagggtctac gggctcacca tccactgggg atgtgccctg gaggcagcca agctcttttg      1020
gagaatcctg gaagaaaaga aattgtccag aatgccacca tcccacctct gaagactcca      1080
gtgactgctg cctcccccca caagaactcc cttctccctc cagccaatga atcaatgtgc      1140
tccttcataa gccattgctt ctccctcact tctttctca aagaagcatg aggtgagaga      1200
aagccacaaa gtcagtgcct ggagaagggt tcagcccaac atggggcccc tctcatcact      1260
gaaatccctc taccttctct ggggtctggca ttataaagaa cagctgaggc tgtcattcca      1320
tgagtcttca gaagaaagga cagctcagaa aatcaaagag gccaaactgcc cagagccaca      1380
gaaaatggag gataattgag gctaagtaac ctgattacaa gttgtactaa catattaaag      1440
gttctgaaaa gtctgcagc aaagacaact atctgatgtt gtttaaccca gtgcttgcta      1500
aacctatctg gctatggaac tcttttgccc agagcaccca tgaatgccat gacacaaatc      1560
tgagaaaatg ctggaa
1576

```

<210> 6
 <211> 1345
 <212> DNA
 <213> Homo sapiens

<220>
 <221> 5'UTR
 <222> 1..113

<220>
 <221> CDS
 <222> 114..959

<220>
 <221> 3'UTR
 <222> 960..1345

<220>
 <221> polyA_signal
 <222> 1321..1326

<400> 6
 gaaacccacg cagcctcctg gattcttccc cgtccctccc tctgtcctgg ggctgtgacc 60
 tcctccatgt tattcacagg gtctcagcac gattcatctc aaaggctgct gca atg 116
 Met

2002247939

1																					
cgt	gtg	gtg	gtg	att	gga	gca	gga	gtc	atc	ggg	ctg	tcc	acc	gcc	ctc	164					
Arg	Val	Val	Val	Ile	Gly	Ala	Gly	Val	Ile	Gly	Leu	Ser	Thr	Ala	Leu						
5								10				15									
tgc	atc	cat	gag	cgc	tac	cac	tca	gtc	ctg	cag	cca	ctg	gac	ata	aag	212					
Cys	Ile	His	Glu	Arg	Tyr	His	Ser	Val	Leu	Gln	Pro	Leu	Asp	Ile	Lys						
20								25				30									
gtc	tac	gcg	gac	cgc	ttc	acc	cca	ctc	acc	acc	acc	gac	gtg	gct	gcc	260					
Val	Tyr	Ala	Asp	Arg	Phe	Thr	Pro	Leu	Thr	Thr	Thr	Asp	Val	Ala	Ala						
35								40				45									
ggc	ctc	tgg	cag	ccc	tac	ctt	tct	gac	ccc	aac	aac	cca	cag	gag	gcg	308					
Gly	Leu	Trp	Gln	Pro	Tyr	Leu	Ser	Asp	Pro	Asn	Asn	Pro	Gln	Glu	Ala						
50				55				60				65									
gac	tgg	agc	caa	cag	acc	ttt	gac	tat	ctc	ctg	agc	cat	gtc	cat	tct	356					
Asp	Trp	Ser	Gln	Gln	Thr	Phe	Asp	Tyr	Leu	Ser	His	Val	His	Ser							
				70				75				80									
ccc	aac	gct	gaa	aac	ctg	ggc	ctg	ttc	cta	atc	tcg	ggc	tac	aac	ctc	404					
Pro	Asn	Ala	Glu	Asn	Leu	Gly	Leu	Phe	Leu	Ile	Ser	Gly	Tyr	Asn	Leu						
				85				90				95									
ttc	cat	gaa	gcc	att	ccg	gtg	gca	aga	gaa	ggc	gca	gac	gtg	att	gtc	452					
Phe	His	Glu	Ala	Ile	Pro	Val	Ala	Arg	Glu	Gly	Ala	Asp	Val	Ile	Val						
100								105				110									
aac	tgc	act	ggg	gta	tgg	gct	ggg	gcg	cta	caa	cga	gac	ccc	ctg	ctg	500					
Asn	Cys	Thr	Gly	Val	Trp	Ala	Gly	Ala	Leu	Gln	Arg	Asp	Pro	Leu	Leu						
115				120				125													
cag	cca	ggc	cgg	ggg	cag	atc	atg	aag	gtg	gac	gcc	cct	tgg	atg	aag	548					
Gln	Pro	Gly	Arg	Gly	Gln	Ile	Met	Lys	Val	Asp	Ala	Pro	Trp	Met	Lys						
130				135				140				145									
cac	ttc	att	ctc	acc	cat	gac	cca	gag	aga	ggc	atc	tac	aat	tcc	ccg	596					
His	Phe	Ile	Leu	Thr	His	Asp	Pro	Glu	Arg	Gly	Ile	Tyr	Asn	Ser	Pro						
				150				155				160									
tac	atc	atc	cca	ggg	acc	cag	aca	gtt	act	ctt	gga	ggc	atc	ttc	cag	644					
Tyr	Ile	Ile	Pro	Gly	Thr	Gln	Thr	Val	Thr	Leu	Gly	Gly	Ile	Phe	Gln						
165								170				175									
ttg	gga	aac	tgg	agt	gaa	cta	aac	aat	atc	cag	gac	cac	aac	acc	att	692					
Leu	Gly	Asn	Trp	Ser	Glu	Leu	Asn	Asn	Ile	Gln	Asp	His	Asn	Thr	Ile						
180								185				190									
tgg	gaa	ggc	tgc	tgc	aga	ctg	gag	ccc	aca	ctg	aag	aat	gca	aga	att	740					
Trp	Glu	Gly	Cys	Cys	Arg	Leu	Glu	Pro	Thr	Leu	Lys	Asn	Ala	Arg	Ile						
195				200				205													
att	ggt	gaa	cga	act	ggc	ttc	cgg	cca	gta	cgc	ccc	cag	att	cgg	cta	788					
Ile	Gly	Glu	Arg	Thr	Gly	Phe	Arg	Pro	Val	Arg	Pro	Gln	Ile	Arg	Leu						
210				215				220				225									
gaa	aga	gaa	cag	ctt	cgc	act	gga	cct	tca	aac	aca	gag	gtc	atc	cac	836					
Glu	Arg	Glu	Gln	Leu	Arg	Thr	Gly	Pro	Ser	Asn	Thr	Glu	Val	Ile	His						
				230				235				240									
aac	tat	ggc	cat	gga	ggc	tac	ggg	ctc	acc	atc	cac	tgg	gga	tgt	gcc	884					
Asn	Tyr	Gly	His	Gly	Gly	Tyr	Gly	Leu	Thr	Ile	His	Trp	Gly	Cys	Ala						
245				250				255				260									
ctg	gag	gca	gcc	aag	ctc	ttt	ggg	aga	atc	ctg	gaa	gaa	aag	aaa	ttg	932					
Leu	Glu	Ala	Ala	Lys	Leu	Phe	Gly	Arg	Ile	Leu	Glu	Glu	Lys	Lys	Leu						
260				265				270													
tcc	aga	atg	cca	cca	tcc	cac	ctc	tga	agactccagtgactgctgcc							979					
Ser	Arg	Met	Pro	Pro	Ser	His	Leu	*													
275				280																	
tccccccaca	agaactccct				tctccccctca				gccaatgaat				caatgtgctc				cttcataagc	1039			
cattgcttct				ccctcacttc				tttctctcaaa				gaagcatgag				gtgagagaaa				gccacaaagt	1099
cagtgccctgg				agaagggttc				atgcccaacat				ggggccctc				tcattcactga				aatccctcta	1159
ccttctctgg				gtctggcatt				ataaagaaca				gctgaggctg				tcttccatg				agtcttcaga	1219
agaaaggacag				gctcagaaaa				tcaaagaggc				caactgcccc				gagccacaga				aaatggagga	1279

2002247939 08 Oct 2004

taattgaggc taagtaacct gattacaagt tgtactaaca tattaaaggt tctgaaaagt 1339
cctgca 1345

<210> 7
<211> 347
<212> PRT
<213> Homo sapiens

<400> 7
Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
1 5 10 15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
20 25 30
Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
35 40 45
Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
50 55 60
Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His
65 70 75 80
Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn
85 90 95
Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys Asp Thr Val Leu
100 105 110
Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met Phe Pro Asp Tyr
115 120 125
Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu Gly Lys Asn Tyr
130 135 140
Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly Val Lys Phe Phe
145 150 155 160
Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg Glu Gly Ala Asp
165 170 175
Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp
180 185 190
Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro
195 200 205
Trp Met Lys His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr
210 215 220
Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly
225 230 235 240
Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His
245 250 255
Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn
260 265 270
Ala Arg Ile Ile Gly Glu Arg Thr Gly Phe Arg Pro Val Arg Pro Gln
275 280 285
Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu
290 295 300
Val Ile His Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp
305 310 315 320
Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu
325 330 335
Lys Lys Leu Ser Arg Met Pro Pro Ser His Leu
340 345

<210> 8
<211> 220
<212> PRT
<213> Homo sapiens

<400> 8
Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala

2002247939 08 Oct 2004

1	Leu	Cys	Ile	His	5	Glu	Arg	Tyr	His	10	Ser	Val	Leu	Gln	Pro	15	Leu	Asp	Ile
				20						25						30			
	Lys	Val	Tyr	Ala	Asp	Arg	Phe	Thr	Pro	Leu	Thr	Thr	Thr	Thr	Thr	Asp	Val	Ala	
			35					40						45					
	Ala	Gly	Leu	Trp	Gln	Pro	Tyr	Leu	Ser	Asp	Pro	Asn	Asn	Pro	Gln	Glu			
			50				55					60							
	Ala	Asp	Trp	Ser	Gln	Gln	Thr	Phe	Asp	Tyr	Leu	Leu	Ser	His	Val	His			
			65			70					75					80			
	Ser	Pro	Asn	Ala	Glu	Asn	Leu	Gly	Leu	Phe	Leu	Ile	Ser	Gly	Tyr	Asn			
					85					90					95				
	Leu	Phe	His	Glu	Ala	Ile	Pro	Asp	Pro	Ser	Trp	Lys	Asp	Thr	Val	Leu			
				100					105						110				
	Gly	Phe	Arg	Lys	Leu	Thr	Pro	Arg	Glu	Leu	Asp	Met	Phe	Pro	Asp	Tyr			
			115					120						125					
	Gly	Tyr	Gly	Trp	Phe	His	Thr	Ser	Leu	Ile	Leu	Glu	Gly	Lys	Asn	Tyr			
			130				135					140							
	Leu	Gln	Trp	Leu	Thr	Glu	Arg	Leu	Thr	Glu	Arg	Gly	Val	Lys	Phe	Phe			
						150					155					160			
	Gln	Arg	Lys	Val	Glu	Ser	Phe	Glu	Glu	Val	Ala	Arg	Glu	Gly	Ala	Asp			
					165					170					175				
	Val	Ile	Val	Asn	Cys	Thr	Gly	Val	Trp	Ala	Gly	Ala	Leu	Gln	Arg	Asp			
				180					185					190					
	Pro	Leu	Leu	Gln	Pro	Gly	Arg	Gly	Gln	Ile	Met	Lys	Asp	Pro	Asp	Ser			
			195					200					205						
	Tyr	Ser	Trp	Arg	His	Leu	Pro	Val	Gly	Lys	Leu	Glu							
			210				215					220							

<210> 9
 <211> 78
 <212> PRT
 <213> Homo sapiens

<400> 9

Met	Arg	Val	Val	Val	Ile	Gly	Ala	Gly	Val	Ile	Gly	Leu	Ser	Thr	Ala
1				5					10					15	
Leu	Cys	Ile	His	Glu	Arg	Tyr	His	Ser	Val	Leu	Gln	Pro	Leu	Asp	Ile
			20					25					30		
Lys	Val	Tyr	Ala	Asp	Arg	Phe	Thr	Pro	Leu	Thr	Thr	Thr	Asp	Val	Ala
			35				40					45			
Ala	Gly	Leu	Trp	Gln	Pro	Tyr	Leu	Ser	Asp	Pro	Asn	Asn	Pro	Gln	Glu
			50			55					60				
Ala	Thr	Leu	Pro	Gly	Arg	Thr	Gln	Phe	Trp	Asp	Phe	Gly	Ser		
					70					75					

<210> 10
 <211> 281
 <212> PRT
 <213> Homo sapiens

<400> 10

Met	Arg	Val	Val	Val	Ile	Gly	Ala	Gly	Val	Ile	Gly	Leu	Ser	Thr	Ala
1				5					10					15	
Leu	Cys	Ile	His	Glu	Arg	Tyr	His	Ser	Val	Leu	Gln	Pro	Leu	Asp	Ile
			20					25					30		
Lys	Val	Tyr	Ala	Asp	Arg	Phe	Thr	Pro	Leu	Thr	Thr	Thr	Asp	Val	Ala
			35				40					45			
Ala	Gly	Leu	Trp	Gln	Pro	Tyr	Leu	Ser	Asp	Pro	Asn	Asn	Pro	Gln	Glu
			50			55					60				
Ala	Asp	Trp	Ser	Gln	Gln	Thr	Phe	Asp	Tyr	Leu	Leu	Ser	His	Val	His
					70					75					80

2002247939 08 Oct 2004

```

Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn
      85                      90                      95
Leu Phe His Glu Ala Ile Pro Val Ala Arg Glu Gly Ala Asp Val Ile
      100                      105                      110
Val Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp Pro Leu
      115                      120                      125
Leu Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro Trp Met
      130                      135                      140
Lys His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr Asn Ser
      145                      150                      155                      160
Pro Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly Ile Phe
      165                      170                      175
Gln Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His Asn Thr
      180                      185                      190
Ile Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn Ala Arg
      195                      200                      205
Ile Ile Gly Glu Arg Thr Gly Phe Arg Pro Val Arg Pro Gln Ile Arg
      210                      215                      220
Leu Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu Val Ile
      225                      230                      235                      240
His Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp Gly Cys
      245                      250                      255
Ala Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu Lys Lys
      260                      265                      270
Leu Ser Arg Met Pro Pro Ser His Leu
      275                      280

```

<210> 11

<211> 456

<212> DNA

<213> Homo sapiens

<220>

<221> allele

<222> 152

<223> 99-16105-152 : polymorphic base A or G

<220>

<221> misc_binding

<222> 140..164

<223> 99-16105-152.probe

<220>

<221> primer_bind

<222> 133..151

<223> 99-16105-152.mis

<220>

<221> primer_bind

<222> 153..171

<223> 99-16105-152.mis complement

<220>

<221> primer_bind

<222> 1..19

<223> 99-16105.pu

<220>

<221> primer_bind

<222> 437..456

<223> 99-16105.rp complement

2002247939 08 Oct 2004

```

<400> 11
cgctttgttg tattctttgt tatttatcca ttttgccaaa ttatctgcaa gtagaaatat    60
cgaaataaga agctcttttag caattttactt tggatattgg ttttcttttg aaggacagtt    120
attaaaatag cttgtaggat tactcatttt crtthtttctt ctttttaaat ataaagcaat    180
gtcatcactt ttttcctgt attatatttc tcctcaataa ttgatatgct acattaaagg    240
aacacaaaaat ggtcttaatt atgcaataat gatcaaggca aagagtgttt cctgggaact    300
aatggttgcc tgagaggagg tgatggcttg aggtccagct gggtattaag ccgcaggaaa    360
tgctgcaggc caagatttgt attatttctc tgagatgaaa atgaacccaa aaaaaggcaa    420
aatgggtttt tctccactaa tgggtaaaat gaactc                                456

```

```

<210> 12
<211> 463
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> 215
<223> 99-5919-215 : polymorphic base A or G

```

```

<220>
<221> misc_binding
<222> 203..227
<223> 99-5919-215.probe

```

```

<220>
<221> primer_bind
<222> 196..214
<223> 99-5919-215.mis

```

```

<220>
<221> primer_bind
<222> 216..234
<223> 99-5919-215.mis complement

```

```

<220>
<221> primer_bind
<222> 11..29
<223> 99-5919.pu

```

```

<220>
<221> primer_bind
<222> 445..465
<223> 99-5919.rp complement

```

```

<400> 12
tttctcttga ctacagcaat gcagatttca attctgccat tgaattccca gacatattcg    60
tcatcccat tttcatcccc caccacctg ccattttctt cgtgttaact tgttttcctg    120
actcacagaa atcacctttt cctgtataca tttttaggat gtcagacttt attctaataga    180
tttctcctag ttgcccccca aaattgtatt ctacrgtgtg attttaaagc tgaattttca    240
agatgatatt tcatatctat attttcacia gcttttcttc tatgaatgtt attgtcagct    300
gtcagggtgt gagatggtac ttgatactac attctttcca agctgttgcc tgaatcggtt    360
taagacaaaag tcattactag gctgtaaaact gttgctctgc aaaattgagc agcacgtatt    420
taaccactca tacttcttag ctctccaaca ctttgagtcr ata                                463

```

```

<210> 13
<211> 742
<212> DNA
<213> Homo sapiens

```

2002247939 08 Oct 2004

<220>
<221> 5'UTR
<222> 1..46

<220>
<221> CDS
<222> 47..508

<220>
<221> 3'UTR
<222> 509..742

<220>
<221> polyA_signal
<222> 718..723

<220>
<221> allele
<222> 21
<223> 8-135-112 : polymorphic base C or T

<220>
<221> allele
<222> 75
<223> 8-135-166 : polymorphic base A or C

<220>
<221> allele
<222> 135
<223> 99-16038-118 : polymorphic base A or G

<220>
<221> allele
<222> 194
<223> 8-137-152 : polymorphic base G or T

<220>
<221> allele
<222> 224
<223> 8-137-182 : polymorphic base A or G

<220>
<221> allele
<222> 390
<223> 8-130-220 : polymorphic base A or C

<220>
<221> allele
<222> 406
<223> 8-130-236 : polymorphic base A or G

<220>
<221> allele
<222> 578
<223> 8-131-199 : polymorphic base A or C

<220>
<221> allele
<222> 641
<223> 8-132-97 : polymorphic base C or T

```
<220>  
<221> allele  
<222> 723  
<223> 8-132-179 : polymorphic base A or T
```

```
<210> 14
<211> 153
<212> PRT
<213> Homo sapiens
```

```
<400> 14
Met Leu Glu Lys Leu Met Gly Ala Asp Xaa Leu Gln Leu Phe Arg Ser
1          5          10          15
Arg Tyr Thr Leu Gly Lys Ile Tyr Phe Ile Gly Phe Gln Xaa Ser Ile
          20          25          30
Leu Leu Ser Lys Ser Glu Asn Ser Leu Asn Ser Ile Ala Lys Glu Thr
          35          40          45
```

2002247939 08 Oct 2004

Glu Xaa Gly Arg Glu Thr Val Thr Arg Lys Glu Xaa Trp Lys Arg Arg
 50 55 60
 His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser
 65 70 75 80
 Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu
 85 90 95
 Glu Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu
 100 105 110
 Phe Leu Xaa Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg
 115 120 125
 Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His
 130 135 140
 Lys Glu Ile Thr Ser Thr Lys Ala Glu
 145 150

<210> 15
 <211> 476
 <212> DNA
 <213> Homo sapiens

<400> 15
 cat gag gac ggc tat ttg gaa atg gca cag agg cat tta cag aga tca 48
 His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser
 1 5 10 15
 tta tgt cct tgg gtc tct tac ctt cct cag ccc tat gca gag ctt gaa 96
 Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu
 20 25 30
 gaa gta agc agc cat gtt gga aaa gtc ttc atg gca aga aac tat gag 144
 Glu Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu
 35 40 45
 ttc ctt gcc tat gag gcc tct aag gac cgc agg cag cct cta gaa cga 192
 Phe Leu Ala Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg
 50 55 60
 atg tgg acc tgc aac tac aac cag caa aaa gac cag tca tgc aac cac 240
 Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His
 65 70 75 80
 aag gaa ata act tct acc aaa gct gaa tga gtttggaagc agattcttcc 290
 Lys Glu Ile Thr Ser Thr Lys Ala Glu *
 85 90
 cagccaatcc ttctgatgac aatgtagtct ggccaacatc ttactggac tctgacggac 350
 tctgtgtctg ggaccagct gataacacgt ggtgatggga ttgtatttgc aactctctg 410
 tcagtaagt ataaaatgcc atttctatgc acccacctgg cctgtgtgac tgggagaatc 470
 tctctt 476

<210> 16
 <211> 89
 <212> PRT
 <213> Homo sapiens

<400> 16
 His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser
 1 5 10 15
 Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu
 20 25 30
 Glu Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu
 35 40 45
 Phe Leu Ala Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg
 50 55 60
 Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His
 65 70 75 80
 Lys Glu Ile Thr Ser Thr Lys Ala Glu

2002247939 08 Oct 2004

85

<210> 17
 <211> 1633
 <212> DNA
 <213> Homo sapiens

<400> 17

```

ttgggggtcca ttgcaacccg aggcgagact agagttccca agcgagaagg gaagaggcag      60
tgggtgcacg tggaaggcgg acagagggct ggaaacaaga cgctccagaa tcaggagctt      120
cccctcagga aatagcatcc tgtgtccccg cactgcagtt gtctggctctc tccagcagtt      180
tggtacttcc ggctgctgca atg cgt gtg gtg gtg att gga gca gga gtc atc      233
          Met Arg Val Val Val Ile Gly Ala Gly Val Ile
                1              5              10
ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc ctg      281
Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val Leu
                15              20              25
cag cca ctg cac ata aag gtc tac gcg gac cgc ttc acc cca ctc acc      329
Gln Pro Leu His Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr
                30              35              40
acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac ccc      377
Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro
                45              50              55
aac aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat ctc      425
Asn Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu
                60              65              70              75
ctg agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc cta      473
Leu Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu
                80              85              90
atc tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc tgg      521
Ile Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp
                95              100              105
aag gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg gat      569
Lys Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp
                110              115              120
atg ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att ctg      617
Met Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu
                125              130              135
gag gga aag aac tat cta cag tgg ctg act gaa agg tta act gag agg      665
Glu Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg
                140              145              150              155
gga gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg gca      713
Gly Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala
                160              165              170
aga gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct ggg      761
Arg Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly
                175              180              185
gcg cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc atg      809
Ala Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met
                190              195              200
aag gtg gac gcc cct tgg atg aag cac ttc att ctc acc cat gac cca      857
Lys Val Asp Ala Pro Trp Met Lys His Phe Ile Leu Thr His Asp Pro
                205              210              215
gag aga ggc atc tac aat tcc ccg tac atc atc cca ggg acc cag aca      905
Glu Arg Gly Ile Tyr Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln Thr
                220              225              230              235
gtt act ctt gga ggc atc ttc cag ttg gga aac tgg agt gaa cta aac      953
Val Thr Leu Gly Gly Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu Asn
                240              245              250
aat atc cag gac cac aac acc att tgg gaa ggc tgc tgc aga ctg gag      1001
Asn Ile Gln Asp His Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu Glu

```

ccc aca ctg	255	gca aga att	260	gaa gca act	265	ggc ttc cgg	1049
Pro Thr Leu	Lys Asn Ala Arg	Ile Ile Gly	Glu Ala Thr	Gly Phe Arg			
270		275		280			
cca gta cgc	ccc cag att	cgg cta gaa	aga gaa cag	ctt cgc act	gga	1097	
Pro Val Arg	Pro Gln Ile Arg	Leu Glu Arg	Glu Gln Leu	Arg Thr Gly			
285		290		295			
cct tca aac	aca gag gtc	atc cac aac	tat ggc cat	gga ggc tac	ggg	1145	
Pro Ser Asn	Thr Glu Val Ile	His Asn Tyr	Gly His Gly	Gly Tyr Gly			
300		305		310		315	
ctc acc atc	cac tgg gga	tgt gcc ctg	gag gca gcc	aag ctc ttt	ggg	1193	
Leu Thr Ile	His Trp Gly	Cys Ala Leu	Glu Ala Ala	Lys Leu Phe	Gly		
	320		325		330		
aga atc ctg	gaa gaa aag	aaa ttg tcc	aga atg cca	cca tcc cac	ctc	1241	
Arg Ile Leu	Glu Glu Lys Lys	Leu Ser Arg	Met Pro Pro	Ser His Leu			
	335		340		345		
tga agactccagt	gactgctgcc	tccccccaca	agaactccct	tctcccctca		1294	
*							
gccaatgaat	caatgtgctc	cttcataagc	cattgcttct	ccctcacttc	tttcctcaaa	1354	
gaagcatgag	gtgagagaaa	gccacraagt	cagtgcctgg	agaagggttc	agcccaacat	1414	
ggggcccttc	tcatcactga	aatccctcta	ccttctctgg	gtctggcatt	ataaagaaca	1474	
gctgaggctg	tcattccatg	agtcttcaga	agaaaggaga	gctcagaaag	tcaaagaggc	1534	
caactgccca	gagccacaga	aaatggagga	taattggaggc	taagtaacct	gattacaagt	1594	
tgtactaaca	tattaaaggt	tctgaaaagt	cctgcaaaa			1633	

<210> 18

<211> 347

<212> PRT

<213> Homo sapiens

<400> 18

Met Arg Val	Val Val	Ile Gly	Ala Gly	Val Ile	Gly Leu	Ser Thr	Ala
1	5		10		15		
Leu Cys Ile	His Glu	Arg Tyr	His Ser	Val Leu	Gln Pro	Leu Asp	Ile
	20		25		30		
Lys Val Tyr	Ala Asp	Arg Phe	Thr Pro	Leu Thr	Thr Thr	Asp Val	Ala
	35		40		45		
Ala Gly Leu	Trp Gln	Pro Tyr	Leu Ser	Asp Pro	Asn Asn	Pro Gln	Glu
	50		55		60		
Ala Asp Trp	Ser Gln	Gln Thr	Phe Asp	Tyr Leu	Leu Ser	His Val	His
	65		70		75		80
Ser Pro Asn	Ala Glu	Asn Leu	Gly Leu	Phe Leu	Ile Ser	Gly Tyr	Asn
	85		90		95		
Leu Phe His	Glu Ala	Ile Pro	Asp Pro	Ser Trp	Lys Asp	Thr Val	Leu
	100		105		110		
Gly Phe Arg	Lys Leu	Thr Pro	Arg Glu	Leu Asp	Met Phe	Pro Asp	Tyr
	115		120		125		
Gly Tyr Gly	Trp Phe	His Thr	Ser Leu	Ile Leu	Glu Gly	Lys Asn	Tyr
	130		135		140		
Leu Gln Trp	Leu Thr	Glu Arg	Leu Thr	Glu Arg	Gly Val	Lys Phe	Phe
	145		150		155		160
Gln Arg Lys	Val Glu	Ser Phe	Glu Glu	Val Ala	Arg Glu	Gly Ala	Asp
	165		170		175		
Val Ile Val	Asn Cys	Thr Gly	Val Trp	Ala Gly	Ala Leu	Gln Arg	Asp
	180		185		190		
Pro Leu Leu	Gln Pro	Gly Arg	Gly Gln	Ile Met	Lys Val	Asp Ala	Pro
	195		200		205		
Trp Met Lys	His Phe	Ile Leu	Thr His	Asp Pro	Glu Arg	Gly Ile	Tyr
	210		215		220		
Asn Ser Pro	Tyr Ile	Ile Pro	Gly Thr	Gln Thr	Val Thr	Leu Gly	Gly

2002247939 08 Oct 2004

225		230		235		240
Ile Phe Gln Leu Gly	Asn Trp Ser Glu Leu	Asn Asn Ile Gln Asp His				
	245		250			255
Asn Thr Ile Trp Glu Gly Cys Cys Arg	Leu Glu Pro Thr Leu Lys Asn					
	260		265			270
Ala Arg Ile Ile Gly Glu Ala Thr Gly Phe Arg Pro Val Arg Pro Gln						
	275		280			285
Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu						
	290		295			300
Val Ile His Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp						
	305		310			315
Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu						
	325		330			335
Lys Lys Leu Ser Arg Met Pro Pro Ser His Leu						
	340		345			

<210> 19

<211> 1200

<212> DNA

<213> Homo sapiens

<400> 19

atg gac aca gca cgg att gca gtt gtc ggg gca ggt gtg gtg ggg ctc	48
Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu	
1 5 10 15	
tcc acg gct gtg tgc atc tcc aaa ctg gtg ccc cga tgc tcc gtt acc	96
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr	
20 25 30	
atc att tca gac aag ttt act cca gat acc acc agt gat gtg gca gcc	144
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala	
35 40 45	
gga atg ctt att cct cac act tat cca gat aca ccc att cac acg cag	192
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln	
50 55 60	
aag cag tgg ttc aga gaa acc ttt aat cac ctc ttt gca att gcc aat	240
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn	
65 70 75 80	
tct gca gaa gct gga gat gct ggt gtt cat ttg gta tca ggt tgg cag	288
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Trp Gln	
85 90 95	
ata ttt cag agc act ccg act gaa gaa gtg cca ttc tgg gct gac gtg	336
Ile Phe Gln Ser Thr Pro Thr Glu Glu Val Pro Phe Trp Ala Asp Val	
100 105 110	
gtt ctg gga ttt cga aag atg act gag gct gag ctg aag aaa ttc ccc	384
Val Leu Gly Phe Arg Lys Met Thr Glu Ala Glu Leu Lys Lys Phe Pro	
115 120 125	
cag tat gtg ttt ggt cag gct ttt aca acc ctg aaa tgt gaa tgc cct	432
Gln Tyr Val Phe Gly Gln Ala Phe Thr Thr Leu Lys Cys Glu Cys Pro	
130 135 140	
gcc tac ctc ccg tgg ttg gag aaa agg ata aag gga agt gga ggc tgg	480
Ala Tyr Leu Pro Trp Leu Glu Lys Arg Ile Lys Gly Ser Gly Gly Trp	
145 150 155 160	
aca ctc act ccg cga ata gaa gac ctg tgg gaa ctt cat ccg tcc ttt	528
Thr Leu Thr Arg Arg Ile Glu Asp Leu Trp Glu Leu His Pro Ser Phe	
165 170 175	
gac atc gtg gtc aac tgt tca ggc ctt gga agc aga cag ctt gca gga	576
Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser Arg Gln Leu Ala Gly	
180 185 190	
gac tca aag att ttc cct gta agg ggc caa gtc ctc caa gtt cag gct	624
Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val Leu Gln Val Gln Ala	

2002247939 08 Oct 2004

ccc tgg gtg gag cat ttt atc cga gat ggc agt ggg ctg aca tat att	672
Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser Gly Leu Thr Tyr Ile	
210 215 220	
tat cct ggt aca tcc cat gta acc cta ggt gga act agg caa aaa ggg	720
Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly Thr Arg Gln Lys Gly	
225 230 235 240	
gac tgg aat ctg tcc ccg gat gca gaa aat agc aga gag att ctt tcc	768
Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser Arg Glu Ile Leu Ser	
245 250 255	
cga tgc tgt gct ctg gag ccc tcc ctc cac gga gcc tgc aac atc agg	816
Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly Ala Cys Asn Ile Arg	
260 265 270	
gag aag gtg ggc ttg agg ccc tac agg cca ggc gtg cga ctg cag aca	864
Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly Val Arg Leu Gln Thr	
275 280 285	
gag ctc ctt gcg cga gat gga cag agg ctg cct gta gtc cac cac tat	912
Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro Val Val His His Tyr	
290 295 300	
ggc cat ggg agt ggg ggc atc tca gtg cac tgg ggc act gct ctg gag	960
Gly His Gly Ser Gly Gly Ile Ser Val His Trp Gly Thr Ala Leu Glu	
305 310 315 320	
gcc gcc agg ctg gtg agc gag tgt gtc cat gcc ctc agg acc ccc att	1008
Ala Ala Arg Leu Val Ser Glu Cys Val His Ala Leu Arg Thr Pro Ile	
325 330 335	
ccc aag tca aac ctg tag atgacataaa atgacagcaa agagactgag	1056
Pro Lys Ser Asn Leu *	
340	
agactgttga tcaaagcaca gaacaggttc aaataacttt tccactgcat gaaagttaa	1116
ttagacattt ctttgttttc aacattagaa gtggtgtaac atgtaagctg agcacggtag	1176
catgcctata gtcccagcta cttg	1200
<210> 20	
<211> 1023	
<212> DNA	
<213> Homo sapiens	
<400> 20	
atg gac aca gca cgg att gca gtt gtc ggg gca ggt gtg gtg ggg ctc	48
Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu	
1 5 10 15	
tcc acg gct gtg tgc atc tcc aaa ctg gtg ccc cga tgc tcc gtt acc	96
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr	
20 25 30	
atc att tca gac aag ttt act cca gat acc acc agt gat gtg gca gcc	144
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala	
35 40 45	
gga atg ctt att cct cac act tat cca gat aca ccc att cac acg cag	192
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln	
50 55 60	
aag cag tgg ttc aga gaa acc ttt aat cac ctc ttt gca att gcc aat	240
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn	
65 70 75 80	
tct gca gaa gct gga gat gct ggt gtt cat ttg gta tca ggg ata aag	288
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Ile Lys	
85 90 95	
gga agt gga ggc tgg aca ctc act cgg cga ata gaa gac ctg tgg gaa	336
Gly Ser Gly Gly Trp Thr Leu Thr Arg Arg Ile Glu Asp Leu Trp Glu	
100 105 110	
ctt cat ccg tcc ttt gac atc gtg gtc aac tgt tca ggc ctt gga agc	384

2002247939 08 Oct 2004

```

Leu His Pro Ser Phe Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser
   115                               120               125
aga cag ctt gca gga gac tca aag att ttc cct gta agg ggc caa gtc      432
Arg Gln Leu Ala Gly Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val
   130                               135               140
ctc caa gtt cag gct ccc tgg gtg gag cat ttt atc cga gat ggc agt      480
Leu Gln Val Gln Ala Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser
   145                               150               155               160
ggg ctg aca tat att tat cct ggt aca tcc cat gta acc cta ggt gga      528
Gly Leu Thr Tyr Ile Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly
   165                               170               175
act agg caa aaa ggg gac tgg aat ctg tcc ccg gat gca gaa aat agc      576
Thr Arg Gln Lys Gly Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser
   180                               185               190
aga gag att ctt tcc cga tgc tgt gct ctg gag ccc tcc ctc cac gga      624
Arg Glu Ile Leu Ser Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly
   195                               200               205
gcc tgc aac atc agg gag aag gtg ggc ttg agg ccc tac agg cca ggc      672
Ala Cys Asn Ile Arg Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly
   210                               215               220
gtg cga ctg cag aca gag ctc ctt gcg cga gat gga cag agg ctg cct      720
Val Arg Leu Gln Thr Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro
   225                               230               235               240
gta gtc cac cac tat ggc cat ggg agt ggg ggc atc tca gtg cac tgg      768
Val Val His His Tyr Gly His Gly Ser Gly Gly Ile Ser Val His Trp
   245                               250               255
ggc act gct ctg gag gcc gcc agg ctg gtg agc gag tgt gtc cat gcc      816
Gly Thr Ala Leu Glu Ala Ala Arg Leu Val Ser Glu Cys Val His Ala
   260                               265               270
ctc agg acc ccc att ccc aag tca aac ctg tag atgacataaa atgacagcaa      869
Leu Arg Thr Pro Ile Pro Lys Ser Asn Leu *
   275                               280
agagactgag agactgttga tcaaagcaca gaacagggttc aaataacttt tccactgcat      929
gaaagttaa ttagacattt ctttgttttc aacattagaa gtggtgtaac atgtaagctg      989
agcacggtag catgcctata gtcccagcta cttg      1023

```

<210> 21
 <211> 341
 <212> PRT
 <213> Homo sapiens

```

<400> 21
Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu
 1                               5               10               15
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr
 20                               25               30
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala
 35                               40               45
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln
 50                               55               60
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn
 65                               70               75               80
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Trp Gln
 85                               90               95
Ile Phe Gln Ser Thr Pro Thr Glu Glu Val Pro Phe Trp Ala Asp Val
100                               105               110
Val Leu Gly Phe Arg Lys Met Thr Glu Ala Glu Leu Lys Lys Phe Pro
115                               120               125
Gln Tyr Val Phe Gly Gln Ala Phe Thr Thr Leu Lys Cys Glu Cys Pro
130                               135               140
Ala Tyr Leu Pro Trp Leu Glu Lys Arg Ile Lys Gly Ser Gly Gly Trp

```

2002247939 08 Oct 2004

145	Thr	Leu	Thr	Arg	Arg	150	Ile	Glu	Asp	Leu	Trp	155	Glu	Leu	His	Pro	Ser	160	Phe
					165							170						175	
Asp	Ile	Val	Val	Asn	Cys	Ser	Gly	Leu	Gly	Ser	Arg	Gln	Leu	Ala	Gly				
			180						185						190				
Asp	Ser	Lys	Ile	Phe	Pro	Val	Arg	Gly	Gln	Val	Leu	Gln	Val	Gln	Ala				
		195					200					205							
Pro	Trp	Val	Glu	His	Phe	Ile	Arg	Asp	Gly	Ser	Gly	Leu	Thr	Tyr	Ile				
		210				215					220								
Tyr	Pro	Gly	Thr	Ser	His	Val	Thr	Leu	Gly	Gly	Thr	Arg	Gln	Lys	Gly				
		225			230					235									
Asp	Trp	Asn	Leu	Ser	Pro	Asp	Ala	Glu	Asn	Ser	Arg	Glu	Ile	Leu	Ser				
				245					250						255				
Arg	Cys	Cys	Ala	Leu	Glu	Pro	Ser	Leu	His	Gly	Ala	Cys	Asn	Ile	Arg				
			260						265					270					
Glu	Lys	Val	Gly	Leu	Arg	Pro	Tyr	Arg	Pro	Gly	Val	Arg	Leu	Gln	Thr				
		275					280					285							
Glu	Leu	Ala	Arg	Asp	Gly	Gln	Arg	Leu	Pro	Val	Val	His	His	Tyr					
		290			295				300										
Gly	His	Gly	Ser	Gly	Gly	Ile	Ser	Val	His	Trp	Gly	Thr	Ala	Leu	Glu				
		305			310				315										
Ala	Ala	Arg	Leu	Val	Ser	Glu	Cys	Val	His	Ala	Leu	Arg	Thr	Pro	Ile				
				325					330						335				
Pro	Lys	Ser	Asn	Leu															
				340															

<210> 22
 <211> 282
 <212> PRT
 <213> Homo sapiens

<400> 22

Met	Asp	Thr	Ala	Arg	Ile	Ala	Val	Val	Gly	Ala	Gly	Val	Val	Gly	Leu
1			5						10					15	
Ser	Thr	Ala	Val	Cys	Ile	Ser	Lys	Leu	Val	Pro	Arg	Cys	Ser	Val	Thr
			20					25					30		
Ile	Ile	Ser	Asp	Lys	Phe	Thr	Pro	Asp	Thr	Thr	Ser	Asp	Val	Ala	Ala
			35				40					45			
Gly	Met	Leu	Ile	Pro	His	Thr	Tyr	Pro	Asp	Thr	Pro	Ile	His	Thr	Gln
	50					55					60				
Lys	Gln	Trp	Phe	Arg	Glu	Thr	Phe	Asn	His	Leu	Phe	Ala	Ile	Ala	Asn
	65				70					75				80	
Ser	Ala	Glu	Ala	Gly	Asp	Ala	Gly	Val	His	Leu	Val	Ser	Gly	Ile	Lys
				85				90					95		
Gly	Ser	Gly	Gly	Trp	Thr	Leu	Thr	Arg	Arg	Ile	Glu	Asp	Leu	Trp	Glu
			100					105					110		
Leu	His	Pro	Ser	Phe	Asp	Ile	Val	Val	Asn	Cys	Ser	Gly	Leu	Gly	Ser
			115				120					125			
Arg	Gln	Leu	Ala	Gly	Asp	Ser	Lys	Ile	Phe	Pro	Val	Arg	Gly	Gln	Val
			130				135				140				
Leu	Gln	Val	Gln	Ala	Pro	Trp	Val	Glu	His	Phe	Ile	Arg	Asp	Gly	Ser
				145			150			155					160
Gly	Leu	Thr	Tyr	Ile	Tyr	Pro	Gly	Thr	Ser	His	Val	Thr	Leu	Gly	Gly
				165				170						175	
Thr	Arg	Gln	Lys	Gly	Asp	Trp	Asn	Leu	Ser	Pro	Asp	Ala	Glu	Asn	Ser
			180					185					190		
Arg	Glu	Ile	Leu	Ser	Arg	Cys	Cys	Ala	Leu	Glu	Pro	Ser	Leu	His	Gly
			195				200					205			
Ala	Cys	Asn	Ile	Arg	Glu	Lys	Val	Gly	Leu	Arg	Pro	Tyr	Arg	Pro	Gly
		210				215					220				
Val	Arg	Leu	Gln	Thr	Glu	Leu	Leu	Ala	Arg	Asp	Gly	Gln	Arg	Leu	Pro

2002247939 08 Oct 2004

225 230 235 240
 Val Val His His Tyr Gly His Gly Ser Gly Gly Ile Ser Val His Trp
 245 250 255
 Gly Thr Ala Leu Glu Ala Ala Arg Leu Val Ser Glu Cys Val His Ala
 260 265 270
 Leu Arg Thr Pro Ile Pro Lys Ser Asn Leu
 275 280

<210> 23
 <211> 47
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> oligonucleotide 24-1443-126

<400> 23
 tacggcttag taagttggag aacyaggatc agaagacagg tctgcct

47

<220>
 <221> allele
 <222> 24
 <223> polymorphic base C or T

<210> 24
 <211> 47
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> oligonucleotide 24-1457-52

<220>
 <221> allele
 <222> 24
 <223> polymorphic base C or A

<400> 24
 tctgagatgc cctgtgtcc tctmagggag tagtggctga gcatttc

47

<210> 25
 <211> 47
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> oligonucleotide 27-93-181

<220>
 <221> allele
 <222> 24
 <223> polymorphic base C or T

<400> 25
 cccagctctg ccactggcga gctytgtggc cttgggcaag ttactcc

47

<210> 26
 <211> 47
 <212> DNA
 <213> Artificial Sequence

2002247939 08 Oct 2004

50

<220>

<223> oligonucleotide 24-1461-256

<220>

<221> allele

<222> 24

<223> polymorphic base A or G

<400> 26

gatggctctg gcattttcag ggarcagtca tgtctgatct caagttc

47