

(19) **DANMARK**

(10) **DK/EP 2906696 T3**



Patent- og
Varemærkestyrelsen

(12) Oversættelse af
europæisk patentskrift

-
- (51) Int.Cl.: **C 12 N 15/113 (2010.01)** **A 61 P 25/28 (2006.01)**
- (45) Oversættelsen bekendtgjort den: **2019-11-25**
- (80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2019-10-09**
- (86) Europæisk ansøgning nr.: **13846313.8**
- (86) Europæisk indleveringsdag: **2013-10-15**
- (87) Den europæiske ansøgnings publiceringsdag: **2015-08-19**
- (86) International ansøgning nr.: **US2013065067**
- (87) Internationalt publikationsnr.: **WO2014062686**
- (30) Prioritet: **2012-10-15 US 201261714139 P**
- (84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**
- (73) Patenthaver: **Ionis Pharmaceuticals, Inc., 2855 Gazelle Court, Carlsbad, CA 92010, USA**
Ludwig Institute for Cancer Research, 666 Third Avenue, New York, NY 10017, USA
The Regents of the University of California, 111 Franklin Street, 5th Floor, Oakland, CA 94607-5200, USA
- (72) Opfinder: **RIGO, Frank, 2855 Gazelle Court, Carlsbad, CA 92010, Spanien**
CLEVELAND, Don W., 9500 Gilman Drive, Mail Code:0670, CMM-East, Room 3053, La Jolla, CA 92093-0670, USA
LAGIER-TOURENNE, Clotilde, 9500 Gilman Drive, Mail: 0670, CMM-East, Room 3053, La Jolla, CA 92093-0670, USA
RAVITS, John M., 1770 O'Farrell Street 13, San Francisco, CA 94080, USA
BENNETT, C. Frank, 2855 Gazelle Court, Carlsbad, CA 92010, USA
FREIER, Susan M., 2855 Gazelle Court, Carlsbad, CA 92010, USA
BAUGHN, Michael W., 9500 Gilman Drive 0624, La Jolla, CA 92093-0624, USA
- (74) Fuldmægtig i Danmark: **Plougmann Vingtoft A/S, Strandvejen 70, 2900 Hellerup, Danmark**
- (54) Benævnelse: **FREMGANGSMÅDER TIL MODULERING AF C9ORF72-EKSPRESSION**
- (56) Fremdragne publikationer:
WO-A1-2012/092367
WO-A1-2013/030588
WO-A2-2012/012467
US-A1- 2012 214 865
DONNELLY CHRISTOPHER J ET AL: "Development of C9orf72 ALS Biomarkers and Therapeutics", ANNALS OF NEUROLOGY & 137TH ANNUAL MEETING OF THE AMERICAN-NEUROLOGICAL-ASSOCIATION (ANA); BOSTON, MA, USA; OCTOBER 07 -10, 2012, vol. 72, no. Suppl. 16, M1415, 7 October 2012 (2012-10-07), pages S67-S68, XP002757118,
A. Madsen: "Antisense Against C9ORF72", MDA/ALS Newsmagazine, 1 July 2012 (2012-07-01), pages 1-2,

Fortsættes ...

XP055251577, Retrieved from the Internet: URL:<http://alsn.mda.org/article/antisense-against-c9orf72> [retrieved on 2016-02-19]

ALANE RENTON ET AL: "A Hexanucleotide Repeat Expansion in the Cause of Chromosome 9p21-Linked ALS-FTD", NEURON, CELL PRESS, US, vol. 72, no. 2, 14 September 2011 (2011-09-14), pages 257-268, XP028322561, ISSN: 0896-6273, DOI: 10.1016/J.NEURON.2011.09.010 [retrieved on 2011-09-17]

MARIELLY DEJESUS-HERNANDEZ ET AL: "Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of Causes Chromosome 9p-Linked FTD and ALS", NEURON, CELL PRESS, US, vol. 72, no. 2, 14 September 2011 (2011-09-14), pages 245-256, XP028322560, ISSN: 0896-6273, DOI: 10.1016/J.NEURON.2011.09.011 [retrieved on 2011-09-17]

PETER K TODD ET AL: "RNA mediated neurodegeneration in repeat expansion disorders", ANNALS OF NEUROLOGY., vol. 67, no. 3, 23 December 2009 (2009-12-23), pages NA-NA, XP055269539, BOSTON, US ISSN: 0364-5134, DOI: 10.1002/ana.21948

S. A. M. MULDER ET AL: "Triplet-repeat oligonucleotide-mediated reversal of RNA toxicity in myotonic dystrophy", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, vol. 106, no. 33, 18 August 2009 (2009-08-18), pages 13915-13920, XP055007026, ISSN: 0027-8424, DOI: 10.1073/pnas.0905780106

RENTON, AE A.: 'Hexanucleotide Repeat Expansion In C9ORF72 Is The Cause Of Chromosome 9p21-Linked ALS-FTD.' NEURON vol. 72, no. 2, 14 September 2011, pages 257 - 268, XP028322561

DEJESUS-HERNANDEZ, M.: 'Expanded GGGGCC Hexanucleotide Repeat In Noncoding Region Of C9ORF72 Causes Chromosome 9p-Linked FTD And ALS.' NEURON. vol. 72, no. 2, 14 September 2011, pages 245 - 256, XP028322560

SHARON J SHA ET AL: "Treatment implications of C9ORF72", ALZHEIMERS RES THER, vol. 4, no. 6, 16 February 2012 (2012-02-16), page 46, XP055302455, London, UK ISSN: 1758-9193, DOI: 10.1371/journal.pone.0029522

DESCRIPTION

Field

[0001] Disclosed herein are methods for reducing expression of C9ORF72 mRNA and protein in an animal. Such methods are useful to treat, prevent, or ameliorate neurodegenerative diseases, including amyotrophic lateral sclerosis (ALS), frontotemporal dementia (FTD), corticalbasal degeneration syndrome (CBD), atypical Parkinsonian syndrome, and olivopontocerellar degeneration (OPCD).

Background

[0002] Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease characterized clinically by progressive paralysis leading to death from respiratory failure, typically within two to three years of symptom onset (Rowland and Shneider, *N. Engl. J. Med.*, 2001, 344, 1688-1700). ALS is the third most common neurodegenerative disease in the Western world (Hirtz et al., *Neurology*, 2007, 68, 326-337), and there are currently no effective therapies. Approximately 10% of cases are familial in nature, whereas the bulk of patients diagnosed with the disease are classified as sporadic as they appear to occur randomly throughout the population (Chio et al., *Neurology*, 2008, 70, 533-537). There is growing recognition, based on clinical, genetic, and epidemiological data, that ALS and frontotemporal dementia (FTD) represent an overlapping continuum of disease, characterized pathologically by the presence of TDP-43 positive inclusions throughout the central nervous system (Lillo and Hodges, *J. Clin. Neurosci.*, 2009, 16, 1131-1135; Neumann et al., *Science*, 2006, 314, 130-133).

[0003] To date, a number of genes have been discovered as causative for classical familial ALS, for example, SOD1, TARDBP, FUS, OPTN, and VCP (Johnson et al., *Neuron*, 2010, 68, 857-864; Kwiatkowski et al., *Science*, 2009, 323, 1205-1208; Maruyama et al., *Nature*, 2010, 465, 223-226; Rosen et al., *Nature*, 1993, 362, 59-62; Sreedharan et al., *Science*, 2008, 319, 1668-1672; Vance et al., *Brain*, 2009, 129, 868-876). Recently, linkage analysis of kindreds involving multiple cases of ALS, FTD, and ALS-FTD had suggested that there was an important locus for the disease on the short arm of chromosome 9 (Boxer et al., *J. Neurol. Neurosurg. Psychiatry*, 2011, 82, 196-203; Morita et al., *Neurology*, 2006, 66, 839-844; Pearson et al. *J. Neurol.*, 2011, 258, 647-655; Vance et al., *Brain*, 2006, 129, 868-876). The chromosome 9p21ALS-FTD locus is the last major autosomal-dominant gene whose mutation is causative of ALS. The ALS-FTD causing mutation is a large hexanucleotide (GGGGCC) repeat expansion in the first intron of the C9ORF72 gene (Renton et al., *Neuron*, 2011, 72, 257-268; DeJesus-Hernandez et al., *Neuron*, 2011, 72, 245-256). A founder haplotype, covering the C9ORF72 gene, is present in the majority of cases linked to this region (Renton et al., *Neuron*, 2011, 72, 257-268). This locus on chromosome 9p21 accounts for nearly half of familial ALS and nearly one-quarter of all ALS cases in a cohort of 405 Finnish patients

(Laaksovirta et al, Lancet Neurol., 2010, 9, 978-985).

[0004] A founder haplotype, covering the C9ORF72 gene, is present in the majority of cases linked to this region.

[0005] There are currently no effective therapies to treat such neurodegenerative diseases. Therefore, it is an object to provide compositions and methods for the treatment of such neurodegenerative diseases.

[0006] WO 2013/030588 describes methods for diagnosing a neurodegenerative disease. US 2012/214865 describes methods for slowing familial ALS disease progression. Donnelly et al., 2012, Annals of Neurology., 72(16), S67-S68 reports on development of C9orf72 ALS biomarkers and therapeutics. Madsen, 2012, MDA/ALS newsmagazine, 1-2 reports on antisense against C9ORF72. WO 2012/092367 describes nucleic acid binding compounds, methods of making, and use thereof. Renton, 2011, Neuron, 72(2), 257-268 reports that a hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-Linked ALS-FTD. DeJesus-Hernandez, 2011, Neuron, 72(2), 245-256 reports that expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. WO 2012/012467 describes modulation of nuclear-retained RNA. Sha et al., 2012, Alzheimers Res Ther., 4(6), 46 reports on treatment implications of C9ORF72.

Summary of the invention

[0007] The invention provides a compound comprising a modified antisense oligonucleotide consisting of 12 to 30 linked nucleosides, wherein the modified antisense oligonucleotide is at least 90% complementary to a C9ORF72 nucleic acid, for use in treating a neurodegenerative disease by reducing the percent of cells with nuclear foci and/or reducing the number of nuclear foci per cell in an animal identified as having foci, wherein the modified antisense oligonucleotide comprises at least one phosphorothioate internucleoside linkage, at least one 5-methylcytosine, at least one bicyclic sugar or at least one 2'-O-methoxyethyl group.

Summary of the disclosure

[0008] Described herein are methods for modulating levels of C9ORF72 mRNA and protein in cells, tissues, and animals. In certain instances, C9ORF72 specific inhibitors modulate expression of C9ORF72 mRNA and protein. In certain instances, C9ORF72 specific inhibitors are nucleic acids, proteins, or small molecules.

[0009] In the invention, modulation occurs in a cell or tissue in an animal. In certain embodiments, the animal is a human. In certain embodiments, C9ORF72 mRNA levels are reduced. In certain embodiments, C9ORF72 protein levels are reduced. In certain

embodiments, certain C9ORF72 mRNA variants are preferentially reduced. In certain embodiments, the C9ORF72 mRNA variants preferentially reduced are variants containing intron 1. In certain embodiments, intron 1 contains a hexanucleotide repeat expansion. In certain embodiments, the hexanucleotide repeat expansion is associated with a C9ORF72 associated disease. In certain embodiments, the hexanucleotide repeat expansion is associated with a C9ORF72 hexanucleotide repeat expansion associated disease. In certain embodiments, the hexanucleotide repeat expansion comprises at least 30 GGGGCC repeats. In certain embodiments, the hexanucleotide repeat expansion is associated with nuclear foci. In certain instances, the methods described herein are useful for reducing C9ORF72 mRNA, C9ORF72 protein levels, and nuclear foci. Such reduction can occur in a time-dependent manner or in a dose-dependent manner.

[0010] Also described are methods useful for preventing, treating, and ameliorating diseases, disorders, and conditions associated with C9ORF72. Such diseases, disorders, and conditions associated with C9ORF72 can be neurodegenerative diseases. The neurodegenerative disease can be ALS, FTD, corticalbasal degeneration syndrome (CBD), atypical Parkinsonian syndrome, and olivopontocerellar degeneration (OPCD).

[0011] Such diseases, disorders, and conditions can have one or more risk factors, causes, or outcomes in common. Certain risk factors and causes for development of a neurodegenerative disease, and, in particular, ALS and FTD, include genetic predisposition and older age.

[0012] In certain instances, methods of treatment include administering a C9ORF72 specific inhibitor to an individual in need thereof. In these instances the C9ORF72 specific inhibitor can be a nucleic acid. In these instances the nucleic acid can be an antisense compound. In these instances, the antisense compound can be a single-stranded antisense oligonucleotide. In these instances, the single-stranded antisense oligonucleotide can be complementary to a C9ORF72 nucleic acid.

Detailed Description

[0013] It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed. Herein, the use of the singular includes the plural unless specifically stated otherwise. As used herein, the use of "or" means "and/or" unless stated otherwise. Additionally, as used herein, the use of "and" means "and/or" unless stated otherwise. Furthermore, the use of the term "including" as well as other forms, such as "includes" and "included", is not limiting. Also, terms such as "element" or "component" encompass both elements and components comprising one unit and elements and components that comprise more than one subunit, unless specifically stated otherwise.

[0014] The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described.

Definitions

[0015] Unless specific definitions are provided, the nomenclature utilized in connection with, and the procedures and techniques of, analytical chemistry, synthetic organic chemistry, and medicinal and pharmaceutical chemistry described herein are those well known and commonly used in the art. Standard techniques may be used for chemical synthesis, and chemical analysis.

[0016] Unless otherwise indicated, the following terms have the following meanings:

"2'-O-methoxyethyl group" (also 2'-MOE and 2'-OCH₂CH₂-OCH₃ and MOE) refers to an O-methoxy-ethyl modification of the 2' position of a furanosyl ring. A 2'-O-methoxyethyl modified sugar is a modified sugar.

"2'-MOE nucleoside" (also 2'-O-methoxyethyl nucleoside) means a nucleoside comprising a 2'-O-methoxyethyl group.

"5-methylcytosine" means a cytosine modified with a methyl group attached to the 5' position. A 5-methylcytosine is a modified nucleobase.

[0017] "About" means within $\pm 7\%$ of a value. For example, if it is stated, "the compounds affected at least about 70% inhibition of C9ORF72", it is implied that the C9ORF72 levels are inhibited within a range of 63% and 77%.

[0018] "Administered concomitantly" refers to the co-administration of two pharmaceutical agents in any manner in which the pharmacological effects of both are manifest in the patient at the same time. Concomitant administration does not require that both pharmaceutical agents be administered in a single pharmaceutical composition, in the same dosage form, or by the same route of administration. The effects of both pharmaceutical agents need not manifest themselves at the same time. The effects need only be overlapping for a period of time and need not be coextensive.

[0019] "Administering" means providing a pharmaceutical agent to an animal, and includes, but is not limited to administering by a medical professional and self-administering.

[0020] "Amelioration" or "ameliorate" or "ameliorating" refers to a lessening of at least one indicator, sign, or symptom of a disease, disorder, or condition. The severity of indicators may be determined by subjective or objective measures, which are known to those skilled in the art.

[0021] "Animal" refers to a human or non-human animal, including, but not limited to, mice, rats, rabbits, dogs, cats, pigs, and non-human primates, including, but not limited to, monkeys

and chimpanzees.

[0022] "Antibody" refers to a molecule characterized by reacting specifically with an antigen in some way, where the antibody and the antigen are each defined in terms of the other. Antibody may refer to a complete antibody molecule or any fragment or region thereof, such as the heavy chain, the light chain, Fab region, and Fc region.

[0023] "Antisense activity" means any detectable or measurable activity attributable to the hybridization of an antisense compound to its target nucleic acid. In certain embodiments, antisense activity is a decrease in the amount or expression of a target nucleic acid or protein encoded by such target nucleic acid.

[0024] "Antisense compound" means an oligomeric compound that is capable of undergoing hybridization to a target nucleic acid through hydrogen bonding. Examples of antisense compounds include single-stranded and double-stranded compounds, such as, antisense oligonucleotides, siRNAs, shRNAs, ssRNAs, and occupancy-based compounds. Antisense mechanisms include, without limitation, RNase H mediated antisense; RNAi mechanisms, which utilize the RISC pathway and include, without limitation, siRNA, ssRNA and microRNA mechanisms; and occupancy based mechanisms, including, without limitation uniform modified oligonucleotides. Certain antisense compounds may act through more than one such mechanisms and/or through additional mechanisms.

[0025] "Antisense inhibition" means reduction of target nucleic acid levels or target protein levels in the presence of an antisense compound complementary to a target nucleic acid compared to target nucleic acid levels or target protein levels in the absence of the antisense compound. Inhibition may be by any means including RNase H degradation, such as with a gapmer, and steric blockage/occupancy based mechanisms, such as with a uniformly modified oligonucleotide.

[0026] "Antisense oligonucleotide" means a single-stranded oligonucleotide having a nucleobase sequence that permits hybridization to a corresponding segment of a target nucleic acid.

[0027] "Bicyclic sugar" means a furanosyl ring modified by the bridging of two atoms. A bicyclic sugar is a modified sugar.

[0028] "Bicyclic nucleoside" (also BNA) means a nucleoside having a sugar moiety comprising a bridge connecting two carbon atoms of the sugar ring, thereby forming a bicyclic ring system. In certain embodiments, the bridge connects the 4'-carbon and the 2'-carbon of the sugar ring.

[0029] "C9ORF72 associated disease" means any disease associated with any C9ORF72 nucleic acid or expression product thereof. Such diseases may include a neurodegenerative disease. Such neurodegenerative diseases may include ALS and FTD.

[0030] "C9ORF72 hexanucleotide repeat expansion associated disease" means any disease associated with a C9ORF72 nucleic acid containing a hexanucleotide repeat expansion. In certain embodiments, the hexanucleotide repeat expansion may comprise GGGGCC, GGGGGG, GGGGGC, or GGGGCG repeated at least 30 times. Such diseases may include a neurodegenerative disease. Such neurodegenerative diseases may include ALS and FTD.

[0031] "C9ORF72 nucleic acid" means any nucleic acid encoding C9ORF72. For example, in certain embodiments, a C9ORF72 nucleic acid includes a DNA sequence encoding C9ORF72, an RNA sequence transcribed from DNA encoding C9ORF72 (including genomic DNA comprising introns and exons), and an mRNA sequence encoding C9ORF72. "C9ORF72 mRNA" means an mRNA encoding a C9ORF72 protein.

[0032] "C9ORF72 specific inhibitor" refers to any agent capable of specifically inhibiting the expression of C9ORF72 mRNA and/or C9ORF72 protein at the molecular level. For example, C9ORF72 specific inhibitors include nucleic acids (including antisense compounds), siRNAs, aptamers, antibodies, peptides, small molecules, and other agents capable of inhibiting the expression of C9ORF72 mRNA and/or C9ORF72 protein. Similarly, C9ORF72 specific inhibitors may affect other molecular processes in an animal.

[0033] "Cap structure" or "terminal cap moiety" means chemical modifications, which have been incorporated at either terminus of an antisense compound.

[0034] "cEt" or "constrained ethyl" means a bicyclic nucleoside having a sugar moiety comprising a bridge connecting the 4'-carbon and the 2'-carbon, wherein the bridge has the formula: 4'-CH(CH₃)-O-2'.

[0035] "Constrained ethyl nucleoside" (also cEt nucleoside) means a nucleoside comprising a bicyclic sugar moiety comprising a 4'-CH(CH₃)-O-2' bridge.

[0036] "Chemically distinct region" refers to a region of an antisense compound that is in some way chemically different than another region of the same antisense compound. For example, a region having 2'-O-methoxyethyl nucleosides is chemically distinct from a region having nucleosides without 2'-O-methoxyethyl modifications.

[0037] "Chimeric antisense compound" means an antisense compound that has at least two chemically distinct regions.

[0038] "Co-administration" means administration of two or more pharmaceutical agents to an individual. The two or more pharmaceutical agents may be in a single pharmaceutical composition, or may be in separate pharmaceutical compositions. Each of the two or more pharmaceutical agents may be administered through the same or different routes of administration. Co-administration encompasses parallel or sequential administration.

[0039] "Complementarity" means the capacity for pairing between nucleobases of a first

nucleic acid and a second nucleic acid.

[0040] "Contiguous nucleobases" means nucleobases immediately adjacent to each other.

[0041] "Diluent" means an ingredient in a composition that lacks pharmacological activity, but is pharmaceutically necessary or desirable. For example, the diluent in an injected composition may be a liquid, e.g. saline solution.

[0042] "Dose" means a specified quantity of a pharmaceutical agent provided in a single administration, or in a specified time period. In certain instances, a dose may be administered in one, two, or more boluses, tablets, or injections. For example, in certain instances where subcutaneous administration is desired, the desired dose requires a volume not easily accommodated by a single injection, therefore, two or more injections may be used to achieve the desired dose. In certain instances, the pharmaceutical agent is administered by infusion over an extended period of time or continuously. Doses may be stated as the amount of pharmaceutical agent per hour, day, week, or month.

[0043] "Effective amount" means the amount of pharmaceutical agent sufficient to effectuate a desired physiological outcome in an individual in need of the pharmaceutical agent. The effective amount may vary among individuals depending on the health and physical condition of the individual to be treated, the taxonomic group of the individuals to be treated, the formulation of the composition, assessment of the individual's medical condition, and other relevant factors.

[0044] "Expression" means conversion of the information from a C9ORF72 gene into mRNA via transcription and then to protein via translation. Expression may result in a phenotypic manifestation of the C9ORF72 gene.

[0045] "Fully complementary" or "100% complementary" means each nucleobase of a first nucleic acid has a complementary nucleobase in a second nucleic acid. In certain embodiments, a first nucleic acid is an antisense compound and a target nucleic acid is a second nucleic acid.

[0046] "Gapmer" means a chimeric antisense compound in which an internal region having a plurality of nucleosides that support RNase H cleavage is positioned between external regions having one or more nucleosides, wherein the nucleosides comprising the internal region are chemically distinct from the nucleoside or nucleosides comprising the external regions. The internal region may be referred to as a "gap" and the external regions may be referred to as the "wings."

[0047] "Gap-narrowed" means a chimeric antisense compound having a gap segment of 9 or fewer contiguous 2'-deoxyribonucleosides positioned between and immediately adjacent to 5' and 3' wing segments having from 1 to 6 nucleosides.

[0048] "Gap-widened" means a chimeric antisense compound having a gap segment of 12 or more contiguous 2'-deoxyribonucleosides positioned between and immediately adjacent to 5' and 3' wing segments having from 1 to 6 nucleosides.

[0049] "Hexanucleotide repeat expansion" means a series of six bases (for example, GGGGCC, GGGGGG, GGGGCG, or GGGGGC) repeated at least twice. In certain embodiments, the hexanucleotide repeat expansion may be located in intron 1 of a C9ORF72 nucleic acid. In certain embodiments, a pathogenic hexanucleotide repeat expansion includes at least 30 repeats of GGGGCC, GGGGGG, GGGGCG, or GGGGGC in a C9ORF72 nucleic acid and is associated with disease. In certain embodiments, the repeats are consecutive. In certain embodiments, the repeats are interrupted by 1 or more nucleobases. In certain embodiments, a wild-type hexanucleotide repeat expansion includes 23 or fewer repeats of GGGGCC, GGGGGG, GGGGCG, or GGGGGC in a C9ORF72 nucleic acid. In certain embodiments, the repeats are consecutive. In certain embodiments, the repeats are interrupted by 1 or more nucleobases.

[0050] "Hybridization" means the annealing of complementary nucleic acid molecules. In the invention, complementary nucleic acid molecules include an antisense compound and a target nucleic acid.

[0051] "Identifying an animal having a C9ORF72 associated disease" means identifying an animal having been diagnosed with a C9ORF72 associated disease or predisposed to develop a C9ORF72 associated disease. Individuals predisposed to develop a C9ORF72 associated disease include those having one or more risk factors for developing a C9ORF72 associated disease, including, having a personal or family history or genetic predisposition of one or more C9ORF72 associated diseases. Such identification may be accomplished by any method including evaluating an individual's medical history and standard clinical tests or assessments, such as genetic testing.

[0052] "Immediately adjacent" means there are no intervening elements between the immediately adjacent elements.

[0053] "Individual" means a human or non-human animal selected for treatment or therapy.

[0054] "Inhibiting C9ORF72" means reducing expression of C9ORF72 mRNA and/or protein levels in the presence of a C9ORF72 specific inhibitor, including a C9ORF72 antisense oligonucleotide, as compared to expression of C9ORF72 mRNA and/or protein levels in the absence of a C9ORF72 specific inhibitor, such as a C9ORF72 antisense oligonucleotide.

[0055] "Internucleoside linkage" refers to the chemical bond between nucleosides.

[0056] "Linked nucleosides" means adjacent nucleosides which are bonded together.

[0057] "Mismatch" or "non-complementary nucleobase" refers to the case when a nucleobase

of a first nucleic acid is not capable of pairing with the corresponding nucleobase of a second or target nucleic acid.

[0058] "Modified internucleoside linkage" refers to a substitution or any change from a naturally occurring internucleoside bond (*i.e.*, a phosphodiester internucleoside bond).

[0059] "Modified nucleobase" refers to any nucleobase other than adenine, cytosine, guanine, thymidine, or uracil. An "unmodified nucleobase" means the purine bases adenine (A) and guanine (G), and the pyrimidine bases thymine (T), cytosine (C), and uracil (U).

[0060] "Modified nucleotide" means a nucleotide having, independently, a modified sugar moiety, modified internucleoside linkage, or modified nucleobase. A "modified nucleoside" means a nucleoside having, independently, a modified sugar moiety or modified nucleobase.

[0061] "Modified oligonucleotide" means an oligonucleotide comprising a modified internucleoside linkage, a modified sugar, or a modified nucleobase.

[0062] "Modified sugar" refers to a substitution or change from a natural sugar.

[0063] "Motif" means the pattern of chemically distinct regions in an antisense compound.

[0064] "Naturally occurring internucleoside linkage" means a 3' to 5' phosphodiester linkage.

[0065] "Natural sugar moiety" means a sugar found in DNA (2'-H) or RNA (2'-OH).

[0066] "Nucleic acid" refers to molecules composed of monomeric nucleotides. A nucleic acid includes ribonucleic acids (RNA), deoxyribonucleic acids (DNA), single-stranded nucleic acids, double-stranded nucleic acids, small interfering ribonucleic acids (siRNA), and microRNAs (miRNA).

[0067] "Nucleobase" means a heterocyclic moiety capable of pairing with a base of another nucleic acid.

[0068] "Nucleobase sequence" means the order of contiguous nucleobases independent of any sugar, linkage, or nucleobase modification.

[0069] "Nucleoside" means a nucleobase linked to a sugar.

[0070] "Nucleoside mimetic" includes those structures used to replace the sugar or the sugar and the base and not necessarily the linkage at one or more positions of an oligomeric compound such as for example nucleoside mimetics having morpholino, cyclohexenyl, cyclohexyl, tetrahydropyranyl, bicyclo, or tricyclo sugar mimetics, *e.g.*, non furanose sugar units. Nucleotide mimetic includes those structures used to replace the nucleoside and the linkage at one or more positions of an oligomeric compound such as for example peptide

nucleic acids or morpholinos (morpholinos linked by -N(H)-C(=O)-O- or other non-phosphodiester linkage). Sugar surrogate overlaps with the slightly broader term nucleoside mimetic but is intended to indicate replacement of the sugar unit (furanose ring) only. The tetrahydropyranyl rings provided herein are illustrative of an example of a sugar surrogate wherein the furanose sugar group has been replaced with a tetrahydropyranyl ring system.

[0071] "Nucleotide" means a nucleoside having a phosphate group covalently linked to the sugar portion of the nucleoside.

[0072] "Oligomeric compound" or "oligomer" means a polymer of linked monomeric subunits which is capable of hybridizing to at least a region of a nucleic acid molecule.

[0073] "Oligonucleotide" means a polymer of linked nucleosides each of which can be modified or unmodified, independent one from another.

[0074] "Parenteral administration" means administration through injection or infusion. Parenteral administration includes subcutaneous administration, intravenous administration, intramuscular administration, intraarterial administration, intraperitoneal administration, or intracranial administration, e.g., intrathecal or intracerebroventricular administration.

[0075] "Peptide" means a molecule formed by linking at least two amino acids by amide bonds. Peptide refers to polypeptides and proteins.

[0076] "Pharmaceutical agent" means the substance or substances in a pharmaceutical composition that provide a therapeutic benefit when administered to an individual. For example, in certain instances an antisense oligonucleotide targeted to C9ORF72 is a pharmaceutical agent.

[0077] "Pharmaceutical composition" means a mixture of substances suitable for administering to an individual. For example, a pharmaceutical composition may comprise one or more pharmaceutical agents and a sterile aqueous solution.

[0078] "Pharmaceutically acceptable derivative" encompasses pharmaceutically acceptable salts, conjugates, prodrugs or isomers of the compounds described herein.

[0079] "Pharmaceutically acceptable salts" means physiologically and pharmaceutically acceptable salts of antisense compounds, *i.e.*, salts that retain the desired biological activity of the parent oligonucleotide and do not impart undesired toxicological effects thereto.

[0080] "Phosphorothioate linkage" means a linkage between nucleosides where the phosphodiester bond is modified by replacing one of the non-bridging oxygen atoms with a sulfur atom. A phosphorothioate linkage (P=S) is a modified internucleoside linkage.

[0081] "Portion" means a defined number of contiguous (*i.e.*, linked) nucleobases of a nucleic

acid. In certain embodiments, a portion is a defined number of contiguous nucleobases of a target nucleic acid. In certain embodiments, a portion is a defined number of contiguous nucleobases of an antisense compound.

[0082] "Prevent" or "preventing" refers to delaying or forestalling the onset or development of a disease, disorder, or condition for a period of time from minutes to indefinitely. Prevent also means reducing risk of developing a disease, disorder, or condition.

[0083] "Prodrug" means a therapeutic agent that is prepared in an inactive form that is converted to an active form within the body or cells thereof by the action of endogenous enzymes or other chemicals or conditions.

[0084] "Side effects" means physiological responses attributable to a treatment other than the desired effects. In certain embodiments, side effects include injection site reactions, liver function test abnormalities, renal function abnormalities, liver toxicity, renal toxicity, central nervous system abnormalities, myopathies, and malaise.

[0085] "Single-stranded oligonucleotide" means an oligonucleotide which is not hybridized to a complementary strand.

[0086] "Specifically hybridizable" refers to an antisense compound having a sufficient degree of complementarity between an antisense oligonucleotide and a target nucleic acid to induce a desired effect, while exhibiting minimal or no effects on non-target nucleic acids under conditions in which specific binding is desired, *i.e.*, under physiological conditions in the case of *in vivo* assays and therapeutic treatments.

[0087] "Targeting" or "targeted" means the process of design and selection of an antisense compound that will specifically hybridize to a target nucleic acid and induce a desired effect.

[0088] "Target nucleic acid," "target RNA," and "target RNA transcript" all refer to a nucleic acid capable of being targeted by antisense compounds.

[0089] "Target segment" means the sequence of nucleotides of a target nucleic acid to which an antisense compound is targeted. "5' target site" refers to the 5'-most nucleotide of a target segment. "3' target site" refers to the 3'-most nucleotide of a target segment.

[0090] "Therapeutically effective amount" means an amount of a pharmaceutical agent that provides a therapeutic benefit to an individual.

[0091] "Treat" or "treating" refers to administering a pharmaceutical composition to effect an alteration or improvement of a disease, disorder, or condition.

[0092] "Unmodified nucleotide" means a nucleotide composed of naturally occurring nucleobases, sugar moieties, and internucleoside linkages. In certain embodiments, an

unmodified nucleotide is an RNA nucleotide (*i.e.* β -D-ribonucleosides) or a DNA nucleotide (*i.e.* β -D-deoxyribonucleoside).

Certain Embodiments

[0093] Herein disclosed are methods for decreasing C9ORF72 mRNA and protein expression.

[0094] Herein disclosed are methods for the treatment, prevention, or amelioration of diseases, disorders, and conditions associated with C9ORF72 in an individual in need thereof. Also contemplated are methods for the preparation of a medicament for the treatment, prevention, or amelioration of a disease, disorder, or condition associated with C9ORF72. C9ORF72 associated diseases, disorders, and conditions include neurodegenerative diseases. The neurodegenerative disease may be ALS or FTD. The neurodegenerative disease may be familial or sporadic.

[0095] Herein disclosed is the use of a C9ORF72 specific inhibitor for treating, preventing, or ameliorating a C9ORF72 associated disease. Herein disclosed is the use of a C9ORF72 specific inhibitor for treating, preventing, or ameliorating a C9ORF72 hexanucleotide repeat expansion associated disease. The hexanucleotide repeat expansion may comprise GGGGCC, GGGGGG, GGGGGC, or GGGGCG. In certain instances, C9ORF72 specific inhibitors are nucleic acids (including antisense compounds), peptides, antibodies, small molecules, and other agents capable of inhibiting the expression of C9ORF72 mRNA and/or C9ORF72 protein.

[0096] Described herein are methods comprising administering a C9ORF72 or C9ORF72 homolog specific inhibitor to an animal.

[0097] Described herein are methods comprising:

identifying an animal having a C9ORF72 associated disease; and

administering a C9ORF72 specific inhibitor.

[0098] In certain embodiments, the animal is a human.

[0099] In certain instances, the C9ORF72 specific inhibitor is any of an antisense compound, an aptamer, an antibody, a peptide, or a small molecule.

[0100] In certain embodiments, the antisense compound comprises a single-stranded antisense oligonucleotide complementary to a C9ORF72 nucleic acid.

[0101] In certain embodiments, the C9ORF72 nucleic acid is a human C9ORF72 nucleic acid.

[0102] In certain embodiments, the animal has a C9ORF72 associated disease.

[0103] In certain embodiments, the C9ORF72 associated disease is a C9ORF72 hexanucleotide repeat expansion associated disease.

[0104] In certain embodiments, the C9ORF72 associated disease is amyotrophic lateral sclerosis (ALS) or frontotemporal dementia (FTD).

[0105] In certain embodiments, the C9ORF72 hexanucleotide repeat expansion associated disease is amyotrophic lateral sclerosis (ALS) or frontotemporal dementia (FTD).

[0106] In certain embodiments, the amyotrophic lateral sclerosis (ALS) is familial ALS or sporadic ALS.

[0107] In certain embodiments, the single-stranded antisense oligonucleotide comprises at least one modification.

[0108] In certain embodiments, the single-stranded antisense oligonucleotide is specifically hybridizable to a human C9ORF72 nucleic acid.

[0109] In certain instances, the single-stranded antisense oligonucleotide is at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% complementary to an equal length portion of a human C9ORF72 nucleic acid.

[0110] In certain embodiments, the single-stranded antisense oligonucleotide is 100% complementary to an equal length portion of a human C9ORF72 nucleic acid.

[0111] In certain embodiments, the single-stranded antisense oligonucleotide comprises at least one modified internucleoside linkage.

[0112] In certain embodiments, each internucleoside linkage of the single-stranded antisense oligonucleotide is a modified internucleoside linkage.

[0113] In certain embodiments, the modified internucleoside linkage is a phosphorothioate internucleoside linkage.

[0114] In certain embodiments, the single-stranded antisense oligonucleotide comprises at least one modified nucleoside.

[0115] In certain embodiments, the single-stranded antisense oligonucleotide comprises at least one modified nucleoside having a modified sugar.

[0116] In certain embodiments, the single-stranded antisense oligonucleotide comprises at

least one modified nucleoside comprising a bicyclic sugar.

[0117] In certain embodiments, the bicyclic sugar comprises a 4' to 2' bridge selected from among: 4'-(CH₂)_n-O-2' bridge, wherein n is 1 or 2; and 4'-CH₂-O-CH₂-2'.

[0118] In certain embodiments, the bicyclic sugar comprises a 4'-CH(CH₃)-O-2' bridge.

[0119] In certain embodiments, at least one modified nucleoside having a modified sugar comprises a non-bicyclic 2'-modified modified sugar moiety.

[0120] In certain embodiments, the 2'-modified sugar moiety comprises a 2'-O-methoxyethyl group.

[0121] In certain embodiments, the 2'-modified sugar moiety comprises a 2'-O-methyl group.

[0122] In certain embodiments, the at least one modified nucleoside having a modified sugar comprises a sugar surrogate.

[0123] In certain embodiments, the sugar surrogate is a morpholino.

[0124] In certain embodiments, the sugar surrogate is a peptide nucleic acid.

[0125] In certain embodiments, each nucleoside is modified.

[0126] In certain embodiments, the single-stranded antisense oligonucleotide comprises at least one modified nucleobase.

[0127] In certain embodiments, the modified nucleobase is a 5'-methylcytosine.

[0128] In certain embodiments, the single-stranded antisense oligonucleotide comprises:

a gap segment consisting of linked deoxynucleosides;

a 5' wing segment consisting of linked nucleosides;

a 3' wing segment consisting of linked nucleosides;

wherein the gap segment is positioned immediately adjacent to and between the 5' wing segment and the 3' wing segment and wherein each nucleoside of each wing segment comprises a modified sugar.

[0129] In certain embodiments the single-stranded antisense oligonucleotide comprises:

a gap segment consisting of ten linked deoxynucleosides;

a 5' wing segment consisting of five linked nucleosides;

a 3' wing segment consisting of five linked nucleosides;

wherein the gap segment is positioned immediately adjacent and between the 5' wing segment and the 3' wing segment, wherein each nucleoside of each wing segment comprises a 2'-O-methoxyethyl sugar; and wherein each internucleoside linkage is a phosphorothioate linkage.

[0130] In certain embodiments, the single-stranded antisense oligonucleotide consists of 15 linked nucleosides.

[0131] In certain embodiments, the single-stranded antisense oligonucleotide consists of 16 linked nucleosides.

[0132] In certain embodiments, the single-stranded antisense oligonucleotide consists of 17 linked nucleosides.

[0133] In certain embodiments, the single-stranded antisense oligonucleotide consists of 18 linked nucleosides.

[0134] In certain embodiments, the single-stranded antisense oligonucleotide consists of 19 linked nucleosides.

[0135] In certain embodiments, the single-stranded antisense oligonucleotide consists of 20 linked nucleosides.

[0136] In certain embodiments, the single-stranded antisense oligonucleotide consists of 21 linked nucleosides.

[0137] In certain embodiments, the single-stranded antisense oligonucleotide consists of 22 linked nucleosides.

[0138] In certain embodiments, the single-stranded antisense oligonucleotide consists of 23 linked nucleosides.

[0139] In certain embodiments, the single-stranded antisense oligonucleotide consists of 24 linked nucleosides.

[0140] In certain embodiments, the single-stranded antisense oligonucleotide consists of 25 linked nucleosides.

[0141] In certain embodiments, the single-stranded antisense oligonucleotide complementary

to a C9ORF72 nucleic acid is complementary to any of an intron, an exon, a splice junction, an exon:exon junction, a repeat, the 3' UTR, or the 5' UTR of the C9ORF72 nucleic acid.

[0142] In certain embodiments, the intron is any of introns 1, intron 2, intron 3, intron 4, intron 5, intron 6, intron 7, intron 8, intron 9, or intron 10.

[0143] In certain embodiments, the exon is any of exon 1a, exon 1b, exon 1c, exon 1d, exon 1d, exon 2, exon 3, exon 4, exon 5, exon 6, exon 7, exon 8, exon 9, exon 10, or exon 11.

[0144] In certain embodiments, the repeat is a hexanucleotide repeat.

[0145] In certain embodiments, the hexanucleotide repeat is any of GGGGCC, GGGGGG, GGGGCG, or GGGGGC.

[0146] In certain embodiments, expression of C9ORF72 RNA is reduced.

[0147] In certain embodiments, expression of C9ORF72 RNA is reduced in a nucleus.

[0148] In certain embodiments, expression of C9ORF72 protein is reduced.

[0149] In certain embodiments, expression of C9ORF72 protein is reduced in a nucleus.

[0150] In certain embodiments, expression is reduced in any of fibroblasts, immortalized lymphoblasts, spinal motor neurons, purkinje and granular cells in the cerebellum, cortical neurons, astrocytes, or microglia.

[0151] In certain embodiments, expression of C9ORF72 isoform V1 RNA is reduced.

[0152] In certain embodiments, expression of C9ORF72 isoform V1 protein is reduced.

[0153] In certain embodiments, expression of C9ORF72 isoform V2 RNA is reduced.

[0154] In certain embodiments, expression of C9ORF72 isoform V2 protein is reduced.

[0155] In certain embodiments, expression of C9ORF72 isoform V3 RNA is reduced.

[0156] In certain embodiments, expression of C9ORF72 isoform V3 protein is reduced.

[0157] In certain embodiments, expression of C9ORF72 isoform V2 RNA and expression of C9ORF72 isoform V3 RNA is reduced and C9ORF72 isoform V1 RNA is not reduced.

[0158] In certain embodiments, expression of C9ORF72 isoform V2 protein and expression of C9ORF72 isoform V3 protein is reduced and C9ORF72 isoform V1 protein is not reduced.

[0159] In certain instances, C9ORF72 RNA foci are reduced.

[0160] In the invention, C9ORF72 RNA foci are reduced in a nucleus.

[0161] In certain instances, compositions described herein are administered by parenteral administration.

[0162] In certain instances, the parenteral administration is any of injection or infusion.

[0163] In certain instances, the parenteral administration is any of intrathecal administration or intracerebroventricular administration.

[0164] In certain embodiments, at least one symptom of a C9ORF72 associated disease is ameliorated or prevented.

[0165] In certain embodiments, at least one symptom of a C9ORF72 hexanucleotide repeat expansion associated disease is ameliorated or prevented.

[0166] In certain embodiments, progression of at least one symptom of a C9ORF72 associated disease is slowed.

[0167] In certain embodiments, progression of at least one symptom of a C9ORF72 hexanucleotide repeat expansion associated disease is slowed.

[0168] In certain embodiments, the at least one symptom is any of motor function, respiration, muscle weakness, fasciculation and cramping of muscles, difficulty in projecting the voice, shortness of breath, difficulty in breathing and swallowing, inappropriate social behavior, lack of empathy, distractibility, changes in food preferences, agitation, blunted emotions, neglect of personal hygiene, repetitive or compulsive behavior, and decreased energy and motivation.

Antisense Compounds

[0169] Oligomeric compounds include, but are not limited to, oligonucleotides, oligonucleosides, oligonucleotide analogs, oligonucleotide mimetics, antisense compounds, antisense oligonucleotides, and siRNAs. An oligomeric compound may be "antisense" to a target nucleic acid, meaning that it is capable of undergoing hybridization to a target nucleic acid through hydrogen bonding.

[0170] In certain embodiments, an antisense compound has a nucleobase sequence that, when written in the 5' to 3' direction, comprises the reverse complement of the target segment of a target nucleic acid to which it is targeted. In certain such embodiments, an antisense oligonucleotide has a nucleobase sequence that, when written in the 5' to 3' direction, comprises the reverse complement of the target segment of a target nucleic acid to which it is

targeted.

[0171] In the invention, an antisense compound targeted to a C9ORF72 nucleic acid is 12 to 30 subunits in length. In certain instances, the antisense compound is 8 to 80, 12 to 50, 15 to 30, 18 to 24, 19 to 22, or 20 linked subunits. In certain instances, the antisense compounds are 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, or 80 linked subunits in length, or a range defined by any two of the above values. In the invention, the antisense compound is a modified antisense oligonucleotide, and the linked subunits are nucleosides.

[0172] In certain embodiments antisense oligonucleotides targeted to a C9ORF72 nucleic acid may be shortened or truncated. For example, a single subunit may be deleted from the 5' end (5' truncation), or alternatively from the 3' end (3' truncation). A shortened or truncated antisense compound targeted to a C9ORF72 nucleic acid may have two subunits deleted from the 5' end, or alternatively may have two subunits deleted from the 3' end, of the antisense compound. Alternatively, the deleted nucleosides may be dispersed throughout the antisense compound, for example, in an antisense compound having one nucleoside deleted from the 5' end and one nucleoside deleted from the 3' end.

[0173] When a single additional subunit is present in a lengthened antisense compound, the additional subunit may be located at the 5' or 3' end of the antisense compound. When two or more additional subunits are present, the added subunits may be adjacent to each other, for example, in an antisense compound having two subunits added to the 5' end (5' addition), or alternatively to the 3' end (3' addition), of the antisense compound. Alternatively, the added subunits may be dispersed throughout the antisense compound, for example, in an antisense compound having one subunit added to the 5' end and one subunit added to the 3' end.

[0174] It is possible to increase or decrease the length of an antisense compound, such as an antisense oligonucleotide, and/or introduce mismatch bases without eliminating activity. For example, in Woolf et al. (Proc. Natl. Acad. Sci. USA 89:7305-7309, 1992), a series of antisense oligonucleotides 13-25 nucleobases in length were tested for their ability to induce cleavage of a target RNA in an oocyte injection model. Antisense oligonucleotides 25 nucleobases in length with 8 or 11 mismatch bases near the ends of the antisense oligonucleotides were able to direct specific cleavage of the target mRNA, albeit to a lesser extent than the antisense oligonucleotides that contained no mismatches. Similarly, target specific cleavage was achieved using 13 nucleobase antisense oligonucleotides, including those with 1 or 3 mismatches.

[0175] Gautschi et al (J. Natl. Cancer Inst. 93:463-471, March 2001) demonstrated the ability of an oligonucleotide having 100% complementarity to the bcl-2 mRNA and having 3 mismatches to the bcl-xL mRNA to reduce the expression of both bcl-2 and bcl-xL *in vitro* and *in vivo*. Furthermore, this oligonucleotide demonstrated potent anti-tumor activity *in vivo*.

[0176] Maher and Dolnick (Nuc. Acid. Res. 16:3341-3358,1988) tested a series of tandem 14 nucleobase antisense oligonucleotides, and a 28 and 42 nucleobase antisense oligonucleotides comprised of the sequence of two or three of the tandem antisense oligonucleotides, respectively, for their ability to arrest translation of human DHFR in a rabbit reticulocyte assay. Each of the three 14 nucleobase antisense oligonucleotides alone was able to inhibit translation, albeit at a more modest level than the 28 or 42 nucleobase antisense oligonucleotides.

Antisense Compound Motifs

[0177] In certain embodiments, antisense compounds targeted to a C9ORF72 nucleic acid have chemically modified subunits arranged in patterns, or motifs, to confer to the antisense compounds properties such as enhanced inhibitory activity, increased binding affinity for a target nucleic acid, or resistance to degradation by *in vivo* nucleases.

[0178] Chimeric antisense compounds typically contain at least one region modified so as to confer increased resistance to nuclease degradation, increased cellular uptake, increased binding affinity for the target nucleic acid, and/or increased inhibitory activity. A second region of a chimeric antisense compound may optionally serve as a substrate for the cellular endonuclease RNase H, which cleaves the RNA strand of an RNA:DNA duplex.

[0179] Antisense compounds having a gapmer motif are considered chimeric antisense compounds. In a gapmer an internal region having a plurality of nucleotides that supports RNaseH cleavage is positioned between external regions having a plurality of nucleotides that are chemically distinct from the nucleosides of the internal region. In the case of an antisense oligonucleotide having a gapmer motif, the gap segment generally serves as the substrate for endonuclease cleavage, while the wing segments comprise modified nucleosides. In certain embodiments, the regions of a gapmer are differentiated by the types of sugar moieties comprising each distinct region. The types of sugar moieties that are used to differentiate the regions of a gapmer may in some embodiments include β -D-ribonucleosides, β -D-deoxyribonucleosides, 2'-modified nucleosides (such 2'-modified nucleosides may include 2'-MOE, and 2'-O-CH₃, among others), and bicyclic sugar modified nucleosides (such bicyclic sugar modified nucleosides may include those having a 4'-(CH₂)_n-O-2' bridge, where n=1 or n=2 and 4'-CH₂-O-CH₂-2'). Preferably, each distinct region comprises uniform sugar moieties. The wing-gap-wing motif is frequently described as "X-Y-Z", where "X" represents the length of the 5' wing region, "Y" represents the length of the gap region, and "Z" represents the length of the 3' wing region. As used herein, a gapmer described as "X-Y-Z" has a configuration such that the gap segment is positioned immediately adjacent to each of the 5' wing segment and the 3' wing segment. Thus, no intervening nucleotides exist between the 5' wing segment and gap segment, or the gap segment and the 3' wing segment. Any of the antisense compounds described herein can have a gapmer motif. In some embodiments, X and Z are the same, in other embodiments they are different. In a preferred embodiment, Y is between 8 and 15

nucleotides. X, Y or Z can be any of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30 or more nucleotides. Thus, gapmers described herein include, but are not limited to, for example 5-10-5, 5-10-4, 4-10-4, 4-10-3, 3-10-3, 2-10-2, 5-9-5, 5-9-4, 4-9-5, 5-8-5, 5-8-4, 4-8-5, 5-7-5, 4-7-5, 5-7-4, or 4-7-4.

[0180] In certain embodiments, the antisense compound has a "wingmer" motif, having a wing-gap or gap-wing configuration, i.e. an X-Y or Y-Z configuration as described above for the gapmer configuration. Thus, wingmer configurations described herein include, but are not limited to, for example 5-10, 8-4, 4-12, 12-4, 3-14, 16-2, 18-1, 10-3, 2-10, 1-10, 8-2, 2-13, 5-13, 5-8, or 6-8.

[0181] In certain embodiments, antisense compounds targeted to a C9ORF72 nucleic acid possess a 5-10-5 gapmer motif.

[0182] In certain embodiments, antisense compounds targeted to a C9ORF72 nucleic acid possess a 5-10-4 gapmer motif.

[0183] In certain embodiments, antisense compounds targeted to a C9ORF72 nucleic acid possess a 4-10-4 gapmer motif.

[0184] In certain embodiments, antisense compounds targeted to a C9ORF72 nucleic acid possess a 4-10-3 gapmer motif.

[0185] In certain embodiments, antisense compounds targeted to a C9ORF72 nucleic acid possess a 5-9-5 gapmer motif.

[0186] In certain embodiments, an antisense compound targeted to a C9ORF72 nucleic acid has a gap-narrowed motif. In certain embodiments, a gap-narrowed antisense oligonucleotide targeted to a C9ORF72 nucleic acid has a gap segment of 9, 8, 7, or 6 2'-deoxynucleotides positioned immediately adjacent to and between wing segments of 5, 4, 3, 2, or 1 chemically modified nucleosides. In certain embodiments, the chemical modification comprises a bicyclic sugar. In certain embodiments, the bicyclic sugar comprises a 4' to 2' bridge selected from among: 4'-(CH₂)_n-O-2' bridge, wherein n is 1 or 2; and 4'-CH₂-O-CH₂-2'. In certain embodiments, the bicyclic sugar is comprises a 4'-CH(CH₃)-O-2' bridge. In certain embodiments, the chemical modification comprises a non-bicyclic 2'-modified sugar moiety. In certain embodiments, the non-bicyclic 2'-modified sugar moiety comprises a 2'-O-methylethyl group or a 2'-O-methyl group.

[0187] In certain embodiments, an antisense compound targeted to a C9ORF72 nucleic acid is uniformly modified. In certain embodiments, the antisense compound comprises 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleosides. In certain embodiments, each nucleoside is chemically modified. In certain embodiments, the chemical modification comprises a non-bicyclic 2'-modified sugar moiety. In certain embodiments, the 2'-modified sugar moiety comprises a 2'-O-methoxyethyl group. In certain embodiments, the 2'-modified

sugar moiety comprises a 2'-O-methyl group. In certain embodiments, uniformly modified antisense compound may target C9ORF72, or any portion thereof, such as a hexanucleotide repeat expansion. In certain embodiments, targeting the hexanucleotide repeat expansion with a uniformly modified antisense compound reduced the repeat RNA by blocking the interaction with RNA binding proteins. In certain embodiments, this results in the toxic RNA being absent from foci and being degraded instead.

Target Nucleic Acids, Target Regions and Nucleotide Sequences

[0188] Nucleotide sequences that encode C9ORF72 include, without limitation, the following: the complement of GENBANK Accession No. NM_001256054.1 (incorporated herein as SEQ ID NO: 1), GENBANK Accession No. NT_008413.18 truncated from nucleobase 27535000 to 27565000 (incorporated herein as SEQ ID NO: 2), GENBANK Accession No. BQ068108.1 (incorporated herein as SEQ ID NO: 3), GENBANK Accession No. NM_018325.3 (incorporated herein as SEQ ID NO: 4), GENBANK Accession No. DN993522.1 (incorporated herein as SEQ ID NO: 5), GENBANK Accession No. NM_145005.5 (incorporated herein as SEQ ID NO: 6), GENBANK Accession No. DB079375.1 (incorporated herein as SEQ ID NO: 7), GENBANK Accession No. BU194591.1 (incorporated herein as SEQ ID NO: 8), Sequence Identifier 4141_014_A (incorporated herein as SEQ ID NO: 9), and Sequence Identifier 4008_73_A (incorporated herein as SEQ ID NO: 10).

[0189] It is understood that the sequence set forth in each SEQ ID NO in the Examples contained herein is independent of any modification to a sugar moiety, an internucleoside linkage, or a nucleobase. As such, antisense compounds defined by a SEQ ID NO may comprise, independently, one or more modifications to a sugar moiety, an internucleoside linkage, or a nucleobase. Antisense compounds described by Isis Number (Isis No) indicate a combination of nucleobase sequence and motif.

[0190] In certain embodiments, a target region is a structurally defined region of the target nucleic acid. For example, a target region may encompass a 3' UTR, a 5' UTR, an exon, an intron, an exon/intron junction, a coding region, a translation initiation region, translation termination region, or other defined nucleic acid region. The structurally defined regions for C9ORF72 can be obtained by accession number from sequence databases such as NCBI. In certain embodiments, a target region may encompass the sequence from a 5' target site of one target segment within the target region to a 3' target site of another target segment within the same target region.

[0191] Targeting includes determination of at least one target segment to which an antisense compound hybridizes, such that a desired effect occurs. In certain embodiments, the desired effect is a reduction in mRNA target nucleic acid levels. In certain embodiments, the desired effect is reduction of levels of protein encoded by the target nucleic acid or a phenotypic change associated with the target nucleic acid.

[0192] A target region may contain one or more target segments. Multiple target segments within a target region may be overlapping. Alternatively, they may be non-overlapping. In certain embodiments, target segments within a target region are separated by no more than about 300 nucleotides. In certain embodiments, target segments within a target region are separated by a number of nucleotides that is, is about, is no more than, is no more than about, 250, 200, 150, 100, 90, 80, 70, 60, 50, 40, 30, 20, or 10 nucleotides on the target nucleic acid, or is a range defined by any two of the preceding values. In certain embodiments, target segments within a target region are separated by no more than, or no more than about, 5 nucleotides on the target nucleic acid. In certain embodiments, target segments are contiguous. Contemplated are target regions defined by a range having a starting nucleic acid that is any of the 5' target sites or 3' target sites listed herein.

[0193] Suitable target segments may be found within a 5' UTR, a coding region, a 3' UTR, an intron, an exon, or an exon/intron junction. Target segments containing a start codon or a stop codon are also suitable target segments. A suitable target segment may specifically exclude a certain structurally defined region such as the start codon or stop codon.

[0194] The determination of suitable target segments may include a comparison of the sequence of a target nucleic acid to other sequences throughout the genome. For example, the BLAST algorithm may be used to identify regions of similarity amongst different nucleic acids. This comparison can prevent the selection of antisense compound sequences that may hybridize in a non-specific manner to sequences other than a selected target nucleic acid (i.e., non-target or off-target sequences).

[0195] There may be variation in activity (e.g., as defined by percent reduction of target nucleic acid levels) of the antisense compounds within a target region. In certain embodiments, reductions in C9ORF72 mRNA levels are indicative of inhibition of C9ORF72 expression. Reductions in levels of a C9ORF72 protein are also indicative of inhibition of target mRNA expression. Reduction in the presence of expanded C9ORF72 RNA foci are indicative of inhibition of C9ORF72 expression. Further, phenotypic changes are indicative of inhibition of C9ORF72 expression. For example, improved motor function and respiration may be indicative of inhibition of C9ORF72 expression.

Hybridization

[0196] Hybridization occurs between an antisense compound disclosed herein and a C9ORF72 nucleic acid. The most common mechanism of hybridization involves hydrogen bonding (e.g., Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding) between complementary nucleobases of the nucleic acid molecules.

[0197] Hybridization can occur under varying conditions. Stringent conditions are sequence-dependent and are determined by the nature and composition of the nucleic acid molecules to

be hybridized.

[0198] Methods of determining whether a sequence is specifically hybridizable to a target nucleic acid are well known in the art. The antisense compounds provided herein are specifically hybridizable with a C9ORF72 nucleic acid.

Complementarity

[0199] An antisense compound and a target nucleic acid are complementary to each other when a sufficient number of nucleobases of the antisense compound can hydrogen bond with the corresponding nucleobases of the target nucleic acid, such that a desired effect will occur (e.g., antisense inhibition of a target nucleic acid, such as a C9ORF72 nucleic acid).

[0200] Non-complementary nucleobases between an antisense compound and a C9ORF72 nucleic acid may be tolerated provided that the antisense compound remains able to specifically hybridize to a target nucleic acid. Moreover, an antisense compound may hybridize over one or more segments of a C9ORF72 nucleic acid such that intervening or adjacent segments are not involved in the hybridization event (e.g., a loop structure, mismatch or hairpin structure).

[0201] In certain instances, the antisense compounds disclosed herein, or a specified portion thereof, are, or are at least, 70%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% complementary to a C9ORF72 nucleic acid, a target region, target segment, or specified portion thereof. Percent complementarity of an antisense compound with a target nucleic acid can be determined using routine methods.

[0202] For example, an antisense compound in which 18 of 20 nucleobases of the antisense compound are complementary to a target region, and would therefore specifically hybridize, would represent 90 percent complementarity. In this example, the remaining noncomplementary nucleobases may be clustered or interspersed with complementary nucleobases and need not be contiguous to each other or to complementary nucleobases. As such, an antisense compound which is 18 nucleobases in length having 4 (four) noncomplementary nucleobases which are flanked by two regions of complete complementarity with the target nucleic acid would have 77.8% overall complementarity with the target nucleic acid. Percent complementarity of an antisense compound with a region of a target nucleic acid can be determined routinely using BLAST programs (basic local alignment search tools) and PowerBLAST programs known in the art (Altschul et al., J. Mol. Biol., 1990, 215, 403 410; Zhang and Madden, Genome Res., 1997, 7, 649 656). Percent homology, sequence identity or complementarity, can be determined by, for example, the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison Wis.), using default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482 489).

[0203] In certain embodiments, the antisense compounds provided herein, or specified portions thereof, are fully complementary (*i.e.*, 100% complementary) to a target nucleic acid, or specified portion thereof. For example, an antisense compound may be fully complementary to a C9ORF72 nucleic acid, or a target region, or a target segment or target sequence thereof. As used herein, "fully complementary" means each nucleobase of an antisense compound is capable of precise base pairing with the corresponding nucleobases of a target nucleic acid. For example, a 20 nucleobase antisense compound is fully complementary to a target sequence that is 400 nucleobases long, so long as there is a corresponding 20 nucleobase portion of the target nucleic acid that is fully complementary to the antisense compound. Fully complementary can also be used in reference to a specified portion of the first and /or the second nucleic acid. For example, a 20 nucleobase portion of a 30 nucleobase antisense compound can be "fully complementary" to a target sequence that is 400 nucleobases long. The 20 nucleobase portion of the 30 nucleobase oligonucleotide is fully complementary to the target sequence if the target sequence has a corresponding 20 nucleobase portion wherein each nucleobase is complementary to the 20 nucleobase portion of the antisense compound. At the same time, the entire 30 nucleobase antisense compound may or may not be fully complementary to the target sequence, depending on whether the remaining 10 nucleobases of the antisense compound are also complementary to the target sequence.

[0204] The location of a non-complementary nucleobase may be at the 5' end or 3' end of the antisense compound. Alternatively, the non-complementary nucleobase or nucleobases may be at an internal position of the antisense compound. When two or more non-complementary nucleobases are present, they may be contiguous (*i.e.*, linked) or non-contiguous. In one embodiment, a non-complementary nucleobase is located in the wing segment of a gapmer antisense oligonucleotide.

[0205] In certain instances, antisense compounds that are, or are up to 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleobases in length comprise no more than 4, no more than 3, no more than 2, or no more than 1 non-complementary nucleobase(s) relative to a target nucleic acid, such as a C9ORF72 nucleic acid, or specified portion thereof.

[0206] In certain instances, antisense compounds that are, or are up to 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleobases in length comprise no more than 6, no more than 5, no more than 4, no more than 3, no more than 2, or no more than 1 non-complementary nucleobase(s) relative to a target nucleic acid, such as a C9ORF72 nucleic acid, or specified portion thereof.

[0207] The antisense compounds provided herein also include those which are complementary to a portion of a target nucleic acid. As used herein, "portion" refers to a defined number of contiguous (*i.e.* linked) nucleobases within a region or segment of a target nucleic acid. A "portion" can also refer to a defined number of contiguous nucleobases of an antisense compound. In certain instances, the antisense compounds, are complementary to at least an 8 nucleobase portion of a target segment. In certain instances, the antisense compounds are complementary to at least a 9 nucleobase portion of a target segment. In

certain instances, the antisense compounds are complementary to at least a 10 nucleobase portion of a target segment. In certain embodiments, the antisense compounds, are complementary to at least an 11 nucleobase portion of a target segment. In certain embodiments, the antisense compounds, are complementary to at least a 12 nucleobase portion of a target segment. In certain embodiments, the antisense compounds, are complementary to at least a 13 nucleobase portion of a target segment. In certain embodiments, the antisense compounds, are complementary to at least a 14 nucleobase portion of a target segment. In certain embodiments, the antisense compounds, are complementary to at least a 15 nucleobase portion of a target segment. Also contemplated are antisense compounds that are complementary to at least a 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more nucleobase portion of a target segment, or a range defined by any two of these values.

Identity

[0208] The antisense compounds provided herein may also have a defined percent identity to a particular nucleotide sequence, SEQ ID NO, or compound represented by a specific Isis number, or portion thereof. As used herein, an antisense compound is identical to the sequence disclosed herein if it has the same nucleobase pairing ability. For example, a RNA which contains uracil in place of thymidine in a disclosed DNA sequence would be considered identical to the DNA sequence since both uracil and thymidine pair with adenine. Shortened and lengthened versions of the antisense compounds described herein as well as compounds having non-identical bases relative to the antisense compounds provided herein also are contemplated. The non-identical bases may be adjacent to each other or dispersed throughout the antisense compound. Percent identity of an antisense compound is calculated according to the number of bases that have identical base pairing relative to the sequence to which it is being compared.

[0209] In certain embodiments, the antisense compounds, or portions thereof, are at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to one or more of the antisense compounds or SEQ ID NOs, or a portion thereof, disclosed herein.

[0210] In certain embodiments, a portion of the antisense compound is compared to an equal length portion of the target nucleic acid. In certain embodiments, an 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleobase portion is compared to an equal length portion of the target nucleic acid.

[0211] In certain embodiments, a portion of the antisense oligonucleotide is compared to an equal length portion of the target nucleic acid. In certain embodiments, an 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleobase portion is compared to an equal length portion of the target nucleic acid.

Modifications

[0212] A nucleoside is a base-sugar combination. The nucleobase (also known as base) portion of the nucleoside is normally a heterocyclic base moiety. Nucleotides are nucleosides that further include a phosphate group covalently linked to the sugar portion of the nucleoside. For those nucleosides that include a pentofuranosyl sugar, the phosphate group can be linked to the 2', 3' or 5' hydroxyl moiety of the sugar. Oligonucleotides are formed through the covalent linkage of adjacent nucleosides to one another, to form a linear polymeric oligonucleotide. Within the oligonucleotide structure, the phosphate groups are commonly referred to as forming the internucleoside linkages of the oligonucleotide.

[0213] Modifications to antisense compounds encompass substitutions or changes to internucleoside linkages, sugar moieties, or nucleobases. Modified antisense compounds are often preferred over native forms because of desirable properties such as, for example, enhanced cellular uptake, enhanced affinity for nucleic acid target, increased stability in the presence of nucleases, or increased inhibitory activity.

[0214] Chemically modified nucleosides may also be employed to increase the binding affinity of a shortened or truncated antisense oligonucleotide for its target nucleic acid. Consequently, comparable results can often be obtained with shorter antisense compounds that have such chemically modified nucleosides.

Modified Internucleoside Linkages

[0215] The naturally occurring internucleoside linkage of RNA and DNA is a 3' to 5' phosphodiester linkage. Antisense compounds having one or more modified, i.e. non-naturally occurring, internucleoside linkages are often selected over antisense compounds having naturally occurring internucleoside linkages because of desirable properties such as, for example, enhanced cellular uptake, enhanced affinity for target nucleic acids, and increased stability in the presence of nucleases.

[0216] Oligonucleotides having modified internucleoside linkages include internucleoside linkages that retain a phosphorus atom as well as internucleoside linkages that do not have a phosphorus atom. Representative phosphorus containing internucleoside linkages include, but are not limited to, phosphodiesters, phosphotriesters, methylphosphonates, phosphoramidate, and phosphorothioates. Methods of preparation of phosphorous-containing and non-phosphorous-containing linkages are well known.

[0217] In certain embodiments, antisense compounds targeted to a C9ORF72 nucleic acid comprise one or more modified internucleoside linkages. In certain embodiments, the modified internucleoside linkages are interspersed throughout the antisense compound. In certain embodiments, the modified internucleoside linkages are phosphorothioate linkages. In certain

embodiments, each internucleoside linkage of an antisense compound is a phosphorothioate internucleoside linkage.

Modified Sugar Moieties

[0218] Antisense compounds can optionally contain one or more nucleosides wherein the sugar group has been modified. Such sugar modified nucleosides may impart enhanced nuclease stability, increased binding affinity, or some other beneficial biological property to the antisense compounds. In certain embodiments, nucleosides comprise chemically modified ribofuranose ring moieties. Examples of chemically modified ribofuranose rings include without limitation, addition of substituent groups (including 5' and 2' substituent groups, bridging of non-geminal ring atoms to form bicyclic nucleic acids (BNA), replacement of the ribosyl ring oxygen atom with S, N(R), or C(R₁)(R₂) (R, R₁ and R₂ are each independently H, C₁-C₁₂ alkyl or a protecting group) and combinations thereof. Examples of chemically modified sugars include 2'-F-5'-methyl substituted nucleoside (see PCT International Application WO 2008/101157 Published on 8/21/08 for other disclosed 5',2'-bis substituted nucleosides) or replacement of the ribosyl ring oxygen atom with S with further substitution at the 2'-position (see published U.S. Patent Application US2005-0130923, published on June 16, 2005) or alternatively 5'-substitution of a BNA (see PCT International Application WO 2007/134181 Published on 11/22/07 wherein LNA is substituted with for example a 5'-methyl or a 5'-vinyl group).

[0219] Examples of nucleosides having modified sugar moieties include without limitation nucleosides comprising 5'-vinyl, 5'-methyl (R or S), 4'-S, 2'-F, 2'-OCH₃, 2'-OCH₂CH₃, 2'-OCH₂CH₂F and 2'-O(CH₂)₂OCH₃ substituent groups. The substituent at the 2' position can also be selected from allyl, amino, azido, thio, O-allyl, O-C₁-C₁₀ alkyl, OCF₃, OCH₂F, O(CH₂)₂SCH₃, O(CH₂)₂-O-N(R_m)(R_n), O-CH₂-C(=O)-N(R_m)(R_n), and O-CH₂-C(=O)-N(R_l)-(CH₂)₂-N(R_m)(R_n), where each R_l, R_m and R_n is, independently, H or substituted or unsubstituted C₁-C₁₀ alkyl.

[0220] As used herein, "bicyclic nucleosides" refer to modified nucleosides comprising a bicyclic sugar moiety. Examples of bicyclic nucleosides include without limitation nucleosides comprising a bridge between the 4' and the 2' ribosyl ring atoms. In certain embodiments, antisense compounds provided herein include one or more bicyclic nucleosides comprising a 4' to 2' bridge. Examples of such 4' to 2' bridged bicyclic nucleosides, include but are not limited to one of the formulae: 4'-(CH₂)-O-2' (LNA); 4'-(CH₂)-S-2'; 4'-(CH₂)₂-O-2' (ENA); 4'-CH(CH₃)-O-2' and 4'-CH(CH₂OCH₃)-O-2' (and analogs thereof see U.S. Patent 7,399,845, issued on July 15, 2008); 4'-C(CH₃)(CH₃)-O-2' (and analogs thereof see published International Application WO/2009/006478, published January 8, 2009); 4'-CH₂-N(OCH₃)-2' (and analogs thereof see published International Application WO/2008/150729, published December 11, 2008); 4'-CH₂-O-N(CH₃)-2' (see published U.S. Patent Application US2004-0171570, published

September 2, 2004); 4'-CH₂-N(R)-O-2', wherein R is H, C₁-C₁₂ alkyl, or a protecting group (see U.S. Patent 7,427,672, issued on September 23, 2008); 4'-CH₂-C(H)(CH₃)-2' (see Chattopadhyaya et al., J. Org. Chem., 2009, 74, 118-134); and 4'-CH₂-C(=CH₂)-2' (and analogs thereof see published International Application WO 2008/154401, published on December 8, 2008).

[0221] Further reports related to bicyclic nucleosides can also be found in published literature (see for example: Singh et al., Chem. Commun., 1998, 4, 455-456; Koshkin et al., Tetrahedron, 1998, 54, 3607-3630; Wahlestedt et al., Proc. Natl. Acad. Sci. U. S. A., 2000, 97, 5633-5638; Kumar et al., Bioorg. Med. Chem. Lett., 1998, 8, 2219-2222; Singh et al., J. Org. Chem., 1998, 63, 10035-10039; Srivastava et al., J. Am. Chem. Soc., 2007, 129(26) 8362-8379; Elayadi et al., Curr. Opinion Invest. Drugs, 2001, 2, 558-561; Braasch et al., Chem. Biol., 2001, 8, 1-7; and Orum et al., Curr. Opinion Mol. Ther., 2001, 3, 239-243; U.S. Patent Nos. 6,268,490; 6,525,191; 6,670,461; 6,770,748; 6,794,499; 7,034,133; 7,053,207; 7,399,845; 7,547,684; and 7,696,345; U.S. Patent Publication No. US2008-0039618; US2009-0012281; U.S. Patent Serial Nos. 60/989,574; 61/026,995; 61/026,998; 61/056,564; 61/086,231; 61/097,787; and 61/099,844; Published PCT International applications WO 1994/014226; WO 2004/106356; WO 2005/021570; WO 2007/134181; WO 2008/150729; WO 2008/154401; and WO 2009/006478. Each of the foregoing bicyclic nucleosides can be prepared having one or more stereochemical sugar configurations including for example α -L-ribofuranose and β -D-ribofuranose (see PCT international application PCT/DK98/00393, published on March 25, 1999 as WO 99/14226).

[0222] In certain embodiments, bicyclic sugar moieties of BNA nucleosides include, but are not limited to, compounds having at least one bridge between the 4' and the 2' position of the pentofuranosyl sugar moiety wherein such bridges independently comprises 1 or from 2 to 4 linked groups independently selected from -[C(R_a)(R_b)]_n-, -C(R_a)=C(R_b)-, -C(R_a)=N-, -C(=O)-, -C(=NR_a)-, -C(=S)-, -O-, -Si(R_a)₂-, -S(=O)_x-, and -N(R_a)-;

wherein:

x is 0, 1, or 2;

n is 1, 2, 3, or 4;

each R_a and R_b is, independently, H, a protecting group, hydroxyl, C₁-C₁₂ alkyl, substituted C₁-C₁₂ alkyl, C₂-C₁₂ alkenyl, substituted C₂-C₁₂ alkenyl, C₂-C₁₂ alkynyl, substituted C₂-C₁₂ alkynyl, C₅-C₂₀ aryl, substituted C₅-C₂₀ aryl, heterocycle radical, substituted heterocycle radical, heteroaryl, substituted heteroaryl, C₅-C₇ alicyclic radical, substituted C₅-C₇ alicyclic radical, halogen, OJ₁, NJ₁J₂, SJ₁, N₃, COOJ₁, acyl (C(=O)-H), substituted acyl, CN, sulfonyl (S(=O)₂-J₁), or sulfoxyl (S(=O)-J₁); and

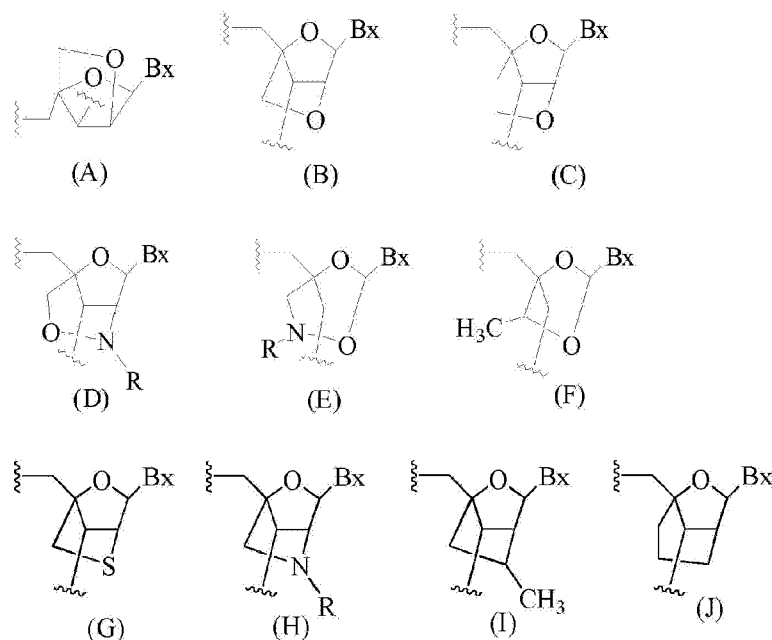
each J₁ and J₂ is, independently, H, C₁-C₁₂ alkyl, substituted C₁-C₁₂ alkyl, C₂-C₁₂ alkenyl, substituted C₂-C₁₂ alkenyl, C₂-C₁₂ alkynyl, substituted C₂-C₁₂ alkynyl, C₅-C₂₀ aryl, substituted

C₅-C₂₀ aryl, acyl (C(=O)-H), substituted acyl, a heterocycle radical, a substituted heterocycle radical, C₁-C₁₂ aminoalkyl, substituted C₁-C₁₂ aminoalkyl or a protecting group.

[0223] In certain embodiments, the bridge of a bicyclic sugar moiety is $-\text{C}(\text{R}_a)(\text{R}_b)]_n-$, $-\text{C}(\text{R}_a)(\text{R}_b)]_n\text{-O-}$, $-\text{C}(\text{R}_a\text{R}_b)\text{-N(R)-O-}$ or $-\text{C}(\text{R}_a\text{R}_b)\text{-O-N(R)-}$. In certain embodiments, the bridge is 4'-CH₂-2', 4'-(CH₂)₂-2', 4'-(CH₂)₃-2', 4'-CH₂-O-2', 4'-(CH₂)₂-O-2', 4'-CH₂-O-N(R)-2' and 4'-CH₂-N(R)-O-2' wherein each R is, independently, H, a protecting group or C₁-C₁₂ alkyl.

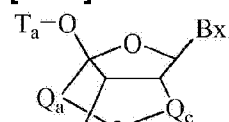
[0224] In certain embodiments, bicyclic nucleosides are further defined by isomeric configuration. For example, a nucleoside comprising a 4'-2' methylene-oxy bridge, may be in the α -L configuration or in the β -D configuration. Previously, α -L-methyleneoxy (4'-CH₂-O-2') BNA's have been incorporated into antisense oligonucleotides that showed antisense activity (Frieden et al., Nucleic Acids Research, 2003, 21, 6365-6372).

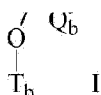
[0225] In certain embodiments, bicyclic nucleosides include, but are not limited to, (A) α -L-methyleneoxy (4'-CH₂-O-2') BNA, (B) β -D-methyleneoxy (4'-CH₂-O-2') BNA, (C) ethyleneoxy (4'-(CH₂)₂-O-2') BNA, (D) aminooxy (4'-CH₂-O-N(R)-2') BNA, (E) oxyamino (4'-CH₂-N(R)-O-2') BNA, and (F) methyl(methyleneoxy) (4'-CH(CH₃)-O-2') BNA, (G) methylene-thio (4'-CH₂-S-2') BNA, (H) methylene-amino (4'-CH₂-N(R)-2') BNA, (I) methyl carbocyclic (4'-CH₂-CH(CH₃)-2') BNA, and (J) propylene carbocyclic (4'-(CH₂)₃-2') BNA as depicted below.



wherein Bx is the base moiety and R is independently H, a protecting group or C₁-C₁₂ alkyl.

[0226] In certain embodiments, bicyclic nucleosides are provided having Formula I:





wherein:

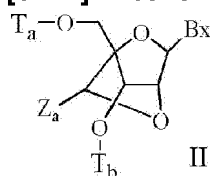
Bx is a heterocyclic base moiety;

$-Q_a-Q_b-Q_c-$ is $-\text{CH}_2-\text{N}(\text{R}_c)-\text{CH}_2-$, $-\text{C}(=\text{O})-\text{N}(\text{R}_c)-\text{CH}_2-$, $-\text{CH}_2-\text{O}-\text{N}(\text{R}_c)-$, $-\text{CH}_2-\text{N}(\text{R}_c)-\text{O}-$ or $-\text{N}(\text{R}_c)-\text{O}-\text{CH}_2-$;

R_c is C_1 - C_{12} alkyl or an amino protecting group; and

T_a and T_b are each, independently H, a hydroxyl protecting group, a conjugate group, a reactive phosphorus group, a phosphorus moiety or a covalent attachment to a support medium.

[0227] In certain embodiments, bicyclic nucleosides are provided having Formula II:



wherein:

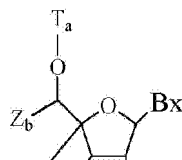
Bx is a heterocyclic base moiety;

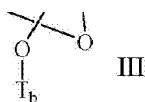
T_a and T_b are each, independently H, a hydroxyl protecting group, a conjugate group, a reactive phosphorus group, a phosphorus moiety or a covalent attachment to a support medium;

Z_a is C_1 - C_6 alkyl, C_2 - C_6 alkenyl, C_2 - C_6 alkynyl, substituted C_1 - C_6 alkyl, substituted C_2 - C_6 alkenyl, substituted C_2 - C_6 alkynyl, acyl, substituted acyl, substituted amide, thiol or substituted thio.

[0228] In one embodiment, each of the substituted groups is, independently, mono or poly substituted with substituent groups independently selected from halogen, oxo, hydroxyl, OJ_c , NJ_cJ_d , SJ_c , N_3 , $\text{OC}(=\text{X})\text{J}_c$, and $\text{NJ}_e\text{C}(=\text{X})\text{NJ}_c\text{J}_d$, wherein each J_c , J_d and J_e is, independently, H, C_1 - C_6 alkyl, or substituted C_1 - C_6 alkyl and X is O or NJ_c .

[0229] In certain embodiments, bicyclic nucleosides are provided having Formula III:





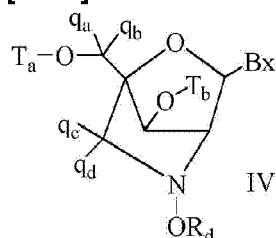
wherein:

Bx is a heterocyclic base moiety;

T_a and T_b are each, independently H, a hydroxyl protecting group, a conjugate group, a reactive phosphorus group, a phosphorus moiety or a covalent attachment to a support medium;

Z_b is C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, substituted C₁-C₆ alkyl, substituted C₂-C₆ alkenyl, substituted C₂-C₆ alkynyl or substituted acyl (C(=O)-).

[0230] In certain embodiments, bicyclic nucleosides are provided having Formula IV:



wherein:

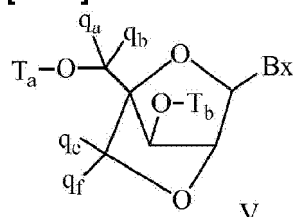
Bx is a heterocyclic base moiety;

T_a and T_b are each, independently H, a hydroxyl protecting group, a conjugate group, a reactive phosphorus group, a phosphorus moiety or a covalent attachment to a support medium;

R_d is C₁-C₆ alkyl, substituted C₁-C₆ alkyl, C₂-C₆ alkenyl, substituted C₂-C₆ alkenyl, C₂-C₆ alkynyl or substituted C₂-C₆ alkynyl;

each q_a, q_b, q_c and q_d is, independently, H, halogen, C₁-C₆ alkyl, substituted C₁-C₆ alkyl, C₂-C₆ alkenyl, substituted C₂-C₆ alkenyl, C₂-C₆ alkynyl or substituted C₂-C₆ alkynyl, C₁-C₆ alkoxy, substituted C₁-C₆ alkoxy, acyl, substituted acyl, C₁-C₆ aminoalkyl or substituted C₁-C₆ aminoalkyl;

[0231] In certain embodiments, bicyclic nucleosides are provided having Formula V:



wherein:

Bx is a heterocyclic base moiety;

T_a and T_b are each, independently H, a hydroxyl protecting group, a conjugate group, a reactive phosphorus group, a phosphorus moiety or a covalent attachment to a support medium;

q_a, q_b, q_e and q_f are each, independently, hydrogen, halogen, C₁-C₁₂ alkyl, substituted C₁-C₁₂ alkyl, C₂-C₁₂ alkenyl, substituted C₂-C₁₂ alkenyl, C₂-C₁₂ alkynyl, substituted C₂-C₁₂ alkynyl, C₁-C₁₂ alkoxy, substituted C₁-C₁₂ alkoxy, OJ_j, SJ_j, SOJ_j, SO₂J_j, NJ_jJ_k, N₃, CN, C(=O)OJ_j, C(=O)NJ_jJ_k, C(=O)J_j, O-C(=O)NJ_jJ_k, N(H)C(=NH)NJ_jJ_k, N(H)C(=O)NJ_jJ_k or N(H)C(=S)NJ_jJ_k;

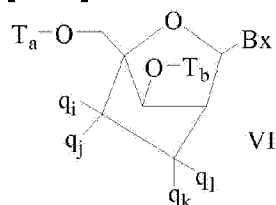
or q_e and q_f together are =C(q_g)(q_h);

q_g and q_h are each, independently, H, halogen, C₁-C₁₂ alkyl or substituted C₁-C₁₂ alkyl.

[0232] The synthesis and preparation of the methyleneoxy (4'-CH₂-O-2') BNA monomers adenine, cytosine, guanine, 5-methyl-cytosine, thymine and uracil, along with their oligomerization, and nucleic acid recognition properties have been described (Koshkin et al., Tetrahedron, 1998, 54, 3607-3630). BNAs and preparation thereof are also described in WO 98/39352 and WO 99/14226.

[0233] Analogs of methyleneoxy (4'-CH₂-O-2') BNA and 2'-thio-BNAs, have also been prepared (Kumar et al., Bioorg. Med. Chem. Lett., 1998, 8, 2219-2222). Preparation of locked nucleoside analogs comprising oligodeoxyribonucleotide duplexes as substrates for nucleic acid polymerases has also been described (Wengel et al., WO 99/14226). Furthermore, synthesis of 2'-amino-BNA, a novel conformationally restricted high-affinity oligonucleotide analog has been described in the art (Singh et al., J. Org. Chem., 1998, 63, 10035-10039). In addition, 2'-amino- and 2'-methylamino-BNA's have been prepared and the thermal stability of their duplexes with complementary RNA and DNA strands has been previously reported.

[0234] In certain embodiments, bicyclic nucleosides are provided having Formula VI:



wherein:

Bx is a heterocyclic base moiety;

T_a and T_b are each, independently H, a hydroxyl protecting group, a conjugate group, a

reactive phosphorus group, a phosphorus moiety or a covalent attachment to a support medium;

each q_i , q_j , q_k and q_l is, independently, H, halogen, C_1 - C_{12} alkyl, substituted C_1 - C_{12} alkyl, C_2 - C_{12} alkenyl, substituted C_2 - C_{12} alkenyl, C_2 - C_{12} alkynyl, substituted C_2 - C_{12} alkynyl, C_1 - C_{12} alkoxyl, substituted C_1 - C_{12} alkoxyl, OJ_j , SJ_j , SOJ_j , SO_2J_j , NJ_jJ_k , N_3 , CN , $C(=O)OJ_j$, $C(=O)NJ_jJ_k$, $C(=O)J_j$, $O-C(=O)NJ_jJ_k$, $N(H)C(=NH)NJ_jJ_k$, $N(H)C(=O)NJ_jJ_k$ or $N(H)C(=S)NJ_jJ_k$; and

q_i and q_j or q_l and q_k together are $=C(q_g)(q_h)$, wherein q_g and q_h are each, independently, H, halogen, C_1 - C_{12} alkyl or substituted C_1 - C_{12} alkyl.

[0235] One carbocyclic bicyclic nucleoside having a 4'-(CH_2)₃-2' bridge and the alkenyl analog bridge 4'-CH=CH-CH₂-2' have been described (Freier et al., Nucleic Acids Research, 1997, 25(22), 4429-4443 and Albaek et al., J. Org. Chem., 2006, 71, 7731-7740). The synthesis and preparation of carbocyclic bicyclic nucleosides along with their oligomerization and biochemical studies have also been described (Srivastava et al., J. Am. Chem. Soc., 2007, 129(26), 8362-8379).

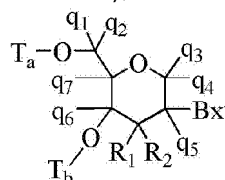
[0236] As used herein, "4'-2' bicyclic nucleoside" or "4' to 2' bicyclic nucleoside" refers to a bicyclic nucleoside comprising a furanose ring comprising a bridge connecting two carbon atoms of the furanose ring connects the 2' carbon atom and the 4' carbon atom of the sugar ring.

[0237] As used herein, "monocyclic nucleosides" refer to nucleosides comprising modified sugar moieties that are not bicyclic sugar moieties. In certain embodiments, the sugar moiety, or sugar moiety analogue, of a nucleoside may be modified or substituted at any position.

[0238] As used herein, "2'-modified sugar" means a furanosyl sugar modified at the 2' position. In certain embodiments, such modifications include substituents selected from: a halide, including, but not limited to substituted and unsubstituted alkoxy, substituted and unsubstituted thioalkyl, substituted and unsubstituted amino alkyl, substituted and unsubstituted alkyl, substituted and unsubstituted allyl, and substituted and unsubstituted alkynyl. In certain embodiments, 2' modifications are selected from substituents including, but not limited to: $O[(CH_2)_nO]_mCH_3$, $O(CH_2)_nNH_2$, $O(CH_2)_nCH_3$, $O(CH_2)_nF$, $O(CH_2)_nONH_2$, $OCH_2C(=O)N(H)CH_3$, and $O(CH_2)_nON[(CH_2)_nCH_3]_2$, where n and m are from 1 to about 10. Other 2'- substituent groups can also be selected from: C_1 - C_{12} alkyl, substituted alkyl, alkenyl, alkynyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH₃, OCN, Cl, Br, CN, F, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, an RNA cleaving group, a reporter group, an intercalator, a group for improving pharmacokinetic properties, or a group for improving the pharmacodynamic properties of an antisense compound, and other substituents having similar properties. In

certain embodiments, modified nucleosides comprise a 2'-MOE side chain (Baker et al., J. Biol. Chem., 1997, 272, 11944-12000). Such 2'-MOE substitution have been described as having improved binding affinity compared to unmodified nucleosides and to other modified nucleosides, such as 2'-O-methyl, O-propyl, and O-aminopropyl. Oligonucleotides having the 2'-MOE substituent also have been shown to be antisense inhibitors of gene expression with promising features for *in vivo* use (Martin, Helv. Chim. Acta, 1995, 78, 486-504; Altmann et al., Chimia, 1996, 50, 168-176; Altmann et al., Biochem. Soc. Trans., 1996, 24, 630-637; and Altmann et al., Nucleosides Nucleotides, 1997, 16, 917-926).

[0239] As used herein, a "modified tetrahydropyran nucleoside" or "modified THP nucleoside" means a nucleoside having a six-membered tetrahydropyran "sugar" substituted in for the pentofuranosyl residue in normal nucleosides (a sugar surrogate). Modified THP nucleosides include, but are not limited to, what is referred to in the art as hexitol nucleic acid (HNA), anitol nucleic acid (ANA), manitol nucleic acid (MNA) (see Leumann, Bioorg. Med. Chem., 2002, 10, 841-854), fluoro HNA (F-HNA) or those compounds having Formula VII:



VII

wherein independently for each of said at least one tetrahydropyran nucleoside analog of Formula VII:

Bx is a heterocyclic base moiety;

T_a and T_b are each, independently, an internucleoside linking group linking the tetrahydropyran nucleoside analog to the antisense compound or one of T_a and T_b is an internucleoside linking group linking the tetrahydropyran nucleoside analog to the antisense compound and the other of T_a and T_b is H, a hydroxyl protecting group, a linked conjugate group or a 5' or 3'-terminal group;

q₁, q₂, q₃, q₄, q₅, q₆ and q₇ are each independently, H, C₁-C₆ alkyl, substituted C₁-C₆ alkyl, C₂-C₆ alkenyl, substituted C₂-C₆ alkenyl, C₂-C₆ alkynyl or substituted C₂-C₆ alkynyl; and each of R₁ and R₂ is selected from hydrogen, hydroxyl, halogen, substituted or unsubstituted alkoxy, NJ₁J₂, SJ₁, N₃, OC(=X)J₁, OC(=X)NJ₁J₂, NJ₃C(=X)NJ₁J₂ and CN, wherein X is O, S or NJ₁ and each J₁, J₂ and J₃ is, independently, H or C₁-C₆ alkyl.

[0240] In certain embodiments, the modified THP nucleosides of Formula VII are provided wherein q₁, q₂, q₃, q₄, q₅, q₆ and q₇ are each H. In certain embodiments, at least one of q₁, q₂, q₃, q₄, q₅, q₆ and q₇ is other than H. In certain embodiments, at least one of q₁, q₂, q₃, q₄, q₅, q₆ and q₇ is methyl. In certain embodiments, THP nucleosides of Formula VII are provided

wherein one of R_1 and R_2 is fluoro. In certain embodiments, R_1 is fluoro and R_2 is H; R_1 is methoxy and R_2 is H, and R_1 is H and R_2 is methoxyethoxy.

[0241] As used herein, "2'-modified" or "2'-substituted" refers to a nucleoside comprising a sugar comprising a substituent at the 2' position other than H or OH. 2'-modified nucleosides, include, but are not limited to, bicyclic nucleosides wherein the bridge connecting two carbon atoms of the sugar ring connects the 2' carbon and another carbon of the sugar ring; and nucleosides with non-bridging 2'substituents, such as allyl, amino, azido, thio, O-allyl, O- C_1 - C_{10} alkyl, $-OCF_3$, O- $(CH_2)_2$ -O- CH_3 , 2'-O $(CH_2)_2SCH_3$, O- $(CH_2)_2$ -O-N(R_m)(R_n), or O- CH_2 -C(=O)-N(R_m)(R_n), where each R_m and R_n is, independently, H or substituted or unsubstituted C_1 - C_{10} alkyl. 2'-modified nucleosides may further comprise other modifications, for example at other positions of the sugar and/or at the nucleobase.

[0242] As used herein, "2'-F" refers to a nucleoside comprising a sugar comprising a fluoro group at the 2' position.

[0243] As used herein, "2'-OMe" or "2'-OCH₃" or "2'-O-methyl" each refers to a nucleoside comprising a sugar comprising an -OCH₃ group at the 2' position of the sugar ring.

[0244] As used herein, "MOE" or "2'-MOE" or "2'-OCH₂CH₂OCH₃" or "2'-O-methoxyethyl" each refers to a nucleoside comprising a sugar comprising a -OCH₂CH₂OCH₃ group at the 2' position of the sugar ring.

[0245] As used herein, "oligonucleotide" refers to a compound comprising a plurality of linked nucleosides. In the invention, one or more of the plurality of nucleosides is modified. In certain embodiments, an oligonucleotide comprises one or more ribonucleosides (RNA) and/or deoxyribonucleosides (DNA).

[0246] Many other bicyclo and tricyclo sugar surrogate ring systems are also known in the art that can be used to modify nucleosides for incorporation into antisense compounds (see for example review article: Leumann, Bioorg. Med. Chem., 2002, 10, 841-854). Such ring systems can undergo various additional substitutions to enhance activity.

[0247] Methods for the preparations of modified sugars are well known to those skilled in the art.

[0248] In nucleotides having modified sugar moieties, the nucleobase moieties (natural, modified or a combination thereof) are maintained for hybridization with an appropriate nucleic acid target.

[0249] In certain embodiments, antisense compounds comprise one or more nucleosides having modified sugar moieties. In certain embodiments, the modified sugar moiety is 2'-MOE. In certain embodiments, the 2'-MOE modified nucleosides are arranged in a gapmer motif. In

certain embodiments, the modified sugar moiety is a bicyclic nucleoside having a (4'-CH(CH₃)-O-2') bridging group. In certain embodiments, the (4'-CH(CH₃)-O-2') modified nucleosides are arranged throughout the wings of a gapmer motif.

Compositions and Methods for Formulating Pharmaceutical Compositions

[0250] Antisense oligonucleotides may be admixed with pharmaceutically acceptable active or inert substances for the preparation of pharmaceutical compositions or formulations. Compositions and methods for the formulation of pharmaceutical compositions are dependent upon a number of criteria, including, but not limited to, route of administration, extent of disease, or dose to be administered.

[0251] An antisense compound targeted to a C9ORF72 nucleic acid can be utilized in pharmaceutical compositions by combining the antisense compound with a suitable pharmaceutically acceptable diluent or carrier. A pharmaceutically acceptable diluent includes phosphate-buffered saline (PBS). PBS is a diluent suitable for use in compositions to be delivered parenterally. Accordingly, in one instance, employed in the methods described herein is a pharmaceutical composition comprising an antisense compound targeted to a C9ORF72 nucleic acid and a pharmaceutically acceptable diluent. In certain instances, the pharmaceutically acceptable diluent is PBS. In the invention, the antisense compound is a modified antisense oligonucleotide.

[0252] Pharmaceutical compositions comprising antisense compounds encompass any pharmaceutically acceptable salts, esters, or salts of such esters, or any other oligonucleotide which, upon administration to an animal, including a human, is capable of providing (directly or indirectly) the biologically active metabolite or residue thereof. Accordingly, for example, the disclosure is also drawn to pharmaceutically acceptable salts of antisense compounds, prodrugs, pharmaceutically acceptable salts of such prodrugs, and other bioequivalents. Suitable pharmaceutically acceptable salts include, but are not limited to, sodium and potassium salts.

[0253] A prodrug can include the incorporation of additional nucleosides at one or both ends of an antisense compound which are cleaved by endogenous nucleases within the body, to form the active antisense compound.

Conjugated Antisense Compounds

[0254] Antisense compounds may be covalently linked to one or more moieties or conjugates which enhance the activity, cellular distribution or cellular uptake of the resulting antisense oligonucleotides. Typical conjugate groups include cholesterol moieties and lipid moieties. Additional conjugate groups include carbohydrates, phospholipids, biotin, phenazine, folate,

phenanthridine, anthraquinone, acridine, fluoresceins, rhodamines, coumarins, and dyes.

[0255] Antisense compounds can also be modified to have one or more stabilizing groups that are generally attached to one or both termini of antisense compounds to enhance properties such as, for example, nuclease stability. Included in stabilizing groups are cap structures. These terminal modifications protect the antisense compound having terminal nucleic acid from exonuclease degradation, and can help in delivery and/or localization within a cell. The cap can be present at the 5'-terminus (5'-cap), or at the 3'-terminus (3'-cap), or can be present on both termini. Cap structures are well known in the art and include, for example, inverted deoxy abasic caps. Further 3' and 5'-stabilizing groups that can be used to cap one or both ends of an antisense compound to impart nuclease stability include those disclosed in WO 03/004602 published on January 16, 2003.

Cell culture and antisense compounds treatment

[0256] The effects of antisense compounds on the level, activity or expression of C9ORF72 nucleic acids can be tested *in vitro* in a variety of cell types. Cell types used for such analyses are available from commercial vendors (e.g. American Type Culture Collection, Manassus, VA; Zen-Bio, Inc., Research Triangle Park, NC; Clonetics Corporation, Walkersville, MD) and are cultured according to the vendor's instructions using commercially available reagents (e.g. Invitrogen Life Technologies, Carlsbad, CA). Illustrative cell types include, but are not limited to, HepG2 cells, Hep3B cells, and primary hepatocytes.

In vitro testing of antisense oligonucleotides

[0257] Described herein are methods for treatment of cells with antisense oligonucleotides, which can be modified appropriately for treatment with other antisense compounds.

[0258] In general, cells are treated with antisense oligonucleotides when the cells reach approximately 60-80% confluency in culture.

[0259] One reagent commonly used to introduce antisense oligonucleotides into cultured cells includes the cationic lipid transfection reagent LIPOFECTIN (Invitrogen, Carlsbad, CA). Antisense oligonucleotides are mixed with LIPOFECTIN in OPTI-MEM 1 (Invitrogen, Carlsbad, CA) to achieve the desired final concentration of antisense oligonucleotide and a LIPOFECTIN concentration that typically ranges 2 to 12 ug/mL per 100 nM antisense oligonucleotide.

[0260] Another reagent used to introduce antisense oligonucleotides into cultured cells includes LIPOFECTAMINE (Invitrogen, Carlsbad, CA). Antisense oligonucleotide is mixed with LIPOFECTAMINE in OPTI-MEM 1 reduced serum medium (Invitrogen, Carlsbad, CA) to achieve the desired concentration of antisense oligonucleotide and a LIPOFECTAMINE

concentration that typically ranges 2 to 12 ug/mL per 100 nM antisense oligonucleotide.

[0261] Another technique used to introduce antisense oligonucleotides into cultured cells includes electroporation.

[0262] Cells are treated with antisense oligonucleotides by routine methods. Cells are typically harvested 16-24 hours after antisense oligonucleotide treatment, at which time RNA or protein levels of target nucleic acids are measured by methods known in the art and described herein. In general, when treatments are performed in multiple replicates, the data are presented as the average of the replicate treatments.

[0263] The concentration of antisense oligonucleotide used varies from cell line to cell line. Methods to determine the optimal antisense oligonucleotide concentration for a particular cell line are well known in the art. Antisense oligonucleotides are typically used at concentrations ranging from 1 nM to 300 nM when transfected with LIPOFECTAMINE. Antisense oligonucleotides are used at higher concentrations ranging from 625 to 20,000 nM when transfected using electroporation.

RNA Isolation

[0264] RNA analysis can be performed on total cellular RNA or poly(A)+mRNA. Methods of RNA isolation are well known in the art. RNA is prepared using methods well known in the art, for example, using the TRIZOL Reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's recommended protocols.

Analysis of inhibition of target levels or expression

[0265] Inhibition of levels or expression of a C9ORF72 nucleic acid can be assayed in a variety of ways known in the art. For example, target nucleic acid levels can be quantitated by, e.g., Northern blot analysis, competitive polymerase chain reaction (PCR), or quantitative real-time PCR. RNA analysis can be performed on total cellular RNA or poly(A)+mRNA. Methods of RNA isolation are well known in the art. Northern blot analysis is also routine in the art. Quantitative real-time PCR can be conveniently accomplished using the commercially available ABI PRISM 7600, 7700, or 7900 Sequence Detection System, available from PE-Applied Biosystems, Foster City, CA and used according to manufacturer's instructions.

Quantitative Real-Time PCR Analysis of Target RNA Levels

[0266] Quantitation of target RNA levels may be accomplished by quantitative real-time PCR using the ABI PRISM 7600, 7700, or 7900 Sequence Detection System (PE-Applied

Biosystems, Foster City, CA) according to manufacturer's instructions. Methods of quantitative real-time PCR are well known in the art.

[0267] Prior to real-time PCR, the isolated RNA is subjected to a reverse transcriptase (RT) reaction, which produces complementary DNA (cDNA) that is then used as the substrate for the real-time PCR amplification. The RT and real-time PCR reactions are performed sequentially in the same sample well. RT and real-time PCR reagents are obtained from Invitrogen (Carlsbad, CA). RT real-time-PCR reactions are carried out by methods well known to those skilled in the art.

[0268] Gene (or RNA) target quantities obtained by real time PCR are normalized using either the expression level of a gene whose expression is constant, such as cyclophilin A, or by quantifying total RNA using RIBOGREEN (Invitrogen, Inc. Carlsbad, CA). Cyclophilin A expression is quantified by real time PCR, by being run simultaneously with the target, multiplexing, or separately. Total RNA is quantified using RIBOGREEN RNA quantification reagent (Invitrogen, Inc. Eugene, OR). Methods of RNA quantification by RIBOGREEN are taught in Jones, L.J., et al, (Analytical Biochemistry, 1998, 265, 368-374). A CYTOFLUOR 4000 instrument (PE Applied Biosystems) is used to measure RIBOGREEN fluorescence.

[0269] Probes and primers are designed to hybridize to a C9ORF72 nucleic acid. Methods for designing real-time PCR probes and primers are well known in the art, and may include the use of software such as PRIMER EXPRESS Software (Applied Biosystems, Foster City, CA).

Analysis of Protein Levels

[0270] Antisense inhibition of C9ORF72 nucleic acids can be assessed by measuring C9ORF72 protein levels. Protein levels of C9ORF72 can be evaluated or quantitated in a variety of ways well known in the art, such as immunoprecipitation, Western blot analysis (immunoblotting), enzyme-linked immunosorbent assay (ELISA), quantitative protein assays, protein activity assays (for example, caspase activity assays), immunohistochemistry, immunocytochemistry or fluorescence-activated cell sorting (FACS). Antibodies directed to a target can be identified and obtained from a variety of sources, such as the MSRS catalog of antibodies (Aerie Corporation, Birmingham, MI), or can be prepared via conventional monoclonal or polyclonal antibody generation methods well known in the art. Antibodies useful for the detection of mouse, rat, monkey, and human C9ORF72 are commercially available.

In vivo testing of antisense compounds

[0271] Antisense compounds, for example, antisense oligonucleotides, are tested in animals to assess their ability to inhibit expression of C9ORF72 and produce phenotypic changes, such as, improved motor function and respiration. In certain embodiments, motor function is

measured by rotarod, grip strength, pole climb, open field performance, balance beam, hindpaw footprint testing in the animal. In certain embodiments, respiration is measured by whole body plethysmograph, invasive resistance, and compliance measurements in the animal. Testing may be performed in normal animals, or in experimental disease models. For administration to animals, antisense oligonucleotides are formulated in a pharmaceutically acceptable diluent, such as phosphate-buffered saline. Administration includes parenteral routes of administration, such as intraperitoneal, intravenous, and subcutaneous. Calculation of antisense oligonucleotide dosage and dosing frequency is within the abilities of those skilled in the art, and depends upon factors such as route of administration and animal body weight. Following a period of treatment with antisense oligonucleotides, RNA is isolated from CNS tissue or CSF and changes in C9ORF72 nucleic acid expression are measured.

Targeting C9ORF72

[0272] Antisense oligonucleotides described herein may hybridize to a C9ORF72 nucleic acid in any stage of RNA processing. For example, described herein are antisense oligonucleotides that are complementary to a pre-mRNA or a mature mRNA. Additionally, antisense oligonucleotides described herein may hybridize to any element of a C9ORF72 nucleic acid. For example, described herein are antisense oligonucleotides that are complementary to an exon, an intron, the 5' UTR, the 3' UTR, a repeat region, a hexanucleotide repeat expansion, a splice junction, an exon:exon splice junction, an exonic splicing silencer (ESS), an exonic splicing enhancer (ESE), exon 1a, exon 1b, exon 1c, exon 1d, exon 1e, exon 2, exon 3, exon 4, exon 5, exon 6, exon 7, exon 8, exon 9, exon 10, exon 11, intron 1, intron 2, intron 3, intron 4, intron 5, intron 6, intron 7, intron 8, intron 9, or intron 10 of a C9ORF72 nucleic acid.

[0273] In certain embodiments, antisense oligonucleotides described herein hybridize to all variants of C9ORF72. In certain embodiments, the antisense oligonucleotides described herein selectively hybridize to certain variants of C9ORF72. In certain embodiments, the antisense oligonucleotides described herein selectively hybridize to variants of C9ORF72 containing a hexanucleotide repeat expansion. In certain embodiments, such variants of C9ORF72 containing a hexanucleotide repeat expansion include SEQ ID NO: 1-3 and 6-10. In certain embodiments, such hexanucleotide repeat expansion comprises at least 30 repeats of any of GGGGCC, GGGGGG, GGGGGC, or GGGGCG.

[0274] In certain embodiments, the antisense oligonucleotides described herein inhibit expression of all variants of C9ORF72. In certain embodiments, the antisense oligonucleotides described herein inhibit expression of all variants of C9ORF72 equally. In certain embodiments, the antisense oligonucleotides described herein preferentially inhibit expression of certain variants of C9ORF72. In certain embodiments, the antisense oligonucleotides described herein preferentially inhibit expression of variants of C9ORF72 containing a hexanucleotide repeat expansion. In certain embodiments, such variants of C9ORF72 containing a hexanucleotide repeat expansion include SEQ ID NO: 1-3 and 6-10. In certain embodiments, such hexanucleotide repeat expansion comprises at least 30 repeats of any of GGGGCC,

GGGGGG, GGGGGC, or GGGGCG. In certain embodiments, the hexanucleotide repeat expansion forms nuclear foci. Antisense oligonucleotides described herein are useful for reducing nuclear foci. Nuclear foci may be reduced in terms of percent of cells with foci as well as number of foci per cell.

[0275] Based on earlier studies directed to repeat expansions, it is not possible to predict if antisense oligonucleotides targeting C9ORF72 outside of the hexanucleotide repeat expansion would successfully inhibit expression of C9ORF72 for two reasons. First, the C9ORF72 repeat expansion is located in an intron and it is not known if the RNA in the foci contains only the repeats or also the flanking intronic sequence. For example, an earlier study on myotonic dystrophy type 2 (DM2), which is a disease caused by a CCTG expansion mutation in intron 1 of the ZNF9 gene, determined that large DM2 expansions did not prevent allele-specific pre-mRNA splicing, nuclear export of the transcripts, or steady-state mRNA or protein levels. The study further demonstrated that the ribonuclear inclusions found associated with the disease are enriched for the CCUG expansion, but not the flanking intronic sequences. These data suggest that the downstream molecular effects of the DM2 mutation may be triggered by the accumulation of CCUG repeat tract alone. Therefore, this study implies that targeting the CCUG repeat expansion alone would lead to amelioration of the disease, since targeting the flanking sequences, especially the region downstream of the repeat expansion, would not affect the formation of ribonuclear inclusions (Margolis et al. Hum. Mol. Genet., 2006, 15:1808-1815). Second, it is not known how fast intron 1 of C9ORF72, which contains the repeats, is excised and accumulates in foci. Thus, it is not possible to predict if targeting the pre-mRNA would result in elimination of the repeat RNA and foci.

C9ORF72 Features

[0276] Antisense oligonucleotides described herein may hybridize to any C9ORF72 variant at any state of processing within any element of the C9ORF72 gene. For example, antisense oligonucleotides described herein may hybridize to an exon, an intron, the 5' UTR, the 3' UTR, a repeat region, a hexanucleotide repeat expansion, a splice junction, an exon:exon splice junction, an exonic splicing silencer (ESS), an exonic splicing enhancer (ESE), exon 1a, exon 1b, exon 1c, exon 1d, exon 1e, exon 2, exon 3, exon 4, exon 5, exon 6, exon 7, exon 8, exon 9, exon 10, exon 11, intron 1, intron 2, intron 3, intron 4, intron 5, intron 6, intron 7, intron 8, intron 9, or intron 10. For example, antisense oligonucleotides may target any of the exons characterized below in Tables 1-5 for the various C9ORF72 variants described below. Antisense oligonucleotides described herein may also target variants not characterized below and such variants are characterized in GENBANK. Moreover, antisense oligonucleotides described herein may also target elements other than exons and such elements are characterized in GENBANK.

Table 1

Functional Segments for NM_001256054.1 (SEQ ID NO: 1)				
Exon Number	mRNA start site	mRNA stop site	Start site in reference to SEQ ID NO: 2	Stop site in reference to SEQ ID NO: 2
exon 1C	1	158	1137	1294
exon 2	159	646	7839	8326
exon 3	647	706	9413	9472
exon 4	707	802	12527	12622
exon 5	803	867	13354	13418
exon 6	868	940	14704	14776
exon 7	941	1057	16396	16512
exon 8	1058	1293	18207	18442
exon 9	1294	1351	24296	24353
exon 10	1352	1461	26337	26446
exon 11	1462	3339	26581	28458

Table 2

Functional Segments for NM_018325.3 (SEQ ID NO: 4)				
Exon Number	mRNA start site	mRNA stop site	Start site in reference to SEQ ID NO: 2	Stop site in reference to SEQ ID NO: 2
exon 1B	1	63	1510	1572
exon 2	64	551	7839	8326
exon 3	552	611	9413	9472
exon 4	612	707	12527	12622
exon 5	708	772	13354	13418
exon 6	773	845	14704	14776
exon 7	846	962	16396	16512
exon 8	963	1198	18207	18442
exon 9	1199	1256	24296	24353
exon 10	1257	1366	26337	26446
exon 11	1367	3244	26581	28458

Table 3

Functional Segments for NM_145005.5 (SEQ ID NO: 6)				
Exon Number	mRNA start site	mRNA stop site	Start site in reference to SEQ ID NO: 2	Stop site in reference to SEQ ID NO: 2
exon 1A	1	80	1137	1216
exon 2	81	568	7839	8326
exon 3	569	628	9413	9472

Functional Segments for NM_145005.5 (SEQ ID NO: 6)				
Exon Number	mRNA start site	mRNA stop site	Start site in reference to SEQ ID NO: 2	Stop site in reference to SEQ ID NO: 2
exon 4	629	724	12527	12622
exon 5B (exon 5 into intron 5)	725	1871	13354	14500

Table 4

Functional Segments for DB079375.1 (SEQ ID NO: 7)				
Exon Number	mRNA start site	mRNA stop site	Start site in reference to SEQ ID NO: 2	Stop site in reference to SEQ ID NO: 2
exon 1E	1	35	1135	1169
exon 2	36	524	7839	8326
exon 3 (EST ends before end of full exon)	525	562	9413	9450

Table 5

Functional Segments for BU194591.1 (SEQ ID NO: 8)				
Exon Number	mRNA start site	mRNA stop site	Start site in reference to SEQ ID NO: 2	Stop site in reference to SEQ ID NO: 2
exon 1D	1	36	1241	1279
exon 2	37	524	7839	8326
exon 3	525	584	9413	9472
exon 4	585	680	12527	12622
exon 5B (exon 5 into intron 5)	681	798	13354	13465

Certain Indications

[0277] In certain instances, disclosed herein are methods of treating an individual comprising administering one or more pharmaceutical compositions described herein. In certain instances, the individual has a neurodegenerative disease. In certain instances, the individual can be at risk for developing a neurodegenerative disease, including, but not limited to, ALS or FTD. The individual may have been identified as having a C9ORF72 associated disease. The individual may have been identified as having a C9ORF72 hexanucleotide repeat expansion associated disease. Disclosed herein are methods for prophylactically reducing C9ORF72 expression in an individual. Certain instances include treating an individual in need thereof by administering

to an individual a therapeutically effective amount of an antisense compound targeted to a C9ORF72 nucleic acid.

[0278] In one instance, administration of a therapeutically effective amount of an antisense compound targeted to a C9ORF72 nucleic acid is accompanied by monitoring of C9ORF72 levels in an individual, to determine an individual's response to administration of the antisense compound. An individual's response to administration of the antisense compound may be used by a physician to determine the amount and duration of therapeutic intervention.

[0279] In certain instances, administration of an antisense compound targeted to a C9ORF72 nucleic acid results in reduction of C9ORF72 expression by at least 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 or 99%, or a range defined by any two of these values. In certain instances, administration of an antisense compound targeted to a C9ORF72 nucleic acid results in improved motor function and respiration in an animal. In certain instances, administration of a C9ORF72 antisense compound improves motor function and respiration by at least 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 or 99%, or a range defined by any two of these values.

[0280] In certain instances, pharmaceutical compositions comprising an antisense compound targeted to C9ORF72 are used for the preparation of a medicament for treating a patient suffering or susceptible to a neurodegenerative disease including ALS and FTD.

Certain Combination Therapies

[0281] In certain instances, one or more pharmaceutical compositions described herein are co-administered with one or more other pharmaceutical agents. In certain instances, such one or more other pharmaceutical agents are designed to treat the same disease, disorder, or condition as the one or more pharmaceutical compositions described herein. In certain instances, such one or more other pharmaceutical agents are designed to treat a different disease, disorder, or condition as the one or more pharmaceutical compositions described herein. In certain instances, such one or more other pharmaceutical agents are designed to treat an undesired side effect of one or more pharmaceutical compositions described herein. In certain instances, one or more pharmaceutical compositions described herein are co-administered with another pharmaceutical agent to treat an undesired effect of that other pharmaceutical agent. In certain instances, one or more pharmaceutical compositions described herein are co-administered with another pharmaceutical agent to produce a combinational effect. In certain instances, one or more pharmaceutical compositions described herein are co-administered with another pharmaceutical agent to produce a synergistic effect.

[0282] In certain instances, one or more pharmaceutical compositions described herein and one or more other pharmaceutical agents are administered at the same time. In certain instances, one or more pharmaceutical compositions described herein and one or more other pharmaceutical agents are administered at different times. In certain instances, one or more

pharmaceutical compositions described herein and one or more other pharmaceutical agents are prepared together in a single formulation. In certain instances, one or more pharmaceutical compositions described herein and one or more other pharmaceutical agents are prepared separately.

[0283] In certain instances, pharmaceutical agents that may be co-administered with a pharmaceutical composition described herein include Riluzole (Rilutek), Lioresal (Lioresal), and Dexamipexole.

[0284] In certain instances, pharmaceutical agents that may be co-administered with a C9ORF72 specific inhibitor described herein include, but are not limited to, an additional C9ORF72 inhibitor. In certain instances, the co-administered pharmaceutical agent is administered prior to administration of a pharmaceutical composition described herein. In certain instances, the co-administered pharmaceutical agent is administered following administration of a pharmaceutical composition described herein. In certain instances the co-administered pharmaceutical agent is administered at the same time as a pharmaceutical composition described herein. In certain instances the dose of a co-administered pharmaceutical agent is the same as the dose that would be administered if the co-administered pharmaceutical agent was administered alone. In certain instances the dose of a co-administered pharmaceutical agent is lower than the dose that would be administered if the co-administered pharmaceutical agent was administered alone. In certain instances the dose of a co-administered pharmaceutical agent is greater than the dose that would be administered if the co-administered pharmaceutical agent was administered alone.

[0285] In certain instances, the co-administration of a second compound enhances the effect of a first compound, such that co-administration of the compounds results in an effect that is greater than the effect of administering the first compound alone. In other instances, the co-administration results in effects that are additive of the effects of the compounds when administered alone. In certain instances, the co-administration results in effects that are supra-additive of the effects of the compounds when administered alone. In certain instances, the first compound is an antisense compound. In certain instances, the second compound is an antisense compound.

EXAMPLES

Non-limiting disclosure

[0286] While certain compounds, compositions, and methods described herein have been described with specificity in accordance with certain embodiments, the following examples serve only to illustrate the compounds described herein and are not intended to limit the same.

Example 1: Antisense inhibition of human C9ORF72 in HepG2 cells

[0287] Antisense oligonucleotides were designed targeting a C9ORF72 nucleic acid and were tested for their effects on C9ORF72 mRNA *in vitro*. The antisense oligonucleotides were tested in a series of experiments that had similar culture conditions. The results for each experiment are presented in separate tables shown below. Cultured HepG2 cells at a density of 20,000 cells per well were transfected using electroporation with 7,000 nM antisense oligonucleotide. After a treatment period of approximately 24 hours, RNA was isolated from the cells and C9ORF72 mRNA levels were measured by quantitative real-time PCR. Human primer probe set RTS3750 (forward sequence TGTGACAGTTGGAATGCAGTGA, designated herein as SEQ ID NO: 15; reverse sequence GCCACTTAAAGCAATCTCTGTCTTG, designated herein as SEQ ID NO: 16; probe sequence TCGACTCTTTGCCCACCGCCA, designated herein as SEQ ID NO: 17) was used to measure mRNA levels. C9ORF72 mRNA levels were adjusted according to total RNA content, as measured by RIBOGREEN®. Results are presented as percent inhibition of C9ORF72, relative to untreated control cells.

[0288] The antisense oligonucleotides in Tables 6-8 were designed as 5-10-5 MOE gapmers. The gapmers are 20 nucleosides in length, wherein the central gap segment comprises ten 2'-deoxynucleosides and is flanked by wing segments on both the 5' end and on the 3' end comprising five nucleosides each. Each nucleoside in the 5' wing segment and each nucleoside in the 3' wing segment has a MOE modification. The internucleoside linkages throughout each gapmer are phosphorothioate linkages. All cytosine residues throughout each gapmer are 5-methylcytosines. "Start site" indicates the 5'-most nucleoside to which the antisense oligonucleotide is targeted in the human gene sequence. "Stop site" indicates the 3'-most nucleoside to which the antisense oligonucleotide is targeted human gene sequence. Each antisense oligonucleotide listed in Tables 6-7 is targeted to the either human C9ORF72 mRNA sequence, designated herein as SEQ ID NO: 1 (GENBANK Accession No. NM_001256054.1) or the human C9ORF72 genomic sequence, designated herein as SEQ ID NO: 2 (the complement of GENBANK Accession No. NT_008413.18 truncated from nucleosides 27535000 to 27565000), or both. 'n/a' indicates that the antisense oligonucleotide did not target that particular gene sequence. The antisense oligonucleotides of Table 8 are targeted to either SEQ ID NO: 3 (GENBANK Accession No. BQ068108.1) or SEQ ID NO: 4 (GENBANK Accession No. NM_018325.3).

Table 6

ISIS No	% inhibition
576816	78
576860	79

Table 7

ISIS No	% inhibition
576974	56

Table 8

ISIS No	% inhibition
576816	90
577061	28
577065	0
577083	70

Example 2: Dose-dependent antisense inhibition of human C9ORF72 in HepG2 cells

[0289] Antisense oligonucleotides from the study described above exhibiting significant *in vitro* inhibition of C9ORF72 mRNA were selected and tested at various doses in HepG2 cells. The antisense oligonucleotides were tested in a series of experiments that had similar culture conditions. The results for each experiment are presented in separate tables shown below. Cells were plated at a density of 20,000 cells per well and transfected using electroporation with 82.3 nM, 246.9 nM, 740.7 nM, 2,222.2 nM, 6,666.7 nM, or 20,000 nM concentrations of antisense oligonucleotide. After a treatment period of approximately 16 hours, RNA was isolated from the cells and C9ORF72 mRNA levels were measured by quantitative real-time PCR. Human C9ORF72 primer probe set RTS3750 was used to measure mRNA levels. C9ORF72 mRNA levels were adjusted according to total RNA content, as measured by RIBOGREEN®. Results are presented as percent inhibition of C9ORF72, relative to untreated control cells.

[0290] The half maximal inhibitory concentration (IC₅₀) of each oligonucleotide is also presented in Tables 9-10. As illustrated, C9ORF72 mRNA levels were reduced in a dose-dependent manner in the antisense oligonucleotide treated cells.

Table 9

ISIS No	82.3 nM	246.9 nM	740.7 nM	2222.2 nM	6666.7 nM	20000.0 nM	IC ₅₀ (μM)
576816	5	23	49	76	91	96	0.9
576860	9	24	54	70	83	87	1.0

Table 10

ISIS No	82.3 nM	246.9 nM	740.7 nM	2222.2 nM	6666.7 nM	20000.0 nM	IC ₅₀ (μM)
576974	25	45	65	70	65	78	0.5
576816	18	36	53	82	91	95	0.6
577083	0	0	24	50	73	74	3.0

Example 3: Dose-dependent antisense inhibition of human C9ORF72 in HepG2 cells

[0291] Antisense oligonucleotides from the study described above exhibiting significant *in vitro* inhibition of C9ORF72 mRNA were selected and tested at various doses in HepG2 cells. The antisense oligonucleotides were tested in a series of experiments that had similar culture conditions. The results for each experiment are presented in separate tables shown below. Cells were plated at a density of 20,000 cells per well and transfected using electroporation with 246.9 nM, 740.7 nM, 2,222.2 nM, 6,666.7 nM, or 20,000 nM concentrations of antisense oligonucleotide. After a treatment period of approximately 16 hours, RNA was isolated from the cells and C9ORF72 total mRNA levels, as well as mRNA levels of the exon 1 transcript, were measured by quantitative real-time PCR. Human C9ORF72 primer probe set RTS3750 was used to measure total C9ORF72 mRNA levels. Primer probe set RTS3905 (forward sequence GGGTCTAGCAAGAGCAGGTG, designated herein as SEQ ID NO: 18; reverse sequence GTCTTGGCAACAGCTGGAGAT, designated herein as SEQ ID NO: 19; probe sequence TGATGTGCGACTCTTTGCCACCGC, designated herein as SEQ ID NO: 20) was used to measure exon 1 message transcript. C9ORF72 mRNA levels were adjusted according to total RNA content, as measured by RIBOGREEN®. Results are presented as percent inhibition of C9ORF72, relative to untreated control cells.

[0292] The half maximal inhibitory concentration (IC₅₀) of each oligonucleotide is also presented in Tables 11 and 12. As illustrated, C9ORF72 mRNA levels were reduced in a dose-dependent manner in the antisense oligonucleotide treated cells. 'n.d.' indicates that there is no data for that particular dose.

Table 11

% inhibition of total C9ORF72 mRNA levels						
ISIS No	246.9 nM	740.7 nM	2222.2 nM	6666.7 nM	20000.0 nM	IC ₅₀ (μM)
576816	29	53	84	90	92	0.60
576860	25	53	72	86	85	0.80
576974	36	49	64	65	68	0.95
577061	0	3	0	4	0	>20.00
577065	7	0	1	6	0	>20.00
577083	0	19	55	71	75	3.35

Table 12

% inhibition of C9ORF72 exon 1 mRNA levels						
ISIS No	246.9 nM	740.7 nM	2222.2 nM	6666.7 nM	20000.0 nM	IC ₅₀ (μM)
576816	45	78	93	n.d.	n.d.	0.26
576860	43	36	71	95	91	0.66
577061	8	36	70	67	76	2.03
577065	20	55	67	82	62	1.06
577083	30	43	85	88	92	0.71

Example 4: Effect of antisense inhibition of human C9ORF72 in ALS patient primary fibroblasts

[0293] Antisense oligonucleotides, from the studies described above and targeting various regions of the C9ORF72 gene (SEQ ID NO: 2), were tested in ALS patient primary fibroblasts. The target location of each antisense oligonucleotide is specified in Table 13 below. As observed from the Table, several of the antisense oligonucleotides target intron 1 and, therefore, the pre-mRNA sequence of C9ORF72. Intron 1 contains a hexanucleotide repeat (GGGGCC), which is expanded to at least 30 repeats in a pathogenic cell (Renton et al., Neuron, 2011, 72, 257-268).

Table 13

Location of antisense oligonucleotides targeting SEQ ID NO: 1	
ISIS No	Location
577061	Intron 1
577065	Intron 1
577083	Intron 1
576816	Exon 2
576860	Intron 5
576974	Exon 11

[0294] The effect of treatment of antisense oligonucleotide targeting C9ORF72 was assessed in primary fibroblasts derived from ALS patients. Primary fibroblasts derived from skin biopsy of a human ALS patient (Dr John Ravits lab) were plated in DMEM medium supplemented with 20% fetal bovine serum. The cells were treated with the antisense oligonucleotides targeting C9ORF72 at 25 nM concentration for 4 hrs. A set of cells was treated with a control oligonucleotide (ASO control), with no known human target, and served as a negative control. After 24 hrs of transfection, the cells were harvested in TriZol reagent for RNA quantification, or were fixed with 4% paraformaldehyde for *in situ* hybridization.

RNA analysis

[0295] The cells were harvested and levels of certain mRNA variants of C9ORF72, as well as total C9ORF72 mRNA were analyzed. The mRNA variants measured were SEQ ID NO: 1 (variant 3) and SEQ ID NO: 6 (variant 2). As presented in Tables 1 and 3 listed above, SEQ ID NOs 1 and 6 encode portions of the C9ORF72 sequence that include intron 1. Because intron 1 includes the hexanucleotide repeat, reducing mRNA levels of these variants are therapeutic in pathogenic cells (cells containing hexanucleotide repeat expansions of more than 30

repeats). Certain other mRNA variants do not include the hexanucleotide repeat; for example, SEQ ID NO: 4. Therefore, it is possible to preferentially inhibit certain C9ORF72 variants by designing antisense oligonucleotides to target certain regions of the target C9ORF72 nucleic acid, such as, for example intron 1.

[0296] The levels of SEQ ID NO: 1 and 6 (variants 3 and 2, respectively) were measured in combination by primer probe set RTS3905, which targets exon 1. The level of total C9orf72 mRNA was measured by primer probe set RTS3750, which targets exon 2. The data is presented in Table 14. As compared to PBS control, Isis 577061 and Isis 577065 (each targeting Intron 1) reduced expression of variant 2/3 RNA by over 90% and reduced total RNA by 18% and 8%, respectively. Isis 577083 (targeting intron 1), Isis 576816 (targeting exon 2), Isis 576860 (targeting exon 5), and Isis 576974 (targeting exon 11) achieved between 75% and 96% inhibition of variant 2/3 RNA expression and total RNA expression. The ASO control did not significantly inhibit RNA levels, as expected.

Table 14

Percent inhibition, compared to the PBS control, of RNA levels of C9ORF72 in primary fibroblasts			
Oligo ID	Location	Variant 2/3	Total RNA
ASO control	n/a	13	3
577061	Intron 1	95	18
577065	Intron 1	96	8
577083	Intron 1	89	81
576816	Exon 2	96	95
576860	Intron 5	81	75
576974	Exon 11	85	77

Analysis of C9ORF72 RNA foci

[0297] In recent years, intracellular accumulation of expanded nucleotide repeats as RNA foci in the nucleus of affected cells has emerged as an important disease mechanism in frontotemporal dementia and ALS patients (DeJesus-Hernandez, M. et al., Neuron. 2011. 72: 245-56). C9ORF72 nuclear foci were detected using FISH with (5'TYE563-CCCCGGCCCCGGCCCC) (SEQ ID NO: 23) LNA probe (Exiqon). The data for the percentage of cells with foci is presented in Table 15. The results indicate that antisense oligonucleotides targeting C9ORF72 RNA decreased the percentage of cells containing C9ORF72 RNA foci.

[0298] Single and multiple RNA foci are detected in patient fibroblasts. The numbers of cells with single or multiple foci from fibroblasts that were treated with ISIS 576816 or the ASO control were counted. The data is presented in Table 16, and indicates that the antisense oligonucleotide targeting C9ORF72 reduced the number of foci per cell compared to the ASO

control.

Table 15

Percent of primary fibroblasts with foci	
	% cells
PBS	37
ASO control	38
577061	9
577065	6
577083	6
576816	9
576860	12
576974	12

Table 16

Number of cells with single of multiple foci		
Foci#	576816-treated	ASO Control
0	82	66
1	14	10
2	2	9
3	2	5
4	0	5
5	0	2
6	0	0
7	0	0
8	0	2
9	0	0
10+	0	1

Example 5: Effect of siRNA inhibition of human C9ORF72 in ALS patient primary fibroblasts

[0299] siRNA targeting C9ORF72 and control siRNA targeting GAPDH were obtained from Thermo Scientific Dharmacon (On-Target plus Smartpool). The effect of treatment of siRNA in ALS patient fibroblasts was assessed.

[0300] Primary fibroblasts derived from a skin biopsy of a human ALS patients were plated in DMEM medium supplemented with 20% fetal bovine serum. The cells were treated with the

siRNA targeting C9ORF72 at 40 nM concentration for 4 hrs. A set of cells was also treated with siRNA targeting GAPDH at 40 nM concentration for 4 hrs and served as the control. After transfection for 24 hrs, the cells were harvested in TriZol reagent for RNA quantification, or were fixed with 4% paraformaldehyde for *in situ* hybridization.

[0301] Total C9ORF72 RNA levels were measured using a primer set (Forward: TCCTGTAATGGAAGCTGCTTTCA, designated herein as SEQ ID NO: 21; Reverse: GGTATCTGCTTCATCCAGCTTT, designated herein as SEQ ID NO: 22). siRNA treatment resulted in 73% inhibition of total C9ORF72 RNA levels compared to the control.

[0302] C9ORF72 nuclear foci were detected using FISH with (5'TYE563-CCCCGGCCCCGGCCCC)(SEQ ID NO: 23) LNA probe (Exiqon). The data is presented in Table 17. The results indicate that siRNA treatment against C9ORF72 does not reduce nuclear foci in the fibroblasts compared to the control.

Table 17

Percentage of cells with foci	
	% cells
siRNA-GAPDH	28
siRNA-C9ORF72	31

Example 6: In vivo rodent inhibition and tolerability with treatment of C9ORF72 antisense oligonucleotides

[0303] In order to assess the tolerability of inhibition of C9orf72 expression *in vivo*, antisense oligonucleotides targeting a murine C9orf72 nucleic acid were designed and assessed in mouse and rat models.

[0304] ISIS 571883 was designed as a 5-10-5 MOE gapmer, 20 nucleosides in length, wherein the central gap segment comprises ten 2'-deoxynucleosides and is flanked by wing segments on both the 5' end and on the 3' end comprising five nucleosides each. Each nucleoside in the 5' wing segment and each nucleoside in the 3' wing segment has a MOE modification. The internucleoside linkages are phosphorothioate linkages. All cytosine residues throughout the gapmer are 5-methylcytosines. ISIS 571883 has a target start site of nucleoside 33704 on the murine C9ORF72 genomic sequence, designated herein as SEQ ID NO: 11 (the complement of GENBANK Accession No. NT_166289.1 truncated from nucleosides 3587000 to 3625000).

[0305] ISIS 603538 was designed as a 5-10-5 MOE gapmer, 20 nucleosides in length, wherein the central gap segment comprises ten 2'-deoxynucleosides and is flanked by wing segments on both the 5' end and on the 3' end comprising five nucleosides each. Each nucleoside in the 5' wing segment and each nucleoside in the 3' wing segment has a MOE modification. The internucleoside linkages are either phosphorothioate linkages or phosphate ester linkages (Gs

Ao Co Co Gs Cs Ts Ts Gs As Gs Ts Ts Ts Gs Co Co Ao Cs A; wherein 's' denotes a phosphorothioate internucleoside linkage, 'o' denotes a phosphate ester linkage; and A, G, C, T denote the relevant nucleosides). All cytosine residues throughout the gapmer are 5-methylcytosines. ISIS 603538 has a target start site of nucleoside 2872 on the rat C9ORF72 mRNA sequence, designated herein as SEQ ID NO: 12 (GENBANK Accession No. NM_001007702.1).

Mouse experiment 1

[0306] Groups of 4 C57BL/6 mice each were injected with 50 µg, 100 µg, 300 µg, 500 µg, or 700 µg of ISIS 571883 administered via an intracerebroventricular bolus injection. A control group of four C57/BL6 mice were similarly treated with PBS. Animals were anesthetized with 3% isoflurane and placed in a stereotactic frame. After sterilizing the surgical site, each mouse was injected -0.2 mm antero-posterior from the bregma and 3 mm dorsoventral to the bregma with the above-mentioned doses of ISIS 571883 using a Hamilton syringe. The incision was closed with sutures. The mice were allowed to recover for 14 days, after which animals were euthanized according to a humane protocol approved by the Institutional Animal Care and Use Committee. Brain and spinal cord tissue were harvested and snap frozen in liquid nitrogen. Prior to freezing, brain tissue was cut transversely five sections using a mouse brain matrix.

RNA analysis

[0307] RNA was extracted from a 2-3 mm brain section posterior to the injection site, from brain frontal cortex and from the lumbar section of the spinal cord tissue for analysis of C9orf72 mRNA expression. C9orf72 mRNA expression was measured by RT-PCR. The data is presented in Table 18. The results indicate that treatment with increasing doses of ISIS 571883 resulted in dose-dependent inhibition of C9orf72 mRNA expression.

[0308] The induction of the microglial marker AIF-1 as a measure of CNS toxicity was also assessed. The data is presented in Table 19. The results indicate that treatment with increasing doses of ISIS 571883 did not result in significant increases in AIF-1 mRNA expression. Hence, the injection of ISIS 571883 was deemed tolerable in this model.

Table 18

Percentage inhibition of C9orf72 mRNA expression compared to the PBS control			
Dose (µg)	Posterior brain	Cortex	Spinal cord
50	22	8	46
100	22	12	47
300	55	47	67
500	61	56	78

Percentage inhibition of C9orf72 mRNA expression compared to the PBS control			
Dose (μ g)	Posterior brain	Cortex	Spinal cord
700	65	65	79

Table 19

Percentage expression of AIF-1 mRNA expression compared to the PBS control		
Dose (μ g)	Posterior brain	Spinal cord
50	102	89
100	105	111
300	107	98
500	131	124
700	122	116

Mouse experiment 2

[0309] Groups of 4 C57BL/6 mice each were injected with 500 μ g of ISIS 571883 administered via an intracerebroventricular bolus injection in a procedure similar to that described above. A control group of four C57/BL6 mice were similarly treated with PBS. The mice were tested at regular time points after ICV administration.

Behavior analysis

[0310] Two standard assays to assess motor behavior were employed; the rotarod assay and grip strength assay. In case of the rotarod assays, the time of latency to fall was measured. The data for the assays is presented in Tables 20 and 21. The results indicate that there were no significant changes in the motor behavior of the mice as a result of antisense inhibition of ISIS 571883 or due to the ICV injection. Hence, antisense inhibition of C9orf72 was deemed tolerable in this model.

Table 20

Latency to fall (sec) in the rotarod assay		
Weeks after injection	PBS	ISIS 571883
0	66	66
4	91	70
8	94	84

Table 21

Mean hindlimb grip strength (g) in the grip strength assay		
Weeks after injection	PBS	ISIS 571883
0	57	63
1	65	51
2	51	52
3	51	51
4	59	72
5	60	64
6	61	72
7	67	68
8	66	70
9	63	61
10	48	46

Rat experiment

[0311] Groups of 4 Sprague-Dawley rats each were injected with 700 µg, 1,000 µg, or 3,000 µg of ISIS 603538 administered via an intrathecal bolus injection. A control group of four Sprague-Dawley rats were similarly treated with PBS. Animals were anesthetized with 3% isofluorane and placed in a stereotactic frame. After sterilizing the surgical site, each rat was injected with 30 µL of ASO solution administered via 8 cm intrathecal catheter 2 cm into the spinal canal with a 50 µL flush. The rats were allowed to recover for 4 weeks, after which animals were euthanized according to a humane protocol approved by the Institutional Animal Care and Use Committee.

RNA analysis

[0312] RNA was extracted from a 2-3 mm brain section posterior to the injection site, from brain frontal cortex, and from the cervical and lumbar sections of the spinal cord tissue for analysis of C9orf72 mRNA expression. C9orf72 mRNA expression was measured by RT-PCR. The data is presented in Table 22. The results indicate that treatment with increasing doses of ISIS 603538 resulted in dose-dependent inhibition of C9orf72 mRNA expression.

[0313] The induction of the microglial marker AIF-1 as a measure of CNS toxicity was also assessed. The data is presented in Table 23. The results indicate that treatment with increasing doses of ISIS 603538 did not result in significant increases in AIF-1 mRNA expression. Hence, the injection of ISIS 603538 was deemed tolerable in this model.

Table 22

Percentage inhibition of C9orf72 mRNA expression compared to the PBS control				
Dose (μ g)	Brain (1 mm section)	Cortex	Spinal cord (lumbar)	Spinal cord (cervical)
700	21	4	86	74
1000	53	49	88	82
3000	64	62	88	80

Table 23

Percentage expression of AIF-1 mRNA expression compared to the PBS control				
Dose (μ g)	Brain (1 mm section)	Cortex	Spinal cord (lumbar)	Spinal cord (cervical)
700	97	119	98	89
1000	105	113	122	96
3000	109	141	156	115

Body weight analysis

[0314] Body weights of the rats were measured at regular time point intervals. The data is presented in Table 24. The results indicate that treatment with increasing doses of ISIS 603538 did not have any significant changes in the body weights of the rats.

Table 24

Body weights of the rats (% initial body weight)						
	Dose (μ g)	Week 1	Week 2	Week 3	Week 4	Week 5
PBS		100	94	103	105	109
ISIS 603538	700	100	94	98	103	107
	1000	100	95	97	101	103
	3000	100	92	98	102	105

SEQUENCE LISTING

[0315]

<110> Isis Pharmaceuticals, Inc.
 Ludwig Institute for Cancer Research
 The Regents of the University of California

<120> METHODS FOR MODULATING G9ORF72 EXPRESSION

<130> BIOLC1214WO

<150> 61/714,139

<151> 2012-10-15

<160> 35

<170> FastSEQ for Windows Version 4.0

 $\langle 210 \rangle$ 1

<211> 3339

<212> DNA

<213> Homo sapiens

$\langle 220 \rangle$

<221> CDS

<222> (203)...(1648)

<400> 1

acgtaaaccta	cggtgtcccgcg	ctaggaaaga	gaggtgcgcgc	aaacagcgac	aagttccgcc	60										
cacgtaaaag	atgacgcttg	gtgtgtcagc	cgccccctgct	gcccggttgc	ttctcttttg	120										
ggg'gcggggt	ctagcaagag	caggtgtggg	tttaggagat	atctccggag	catttgata	180										
atgtgacagt	tggaatgcag	tg	atg	tgc	act	ctt	tgc	cca	ccg	cca	tct	cca	232			
		Met	Ser	Thr	Leu	Cys	Pro	Pro	Pro	Ser	Pro					
		1				5						10				
gct	gtt	gcc	aag	aca	gag	att	gct	tta	agt	ggc	aaa	tca	cct	tta	tta	280
Ala	Val	Ala	Lys	Thr	Glu	Ile	Ala	Leu	Ser	Gly	Lys	Ser	Pro	Leu	Leu	
				15					20					25		
gca	gct	act	ttt	gct	tac	tgg	gac	aat	att	ctt	ggg	cct	aga	gta	agg	328
Ala	Ala	Thr	Phe	Ala	Tyr	Trp	Asp	Asn	Ile	Leu	Gly	Pro	Arg	Val	Arg	
			30					35					40			
cac	att	tgg	gct	cca	aag	aca	gaa	cag	gta	ctt	ctc	agt	gat	gga	gaa	376
His	Ile	Trp	Ala	Pro	Lys	Thr	Glu	Gln	Val	Leu	Leu	Ser	Asp	Gly	Glu	
		45					50					55				
ata	act	ttt	ctt	gcc	aac	cac	act	cta	aat	gga	gaa	atc	ctt	cga	aat	424
Ile	Thr	Phe	Leu	Ala	Asn	His	Thr	Leu	Asn	Gly	Glu	Ile	Leu	Arg	Asn	
		60				65					70					
gca	gag	agt	ggg	gct	ata	gat	gta	aag	ttt	ttt	gtc	ttg	tct	gaa	aag	472
Ala	Glu	Ser	Gly	Ala	Ile	Asp	Val	Lys	Phe	Phe	Val	Leu	Ser	Glu	Lys	
	75				80					85					90	
gga	gtg	att	att	gtt	tca	tta	atc	ttt	gat	gga	aac	tgg	aat	ggg	gat	520
Gly	Val	Ile	Ile	Val	Ser	Leu	Ile	Phe	Asp	Gly	Asn	Trp	Asn	Gly	Asp	
				95					100				105			
cgc	agc	aca	tat	qga	cta	tca	att	ata	ctt	cca	cag	aca	gaa	ctt	agt	568
Arg	Ser	Thr	Tyr	Gly	Leu	Ser	Ile	Ile	Leu	Pro	Gln	Thr	Glu	Leu	Ser	
			110					115					120			
ttc	tac	ctc	cca	ctt	cat	aga	gtg	tgt	gtt	gat	aga	tta	aca	cat	ata	616
Phe	Tyr	Leu	Pro	Leu	His	Arg	Val	Cys	Val	Asp	Arg	Leu	Thr	His	Ile	
		125				130					135					

atc cgg aaa gga aga ata tgg atg cat aag gaa aga caa gaa aat gtc Ile Arg Lys Gly Arg Ile Trp Met His Lys Glu Arg Gln Glu Asn Val 140 145 150	664
cag aag att atc tta gaa ggc aca gag aga atg gaa gat cag ggt cag Gln Lys Ile Ile Leu Glu Gly Thr Glu Arg Met Glu Asp Gln Gly Gln 155 160 165 170	712
agt att att cca atg ctt act gga gaa gtg att cct gta atg gaa ctg Ser Ile Ile Pro Met Leu Thr Gly Glu Val Ile Pro Val Met Glu Leu 175 180 185	760
ctt tca tct atg aaa tca cac agt gtt cct gaa gaa ata gat ata gct Leu Ser Ser Met Lys Ser His Ser Val Pro Glu Glu Ile Asp Ile Ala 190 195 200	808
gat aca gta ctc aat gat gat gat att ggt gac agc tgt cat gaa ggc Asp Thr Val Leu Asn Asp Asp Asp Ile Gly Asp Ser Cys His Glu Gly 205 210 215	856
ttt ctt ctc aat gcc atc agc tca cac ttg caa acc tgt ggc tgt tcc Phe Leu Leu Asn Ala Ile Ser Ser His Leu Gln Thr Cys Gly Cys Ser 220 225 230	904
gtt gta gta ggt agc agt gca gag aaa gta aat aag ata gtc aga aca Val Val Val Gly Ser Ser Ala Glu Lys Val Asn Lys Ile Val Arg Thr 235 240 245 250	952
tta tgc ctt ttt ctg act cca gca gag aga aaa tgc tcc agg tta tgt Leu Cys Leu Phe Leu Thr Pro Ala Glu Arg Lys Cys Ser Arg Leu Cys 255 260 265	1000
gaa gca gaa tca tca ttt aaa tat gag tca ggg ctc ttt gta caa ggc Glu Ala Glu Ser Ser Phe Lys Tyr Glu Ser Gly Leu Phe Val Gln Gly 270 275 280	1048
ctg cta aag gat tca act gga agc ttt gtg ctg cct ttc cgg caa gtc Leu Leu Lys Asp Ser Thr Gly Ser Phe Val Leu Pro Phe Arg Gln Val 285 290 295	1096
atg tat gct cca tat ccc acc aca cac ata gat gtg gat gtc aat act Met Tyr Ala Pro Tyr Pro Thr Thr His Ile Asp Val Asp Val Asn Thr 300 305 310	1144
gtg aag cag atg cca ccc tgt cat gaa cat att tat aat cag cgt aga Val Lys Gln Met Pro Pro Cys His Glu His Ile Tyr Asn Gln Arg Arg 315 320 325 330	1192
tac atg aga tcc gag ctg aca gcc ttc tgg aga gcc act tca gaa gaa Tyr Met Arg Ser Glu Leu Thr Ala Phe Trp Arg Ala Thr Ser Glu Glu 335 340 345	1240
gac atg gct cag gat acg atc atc tac act gac gaa agc ttt act cct Asp Met Ala Gln Asp Thr Ile Ile Tyr Thr Asp Glu Ser Phe Thr Pro 350 355 360	1288
gat ttg aat att ttt caa gat gtc tta cac aga gac act cta gtg aaa Asp Leu Asn Ile Phe Gln Asp Val Leu His Arg Asp Thr Leu Val Lys 365 370 375	1336
gcc ttc ctg gat cag gtc ttt cag ctg aaa cct ggc tta tct ctc aga Ala Phe Leu Asp Gln Val Phe Gln Leu Lys Pro Gly Leu Ser Leu Arg 380 385 390	1384
agt act ttc ctt gca cag ttt cta ctt gtc ctt cac aga aaa gcc ttg Ser Thr Phe Leu Ala Gln Phe Leu Leu Val Leu His Arg Lys Ala Leu 395 400 405 410	1432
aca cta ata aaa tat ata gaa gac gat acg cag aag gga aaa aag ccc	1480

Thr Leu Ile Lys Tyr Ile Glu Asp Asp Thr Gln Lys Gly Lys Lys Pro
 415 420 425

ttt aaa tct ctt cgg aac ctg aag ata gac ctt gat tta aca gca gag 1528
 Phe Lys Ser Leu Arg Asn Leu Lys Ile Asp Leu Asp Leu Thr Ala Glu
 430 435 440

ggc gat ctt aac ata ata atg gct ctg gct gag aaa att aaa cca ggc 1576
 Gly Asp Leu Asn Ile Ile Met Ala Leu Ala Glu Lys Ile Lys Pro Gly
 445 450 455

cta cac tct ttt atc ttt gga aga cct ttc tac act agt gtg caa gaa 1624
 Leu His Ser Phe Ile Phe Gly Arg Pro Phe Tyr Thr Ser Val Gln Glu
 460 465 470

cga gat gtt cta atg act ttt taa atgtgtaact taataagcct attccatcac 1678
 Arg Asp Val Leu Met Thr Phe
 475 480

aatcatgac gctggtaaag tagctcagtg gtgtggggaa acgttcccct ggatcatact 1738
 ccagaattct gctctcagca attgcagtta agtaagttac actacagttc tcacaagagc 1798
 ctgtgagggg atgtcaggtg catcattaca ttgggtgccc cttttcctag atttatgctt 1858
 ttgggataca gacctatgtt tacaatataa taaatatat tgctatcttt taaagatata 1918
 ataataggat gtaaaactga ccacaactac tgtttttttg aaatacatga ttcatggttt 1978
 acatgtgtca aggtgaaatc tgagttggct ttacagata gttgactttc tatcttttgg 2038
 cattcttttg tgtgtagaat tactgtaata cttctgcaat caactgaaaa ctagagcctt 2098
 taaatgattt caattccaca gaaagaaagt gagcttgaac ataggatgag ctttagaaaag 2158
 aaaattgac aagcagatgt ttaattggaa ttgattacta gatcctactt tgtggattta 2218
 gtccctggga ttcatgtctg agaaatgtct aatagttctc tatagtcctt gttcctgggtg 2278
 aaccacagtt aggtgtgttt gtttatttta ttgttcttgc tattgttgat attctatgta 2338
 gttgagctct gtaaaaggaa attgtatttt atgttttagt aattgttgcc aactttttta 2398
 attaattttc attatttttg agccaaattg aaatgtgcac ctctgtgccc ttttttctcc 2458
 ttagaaaaac taattacttg gaacaagtgc agatttcaact ggtcagtcac ttctatcttg 2518
 ttttctctct gctaagtctt accatgtacc tgctttggca atcattgcaa ctctgagatt 2578
 ataaaatgcc ttagagaata tactaactaa taagatcttc ttttcagaaa cagaaaaatag 2638
 ttctttgagt acttccttct tgcatttctg cctatgtctt tgaagttgtt gctgtttgcc 2698
 tgcaataggc tataaggaat agcaggagaa attttactga agtgcgtgtt tctaggtgac 2758

tactttggca gagctaagtt atcttttgtt ttcttaatgc gtttggacca ttttgctggc 2818
 tataaaataa ctgattaata taattctaac acaatgttga cattgtagtt acacaaacac 2878
 aaataaatat tttattttaa attctggaag taatataaaa gggaaaatat atttataaga 2938
 aagggataaa ggtaatagag ccttctgccc cccacccac caaatttaca caacaaaatg 2998
 acatgttcga atgtgaaagg tcataatagc tttcccatca tgaatcagaa agatgtggac 3058
 agottgatgt tttagacaac cactgaacta gatgactgtt gtactgtagc tcagtcattt 3118
 aaaaaatata taaatactac cttgtagtgt cccatactgt gttttttaca tggtagattc 3178
 ttattttagt gctaaactgt tattttcttt ggctggttta ttgtactgtt atacagaatg 3238
 taagttgtac agtgaaataa gttattaaag catgtgtaaa cattgttata tatcttttct 3298
 cctaaatgga gaattttgaa taaaatatat ttgaaatttt g 3339

<210> 2

<211> 30001

<212> DNA

<213> Homo sapiens

<400> 2

caaagaaaag ggggaggttt tgttaaaaaa gagaatgtt acatagtgtt ctttgagaaa 60
 attcattggc actattaagg atctgaggag ctggtgagtt tcaactgggt agtgatgggt 120
 gtagataaaa ttagagctgc agcaggtcat tttagcaact attagataaa actgggtctc 180
 ggtcacaacg ggcagttgca gcagctggac ttggagagaa ttacactgtg ggagcagtgt 240
 catttgtcct aagtgtcttt ctacccccta ccccaactat tttagttggg tataaaaaa 300
 atgacccaat ttgtatgac aactttcaca aagcatagaa cagtaggaaa aggggtctgt 360
 tctgcagaag gtgtagacgt tgagagccat tttgtgtatt tattcctccc tttcttctc 420
 ggtgaatgat taaaacgttc tgtgtgattt ttagtgtatg aaaagattaa atgtacttca 480
 ctgtagtaag tgccatctca caattgcaga tcaaaaggca cacagtttaa aaaacctttg 540
 tttttttaca catctgagtg gtgtaaatgc tactcatctg tagtaagtgg aatctatata 600
 cctgcagacc aaaagacgca aggtttcaca aatctttgtg ttttttacac atcaaacaga 660
 atgtatgctt ttttttctgt ttttttctgt ttttttctgt ttttttctgt ttttttctgt 720

```

atggtatgctt tttcaaaagc taaaaaaaaa caactcatcc acatattgca attagcaaaa 720
atgacatttc ccagtgtaga aatcatgctt gagagaattc ttacatgtaa aggcataaatt 780
gcgatgactt tgcaggggac cgtgggattc ccgcccgcag tgcgggagct gtcccctacc 840
agggtttgca gtggagtttt gaatgcactt aacagtgtct tacggtaaaa acaaaatttc 900
atccaccaat tatgtgttga gcgcccactg cctaccaagc acaaacaaaa ccattcaaaa 960
ccacgaaatc gtcttcactt tctccagatc cagcagcctc cctattaag gttcgcacac 1020
gctattgcgc caacgctcct ccagagcggg tcttaagata aaagaacagg acaagttgcc 1080
ccgccccatt tcgctagcct cgtgagaaaa cgtcatcgca catagaaaac agacagacgt 1140
aacctacggt gtcccgcctag gaaagagagg tgcgtcaaac agcgacaagt tccgcccacg 1200
taaaagatga cgcttggtgt gtcagccgtc cctgctgcc gggtgcttct cttttggggg 1260
cggggtctag caagagcagg tgtgggttta ggaggtgtgt gtttttgttt tccccaccct 1320
ctctcccac tacttgctct cacagtactc gctgagggtg aacaagaaaa gacctgataa 1380
agattaacca gaagaaaaca aggagggaaa caaccgcagc ctgtagcaag ctctggaact 1440
caggagtgc gcgctagggg ccggggccgg ggccggggcg tggtcggggc ggcccgggg 1500
gcgggcccgg ggccggggctg ccgttgccgt gcttgcgcc ccggcgggcg aggcgcaggc 1560
ggtgccgagt ggtgagtgga ggaggcggca tccgtggcgg tggctgtttg gggttccgg 1620
gcgggaaga ggccggggtg gaagcggggg ctctctcag agctcgacgc atttttactt 1680
tccctctcat ttctctgacc gaagctgggt gtcgggcttt cgcctctagc gactggtgga 1740
attgctgca tccgggcccc gggttcccg gcggcgccg ccggcgccgc ggccgaggga 1800
caagggatgg ggatctggcc tcttccctgc tttcccgccc tcagtaccgc agctgtctcc 1860
ttcccgggga ccgctggga gcgctgcgc tgcgggctcg agaaaaggga gctcgggta 1920
ctgagaggcc tcgctgggg gaaggccgga ggggtggcgg ccgcgcgctt ctgcccacca 1980
agtccgggtt ccgtaggaac ccgagacggt ccttgcggc gagagatca tgcgggatga 2040
gatgggggtg tggagacgcc tgcacaattt cagcccaagc ttctagagag tgggtgatga 2100
ttgcatatga gggcagcaat gcaagtcggt gtgctcccca ttctgtggga catgacctgg 2160
ttgcttcaca gctccgagat gacacagact tgcttaaagg aagtgactat tgtgacttgg 2220
gcatcacttg actgatgta atcagttgtc taaagaagtg cacagattac atgtccgtgt 2280
gctcattggg tctatctggc cgcgttgaac accaccaggc tttgtattca gaaacaggag 2340
ggaggtcctg cactttccca ggagggtgg ccctttcaga tgcaatcgag attgttaggc 2400
tctgggagag tagttgcctg gttgtggcag ttggtaaatt tctattcaaa cagttgccat 2460

```

```

gcaccagttg ttcacaacaa ggttacgtaa totgtctggo attacttcta cttttgtaca 2520
aaggatcaaa aaaaaaaaaa atactgttaa gatatgattt ttctcagact ttgggaaact 2580
tttaacataa tctgtgaata tcacagaaac aagactatca tatagggat attaataacc 2640
tggagtcaaa atacttgaaa tacggtgtca tttagacagg gcattgttgt caccacctct 2700
gccaaaggct gccactttag gaaaaccctg aatcagttgg aaactgctac atgctgatag 2760
tacatctgaa acaagaacga gagtaattac cacattccag atgttccact aagccagcat 2820
ttacctgctc caggaaaaaa ttacaagcac cttatgaagt tgataaaaata ttttgtttgg 2880
ctatgtttgg actccacaat ttgctttcag aqaaacaaag taaaccaagg aggacttctg 2940
tttttcaagt ctgccctcgg gttctattct acgttaatta gatagttccc aggaggacta 3000
ggttagccta cctattgtct gagaaaactg gaactgtgag aaatggccag atagtgatat 3060
gaacttcacc ttccagttct cctgatgtt gaagattgag aaagtgttgt gaactttctg 3120
gtactgtaaa cagttcactg tccttgaagt ggtcctgggc agctcctgtt gtggaaagt 3180
gacggtttag gatcctgctt ctctttgggc tgggagaaaa taaacagcat gggtacaagt 3240
attgagagcc aggttgagga aggtggctta cactgtaat gccagagctt tgggaggcgg 3300
aggcaagagg atcacttgaa gccaggagtt caagctcaac ctgggcaacg tagacctgt 3360
ctctacaaaa aattaaaaac ttageccggc gtggtgatgt gcacctgtag tctagctac 3420
ttgggaggct gaggcaggag ggtcatttga gcccaagagt ttgaagttac cgagagctat 3480
gatcctgcca gtgcattcca gctggatga caaacgaga cctgtctct aaaaaacaag 3540
aagtgagggc tttatgattg tagaattttc actacaatag cagtggacca accaccttcc 3600
taaataccaa tcagggaaga gatggttgat tttttaacag acgtttaaag aaaaagcaaa 3660
acctcaact tagcactcta ctaacagttt tagcagatgt taattaatgt aatcatgtct 3720
gcatgtatgg gattatttcc agaaagtgtg ttgggaaacc tctcatgaac cctgtgagca 3780
agccaccgtc tcaactcaatt tgaactcttg cttccctcaa aagactggct aatgttttgt 3840
aactctctgg agtagacagc actacatgta cgtaagatag gtacataaac aactatttgt 3900
tttgagctga ttttttcag ctgcatttgc atgtatggat ttctctcacc aaagacgatg 3960
acttcaagta ttagtaaaat aattgtacag ctctcctgat tatacttctc tgtgacattt 4020
catttccagg gctatttctt ttggtaggat ttaaaaacta gcaattcagt atgatctttg 4080
tcttccattt tctttcttat tctttttgtt ttgttttttt tcttgaggca 4140
gagtctctct ctgtcgcaca ggctggagtg cagtggcgcc atctcagctc attgcaacct 4200
ctgccacctc cgggttcaag agattctcct gctcagcct ccgagtagc tgggattaca 4260
ggtgtccacc accacaaccg gctaattttt tgtattttta gtagagggtg ggtttcacca 4320
tgttggccag gctggctttg agctcctgac ctccagtgat ccacctgctc cggcctacca 4380
aagagctggg ataacagggt tgaccacca tgcccgccc atttttttt tcttattctg 4440
ttaggagtga gagtgtaat agcagtataa tagttcaatt ttcacaacgt ggtaaaaagt 4500
tccctataat tcaatcagat tttgctccag ggttcagttc tgttttagga aatactttta 4560
ttttcagttt aatgatgaaa tattagagtt gtaatatgtc ctctatgatt atccaccttt 4620
ttaacctaaa agaataaag aaaaatatgt ttgcaatata attttatggg tgtatgttaa 4680
cttaattcat tatgttggcc tccagtttgc tgttgttagt tatgacagca gtagtgtcat 4740
taccatttca attcaagatta cattcctata tttgatcatt gtaaacctac tcttacatt 4800

```

gtattaaaaa	cagtggatat	tttaaagaag	ctgtacggct	tatatctagt	gctgtctctt	4860
aagactatta	aattgataca	acatatTTTaa	aagtaaatat	tacctaaatg	aatttttgaa	4920
attacaaata	cacgtgttaa	aactgtcggt	gtgttcaacc	atttctgtac	atacttagag	4980
ttaactgttt	tgccaggtc	tgtatgccta	ctcataatat	gataaaagca	ctcatctaat	5040
gctctgtaaa	tagaagctcg	tgttttccat	cagactgaac	tctcttgaca	agatgtggat	5100
gaatttcttt	aagtaaaatt	tttactcttg	tcatacattt	acagatcaaa	tgttagctcc	5160
caaagcaatc	atatggcaaa	gatagggtata	tcatagtttg	ccatttagct	gctttgtatt	5220
gctattatta	taaatagact	tcacagtttt	agacttgctt	aggtgaaatt	gcaattcttt	5280
ttactttcag	tcttagataa	caagtcttca	attatagtac	aatcacacat	tgcttaggaa	5340
tgcatcatta	ggcgattttg	tcattatgca	aacatcatag	agtgtactta	cacaaaccta	5400
gatagtatag	ctctttatgta	cttaggcctg	atggtatagt	ctgttgctcc	tagggcacia	5460
acctgtacaa	cgtttactgt	actgaatact	atagacagtt	gtaacacagt	ggtaaatat	5520
tatctaaata	tatgcaaaca	gagaaaaggt	acagtaaaag	tatggtataa	aagataatgg	5580
tatacctgtg	taggccactt	accacgaatg	gagcttgcat	gactagaagt	tgtctgtggg	5640
gagtcaagtga	gtgagtggtg	aattaatgtg	aaggcctaga	acactgtaca	ccactgtaga	5700
ctataataac	agtaocgtga	agctacaccca	aatttatctt	aacagttttt	cttcaataaa	5760
aaattataac	tttttaacct	tgtaaacttt	tttaattttt	aactttttaa	atacttagct	5820
tgaaacacaa	atatactgta	tagctataca	aaaataattt	ttcttgtgat	cttattctca	5880
gaagcttttt	totattttot	attttaaatt	ttttttttta	cttggttagtc	gtttttgtta	5940

aaaaactaaaa	cacacacact	ttcacctagc	catagacagg	attaggatac	tcaagtatac	6000
tcccttccac	ctcactgcct	tcacctcca	catcttgctc	cactggaagg	tttttagggg	6060
caataacaca	catgtagctg	tcacctatga	taacagtgct	ttctgttgaa	tacctoctga	6120
aggcattgcc	tgaggctgtt	ttacatttaa	cttaaaaaaa	aaaaaagtag	aaggagtgca	6180
ctctaaaaat	acaataaaag	gcatagtata	gtgaatacac	aaaccagcaa	tgtagtattt	6240
tattatcaag	tgtgtgtcac	tgtaataatt	tgtgtgtcta	tacttttaaat	actatgcaaa	6300
atagtactaa	gaccttatga	tggttacagt	gtcactaagg	caatagcata	ttttcaggtc	6360
cattgtaatc	taatgggact	accatcatat	atgcaqtcta	ccattgactg	aaacgttaca	6420
tggcacataa	ctgtatttgc	aagaatgatt	tgtttttacat	taatatcaca	taggatgtac	6480
cttttttagag	tggtagtgtt	atgtggatta	agatgtacaa	gttgagcaag	gggaccaaga	6540
gcctcgggtt	ctgtcttggg	tgtgagcgtt	tatgttcttc	tcctcatgtc	tgttttctca	6600
ttaaattcaa	aggettgaac	gggcoctatt	tagcccttct	gttttctacg	tgttctaaat	6660
aactaaagct	tttaaaattc	agccatttag	tgtagaactc	tccttgcaag	gatgaaatgc	6720
tgtatttggt	tcttggtcag	catattaaat	attttttatc	ttgtcttgat	acttcaatgt	6780
cgttttaaac	atcaggatcg	ggcttcagta	ttctcataac	caggaggttc	actgaggata	6840
caggactggt	tgcceatttt	ttgttatggc	tcacagattg	gtgtatttcc	atgtcttttt	6900
tttttttttt	ttttttgacc	tttttagoggc	tttaaaagta	ttctgttgtt	aggtgttgtt	6960
ttacttttct	aagattactt	aacaaagcac	cacaaactga	gtggctttaa	acaacagcaa	7020
tttattctct	cacaattcta	gaagctagaa	gtccgaaatc	aaagtgttga	cagggggcatg	7080
atcttcaaga	gagaagactc	ttctcttgcc	tcttctcggc	ttctggttgt	taccagcaat	7140
cctgagtgtt	ctcttcttgc	ctttagtgtt	acaacaatca	gtatctgcct	tttgtttcca	7200
catggctgtc	taccatttgt	ctctgtgtct	ccaaatctct	ctccttataa	acacagcagt	7260
tattggatta	ggcccccactc	taatccagta	tgaccccat	ttaacatgat	tacacttatt	7320
tctagataag	gtcacattca	cgtacaccaa	gggttaggaa	ttgaacatat	ctttttgggg	7380
gacacaattc	aaccacaag	tgtcagctc	tagctgagcc	tttcccttcc	tgtttttctc	7440
cttttttagt	gctatggggt	aggggccaaa	tctccagctc	tactagaatt	gcacatggac	7500
tggatatattg	ggaatactgc	gggtctattc	tatgagcttt	agtatgtaac	atttaatatc	7560
agtgtaaaga	agcccttttt	taagttattt	ctttgaattt	ctaaatgtat	gccttgaata	7620
taagtaacaa	gtttaccatgt	cttgtaaaat	gatcatatca	acaaacattt	aatgtgcacc	7680
tactgtgcta	gttgaatgtc	tttatctctga	taggagataa	caggattcca	catctttgac	7740
ttaagaggac	atgaacaaata	tgtctaaatc	attttggggt	tttagtgata	tctttaaatt	7800
gctgaaccta	atcatttgggt	tcatatgtca	ttgttttagat	atctccggag	cattttggata	7860
atgtgacagt	tggaatgcag	tgtatgtcac	tctttgccc	ccgccatctc	cagctgttgc	7920
caagacagag	attgtcttaa	gtggcaaatc	acctttatta	gcagctactt	ttgcttactg	7980
ggacaatat	cttggtccta	gagtaaggca	cattttgggt	ccaaagacag	aacaggtact	8040
tctcagtgat	ggagaaataa	cttttcttgc	caaccacat	ctaaatggag	aaactcttct	8100
aaatgcagag	agtggtgcta	ttagtgtaaa	gttttttgtc	ttgtctgaaa	agggagtgat	8160
tattgtttca	ttaatctttg	atggaaaactg	gaatggggat	cgagcacat	atggactatc	8220
aattataact	ccacagacag	aacttagttt	ctacctcca	cttcatagag	tgtgtgttga	8280
tagattaaac	catataatcc	ggaaaggaag	aatatggatg	cataaggtaa	gtgatttttc	8340
agcttattaa	tcatgtttaa	ctactctgtg	aaagcttatt	ttctgtgata	tataaatctt	8400
attttttttaa	ttatatgcag	tgaacatcaa	acaataaatg	ttatttattt	tgcatttacc	8460
ctattagata	caaatacatc	tggctcgata	cctgtcatct	tcatattaac	tgtggaaggt	8520
acgaaatggg	agctccacat	tatagatgaa	aagctaaagc	ttagacaaat	aaagaaactt	8580
ttagaccctg	gattcttctt	gggagccttt	gactctaata	ccttttgatt	ccctttcatt	8640
gcacaattct	gtcttttget	tactactatg	tgtaaagtata	acagttccaa	gtaaatgttt	8700
cataaagctg	tggctcatgta	gccttttggtc	tcttttaact	ctttgcgaag	ttcccagggt	8760
cataaaatga	ggaggttgaa	tggaatgggt	cccaagagaa	ttccttttaa	tcttacagaa	8820
attattgttt	tcttaaatcc	tgtagtggaa	tatataatgc	tatttacatt	tcagtatagt	8880
tttgatgtat	ctaaagaaca	cattgaaatt	tacttctgtg	gttccaggtt	gatactaacc	8940
ttacagttaa	tttaagacatt	ccatttttaa	ccatttttaa	ccatttttaa	ccatttttaa	9000

tgaaaglcca	lcaagcaalla	ccagulllcaa	aaggculllg	cccaatagla	aggaaaaaa	9000
atatcttitta	aaagaataat	tttttactat	gtttgcagge	ttacttcctt	ttttctcaca	9060
ttatgaaact	cttaaaatca	ggagaatcct	ttaaacaaca	tcataatggt	taatttgaaa	9120
agtgcgaagtc	attcttttcc	tttttgaaac	tatgcagatg	ttacattgac	tgttttctgt	9180
gaagtatatct	ttttttoact	gcagaataaa	ggttgttttg	atcttatattt	gtattgttta	9240
tgagaacatg	catttgttgg	gttaattttcc	tacccctgccc	cccatTTTTT	ccctaaagta	9300
gaaagtattt	ttcttgtgaa	ctaaatttact	acacaagaac	atgtctattg	aaaaataagc	9360
aagtatcaaa	atggtgtggg	tgtttttttt	aaataaattt	tccttgctc	aggaaagaca	9420

agaaaatgtc	cagaagatta	tottagaagg	cacagagaga	atggaagatc	aggtatatgc	9480
aaattgcata	ctgtcaaatg	tttttctcac	agcatgtatc	tgtataaggt	tgatggctac	9540
atttgtcaag	gccttgagga	catacgaata	agcctttaat	ggagctttta	tgaggggtga	9600
cagaataaac	tggaggaaga	tttccatata	ttaaacccaa	agagttaaat	cagtaaacaa	9660
aggaaaaatag	taattgcata	tacaaatttaa	tatttgctcc	cttttttttt	ctgtttgccc	9720
agaataaatt	ttggataact	tgttcatagt	aaaaataaaa	aaaattgtct	ctgatattgt	9780
ctttaaggta	ctacttctcg	aacctttccc	tagaagttagc	tgtaacagaa	ggagagcata	9840
tgtacccctg	aggtatctgt	ctggggtgta	ggcccaggtc	cacacaatat	ttcttctaag	9900
tottatgttg	tatcgttaag	actcatgcaa	tttacatttt	attccataac	tatttttagta	9960
ttaaaatttg	tcagtgtatc	ttcttaccct	ctcctctagg	aaaatgtgcc	atgtttatcc	10020
cttggtcttg	aatgcccctc	aggaacagac	actaagagtt	tgagaagcat	ggttacaagg	10080
gtgtggcttc	ccctgoggaa	actaagtaca	gactatttca	ctgtaaagca	gagaagttct	10140
tttgaaggag	aatctccagt	gaagaaagag	ttcttcaact	ttacttccat	ttcctcttgt	10200
gggtgaccct	caatgtctct	tgtaaaactc	caatatTTTt	aacatggctg	ttttgcccct	10260
ctttgcttct	ttttagcatg	aatgagacag	atgatacttt	aaaaaagtaa	ttaaaaaaa	10320
aaacttgtga	aaatacatgg	ccataataca	gaaccaata	caatgatctc	ctttaccaaa	10380
ttgttatgtt	tgtacttttg	tagatagctt	tccaattcag	agacagttat	ctgtgtgtaa	10440
ggctcgaact	acacaagaaa	gatttccctt	tacccaaaga	atcccagtoc	ttatttgctg	10500
gtcaataaagc	agggtcccca	ggaatggggt	aactttcagc	accctctaac	ccactagtta	10560
ttagtagact	aattaagtaa	acttatcgca	agttgaggaa	acttagaacc	aactaaaatt	10620
ctgcttttac	tgggattttg	ttttttcaaa	ccagaaacct	ttacttaagt	tgactactat	10680
taatgaattt	tgggtctctc	tttaagtgtc	cttcttaaaa	atgttatctt	actgctgaga	10740
agttcaagtt	tgggaagtac	aaggaggaa	agaaacttaa	gagattttct	tttagagcct	10800
cttctgtatt	tagccctgta	ggattttttt	tttttttttt	ttttttgggtg	ttgttgagct	10860
tcagttaggc	tattcattca	cttatactga	taatgtctga	gatactgtga	atgaaatact	10920
atgtatgctt	aaacctaaqa	ggaaatatTT	tcccaaaatt	attcttcccg	aaaaggagga	10980
gttgcccttt	gattgagttc	ttgcaaactc	cacaacgact	ttattttgaa	caatactgtt	11040
tggggatgat	gcattagttt	gaaacaactt	cagttgtagc	tgtcatctga	taaaattgct	11100
tcacagggaa	ggaaatttaa	cacggatcta	gtcattatct	ttgttagatt	gaatgtgtga	11160
attgtaattg	taaacaggca	tgataattat	tacttttaaa	actaaaaaca	gtgaatagtt	11220
agttgtggag	gttactaaag	gatgggtttt	ttttaaataa	aactttcagc	attatgcaaa	11280
tgggcatatg	gcttaggata	aaactttccag	aagttagcatc	acattttaaat	tctcaagcaa	11340
cttaataata	tggggctctg	aaaaactggt	taaggttact	ccaaaaatgg	ccctgggtct	11400
gacaaagatt	ctaaacttaa	gatgcttatg	aagactttga	gtaaaatcat	ttcataaaa	11460
aagttaggaa	aaacaactag	tattaaattc	atcttaataa	atgtatgatt	taaaaaatat	11520
gttttagctaa	aaatgcatag	tcatttgaca	atttcattta	tatctcaaaa	aatttactta	11580
accaagttgg	tcacaaaact	gatgagactg	gtgggtggtag	tgaataaatg	agggaccatc	11640
catatttgag	acacttttaca	tttgtgatgt	gttatactga	attttcagtt	tgattctata	11700
gactacaaat	ttcaaaatta	caattttcaag	atgtaataag	tagtaatatc	tgaaatagc	11760
tctaaaggga	atttttctgt	tttattgatt	cttaaaatat	atgtgtgat	tttgatttgc	11820
atttgggtag	attatacttt	tatgagtatg	gaggttaggt	attgattcaa	gttttccctta	11880
cctatttggg	aaggattttca	aagtcttttt	gtgcttggtt	ttctcatttt	ttaaatatga	11940
aatatattga	tgacctttta	caaatttttt	ttatctcaaa	ttttaaagga	gatcttttct	12000
aaaagagga	tgaatcatta	atcattgcat	gttaacagta	acgataaacc	aatgattcca	12060
tactctctaa	agaataaaa	tgagctttag	ggccgggcat	ggtcagaaat	ttgacaccaa	12120
cctggccaac	atggcgaaac	cccgctctca	ctaaaaatac	aaaatcagc	cgggcatggt	12180
ggcggcacct	atagtccag	ctacttgga	ggatgagaca	ggagagtcac	ttgaacctgg	12240
gaggagaggt	tgcaagttagc	tgagatcacg	ccattgcact	ccagcctgag	caatgaaagc	12300
aaaactccat	ctcaaaaaaa	aaaaaagaaa	agaaagaata	aaagttagct	ttggatttga	12360
tataaactct	ttagacatgt	agtagacttg	tttgatactg	tgtttgaaca	aattacgaag	12420
tattttctatc	aaagaatgtt	attgtttgat	gttattttta	ttttttattg	cccagcttct	12480
ctcatattac	gtgattttct	tcacttcatg	tcactttatt	gtgcagggtc	agagtattat	12540
tccaatgctt	actggagaag	tgattcctgt	aatggaactg	cttccatcta	tgaatcaca	12600
cagtgttctc	gaagaaatat	atgtaagttt	aaatgagagc	aattatacac	tttatgagtt	12660
ttttggggtt	atagtattat	tatgtatat	attaatatc	taattttta	agtaaggact	12720
ttgtcataca	tactattcac	atacagtatt	agccacttta	gcaataaagc	acacacaaaa	12780
tcttggtttt	tatggcaaaa	cagaggcatt	tttgatcag	gatgacaaaa	ttaaattcat	12840
tttgtttatt	tcattacttt	tataattcct	aaaagtggga	ggatccagc	tottatagga	12900

gcaattaata	tttaatgtag	tgtcttttga	aacaaaactg	tgtgcaaaag	tagtaaccat	12960
taatggaagt	ttacttgtag	tcacaaattt	agtttccctta	atcatttgtt	gaggacgttt	13020
tcaatgagga	atattgagga	tttaaggaata	acttttagga	actatttctta	ttcttttctg	13080

tyaaccacac	accatgagty	ccaagayata	cccttagyaa	accatctcty	ccgtctctcty	13000
atctttgtcat	ttaggttagt	ctcctgattc	tgacagctca	gaagaggaa	ttgttcttct	13140
aaaaattggt	taacctgctt	gaccagcttt	cacatttggt	cttctgaagt	ttatggtagt	13200
gcacagagat	tgttttttgg	ggagtcttga	ttctcggaaa	tgaaggcagt	gtgttatatt	13260
gaatccagac	ttccgaaaac	ttgtatatta	aaagtgttat	ttcaacacta	tgttacagcc	13320
aqactaatit	ttttatitit	tgatgcattt	tagatagctg	atacagtaet	caatgatgat	13380
gatattgggt	acagctgtca	tgaaggcttt	cttctcaagt	aagaattttt	cttttcataa	13440
aagctggatg	aagcagatac	catcttatgc	tcacctatga	caagatttgg	aagaaagaaa	13500
ataacagact	gtctacttag	attgttctag	ggacattacg	tatttgaact	gttgcttaaa	13560
tttgtgttat	ttttcactca	ttatatctct	atatataatt	ggtgttattc	catttgctat	13620
ttaaagaaac	cgagtttcca	tcccagacaa	gaaatcatgg	cccttgctt	gattctgggt	13680
tcttgtttta	cttctcatta	aagctaacag	aatcctttca	tattaagttg	tactgtagat	13740
gaacttaagt	tatttaggcg	tagaacaata	ttattcatat	tttaactgat	ctttttccat	13800
ccagcagtg	agttagtagt	ttaaagagtt	gtgaccttaa	accagactcc	ctggattaat	13860
gctgtgtacc	cgtgggcaag	gtgcttgaat	tctctatata	cctatttctt	catctgtaaa	13920
atggcaataa	tagtaatagt	acctaattgt	taggggtgtt	ataagcattg	agtaagataa	13980
ataatataaa	gcacttagaa	cagtgcctgg	aacataaaaa	cacttaataa	tagctcatag	14040
ctaacatttc	ctattttacat	ttcttctaga	aatagccagt	atttgttgag	tgccctacatg	14100
ttagtctctt	tactagttag	tttcatgtta	ttatcttata	ttctgtttta	aagtttcttc	14160
acagttacag	atcttcatga	aattttactt	ttataaaaag	agaagtaaaa	gtataaaagta	14220
ttcactttta	tggtcacagt	cttttctctt	aggctcatga	tggaagtatca	gaggcatgag	14280
tgtgtttaac	ctaagagcct	taatggcttg	aatcagaagc	acttttagtcc	tgtatctgtt	14340
cagtgtcagc	cttccataca	tcatttttaa	tcccatttga	ctttaagtaa	gtcacttaat	14400
ctctctacat	gttcaattct	tcagctataa	aatgatgtga	tttcaataaa	taaatacatt	14460
aattaaatga	tattatactg	actaattggg	ctgttttaag	gctcaataag	aaaatttctg	14520
tgaaggtctt	ctagaaaatg	taggttccct	tacaaaataa	agataacatt	gtgcttatag	14580
cttcggtgtt	tatcatataa	agctattctg	agttatttga	agagctcacc	tacttttttt	14640
tgtttttagt	ttgttaaatt	gttttatagg	caatgttttt	aatctgtttt	ctttaactta	14700
cagtgccatc	agctcacact	tgcaaacctg	tggtctgttc	gtttagtagt	gtagcagtg	14760
agagaaagta	aataaggtag	tttattttat	aatctagcaa	atgatttgac	tctttaagac	14820
tgatgatata	tcatggattg	tcattttaat	ggttaggttg	aattaaaatg	atctagtagt	14880
ataaggaggc	aatgtaatct	catcaaatg	ctaagacacc	ttgtggcaac	agtgagtttg	14940
aaataaaact	agtaagaatc	atcttatcag	ttattttgat	agctcggaaa	taccagtgct	15000
agtagtgtat	aaatggtttt	gagaatatac	taaaatcaga	tatataaaaa	aaattactct	15060
tctatttccc	aatgttatct	ttaacaaatc	tgaagatagt	catgtacttt	tggttagtagt	15120
tccaaagaaa	tggtatttgt	ttattcatct	tgatttccatt	gtcttcgctt	tccttctaaa	15180
tctgtccctt	ctagggagct	attgggatta	agtgtgcatt	gattattata	ctttattcag	15240
taatgtttct	gaccttttcc	ttcagtgcta	cttgagttaa	ttagggatta	atgaacagtt	15300
acatttccaa	gcattagcta	ataaactaaa	ggatttttga	cttttcttca	ctgaccatta	15360
gttagaaaga	gttcagagat	aagtatgtgt	atctttcaat	ttcagcaaac	ctaatttttt	15420
aaaaaaaagt	ttacatagga	aatatgtttg	aaatgatact	ttacaaagat	attcataatt	15480
tttttttgta	atcagctact	ttgtatatct	acatgagcct	taatttatat	ttctcatata	15540
accatttatg	agagcttagt	atacctgtgt	catttatattg	catctacgaa	ctagtgcact	15600
tattctctct	gttacctcaa	acaggtggct	ttccatctgt	gatctccaaa	gccttaggtt	15660
gcacagagtg	actgcccagc	tgctttatga	agggagaaa	gctccatagt	tggaagtgtt	15720
tttttttttt	tttttaaacat	ttttcccatc	ctccatcctc	ttgagggaga	atagcttacc	15780
ttttatcttg	ttttaatttg	agaaagaagt	tgccaccact	ctaggttgaa	aaccactcct	15840
ttaacataat	aactgtggat	atgggttgaa	tttcaagata	gttaccatgc	tttttatttt	15900
tcctaataga	gctgtaggct	aaatattatt	agaatcagat	ttctaaatcc	cacccaatga	15960
cctgcttatt	ttaaatcaaa	ttcaataatt	aattctcttc	tttttgagg	atctggacat	16020
tctttgatat	ttcttacaac	gaattttcatg	tgtagaccca	ctaaacagaa	gctataaaaag	16080
ttgcatgggt	aaataagctc	gagaaagctc	gcagatgata	taattcacct	gaagagtcac	16140
agtatgtagc	caaatgttaa	agggttttgag	atgccatata	gtaaatttac	caagcatttt	16200
ctaaatttat	ttgaccacag	aatccctatt	ttaaagcaaca	actgttacat	cccatggatt	16260
ccaggtgact	aaagaatact	tatttcttag	gatatgtttt	attgataata	acaattaaaa	16320
tttcagatat	ctttcataag	caaatcagtg	gtctttttac	ttcatgtttt	aatgctaaaa	16380

tattttcttt	tatagatagt	cagaacatta	tgcttttttc	tgactccagc	agagagaaaa	16440
tgctccaggt	tatgtgaagc	agaatcatca	tttaaatatg	agtcagggtc	ctttgtacaa	16500
ggcctgctaa	aggtatagtt	tctagttatc	acaagtgaag	ccacttttct	aaaatcattt	16560
ttgagactct	ttatagacaa	atcttaataa	ttagcattta	atgtatctca	tattgacatg	16620
cccagagact	gacttctctt	acacagttct	gcacatagac	tatatgtctt	atggatttat	16680
agttagtatc	atcagtgaag	caccatagaa	taccttttgt	gttccagggt	ggtccctgtt	16740
cctacatgtc	tagctccagc	actttttttt	ttttaacaca	tgcttaaatc	aggttgacaa	16800
tcaaaaaata	gatcatttct	ttttaactaa	atagatttga	attttattga	aaaaaaattt	16860
laaacalccl	laagaagccl	alaggallla	agcaallccl	alglalglgl	acaaaaalal	16920
atatatttct	atatataata	tatattagaa	aaaaattgta	tttttctttt	atttgagctt	16980
actgtcaagg	agcaaaacag	agaaatgtaa	attagcaatt	atttataata	cttaaaggga	17040
agaaagtgtg	tcaccttggt	gaatctatta	ttgttatctt	aattatagtc	ccaagacgtg	17100
aagaaatagc	tttccctaag	gttatgtgat	tgctctcatg	tgactacttt	cttgaggatg	17160
tagccacggc	aaaatgaaat	aaaaaaattt	aaaaattgtt	gcaaatacaa	gttatattag	17220

gctttttgtgc	attttcaata	atgtgctgct	atgaactcag	aatgatagta	tttaaata	17280
gaaactagtt	aaaggaaacg	tagtttctat	ttgagttata	catatctgta	aattagaact	17340
tctctgttta	aaggcataat	aaagtgccta	atacttttgt	ttcctcagca	cctctcatt	17400
taattatata	attttagttc	tgaaggagg	ctataccaga	tgccatagagg	aaatttcaaa	17460
actatgatct	aatgaaaaaa	tatttaatat	ttctccatgc	aaatacaaat	catatagttt	17520
tccagaaaaat	accttttgaca	ttatacaaaag	atgattatca	cagcattata	atagtaaaaa	17580
aatggaaata	gcctctttct	tctgtttctgt	tcatagcaca	gtgcctcata	cgcagtaggt	17640
tattattaca	tggtaaactgg	ctaccccaac	tgattaggaa	agaagtaaat	ttgttttata	17700
aaaatacata	ctcatttgagg	tgcatagaat	aattaaagaa	ttaaaagaca	cttgtaattt	17760
tgaatccagt	gaatacccac	tggtaatatt	tggtatatct	ctttctagtc	tttttttccc	17820
ttttgcatgt	attttcttta	agactcccac	ccccactgga	tcatctctgc	atgttctaat	17880
ctgccttttt	cacagcaqat	tctaaqcctc	tttgaatatc	aacacaaact	tcaacaactt	17940
catctataga	tgccaaaata	taaatttcatt	tttatttact	taaccacttc	ctttggatgc	18000
ttaggtcatt	ctgatgtttt	gctattgaaa	ccaatgctac	actgaacact	tctgtcacta	18060
aaacttttgc	cacactcatg	aatagctttc	taggataaat	ttttagagat	ggatttgcta	18120
aatcagagac	catttttttaa	aattaaaaaa	caattattca	tatcgtttgg	catgtaagac	18180
agtaaatttt	ccttttattt	tgacaggatt	caactggaag	ctttgtgctg	cctttccggc	18240
aagtcatgta	tgtcccatat	cccaccacac	acatagatgt	ggatgtcaat	actgtgaagc	18300
agatgccacc	ctgtcatgaa	catattttata	atcagcgtag	atacatgaga	tccgagctga	18360
cagccttctg	gtagccact	tcagaagaag	acatggctca	ggatacgcac	atctacactg	18420
acgaaagctt	tactcctgat	ttgtacgtaa	tgctctgcct	gctgggtactg	tagtcaagca	18480
atatgaaatt	gtgtctttta	cgaataaaaa	caaaacagaa	gttgcattta	aaaagaaaga	18540
aatattacca	gcagaattat	gottgaagaa	acatttaatc	aagcattttt	ttctttaaag	18600
ttcttctttt	tccatacaat	tgtgttttacc	ctaaaaatagg	taagattaac	cottaaagta	18660
aatatttaac	tatttgttta	ataaatatat	attgagctoc	taggcactgt	totaggtacc	18720
gggcttaata	gtggccaacc	agacagcccc	agccccagcc	cctacattgt	gtatagtcta	18780
ttatgtaaca	gttattgaat	ggactttatta	acaaaaacca	agaagtaatt	ctaagtcttt	18840
tttttcttga	catatgaata	taaaatatag	caaaactgtt	aaaatatatt	aatggaacat	18900
ttttttactt	tgcattttat	attgtttattc	acttcttatt	tttttttaaa	aaaaaaagcc	18960
tgaacagtaa	attcaaaaagg	aaaagtaatg	ataattaatt	gttgagcatg	gacccaactt	19020
gaaaaaaaaa	atgatgatga	taaatctata	atcctaaaaa	cctaagtaaa	cacttaaaag	19080
atgtttctgaa	atcaggaaaa	gaattatagt	atacttttgt	gtttctcttt	tatcagttga	19140
aaaaaggcac	agtagctcat	gctgttaaga	acagagcttt	gggagtgcac	ggcaggcgga	19200
tcacttgagg	ccaggagttc	cagaccagcc	tgggcaacat	agtgaaccc	catctctaca	19260
aaaaataaaa	aagaattatt	ggaatgtgtt	tctgtgtgoc	tgtaatccta	gctattccga	19320
aagctgaggc	agaggagctc	tttgagccca	ggagttttgag	gttacaggga	gttatgatgt	19380
gccagtgtac	tccagcctgg	ggaacaccga	gactctgtct	tatttaaaaa	aaaaaaaaaa	19440
aaaatgcttg	caataatgcc	tggcacatag	aaggtaacag	taagtgttaa	ctgtaataac	19500
ccaggctctaa	gtgtgttaagg	caatagaaaa	attggggcaa	ataagcctga	cctatgtatc	19560
tacagaatca	gtttgagctt	aggtaacaga	cctgtggagc	accagtaatt	acacagtaag	19620
tgttaaccac	aagcatagaa	taggaatatc	ttgttcaagg	gacccccagc	cttatacatc	19680
tcaaggtgca	gaagatgac	ttaatatagg	accatttttt	tcctagttct	ccagagtttt	19740
tattggttct	tgagaaagta	gtaggggaat	gttttagaaa	atgaattggg	ccaactgaaa	19800
ttacatgtca	gtaagttttt	atatattggg	aaattttagt	agacatgtag	aagttttcta	19860

attaatctgt	gccttgaaac	attttctttt	ttcctaaagt	gcttagtatt	ttttccgttt	19920
tttgattggg	tacttgggag	ctttttttgag	gaaatttagt	gaactgcaga	atgggtttgc	19980
aaaccatttg	tatttttgtt	ttgtttttta	gaggatgtat	gtgtatttta	acattttctta	20040
atcattttta	gccagctatg	ttgtttttgc	tgatttgaca	aactacagtt	agacagctat	20100
tctcattttg	ctgatcatga	caaaaataata	tctgaattt	ttaaattttg	catccagctc	20160
taaattttct	aaacataaaa	ttgtccaaaa	aatagtattt	tcagccacta	gatttgttgt	20220
taagtctatt	gtcacagagt	catttttactt	ttaaqtatat	gtttttacat	gttaattatg	20280
tttgttattt	ttaatttttaa	cttttttaaaa	taattccagt	cactgccaat	acatgaaaaa	20340
lllgllcactg	gaalllllll	lllgalllll	alllllagll	calglglaca	lglgcagglg	20400
tgttatacac	gtaaattgcg	tgtcatgagg	gtttgtgtga	caggtgattt	cattaccocag	20460
gtaataagca	tagtaccocaa	taggtagttt	tttgatcctc	acccttctcc	cacctcaag	20520
taggcctggg	tgtgtctgtt	tccttctttg	tgtccatgta	tactcagtgt	ttagctccca	20580
cttagaagtg	agaacatgcg	gtagttgggt	ttctgttcc	ggattagttc	acttaggata	20640
atgacctcta	gctccatctg	gtttttatgg	ctgcatagta	ttccatgggt	tatatgtatc	20700
acattttctt	tatccagctc	accatttgata	ggcatttagg	ttgattccct	gtctttgtta	20760
tcatgaatag	tgtgtgtgat	aacatacaca	tgcatgtgtc	tttatggtag	aaaaatttgt	20820
attcctttag	gtacatatag	aataatgggg	ttgctagggg	gaatggtagt	tctattttca	20880
gttatttgag	aaatcttcaa	actgcttttc	ataatagcta	aactaattta	cagtcccgcc	20940
agcagtgtat	aagtgttccc	ttttctccac	aaccttgcca	acatctgtga	ttttttgact	21000
ttttaataat	agccatttct	agagaattga	tttgcatttc	tctattagtg	atattaagca	21060
ttttttcata	tgtcttttag	ctgtctgtat	atatcttctc	gaaaaatttt	catgtccctt	21120
gcccagtttg	tagtgggggtg	gggtgttttt	tgtttgttaa	ttagttttaa	gttccctcca	21180
gattctgcat	atccctttgt	tggtatcatg	gtttgcagat	atttttctcc	cattgtgtag	21240
gttgtctttt	actctgttga	tagtttcttt	tgccatgcag	gagctogtta	ggccccattt	21300
gtgtttgttt	ttgttgacgt	tgtttttggc	gtcttcatca	taaaatctgt	gccagggcct	21360
atctccadac	tgtattttcc	taoattatct	tccaagtttt	ttacaatttt	aatttttacc	21420

tttatgtctt	taatccatct	tgagttgatt	tttgtatatg	gcacaaggaa	ggggtccagt	21480
ttcactccaa	ttcctatggc	tagcaattat	cccagcacca	tttatgtaat	acggagtcct	21540
ttccccattg	cttgtttttt	gtcaactttg	ttgaagatca	gatgggttga	agtgtgtggc	21600
tttattttct	ggctctctat	tctccattgg	tctatgtgtc	tgtttttata	acagtaccct	21660
gctgtttcagg	ttcctatagc	cttttagtat	aaaatcggtc	aatgtgatgc	ctccagcttt	21720
gttctttttg	cttaggattg	ctttggctat	ttgggctcct	ttttgggtcc	atattaattt	21780
taaaacagtt	ttttctggtt	ttgtgaagga	tatcattggg	agtttatagg	aatagcattg	21840
aatctgtaga	ttgctttggg	cagtatggcc	attttaacaa	tattaattct	tctatcttat	21900
gaatatggaa	tggtttttcca	tgtgtttgtg	tcatctcttt	atacctgatg	tataaagaaa	21960
agctgggtatt	attcctactc	aatctgttcc	aaaaaattga	ggaggaggaa	ctcttcccta	22020
atgaggccag	cctcattctg	ataccaaaac	ctggcagaga	cacacagaaa	aaaagaaaac	22080
ttcaggccaa	tatccttgat	gaatatagat	gcaaaaaatc	tcaacaaaat	actagcaaac	22140
caaatccagc	agcacatcaa	aaagctgata	tactttgatc	aagtaggctt	tatccctggg	22200
atgcaagggt	ggttcaacat	acacaaatca	ataagtgtga	ttcctcactt	aaacagagct	22260
aaaaacaaaa	accacaagat	tatctcaata	ggtagagaaa	aggttgtcaa	taaaatttaa	22320
catcctccat	gttaaaaacc	ttcagttagt	cagggttagt	gactcacacc	tgtaatccca	22380
gcactttggg	aggccaaggc	gggcatactt	cttaagccca	ggagttcaag	acgagcctag	22440
gcagcatggt	gaaaccccat	ctctacaaaa	aaaaaaaaaa	aaaaaaatta	gcttgggtatg	22500
gtgacatgca	cctatagtcc	cagctattca	ggaggttgag	gtgggaggat	tgtttgagcc	22560
cgggaggcag	aggtttggcag	cgagctgaga	tcatgccacc	gcactccagc	ctgggcaacg	22620
gagtgagacc	ctgtctcaaa	aaagaaaaat	cacaaacaat	cctaaacaaa	ctaggcattg	22680
aaggaaacatg	cctcaaaaaa	ataagaacca	tctatgacag	acccatagcc	aatatcctac	22740
caaattgggca	aaagctggaa	gtattctcct	tgagaaccgt	aacaagacaa	ggatgtccac	22800
tctcaccact	ccttttcagc	atagttcttg	aagtctctagc	cagagcaatc	aggaaagaga	22860
aagaaagaaa	gacattcaga	taggaagaga	agaagtcaaa	ctatttctgt	ttgcaggcag	22920
tataattctg	tactagaaaa	atctcatagt	ctctgccagg	aaactcctaa	atctgtttaa	22980
aatttccagca	aagttttggc	attctctata	ctccaacacc	ttccaaagtg	agagcaaaaat	23040
caagaacaca	gtcccattca	caatagccgc	aaaacgaata	aaatacctag	gaatccagct	23100
aaccagggag	gtgaaagatc	tctatgagaa	ttacaaaaca	ctgtgaaag	aatccagaga	23160
tgacacaaac	aaattggaaat	gttctttttt	aacaccttgc	tttatctaat	tcacttatga	23220
tgaagatact	cattcagtggt	aacaggtata	ataagtcacc	tcgattaaat	ataagcctta	23280
ttctctttcc	agagcccaag	aaggggcact	atcagtgcgc	agtcataaat	gaagaaatgc	23340

taatatTTTT	cccttttacg	gtttctttct	tctgtagtgt	ggtacactcg	tttcttaaga	23400
taaggaaaact	tgaactacct	tctgttttgc	ttctacacat	acctattctc	tttttttgcc	23460
actctgggtca	ggtataggat	gatecctacc	actttcagtt	aaaaactcct	cctcttacta	23520
aatgttctct	tacctctctg	cctgagtaga	acctaggga	aatggaagag	aaaaagatga	23580
aaggagggtg	gggctctggg	agggaaataag	tagtctctgt	tgtttgtgtg	tttgcttttag	23640
cacctgtctat	atcctagggt	ctgtgttagg	cacacattat	tttazgtggc	cattatatta	23700
ctactactca	ctctgtgtcgt	tgccaaagta	tgtagtactt	tcttggtatg	ttggttctatg	23760
ttactttacag	atggtgggct	tggttaggca	aacctcagtg	ataatcatcg	gagtgtgttc	23820
lclaalclca	clcaaallll	llllcacall	llllgllllg	llllglllll	lgallgllagl	23880
ggcttatttt	tggtgtctgt	ttgttttttg	tttttttttg	agatggcaag	aattggtagt	23940
tttattttatt	aattgcttaa	gggtctctac	tttttttaaa	agatgagagt	agtaaaatag	24000
attgatagat	acatacatac	ccttactggg	gactgcttat	attcttttaga	gaaaaaatta	24060
catatttagcc	tgacaaaacac	cagtaaaaatg	taaaataatc	cttgagttaa	taaaatgaatg	24120
tatattttgt	gtctccaaat	atatatatct	atattctttac	aaatgtgttt	atatgtaata	24180
tcaattttata	agaactttaaa	atgtttggctc	aagtgaggga	ttgtggaagg	tagcattata	24240
tggccatttc	aacatttgaa	ctttttttctt	ttcttcattt	tcttcttttc	ttcaggaata	24300
ttttttcaaga	tgtcttacac	agagacactc	tagtgaaagc	cttctctgat	caggtaaatg	24360
ttgaaacttga	cttgtctcaga	gtgaatgata	tgacatgttt	tcttttttta	tatatcctac	24420
aatgcctgtt	ctatatatatt	atattccctt	ggatcatgcc	ccagagttct	gtcagcaaat	24480
tgcagtttaag	ttagttaacac	tacagttctc	agaagagtct	gtgaggcat	gtcaagtga	24540
tcattacatt	ggttgctctt	tgctctagat	ttatgcttcg	ggaattcaga	cctttgttta	24600
caatataata	aataattattg	ctatctttta	aagatataat	aataagatat	aaagttgacc	24660
acaactactg	ttttttgaaa	catagaattc	ctggtttaca	tgtatocaaag	tgaaatctga	24720
cttagctttt	acagatataa	tatatacata	tatatatcct	gcaatgcttg	tactatatat	24780
gtagtacaag	tatatatata	tgtttgtgtg	tgtatatata	tatagtacga	gcataatatac	24840
atattaccag	cattgttagga	tatatatatg	tttatatatt	aaaaaaaaag	tataaactta	24900
aaacccclall	alglallglg	gaglalalgl	lalalalgal	alglaaaaa	lalaacalal	24960
actctatgat	agagtgtaat	atatttttta	tatatatttt	aacattttata	aatgataga	25020
attaagaatt	gagtcctaat	ctgtttttatt	aggtgctttt	tgtagtgtct	ggcttttcta	25080
aagtgtctaa	atgatttttc	cttttgactt	attaatgggg	aagagcctgt	atattaacaa	25140
ttaagagtgc	agcattccat	acgtcaaaaca	acaaacattt	taattcaagc	attaacctat	25200
aacaagtaag	tttttttttt	tttttttgaga	aaggagggtt	gtttattttgc	ctgaaatgac	25260
tcaaaaaatat	ttttgaaaca	tagtgtactt	atttaaaata	catcttttatt	gtttcattct	25320
tttaaaaaaat	atctacttaa	ttacacagtt	gaaggaaato	gtagattata	tggaacttat	25380
ttcttaatat	attacagttt	gttataataa	cattctgggg	atcaggccag	gaaactgtgt	25440
catagataaa	gctttgaaat	aatgagatcc	ttatgtttac	tagaaatttt	ggattgagat	25500
ctatgaggtc	tgtgacatat	tgccaagtgc	aaggaaaatt	cgtaggcctg	gaatttcattg	25560
attctctctct	ttcctctctct	ttcctctctct	ttcctctctct	ttcctctctct	ttcctctctct	25620

```

cccccaagc cgacataaaa ccccccacac cccccacccc atcatalgca cacallctac 25620
tccatccacc ccactocacc ccccgcaaaa gtacaggtat atgaatgtct caaaaccata 25680
ggctcatctt ctaggagctt caatgttatt tgaagatttg gccagaaaaa attaatgtat 25740
acgaaataac ttatgtatga gttttaaaag tgaagtaaac atggatgtat tctgaagtag 25800
aatgcaaaat ttgaatgcat ttttaaagat aaattagaaa acttctaaaa actgtcagat 25860
tgtctgggcc tgggtgctta tgccgttaat cccagcactt tgggagtcct aggtgggtgg 25920
atcacaaagt caggagatcg agaccatcct gccaacatgg tgaacccccg tctctactaa 25980
gtatacaaaa attagctggg cgtggcagcg tgtgcctgta atcccagcta cctgggaggc 26040
tgaggcagga gaatcgcttg aaccacaggag gtgtaggttg cagtgaagtc agatcgcgcc 26100
actgcacttt agcctgggtg cagagctaga ctccgtctca aaaaaaaaaa aaaatatcag 26160
attgttccta cacctagtgc ttctatacca cactcctgtt agggggcctc agtggaatg 26220
gttaaggaga tgtttagtgt gtattgtctg ccaagcactg tcaacactgt catagaaact 26280
tctgtacgag tagaatgtga gaaaattatg tgttgaaatg gttcctctcc ctgcaggctc 26340
ttcagctgaa accctggctta tctctcagaa gtactttcct tgcacagttt ctacttgtcc 26400
ttcacagaaa agccttgaca ctaataaaat atatagaaga cgatacgtga gtaaaactcc 26460
tacacggaag aaaaaccttt gtacattggt tttttgtttt gtttcctttg tacattttct 26520
atatcataat ttttgcgctt cttttttttt tttttttttt tttttttcca ttatttttag 26580
gcagaaggga aaaaagccct ttaaactctc tcggaacctg aagatagacc ttgatttaac 26640
agcagagggc gatcttaaca taataatggc tctggctgag aaaattaaac caggctaca 26700
ctcttttatac tttggaagac ctttctacac tagtgtgcaa gaacgagatg ttctaataac 26760
tttttaaatg tgtaaactta taagcctatt ccatacaaat catgatecct ggtaaagtag 26820

```

```

ctcagtggtg tggggaaacg ttcccctgga tcatactcca gaattctgct ctcagcaatt 26880
gcagttaagt aagttacact acagttctca caagagcctg tgaggggatg tcaggtgcct 26940
cattacattg ggtgtctctt ttccctagatt tatgcttttg ggatacagac ctatgtttac 27000
aatataataa atattattgc tatcttttaa agatataata ataggatgta aacttgacca 27060
caactactgt ttttttgaaa tacatgattc atggtttaca tgtgtcaagg tgaactctga 27120
gttggctttt acagatagtt gaactttctat tctttggcat tctttgggtg gtagaattac 27180
tgtaataact ctgcaatcaa ctgaaaacta gagcctttta atgatttcaa ttccacagaa 27240
agaaagttag cttgaacata ggatgagctt tagaaagaaa attgatcaag cagatgttta 27300
attggaattg attattagat cctactttgt ggatttagtc cctgggattc agtctgtaga 27360
aatgtctaat agttctctat agtccttggt cctgggtgaac cacagttagg gtgttttgtt 27420
tattttattg ttcttgctat tgttgatatt ctatgtagtt gagctctgta aaaggaaatt 27480
gtattttatg ttttagtaat tgttgccaac tttttaaatt aattttcatt atttttgagc 27540
caaatgaaa tgtgcacctc ctgtgccttt tttctcctta gaaaatctaa ttacttgga 27600
caagttcaga tttcactggt cagtcatttt catcttggtt tcttcttgct aagtcttacc 27660
atgtacctgc tttggcaatc attgcaactc tgagattata aaatgcctta gagaatatac 27720
taactaataa gatctttttt tcagaaacag aaaatagttc cttgagtaact tccctcttgc 27780
atttctgcct atgtttttga agttgttgct gtttgccctg aataggctat aaggaatagc 27840
aggagaaatt ttactgaagt gctgttttcc taggtgctac tttggcagag ctaagttatc 27900
ttttgttttc ttactgcgtt tggaccattt tcttggtctat aaaataactg attaatataa 27960
ttctaacaca atgttgacat tgtagttaca caaacacaaa taaatatatt atttaaaatt 28020
ctggaagtaa tataaaaggg aaaatatatt tataagaaag ggataaaggc aatagagccc 28080
ttctgcccc caccaccaa atttacacaa caaatgaca tgttcgaatg tgaagggtca 28140
taatagcttt cccatcatga atcagaaaga tgtggacagc ttgatgtttt agacaaccac 28200
tgaactagat gactgttgta ctgtagctca gtcattttaa aaatatataa atactacctt 28260
gtagtgtccc atactgtgtt ttttacctgg tagattctta tttaatgtct aactggttat 28320
ttcttttggc tggttttatt tactgttata cagaatgtaa gttgtacagt gaaataagtt 28380
attaaagcat gtgtaaacat tgttatatat cttttctcct aaatggagaa ttttgaataa 28440
aatatatttg aaattttgcc tctttcagtt gttcattcag aaaaaaatac tatgatattt 28500
gaagactgat cagcttctgt tcagctgaca gtcctgctgg atctaaactt tttttaaaat 28560
taattttgtc ttttcaaaga aaaaatattt aaagaagcct tataatatata tcttatgtta 28620
aaaaaacttt ctgcttaact ctctggattt cattttgatt tttcaaatta tatattaata 28680
tttcaaagt aaaatactat tttagataaat tgttttttaa cattcttatt attataatat 28740
taataatacc taaactgaag ttattcatcc caggtatcta atacatgtat ccaaagtaaa 28800
aatccaagga atctgaacac tttcatctgc aaagctagga atagggttga cattttcact 28860
ccaagaaaaa gttttttttt gaaaatagaa tagttgggat gagaggtttc tttaaaagaa 28920
gactaactga tcacattact atgattctca aagaagaaac caaaacttca tataatacta 28980
taaagtaaat ataaaatagt tccttctata gtatatttct ataatgctac agtttaaaaa 29040
gatcactctt atataatact attttgattt tgatgtagaa ttgcacaaat tgatatttct 29100
cctatgatct gcagggtata gcttaaagta acaaaacacag tcaaccacct ccattttaa 29160
cacagtaaca ctatgggact agttttatta cttccatttt acaaatgagg aaactaaagc 29220
ttaaagatgt gtaatacacc gcccaaggtc acacagctgg taaagggtga tttcatccca 29280
gacagttaca gtcattgcca tgggcacagc tcctaactta gtaactccat gtaactggta 29340
ctcagtgtag ctgaattgaa aggagagtaa ggaagcaggt ttacaggtc tacttgcaact 29400
attcagagcc cgagtgtgaa tccctgctgt gctgcttggg gaagttactt aacctatgca 29460
aggttcattt tgtaaatatt ggaaatggag tgataataac tacttcacca gaggatttaa 29520
tgagacctta tacgatcctt agttcagta ctagctagtg cttaataaat gctttttct 29580
ccaatctgac aatctccagc ttgtaattgq ggcatttga acatttata tgattattgq 29640

```

```

catggtaggt taaagctgtc atcttgctgt ttctatttgg ttctttttgt ttctctctta 29700
cttttggatt tttttattct actatgtctt ttctatttgc ttattaacta tactctttga 29760
tttatttttag tggttgtttt aggggttatac ctctttctaa ttaccagtt tataaccagt 29820
ttatatacta ctigacatat agcttaagaa acttaactgt gttgtctttt tgctgttatg 29880
gtcttaacgt ttttatttct acaaacatta taaactccac actttattgt tttttaattt 29940
tacttataca gtcaattatc ttttaaagat atttaaatat aaacattcaa aacaccccaa 30000
t 30001

```

<210> 3

<211> 1031

<212> DNA

<213> Homo sapiens

<400> 3

```

attcccggga tacgtaacct acggtgtccc gctaggaaag agaggtgcgt caaacagcga 60
caagttccgc ccacgtaaaa gatgacgctt ggtgtgtcag ccgtccctgc tgcccggttg 120
cttctctttt gggggcgggg tctagcaaga gcaggtgtgg gtttaggaga tatctccgga 180
gcatttggtt aatgtgacag ttggaatgca gtgatgtcga ctctttgccc accgccatct 240
ccagctgttg ccaagacaga gattgcttta agtggcaaat cacctttatt agcagctact 300
tttgcttact gggacaatat tcttggtcct agagtaaggc acatttgggc tccaaagaca 360
gaacaggtac ttctcagtga tggagaaata acttttcttg ccaaccacac tctaaatgga 420
gaaatccttc gaaatgcaga gagtgggtgt atagatgtaa agttttttgt cttgtctgaa 480
aagggagtga ttattgtttc attaatcttt gatggaaact ggaatgggga tcgcagcaca 540
tatggactat caattatact tccacagaca gaacttagtt tctacctccc acttcataga 600
gtgtgtgttg atagattaac acatataatc cggaaaggaa gaatatggat gcataaggaa 660
agacaagaaa aatgtccaga agattatctt agaaggcaca gagagaatgg aagatcaggg 720
tcagagtatt attccaatgc ttactggaga agtgattcct gtaatggaaa ctgctttcct 780
ctatgaaatt ccccggtt cctggaggaa atagatatag gctgatacag ttaccaatg 840
atggatgaat attgggggac cgcttggtca ttgaaaggct ttcttttctc caggaaagaa 900
atttttttcc ttttccataa aaagcttggg aatggaagac aacaattccc attctttttt 960
tgcgttccac ccctatgtga caacagaaat ttttggggaa acaacaacga aaaaatttta 1020
tccgcgcgc a 1031

```

<210> 4

<211> 3244

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (108) ... (1553)

<400> 4

```

ggggggggct gcggttgccg tgccctgcgc cgcggcgccg gaggcgcagg cgggtggcgag 60
tggatatctc cggagcattt ggataatgtg acagttggaa tgcagtg atg tcg act 116
                                         Met Ser Thr
                                         1

ctt tgc cca ccg cca tct cca gct gtt gcc aag aca gag att gct tta 164
Leu Cys Pro Pro Pro Ser Pro Ala Val Ala Lys Thr Glu Ile Ala Leu
      5              10              15

agt ggc aaa tca cct tta tta gca gct act ttt gct tac tgg gac aat 212
Ser Gly Lys Ser Pro Leu Leu Ala Ala Thr Phe Ala Tyr Trp Asp Asn
      20              25              30              35

att ctt ggt cct aga gta agg cac att tgg gct cca aag aca gaa cag 260
Ile Leu Gly Pro Arg Val Arg His Ile Trp Ala Pro Lys Thr Glu Gln
      40              45              50

gta ctt ctg agt gat gga gaa ata act ttt ctt gcc aac cac act cta 308

```

Val Leu Leu Ser Asp Gly Glu Ile Thr Phe Leu Ala Asn His Thr Leu	
55 60 65	
aat gga gaa atc ctt cga aat gca gag agt ggt gct ata gat gta aag	356
Asn Gly Glu Ile Leu Arg Asn Ala Glu Ser Gly Ala Ile Asp Val Lys	
70 75 80	
ttt ttt gtc ttg tct gaa aag gga gtg att att gtt tca tta atc ttt	404
Phe Phe Val Leu Ser Glu Lys Gly Val Ile Ile Val Ser Leu Ile Phe	
85 90 95	
gat gga aac tgg aat ggg gat cgc agc aca tat gga cta tca att ata	452
Asp Gly Asn Trp Asn Gly Asp Arg Ser Thr Tyr Gly Leu Ser Ile Ile	
100 105 110 115	
ctt cca cag aca gaa ctt agt ttc tac ctc cca ctt cat aga gtg tgt	500
Leu Pro Gln Thr Glu Leu Ser Phe Tyr Leu Pro Leu His Arg Val Cys	
120 125 130	
gtt gat aga tta aca cat ata atc cgg aaa gga aga ata tgg atg cat	548
Val Asp Arg Leu Thr His Ile Ile Arg Lys Gly Arg Ile Trp Met His	
135 140 145	
aag gaa aga caa gaa aat gtc cag aag att atc tta gaa ggc aca gag	596
Lys Glu Arg Gln Glu Asn Val Gln Lys Ile Ile Leu Glu Gly Thr Glu	
150 155 160	
aga atg gaa gat cag ggt cag agt att att cca atg ctt act gga gaa	644
Arg Met Glu Asp Gln Gly Gln Ser Ile Ile Pro Met Leu Thr Gly Glu	
165 170 175	
gtg att cct gta atg gaa ctg ctt tca tct atg aaa tca cac agt gtt	692
Val Ile Pro Val Met Glu Leu Leu Ser Ser Met Lys Ser His Ser Val	
180 185 190 195	
cct gaa gaa ata gat ata gct gat aca gta ctc aat gat gat gat att	740
Pro Glu Glu Ile Asp Ile Ala Asp Thr Val Leu Asn Asp Asp Asp Ile	
200 205 210	
ggt gac agc tgt cat gaa ggc ttt ctt ctc aat gcc atc agc tca cac	788
Gly Asp Ser Cys His Glu Gly Phe Leu Leu Asn Ala Ile Ser Ser His	
215 220 225	
ttg caa acc tgt ggc tgt tcc gtt gta gta ggt agc agt gca gag aaa	836
Leu Gln Thr Cys Gly Cys Ser Val Val Val Gly Ser Ser Ala Glu Lys	
230 235 240	
gta aat aag ata gtc aga aca tta tgc ctt ttt ctg act cca gca gag	884
Val Asn Lys Ile Val Arg Thr Leu Cys Leu Phe Leu Thr Pro Ala Glu	
245 250 255	
aga aaa tgc tcc agg tta tgt gaa gca gaa tca tca ttt aaa tat gag	932
Arg Lys Cys Ser Arg Leu Cys Glu Ala Glu Ser Ser Phe Lys Tyr Glu	
260 265 270 275	
tca ggg ctc ttt gta caa ggc ctg cta aag gat tca act gga agc ttt	980
Ser Gly Leu Phe Val Gln Gly Leu Leu Lys Asp Ser Thr Gly Ser Phe	
280 285 290	
gtg ctg cct ttc cgg caa gtc atg tat gct cca tat ccc acc aca cac	1028
Val Leu Pro Phe Arg Gln Val Met Tyr Ala Pro Tyr Pro Thr Thr His	
295 300 305	
ata gat gtg gat gtc aat act gtg aag cag atg cca ccc tgt cat gaa	1076
Ile Asp Val Asp Val Asn Thr Val Lys Gln Met Pro Pro Cys His Glu	
310 315 320	
cat att tat aat cag cgt aga tac atg aga tcc gag ctg aca gcc ttc	1124

His Ile Tyr Asn Gln Arg Arg Tyr Met Arg Ser Glu Leu Thr Ala Phe
 325 330 335
 Lgg aga gcc acf lca gaa gaa gac alg gcl cag gal acg alc alc lac 1172
 Trp Arg Ala Thr Ser Glu Glu Asp Met Ala Gln Asp Thr Ile Ile Tyr
 340 345 350 355
 act gac gaa agc ttt act cct gat ttg aat att ttt caa gat gtc tta 1220
 Thr Asp Glu Ser Phe Thr Pro Asp Leu Asn Ile Phe Gln Asp Val Leu
 360 365 370
 cac aga gac act cta gtg aaa gcc ttc ctg gat cag gtc ttt cag ctg 1268
 His Arg Asp Thr Leu Val Lys Ala Phe Leu Asp Gln Val Phe Gln Leu
 375 380 385
 aaa cct ggc tta tct ctc aga agt act ttc ctt gca cag ttt cta ctt 1316
 Lys Pro Gly Leu Ser Leu Arg Ser Thr Phe Leu Ala Gln Phe Leu Leu
 390 395 400
 gtc ctt cac aga aaa gcc ttg aca cta ata aaa tat ata gaa gac gat 1364
 Val Leu His Arg Lys Ala Leu Thr Leu Ile Lys Tyr Ile Glu Asp Asp
 405 410 415
 acg cag aag gga aaa aag ccc ttt aaa tct ctt cgg aac ctg aag ata 1412
 Thr Gln Lys Gly Lys Lys Pro Phe Lys Ser Leu Arg Asn Leu Lys Ile
 420 425 430 435
 gac ctt gat tta aca gca gag ggc gat ctt aac ata ata atg gct ctg 1460
 Asp Leu Asp Leu Thr Ala Glu Gly Asp Leu Asn Ile Ile Met Ala Leu
 440 445 450
 gct gag aaa att aaa cca ggc cta cac tct ttt atc ttt gga aga cct 1508
 Ala Glu Lys Ile Lys Pro Gly Leu His Ser Phe Ile Phe Gly Arg Pro
 455 460 465
 ttc tac act agt gtg caa gaa cga gat gtt cta atg act ttt taa 1553
 Phe Tyr Thr Ser Val Gln Glu Arg Asp Val Leu Met Thr Phe
 470 475 480
 atgtgtaact taataagcct attccatcac aatcatgac gctggtaaag tagctcagtg 1613
 gtgtgqggaa acgttccctt ggatcatact ccagaattct gctctcagca attgcagtta 1673
 agtaagttac actacagttc tcacaagagc ctgtgagggg atgtcagggtg catcattaca 1733
 ttgggtgtct cttttcctag atttatgctt ttgggataca gacctatgtt tacaatataa 1793
 taaatattat tgctatcttt taaagatata ataataggat gtaaaacttga ccacaactac 1853
 tgtttttttg aaatacatga ttcattggtt acatgtgtca aggtgaaatc tgagttggct 1913
 ttacagata gttgactttc tatcttttgg cattcttggg tgtgtagaat tactgtaata 1973
 cttctgcaat caactgaaaa ctagagcett taaatgattt caattccaca gaaagaaagt 2033
 gagcttgaac ataggatgag ctttagaaaag aaaattgato aagcagatgt ttaattggaa 2093
 ttgattatta gatctacttt tgtggattta gtccctggga ttcagtctgt agaaatgtct 2153
 aatagttctc tatagtcttt gtctctgggt aaccacagtt aggggtgttt gtttatttta 2213
 ttgttcttgc tattgttgat attctatgta gttgagctct gtaaaaggaa attgtatttt 2273
 atgttttagt aattgttgcc aactttttta attaatcttc attatttttg agccaaattg 2333
 aaatgtgcac ctctgtgccc tttttctccc ttagaaaatc taattacttg gaacaagttc 2393
 agatttcaact ggtagctcat ttctcatctt ttttctcttc gctaagctctt accatgtacc 2453
 tgctttggca atcattgcaa ctctgagatt ataaaatgcc ttagagaata tactaactaa 2513
 taagatcttt ttttcagaaa cagaaaatag ttctttgagt acttctctct tgcattttctg 2573
 cctatgtttt tgaagtttgt gtgttttgcc tgcaataggc tataaggaat agcaggagaa 2633
 atttttactga agtgctgttt tcttaggtgc tactttggca gagctaagtt atcttttgtt 2693
 ttcttaatgc gtttggaacca ttttgcctgg tataaaataa ctgattaata taattctaac 2753
 acaatgttga cattgtagt acacaaacac aaataaatat tttattttaa attctggaag 2813
 taatataaaa gggaaaatat atttataaga aagggataaa ggtaatatag cctttctgccc 2873
 cccacccac caaattttaca caacaaaatg acatgttcga atgtgaaagg tcataatagc 2933
 tttcccatca tgaatcagaa agatgtggac agettgatgt tttagacaac cactgaacta 2993
 gatgactgtt gtactgtagc tcagtcattt aaaaaatata taaatactac cttgtagtgt 3053
 cccatactgt gttttttaca tggtagattc ttatttaagt gctaactggg tattttcttt 3113
 ggctggttta ttgtactgtt atacagaatg taagttgtac agtgaaataa gttatttaaag 3173
 catgtgtaaa cattgttata tatcttttct cctaaatgga gaattttgaa taaaatatat 3233
 ttgaaatttt g 3244

<210> 5
 <211> 761
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> 693, 722
 <223> n = A,T,C or G

<400> 5
 caogaggctt tgatatttct tacaacgaat ttcatgtgta gaocccactaa acagaagcta 60
 taaaagttgc atgggtcaa at aagtctgaga aagtctgcag atgatataat tcacctgaag 120
 agtcacagta tgtagccaaa tggttaaagg tttgagatgc catcacagtaa atttaccaag 180
 catttttctaa atttatttga ccacagaatc cctatttttaa gcaacaactg ttacatccca 240
 tggattccag gtgactaaag aatacttatt tcttaggata tgttttattg ataataacaa 300
 ttaaaatttc agatatcttt cataagcaaa tcagtgggtct ttttacttca tgttttaatg 360
 ctaaaatatt ttcttttata gatagtcaga acattatgcc tttttctgac tccagcagag 420
 agaaaatgct ccagggttatg tgaagcagaa tcatcattta aatatgagtc agggctcttt 480
 gtacaaggcc tgctaaagga ttcaactgga agctttgtgc tgcttttccg gcaagtcattg 540
 tatgctccat atcccaccac acacatagat gtggatgtca atactgtgaa gcagatgcc 600
 cctgtctatg aacatatatta taatcagcgt agatacatga gatccgagct gacagccttc 660
 tggagagcca cttcagaaga agacatggct cangatacga tcatctacac tgacgaaagc 720
 tntactcctg atttgaatat ttttcaagat gtcttacaca g 761

<210> 6
 <211> 1901
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (125)...(793)

<400> 6
 acgtaaccta cgggtgtcccg ctaggaaaga gaggtgcgtc aaacagcgac aagttccgcc 60
 cactgaaaag atgaacgttg atatctccgg agcatttgga taatgtgaca gttggaatgc 120
 agtg atg tgg act ctt tgc cca ccg cca tct cca gct gtt gcc aag aca 169
 Met Ser Thr Leu Cys Pro Pro Pro Ser Pro Ala Val Ala Lys Thr
 1 5 10 15
 gag att gct tta agt ggc aaa tca cct tta tta gca gct act ttt gct 217
 Glu Ile Ala Leu Ser Gly Lys Ser Pro Leu Leu Ala Ala Thr Phe Ala
 20 25 30
 tac tgg gac aat att ctt ggt cct aga gta agg cac att tgg gct cca 265
 Tyr Trp Asp Asn Ile Leu Gly Pro Arg Val Arg His Ile Trp Ala Pro
 35 40 45
 aag aca gaa cag gta ctt ctc agt gat gga gaa ata act ttt ctt gcc 313
 Lys Thr Glu Gln Val Leu Leu Ser Asp Gly Glu Ile Thr Phe Leu Ala
 50 55 60
 aac cac act cta aat gga gaa atc ctt cga aat gca gag agt ggt gct 361
 Asn His Thr Leu Asn Gly Glu Ile Leu Arg Asn Ala Glu Ser Gly Ala
 65 70 75
 ata gat gta aag ttt ttt gtc ttg tct gaa aag gga gta att att gtt 409

```

aca gac gca aag ccc ccc gcc ccg ccc gaa aag gga gcg acc acc gcc 703
Ile Asp Val Lys Phe Phe Val Leu Ser Glu Lys Gly Val Ile Ile Val
80 85 90 95

tca tta atc ttt gat gga aac tgg aat ggg gat cgc agc aca tat gga 457
Ser Leu Ile Phe Asp Gly Asn Trp Asn Gly Asp Arg Ser Thr Tyr Gly
100 105 110

cta tca att ata ctt cca cag aca gaa ctt agt ttc tac ctc cca ctt 505
Leu Ser Ile Ile Leu Pro Gln Thr Glu Leu Ser Phe Tyr Leu Pro Leu
115 120 125

cat aga gtg tgt gtt gat aga tta aca cat ata atc cgg aaa gga aga 553
His Arg Val Cys Val Asp Arg Leu Thr His Ile Ile Arg Lys Gly Arg
130 135 140

ata tgg atg cat aag gaa aga caa gaa aat gtc cag aag att atc tta 601
Ile Trp Met His Lys Glu Arg Gln Glu Asn Val Gln Lys Ile Ile Leu
145 150 155

gaa ggc aca gag aga atg gaa gat cag ggt cag agt att att cca atg 649
Glu Gly Thr Glu Arg Met Glu Asp Gln Gly Gln Ser Ile Ile Pro Met
160 165 170 175

ctt act gga gaa ctg att cct gta atg gaa ctg ctt tca tct atg aaa 697
Leu Thr Gly Glu Val Ile Pro Val Met Glu Leu Leu Ser Ser Met Lys
180 185 190

tca cac agt gtt cct gaa gaa ata gat ata gct gat aca gta ctc aat 745
Ser His Ser Val Pro Glu Glu Ile Asp Ile Ala Asp Thr Val Leu Asn
195 200 205

gat gat gat att ggt gac agc tgt cat gaa ggc ttt ctt ctc aag taa 793
Asp Asp Asp Ile Gly Asp Ser Cys His Glu Gly Phe Leu Leu Lys
210 215 220

gaatTTTTtct tttcataaaa gctggatgaa gcagatacca tcttatgctc acctatgaca 853
agatttggaa gaaagaaaaat aacagactgt ctacttagat tgttctaggg acattacgta 913
tttgaactgt tgcctaaatt tgtgttattt ttcactcatt atatttctat atatatattgg 973
tgttattcca tttgctattt aaagaaaccg agtttccatc ccagacaaga aatcatggcc 1033
ccttgcttga ttctggtttc ttgttttact tctcattaaa gctaacagaa tcttttcata 1093
ttaagttgta ctgtagatga acttaagtta tttaggcgta gaacaaaatt attcatattt 1153
atactgatct ttttccatcc agcagtggag tttagtactt aagagtttgt gcccttaaac 1213

cagactccct ggattaatgc tgtgtaccgc tgggcaaggt gctgaattc tctatacacc 1273
tatttccctca tctgtaaaat ggcaataata gtaatagtag ctaatgtgta gggttgttat 1333
aagcattgag taagataaat aatataaagc acttagaaca gtgcctggaa cataaaaaaca 1393
cttaataata gctcatagct aacatttcct atttacattt cttctagaaa tagccagtat 1453
ttgttgagtg cctacatggt agttccttta ctagttgctt tacatgtatt atcttatatt 1513
ctgtttttaa gtttcttcac agttacagat tttcatgaaa ttttactttt aataaaagag 1573
aagtaaaagt ataaagtatt cacttttatg ttcacagtct tttccttttag gctcatgatg 1633
gagtatcaga ggcattgagt tgtttaacct aagagcctta atggccttgaa tcagaagcac 1693
tttagtcctg tatctgttca gtgtcagcct ttcatacatc attttaaatc ccatttgact 1753
ttaagtaagt cacttaactc ctctacatgt caatttcttc agctataaaa tgatgggtatt 1813
tcaataaata aatacattaa ttaaagtata ttatactgac taattgggct gttttaaggc 1873
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1901

```

<210> 7

<211> 562

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> 166

$\langle 223 \rangle$ n = A,T,C or G

<400> 7

agacgtaacc	tacggtgtcc	cgctaggaaa	gagagatatt	tcgggagcat	ttggataaatg	60
tgacagttgg	aatgcagtga	tgtcgactct	ttgccacccg	ccatctccag	ctgttgccaa	120
gacagagatt	gctttaagtg	gcaaatcacc	tttattagca	gctacntttt	gcttactggg	180
acaatattct	tggtcctaga	gtaaggcaca	tttgggctcc	aaagacagaa	caggtaacttc	240
tcagtgatgg	agaaataact	tttcttgcca	accacactct	aaatggagaa	atccttcgaa	300
atgcagcatt	tggtgctata	gatgtaaaatg	tttttgtctt	gtctgaaaag	ggagtgatta	360
ttgtttcatt	aactcttgat	ggaactctga	atggggatcg	cagcacatat	ggactatcaa	420
ttatacttcc	acagacagaa	cttagtttct	acctccaact	tcatagagtg	tgtgttgata	480
gattaacaca	tataatccgg	aaaggaagaa	tatggatgca	taaggaaaaga	caagaaaatg	540
tccagaagat	tatcttagaa	gg				562

 $\langle 210 \rangle$ 8

<211> 798

<212> DNA

<213> Homo sapiens

$\langle 220 \rangle$

<221> CDS

<222> (81) ... (590)

$\langle 400 \rangle$ 8

gggtctcttt ttgggggcgg ggtctagcaa gaggagatat ctccggagca tttggataat 60
gtgacagttg gaatgcagtg atg tcg act ctt tgc cca ccg cca tet cca gct 113

Met Ser Thr Leu Cys Pro Pro Pro Ser Pro Ala
1 5 10

ggt gcc aag aca gag att gct tta agt ggc aaa tca cct tta tta gca 161
Val Ala Lys Thr Glu Ile Ala Leu Ser Gly Lys Ser Pro Leu Leu Ala
15 20 25

gct act ttt gct tac tgg gac aat att ctt ggt cct aga gta agg cac 209
Ala Thr Phe Ala Tyr Trp Asp Asn Ile Leu Gly Pro Arg Val Arg His
30 35 40

att tgg gct cca aag aca gaa cag gta ctt ctc agt gat gga gaa ata 257
Ile Trp Ala Pro Lys Thr Glu Gln Val Leu Leu Ser Asp Gly Glu Ile
45 50 55

act ttt ctt gcc aac cac act cta aat gga gaa atc ctt cga aat gca 305
Thr Phe Leu Ala Asn His Thr Leu Asn Gly Glu Ile Leu Arg Asn Ala
60 65 70 75

gag agt ggt gct ata gat gta aag ttt ttt gtc ttg tct gaa aag gga 353
Glu Ser Gly Ala Ile Asp Val Lys Phe Phe Val Leu Ser Glu Lys Gly
80 85 90

gtg att att gtt tca tta atc ttt gat gga aac tgg aat ggg gat cgc 401
Val Ile Ile Val Ser Leu Ile Phe Asp Gly Asn Trp Asn Gly Asp Arg
95 100 105

agc aca tat gga cta tca att ata ctt cca cag aca gaa ctt agt ttc 449
 Ser Thr Tyr Gly Leu Ser Ile Ile Leu Pro Gln Thr Glu Leu Ser Phe
 110 115 120

tac ctc coa ctt cat aga gtg tgt gtt gat aga tta aca cat ata atc 497
 Tyr Leu Pro Leu His Arg Val Cys Val Asp Arg Leu Thr His Ile Ile
 125 130 135

cgg aaa gga aga ata tgg atg cat aag gaa aga caa gaa aat gtc cag 545

```

Arg Lys Gly Arg Ile Trp Met His Lys Glu Arg Gln Glu Asn Val Gln
140                               145                               150                               155

aag att atc tta gaa ggc aca gag aga atg gaa gat cag ggt cag          590
Lys Ile Ile Leu Glu Gly Thr Glu Arg Met Glu Asp Gln Gly Gln
                160                165                170

agtattatttc caatgcttac tggagaagtg attcctgtaa tgggaactgct ttcatttatg 650
aatcacaca gtgttcctga agaaatagat atagctgata cagtactcca tgatgatgat 710
atttggtgac agctgtcatg aaaggctttc ttctcaagta ggaatttttt cttttcataa 770
aagctgggat gaagccagat tcccatct          798

```

<210> 9

<211> 169

<212> DNA

<213> Homo sapiens

<400> 9

```

aaacagcgac aagttccgcc cacgtaaaag atgatgcttg gtgtgtcagc cgtccctgct 60
gcccggttgc ttctcttttg ggggcggggg ctagcaagag cagatatctc cggagcattt 120
ggataatgtg acagttggaa tgcggtgatg tcgactcttt gccacccgc          169

```

<210> 10

<211> 176

<212> DNA

<213> Homo sapiens

<400> 10

```

aaaacgtcat cgcacataga aaacagacag acgtaaccta cgggtgtccg ctaggaaaga 60
gagggtgcgtc aaacagcgac aagttccgcc cacgtaaaag atgacgcttg atatctccg 120
agcatttggg taatgtgaca gttggaatgc agtgatgtcg actctttgac caccgc          176

```

<210> 11

<211> 38001

<212> DNA

<213> Mus musculus

<400> 11

```

tgtctctagg taaaattttg aaggaaaaaa aaaacactaa gaaggatatat tccctcaaag 60
ttccagtcctt attctgaagt gtaatgttat gttagtttga ctccacagaca ggtttttaaag 120
aagggtcttac ttcaagagga caccaaaaca ataccttcta ttccatagtgg gctctggaat 180
cacagaaaac tgacccaatc aattacattg atagctcttg cttactacag acaagcaaatt 240
tatcttaagt gtgcatgcat gcgctgtgat gtgtgttagt acctaacacc cacctgggaa 300
cttttcagct ttccagtgtg ggatatagta taaacgtcta ttccctcgtgt tgtggattag 360
ctgactgggc tcactcagct gccttcctta cctgcaaaact caccactttt gactacagca 420
tcgcactctt aacctagacc ttccaaacat ggctctatgc tatttctgtg tgtctggatg 480
tatttttaac tctcagatgt ataacttcatt tatgagatat acatctgaag accacgggtac 540
aaaacactgt aagaacttga tagaatgaca actgctaggt aaaaaaaaaa aaaaaaaaaa 600
aaaaaaaaaa aaaaaaaaaa atacaatacc tgggtgagagt tctatttttta ccgaagggtg 660
tattgatagg tattctgtta ttaatgcctt tcttttccct ataaatgatg aaaagttgct 720
ggaaaataat aaacactact catctgtagt gaaaagccac aatacagtta caaaccaatc 780
aatcaatcaa taaactagac gtcattgtgt tcttttccca aaggttaaaa aacaaagtgc 840
actgtgctat ttggcaaaaa tgacgttttag aagaaaacac ggtgactacg cacagagggg 900
gggggaatca ttgtgcttgt tgcggagtga acacgtacag tgtgcacgca gacttacggc 960
atttaaccgt gtcattaggga ccaaaggaaa tccactcact cactaaatat ttgttgagca 1020
cccactacct gccaaactcc aaacaaaaca aagcaaaact acttacaacc acaaaactacg 1080
cttcgtaacc tagatagata acgcagggtga cactatctat ctaggttgag ctcagctctg 1140
cccattgctt tctgagcggc ctcttggaag aagaagctaca aagcccatga cagcctccgc 1200
ctggccagct gccactggca tctcaaggct ggcaaaagcaa agtgaaagcg ccaaccggca 1260
acttacggag tcccacgagg gaaccgcggc gcgtcaagca gagacgagtt ccgccacgt 1320

```

```

gaaagatggc gttttagtagt acagccatcc caattgccct ttccttctag gtggaaagt 1380
gtgtctagac agtccaggga ggggtgtgca gggaggtgcg ttttggttgc ctccagctgc 1440
aacttaactc cacaacggtg accaaggaca aaagaaggaa acaagactgc agagatccgc 1500
accggggagc cctgcagatt ctgggtctgc tgtggactgg gggcgggact gcgactgggc 1560
ggcgctgggg gcggtgtccg ggogggggcg tcccgggcg gggcccgag cgggctgcgc 1620
ttgcggtcce tgccgcggcg gtgaaggcgc agcagcgcg agtgggtgag tgagacgcgc 1680
ggcgggaggg gggtctctgc cacgggtcggc tcggggggcg gccggctccg ggtaccagcg 1740
gggttttttt ctcttcagag gtgaactcct cctgttcccc cgggcgaaag agcccttggc 1800
cttgccaggag ttgcgggggc cgcggcggtg cggaggggat ggggatgggc ctcatctttg 1860
ctgtccgccc gcgtcccccg atcccgaccc ggagcgtctc ccgggcccct gagggaaacc 1920
tcggggagta cggcgagcgc gggcccccac gccacaagcc tgggcccag gggcctggcc 1980
cggcgacagc tgggtgggtc tggaaccag tcagggtctc cgagggtccc cgcccgggag 2040
gagaaagcgc gggtgggatg gagtaaggac ggacagaaca acacgcaggc aggatattgc 2100
agaagtttgc aaggagtgcg gatgccact tacatgggct gctactctta ccaggttgtt 2160
ccccagttct gtgggacgtg acctggttgc ctccagctc cgcggttgta cagacttatt 2220
aaaggaaagt accattgtga cttgggcctc acttgactga tggtaatcag ttgcagagag 2280
agaagtgcac tgattaagtc tgtccacaca gggctctgtc gcccaggagt gcatttgcct 2340
gggagggatt ggttgcgctt tctgggtgtg ggaactattg gctctttag agttttgtcc 2400
cggcagtagg ataaatttct tgttacactg ttcccgttcg tcaccagttg agaaaaacgc 2460
gtacacagtc tgtctcagta gtacttttac tttatattaa gggcccaaaa gggactggaa 2520
aatactttta gatagaatcg ttagtccact tggaaaactt aaaatatgag agagagaggg 2580
ggggggggaga gagagagaga gagagagaga gaaaggaagg aagaaggagg aagaggagga 2640
ggaaagagat tgagattatg ttaataatat ggaatcagaa tatttgaaat atagtaagcg 2700
tcccctcagt taaagaggac attccaggag gccccagta tagcctgaaa tctcaggaaa 2760
cgctacata caccatcgt gtggatata gtgttttccc ttcattacat ttcatacaca 2820
gatgttaaag tttagaaagt aggcacaata agagattaca aataactgat aataaagtcg 2880
agccattgca gctgctctgt aaaagtccgt tgaatgtgat cgttttgtgt ttcaaagtaa 2940
cttactgtac ttcacccctg ttaagcaaaa caagattcac ctgaacgcag gcaccttggg 3000
accttggcag acaccagatc tgataaccaa gaggatggag aagtagtggc agacagtgtg 3060
gagagcatga atatgctaga caaaaggggt aatcataacc taggagcaga aagcaggtat 3120

```

```

ttcatcatcc tccacagtaa aaacctatgt cacgtaaaaa acctacaagt agtttttctt 3180
ttactctttt tgaatgaaag cttgctacag gcactgaaag ttaaaaaaat ctgtggatca 3240
ggaggaaacg gggttttctg tctgagtcac tgctgactag cacctcagtg accttggca 3300
ctgtgggaaa cccagagtc agttggaaac ttcgaaacta aaggtgacgg tgttcttatt 3360
tcatagaaca caaaaaataa gaggggttac agcctgcgct gcagactgga cattcaacaa 3420
gcatttaaat ttctgggaga caaatgtaaa tataacttta aaagttggta aaatactctg 3480
tttgctatg ttggccatcc aatgtttgct tttagaaat gactgaatgg ataaaacgtc 3540
tatcttttga gcttgcctta gacccccatg ttgagtgaat actgtccaaq tgtaggtta 3600
gocggcctga gaaacttggg tctaggcaag atggcacagt ccgggtgtca tgagtatgca 3660
tgtgagtttt ggctgaaatt gaacatttgt agagaatgac aaaggctggg ctggcaagta 3720
gtccactgtc ttacagtggt tcttggttag ttctgtttg gcagagaggg ctggttgatg 3780
gctgtcctgc cctcttccc acaagtggaa gccttatggt ataattcttg atccagtag 3840
cagtaggcaa atgaacttcc tcaaagcagc ctggaaagct gatttttttt tctttcttcc 3900
tctttttttt ttttttttca caaggttaaa gaaaaaacia agggcttcaa atgtgccagt 3960
ctgctaacag tgttaacatg tttattaaca taataaaact ttattagttt ttggaagtat 4020
tggttaagcc ctctgaccc ctgaactcgg tttatagagt gatgagtggt agcctcactc 4080
tggtttggac tctggtctct ctcagaagac tctgtggcta atgttaacct totgaagtag 4140
ccagaaaaaa tataagcaaa agtctgtgag gttgaaatga attttttggc cacatttcta 4200
tatgggttcc caccaatgct aacttcaggt gttagtaata tcagactcac agcttccctg 4260
attacacttc gctataagac tttatttttt aggtcatagg aatttccct tttcatgat 4320
tctaaatca tgaataaaca tagtctaaaa atacgggtatt cctgaaataa acaatttcta 4380
agttttaagc tgcgtgctat tctgaacagt ctgatgccct ctgtagctt ttactgtgtc 4440
ctaccccggt catggttgat tctttgttcc aaacatctgt ctgttgtatc cacactggat 4500
tgcaccacct gcgtgctagt cagtcactca gacattttag ttataaggta gcttatattt 4560
actccttatt ttatttaata atggcctcat agcaaggcgg taatgatact ggtaatttgg 4620
qtttgcttaa gaggaacat gaaqtaqtct taaatgaaaa ggtgaaaatt cccactatag 4680
tttgaggggg gaggtatata tggtagtact acgattcacg gtaagactaa atcttctgtg 4740
aaattatgaa ggagaaaaag ttacactggg ctggtcttgc tgttgatta attttatagt 4800
tataaccact gtacatgata aataacccta aaacaatgaa ttgttaggtg gatggcataa 4860
tctgaaaacc atgttctgag cagttgatgg cagcaggtcg tgctggaagt gttaggcata 4920
tttatagatt tcagcccaag ttctgaagag ctggagaga tggctcagtg gttgaagagt 4980
cttgcctattg cagaggacct aggttctctc acaggcacca ggcaagcgtg ggacacactg 5040
agatacatat agacaaaaca taaaattaaa taaattgtgc ataataatac tagtaataa 5100
tgagtaaaat aaggataaat acacatcata attaaataaa taaattgtaa agttccctag 5160
aagtgagggt caccaagcca ttcacaagat ggctgcgctg atgcagggat atatgtgaac 5220
tagaaaaagg tcaaaacttaa cagagaagtt ccaaggcatg ctactgcagg cttggctagc 5280
atgcttgacc tgcagaaatg ctgacggcca ctgggaggtt ttcacaaatg aggaattaga 5340
agaacttttt ttactaatct ccagaaaaaa aaaaggggaag aagaaactga agcagcctgt 5400

```

```

gatgtggacc agaaacgcag tgacagtaac atgtgtgaca ttgcaaaggg atgaaaggac 5460
agagctgtgg aatacagacc tcaggtggag ctcagcatag agtcattcgg ggattatgcc 5520
tgctgcagca acaaaaggat gagctcaaaa gagacaccga ctctgaatg cagtggggtg 5580
ttgttttgtt ttgtttcaaa tgaattgggc agaaaacttt ccagctgtgg aagcttctga 5640
accgtccott gctgtgaca tctaagcgtc cgctgtgtcc cagctcagtg atctagggtc 5700
ttccaaacag atggtccggg gctgagcact ttgaatctca atctgagtt tctaccacgc 5760
ctttggccat ttaattccca gataaaagac acatacaacc ttcatattta taataaacct 5820
tagtcagcac aagagctgag caaatatctg tctctatgc tattatatct attaccagc 5880
caataacccc attctataat ttgctgtgct tcatctgggc tgctcttaac ttcagtcagc 5940
cagcccacgt ggccattatt ttaagatttt tttaccccat agtgtcttct cactttactt 6000
tacatttttc tctctctct catggttctc ctctgacccc aagcctagga accctaaacc 6060
ccacccatgt ctctctgccc catctatttg ctgtaggcat ctctattcac caatcaggat 6120
aacttgaggg caagggttaag tagtctctcg ggtctaggtg ctgtctctgg gagcaaccag 6180
tatttagcat agcaaaagac cagacctcca caatgatcac tctgaccatc ggggcagaag 6240
gcacctacta gccctgtgcc ctcacctcac tttgttgaat cacatcttat cctgtagtgt 6300
gtatcactgc ctgttatcac aggaaaaagt gaggccatc aaataagatg tttcagaaaag 6360
agaccatgtt catataatta tcaattctgt aagcttttaa tgggtatatt ttgttattaa 6420
tctctttgtt cctattttgc aaattatacc ttacagtaaa tatatatgca tccaatgggg 6480
tctttgaatt cctcccgagg gaggtaggag actctttgag gatgggctgc atttaaagct 6540
aaacaacgca acatgaacct tagtccttat agatagccta gagatgagac taataaaaag 6600

```

```

aaatgggata taatgcttta agtttcccaa tcaagcttaa agcttttctt ataaatcttt 6660
aagattatgc tctggggctc aatactgctt caagaagggc tttctttttg tatttagaat 6720
tattcacctt tttaaacaaa aggagaaaat ggaatagaaa tatgtttgca acataatttt 6780
atgactatgt gtttatttct cgtgttctgt gggcctgcag tttgctgctg ttaatgagga 6840
caacagtggc accaatacac tttccactca gattacatct tctgttccct ttctgaaagc 6900
tgccctctcc actgggcccc aaagagtcag tatcttaaac aagctgtaca acttagataa 6960
ccatggtctc ttcagactag ttaattgaca tatattaaaa agtaaatagt accaaagtga 7020
atctctgaaa ttaaaaaatg acatttaaaa actctaggta aactatttct taagcttaag 7080
tgttttgcca agttctgtaa tcataatatg atagaaacgc tcaactcagca ttctaaatat 7140
agaagtact ccttcgcatg acactcta atcttgataag gtggagaaaag agagagagag 7200
agggggagag acagaaaaat tgggtggtca aggaccattt gagggaaatta gttatgttct 7260
tccgtctctc tcgtgactta ggggttgaat acagtcaatt agctcgggtg atggctgtcc 7320
tgttgaaaag tctgcccagc agagcaaaata gactttttta tttacatgga catccgtttg 7380
tgactaatct aatgttcaact cccaaagtaa tcacacagac agagaggtag ctctcttcag 7440
tactcttacc ttacatgaat cctaccattt tgttattttt tttccacttt aaatctttga 7500
ttatgtgttt ttaattagaa aatttgcata caaatttcca tacagtatgt agaattgact 7560
gtgtttgaat ggttgaagat ccacatgtgt aaccctagct ctggactggc totgagcttg 7620
tttgtcttcc tcttttgtgt tctgagtaac tgaaactctt tcatttttagc agcttagtat 7680
gogcccttca cattgtgtgt ctgctgtctg cactaacatt actcctttgc ttatgttccc 7740
cttcctgatt cagtgtcaat ttaagcagta gtactggacc tcagtacott agcccgagct 7800
cactgagggt acagggtctg ggtctgtctg ctgtcttttg agcttaacct tttttaatgt 7860
tttatgggat ttctgtgccc aggtttgggg gttttgtttt gttttgtttt ttgttttttg 7920
tttttttaa tttctagga acacctagaa aacacaaact aggaaactta aaagagcagc 7980
gtcttggtcc ctgcgttcta gaaagtccaa gctaatgcc agtgtcatgg ttgtcaggaa 8040
catgagctc tgaaggcttc ttgggaaacc tttcttgtct caacacctct ggtggcaagc 8100
agtagtccat ggtactctct ctgtccacgg tcagcatccc agtccctgcc ctttatcttt 8160
gtgcagccga ccagctttgc tttagtctgt ctcttctctc ggtctccttc cccgtctctc 8220
ttaagcacag cagtcattgg attagagccc atccttccct cggatggccc atttgacct 8280
attttacgta tttgtaacta aggtcccatt tacttacaca gggccctccc ctctctgttt 8340
tgttctttag ctgaaatggt ttggagacca aatatccaat cattacaatt gtgcacaagc 8400
tatgttcaat tggaggtaat aaaggctcat tctttgttcc tattgggatg tgacattttt 8460
ctaagtcaact tggggtttga tagatatctt taaatggctg aacctgatca ctgttctttt 8520
gtatgtccct gtttagctat tgcaagcgtt cggataatgt gagacctgga atgcagtgag 8580
acctgggatg cagggatgtc gactatctgc ccccaccat ctctgtctgt tgccaagaca 8640
gagattgctt taagtgttga atcacccttg ttggcggcta cctttgttta ctgggataat 8700
attcttggtc ctagagtaag gcatatttgg gctccaaaga cagaccaagt gcttctcagt 8760
gatggagaaa taacttttct tgccaaccac actctaaatg gagaattct tcgaaatgca 8820
gagagtgggg ctatagatgt aaaaattttt gtcttatctg aaaaagggtt aattattgtt 8880
tcattaatct tcgacggaaa ctggaatgga cttatggact atcaattata atcaattata 8940
ctgccgcaga cagagctgag cttctacctc ccacttcaca gagtgtgtgt tgacaggcta 9000
acacacatta ttcgaaaagg aagaatatgg atgcataagg taaggggctt ttgagcttga 9060
tcattgttag ctggccaatg aaagtttttt tctggtacag ttacacttaa gttttgaaa 9120
ttatatgctg ctaaacaccg acagctgtta tgttgtgtct cctgggcaca gaaagccctg 9180
ctctcatgoc tggggtcttc acagctctaa tggaagtaa gatcttataa acatttgttc 9240
tgagtttgtt ctggaagctg tgactctacc ttcttgtttt cctttccctg tgtgactttg 9300
tctttgtctt acaacagtgc aaaagtataa atattctcag attttgataa gctgtcagcc 9360
acacagcctt agtaactaag ctgctgtccc acgctcccag ttctgtataa cgaggatgga 9420
ccaattagat tctaaggagt tatctcttcc aatttgcaaa tttagctaaa ggaaatattg 9480
ttttctctct atatttacat tgcttttcat tttcagcata tctaaagaac aaacctaat 9540
ctccttctca cttctgaatt taataaatc ctaaaaatcc attaaaaac qactaattct 9600

```

ataaggcctc	taacctacaa	agggaagtag	cattttgaaa	agaatagttt	tctctattat	9660
acctattcat	gcagacttcc	tcccttattt	ctgacatact	taacaaaaat	catttagatt	9720
caaacagttt	agctgcaggt	gatattacag	acaagtaatc	ccagtgtctc	atctagtctg	9780
aggcaaaagg	atttgagctc	agtgccagcc	tgttctatct	acctgggtgag	ttccagtcct	9840
ataaataaac	aaactaaaaac	aacogttcc	ctgttccoa	gatgcgagtc	gatcttgttt	9900
gattttaaata	gtgtgttaatt	attttctttt	gaagctgcag	gtgttatgtg	ggctgtttta	9960
gactaaattc	tctcttttact	gtggagtaaa	gggtgctgtg	attgtatttc	atgttctctg	10020
cgagagcttg	aacttggttg	gctaatoctc	tgtctccatc	ctgtctcccc	acctgcgtaa	10080

aaagtatttt	cctgtgagct	gtacatgata	gagcatatct	acattgaaaa	atgaacgagc	10140
atcaaaatgg	atttgttaaa	gtaaaatttc	tttttcttag	gaaagacaag	aaaatgtcca	10200
gaaaattgtc	ttggaaggca	cagagaggat	ggaagatcag	gtacagtga	tatcacatgc	10260
tgcctgtggc	aggtcctctt	tgttatgttc	ggataaaagt	tgttgggtac	ttctggtaag	10320
gacctgagga	tacattcatt	tgacggaagg	agcctgaaaa	tgagtattct	tgtaagctg	10380
tatagaatga	actgaataaa	aatttctgca	gcctaagttt	gaatttttaa	aaaatttaat	10440
tacatctaca	aattagttat	tggtccacct	ttttcaatca	gcaagaatat	gtttgaggtc	10500
atttattttg	agttaaaattg	catgcagttt	atttatttta	ttgaaaataq	gttttttaaa	10560
ctatattttc	tgttatgggt	ttccctcct	ctgaatcctc	ctagaacctc	cacctaccca	10620
aatctatata	tgttctttct	ctctctcatt	aggatacaat	caggcatgta	aaataatagt	10680
agtagtagta	gtaataataa	tgtaaaataa	gttaaagtaa	aaacaaacca	gagtaggaca	10740
acataaatag	aagtagaaaa	gagccaaaata	agaaattcaa	gaaacacata	tagacacaga	10800
cacaatat	gcatacacag	aaattgcata	aaaccgcaag	actggaaaacc	ataatatgta	10860
tgtaatgttg	agtgggaagg	cctgacagca	cagttagtaa	agcactttca	aaaacaccac	10920
tgactttgtg	ttgtgtttgc	tgtctgtgtg	tgctgaaggc	tggtccttaga	gagtggtgtg	10980
tataccagag	aagacttaca	taaacactta	gcttttctct	tgtgaacctga	tagcaattgg	11040
aaatagtgtc	tggtctaggc	attccggctt	attgccactt	cccctcagca	ctgaggcccc	11100
atctgaatcg	gatccgtgca	acccttgtgc	atatgcagtt	ttaaaagtta	tcccttctgc	11160
aactatgtc	acaggagttg	cgttcttaag	ggagtgagea	cacccctgag	gcattggctc	11220
aggggtgcag	agccagccat	aggoacagtt	ttttttaaaa	ggtttatggt	gtagttttga	11280
aaactcaatt	tatgtgtatt	tgtggcagat	tgtttgaatg	ttgaaatttg	ccagtaacat	11340
cttttatctt	cttcccttta	gcttggcatg	cccccacccc	tcatttctcc	ttgtcaaaact	11400
ccagtaatta	aacatggcta	tgtggccttt	tctctcattt	tccttagcat	ggctaaggag	11460
aatgggactt	aaaaaataat	aatcatcatt	ttaagtatgt	ctgagggttt	gaggatatag	11520
tggtagaata	cttgcctagc	ttccatagct	tgatcctaca	tttgatccct	ggcaaaacac	11580
acacacacac	acatatacac	acacataaaa	tgacttttat	aaagttaqtg	tqctgtqctg	11640
tgatgaacag	tgccatagga	aatattcttg	gaaaagacct	gaaactaaat	gctctaaaag	11700
gtctaatctt	tacttgcctg	ctgatcgtta	agcagagtct	ccaagtataa	agtcactttc	11760
accaacctct	gcactggatt	tctggagtaa	ttagggagag	tcatttcaat	ataagaaaat	11820
ttagtaccaa	ataaaatttt	cattcagtg	aattttgttt	ttgaaagtaa	gagccactg	11880
tggtgttttg	aatatgcttg	gcccaggag	tgtcctgtaa	gatttttgtt	gttgttgaa	11940
tccattgaga	cttatgttga	caataaaatgc	ctgagagtcc	atgtctaaaa	tgtgttacct	12000
gtctgaaccc	aacggagata	aaacttacca	tttctgaaaa	ggatgaggtg	ttttatttac	12060
atagctgatg	taattgtgctt	gcaacagctc	tattatgaat	cttaataacta	cttcagtata	12120
tcacagcact	tcaggaaatt	taacatacat	tgttttaatt	catgtcttaa	ttgtatttgt	12180
aaacagacat	ttcagcagtt	actctaaaaa	gtagaaataa	tgagtggttg	cttctgggtc	12240
ttaggtgaa	atattgaaat	gataaaattt	tctgggctgg	agagatggct	cagaggttaa	12300
gagcactgac	tgtctctcca	gagatcctga	gttcaattcc	cagcaaccac	atggttagctc	12360
acaaccatct	gtaatgggga	tctgatgcc	tcttctggtg	tgtctgaaga	caactacagt	12420
gaactcatac	aaataaaaaat	aaataaatct	ttttttaaaa	atctatatct	gcattaggcat	12480
ttctagatta	ggataaaatt	tccaaaggaa	ataagcactc	ccatgataag	ggcattggaa	12540
atgaagcccc	agccccacc	cccgtctgc	acgtgtgttg	aggtatgagt	ctagggcctc	12600
cttatacatg	ccaggcagct	gttctgtcac	caagtggaa	ataatcctca	accttaatt	12660
tgaggttcta	actttaaaat	agatgtgagg	ggtttaaata	atcatttcat	gaaacttaa	12720
tgagcaagtt	tattactgag	gtgagtataa	gtaattgata	attttaata	tatttagctg	12780
agattgatag	acacttggca	atgtcagcat	cttatttagg	tgatcataaa	ctgatgggag	12840
aaatggtaaa	tgtaggggg	gtcgcacacc	gtcgcacacc	gcagttatgc	tgcaaacaa	12900
atgccgggaa	atagaaaattc	aaggtcttgt	tttgcgggtg	cagactcttc	tgtctcactg	12960
attctatgtg	gtaacttcag	tatgcatttg	gatagattat	gtcccatttt	gaatgtggaa	13020
gctggctgtt	gagaggagac	ttcctggtga	attccttttt	ctaagcatta	ccatctgtct	13080
tagtcagggt	ttctattcct	gcacaaacat	tatgaccaag	aagcacttgg	ggaggaaaag	13140
gtttattoag	cttaacattc	cacactgctg	ttcatcacca	aggaagtgcag	gactggaact	13200
taagcaggtc	aggaagcagg	agctgatgca	gaggccaagg	agggatgttc	tttactggct	13260
tgttccctg	gcttgcctag	cctgctgtct	tatagaaccc	aagactacta	gcttagggat	13320
ggcaccaccc	acaatgggcc	ctccccctt	gatcactaat	tgagaaaatg	ccccacagct	13380
ggatctcatg	gaggcatttc	ctcaactgaa	actcctttct	ctgtgataac	tccagcctgt	13440
gtcaagttga	cacacaaaac	cagccagtac	aacatctttt	cacatttaat	ttttctcact	13500
ttaaacgtgg	cctttaacaa	gcgcttataa	aaatgcttaa	gcttaaatgt	tatttaagct	13560

taatatactt	aatatacagc	actgtagctt	aaatgttgca	tgtgagagta	tatgataagc	13620
catgctcacc	aaggaaaaa	agcttaaaag	gcataaaaa	cctgacagcg	gtttctgagt	13680

```

gggaggctcg gggactgtgc tgagcaatcc caaccaaggg tgttttactc tetgectcca 13740
tttgaatagt ttttcctgca caacctaccc accctgtgat ttctgttact cgattatgtt 13800
tgatctaggg tcagagtatc attcccatgc ttactgggga agtcatttct gtaatggagc 13860
tgcttgcatc tatgaaatcc cacagtgttc ctgaagacat tgatgtaagt gtcagtgtatc 13920
ttttatgggt tcccttgagt ggtgagtggg tggtatgttg gtcatgtgc gtgtgtgtgc 13980
ttgcatactg ggaattgaac ccaagtcctc aggaagagca gccgtgtctc ttaagcactg 14040
agccatctct tcagaacctc ttccaccagt ttctttgacc atttgttgag aatattccag 14100
tcacacattt tccgtgagta aatctctcta atgctgattt gtcattaagc tcagtctcct 14160
aattctgata gctaagaagg gtaaattatt aaaaagtgc ctttactctt cctggccaat 14220
tcccctttgt tcttctgaaa agtgcataga cagcatcact ttatagatca ccttgatgct 14280
cgtgagaggg ctggctcgtg ctggctctag acttcggcac acttattaag agttctccca 14340
acactgtaaa cagactaatt tttatattgt gcattttaga tagctgatac agtgctcaat 14400
gatgatgaca ttggtgacag ctgtcacgaa ggctttcttc tcaagtaaga attttacttc 14460
ttttctgaa tgcataagta agcagattaa aaatcttaat gctcaccat gacaagattt 14520
acagggaaaa gatgtagaa aacctacttc ctccaattat ttagggtcaa catggcacat 14580
ttgagcttac accgtgtgtt ctccaccata atacttgaca ttactcttcc 14640
cacagtctaa aaaggcagag ttccgtagt acccaggga gttctggtct gtgtttgggt 14700
ctgggtttct ctttcaattc tcaactaagta taaccttag gaatctatca agttgagttg 14760
cattttaaat tctgtgaat tcttcaggtc tagaaatgga aatcattcat attttagact 14820
gacatttttc atcttcttgt gtaatttaac atttaagaac ttgagctcta atatcagact 14880
gtctagttta caactgggaa aacttggtga agctacccaa agctgaacct ccattttctt 14940
acctgtgaaa tgtgaacagt gataacagct agtttcttgg gtcctgttag gcaccaaag 15000
acaggataat ataaagcacc taggacagtg gagccaatga gccaggagcc agtggtgccct 15060
attatatctg ctctaagaaa gacagttaagt qgaataqcca atactgactg tcttaqtcaq 15120
gctttctatt cctgaacaaa aaacatcatg accaagaagc aagctgggga ggaaagggtt 15180
tattcagctt acacttccac gttgctgttc atccacaaag gaagtcagga ctggaactca 15240
gatcaggaaa caggagcaga tgcagaggcc atggaggaat gttacttact agcttgcttt 15300
cttatagacc ccaagactac cagccagag atgggtccac ccacaaggga cctgcccccc 15360
ttgatcacta attgagaaaa tgccccacag ctggatctca tggaggcatt tccccaaactg 15420
aaactccttt ctctatgata actccagcct gtttcaagtt gacacaaaaac cagccagtac 15480
gtgaccgag cagctgtgtt ttctctgca gggctgtgtt ctctgtttgt cctcatctc 15540
ctgttttagt ctcttttaca gttacagact gtcacagta acgagagaga agtgaatagg 15600
attttgttaa agtgtttact totatgtcac attcccttcc tataataagc tcacagtga 15660
ataccaggtg accgtgttta acggcatcta ttacctact ggggtatctt tttccttaa 15720
atggatttaa ttttatgtgt gtttgaatac ctgcataatgt gtatgtacac catatttatg 15780
tatgctgtgt actgaaaaa gggaaaagag ggctttggct ttcttgaaac tagatggttg 15840
tgagcttcca tgtgggttct ggattgtctc gcaagcagac gcaggcacac tttagcagtg 15900
agecgtctct gtcccgagtt gtcttaagac tccctaaaaa atgcagggtt 15960
ttaccogaat aaaagatgac atcatgcaga tggctttggg gttcatcaag ctctgtgtg 16020
ttgtcctaac ctgtctgggc tttgtcgttg tgaagctgta actcgtcaa tgttttcctt 16080
acctacagtg ccacagctc acacctgcag acctgtggtt gttccgttgt agttggcagc 16140
agtgcagaga aagtaataa ggttaattcgt tctacagttg aacatgatct gacttttate 16200
atcatagca tctcatcat tatcatctaa acagtaggct gcaattgaaa taaccccata 16260
gtataaggaa gcaatgtaat tttaccaa atctctgaca cctctagca gaactgactc 16320
taatagaatg agtaagaatt caattaccaa attaattttg atactctttt ttatttttgt 16380
tattactttt ttattttatt ttaattaggt attttcttca ttacatttc caatgctatc 16440
ccaaaagttt ccataacct cccaccact cccactcccc tatccacca cteccctttg 16500
gccttgggtt tcaactgtac tgagacatat aaaaattgca agaccaatgg gactctctt 16560
ccaatgatgg ccaactagac catcttctga tacatatgca gctagagaca cgagctccag 16620
ggggtactgg ttagttcata ttgttgttcc acctaaaggg ttgcagacc ctttagctcc 16680
ttaggtactt tctctagctc ctccattggg gccctgtga tccatocaat agctgactgt 16740
gagcatccac ttctctgttt gctaggcccc agcatagcct cacaagagac agctatatca 16800
gggtcctttt agcaaaatct tgcagtgtg tgcaatggtg tcagcgtttg gaagctgatt 16860
atgagatgga tcccaggat ggcagtatct agatcgtcca tctttctgtc tcagttccaa 16920
actttgtctc tgtaactcct tccatgggtg ttttgttccc aattctaaga agggacaaag 16980
tgtcacact ttggttttca ttctcttga atttcatgtg ttttgc aaat tgtatcttat 17040

```

```

atcttgggta tctaagttt ctgggcta atccacttat cagtgagtag atattgtgtg 17100
agttcctttg tgattgggtt acctcactca ggatgatgcc ctccaagtcc atccatttgc 17160
ctaggaattt cataaattca ttctttttaa tagctgagta gtaactcatt gtataaatgt 17220
accactttt ctgtattccat tctctgttg aaggacatct ggttctttc cagcttctgg 17280
ctattataaa taaggctgct atgaacatag tggagcatgt gacctcttca cgggttgga 17340
catcttctgg atatatgcc aggagaggtt ttgtgggac ctccggtagt actatgtcca 17400
atttctgag gaacggccag actgatttcc agagtgttg tacaagcttg caattccacg 17460
aacaatggag gagtattcct atttctccac atcctcgcca gcatctgctg tcacctgaat 17520
llllcatcgl agccallctg aclygllylga gylglaalcl caggylllyll llgalllyca 17580
tttacctgat gcttaaggat gctgagttt ttttccaggt gcttctctgc cattcgggat 17640
tctcaggtg agaattcttg gtttagctct gagccccatt tttaatggg ttatttgatt 17700
ttctggagtc cacttcttg agttctttat atatatgga tattagctcc ctatctgatt 17760
taggataggt aaagatcctt tccaaatctg ttgggtgacct ttttgtctta ttgatgggtg 17820
cttttgcctt acagagctt tccaatttta tccagtagca ttctctcatt ctccctctta 17880

```

```

ccccgccc acagaagccc tgcacccccc tgaagcacc cctgcagac cccgccccca 17880
cagcacaaagc cattgatggtt ctatttcagga attttttccc tgagccaata tcttcgaggc 17940
tgttccccac tctctctctc ataagcttca ctgtctctgg ttttatgtgg agttccttga 18000
tccacatgga tttgacatta gtacaaggaa ataggaatgg attaatgtgc attcttctac 18060
atgatatccg ccagttgtgc tagcaccatt tgttgaaaat gcttttttcc actggatggt 18120
tttagctccc ttgtcaaaaga tcaagtgaac ataggtgtgt gggttcattt ctgggtcttc 18180
aattctatcc cattgggtcta ctgtctgtta tataccaata ccatgoagtt tttatcacaa 18240
ttgcccgtga gtacagcttt aggtcaggca tgggtattcc accagaggat cttttatcct 18300
tgagaagagt ttttgccttc ctagggtttt tgttattcca gatgaatttg catattgccc 18360
tttctaattc gttgaagaat tgagtttgaa ttttgatggg gattgcattg aatctgtaga 18420
ttgcttttgg caagatagcc atttttacaa tgttgatcct gccaatccat gagcatggga 18480
gatctttcca tctcttgaga tcttctttta tttctttctt cagagacttt aagttcttgt 18540
catcacgata tttcacttcc tttagagtcac gccaaagtat tttatattat ttgtgactat 18600
tgagaagggg gttgttttcc taatttcttt ctacagcctgt ttatcctttg tatagagaaa 18660
ggccattact tgtttgagtt aattttatat ccagctactt cattgaagct gtttatcaga 18720
tttagagggt ctctgggtga attcttaggg tcacttatat atactacat atcatctgca 18780
aaaagtgata ttttgacttc ttcctttcca atttgtatcc cettgatctc ctcttgttat 18840
cgaattgctc tggctaagac ttcaagtaca ttgttgaaata gggggaaga aagtggacag 18900
ccttgtctag tccctgattt tagtgggggt gcttccagct tctcaccatt tactttgatg 18960
ttggctactg gtttgcctga gattgctttt atcatgttta ggtatgggac ttgaattcct 19020
gatctttcca agacttttat catgaatggg tgttggattt tgacaaatgc tttctcctca 19080
tctaacgaga tgatcatgtg gtttttgtct ttgagtttat ataatggatt acattgatgg 19140
atcttcgtat attgaacct ctctgcaccc ctggaataaa acctacttgg tcaggatgga 19200
tgattgtttt tagtgatctt tggattcagt tagtgagaat tttactgagt atttttgcac 19260
caatattcat aagggaaatt ggtctgaagt tctctatctt tgttggttct ttctgtggtt 19320
taggtatcag agtaattgtg gttcataga atgagttggg tagagtacct totgcttctg 19380
ttttgtggaa tagtttgtga agaactggaa ttgatcttc ttgaaggtc tgatagaact 19440
ctgcactaaa cccacttggt cctgggattt tttttggttg ggagactatt aatgactgct 19500
tctatttctt taggggatat aggactgttt agatcattaa cctgatcttg atttaacttt 19560
ggtagctggg atctgtctag aaacttgtcc atttcatcca ggttctccag ttttgttgag 19620
tatagccttt tgtagaagga tctgatgggt ttttgattt cttcaggatc tgttgttatg 19680
tctccctttt cattctgat ttgtttaatt agaatacttt cctgttggcc tctagtgaat 19740
ctggctaagg gtttatctat ctgtttgatt ttctctaaga accagctcct tgattggttg 19800
attctttgaa tagttctctt tgtttccact tggttgattt caccctgag tttgattgtt 19860
tctgtccgtc tactctctt gggtgaaatt gcttctttt gttctagagc ttttaggtgt 19920
gttgtcaage tgctaattgt tgctctctct agtttctttt tggaggcact cagagctatg 19980
agttttccc ttagaatgc tctcattgtg tcccataagt ttgggtatgt agtggcttca 20040
ttttcattaa actccaaaaa gtccttaatt tctttcttca ttcttctctt gaccaaggta 20100
tcattgagaa gactgttgtt cagtttccac gtgaatgttg gctttctatt atttattttg 20160
ttattgaaga tccagcttag tccatggtga tctgatagga tgcatgggac aatttcaata 20220
tttttgata tgttagggct tgttttctg accaattatg tggteaattt tggagaagggt 20280
accatgaggt gctgagaaga aggtatatcc ttttgittta ggataaaatg ttctgtagat 20340
atctgtcaga tccatttgtt tcattacttc tgttagtttc actgtgtccc tgtttagttt 20400
ctgtttccac gatctgtcca ttggtgaaag tggccatctt tatagtcaat gaagacatac 20460
aaatacatat tcatatcaac tggacaacac ctaatttctt tttcaatggt ttacatggaa 20520

```

```

ataagttagg ggttgttatt tgcattacaa agttactcat ccttttctt cttttctttt 20580
ttttttttt ttttttttg agaacaagcc tgtgtactta tatgaacttt aatttgccaa 20640
attcataatt ctatttcaat catttatgac agaatgctaa aactctcatt atatttttag 20700
taggcattta gagctgttat gtgtaacccc aaaaagtagc tttccacttg agatgctgaa 20760
ggccttgggt tccgtgggct gtcacatcatg ttggctgtat gaaaagagaa aggcctccatt 20820
gtttgggcat cacttaaata ttttttccac ttcatcttc ttttaggtta agtagcttgt 20880
ccttgatcat ttcatttttg agagacaact tcccactact ctagttagaa aqtgctgtct 20940
tgacgtgtc tctggctgtg gtcagagctc agcagagctg cacagctggg tacctttctc 21000
lgtagagctc laggcacac cltcllactg ggcacalll claalccac callcactlg 21060
ttccccatga aagttagtag ggttttttct gtggaagatt ttgggcagtc ctgttgccac 21120
tttgcatcag acaatagttc cctcattgaa acacgcagtt tattctccag agcggctctgc 21180
ccactccaaa ggcagtaggt gctgggtaga gatagccaa gtatcacact aggcctatgac 21240
tgctcactca gatcactcgg atgaagctt catggccaaa tacagttgag aaagaacaaa 21300
tattcttcac tttagagagca acaagagtta ttcaagtgt acaagttctg agattccatg 21360
cagttgattt accagctact tctaaactt aactggccac aaatccctt tgtaagcagt 21420
atgttgtttt gaccatgccc ctgtcaaaag atactcetta cttgggaact gttttaatga 21480
tggcaacaaa aatttctatt taaatttatt tcaaaagcaa gcaagatct ttttacttca 21540
cattccaatg ttgactttt tctctagat agtaagaacg ctgtgccttt ttctgacacc 21600
agcagagagg aaatgctcca ggtgtgtgta agcagaatcg tcttttaagt acgaatcggg 21660
actctttgtg caaggcttgc taaaggtaaa cttgcccagc atttatcatg tgtgacgcaa 21720
caagtagaga tggagggtac aaataatcac tgagaggett tggaaagtat attgttagca 21780
tttaatgtct catagtttta gttgtctggg tactggtttg ttttcatcat tctgagcatg 21840
aagtgtatgt cttagggtat tatagttcgt atcatgtatg aaacaocatg gggtaatat 21900
tatatttcac ttggttccct ctagctatgt gtctggcccc agtgccttcc ttgtaaatgc 21960
atgcttgaat cagactgagc tgatatgata atgttgatgc tcttttgtct tactgagtg 22020

```

ctatqaatat	gcaccatact	tactcattgt	aaqaaattaa	aatgtctctt	aaqqatgtaa	22080
acatagcaaa	atgaagcaaa	acaaaagcga	tgctgtttta	ggtaccctaa	ctgacctgt	22140
gtattcaagg	agcattcccta	cttctgtgat	gcaaaagctg	tctacactgg	gcagatctac	22200
aaccagcatt	aaaccaaata	gggaatcact	gaaatcacgt	tatcaaagat	gagaaacaa	22260
ataataatgt	ctactttcac	ggcttttatt	caggtctagt	gctataagtt	tttgccaaaa	22320
caaaaatgaa	aacatagact	ctgggctgag	gctttccctt	agcagaaaa	tgcttacttg	22380
ttgtgtccgg	ccagcagatc	acagcctggg	ttctagcctg	gaaagggcatt	ttggaaacct	22440
ggaagagaag	aggggctagg	taacgagaga	aagaacggag	ccaagtcaaa	agcaactctg	22500
atcaaagctc	aattttacta	tatcagcacg	cagttataaa	ggaggggaa	ggggggccaa	22560
tagcaaggcg	gcaggttcca	gcagtggcg	tgccagaccg	attgagccgg	caagctcctt	22620
ccagggtgaa	acagtggagc	cctaaggctg	ggggagggga	ggctacactt	agcatgcctg	22680
atgccctaga	tgccacctaa	atgacaaatc	cagtcacagta	caggatgtag	agcaccctcc	22740
cccaaaaaat	tatttttttt	gtataccaga	aatgaaattg	ctgagaaaaa	aaaatgaaga	22800
ccataattat	actcccagta	gctacaaact	aaacagcccc	atagatgaag	tgagtgatgt	22860
ctgctgtgac	aattatgaaa	tgaaagaagt	aaagatgaac	aaatgaagg	aagacatcca	22920
gtactcagga	ctgaaagact	gctgctaaaa	tgccctatcca	acccagagct	ctctgcagac	22980
tctggacaga	tccgctctag	atgtgaagat	ggctcttttt	tttttttttt	ttttttggtt	23040
tttcgagaca	gggtttctct	gtgtagccct	ggctgtccag	gaactcactc	tgtagaccag	23100
gctggcctcg	aactcagaaa	tccgcctgcc	tctgcctccc	gagtgtctgg	attaaaggcg	23160
tgccaccacca	tacttggtct	tttgtgaaga	tgctcttaac	agaactagaa	agaagtacc	23220
cttggtttgc	tgccctctct	atgcagtatc	cccaaaggct	cgcatgcact	gaacatttca	23280
tcttacctgg	tgccactggt	gggaagtgat	ggaaatgcga	ggaattgtag	cctcgttgag	23340
atgtttctca	ttaaggcaact	gggggcatac	ctatggagca	tacagttaga	acctggtttg	23400
caacctctcc	cctctccatc	caggtctctc	cctgtgcacc	tgcccttggt	gttctgccac	23460
tccatgaacc	caaagtaaa	tggaactatg	ccttagactg	taacagttag	tcagaagaaa	23520
catttctctc	ttaaagctga	gttttctggg	tgctttgtca	tgtaaatgga	gtctgattag	23580
tacagaccct	gagttaggcg	ggcaatctta	tgccagaaaca	tcaagctgg	tagcatagac	23640
atacctaatt	tcacaataga	cactgatgga	ctcagctctg	agtacttaca	gtaagaatat	23700
acagcagaga	tacggagctc	tcttacagtg	gtgctctggg	agaactggcc	gtcctgtgaa	23760
gaaaagccag	agtggctcat	tctcaccaga	cacaaactga	gctcataaga	cgcttgaacc	23820
tgagatctct	ctcagcagcc	actagaagaa	cacttaggag	aaaccattca	acacgtcagt	23880
ctggggaaaa	gggtggtttt	gggtttggtt	ttggtttttt	agtatatctc	ccaaatcaaa	23940
aacaacaaaa	cccaaacttg	acagatgaca	tcacactgca	aagcttttgc	acaaccaaga	24000

aagcaacctg	cagagtgcag	taataaccca	cagaaggaga	ggagatactt	gtgggcagtt	24060
catcacacag	gtcaatataa	gcaagtactg	atagtgtggc	catctocaaa	gaagatatga	24120
aaataactgg	tatatatgaa	gtagtactta	gcattgtctg	gtatatggta	aattcaaaac	24180
catgatgaga	tattgcccac	cttagatgga	tattatcaaa	acaacatcaa	aaagtgcaca	24240
atgctttcaa	ggatatggg	aaagtgtact	tgccaggaatt	taaatatta	atttgccatt	24300
caagaggata	ggatggcagt	ttaaattaaa	aaactagaag	tggtagagca	gtcgcctaga	24360
acatacaagg	ttcagcacta	taataaatga	gcaattagac	atttgaaqca	acaatctcac	24420
cactaggcaa	gtcctaaaa	aaatggactc	gcttcttctt	cttcgggaaa	acaccaaattg	24480
gcagalgacy	ccgglgcagc	gggagggccc	agaggacclg	ggggclcagg	allaggagyc	24540
cgcggcgct	tccacggagg	atcgggcagc	ggctcttagg	gccgtggctg	tgcccgaggc	24600
cgtggccgtg	gtcagggccg	cggggctcgt	ggaggtaaa	ctgaagacaa	ggagtggatc	24660
cccgtcacca	agctgggccc	cctgggttaag	gacatgaaga	tcaagtcctt	ggaggagatc	24720
tacctgttct	cctctgncat	taaggagctc	gagatcattg	atttcttctt	gggtgcgtcc	24780
ctaaaggatg	aggtttctgaa	aatcatgcca	gtgcagaagc	agactcgggc	tgcccgagcg	24840
accaggttca	aggctttcgt	cgtctattgg	gactacaaatg	gtccggttgg	tcttggtgtt	24900
aagtgtccca	aggaggttgc	tactgccatc	cgaggggcca	tcactcttgg	caagctttcc	24960
atcgtccctg	tgccgagagg	ctaactgggg	aacaagattg	gcaagcccca	cactgttcca	25020
tgcaagggtg	caggccgctg	tggtctctgt	ctgggtcgtc	tcactccctg	cccagagagc	25080
actggcattg	tctctgctcc	tgaagctcct	gatgatggcc	ggtatagatg	actgtctcac	25140
ttcagccaga	ggctgcactg	ccacctggg	caactttgt	aaggccacct	ttgatgccat	25200
ctccaagact	tacagctacc	tgacccccga	cctctggaaa	gagactgtct	tcaccaagtc	25260
tccttatcag	gaattcacgg	atcatcttgt	gaaaacccac	accagagtct	ctgttcagag	25320
gaccaggtct	ccagctgtgg	ctaccacata	aggggtttta	tatgagaaaa	ataaaagaat	25380
taagtctgct	gaaaaaaaaa	aaaaaagaaa	gaaagaaaga	aaagaaatgg	actcggtagt	25440
tggaatgaagc	ccaggcacct	tcactgttgt	tgccagcagc	gtcaccatgc	aggatcagtc	25500
taaacgcccc	tcacacaaatg	aatggtacat	agccacagtg	aagtgtttga	ccacaaaaag	25560
gaaaglcagl	lglgalaagl	gaaacaagcc	aggcacagaa	agataaaalg	lgcaglllal	25620
cattatgtgt	aaaggctaaa	acgttttatct	catacaagta	gaaggtaaat	acggagacta	25680
ccagaactta	taaagagttc	taggaaaaag	ctatagagag	gctcaggggt	gaataactaa	25740
aattatacct	aaaataacta	aaaggatagc	ttacaatat	ctgtagcact	gtagaataat	25800
tgtgacagtt	tggtgtattt	ttctgggttt	tgtagtggtg	agagaaagta	tgtggacaga	25860
gggtgatata	aagtgtctga	ctctgcactg	cattattttta	ggcagggctc	ctctctaacc	25920
attgaatgga	ctggctaggg	agtggtgccc	taacatctac	ctgtccgtac	atctcccaat	25980
actaggttat	aagtacactg	ggtttttaagt	acaggctata	ggtatagata	taggctacag	26040
gtatagatat	aggctgctgc	aactgattac	atgggtgctg	ggaacctaac	atagggttgg	26100
tctcatgttt	tacacagaaa	tcagtactgt	gctactgag	tcatttcccc	agttctagta	26160
tttttttttt	aaataacctag	taatttqaat	tttqaatttt	cctaacaaaa	qaaaatqata	26220

actatctgag	atgctagtta	tgataccctg	agtgaatcac	actttgtgtg	catgtactga	26280
aattcattgt	accctgaaaa	tacaaaaatt	gctctgtgtt	gattggotag	atgcatgtgt	26340
attagtcagc	aatctctaga	gtaataaaac	ttagatatat	gggatgtatt	agacttttgg	26400
ccttacaggc	caagatccag	ctaattccatc	agtggcaggc	tgtgaacagt	aagtotaaga	26460
atccaatagt	tgttcagtc	acaaggccgg	gtggctcagc	tgcctttctg	atacagtgga	26520
atcccaaaga	aataggcgcc	aaagctagtg	aggaatggtc	ttgctagcaa	agcgaaggtg	26580
aaggtataca	ggcagaagac	agagaccttc	tttttccgtg	tccttatata	ggctctagc	26640
agacaagtg	gccagacta	gatgtgaatt	aaatgttttg	ggtttggttt	ggtttgattt	26700
ggtttggttt	ggtttggttt	ggtttggttt	ggtttggttt	ggtttgctt	ttcgagacag	26760
ggtttctctg	tatagccctg	gctgtcctgg	gtttagacc	aggctggcct	caaaactcaga	26820
aatctcttgc	ctctgcttcc	caagtgtctg	gattaaaggc	gtgcacacca	ctacgcccg	26880
ctcaatagca	ttaaatggca	tgtcttttcc	tatctcaaat	gatctggatt	aaaagagtg	26940
cttctactct	caaaggtctg	gattagaagt	ggatctttct	acttcagatt	aagttaaact	27000
ctctcacagg	tgtgcccctc	acttttggat	ttttgtctct	agatggagtc	aacatgacaa	27060
ccaaaagtaa	ctattacaag	tcacccaat	atcaacttga	tacacaatca	tatctcttta	27120
tgtcataatt	aattttccaaa	tgaaaacaat	aacctatgca	taaaaacacc	taaacatgaa	27180
taactattcc	acatacaatc	agaaatgcat	tcattatata	tttaaccaag	tcctaattat	27240
gctaacgtg	atataactat	tottcataca	acagcaaaca	tgataaattt	acaatatggtg	27300
gcaatgtctt	attctttttaa	tattctcaac	ttaaatatga	taaccattga	tgttatctta	27360
atgtgtgtta	tatcatatga	taaagaaatt	gatgaagaa	agcacaagt	tcgtgataaa	27420
tgtctttctta	agaaaatagg	acagaaaactc	tgtcaattat	aatcatcttt	tcgtcaacta	27480

gtcatgtggc	cttagtattt	ataactacct	tctctgtcta	aaccttttgg	tatttttctc	27540
accottggca	agaacctcag	cagggtcttg	ctcttttct	ggaggagtga	cccatacctt	27600
catttcttac	atgtatgtgc	cottttgtcat	cctgctctga	ccagggttgt	gtaacattga	27660
ctttaatcac	aggacatcgt	agcaccaca	catgccccga	aggatctctt	gccttataga	27720
cataacccctt	cttacctcca	tagtggggag	cgagtcocag	tctctcttgg	tgtctgcata	27780
cagtcacgcc	tcctaacact	gttattctct	tcttagccgg	ttgacttaag	ggcatcagaa	27840
ggccaaagt	gccagagga	aatctgagct	tccagttcaa	tgaatgta	gttgttctag	27900
gcaagcagaa	ctgaaggct	caggaatagg	aagcaaacac	ttcccatgga	tcactacagg	27960
glgagaglg	glagaaall	lclclllcl	accalllgac	lcllggacl	alggalclcg	28020
gtatcaaa	aaatgtctca	tattattgtac	actgattcag	agctgcctt	ctggaaaacc	28080
ctgccccagc	cottcatact	gttgccatca	aattgtcacc	tgtgtcttcc	tgttaccac	28140
ttttgtctcg	gttaggggtta	ctattgtctgt	gaggaaacac	catgagcacc	aaagcaactt	28200
ggggagaaat	gggtttattc	agcttattcgc	tctacatcac	agctcatcat	caaaggaagt	28260
cagaacagga	gctcaagtc	ggcaggaatc	tggaggccgt	ggaggaaagc	tgtcagctgg	28320
ctgcctccct	aggtctgtct	agactgctta	tagaactcag	gaccaccagc	tccagggttg	28380
ccccaccccg	caatggattg	ggcctccct	caggaatcac	aattgcccc	cagacttacc	28440
tacagcctag	gcattttgga	ggctttgagt	ctgcctctc	tctgatgatt	ctagcttttg	28500
tcaagttgaa	gcaaaagtag	acaggcctta	aactcacac	aaccacctg	ctcctaattt	28560
ctgagtgcta	atatttatac	atatttaaaa	ttaaatataa	catataaagg	gcaatagaaa	28620
ggaactagatt	catgtaatgg	atacaagtta	tggaaagtgt	gtgtgtgtgt	gtctgtctgt	28680
ctgtgtgtgt	gtttctgatt	taattctgtc	atgattttt	tcttgtaggt	ggtaggtgag	28740
tgcattggaat	acatttgata	ctgaaagggt	aaattgaatg	tggagcctca	cagcttctgt	28800
tcacatgcc	tatgataacc	gtagaaattc	atggattagt	atagacgttg	agctctggtta	28860
attttgggtg	gtgatattta	tatatattatg	tatatatata	tgtgtgtgta	tgtatgtatg	28920
tatgtatata	tatattatgta	tatgtgtata	tatatatata	tatatatata	tatgcaagat	28980
ttcttataat	taagtttaca	aaattaaaaa	ctatctttaa	aattgaattc	ttgcaataaa	29040
aaalllagcl	lllgglgall	ggallcllaa	lalggllgal	glllacclag	aaagllaaaa	29100
gcoctgagtt	cagctctccac	tttcacccccc	aaaatgaaaa	tcagcttttg	ggtttcagat	29160
catagagctca	gaattaaaga	aaacacattt	ctaactttgc	ttttacaaat	cttaattttta	29220
coaatttctc	ttaaagtcac	aatgagatc	acagctcttc	ctagcacccc	ttgttcaatt	29280
agataaatgtg	atttctgaaa	gagctccctc	tacacagggc	acagggcagg	tgcaaaactg	29340
tgattgggtg	aaatacctgc	gagctctcca	agcaaaagcca	ggcctatttg	ctttagctgc	29400
cacatcggt	tcttagaccc	gacatccct	cccactgtta	tctcctctaa	ttccttccaa	29460
ccccacaaca	ctaggtagga	gagaaagaag	ggttagtggtg	gaagtttgca	cacatctttt	29520
tgaactattt	cctactgatt	aggggtgtta	ggtccttgag	acaagtcocag	tcttcattgt	29580
caggatatct	ccaacttctt	cttctcatct	ctttgctcac	aaagtttctc	acaagttgat	29640
aaactacaac	aacaggaacc	agcagtagca	aggacatcag	agttgtatag	ctttccagaa	29700
aatactttga	tatacagtaa	ttatcctagc	ctttaagagt	gaagatttg	gcagcctctg	29760
tgtttctacac	tcagcataat	acctgtgata	ctgcaggtat	ttgtctgatg	gtaagtggtc	29820
gccagctcac	ctagaaagag	gtaaatactt	ttctattaac	atacatattc	atttagatat	29880
aggaagaaga	taaaacaatg	gagaaaggca	gtcataaatt	tacagaccag	caagtaaaccg	29940
cattaacttg	gcattaggtct	ttgtagtctt	tttctgcagt	gcgtatttcc	tgcagtgcc	30000
acacccataca	gttgattgtt	acgtggcaatg	ttctgaccca	ctttttatgg	tatactgtgt	30060
actgtcactg	tcaacacaaa	tggttagtgc	tggtattttta	tacagtatca	gcttgaaggt	30120
lclllclgaa	caagccclgl	accagallca	caggaaalaly	calclcllal	callaclala	30180
ttcttttaac	aattgcttct	ctcagttggc	atgtggtcag	tgagttctct	cttcttcttg	30240
acaggatgca	acaggcagtt	ttgtcttacc	cttccggcaa	gttatgtatg	ccccgtacc	30300
caccacgcac	attgattgtg	atgtcaaac	tgtcaagcag	atgcacccgt	gtcatgaaca	30360
tattttctact	gaagagagat	caatgaggtg	agagatgaca	gcttatattga	gggcaactta	30420

tatttataat	caacgcagat	acatgaggcc	agagctgaca	gccttctgga	gggcacaccc	30420
agaagaggac	atggcgccag	acacccatcat	ctacacagat	gagagcttca	ctcctgattt	30480
gtatgtgacg	cttggcctta	gggtgtcattg	ttaaacaaca	taaaacttct	catttatgag	30540
taaaaacagt	gcaagttgta	ttttaaagaa	aagaaatatg	acaagccat	actcaggcac	30600
tttttcttta	ttttcttaac	tttaagggtt	tttttttttt	aagatttatt	tattattata	30660
totaagtaca	ctgtagctgt	cttcagacac	accagaagag	ggcgtcagat	ctcattacaa	30720
atggtttgta	gccaccatgt	ggttgctggg	atttgaactc	aggacctttg	gaagagcagt	30780
cagtgccttt	acctgctgag	ccatcttgcc	acccccaact	ttacattttt	tatactatta	30840
tttttagaca	gtctcactgg	gcctaataac	ttacataggt	ggcctggaac	tcactatata	30900
gatcaggcta	gccttcaact	cccagatata	cacctgcctc	tgccaacca	atacttggtt	30960

taaaggcgtg	tgccctccta	cctagcctaa	atcttcattt	ctttaaatac	tgttttgcta	31020
agataggtaa	agatttctct	ttaaaaataa	atacttagca	aatatatacc	gatctccta	31080
ttacttaatg	aagggccagc	ttaatagtta	tcagtcagtt	atcagtgcca	gccccactg	31140
ctgggaattt	agtgtataac	gttcattgta	tggttagactg	aagtaattct	aagtattttt	31200
ttcttgggtg	tgactatcaa	acacagaaaa	gtatttgaaa	tttataaaga	gaacagggtt	31260
tttctttgca	ttttatattt	ttctatttat	ttcttaccag	aagatgagag	cagcaaagta	31320
aaaggcagta	agtgcctgat	ggtttggagg	aacttgggat	tttaattata	aaacttcaag	31380
aaagcatttc	aatggtgttc	tagagtctaa	aaaagaatag	tgagacccta	ttcctgttct	31440
clccgalcaa	ccaagagccl	gaalggllgc	laglcccllag	lalacacclga	aaagacgcta	31500
agtgtggtca	tcocggttgg	aggcctttag	gaagcagtgga	ccctggacca	atgggtgtca	31560
cogtgtgtct	gaagaagaaa	gcagagctga	aacaagaggo	gcagtgtagg	gacaccagca	31620
gccacagtaa	actgctgccc	agaggtecc	gtgtggggct	gcagaattaa	aagaacccat	31680
tctacacagc	tctgctgtgc	tctgtttagt	ctgagaaagg	ttgagaggaa	ttgtttcaga	31740
agaggtaatg	ttcaaatgta	actcttatgt	cactagtcca	catactggca	atcttgga	31800
acatagaaa	ttctcactg	agtctgcgtg	cctgcgtctt	cctcgtgact	aatatacttg	31860
aagtcctgtt	tattttttta	gttgattgtt	tagaatctct	tctcaggaaa	tgaggtaaac	31920
ttgaatggat	ttgcaccatg	ttagtgtttt	tgttttgaat	atgtttgttt	ggaagatttg	31980
aagaaaaagc	aattgttcag	ctattctggc	atgacaaaat	catgtcatga	attttagaat	32040
tttattttcca	gttttaagta	aatgttttga	atataaaatt	gtcagaaata	ttttcagcca	32100
caagattata	tcttctatta	ttgtgggctc	atgatagtat	cagtgtgggt	tataataat	32160
tcacttttga	gttggggagg	tttgagggtt	cagattcagg	gactcacaca	ctgggcaatt	32220
actgtaccac	tatgcagttg	cttattagta	ccacagagta	attcccagtt	aagttacttt	32280
taatttttaac	ctttttaaga	taaaagcagt	ctgatgatac	attaaagtgc	gacatttcct	32340
tgaagatagt	ctttcctttt	ccagcttttg	tgatccagat	ctcattcagt	aaagcagaaa	32400
ttgggaaata	gtggacttaa	gttctaaggg	acccacaaac	cccgtagctg	tgctgtccgt	32460
tttcagccag	taaccatgaa	gtgctggcgt	cccttcacgc	gcccccttct	ccattttggtg	32520
caclcalccc	lcaaggclga	gagggclgcl	gcclclclgl	clalllcccl	cllccccatg	32580
gttctctggc	agtgatgttg	tgatctctac	catctgagtc	ttgctttgca	tttatcttac	32640
tgtgaaaaat	gttatatttt	ccctctgaca	tgaatataat	agcctaggga	aagacagaag	32700
taaaacactg	aaagggaatg	ggggctgaga	aaaaaacagt	cattagcttc	tgtctggcca	32760
gcctgctgaa	gtgggtcacc	tcagttggcc	attttgtctg	aacgttacat	gccagccaac	32820
cttagctgag	gtatgaataa	gttatgctgc	tggtctatac	ttacagatgg	taagtctctt	32880
gacctgagge	aaacgtgtaa	gggtgacggt	ctaaacacac	tgatggacag	gcacatgccc	32940
tgccctggata	gcctcaaaac	acaaacagtg	tacaaatgta	cccttgcggt	aaagtggatc	33000
tatgtgcgtt	tgtgtttatt	ttctgtgcat	taagtatgta	tatgtatgtg	tgtttatatt	33060
gtgcacattg	agtatatgca	tggtgtttta	cactgaatac	tgaacccacg	gcctcctgca	33120
aactaagtat	gctattccaaa	tgacacacac	tgtcttttta	cacatctggt	tataaaaact	33180
caactttttt	actagagcaa	gaagttgttg	aatgtaactc	tgtaaaacog	tttaatatct	33240
gaaccttttt	cttcttagga	atattttcca	agatgtctta	cacagagaca	ctctagtga	33300
agccttcctg	gatcaggtaa	atatgatgac	acccattgcc	agacaaaaga	acatcatata	33360
ttttctttta	aaatatgtcc	cacagtgcct	acagaatata	taaaagcac	caaagaatta	33420
aagtgcctga	ggcctttcta	taagtctgaa	acggattcct	ctttgaatta	ttaatgggaa	33480
atagcctgta	tattaacctg	ttaaagcagca	ttctccatcc	tagtggctgc	ttcagggtcca	33540
accctctgcc	tttagaattt	ttgtgggttg	tgaagacagg	ggtgtgcttt	catttgtgtt	33600
aallgaallg	aaaalallcl	laaaaccllag	gllgcllclg	cllaaalgyl	agcalccclla	33660
ttgtctctgt	ttttaaaggt	atctgatgag	taaacatctg	gagatgggtac	tggtattctat	33720
ggaacttggt	tctatacgtat	agcagagctt	tgtcataata	gcctgctggg	aatcaggcca	33780
agatcctgtg	ccatagacat	agagttgaga	tgaggagaac	ctcgtgttca	ctgggacttg	33840
tggtgtctggg	tctgtgtgag	gtgaggacag	cctgtaatcc	caagtctctg	aagctgaaaa	33900
gtccctcctc	ctactccaca	caacctgaag	tcattgactt	agttatttcc	ataataaaat	33960
aaggagatat	tttaaggtag	aatacaagat	ctaagtgcac	ttaactaggg	aatctgaaaa	34020
ggggacagtg	ggtttccaga	catttgccgc	taccagagtc	ttgccctttg	gaaatcggaa	34080
gaaatggctg	taactgggtg	tgtgtgtcag	atcctgtcaa	caatgtcgcg	gaagctgcac	34140
tgtcttgtgt	ccctgcagg	cttccatttg	aagcctggcc	tgtctctcag	gagtaacttc	34200
cttgacacag	tcctcctcat	tcttcacaga	aaagccttga	cactaatcaa	gtacatcgag	34260
gatgatacgt	gagtcctgct	cctctagagg	aaagccttta	tgcatcgaca	gttgcctgtt	34320
gttccctttg	aacattgtct	gtattataat	gggggggttt	ttgtctcttt	tgttttgttt	34380
ataggcagaa	ggggaaaaag	ccctttaagt	ctcttcggaa	cctgaagata	gatcttgatt	34440

tacagcagaa	ggggcattct	aacataataa	tgagctatag	tgagaaattt	agggcaggcc	34500
------------	------------	------------	------------	------------	------------	-------

```

taacagcaga gggcgacccc aacataataa tggctcctagc tggagaaact aagccagggc 34500
tacactcttt catcttttggg agacctttct acactagtgt acaagaacgt gatgttctaa 34560
tgaccttttg accgtgtggg ttgtctgtgc tgtctcttca cagtcacacc tgctgtttaca 34620
gtgtctcagc agtgtgtggg cacatccttc ctcccaggtc ctgctgcagg acagggtaca 34680
ctacacttgt cagtagaagt ctgtacctga tgtcagggtgc atcgtttacag tgaatgactc 34740
ttcctagaat agatgtactc ttttagggcc ttatgtttac aattatccta agtactattg 34800
ctgtctttta aagatatqaa tgatggaata tacacttgac cataactgct gattggtttt 34860
ttgttttgtt ttgtttgttt tottggaaac ttatgattcc tggttttacat gtaccacact 34920
gaaacccclcg ltagcllllac agalaaaglg lgagllgac lclgcccccl clglgllclg 34980
tggtatgtcc gattactttc gccacagcta aacattagag catttaaagt ttgcagttcc 35040
tcagaaagga acttagtctg actacagatt agttcttgag agaagacact gatagggcag 35100
agctgtaggg gaaatcagtt gttagccctt cctttataga cgtagtcctt cagattcggg 35160
ctgtacagaa atgcogaggg gtcatgcatg ggcctgagt atcgtgacct gtgacaagtt 35220
ttttgttggg ttattgtagt tctgtcaaag aaagtggcat ttgtttttat aattgttgcc 35280
aacttttaag gtttaatttc attatttttg agccgaatta aaatgocac ctctgtgcc 35340
tttcccaatc ttgaaaaata taatttcttg gcagagggtc agatttcagg gccacgtcac 35400
tttcacttga ccacccttg cacggctgcc gtgtgcctgg cttagattag aagtccttgt 35460
taagtatgtc agagtacatt cgtcgataag atctttgaag agcagggaag cgtcttgcc 35520
ctttcctttg gttctgcct gtactctggt gtttccctg tcacctgcat cataggaaca 35580
gcagagaaat ctgaccaggt gctatttttc taggtgctac tatggcaaac tcaagtggtc 35640
tgtttctgtt cctgtaacgt tcgactatct cgctagctgt gaagtactga ttagtgagg 35700
tctgtgcaac agcagtgtag gagtatacac aaacacaaat atgtgtttct atttaaaact 35760
gtggacttag cataaaaagg gagaatata ttatttttta caaaagggat aaaaatgggc 35820
ccgcttctc acccaaccga tttagcgaga aaaagcttct tattctgaaa ggtcacgggtg 35880
gctttggcat tacaatcag aacaacacac actgacctg atggtctgtg aactaactgc 35940
aaggcactcc gtcatggtaa gcgagtaggt ccacctcct agtgtgccgc tcattgcttt 36000
acacagcaga alcllalllg aglgclall gllglclll clglclllac glglglall 36060
agaaaatgta agctgtacag tgaataagtt attgaagcat gtgtaaacac tgttatatat 36120
cttttctct agatgggaa ttttgaataa aatacctttg aaattctgtg tatgttttag 36180
ttcattattt agggaaaacg ctgctgtgaa agggggcgtg atcagcttcc tattctgcga 36240
cagtcgtgtt gaacggaacc cattggtttt catcttctgt cccccccct tggtttttct 36300
agacagggtt tctctgtata gccctggctg tcttgacct cactctgtag accaggctgg 36360
cctcgaaact agaaatctac ctgctctgc ctcccaagt ctgggaggca gttgcccac 36420
caactagtct tcttttttca aagaagatat ttaaagctaa cgaataatgc tagactctta 36480
catcttaaaa aaaaaagaag agaaaaagaa agaaaaagta atcacactgc ccagtgtgta 36540
gtgcatgctt ctacttccgg tcttggggag atgggggcag gatgagacgc tccagaccgg 36600
cttccaatac agagtccaag acccaactg ctacgtgagg ctacaccagc ctgcctttaa 36660
aaacataaag ctaaagcttt cttcttaact tccagtattg caccttgatt ccccttcaa 36720
atttcacata caaaataatt cttaaattct ttttgaaaa atgttctact gaggccagag 36780
agacagttcg cttggtaaag gtgcctgttg ccaaactgta taactctgag taaatcatag 36840
ccccacatgg gggaggaaga aacccccgca gcttgccctc tgatgccatg tatgactaa 36900
aacacgcacg tgtgtgcgca cacatttttt aagttcctat tacattgata gtaatatata 36960
ttaaactgat ttattctccc caagtcattg atacgggtgt ccaacgtaaa atccagcggc 37020
tgaacaaagc acttttaggc gctttaagtt ggaaagcaag aaacggagat tgacactgtc 37080
acccaagag aaaacclcll claglagcga galcggclgl ggaglgaaag lgclcagag 37140
ctgggaacgc acacagctca gtagtgata gcatcccca gcctcaactc ctaacactgg 37200
gaaagcgtag ggctctcaga tgaggaaaca aaaccataca aagctgctgc aagctaaaca 37260
gaaaaatagt ggcattacac taactgttgt ggaattgtac agaccgattc tctcccaat 37320
ctgccagtg tgggcggtt gagagaatga agagagctac tggcctcagg taacagtgt 37380
tcccacagga ctgtctcagg ctgccaccac cataaatagc attttagac tgacagagct 37440
aaggcttgac acacagccaa aagctactca cattccatt catccccagc tgttetgtca 37500
tcgctaagca cagagcattc agcacagctc ttccctgtgg tgggtactca gcaactgtga 37560
gttgaaagga ttgaaaaaac tcaagactat gttctcaaac atttttttaa gctcttttta 37620
aaaccacott agaatgaaag cttttgactt cttattaaca tgcactaact tcatatacac 37680
atttagtgtt attgtacagg cacgaagcat actctggtea gaacctgtct cctttgggtc 37740
accctcccca ccgttttcag cttctattcc accctccata cgtctoaaga tccacatgtg 37800
agagggaaca ctcagagcct tgtctttctg tatctgggat atctcaacta acatgatatt 37860
ctccagttct gttccatcca tttcattgca aagagcaaga tttcaactca cagccaaata 37920

```

```

acacatttgt ccatgtatat ccgtattttt ccttattcat ctgttgaatg gcacaagact 37980
gatatcatgg gtaatatcta t                                     38001

```

<210> 12

<211> 3435

<212> DNA

<213> Rattus norvegicus

<220>

<221> CDS

<222> (370) ... (1815)

<400> 12

```

cgttttagt gtcagccatc ccaattgcct gttccttctc tgtgggagtg gtgtctagac 60
agtcacaggca gggatatgcta ggcaggtgcg ttttggttgc ctacagatcg aacttgactc 120
cataacggtg accaaagaca aaagaaggaa accagattaa aaagaaccgg acacagaccc 180
ctgcagaatc tggagcggcc gtggttgggg gcggggctac gacggggcgg actcgggggc 240
gtgggagggc ggggcccggg cggggcccgg agccggctgc ggttgcggtc cctgcgccgg 300
cggtgaaggc gcagcggcgg cgagtggcta ttgcaagcgt ttggataatg tgagacctgg 360
gatgcaggg atg tgc act atc tgc ccc cca cca tct cct gct gtt gcc aag 411
      Met Ser Thr Ile Cys Pro Pro Pro Ser Pro Ala Val Ala Lys
          1                5                10

aca gag att gct tta agt ggt gaa tca ccc ttg ttg gcg gct acc ttt 459
Thr Glu Ile Ala Leu Ser Gly Glu Ser Pro Leu Leu Ala Ala Thr Phe
    15                20                25                30

gct tac tgg gat aat att ctt ggt cct aga gta agg cac att tgg gct 507
Ala Tyr Trp Asp Asn Ile Leu Gly Pro Arg Val Arg His Ile Trp Ala
          35                40                45

cca aag aca gac caa gta ctc ctc agt gat gga gaa atc act ttt ctt 555
Pro Lys Thr Asp Gln Val Leu Leu Ser Asp Gly Glu Ile Thr Phe Leu
          50                55                60

gcc aac cac act ctg aat gga gaa att ctt cgg aat gcg gag agt ggg 603
Ala Asn His Thr Leu Asn Gly Glu Ile Leu Arg Asn Ala Glu Ser Gly
          65                70                75

gca ata gat gta aag ttt ttt gtc tta tct gaa aag ggc gtc att att 651
Ala Ile Asp Val Lys Phe Phe Val Leu Ser Glu Lys Gly Val Ile Ile
          80                85                90

gtt tca tta atc ttc gac ggg aac tgg aac gga gat cgg agc act tac 699
Val Ser Leu Ile Phe Asp Gly Asn Trp Asn Gly Asp Arg Ser Thr Tyr
          95                100                105                110

gga cta tca att ata ctg ccg cag acg gag ctg agt ttc tac ctc cca 747
Gly Leu Ser Ile Ile Leu Pro Gln Thr Glu Leu Ser Phe Tyr Leu Pro
          115                120                125

ctg cac aga gtg tgt gtt gac agg cta acg cac atc att cga aaa gga 795
Leu His Arg Val Cys Val Asp Arg Leu Thr His Ile Ile Arg Lys Gly
          130                135                140

agg ata tgg atg cac aag gaa aga caa gaa aat gtc cag aaa att gtc 843

Arg Ile Trp Met His Lys Glu Arg Gln Glu Asn Val Gln Lys Ile Val
    145                150                155

ttg gaa ggc acc gag agg atg gaa gat cag ggc cag agt atc atc cct 891
Leu Glu Gly Thr Glu Arg Met Glu Asp Gln Gly Gln Ser Ile Ile Pro
    160                165                170

atg ctt act ggg gag gtc atc cct gtg atg gag ctg ctt gcg tct atg 939
Met Leu Thr Gly Glu Val Ile Pro Val Met Glu Leu Leu Ala Ser Met
    175                180                185                190

aga tca cac agt gtt cct gaa gac ctc gat ata gct gat aca gta ctc 987
Arg Ser His Ser Val Pro Glu Asp Leu Asp Ile Ala Asp Thr Val Leu
          195                200                205

aat gat gat gac att ggt gac agc tgt cat gaa ggc ttt ctt ctc aat 1035
Asn Asp Asp Asp Ile Gly Asp Ser Cys His Glu Gly Phe Leu Leu Asn
          210                215                220

```

210	215	220	
gcc atc agc tca cat ctg cag acc tgc ggc tgc tct gtg gtg gta ggc	1083		
Ala Ile Ser Ser His Leu Gln Thr Cys Gly Cys Ser Val Val Val Gly			
225	230	235	
agc agt gca gag aaa gta aat aag ata gta aga aca ctg tgc ctt ttt	1131		
Ser Ser Ala Glu Lys Val Asn Lys Phe Val Arg Thr Leu Cys Leu Phe			
240	245	250	
ctg aca cca gca gag agg aag tgc tcc agg ctg tgt gaa gcc gaa tgc	1179		
Leu Thr Pro Ala Glu Arg Lys Cys Ser Arg Leu Cys Glu Ala Glu Ser			
255	260	265	270
tcc ttt aaa tac gaa tct gga ctc ttt gta caa ggc ttg cta aag gat	1227		
Ser Phe Lys Tyr Glu Ser Gly Leu Phe Val Gln Gly Leu Leu Lys Asp			
275	280	285	
gcg act ggc agt ttt gta cta cct ttc cgg caa gtt atg tat gcc cct	1275		
Ala Thr Gly Ser Phe Val Leu Pro Phe Arg Gln Val Met Tyr Ala Pro			
290	295	300	
tat ccc acc aca cac atc gat gtg gat gtc aac act gtc aag cag atg	1323		
Tyr Pro Thr Thr His Ile Asp Val Asp Val Asn Thr Val Lys Gln Met			
305	310	315	
cca ccg tgt cat gaa cat att tat aat caa cgc aga tac atg agg tca	1371		
Pro Pro Cys His Glu His Ile Tyr Asn Gln Arg Arg Tyr Met Arg Ser			
320	325	330	
gag ctg aca gcc ttc tgg agg gca act tca gaa gag gac atg gct cag	1419		
Glu Leu Thr Ala Phe Trp Arg Ala Thr Ser Glu Glu Asp Met Ala Gln			
335	340	345	350
gac acc atc atc tac aca gat gag agc ttc act cct gat ttg aat att	1467		
Asp Thr Ile Ile Tyr Thr Asp Glu Ser Phe Thr Pro Asp Leu Asn Ile			
355	360	365	
ttc caa gat gtc tta cac aga gac act cta gtg aaa gcc ttt ctg gat	1515		
Phe Gln Asp Val Leu His Arg Asp Thr Leu Val Lys Ala Phe Leu Asp			
370	375	380	
cag gtc ttc cat ttg aag cct ggc ctg tct ctc agg agt act ttc ctt	1563		
Gln Val Phe His Leu Lys Pro Gly Leu Ser Leu Arg Ser Thr Phe Leu			
385	390	395	
gca cag ttc ctc ctc att ctt cac aga aaa gcc ttg aca cta atc aag	1611		
Ala Gln Phe Leu Leu Ile Leu His Arg Lys Ala Leu Thr Leu Ile Lys			
400	405	410	
tac ata gag gat gac acg cag aag ggg aaa aag ccc ttt aag tct ctt	1659		
Tyr Ile Glu Asp Asp Thr Gln Lys Gly Lys Lys Pro Phe Lys Ser Leu			
415	420	425	430
cgg aac ctg aag ata gat ctt gat tta aca gca gag ggc gac ctt aac	1707		
Arg Asn Leu Lys Ile Asp Leu Asp Leu Thr Ala Glu Gly Asp Leu Asn			
435	440	445	
ata ata atg gct cta gct gag aaa att aag cca ggc cta cac tct ttc	1755		
Ile Ile Met Ala Leu Ala Glu Lys Ile Lys Pro Gly Leu His Ser Phe			
450	455	460	
atc ttc ggg aga cct ttc tac act agt gtc caa gaa cgt gat gtt cta	1803		
Ile Phe Gly Arg Pro Phe Tyr Thr Ser Val Gln Glu Arg Asp Val Leu			
465	470	475	
atg act ttt taa acatgtgggtt tgctccgtgt gtctcatgac agtcacactt	1855		
Met Thr Phe			
480			

```

getgttacag tgtctcagcg ctttggacac atccttctct cagggctcctg ccgcaggaca 1915
cgttacacta cacttgtcag tagaggtctg taccagatgt caggtacatc gttgtagtga 1975
atgtctcttt tcctagaacta gatgtaccct cgtagggact tatgtttaca accctcctaa 2035
gtactagtgc tgtcttgtta ggatacgaat gaagggatgt aaacttcacc acaactgctg 2095
gtttggtttt ttgtttttgt tttttgaaac ttataattca tggtttacat gcatcacact 2155
gaaaccctag ttagcttttt acaggttaagc tgtgagttga ctgctgtctc ctgtgttctc 2215
tggcctgtac gatctgtggc gtgtaggac accttttgca caactaaaaa ctaaagcact 2275
ttgttttgag ttctacagaa agcaacttag tctgtctgca gattcgtttt tgaaagaaga 2335
catgagaaaag cggagttttt ggtgaagtca gttgttgat cttcctttat agacttagtc 2395
cttttagatgt ggtctgtata gacatgccca accatcatgc atgggcactg aatatcgtga 2455
actgtgggtat gctttttgtt ggtttattgt acttctgtca aagaaagtgg cattggtttt 2515
tataattgtt gccaaagttt aagggttaatt ttcatatttt ttgagccaaa ttaaaaatgtg 2575
cacctcctgt gcctttccca atcttggaaa atataatttc ttggcagaag gtcagatttc 2635
agggcccagt cactttcgtc tgacttcctt ttgcacagtc cgccatgggc ctggcttaga 2695
agtctctgta aactatgcc aagagtacat tctgtgataa aatcttcttt gcagagcagg 2755
agagcttctt gcctctttcc tttcatttct gcttggactt tgggtgtctc cacgttccct 2815
gcacctaag gacagcagga gaactctgac ccagtgcta tttctctagg tgctattgtg 2875
gcaaactcaa gcggtccgtc tctgtccctg taacgttctg acctgtctgg ctgtgaagta 2935
ctgactggta aagctccgtg ctacagcagt gtagggtata cacaacaca agtaagtgtt 2995
ttatttaaaa ctgtggactt agcataaaaaa gggagactat atttattttt tacaaaaggg 3055
ataaaaatgg aaccttttcc tcacccacca gatttagtca gaaaaaaca ttctattctg 3115
aaaggtcaca gtgggtttga catgacacat cagaacaacg cacactgtcc atgatggctt 3175
atgaactcca agtcaactcca tcatggtaaa tgggtagatc cctccttcta gtgtgccaca 3235
ccattgcttc ccacagtaga atcttattta agtgctaagt gttgtctctg ctggtttact 3295
ctgttggttt agagaatgta agttgtatag tgaataagtt attgaagcat gtgtaaacac 3355
tgttatacat cttttctcct agatggggaa tttggaataa aataccttta aaattcaaaa 3415
aaaaaaaaa aaaaaaaaaa 3435

```

<210> 13

<400> 13

000

<210> 14

<400> 14

000

<210> 15

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 15

tgtgacagtt ggaatgcagt ga 22

<210> 16

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 16
gccacttaaa gcaatctctg tcttg 25

<210> 17
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 17
tcgactcttt gccaccgcc a 21

<210> 18
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 18
gggtctagca agagcaggtg 20

<210> 19
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 19
gtcttgcaa cagctggaga t 21

<210> 20
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 20
tgatgtcgac tcttgccca ccgc 24

<210> 21

<211> 22
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer

 <400> 21
 tcctgtaatg gaactgcttt ca 22

 <210> 22
 <211> 22
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer

 <400> 22
 ggtatctgct tcatccagct tt 22

 <210> 23
 <211> 16
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Probe

 <400> 23
 ccccggtcccc ggcccc 16

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- [WO2013030588A \[0006\]](#)
- [US2012214865A \[0006\]](#)
- [WO2012092367A \[0006\]](#)
- [WO2012012467A \[0006\]](#)
- [WO2008101157A \[0218\]](#)
- [US20050130923A \[0218\]](#)
- [WO2007134181A \[0218\]](#)
- [US7399845B \[0220\] \[0221\]](#)
- [WO2009006478A \[0220\]](#)
- [WO2008150729A \[0220\]](#)
- [US20040171570A \[0220\]](#)
- [US7427672B \[0220\]](#)
- [WO2008154401A \[0220\]](#)
- [US6268490B \[0221\]](#)
- [US6525191B \[0221\]](#)
- [US6670461B \[0221\]](#)
- [US6770748B \[0221\]](#)
- [US6794499B \[0221\]](#)
- [US7034133B \[0221\]](#)
- [US7053207B \[0221\]](#)
- [US7547684B \[0221\]](#)
- [US7696345B \[0221\]](#)
- [US20080039618A \[0221\]](#)
- [US20090012281A \[0221\]](#)
- [US60989574B \[0221\]](#)
- [US61026995B \[0221\]](#)
- [US61026998B \[0221\]](#)
- [US61056564B \[0221\]](#)
- [US61086231B \[0221\]](#)
- [US61097787B \[0221\]](#)
- [US61099844B \[0221\]](#)
- [WO1994014226A \[0221\]](#)
- [WO2004106356PCT \[0221\]](#)
- [WO2005021570PCT \[0221\]](#)
- [WO2007134181PCT \[0221\]](#)
- [WO2008150729PCT \[0221\]](#)
- [WO2008154401PCT \[0221\]](#)
- [WO2009006478PCT \[0221\]](#)
- [DK9800393W \[0221\]](#)
- [WO9914226A \[0221\] \[0232\] \[0233\]](#)
- [WO9839352A \[0232\]](#)
- [WO03004602A \[0255\]](#)
- [WO61714139A \[0315\]](#)

Non-patent literature cited in the description

- ROWLANDSHNEIDERN. Engl. J. Med., 2001, vol. 344, 1688-1700 [0002]
- HIRTZ et al. Neurology, 2007, vol. 68, 326-337 [0002]
- CHIO et al. Neurology, 2008, vol. 70, 533-537 [0002]
- LILLOHODGESJ. Clin. Neurosci., 2009, vol. 16, 1131-1135 [0002]
- NEUMANN et al. Science, 2006, vol. 314, 130-133 [0002]
- JOHNSON et al. Neuron, 2010, vol. 68, 857-864 [0003]
- KWIATKOWSKI et al. Science, 2009, vol. 323, 1205-1208 [0003]
- MARUYAMA et al. Nature, 2010, vol. 465, 223-226 [0003]
- ROSEN et al. Nature, 1993, vol. 362, 59-62 [0003]
- SREEDHARAN et al. Science, 2008, vol. 319, 1668-1672 [0003]
- VANCE et al. Brain, 2009, vol. 129, 868-876 [0003]
- BOXER et al. J. Neurol. Neurosurg. Psychiatry, 2011, vol. 82, 196-203 [0003]
- MORITA et al. Neurology, 2006, vol. 66, 839-844 [0003]
- PEARSON et al. J. Nerol., 2011, vol. 258, 647-655 [0003]
- VANCE et al. Brain, 2006, vol. 129, 868-876 [0003]
- RENTON et al. Neuron, 2011, vol. 72, 257-268 [0003] [0003] [0293]
- DEJESUS-HEMANDEZ et al. Neuron, 2011, vol. 72, 245-256 [0003]
- LAAKSOVIRTA et al. Lancet Neurol., 2010, vol. 9, 978-985 [0003]
- DONNELLY et al. Annals of Neurology., 2012, vol. 72, 1667-68 [0006]
- RENTONNeuron, 2011, vol. 72, 2257-268 [0006]
- DEJESUS-HERNANDEZNeuron, 2011, vol. 72, 2245-256 [0006]
- SHA et al. Alzheimers Res Ther., 2012, vol. 4, 646- [0006]
- WOOLF et al. Proc. Natl. Acad. Sci. USA, 1992, vol. 89, 7305-7309 [0174]
- GAUTSCHI et al. J. Natl. Cancer Inst., 2001, vol. 93, 463-471 [0175]
- MAHERDOLNICKNuc. Acid. Res., 1988, vol. 16, 3341-3358 [0176]
- ALTSCHUL et al. J. Mol. Biol., 1990, vol. 215, 403-410 [0202]
- ZHANGMADDENGenome Res., 1997, vol. 7, 649-656 [0202]
- SMITHWATERMANAdv. Appl. Math., 1981, vol. 2, 482-489 [0202]
- CHATTOPADHYAYA et al. J. Org. Chem., 2009, vol. 74, 118-134 [0220]
- SINGH et al. Chem. Commun., 1998, vol. 4, 455-456 [0221]
- KOSHKIN et al. Tetrahedron, 1998, vol. 54, 3607-3630 [0221] [0232]
- WAHLESTEDT et al. Proc. Natl. Acad. Sci. U. S. A., 2000, vol. 97, 5633-5638 [0221]
- KUMAR et al. Bioorg. Med. Chem. Lett., 1998, vol. 8, 2219-2222 [0221] [0233]
- SINGH et al. J. Org. Chem., 1998, vol. 63, 10035-10039 [0221] [0233]
- SRIVASTAVA et al. J. Am. Chem. Soc., 2007, vol. 129, 268362-8379 [0221] [0235]
- ELAYADI et al. Curr. Opinion Invest. Drugs, 2001, vol. 2, 558-561 [0221]
- BRAASCH et al. Chem. Biol., 2001, vol. 8, 1-7 [0221]
- ORUM et al. Curr. Opinion Mol. Ther., 2001, vol. 3, 239-243 [0221]
- FRIEDEN et al. Nucleic Acids Research., 2003, vol. 21, 6365-6372 [0224]

- **FREIER et al.** Nucleic Acids Research, 1997, vol. 25, 2244-2944 [\[0235\]](#)
- **ALBAEK et al.** J. Org. Chem., 2006, vol. 71, 7731-7740 [\[0235\]](#)
- **BAKER et al.** J. Biol. Chem., 1997, vol. 272, 11944-12000 [\[0238\]](#)
- **MARTIN** Helv. Chim. Acta, 1995, vol. 78, 486-504 [\[0238\]](#)
- **ALTMANN et al.** Chimia, 1996, vol. 50, 168-176 [\[0238\]](#)
- **ALTMANN et al.** Biochem. Soc. Trans., 1996, vol. 24, 630-637 [\[0238\]](#)
- **ALTMANN et al.** Nucleosides Nucleotides, 1997, vol. 16, 917-926 [\[0238\]](#)
- **LEUMANN** Bioorg. Med. Chem., 2002, vol. 10, 841-854 [\[0239\]](#) [\[0246\]](#)
- **JONES, L.J. et al.** Analytical Biochemistry, 1998, vol. 265, 368-374 [\[0268\]](#)
- **MARGOLIS et al.** Hum. Mol. Genet., 2006, vol. 15, 1808-1815 [\[0275\]](#)
- **DEJESUS-HERNANDEZ, M. et al.** Neuron, 2011, vol. 72, 245-56 [\[0297\]](#)

Patentkrav

1. Forbindelse omfattende et modificeret antisense-oligonukleotid bestående af 12 til 30 bundne nukleosider, hvor det modificerede antisense-oligonukleotid er mindst 90% komplementært med en C9ORF72 nukleinsyre, til anvendelse i
5 behandling af en neurodegenerativ sygdom ved at reducere procentandelen af cellerne med nukleære foci og/eller reducere antallet af nukleære foci pr. celle i et dyr identificeret som havende foci, hvor det modificerede antisense-oligonukleotid omfatter mindst én phosphorthioat-internukleosidbinding, mindst én 5-methyl-cytosin, mindst ét bicyklisk sukker eller mindst én 2'-O-methoxyethylgruppe.
10
2. Forbindelsen til anvendelse ifølge krav 1, hvor sygdommen er associeret med C9ORF72-hexanucleotid-repeat ekspansion.
3. Forbindelsen til anvendelse ifølge krav 2, hvor sygdommen er hvilke som helst
15 af amyotrofisk lateral sclerose (ALS), frontotemporal demens (FTD), corticalbasal degenerationssyndrom (CBD), atypisk Parkinson-syndrom, eller olivopontocerebellær degeneration (OPCD).
4. Forbindelsen til anvendelse ifølge et hvilket som helst af kravene 1-3, hvor
20 nukleobasesekvensen af det modificerede antisense-oligonukleotid er 100% komplementært med en C9ORF72 nukleinsyre med samme længde.
5. Forbindelsen til anvendelse ifølge et hvilket som helst af kravene 1-4, hvor nukleobasesekvensen af det modificerede antisenseoligonukleotid er
25 komplementært med en hexanukleotid-repeat ekspansion, eventuelt hvor hexanukleotid-repeat ekspansionen er hvilken som helst af GGGGCC, GGGGGG, GGGGGC, eller GGGGCG.
6. Forbindelsen til anvendelse ifølge et hvilket som helst af kravene 1-5, hvor
30 C9ORF72-nukleinsyren har nukleobasesekvensen af hvilke som helst af SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, eller SEQ ID NO: 10.

7. Forbindelsen til anvendelse ifølge et hvilket som helst af kravene 1-6, hvor det modificerede antisense-oligonukleotid er et enkelt-strengt oligonukleotid.

8. Forbindelsen til anvendelse ifølge krav 7, hvor det enkelt-strengede
5 modificerede antisense-oligonukleotid omfatter:

et gap-segment bestående af bundne deoxynukleosider;

et 5'-wing-segment bestående af bundne nukleosider;

et 3'-wing-segment bestående af bundne nukleosider;

10 hvor gap-segmentet er positioneret umiddelbart grænsende til og mellem
5'-wing-segmentet og 3'-wing-segmentet og hvor hvert nukleosid af hvert
wing-segment omfatter et modificeret sukker.

9. Forbindelsen til anvendelse ifølge et hvilket som helst af kravene 1-8, hvor det
modificerede antisense-oligonukleotid omfatter mindst én phosphorthioat-
internukleosidbinding.

15

10. Forbindelsen til anvendelse ifølge krav 9, hvor hver internukleosidbinding er
en phosphorthioat-internukleosidbinding.

11. Forbindelsen til anvendelse ifølge et hvilket som helst af kravene 1-10, hvor
20 det modificerede antisense oligonukleotid omfatter mindst én 5-methylcytosin.

12. Forbindelsen til anvendelse ifølge et hvilket som helst af kravene 1-11, hvor
det modificerede antisense-oligonukleotid omfatter mindst ét modificeret
nukleosid omfattende et bicyklisk sukker.

25

13. Forbindelsen til anvendelse ifølge krav 12, hvor det bicykliske sukker omfatter
en kemisk bro mellem 4'- og 2'-positionerne af sukkeret.

- 14.** Forbindelsen til anvendelse ifølge krav 13, hvor den kemiske bro er a) 4'-CH(R)-O-2' og hvor R er methyl; b) 4'-CH(R)-O-2' og hvor R er H; eller c) 4'-CH(R)-O-2' og hvor R er -CH₂-O-CH₃.
- 5 **15.** Forbindelsen til anvendelse ifølge et hvilket som helst af kravene 1-14, hvor det modificerede antisense-oligonukleotid omfatter mindst ét modificeret nukleosid omfattende en 2'-O-methoxyethylgruppe.
- 16.** Forbindelsen til anvendelse ifølge et hvilket som helst af kravene 1-15, hvor
- 10 forbindelsen er en konjugeret forbindelse.