

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
27 May 2004 (27.05.2004)

PCT

(10) International Publication Number
WO 2004/043394 A2

(51) International Patent Classification⁷: **A61K**

(21) International Application Number:
PCT/US2003/035903

(22) International Filing Date:
12 November 2003 (12.11.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/293,864 11 November 2002 (11.11.2002) US

(71) Applicant (*for all designated States except US*): **ISIS PHARMACEUTICALS, INC.** [US/US]; 2292 Faraday Avenue, Carlsbad, CA 92008 (US).

(72) Inventor; and

(75) Inventor/Applicant (*for US only*): **DOBIE, Kenneth, W.** [US/US]; 703 Stratford Court, #4, Delmar, CA 92014 (US).

(74) Agent: **LEGAARD, Paul, K.**; Cozen O'Connor, 1900 Market Street, Philadelphia, PA 19103 (US).

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

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

Published:

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

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

(54) Title: MODULATION OF HUNTINGTIN INTERACTING PROTEIN 1 EXPRESSION

(57) Abstract: Compounds, compositions and methods are provided for modulating the expression of huntingtin interacting protein 1. The compositions comprise oligonucleotides, targeted to nucleic acid encoding huntingtin interacting protein 1. Methods of using these compounds for modulation of huntingtin interacting protein 1 expression and for diagnosis and treatment of disease associated with expression of huntingtin interacting protein 1 are provided.



WO 2004/043394 A2

MODULATION OF HUNTINGTIN INTERACTING PROTEIN 1 EXPRESSION

FIELD OF THE INVENTION

The present invention provides compositions and methods for modulating the expression of huntingtin interacting protein 1. In particular, this invention relates to compounds, particularly oligomeric compounds such as oligonucleotide compounds, which, in some embodiments, hybridize with nucleic acid molecules encoding huntingtin interacting protein 1. Such compounds are shown herein to modulate the expression of huntingtin interacting protein 1.

BACKGROUND OF THE INVENTION

Huntington's disease is a debilitating condition characterized by shaky movements, impaired cognitive and emotional functions and eventually leads to dementia and death. Huntington's disease is caused by the death of a specific group of nerve cells which subsequently alters the brain's ability to coordinate movement. This inherited neurological disease is caused by the abnormal lengthening of a CAG repeat in the gene encoding the huntingtin protein, resulting in long stretches of glutamine within the protein. Mutant huntingtin triggers apoptosis, and one biochemical mechanism through which it does so involves several huntingtin interacting proteins and the protein-cleaving enzymes called caspases. In normal nerve cells, huntingtin can form a complex with several proteins, including huntingtin interacting protein 1 (also called HIP1 and HIP-1), but the mutant huntingtin with a longer glutamine tract has a weaker interaction with huntingtin interacting protein 1. Huntingtin interacting protein 1 instead interacts with huntingtin interacting protein 1 protein interactor (HIPPI) and subsequently induces an apoptotic cascade involving caspase-8 and caspase-3, caspases which have been implicated in neuronal death (Davies and Ramsden, *Mol. Pathol.*, 2001, 54, 409-413; and Mattson, *Nature*, 2002, 415, 377-379).

The gene encoding huntingtin interacting protein 1 was cloned in 1997 and is ubiquitously expressed in different brain regions at low levels (Wanker et al., *Hum. Mol. Genet.*, 1997, 6, 487-495). The gene has been mapped close to the Elastin Locus on chromosome 7q11.23, a region which when deleted produces Williams-Beuren syndrome (Wedemeyer et al., *Genomics*, 1997, 46, 313-315). The genomic DNA gives rise to two alternative splice forms termed HIP1-1 and HIP1-2 which differ in their 5-prime sequence (Chopra et al., *Mamm. Genome*, 2000, 11, 1006-1015). Claimed and disclosed in PCT Publication WO 97/18825 is a

cDNA molecule comprising the sequence of huntingtin interacting protein 1 (Kalchman and Hayden, 1997).

Like HIPPI, huntingtin interacting protein 1 contains a pseudo-death effector domain (DED). The DED is a small protein-protein interaction domain that facilitates the assembly of protein components required for the execution of various cell-death pathways and is also found in caspase-8. Overexpression of mutant huntingtin has been shown to induce apoptosis in a caspase-8 dependent manner, and the heterodimer formed between huntingtin interacting protein 1 and HIPPI recruits and activates procaspase-8. Since huntingtin interacting protein 1 has a higher affinity for HIPPI than mutant huntingtin, diseased brains contain higher levels of the huntingtin interacting protein 1/ HIPPI complex, thereby initiating the apoptotic cascade and suggesting a potential molecular basis for the pathogenesis of Huntington's disease (Gervais et al., *Nat Cell Biol*, 2002, 4, 95-105).

This toxic gain-of-function in Huntington's disease resulting from abnormal interactions between mutated huntingtin, huntingtin interacting protein 1, and HIPPI is only one role that huntingtin interacting protein 1 holds. In normal cells, huntingtin interacting protein 1 binds to clathrin and the clathrin adaptor protein 2 (AP2). Clathrin-mediated endocytosis is a major pathway for internalization of macromolecules into the cytoplasm, thus huntingtin interacting protein 1 is component of the endocytic machinery (Metzler et al., *J. Biol. Chem.*, 2001, 276, 39271-39276; Mishra et al., *J. Biol. Chem.*, 2001, 276, 46230-46236; Rao et al., *Mol. Cell. Biol.*, 2001, 21, 7796-7806; and Waelter et al., *Hum. Mol. Genet.*, 2001, 10, 1807-1817).

Huntingtin interacting protein 1 may contribute to the pathogenesis of hematopoietic malignancy. The fusion of the huntingtin interacting protein 1 gene to the platelet-derived growth factor beta receptor (PDGFBR) gene arises via the t(5:7)(q33;q11.2) chromosomal translocation and has been identified in leukemic cells of a patient with chronic myelomonocytic leukemia (CMML) (Ross et al., *Blood*, 1998, 91, 4419-4426). The resultant protein is constitutively tyrosine-phosphorylated and transforms the murine hematopoietic cell line Ba/F3 to interleukin-3-independent growth (Ross and Gilliland, *J. Biol. Chem.*, 1999, 274, 22328-22336). Furthermore, this fusion protein associates with SHIP1 preventing SHIP1 from binding its substrates phosphatidylinositol-3,4,5-triphosphate and inositol-1,3,4,5-tetraphosphate, an action that may alter the levels of these signal transduction molecules and resulting in activation of cellular proliferation or survival (Saint-Dic et al., *J. Biol. Chem.*, 2001, 276, 21192-21198).

Currently, there are no known therapeutic agents that effectively inhibit the synthesis of huntingtin interacting protein 1. To date, investigative strategies aimed at modulating huntingtin interacting protein 1 function have involved the use of inactive mutants to elucidate the role of

-3-

the DED (Hackam et al., J. Biol. Chem., 2000, 275, 41299-41308), regions required for transforming Ba/F3 cells (Ross and Gilliland, J. Biol. Chem., 1999, 274, 22328-22336), and regions required for binding clathrin and AP2 and effecting endocytosis (Metzler et al., J. Biol. Chem., 2001, 276, 39271-39276; and Waelter et al., Hum. Mol. Genet., 2001, 10, 1807-1817). In addition, mice with targeted deletions of huntingtin interacting protein 1 develop normally into adulthood, but have increased apoptosis of postmeiotic spermatids, indicating that huntingtin interacting protein 1 is required for differentiation, proliferation, or survival of spermatogenic progenitors (Rao et al., Mol. Cell. Biol., 2001, 21, 7796-7806).

Consequently, there remains a long felt need for agents capable of effectively inhibiting huntingtin interacting protein 1 function.

Antisense technology is emerging as an effective means for reducing the expression of specific gene products and may therefore prove to be uniquely useful in a number of therapeutic, diagnostic, and research applications for the modulation of huntingtin interacting protein 1 expression.

The present invention provides compositions and methods for modulating huntingtin interacting protein 1 expression.

SUMMARY OF THE INVENTION

The present invention is directed to compounds, especially nucleic acid and nucleic acid-like oligomers, which are targeted to a nucleic acid encoding huntingtin interacting protein 1, and which modulate the expression of huntingtin interacting protein 1. Pharmaceutical and other compositions comprising the compounds of the invention are also provided. Further provided are methods of screening for modulators of huntingtin interacting protein 1 and methods of modulating the expression of huntingtin interacting protein 1 in cells, tissues or animals comprising contacting said cells, tissues or animals with one or more of the compounds or compositions of the invention. Methods of treating an animal, particularly a human, suspected of having or being prone to a disease or condition associated with expression of huntingtin interacting protein 1 are also set forth herein. Such methods comprise administering a therapeutically or prophylactically effective amount of one or more of the compounds or compositions of the invention to the person in need of treatment.

DETAILED DESCRIPTION OF THE INVENTION

The present invention employs compounds, including oligomers such as oligonucleotides and similar species for use in modulating the function or effect of nucleic acid

molecules encoding huntingtin interacting protein 1. This is accomplished by providing oligonucleotides that specifically hybridize with one or more nucleic acid molecules encoding huntingtin interacting protein 1. As used herein, the terms "target nucleic acid" and "nucleic acid molecule encoding huntingtin interacting protein 1" have been used for convenience to encompass DNA encoding huntingtin interacting protein 1, RNA (including pre-mRNA and mRNA or portions thereof) transcribed from such DNA, and also cDNA derived from such RNA. The hybridization of a compound of this invention with its target nucleic acid is generally referred to as "antisense." Consequently, a mechanism believed to be included in the practice of some embodiments of the invention is referred to herein as "antisense inhibition." Such antisense inhibition is typically based upon hydrogen bonding-based hybridization of oligonucleotide strands or segments such that at least one strand or segment is cleaved, degraded, or otherwise rendered inoperable. In this regard, specific nucleic acid molecules and their functions can be targeted for such antisense inhibition.

The functions of DNA to be interfered with include, but are not limited to, replication and transcription. Replication and transcription, for example, can be from an endogenous cellular template, a vector, a plasmid construct or otherwise. Functions of RNA to be interfered with also include functions such as, for example, translocation of the RNA to a site of protein translation, translocation of the RNA to sites within the cell which are distant from the site of RNA synthesis, translation of protein from the RNA, splicing of the RNA to yield one or more RNA species, and catalytic activity or complex formation involving the RNA which may be engaged in or facilitated by the RNA. One result of such interference with target nucleic acid function is modulation of the expression of huntingtin interacting protein 1. In the context of the present invention, "modulation" and "modulation of expression" mean either an increase (stimulation) or a decrease (inhibition) in the amount or levels of a nucleic acid molecule encoding the gene, e.g., DNA or RNA. Inhibition is often a desired form of modulation of expression and mRNA is often a desired target nucleic acid.

In the context of this invention, "hybridization" means the pairing of complementary strands of oligomeric compounds. In the present invention, one mechanism of pairing involves hydrogen bonding, which may be Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding, between complementary nucleoside or nucleotide bases (nucleobases) of the strands of oligomeric compounds. For example, adenine and thymine are complementary nucleobases that pair through the formation of hydrogen bonds. Hybridization can occur under varying circumstances.

The compounds of the invention are specifically hybridizable when binding of the compound to the target nucleic acid interferes with the normal function of the target nucleic acid to cause a loss of activity. In some embodiments, there may be a sufficient degree of complementarity to avoid non-specific binding of the compound to non-target nucleic acid sequences under conditions in which specific binding is desired, i.e., under physiological conditions in the case of *in vivo* assays or therapeutic treatment, and under conditions in which assays are performed in the case of *in vitro* assays.

In the present invention the phrase “stringent hybridization conditions” or “stringent conditions” refers to conditions under which a compound of the invention will hybridize to its target sequence, but to a minimal number of other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances and in the context of this invention, “stringent conditions” under which oligomeric compounds hybridize to a target sequence are determined by the nature and composition of the oligomeric compounds and the assays in which they are being investigated.

“Complementary,” as used herein, refers to the capacity for precise pairing between two nucleobases of an oligomeric compound. For example, if a nucleobase at a certain position of an oligonucleotide (an oligomeric compound), is capable of hydrogen bonding with a nucleobase at a certain position of a target nucleic acid, the target nucleic acid being a DNA, RNA, or oligonucleotide molecule, then the position of hydrogen bonding between the oligonucleotide and the target nucleic acid is considered to be a complementary position. The oligonucleotide and the further DNA, RNA, or oligonucleotide molecule are complementary to each other when a sufficient number of complementary positions in each molecule are occupied by nucleobases which can hydrogen bond with each other. Thus, “specifically hybridizable” and “complementary” are terms which are used to indicate a sufficient degree of precise pairing or complementarity over a sufficient number of nucleobases such that stable and specific binding occurs between the oligonucleotide and a target nucleic acid.

It is understood in the art that the sequence of a compound need not be 100% complementary to that of its target nucleic acid to be specifically hybridizable. Moreover, an oligonucleotide may hybridize over one or more segments such that intervening or adjacent segments are not involved in the hybridization event (e.g., a loop structure or hairpin structure). The compounds of the present invention can comprise at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence complementarity to a target region within the target nucleic acid sequence to which they are targeted. For example, a compound in which 18 of 20 nucleobases of the 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, a 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 and would fall within the scope of the present invention. Percent complementarity of a 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; and 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 WI), using default settings, which uses the algorithm of Smith and Waterman (*Adv. Appl. Math.*, 1981, 2, 482-489). In some embodiments, homology, sequence identity or complementarity, between the oligomeric compound and target is between about 50% to about 60%, between about 60% to about 70%, between about 70% and about 80%, or between about 80% and about 90%. In other embodiments, homology, sequence identity or complementarity, is about 90%, about 92%, about 94%, about 95%, about 96%, about 97%, about 98%, or about 99%.

According to the present invention, compounds include antisense oligomeric compounds, antisense oligonucleotides, ribozymes, external guide sequence (EGS) oligonucleotides, alternate splicers, primers, probes, and other oligomeric compounds that hybridize to at least a portion of the target nucleic acid. As such, these compounds may be introduced in the form of single-stranded, double-stranded, circular or hairpin oligomeric compounds and may contain structural elements such as internal or terminal bulges or loops. Once introduced to a system, the compounds of the invention may elicit the action of one or more enzymes or structural proteins to effect modification of the target nucleic acid.

One non-limiting example of such an enzyme is RNase H, a cellular endonuclease which cleaves the RNA strand of an RNA:DNA duplex. It is known in the art that single-stranded antisense compounds which are "DNA-like" elicit RNase H. Activation of RNase H, therefore, results in cleavage of the RNA target, thereby greatly enhancing the efficiency of oligonucleotide-mediated inhibition of gene expression. Similar roles have been postulated for other ribonucleases such as those in the RNase III and ribonuclease L family of enzymes.

While one form of an antisense compound is a single-stranded antisense oligonucleotide, in many species the introduction of double-stranded structures, such as double-stranded RNA (dsRNA) molecules, has been shown to induce potent and specific antisense-mediated reduction of the function of a gene or its associated gene products. This phenomenon occurs in both plants and animals and is believed to have an evolutionary connection to viral defense and transposon silencing.

The first evidence that dsRNA could lead to gene silencing in animals came in 1995 from work in the nematode, *Caenorhabditis elegans* (Guo and Kempheus, Cell, 1995, 81, 611-620). Montgomery et al. have shown that the primary interference effects of dsRNA are posttranscriptional (Montgomery et al., Proc. Natl. Acad. Sci. USA, 1998, 95, 15502-15507). The posttranscriptional antisense mechanism defined in *Caenorhabditis elegans* resulting from exposure to double-stranded RNA (dsRNA) has since been designated RNA interference (RNAi). This term has been generalized to mean antisense-mediated gene silencing involving the introduction of dsRNA leading to the sequence-specific reduction of endogenous targeted mRNA levels (Fire et al., Nature, 1998, 391, 806-811). Recently, it has been shown that it is, in fact, the single-stranded RNA oligomers of antisense polarity of the dsRNAs which are the potent inducers of RNAi (Tijsterman et al., Science, 2002, 295, 694-697).

The oligonucleotides of the present invention also include variants in which a different base is present at one or more of the nucleotide positions in the oligonucleotide. For example, if the first nucleotide is an adenosine, variants may be produced which contain thymidine, guanosine or cytidine at this position. This may be done at any of the positions of the oligonucleotide. These oligonucleotides are then tested using the methods described herein to determine their ability to inhibit expression of huntingtin interacting protein 1 mRNA.

In the context of this invention, the term "oligomeric compound" refers to a polymer or oligomer comprising a plurality of monomeric units. In the context of this invention, the term "oligonucleotide" refers to an oligomer or polymer of ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) or mimetics, chimeras, analogs and homologs thereof. This term includes oligonucleotides composed of naturally occurring nucleobases, sugars and covalent internucleoside (backbone) linkages as well as oligonucleotides having non-naturally occurring portions that function similarly. Such modified or substituted oligonucleotides are often favorable over native forms because of desirable properties such as, for example, enhanced cellular uptake, enhanced affinity for a target nucleic acid and increased stability in the presence of nucleases.

While oligonucleotides are one form of the compounds of this invention, the present invention comprehends other families of compounds as well, including but not limited to oligonucleotide analogs and mimetics such as those described herein.

The compounds in accordance with this invention can comprise from about 8 to about 80 nucleobases (i.e. from about 8 to about 80 linked nucleosides). One of ordinary skill in the art will appreciate that the invention embodies compounds of 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 nucleobases in length.

In one embodiment, the compounds of the invention are 12 to 50 nucleobases in length. One having ordinary skill in the art will appreciate that this embodies compounds of 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, or 50 nucleobases in length.

In another embodiment, the compounds of the invention are 15 to 30 nucleobases in length. One having ordinary skill in the art will appreciate that this embodies compounds of 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleobases in length.

In other embodiments, the compounds are oligonucleotides from about 12 to about 50 nucleobases or from about 15 to about 30 nucleobases.

Antisense compounds 8-80 nucleobases in length comprising a stretch of at least eight (8) consecutive nucleobases selected from within the illustrative compounds are considered to be suitable compounds as well.

Exemplary compounds include oligonucleotide sequences that comprise at least the 8 consecutive nucleobases from the 5'-terminus of one of the illustrative compounds (the remaining nucleobases being a consecutive stretch of the same oligonucleotide beginning immediately upstream of the 5'-terminus of the compound that is specifically hybridizable to the target nucleic acid and continuing until the oligonucleotide contains about 8 to about 80 nucleobases). Similarly, compounds are represented by oligonucleotide sequences that comprise at least the 8 consecutive nucleobases from the 3'-terminus of one of the illustrative compounds (the remaining nucleobases being a consecutive stretch of the same oligonucleotide beginning immediately downstream of the 3'-terminus of the compound that is specifically hybridizable to the target nucleic acid and continuing until the oligonucleotide contains about 8 to about 80 nucleobases). One having skill in the art armed with the compounds illustrated herein will be able, without undue experimentation, to identify additional compounds.

“Targeting” a compound to a particular nucleic acid molecule, in the context of this invention, can be a multistep process. The process can begin with the identification of a target nucleic acid whose function is to be modulated. This target nucleic acid may be, for example, a cellular gene (or mRNA transcribed from the gene) whose expression is associated with a particular disorder or disease state, or a nucleic acid molecule from an infectious agent. In the present invention, the target nucleic acid molecule encodes huntingtin interacting protein 1.

The targeting process can also include determination of at least one target region, segment, or site within the target nucleic acid for the antisense interaction to occur such that the desired effect, e.g., modulation of expression, will result. Within the context of the present invention, the term “region” is defined as a portion of the target nucleic acid having at least one identifiable structure, function, or characteristic. Within regions of target nucleic acids are segments. “Segments” are defined as smaller or sub-portions of regions within a target nucleic acid. “Sites,” as used in the present invention, are defined as positions within a target nucleic acid.

Since, as is known in the art, the translation initiation codon is typically 5'-AUG (in transcribed mRNA molecules; 5'-ATG in the corresponding DNA molecule), the translation initiation codon is also referred to as the “AUG codon,” the “start codon” or the “AUG start codon.” A minority of genes have a translation initiation codon having the RNA sequence 5'-GUG, 5'-UUG or 5'-CUG, and 5'-AUA, 5'-ACG and 5'-CUG have been shown to function *in vivo*. Thus, the terms “translation initiation codon” and “start codon” can encompass many codon sequences, even though the initiator amino acid in each instance is typically methionine (in eukaryotes) or formylmethionine (in prokaryotes). It is also known in the art that eukaryotic and prokaryotic genes may have two or more alternative start codons, any one of which may be preferentially utilized for translation initiation in a particular cell type or tissue, or under a particular set of conditions. In the context of the invention, “start codon” and “translation initiation codon” refer to the codon or codons that are used *in vivo* to initiate translation of an mRNA transcribed from a gene encoding huntingtin interacting protein 1, regardless of the sequence(s) of such codons. It is also known in the art that a translation termination codon (or “stop codon”) of a gene may have one of three sequences, i.e., 5'-UAA, 5'-UAG and 5'-UGA (the corresponding DNA sequences are 5'-TAA, 5'-TAG and 5'-TGA, respectively).

The terms “start codon region” and “translation initiation codon region” refer to a portion of such an mRNA or gene that encompasses from about 25 to about 50 contiguous nucleotides in either direction (i.e., 5' or 3') from a translation initiation codon. Similarly, the terms “stop codon region” and “translation termination codon region” refer to a portion of such

an mRNA or gene that encompasses from about 25 to about 50 contiguous nucleotides in either direction (i.e., 5' or 3') from a translation termination codon. Consequently, the "start codon region" (or "translation initiation codon region") and the "stop codon region" (or "translation termination codon region") are all regions which may be targeted effectively with the compounds of the present invention.

The open reading frame (ORF) or "coding region," which is known in the art to refer to the region between the translation initiation codon and the translation termination codon, is also a region which may be targeted effectively. Within the context of the present invention, a suitable region is the intragenic region encompassing the translation initiation or termination codon of the open reading frame (ORF) of a gene.

Other target regions include the 5' untranslated region (5'UTR), known in the art to refer to the portion of an mRNA in the 5' direction from the translation initiation codon, and thus including nucleotides between the 5' cap site and the translation initiation codon of an mRNA (or corresponding nucleotides on the gene), and the 3' untranslated region (3'UTR), known in the art to refer to the portion of an mRNA in the 3' direction from the translation termination codon, and thus including nucleotides between the translation termination codon and 3' end of an mRNA (or corresponding nucleotides on the gene). The 5' cap site of an mRNA comprises an N7-methylated guanosine residue joined to the 5'-most residue of the mRNA via a 5'-5' triphosphate linkage. The 5' cap region of an mRNA is considered to include the 5' cap structure itself as well as the first 50 nucleotides adjacent to the cap site. The 5' cap region can be targeted.

Although some eukaryotic mRNA transcripts are directly translated, many contain one or more regions, known as "introns," which are excised from a transcript before it is translated. The remaining (and therefore translated) regions are known as "exons" and are spliced together to form a continuous mRNA sequence. Targeting splice sites, i.e., intron-exon junctions or exon-intron junctions, may also be particularly useful in situations where aberrant splicing is implicated in disease, or where an overproduction of a particular splice product is implicated in disease. Aberrant fusion junctions due to rearrangements or deletions are also suitable target sites. mRNA transcripts produced via the process of splicing of two (or more) mRNAs from different gene sources are known as "fusion transcripts." It is also known that introns can be effectively targeted using antisense compounds targeted to, for example, DNA or pre-mRNA.

It is also known in the art that alternative RNA transcripts can be produced from the same genomic region of DNA. These alternative transcripts are generally known as "variants." More specifically, "pre-mRNA variants" are transcripts produced from the same genomic DNA

that differ from other transcripts produced from the same genomic DNA in either their start or stop position and contain both intronic and exonic sequence.

Upon excision of one or more exon or intron regions, or portions thereof during splicing, pre-mRNA variants produce smaller "mRNA variants." Consequently, mRNA variants are processed pre-mRNA variants and each unique pre-mRNA variant must always produce a unique mRNA variant as a result of splicing. These mRNA variants are also known as "alternative splice variants." If no splicing of the pre-mRNA variant occurs then the pre-mRNA variant is identical to the mRNA variant.

It is also known in the art that variants can be produced through the use of alternative signals to start or stop transcription and that pre-mRNAs and mRNAs can possess more than one start codon or stop codon. Variants that originate from a pre-mRNA or mRNA that use alternative start codons are known as "alternative start variants" of that pre-mRNA or mRNA. Those transcripts that use an alternative stop codon are known as "alternative stop variants" of that pre-mRNA or mRNA. One specific type of alternative stop variant is the "polyA variant" in which the multiple transcripts produced result from the alternative selection of one of the "polyA stop signals" by the transcription machinery, thereby producing transcripts that terminate at unique polyA sites. Within the context of the invention, the types of variants described herein are also suitable target nucleic acids.

Locations on the target nucleic acid to which the compounds hybridize are hereinbelow referred to as "suitable target segments." As used herein, the term "suitable target segment" is defined as at least an 8-nucleobase portion of a target region to which an active compound is targeted. While not wishing to be bound by theory, it is presently believed that these target segments represent portions of the target nucleic acid which are accessible for hybridization.

While the specific sequences of particular suitable target segments are set forth herein, one of skill in the art will recognize that these serve to illustrate and describe particular embodiments within the scope of the present invention. Additional suitable target segments may be identified by one having ordinary skill.

Once one or more suitable target regions, segments or sites have been identified, compounds are chosen which are sufficiently complementary to the target, i.e., hybridize sufficiently well and with sufficient specificity, to give the desired effect.

The oligomeric compounds are also targeted to or not targeted to regions of the target nucleobase sequence (e.g., such as those disclosed in Example 13) comprising nucleobases 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 601-650, 651-700, 701-750, 751-800, 801-850, 851-900, 901-950, 951-1000, 1001-

1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, 2001-2050, 2051-2100, 2101-2150, 2151-2200, 2201-2250, 2251-2300, 2301-2350, 2351-2400, 2401-2450, 2451-2500, 2501-2550, 2551-2600, 2601-2650, 2651-2700, 2701-2750, 2751-2800, 2801-2850, 2851-2900, 2901-2950, 2951-3000, 3001-3050, 3051-3100, 3101-3150, 3151-3200, 3201-3250, 3251-3300, 3301-3350, 3351-3400, 3401-3450, 3451-3500, 3501-3550, 3551-3600, 3601-3650, 3751-3700, 3701-3750, 3751-3800, 3801-3850, 3851-3900, 3901-3950, 3951-4000, 4001-4050, 4051-4100, 4101-4150, 4151-4200, 4201-4250, 4251-4300, 4301-4350, 4351-4400, 4401-4450, 4451-4500, 4501-4550, 4551-4600, 4601-4650, 4751-4700, 4701-4750, 4751-4800, 4801-4850, 4851-4900, 4901-4950, 4951-5000, 5001-5050, 5051-5100, 5101-5150, 5151-5200, 5201-5250, 5251-5300, 5301-5350, 5351-5400, 5401-5450, 5451-5500, 5501-5550, 5551-5600, 5601-5650, 5751-5700, 5701-5750, 5751-5800, 5801-5850, 5851-5900, 5901-5950, 5951-6000, 6001-6050, 6051-6100, 6101-6150, 6151-6200, 6201-6250, 6251-6300, 6301-6350, 6351-6400, 6401-6450, 6451-6500, 6501-6550, 6551-6600, or 6601-6626, or any combination thereof.

In a further embodiment, the “suitable target segments” identified herein may be employed in a screen for additional compounds that modulate the expression of huntingtin interacting protein 1. “Modulators” are those compounds that decrease or increase the expression of a nucleic acid molecule encoding huntingtin interacting protein 1 and which comprise at least an 8-nucleobase portion which is complementary to a suitable target segment. The screening method can comprise, for example, the steps of contacting a target segment of a nucleic acid molecule encoding huntingtin interacting protein 1 with one or more candidate modulators, and selecting for one or more candidate modulators which decrease or increase the expression of a nucleic acid molecule encoding huntingtin interacting protein 1. Once it is shown that the candidate modulator or modulators are capable of modulating (e.g. either decreasing or increasing) the expression of a nucleic acid molecule encoding huntingtin interacting protein 1, the modulator may then be employed in further investigative studies of the function of huntingtin interacting protein 1, or for use as a research, diagnostic, or therapeutic agent in accordance with the present invention.

The suitable target segments of the present invention may be also be combined with their respective complementary compounds of the present invention to form stabilized double-stranded (duplexed) oligonucleotides. Such double stranded oligonucleotide moieties have been shown in the art to modulate target expression and regulate translation as well as RNA processing via an antisense mechanism. Moreover, the double-stranded moieties may be subject

to chemical modifications (Fire et al., *Nature*, 1998, 391, 806-811; Timmons and Fire, *Nature* 1998, 395, 854; Timmons et al., *Gene*, 2001, 263, 103-112; Tabara et al., *Science*, 1998, 282, 430-431; Montgomery et al., *Proc. Natl. Acad. Sci. USA*, 1998, 95, 15502-15507; Tuschl et al., *Genes Dev.*, 1999, 13, 3191-3197; Elbashir et al., *Nature*, 2001, 411, 494-498; and Elbashir et al., *Genes Dev.* 2001, 15, 188-200). For example, such double-stranded moieties have been shown to inhibit the target by the classical hybridization of antisense strand of the duplex to the target, thereby triggering enzymatic degradation of the target (Tijsterman et al., *Science*, 2002, 295, 694-697).

The compounds of the present invention can also be applied in the areas of drug discovery and target validation. The present invention comprehends the use of the compounds and suitable target segments identified herein in drug discovery efforts to elucidate relationships that exist between huntingtin interacting protein 1 and a disease state, phenotype, or condition. These methods include, for example, detecting or modulating huntingtin interacting protein 1 comprising contacting a sample, tissue, cell, or organism with the compounds of the present invention, measuring the nucleic acid or protein level of huntingtin interacting protein 1 and/or a related phenotypic or chemical endpoint at some time after treatment, and optionally comparing the measured value to a non-treated sample or sample treated with a further compound of the invention. These methods can also be performed in parallel or in combination with other experiments to determine the function of unknown genes for the process of target validation or to determine the validity of a particular gene product as a target for treatment or prevention of a particular disease, condition, or phenotype.

The compounds of the present invention can be utilized for diagnostics, therapeutics, prophylaxis and as research reagents and kits. Furthermore, antisense oligonucleotides, which are able to inhibit gene expression with exquisite specificity, are often used by those of ordinary skill to elucidate the function of particular genes or to distinguish between functions of various members of a biological pathway.

For use in kits and diagnostics, the compounds of the present invention, either alone or in combination with other compounds or therapeutics, can be used as tools in differential and/or combinatorial analyses to elucidate expression patterns of a portion or the entire complement of genes expressed within cells and tissues.

As one nonlimiting example, expression patterns within cells or tissues treated with one or more compounds are compared to control cells or tissues not treated with compounds and the patterns produced are analyzed for differential levels of gene expression as they pertain, for example, to disease association, signaling pathway, cellular localization, expression level, size,

structure or function of the genes examined. These analyses can be performed on stimulated or unstimulated cells and in the presence or absence of other compounds that affect expression patterns.

Examples of methods of gene expression analysis known in the art include DNA arrays or microarrays (Brazma and Vilo, *FEBS Lett.*, 2000, 480, 17-24; Celis, et al., *FEBS Lett.*, 2000, 480, 2-16), SAGE (serial analysis of gene expression)(Madden, et al., *Drug Discov. Today*, 2000, 5, 415-425), READS (restriction enzyme amplification of digested cDNAs) (Prashar and Weissman, *Methods Enzymol.*, 1999, 303, 258-72), TOGA (total gene expression analysis) (Sutcliffe, et al., *Proc. Natl. Acad. Sci. U. S. A.*, 2000, 97, 1976-81), protein arrays and proteomics (Celis, et al., *FEBS Lett.*, 2000, 480, 2-16; Jungblut, et al., *Electrophoresis*, 1999, 20, 2100-10), expressed sequence tag (EST) sequencing (Celis, et al., *FEBS Lett.*, 2000, 480, 2-16; Larsson, et al., *J. Biotechnol.*, 2000, 80, 143-57), subtractive RNA fingerprinting (SuRF) (Fuchs, et al., *Anal. Biochem.*, 2000, 286, 91-98; Larson, et al., *Cytometry*, 2000, 41, 203-208), subtractive cloning, differential display (DD) (Jurecic and Belmont, *Curr. Opin. Microbiol.*, 2000, 3, 316-21), comparative genomic hybridization (Carulli, et al., *J. Cell Biochem. Suppl.*, 1998, 31, 286-96), FISH (fluorescent in situ hybridization) techniques (Going and Gusterson, *Eur. J. Cancer*, 1999, 35, 1895-904) and mass spectrometry methods (To, *Comb. Chem. High Throughput Screen*, 2000, 3, 235-41).

The compounds of the invention are useful for research and diagnostics, because these compounds hybridize to nucleic acids encoding huntingtin interacting protein 1. For example, oligonucleotides that are shown to hybridize with such efficiency and under such conditions as disclosed herein as to be effective huntingtin interacting protein 1 inhibitors will also be effective primers or probes under conditions favoring gene amplification or detection, respectively. These primers and probes are useful in methods requiring the specific detection of nucleic acid molecules encoding huntingtin interacting protein 1 and in the amplification of said nucleic acid molecules for detection or for use in further studies of huntingtin interacting protein 1. Hybridization of the antisense oligonucleotides, particularly the primers and probes, of the invention with a nucleic acid encoding huntingtin interacting protein 1 can be detected by means known in the art. Such means may include conjugation of an enzyme to the oligonucleotide, radiolabelling of the oligonucleotide or any other suitable detection means. Kits using such detection means for detecting the level of huntingtin interacting protein 1 in a sample may also be prepared.

The specificity and sensitivity of antisense is also harnessed by those of skill in the art for therapeutic uses. Antisense compounds have been employed as therapeutic moieties in the

treatment of disease states in animals, including humans. Antisense oligonucleotide drugs, including ribozymes, have been safely and effectively administered to humans and numerous clinical trials are presently underway. It is thus established that antisense compounds can be useful therapeutic modalities that can be configured to be useful in treatment regimes for the treatment of cells, tissues and animals, especially humans.

For therapeutics, an animal, preferably a human, suspected of having a disease or disorder which can be treated by modulating the expression of huntingtin interacting protein 1 is treated by administering antisense compounds in accordance with this invention. For example, in one non-limiting embodiment, the methods comprise the step of administering to the animal in need of treatment, a therapeutically effective amount of a huntingtin interacting protein 1 inhibitor. The huntingtin interacting protein 1 inhibitors of the present invention effectively inhibit the activity of the huntingtin interacting protein 1 protein or inhibit the expression of the huntingtin interacting protein 1 protein. In one embodiment, the activity or expression of huntingtin interacting protein 1 (protein and/or mRNA) in an animal is inhibited by at least 10%, by at least 20%, by at least 25%, by at least 30%, by at least 40%, by at least 50%, by at least 60%, by at least 70%, by at least 75%, by at least 80%, by at least 85%, by at least 90%, by at least 95%, by at least 98%, by at least 99%, or by 100%.

For example, the reduction of the expression of huntingtin interacting protein 1 may be measured in serum, adipose tissue, liver or any other body fluid, tissue or organ of the animal. In some embodiments, the cells contained within said fluids, tissues or organs being analyzed contain a nucleic acid molecule encoding huntingtin interacting protein 1 protein and/or the huntingtin interacting protein 1 protein itself.

The compounds of the invention can be utilized in pharmaceutical compositions by adding an effective amount of a compound to a suitable pharmaceutically acceptable diluent or carrier. Use of the compounds and methods of the invention may also be useful prophylactically.

As is known in the art, a nucleoside is a base-sugar combination. The base portion of the nucleoside is normally a heterocyclic base. The two most common classes of such heterocyclic bases are the purines and the pyrimidines. 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 either the 2', 3' or 5' hydroxyl moiety of the sugar. In forming oligonucleotides, the phosphate groups covalently link adjacent nucleosides to one another to form a linear polymeric compound. In turn, the respective ends of this linear polymeric compound can be further joined to form a circular compound, however, linear compounds are generally favorable. In addition, linear compounds may have

internal nucleobase complementarity and may therefore fold in a manner as to produce a fully or partially double-stranded compound. Within oligonucleotides, the phosphate groups are commonly referred to as forming the internucleoside backbone of the oligonucleotide. The normal linkage or backbone of RNA and DNA is a 3' to 5' phosphodiester linkage.

Modified Internucleoside Linkages (Backbones)

Specific examples of compounds useful in this invention include oligonucleotides containing modified backbones or non-natural internucleoside linkages. As defined in this specification, oligonucleotides having modified backbones include those that retain a phosphorus atom in the backbone and those that do not have a phosphorus atom in the backbone. For the purposes of this specification, and as sometimes referenced in the art, modified oligonucleotides that do not have a phosphorus atom in their internucleoside backbone can also be considered to be oligonucleosides.

Modified oligonucleotide backbones containing a phosphorus atom therein include, for example, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates, 5'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein one or more internucleotide linkages is a 3' to 3', 5' to 5' or 2' to 2' linkage. Oligonucleotides having inverted polarity comprise a single 3' to 3' linkage at the 3'-most internucleotide linkage i.e. a single inverted nucleoside residue which may be abasic (the nucleobase is missing or has a hydroxyl group in place thereof). Various salts, mixed salts and free acid forms are also included.

Representative United States patents that teach the preparation of the above phosphorus-containing linkages include, but are not limited to, U.S.: 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; 5,194,599; 5,565,555; 5,527,899; 5,721,218; 5,672,697; and 5,625,050, certain of which are commonly owned with this application, and each of which is herein incorporated by reference.

Modified oligonucleotide backbones that do not include a phosphorus atom therein have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or more short chain

heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; riboacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH₂ component parts.

Representative United States patents that teach the preparation of the above oligonucleosides include, but are not limited to, U.S.: 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437; 5,792,608; 5,646,269; and 5,677,439, certain of which are commonly owned with this application, and each of which is herein incorporated by reference.

Modified sugar and internucleoside linkages-Mimetics

In other oligonucleotide mimetics, both the sugar and the internucleoside linkage (i.e. the backbone), of the nucleotide units are replaced with novel groups. The nucleobase units are maintained for hybridization with an appropriate target nucleic acid. One such compound, an oligonucleotide mimetic that has been shown to have excellent hybridization properties, is referred to as a peptide nucleic acid (PNA). In PNA compounds, the sugar-backbone of an oligonucleotide is replaced with an amide containing backbone, in particular an aminoethylglycine backbone. The nucleobases are retained and are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone. Representative United States patents that teach the preparation of PNA compounds include, but are not limited to, U.S.: 5,539,082; 5,714,331; and 5,719,262, each of which is herein incorporated by reference. Further teaching of PNA compounds can be found in Nielsen *et al.*, Science, 1991, 254, 1497-1500.

In some embodiments of the invention are oligonucleotides with phosphorothioate backbones and oligonucleosides with heteroatom backbones, and in particular -CH₂-NH-O-CH₂-, -CH₂-N(CH₃)-O-CH₂- (known as a methylene (methylimino) or MMI backbone), -CH₂-O-N(CH₃)-CH₂-, -CH₂-N(CH₃)-N(CH₃)-CH₂- and -O-N(CH₃)-CH₂-CH₂- (wherein the native phosphodiester backbone is represented as -O-P-O-CH₂-) of the above referenced U.S. patent 5,489,677, and the amide backbones of the above referenced U.S. patent 5,602,240. Also suitable are oligonucleotides having morpholino backbone structures of the above-referenced U.S. patent 5,034,506.

Modified sugars

Modified oligonucleotides may also contain one or more substituted sugar moieties. Oligonucleotides comprise one of the following at the 2' position: OH; F; O-, S-, or N-alkyl; O-, S-, or N-alkenyl; O-, S- or N-alkynyl; or O-alkyl-O-alkyl, wherein the alkyl, alkenyl and alkynyl may be substituted or unsubstituted C₁ to C₁₀ alkyl or C₂ to C₁₀ alkenyl and alkynyl. Particular moieties also include O[(CH₂)_nO]_mCH₃, O(CH₂)_nOCH₃, O(CH₂)_nNH₂, O(CH₂)_nCH₃, O(CH₂)_nONH₂, and O(CH₂)_nON[(CH₂)_nCH₃]₂, where n and m are from 1 to about 10. Other oligonucleotides comprise one of the following at the 2' position: C₁ to C₁₀ lower alkyl, substituted lower alkyl, alkenyl, alkynyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH₃, OCN, Cl, Br, CN, 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 the pharmacokinetic properties of an oligonucleotide, or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. Another modification includes 2'-methoxyethoxy (2'-O-CH₂CH₂OCH₃, also known as 2'-O-(2-methoxyethyl) or 2'-MOE) (Martin et al., *Helv. Chim. Acta*, 1995, 78, 486-504) i.e., an alkoxyalkoxy group. Another modification includes 2'-dimethylaminoethoxy, i.e., a O(CH₂)₂ON(CH₃)₂ group, also known as 2'-DMAOE, as described in examples hereinbelow, and 2'-dimethylaminoethoxyethoxy (also known in the art as 2'-O-dimethyl-amino-ethoxy-ethyl or 2'-DMAEOE), i.e., 2'-O-CH₂-O-CH₂-N(CH₃)₂, also described in examples hereinbelow.

Other modifications include 2'-methoxy (2'-O-CH₃), 2'-aminopropoxy (2'-OCH₂CH₂CH₂NH₂), 2'-allyl (2'-CH₂-CH=CH₂), 2'-O-allyl (2'-O-CH₂-CH=CH₂) and 2'-fluoro (2'-F). The 2'-modification may be in the arabino (up) position or ribo (down) position. One 2'-arabino modification is 2'-F. Similar modifications may also be made at other positions on the oligonucleotide, particularly the 3' position of the sugar on the 3' terminal nucleotide or in 2'-5' linked oligonucleotides and the 5' position of 5' terminal nucleotide. Oligonucleotides may also have sugar mimetics such as cyclobutyl moieties in place of the pentofuranosyl sugar. Representative United States patents that teach the preparation of such modified sugar structures include, but are not limited to, U.S.: 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633; 5,792,747; and 5,700,920, certain of which are commonly owned with the instant application, and each of which is herein incorporated by reference in its entirety.

Another modification of the sugar includes Locked Nucleic Acids (LNAs) in which the 2'-hydroxyl group is linked to the 3' or 4' carbon atom of the sugar ring, thereby forming a bicyclic sugar moiety. The linkage is preferably a methylene $(-\text{CH}_2-)_n$ group bridging the 2' oxygen atom and the 4' carbon atom wherein n is 1 or 2. LNAs and preparation thereof are described in WO 98/39352 and WO 99/14226.

Natural and Modified Nucleobases

Oligonucleotides may also include nucleobase (often referred to in the art simply as "base") modifications or substitutions. As used herein, "unmodified" or "natural" nucleobases include the purine bases adenine (A) and guanine (G), and the pyrimidine bases thymine (T), cytosine (C) and uracil (U). Modified nucleobases include other synthetic and natural nucleobases such as 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-halouracil and cytosine, 5-propynyl $(-\text{C}\equiv\text{C}-\text{CH}_3)$ uracil and cytosine and other alkynyl derivatives of pyrimidine bases, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 2-F-adenine, 2-amino-adenine, 8-azaguanine and 8-azaadenine, 7-deazaguanine and 7-deazaadenine and 3-deazaguanine and 3-deazaadenine. Additional modified nucleobases include tricyclic pyrimidines such as phenoxazine cytidine(1H-pyrimido[5,4-b][1,4]benzoxazin-2(3H)-one), phenothiazine cytidine (1H-pyrimido[5,4-b][1,4]benzothiazin-2(3H)-one), G-clamps such as a substituted phenoxazine cytidine (e.g. 9-(2-aminoethoxy)-H-pyrimido[5,4-b][1,4]benzoxazin-2(3H)-one), carbazole cytidine (2H-pyrimido[4,5-b]indol-2-one), pyridoindole cytidine (H-pyrido[3',2':4,5]pyrrolo[2,3-d]pyrimidin-2-one). Modified nucleobases may also include those in which the purine or pyrimidine base is replaced with other heterocycles, for example 7-deaza-adenine, 7-deazaguanosine, 2-aminopyridine and 2-pyridone. Additional nucleobases include those disclosed in United States Patent No. 3,687,808, those disclosed in The Concise Encyclopedia Of Polymer Science And Engineering, pages 858-859, Kroschwitz, J.I., ed. John Wiley & Sons, 1990, those disclosed by Englisch et al., Angewandte Chemie, International Edition, 1991, 30, 613, and those disclosed by Sanghvi, Y.S., Chapter 15, Antisense Research and Applications, pages 289-302, Croke, S.T. and Lebleu, B., ed., CRC Press, 1993. Certain of these nucleobases are particularly useful for increasing the binding affinity of the compounds of the invention. These include 5-substituted pyrimidines, 6-azapyrimidines and N-2, N-6 and O-6

substituted purines, including 2-aminopropyladenine, 5-propynyluracil and 5-propynylcytosine. 5-methylcytosine substitutions have been shown to increase nucleic acid duplex stability by 0.6-1.2 °C and are presently suitable base substitutions, even more particularly when combined with 2'-O-methoxyethyl sugar modifications.

Representative United States patents that teach the preparation of certain of the above noted modified nucleobases as well as other modified nucleobases include, but are not limited to, the above noted U.S. 3,687,808, as well as U.S.: 4,845,205; 5,130,302; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,594,121, 5,596,091; 5,614,617; 5,645,985; 5,830,653; 5,763,588; 6,005,096; and 5,681,941, certain of which are commonly owned with the instant application, and each of which is herein incorporated by reference, and United States patent 5,750,692, which is commonly owned with the instant application and also herein incorporated by reference.

Conjugates

Another modification of the oligonucleotides of the invention involves chemically linking to the oligonucleotide one or more moieties or conjugates which enhance the activity, cellular distribution or cellular uptake of the oligonucleotide. These moieties or conjugates can include conjugate groups covalently bound to functional groups such as primary or secondary hydroxyl groups. Conjugate groups of the invention include intercalators, reporter molecules, polyamines, polyamides, polyethylene glycols, polyethers, groups that enhance the pharmacodynamic properties of oligomers, and groups that enhance the pharmacokinetic properties of oligomers. Typical conjugate groups include cholesterols, lipids, phospholipids, biotin, phenazine, folate, phenanthridine, anthraquinone, acridine, fluoresceins, rhodamines, coumarins, and dyes. Groups that enhance the pharmacodynamic properties, in the context of this invention, include groups that improve uptake, enhance resistance to degradation, and/or strengthen sequence-specific hybridization with the target nucleic acid. Groups that enhance the pharmacokinetic properties, in the context of this invention, include groups that improve uptake, distribution, metabolism or excretion of the compounds of the present invention. Representative conjugate groups are disclosed in International Patent Application PCT/US92/09196, filed October 23, 1992, and U.S. Patent 6,287,860, the entire disclosure of which are incorporated herein by reference.

Conjugate moieties include but are not limited to lipid moieties such as a cholesterol moiety, cholic acid, a thioether, e.g., hexyl-S-tritylthiol, a thiocholesterol, an aliphatic chain, e.g., dodecandiol or undecyl residues, a phospholipid, e.g., di-hexadecyl-rac-glycerol or triethyl-ammonium 1,2-di-O-hexadecyl-rac-glycero-3-H-phosphonate, a polyamine or a polyethylene

glycol chain, or adamantane acetic acid, a palmityl moiety, or an octadecylamine or hexylamino-carbonyl-oxycholesterol moiety. Oligonucleotides of the invention may also be conjugated to active drug substances, for example, aspirin, warfarin, phenylbutazone, ibuprofen, suprofen, fenbufen, ketoprofen, (*S*)-(+)-pranoprofen, carprofen, dansylsarcosine, 2,3,5-triiodobenzoic acid, flufenamic acid, folic acid, a benzothiadiazide, chlorothiazide, a diazepam, indomethacin, a barbiturate, a cephalosporin, a sulfa drug, an antidiabetic, an antibacterial or an antibiotic. Oligonucleotide-drug conjugates and their preparation are described in United States Patent Application 09/334,130 (filed June 15, 1999) which is incorporated herein by reference in its entirety.

Representative United States patents that teach the preparation of such oligonucleotide conjugates include, but are not limited to, U.S.: 4,828,979; 4,948,882; 5,218,105; 5,525,465; 5,541,313; 5,545,730; 5,552,538; 5,578,717; 5,580,731; 5,580,731; 5,591,584; 5,109,124; 5,118,802; 5,138,045; 5,414,077; 5,486,603; 5,512,439; 5,578,718; 5,608,046; 4,587,044; 4,605,735; 4,667,025; 4,762,779; 4,789,737; 4,824,941; 4,835,263; 4,876,335; 4,904,582; 4,958,013; 5,082,830; 5,112,963; 5,214,136; 5,082,830; 5,112,963; 5,214,136; 5,245,022; 5,254,469; 5,258,506; 5,262,536; 5,272,250; 5,292,873; 5,317,098; 5,371,241; 5,391,723; 5,416,203; 5,451,463; 5,510,475; 5,512,667; 5,514,785; 5,565,552; 5,567,810; 5,574,142; 5,585,481; 5,587,371; 5,595,726; 5,597,696; 5,599,923; 5,599,928; and 5,688,941, certain of which are commonly owned with the instant application, and each of which is herein incorporated by reference.

Chimeric compounds

It is not necessary for all positions in a given compound to be uniformly modified, and in fact more than one of the aforementioned modifications may be incorporated in a single compound or even at a single nucleoside within an oligonucleotide.

The present invention also includes antisense compounds that are chimeric compounds. "Chimeric" antisense compounds or "chimeras," in the context of this invention, are antisense compounds, particularly oligonucleotides, which contain two or more chemically distinct regions, each made up of at least one monomer unit, i.e., a nucleotide in the case of an oligonucleotide compound. These oligonucleotides typically contain at least one region wherein the oligonucleotide is modified so as to confer upon the oligonucleotide increased resistance to nuclease degradation, increased cellular uptake, increased stability and/or increased binding affinity for the target nucleic acid. An additional region of the oligonucleotide may serve as a substrate for enzymes capable of cleaving RNA:DNA or RNA:RNA hybrids. By way of example, RNase H is a cellular endonuclease which cleaves the RNA strand of an RNA:DNA

duplex. Activation of RNase H, therefore, results in cleavage of the RNA target, thereby greatly enhancing the efficiency of oligonucleotide-mediated inhibition of gene expression. The cleavage of RNA:RNA hybrids can, in like fashion, be accomplished through the actions of endoribonucleases, such as RNaseL which cleaves both cellular and viral RNA. Cleavage of the RNA target can be routinely detected by gel electrophoresis and, if necessary, associated nucleic acid hybridization techniques known in the art.

Chimeric antisense compounds of the invention may be formed as composite structures of two or more oligonucleotides, modified oligonucleotides, oligonucleosides and/or oligonucleotide mimetics as described above. Such compounds have also been referred to in the art as hybrids or gapmers. Representative United States patents that teach the preparation of such hybrid structures include, but are not limited to, U.S.: 5,013,830; 5,149,797; 5,220,007; 5,256,775; 5,366,878; 5,403,711; 5,491,133; 5,565,350; 5,623,065; 5,652,355; 5,652,356; and 5,700,922, certain of which are commonly owned with the instant application, and each of which is herein incorporated by reference in its entirety.

The compounds of the invention may also be admixed, encapsulated, conjugated or otherwise associated with other molecules, molecule structures or mixtures of compounds, as for example, liposomes, receptor-targeted molecules, oral, rectal, topical or other formulations, for assisting in uptake, distribution and/or absorption. Representative United States patents that teach the preparation of such uptake, distribution and/or absorption-assisting formulations include, but are not limited to, U.S.: 5,108,921; 5,354,844; 5,416,016; 5,459,127; 5,521,291; 5,543,158; 5,547,932; 5,583,020; 5,591,721; 4,426,330; 4,534,899; 5,013,556; 5,108,921; 5,213,804; 5,227,170; 5,264,221; 5,356,633; 5,395,619; 5,416,016; 5,417,978; 5,462,854; 5,469,854; 5,512,295; 5,527,528; 5,534,259; 5,543,152; 5,556,948; 5,580,575; and 5,595,756, each of which is herein incorporated by reference.

The compounds of the invention encompass any pharmaceutically acceptable salts, esters, or salts of such esters, or any other compound which, upon administration to an animal, including a human, is capable of providing (directly or indirectly) the biologically active metabolite or residue thereof.

The term "pharmaceutically acceptable salts" refers to physiologically and pharmaceutically acceptable salts of the compounds of the invention: i.e., salts that retain the desired biological activity of the parent compound and do not impart undesired toxicological effects thereto. For oligonucleotides, suitable examples of pharmaceutically acceptable salts and their uses are further described in U.S. Patent 6,287,860, which is incorporated herein in its entirety.

The present invention also includes pharmaceutical compositions and formulations that include the compounds of the invention. The pharmaceutical compositions of the present invention may be administered in a number of ways depending upon whether local or systemic treatment is desired and upon the area to be treated. Administration may be topical (including ophthalmic and to mucous membranes including vaginal and rectal delivery), pulmonary, e.g., by inhalation or insufflation of powders or aerosols, including by nebulizer; intratracheal, intranasal, epidermal and transdermal), oral or parenteral. Parenteral administration includes intravenous, intraarterial, subcutaneous, intraperitoneal or intramuscular injection or infusion; or intracranial, e.g., intrathecal or intraventricular, administration. Oligonucleotides with at least one 2'-O-methoxyethyl modification are believed to be particularly useful for oral administration. Pharmaceutical compositions and formulations for topical administration may include transdermal patches, ointments, lotions, creams, gels, drops, suppositories, sprays, liquids and powders. Conventional pharmaceutical carriers, aqueous, powder or oily bases, thickeners and the like may be necessary or desirable. Coated condoms, gloves and the like may also be useful.

The pharmaceutical formulations of the present invention, which may conveniently be presented in unit dosage form, may be prepared according to conventional techniques well known in the pharmaceutical industry. Such techniques include the step of bringing into association the active ingredients with the pharmaceutical carrier(s) or excipient(s). In general, the formulations are prepared by uniformly and intimately bringing into association the active ingredients with liquid carriers or finely divided solid carriers or both, and then, if necessary, shaping the product.

The compositions of the present invention may be formulated into any of many possible dosage forms such as, but not limited to, tablets, capsules, gel capsules, liquid syrups, soft gels, suppositories, and enemas. The compositions of the present invention may also be formulated as suspensions in aqueous, non-aqueous or mixed media. Aqueous suspensions may further contain substances that increase the viscosity of the suspension including, for example, sodium carboxymethylcellulose, sorbitol and/or dextran. The suspension may also contain stabilizers.

Pharmaceutical compositions of the present invention include, but are not limited to, solutions, emulsions, foams and liposome-containing formulations. The pharmaceutical compositions and formulations of the present invention may comprise one or more penetration enhancers, carriers, excipients or other active or inactive ingredients.

Emulsions are typically heterogenous systems of one liquid dispersed in another in the form of droplets usually exceeding 0.1 μm in diameter. Emulsions may contain additional components in addition to the dispersed phases, and the active drug that may be present as a

solution in either the aqueous phase, oily phase or itself as a separate phase. Microemulsions are included as an embodiment of the present invention. Emulsions and their uses are well known in the art and are further described in U.S. Patent 6,287,860, which is incorporated herein in its entirety.

Formulations of the present invention include liposomal formulations. As used in the present invention, the term "liposome" means a vesicle composed of amphiphilic lipids arranged in a spherical bilayer or bilayers. Liposomes are unilamellar or multilamellar vesicles which have a membrane formed from a lipophilic material and an aqueous interior that contains the composition to be delivered. Cationic liposomes are positively charged liposomes that are believed to interact with negatively charged DNA molecules to form a stable complex. Liposomes that are pH-sensitive or negatively-charged are believed to entrap DNA rather than complex with it. Both cationic and noncationic liposomes have been used to deliver DNA to cells.

Liposomes also include "sterically stabilized" liposomes, a term which, as used herein, refers to liposomes comprising one or more specialized lipids that, when incorporated into liposomes, result in enhanced circulation lifetimes relative to liposomes lacking such specialized lipids. Examples of sterically stabilized liposomes are those in which part of the vesicle-forming lipid portion of the liposome comprises one or more glycolipids or is derivatized with one or more hydrophilic polymers, such as a polyethylene glycol (PEG) moiety. Liposomes and their uses are further described in U.S. Patent 6,287,860, which is incorporated herein in its entirety.

The pharmaceutical formulations and compositions of the present invention may also include surfactants. The use of surfactants in drug products, formulations and in emulsions is well known in the art. Surfactants and their uses are further described in U.S. Patent 6,287,860, which is incorporated herein in its entirety.

In one embodiment, the present invention employs various penetration enhancers to affect the efficient delivery of nucleic acids, particularly oligonucleotides. In addition to aiding the diffusion of non-lipophilic drugs across cell membranes, penetration enhancers also enhance the permeability of lipophilic drugs. Penetration enhancers may be classified as belonging to one of five broad categories, *i.e.*, surfactants, fatty acids, bile salts, chelating agents, and non-chelating non-surfactants. Penetration enhancers and their uses are further described in U.S. Patent 6,287,860, which is incorporated herein in its entirety.

One of skill in the art will recognize that formulations are routinely designed according to their intended use, *i.e.* route of administration.

Formulations for topical administration include those in which the oligonucleotides of the invention are in admixture with a topical delivery agent such as lipids, liposomes, fatty acids, fatty acid esters, steroids, chelating agents and surfactants. Suitable lipids and liposomes include neutral (e.g. dioleoylphosphatidyl DOPE ethanolamine, dimyristoylphosphatidyl choline DMPC, distearoylphosphatidyl choline) negative (e.g. dimyristoylphosphatidyl glycerol DMPG) and cationic (e.g. dioleoyltetramethylaminopropyl DOTAP and dioleoylphosphatidyl ethanolamine DOTMA).

For topical or other administration, oligonucleotides of the invention may be encapsulated within liposomes or may form complexes thereto, in particular to cationic liposomes. Alternatively, oligonucleotides may be complexed to lipids, in particular to cationic lipids. Fatty acids and esters, pharmaceutically acceptable salts thereof, and their uses are further described in U.S. Patent 6,287,860, which is incorporated herein in its entirety. Topical formulations are described in detail in United States patent application 09/315,298 filed on May 20, 1999, which is incorporated herein by reference in its entirety.

Compositions and formulations for oral administration include powders or granules, microparticulates, nanoparticulates, suspensions or solutions in water or non-aqueous media, capsules, gel capsules, sachets, tablets or minitables. Thickeners, flavoring agents, diluents, emulsifiers, dispersing aids or binders may be desirable. Oral formulations are those in which oligonucleotides of the invention are administered in conjunction with one or more penetration enhancers surfactants and chelators. Surfactants include fatty acids and/or esters or salts thereof, bile acids and/or salts thereof. Bile acids/salts and fatty acids and their uses are further described in U.S. Patent 6,287,860, which is incorporated herein in its entirety. Also suitable are combinations of penetration enhancers, for example, fatty acids/salts in combination with bile acids/salts. A particularly suitable combination is the sodium salt of lauric acid, capric acid and UDCA. Further penetration enhancers include polyoxyethylene-9-lauryl ether, polyoxyethylene-20-cetyl ether. Oligonucleotides of the invention may be delivered orally, in granular form including sprayed dried particles, or complexed to form micro or nanoparticles. Oligonucleotide complexing agents and their uses are further described in U.S. Patent 6,287,860, which is incorporated herein in its entirety. Oral formulations for oligonucleotides and their preparation are described in detail in United States applications 09/108,673 (filed July 1, 1998), 09/315,298 (filed May 20, 1999) and 10/071,822, filed February 8, 2002, each of which is incorporated herein by reference in their entirety.

Compositions and formulations for parenteral, intrathecal or intraventricular administration may include sterile aqueous solutions which may also contain buffers, diluents

and other suitable additives such as, but not limited to, penetration enhancers, carrier compounds and other pharmaceutically acceptable carriers or excipients.

Certain embodiments of the invention provide pharmaceutical compositions containing one or more oligomeric compounds and one or more other chemotherapeutic agents that function by a non-antisense mechanism. Examples of such chemotherapeutic agents include, but are not limited to, cancer chemotherapeutic drugs such as daunorubicin, daunomycin, dactinomycin, doxorubicin, epirubicin, idarubicin, esorubicin, bleomycin, mafosfamide, ifosfamide, cytosine arabinoside, bis-chloroethylnitrosurea, busulfan, mitomycin C, actinomycin D, mithramycin, prednisone, hydroxyprogesterone, testosterone, tamoxifen, dacarbazine, procarbazine, hexamethylmelamine, pentamethylmelamine, mitoxantrone, amsacrine, chlorambucil, methylcyclohexylnitrosurea, nitrogen mustards, melphalan, cyclophosphamide, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-azacytidine, hydroxyurea, deoxycoformycin, 4-hydroxyperoxycyclophosphoramide, 5-fluorouracil (5-FU), 5-fluorodeoxyuridine (5-FUdR), methotrexate (MTX), colchicine, taxol, vincristine, vinblastine, etoposide (VP-16), trimetrexate, irinotecan, topotecan, gemcitabine, teniposide, cisplatin and diethylstilbestrol (DES). When used with the compounds of the invention, such chemotherapeutic agents may be used individually (*e.g.*, 5-FU and oligonucleotide), sequentially (*e.g.*, 5-FU and oligonucleotide for a period of time followed by MTX and oligonucleotide), or in combination with one or more other such chemotherapeutic agents (*e.g.*, 5-FU, MTX and oligonucleotide, or 5-FU, radiotherapy and oligonucleotide). Anti-inflammatory drugs, including but not limited to nonsteroidal anti-inflammatory drugs and corticosteroids, and antiviral drugs, including but not limited to ribivirin, vidarabine, acyclovir and ganciclovir, may also be combined in compositions of the invention. Combinations of antisense compounds and other non-antisense drugs are also within the scope of this invention. Two or more combined compounds may be used together or sequentially.

In another related embodiment, compositions of the invention may contain one or more antisense compounds, particularly oligonucleotides, targeted to a first nucleic acid and one or more additional compounds targeted to a second nucleic acid target. Alternatively, compositions of the invention may contain two or more compounds targeted to different regions of the same nucleic acid target. Numerous examples of compounds are known in the art. Two or more combined compounds may be used together or sequentially.

The formulation of therapeutic compositions and their subsequent administration (dosing) is believed to be within the skill of those in the art. Dosing is dependent on severity and responsiveness of the disease state to be treated, with the course of treatment lasting from several days to several months, or until a cure is effected or a diminution of the disease state is achieved.

Optimal dosing schedules can be calculated from measurements of drug accumulation in the body of the patient. Persons of ordinary skill can easily determine optimum dosages, dosing methodologies and repetition rates. Optimum dosages may vary depending on the relative potency of individual oligonucleotides, and can generally be estimated based on EC₅₀s found to be effective in *in vitro* and *in vivo* animal models. In general, dosage is from 0.01 µg to 100 g per kg of body weight, and may be given once or more daily, weekly, monthly or yearly, or even once every 2 to 20 years. Persons of ordinary skill in the art can easily estimate repetition rates for dosing based on measured residence times and concentrations of the drug in bodily fluids or tissues. Following successful treatment, it may be desirable to have the patient undergo maintenance therapy to prevent the recurrence of the disease state, wherein the oligonucleotide is administered in maintenance doses, ranging from 0.01 µg to 100 g per kg of body weight, once or more daily, to once every 20 years.

While the present invention has been described with specificity in accordance with certain of its embodiments, the following examples serve only to illustrate the invention and are not intended to limit the same. Each of the references, patents, international publications, GenBank accession numbers, and the like recited in the present application are incorporated herein by reference in its entirety.

EXAMPLES

Example 1: Synthesis of Nucleoside Phosphoramidites

The following compounds, including amidites and their intermediates were prepared as described in US Patent 6,426,220 and published PCT WO 02/36743; 5'-O-Dimethoxytrityl-thymidine intermediate for 5-methyl dC amidite, 5'-O-Dimethoxytrityl-2'-deoxy-5-methylcytidine intermediate for 5-methyl-dC amidite, 5'-O-Dimethoxytrityl-2'-deoxy-N⁴-benzoyl-5-methylcytidine penultimate intermediate for 5-methyl dC amidite, [5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-deoxy-N⁴-benzoyl-5-methylcytidin-3'-O-yl]-2-cyanoethyl-*N,N*-diisopropylphosphoramidite (5-methyl dC amidite), 2'-Fluorodeoxyadenosine, 2'-Fluorodeoxyguanosine, 2'-Fluorouridine, 2'-Fluorodeoxycytidine, 2'-O-(2-Methoxyethyl) modified amidites, 2'-O-(2-methoxyethyl)-5-methyluridine intermediate, 5'-O-DMT-2'-O-(2-methoxyethyl)-5-methyluridine penultimate intermediate, [5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-methoxyethyl)-5-methyluridin-3'-O-yl]-2-cyanoethyl-*N,N*-diisopropylphosphoramidite (MOE T amidite), 5'-O-Dimethoxytrityl-2'-O-(2-methoxyethyl)-5-methylcytidine intermediate, 5'-O-dimethoxytrityl-2'-O-(2-methoxyethyl)-N⁴-benzoyl-5-methylcytidine penultimate intermediate, [5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-

-28-

methoxyethyl)-N⁴-benzoyl-5-methylcytidin-3'-O-yl]-2-cyanoethyl-N,N-diisopropylphosphoramidite (MOE 5-Me-C amidite), [5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-methoxyethyl)-N⁶-benzoyladenodin-3'-O-yl]-2-cyanoethyl-N,N-diisopropylphosphoramidite (MOE A amidite), [5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-methoxyethyl)-N⁴-isobutyrylguanosin-3'-O-yl]-2-cyanoethyl-N,N-diisopropylphosphoramidite (MOE G amidite), 2'-O-(Aminooxyethyl) nucleoside amidites and 2'-O-(dimethylaminooxyethyl) nucleoside amidites, 2'-(Dimethylaminooxyethoxy) nucleoside amidites, 5'-O-tert-Butyldiphenylsilyl-O²-2'-anhydro-5-methyluridine, 5'-O-tert-Butyldiphenylsilyl-2'-O-(2-hydroxyethyl)-5-methyluridine, 2'-O-([2-phthalimidoxy)ethyl]-5'-*t*-butyldiphenylsilyl-5-methyluridine, 5'-O-*tert*-butyldiphenylsilyl-2'-O-[(2-formadoximinoxy)ethyl]-5-methyluridine, 5'-O-*tert*-Butyldiphenylsilyl-2'-O-[N,N dimethylaminoxyethyl]-5-methyluridine, 2'-O-(dimethylaminoxyethyl)-5-methyluridine, 5'-O-DMT-2'-O-(dimethylaminoxyethyl)-5-methyluridine, 5'-O-DMT-2'-O-(2-N,N-dimethylaminoxyethyl)-5-methyluridine-3'-[(2-cyanoethyl)-N,N-diisopropylphosphoramidite], 2'-(Aminooxyethoxy) nucleoside amidites, N²-isobutyryl-6-O-diphenylcarbamoyl-2'-O-(2-ethylacetyl)-5'-O-(4,4'-dimethoxytrityl)guanosine-3'-[(2-cyanoethyl)-N,N-diisopropylphosphoramidite], 2'-dimethylaminoethoxyethoxy (2'-DMAEOE) nucleoside amidites, 2'-O-[2(2-N,N-dimethylaminoethoxy)ethyl]-5-methyl uridine, 5'-O-dimethoxytrityl-2'-O-[2(2-N,N-dimethylaminoethoxy)-ethyl]-5-methyl uridine and 5'-O-Dimethoxytrityl-2'-O-[2(2-N,N-dimethylaminoethoxy)-ethyl]-5-methyl uridine-3'-O-(cyanoethyl-N,N-diisopropyl)phosphoramidite.

Example 2: Oligonucleotide and oligonucleoside synthesis

The antisense compounds used in accordance with this invention may be conveniently and routinely made through the well-known technique of solid phase synthesis. Equipment for such synthesis is sold by several vendors including, for example, Applied Biosystems (Foster City, CA). Any other means for such synthesis known in the art may additionally or alternatively be employed. It is well known to use similar techniques to prepare oligonucleotides such as the phosphorothioates and alkylated derivatives.

Oligonucleotides: Unsubstituted and substituted phosphodiester (P=O) oligonucleotides are synthesized on an automated DNA synthesizer (Applied Biosystems model 394) using standard phosphoramidite chemistry with oxidation by iodine.

Phosphorothioates (P=S) are synthesized similar to phosphodiester oligonucleotides with the following exceptions: thiation was effected by utilizing a 10% w/v solution of 3, H-1,2-benzodithiole-3-one 1,1-dioxide in acetonitrile for the oxidation of the phosphite linkages. The

thiation reaction step time was increased to 180 sec and preceded by the normal capping step. After cleavage from the CPG column and deblocking in concentrated ammonium hydroxide at 55°C (12-16 hr), the oligonucleotides were recovered by precipitating with >3 volumes of ethanol from a 1 M NH₄OAc solution. Phosphinate oligonucleotides are prepared as described in U.S. Patent 5,508,270, herein incorporated by reference.

Alkyl phosphonate oligonucleotides are prepared as described in U.S. Patent 4,469,863, herein incorporated by reference.

3'-Deoxy-3'-methylene phosphonate oligonucleotides are prepared as described in U.S. Patents 5,610,289 or 5,625,050, herein incorporated by reference.

Phosphoramidite oligonucleotides are prepared as described in U.S. Patent, 5,256,775 or U.S. Patent 5,366,878, herein incorporated by reference.

Alkylphosphonothioate oligonucleotides are prepared as described in published PCT applications PCT/US94/00902 and PCT/US93/06976 (published as WO 94/17093 and WO 94/02499, respectively), herein incorporated by reference.

3'-Deoxy-3'-amino phosphoramidate oligonucleotides are prepared as described in U.S. Patent 5,476,925, herein incorporated by reference.

Phosphotriester oligonucleotides are prepared as described in U.S. Patent 5,023,243, herein incorporated by reference.

Borano phosphate oligonucleotides are prepared as described in U.S. Patents 5,130,302 and 5,177,198, both herein incorporated by reference.

Oligonucleosides: Methylenemethylimino linked oligonucleosides, also identified as MMI linked oligonucleosides, methylenedimethylhydrazo linked oligonucleosides, also identified as MDH linked oligonucleosides, and methylenecarbonylamino linked oligonucleosides, also identified as amide-3 linked oligonucleosides, and methyleneaminocarbonyl linked oligonucleosides, also identified as amide-4 linked oligonucleosides, as well as mixed backbone compounds having, for instance, alternating MMI and P=O or P=S linkages are prepared as described in U.S. Patents 5,378,825, 5,386,023, 5,489,677, 5,602,240 and 5,610,289, all of which are herein incorporated by reference.

Formacetal and thioformacetal linked oligonucleosides are prepared as described in U.S. Patents 5,264,562 and 5,264,564, herein incorporated by reference.

Ethylene oxide linked oligonucleosides are prepared as described in U.S. Patent 5,223,618, herein incorporated by reference.

Example 3: RNA Synthesis

In general, RNA synthesis chemistry is based on the selective incorporation of various protecting groups at strategic intermediary reactions. Although one of ordinary skill in the art will understand the use of protecting groups in organic synthesis, a useful class of protecting groups includes silyl ethers. In particular bulky silyl ethers are used to protect the 5'-hydroxyl in combination with an acid-labile orthoester protecting group on the 2'-hydroxyl. This set of protecting groups is then used with standard solid-phase synthesis technology. It is important to lastly remove the acid labile orthoester protecting group after all other synthetic steps. Moreover, the early use of the silyl protecting groups during synthesis ensures facile removal when desired, without undesired deprotection of 2' hydroxyl.

Following this procedure for the sequential protection of the 5'-hydroxyl in combination with protection of the 2'-hydroxyl by protecting groups that are differentially removed and are differentially chemically labile, RNA oligonucleotides were synthesized.

RNA oligonucleotides are synthesized in a stepwise fashion. Each nucleotide is added sequentially (3' - to 5'-direction) to a solid support-bound oligonucleotide. The first nucleoside at the 3'-end of the chain is covalently attached to a solid support. The nucleotide precursor, a ribonucleoside phosphoramidite, and activator are added, coupling the second base onto the 5'-end of the first nucleoside. The support is washed and any unreacted 5'-hydroxyl groups are capped with acetic anhydride to yield 5'-acetyl moieties. The linkage is then oxidized to the more stable and ultimately desired P(V) linkage. At the end of the nucleotide addition cycle, the 5'-silyl group is cleaved with fluoride. The cycle is repeated for each subsequent nucleotide.

Following synthesis, the methyl protecting groups on the phosphates are cleaved in 30 minutes utilizing 1 M disodium-2-carbamoyl-2-cyanoethylene-1,1-dithiolate trihydrate (S_2Na_2) in DMF. The deprotection solution is washed from the solid support-bound oligonucleotide using water. The support is then treated with 40% methylamine in water for 10 minutes at 55 °C. This releases the RNA oligonucleotides into solution, deprotects the exocyclic amines, and modifies the 2'- groups. The oligonucleotides can be analyzed by anion exchange HPLC at this stage.

The 2'-orthoester groups are the last protecting groups to be removed. The ethylene glycol monoacetate orthoester protecting group developed by Dharmacon Research, Inc. (Lafayette, CO), is one example of a useful orthoester protecting group which, has the following important properties. It is stable to the conditions of nucleoside phosphoramidite synthesis and oligonucleotide synthesis. However, after oligonucleotide synthesis the oligonucleotide is treated with methylamine that not only cleaves the oligonucleotide from the solid support but also removes the acetyl groups from the orthoesters. The resulting 2-ethyl-hydroxyl substituents on

the orthoester are less electron withdrawing than the acetylated precursor. As a result, the modified orthoester becomes more labile to acid-catalyzed hydrolysis. Specifically, the rate of cleavage is approximately 10 times faster after the acetyl groups are removed. Therefore, this orthoester possesses sufficient stability in order to be compatible with oligonucleotide synthesis and yet, when subsequently modified, permits deprotection to be carried out under relatively mild aqueous conditions compatible with the final RNA oligonucleotide product.

Additionally, methods of RNA synthesis are well known in the art (Scaringe, S. A. Ph.D. Thesis, University of Colorado, 1996; Scaringe, S. A., et al., *J. Am. Chem. Soc.*, 1998, 120, 11820-11821; Matteucci, M. D. and Caruthers, M. H. *J. Am. Chem. Soc.*, 1981, 103, 3185-3191; Beaucage, S. L. and Caruthers, M. H. *Tetrahedron Lett.*, 1981, 22, 1859-1862; Dahl, B. J., et al., *Acta Chem. Scand.*, 1990, 44, 639-641; Reddy, M. P., et al., *Tetrahedron Lett.*, 1994, 25, 4311-4314; Wincott, F. et al., *Nucleic Acids Res.*, 1995, 23, 2677-2684; Griffin, B. E., et al., *Tetrahedron*, 1967, 23, 2301-2313; and Griffin, B. E., et al., *Tetrahedron*, 1967, 23, 2315-2331).

RNA antisense compounds (RNA oligonucleotides) of the present invention can be synthesized by the methods herein or purchased from Dharmacon Research, Inc (Lafayette, CO). Once synthesized, complementary RNA antisense compounds can then be annealed by methods known in the art to form double stranded (duplexed) antisense compounds. For example, duplexes can be formed by combining 30 μ l of each of the complementary strands of RNA oligonucleotides (50 μ M RNA oligonucleotide solution) and 15 μ l of 5X annealing buffer (100 mM potassium acetate, 30 mM HEPES-KOH pH 7.4, 2 mM magnesium acetate) followed by heating for 1 minute at 90°C, then 1 hour at 37°C. The resulting duplexed antisense compounds can be used in kits, assays, screens, or other methods to investigate the role of a target nucleic acid.

Example 4: Synthesis of Chimeric Oligonucleotides

Chimeric oligonucleotides, oligonucleosides or mixed oligonucleotides/oligonucleosides of the invention can be of several different types. These include a first type wherein the "gap" segment of linked nucleosides is positioned between 5' and 3' "wing" segments of linked nucleosides and a second "open end" type wherein the "gap" segment is located at either the 3' or the 5' terminus of the oligomeric compound. Oligonucleotides of the first type are also known in the art as "gapmers" or gapped oligonucleotides. Oligonucleotides of the second type are also known in the art as "hemimers" or "wingmers."

-32-

[2'-O-Me]--[2'-deoxy]--[2'-O-Me] Chimeric Phosphorothioate Oligonucleotides

Chimeric oligonucleotides having 2'-O-alkyl phosphorothioate and 2'-deoxy phosphorothioate oligonucleotide segments are synthesized using an Applied Biosystems automated DNA synthesizer Model 394, as above. Oligonucleotides are synthesized using the automated synthesizer and 2'-deoxy-5'-dimethoxytrityl-3'-O-phosphoramidite for the DNA portion and 5'-dimethoxytrityl-2'-O-methyl-3'-O-phosphoramidite for 5' and 3' wings. The standard synthesis cycle is modified by incorporating coupling steps with increased reaction times for the 5'-dimethoxytrityl-2'-O-methyl-3'-O-phosphoramidite. The fully protected oligonucleotide is cleaved from the support and deprotected in concentrated ammonia (NH₄OH) for 12-16 hr at 55°C. The deprotected oligo is then recovered by an appropriate method (precipitation, column chromatography, volume reduced *in vacuo* and analyzed spectrophotometrically for yield and for purity by capillary electrophoresis and by mass spectrometry).

[2'-O-(2-Methoxyethyl)]--[2'-deoxy]--[2'-O-(Methoxyethyl)] Chimeric Phosphorothioate Oligonucleotides

[2'-O-(2-methoxyethyl)]--[2'-deoxy]--[2'-O-(methoxyethyl)] chimeric phosphorothioate oligonucleotides were prepared as per the procedure above for the 2'-O-methyl chimeric oligonucleotide, with the substitution of 2'-O-(methoxyethyl) amidites for the 2'-O-methyl amidites.

[2'-O-(2-Methoxyethyl)Phosphodiester]--[2'-deoxy Phosphorothioate]--[2'-O-(2-Methoxyethyl) Phosphodiester] Chimeric Oligonucleotides

[2'-O-(2-methoxyethyl phosphodiester)]--[2'-deoxy phosphorothioate]--[2'-O-(methoxyethyl) phosphodiester] chimeric oligonucleotides are prepared as per the above procedure for the 2'-O-methyl chimeric oligonucleotide with the substitution of 2'-O-(methoxyethyl) amidites for the 2'-O-methyl amidites, oxidation with iodine to generate the phosphodiester internucleotide linkages within the wing portions of the chimeric structures and sulfurization utilizing 3,4-dihydro-2H-benzothiole-3-one 1,1-dioxide (Beaucage Reagent) to generate the phosphorothioate internucleotide linkages for the center gap.

Other chimeric oligonucleotides, chimeric oligonucleosides and mixed chimeric oligonucleotides/oligonucleosides are synthesized according to United States patent 5,623,065, herein incorporated by reference.

Example 5: Design and screening of duplexed antisense compounds targeting Huntingtin Interacting Protein 1

In accordance with the present invention, a series of nucleic acid duplexes comprising the antisense compounds of the present invention and their complements can be designed to target huntingtin interacting protein 1. The nucleobase sequence of the antisense strand of the duplex comprises at least an 8-nucleobase portion of an oligonucleotide in Table 1. The ends of the strands may be modified by the addition of one or more natural or modified nucleobases to form an overhang. The sense strand of the dsRNA is then designed and synthesized as the complement of the antisense strand and may also contain modifications or additions to either terminus. For example, in one embodiment, both strands of the dsRNA duplex would be complementary over the central nucleobases, each having overhangs at one or both termini.

For example, a duplex comprising an antisense strand having the sequence CGAGAGGCGGACGGACCG (SEQ ID NO:166) and having a two-nucleobase overhang of deoxythymidine(dT) would have the following structure:

cgagagggcggacgggaccgTT	(SEQ ID NO:167)	Antisense Strand
TTgctctccgcctgccctggc	(SEQ ID NO:168)	Complement

RNA strands of the duplex can be synthesized by methods disclosed herein or purchased from Dharmacon Research Inc., (Lafayette, CO). Once synthesized, the complementary strands are annealed. The single strands are aliquoted and diluted to a concentration of 50 μ M. Once diluted, 30 μ L of each strand is combined with 15 μ L of a 5X solution of annealing buffer. The final concentration of said buffer is 100 mM potassium acetate, 30 mM HEPES-KOH pH 7.4, and 2 mM magnesium acetate. The final volume is 75 μ L. This solution is incubated for 1 minute at 90°C and then centrifuged for 15 seconds. The tube is allowed to sit for 1 hour at 37°C at which time the dsRNA duplexes are used in experimentation. The final concentration of the dsRNA duplex is 20 μ M. This solution can be stored frozen (at, for example, -20°C) and freeze-thawed up to 5 times.

Once prepared, the duplexed antisense compounds are evaluated for their ability to modulate huntingtin interacting protein 1 expression.

When cells reached 80% confluency, they are treated with duplexed antisense compounds of the invention. For cells grown in 96-well plates, wells are washed once with 200 μ L OPTI-MEM-1 reduced-serum medium (Gibco BRL) and then treated with 130 μ L of OPTI-MEM-1 containing 12 μ g/mL LIPOFECTIN (Gibco BRL) and the desired duplex antisense compound at a final concentration of 200 nM. After 5 hours of treatment, the medium is replaced with fresh medium. Cells are harvested 16 hours after treatment, at which time RNA is isolated and target reduction measured by RT-PCR.

Example 6: Oligonucleotide Isolation

After cleavage from the controlled pore glass solid support and deblocking in concentrated ammonium hydroxide at 55°C for 12-16 hours, the oligonucleotides or oligonucleosides are recovered by precipitation out of 1 M NH₄OAc with >3 volumes of ethanol. Synthesized oligonucleotides were analyzed by electrospray mass spectroscopy (molecular weight determination) and by capillary gel electrophoresis and judged to be at least 70% full length material. The relative amounts of phosphorothioate and phosphodiester linkages obtained in the synthesis was determined by the ratio of correct molecular weight relative to the -16 amu product (+/-32 +/-48). For some studies oligonucleotides were purified by HPLC, as described by Chiang et al., J. Biol. Chem. 1991, 266, 18162-18171. Results obtained with HPLC-purified material were similar to those obtained with non-HPLC purified material.

Example 7: Oligonucleotide Synthesis - 96 Well Plate Format

Oligonucleotides were synthesized via solid phase P(III) phosphoramidite chemistry on an automated synthesizer capable of assembling 96 sequences simultaneously in a 96-well format. Phosphodiester internucleotide linkages were afforded by oxidation with aqueous iodine. Phosphorothioate internucleotide linkages were generated by sulfurization utilizing 3,4-dihydro-2H-benzothiole-3-one 1,1-dioxide (Beaucage Reagent) in anhydrous acetonitrile. Standard base-protected beta-cyanoethyl-diisopropyl phosphoramidites were purchased from commercial vendors (e.g. PE-Applied Biosystems, Foster City, CA, or Pharmacia, Piscataway, NJ). Non-standard nucleosides are synthesized as per standard or patented methods. They are utilized as base protected beta-cyanoethyl-diisopropyl phosphoramidites.

Oligonucleotides were cleaved from support and deprotected with concentrated NH₄OH at elevated temperature (55-60°C) for 12-16 hours and the released product then dried *in vacuo*. The dried product was then re-suspended in sterile water to afford a master plate from which all analytical and test plate samples are then diluted utilizing robotic pipettors.

Example 8: Oligonucleotide Analysis – 96-Well Plate Format

The concentration of oligonucleotide in each well was assessed by dilution of samples and UV absorption spectroscopy. The full-length integrity of the individual products was evaluated by capillary electrophoresis (CE) in either the 96-well format (Beckman P/ACE™ MDQ) or, for individually prepared samples, on a commercial CE apparatus (e.g., Beckman P/ACE™ 5000, ABI 270). Base and backbone composition was confirmed by mass analysis of the compounds utilizing electrospray-mass spectroscopy. All assay test plates were diluted from

the master plate using single and multi-channel robotic pipettors. Plates were judged to be acceptable if at least 85% of the compounds on the plate were at least 85% full length.

Example 9: Cell culture and oligonucleotide treatment

The effect of antisense compounds on target nucleic acid expression can be tested in any of a variety of cell types provided that the target nucleic acid is present at measurable levels. This can be routinely determined using, for example, PCR or Northern blot analysis. The following cell types are provided for illustrative purposes, but other cell types can be routinely used, provided that the target is expressed in the cell type chosen. This can be readily determined by methods routine in the art, for example Northern blot analysis, ribonuclease protection assays, or RT-PCR.

T-24 cells: The human transitional cell bladder carcinoma cell line T-24 was obtained from the American Type Culture Collection (ATCC) (Manassas, VA). T-24 cells were routinely cultured in complete McCoy's 5A basal media (Invitrogen Corporation, Carlsbad, CA) supplemented with 10% fetal calf serum (Invitrogen Corporation, Carlsbad, CA), penicillin 100 units per mL, and streptomycin 100 micrograms per mL (Invitrogen Corporation, Carlsbad, CA). Cells were routinely passaged by trypsinization and dilution when they reached 90% confluence. Cells were seeded into 96-well plates (Falcon-Primaria #353872) at a density of 7000 cells/well for use in RT-PCR analysis.

For Northern blotting or other analysis, cells may be seeded onto 100 mm or other standard tissue culture plates and treated similarly, using appropriate volumes of medium and oligonucleotide.

A549 cells: The human lung carcinoma cell line A549 was obtained from the American Type Culture Collection (ATCC) (Manassas, VA). A549 cells were routinely cultured in DMEM basal media (Invitrogen Corporation, Carlsbad, CA) supplemented with 10% fetal calf serum (Invitrogen Corporation, Carlsbad, CA), penicillin 100 units per mL, and streptomycin 100 micrograms per mL (Invitrogen Corporation, Carlsbad, CA). Cells were routinely passaged by trypsinization and dilution when they reached 90% confluence.

NHDF cells: Human neonatal dermal fibroblast (NHDF) were obtained from the Clonetics Corporation (Walkersville, MD). NHDFs were routinely maintained in Fibroblast Growth Medium (Clonetics Corporation, Walkersville, MD) supplemented as recommended by the supplier. Cells were maintained for up to 10 passages as recommended by the supplier.

HEK cells: Human embryonic keratinocytes (HEK) were obtained from the Clonetics Corporation (Walkersville, MD). HEKs were routinely maintained in Keratinocyte Growth

Medium (Clonetics Corporation, Walkersville, MD) formulated as recommended by the supplier. Cells were routinely maintained for up to 10 passages as recommended by the supplier.

Treatment with antisense compounds: When cells reached 65-75% confluency, they were treated with oligonucleotide. For cells grown in 96-well plates, wells were washed once with 100 μ L OPTI-MEM™-1 reduced-serum medium (Invitrogen Corporation, Carlsbad, CA) and then treated with 130 μ L of OPTI-MEM™-1 containing 3.75 μ g/mL LIPOFECTIN™ (Invitrogen Corporation, Carlsbad, CA) and the desired concentration of oligonucleotide. Cells are treated and data are obtained in triplicate. After 4-7 hours of treatment at 37°C, the medium was replaced with fresh medium. Cells were harvested 16-24 hours after oligonucleotide treatment.

The concentration of oligonucleotide used varies from cell line to cell line. To determine the optimal oligonucleotide concentration for a particular cell line, the cells are treated with a positive control oligonucleotide at a range of concentrations. For human cells the positive control oligonucleotide is selected from either ISIS 13920 (**TCCGTCATCGCTCCTCAGGG**, SEQ ID NO:1) which is targeted to human H-ras, or ISIS 18078, (**GTGCGCGAGCCCGAAATC**, SEQ ID NO:2) which is targeted to human Jun-N-terminal kinase-2 (JNK2). Both controls are 2'-O-methoxyethyl gapmers (2'-O-methoxyethyls shown in bold) with a phosphorothioate backbone. For mouse or rat cells the positive control oligonucleotide is ISIS 15770, **ATGCATTCTGCCCAAGGA**, SEQ ID NO: 3, a 2'-O-methoxyethyl gapmer (2'-O-methoxyethyls shown in bold) with a phosphorothioate backbone which is targeted to both mouse and rat c-raf. The concentration of positive control oligonucleotide that results in 80% inhibition of c-H-ras (for ISIS 13920), JNK2 (for ISIS 18078) or c-raf (for ISIS 15770) mRNA is then utilized as the screening concentration for new oligonucleotides in subsequent experiments for that cell line. If 80% inhibition is not achieved, the lowest concentration of positive control oligonucleotide that results in 60% inhibition of c-H-ras, JNK2 or c-raf mRNA is then utilized as the oligonucleotide screening concentration in subsequent experiments for that cell line. If 60% inhibition is not achieved, that particular cell line is deemed as unsuitable for oligonucleotide transfection experiments. The concentrations of antisense oligonucleotides used herein are from 50 nM to 300 nM.

Example 10: Analysis of oligonucleotide inhibition of huntingtin interacting protein 1 expression

Antisense modulation of huntingtin interacting protein 1 expression can be assayed in a variety of ways known in the art. For example, huntingtin interacting protein 1 mRNA levels can

be quantitated by, e.g., Northern blot analysis, competitive polymerase chain reaction (PCR), or real-time PCR (RT-PCR). Real-time quantitative PCR is presently favorable. RNA analysis can be performed on total cellular RNA or poly(A)+ mRNA. A method of RNA analysis of the present invention is the use of total cellular RNA as described in other examples herein. Methods of RNA isolation are well known in the art. Northern blot analysis is also routine in the art. Real-time quantitative (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.

Protein levels of huntingtin interacting protein 1 can be quantitated in a variety of ways well known in the art, such as immunoprecipitation, Western blot analysis (immunoblotting), enzyme-linked immunosorbent assay (ELISA) or fluorescence-activated cell sorting (FACS). Antibodies directed to huntingtin interacting protein 1 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.

Example 11: Design of phenotypic assays and *in vivo* studies for the use of huntingtin interacting protein 1 inhibitors

Phenotypic assays

Once huntingtin interacting protein 1 inhibitors have been identified by the methods disclosed herein, the compounds are further investigated in one or more phenotypic assays, each having measurable endpoints predictive of efficacy in the treatment of a particular disease state or condition.

Phenotypic assays, kits and reagents for their use are well known to those skilled in the art and are herein used to investigate the role and/or association of huntingtin interacting protein 1 in health and disease. Representative phenotypic assays, which can be purchased from any one of several commercial vendors, include those for determining cell viability, cytotoxicity, proliferation or cell survival (Molecular Probes, Eugene, OR; PerkinElmer, Boston, MA), protein-based assays including enzymatic assays (Panvera, LLC, Madison, WI; BD Biosciences, Franklin Lakes, NJ; Oncogene Research Products, San Diego, CA), cell regulation, signal transduction, inflammation, oxidative processes and apoptosis (Assay Designs Inc., Ann Arbor, MI), triglyceride accumulation (Sigma-Aldrich, St. Louis, MO), angiogenesis assays, tube formation assays, cytokine and hormone assays and metabolic assays (Chemicon International Inc., Temecula, CA; Amersham Biosciences, Piscataway, NJ).

In one non-limiting example, cells determined to be appropriate for a particular phenotypic assay (i.e., MCF-7 cells selected for breast cancer studies; adipocytes for obesity studies) are treated with huntingtin interacting protein 1 inhibitors identified from the *in vitro* studies as well as control compounds at optimal concentrations which are determined by the methods described above. At the end of the treatment period, treated and untreated cells are analyzed by one or more methods specific for the assay to determine phenotypic outcomes and endpoints.

Phenotypic endpoints include changes in cell morphology over time or treatment dose as well as changes in levels of cellular components such as proteins, lipids, nucleic acids, hormones, saccharides or metals. Measurements of cellular status which include pH, stage of the cell cycle, intake or excretion of biological indicators by the cell, are also endpoints of interest.

Analysis of the genotype of the cell (measurement of the expression of one or more of the genes of the cell) after treatment is also used as an indicator of the efficacy or potency of the huntingtin interacting protein 1 inhibitors. Hallmark genes, or those genes suspected to be associated with a specific disease state, condition, or phenotype, are measured in both treated and untreated cells.

In vivo studies

The individual subjects of the *in vivo* studies described herein are warm-blooded vertebrate animals, which includes humans.

The clinical trial is subjected to rigorous controls to ensure that individuals are not unnecessarily put at risk and that they are fully informed about their role in the study.

To account for the psychological effects of receiving treatments, volunteers are randomly given placebo or huntingtin interacting protein 1 inhibitor. Furthermore, to prevent the doctors from being biased in treatments, they are not informed as to whether the medication they are administering is a huntingtin interacting protein 1 inhibitor or a placebo. Using this randomization approach, each volunteer has the same chance of being given either the new treatment or the placebo.

Volunteers receive either the huntingtin interacting protein 1 inhibitor or placebo for eight week period with biological parameters associated with the indicated disease state or condition being measured at the beginning (baseline measurements before any treatment), end (after the final treatment), and at regular intervals during the study period. Such measurements include the levels of nucleic acid molecules encoding huntingtin interacting protein 1 or huntingtin interacting protein 1 protein levels in body fluids, tissues or organs compared to pre-treatment levels. Other measurements include, but are not limited to, indices of the disease state

or condition being treated, body weight, blood pressure, serum titers of pharmacologic indicators of disease or toxicity as well as ADME (absorption, distribution, metabolism and excretion) measurements.

Information recorded for each patient includes age (years), gender, height (cm), family history of disease state or condition (yes/no), motivation rating (some/moderate/great) and number and type of previous treatment regimens for the indicated disease or condition.

Volunteers taking part in this study are healthy adults (age 18 to 65 years) and roughly an equal number of males and females participate in the study. Volunteers with certain characteristics are equally distributed for placebo and huntingtin interacting protein 1 inhibitor treatment. In general, the volunteers treated with placebo have little or no response to treatment, whereas the volunteers treated with the huntingtin interacting protein 1 inhibitor show positive trends in their disease state or condition index at the conclusion of the study.

Example 12: RNA Isolation

Poly(A)+ mRNA isolation

Poly(A)+ mRNA was isolated according to Miura et al., (Clin. Chem., 1996, 42, 1758-1764). Other methods for poly(A)+ mRNA isolation are routine in the art. Briefly, for cells grown on 96-well plates, growth medium was removed from the cells and each well was washed with 200 μ L cold PBS. 60 μ L lysis buffer (10 mM Tris-HCl, pH 7.6, 1 mM EDTA, 0.5 M NaCl, 0.5% NP-40, 20 mM vanadyl-ribonucleoside complex) was added to each well, the plate was gently agitated and then incubated at room temperature for five minutes. 55 μ L of lysate was transferred to Oligo d(T) coated 96-well plates (AGCT Inc., Irvine CA). Plates were incubated for 60 minutes at room temperature, washed 3 times with 200 μ L of wash buffer (10 mM Tris-HCl pH 7.6, 1 mM EDTA, 0.3 M NaCl). After the final wash, the plate was blotted on paper towels to remove excess wash buffer and then air-dried for 5 minutes. 60 μ L of elution buffer (5 mM Tris-HCl pH 7.6), preheated to 70°C, was added to each well, the plate was incubated on a 90°C hot plate for 5 minutes, and the eluate was then transferred to a fresh 96-well plate.

Cells grown on 100 mm or other standard plates may be treated similarly, using appropriate volumes of all solutions.

Total RNA Isolation

Total RNA was isolated using an RNEASY 96™ kit and buffers purchased from Qiagen Inc. (Valencia, CA) following the manufacturer's recommended procedures. Briefly, for cells grown on 96-well plates, growth medium was removed from the cells and each well was washed with 200 μ L cold PBS. 150 μ L Buffer RLT was added to each well and the plate vigorously

-40-

agitated for 20 seconds. 150 μ L of 70% ethanol was then added to each well and the contents mixed by pipetting three times up and down. The samples were then transferred to the RNEASY 96TM well plate attached to a QIAVACTM manifold fitted with a waste collection tray and attached to a vacuum source. Vacuum was applied for 1 minute. 500 μ L of Buffer RW1 was added to each well of the RNEASY 96TM plate and incubated for 15 minutes and the vacuum was again applied for 1 minute. An additional 500 μ L of Buffer RW1 was added to each well of the RNEASY 96TM plate and the vacuum was applied for 2 minutes. 1 mL of Buffer RPE was then added to each well of the RNEASY 96TM plate and the vacuum applied for a period of 90 seconds. The Buffer RPE wash was then repeated and the vacuum was applied for an additional 3 minutes. The plate was then removed from the QIAVACTM manifold and blotted dry on paper towels. The plate was then re-attached to the QIAVACTM manifold fitted with a collection tube rack containing 1.2 mL collection tubes. RNA was then eluted by pipetting 140 μ L of RNase free water into each well, incubating 1 minute, and then applying the vacuum for 3 minutes.

The repetitive pipetting and elution steps may be automated using a QIAGEN Bio-Robot 9604 (Qiagen, Inc., Valencia CA). Essentially, after lysing of the cells on the culture plate, the plate is transferred to the robot deck where the pipetting, DNase treatment and elution steps are carried out.

Example 13: Real-time Quantitative PCR Analysis of huntingtin interacting protein 1 mRNA Levels

Quantitation of huntingtin interacting protein 1 mRNA levels was accomplished by real-time quantitative PCR using the ABI PRISMTM 7600, 7700, or 7900 Sequence Detection System (PE-Applied Biosystems, Foster City, CA) according to manufacturer's instructions. This is a closed-tube, non-gel-based, fluorescence detection system that allows high-throughput quantitation of polymerase chain reaction (PCR) products in real-time. As opposed to standard PCR in which amplification products are quantitated after the PCR is completed, products in real-time quantitative PCR are quantitated as they accumulate. This is accomplished by including in the PCR reaction an oligonucleotide probe that anneals specifically between the forward and reverse PCR primers, and contains two fluorescent dyes. A reporter dye (e.g., FAM or JOE, obtained from either PE-Applied Biosystems, Foster City, CA, Operon Technologies Inc., Alameda, CA or Integrated DNA Technologies Inc., Coralville, IA) is attached to the 5' end of the probe and a quencher dye (e.g., TAMRA, obtained from either PE-Applied Biosystems, Foster City, CA, Operon Technologies Inc., Alameda, CA or Integrated DNA Technologies Inc., Coralville, IA) is attached to the 3' end of the probe. When the probe and dyes are intact, reporter

dye emission is quenched by the proximity of the 3' quencher dye. During amplification, annealing of the probe to the target sequence creates a substrate that can be cleaved by the 5'-exonuclease activity of Taq polymerase. During the extension phase of the PCR amplification cycle, cleavage of the probe by Taq polymerase releases the reporter dye from the remainder of the probe (and hence from the quencher moiety) and a sequence-specific fluorescent signal is generated. With each cycle, additional reporter dye molecules are cleaved from their respective probes, and the fluorescence intensity is monitored at regular intervals by laser optics built into the ABI PRISM™ Sequence Detection System. In each assay, a series of parallel reactions containing serial dilutions of mRNA from untreated control samples generates a standard curve that is used to quantitate the percent inhibition after antisense oligonucleotide treatment of test samples.

Prior to quantitative PCR analysis, primer-probe sets specific to the target gene being measured are evaluated for their ability to be "multiplexed" with a GAPDH amplification reaction. In multiplexing, both the target gene and the internal standard gene GAPDH are amplified concurrently in a single sample. In this analysis, mRNA isolated from untreated cells is serially diluted. Each dilution is amplified in the presence of primer-probe sets specific for GAPDH only, target gene only ("single-plexing"), or both (multiplexing). Following PCR amplification, standard curves of GAPDH and target mRNA signal as a function of dilution are generated from both the single-plexed and multiplexed samples. If both the slope and correlation coefficient of the GAPDH and target signals generated from the multiplexed samples fall within 10% of their corresponding values generated from the single-plexed samples, the primer-probe set specific for that target is deemed multiplexable. Other methods of PCR are also known in the art.

PCR reagents were obtained from Invitrogen Corporation, (Carlsbad, CA). RT-PCR reactions were carried out by adding 20 μ L PCR cocktail (2.5x PCR buffer minus $MgCl_2$, 6.6 mM $MgCl_2$, 375 μ M each of dATP, dCTP, dGTP and dTTP, 375 nM each of forward primer and reverse primer, 125 nM of probe, 4 Units RNase inhibitor, 1.25 Units PLATINUM® Taq, 5 Units MuLV reverse transcriptase, and 2.5x ROX dye) to 96-well plates containing 30 μ L total RNA solution (20-200 ng). The RT reaction was carried out by incubation for 30 minutes at 48°C. Following a 10 minute incubation at 95°C to activate the PLATINUM® Taq, 40 cycles of a two-step PCR protocol were carried out: 95°C for 15 seconds (denaturation) followed by 60°C for 1.5 minutes (annealing/extension).

Gene target quantities obtained by real time RT-PCR are normalized using either the expression level of GAPDH, a gene whose expression is constant, or by quantifying total RNA

-42-

using RiboGreen™ (Molecular Probes, Inc. Eugene, OR). GAPDH expression is quantified by real time RT-PCR, by being run simultaneously with the target, multiplexing, or separately. Total RNA is quantified using RiboGreen™ RNA quantification reagent (Molecular Probes, Inc. Eugene, OR). Methods of RNA quantification by RiboGreen™ are taught in Jones, L.J., et al, (Analytical Biochemistry, 1998, 265, 368-374).

In this assay, 170 µL of RiboGreen™ working reagent (RiboGreen™ reagent diluted 1:350 in 10mM Tris-HCl, 1 mM EDTA, pH 7.5) is pipetted into a 96-well plate containing 30 µL purified, cellular RNA. The plate is read in a CytoFluor 4000 (PE Applied Biosystems) with excitation at 485nm and emission at 530nm.

Probes and primers to human huntingtin interacting protein 1 were designed to hybridize to a human huntingtin interacting protein 1 sequence, using published sequence information (GenBank accession number NM_005338.3, incorporated herein as SEQ ID NO:4). For human huntingtin interacting protein 1 the PCR primers were:

forward primer: TGACCGAGGCCTGTAAGCA (SEQ ID NO:5)

reverse primer: TTCTCAAGGCTTCCCTCTTCCT (SEQ ID NO:6)

and the PCR probe was:

FAM-TGGCAGGGAAACCCTCGCCTACC-TAMRA (SEQ ID NO:7)

where FAM is the fluorescent dye and TAMRA is the quencher dye. For human GAPDH the PCR primers were:

forward primer: GAAGGTGAAGGTCGGAGTC (SEQ ID NO:8)

reverse primer: GAAGATGGTGATGGGATTTC (SEQ ID NO:9)

and the PCR probe was:

5' JOE-CAAGCTTCCCGTTCTCAGCC- TAMRA 3' (SEQ ID NO:10)

where JOE is the fluorescent reporter dye and TAMRA is the quencher dye.

Example 14: Northern blot analysis of huntingtin interacting protein 1 mRNA levels

Eighteen hours after antisense treatment, cell monolayers were washed twice with cold PBS and lysed in 1 mL RNAZOL™ (TEL-TEST "B" Inc., Friendswood, TX). Total RNA was prepared following manufacturer's recommended protocols. Twenty micrograms of total RNA was fractionated by electrophoresis through 1.2% agarose gels containing 1.1% formaldehyde using a MOPS buffer system (AMRESCO, Inc. Solon, OH). RNA was transferred from the gel to HYBOND™-N+ nylon membranes (Amersham Pharmacia Biotech, Piscataway, NJ) by overnight capillary transfer using a Northern/Southern Transfer buffer system (TEL-TEST "B" Inc., Friendswood, TX). RNA transfer was confirmed by UV visualization. Membranes were

fixed by UV cross-linking using a STRATALINKER™ UV Crosslinker 2400 (Stratagene, Inc, La Jolla, CA) and then probed using QUICKHYB™ hybridization solution (Stratagene, La Jolla, CA) using manufacturer's recommendations for stringent conditions.

To detect human huntingtin interacting protein 1, a human huntingtin interacting protein 1 specific probe was prepared by PCR using the forward primer TGACCGAGGCTGTAAAGCA (SEQ ID NO:5) and the reverse primer TTCTCAAGGCTTCCCTCTTCCT (SEQ ID NO:6). To normalize for variations in loading and transfer efficiency membranes were stripped and probed for human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) RNA (Clontech, Palo Alto, CA).

Hybridized membranes were visualized and quantitated using a PHOSPHORIMAGER™ and IMAGEQUANT™ Software V3.3 (Molecular Dynamics, Sunnyvale, CA). Data was normalized to GAPDH levels in untreated controls.

Example 15: Antisense inhibition of human huntingtin interacting protein 1 expression by chimeric phosphorothioate oligonucleotides having 2'-MOE wings and a deoxy gap

In accordance with the present invention, a series of antisense compounds were designed to target different regions of the human huntingtin interacting protein 1 RNA, using published sequences (GenBank accession number NM_005338.3, incorporated herein as SEQ ID NO:4, the complement of nucleotides 2843247 to 2908700 of the sequence with GenBank accession number NT_007867.8, representing a genomic sequence, incorporated herein as SEQ ID NO:11). The compounds are shown in Table 1. "Target site" indicates the first (5'-most) nucleotide number on the particular target sequence to which the compound binds. All compounds in Table 1 are chimeric oligonucleotides ("gapmers") 20 nucleotides in length, composed of a central "gap" region consisting of ten 2'-deoxynucleotides, which is flanked on both sides (5' and 3' directions) by five-nucleotide "wings." The wings are composed of 2'-methoxyethyl (2'-MOE)nucleotides. The internucleoside (backbone) linkages are phosphorothioate (P=S) throughout the oligonucleotide. All cytidine residues are 5-methylcytidines. The compounds were analyzed for their effect on human huntingtin interacting protein 1 mRNA levels by quantitative real-time PCR as described in other examples herein. Data are averages from three experiments in which A549 cells were treated with the antisense oligonucleotides of the present invention. The positive control for each datapoint is identified in the table by sequence ID number. If present, "N.D." indicates "no data."

-44-

Table 1

Inhibition of human huntingtin interacting protein 1 mRNA levels by chimeric phosphorothioate oligonucleotides having 2'-MOE wings and a deoxy gap

ISIS #	REGION	TARGET SEQ ID NO	TARGET SITE	SEQUENCE	% INHIB	SEQ ID NO	CONTROL SEQ ID NO
251581	5'UTR	4	6	cgtaataatggccttattga	82	14	1
251582	5'UTR	4	11	tcctgcgtataatggcctt	84	15	1
251583	5'UTR	4	16	ccacttctcgttataatg	90	16	1
251584	Start Codon	4	237	catctgctcatgtcactca	97	17	1
251585	Start Codon	4	243	gccccacatctgctcatgt	92	18	1
251586	Coding	4	248	aggtggccccacatcctgct	87	19	1
251587	Coding	4	253	cgctcaggtggccccacatc	75	20	1
251588	Coding	4	260	taccctcgtcaggtggcc	88	21	1
251589	Coding	4	322	tttgggtgtgtactccatc	73	22	1
251590	Coding	4	327	gggattttgggtgtgtact	57	23	1
251591	Coding	4	332	aacctgggattttgggtgtg	70	24	1
251592	Coding	4	378	tccagctcgtccagctggc	85	25	1
251593	Coding	4	408	taactggaaaaagtgttca	74	26	1
251594	Coding	4	653	aactgctccatgaagcggtc	0	27	1
251595	Coding	4	674	aacagatcttcaacttgt	71	28	1
251596	Coding	4	857	tccatgaggtcatccttctc	78	29	1
251597	Coding	4	862	ccatgtccatgaggtcatcc	78	30	1
251598	Coding	4	902	tcatcaaaactgtgtcaaa	93	31	1
251599	Coding	4	1025	tgtgccttcaatccactgat	89	32	1
251600	Coding	4	1322	ttctccgcagcaggtcagc	81	33	1
251601	Coding	4	1343	acctgtttgtcacctctgc	80	34	1
251602	Coding	4	1524	caggtgccttgagaaacct	77	35	1
251603	Coding	4	1901	gggcaggccagatactggct	66	36	1
251604	Coding	4	1906	cttctgggcaggccagatac	53	37	1
251605	Coding	4	2044	tgccatactgcttacaggcc	77	38	1
251606	Coding	4	2085	tcccttctcctcaggagg	88	39	1
251607	Coding	4	2161	gcaggagctctcgcgatg	21	40	1
251608	Coding	4	2427	ggatgctgtacctgcccgc	89	41	1
251609	Coding	4	2526	agctgcatccaccatgacag	75	42	1
251610	Coding	4	2616	tcagccacaagctgggctg	62	43	1
251611	Coding	4	2673	agaggcctgctgacgtggg	47	44	1
251612	Coding	4	2678	ccccgagaggcctgctgcag	90	45	1
251613	Coding	4	2683	tcactccccgagaggcctgc	94	46	1
251614	Coding	4	2688	ctgggtcactccccgagagg	89	47	1
251615	Coding	4	2693	gtggcctgggtcactccccg	88	48	1
251616	Coding	4	2723	ccggaaatggtgagggccac	78	49	1
251617	Coding	4	2728	atttgcggaaatggtgag	48	50	1
251618	Coding	4	2771	gtcatgctgagaagtccat	87	51	1
251619	Coding	4	2916	ttctcccagcctcagcaa	95	52	1
251620	Stop Codon	4	2986	gggtgttgggttggctctat	75	53	1
251621	3'UTR	4	3099	tcggcactgggtaatggcag	85	54	1
251622	3'UTR	4	3219	ggcagcactggccagcctgg	87	55	1
251623	3'UTR	4	3339	tcctctattaaggataccca	97	56	1
251624	3'UTR	4	3446	tgctcacaagtgtgtcaaa	81	57	1
251625	3'UTR	4	3797	tgaccctggagcatggactg	88	58	1
251626	3'UTR	4	3854	aaaggcactcactctccttc	92	59	1
251627	3'UTR	4	4038	ggacagttcattccggcagg	83	60	1
251628	3'UTR	4	4142	tcaagaggatgccaaaggcag	91	61	1
251629	3'UTR	4	4230	ccaagtatagggtcttccc	92	62	1

-45-

251630	3'UTR	4	4304	tgattgctccaagcatctct	91	63	1
251631	3'UTR	4	4312	tgaagtctgattgctccaa	89	64	1
251632	3'UTR	4	4444	ctgacccaagagctccaaat	91	65	1
251633	3'UTR	4	4491	tggctgaaaggagttggagc	75	66	1
251634	3'UTR	4	4551	agctgttcatgtgccctctg	83	67	1
251635	3'UTR	4	4624	gatcagaaggctactaaat	78	68	1
251636	3'UTR	4	4736	ccgtcatgtagcaaaccta	92	69	1
251637	3'UTR	4	4749	gaagtctcacaaccgctcat	94	70	1
251638	3'UTR	4	4818	atgcacagagaggagttgg	78	71	1
251639	3'UTR	4	4890	atggaggtcacacgtctgag	93	72	1
251640	3'UTR	4	5928	gctcttttttagagacagga	94	73	1
251641	3'UTR	4	6016	acactgaattagcctctgct	92	74	1
251642	3'UTR	4	6085	gggtagccattctaactga	88	75	1
251643	3'UTR	4	6190	ttaagatgtgattccggtt	89	76	1
251644	3'UTR	4	6243	cactagtgtgctcagtgac	88	77	1
251645	exon: intron junction	11	7413	aggaaactcacgttcgggtgt	63	78	1
251646	exon: intron junction	11	13032	agagactcactttggtgtgg	69	79	1
251647	intron	11	22819	taatcaagttcaatgatcac	77	80	1
251648	intron	11	26634	gccataaaaggcctgagct	19	81	1
251649	exon: intron junction	11	37726	ttggactcacttctcatcct	94	82	1
251650	intron: exon junction	11	42028	cctctgcattctgcaaaaga	77	83	1
251651	intron	11	45218	ttcatcctcgttaattaagc	65	84	1
251652	intron	11	52996	ctctgctgatattcacagga	66	85	1
251653	genomic	11	246	catggcaattaaagcccga	79	86	1
251654	genomic	11	256	ggcacaacaacatggcaatt	63	87	1
251655	genomic	11	303	gcattggctgtgccagctg	52	88	1
251656	genomic	11	497	ggatgagatgaataagcctc	10	89	1
251657	intron	11	36494	ggtgtcttcatagcccct	78	90	1
251658	intron	11	36622	tgtgtggttggcagctta	78	91	1

As shown in Table 1, SEQ ID NOs 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 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, 80, 82, 83, 84, 85, 86, 87, 88, 90 and 91 demonstrated at least 45% inhibition of human huntingtin interacting protein 1 expression in this assay and are therefore suitable. More suitable are SEQ ID NOs 17, 56 and 52. The target regions to which these sequences are complementary are herein referred to as "suitable target segments" and are therefore suitable for targeting by compounds of the present invention. These suitable target segments are shown in Table 2. The sequences represent the reverse complement of the suitable antisense compounds shown in Table 1. "Target site" indicates the first (5'-most) nucleotide number on the particular target nucleic acid to which the oligonucleotide binds. Also shown in Table 2 is the species in which each of the suitable target segments was found.

-46-

Table 2
Sequence and position of suitable target segments identified
in huntingtin interacting protein 1

SITE ID	TARGET SEQ ID NO	TARGET SITE	SEQUENCE	REV COMP OF SEQ ID	ACTIVE IN	SEQ ID NO
168097	4	6	tcaataaggccattaatac	14	<i>H. sapiens</i>	92
168098	4	11	aaggccattaatacgcagga	15	<i>H. sapiens</i>	93
168099	4	16	cattaatacgcaggagtgg	16	<i>H. sapiens</i>	94
168100	4	237	tgagtacatgagcaggatg	17	<i>H. sapiens</i>	95
168101	4	243	acatgagcaggatgtggggc	18	<i>H. sapiens</i>	96
168102	4	248	agcaggatgtggggccacct	19	<i>H. sapiens</i>	97
168103	4	253	gatgtggggccacctgagcg	20	<i>H. sapiens</i>	98
168104	4	260	ggccacctgagcgaggggta	21	<i>H. sapiens</i>	99
168105	4	322	gatggagtaccacacaaaa	22	<i>H. sapiens</i>	100
168106	4	327	agtaccacacaaaaatccc	23	<i>H. sapiens</i>	101
168107	4	332	cacacaaaaatcccagggt	24	<i>H. sapiens</i>	102
168108	4	378	gccagctggacgaggtcga	25	<i>H. sapiens</i>	103
168109	4	408	tgaacaactttccagttta	26	<i>H. sapiens</i>	104
168111	4	674	acaagtgtgaaagatctgt	28	<i>H. sapiens</i>	105
168112	4	857	gagaaggatgacctcatgga	29	<i>H. sapiens</i>	106
168113	4	862	ggatgacctcatgacatgg	30	<i>H. sapiens</i>	107
168114	4	902	ttgacaacaagtgtgatga	31	<i>H. sapiens</i>	108
168115	4	1025	atcagtgattgaaggcaca	32	<i>H. sapiens</i>	109
168116	4	1322	gctgacctgctcgggaagaa	33	<i>H. sapiens</i>	110
168117	4	1343	gcagaggtgaccaaacaggt	34	<i>H. sapiens</i>	111
168118	4	1524	agttctgcaaggcagcctg	35	<i>H. sapiens</i>	112
168119	4	1901	agccagtatctggcctgccc	36	<i>H. sapiens</i>	113
168120	4	1906	gtatctggcctgcccagaag	37	<i>H. sapiens</i>	114
168121	4	2044	ggcctgtaagcagtatggca	38	<i>H. sapiens</i>	115
168122	4	2085	cctccctggagggaagaggga	39	<i>H. sapiens</i>	116
168124	4	2427	gcggcagggtgtacagatcc	41	<i>H. sapiens</i>	117
168125	4	2526	ctgcatggtgatgcatgct	42	<i>H. sapiens</i>	118
168126	4	2616	cagccagctgtggctgca	43	<i>H. sapiens</i>	119
168127	4	2673	cccagctgcagcaggcctct	44	<i>H. sapiens</i>	120
168128	4	2678	ctgcagcaggcctctcgggg	45	<i>H. sapiens</i>	121
168129	4	2683	gcagcctctcggggagtga	46	<i>H. sapiens</i>	122
168130	4	2688	cctctcggggagtgaaccag	47	<i>H. sapiens</i>	123
168131	4	2693	cggggagtgaaccaggccac	48	<i>H. sapiens</i>	124
168132	4	2723	gtggcctcaaccatttccgg	49	<i>H. sapiens</i>	125
168133	4	2728	ctcaaccatttccggcaaat	50	<i>H. sapiens</i>	126
168134	4	2771	atggacttctcaagcatgac	51	<i>H. sapiens</i>	127
168135	4	2916	ttgctgagggtgggaagaa	52	<i>H. sapiens</i>	128
168136	4	2986	atagagccaaaccaacacc	53	<i>H. sapiens</i>	129
168137	4	3099	ctgccattaccagtccga	54	<i>H. sapiens</i>	130
168138	4	3219	ccaggctggcagtgctgcc	55	<i>H. sapiens</i>	131
168139	4	3339	tgggtatccttaatagagga	56	<i>H. sapiens</i>	132
168140	4	3446	ttgcacaactgtgagca	57	<i>H. sapiens</i>	133
168141	4	3797	cagtcctatgctccagggtca	58	<i>H. sapiens</i>	134
168142	4	3854	gaaggagagtgtgctccttt	59	<i>H. sapiens</i>	135
168143	4	4038	cctgccggaatgaactgtcc	60	<i>H. sapiens</i>	136
168144	4	4142	ctgcctggcatcctctga	61	<i>H. sapiens</i>	137
168145	4	4230	gggaagaaccctacttgg	62	<i>H. sapiens</i>	138
168146	4	4304	agagatgctggagcaatca	63	<i>H. sapiens</i>	139
168147	4	4312	ttggagcaatcagaactca	64	<i>H. sapiens</i>	140
168148	4	4444	attggagctcttgggtcag	65	<i>H. sapiens</i>	141
168149	4	4491	gctccaactccttcagcca	66	<i>H. sapiens</i>	142
168150	4	4551	cagaggccacatgaacagct	67	<i>H. sapiens</i>	143

168151	4	4624	atttaagtgccttctgatc	68	<i>H. sapiens</i>	144
168152	4	4736	taggttttgctacatgacgg	69	<i>H. sapiens</i>	145
168153	4	4749	atgacgggtgtgagacttc	70	<i>H. sapiens</i>	146
168154	4	4818	ccaactccctctctgtgcat	71	<i>H. sapiens</i>	147
168155	4	4890	ctcagacgtgtgacctccat	72	<i>H. sapiens</i>	148
168156	4	5928	tcctgtctctaaaaagaagc	73	<i>H. sapiens</i>	149
168157	4	6016	agcagaggctaattcagtg	74	<i>H. sapiens</i>	150
168158	4	6085	tcagattagaatggctacc	75	<i>H. sapiens</i>	151
168159	4	6190	aaacgggaatcacatcttaa	76	<i>H. sapiens</i>	152
168160	4	6243	gtcactgagcatcactagt	77	<i>H. sapiens</i>	153
168161	11	7413	acacccgaacgtgagttcct	78	<i>H. sapiens</i>	154
168162	11	13032	ccacaccaaatgagtgctct	79	<i>H. sapiens</i>	155
168163	11	22819	gtgatcattgaacttgatta	80	<i>H. sapiens</i>	156
168165	11	37726	aggatgagaagtgagtccaa	82	<i>H. sapiens</i>	157
168166	11	42028	tctttgcagaatgcagagg	83	<i>H. sapiens</i>	158
168167	11	45218	gcttaattaacgaggatgaa	84	<i>H. sapiens</i>	159
168168	11	52996	tcctgtagatatcagcagag	85	<i>H. sapiens</i>	160
168169	11	132	tggggcttfaattgccatg	86	<i>H. sapiens</i>	161
168170	11	142	aattgccatgtgtgtgcc	87	<i>H. sapiens</i>	162
168171	11	188	cagctgggcacagccaatgc	88	<i>H. sapiens</i>	163
168173	11	70	atggggctgatgaagacacc	90	<i>H. sapiens</i>	164
168174	11	198	taagcatgcccacacaca	91	<i>H. sapiens</i>	165

As these “suitable target segments” have been found by experimentation to be open to, and accessible for, hybridization with the antisense compounds of the present invention, one of skill in the art will recognize or be able to ascertain, using no more than routine experimentation, further embodiments of the invention that encompass other compounds that specifically hybridize to these suitable target segments and consequently inhibit the expression of huntingtin interacting protein 1.

According to the present invention, antisense compounds include antisense oligomeric compounds, antisense oligonucleotides, ribozymes, external guide sequence (EGS) oligonucleotides, alternate splicers, primers, probes, and other short oligomeric compounds that hybridize to at least a portion of the target nucleic acid.

Example 16: Western blot analysis of huntingtin interacting protein 1 protein levels

Western blot analysis (immunoblot analysis) is carried out using standard methods. Cells are harvested 16-20 h after oligonucleotide treatment, washed once with PBS, suspended in Laemmli buffer (100 ul/well), boiled for 5 minutes and loaded on a 16% SDS-PAGE gel. Gels are run for 1.5 hours at 150 V, and transferred to membrane for western blotting. Appropriate primary antibody directed to huntingtin interacting protein 1 is used, with a radiolabeled or fluorescently labeled secondary antibody directed against the primary antibody species. Bands are visualized using a PHOSPHORIMAGER™ (Molecular Dynamics, Sunnyvale CA).

What is claimed is:

1. A compound 8 to 80 nucleobases in length targeted to a nucleic acid molecule encoding huntingtin interacting protein 1, wherein the compound specifically hybridizes with the nucleic acid molecule encoding huntingtin interacting protein 1 (SEQ ID NO: 4) and inhibits the expression of huntingtin interacting protein 1.
2. The compound of claim 1 comprising 12 to 50 nucleobases in length.
3. The compound of claim 2 comprising 15 to 30 nucleobases in length.
4. The compound of claim 1 comprising an oligonucleotide.
5. The compound of claim 4 comprising an antisense oligonucleotide.
6. The compound of claim 4 comprising a DNA oligonucleotide.
7. The compound of claim 4 comprising an RNA oligonucleotide.
8. The compound of claim 4 comprising a chimeric oligonucleotide.
9. The compound of claim 4 wherein at least a portion of the compound hybridizes with RNA to form an oligonucleotide-RNA duplex.
10. The compound of claim 1 having at least 70% complementarity with a nucleic acid molecule encoding huntingtin interacting protein 1 (SEQ ID NO: 4) the compound specifically hybridizing to and inhibiting the expression of huntingtin interacting protein 1.
11. The compound of claim 1 having at least 80% complementarity with a nucleic acid molecule encoding huntingtin interacting protein 1 (SEQ ID NO: 4) the compound specifically hybridizing to and inhibiting the expression of huntingtin interacting protein 1.
12. The compound of claim 1 having at least 90% complementarity with a nucleic acid molecule encoding huntingtin interacting protein 1 (SEQ ID NO: 4) the compound specifically hybridizing to and inhibiting the expression of huntingtin interacting protein 1.
13. The compound of claim 1 having at least 95% complementarity with a nucleic acid molecule encoding huntingtin interacting protein 1 (SEQ ID NO: 4) the compound specifically hybridizing to and inhibiting the expression of huntingtin interacting protein 1.
14. The compound of claim 1 having at least one modified internucleoside linkage, sugar moiety, or nucleobase.
15. The compound of claim 1 having at least one 2'-O-methoxyethyl sugar moiety.
16. The compound of claim 1 having at least one phosphorothioate internucleoside linkage.
17. The compound of claim 1 having at least one 5-methylcytosine.
18. The compound of claim 1 wherein the compound specifically hybridizes to the 5' untranslated region, the start codon region, the coding region, the stop codon region, or the 3' untranslated region of the nucleic acid molecule encoding huntingtin interacting protein 1.

19. The compound of claim 1 wherein the compound specifically hybridizes to the 5' untranslated region.
20. The compound of claim 1 wherein the compound specifically hybridizes to the start codon region.
21. The compound of claim 1 wherein the compound specifically hybridizes to the coding region.
22. The compound of claim 1 wherein the compound specifically hybridizes to the stop codon region.
23. The compound of claim 1 wherein the compound specifically hybridizes to the 3' untranslated region.
24. The compound of claim 1 wherein the compound comprises SEQ ID NO:17, 52, 56, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 53, 54, 55, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 82, 83, 84, 85, 86, 87, 88, 90, or 91.
25. The compound of claim 1 wherein the compound comprises SEQ ID NO:17, 52, or 56.
26. A method of inhibiting the expression of huntingtin interacting protein 1 in cells or tissues comprising contacting the cells or tissues with the compound of claim 1 so that expression of huntingtin interacting protein 1 is inhibited.
27. A method of screening for a modulator of huntingtin interacting protein 1, the method comprising the steps of:
 - contacting a suitable target segment of a nucleic acid molecule encoding huntingtin interacting protein 1 with one or more candidate modulators of huntingtin interacting protein 1;
 - and
 - identifying one or more modulators of huntingtin interacting protein 1 expression which modulate the expression of huntingtin interacting protein 1.
28. The method of claim 27 wherein the modulator of huntingtin interacting protein 1 expression comprises an oligonucleotide, an antisense oligonucleotide, a DNA oligonucleotide, an RNA oligonucleotide, an RNA oligonucleotide having at least a portion of said RNA oligonucleotide capable of hybridizing with RNA to form an oligonucleotide-RNA duplex, or a chimeric oligonucleotide.
29. A diagnostic method for identifying a disease state comprising identifying the presence of huntingtin interacting protein 1 in a sample using at least one of the primers comprising SEQ ID NOs: 5 or 6, or the probe comprising SEQ ID NO: 7.
30. A kit or assay device comprising the compound of claim 1.

31. A method of treating an animal having a disease or condition associated with huntingtin interacting protein 1 comprising administering to the animal a therapeutically or prophylactically effective amount of the compound of claim 1 so that expression of huntingtin interacting protein 1 is inhibited.

32. The method of claim 31 wherein the disease or condition involves dysregulation of cellular apoptosis.

SEQUENCE LISTING

<110> Kenneth W. Dobie

<120> MODULATION OF HUNTINGTIN INTERACTING PROTEIN 1 EXPRESSION

<130> ISIS0005-500WO (RTS-0432WO)

<150> 10/293,864

<151> 2002-11-11

<160> 168

<210> 1

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 1

tccgtcatcg ctctcaggg

20

<210> 2

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 2

gtgcgcgcgga gcccgaatc

20

<210> 3

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 3

atgcattctg cccccaagga

20

<210> 4

<211> 6626

<212> DNA

<213> H. sapiens

<220>

<221> CDS

<222> (245)...(2989)

<400> 4

cagcatcaat aaggccatta atacgcagga agtggctgta aaggaaaaac acgccagaac 60
 gtgcatactg ggcacccacc atgagaaagg ggcacagacc ttctggctctg ttgtcaaccg 120
 cctgcctctg tctagcaacg cagtgcctctg ctggaagttc tgccatgtgt tccacaaact 180
 cctccgagat ggacaccoga acgtcctgaa ggactctctg agatacagaa atgaattgag 240
 tgac atg agc agg atg tgg ggc cac ctg agc gag ggg tat ggc cag ctg 289
 Met Ser Arg Met Trp Gly His Leu Ser Glu Gly Tyr Gly Gln Leu
 1 5 10 15

tgc agc atc tac ctg aaa ctg cta aga acc aag atg gag tac cac acc 337
 Cys Ser Ile Tyr Leu Lys Leu Leu Arg Thr Lys Met Glu Tyr His Thr
 20 25 30

aaa aat ccc agg ttc cca ggc aac ctg cag atg agt gac cgc cag ctg 385
 Lys Asn Pro Arg Phe Pro Gly Asn Leu Gln Met Ser Asp Arg Gln Leu
 35 40 45

gac gag gct gga gaa agt gac gtg aac aac ttt ttc cag tta aca gtg 433
 Asp Glu Ala Gly Glu Ser Asp Val Asn Asn Phe Phe Gln Leu Thr Val
 50 55 60

gag atg ttt gac tac ctg gag tgt gaa ctc aac ctc ttc caa aca gta 481
 Glu Met Phe Asp Tyr Leu Glu Cys Glu Leu Asn Leu Phe Gln Thr Val
 65 70 75

ttc aac tcc ctg gac atg tcc cgc tct gtg tcc gtg acg gca gca ggg 529
 Phe Asn Ser Leu Asp Met Ser Arg Ser Val Ser Val Thr Ala Ala Gly
 80 85 90 95

cag tgc cgc ctc gcc ccg ctg atc cag gtc atc ttg gac tgc agc cac 577
 Gln Cys Arg Leu Ala Pro Leu Ile Gln Val Ile Leu Asp Cys Ser His
 100 105 110

ctt tat gac tac act gtc aag ctt ctc ttc aaa ctc cac tcc tgc ctc 625
 Leu Tyr Asp Tyr Thr Val Lys Leu Leu Phe Lys Leu His Ser Cys Leu
 115 120 125

cca gct gac acc ctg caa ggc cac cgg gac cgc ttc atg gag cag ttt 673
 Pro Ala Asp Thr Leu Gln Gly His Arg Asp Arg Phe Met Glu Gln Phe
 130 135 140

aca aag ttg aaa gat ctg ttc tac cgc tcc agc aac ctg cag tac ttc 721
 Thr Lys Leu Lys Asp Leu Phe Tyr Arg Ser Ser Asn Leu Gln Tyr Phe
 145 150 155

aag cgg ctc att cag atc ccc cag ctg cct gag aac cca ccc aac ttc 769
 Lys Arg Leu Ile Gln Ile Pro Gln Leu Pro Glu Asn Pro Pro Asn Phe
 160 165 170 175

ctg cga gcc tca gcc ctg tca gaa cat atc agc cct gtg gtg gtg atc 817
 Leu Arg Ala Ser Ala Leu Ser Glu His Ile Ser Pro Val Val Val Ile
 180 185 190

cct gca gag gcc tca tcc ccc gac agc gag cca gtc cta gag aag gat 865
 Pro Ala Glu Ala Ser Ser Pro Asp Ser Glu Pro Val Leu Glu Lys Asp

195					200					205						
gac	ctc	atg	gac	atg	gat	gcc	tct	cag	cag	aat	tta	ttt	gac	aac	aag	913
Asp	Leu	Met	Asp	Met	Asp	Ala	Ser	Gln	Gln	Asn	Leu	Phe	Asp	Asn	Lys	
		210					215					220				
ttt	gat	gac	atc	ttt	ggc	agt	tca	ttc	agc	agt	gat	ccc	ttc	aat	ttc	961
Phe	Asp	Asp	Ile	Phe	Gly	Ser	Ser	Phe	Ser	Ser	Asp	Pro	Phe	Asn	Phe	
	225					230					235					
aac	agt	caa	aat	ggg	gtg	aac	aag	gat	gag	aag	gac	cac	tta	att	gag	1009
Asn	Ser	Gln	Asn	Gly	Val	Asn	Lys	Asp	Glu	Lys	Asp	His	Leu	Ile	Glu	
240					245					250					255	
cga	cta	tac	aga	gag	atc	agt	gga	ttg	aag	gca	cag	cta	gaa	aac	atg	1057
Arg	Leu	Tyr	Arg	Glu	Ile	Ser	Gly	Leu	Lys	Ala	Gln	Leu	Glu	Asn	Met	
				260					265					270		
aag	act	gag	agc	cag	cgg	gtt	gtg	ctg	cag	ctg	aag	ggc	cac	gtc	agc	1105
Lys	Thr	Glu	Ser	Gln	Arg	Val	Val	Leu	Gln	Leu	Lys	Gly	His	Val	Ser	
			275					280						285		
gag	ctg	gaa	gca	gat	ctg	gcc	gag	cag	cag	cac	ctg	cgg	cag	cag	gcg	1153
Glu	Leu	Glu	Ala	Asp	Leu	Ala	Glu	Gln	Gln	His	Leu	Arg	Gln	Gln	Ala	
		290					295					300				
gcc	gac	gac	tgt	gaa	ttc	ctg	cgg	gca	gaa	ctg	gac	gag	ctc	agg	agg	1201
Ala	Asp	Asp	Cys	Glu	Phe	Leu	Arg	Ala	Glu	Leu	Asp	Glu	Leu	Arg	Arg	
	305					310					315					
cag	cgg	gag	gac	acc	gag	aag	gct	cag	cgg	agc	ctg	tct	gag	ata	gaa	1249
Gln	Arg	Glu	Asp	Thr	Glu	Lys	Ala	Gln	Arg	Ser	Leu	Ser	Glu	Ile	Glu	
320					325					330					335	
agg	aaa	gct	caa	gcc	aat	gaa	cag	cga	tat	agc	aag	cta	aag	gag	aag	1297
Arg	Lys	Ala	Gln	Ala	Asn	Glu	Gln	Arg	Tyr	Ser	Lys	Leu	Lys	Glu	Lys	
			340					345						350		
tac	agc	gag	ctg	gtt	cag	aac	cac	gct	gac	ctg	ctg	cgg	aag	aat	gca	1345
Tyr	Ser	Glu	Leu	Val	Gln	Asn	His	Ala	Asp	Leu	Leu	Arg	Lys	Asn	Ala	
			355					360					365			
gag	gtg	acc	aaa	cag	gtg	tcc	atg	gcc	aga	caa	gcc	cag	gta	gat	ttg	1393
Glu	Val	Thr	Lys	Gln	Val	Ser	Met	Ala	Arg	Gln	Ala	Gln	Val	Asp	Leu	
		370				375						380				
gaa	cga	gag	aaa	aaa	gag	ctg	gag	gat	tcg	ttg	gag	cgc	atc	agt	gac	1441
Glu	Arg	Glu	Lys	Lys	Glu	Leu	Glu	Asp	Ser	Leu	Glu	Arg	Ile	Ser	Asp	
	385					390					395					
cag	ggc	cag	cgg	aag	act	caa	gaa	cag	ctg	gaa	gtt	cta	gag	agc	ttg	1489
Gln	Gly	Gln	Arg	Lys	Thr	Gln	Glu	Gln	Leu	Glu	Val	Leu	Glu	Ser	Leu	
400					405					410					415	
aag	cag	gaa	ctt	gcc	aca	agc	caa	cgg	gag	ctt	cag	gtt	ctg	caa	ggc	1537
Lys	Gln	Glu	Leu	Ala	Thr	Ser	Gln	Arg	Glu	Leu	Gln	Val	Leu	Gln	Gly	
				420					425					430		

agc ctg gaa act tct gcc cag tca gaa gca aac tgg gca gcc gag ttc 1585
 Ser Leu Glu Thr Ser Ala Gln Ser Glu Ala Asn Trp Ala Ala Glu Phe
 435 440 445

gcc gag cta gag aag gag cgg gac agc ctg gtg agt ggc gca gct cat 1633
 Ala Glu Leu Glu Lys Glu Arg Asp Ser Leu Val Ser Gly Ala Ala His
 450 455 460

agg gag gag gaa tta tct gct ctt cgg aaa gaa ctg cag gac act cag 1681
 Arg Glu Glu Glu Leu Ser Ala Leu Arg Lys Glu Leu Gln Asp Thr Gln
 465 470 475

ctc aaa ctg gcc agc aca gag gaa tct atg tgc cag ctt gcc aaa gac 1729
 Leu Lys Leu Ala Ser Thr Glu Glu Ser Met Cys Gln Leu Ala Lys Asp
 480 485 490 495

caa cga aaa atg ctt ctg gtg ggg tcc agg aag gct gcg gag cag gtg 1777
 Gln Arg Lys Met Leu Leu Val Gly Ser Arg Lys Ala Ala Glu Gln Val
 500 505 510

ata caa gac gcc ctg aac cag ctt gaa gaa cct cct ctc atc agc tgc 1825
 Ile Gln Asp Ala Leu Asn Gln Leu Glu Glu Pro Pro Leu Ile Ser Cys
 515 520 525

gct ggg tct gca gat cac ctc ctc tcc acg gtc aca tcc att tcc agc 1873
 Ala Gly Ser Ala Asp His Leu Leu Ser Thr Val Thr Ser Ile Ser Ser
 530 535 540

tgc atc gag caa ctg gag aaa agc tgg agc cag tat ctg gcc tgc cca 1921
 Cys Ile Glu Gln Leu Glu Lys Ser Trp Ser Gln Tyr Leu Ala Cys Pro
 545 550 555

gaa gac atc agt gga ctt ctc cat tcc ata acc ctg ctg gcc cac ttg 1969
 Glu Asp Ile Ser Gly Leu Leu His Ser Ile Thr Leu Leu Ala His Leu
 560 565 570 575

acc agc gac gcc att gct cat ggt gcc acc acc tgc ctc aga gcc cca 2017
 Thr Ser Asp Ala Ile Ala His Gly Ala Thr Thr Cys Leu Arg Ala Pro
 580 585 590

cct gag cct gcc gac tca ctg acc gag gcc tgt aag cag tat ggc agg 2065
 Pro Glu Pro Ala Asp Ser Leu Thr Glu Ala Cys Lys Gln Tyr Gly Arg
 595 600 605

gaa acc ctc gcc tac ctg gcc tcc ctg gag gaa gag gga agc ctt gag 2113
 Glu Thr Leu Ala Tyr Leu Ala Ser Leu Glu Glu Glu Gly Ser Leu Glu
 610 615 620

aat gcc gac agc aca gcc atg agg aac tgc ctg agc aag atc aag gcc 2161
 Asn Ala Asp Ser Thr Ala Met Arg Asn Cys Leu Ser Lys Ile Lys Ala
 625 630 635

atc ggc gag gag ctc ctg ccc agg gga ctg gac atc aag cag gag gag 2209
 Ile Gly Glu Glu Leu Leu Pro Arg Gly Leu Asp Ile Lys Gln Glu Glu
 640 645 650 655

ctg ggg gac ctg gtg gac aag gag atg gcg gcc act tca gct gct att	2257
Leu Gly Asp Leu Val Asp Lys Glu Met Ala Ala Thr Ser Ala Ala Ile	
660 665 670	
gaa act gcc acg gcc aga ata gag gag atg ctc agc aaa tcc cga gca	2305
Glu Thr Ala Thr Ala Arg Ile Glu Glu Met Leu Ser Lys Ser Arg Ala	
675 680 685	
gga gac aca gga gtc aaa ttg gag gtg aat gaa agg atc ctt ggt tgc	2353
Gly Asp Thr Gly Val Lys Leu Glu Val Asn Glu Arg Ile Leu Gly Cys	
690 695 700	
tgt acc agc ctc atg caa gct att cag gtg ctc atc gtg gcc tct aag	2401
Cys Thr Ser Leu Met Gln Ala Ile Gln Val Leu Ile Val Ala Ser Lys	
705 710 715	
gac ctc cag aga gag att gtg gag agc ggc agg ggt aca gca tcc cct	2449
Asp Leu Gln Arg Glu Ile Val Glu Ser Gly Arg Gly Thr Ala Ser Pro	
720 725 730 735	
aaa gag ttt tat gcc aag aac tct cga tgg aca gaa gga ctt atc tca	2497
Lys Glu Phe Tyr Ala Lys Asn Ser Arg Trp Thr Glu Gly Leu Ile Ser	
740 745 750	
gcc tcc aag gct gtg ggc tgg gga gcc act gtc atg gtg gat gca gct	2545
Ala Ser Lys Ala Val Gly Trp Gly Ala Thr Val Met Val Asp Ala Ala	
755 760 765	
gat ctg gtg gta caa ggc aga ggg aaa ttt gag gag cta atg gtg tgt	2593
Asp Leu Val Val Gln Gly Arg Gly Lys Phe Glu Glu Leu Met Val Cys	
770 775 780	
tct cat gaa att gct gct agc aca gcc cag ctt gtg gct gca tcc aag	2641
Ser His Glu Ile Ala Ala Ser Thr Ala Gln Leu Val Ala Ala Ser Lys	
785 790 795	
gtg aaa gct gat aag gac agc ccc aac cta gcc cag ctg cag cag gcc	2689
Val Lys Ala Asp Lys Asp Ser Pro Asn Leu Ala Gln Leu Gln Gln Ala	
800 805 810 815	
tct cgg gga gtg aac cag gcc act gcc ggc gtt gtg gcc tca acc att	2737
Ser Arg Gly Val Asn Gln Ala Thr Ala Gly Val Val Ala Ser Thr Ile	
820 825 830	
tcc gcc aaa tca cag atc gaa gag aca gac aac atg gac ttc tca agc	2785
Ser Gly Lys Ser Gln Ile Glu Glu Thr Asp Asn Met Asp Phe Ser Ser	
835 840 845	
atg acg ctg aca cag atc aaa cgc caa gag atg gat tct cag gtt agg	2833
Met Thr Leu Thr Gln Ile Lys Arg Gln Glu Met Asp Ser Gln Val Arg	
850 855 860	
gtg cta gag cta gaa aat gaa ttg cag aag gag cgt caa aaa ctg gga	2881
Val Leu Glu Leu Glu Asn Glu Leu Gln Lys Glu Arg Gln Lys Leu Gly	
865 870 875	
gag ctt cgg aaa aag cac tac gag ctt gct ggt gtt gct gag ggc tgg	2929

Glu Leu Arg Lys Lys His Tyr Glu Leu Ala Gly Val Ala Glu Gly Trp
 880 885 890 895
 gaa gaa gga aca gag gca tct cca cct aca ctg caa gaa gtg gta acc 2977
 Glu Glu Gly Thr Glu Ala Ser Pro Pro Thr Leu Gln Glu Val Val Thr
 900 905 910
 gaa aaa gaa tag agccaaacca acaccccata tgtcagtgt aatccttgtt 3029
 Glu Lys Glu *

acctatctcg tgtgtgttat ttccccagcc acaggccaaa tccttggagt cccaggggca 3089
 gccacaccac tgccattaoc cagtgcagag gacatgcatg acacttccca aagactccct 3149
 ccatagcgac accctttctg tttggacca tggatcatctc tgttcttttc cgcctccct 3209
 agtttagcatc caggctggcc agtgctgccc atgagcaagc ctaggtagca agaggggtgg 3269
 tggggggcag gggcactcaa cagagaggac caacatccag tcctgctgac tatttgacc 3329
 ccacaacaat gggatcctt aatagaggag ctgcttgttg tttgttgaca gcttgaaag 3389
 ggaagatctt atgocctttc tttctgttt tcttctcagt ctttctcagt tcatcatttg 3449
 cacaaacttg tgagcatcag agggctgatg gattccaaac caggacacta ccctgagatc 3509
 tgcacagtca gaaggacggc aggagtgtcc tggctgtgaa tgccaaagcc attctcccc 3569
 tctttgggca gtgccatgga tttccactgc ttcttatggg ggttgggtgg gttttttgg 3629
 tttgtttttt ttttttaagt ttcactcaca tagccaactc tcccaaaggg cacaccctg 3689
 gggctgagtc tccagggccc cccaactgtg gtagctccag cgatgggtgct gccagggcct 3749
 ctoggtgctc catctccgcc tccacactga ccaagtgtg gccacccag tccatgctcc 3809
 agggctcaggc ggagctgctg agtgacagct ttcctcaaaa agcagaagga gagtgagtgc 3869
 ctttccctcc taaagctgaa tcccggcgga aagcctctgt ccgcctttac aagggagaag 3929
 acaacagaaa gagggacaag agggttcaca cagcccagtt cccgtgacga ggctcaaaaa 3989
 cttgatcaca tgcctgaatg gagctgggtga gatcaacaac actacttccc tgccggaatg 4049
 aactgtccgt gaatggtctc tgtcaagcgg gccgtctccc ttggccaga gacggagtgt 4109
 gggagtgatt cccaactcct ttctgcagac gtctgccttg gcatcctctt gaataggaag 4169
 atcgttccac tttctacgca attgacaaac ccggaagatc agatgcaatt gctcccatca 4229
 gggaagaacc ctatacttgg tttgctaccc ttagtattta ttactaacct cccttaagca 4289
 gcaacagcct acaaagagat gcttggagca atcagaactt cagggtgtgac tctagcaaag 4349
 ctcatctttc tgcccggcta catcagcctt caagaatcag aagaaagcca aggtgctgga 4409
 ctgttactga cttggatccc aaagcaagga gatcatttgg agctcttggg tcagagaaaa 4469
 tgagaaagga cagagccagc ggctccaact cctttcagcc acatgcccc ggcctctcgt 4529
 gccctgtgga caggatgagg acagagggca catgaacagc ttgccagga tgggcagccc 4589
 aacagcactt ttctctttot agatggaccc cagcatttaa gtgaccttct gatcttggga 4649
 aaacagcgtc ttctttttt atctatagca actcattggg ggtagccatc aagcacttcc 4709
 caggatctgc tccaacagaa tattgctagc ttttgctaca tgacgggttg tgagacttct 4769
 gtttgatcac tgtgaaccaa cccccatctc cctagcccac cccctcccc aactccctct 4829
 ctgtgcattt tctaagtggg acattcaaaa aactctctcc caggacctg gatgaccata 4889
 ctacagcgtg tgaacctocat actgggttaa ggaagtatca gcactagaaa ttgggcagtc 4949
 ttaatgttga atgctgcttt ctgcttagta tttttttgat tcaaggctca gaaggatgg 5009
 tgcgtggctt ccctgtocca gttgtggcaa ctaaaccaat cgggtgtgtc ttgatgcggg 5069
 tcaacatttc caaaagtggc tagtcctcac ttctagatct cagccattct aactcatatg 5129
 ttoccaatta ccaaggggtg gccgggcaca gtggctcagc cctgtaatcc cagcactttg 5189
 agaggctgag gtggtaggat cacctgaggt caggagttca agaccagcct gtccaacatg 5249
 gtgaaacccc catctctact aaaaatacca aaaattagcc gagcgtagtg acgggtgccc 5309
 gtaatccag ctactcagga ggctgagaca ggagaatcac ctgaacccca gaggcagagg 5369
 ttgcagtgag ctgagatcac gccattgtac tccagcctgg gcaacaagag caaaactccg 5429
 tctcaaaaaa aaaaaaaaaat tacaatggg gcaaacagtc tagtgaatg gatcaaatg 5489
 aattctctg cccagccggg cacagtggcg catgctgta atcccagaac tttgggaggc 5549
 caaggcggg gattgtctt agctcaggag tttagacca ggctgggcat catagcaaga 5609
 cctcatctct actaaaattc aaaaacaaaa ttagccgggc atgatggtgc atgcctgtag 5669
 tctcagctag ttggggagct aaggtgggag aattgcttga gcttgggaag tcgaggctgc 5729
 agtcagccct gattgtgcca gtgcactccg gcctgggtga cagagtgaga cccgtgctca 5789

```

aaaaaaaaa gattctgtgt cagagcccag cccaggagtt tgaggctgca atgagccatg 5849
atttcccact gcactccago ctgagtgaca gagcgagact ccatctcttt aaaaacaaac 5909
aaaaaattat ctgaatgato ctgtctctaa aaagaagcca cagaaatggt taaaaacttc 5969
atcgacttag cctgagtcac aacggttaag aaagcactta aacagaagca gaggctaatt 6029
cagtgtcaca tgaggaagta gctgtcagat gtcacataat tactttcgta atagctcaga 6089
ttagaatggc taccocattc tctagacaaa atcaaattgt cctattgtga ctcttctaaa 6149
aatgaagatg aagagctatt taatgacaca ccttggatta aaacgggaat cacatcttaa 6209
agctaaaaat gaacctgcaa gccttctaaa tgagtcactg agcatcacta gtgacaagtc 6269
tcgggtgagc gtaaatgggt catgacaaga tgggacagca acaaaatcat ggcttaggat 6329
cgacaagaag ttaaaaaaca gctgcatctg ttacttaagt ttgtaagaca gtgccctgag 6389
acctctagag aaaagatggt tgtttacata agagaaagaa ggccagacat ggtgtctcac 6449
acgtttaatc ccagcacttt gggaggcagg ggcggtgga tcacctgagg tcaggagttc 6509
aagactagcc tggccaacat ggtgaaacc cgtctctact aaaaatacaa aaattagccg 6569
ggcatggtgg caggcgccca taatcccagc tactggggag gctgaggcag gagaatc 6626

```

<210> 5
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220>

<223> PCR Primer

<400> 5
 tgaccgaggc ctgtaagca 19

<210> 6
 <211> 22
 <212> DNA
 <213> Artificial Sequence

<220>

<223> PCR Primer

<400> 6
 ttctcaaggc ttccctcttc ct 22

<210> 7
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>

<223> PCR Probe

<400> 7
 tggcagggaa accctcgct acc 23

<210> 8
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220>

<223> PCR Primer

<400> 8
 gaaggtgaag gtcggagtc 19

<210> 9
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> PCR Primer

<400> 9
 gaagatgggtg atgggatttc 20

<210> 10
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> PCR Probe

<400> 10
 caagcttccc gttctcagcc 20

<210> 11
 <211> 65454
 <212> DNA
 <213> H. sapiens

<220>

<400> 11
 tgatgtcatt aaggaagtac aaggggaattt gcttagaagt tggaaaatgc ccaagagtgt 60
 gggaaaacaa agacttagtg accaccgccg gtgctggcca gccggagaag gctctgtgga 120
 aggtttggag gggagagagg ggcagctgga tgctcttggg ccacggtcgc tccctgatct 180
 ctgcgcctct tcctcctgct ccgggagaaa taatgtttcc ctgggggatg aaaagcatct 240
 ctttgtgceg gctttaattg ccatgtttgt gtgccaaggg agtgagtggc aggcgggagc 300
 agcagctggg cacagccaat gccaggcagt ggtgccact ccctcaggac ggcccagcca 360
 gctggctcct gggagcgtg cccacctctg cccccagctg ggccctgca gaggaaccga 420
 ccaccctggg ggctggggga ggttggtgct aggaggagaa aggggcgggc atctgggagg 480
 gtctcagcca ctctcagagg cttattcctc tcctcctcct ttccctcccc ccttcttgtt 540
 tttcagactg tcagcatcaa taaggccatt aatcgcagg aagtggctgt aaaggaaaaa 600
 cacgccagaa atatcctttt ggatgttgct tggagaccg accctgaggg aggtcagctc 660
 atggggactg aggtcagggc caggctgcct tgctcagctc caggaagggg caaccctgca 720
 caggccaggt ccctgcagct tctgatgacg gcagcttctc agagagggct ggctgcagag 780
 accacagacc ttcaggggtg cagacaccaa aaaggctgtg gagcccaggc ctttcaactt 840
 gccaaagatc ctgctccttt ccttaaggac ttaagcactc ctttttttct ttttccaaaa 900
 ggggtcttgc cgtgttgccc aggctggagt gcaatggcgt gatcatagct cactgcagcc 960
 tcaaactcct gggttcaogc aatcctctcg tctcagcctc ccgagtagct gggactacag 1020
 gtgtgcacca ctatgcctgg ctaattttatt ttatgatttt tagagatggg gtattgctca 1080
 ttgcccaagc tggcctcaag caatcctccc tcctctgtaa ccccaaagtg ctggaattac 1140

aggggagagc cactgcacct ggccgactca agctttgtag aacctcatag tcacttgaaa 1200
 gttactttcc tttgagagac ctccctggggg tcaggagggga tcttcaccta tattcaaagc 1260
 cctccagggtc ctttctttgc ctttacagga acacagggac cactcccctg ggggttgc 1320
 aatcaatagt tatctccttt tctgagcatg aaagcaaaaa aaaaagaaaa agagagtttt 1380
 tttttttttt cttttttgag acagagtccc actctgttgc ccaggctgga gggcagtg 1440
 atgttctcgg ctcaactgtaa cctctgcctc ctggattcaa gcgattcttg tgcctcagct 1500
 tcccaagtag ctgggatgac aggcgtgtgc caccacaccc ggctaatttt tgtaatttta 1560
 gtaggaccgg ggttttgcca tgttgccag gctggccttg aacgcctggc ctcagcctcc 1620
 taaagtgctg agatcacagg catgagccac catgcctggc cgaaaaaag aaagtcttag 1680
 cttcagaggt tggttggcct taaactgagg caggggctct ctacctcgt gaacaggttc 1740
 aacctatgcc agggggaaga agacaagagc cttgaagtgg attagggat ggggtgacttg 1800
 aaagccctct gtaagcccac cacaccagc agcagcctgt ggctttgtag agaggtgag 1860
 gacctgctg gctgatgaat ccctaggaat ctgcctttga gttgcagaat ccaggaactg 1920
 gagcgttaa tccccaaagg ctgaaggaga gagtctgcca ggggggtggc taagctttta 1980
 actctctgtg tgctgggcca gagcaagggt gagttccggg caagcagagt tgggactttt 2040
 tttttttttt ttttgagaca gagtcttgc ctgtgccta ggctggagtg cagtggcatg 2100
 atctcagctc accgcaacct ccgccttcca ggctcaagag atccacctgc ctcaacctc 2160
 caagtagctg ggaccacagg tgtgtatcac taccctagc taatttttgt attttttagta 2220
 gagacggggg tttgccaatg tgctcaggct ggtttcaaac tcctgagctc aagcaatcct 2280
 cctgccttgg cctcccaaag tgctggaatt acaggcatga gccaccgccc ccaacttcgg 2340
 aggccgcacc cagcctcaga ggctggactt ctgtctgcag tgggaagctt ctcctcacct 2400
 gggctcttgg ttactgtgtc acagccaggg aagccagagc agcttgtccc gaagccttct 2460
 agcctcaacc cctagaggcg accctccagg ttagatagcc agagaagccc agcattgctt 2520
 agtgtgatat atgaaactgct gtttgctaga gggaaagcag cttgccttta gtgggaagat 2580
 gctttggctg gaattaggat ggcggctgca gaacctctct ggggtaacct agaagtccag 2640
 ccctgtgcca gctagctag atcttgcct ttttttcccc agaaagacac aggccctt 2700
 gtttcctgca tcacgtttgg tacccttggc atgattcaaa gcaaaccaag aaaagcctt 2760
 attttaaacg gccagtgga ttttctcaga ggaaatgact cagccctgac ccttacctac 2820
 gaataatttg tgatcaatct gataaaagat atgagtaagg gagagacagt tatcatcagg 2880
 cagtgtcctg ttcaggacaa gacatgggtg cagagaggcc aggcctgcag cggcagggtc 2940
 agtcacatg ctgccgggag gctccctggg aggggggaa gggccctgca aggagattct 3000
 gacattctgc cactggaact ctccggctca ctctgggacc ttgagtgggg cctccatgag 3060
 tcttaggtca ctttctgttc aatcaaaagg tggacaaaga agatttcagg ccagggtgtg 3120
 tggctcatgc ctataatccc agcactttgg aagggtggagg cggatggatg acctgaggtc 3180
 aggagttcga gaccagcctg gccaacatgg taaaaccccg tcctactgaa aaaacaaaa 3240
 ttagccaggt gtgggtggcag gtacctgtaa tcccagctac ttgggaggct gtggcaggag 3300
 aatcacttga acctgggagg cagaggttgc agtgagccga gatcgtacca ctgtactcca 3360
 gcctgggcaa caagagcaaa actctgtttc aaaaagaaaa agatttcaaa gatttaacag 3420
 tccagcccta acattctatg agtctctgat tcataatttc taaatggtga ataaatggtg 3480
 aaaaatcaac tccatacagt agggtttgaa atgcctaatt aaagatatoc oaaagctttt 3540
 tttttttttt ttttttttga ggcagggcca tgcaactgtc cccaggctgg agtgcagtga 3600
 cacaatcagt tggaggctca ctgcagcctc caacttttag gctcaagtga tcctcccacc 3660
 tcagcctcct gactagctgg gactataggg acatgccatc atgcctggct aattttttat 3720
 agagatgggg tctcactgtg gccaggctg gtcttgaact cctagctagg ctcaagcagt 3780
 cctcccctg ggcctcccaa agtgctgcaa ttacaggcgt gagccaccat gccagcgcga 3840
 tttatctatt atcaggctct ttatcagctg ccttagatac aaactgaata agcattagga 3900
 agtctgactt cagaaatgct gcttacacat ttatcacaga tgctaaatg gtgaacagg 3960
 tggcccgcga ccttgggtatt cactcctgga gggctgggga ctgagtctta tttttcttat 4020
 tttctttttt ttctccctga atggatttta caggctttga cacatacttg gggcataaaa 4080
 atgcttatat gtgaaagaat gccagaaagg cattatttct gaattatttc agaaatogct 4140
 cctacaccaa ggccaggcac agtggctcac acctgtaatc ccagaacttt gggaggccaa 4200
 ggtgggtgga tctcttgagg ccaggagttc gagaccagcc tgagcatcat ggtgaaacct 4260
 cgtctctact aaaaaaaca aaaattagcc gggcatggtg gcgggtgctt gtgatcccag 4320
 ctacttggga ggctgaggca cgagaatcac ttgaacctgg gaggtggagg ttgcagtgag 4380
 ccaagattgt gccactgcac tccagcctgg gcaacagagc gagaccctgt ctaaaaaaaa 4440
 aatagaaaga aagaacgaac gagagagaga gggaggggag gaaggaagg agagagagaa 4500
 agaaagaaag gaaggaagga aggaaggaag gaaggaagga agaaagaaag aaagaaagaa 4560

agaaaagaaag	aaagaaagaa	agaaaagaaag	aaagaaagaa	agaaaggact	cctacaccaa	4620
aagctgattt	tagcttccag	acccaaacat	ctgtttaagc	ccaccctct	ctaaatgaag	4680
atgtcatcga	ggaaaggagc	tttatgggat	cttccagttc	agaagcaaaa	taacatgtgc	4740
tctgggtgag	tcatttagtt	tcctacctga	gatgagtcct	tctgccctaa	atgacagtgc	4800
acttgcgacc	tagccttgct	gagtaaagcc	taaagttocg	ggcacagagt	tgtcctaaca	4860
actgacagtg	cccagttact	gggtctgtga	atcttgttga	agcacacaat	tacaaaagtg	4920
gaaattctct	cctcactcca	aaaagagctc	actaatgaca	gagatctgga	atagccgaac	4980
ttaaaatcat	gtgaggccag	agggacatth	gtttatthtt	gagatggagt	ctcgctctgt	5040
cgcccaggct	ggagtgcagt	gggtcaatct	tggctcactg	caacctccat	ctcccaggth	5100
caagtgattc	tcctgcctca	gcctcccagag	taggtgggat	tacaggcatg	caccaccatg	5160
cccggctaata	ttttgtactt	ttagtagaga	cggggthtca	ccatgttggc	caggctggtc	5220
ttgatctcct	gacctogtga	tctgcccacc	toggcctccc	agagtcctgt	gattacagcc	5280
gtgagccacc	acgcccaccc	ctgatggaca	tttattacat	gcaaactgac	aactatctgc	5340
agagggtggg	tgtgtttgga	atcaggcctg	gggagaagtg	gttccaggat	cccataatg	5400
gggaattcaa	ggtgtggaga	gattgaaggt	gacctgaac	agggccaggc	aggtttatgg	5460
aagaagtccg	gtgttcaggt	ttatggaaag	ggctthtgtt	gttgctgttg	ttgttgaggc	5520
agagthtttc	tcgttgccca	ggctggagtg	caatggcatg	atctcggctc	actgcaatct	5580
ctgctcccgc	ggttcaagca	attctcctgc	ctcagcctcc	tgagttagctg	ggattacag	5640
catgctccoc	cacactgggc	taathttgtg	ttthtagtag	agacggggtt	tcaccacatt	5700
ggtcaggctg	atctcaaaact	ctgacctcag	gtgatccgcc	cgcctcagcc	tcccaaagtg	5760
ctgggattgg	aggtgtgagc	caccgcgctc	cacctgatg	gacathttat	acatgcaaac	5820
tgacgactat	ctgcagggag	tgggtgtgtt	tggaaatagg	cctggggaga	agcggthcca	5880
ggattccata	tatggggaat	ttaaggtgtg	gagagattga	aggtgccct	gaacagggcc	5940
aggtgggtht	acggaagaag	tggggtgttg	aggtthacag	aaagggcatt	ttaaacaaag	6000
attgctgctg	tcaaggcaag	caaaaggctg	gctaaatgga	aaggggctgg	gagatgctgc	6060
ctggaggagg	agccaagggg	tcaaggtgca	ggactccctg	cagggacca	ggtccttgat	6120
gaggagcaga	agcacactct	cactthtttt	thttthtttt	thttgagatg	gagtcttgct	6180
ctgtcaccca	ggctggagtg	cagtggcgtg	atcctggctc	actgcaacct	ccacctcccg	6240
gattcaagca	attctcttgc	ctcagcctcc	tgagaagctg	ggattacagg	cgcccaccgc	6300
cacgcccagc	taathttgtg	thtttagtag	agatggagtt	tcgcccgtgtt	ggtcaggtgc	6360
atcttaaaact	cctgagctca	ggtgatccac	ccacctggc	ccccaaagt	gctgggatta	6420
taggcgtgag	tctctgcacc	tggcttgcac	accttactt	ttaataatgc	ataagaaaga	6480
acctgaggaa	gcccacaaat	gttgtthtaa	agttaggtgt	cataaggcca	gacacagtgg	6540
ttcatgcctg	taaccccagc	actthtgogag	gccaagaggg	gaggaccact	tgaggccagg	6600
agcttgagac	cagcctgggc	aacatagcaa	gatctcatct	ctaccaaata	thtaaaatt	6660
agccaggcac	agtgggtgat	gcctatagcc	ccagctactt	gggaggctga	ggtaggagga	6720
ttgcctgagc	ccaggagtht	gaggttacag	tgagctatga	tcacaccact	gcactccagt	6780
ctgggtgaca	gagtaagacc	ctatcttaac	caccacaaca	aatagctgta	gthattcccg	6840
aaggacacat	catcaactca	gcattctcgt	agaaaggaca	acccaatacc	accgtggcct	6900
acagggtatt	tgtgcaaat	agaaaaagac	acacctctct	ctcaagacgc	ttacatctcg	6960
ggaaattgtc	tactcaagtg	gattthatta	aaataagtat	tcccatctgc	cctaaaactc	7020
ccagaaggaa	tgtacagtht	tgtgggtgtct	tagctctgag	cggcgctccc	tccaaggcac	7080
tgccaaccac	gacttgcact	actgctcgtt	cccctgcagt	ggctgtgtgc	aggtthgtc	7140
ccctgcagtg	gctgggtgca	gcctctatcc	cctgtgatgg	ctgggtgcag	gctcagthcc	7200
ctgcagtgcc	tgggtgcagg	ctcagthccc	tgcagtgact	ggcagagccc	atccgthgtt	7260
ttccataacc	ccccctcacc	gtgcatactg	ggcaccaccc	atgagaaagg	ggcacagacc	7320
ttctgggtctg	ttgtcaaccg	cctgcctctg	tctagcaaac	cagthgctctg	ctggaagtht	7380
tgccatgtgt	tccacaaact	cctccgagat	ggacaccocga	acgtgagtht	ctggggctat	7440
ggggtggcag	ggagccaggg	atcctthgtg	ggaaagtgc	tgctggggcc	acagaggctg	7500
ctgtcctctc	ccacactgcc	ccctthctct	ggcctthggg	thcccattha	gagthgggtg	7560
agtctccctc	ccatcagcct	gtccccctgc	cctaggtthca	cctccgcctc	tctccatcct	7620
ccttccctth	gtctthctth	cctthctccat	tctctcaocag	gctgtthctth	tgcccctgca	7680
ggtctggaag	gactctctga	gatacagaaa	tgaattgagt	gacatgagca	ggatgtgggt	7740
gagthttgag	atgtactcag	gagccacctg	cttctcctth	cctthctcag	aggccacaga	7800
gcaaggactg	gagggtaaaa	taaatctatc	tctctcagct	tgttgaggat	thctcccctg	7860
ggtgccagag	acggactaaag	gatgccccca	aagagthacca	tgatathccat	agthctgcct	7920
ggggtctctg	gacgctthcag	gaaagthcca	gggcctggga	gctthctcctg	ccaatgaaaa	7980

ctcataaact catgtcacaa agctgcccac ggttgggatc cctcagcagt agctacatgg 8040
ccatgtttcc tgcctttttt tttttttttt tggtagagac ggggtcctgc tatgctgccc 8100
aggctggtct tgaactcctg gcctcaagtg atcctcccac ctcagcctcc caaagtgtcg 8160
ggatcacaca catgagccac tgtgccacgc ccacagttcc tacttctctt ctccctccac 8220
tggccocggt ggagtcatag tgcattgaga attagaatta gcatacccaa ttcttaatgc 8280
cagagctttt tctttcctgt ttagagatta tcagtgcctt aggatgaggg gggaggtagg 8340
tggcaatata acagtcaatg gaaaagagca tttctgtaca cacaccacaa tcaccccaga 8400
ttctgatgtc cgtgggtgtg tataatggaga ggggtaatct gcattttgaa gaatctctct 8460
tggccaggcg cagtggctca cgcctgtaat cctagcactt taggaggccg aggcaagcag 8520
atcacctgag gccaggagtt tgagactaac ctggccaaca tggtgaaacc cgggctctat 8580
taaaaataca aaaattagct ggggtgtggtg gcgtgtgcct gtaatcccag ctactcggga 8640
ggctgaggca ggagaatcgc ttgaacctag gaggcgaagg ttgtggtgag ccaagattgg 8700
gccactgcac tccagcctgg gtgacagagt gagactctgt ctcaaaaaaa aaaaaaaaaat 8760
ctccctgggt gattctgata tgttcttttg cctctaccag ttgagaacta ctacttttagc 8820
ctttttgcac agtcttaoct gtatctgtct gcacctataa aatggtggca actatgtagt 8880
cctatctacc tatcataggt gaggaaacac agggccagaa agggaagcag cttgtcctaa 8940
agtcacaaaa ttaggtacca gaacaaagtt gagagttgtc ccagatttc ccaactgtgt 9000
gctagaaaaac gttccctgta acatgtaaag ttttaggtctt tttccttttt gttagaattt 9060
gcctgggtata tatttctoca tctttttact ttaagcattc ctatattttt atgttttaga 9120
aataaacact ataaataaca tttggttgga ttttgtgtgg tttgtggttg ttgttttttt 9180
atccaggatg gtagtctttg tcttttaact gggcatttag tctatttaca tttattatgc 9240
tacgtggtat gtttagattt atgtctccca ttttattcta tgttttctat ttgtcccatc 9300
tgttctactt ctatttctct cattccttgc cttcttttgg attggtctgg tgtttttttc 9360
tcattctggt ttttctcoat tactatttgg gaggttaaaa aataaatcta ctttttttgt 9420
ttttggcatg ttctctaata attacaaaat gcttgcttgt tttatctagg tctaccttc 9480
gtatctttgt ttctgtata acacaaagac tttagaatat tttagttcca ttcagtcctt 9540
caatttatct gtttagcataa ccatgtattt tggttttggg tatgtttatc tttaaattct 9600
gtaagttatt gttttatttt ccttgtttta ttctgtgctc tccatagacc tttcatctga 9660
gatcaacttc tttctgctcg aattatttcc tttctaactt ccataattga gagcctactg 9720
tggcgtgaca ctgtgtttgc ttgtttaaga atgtcgtttt tggccaggcg cagtggctca 9780
tgcctgtaat ccagcactt tgggaggcgg agggcggcgg atcacgaggt caggagattg 9840
agaccatcct ggctaacaag gtgaaacccc cgtctctact aaaaaatac agaaaattag 9900
ccgggcgtgg tgggtgggtgc ctgtagtccc agctactcag gaggctgagg caggagaaag 9960
gcgtgaaccc gggaggcgga gttcacagtg agccgagatc gcaccactgc actccagcct 10020
aggtagacaga gagagacttc gtctcaaaaa aaaacaaaaa aaaccacaaa aaaaagtctt 10080
tttttttttt tttttttttt tttttttttt ggagacagag tcttgctctg ttgcctgggc 10140
cagagtacac tggcataatc tcagctcact gcacccttcg ccaccaggt tgaagcgatt 10200
ctcctgcctc agcctcccga gtagctagaa ttacagtgcc taccaccaca cccagctaat 10260
ttttatattt tagtggagac acggttttgc catgttggcc aggctggtct cgaactcctg 10320
acctcaagtg atcttccctc ctoggcctcc caaagtgttg ggattacagg catcagccat 10380
cgcgccacgc cccttttcag attttggcta gacctgtgct agactttctc atcctacccc 10440
tcatgtctct aaaccactct ttcagttttt tctacctatt tgtctttgct gtattctgga 10500
taatttttat caactctagc cttcagtttg cctcctctcc agttgggtct aatgtgttat 10560
taagcttctc cataaggttt tcaatttcoat ttagtacagt tttcttttct tttcttcttc 10620
ttcttcttct tttttttttt ttttaagacag agttttgctc ttgttgccca ggctggaggg 10680
caatggcacg atctctgctc actgcaacct ccacctcca ggttcaagtg attctcctgt 10740
ctcagcctcc caagtagctg ggattacagg catgtgccac cagcccagc taattttgta 10800
tgtttaatag agatggggtt tcaccatggt ggctaggctg gtttcgaact cctaacctca 10860
ggtgatccaa tgcctcagcc tctcaaatg ctaggattac aggcatgagc caccacacct 10920
ggccaatta gtacagtttt catttcttga agttctatct atttattttt tttcaaatct 10980
gcttgttttt tgcgggggaa ggtggttgtt gttgotgttt gagacagggc ctcaactctgt 11040
caccaggct ggagtgcagt ggcgcgatca tggcttattg cagcctcaac ctccctgagct 11100
caggtgatcc ttccacctca gccttccaag tagctgggaa aacaggtgtg caccacaaa 11160
cccagctaat tttttagacg gcaggtttcc gccatgttgc caaggggtgg ctcgaacctc 11220
tggccttaag cgatctgcca cctcagcctc ocaaagcact gggattacag gcaggaggag 11280
ccactgtgtc tggccctgct tctcctcctt tacaatttat tatgcctgc atatatattt 11340
aaacagtctc ttatttcttt agatatatga accaaagtca tctcatatat gttttgtaat 11400

tctattatct aaagtcttta gagatthgtt tctgtcctcc gtcccttttg gtgcattttg 11460
cccagctgtg ttagttttct ttatgttcgg ttacttttcta ccctaagctg ctcatcttcc 11520
ttggaattht atctgtggag aattctthtga agtctgtaat gaaggcaggt tccctccagag 11580
agaattggca tttgcttcga ttgagatcac tcgaaattaa attctcagct taaggthttg 11640
tgagthtagg gcagatthtc cttccccaat cacagcggtg atthtgaggcg gtagaattct 11700
tcatagtcct tagggaggac ttcctcagtg caggagtaca agcgtthgtt actthattat 11760
ttatctthtao tttccctata ccaaggaggc agtctthttg gtgtgtgttg gtggthgttg 11820
tgggtthttg agacagthtc ttgctctgtc accaaggctg gaggatggg gacatgatca 11880
cagctcactg cagcctcaac ttcctgggct caagtgatcc tcccacttca gcctctgaa 11940
tagctgagac cacagggtgca caccacatg cccagctaat thtttaattt thttgtagaga 12000
tggggthttcg ccatgthggc caggctgggc tcaaattcct gagctcaagc aatcctccct 12060
cctcggcctc ccaaagtgct gggattacag gcatgagcca ctgcacctgg ccagactcaa 12120
cattcttaag taggaagctt tgtccagaat ttatcaaagg cagtccttat gthttgtacaa 12180
aagatthgtc tataaggaat gthcattgaa atggtthtat atagcaaaaa thtcaactgg 12240
thtgaattgg ctthcagctca thtgaattta ttgagtcagt tcaataaaat acaataatac 12300
tggtcagtca aaaaaataa tgaagaaaag thtaagaagt tgaataaaat tgcagcatat 12360
taagcaagaa aaatagattg atctatthtt acatthaaaga ataaaggtag atggctgggt 12420
gcggtggctc acgcctgtaa tcccagcact thgggaggca gaggcgggca gatcacctga 12480
gatcagaagt tcgagactag tctggccaat atggtgaaac cccatctcta ctaaaagcta 12540
caaaaattag ctgggcgtgg tggctcacac ctgtagtgtg atcctagcct thtggggggc 12600
caaagcaggt ggatcactca aggccaggag thttgagacca gcctgggtaa catgggtgaaa 12660
ccatatctct agcaaaaaata caaaaattga ccaggcgtgg thgtgctgtc ctgtaatccc 12720
agctactcgg gaggctgagg caggagaatc actthgaacc aggaggtgga ggctgcagtg 12780
agccgagatc atgccactgc actccagcct gggcaacaga gctagacccc atctctaaat 12840
aaataaaata aaggggagaa aacaaagaag atagattctt taccagagaa thcctggctg 12900
cagatctctt gactgthtat thctthgttgt tgaactctgt tcccctcctc thcctaaaag 12960
ggccacctga gcgaggggta tggccagctg tgcagcatct acctgaaact gctaagaacc 13020
aagatggagt accacaccaa agtgagtctc tgcggacagt tctgcccga cccgcccctc 13080
ccctgctcca tcccttcagc cctccctgg gctcatthgt cagctctthc aggtaataga 13140
cagcccaggc thctgaggaa gtgtgcacat catgtacca agctgtgaga gaggaaaagc 13200
accgccaggc ccacggggtg tgacgaaggc tgggathttg gccctgtctc thtgcaccct 13260
thgttcccca thttgcagct agaaaagaaac ctggccaggc acggthggctc acacctataa 13320
tcttagctct thcggaggcc caggcggatg gatcaccoga ggtcaggggt tcaagactgg 13380
cctggacaac atggcaaaac cccgtctcta ctaaaattac aaaaattacc catggthgtg 13440
atgcatgcct gtagtcccag ctactcggga ggctgaggca ggagaatcgc thgaacctga 13500
gaccaggagg cagaggtthc agtgagccaa aacctgcca ctgcactcca gcctgggaga 13560
caaatgaga ctctgtctca aaaaaaaaaa caaaaaaaag aaagagtaac cctaggcact 13620
atattgcctt atgaaaatgt tacagaagca thgathttac thgttaataaa aacaattcct 13680
ggctgggcac agtggctcat gcctgtaatc tgagcattat gggaggtcga ggctggagga 13740
thgcttaaga ccaggagtht gagaccagcc tgggcatag caagatcctg thctacaaa 13800
aaaattthaa aattagccgg gcgtggtagt gtgcaactgt agtcccagct actcgggagg 13860
ctgaggtggc aggatacctt gagcctggga agtcaaggat gcagtgagct atgatcgc 13920
cactgccctc ccgcctgggc aacagagtga ggcctgttht cacagaaaaa agagagaaga 13980
aggtgacaca atggatagtg tgaggtgggg gaggcgagca gctgcccgggt gtacaggggt 14040
ctggcaactg tgggagacaa cggggatcag gcagaggggt gggcacatag gaggcaactg 14100
aaaccagaac tgaggctggg ccaggccagg gtgtcagagg ctgaggcggg gactcgggct 14160
gccctggggc tggattgcaa ggogaagaag caggagggag aggctgagag thtggagggg 14220
agggggcttg tagcgtgtga thggagggag gaccgggctc tggtaaggt thcatgagga 14280
ggcggggaaa ggtgagaact ggaththaaga ththttgggg thgggcgcag thgctcacgc 14340
ctgtcatccc agcathttag gaggccaagg cgggcggatc gtctgaggct thgagthcaa 14400
gaccagcctg gccaacatgg caaaactccg thctacthth aatacaaaa agaaattagc 14460
thacgtctgt agtcccagct actcgggagg ctgaggcagg agaatcgtt gaacctggga 14520
ggcggaggtt gcagcgagcc gagatgctgc ctctgcactc cagactgggc gacagagcaa 14580
gactcaatct caaaaaaaaaa aaaaaaaaaa cgtthttggga gtagagattg aaagtthta 14640
atthgagctgg atgtgggtgc thcacactct aatctccaca caatggaagg ccaaggcagg 14700
aggatcactt gaggccagga gthggagacc agcctgacca acatagthaa accccatct 14760
ctactaaaaa cacaaaaatt agctgggcgt ggtgacacgc acctgthac ccagctactc 14820

aggaggctga ggcttgagaa tggcttgaac ccaggaggcg gaggttgacag tgagccgaga 14880
 tggcgccact gcactocage ctgggtgaca gagcaagact ctgtctcaaa aaaaaaaaaa 14940
 aaagtgttaa atggaaagtg ggaggaggaa gggcctaggg taggatttta ggtggaaaag 15000
 gaaagaaagg aggggtgtaga gaaagcaaga ctgggaagaa ggttaaggag gaagaggact 15060
 gggatgggag ctggggcagt ggggctgagg ggctgataga ggccaaaggg gtccctttgc 15120
 cttgtatata ggaagacagc agagttcagt ggcatcagaa gcagctgggtg ccaatagcca 15180
 gtcccctagc ttttaatagc atggaagggc agaagggaaat ggtcacgtgg gcttggagaa 15240
 ctggggctga agaggatgca gctggccttg cttccagagg ctctcgagct cactgctgtc 15300
 ttgtttttcc tggttttaga aagcagagaa aaaggggaag ctggtattgt ccaagtgtgg 15360
 aggagcaaag gaottttoca gtttttagca attagcgata caagagtggg gaaggaggct 15420
 gggcagagtg gctcacacct gcaatcccag catttgggga ggccaaggct tgaggtgaga 15480
 ccagcctggg caacacagtg agaccctgtc tcttaaaaaa aatttttttt ggccaggcgc 15540
 ggtggctcac gcttgtaatc ccagcactct gggaggccga ggcgggtgga tcacctgagg 15600
 tcaggagttc aagaccagcc tgggcaacat ggtgaaaccc catctccact aaaaatacaa 15660
 aaaaattagc cgggcatggt ggtgctgccc tgttaatccc agctactcag gaggctgagg 15720
 caggagaatc gottgaaccc aaggggctga ggttgacagt agctgagatc atgccattgc 15780
 actccagcct gggcaacaga gtgagactcc gtctcgaaaa caacaacaaa aatatcccgc 15840
 acacattaaa gaaaaattca ttctctggct gggcgtgggtg gctcacacct gtaatcccag 15900
 cactctggga ggccaaggcg ggcggatcac ctgaggtcag gagttcaaga ccagcctggc 15960
 caatgtggcg aaaccccctc tctactaac acacacacac acacacacac acacacacac 16020
 acacacacac acacacacaa attagcccgc atgggtgtgt atgcctgtag taccagctac 16080
 tgaggaggct gaggtggagg atcacttgag cccaggaatt cgaggctgca gtgagtgatg 16140
 atcatgccac tgcactccag cgtgggcaac agtgaggcca tgtctctaaa aaataaaaaa 16200
 agagagagaa acaagagaga gctggtttcc ccctcctctg ccatgtaagc aacgtaatca 16260
 ggatgaagcg cctcaccagg caccgaatct gtcaacaccc tgactttgaa ctccctggcc 16320
 togagaactg aaagacactc tctatgggtt aagccacca gtgcatggta tcttgttata 16380
 actgcccagc ctgactgaga cggacgttca ggacagagag cgtgaatgca tagtgacacc 16440
 agotgtgagt ctttctccag ggacagtcgg cagccggccc taggtgcaga gccgatgaca 16500
 aggaccagg ctctcagcag gtcttccaag cagtgtggtg gaaaggcagg cagggtgtgg 16560
 ggaagtggag ccaggccacc agtcatgatg tcaagactga gccaggaagc aaaggcaggc 16620
 agagagatgg ggaggagagg gagcaggagg ggactggcca tctctgagac agaagcgtga 16680
 gtagtgggtg gacttgaggg caggagagga ctgaaaggc agaggcctgg gcgatgcagc 16740
 cagagaggga gatgctggtg tggggaggct tgggcaggga tgttttaggt gatggcagag 16800
 tctggagtgg ggatggagta gagtgaaagg tgctgaaatt gaggtcagag gttgtaactc 16860
 cagcgtgtat gacttggggc aaaggaaaac tgtgtcccag gtgccaggtc aagcctcaga 16920
 ggccttggca ccatggagcc caggagcaaa gtctgcaatg gggatttttt tttctttttt 16980
 ttgcggtggg ggagacggag tctcactctt gcccaggctg gagtgacagt gtgcatctc 17040
 agctcactgc aacctccgcc tcccagcgtc aagcaattct gcctcaacct cccaaggagc 17100
 tgggactaca ggcgtgcacc accacacccg gctaattttt gtatttttag ttgagacagg 17160
 gttttgccat gttggccagc ctctgtctcta attcctgacc ccaagtgatc catctgcctt 17220
 ggccccaaa gtgctgggat tacagggtgtg agccaccgca cctggcctag agcctaaagt 17280
 atccacttgt acatgtagat gccccacgat ggaatggcca cccatctctg tggccttttc 17340
 cctttgccac agggacaaac cacacagatg acaggatcat gctggctgta gatactcagc 17400
 aatgattgat gataccagcg atttttcttt ttttcttttt tgtttgtttt gaggtagggt 17460
 ctcaactctg aaccaagctg gagtgcagtg gcctgaaact gtaaacttga actcccgggc 17520
 tcaagcaatc ctcccacctc ggctcccaa gttagctgga ccacaggcgt gtgccaccac 17580
 gccggctga gagagggtc tcatgtctt ctgcccgtgac tcccttctc tgccctccct 17640
 ccagaatccc aggttcccag gcaacctgca gatgagtgac cgccagctgg acgaggctgg 17700
 agaaagtgac gtgaacaact tgaagtggc tcctgccctg agcccaggga gggagaaagc 17760
 ttttgtgaat gctgacactt ctcataaggg tcatggaggg cctgatgggg ggaggccgtg 17820
 gctgggatgg ggaccaaac ccctgggtga cttggccttg gggctactta tttattgggtg 17880
 gtgectcatc cagaaccctt gcctggctat tcacaccccc aaagctttcc tgtctgtctc 17940
 gctttctgcc ctctgactcc aacctggta cctatcctct ccctctgtca ctgtoactgt 18000
 ttttgttttt gttttttgag aggcagtctt gccctgttac ccaggctgga gtgcagtgac 18060
 ttgatctcgg ctactgcag cctccacctc ccaggttcaa gtgattctcc tgccctcagcc 18120
 tcccagtag ctgggattac aggcattgag catcacaccc agctaatttt tatattttta 18180
 gtagagatgg ggtttcgccg cgttggccag gctggctctg aactcctgac ctccaggtgat 18240

ccgcccacct	cggcctccca	aagtgtctggg	attacaggcg	tgagccaccg	ccccgacct	18300
gtcgctgtca	ctggtgactt	caccaggctg	catggccata	ataccacaaa	ggctaagact	18360
tggagctgga	gttgtgtgtg	tgtttgcgca	tgacatgag	cattggagac	tggagtagcg	18420
tagagcgtgg	gggaggggac	aggtaacaga	ccggcctcag	gctgtggagt	gtaagctctc	18480
tttctctctg	ggctccagttt	ccagttaaca	gtggagatgt	ttgactacct	ggagtgtgaa	18540
ctcaacctct	tccaaacagg	tgagtctctt	ccctcccgtc	taaccacagg	tctcatggga	18600
actacctaat	tcctagtcct	cctctccctg	caaagtgtgc	agcacaaggg	gtagggaaat	18660
ggagacattc	acaccccac	tctggtctct	ccaacctcgt	tgacgggagg	gactgaaact	18720
cttcagtatt	tttcttttta	agagacaagg	tctcggccgg	gtgcagtggc	tcgcacctgt	18780
aatcctagca	ctttggggagg	ctaaggtggg	cccatcactt	gaggccagaa	gttcaagacc	18840
agcctggcca	acatgggtgaa	accctgtctc	tactaaaaat	ataaaaaatta	gccgggcatg	18900
gtggcacatg	cctgtaatcc	cagttacttg	ggaggctgag	gcaggagaat	tgcttgaacc	18960
caggaggtcg	aggttgcagt	gagccaagat	catgccattg	cactccagcc	tgagtgatac	19020
agcaagactt	catctcaaaa	gaaaaaagag	agagagagag	agaaggtctc	actctgtcgc	19080
ccaggctgga	ttgcagtggc	atgttacggc	tcactgcaac	ctcaaaactac	taggctcaaa	19140
tgatcctccc	acatcagcct	cccaagtagc	tgggactaca	ggcacgcacc	accacatctg	19200
gctgattttt	tatatTTTTT	gcagagatag	gggtctcact	gtgttgccca	ggctgttctt	19260
gaactaccag	cctcaaggtg	tcttcccAAC	tcagctctccc	aaagttctgg	gattacaggc	19320
gtgagctacc	acacctggcc	tcctagatat	tttccaagcg	ggtggttca	ogtctgagca	19380
gaaagaccac	acactgcccc	ttcctactgg	cctctctcct	gaagcgcagg	cctccaaaag	19440
cccaaagaca	ggtcttacct	ctttgccagc	cactggactg	tagccatggc	cctggccagt	19500
ccttggtcca	gtcttgtctg	ctactcggaa	ctgggctggc	tcatggggcac	gatttctact	19560
gaggaggaag	ggctcatgtc	tgttgttgag	ttggcagcca	tgggaaaca	gctggctcagg	19620
attggttctg	accaggact	ggttgtcacc	caagtccaca	tgagaagctc	acacagtcct	19680
gttccactag	cagggggagt	ggccacactt	ccaggctcct	gcttgccaag	aactgtgggtg	19740
cctccagacc	atcagcccc	aatgcttccc	tgacaagcct	cagctgggtcc	ccttctctcc	19800
agcgcgtctc	ctccccacat	ctcccagcag	gtccccctcc	tttccctggg	ggatgtaagc	19860
accagcgttc	tgagactcag	ggcgttccat	gacaagcatc	caactctgaa	agacaacggt	19920
ccacattgca	ggtctccata	cgcaagcttg	gtggtctcat	gcccagcct	cgtggocctc	19980
tctgcatgtg	gacaggcagg	accactctt	ggggttggga	ctctctgacc	cogaatttcc	20040
ttcagcttcc	ttcagcttgt	tcactctgtc	agaaatgcct	actactggcc	tcttttccag	20100
aatattctgt	tttgtgttgt	acgatggctg	ggacagggca	acaaattctg	taatacattg	20160
accagtcttc	agcacaggat	gtggtcagaa	aaacaccaga	aacaggccag	gcacagtggc	20220
tcatgcctgt	aatcctagca	ctttggaagg	ccaagggggg	cggatcacct	gaagtcagga	20280
gttcaagacc	agcctgggca	acatgggtgaa	accctgtctc	tactaaaaat	acaaaaatta	20340
gccagggtgtg	gtgggtgagca	cctgtaatcc	caactactca	ggaggctgag	gcaggagaat	20400
ggcttgaact	caggaggcgg	agtttgcagt	gagccgagat	gcagtgagca	agattctgtc	20460
tcagaaaaaa	gaaagaccaa	caccagaaac	agctggaaac	tgcggtttgt	gcgagaaagg	20520
gacattcagc	cgactagga	gctatttgag	caggggtgta	tacaaaagtc	agtagagaaa	20580
taaaaataaa	acgcagaatc	tagtctatgc	cctaagctat	tcccagtcag	actgaaagca	20640
gtggtgtaga	tgcatattt	ttttaaaaaa	tgcatggggg	ccgggocgag	tggctcatgc	20700
ctgtaatttc	agcacgttgg	gaggcgaagg	cagttggatc	atgaggctcag	gagttcaaga	20760
ccagcctggc	caacatggtg	aaacgccatc	tctattgaaa	atacaaaaat	tagccaggcg	20820
tgattgatgg	tgtgtgcctg	taatcccagc	tactcagatg	gctgaggcag	gagaattgct	20880
tgaacccggg	aggcagaggt	tgacgtgagc	caacatcaca	ccactgcact	ccagcccggg	20940
tgacagagca	agactccttg	aaaaaaaaaa	tgctttggaa	atgtttgcat	aaaagtatat	21000
aaacaagtat	acattgttgt	ggctaacttt	tcactatctt	gattttgggtg	ggaacaacgt	21060
atgtggataa	ttcagaattc	attogcattt	agctttcact	tttttgttgt	tgtttctttt	21120
gagatagtct	ggttgtgtca	cccagctgg	aatgcagtgg	tgcgattccg	gctcactgca	21180
acctccgctt	ctggggctca	agccatcctc	ctacctcagc	ctcccagta	gctgggacta	21240
cgggcacgca	ccacatgcc	cagctaattt	ttgtgtactt	tttttagaga	caaggcttca	21300
ttatgttgcc	caggctggtc	tcaaaactct	gagctcaagt	gatgtgcctg	ccttttccctc	21360
ccaaagtgtc	gggattattg	gcatgagcca	ccacgccag	cctgcattta	gctttggaat	21420
gtgctagatc	tgaatgcagg	cgtgagtgag	attcttagaa	ttgtcatcct	tgggaaagtg	21480
cagcctagaa	gttgacgttc	ggctgacaca	agctctctcc	ccatggctct	cacctcttct	21540
tggagccggc	tgctgcctg	cctcccgtg	ctgtgtgtt	tagggataag	caagacaaa	21600
gtttcaacct	ttgggtcatc	cttccagttc	ctgacagacc	gtggaccatg	agtaagcggg	21660

gatggtgatg aggtccaaaa ggacttggtt agctgagaga atgagtaagg ggtagagggtg 21720
 gtgagctcag ccatcttttt tgtgocacat ctcatgacct cttagctctgg ttttttgtgt 21780
 gtttgtttgt tttgttttga gagagtctcg ctctgttgcc cagactggag tgcagtgcca 21840
 ccatctcagc tcgctgcaac ctccgcctcc caggctcaag caattctcct gtctcagcct 21900
 cctgagtagc tgggattaca ggtgtgtgcc accacgcccg gctaattttt gtacttttag 21960
 tagagacagg gtttcaccat gttggccagg ctagtctcga actcctggcc tcaagtgatc 22020
 tgcctgcctt ggccctccaa ggtgcttgga ttataggcat gagccaccgt gcccagcctg 22080
 ctctggtttt atatgcattg ttcattaata ttgttaccac gagatgctga cctactcatc 22140
 tctaaattcc ctctgttcta gtccagtgcc tgggtaatag caggccctca accataaaag 22200
 aatgcttgct ttttgaaaa cagcaggcta ccagggtgcag tggctcatgc ctataaacc 22260
 agcacttttag gaagccaagg tgggaagatt gcttgaatcc aggagttata gtcgggtagc 22320
 tctggtccag aagactctaa gtaattcacc cttagagttt ccagagcatg tagctgaaag 22380
 ttaacactat gtaatgaatt ctattcacct atttcatcat cactgctgtt tttcttttct 22440
 ctctctctog tttttttttt tttttttttt ttttggagat gggatctcac tctgtcgccg 22500
 aggctggagt actgtggcac aatcatagct cactgcagcc ttgacctccc tggctcctcaag 22560
 ggatcctccc acctcagcct cctgagtagc tgggaccagg gatgtgtgcc atcacacca 22620
 gctaattttt taattttttg tagagatggg gtcttgctat atgaccaggc tggctctgaa 22680
 ctctgagctc caagtgatcc tccacctca gcctccaaa gtgttgggat tataggcatg 22740
 agccaccaac cccagcctcc tgtgctttaa gaagcaccgg cagtgtgcac tgtccaatag 22800
 ggcaactcact agccccatgt gatcattgaa cttgattaca attaaataac atttaaaagt 22860
 caggctcctca gctgcaactag tcctatttca ggcataattc agaaagccac atatggctaa 22920
 tatattaaaa agtttataatg aacattttca tcattgtaga attctattgg atagcactat 22980
 tgtgggtgat tctgatgtgt gctaaaactt gacaaccact gagcttggac atacagtttt 23040
 caaccctgga tgcagattag aatctacctg gggcgttttt tttgtttttg ttttctctga 23100
 aacggagcct tgctctgtcg ttcaggctgg agtgcagtg tgggatcttg gctcactgca 23160
 gcctccaagc aattcttgtg cctcagcctc ctgagtagct gggactacag gtgcccgcta 23220
 ccacacctgg ccaatttttg tatttttagt agagatgggg tttccccatg ttggccaggc 23280
 tggcctcgaa ctcttgacct caagcagtc tcccacctca gcctccaaa gtgctgggat 23340
 tacaggcatg agccacagca cctggcctaa cgggggacag caggggaaga cttttgacac 23400
 ccactgacat cccaacatac agcagaccaa ttgaaacaga tgggtggggcg gttctggggg 23460
 ttgcaggaag acttttagtg ttcagtttcc ttttccagc ctaaccagtc tcaaagccaa 23520
 agtcaagaac cactgcacaa ttccatgcac ttttccagc attctcattg aagtcctctg 23580
 agacccttac gaggtatgca ctgttatagc cacatgtgct caacagaaga gcatcaggag 23640
 gcacagagtg actgagcaac ttgtccgtga tgacactgct cttagatgtt ctggatgcca 23700
 aagggcattc ctggaaggta tcaggaatcg gggccaggca gggaaatgga acagcccgtg 23760
 gttcctgcca gttggagcct gaggaagccg gtagacttgg ggatggctgc cttgcagtga 23820
 cagctgcttt taatttgtgt caagaagaaa agctgccttt taattagaac agttcgggat 23880
 tttttttttt ttcagctgcc agactaacc agttaccagg gattagtgca gtcgggaagc 23940
 ggaatacagg cagtgtctagg ctccctctg ggccagttcc actcacagaa ggaccgccc 24000
 gagagtcatt tcattgcccg ccgctgagc agtctgcaga ggcagtggt ggccagtgca 24060
 agaaagaaaa tcagcctggc cagtgggtgt agctgagtg ctgcagccag acggccagtt 24120
 gcttggctca ggatgggggg caggaggta gtgtggagg agccccagcg tgagaaattg 24180
 agggagcaac tttgtctctg aaactcaaga ggggtgaggc aagacttcta tttttgttga 24240
 caggtgcccc aagtttcagc cctgaaccct gacatactt cctctccca agtcctttt 24300
 cttatttttc ttgttttgta ttgttttgt ttgttttgg ttcttttgag atggagtctc 24360
 actctgtcac ccaggctgga gtgcagtggc accgtctcag ctcaactgcaa cctccaactt 24420
 ccaggctcaa gctattctcc tgctcagcc tcccaagtag ctgggtttac aggcacacgc 24480
 caccacacc agctaatttt tgtattttag tagagacagg gttttgccat gttggccagg 24540
 ctggtctcga actactgacc tcaggcgatc caccgcctc ggctcccaa agtgctggga 24600
 ttacaggcat gagccactgt gcctggcttg tttgtttta ttattttatg aaacagagtc 24660
 tcaactctgt gcccaggctg gagtgcagt gtgggatctt ggctcactgc agcctccacc 24720
 tcttgggttc aagcgattct cctgcctcag ctcccaagt agctgggatt acaggcgtga 24780
 gccatcacgc ctggctaatt tttgtatttt cagtagagat aaggttttgc catgttgctc 24840
 aggtggtct caaactcagg ctccaggcag ccccccacct cggcctcaca aagtgotggg 24900
 attacaggc tgagccact cacctggctg agaccagtc tcgtaaagag gaaaaaaatg 24960
 aagatgagc tctcctcca tgtccgcat ggtaggagct ctcccaggt atgagcaaat 25020
 ctcttcttcc ctgagctatc aggggtctct gcacagagta gcgcttagtt caccagtc 25080

tgatgcactg	aatattctga	ctctcaggag	caaaggctcc	tggctcaacc	atcattggcc	25140
ccatcagctc	ttgggctctc	caacaagtct	gttaactcac	cacctctcaa	caaaaccact	25200
tttttttttt	tttttttgag	atggtatctt	actcttgtca	cccaggctgg	agtgcagtgg	25260
tgcaatctcg	gctcactgca	acctctgcct	cctgggttca	agtgattctc	ctgcctcagc	25320
ctcccagagta	gctgggatta	caggcatgtg	ccaccatgcc	cagctaattt	ttgtattttt	25380
agtagagaca	gggttttccc	atgttggaga	ggctggctc	gaactcctga	cctcaggcga	25440
tccacccgcc	tcagcctcct	aaagtgctgg	gattacaggc	gtgagccacc	acaccctgcc	25500
aaaaccactt	tttttttttc	gagactgagt	tttgctctgt	tgctggagtg	caatagcatg	25560
atcttggtctc	actgcaacct	ctgocctctg	ggttcaagcg	attctcctgc	ctcagcctcc	25620
tgagtagctg	ggattacagg	catgcgccac	catgctcggc	taacaaaacc	acttcctgac	25680
tttgtgagag	ccttcaagtt	actaaccttc	tgtttcttca	cctgtttaat	ggggatacgt	25740
ttacctatct	catgggagtg	ttgtgaaggt	taaataaatt	agatgaggta	aagcacgcac	25800
agaatcgggtc	cttgggtgat	gttggacccc	tgcctctgcc	cctctgaaga	ggctgcctgt	25860
aatcccctgg	ctctaccacc	tttctccctc	acttttattt	cctagtattc	aactccctgg	25920
acatgtcccg	ctctgtgtcc	gtgacggcag	cagggcagtg	ccgcctcgcc	ccgtgatcc	25980
aggtcactct	ggactgcagc	ccactttatg	actacactgt	caagcttctc	ttcaaactcc	26040
actcctgtga	gtaccgcggg	ccagatcttc	ttacatgaga	ttcaggccag	agggaggatc	26100
ccagcctgag	gatgtcccca	gagaaacgca	gtccttctca	gtgccttttg	ctgtctgctt	26160
ctgttccaaa	aggccccgga	gcttctgacc	attgtgagga	taaaagagca	gggccccaggc	26220
tttggtgacc	ccagtaaagc	ccctggcttg	ccactcttgc	gtcccagtgt	tacaggatct	26280
ttgggggtgtc	cgttttctgg	ctggaaacct	ctggggccag	tgggtgccttt	gcccagatcc	26340
ttgttcggca	tccaggaaga	atgaggtatg	cagacaagtg	gaggggtggac	aagatgaaga	26400
ggagctttat	tgagtattag	aacagctcag	aggagactgc	agtgggtacc	gctctctgtc	26460
tgtaggcagg	ttgtcctgtc	gagtgttcag	ctctcagcag	aaaagaggcc	atggagtggg	26520
tagctcctct	ctgcagctga	ttatcctagc	atctctgcag	gtctctgaag	cctcagcaga	26580
gagggtagct	cctctctgta	gctggctgtc	ccatctctgc	tcagctctgg	ctgagctcag	26640
gccttttatg	ggcctcagag	gggaggaaat	gcaccacgat	tgggtccatgg	gcaggcccag	26700
aaagggcacc	ccaagttccc	actccggctc	gtgggattgg	cagcccggcc	cccaccgggg	26760
acccgcctct	tcaccacagga	atctgtctgc	ctcccgtctc	catgcatggc	aacagggtcc	26820
agccccaaact	ttgctctaag	atcagagtgg	gtgcccagac	cagggagaag	ccaggcagcg	26880
ggaacaggtt	cagaagaggg	gagggagagc	cttctcaggc	cctgaagagt	acagggatgc	26940
ctgagtctgc	agctggggct	gggggcgggt	goggggcagc	agggctgcca	cccactccat	27000
ggagtgggag	gccccagctc	gcagctgtgt	tttgggtggc	tgcagctgca	cctgggaagg	27060
tagggttcct	gcctgctccg	gacctcaag	agcacaggga	ggctcagatc	tgcagtcaca	27120
acttgggcgg	ctacagcccc	atccagcagg	gtgaggcttc	tgcctgttcc	atggagtgtg	27180
cagccctggc	catgcctccc	tgctgcagct	ggcatgatgg	cagtggcagg	ccatctggac	27240
tggcagctgc	catcaccagg	atggctcctc	ctgtctgcct	ctagcttgag	cacgtcacca	27300
ttcaacaagt	atgtcgtgca	tcttacgaac	agcaaaaata	tcagcagacc	tggcctagac	27360
ctacatgagc	taacataatg	atttccagac	catttatcta	cacaatccct	ttgtgcaaat	27420
ggaattttat	ggagaaatat	aaattataaa	acacagctgc	tcattaaatg	gagacctatc	27480
ccttacctag	cttcagttcc	ctgggggtgc	cttagcagaa	tataatagtg	ttggaagttc	27540
tggctctaca	gaagtcttta	cacttttttt	tttttttttt	ctgagacaga	gtctcactct	27600
ctcaccaggg	ctagagtgca	atggcatgat	ctoggctcac	tgcaacctct	gcctaccagg	27660
ttcaagcaat	tcttgtgcct	tagcctocca	agtagctggg	accacaggtc	tgtgctggca	27720
caccagccta	attttttght	gttgttghtt	gaatttttag	tagagatggg	gtttcccat	27780
gttggccagg	ctggctctga	actcctggcc	tcaagtgatc	cgcccacctc	agcctcctaa	27840
agtgctagga	ttacaggtat	gaccoaccac	accgcaccag	tcttataact	ttgtgtcctt	27900
ctggtttaaa	atgatgtgtc	tgggtggggc	acagtgggtc	acgcatggaa	tcccatcact	27960
ttgggaggct	gatgcaggca	gatcgcttga	gctcaggggt	tggagaccag	cctgagcaac	28020
atagtgaat	ccctgtctgt	acaaaaaata	caaaaattag	ccaggcatgg	tgacatgcgc	28080
ctgtagtccc	agctacctgg	gaggctaagg	cgagaggatc	gcttgagccc	agaaggtcaa	28140
ggctacagtg	agctgtgttt	gtgccactgc	actccagcct	gggcaacaga	acgagaccgt	28200
gctaataaaa	aataaaaata	ataaaaataa	tagtataaag	acagtatact	ctgtaactgc	28260
ggttattgac	aagtaatgaa	tctctttatt	tgacagctac	aagggaaagg	tgtagcatcg	28320
cctagatgct	aacctaggtat	ggttttgtgg	aagggagatt	ggaattagct	gcccagggat	28380
tagaagggtc	ttgocggcaag	taggcctgcc	tgaagggagc	tgagaggggtg	gagtgggccc	28440
ccaagagaag	ggaccagctg	cttagagtgc	tgaatacaga	tgaacttttc	attgtttccc	28500

ctggaaactt ttcattgttt ccaagtattc tgttgetcag tagaacatca ccaactgcctt 28560
 tttgtaaaaat ggggagttaa gctgggcatg gtgggtgcaca cctgcagtcc cagttactct 28620
 ggaggctgag ggagttatga gaagagacag gtaattaagt gtttggggtt aaaggccttc 28680
 tttgaaggca gggcaatacc caagaccac actgcctctg ggctgaggaa ggggtaggag 28740
 ggatactctt tacaatacat agttttttaa tttactttat ttattttatt atttttctga 28800
 aacagggtct cgctctgttg cccaggctgg agtgcagtgg cacaatctca tccoactgca 28860
 acctccatct cccaggttca agcaattctc ctacctcagc ctcccaagta gctgagatta 28920
 taggtgcccc ccaccatgcc tggctaattt ttgtattggt tttagtagag acagggtttc 28980
 atcctgttgg ccaggcttgt ctcgaaactc tgagctcaag tgatcctcct gcctcggcct 29040
 cccaaagtgc tgggattgca attgtgagcc actgcagctg gccctctttt ctaaaaaata 29100
 acattttatt ttgttttaga tctacagaaa agttataaaa atagtaccga gagttctcgt 29160
 ataccagtt ccattttccc tttgttcgta tattagtagt aacattttgc cacaactggc 29220
 cgggcacggg ggctcacgcc tgtaatccca gcactttggg aggccgaggt gggcggatca 29280
 tgaggtcagg agatcgagac catcctggcc aacatggtga aaccctgtct ctactaaaaa 29340
 tacaanaaat tagccgggag tgggtggcggg cgctgtagt cccagctgct caggaggctg 29400
 aggcagcaga atgggtgag cctgggaggt ggagcttga gtgagccgag atcgtgccac 29460
 tgcaactccag cctgggtgac agagcgagac atcgtctcga aaaaaaaaaa aaagaaactg 29520
 gtcacaacta attaaccagt gttgatacct tattattaat taaagtctgt acttaaggca 29580
 gggcacgggt actcacacct gtaatcccag cactttggga ggtcaaggca ggcggatcac 29640
 gaggtcaaga gattgaaacc atcctggcca acacggcgaa accctgtatc tgctaaaaat 29700
 acaaaaaatta gcggggtgtg gtggcgcgcg cctgtagtcc cagctgcttg ggagctgagg 29760
 caggagaatt gcttgaacc gggaggcgga ggttgtgtg agctgagatt gtgccactgc 29820
 cctccagcct ggcaacagag tgagattctg tctcaaaaaa aaaaaaaagt tctacttta 29880
 ttcagattgt cacagttttt accgaatgtc cttttctgcc ccagaatctc acccgtaca 29940
 tgcatgtgac atttagttgt gatgtctcct taggttctc ttggctatga cgtttcactt 30000
 ctcaaacttc cttattttt gatgactttg acagttctga gttttttgta gaatgttct 30060
 gaattctggt tgctgatgt ttttctccgg atttaactgg ggttaagggt tcttgggaga 30120
 aagatcattt aggtcaagt ccatattcat tcaactgtca cttctgccat ttttgtttt 30180
 ttttgtttg tttgtttca agacagggtc toctctgtc acccaggctg gagtgcagt 30240
 gtgcgattgt agctcaactc agcctcaaac tctgggctt aaaagatcct cccacctcaa 30300
 cctcttgagt agctaagact acagtgcaca tcaccacacc tggctaattt ttatttctta 30360
 atttttgtag agatgggggt ctcaactatgt tgcccaggct ggtctcaaac tctgggtac 30420
 aagcgatgct cccactcag cctcccaag tgotgggatt acacatggga gccactgagc 30480
 ccagccatct tttgccatt caataatcat gttttataca tgtgttccct gattttatcc 30540
 agtgaatgta tgaccattt ccaataaca aatttaaat taaacaaata aagcaaacac 30600
 taaaaggatt gctgagctgt gatctacagc tgcaatatgc aggttcgcac atctgagagg 30660
 ttacagaacc tcagagaact ggctctgtag gggaggaggc tgcctaaga ggttactgg 30720
 gtgtgcagtg tcatagagtt agctaacca caagccggga ctggaactca gatcccaagc 30780
 ctcaatccct tgttctttt tttgacagag tctcaactgt gcccaggctg gagtgcagt 30840
 gtgcaatctc agctcaactg aacttccgcc tgctgggtt aagcaattt cctgcctcag 30900
 cctccaagt agctgggatt acaggcgcac gccaccgagc cggctaatt tctgtattt 30960
 tagtagagac atggtttca catattggtc aggtggtct cgaactcctg acctcaagt 31020
 attctcctgc ctcaagcctc caaagtgtg ggattacatg cgtgagccac tgcacctggc 31080
 cttcaatccc ttgttctttg ccttctgaca gccatcctt tgggcagggt ctgaggatct 31140
 cttccagact aatatcagcc ctgagcctca acctaaaaca aatgcaggca atagctgaga 31200
 aagaactgct ccctcctggg cctgccatgg tccagcctgc tactgggtca ggtcaccatt 31260
 tcttctctag tccaggacac aatttccacc ctattgtgcc cggcatggct tgtccttcag 31320
 gaactaattg aagtgaagg atttggagga taagcgtctc cacactcctg ctggttctg 31380
 ctgggctccc ttggttacca gacctcggga cagctctagg ccagctgtgg ccctggcag 31440
 tgctggccac atgccccagg gtagctgggc ccctccccct cgagagcccc gctgtggctt 31500
 ccctgocctc tggccccct cccctctcac actctttcca atttcttcca ggctcccag 31560
 ctgacacct gcaaggccac cgggaccgct tcatggagca gtttcaaag taagtggttc 31620
 aagtaacagg aatggagggtg aattcaagag cgctataaaa tcattaagcc atgcaagtgc 31680
 actgtgagag gtggctggac acagtgggtc atgcctgtaa cccaactct ttgggaggct 31740
 gagggggag gatcacttga ggccaagagt ttgaccagag gctgggcaac atagcgagac 31800
 cccatctota cagaaaaaaaa atgtaaaaca aattagctgg gcacggtgg atgcaactgt 31860
 gatccagacc acttgagagg ctgaggcagg aggatggcct gacgctagga gtccgaggct 31920

gcagtgagct agattgtgcc actggacttc agtctatgtg acaaagtgag accctgtctc 31980
 taaaaaaaaa ttttttttaa atacttgtgg gcgtagaggt atcttcgggc tggagtgcaa 32040
 tggcaatggc acgatctcgg ctcaactgcaa cctccgcctc ctgggttcaa gtgattctcc 32100
 tgcatacagc tctcgagtag ctgagactac aggcacacac caccatgcc agctaatttt 32160
 tgtattttta gtagagatag ggtttcacca tgttggccag gatggctctg atctcttgac 32220
 cttgtgatct gccacctca gccctccaaa gtgttgggat tacaggcgtg agccactgca 32280
 cccgacctga ggttatcttc attctaagggt gataaagtga cagaaaatag ctcaaacaga 32340
 ggcaaatcta cctatagcta gggtaactat tggttatcta ataaatccca tagaagcata 32400
 tagaagaaag agatttcata cattcattag ttgcctgctt atgccccaga cttcatctc 32460
 ttctcttttc tgtctcccta agttctggtc ttatagcatt tatagactga taaccctaa 32520
 atatgcatac ctagccagac cttacttctg aatgccagcc tcttgtcccc aatttgacat 32580
 ctgtatagca tacttaaatc tcaaaactgta ctttaggttt tctcaaaaaa aaaaaacaaa 32640
 aaaaaacctt tctctctcca ttctgttttt gtttgtttgt ttttgttttt gtttttgaga 32700
 cagagtctcg ctctgtcacc caggcttgag tgcgggtggt tgatcttggc ttactgcaac 32760
 ctacagctcc ccatctccag tgattctcct gcctcagcct cttcagtagc taggattaca 32820
 ggcgccacc accatgcccg actaattttt gtatttttag tagagacagg atttcacat 32880
 gttggccggg ctggtctgga actcctgacc tcaagtgatc tgcctccaa agttaggga 32940
 ttaccggcgc gactcactgt gccagcccc tcccattct tttctatctc aacaataact 33000
 gccatcatcc atttagcca gagatctgga agtcaccatt ttcctcatcc cccatgtcca 33060
 atccactagc aggtttggtt gataccctac aaatatgacc tgcttctgag gtgggatggg 33120
 agactggact ctggaggcag ggcttgggca tccgaccaa ttaaggacaa gtaaacagg 33180
 gaaagggcgg aagcacctct gcataagaca cacctaccag tgtgcagtga cagtttacca 33240
 ttgctatggc aacatccgga agtttagcgc ctttccatg gcaataacct gacaatctgg 33300
 aatttaccac tttttttcta gaaatttctg cataatctgc cccttaattt acatataatt 33360
 aaaagtgggt ataaatgtaa cagctgctat tctgggctca ctgcgtatgg agtagccctg 33420
 cgctgcaagg aacaggacct ccgctgctgg ctgtgcactg ccgcctcaat aaaagttgct 33480
 aacaccagcc ggggtgcagtg gctcatgcct gtaatcccag cactttggga ggccgaggtg 33540
 ggcagatcac ttgaggtcag gagttcaaga ccagcctggc caacatgggtg aaaccccgtc 33600
 tctactaaaa aattagcagg gcatggtggt gcacacctgt aatcccagct actcgggagg 33660
 ctgaggcaga agaatcgctt gaaccagga ggtggagggt gcagtgagcc aagatcacat 33720
 cactgcactc aggcctgggc gtcagaatga gattccatct caaaacaaaa acaaaaacaa 33780
 aagatcttta cttaggctgg gtgcagtggc tcacacctgt aatcccagca ctttgggagg 33840
 ccgaagcaag tggatcgctt gaggtcagga gtttgagagc agcctggcta acatgatgaa 33900
 accctgtctc tactgtaaat acaaaaatta gccctgtaat cccagctact caggaggctg 33960
 aggcaggaga attgcttgaa ctcaagggtg agagggttgc atgagctgag atcatgccac 34020
 tgcactacag cttgggcaac agagcaagac tccatccaaa aaaaaaaaaa ggaaaggagg 34080
 ggaggggagg cgaaccagg caaaagggaa ggataataga gtgagagttg ggaggacaga 34140
 gggggctgag aagggcctcc tgacttagct gtgtaaaagt gcaactctct gtttttctctg 34200
 ttaatttctc cacccttgac cataactccc aatgccaag tttcagcttc tctgcctggt 34260
 tccccaggag caagcgaagg acaagtgtgt ggttgtacc acctcctggt acccctgttc 34320
 catttgagaa gccgagtaca tattctgtga taaatggctt ctttagtcca aggagacagg 34380
 ccctgaggct gatctcctgg aatgtcaaat tccggtagct tgtgaaagca gttgagcctg 34440
 ttactgttgg ctccaaacca gcaagatggg gatgtgtggt caggtgattt catttcttt 34500
 catttgcttc tagcttagaa aaacatctcc cctgagtagc cactttgctg ttccacacag 34560
 acacggctgc tgtctgagac ctcaactgtcc ttgaattcag gggaagcagc agtgatgtca 34620
 cagagacaga gacaggccta agtcttggga totgacaggc tgtaacgaga ccctgggggt 34680
 ttcgtttatt ggactaagct ggagaaagtt cattgttgc ggccaaagga ttttaaggccc 34740
 atgtaccctt ttcatctcag aatagatgta gtttaactcca gctactcagg aggctgaggt 34800
 gggaggatcg cttgagccca ggaggttgag attgcagcaa gccttgatcg ttccactgca 34860
 cttcagcctg ggtgacacag caagaccttg tctctaaaaa aattaaaaca ggccaagcat 34920
 ggtggctctc acctataatc ccagcaacttt ggaaggccaa ggtgggtaga tggcttgagc 34980
 ccaggagttt gagaccagcc tggacaacat agagagacc catctttatc aaaaacacac 35040
 aaaaaattag ccagggtgtg tggcacacac ctgtactccc agctacttgg aggctgaggt 35100
 gagaggatcc cttgggcca gggaggtcaa ggctgcagtg agctatgatt ttgccactat 35160
 actccagcct gggcgacata ctaagacca atctcaaaaa aatatttaca aataggcag 35220
 gcacattggc tcacgcctat aacccaaca ttttgggagg ccgagacggg cggatcacct 35280
 gaggtcggga gttcgaaacc agcctgacca acatggagaa accccatctc tactaaagat 35340

acaaaattac ccaggcatga tggcaocatgc ttgtaatccc agcaactcag caggctgagg 35400
 caggagaatc ccttgaacct ggaaggtgga gcttgcagtg agccgagatc gcgccattgc 35460
 actccagcct gggcaataag agcgacactc catctcaaaa aataaacaaa caaaataaaa 35520
 atgaagaata tataagtctg gattctctta ggtttcttgg tcttcagtggt ttagattcaa 35580
 aatagagatt tgagaatgag aaaaagatgt tgatthttgcc agatgtggtg gctcatgcct 35640
 gtaatcccag cactttggga ggccaaggca gtaggattgc ttgaagccag gagtttgaga 35700
 ccagcctggg caacacagca agagcctggt tctacacaaa aattaaaaat tagccaggca 35760
 tggtggtgca tccctgtaat tctagctact caggaggctg aggcagaaga ctgcttgagc 35820
 ccaggaattc aaggctgcaa tgagctctga ttgcaccact gcactctagc ctgggtgaca 35880
 gagcaagacc ctgcctctac aaaaagagag agagagagaa aaaaaatatt ggthttctca 35940
 tgtgcatatc cacccaagat taggacatag taccacacga ttgtaattct ccaagtgttt 36000
 atgaacggc tgccctgtcc ctggcattat gggaaacagc tcagcggagc tcagtcccca 36060
 cttcccagga acttatagtc ttaattgtgg agagaacaaa tagtccaaaa acaaagagaa 36120
 gtaatgaaag aagagagcta cctttgtagt aaaattccct tcccttttcc tatggttaga 36180
 tggcttatcc tcagagaaaag tcaatttctt ttaatgttac tctaaaccaa acactcctcg 36240
 gaagttaaat ctgcaaaaata acaacatgtg ttctacagca atgaagaaat tagctthttg 36300
 attatctact cttccaaagg atthttagca tgaattactc thtctcacc agcagggtac 36360
 tgaggaccag cthtaaatata tctctgagc catcataatt gaatcccagc accatggagt 36420
 ttatctcctt gacagcctgt gcctthtgggc tggggagggg gcaggaaagc cagggtgctg 36480
 ctctgtcccc tacatggggc tgatgaagac acccagoccc cctcagggtc ttctccacc 36540
 ctagggtgaa agatctgttc taccgctcca gcaacctgca gtacttcaag cggctcattc 36600
 agatcccca gctgcctgag gtaagcatgc ccaaccacac accctcggca ctgcagaggc 36660
 cccaggtact ctcttaaggg ccggcggggc ctggcaagca agcactattt gaggatgtgt 36720
 ctccgtcttc agaaccacc caacttctg cagcctcag ccctgtcaga acatctcagc 36780
 cctgtggtgg tgatccctgc agaggcctca tccccgcaga gcgagccagt cctagagaag 36840
 gatgacctca tggacatgga tgctctcag cagggtgagga ccacttggga gagaaacttg 36900
 gcctthctc tcacctgcaa gtacagggga gaggtgggg gagacctgg ccaaagcca 36960
 ttgactctaa ccaggttcag gcttctcttc attcacctag caccaaccta ggctgagcat 37020
 cctgctctgt ggctcccaga agggtcataa atgggtgcag tggctcacag ctgtaatccc 37080
 aacatthtg aaggccaagg tgggaagact gcttgaggcc agttcaagac cagcctgggc 37140
 aacatagtaa ggcccactc ttacaaaatt agccagatgt ggtggcacat acctgtggtc 37200
 ccagctatgt ggaagctga gatgggagga tcaacttgagc ctgggaggcc aaggctacag 37260
 tgagctgtta tcatgccact gcactccagc ctaggcaaga gaacaagaac caggccctaa 37320
 aatataaat acataaaaatt aaaataaaaa aattatccag gtatggtggc acacacctat 37380
 acttccagct actcaggagg ctgaaacagg aggatcactt gagccaggag ttggaggctg 37440
 cagtgagcta tgattgcacc actgtactct agcctgggca gtagagcgag aacctgtctc 37500
 aaaaaatta aaaaaattaa acaggtctga accgtthaat tcgagaaagg gggcattctc 37560
 ccatatcact caactgacct acacacagaa ttctctggct ctctgactta ttctcactcc 37620
 tthttggtca accacagaat ttatthgaca acaagthtga tgacatcttt ggcagthcat 37680
 tcagcagtg tcccttcaat ttcaacagtc aaaatggtgt gaacaaggat gagaagttag 37740
 tccaagctgg gttcaagcag atggttcagg agctaagtha agccatggtc tgcctcaaaa 37800
 cactaaccaa agaggaattc ttaatgatac tggggcttct tagatacaga acatctthaa 37860
 gggthtgggg caatggctta tgctgtaat cccaacatgt ggggaggatg aggtaggagg 37920
 atthtttaag gccaggagtt taagaccaag cthtgggcaac atagcaagat cccatcttht 37980
 ttaataaaaa gtaaaaaaat tagctgggca ggtggtacac acctgtagtc ccagtaactc 38040
 aggaggctga agtgggaaga tcatttgagc ctgggatatac aaggctgcag tgagctatga 38100
 tcgtgccact acacttcagc ctgggcgaca aagccagact gtctctaaaa caaaaccgaa 38160
 aacacacaca aaaaaggaat gtcttgacct tcaaatattg gccctthta tctcagaaga 38220
 aatcaataa ccatggattt atgagtatta gattagatc tggtaacatt tagagtataa 38280
 thtatggcat thcaaagaat tgtcccaaaa ttaataccag cththtaatt cctcccctga 38340
 gctcacaatt aaaaacagag ggatagaagc actatgaaag caaactcatt ccccttctct 38400
 tcccaggac cacttaattg agcgactata cagagagatc agtggattga aggcacagct 38460
 agaaaacatg aagactgagg tataactthg atctgctctg cctthtgcgt tcaccaaacc 38520
 acggtagatt tgaatgthaa atthgcatca cactagccag gcacagtggc tcacacctgt 38580
 aatcctagca cthtgggag ccaaggcagg aggtaccct gaggtcggga gthtcagacc 38640
 agcctgggca acagggtgaa acccccgtct tcaataaaaa tgcaataatt agccgggtgt 38700
 gthtggcaggc acctgtaatc ccagctactc ggggaagctga ggcatgagaa thgctthaac 38760

ttgggaggca gaggttgacag tgaactgaga tgcgtgtcact gcactccagc ctggggcgata 38820
 gaacaagact ctgtctcaaa aaaaaaaaaa aaaattgcat cacctagaca gtttttggac 38880
 cacatcctta gggtaggtta attagtagaa acaagccagg cgtgggtggg tgcacctgtg 38940
 atcccagcta ctcaggaggc tgaggtagga ggatcactca agcccaggag ttttgaatcc 39000
 agcctgggca acatagcgag acctgtgtct ctaaaaccaa ataaaaactg gtagaaatgc 39060
 taaatccaaa agaaccaggt cttcccgtgt tctctgtcat aatgcatttc cattttacat 39120
 gccatgagga cagcatttag gccaggcaca gtatccatgt gattgatacc cagagatgac 39180
 cttggtcccc acgaggctga gaagctgtgg agaacaagac cagaccttct cacatggcga 39240
 taaggagaga ttggtttgcc gggcacagtg gotcacgcct gtaacccag cactttggaa 39300
 ggctcaggca gaaggatcac ttgagcccag cagttcaaga ctgacttggg caacatagca 39360
 agaccccatc tctacaaaaa acttttttaa aattagctgg gcatgggtgg gcatacctgt 39420
 aatgccagct acttgggagg ctgaagctag aggatctctt gagcctagga ggttgaggct 39480
 gcagtgagcc ataatacact cactgcactc cagtctaggt gatagtgata ccctgtccct 39540
 cccctcccc aaccgcaaaa aaaaaagag aagagacagc tgtgagcaat gttggtgcta 39600
 tattaacaac cgcccctcac agtggctggg caggtgaact caccacaggg gacagccact 39660
 gctaccccca gccactctaa agaggaccac aattccccgg ccatcatccc ctgttattgt 39720
 tgttgattga ggggctccta atgaccagat ggtccaaacc tcctgggacg tggagagtgt 39780
 acttagggga atcaggtatt tacttggaaag catggtagga cccgcttctc cggcccatgc 39840
 cagtgtgctg cagctgaag gccacgtcag cgagctggaa gcagatctgg ccgagcagca 39900
 gcacctgcgg cagcagcgg ccgacgactg tgaattcctg ogggcagaac tggacgagct 39960
 caggaggcag cgggaggaca ccgagaaggc tcagcggagc ctgtctgaga tagaaagtga 40020
 gcgggtgggtg ggggcggggg cgggccccgg gggcaggcgc gggcagcaga gccagctgg 40080
 actcaggatg cggcacagag gctgggtggg ggagaccag actgtctttc tagaaacagc 40140
 taggaatgcc agctcttagc ctcaagtccag agggcgggtt tgagtcctgc aggcacaaag 40200
 ctgtgtgca ggcaccgtg agcacagaga tgggaaatac agaagacagc actgaagtgc 40260
 ttgccctggg cagtggatac taagggagaa ggagagagga ttgaggtoct aacaaagaaa 40320
 ttgacataaa agctgagggc acagtggcac atgcctgtaa tcctagcact ctgggaggcc 40380
 aggatggaag gactgcttga ggccaggagt tcaagaccaa cccgggcaac atagcaagc 40440
 cccatctcta gaaaaataa aaaagactag cgggcatgg tggcacgtgc ctgtagctcc 40500
 agctactctt gaggctgaga caggaggatc gcttgagccc aggaattcaa ggctgcagtg 40560
 agctgattgc accactgcac tccggcctct gcaacagaga gccttgtctc taaaaaattt 40620
 caaataaatt tgttaaaaag ctgaaaagccc ccatagaata aatgccatgt atgtaagtgc 40680
 tgaagaggca tgaatccctg gcttgattat ttatttgttt tatttttttg aggtggagtc 40740
 tcaactctgtc gccaggctg gaatgcagtg gtgtgatctg ggctcactgc aacctctacc 40800
 tcccaggttt aagggtattg cctgcctcag cctcccaggt agctgggatt acaggcacac 40860
 accaccatgt ccagctgatt tttgtattta tagtagagat ggggtttcac catgttggtc 40920
 aggctggctc tgaactcctg acctcaggtg atcccaccac cttggcctcc caaagtgtg 40980
 ggattatagg catgagccac ctgccttggc cctggcttg atttttaact agttgaattc 41040
 tttttaaaga tgacttcctg gagaagtttc tgtaagttag actttcaagg gagaaaagca 41100
 tatatgcatt cctaaatcaa aggaagatct ttggccgggc acagtggctc atgcctataa 41160
 tcccactgag atgggaggat cacctgagcc caggagtttg aaaccagcct ggaaaacata 41220
 gtgagaccct gtctctacaa aaattagccg ggtgtggtgg cgtgtgcccc tgaacccagc 41280
 tactcaggag actgggggtg gaggatgact tgagcccaag gagggtcaagg ctgcaagtga 41340
 cagtgattgt gccactgcac cccagcctgg gtgacagagc aagactgtct caaaaacaaa 41400
 caaggaggac cttctagggc ccctggetca ttgcaaggaa ggcaagggtc cctgctaggt 41460
 tagactcctc accttggctc tttacaatac agggaaagct caagccaatg aacagcgata 41520
 tagcaagcta aaggagaagt acagcgagct ggttcagaac cacgctgacc tgctgoggaa 41580
 ggtaagacct tcagcccctg tcaccatcct gcaggccctg cacctctagg gagagagcgg 41640
 ctcaggcctg tggcttcccc ggggcccagca acccctacat tgatctctaa ggcattgccg 41700
 tcatctcggg aaccacacct tttcaggctt ccttgctct gtgtcttggg ctgtgtcctg 41760
 ggtgccaatc ccatgtaggt caccacactt ctttattatt ttgtaaatat ttgagcatca 41820
 agcatctgaa ggtgatggct ctgttccggc tgtgggttag aaagtgattc ctgtgtctga 41880
 ctctagggca cgcacagcct gagtatgatt gtctagaag gaggatgtcc tctaagcctg 41940
 ggatctcctg gttcaagaca ctgttcttct tttgcagaat gcagaggtga ccaaacaggt 42000
 gtccatggcc agacaagccc aggtagattt ggaacgagag aaaaaagagc tggagattc 42060
 gttggagcgc atcagtgacc agggccagcg gaaggtgagt gggacgagga gcaactcggga 42120

aatgagggag ggggctggtg agttgggtggc gggggctttg tggccttctg ctccatgggc 42240
agttctgtgg gtcggttggc atcacacagc agggagcaca ccatggtggc cacacagaac 42300
agcaaaacac cgtcaccagt gcatacacga agtaatggaa ggaccaagt tgctgggggt 42360
taatcaaca aggctttcca aggatgagggc tgggtggagag tttgaggaaa ggaaaaacaca 42420
tgggtggcag agagagaggg atgagcgttt ttttttgttt gtttgttttg tttgttttgt 42480
tttgagacg gcagtaatcc caggactttg ggaggctgag ggggggtggtg atcaccaagt 42540
caggagttcg agaccagcct ggccaatatg gtgaaacccc atctctacta aaaatacaaa 42600
agttagccag gtgtgatggc gtgtgcccgt agtcccagct gcttgggagg ctgaggcaca 42660
agaatctctt gaaccttgga ggtggagggt gcagtgagct gagactgcac cactgcactc 42720
cagcctgggc aacaacagcg agactccgtc tcaaaaataa ataaataaat aaataaataa 42780
ataaataata aaaaatcctt ccttgacatt ccacattttc agtacactct ggggtcttgc 42840
ctcctcctcc cacaccctgc acttttttgg ggggggtgtaa cttacatata gtagagtta 42900
cagatctctt gttgatcgct tgggacgttt ttacatttt atattctttg tcaactgtcac 42960
ccagatcaga gtccctctgt ttttcttctc tttcagactc aagaacagct ggaagtcta 43020
gagagcttga agcaggaact tgccacaagc caacgggagc ttcaggttct gcaaggcagc 43080
ctggaactt ctgccaggt aaatacctcc tttttttttt tggagataga gtctcttct 43140
gtcactcagg ttggaatgca gtggtgcat cgcagctcac tgcagcctcc acctccctgg 43200
gctcaggtga tctcccccac tttagccccc gagtaactgg gactacaggc acacaccct 43260
acacctggct agtttttgta ttttctttgg tagagacggg gtttactat gttgccag 43320
ctggtctoga actccagggc tcaagtgatc ctcccacctc agcctcctaa agtgtggga 43380
ttacagacat gagccttcat gcccggcctc ctttcttggc ccaccctgg ctttggcgtt 43440
gctgtcatcc accatccttg gcttgccaa gtcagcccc actgcaatca gtgtgtccc 43500
gggagggaaat cagagtggca ggttaaagag ccatcacctt cccagtcctt gcaaccgggt 43560
ggtgggttg acctctggga agtagggact gtttaactca accagcgtct cctctttcc 43620
ttgtggtcac ctttgagtc agaagcaaac tgggcagccg agttcgcca gctagagaag 43680
gagcgggaca gcctggtgag tggcgagct catagggagg aggaattatc tgctctcgg 43740
aaagaactgc aggacactca gctcaaactg gccagcacag aggcaagtca cggacatgga 43800
cacgagcgag cacctgtgaa tccccaccga gggcctctgc gcatgcacgg aggtcgggag 43860
gacccgggg ctgctgagaa ggggtttggg gccttggcct gattgtgcag acattctgta 43920
ggtgtaatgc cagcaggccc tgcatgtcct gcagagtcca tgagggaaag caaactgctg 43980
tctttttgt aggagaagag aagtttgga tctcctcca gccatgagaa gcgggcgag 44040
tgatccatcc ccacagcctc tgtgggagc cgcacatcgt gggcagccgc acatcctgtg 44100
cacacagtta aagcgcctt ctctgtctca ggcttactgg cttggacctc attggcctag 44160
acttgagota agatgctaag agccccagcc aggtcatcct gctcaggttc attatggagt 44220
ctagggcaga ctctcacctc cctggaccat ttttaggaat ctatgtgcca gcttgccaaa 44280
gaccaacgaa aaatgcttct ggtggggctc aggaaggctg cggagcaggt gatacaagac 44340
gccctgaacc agcttgaaga acctcctctc atcagctgcg ctgggtctgc aggtacactt 44400
gcaattgccc agctggcagg gccaggtcc ttacagcctg agactctgtt gatgttgaat 44460
ctcatgtgag acttagctca ggggctctca gccagcagc atgtcagcat taccttaggg 44520
gogcccaggc cccatcctag atcagttaca tgtgaaact ctgtcatta gtgcctatac 44580
actagtattt tagtattttc tcccccccc ccccccgcc cccccgctt ttgagatagg 44640
gtcttactct gttgctcagg ctagagtgca gtgccgtggc tcaactgcaac ctccgcttcc 44700
tgggtcaca caatcctccc cactcattct cccaagtagc tgggactaca ggacgcaac 44760
accacgcca gctaattttt ggtttcgggt tgtgtctcag actggtcttg aagacctgg 44820
ctcaggctat ccaccacct tggcctccca aagtgtggtg attacaggca tgagccacca 44880
tgocctggcct tggctaattt ttttaatttt ggtagagagc aggtctcact ctattgcca 44940
ggatggtccc aaacttctga actcccacct tggcctctca aagtgtctgga attacaggca 45000
tgagacacca cgccaggccc agcagtagta ttttctaaag cccccagggt tgatagctac 45060
tgctttagac agtgggtcac actgataaaa caccaccag gagcagaca ttctgtttct 45120
ctaaggaaaa aactatggtc tgaatcaaga ggtgattact catgaacttc acgtctaggc 45180
aggaagctta cccactaggt aagctcctcc attcagtgct taattaacga ggatgaagcc 45240
agctatgaga acttgctctg accttgccct gtgttccctc tcacagatca cctcctctcc 45300
acggtcacat ccatttccag ctgcatcgag caactggaga aaagctggag ccagtatctg 45360
gcctgcccag aaggtaagaa tggccaagga cagtctctgt cggctagtga tggccagaca 45420
gggttcagaa gcacctgaat gcggggatag tgacaggtcc ctctgcatca agaaaggcat 45480
gtaggcaact catacaagaa aggcagtag gcaactcata aaacgggagg agaggggatg 45540
aaagtgtcac catcaaccag acctgagaaa cttctcttcc caatcctggc agacatcagt 45600

ggacttctcc attccataac cctgctggcc cacttgacca gcgacgccat tgctcatggt 45660
 gccaccacct gcctcagagc cccacctgag cctgcccact gtgagtactg gggcatgagg 45720
 ggctgttcat ggaccagggg agcagggggc ctttaaaagt ctctgttggg ccgggcgag 45780
 tggctcatgc ctgtaactcc agcactttgg gaggctgagg cgggcagatc acttgaggtc 45840
 gggagttcaa gaacagcctg gccaacatgg caaaacccca tctctactaa agatacaaaa 45900
 acgatgggcc agatgcaatg gttcacgcct gtaatcgccg taacactttg ggaggacgag 45960
 gtgggcagat cacctgaggt caggagttcg agaccagcct ggccaacatg gcgaaacccc 46020
 gtctctacta aaaatacaaaa aattagccgg gcatggtggc gcacccgtaa tcccagctac 46080
 ttgggaggcc gaggcaggag aatcgcttga actcaggagg cggagtttgc agtgagccga 46140
 gatggcgcca ctgcaactca gcctgggcaa caagagcgag actccatctc aaaaaaaaaag 46200
 tgtctattgc cttgtatctc cagcaactgac cgaggcctgt aagcagtatg gcagggaaac 46260
 cctcgcctac ctggcctccc tggaggaaga gggaagcctt gagaatgccg acagcacagc 46320
 catgaggaac tgcctgagca agatcaaggc catcggcgag gtacttggag tagtatcatt 46380
 gaggagcatt gttattcttc tgggtgtgcg tgctggtgaa tggccaggga atcggtgatg 46440
 ttctgagcta gttctttctg cacttagaac ttgattctag aaagagattg ttaaaattgg 46500
 aaaatctggc cgggtgcagt gatttatgcg tgtaatccca gcactttggg aggccgagtc 46560
 aggaggatca cttgaggcta gacgggatgg ctcacgcctg taatcccagc actttgggag 46620
 gtgagggtgg gcagatcact tgaggccaag agctagagac cagcctggcc aacagggtga 46680
 aacctgtct ctactaaaaa tacaaaaaat aaccagcat agtggtgcat gcctataaat 46740
 ccaggctact ggaggtggag tcaagagaat tgcttgacc caggaggtgg agttgcatg 46800
 gcaccgagat catgccattg cactccagcc tgggcataag agtgagatct atctcgaaaa 46860
 aaaaaaggat cacttgatcc cagaagttt agacaggcct aggcaacaaa atgagaccct 46920
 gtctctttta aaattaaaaa aaaaaaaatt taaaaaggaa aaacctcctg aaggactttc 46980
 cagctatatt taattcactg gctgttagct gagaatctac tatgttgtca tcattacact 47040
 aggcaatata aaagaagtat aggcagctcc cctgtctctg aggattttta tctagggtggg 47100
 caggtaatac caaaaatgga aataacactg tgttcattca ttactaaag atgtagtgcc 47160
 gagcatggtg gctcacacct gtaatccctg cactttggga gcccgaggca ggaggatcac 47220
 ttgaggccag gagttcaata tcagcctagg caacatagtg agaccctgtc totacaaaaat 47280
 aaaatTTTTT ttttaattag acaggtgtga tgtcacgcac ctgtagtccc agctactcgg 47340
 gaggctgagg caggaggatc gcttgagccc aggatttggg ggccacagtg agctatgatt 47400
 gtgccactgc actccagcct aggcaacaga gtgagaccct gcctctgcta taaataaata 47460
 tgtcctgtat accagaaatt gggttaagta tatatagga cagagaagac atggactcta 47520
 tggaaaagaa aaataagaaa accatttcta tgcagtacag ttttttttca ttttttcccc 47580
 agtgttatga attggaatat attggttaag atacaaggat agggccaggc acgggtggctc 47640
 acgcctgtaa tccaacatt ttgggaggcc gaggcagggt gatcacctga ggtcaggagt 47700
 tctagaccag cctggccaac atggcaaac cccatctcta ctaaaaatac aaaaaattag 47760
 ccaggcgtga tggtaggcgc ctgtaatccc agctacttag gaggctgagg caggagaatc 47820
 acttgaacct gggaggctga ggttgcaagt agctaagatc gtgcctttgc actccaacct 47880
 gggcaacagg agcgaaactc cactcaaaaa ataaataaaa aataaaaagcc atgaggatag 47940
 actgtctaata tggagttcca gaagaagata ataaaaagac cacatccagc caagggtgggt 48000
 ggatcaattg aggccaggaa tttgagacca gcctgggcaa cttagtggcg ccctgtctct 48060
 acaacaaatt taaaaattgg ccgggtgtga tggtagcac ttttgggtccc agctactcag 48120
 gaggctgagg caggaggatc acctaagcct gggagctcaa gactgcagtg aaccgtgtac 48180
 tctgcagtga attgcagagt gtactctgca gagtggttca cctgcagtga accactgtac 48240
 tccagcctgg gtgagagagt aagaccctgt ctctaaaaag aaaaaataaa aagatcaggg 48300
 ccagggtgtg tggctcacgc ctgtaatccc agcactttgg gaggctgagg tgggtggatc 48360
 acttgaggcc aggagtttga gaccagcctg gccaatatag tgaaactcca tctctattaa 48420
 aaatacaaaa ttagccagtt gtggtggcta attttgaat cccagcacat acctgtaatc 48480
 ccagctactt gagaggctga ggcaggataa tcctttgagc ctgggaggcg gaggctgcag 48540
 tgagccgaga ttgaccact gcactccagc ctgggcgaca cagtgagaca ctgcctcaaa 48600
 aaaaaaaaaa aaaaaaaaaa aaaaaaatta tcatctacat ctttgccctcc ctaaataata 48660
 gagttagttt cacatgatata agactacagg taaatggagt cctactgtat atattcttct 48720
 acaactggct ttttactctc acattctgtg agattaatcc aagaagataa attcattttc 48780
 actgctctct cattatctat tgtgtgaata tcgcttaatt catccgtctg ttctcttggg 48840
 gatggacact tgggttattt ccagtttggg ttgttattaa tttgccccta tgatcatata 48900
 tgcacacatg tcagtttctc tagagtata acatagaatc agaattacta ggtattagga 48960
 tgaatctcac acatagaatg ttgagtgaaa aatcaagtca tggaaaaata tatacagtag 49020

gattcctttt atataaagac toaaacattc aaacttaata gggatgcatc tatgtggagt 49080
taaaaaataa taataataag acaggccagg catagtggca catgtctgaa aatcccagca 49140
ctttgggagg ccgaagcggg aggatcaact gaggccagga gtttgataca gcctgggcta 49200
caaagcaaga ctctgactct acaaaaaagt ttaaaagtta gcagagccta gtggtgtgtg 49260
ccggttgctc cagctactca ggaggctgag gcagggagga tcacttgagc ccaggagtgg 49320
gaggctgcag tggactatga tgcaccact gcactctagc ctgggcaaca ggcttttatg 49380
ttttgtcttt aaaaaataaa aataaaataa gacacaacaa ccacgctctt gaattctgtg 49440
tctcttgggg tagaaggcag ggagtaaggg agatagaatc agggcacata caagttcaaa 49500
taggattcca aaggatttga tgtttcttac agcattcctt aagttttaag ttgggtggtg 49560
gataacatgc gtctttatct ttatcctttg aaccattcat atatatatta taccttagta 49620
tgtataacac atcatgtttt aaaaatcata ggctttgctg tttgtcagac cttagtcca 49680
atccctgccc tgccattttgt taccataact gtggtcaaga cgttctacct ctgatccctg 49740
gcttcctatc tgcacagtag gaataatacc cgcctcatat gatcactgtg gaagcaaagt 49800
taagtcaggg tcaaagtgga agatatcatc tgtatttgaa aaagttgtag gccaggata 49860
gtaattcaat ccagcacaat tgggaggcca agacaggagg atcacttgag ccagggaatt 49920
caagaccagc ctgggcaaca tagtgagacc ccatctctat taaaaaaaa aaaaaacaat 49980
tatccgggct tgggtggtgca cctctgtact ccagctact ctggtggtag aggtgggaga 50040
attgcttgag catggaaggt cgaggctgca gtgagctatg atcatgccac tgcactccat 50100
ctctggtgac agggcaagac tgtgtctcaa aaaaataaaa aattagggcc aggtgtggtg 50160
gcttacacct gtaatccag cactttggga ggcaggcca ggagattac ctgaggtcag 50220
gagtttgaga ccagcctggg caaaatagtg aaacccatct ctactaaaaa tataaaaatt 50280
agctgggcat ggtggtgcaa acctgcaatc ccagctactt gggaggctga ggcaggataa 50340
tcccttgagc ccgggaggcg gaggctgcag tgagccgaga ttgtgccact gtactccagc 50400
ctgggcaaca gattgagact ctgtctcaaa aaaaaaaaaa aagaaaagaa aagaaaagaa 50460
gcaaaagaaa aagaaaaatt gtaaaaaaat acctcctgtg tagctagtac aattgacact 50520
gcacaactaa gggactgggg tcttagtggt ggcggaatgg ccttgagggg ggaggcagcg 50580
atggcattct ggaaggaggt ggatcctgag tgaggaggag catgaggttc cagagtttgt 50640
gtcacagaga ggacaagatg gggctctgcaa aagaccccc acaacccctg tggcttgagc 50700
aaggtgtttg ctgggtggcc tcttgccctg ccactctgta agggttacag atggcagagg 50760
agaagagaca ggaggcccca aggtcagttc agcctttgtg atgtgttcac aggagctcct 50820
gcccagggga ctggacatca agcaggagga gctgggggac ctggtggaca aggagatggc 50880
ggccacttca gctgctattg aaactgccac ggccagaata gaggtaggag gttcctgagc 50940
gatctcctga aacgatgctt ttgcagctgc cctctgcaa cactgctcat taaacatgtc 51000
acagtgcgtt attaaggcca tggcaacccc ctaagacaga aaccagaatt tgccaggcac 51060
agtggctcat gcctgtaacc ccagcacctt gggaggatca cttgagtcca ggagtttgag 51120
accagcctgg acatcatagc aagaccccat ctctacaaaa gataaaataa ttagctgggc 51180
atggtggtgc atgcttatag tcccagctac tccaggagct gcaggaggat cacttgagcc 51240
caggaattgg aggctacagt gagccatgat tgcaccactg cactctagcc tgggtgagag 51300
agtgagaccc tgtctgtaaa aatgttttta aattataaaa aaaaaaaga gagaccaaaa 51360
tctgaactcg ctagtttttc tgagtgatag aaaaatgagaa ggtcctcatg agagaaggtc 51420
ctcatgagag aaggtgatta gtttcctcaa gaatcagata ctttggtgtc agaacatcct 51480
gggttctggc caggcgtggt ggctcacgcc tgtattcca gcactttggg aggccgaggc 51540
aggcagatca tgagtcagg agtttgagac cagcctggac aacatggtga aacctggtct 51600
ctactaaaca taaaaaatt agctggcgct ggtggcgct gctgtaatc ccagccactc 51660
agaaggctga ggcagaagaa tcacttgaac ccaggagcag gaggttgcag tgagccgaga 51720
tctgaccgct cactccagcc tgggtgacag agtgagactc catctcaaaa caaaacaaaa 51780
tgacacggat cacagagcca gccactgca gctgcaactc ccctgaatag gttagagtct 51840
ggattctttt ctgactctct caagaatgtg ggcagggact tggggacttc cagattcagg 51900
tttcccagct accacacgat gttggactga aagtatagta agacattagt ggatccttaa 51960
tattcaaggc acatttagaa accatgcttc tttttcacag gagatgctca gcaaatoccg 52020
agcaggagac acaggagtca aattggaggt gaatgaaagg toggctctgag cggcatggtg 52080
ggacctaggg gagcaggatc tgtcttcctg acattggtct atactttgca tacttattag 52140
ggaattagag gagagcagta gcagccacgg ggaagggtct agttgatgta atcatcaata 52200
acagtattaa tgatgacagc aatattgctg tttcactata gaaaagaaac ccgctggtgta 52260
gctcccgcct gtaatcccag cactttggga ggctgaggag ggaggatcac ctgaggtcgg 52320
gagttcgaga ccagcctgcc caatgtggtg aaaagtcact ttactaaaa atacaaaaat 52380
tagccaggcg tgggtggtgca tgcattgaaat cccagctact ttggaggctg aggcaggaga 52440

atcacttgaa cccgggaggt ggaggttgca gtgagccaag atcgtgcctc tgcactctag 52500
cctgggcgac agagcaagac tctgtctcaa aaaagaagaa aagaagtoct aggaatcaca 52560
atctcaactat tccttaogat ggagaccocac aattagatct ctctcactcc ccagcctctc 52620
cttcoggtgcc tccccgcaga atgttccagc aacctcagca cccttcttac ctccctttcc 52680
cattccaagc ttgccttttg ctaggagtgg ggaagagaac cgtcgttttc attgatcttg 52740
gatcttgatc tcagtgatc ctogacttgt ttgtttggca ggatccttgg ttgctgtacc 52800
agcctcatgc aagctattca ggtgctcadc gtggcctcta aggacctcca gagagagatt 52860
gtggagagcg gcaggggtgag cgtgggtgtg ggccctgggc aggaagagga ggcatcgggtg 52920
acagactccc gctccaacgg actctgtgat gctgccgtct tactctgtgt gtccacctga 52980
gtacagagca gccactoctg tagatatcag cagaggccct ggggagaagt cagagctcca 53040
agacctccc agaggggtggc caggcatgtg tcccaactcc agctccctc gcacaggcag 53100
acattgttgg aacttgctgt gggagccctt tttacattgc aaacctgggt actctgaggg 53160
ggacatgagg acttttgcaa aagaagcaga tcctgaatgc caacttgaga ctggcttttc 53220
acagggagtt cggcctaate tgtttctaac agcagagtca actcaccctt acaaatagac 53280
tctgctgtca gcctcttate aacgcaccac ttctgtgtct cacacctttg ctgggattat 53340
catcatcttg tttttttccc atttcccat atgccgcgc agctcgggag gcgtaagtcc 53400
aggggagggg ttaatgagcc tgagaacggg gccataaagc aagcagccag gcctgctttg 53460
ctggtgcctt ttatcttccct gagagctgat gttttctgtt tctaaaatag aaagatggac 53520
tggacaagga ttatctgcag attctcagac ttccagacac ttccaagca taaaaacatt 53580
tttcaaatcc aagtagaaat atttttttt aatgtattca ctgtgtagtc tctttttttt 53640
ggcaaatcgc agagagtggc tagactcttt tcagatcata acttattcat atgagaatat 53700
ttattcttat atatacaata cagaaatgtc atggtatggt agaaatgcta caggcttttg 53760
agtcagaaaa gcccaagatt tacttccatt tatttctctc atgaccttg gaaaagtctt 53820
tttgtttgtt gttttgttgt tttgggtttt ttgtttgctt gtttgtttgt ttttgagact 53880
gagtctcgct ctgtcgccag gctagaatgc agtggtgcaa tctcggctca ctgcaacctc 53940
tgctcctgg gttcaagcga ttctcctgcc tcagcctccc gagtagctgg gattccaggc 54000
acacaccacg cccagctaate ttttgtattt ttagtagaga cggagtcca ccatgttaga 54060
caggatggtc tcaatctctt gacctcgtga tctgcccgcc ttggcctcgc aaagtgctgg 54120
gattacagcc gtgaaccacc gcgcctagcc agggttttgt tttgtttgt ttgagacaga 54180
atctcgctct gttgctcagg ctagagtga gtgacgcgat ctcggtcac tgcaacctcc 54240
gcctcctagg ttcaagtgat gtcctgcct cagcctcctg agtagctggg atcacaggca 54300
cccaccacca cacctagcta atttttgtat tttaataga gacagggtt caccatgtt 54360
ccaagctgg tcttgaactc ctgacctcaa gtgatccacc tgctcagcc tcccaaagt 54420
ctgggattac aggcgagcgc cactgcaccc agctggaaaa gccatttttg tctctgaatc 54480
ttcttctttt ttgtaaaatg ggaatactaa tgcttatgtc tcagagttac tatgaggatg 54540
atgtgggata atatatgtat aaaagcacct gccatatagt acatgctcaa taaaagggtg 54600
ctattactat ttttatttc cctagggtag agcatcccct aaagagtttt atgccaagaa 54660
ctctcgatgg acagaaggac ttatctcagc ctccaaggct gtgggctggg gagccactgt 54720
catggtgtaa gtatctattg gtaccaaggg tcctcccattg acccctcttc cattgatcca 54780
ctccaaacaa tagctaagga gggaaaaaaa aatctgtccc ttagaaataa actattgatc 54840
aggaagtcaa taggaccgag tttaacaggg agcctggctc tcccaggggg cacagggcag 54900
gcagcctccc ctccctgttt agccaagggc gatgggggtg tctggagggtg ggattgtgga 54960
ggagttgcag ctcatcttgc cgtaacctag tccctcttgt cgttttccat cagggatgca 55020
gctgatctgg tggtaacaagg cgttgtggct tttagaggagc taatgggtgt ttctcatgaa 55080
attgctgcta gcacagccca gcttgtggct gcatccaagg taggacctgg ctggacctcc 55140
taggacgctg gaaggcctgg ttagagagta ctaggctagg ttaaagagta cttggctcgc 55200
ttaggcagta cttggctgag ttagagagta cttggccagg ttagatggca cttggctagg 55260
ttagactgta cttggctagg ttaaagagta ctaagggtt agagagtact tggctaggtt 55320
agatggtact tggctgggtt agagaatact aaaggattag agaatacttg gctacgttag 55380
atagtacttg gtaggttag atggtacttg gttagggttag atggtacttg gctagggttag 55440
agactactta ggagttagag agtacttggc taggttagat agtacttggg tgatgggatg 55500
gaactggcct tgagaacaaa tgatctatcc agggctcagg tggattcgat tccaaataaa 55560
aattcaccct gttaaagggtc ccatactatc ctcttaocaga ggcttttttt tttttttttt 55620
ggagacagag ccttgcctcg tcaaccaggc tggagtgcag tggcgtgatc tcagctcact 55680
gcaacctctg cctcccgggt tcaagtgatt ctctgactc agcctcccga gtagctggga 55740
ctacagacac gtgccaccac gccagctaa ttttttgtat ttttagtgga gacggggtt 55800
caccgcgtta gccaggatgg cctcgatctc ctgaccttgt gatccacca cctcagcctc 55860

ccaaagtgct gggattacag gcgtagacca ccacacctgg ccttagaggc ctttcatgtg 55920
 acagcaagga atgtctgtac ccaggtctgt cccaggtgtg ccacacatgt ccctcgcggc 55980
 tttcctcaca tctgttataa gtggccttag agagtcaccg caatgccaac ttccgattag 56040
 cagtcctggc attcaagaaa cactcgaata aggaataggt tggcatcatg aacaggttac 56100
 ctggcttggc tgtgagatte tacaaccctt ctttattagt gttcaaattt gaatcaggcc 56160
 gggcacgatg gctcactcct gtaatcccag cacttttagg aggctgagg tcgggagttc 56220
 aagaccagcc tggccaacat ggtgaaatcc agtttctact aaaaatacaa aaaattagcc 56280
 ggggtgtgtg gcgggcactt gtaatcccag ctgcttggga ggctgaggca ggagaattgc 56340
 ttgaacctgg gaggtggagg ttgcagttag ccaagatcat gccattgcac tccagcctga 56400
 gcaacaagag caaactctgt ctaaaaaaaa attatttttt gaatcaatat catcaatgcc 56460
 cccaactgat tctcatgtaa agtgtgcccc aaatcctttt taagagtttc acagtatccc 56520
 agagtgttct ttattttgtc ttattttctg tcatcccaaa ttggtttcca caactgggca 56580
 cattgtccaa ataagtgata atgcttagag ttgggatcc accaagcaaa gaagccagga 56640
 ataacttttt atatgataga tatgtcagga gctgactata gtcagcagat tttgagaagc 56700
 tgattgggtg ttgcccgttg gccacacatat gtttgctaag aaccatcaga gcaattatct 56760
 gattcagtc ttgttgcctt aggtgttcta tgaacctaaa tctgctttgt cctggtaggt 56820
 gaaagctgat aaggacagcc ccaacctagc ccagctgcag caggcctctc ggggagtgaa 56880
 ccaggccact gccggcgttg tggcctcaac ctttccggc aaatcacaga tcgaagagac 56940
 aggtagcctt tccaaagggg ccttttctt accaccctg ttgagctctt ctctgcctcc 57000
 tccctgtgta tcccaaccaa atcccacagg actgtgtcta aattctttca ttttttcat 57060
 ctttttattt tttttttctt ttatttctgt ttttgagaca ggttctcact ctgtcgcca 57120
 ggctggagtg cagtgggtacc atctctgctc actgcaacct ccacctccct ggttcacacg 57180
 attctcctgc ctgagcttcc cgagtagcta ggattacagg cgcgccccac cagcctggc 57240
 taatttttgt attttttagt gagacggggt ttactatgt tagccaagct ggtctcgaac 57300
 tcctcacctc aggtgatctg cccacctgg cctcccaaag ttgctggga ttataggtgt 57360
 gagccaccgc gccctgccat tttttttct tttttttaga tggagtctt ctctgttgc 57420
 ccaggctgga gtgcaatggc gcaatcttgg ctcaactgaa cctccgcctc cagggttcaa 57480
 gtgattctcc tgccttagcc tcccaagtag ctggaattac aggcgcccac cactgcaccc 57540
 agctaatttt tgtattttta gtaaagacgg ggtttcacca tgttggocag gctggtctca 57600
 aactcctgac ctcatgatcc acccacctca gcctccaaa gagctaggat tacaggagtg 57660
 agccaccgtt tcccgcccc ctgagttact ttcatgcagt gttgacaaat ggcaagccag 57720
 gcgtttccac agagcattgg cattggctgc ctctcaggtg ccagtcagcc agggtagaat 57780
 ttgatgagac cttcttgttt ccatccttgc agacaacatg gacttctcaa gcatgacgt 57840
 gacacagatc aaacgccaa agatggatc tcaggtgaga gctccatctg taagtctaga 57900
 tagcctctat tgcctagaca tggactctga atattattcc catggagaaa agccagaggg 57960
 aatgaacac aggtgcacct actaaccaa gactgaatgt aaatcttggc ggaagaaaca 58020
 gtggaggtag agaaggtgat tagtgaatca ccttgttttt ttttttttgt ttgtcttgtt 58080
 ttgtttttt tcttttttga gacggagtct cgtcctgtcg cccagactgg agtgcagtg 58140
 cgtgatcttg gctcacccga ctctccacct cccaggttca agcaattctt ctacttcaac 58200
 ctcccaagta gctgggatta cagacgtgcg tcaccacacc tagctaattt tttttttttt 58260
 tttgagacgg agtcttgtc tctcgcccag gctggagtgc agtggtgtga tctcggtca 58320
 ctgcaagctc cctctctcag gttcacgcca ttctcctgcc tcagcctctc tgagtagctg 58380
 gaactacagg cgcccaccac catgctgctg taattttttg ttttttttagt agagacaggg 58440
 tttcaccgtg ttagccagga tggctctgat ctctgacct cgtgatctgc ctgccttggc 58500
 ctcccaaaact gctgggatta caagcgtgag ccaccgtgcc cggccttaat tttttatttt 58560
 taatagagat aaggtttcc catgctggcc aggtgtgtct caaactcctg acctgtgat 58620
 ctgcccacct gggccttcca aagtgtgtgg attacaggcg tgagccaccg gccctttttt 58680
 ttttttaaat aagatagaga cggggtctta ctatgttggc caggctggtc tcaactcctg 58740
 gcctcacaca accctcctgc ctgagcctcc caaagggctg ggattacagg tgtgagccac 58800
 cacaccagc caagatcaga ctcttaactc ttgcccacta ttcttggact aacgttctt 58860
 tacatgttat ccaggatctc tttagtatcca caaacttaat agtcttacta taaactgt 58920
 ttttagactt tggtatgaat tattttgctg tgacttttag ttttaaaaga gaaaactcct 58980
 ttaactcatc aaaggcaatg ccattatttc tctctgttgt tctataaagt catattatgt 59040
 aaaataggta tcagcgagat cctttaaaaa aaaattgaca tagtttttct agaataactg 59100
 gtcacagtc ttggggtgaa tttcaaatg gaattgttaa agtactcaa gatgtgtcac 59160
 tcagcatttt tttttctctt tttgagtcag ggccttgcct tgttcccag gctggagtg 59220
 attggtgtga tcacagctca ctgcagcctc aaactcctgg gctcaggcgg atcctcccac 59280

ctcagcctcc tgagtagctg ggactacaga tacacaccac catgcccagc taatTTTTcc 59340
 ttctTTTTtt tTTTTTTTTg agaccgagtc ttgctctgtt gcctaggctg gagtgagtac 59400
 aatggcacga tcttggttca ctgcaacctc cgcttccag gttcaagcga ttctcctacc 59460
 tcagcctcca aagtagctgg gattacagac atgcaccacc acacctggct aatTTTgtat 59520
 atTTTTtagta gagacggggg ttcattcatgt ttgtcaggct agtctctaac tcctgacccc 59580
 aggtgatcca cccacgTTgg cctcccaaag tgctgggatt acgggcataa gccaccatgc 59640
 cgggccagca aatattcctt gaattaagaa ataaaataaa ttacattaat gagaagtaca 59700
 gctagacaag atcacccttt cctttatagt aagatTTTTc aggctTTTTc ctgtagcagt 59760
 aataagtgtg agttgcacaa atacctgaaa gatatcagaa ccatcatatt tgcttcctac 59820
 ctgggTTTTct agggaaactgg gattgaagaa gggttacaac tagattggta ctaccatac 59880
 agttaggcta aacattagag ttgTTTTtag atagtgattt ctcagaggaa cagaagataa 59940
 gccacacctac catttaggca tttgtaaagc atcattcact tccagtagaa gatcacagtg 60000
 agtaaagcat aagagctact gggatagata aatgaatggg agaatggaaa aattagagca 60060
 aagaagagaa ttgaaattga gggccagata aagtatggaa acggtgaata atatggaaga 60120
 agaggaattc acaatgcagc agaaagggag taatcaaaat aagatagatg taggctgggc 60180
 atgggtgtttc atgcctgtaa tctcagcact tagggaggcc aagacaggag gatctcttga 60240
 ggccaggagt tcaactaccg cctgggcagc agaacaagac cccatctcta taaaaaata 60300
 aataaaaatt aaatctctga gagctgatta agagaaaatc ccacttactt tcccttacag 60360
 gttagggtgc tagagctaga aatgaattg cagaaggagc gtcaaaaact gggagagctt 60420
 cggaaaaagc actacgagct tgctgggtgtt gctgaaggct gggaaagaag taagctgact 60480
 ccaaggatgg gcactaacc gcagcaggga ggggaagctg tgtggccagg gccttcccct 60540
 agaatggaac agaacaagga acccaccaga tgggtggctt cttccttctc ttctctcctc 60600
 atggaacctg tttctcttc cgttgtaa atcctTTaag cctaactTTg ggtgaagtta 60660
 tgggtgattg aaagctTTTg ttgtcacctt ttaagacct aacattgatc ataactcact 60720
 cactgccatt ccctcccagg cccagctagc cttgccact ccctgggaat ttagtttctg 60780
 gccagcagag agttgaaaaa tgaataaata gtggtcagtc gagatctggc agacagggac 60840
 aagtgaatta caaggttctg tctatctaga gccagtgaag ggcaacaatt gccaaaagac 60900
 agaagggatc aggtgctagg cggggcgtgg tggctcacgc ctataatccc agcactTTTg 60960
 gagccgagg cgggcggatc acgaggtcag gagatcgaga ccctctggc taacacggtg 61020
 aaaccccgtc tctactaaaa atacaaaaaa ttagcctggc gtgggtgggg ggcctgtag 61080
 tcccagctac tcaggaggct gaggcaggag aatggcgtga acccgggagg cggagcttgc 61140
 agtgagccga gattgcgcca ctgcactcca gctgggtga cagagcgaga ctccatctca 61200
 aaaaaaaaaa aaaaaaaaaa gaaggatca gctgtgaga aatggagtgg catggagtca 61260
 cattcaggct gccagcccag agacttgcca tccatcacct cgagtctata tgttgttgtc 61320
 tgcatgtcga aagcagtcaa cttgatcaca gccaaagactg ttcaototta gtggacttga 61380
 ctgactctgg gcaccaacca tagacagatc ctagtcatc ctagacaoga aacttgggga 61440
 ctatgtgatt acaccatggt tataattata atgtatTTTT gagccccat ggtggcagg 61500
 tttgttggtt aagtgaatga atgcattcat gagtgcTTTg caacataaat tatcattgtc 61560
 ttttaggaac agaggcatct ccacctacac tgcaagaagt ggtaaccgaa aaagaataga 61620
 gccaaacca caccccatat gtcagtgtaa atccttgTTa cctatctcgt gtgtgttatt 61680
 tcccagcca caggccaaat ccttggagtc ccaggggag ccacaccact gccattacc 61740
 agtgccgagg acatgcatga cacttccaaa gactccctcc atagcgacac cctttctgtt 61800
 tggaccatg gtcattctctg ttctTTTTcc gcctccctag ttagcatcca ggctggccag 61860
 tgctgcccat gagcaagcct aggtacgaag aggggtgggt gggggcaggg ccaactcaaca 61920
 gagaggacca acatccagtc ctgctgacta tttgaccccc acaacaatgg gtatccttaa 61980
 tagaggagct gcttgttgtt tgttgacagc ttggaaaggg aagatcttat gccttttctt 62040
 ttctgTTTTc ttctcagTct tttcagtttc atcatttgca caaacttTtg agcatcagag 62100
 ggctgatgga ttccaaacca ggacactacc ctgagatctg cacagtcaga aggacggcag 62160
 gagtgcctg gctgtgaaat ccaaagccat tctccccctc tttgggcagt gccatggatt 62220
 tccactgctt cttatggtgg ttggttgggt tttttggTTt tgtTTTTTTt ttttaagTTt 62280
 actcacatag ccaactctcc caaagggcac acccctgggg ctgagtctcc agggccccc 62340
 aactgtggta gctccagcga tggtgctgcc caggcctctc ggtgctccat ctccgcctcc 62400
 aactgacca agtgctggcc caccagTcc atgctccagg gtcaggcgga gctgctgagt 62460
 gacagctttc ctcaaaaagc agaaggagag tgagtgcctt tccctcctaa agctgaatcc 62520
 cggcggaaaag cctctgtccg cctttacaag ggagaagaca acagaaagag ggacaagagg 62580
 gttcacacag cccagttccc gtgacgaggc tcaaaaactt gatcacatgc ttgaatggag 62640
 ctggtgagat caacaacact acttccctgc cggaatgaac tgtccgtgaa tggctctgt 62700

caagcggggc gtctcccttg gccagagac ggagtgtggg agtgattccc aactcctttc 62760
 tgcagacgtc tgccttggca tcctcttgaa taggaagatc gttccacott ctacgcaatt 62820
 gacaaaccog gaagatcaga tgcaattgct cccatcaggg aagaacccta tacttggttt 62880
 gctaccctta gtatttatta ctaacctccc ttaagcagca acagcctaca aagagatgct 62940
 tggagcaatc agaacttcag gtgtgactct agcaaggctc atctttctgc cggctacat 63000
 cagccttcaa gaatcagaag aaaggccaag gtgctggact gttactgact tggatcccaa 63060
 agcaaggaga tcatthggag ctcttgggtc agagaaaatg agaaaggaca gagccagcgg 63120
 ctccaactcc tttcagccac atgccccagg ctctcgtctc cctgtggaca ggatgaggac 63180
 agagggcaca tgaacagctt gccagggatg ggcagcccaa cagcactttt cctcttctag 63240
 atggacccca gcatttaagt gaccttctga tcttggaaaa acagcgtctt ccttctttat 63300
 ctatagcaac tcatthgggg tagccatcaa gcacttccca ggatctgctc caacagaata 63360
 ttgctagggt ttgctacatg acgggttgg agacttctgt ttgatcactg tgaaccaacc 63420
 cccatctccc tagccacccc cctcccccaa ctccctctct gtgcattttc taagtgggac 63480
 attcaaaaaa ctctctccca ggacctcgga tgaccatact cagacgtgtg acctccatac 63540
 tgggctaagg aagtatcagc actagaaatt gggcagctct aatgthgaat gctgctttct 63600
 gcttagtatt tttttgattc aaggctcaga aggaatgggt cgtggcttcc ctgtcccagt 63660
 tgtggcaact aaaccaatcg gtgtgttctt gatgcgggtc aacatttcca aaagtggcta 63720
 gtcctcactt ctagatctca gccattctaa ctcatatgtt cccaattacc aaggggtggc 63780
 cgggcacagt ggctcacgcc tgtaatccca ccaacttgag aggctgaggt ggtaggatca 63840
 cctgaggtca ggagttcaag accagcctgt ccaacatggt gaaaccccca tctctactaa 63900
 aaataccaaa aattagccga gcgtagtgac ggggtgccgt aatcccagct actcaggagg 63960
 ctgagacagg agaatcacct gaaccccaga ggcagagggt gcagtgagct gagatcacgc 64020
 cattgtactc cagcctgggg aacaagagca aaactccgtc tcaaaaaaaa aaaaaaatta 64080
 caaatggggc aaacagtcta gtgtaatgga tcaaattaag attctctgcc cagccgggca 64140
 cagtggcgca tgectgtaat cccagaactt tgggaggcca agacgggatg attgcttgag 64200
 ctcaggagtt tgagaccagg ctgggcatca tagcaagacc tcatctctac taaaattcaa 64260
 aaacaaaatt agccgggcat gatggtgcat gcctgtagtc tcagctagtt ggggagctaa 64320
 ggtgggagaa ttgcttgagc ttgggaagtc gaggctgcag tcagccctga ttgtgccagt 64380
 gcactccggc ctgggtgaca gagtgagacc ctgtctcaaa aaaaaaaaga ttctgtgtca 64440
 gagcccagcc caggagtttg aggctgcaat gagccatgat ttcccactgc actccagcct 64500
 gagtgacaga gcgagactcc atctctttaa aaacaaacaa aaaattatct gaatgatcct 64560
 gtctctaaaa agaagccaca gaaatgttta aaaacttcat cgacttagcc tgagtataa 64620
 cggttaagaa agcacttaaa cagaagcaga ggctaattca gtgtcacatg aggaagtagc 64680
 tgtcagatgt cacataatta ctttctgtaat agctcagatt agaathggta cccattctc 64740
 tagacaaaat caaattgtcc tattgtgact cttctaaaaa tgaagatgaa gagctattta 64800
 atgacacacc ttggattaaa acgggaatca catcttaaag ctaaaaatga acctgcaagc 64860
 cttctaaatg agtcaactgag catcactagt gacaagtctc ggggtgagcgt aaatgggtca 64920
 tgacaagatg ggacagcaac aaaatcatgg cttaggatcg acaagaagtt aaaaaacagc 64980
 tgcactctgt acttaagttt gtaagacagt gcoctgagac ctctagagaa aagatgtttg 65040
 tttacataag agaaagaggc cagacatggt gtotcacaag tttaatccca gcactttggg 65100
 aggcaggggc ggggtgatca cctgaggtca ggagttcaag actagcctgg ccaacatggt 65160
 gaaaccccgt ctctactaaa aatacaaaaa ttagccgggc atgggtggcag gcgcctataa 65220
 tcccagctac tggggaggct gaggcaggag aatcacttga acccggggga cagaggttgt 65280
 agtgagccaa gatcgacca ctgcactcca gcctgggtca cagagtgaga ctccatctca 65340
 aaaaaaaaaa agagagagag agagaaagaa atagaagaga agagccatct tggcaggtt 65400
 attttatatc tgagcaagga gtttaaatga gactagttta gattgtctgc tgat 65454

<210> 12
 <211> 516
 <212> DNA
 <213> H. sapiens

<220>

<400> 12
 tctgtggaag gtttggaggg gagagagggg cagctggatg ctcttgggcc acggtcgccc 60

ctgatctctg cgcctcttcc tcctgctccg ggagaaataa tgtttccctg ggggatgaaa 120
 gcatctcttt gtgcgggctt taattgccat gttgttgtgc caagggagtg agtggcggcg 180
 ggaccagcag ctgggcacag ccaatgccag gcagtgggtgc ccactccctc aggacgcca 240
 gccagctggc tcctgggagc gctgcccacc tctgccccca gctgggcgcc tgcaaggaac 300
 cgaccacccg tggggctggg ggaggttggc tggaggagga gaaagggcg ggctctggga 360
 gggctctcagc cactctcaga ggcttattca tctcatcctc ctttccctcc cccttcttgt 420
 ttttcagact gtcagcatca ataaggccat taatacgcag gaagtggctg taaaggaaaa 480
 acacgccaga aatatacctt ggatgttgct tggaag 516

<210> 13
 <211> 219
 <212> DNA
 <213> H. sapiens

<220>

<400> 13
 gaggtcgatc tcaggattta ctcttactg tgttgagcag gaaagttagg tggtctctct 60
 gtcccctaca tggggctgat gaagacacc agcaccctc aggtccttct ccaccctag 120
 gttgaaagat ctgttctacc gctccagcaa cctgcagtac ttcaagcggc tcattcagat 180
 cccccagctg cctgaggtaa gcatgccc aa ccacacacc 219

<210> 14
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 14
 cgtattaatg gccttattga 20

<210> 15
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 15
 tcctgcgtat taatggcctt 20

<210> 16

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

<220>

<223> Antisense Oligonucleotide

<400> 16
 ccacttcctg cgtattaatg 20

<210> 17
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 17
 catcctgctc atgtcactca 20

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

<220>

<223> Antisense Oligonucleotide

<400> 18
 gccccacatc ctgctcatgt 20

<210> 19
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 19
 aggtggcccc acatcctgct 20

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

<220>

<223> Antisense Oligonucleotide

<400> 20

cgctcaggtg gccccacatc 20

<210> 21
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 21
 taccctcgc tcaggtggcc 20

<210> 22
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 22
 ttttggtgtg gtactccatc 20

<210> 23
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 23
 gggatttttg gtgtgtact 20

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

<220>

<223> Antisense Oligonucleotide

<400> 24
 aacctgggat ttttggtgtg 20

<210> 25
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 25

tccagcctcg tccagctggc

20

<210> 26

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 26

taactggaaa aagttgttca

20

<210> 27

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 27

aactgctcca tgaagcggtc

20

<210> 28

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 28

aacagatcct tcaactttgt

20

<210> 29

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 29

tccatgaggt catccttctc

20

<210> 30

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 30
ccatgtccat gaggtcatcc 20

<210> 31

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 31
tcatcaaact tgttgtcaaa 20

<210> 32

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 32
tgtgccttca atccactgat 20

<210> 33

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 33
ttcttccgca gcaggtcagc 20

<210> 34

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 34
acctgtttgg tcacctctgc 20

<210> 35

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

 <220>

 <223> Antisense Oligonucleotide

 <400> 35
 caggctgcct tgcagaacct 20

 <210> 36
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>

 <223> Antisense Oligonucleotide

 <400> 36
 gggcaggcca gatactggct 20

 <210> 37
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>

 <223> Antisense Oligonucleotide

 <400> 37
 cttctgggca ggccagatac 20

 <210> 38
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>

 <223> Antisense Oligonucleotide

 <400> 38
 tgccatactg cttacaggcc 20

 <210> 39
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>

 <223> Antisense Oligonucleotide

 <400> 39

tccctcttcc tccagggagg 20

<210> 40
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 40
gcaggagctc ctcgccgatg 20

<210> 41
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 41
ggatgctgta cccctgccgc 20

<210> 42
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 42
agctgcatcc accatgacag 20

<210> 43
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 43
tgcagccaca agctgggctg 20

<210> 44
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 44
agaggcctgc tgcagctggg 20

<210> 45
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 45
ccccgagagg cctgctgcag 20

<210> 46
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 46
tcactccccg agaggcctgc 20

<210> 47
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 47
ctggttcact ccccgagagg 20

<210> 48
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 48
gtggcctggt tcactccccg 20

<210> 49
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 49

ccggaaatgg ttgaggccac

20

<210> 50

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 50

atttgccgga aatggttgag

20

<210> 51

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 51

gtcatgcttg agaagtccat

20

<210> 52

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 52

ttcttcccag ccctcagcaa

20

<210> 53

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 53

gggtgttggt ttggctctat

20

<210> 54

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

<220>

<223> Antisense Oligonucleotide

<400> 54
tcggcactgg gtaatggcag 20

<210> 55
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 55
ggcagcactg gccagcctgg 20

<210> 56
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 56
tcctctatta aggataccca 20

<210> 57
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 57
tgctcacaag tttgtgcaaa 20

<210> 58
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 58

tgaccctgga gcatggactg 20

<210> 59
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 59
aaaggcactc actctccttc 20

<210> 60
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 60
ggacagttca ttccggcagg 20

<210> 61
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 61
tcaagaggat gccaaaggcag 20

<210> 62
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 62
ccaagtatag ggttcttccc 20

<210> 63
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 63

tgattgctcc aagcatctct

20

<210> 64

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 64

tgaagttctg attgctccaa

20

<210> 65

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 65

ctgacccaag agctccaaat

20

<210> 66

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 66

tggctgaaag gagttggagc

20

<210> 67

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 67

agctgttcat gtgccctctg

20

<210> 68

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 68
gatcagaagg tcacttaaatt 20

<210> 69

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 69
ccgtcatgta gcaaaccta 20

<210> 70

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 70
gaagtctcac aaccgcat 20

<210> 71

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 71
atgcacagag agggagttgg 20

<210> 72

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 72
atggaggtca cacgtctgag 20

<210> 73

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

<220>

<223> Antisense Oligonucleotide

<400> 73
gcttcttttt agagacagga 20

<210> 74
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 74
aactgaatt agcctctgct 20

<210> 75
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 75
ggtagccat tctaactga 20

<210> 76
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 76
ttaagatgtg attcccgttt 20

<210> 77
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 77

cactagtgat gctcagtgac 20

<210> 78
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 78
 aggaactcac gttcgggtgt 20

<210> 79
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 79
 agagactcac tttggtgtgg 20

<210> 80
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 80
 taatcaagtt caatgatcac 20

<210> 81
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 81
 gcccataaaa ggctgagct 20

<210> 82
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 82
 ttggactcac ttctcatcct 20

<210> 83
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 83
 cctctgcatt ctgcaaaaga 20

<210> 84
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 84
 ttcacctctcg ttaattaagc 20

<210> 85
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 85
 ctctgctgat atctacagga 20

<210> 86
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 86
 catggcaatt aaagcccgca 20

<210> 87
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 87

ggcacaacaa catggcaatt

20

<210> 88

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 88

gcattggctg tgcccagctg

20

<210> 89

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 89

ggatgagatg aataagcctc

20

<210> 90

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 90

ggtgtcttca tcagcccat

20

<210> 91

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 91

tgtgtggttg ggcatgctta

20

<210> 92

<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 92
tcaataaggc cattaatagc 20

<210> 93
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 93
aaggccatta atacgcagga 20

<210> 94
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 94
cattaatagc caggaagtgg 20

<210> 95
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 95
tgagtgacat gagcaggatg 20

<210> 96
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 96
acatgagcag gatgtggggc 20

<210> 97
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 97

agcaggatgt ggggccacct 20
<210> 98
<211> 20
<212> DNA
<213> H. sapiens
<220>
<400> 98
gatgtggggc cacctgagcg 20
<210> 99
<211> 20
<212> DNA
<213> H. sapiens
<220>
<400> 99
ggccacctga gcgaggggta 20
<210> 100
<211> 20
<212> DNA
<213> H. sapiens
<220>
<400> 100
gatggagtac cacaccaaaa 20
<210> 101
<211> 20
<212> DNA
<213> H. sapiens
<220>
<400> 101
agtaccacac caaaaatccc 20
<210> 102
<211> 20
<212> DNA
<213> H. sapiens
<220>
<400> 102
cacaccaaaa atcccaggtt 20
<210> 103
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 103
gccagctgga cgaggctgga 20

<210> 104
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 104
tgaacaactt tttccagtta 20

<210> 105
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 105
acaaagttga aagatctggt 20

<210> 106
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 106
gagaaggatg acctcatgga 20

<210> 107
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 107
ggatgacctc atggacatgg 20

<210> 108
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 108
tttgacaaca agtttgatga 20

<210> 109
<211> 20

<212> DNA
<213> H. sapiens

<220>

<400> 109
atcagtggat tgaaggcaca 20

<210> 110
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 110
gctgacctgc tgcggaagaa 20

<210> 111
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 111
gcagaggtga ccaaacaggt 20

<210> 112
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 112
aggttctgca aggcagcctg 20

<210> 113
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 113
agccagtatc tggcctgccc 20

<210> 114
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 114
gtatctggcc tgcccagaag 20

<210> 115
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 115
ggcctgtaag cagtatggca 20

<210> 116
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 116
cctccctgga ggaagaggga 20

<210> 117
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 117
gcggcagggg tacagcatcc 20

<210> 118
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 118
ctgtcatggt ggatgcagct 20

<210> 119
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 119
cagcccagct tgtggctgca 20

<210> 120
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 120
cccagctgca gcaggcctct 20

<210> 121
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 121
ctgcagcagg cctctcgggg 20

<210> 122
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 122
gcaggcctct cggggagtga 20

<210> 123
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 123
cctctcgggg agtgaaccag 20

<210> 124
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 124
cggggagtga accaggccac 20

<210> 125
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 125
gtggcctcaa ccatttccgg 20

<210> 126
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 126
ctcaaccatt tccggcaaat 20

<210> 127
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 127
atggacttct caagcatgac 20

<210> 128
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 128
ttgctgaggg ctgggaagaa 20

<210> 129
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 129
atagagccaa accaacaccc 20

<210> 130
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 130
ctgccattac ccagtgccga 20

<210> 131
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 131
ccaggctggc cagtgctgcc 20

<210> 132

<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 132
tgggtatcct taatagagga 20

<210> 133
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 133
tttgcaaaaa cttgtgagca 20

<210> 134
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 134
cagtccatgc tccagggtca 20

<210> 135
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 135
gaaggagagt gagtgccttt 20

<210> 136
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 136
cctgccgaa tgaactgtcc 20

<210> 137
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 137

ctgccttggc atcctcttga 20

<210> 138
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 138
gggaagaacc ctatacttgg 20

<210> 139
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 139
agagatgctt ggagcaatca 20

<210> 140
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 140
ttggagcaat cagaacttca 20

<210> 141
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 141
at ttggagct cttgggtcag 20

<210> 142
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 142
gctccaactc ctttcagcca 20

<210> 143
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 143

cagagggcac atgaacagct

20

<210> 144

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 144

atttaagtga ctttotgatc

20

<210> 145

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 145

taggttttgc tacatgacgg

20

<210> 146

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 146

atgacggggtt gtgagacttc

20

<210> 147

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 147

ccaactccct ctctgtgcat

20

<210> 148

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 148

ctcagacgtg tgacctccat

20

<210> 149

<211> 20

<212> DNA
<213> H. sapiens

<220>

<400> 149
tcctgtctct aaaaagaagc 20

<210> 150
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 150
agcagaggct aattcagtgt 20

<210> 151
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 151
tcagattaga atggctaccc 20

<210> 152
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 152
aaacgggaat cacatcttaa 20

<210> 153
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 153
gtcactgagc atcactagtgt 20

<210> 154
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 154
acacccgaac gtgagttcct 20

<210> 155
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 155
ccacaccaaa gtgagtctct 20

<210> 156
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 156
gtgatcattg aacttgatta 20

<210> 157
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 157
aggatgagaa gtgagtccaa 20

<210> 158
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 158
tcttttgag aatgcagagg 20

<210> 159
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 159
gcttaattaa cgaggatgaa 20

<210> 160
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 160
tcctgtagat atcagcagag 20

<210> 161
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 161
tgcgggcttt aattgcatg 20

<210> 162
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 162
aattgcatg ttgttgaggc 20

<210> 163
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 163
cagctgggca cagccaatgc 20

<210> 164
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 164
atggggctga tgaagacacc 20

<210> 165
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 165
taagcatgcc caaccacaca 20

<210> 166
<211> 19
<212> DNA
<213> H. sapiens

<220>

<400> 166

cgagaggcgg acgggaccg

19

<210> 167

<211> 21

<212> DNA

<213> H. sapiens

<220>

<400> 167

cgagaggcgg acgggaccgt t

21

<210> 168

<211> 21

<212> DNA

<213> H. sapiens

<220>

<400> 168

cggtcccgtccgcctctcggt

21