

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

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

22 August 2019 (22.08.2019)



(10) International Publication Number

WO 2019/161105 A1

(51) International Patent Classification:

C12N 15/11 (2006.01)

C12N 15/12 (2006.01)

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(21) International Application Number:

PCT/US2019/018076

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(22) International Filing Date:

14 February 2019 (14.02.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/630,565

14 February 2018 (14.02.2018) US

(71) Applicant: **DEEP GENOMICS INCORPORATED**

[CA/CA]; MaRs Centre, 661 University Ave, Suite 480, Toronto, ON M5G 1M1 (CA).

(72) Inventor; and

(71) Applicant (for US only): **SCHMITGES, Frank** [US/US];

MaRs Centre, 661 University Ave, Suite 480, Toronto, ON M5G 1M1 (CA).

(72) Inventors: **MERICO, Daniele**; MaRs Centre, 661 University Ave, Suite 480, Toronto, ON M5G 1M1 (CA). **WIEN-**

HOLDS, Erno; MaRs Centre, 661 University Ave, Suite 480, Toronto, ON M5G 1M1 (CA). **O'HARA, Matthew**;

MaRs Centre, 661 University Ave, Suite 480, Toronto, ON M5G 1M1 (CA).

(74) Agent: **MCDONALD, J., Cooper**; Clark & Elbing LLP,

101 Federal Street, 15th Floor, Boston, MA 02110 (US).

(81) Designated States (unless otherwise indicated, for every

kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

(54) Title: OLIGONUCLEOTIDE THERAPY FOR WILSON DISEASE

(57) Abstract: The present disclosure provides antisense oligonucleotides, compositions, and methods that target ATP7B exon 6 or a flanking intron, thereby modulating splicing of ATP7B pre-mRNA to increase the level of ATP7B mRNA molecules having exon 6, e.g., to provide a therapy for Wilson disease. The present disclosure provides an antisense oligonucleotide including a nucleobase sequence at least 70% complementary to an ATP7B target sequence in exon 6, a 5'-flanking intron, a 3'-flanking intron, or a combination of exon 6 and the 5'-flanking or 3'-flanking intron.



WO 2019/161105 A1

OLIGONUCLEOTIDE THERAPY FOR WILSON DISEASE

FIELD OF THE INVENTION

[0001] The present invention relates to the field of oligonucleotides and their use for the treatment of disease. In particular, the invention pertains to antisense oligonucleotides that may be used in the treatment of Wilson disease.

BACKGROUND

[0002] Wilson disease is a fatal copper homeostasis disorder, typically diagnosed in patients between the ages of 5 and 35, leading to hepatic and neurologic symptoms due to free copper accumulation. The prevalence of Wilson disease is estimated to be 1 in 30,000 individuals. Impaired hepato-biliary excretion of copper may be caused by defects in the trans Golgi copper transporter ATP7B in liver hepatocytes. ATP7B is required for transport of copper from the cytoplasm to the endomembrane compartment, which is followed by the release of copper into the bile via vesicular transport from the Golgi and/or loading of copper onto apoceruloplasmin (CP) for bloodstream transport. Free copper accumulation causes direct oxidative damage to biomolecules (DNA, lipids, mitochondria) and oxidative damage/stress triggering apoptosis.

[0003] Present treatment guidelines for Wilson disease recommend an initial treatment with chelating agents, which sequester free copper and lead to its excretion, followed by maintenance treatment with chelating agents or zinc salts, which block intestinal absorption. Initial treatment with chelating agents (2-12 months) is traditionally required to remove excess copper from patients. Worsening liver condition can cause increased copper release from dead hepatocytes, as such, maintenance treatment may only be started when a stable liver condition is achieved. Maintenance treatment with zinc salts, however, is complicated by adherence difficulties and gastrointestinal side effects. Adverse effects due to chelation therapy further causes discontinuation in 20-30% patients. Chelation therapy is associated with immune reactions, reduced wound healing, neutropenia or thrombocytopenia, lymphadenopathy, proteinuria and nephrotoxicity, long-term liver iron accumulation, and spikes in free copper resulting in neurological damage.

[0004] An ATP7B mutation associated with Wilson disease is a single base pair change in exon 6 from adenosine to cytosine causing a corresponding change in amino acid 645 from methionine to arginine (M645R): chr13:52535985:A:C [hg19/b37]; NCBI Reference Sequence NG_008806.1 (SEQ ID NO: 1):g.54646T>G; NM_0053.3(ATP7B):c.1934T>G (p.Met645Arg). The M645R mutation is presumed to result in partial to complete loss-of-function based on genetic evidence and compound heterozygotes with a truncating variant typically have Wilson disease onset between ages 4-15.

[0005] Recent analysis of the M645R mutated ATP7B protein, expressed using a cDNA-encoding plasmid, indicated that the mutant protein has similar capability to uptake copper in microsomal fractions of sf9 insect cells expressing only the mutant protein, whereas all other pathogenic mutations tested showed a decrease in capability. Further research found a similar result, with the M645R mutated ATP7B protein having no biochemical defect in numerous cellular assays.

[0006] Previous analysis of the M645R and associated mutations may have failed, however, to consider the effect of such a mutation on splicing. Splicing is a natural biological mechanism that occurs within human cells. Splicing may be used to process the primary messenger ribonucleic acid (mRNA) that is transcribed from deoxyribonucleic acid (DNA), before the mRNA is translated into protein. Splicing may

involve removing one or more contiguous segments of mRNA (introns) to conjoin the remaining segments (exons), delimited by pairs of 5' splice sites and 3' splice sites. Alternative splicing, which may be the splicing together of different combinations of exons, may result in multiple mRNA transcripts from a single gene.

- 5 **[0007]** Certain human genetic diseases (e.g., caused by genetic aberrations, such as point mutations), may be caused by aberrant splicing. As such, there is a need for a splicing mediator to treat diseases that are caused by aberrant splicing.

SUMMARY OF THE INVENTION

- 10 **[0008]** In general, the invention provides an oligonucleotide including a nucleobase sequence complementary to a sequence within ATP7B exon6, a flanking intron, or a combination thereof.
- [0009]** In one aspect, the invention provides an antisense oligonucleotide including a nucleobase sequence at least 70% complementary to an ATP7B target sequence in exon 6, a 5'-flanking intron, a 3'-flanking intron, or a combination of exon 6 and the 5'-flanking or 3'-flanking intron.
- 15 **[0010]** In some embodiments, the ATP7B target sequence reduces the binding of a splicing factor to an intronic splicing silencer in the 5'-flanking or 3'-flanking intron.
- [0011]** In some embodiments, the ATP7B target sequence includes at least one nucleotide located among positions 54672-54680 in SEQ ID NO: 1. In some embodiments, the ATP7B target sequence includes at least one nucleotide located among positions 54691-54701 in SEQ ID NO: 1. In some
- 20 **[0012]** In some embodiments, the ATP7B target sequence includes at least one nucleotide located among positions 54492-54506 in SEQ ID NO: 1.
- [0012]** In some embodiments, the ATP7B target sequence includes at least one nucleotide located among positions 54472-54516 in SEQ ID NO: 1. In some embodiments, the ATP7B target sequence includes at least one nucleotide located among positions 54522-54593 in SEQ ID NO: 1. In some
- 25 **[0013]** In some embodiments, the ATP7B target sequence includes at least one nucleotide located among positions 54665-54718 in SEQ ID NO: 1.
- [0013]** In some embodiments, the nucleobase sequence is complementary to a sequence within the 5'-flanking intron. In some embodiments, the ATP7B target sequence is located within the 5'-flanking intron among positions up to 54517 in SEQ ID NO: 1. In some embodiments, the nucleobase sequence
- 30 **[0014]** In some embodiments, the ATP7B target sequence is located within the combination of the 5'-flanking intron and exon 6. In some embodiments, the ATP7B target sequence is located within the combination of the 5'-flanking intron and exon 6 among positions 54562 to 54593 in SEQ ID NO: 1. In some embodiments, nucleobase sequence has at least 70% sequence identity to SEQ ID NO: 11.
- [0015]** In some embodiments, the ATP7B target sequence is located within exon 6 or the
- 40 **[0015]** In some embodiments, the ATP7B target sequence is located within the combination of the 3'-flanking intron and exon 6. In some embodiments, the ATP7B target sequence is

located among positions 54631 to 54677 in SEQ ID NO: 1. In some embodiments, the nucleobase sequence has at least 70% sequence identity to SEQ ID NO: 22, 23, 24, or 25.

[0016] In some embodiments, the ATP7B target sequence is located within the 3'-flanking intron. In some embodiments, the ATP7B target sequence is located among positions 54655 to 54738 in SEQ ID

NO: 1. In some embodiments, the 5'-terminal nucleotide of the oligonucleotide is complementary to neither position 54695 nor position 54696 of SEQ ID NO: 1. In some embodiments, the nucleobase sequence has at least 70% sequence identity to SEQ ID NO: 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, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, or 155. In some embodiments, the nucleobase sequence has at least 70% sequence identity to SEQ ID NOs: 39, 48, 49, 50, 61, 63, 64, 65, 66, 67, 76, 77, 78, 79, 80, 81, 90, 91, 92, 93, 94, 95, 100, 103, 104, 121, and 149.

[0017] In some embodiments, the sequence identity of any nucleobase sequence is at least 80%, 85%, 90%, or 95%, e.g., at least 90%. In some embodiments, the sequence identity is 100%.

[0018] In some embodiments, the antisense oligonucleotide includes at least one modified nucleobase.

[0019] In some embodiments, the antisense oligonucleotide includes at least one modified

internucleoside linkage. In some embodiments, the modified internucleoside linkage is a phosphorothioate linkage. In some embodiments, the phosphorothioate linkage is a stereochemically enriched phosphorothioate linkage. In some embodiments, at least 50% of internucleoside linkages in the antisense oligonucleotide are independently the modified internucleoside linkage. In some

embodiments, at least 70% of the internucleoside linkages in the antisense oligonucleotide are independently the modified internucleoside linkage. In some embodiments, all internucleoside linkages in the antisense oligonucleotide are independently the modified internucleoside linkage.

[0020] In some embodiments, the antisense oligonucleotide includes at least one modified sugar nucleoside. In some embodiments, at least one modified sugar nucleoside is a 2'-modified sugar nucleoside. In some embodiments, at least one 2'-modified sugar nucleoside includes a 2'-modification selected from the group consisting of 2'-fluoro, 2'-methoxy, and 2'-methoxyethoxy. In some

embodiments, the 2'-modified sugar nucleoside includes the 2'-methoxyethoxy modification.

[0021] In some embodiments, all nucleosides in the antisense oligonucleotide are independently the modified sugar nucleosides.

[0022] In certain embodiments of the invention, all internucleoside linkages in the antisense oligonucleotide are phosphorothioate diester linkages, and all nucleosides in the antisense oligonucleotide are 2'-methoxyethoxy modified ribose nucleosides. Such antisense oligonucleotides may also further include a targeting moiety, such as N-acetylgalactosamine or a cluster thereof.

[0023] In some embodiments, at least one modified sugar nucleoside is a bridged nucleic acid. In some embodiments, the bridged nucleic acid is a locked nucleic acid (LNA), ethylene-bridged nucleic acid (ENA), or cEt nucleic acid.

[0024] In some embodiments, the antisense oligonucleotide is a morpholino oligomer.

[0025] In some embodiments, the antisense oligonucleotide further includes a targeting moiety. In some embodiments, the targeting moiety is covalently conjugated at the 5'-terminus of the antisense oligonucleotide. In some embodiments, the targeting moiety is covalently conjugated at the 3'-terminus of the antisense oligonucleotide. In some embodiments, the targeting moiety is covalently conjugated at an internucleoside linkage of the antisense oligonucleotide.

[0026] In some embodiments, the targeting moiety is covalently conjugated through a linker. In some embodiments, the linker is a cleavable linker (e.g., a linker including $-S-S-$, $-C(O)O-$, $-C(O)S-$, $-OC(O)-$, $-SC(O)-$).

[0027] Examples of targeting moieties include N-acetylgalactosamine, glycyrrhetic acid, glycyrrhizin, lactobionic acid, lactoferrin, IgA, a bile acid (e.g., lithocholyltaurine or taurocholic acid), and clusters thereof. In some embodiments, the targeting moiety includes N-acetylgalactosamine. In some embodiments, the targeting moiety is an N-acetylgalactosamine cluster.

[0028] In some embodiments, the antisense oligonucleotide includes at least 12 nucleosides. In some embodiments, the antisense oligonucleotide includes at least 16 nucleosides. In some embodiments, the antisense oligonucleotide includes a total of 50 nucleosides or fewer. In some embodiments, the antisense oligonucleotide includes a total of 30 nucleosides or fewer. In some embodiments, the antisense oligonucleotide includes a total of 20 nucleosides or fewer. In some embodiments, the antisense oligonucleotide includes a total of 16 to 20 nucleosides. In some embodiments, the antisense oligonucleotide includes a total of 16 to 19 nucleosides.

[0029] In another aspect, the invention provides a pharmaceutical composition including the antisense oligonucleotide of the invention and a pharmaceutically acceptable excipient.

[0030] In yet another aspect, the invention provides a method of increasing the level of exon 6-containing ATP7B mRNA molecules in a cell expressing an aberrant ATP7B gene by contacting the cell with the antisense oligonucleotide of the invention. In some embodiments, the cell is in a subject.

[0031] In still another aspect, the invention provides a method of treating Wilson disease in a subject having an aberrant ATP7B gene by administering a therapeutically effective amount of the antisense oligonucleotide of the invention or the pharmaceutical composition of the invention to the subject in need thereof. In some embodiments, the administering step is performed parenterally. In some embodiments, the aberrant ATP7B gene is ATP7B having a g.54646T>G mutation in SEQ ID NO: 1.

[0032] Recognized herein is the need for compositions and methods for treating diseases that may be caused by abnormal splicing resulting from an underlying genetic aberration. In some cases, antisense nucleic acid molecules, such as oligonucleotides, may be used to effectively modulate the splicing of targeted genes in genetic diseases, in order to alter the gene products produced. This approach can be applied in therapeutics to selectively modulate the expression and gene product composition for genes involved in genetic diseases.

[0033] The present disclosure provides compositions and methods that may advantageously use antisense oligonucleotides targeted to and hybridizable with nucleic acid molecules that encode for ATP7B. Such antisense oligonucleotides may target one or more splicing regulatory elements in one or more exons or introns of ATP7B. These splicing regulatory elements modulate splicing of ATP7B ribonucleic acid (RNA).

[0034] In one aspect, the present disclosure provides an ATP7B RNA splice-modulating antisense oligonucleotide having a sequence targeted to one or more splicing regulatory elements adjacent to an

exon of ATP7B. In some embodiments, a genetic aberration of ATP7B includes the M645R mutation. In some embodiments, the M645R mutation results from ATP7B chr13:52535985:A:C [hg19/b37] (g.54646T>G in SEQ ID NO: 1). In some embodiments, the one or more splicing regulatory elements include an exonic splicing silencer element or an intronic splicing silencer element. In some
5 embodiments, the sequence is targeted to an abnormally spliced exon. In some embodiments, the sequence is targeted to an intron adjacent to an abnormally spliced exon. In some embodiments, the antisense oligonucleotide modulates variant splicing to yield an increase in exon inclusion. In some embodiments, the antisense oligonucleotide has a length of 12 to 20 nucleotides. In some embodiments, the antisense oligonucleotide has a length of 12 to 30 nucleotides. In some embodiments, the antisense
10 oligonucleotide has a length of 12 to 50 nucleotides.

[0035] In another aspect, the present disclosure provides a method for modulating splicing of ATP7B RNA in a cell, tissue, or organ of a subject, including bringing the cell, tissue, or organ in contact with an antisense oligonucleotide including one or more sequences targeted to one or more splicing regulatory elements of an abnormally spliced exon or an intron adjacent to the abnormally spliced exon. In some
15 embodiments, the genetic aberration of ATP7B includes M645R. In some embodiments, the M645R results from ATP7B chr13:52535985:A:C [hg19/b37] (g.54646T>G in SEQ ID NO: 1). In some embodiments, the splicing regulatory element is an exonic splicing silencer element or an intronic splicing silencer element. In some embodiments, the sequence is targeted to an abnormally spliced exon. In some embodiments, the sequence is targeted to an intron adjacent to an abnormally spliced exon. In
20 some embodiments, the antisense oligonucleotide modulates variant splicing to yield an increase in exon inclusion (e.g., exon 6 inclusion, e.g., increase by at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, or at least 50%; e.g., up to 100%, up to 90%, up to 80%, up to 70%, up to 60%, up to 50%, as compared to the ratio of exon 6-including ATP7B transcripts to the total number of ATP7B transcript molecules in a cell including ATP7B gene having an
25 exon 6-skipping mutation in the absence of a treatment with an antisense oligonucleotide). In some embodiments, the antisense oligonucleotide has a length of 12 to 20 nucleotides. In some embodiments, the antisense oligonucleotide has a length of 12 to 30 nucleotides. In some embodiments, the antisense oligonucleotide has a length of 12 to 50 nucleotides. In some embodiments, the subject has or is suspected of having a disease, e.g., Wilson disease, and the subject is monitored for a progression or
30 regression of the disease in response to bringing the cell, tissue, or organ in contact with the composition.

[0036] In another aspect, the present disclosure provides a method for treating Wilson disease in a subject, including administering to the subject a therapeutically effective amount of an antisense oligonucleotide including a sequence targeted to a splicing regulatory element of an abnormally spliced exon or an intron adjacent to the abnormally spliced exon. The antisense oligonucleotide modulates
35 splicing of ATP7B RNA. In some embodiments, the genetic aberration of ATP7B includes the M645R mutation. In some embodiments, the M645R mutation results from ATP7B chr13:52535985:A:C [hg19/b37] (g.54646T>G mutant of SEQ ID NO: 1). In some embodiments, the splicing regulatory element includes an exonic splicing silencer element or an intronic splicing silencer element. In some
40 embodiments, the sequence is targeted to an abnormally spliced exon of ATP7B. In some embodiments, the sequence is targeted to an intron adjacent to an abnormally spliced exon of the genetic aberration of ATP7B that modulates variant splicing of ATP7B RNA. In some embodiments, the antisense oligonucleotide modulates splicing to yield an increase in exon inclusion (e.g., increase by at least 5%, at

least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, or at least 50%; e.g., up to 100%, up to 90%, up to 80%, up to 70%, up to 60%, up to 50%, as compared to the ratio of exon 6-including ATP7B transcripts to the total number of ATP7B transcript molecules in a cell including ATP7B gene having an exon 6-skipping mutation in the absence of a treatment with an antisense oligonucleotide). In some embodiments, the antisense oligonucleotide has a length of 12 to 20 nucleotides. In some embodiments, the antisense oligonucleotide has a length of 12 to 30 nucleotides. In some embodiments, the antisense oligonucleotide has a length of 12 to 50 nucleotides. In some embodiments, the subject is monitored for a progression or regression of Wilson disease in response to administering to the subject the therapeutically effective amount of the antisense oligonucleotide. In some embodiments, the antisense oligonucleotide reduces 24-hour urinary copper level in the subject, e.g., by at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, or at least 50%; e.g., up to 90%, up to 80%, up to 70%, up to 60%, up to 50%, as compared to a control subject. In some embodiments, the control subject is the subject prior to therapy with an antisense oligonucleotide of the invention or is a subject suffering from Wilson disease and not receiving an antisense oligonucleotide of the invention. In some embodiments, the antisense oligonucleotide reduces 24-hour urinary copper level in the subject to <100 µg/24 hours (<1.6 µmol/24 hours) (e.g., to ≤40 µg/24 hours (≤0.6 µmol/24 hours)).

[0037] In another aspect, the present disclosure provides a pharmaceutical composition for treatment of Wilson disease including an antisense oligonucleotide and a pharmaceutically acceptable carrier. The antisense oligonucleotide includes a sequence targeted to a splicing regulatory element of an abnormally spliced exon or an intron adjacent to the abnormally spliced exon. The antisense oligonucleotide modulates splicing of ATP7B RNA. In some embodiments, the genetic aberration of ATP7B includes M645R. In some embodiments, the M645R mutation results from ATP7B chr13:52535985:A:C [hg19/b37] (g.54646T>G mutant of SEQ ID NO: 1).

Definitions

[0038] Various terms used throughout the present description may be read and understood as follows, unless the context indicates otherwise: “or” as used throughout is inclusive, as though written “and/or”; singular articles and pronouns as used throughout include their plural forms, and vice versa; similarly, gendered pronouns include their counterpart pronouns so that pronouns should not be understood as limiting anything described herein to use, implementation, performance, etc. by a single gender; “exemplary” should be understood as “illustrative” or “exemplifying” and not necessarily as “preferred” over other embodiments. Further definitions for terms may be set out herein; these may apply to prior and subsequent instances of those terms, as will be understood from a reading of the present description.

[0039] The term “acyl,” as used herein, represents a chemical substituent of formula $-C(O)-R$, where R is alkyl, aryl, arylalkyl, cycloalkyl, heterocyclyl, heterocyclyl alkyl, heteroaryl, or heteroaryl alkyl. An optionally substituted acyl is an acyl that is optionally substituted as described herein for each group R.

[0040] The term “acyloxy,” as used herein, represents a chemical substituent of formula $-OR$, where R is acyl. An optionally substituted acyloxy is an acyloxy that is optionally substituted as described herein for acyl.

[0041] The term "alkane-tetrayl," as used herein, represents a tetravalent, acyclic, straight or branched chain, saturated hydrocarbon group having from 1 to 16 carbons, unless otherwise specified. Alkane-tetrayl may be optionally substituted as described for alkyl.

[0042] The term "alkane-triyl," as used herein, represents a trivalent, acyclic, straight or branched chain, saturated hydrocarbon group having from 1 to 16 carbons, unless otherwise specified. Alkane-triyl may be optionally substituted as described for alkyl.

[0043] The term "alkanoyl," as used herein, represents a chemical substituent of formula $-C(O)-R$, where R is alkyl. An optionally substituted alkanoyl is an alkanoyl that is optionally substituted as described herein for alkyl.

[0044] The term "alkoxy," as used herein, represents a chemical substituent of formula $-OR$, where R is a C_{1-6} alkyl group, unless otherwise specified. An optionally substituted alkoxy is an alkoxy group that is optionally substituted as defined herein for alkyl.

[0045] The term "alkyl," as used herein, refers to an acyclic straight or branched chain saturated hydrocarbon group, which, when unsubstituted, has from 1 to 12 carbons, unless otherwise specified. In certain preferred embodiments, unsubstituted alkyl has from 1 to 6 carbons. Alkyl groups are exemplified by methyl; ethyl; n- and iso-propyl; n-, sec-, iso- and tert-butyl; neopentyl, and the like, and may be optionally substituted, valency permitting, with one, two, three, or, in the case of alkyl groups of two carbons or more, four or more substituents independently selected from the group consisting of: alkoxy; acyloxy; amino; aryl; aryloxy; azido; cycloalkyl; cycloalkoxy; halo; heterocyclyl; heteroaryl; heterocyclylalkyl; heteroarylalkyl; heterocyclioxy; heteroaryloxy; hydroxy; nitro; thiol; silyl; cyano; $=O$; $=S$; and $=NR'$, where R' is H, alkyl, aryl, or heterocyclyl. In some embodiments, a substituted alkyl includes two substituents (oxo and hydroxy, or oxo and alkoxy) to form a group $-L-CO-R$, where L is a bond or optionally substituted C_{1-11} alkylene, and R is hydroxyl or alkoxy. Each of the substituents may itself be unsubstituted or, valency permitting, substituted with unsubstituted substituent(s) defined herein for each respective group.

[0046] The term "alkylene," as used herein, represents a divalent substituent that is a monovalent alkyl having one hydrogen atom replaced with a valency. An optionally substituted alkylene is an alkylene that is optionally substituted as described herein for alkyl.

[0047] The term "aryl," as used herein, represents a mono-, bicyclic, or multicyclic carbocyclic ring system having one or two aromatic rings. Aryl group may include from 6 to 10 carbon atoms. All atoms within an unsubstituted carbocyclic aryl group are carbon atoms. Non-limiting examples of carbocyclic aryl groups include phenyl, naphthyl, 1,2-dihydronaphthyl, 1,2,3,4-tetrahydronaphthyl, fluorenyl, indanyl, indenyl, etc. The aryl group may be unsubstituted or substituted with one, two, three, four, or five substituents independently selected from the group consisting of: alkyl; alkoxy; acyloxy; amino; aryl; aryloxy; azido; cycloalkyl; cycloalkoxy; halo; heterocyclyl; heteroaryl; heterocyclylalkyl; heteroarylalkyl; heterocyclioxy; heteroaryloxy; hydroxy; nitro; thiol; silyl; and cyano. Each of the substituents may itself be unsubstituted or substituted with unsubstituted substituent(s) defined herein for each respective group.

[0048] The term "aryl alkyl," as used herein, represents an alkyl group substituted with an aryl group. The aryl and alkyl portions may be optionally substituted as the individual groups as described herein.

[0049] The term "arylene," as used herein, represents a divalent substituent that is an aryl having one hydrogen atom replaced with a valency. An optionally substituted arylene is an arylene that is optionally substituted as described herein for aryl.

[0050] The term “aryloxy,” as used herein, represents a group –OR, where R is aryl. Aryloxy may be an optionally substituted aryloxy. An optionally substituted aryloxy is aryloxy that is optionally substituted as described herein for aryl.

[0051] The term “ATP7B,” as used herein, represents a nucleic acid (e.g., genomic DNA, pre-mRNA, or mRNA) that is translated and, if genomic DNA, first transcribed, *in vivo* to ATPase copper transporting beta protein. An exemplary genomic DNA sequence comprising the human ATP7B gene is given by SEQ ID NO: 1 (NCBI Reference Sequence: NG_008806.1). SEQ ID NO: 1 provides the sequence for the antisense strand of the genomic DNA of ATP7B (positions 5001-83826 in SEQ ID NO: 1) and other nearby genes. One of skill in the art will recognize that an RNA sequence typically includes uridines instead of thymidines. The term “ATP7B,” as used herein, represents wild-type and mutant versions. An exemplary mutant nucleic acid (e.g., genomic DNA, pre-mRNA, or mRNA) results in ATPase copper transporting beta protein lacking exon 6.

[0052] The term “bicyclic sugar moiety,” as used herein, represents a modified sugar moiety including two fused rings. In certain embodiments, the bicyclic sugar moiety includes a furanosyl ring.

[0053] The expression “C_{x-y},” as used herein, indicates that the group, the name of which immediately follows the expression, when unsubstituted, contains a total of from x to y carbon atoms. If the group is a composite group (e.g., aryl alkyl), C_{x-y} indicates that the portion, the name of which immediately follows the expression, when unsubstituted, contains a total of from x to y carbon atoms. For example, (C₆₋₁₀-aryl)-C₁₋₆-alkyl is a group, in which the aryl portion, when unsubstituted, contains a total of from 6 to 10 carbon atoms, and the alkyl portion, when unsubstituted, contains a total of from 1 to 6 carbon atoms.

[0054] The term “complementary,” as used herein in reference to a nucleobase sequence, refers to the nucleobase sequence having a pattern of contiguous nucleobases that permits an oligonucleotide having the nucleobase sequence to hybridize to another oligonucleotide or nucleic acid to form a duplex structure under physiological conditions. Complementary sequences include Watson-Crick base pairs formed from natural and/or modified nucleobases. Complementary sequences can also include non-Watson-Crick base pairs, such as wobble base pairs (guanosine-uracil, hypoxanthine-uracil, hypoxanthine-adenine, and hypoxanthine-cytosine) and Hoogsteen base pairs.

[0055] The term “contiguous,” as used herein in the context of an oligonucleotide, refers to nucleosides, nucleobases, sugar moieties, or internucleoside linkages that are immediately adjacent to each other. For example, “contiguous nucleobases” means nucleobases that are immediately adjacent to each other in a sequence.

[0056] The term “cycloalkyl,” as used herein, refers to a cyclic alkyl group having from three to ten carbons (e.g., a C₃-C₁₀ cycloalkyl), unless otherwise specified. Cycloalkyl groups may be monocyclic or bicyclic. Bicyclic cycloalkyl groups may be of bicyclo[p.q.0]alkyl type, in which each of p and q is, independently, 1, 2, 3, 4, 5, 6, or 7, provided that the sum of p and q is 2, 3, 4, 5, 6, 7, or 8. Alternatively, bicyclic cycloalkyl groups may include bridged cycloalkyl structures, e.g., bicyclo[p.q.r]alkyl, in which r is 1, 2, or 3, each of p and q is, independently, 1, 2, 3, 4, 5, or 6, provided that the sum of p, q, and r is 3, 4, 5, 6, 7, or 8. The cycloalkyl group may be a spirocyclic group, e.g., spiro[p.q]alkyl, in which each of p and q is, independently, 2, 3, 4, 5, 6, or 7, provided that the sum of p and q is 4, 5, 6, 7, 8, or 9. Non-limiting examples of cycloalkyl include cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, 1-bicyclo[2.2.1]heptyl, 2-bicyclo[2.2.1]heptyl, 5-bicyclo[2.2.1]heptyl, 7-bicyclo[2.2.1]heptyl, and decalinyl.

The cycloalkyl group may be unsubstituted or substituted (e.g., optionally substituted cycloalkyl) with one, two, three, four, or five substituents independently selected from the group consisting of: alkyl; alkoxy; acyloxy; amino; aryl; aryloxy; azido; cycloalkyl; cycloalkoxy; halo; heterocyclyl; heteroaryl; heterocyclylalkyl; heteroarylalkyl; heterocyclyloxy; heteroaryloxy; hydroxy; nitro; thiol; silyl; cyano; =O; =S; =NR', where R' is H, alkyl, aryl, or heterocyclyl. Each of the substituents may itself be unsubstituted or substituted with unsubstituted substituent(s) defined herein for each respective group.

[0057] The term "cycloalkylene," as used herein, represents a divalent substituent that is a cycloalkyl having one hydrogen atom replaced with a valency. An optionally substituted cycloalkylene is a cycloalkylene that is optionally substituted as described herein for cycloalkyl.

[0058] The term "cycloalkoxy," as used herein, represents a group -OR, where R is cycloalkyl. Cycloalkoxy may be an optionally substituted cycloalkoxy. An optionally substituted cycloalkoxy is cycloalkoxy that is optionally substituted as described herein for cycloalkyl.

[0059] The term "duplex," as used herein, represents two oligonucleotides that are paired through hybridization of complementary nucleobases.

[0060] The term "exon 6," as used herein, refers to exon 6 of ATP7B pre-mRNA or genomic DNA, e.g., SEQ ID NO: 2, which corresponds to positions 54582 to 54658 in SEQ ID NO: 1 (hg19/b37 coordinates chr13:52535973-52536049), or a mutant version thereof (e.g., g.54646T>G in SEQ ID NO: 1).

[0061] The term "flanking intron," as used herein, refers to an intron that is adjacent to the 5'- or 3'-end of exon 6 or a mutant thereof. The flanking intron is a 5'-flanking intron or a 3'-flanking intron. The 5'-flanking intron corresponds to the flanking intron that is adjacent to the 5'-end of exon 6 (e.g., intron positions between exon 5 and exon 6 in SEQ ID NO: 1). The 3'-flanking intron corresponds to the flanking intron that is adjacent to the 3'-end of exon 6 (e.g., intron positions between exon 6 and exon 7 in SEQ ID NO: 1).

[0062] The term "genetic aberration," as used herein, generally refers to a mutation or variant in a gene. Examples of genetic aberration may include, but are not limited to, a point mutation (single nucleotide variant or single base substitution), an insertion or deletion (indel), a transversion, a translocation, an inversion, or a truncation. An aberrant ATP7B gene includes one or more mutations causing the splicing of pre-mRNA to skip exon 6.

[0063] The term "halo," as used herein, represents a halogen selected from bromine, chlorine, iodine, and fluorine.

[0064] The term "heteroalkane-tetrayl," as used herein refers to an alkane-tetrayl group interrupted once by one heteroatom; twice, each time, independently, by one heteroatom; three times, each time, independently, by one heteroatom; or four times, each time, independently, by one heteroatom. Each heteroatom is, independently, O, N, or S. In some embodiments, the heteroatom is O or N. An unsubstituted C_{x-y} heteroalkane-tetrayl contains from X to Y carbon atoms as well as the heteroatoms as defined herein. The heteroalkane-tetrayl group may be unsubstituted or substituted (e.g., optionally substituted heteroalkane-tetrayl), as described for heteroalkyl.

[0065] The term "heteroalkane-triyl," as used herein refers to an alkane-triyl group interrupted once by one heteroatom; twice, each time, independently, by one heteroatom; three times, each time, independently, by one heteroatom; or four times, each time, independently, by one heteroatom. Each heteroatom is, independently, O, N, or S. In some embodiments, the heteroatom is O or N. An

unsubstituted C_{X-Y} heteroalkane-triyl contains from X to Y carbon atoms as well as the heteroatoms as defined herein. The heteroalkane-triyl group may be unsubstituted or substituted (e.g., optionally substituted heteroalkane-triyl), as described for heteroalkyl.

[0066] The term “heteroalkyl,” as used herein, refers to an alkyl group interrupted one or more times

by one or two heteroatoms each time. Each heteroatom is independently O, N, or S. None of the heteroalkyl groups includes two contiguous oxygen atoms. The heteroalkyl group may be unsubstituted or substituted (e.g., optionally substituted heteroalkyl). When heteroalkyl is substituted and the substituent is bonded to the heteroatom, the substituent is selected according to the nature and valency of the heteroatom. Thus, the substituent bonded to the heteroatom, valency permitting, is selected from the group consisting of =O, -N(R^{N2})₂, -SO₂OR^{N3}, -SO₂R^{N2}, -SOR^{N3}, -COOR^{N3}, an N protecting group, alkyl, aryl, cycloalkyl, heterocyclyl, or cyano, where each R^{N2} is independently H, alkyl, cycloalkyl, aryl, or heterocyclyl, and each R^{N3} is independently alkyl, cycloalkyl, aryl, or heterocyclyl. Each of these substituents may itself be unsubstituted or substituted with unsubstituted substituent(s) defined herein for each respective group. When heteroalkyl is substituted and the substituent is bonded to carbon, the substituent is selected from those described for alkyl, provided that the substituent on the carbon atom bonded to the heteroatom is not Cl, Br, or I. In some embodiments, carbon atoms are found at the termini of a heteroalkyl group. In some embodiments, heteroalkyl is PEG.

[0067] The term “heteroalkylene,” as used herein, represents a divalent substituent that is a heteroalkyl having one hydrogen atom replaced with a valency. An optionally substituted heteroalkylene is a heteroalkylene that is optionally substituted as described herein for heteroalkyl.

[0068] The term “heteroaryl,” as used herein, represents a monocyclic 5-, 6-, 7-, or 8-membered ring system, or a fused or bridging bicyclic, tricyclic, or tetracyclic ring system; the ring system contains one, two, three, or four heteroatoms independently selected from the group consisting of nitrogen, oxygen, and sulfur; and at least one of the rings is an aromatic ring. Non-limiting examples of heteroaryl groups include benzimidazolyl, benzofuryl, benzothiazolyl, benzothieryl, benzoxazolyl, furyl, imidazolyl, indolyl, isoindazolyl, isoquinolinyl, isothiazolyl, isothiazolyl, isoxazolyl, oxadiazolyl, oxazolyl, purinyl, pyrrolyl, pyridinyl, pyrazinyl, pyrimidinyl, qunazolinyl, quinolinyl, thiadiazolyl (e.g., 1,3,4-thiadiazole), thiazolyl, thienyl, triazolyl, tetrazolyl, dihydroindolyl, tetrahydroquinolyl, tetrahydroisoquinolyl, etc. The term bicyclic, tricyclic, and tetracyclic heteroaryls include at least one ring having at least one heteroatom as described above and at least one aromatic ring. For example, a ring having at least one heteroatom may be fused to one, two, or three carbocyclic rings, e.g., an aryl ring, a cyclohexane ring, a cyclohexene ring, a cyclopentane ring, a cyclopentene ring, or another monocyclic heterocyclic ring. Examples of fused heteroaryls include 1,2,3,5,8,8a-hexahydroindolizine; 2,3-dihydrobenzofuran; 2,3-dihydroindole; and 2,3-dihydrobenzothiophene. Heteroaryl may be optionally substituted with one, two, three, four, or five substituents independently selected from the group consisting of: alkyl; alkoxy; acyloxy; aryloxy; amino; arylalkoxy; cycloalkyl; cycloalkoxy; halogen; heterocyclyl; heterocyclyl alkyl; heteroaryl; heteroaryl alkyl; heterocycloxy; heteroaryloxy; hydroxyl; nitro; thiol; cyano; =O; -NR₂, where each R is independently hydrogen, alkyl, acyl, aryl, arylalkyl, cycloalkyl, heterocyclyl, or heteroaryl; -COOR^A, where R^A is hydrogen, alkyl, aryl, arylalkyl, cycloalkyl, heterocyclyl, or heteroaryl; and -CON(R^B)₂, where each R^B is independently hydrogen, alkyl, aryl, arylalkyl, cycloalkyl, heterocyclyl, or heteroaryl. Each of the substituents may itself be unsubstituted or substituted with unsubstituted substituent(s) defined herein for each respective group.

[0069] The term "heteroarylene," as used herein, represents a divalent substituent that is a heteroaryl having one hydrogen atom replaced with a valency. An optionally substituted heteroarylene is a heteroarylene that is optionally substituted as described herein for heteroaryl.

[0070] The term "heteroaryloxy," as used herein, refers to a structure $-OR$, in which R is heteroaryl.

5 Heteroaryloxy can be optionally substituted as defined for heteroaryl.

[0071] The term "heterocyclyl," as used herein, represents a monocyclic, bicyclic, tricyclic, or tetracyclic ring system having fused or bridging 4-, 5-, 6-, 7-, or 8-membered rings, unless otherwise specified, the ring system containing one, two, three, or four heteroatoms independently selected from the group consisting of nitrogen, oxygen, and sulfur. Heterocyclyl may be aromatic or non-aromatic. An aromatic heterocyclyl is heteroaryl as described herein. Non-aromatic 5-membered heterocyclyl has zero or one double bonds, non-aromatic 6- and 7-membered heterocyclyl groups have zero to two double bonds, and non-aromatic 8-membered heterocyclyl groups have zero to two double bonds and/or zero or one carbon-carbon triple bond. Heterocyclyl groups have a carbon count of 1 to 16 carbon atoms unless otherwise specified. Certain heterocyclyl groups may have a carbon count up to 9 carbon atoms. Non-aromatic heterocyclyl groups include pyrrolinyl, pyrrolidinyl, pyrazolinyl, pyrazolidinyl, imidazolinyl, imidazolidinyl, piperidinyl, homopiperidinyl, piperazinyl, pyridazinyl, oxazolidinyl, isoxazolidinyl, morpholinyl, thiomorpholinyl, thiazolidinyl, isothiazolidinyl, thiazolidinyl, tetrahydrofuranyl, dihydrofuranyl, tetrahydrothienyl, dihydrothienyl, pyranyl, dihydropyranyl, dithiazolyl, etc. The term "heterocyclyl" also represents a heterocyclic compound having a bridged multicyclic structure in which one or more carbons and/or heteroatoms bridges two non-adjacent members of a monocyclic ring, e.g., quinuclidine, tropanes, or diaza-bicyclo[2.2.2]octane. The term "heterocyclyl" includes bicyclic, tricyclic, and tetracyclic groups in which any of the above heterocyclic rings is fused to one, two, or three carbocyclic rings, e.g., a cyclohexane ring, a cyclohexene ring, a cyclopentane ring, a cyclopentene ring, or another heterocyclic ring. Examples of fused heterocyclyls include 1,2,3,5,8a-hexahydroindolizine; 2,3-dihydrobenzofuran; 2,3-dihydroindole; and 2,3-dihydrobenzothiophene. The heterocyclyl group may be unsubstituted or substituted with one, two, three, four or five substituents independently selected from the group consisting of: alkyl; alkoxy; acyloxy; aryloxy; amino; arylalkoxy; cycloalkyl; cycloalkoxy; halogen; heterocyclyl; heterocyclyl alkyl; heteroaryl; heteroaryl alkyl; heterocyclyloxy; heteroaryloxy; hydroxyl; nitro; thiol; cyano; $=O$; $=S$; $-NR_2$, where each R is independently hydrogen, alkyl, acyl, aryl, arylalkyl, cycloalkyl, heterocyclyl, or heteroaryl; $-COOR^A$, where R^A is hydrogen, alkyl, aryl, arylalkyl, cycloalkyl, heterocyclyl, or heteroaryl; and $-CON(R^B)_2$, where each R^B is independently hydrogen, alkyl, aryl, arylalkyl, cycloalkyl, heterocyclyl, or heteroaryl.

[0072] The term "heterocyclyl alkyl," as used herein, represents an alkyl group substituted with a heterocyclyl group. The heterocyclyl and alkyl portions of an optionally substituted heterocyclyl alkyl are optionally substituted as described for heterocyclyl and alkyl, respectively.

[0073] The term "heterocyclylene," as used herein, represents a divalent substituent that is a heterocyclyl having one hydrogen atom replaced with a valency. An optionally substituted heterocyclylene is a heterocyclylene that is optionally substituted as described herein for heterocyclyl.

[0074] The term "heterocyclyloxy," as used herein, refers to a structure $-OR$, in which R is heterocyclyl. Heterocyclyloxy can be optionally substituted as described for heterocyclyl.

[0075] The term "heteroorganic," as used herein, refers to (i) an acyclic hydrocarbon interrupted one or more times by one or two heteroatoms each time, or (ii) a cyclic hydrocarbon including one or more

(e.g., one, two, three, or four) endocyclic heteroatoms. Each heteroatom is independently O, N, or S. None of the heteroorganic groups includes two contiguous oxygen atoms. An optionally substituted heteroorganic group is a heteroorganic group that is optionally substituted as described herein for alkyl.

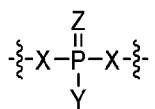
[0076] The term “hydrocarbon,” as used herein, refers to an acyclic, branched or acyclic, linear compound or group, or a monocyclic, bicyclic, tricyclic, or tetracyclic compound or group. The hydrocarbon, when unsubstituted, consists of carbon and hydrogen atoms. Unless specified otherwise, an unsubstituted hydrocarbon includes a total of 1 to 60 carbon atoms (e.g., 1 to 16, 1 to 12, or 1 to 6 carbon atoms). An optionally substituted hydrocarbon is an optionally substituted acyclic hydrocarbon or an optionally substituted cyclic hydrocarbon. An optionally substituted acyclic hydrocarbon is optionally substituted as described herein for alkyl. An optionally substituted cyclic hydrocarbon is an optionally substituted aromatic hydrocarbon or an optionally substituted non-aromatic hydrocarbon. An optionally substituted aromatic hydrocarbon is optionally substituted as described herein for aryl. An optionally substituted non-aromatic cyclic hydrocarbon is optionally substituted as described herein for cycloalkyl. In some embodiments, an acyclic hydrocarbon is alkyl, alkylene, alkane-triyl, or alkane-tetrayl. In certain embodiments, a cyclic hydrocarbon is aryl or arylene. In particular embodiments, a cyclic hydrocarbon is cycloalkyl or cycloalkylene.

[0077] The terms “hydroxyl” and “hydroxy,” as used interchangeably herein, represent -OH.

[0078] The term “hydrophobic moiety,” as used herein, represents a monovalent group covalently linked to an oligonucleotide backbone, where the monovalent group is a bile acid (e.g., cholic acid, taurocholic acid, deoxycholic acid, oleyl lithocholic acid, or oleoyl cholenic acid), glycolipid, phospholipid, sphingolipid, isoprenoid, vitamin, saturated fatty acid, unsaturated fatty acid, fatty acid ester, triglyceride, pyrene, porphyrine, texaphyrine, adamantane, acridine, biotin, coumarin, fluorescein, rhodamine, Texas-Red, digoxigenin, dimethoxytrityl, *t*-butyldimethylsilyl, *t*-butyldiphenylsilyl, cyanine dye (e.g., Cy3 or Cy5), Hoechst 33258 dye, psoralen, or ibuprofen. Non-limiting examples of the monovalent group include ergosterol, stigmasterol, β -sitosterol, campesterol, fucosterol, saringosterol, avenasterol, coprostanol, cholesterol, vitamin A, vitamin D, vitamin E, cardiolipin, and carotenoids. The linker connecting the monovalent group to the oligonucleotide may be an optionally substituted C₁₋₆₀ hydrocarbon (e.g., optionally substituted C₁₋₆₀ alkylene) or an optionally substituted C₂₋₆₀ heteroorganic (e.g., optionally substituted C₂₋₆₀ heteroalkylene), where the linker may be optionally interrupted with one, two, or three instances independently selected from the group consisting of an optionally substituted arylene, optionally substituted heterocyclylene, and optionally substituted cycloalkylene. The linker may be bonded to an oligonucleotide through, e.g., an oxygen atom attached to a 5'-terminal carbon atom, a 3'-terminal carbon atom, a 5'-terminal phosphate or phosphorothioate, a 3'-terminal phosphate or phosphorothioate, or an internucleoside linkage.

[0079] The term “internucleoside linkage,” as used herein, represents a divalent group or covalent bond that forms a covalent linkage between adjacent nucleosides in an oligonucleotide. An internucleoside linkage is an unmodified internucleoside linkage or a modified internucleoside linkage. An “unmodified internucleoside linkage” is a phosphate (-O-P(O)(OH)-O-) internucleoside linkage (“phosphate phosphodiester”). A “modified internucleoside linkage” is an internucleoside linkage other than a phosphate phosphodiester. The two main classes of modified internucleoside linkages are defined by the presence or absence of a phosphorus atom. Non-limiting examples of phosphorus-containing internucleoside linkages include phosphodiester linkages, phosphotriester linkages, phosphorothioate

diester linkages, phosphorothioate triester linkages, phosphorodithioate linkages, boranophosphonate linkages, morpholino internucleoside linkages, methylphosphonates, and phosphoramidate. Non-limiting examples of non-phosphorus internucleoside linkages include methylenemethylimino ($-\text{CH}_2-\text{N}(\text{CH}_3)-\text{O}-\text{CH}_2-$), thiodiester ($-\text{O}-\text{C}(\text{O})-\text{S}-$), thionocarbamate ($-\text{O}-\text{C}(\text{O})(\text{NH})-\text{S}-$), siloxane ($-\text{O}-\text{Si}(\text{H})_2-\text{O}-$), and N,N'-dimethylhydrazine ($-\text{CH}_2-\text{N}(\text{CH}_3)-\text{N}(\text{CH}_3)-$). Phosphorothioate linkages are phosphodiester linkages and phosphotriester linkages in which one of the non-bridging oxygen atoms is replaced with a sulfur atom. In some embodiments, an internucleoside linkage is a group of the following structure:



10 where

Z is O, S, B, or Se;

Y is $-\text{X}-\text{L}-\text{R}^1$;

each X is independently $-\text{O}-$, $-\text{S}-$, $-\text{N}(\text{L}-\text{R}^1)-$, or L;

each L is independently a covalent bond or a linker (e.g., optionally substituted C_{1-60} hydrocarbon

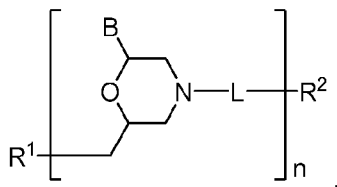
15 linker or optionally substituted C_{2-60} heteroorganic linker);

each R^1 is independently hydrogen, $-\text{S}-\text{S}-\text{R}^2$, $-\text{O}-\text{CO}-\text{R}^2$, $-\text{S}-\text{CO}-\text{R}^2$, optionally substituted C_{1-9} heterocyclyl, a hydrophobic moiety, or a targeting moiety; and

each R^2 is independently optionally substituted C_{1-10} alkyl, optionally substituted C_{2-10} heteroalkyl, optionally substituted C_{6-10} aryl, optionally substituted C_{6-10} aryl C_{1-6} alkyl, optionally substituted C_{1-9} heterocyclyl, or optionally substituted C_{1-9} heterocyclyl C_{1-6} alkyl.

When L is a covalent bond, R^1 is hydrogen, Z is oxygen, and all X groups are $-\text{O}-$, the internucleoside group is known as a phosphate phosphodiester. When L is a covalent bond, R^1 is hydrogen, Z is sulfur, and all X groups are $-\text{O}-$, the internucleoside group is known as a phosphorothioate diester. When Z is oxygen, all X groups are $-\text{O}-$, and either (1) L is a linker or (2) R^1 is not a hydrogen, the internucleoside group is known as a phosphotriester. When Z is sulfur, all X groups are $-\text{O}-$, and either (1) L is a linker or (2) R^1 is not a hydrogen, the internucleoside group is known as a phosphorothioate triester. Non-limiting examples of phosphorothioate triester linkages and phosphotriester linkages are described in US 2017/0037399, the disclosure of which is incorporated herein by reference.

[0080] The term "morpholino," as used herein in reference to a class of oligonucleotides, represents an oligomer of at least 10 morpholino monomer units interconnected by morpholino internucleoside linkages. A morpholino includes a 5' group and a 3' group. For example, a morpholino may be of the following structure:



where

35 n is an integer of at least 10 (e.g., 12 to 50) indicating the number of morpholino units;

each B is independently a nucleobase;

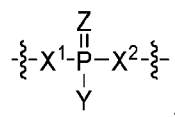
R¹ is a 5' group;

R² is a 3' group; and

L is (i) a morpholino internucleoside linkage or, (ii) if L is attached to R², a covalent bond.

A 5' group in morpholino may be, e.g., hydroxyl, a hydrophobic moiety, phosphate, diphosphate, triphosphate, phosphorothioate, diphosphorothioate, triphosphorothioate, phosphorodithioate, disphorodithioate, triphosphorodithioate, phosphonate, phosphoramidate, a cell penetrating peptide, an endosomal escape moiety, or a neutral organic polymer. A 3' group in morpholino may be, e.g., hydrogen, a hydrophobic moiety, phosphate, diphosphate, triphosphate, phosphorothioate, diphosphorothioate, triphosphorothioate, phosphorodithioate, disphorodithioate, triphosphorodithioate, phosphonate, phosphoramidate, a cell penetrating peptide, an endosomal escape moiety, or a neutral organic polymer.

[0081] The term "morpholino internucleoside linkage," as used herein, represents a divalent group of the following structure:



where

Z is O or S;

X¹ is a bond, $\text{--CH}_2\text{--}$, or --O-- ;

X² is a bond, $\text{--CH}_2\text{--O--}$, or --O-- ; and

Y is --NR_2 , where each R is independently C₁₋₆ alkyl (e.g., methyl), or both R combine together with the nitrogen atom to which they are attached to form a C₂₋₉ heterocyclyl (e.g., N-piperazinyl); provided that both X¹ and X² are not simultaneously a bond.

[0082] The term "nucleobase," as used herein, represents a nitrogen-containing heterocyclic ring found at the 1' position of the ribofuranose/2'-deoxyribofuranose of a nucleoside. Nucleobases are unmodified or modified. 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 5-substituted pyrimidines, 6-azapyrimidines, alkyl or alkynyl substituted pyrimidines, alkyl substituted purines, and N-2, N-6 and O-6 substituted purines, as well as synthetic and natural nucleobases, e.g., 5-methylcytosine, 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-alkyl (e.g., 6-methyl) adenine and guanine, 2-alkyl (e.g., 2-propyl) adenine and guanine, 2-thiouracil, 2-thiothymine, 2-thiocytosine, 5-halouracil, 5-halocytosine, 5-propynyl uracil, 5-propynyl cytosine, 5-trifluoromethyl uracil, 5-trifluoromethyl cytosine, 7-methyl guanine, 7-methyl adenine, 8-azaguanine, 8-azaadenine, 7-deazaguanine, 7-deazaadenine, 3-deazaguanine, 3-deazaadenine. Certain nucleobases are particularly useful for increasing the binding affinity of nucleic acids, e.g., 5-substituted pyrimidines; 6-azapyrimidines; N2-, N6-, and/or O6-substituted purines. Nucleic acid duplex stability can be enhanced using, e.g., 5-methylcytosine. Non-limiting examples of nucleobases include: 2-aminopropyladenine, 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-N-methylguanine, 6-N-methyladenine, 2-propyladenine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-propynyl ($\text{--C}\equiv\text{C--CH}_3$) uracil, 5-propynylcytosine, 6-azouracil, 6-azocytosine, 6-azothymine, 5-ribosyluracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl, 8-aza and other 8-substituted purines, 5-halo, particularly 5-bromo, 5-trifluoromethyl, 5-halouracil, and 5-halocytosine, 7-methylguanine, 7-

methyladenine, 2-F-adenine, 2-aminoadenine, 7-deazaguanine, 7-deazaadenine, 3-deazaguanine, 3-deazaadenine, 6-N-benzoyladenine, 2-N-isobutyrylguanine, 4-N-benzoylcytosine, 4-N-benzoyluracil, 5-methyl 4-N-benzoylcytosine, 5-methyl 4-N-benzoyluracil, universal bases, hydrophobic bases, promiscuous bases, size-expanded bases, and fluorinated bases. Further modified nucleobases include tricyclic pyrimidines, such as 1,3-diazaphenoxazine-2-one, 1,3-diazaphenothiazine-2-one and 9-(2-aminoethoxy)-1,3-diazaphenoxazine-2-one (G-clamp). Modified nucleobases may also include those in which the purine or pyrimidine base is replaced with other heterocycles, for example, 7-deazaadenine, 7-deazaguanine, 2-aminopyridine, or 2-pyridone. Further nucleobases include those disclosed in U.S. Pat. No. 3,687,808; The Concise Encyclopedia of Polymer Science and Engineering, Kroschwitz, J. I., Ed., John Wiley & Sons, 1990, 858-859; Englisch et al., *Angewandte Chemie*, International Edition, 1991, 30, 613; Sanghvi, Y. S., Chapter 15, *Antisense Research and Applications*, Crooke, S. T. and Lebleu, B., Eds., CRC Press, 1993, 273-288; and in Chapters 6 and 15, *Antisense Drug Technology*, Crooke S. T., Ed., CRC Press, 2008, 163-166 and 442-443.

[0083] The term "nucleoside," as used herein, represents sugar-nucleobase compounds and groups known in the art (e.g., modified or unmodified ribofuranose-nucleobase and 2'-deoxyribofuranose-nucleobase compounds and groups known in the art). The sugar may be ribofuranose. The sugar may be modified or unmodified. An unmodified sugar nucleoside is ribofuranose or 2'-deoxyribofuranose having an anomeric carbon bonded to a nucleobase. An unmodified nucleoside is ribofuranose or 2'-deoxyribofuranose having an anomeric carbon bonded to an unmodified nucleobase. Non-limiting examples of unmodified nucleosides include adenosine, cytidine, guanosine, uridine, 2'-deoxyadenosine, 2'-deoxycytidine, 2'-deoxyguanosine, and thymidine. The modified compounds and groups include one or more modifications selected from the group consisting of nucleobase modifications and sugar modifications described herein. A nucleobase modification is a replacement of an unmodified nucleobase with a modified nucleobase. A sugar modification may be, e.g., a 2'-substitution, locking, carbocyclization, or unlocking. A 2'-substitution is a replacement of 2'-hydroxyl in ribofuranose with 2'-fluoro, 2'-methoxy, or 2'-(2-methoxy)ethoxy. A locking modification is an incorporation of a bridge between 4'-carbon atom and 2'-carbon atom of ribofuranose. Nucleosides having a locking modification are known in the art as bridged nucleic acids, e.g., locked nucleic acids (LNA), ethylene-bridged nucleic acids (ENA), and cEt nucleic acids. The bridged nucleic acids are typically used as affinity enhancing nucleosides.

[0084] The term "nucleotide," as used herein, represents a nucleoside bonded to an internucleoside linkage or a monovalent group of the following structure $-X^1-P(X^2)(R^1)_2$, where X^1 is O, S, or NH, and X^2 is absent, =O, or =S, and each R^1 is independently -OH, $-N(R^2)_2$, or $-O-CH_2CH_2CN$, where each R^2 is independently an optionally substituted alkyl, or both R^2 groups, together with the nitrogen atom to which they are attached, combine to form an optionally substituted heterocycl.

[0085] The term "oligonucleotide," as used herein, represents a structure containing 10 or more (e.g., 10 to 50) contiguous nucleosides covalently bound together by internucleoside linkages. An oligonucleotide includes a 5' end and a 3' end. The 5' end of an oligonucleotide may be, e.g., hydroxyl, a targeting moiety, a hydrophobic moiety, 5' cap, phosphate, diphosphate, triphosphate, phosphorothioate, diphosphorothioate, triphosphorothioate, phosphorodithioate, diphosphorodithioate, triphosphorodithioate, phosphonate, phosphoramidate, a cell penetrating peptide, an endosomal escape moiety, or a neutral organic polymer. The 3' end of an oligonucleotide may be, e.g., hydroxyl, a targeting moiety, a

hydrophobic moiety, phosphate, diphosphate, triphosphate, phosphorothioate, diphosphorothioate, triphosphorothioate, phosphorodithioate, disphorodithioate, triphosphorodithioate, phosphonate, phosphoramidate, a cell penetrating peptide, an endosomal escape moiety, or a neutral organic polymer (e.g., polyethylene glycol). An oligonucleotide having a 5'-hydroxyl or 5'-phosphate has an unmodified 5' terminus. An oligonucleotide having a 5' terminus other than 5'-hydroxyl or 5'-phosphate has a modified 5' terminus. An oligonucleotide having a 3'-hydroxyl or 3'-phosphate has an unmodified 3' terminus. An oligonucleotide having a 3' terminus other than 3'-hydroxyl or 3'-phosphate has a modified 3' terminus.

[0086] The term "oxo," as used herein, represents a divalent oxygen atom (e.g., the structure of oxo may be shown as =O).

[0087] The term "pharmaceutically acceptable," as used herein, refers to those compounds, materials, compositions, and/or dosage forms, which are suitable for contact with the tissues of an individual (e.g., a human), without excessive toxicity, irritation, allergic response and other problem complications commensurate with a reasonable benefit/risk ratio.

[0088] The term "protecting group," as used herein, represents a group intended to protect a functional group (e.g., a hydroxyl, an amino, or a carbonyl) from participating in one or more undesirable reactions during chemical synthesis. The term "*O*-protecting group," as used herein, represents a group intended to protect an oxygen containing (e.g., phenol, hydroxyl or carbonyl) group from participating in one or more undesirable reactions during chemical synthesis. The term "*N*-protecting group," as used herein, represents a group intended to protect a nitrogen containing (e.g., an amino or hydrazine) group from participating in one or more undesirable reactions during chemical synthesis. Commonly used *O*- and *N*-protecting groups are disclosed in Wuts, "Greene's Protective Groups in Organic Synthesis," 4th Edition (John Wiley & Sons, New York, 2006), which is incorporated herein by reference. Exemplary *O*- and *N*-protecting groups include alkanoyl, aryloyl, or carbamyl groups such as formyl, acetyl, propionyl, pivaloyl, *t*-butylacetyl, 2-chloroacetyl, 2-bromoacetyl, trifluoroacetyl, trichloroacetyl, phthalyl, *o*-nitrophenoxycetyl, α -chlorobutyryl, benzoyl, 4-chlorobenzoyl, 4-bromobenzoyl, *t*-butyldimethylsilyl, tri-*iso*-propylsilyloxymethyl, 4,4'-dimethoxytrityl, isobutyryl, phenoxyacetyl, 4-isopropylphenoxyacetyl, dimethylformamidino, and 4-nitrobenzoyl.

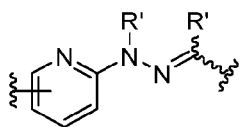
[0089] Exemplary *O*-protecting groups for protecting carbonyl containing groups include, but are not limited to: acetals, acylals, 1,3-dithianes, 1,3-dioxanes, 1,3-dioxolanes, and 1,3-dithiolanes.

[0090] Other *O*-protecting groups include, but are not limited to: substituted alkyl, aryl, and arylalkyl ethers (e.g., trityl; methylthiomethyl; methoxymethyl; benzyloxymethyl; siloxymethyl; 2,2,2-trichloroethoxymethyl; tetrahydropyranyl; tetrahydrofuranyl; ethoxyethyl; 1-[2-(trimethylsilyl)ethoxy]ethyl; 2-trimethylsilylethyl; *t*-butyl ether; *p*-chlorophenyl, *p*-methoxyphenyl, *p*-nitrophenyl, benzyl, *p*-methoxybenzyl, and nitrobenzyl); silyl ethers (e.g., trimethylsilyl; triethylsilyl; triisopropylsilyl; dimethylisopropylsilyl; *t*-butyldimethylsilyl; *t*-butyldiphenylsilyl; tribenzylsilyl; triphenylsilyl; and diphenylmethylsilyl); carbonates (e.g., methyl, methoxymethyl, 9-fluorenylmethyl; ethyl; 2,2,2-trichloroethyl; 2-(trimethylsilyl)ethyl; vinyl, allyl, nitrophenyl; benzyl; methoxybenzyl; 3,4-dimethoxybenzyl; and nitrobenzyl).

[0091] Other *N*-protecting groups include, but are not limited to, chiral auxiliaries such as protected or unprotected D, L or D, L-amino acids such as alanine, leucine, phenylalanine, and the like; sulfonyl-containing groups such as benzenesulfonyl, *p*-toluenesulfonyl, and the like; carbamate forming groups such as benzyloxycarbonyl, *p*-chlorobenzyloxycarbonyl, *p*-methoxybenzyloxycarbonyl, *p*-

nitrobenzyloxycarbonyl, 2-nitrobenzyloxycarbonyl, p-bromobenzyloxycarbonyl, 3,4-dimethoxybenzyloxycarbonyl, 3,5-dimethoxybenzyloxycarbonyl, 2,4-dimethoxybenzyloxycarbonyl, 4-methoxybenzyloxycarbonyl, 2-nitro-4,5-dimethoxybenzyloxycarbonyl, 3,4,5-trimethoxybenzyloxycarbonyl, 1-(p-biphenyl)-1-methylethoxycarbonyl, α,α -dimethyl-3,5-dimethoxybenzyloxycarbonyl, benzhydroxy carbonyl, t-butyloxycarbonyl, diisopropylmethoxycarbonyl, isopropoxycarbonyl, ethoxycarbonyl, methoxycarbonyl, allyloxycarbonyl, 2,2,2-trichloroethoxycarbonyl, phenoxycarbonyl, 4-nitrophenoxy carbonyl, fluorenyl-9-methoxycarbonyl, cyclopentyloxycarbonyl, adamantyloxycarbonyl, cyclohexyloxycarbonyl, phenylthiocarbonyl, and the like, arylalkyl groups such as benzyl, triphenylmethyl, benzyloxymethyl, and the like and silyl groups such as trimethylsilyl, and the like.

10 **[0092]** The term "pyrid-2-yl hydrazone," as used herein, represents a group of the structure:



, where each R' is independently H or optionally substituted C₁₋₆ alkyl. Pyrid-2-yl hydrazone may be unsubstituted (i.e., each R' is H).

[0093] The term "splice site," as used herein, generally refers to a site in a genome corresponding to an end of an intron that may be involved in a splicing procedure. A splice site may be a 5' splice site (e.g., a 5' end of an intron) or a 3' splice site (e.g., a 3' end of an intron). A given 5' splice site may be associated with one or more candidate 3' splice sites, each of which may be coupled to its corresponding 5' splice site in a splicing operation.

[0094] The term "splicing enhancer," as used herein, refers to motifs with positive effects (e.g., causing an increase) on exon inclusion.

20 **[0095]** The term "splicing silencer," as used herein, refers to motifs with negative effects (e.g., causing a decrease) on exon inclusion.

[0096] The term "stereochemically enriched," as used herein, refers to a local stereochemical preference for one enantiomer of the recited group over the opposite enantiomer of the same group. Thus, an oligonucleotide containing a stereochemically enriched internucleoside linkage is an oligonucleotide, in which a stereogenic internucleoside linkage (e.g., phosphorothioate) of predetermined stereochemistry is present in preference to a stereogenic internucleoside linkage (e.g., phosphorothioate) of stereochemistry that is opposite of the predetermined stereochemistry. This preference can be expressed numerically using a diastereomeric ratio for the stereogenic internucleoside linkage (e.g., phosphorothioate) of the predetermined stereochemistry. The diastereomeric ratio for the stereogenic internucleoside linkage (e.g., phosphorothioate) of the predetermined stereochemistry is the molar ratio of the diastereomers having the identified stereogenic internucleoside linkage (e.g., phosphorothioate) with the predetermined stereochemistry relative to the diastereomers having the identified stereogenic internucleoside linkage (e.g., phosphorothioate) with the stereochemistry that is opposite of the predetermined stereochemistry. The diastereomeric ratio for the phosphorothioate of the predetermined stereochemistry may be greater than or equal to 1.1 (e.g., greater than or equal to 4, greater than or equal to 9, greater than or equal to 19, or greater than or equal to 39).

[0097] The term "subject," as used herein, represents a human or non-human animal (e.g., a mammal) that is suffering from, or is at risk of, disease, disorder, or condition, as determined by a qualified professional (e.g., a doctor or a nurse practitioner) with or without known in the art laboratory

test(s) of sample(s) from the subject. A non-limiting example of a disease, disorder, or condition includes Wilson disease (e.g., Wilson disease associated with exon 6 skipping).

[0098] A "sugar" or "sugar moiety," includes naturally occurring sugars having a furanose ring or a structure that is capable of replacing the furanose ring of a nucleoside. Sugars included in the nucleosides of the invention may be non-furanose (or 4'-substituted furanose) rings or ring systems or open systems. Such structures include simple changes relative to the natural furanose ring (e.g., a six-membered ring). Alternative sugars may also include sugar surrogates wherein the furanose ring has been replaced with another ring system such as, e.g., a morpholino or hexitol ring system. Non-limiting examples of sugar moieties useful that may be included in the oligonucleotides of the invention include β -D-ribose, β -D-2'-deoxyribose, substituted sugars (e.g., 2', 5', and bis substituted sugars), 4'-S-sugars (e.g., 4'-S-ribose, 4'-S-2'-deoxyribose, and 4'-S-2'-substituted ribose), bicyclic sugar moieties (e.g., the 2'-O—CH₂-4' or 2'-O—(CH₂)₂-4' bridged ribose derived bicyclic sugars) and sugar surrogates (when the ribose ring has been replaced with a morpholino or a hexitol ring system).

[0099] The term "targeting moiety," as used herein, represents a moiety (e.g., N-acetylgalactosamine or a cluster thereof) that specifically binds or reactively associates or complexes with a receptor or other receptive moiety associated with a given target cell population. An antisense oligonucleotide may contain a targeting moiety. An antisense oligonucleotide including a targeting moiety is also referred to herein as a conjugate. A targeting moiety may include one or more ligands (e.g., 1 to 6 ligands, 1 to 3 ligands, or 1 ligand). The ligand can be an antibody or an antigen-binding fragment or an engineered derivative thereof (e.g., Fcab or a fusion protein (e.g., scFv)). Alternatively, the ligand may be a small molecule (e.g., N-acetylgalactosamine).

[00100] The term "therapeutically effective amount," as used herein, represents the quantity of an antisense oligonucleotide of the invention necessary to ameliorate, treat, or at least partially arrest the symptoms of a disease or disorder (e.g., to increase the level of ATP7B mRNA molecules including exon 6). Amounts effective for this use may depend, e.g., on the severity of the disease and the weight and general state of the subject. Typically, dosages used *in vitro* may provide useful guidance in the amounts useful for *in vivo* administration of the pharmaceutical composition, and animal models may be used to determine effective dosages for treatment of particular disorders. In some embodiments, a therapeutically effective amount of an antisense oligonucleotide of the invention reduces 24-hour urinary copper level in the subject to <100 μ g/24 hours (<1.6 μ mol/24 hours) (e.g., to $\leq$ 40 μ g/24 hours ($\leq$ 0.6 μ mol/24 hours)).

[00101] The term "thiocarbonyl," as used herein, represents a C(=S) group. - Non-limiting example of functional groups containing a "thiocarbonyl" includes thioesters, thioketones, thioaldehydes, thioanhydrides, thioacyl chlorides, thioamides, thiocarboxylic acids, and thiocarboxylates.

[00102] The term "thioheterocyclylene," as used herein, represents a divalent group —S—R', where R' is a heterocyclylene as defined herein.

[00103] The term "thiol," as used herein, represents an —SH group.

[00104] The term "triazolocycloalkenylene," as used herein, refers to the heterocyclylenes containing a 1,2,3-triazole ring fused to an 8-membered ring, all of the endocyclic atoms of which are carbon atoms, and bridgehead atoms are sp²-hybridized carbon atoms. Triazolocycloalkenylenes can be optionally substituted in a manner described for heterocyclyl.

[00105] The term “triazoloheterocyclylene,” as used herein, refers to the heterocyclylenes containing a 1,2,3-triazole ring fused to an 8-membered ring containing at least one heteroatom. The bridgehead atoms in triazoloheterocyclylene are carbon atoms. Triazoloheterocyclylenes can be optionally substituted in a manner described for heterocyclyl.

5 **[00106]** Enumeration of positions within oligonucleotides and nucleic acids, as used herein and unless specified otherwise, starts with the 5'-terminal nucleoside as 1 and proceeds in the 3'-direction.

[00107] The compounds described herein, unless otherwise noted, encompass isotopically enriched compounds (e.g., deuterated compounds), tautomers, and all stereoisomers and conformers (e.g. enantiomers, diastereomers, *E/Z* isomers, atropisomers, etc.), as well as racemates thereof and mixtures of different proportions of enantiomers or diastereomers, or mixtures of any of the foregoing forms as well as salts (e.g., pharmaceutically acceptable salts).

10 **[00108]** Additional aspects and advantages of the present disclosure will become readily apparent to those skilled in this art from the following detailed description, wherein only illustrative embodiments of the present disclosure are shown and described. As will be realized, the present disclosure is capable of other and different embodiments, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.

INCORPORATION BY REFERENCE

20 **[00109]** All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. To the extent publications and patents or patent applications incorporated by reference contradict the disclosure contained in the specification, the specification is intended to supersede and/or take precedence over any such contradictory material.

BRIEF DESCRIPTION OF THE DRAWINGS

25 **[00110]** FIGS. 1A and B show the chr13:52535985:A:C [hg19/b37] variant reduces exon 6 inclusion in ATP7B minigenes. FIG. 1A shows RT-PCR analysis of HEK293T cells transfected with ATP7B minigenes. Exon inclusion (445 bp) and exclusion (368 bp) fragments are indicated by black solid arrowheads for both wildtype minigene (WT) and chr13:52535985:A:C [hg19/b37] (M645R) variant minigene. 100 bp DNA ladder is shown for size reference. FIG. 1B depicts the percentage of exon 6 inclusion in ATP7B minigenes calculated by quantification of the RT-PCR fragments observed in FIG. 1A.

30 **[00111]** FIGS. 2A, 2B, and 2C show antisense oligonucleotide target (“hotspot”) identification by coarse-tilling of ATP7B minigenes. FIG. 2A is a schematic representation of the relative locations of the set of antisense oligonucleotides having the sequences set forth in SEQ ID NOs: 3-34, coarse-tilled across exon 6 and the flanking introns. FIG. 2B shows capillary electrophoresis of RT-PCR products of HEK293T cells transfected with ATP7B wildtype and M645R ATP7B minigenes and antisense oligonucleotides having the sequence set forth in SEQ ID NOs: 3-34. A 100 bp DNA ladder is shown for size reference with the exon 6 inclusion band at 445 bp and exclusion band at 368 bp. FIG. 2C depicts the percentage of exon 6 inclusion in ATP7B wildtype and M645R ATP7B minigenes co-transfected with antisense oligonucleotides having the sequence set forth in SEQ ID NOs: 3-34 calculated by

quantification of the RT-PCR fragments observed in FIG. 2B. Inclusion percentages were normalized according to fragment sizes (n=4 for WT, n=5 for M645R).

[00112] FIG. 3 is a depiction of the forward strand of the human genome surrounding exon 6 of the ATP7B gene, which is in reverse complement (i.e., antisense) with respect to the ATP7B transcript. (SEQ ID NO: 163 CCAGGTAGAGGAAGGGACTTAGATGAGAGCTGGAGTTTATCTTTTGTGTTCTACCTAC and SEQ ID: 164 CTTGTCATTAAGAGAGGGGTGGGGAAAAAGGAGGAAGGTACTTGGTTAAATATGCATTGGCAGAAAGCACTTTTCAGCTTTGGAAATTAGAAAG) (underlined). The location and sequence of antisense oligonucleotides according to the invention are also shown SEQ ID NOs: 28-33, 47-51, 60, 61, 63-67, 76-81, 93-95, 120, 121, 130, 132-137, 147, and 150 and their corresponding effect on the percent of exon 6 spliced in (dPSI).

[00113] FIGS. 4A and 4B show recovery of copper tolerance and ATP7B protein levels upon treatment with a splice modulating antisense oligonucleotide having a sequence set forth in SEQ ID: 29. FIG. 4A depicts the results of a copper sensitivity assay using both HepG2 wild-type (Wt) and 2F3 cells as well as corresponding non-transfected (NT) controls. 2F3 cells are compound heterozygous, with one allele including the M645R mutation and the other allele inactivated by the insertion of a plasmid sequence. The M645R variant (2F3) has lower cell viability with increasing copper concentration as compared to wildtype cells. This phenotype is partially rescued by transfection with an oligonucleotide (SEQ ID NO: 29), which increases the inclusion of exon 6 (error bars representing the standard deviation for the experiment done in triplicate). FIG. 4B depicts a western blot against ATP7B using both HepG2 wild-type (Wt) and 2F3 cells as well as corresponding non-transfected (NT) controls. The mutant (2F3) shows a decrease in ATP7B protein levels. This phenotype is partially rescued by transfection with an oligonucleotide (SEQ ID NO: 29), which increases the inclusion of exon 6.

DETAILED DESCRIPTION

[00114] In general, the present invention provides antisense oligonucleotides, compositions, and methods that target ATP7B exon 6 or a flanking intron. Surprisingly, the inventors have found that altering ATP7B gene splicing to include exon 6 in the transcript may be used to treat Wilson disease and antisense oligonucleotides may be used to alter splicing of the ATP7B gene to include exon 6. The antisense oligonucleotides of the invention may modulate splicing of ATP7B pre-mRNA to increase the level of ATP7B mRNA molecules having exon 6. Accordingly, the antisense oligonucleotides may be used to treat Wilson disease in a subject in need of a treatment therefor. Typically, an antisense oligonucleotide includes a nucleobase sequence at least 70% complementary to an ATP7B target sequence in exon 6, a 5'-flanking intron, a 3'-flanking intron, or a combination of exon 6 and the 5'-flanking or 3'-flanking intron.

[00115] Genetic variants may correspond to changes or modifications in transcription and/or splicing. RNA is initially transcribed from DNA as pre-mRNA, with protein-coding and 5'UTR/3'UTR exons separated by introns. Splicing generally refers to the molecular process, carried out by the spliceosome complexes that may remove introns and adjoins exons, producing a mature mRNA sequence, which is then scanned and translated to protein by the ribosome. The molecular reaction catalyzed by the spliceosome may comprise (i) nucleophilic attack of the branch site adenosine 2'OH onto the outmost base of the intronic donor dinucleotide, with consequent release of the outmost exonic donor base 3'OH;

and (ii) nucleophilic attack of the exonic donor 3'OH onto the outmost exonic acceptor base, with consequent release of the intron lariat and the spliced exons.

[00116] Splicing sequence changes can include the following categories: (a) alteration of a splice site (denominated canonical splice site) or exon recognition sequence required for the proper composition of a gene product, and (b) activation and utilization of an incorrect splice site (denominated cryptic splice site), or incorrect recognition of intronic sequence as an exon (denominated pseudo exon); both may result in the improper composition of a gene product. The splice site recognition signal may be required for spliceosome assembly and can comprise the following structures: (i) highly conserved intronic dinucleotide (AG, GT), immediately adjacent to the exon-intron boundary, and (ii) consensus sequence surrounding the intronic dinucleotide (often delimited to 3 exonic and 6 intronic nucleotides for the donor site, 3 exonic and 20 intronic nucleotides for the acceptor site) and branch site (variable position on the intronic acceptor side); both with lower conservation and more sequence variety.

[00117] In addition to splice site recognition, the exon recognition signal may comprise a plethora of motifs recognized by splicing factors and other RNA binding proteins, some of which may be ubiquitously expressed and some of which may be tissue specific. These motifs may be distributed over the exon body and in the proximal intronic sequence. The term "splicing enhancer" refers to motifs with positive effects (e.g., causing an increase) on exon inclusion, and the term "splicing silencer" refers to motifs with negative effects (e.g., causing a decrease) on exon inclusion. The exon recognition signal may be particularly important for correct splicing in the presence of weak consensus sequence. When a variant weakens the splice site recognition, the exon can be skipped and/or a nearby cryptic splice site which is already fairly strong can be used; especially in the presence of short introns, full intron retention is also a possible outcome. In particular, alteration of the intronic dinucleotide often results in splicing alteration, whereas consensus sequence alteration may be, on average, less impactful and more context-dependent. When the exon recognition signal is weakened, exon skipping may be a more likely outcome, but cryptic splice site use is also possible, especially in the presence of very weak consensus sequence. Variants can also strengthen a weak cryptic splice site in proximity of the canonical splice site, and significantly increase its usage resulting in improper splicing and incorrect gene product (with effects including amino acid insertion/deletion, frameshift, and stop-gain). Finally, variants that are more distant from canonical splice site can induce recognition of an exonic sequence as an intron, again resulting in improper gene product composition; specifically, these variants can increase the strength of the splice sites or the exon recognition signal.

[00118] Antisense oligonucleotides can be used to modulate gene splicing (e.g., by targeting splicing regulatory elements of the gene).

[00119] Antisense oligonucleotides may comprise splice-switching oligonucleotides (SSOs), which may modulate splicing by steric blockage (e.g., to enhance the inclusion of exon 6), preventing the spliceosome assembly or the binding of splicing factors and RNA binding proteins. Blocking the spliceosome assembly proteins may be therapeutically used to cause exon skipping. Blocking binding of specific splicing factors or RNA binding proteins that have an inhibitory effect may be used to produce increased exon inclusion. Specific steric blocker antisense oligonucleotide chemistries may include the modified RNA chemistry with phosphorothioate backbone (PS) with a sugar modification (e.g., 2'-modification) and phosphorodiamidate morpholino (PMO). Exemplary PS backbone sugar modifications may include 2'-O-methyl (2'OMe) and 2'-O-methoxyethyl (2'-MOE), which is also known as 2'-

methoxyethoxy. Other nucleotide modifications may be used, for example, for the full length of the oligonucleotide or for specific bases. The oligonucleotides can be covalently conjugated to a targeting moiety (e.g., a GalNAc cluster), or to a peptide (e.g., a cell penetrating peptide), or to another molecular or multimolecular group (e.g., a hydrophobic moiety or neutral polymer) different from the rest of the oligonucleotide. Antisense oligonucleotides may be used as a single stereoisomer or a combination of stereoisomers.

[00120] The ATP7B gene (ATPase copper transporting beta, OMIM: 606882) may play an important role in pathogenicity of Wilson disease (also known as Wilson's disease). ATP7B is a gene encoding an intracellular trans-Golgi copper transporter. The gene may be expressed in liver hepatocytes and may be required for copper excretion from the bloodstream to the bile and for ceruloplasmin copper loading. Defective copper excretion and/or ceruloplasmin copper loading can lead to toxic effects in the liver and central nervous system. ATP7B homozygous or compound heterozygous loss-of-function may result in the autosomal recessive Wilson Disease (OMIM: 277900).

[00121] Recognizing a need for effective splicing modulation therapies for diseases such as Wilson disease, the present disclosure provides ATP7B splice-modulating antisense oligonucleotides comprising sequences targeted to a splicing regulatory element of an abnormally spliced exon or an intron adjacent to an abnormally spliced exon of ATP7B. The present disclosure also provides methods for modulating splicing of ATP7B RNA in a cell, tissue, or organ of a subject by bringing the cell, tissue, or organ in contact with an antisense oligonucleotide of the invention. An ATP7B splice-modulating antisense oligonucleotide may comprise a nucleobase sequence targeted to a splicing regulatory element of an abnormally spliced exon or an intron adjacent to an abnormally spliced exon of ATP7B. In addition, the present disclosure provides a method for treating Wilson disease in a subject by administering to the subject a therapeutically effective amount of an oligonucleotide of the invention. An ATP7B splice-modulating antisense oligonucleotide may comprise a sequence targeted to a splicing regulatory element of an abnormally spliced exon or an intron adjacent to an abnormally spliced exon of ATP7B.

[00122] Splicing regulatory elements may include, for example, exonic splicing silencer elements or intronic splicing silencer elements. The antisense oligonucleotides may comprise sequences targeted to an exon or an intron adjacent to the exon of ATP7B which modulates variant splicing of ATP7B RNA. The modulation of splicing may result in an increase in exon inclusion. Antisense oligonucleotides may comprise a total of 8 to 50 nucleotides (e.g., 8 to 16 nucleotides, 8 to 20 nucleotides, 12 to 20 nucleotides, 12 to 30 nucleotides, or 12 to 50 nucleotides).

[00123] Genetic aberrations of the ATP7B gene may play an important role in pathogenicity. In particular, an ATP7B M645R genetic aberration, ATP7B chr13:52535985:A:C [hg19/b37] (hg19 coordinates) (g.54646T>G mutant of SEQ ID NO: 1), may result in NM_000053.3 cDNA change 1934T>G and protein sequence Met645Arg (M645R) in exon 6. Genome coordinates may be expressed, for example, with respect to human genome reference hg19/b37. For example, this variant has been reported as pathogenic in patients with Wilson Disease.

[00124] Other exemplary genetic aberrations which are predicted *in silico* to cause a decrease in exon 6 inclusion and which have been observed in the Human Gene Mutation Database (HGMD) include chr13:52535964:T:C (position 54667 in SEQ ID NO: 1; HGMD ID: CS076596) and chr13:52535994:T:C (position 54637 in SEQ ID NO: 1; HGMD ID: CM164020).

[00125] These exemplary genetic aberrations may be targeted with antisense oligonucleotides to increase levels of exon inclusion, and other similar mutations in splicing regulatory elements may be targeted in a similar fashion.

[00126] Different antisense oligonucleotides can be combined for increasing the inclusion of exon 6 of ATP7B. A combination of two antisense oligonucleotides may be used in a method of the invention, such as two antisense oligonucleotides, three antisense oligonucleotides, four different antisense oligonucleotides, or five different antisense oligonucleotides targeting the same or different regions or hotspots.

[00127] An antisense oligonucleotide according to the invention may be indirectly administered using suitable techniques and methods known in the art. It may for example be provided to an individual or a cell, tissue or organ of the individual in the form of an expression vector wherein the expression vector encodes a transcript comprising said oligonucleotide. The expression vector is preferably introduced into a cell, tissue, organ or individual via a gene delivery vehicle. In an embodiment, there is provided a viral based expression vector comprising an expression cassette or a transcription cassette that drives expression or transcription of an antisense oligonucleotide as identified herein. Accordingly, the present invention provides a viral vector expressing an antisense oligonucleotide according to the invention.

[00128] An antisense oligonucleotide according to the invention may be directly administered using suitable techniques and methods known in the art, e.g., using conjugates described herein.

Conjugates

[00129] Oligonucleotides of the invention may include an auxiliary moiety, e.g., a targeting moiety, hydrophobic moiety, cell penetrating peptide, or a polymer. An auxiliary moiety may be present as a 5' terminal modification (e.g., covalently bonded to a 5'-terminal nucleoside), a 3' terminal modification (e.g., covalently bonded to a 3'-terminal nucleoside), or an internucleoside linkage (e.g., covalently bonded to phosphate or phosphorothioate in an internucleoside linkage).

Targeting Moieties

[00130] An oligonucleotide of the invention may include a targeting moiety.

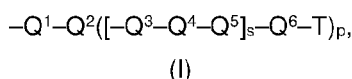
[00131] A targeting moiety is selected based on its ability to target oligonucleotides of the invention to a desired or selected cell population that expresses the corresponding binding partner (e.g., either the corresponding receptor or ligand) for the selected targeting moiety. For example, an oligonucleotide of the invention could be targeted to hepatocytes expressing asialoglycoprotein receptor (ASGP-R) by selecting a targeting moiety containing N-acetylgalactosamine (GalNAc).

[00132] A targeting moiety may include one or more ligands (e.g., 1 to 9 ligands, 1 to 6 ligands, 1 to 3 ligands, 3 ligands, or 1 ligand). The ligand may target a cell expressing asialoglycoprotein receptor (ASGP-R), IgA receptor, HDL receptor, LDL receptor, or transferrin receptor. Non-limiting examples of the ligands include N-acetylgalactosamine, glycyrrhetic acid, glycyrrhizin, lactobionic acid, lactoferrin, IgA, or a bile acid (e.g., lithocholyltaurine or taurocholic acid).

[00133] The ligand may be a small molecule, e.g., a small molecules targeting a cell expressing asialoglycoprotein receptor (ASGP-R). A non-limiting example of a small molecule targeting an asialoglycoprotein receptor is N-acetylgalactosamine. Alternatively, the ligand can be an antibody or an antigen-binding fragment or an engineered derivative thereof (e.g., Fcab or a fusion protein (e.g., scFv)).

[00134] A targeting moiety may be $\text{LinkA}(-\text{T})_p$, where LinkA is a multivalent linker, each T is a ligand (e.g., asialoglycoprotein receptor-targeting ligand (e.g., N-acetylgalactosamine)), and p is an integer from 1 to 9. When each T is N-acetylgalactosamine, the targeting moiety is referred to as a galactosamine cluster. Galactosamine clusters that may be used in oligonucleotides of the invention are known in the art. Non-limiting examples of the galactosamine clusters that may be included in the oligonucleotides of the invention are provided in US 5,994,517; US 7,491,805; US 9,714,421; US 9,867,882; US 9,127,276; US 2018/0326070; US 2016/0257961; WO 2017/100461; and in Sliedregt et al., *J. Med. Chem.*, 42:609-618, 1999. Ligands other than GalNAc may also be used in clusters, as described herein for galactosamine clusters.

[00135] Targeting moiety $\text{LinkA}(-\text{T})_p$ may be a group of formula (I):

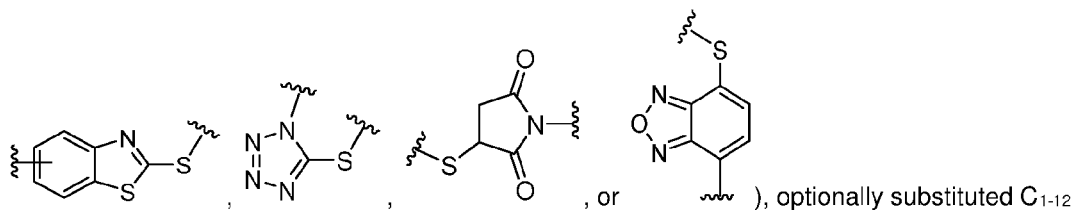
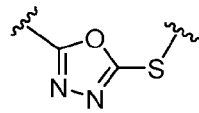


where

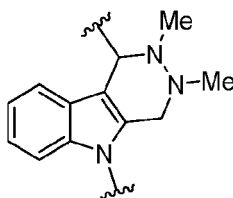
each s is independently an integer from 0 to 20 (e.g., from 0 to 10), where the repeating units are the same or different;

Q^1 is a conjugation linker (e.g., $[-\text{Q}^3-\text{Q}^4-\text{Q}^5]_s-\text{Q}^C$, where Q^C is optionally substituted C_{2-12} heteroalkylene (e.g., a heteroalkylene containing $-\text{C}(\text{O})-\text{N}(\text{H})-$, $-\text{N}(\text{H})-\text{C}(\text{O})-$, $-\text{S}(\text{O})_2-\text{N}(\text{H})-$, $-\text{N}(\text{H})-$

$\text{S}(\text{O})_2-$, or $-\text{S}-\text{S}-$), optionally substituted C_{1-12} thioheterocyclylene (e.g.,

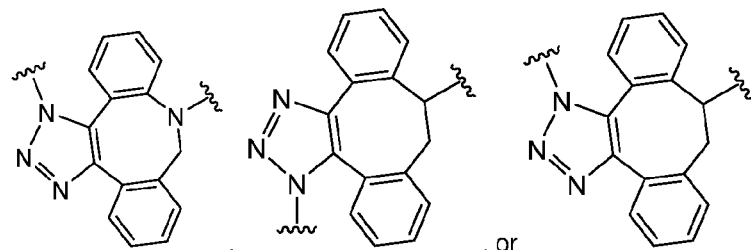
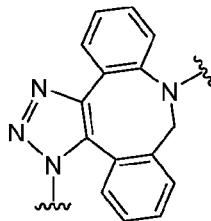


), optionally substituted C_{1-12} heterocyclylene (e.g., 1,2,3-triazole-1,4-diyl or

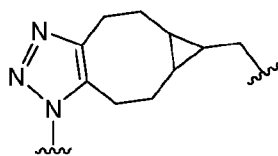


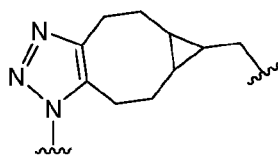
), cyclobut-3-ene-1,2-dione-3,4-diyl,

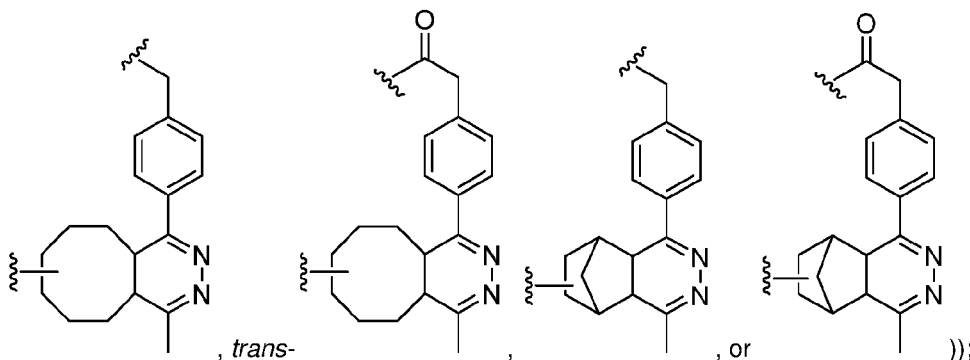
pyrid-2-yl hydrazone, optionally substituted C_{6-16} triazoloheterocyclylene (e.g.,



), optionally substituted C_{8-16}



triazolocycloalkenylene (e.g., , or a dihydropyridazine group (e.g., *trans*-



Q^2 is a linear group (e.g., $[-Q^3-Q^4-Q^5]_s-$), if p is 1, or a branched group (e.g., $[-Q^3-Q^4-Q^5]_s-Q^7([Q^3-Q^4-Q^5]_s-(Q^7)_{p1})_{p2}$, where $p1$ is 0, 1, or 2, and $p2$ is 0, 1, 2, or 3), if p is an integer from 2 to 9;

5 each Q^3 and each Q^6 is independently absent, $-CO-$, $-NH-$, $-O-$, $-S-$, $-SO_2-$, $-OC(O)-$, $-C(O)O-$, $-NHC(O)-$, $-C(O)NH-$, $-CH_2-$, $-CH_2NH-$, $-NHCH_2-$, $-CH_2O-$, or $-OCH_2-$;

each Q^4 is independently absent, optionally substituted C_{1-12} alkylene, optionally substituted C_{2-12} alkenylene, optionally substituted C_{2-12} alkynylene, optionally substituted C_{2-12} heteroalkylene, optionally substituted C_{6-10} arylene, optionally substituted C_{1-9} heteroarylene, or optionally substituted C_{1-9}

10 heterocyclylene;

each Q^5 is independently absent, $-CO-$, $-NH-$, $-O-$, $-S-$, $-SO_2-$, $-CH_2-$, $-C(O)O-$, $-OC(O)-$, $-C(O)NH-$, $-NH-C(O)-$, $-NH-CH(R^a)-C(O)-$, $-C(O)-CH(R^a)-NH-$, $-OP(O)(OH)O-$, or $-OP(S)(OH)O-$;

each Q^7 is independently optionally substituted hydrocarbon or optionally substituted heteroorganic (e.g., C_{1-6} alkane-triyl, optionally substituted C_{1-6} alkane-tetrayl, optionally substituted C_{2-6}

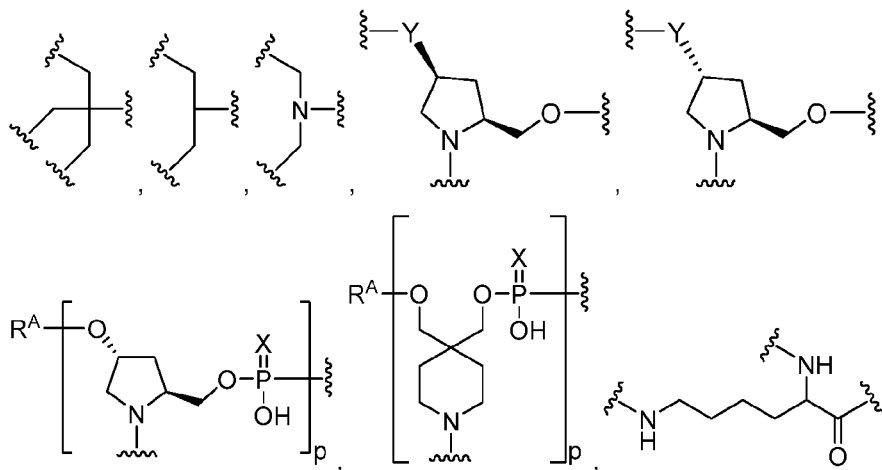
15 heteroalkane-triyl, or optionally substituted C_{2-6} heteroalkane-tetrayl); and

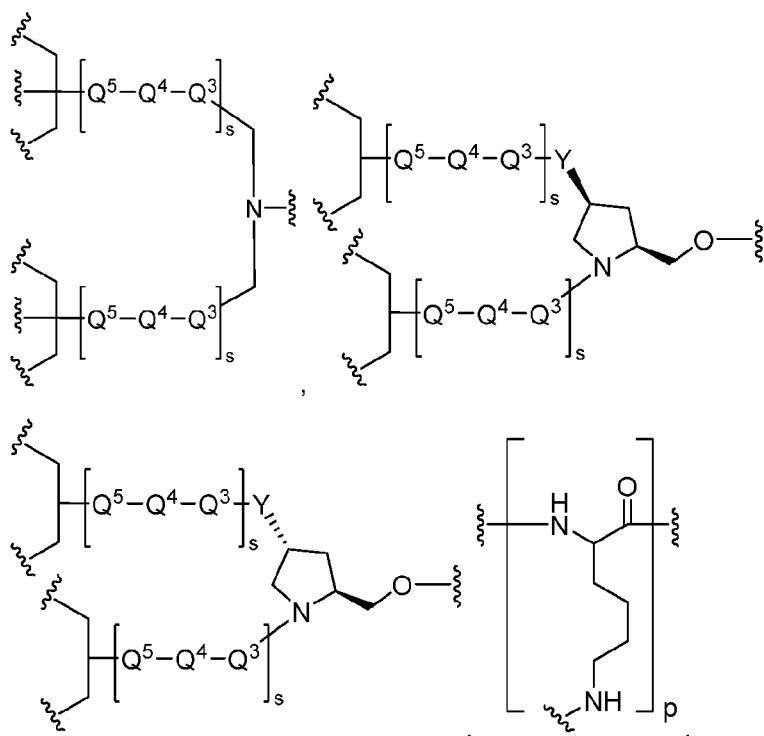
each R^a is independently H or an amino acid side chain;

provided that at least one of Q^3 , Q^4 , and Q^5 is present.

[00136] In some instances, for each occurrence of $[-Q^3-Q^4-Q^5]_s-$, at least one of Q^3 , Q^4 , and Q^5 is present.

20 **[00137]** In some instances, Q^7 may be a structure selected from the group consisting of:





[00138] where R^A is H or oligonucleotide, X is O or S, Y is O or NH, and the remaining variables are as described for formula (I).

- 5 **[00139]** Group –LinkA– may include a poly(alkylene oxide) (e.g., polyethylene oxide, polypropylene oxide, poly(trimethylene oxide), polybutylene oxide, poly(tetramethylene oxide), and diblock or triblock co-polymers thereof). In some embodiments, –LinkA– includes polyethylene oxide (e.g., poly(ethylene oxide) having a molecular weight of less than 1 kDa).

Hydrophobic Moieties

- 10 **[00140]** Advantageously, an oligonucleotide including a hydrophobic moiety may exhibit superior cellular uptake, as compared to an oligonucleotide lacking the hydrophobic moiety. Oligonucleotides including a hydrophobic moiety may therefore be used in compositions that are substantially free of transfecting agents. A hydrophobic moiety is a monovalent group (e.g., a bile acid (e.g., cholic acid, taurocholic acid, deoxycholic acid, oleyl lithocholic acid, or oleyl cholenic acid), glycolipid, phospholipid, sphingolipid, isoprenoid, vitamin, saturated fatty acid, unsaturated fatty acid, fatty acid ester, triglyceride, pyrene, porphyrine, texaphyrine, adamantane, acridine, biotin, coumarin, fluorescein, rhodamine, Texas-Red, digoxigenin, dimethoxytrityl, *t*-butyldimethylsilyl, *t*-butyldiphenylsilyl, cyanine dye (e.g., Cy3 or Cy5), Hoechst 33258 dye, psoralen, or ibuprofen) covalently linked to the oligonucleotide backbone (e.g., 5'-terminus). Non-limiting examples of the monovalent group include ergosterol, stigmasterol, β -sitosterol, campesterol, fucosterol, saringosterol, avenasterol, coprostanol, cholesterol, vitamin A, vitamin D, vitamin E, cardiolipin, and carotenoids. The linker connecting the monovalent group to the oligonucleotide may be an optionally substituted C_{1-60} hydrocarbon (e.g., optionally substituted C_{1-60} alkylene) or an optionally substituted C_{2-60} heteroorganic (e.g., optionally substituted C_{2-60} heteroalkylene), where the linker may be optionally interrupted with one, two, or three instances independently selected from the group consisting of an optionally substituted arylene, optionally substituted heterocyclylene, and optionally substituted cycloalkylene. The linker may be bonded to an oligonucleotide through, e.g., an oxygen atom attached to
- 25

a 5'-terminal carbon atom, a 3'-terminal carbon atom, a 5'-terminal phosphate or phosphorothioate, a 3'-terminal phosphate or phosphorothioate, or an internucleoside linkage.

Cell Penetrating Peptides

[00141] One or more cell penetrating peptides (e.g., from 1 to 6 or from 1 to 3) can be attached to an oligonucleotide disclosed herein as an auxiliary moiety. The CPP can be linked to the oligonucleotide through a disulfide linkage, as disclosed herein. Thus, upon delivery to a cell, the CPP can be cleaved intracellularly, e.g., by an intracellular enzyme (e.g., protein disulfide isomerase, thioredoxin, or a thioesterase) and thereby release the polynucleotide.

[00142] CPPs are known in the art (e.g., TAT or Arg₈) (Snyder and Dowdy, 2005, *Expert Opin. Drug Deliv.* 2, 43-51). Specific examples of CPPs including moieties suitable for conjugation to the oligonucleotides disclosed herein are provided, e.g., in WO 2015/188197; the disclosure of these CPPs is incorporated by reference herein.

[00143] CPPs are positively charged peptides that are capable of facilitating the delivery of biological cargo to a cell. It is believed that the cationic charge of the CPPs is essential for their function.

Moreover, the transduction of these proteins does not appear to be affected by cell type, and these proteins can efficiently transduce nearly all cells in culture with no apparent toxicity. In addition to full-length proteins, CPPs have also been used successfully to induce the intracellular uptake of DNA, antisense polynucleotides, small molecules, and even inorganic 40 nm iron particles suggesting that there is considerable flexibility in particle size in this process.

[00144] In one embodiment, a CPP useful in the methods and compositions of the invention includes a peptide featuring substantial alpha-helicity. It has been discovered that transfection is optimized when the CPP exhibits significant alpha-helicity. In another embodiment, the CPP includes a sequence containing basic amino acid residues that are substantially aligned along at least one face of the peptide. A CPP useful in the invention may be a naturally occurring peptide or a synthetic peptide.

Polymers

[00145] An oligonucleotide of the invention may include covalently attached neutral polymer-based auxiliary moieties. Neutral polymers include poly(C₁₋₆ alkylene oxide), e.g., poly(ethylene glycol) and poly(propylene glycol) and copolymers thereof, e.g., di- and triblock copolymers. Other examples of polymers include esterified poly(acrylic acid), esterified poly(glutamic acid), esterified poly(aspartic acid), poly(vinyl alcohol), poly(ethylene-co-vinyl alcohol), poly(*N*-vinyl pyrrolidone), poly(ethyloxazoline), poly(alkylacrylates), poly(acrylamide), poly(*N*-alkylacrylamides), poly(*N*-acryloylmorpholine), poly(lactic acid), poly(glycolic acid), poly(dioxanone), poly(caprolactone), styrene-maleic acid anhydride copolymer, poly(L-lactide-co-glycolide) copolymer, divinyl ether-maleic anhydride copolymer, *N*-(2-hydroxypropyl)methacrylamide copolymer (HMPA), polyurethane, *N*-isopropylacrylamide polymers, and poly(*N,N*-dialkylacrylamides). Exemplary polymer auxiliary moieties may have molecular weights of less than 100, 300, 500, 1000, or 5000 Da (e.g., greater than 100 Da). Other polymers are known in the art.

Nucleobase Modifications

[00146] Oligonucleotides of the invention may include one or more modified nucleobases.

Unmodified nucleobases include the purine bases adenine (A) and guanine (G), and the pyrimidine bases thymine (T), cytosine (C), and uracil (U). Modified nucleobases include 5-substituted pyrimidines, 6-azapyrimidines, alkyl or alkynyl substituted pyrimidines, alkyl substituted purines, and N-2, N-6 and O-6

substituted purines, as well as synthetic and natural nucleobases, e.g., 5-methylcytosine, 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-alkyl (e.g., 6-methyl) adenine and guanine, 2-alkyl (e.g., 2-propyl) adenine and guanine, 2-thiouracil, 2-thiothymine, 2-thiocytosine, 5-halouracil, 5-halocytosine, 5-propynyl uracil, 5-propynyl cytosine, 5-trifluoromethyl uracil, 5-trifluoromethyl cytosine, 7-methyl guanine, 7-methyl adenine, 8-azaguanine, 8-azaadenine, 7-deazaguanine, 7-deazaadenine, 3-deazaguanine, 3-deazaadenine. Certain nucleobases are particularly useful for increasing the binding affinity of nucleic acids, e.g., 5-substituted pyrimidines; 6-azapyrimidines; N2-, N6-, and/or O6-substituted purines. Nucleic acid duplex stability can be enhanced using, e.g., 5-methylcytosine. Non-limiting examples of nucleobases include: 2-aminopropyladenine, 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-N-methylguanine, 6-N-methyladenine, 2-propyladenine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-propynyl ($\text{—C}\equiv\text{C—CH}_3$) uracil, 5-propynylcytosine, 6-azouracil, 6-azocytosine, 6-azothymine, 5-ribosyluracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl, 8-aza and other 8-substituted purines, 5-halo, particularly 5-bromo, 5-trifluoromethyl, 5-halouracil, and 5-halocytosine, 7-methylguanine, 7-methyladenine, 2-F-adenine, 2-aminoadenine, 7-deazaguanine, 7-deazaadenine, 3-deazaguanine, 3-deazaadenine, 6-N-benzoyl adenine, 2-N-isobutyrylguanine, 4-N-benzoylcytosine, 4-N-benzoyluracil, 5-methyl 4-N-benzoylcytosine, 5-methyl 4-N-benzoyluracil, universal bases, hydrophobic bases, promiscuous bases, size-expanded bases, and fluorinated bases. Further modified nucleobases include tricyclic pyrimidines, such as 1,3-diazaphenoxazine-2-one, 1,3-diazaphenothiazine-2-one and 9-(2-aminoethoxy)-1,3-diazaphenoxazine-2-one (G-clamp). Modified nucleobases may also include those in which the purine or pyrimidine base is replaced with other heterocycles, for example 7-deazaadenine, 7-deazaguanine, 2-aminopyridine and 2-pyridone. Further nucleobases include those disclosed in Merigan et al., U.S. Pat. No. 3,687,808, those disclosed in The Concise Encyclopedia Of Polymer Science And Engineering, Kroschwitz, J. I., Ed., John Wiley & Sons, 1990, 858-859; Englisch et al., Angewandte Chemie, International Edition, 1991, 30, 613; Sanghvi, Y. S., Chapter 15, Antisense Research and Applications, Crooke, S. T. and Lebleu, B., Eds., CRC Press, 1993, 273-288; and those disclosed in Chapters 6 and 15, Antisense Drug Technology, Crooke S. T., Ed., CRC Press, 2008, 163-166 and 442-443.

[00147] The replacement of cytidine with 5-methylcytidine can reduce immunogenicity of oligonucleotides, e.g., those oligonucleotides having CpG units.

[00148] The replacement of one or more guanosines with, e.g., 7-deazaguanosine or 6-thioguanosine, may inhibit the antisense activity reducing G tetraplex formation within antisense oligonucleotides.

Sugar Modifications

[00149] Oligonucleotides of the invention may include one or more sugar modifications in nucleosides. Nucleosides having an unmodified sugar include a sugar moiety that is a furanose ring as found in ribonucleosides and 2'-deoxyribonucleosides.

[00150] Sugars included in the nucleosides of the invention may be non-furanose (or 4'-substituted furanose) rings or ring systems or open systems. Such structures include simple changes relative to the natural furanose ring (e.g., a six-membered ring). Alternative sugars may also include sugar surrogates wherein the furanose ring has been replaced with another ring system such as, e.g., a morpholino or hexitol ring system. Non-limiting examples of sugar moieties useful that may be included in the

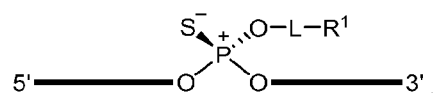
oligonucleotides of the invention include β -D-ribose, β -D-2'-deoxyribose, substituted sugars (e.g., 2', 5', and bis substituted sugars), 4'-S-sugars (e.g., 4'-S-ribose, 4'-S-2'-deoxyribose, and 4'-S-2'-substituted ribose), bridged sugars (e.g., the 2'-O—CH₂-4' or 2'-O—(CH₂)₂-4' bridged ribose derived bicyclic sugars) and sugar surrogates (when the ribose ring has been replaced with a morpholino or a hexitol ring system).

[00151] Typically, a sugar modification may be, e.g., a 2'-substitution, locking, carbocyclization, or unlocking. A 2'-substitution is a replacement of 2'-hydroxyl in ribofuranose with 2'-fluoro, 2'-methoxy, or 2'-(2-methoxy)ethoxy. A locking modification is an incorporation of a bridge between 4'-carbon atom and 2'-carbon atom of ribofuranose. Nucleosides having a sugar with a locking modification are known in the art as bridged nucleic acids, e.g., locked nucleic acids (LNA), ethylene-bridged nucleic acids (ENA), and cEt nucleic acids. The bridged nucleic acids are typically used as affinity enhancing nucleosides.

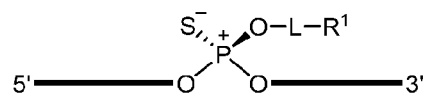
Internucleoside Linkage Modifications

[00152] Oligonucleotides of the invention may include one or more internucleoside linkage modifications. The two main classes of internucleoside linkages are defined by the presence or absence of a phosphorus atom. Non-limiting examples of phosphorus-containing internucleoside linkages include phosphodiester linkages, phosphotriester linkages, phosphorothioate diester linkages, phosphorothioate triester linkages, morpholino internucleoside linkages, methylphosphonates, and phosphoramidate. Non-limiting examples of non-phosphorus internucleoside linkages include methylenemethylimino (—CH₂—N(CH₃)—O—CH₂—), thiodiester (—O—C(O)—S—), thionocarbamate (—O—C(O)(NH)—S—), siloxane (—O—Si(H)₂—O—), and N,N'-dimethylhydrazine (—CH₂—N(CH₃)—N(CH₃)—). Modified linkages, compared to natural phosphodiester linkages, can be used to alter, typically increase, nuclease resistance of the oligonucleotide. Methods of preparation of phosphorous-containing and non-phosphorous-containing internucleoside linkages are known in the art.

[00153] Internucleoside linkages may be stereochemically enriched. For example, phosphorothioate-based internucleoside linkages (e.g., phosphorothioate diester or phosphorothioate triester) may be stereochemically enriched. The stereochemically enriched internucleoside linkages including a stereogenic phosphorus are typically designated *S_P* or *R_P* to identify the absolute stereochemistry of the phosphorus atom. Within an oligonucleotide, *S_P* phosphorothioate indicates the following structure:



Within an oligonucleotide, *R_P* phosphorothioate indicates the following structure:



[00154] The oligonucleotides of the invention may include one or more neutral internucleoside linkages. Non-limiting examples of neutral internucleoside linkages include phosphotriesters, phosphorothioate triesters, methylphosphonates, methylenemethylimino (5'-CH₂—N(CH₃)—O-3'), amide-3 (5'-CH₂—C(=O)—N(H)-3'), amide-4 (5'-CH₂—N(H)—C(=O)-3'), formacetal (5'-O—CH₂—O-3'), and thioformacetal (5'-S—CH₂—O-3'). Further neutral internucleoside linkages include nonionic linkages including siloxane (dialkylsiloxane), carboxylate ester, carboxamide, sulfide, sulfonate ester, and amides

(See for example: Carbohydrate Modifications in Antisense Research; Y. S. Sanghvi and P. D. Cook, Eds., ACS Symposium Series 580; Chapters 3 and 4, 40-65).

Terminal Modifications

5 **[00155]** Oligonucleotides of the invention may include a terminal modification, e.g., a 5'-terminal modification or a 3'-terminal modification.

[00156] The 5' end of an oligonucleotide may be, e.g., hydroxyl, a hydrophobic moiety, a targeting moiety, 5' cap, phosphate, diphosphate, triphosphate, phosphorothioate, diphosphorothioate, triphosphorothioate, phosphorodithioate, diphosphorodithioate, triphosphorodithioate, phosphonate, 10 phosphoramidate, a cell penetrating peptide, an endosomal escape moiety, or a neutral organic polymer. An unmodified 5'-terminus is hydroxyl or phosphate. An oligonucleotide having a 5' terminus other than 5'-hydroxyl or 5'-phosphate has a modified 5' terminus.

[00157] The 3' end of an oligonucleotide may be, e.g., hydroxyl, a targeting moiety, a hydrophobic moiety, phosphate, diphosphate, triphosphate, phosphorothioate, diphosphorothioate, 15 triphosphorothioate, phosphorodithioate, diphosphorodithioate, triphosphorodithioate, phosphonate, phosphoramidate, a cell penetrating peptide, an endosomal escape moiety, or a neutral organic polymer (e.g., polyethylene glycol). An unmodified 3'-terminus is hydroxyl or phosphate. An oligonucleotide having a 3' terminus other than 3'-hydroxyl or 3'-phosphate has a modified 3' terminus.

[00158] The terminal modification (e.g., 5'-terminal modification) may be, e.g., a targeting moiety as 20 described herein.

[00159] The terminal modification (e.g., 5'-terminal modification) may be, e.g., a hydrophobic moiety as described herein.

Complementarity

25 **[00160]** In some embodiments, oligonucleotides of the invention are complementary to an ATP7B target sequence over the entire length of the oligonucleotide. In other embodiments, oligonucleotides are at least 99%, 95%, 90%, 85%, 80%, or 70% complementary to the ATP7B target sequence. In further 30 embodiments, oligonucleotides are at least 80% (e.g., at least 90% or at least 95%) complementary to the ATP7B target sequence over the entire length of the oligonucleotide and include a nucleobase sequence that is fully complementary to an ATP7B target sequence. The nucleobase sequence that is fully complementary may be, e.g., 6 to 20, 10 to 18, or 18 to 20 contiguous nucleobases in length.

[00161] An oligonucleotide of the invention may include one or more (e.g., 1, 2, 3, or 4) mismatched nucleobases relative to the target nucleic acid. In certain embodiments, a splice-switching activity against the target is reduced by such mismatch, but activity against a non-target is reduced by a greater amount. 35 Thus, the off-target selectivity of the oligonucleotides may be improved.

Methods for preparing compositions

[00162] The present disclosure provides methods for preparing or generating compositions provided herein. A nucleic acid molecule, such as an oligonucleotide, comprising a targeted sequence may be 40 generated, for example, by various nucleic acid synthesis approaches. For example, a nucleic acid molecule comprising a sequence targeted to a splice site may be generated by oligomerization of modified and/or unmodified nucleosides, thereby producing DNA or RNA oligonucleotides. Antisense

oligonucleotides can be prepared, for example, by solid phase synthesis. Such solid phase synthesis can be performed, for example, in multi-well plates using equipment available from vendors such as Applied Biosystems (Foster City, CA). It is well known to use similar techniques to prepare oligonucleotides such as the phosphorothioates and alkylated derivatives. Oligonucleotides may be subjected to purification and/or analysis using methods known to those skilled in the art. For example, analysis methods may include capillary electrophoresis (CE) and electrospray-mass spectroscopy.

Pharmaceutical Compositions

[00163] An oligonucleotide of the invention may be included in a pharmaceutical composition. A pharmaceutical composition typically includes a pharmaceutically acceptable diluent or carrier. A pharmaceutical composition may include (e.g., consist of), e.g., a sterile saline solution and an oligonucleotide of the invention. The sterile saline is typically a pharmaceutical grade saline. A pharmaceutical composition may include (e.g., consist of), e.g., sterile water and an oligonucleotide of the invention. The sterile water is typically a pharmaceutical grade water. A pharmaceutical composition may include (e.g., consist of), e.g., phosphate-buffered saline (PBS) and an oligonucleotide of the invention. The sterile PBS is typically a pharmaceutical grade PBS.

[00164] Pharmaceutical compositions may include one or more oligonucleotides and one or more excipients. Excipients may be selected from water, salt solutions, alcohol, polyethylene glycols, gelatin, lactose, amylase, magnesium stearate, talc, silicic acid, viscous paraffin, hydroxymethylcellulose and polyvinylpyrrolidone.

[00165] Pharmaceutical compositions including an oligonucleotide encompass any pharmaceutically acceptable salts of the oligonucleotide. Pharmaceutical compositions including an oligonucleotide, upon administration to a subject (e.g., a human), are capable of providing (directly or indirectly) the biologically active metabolite or residue thereof. Accordingly, for example, the disclosure is also drawn to pharmaceutically acceptable salts of oligonucleotides. Suitable pharmaceutically acceptable salts include, but are not limited to, sodium and potassium salts. In certain embodiments, prodrugs include one or more conjugate group attached to an oligonucleotide, wherein the conjugate group is cleaved by endogenous enzymes within the body.

[00166] Lipid moieties have been used in nucleic acid therapies in a variety of methods. In certain such methods, the nucleic acid, such as an oligonucleotide, is introduced into preformed liposomes or lipoplexes made of mixtures of cationic lipids and neutral lipids. DNA complexes with mono- or poly-cationic lipids may form, e.g., without the presence of a neutral lipid. A lipid moiety may be, e.g., selected to increase distribution of a pharmaceutical agent to a particular cell or tissue. A lipid moiety may be, e.g., selected to increase distribution of a pharmaceutical agent to fat tissue. A lipid moiety may be, e.g., selected to increase distribution of a pharmaceutical agent to muscle tissue.

[00167] Pharmaceutical compositions may include a delivery system. Examples of delivery systems include, but are not limited to, liposomes and emulsions. Certain delivery systems are useful for preparing certain pharmaceutical compositions including those including hydrophobic compounds. Certain organic solvents such as dimethylsulfoxide may be used.

[00168] Pharmaceutical compositions may include one or more tissue-specific delivery molecules designed to deliver the one or more pharmaceutical agents of the present invention to specific tissues or

cell types. For example, pharmaceutical compositions may include liposomes coated with a targeting moiety as described herein.

[00169] Pharmaceutical compositions may include a co-solvent system. Certain co-solvent systems include, e.g., benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. Such co-solvent systems may be used, e.g., for hydrophobic compounds. A non-limiting example of a co-solvent system is the VPD co-solvent system, which is a solution of absolute ethanol including 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant Polysorbate 80™ and 65% w/v polyethylene glycol 300. The proportions of such co-solvent systems may be varied considerably without significantly altering their solubility and toxicity characteristics. Furthermore, the identity of co-solvent components may be varied: for example, other surfactants may be used instead of Polysorbate 80™; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g., polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose.

[00170] Pharmaceutical compositions may be prepared for administration by injection or infusion (e.g., intravenous, subcutaneous, intramuscular, intrathecal, intracerebroventricular, etc.). A

pharmaceutical composition may include, e.g., a carrier and may be formulated, e.g., in aqueous solution, e.g., water or physiologically compatible buffers, e.g., Hanks's solution, Ringer's solution, or physiological saline buffer. Other ingredients may also be included (e.g., ingredients that aid in solubility or serve as preservatives). Injectable suspensions may be prepared, e.g., using appropriate liquid carriers, suspending agents and the like. Certain pharmaceutical compositions for injection are presented in unit dosage form, e.g., in ampoules or in multi-dose containers. Certain pharmaceutical compositions for injection may be, e.g., suspensions, solutions, or emulsions in oily or aqueous vehicles, and may contain excipients (e.g., suspending, stabilizing and/or dispersing agents). Certain solvents suitable for use in pharmaceutical compositions for injection include, but are not limited to, lipophilic solvents and fatty oils, e.g., sesame oil, synthetic fatty acid esters (e.g., ethyl oleate or triglycerides), and liposomes.

Methods of the Invention

[00171] The invention provides methods of using oligonucleotides of the invention.

[00172] A method of the invention may be a method of increasing the level of exon 6-containing ATP7B mRNA molecules in a cell expressing an aberrant ATP7B gene by contacting the cell with the antisense oligonucleotide of the invention.

[00173] A method of the invention may be a method of treating Wilson disease in a subject having an aberrant ATP7B gene by administering a therapeutically effective amount of the antisense oligonucleotide of the invention or a pharmaceutical composition of the invention to the subject in need thereof.

[00174] The oligonucleotide of the invention or the pharmaceutical composition of the invention may be administered to the subject using methods known in the art. For example, the oligonucleotide of the invention or the pharmaceutical composition of the invention may be administered parenterally (e.g., intravenously, intramuscularly, subcutaneously, transdermally, intranasally, or intrapulmonarily) to the subject.

[00175] Dosing is typically dependent on a variety of factors including, e.g., severity and responsiveness of the disease state to be treated. The treatment course may last, e.g., 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. Thus, optimum dosages, dosing methodologies and repetition rates can be established as needed. 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 may be from 0.01 µg to 1 g per kg of body weight, and may be given once or more daily, weekly, monthly, bimonthly, trimonthly, every six months, annually, or biannually. Frequency of dosage may vary. Repetition rates for dosing may be established, for example, 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 1 g per kg of body weight, e.g., once daily, twice daily, three times daily, every other day, weekly, biweekly, monthly, bimonthly, trimonthly, every six months, annually or biannually.

[00176] Methods of treating Wilson disease in a subject in need thereof may also include administering to the subject a pharmaceutically acceptable chelating agent or a pharmaceutically acceptable salt of zinc. Non-limiting examples of pharmaceutically acceptable salts of zinc include zinc acetate, zinc gluconate, and zinc sulfate. Non-limiting examples of pharmaceutically acceptable chelating agents include D-penicillamine, trientine, sodium mercaptosuccinate, dimercaptosuccinic acid, and tetrathiomolybdate. In some embodiments, the method includes administering to the subject a pharmaceutically acceptable salt of zinc (e.g., zinc acetate, zinc gluconate, or zinc sulfate). In one example, an oligonucleotide of the invention and a pharmaceutically acceptable chelating agent or salt of zinc are administered together in the same pharmaceutical composition. In another example, an oligonucleotide of the invention and a pharmaceutically acceptable chelating agent or salt of zinc are administered separately at about the same time (e.g., one minute apart or less, or five minutes apart or less). In some embodiments, an oligonucleotide of the invention and a pharmaceutically acceptable chelating agent or salt of zinc are administered separately via the same route of administration (e.g., intravenous injection). In some embodiments, an oligonucleotide of the invention and a pharmaceutically acceptable chelating agent or salt of zinc are administered separately via different routes of administration (e.g., intravenous injection of an oligonucleotide of the invention and oral administration of a pharmaceutically acceptable chelating agent or salt of zinc).

[00177] In some embodiments, an oligonucleotide of the invention is administered prior to a chelating agent or salt of zinc. In further embodiments, an oligonucleotide of the invention is administered within 1 hour of the chelating agent or salt of zinc administration (e.g., before, e.g., 15 min, 30 min, or 1 hour before). In some embodiments, an oligonucleotide of the invention is administered within 12 hours of the chelating agent or salt of zinc administration (e.g., before, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 hours before). In certain embodiments, an oligonucleotide of the invention is administered within 24 hours of the chelating agent or salt of zinc administration (e.g., before, e.g., 12 or 24 hours before). In particular embodiments, an oligonucleotide of the invention is administered within 1 week of the chelating agent or salt of zinc administration (e.g., before, e.g., 1, 2, 3, 4, 5, or 6 days before). In some embodiments, an oligonucleotide of the invention is administered within 1 month of the chelating agent or salt of zinc administration (e.g., before, e.g., 1, 2, 3, or 4 weeks before).

[00178] In some embodiments, an oligonucleotide of the invention is administered after a chelating agent or salt of zinc. In further embodiments, an oligonucleotide of the invention is administered within 1 hour of the chelating agent or salt of zinc administration (e.g., after, e.g., 15 min, 30 min, or 1 hour after).

In some embodiments, an oligonucleotide of the invention is administered within 12 hours of the chelating agent or salt of zinc administration (e.g., after, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 hours after). In certain embodiments, an oligonucleotide of the invention is administered within 24 hours of the chelating agent or salt of zinc administration (e.g., after, e.g., 12 or 24 hours after). In particular embodiments, an oligonucleotide of the invention is administered within 1 week of the chelating agent or salt of zinc administration (e.g., after, e.g., 1, 2, 3, 4, 5, or 6 days after). In some embodiments, an oligonucleotide of the invention is administered within 1 month of the chelating agent or salt of zinc administration (e.g., after, e.g., 1, 2, 3, or 4 weeks after).

EXAMPLES

[00179] The following materials, methods, and examples are illustrative only and not intended to be limiting.

Materials and Methods

[00180] In general, the practice of the present invention employs, unless otherwise indicated, conventional techniques of chemistry, molecular biology, recombinant DNA technology, and standard techniques in electrophoresis. See, e.g., Sambrook, Fritsch and Maniatis, *Molecular Cloning*: Cold Spring Harbor Laboratory Press (1989) and *Current Protocols in Molecular Biology*, eds. Ausubel et al., John Wiley & Sons (1992).

[00181] Oligonucleotides. All antisense oligonucleotides used were obtained from Integrated DNA Technologies Inc. (USA). All bases in the antisense oligonucleotides were 2'-O-methoxyethyl-modified (MOE) with a full phosphorothioate backbone.

[00182] Construction of ATP7B minigenes. Minigene plasmids for ATP7B exon 6 were designed to contain fragments corresponding to exon 5 to 7 of the genomic locus of ATP7B gene, including complete intronic sequences. To minimize aberrant splicing of the transcribed mRNA fragments, consensus splice acceptor sequences of exon 5 were removed by deleting nine nucleotides from the 5' end of exon 5, likewise consensus splice donor sequences of exon 7 were removed by deleting 40 nucleotides from the 3' end of exon 7. Minigenes were constructed by DNA assembly of PCR fragments that were amplified from HEK293T genomic DNA using KOD Hot Start DNA polymerase (Novagen). For the wild-type minigene, the full fragment was amplified with primers P461

(GATCCACCGGTCGCCACCATGACCTGCGCGTCCTGTGTC) (SEQ ID: 165) and P463 (TCCTCGCCCTTGCTCACCATGGACAGTCCTGGAATGATGTTGTGG) (SEQ ID: 166). For the mutant minigene, the NM_000053.3(ATP7B):c.1934T>G(p.Met645Arg) variant was introduced by site-directed mutagenesis of two overlapping fragments which were amplified with primer combination P461 (SEQ ID: 165) and P465 (CTGCTTTATTTCCCTCTTGTGGTCCAAGTGATGAGC) (SEQ ID: 167) and primer combination P464 (GACCACAAGAGGGAAATAAAGCAGTAGGTAGAACAC) (SEQ ID: 168) and P463 (SEQ ID: 166) respectively. PCR fragments were cloned into mClover2 plasmid that was linearized with primers P459 (ATGGTGAGCAAGGGCGAGGA) (SEQ ID: 169) and P460 (CATGGTGCGACCGGTGGAT) (SEQ ID: 170) using a NEBuilder HiFi DNA Assembly Kit (New England Biolabs) according to manufacturer's instructions. Plasmid DNA was isolated using the Presto Mini Plasmid Kit (Geneaid).

[00183] HepG2 mutant line, 2F3. This line was created from HepG2 cells by inserting the M645R variant using a CRISPR/ Cas 9 according to commonly used molecular biology techniques. The resulting

2F3 line carries the M645R variant at one allele and a large plasmid insertion at the other. This insertion has been assumed to render that allele unable to produce functional protein, thus creating a situation analogous to a compound heterozygote, consistent with what found in Wilson disease patients (Margarit et al, *Clin Genet.*, 68:61-68, 2005).

5 **[00184]** Cell culture. HEK293T cells were grown in Iscove's Modified Dulbecco's Medium (Gibco)

supplemented with 10% (v/v) Cosmic Calf Serum (HyClone), 2mM L-Glutamine (Gibco) and 1% antibiotics (100-U/ml penicillin G and 100-ug/ml streptomycin, Gibco) in a humidified incubator at 37°C with 5% CO₂. Upon reaching confluency the cells were passaged by washing with Phosphate-Buffered

10 Saline followed by Trypsin (Gibco) dissociation and plated in 10 to 20-fold dilution. HepG2 wild-type and mutant 2F3 cells were grown in Dulbecco's Modified Eagle's Medium (Gibco) supplemented with 10% heat inactivated fetal bovine serum (Gibco) in a humidified incubator at 37°C with 5% CO₂. Upon reaching confluency the cells were passaged by washing with Phosphate-Buffered Saline followed by TrypLE (Gibco) dissociation and plated in a culture flask in 2 to 4-fold dilution.

[00185] Transfection of HEK293T cells with minigene plasmids. For transfection of HEK293T cells

15 with minigene plasmids, 300,000 cells were plated in 1 ml of complete medium with antibiotics in 12-well tissue culture plates and incubated at 37°C and 5% CO₂ for 24 hours. Plasmid transfection mixes were made by combining 1,000 ng of plasmid DNA (~ 5 µl) diluted in 50 µl Opti-MEM reduced serum medium (Gibco) with 3.0 µl X-tremeGENE HP DNA Transfection Reagent solution (Roche). Plasmid transfection

20 mixes were incubated at room temperature for 20 minutes and then added to the HEK293T cells, which were subsequently incubated at 37°C and 5% CO₂. After 24 hours transfection media was replaced with 2,000 µl complete media. 48 hours after transfection, cells were lysed for RNA isolation and RT-PCR analysis.

[00186] Co-transfection of HEK293T cells with antisense oligonucleotides and minigene plasmids.

Antisense oligonucleotides and minigene plasmids were co-transfected into HEK293T cells in 96-well

25 format. Antisense oligonucleotide stock solutions (100 µM) were diluted to 5 µM in Opti-MEM reduced serum medium (Gibco). Antisense oligonucleotides were transfected at absolute amounts of 50 pmol of an antisense oligonucleotide per well containing 50,000 HEK293T cells. For this, 10-µl aliquots of 5 µM

antisense oligonucleotide solutions were transferred to the wells of a 96-well tissue culture plate and 10 µl

lipid transfection reagent solution containing 9.7 µl Opti-MEM and 0.3 µl Lipofectamine RNAiMAX

30 (Invitrogen) was added to the wells. Antisense oligonucleotide-lipid complexes in the mixture were formed by gentle mixing through pipetting twice and subsequent tapping of the plate followed by incubation for 20 minutes at room temperature. Next, 50,000 HEK293T cells in 100 µl complete media solution without

antibiotics were added to the antisense oligonucleotide-lipid complexes and incubated for five hours at 37°C and 5% CO₂. After incubation, cells were transfected with minigene plasmids. For this, plasmid

35 transfection mixes were made by combining 5 µl containing 200 ng of plasmid diluted in Opti-MEM with 0.6 µl X-tremeGENE 9 DNA Transfection Reagent solution (Roche). Plasmid transfection mixes were incubated at room temperature for 20 minutes and then added to the HEK293T cells, which were

subsequently incubated at 37°C and 5% CO₂. After 24 hours transfection media was replaced with 200 µl complete media. 48 hours after transfection, cells were lysed for RNA isolation and RT-PCR analysis.

40 **[00187]** Transfection HepG2 wild-type and mutant 2F3 cells. All reagents were used according to

manufacturer's recommendations. Cells were suspended by incubation with TrypLE for 15-20 minutes and diluted to 250000-500000 cells per milliliter in Dulbecco's Modified Eagle's Medium with 10% heat

inactivated fetal calf serum. 50 pmol of oligonucleotide in transfection medium, containing RNAiMAX (Invitrogen) and Opti-MEM (Gibco), was combined with 25000-50000 cells from the culture suspension above. RNA was collected 48 hours after transfection unless otherwise stated.

[00188] RNA preparation. RNA was prepared by using a SingleShot Cell Lysis kit (Bio-Rad) or

5 RNeasy total RNA kit (QIAGEN) according to manufacturer's recommendations, alternatively Direct-zol TM-96 MagBead RNA (Zymo Research) was used according to manufacturer's recommendations except wherein all washes with volumes of 500 μ L only 300 μ L was used and for all rpm speeds of 900 rpm, 1050 rpm was used.

[00189] RT-PCR analysis. Synthesis of first-strand cDNA was performed with the ImProm-II Reverse

10 Transcription System (Promega) according to manufacturer's recommendations with minor modifications. 50-300 ng of purified RNA were incubated in a 96-well PCR plate with 1 μ L Oligo-dT-VN primer (100 μ M, TTTTTTTTTTTTTTTTTT VN) for 5 min at 70°C, followed by rapid cooling for 5 min at 4°C. 14.5- μ L of

reverse transcriptase mixture, containing 20 Units ImProm-II Reverse Transcriptase, reaction buffer, 4 mM MgCl₂, 0.5 mM dNTPs (FroggaBio) and 40 units RNase Inhibitor (Bioshop) was added to the RNA-Oligo-

15 dT-VN reaction and incubated for 5 min at 25°C, 60 min at 42°C and finally cooled to 12°C. Target-specific splicing fragments were amplified by PCR. PCR primers for HepG2 wild-type and mutant 2F3 cell experiments were CCAGCAAAGCCCTTGTTAAG (SEQ ID NO: 156) and GCTCGTTGCTGGGTATCAG (SEQ ID NO: 157). PCR primers for minigene experiments were GATCACAGGGATGACCTGC (SEQ ID NO: 158) and GTTTACGTCGCCGTCCAG (SEQ ID NO: 159). PCR reactions contained 5 μ L first-strand

20 cDNA product, 0.4 μ M forward primer, 0.4 μ M reverse primer, 300 μ M of each dNTP, 25 mM Tricine, 7.0% Glycerol (m/v), 1.6% DMSO (m/v), 2 mM MgCl₂, 85 mM NH₄-acetate (pH8.7), and 1 unit Taq DNA polymerase (FroggaBio) in a total volume of 25 μ L. Fragments were amplified by a touchdown PCR program (95°C for 120 sec; 10 cycles of 95°C for 20 sec, 68°C for 30 sec with a decrement of 1°C per cycle, and 72°C for 60 sec; followed by 30 cycles of 95°C for 20 sec, 58°C for 30 sec, and 72°C for 60

25 sec; 72°C for 180 sec). **[00190]** Capillary electrophoresis. Samples were analyzed using a LabChip GX Touch Nucleic Acid Analyzer (GE) using a DNA 1K Hi Sensitivity LabChip and associated reagents (GE) according to manufacturer's recommendations.

[00191] Western blotting. Cells were grown and transfected as above. After 48 hours the media was

30 removed, and cells were rinsed with DPBS. The cells were suspended with TrypLE (Gibco) and pelleted. The TrypLE was removed and 150 μ L of ice cold RIPA buffer (SIGMA) with 1x HALT protease inhibitor (Pierce Biotechnology) was added to the well. The solution was placed on ice for 10 minutes and then centrifuged at 15000 rcf at 4°C. The supernatant was put into a fresh tube and the pellet was discarded.

Using a protein quantification kit (Pierce) the protein concentrations was determined. Twenty to thirty μ g

35 of lysate protein was heated at 70°C with Nupage buffer (Novex) and loaded onto a 10% Bis-Tris gel (Invitrogen). The gel was run for ~40 minutes at 200V in 1x MOPS buffer (Novex). The gel was removed and transferred to a PVDF membrane (GE) on ice for 90 minutes at 350 mA constant current. After

transfer, the membrane was blocked in TBST-5% milk for 60 minutes at room temperature. After blocking, primary antibodies for GAPDH (Abcam) and ATP7B (Abcam) were added in TBSB-1% milk and

40 refrigerated at 4°C overnight. The membrane was then rinsed with TBST for 5 minutes 5 times.

Secondary antibodies conjugated with horseradish peroxidase (Cell Signalling technology) were added to

the solution for 60 minutes at room temperature. The membrane was then rinsed with TBST for 5 minutes 5 times. The images were recorded with a GE AI600RGB device.

[00192] Copper sensitivity assay. Cells were grown and transfected as stated above. Cells were then transfected with an antisense oligonucleotide or not transfected for control. Forty-eight hours after
5 transfection, CuCl₂ (Sigma) was prepared in deionized water at a concentration of 0.5M. This was then diluted in DMEM supplemented with 10% fetal calf serum to final concentrations of 0.2 mM, 0.5 mM, 0.75 mM, 1.0 mM, and 1.25 mM. A media change was then performed on the cells using the media + CuCl₂. After 48 hours the cell viability was read with a Neo2 instrument (Biotek) using CellTiter-Fluor™ Cell Viability Assay (Promega). Data was normalized to the copper-free treatment within groups.

10 **Example 1 The splicing of ATP7B exon 6 is disrupted in the chr13:52535985:A:C [hg19/b37] variant and can be partially rescued through the use of antisense oligonucleotides**

[00193] To confirm exon 6 skipping in the chr13:52535985:A:C [hg19/b37] (M645R) variant, wild type and variant containing minigenes were constructed containing exons 5-7 and the corresponding introns, 5 and 6. Minigenes were then transfected into HEK293T cells to examine the effect of the M645R variant
15 on splicing. As seen in FIG. 1A, wildtype minigenes showed both exon 6 inclusion, represented by the upper band, and exclusion. M645R mutants, however, showed no exon 6 inclusion indicating the chr13:52535985:A:C [hg19/b37] mutation induces exon 6 skipping. The results of the experiment in FIG 1A were replicated and quantified. As seen in FIG 1B, there is a robust (standard deviations of 5.7% and 1.3% for wildtype and M645R minigenes respectively) decrease in exon 6 inclusion due to the
20 chr13:52535985:A:C [hg19/b37] mutation.

[00194] To examine the ability of antisense oligonucleotides to promote exon 6 inclusion in the M645R variant the minigenes above were co-transfected with antisense oligonucleotides having sequences set forth in SEQ ID NOs: 3-34 (see Table 1). Antisense oligonucleotides were tiled along exon 6 and the surrounding introns. FIG. 2A depicts the location of the targeted ASOs relative to exon 6 and
25 the surrounding introns. FIG. 2B shows the RT-PCR samples measured by capillary electrophoresis. A 100 bp DNA ladder is shown for size reference with the exon 6 inclusion band at 445 bp and exclusion band at 368 bp. These results were quantified and are depicted in FIG. 2C. Observing both FIG. 2B and 2C it is clear that targeting the intronic regions surrounding exon 6 induces exon 6 inclusion of both wildtype and M645R variant minigenes. These observations suggest antisense oligonucleotides targeting
30 these regions or "hotspots" (positions 54522-54593 and 54665-54718 in SEQ ID NO: 1; chr13:52536038-52536109 and chr13:52535966-52535913), e.g., those complementary to a nucleobase sequence in SEQ ID NOs: 3-12 for hotspot 1 and SEQ ID NOs: 28-34 for hotspot 2, may be particularly useful in the treatment of Wilson disease associated with exon 6 skipping (e.g., Wilson disease caused by the M645R mutation).

[00195] Table 1

SEQ ID NO	Sequence
3	GTACTTGGTTAAAATATGCA
4	AGGAAGGTACTTGGTTAAAA
5	AAAAGGAGGAAGGTACTTGG
6	TGGGGAAAAAGGAGGAAGGT
7	AGGGGTGGGGAAAAAGGAGG
8	AAAGAGAGGGGTGGGGAAAA
9	CATTAAAAAGAGAGGGGTGG
10	CTTGTCATTAAAAAGAGAGG
11	AATTTCTTGTCATTAAAA
12	AAAGCCAATTTCTTGTCAT
13	AGCATGAAAGCCAATTCCT
14	AGGGAAGCATGAAAGCCAAT
15	TGGGCCAGGGAAGCATGAAA
16	TTTCTCTGGGCCAGGGAAGC
17	TGGGGTTTCTCTGGGCCAGG
18	GAGCGTTGGGGTTTCTCTGG
19	AGTGATGAGCGTTGGGGTTT
20	GGTCCAAGTGATGAGCGTTG
21	CTTGTGGTCCAAGTGATGAG
22	TTCCCTCTTGTTGGTCCAAGT
23	CTTTATTTCCCTCTTGTTGGT
24	TACTGCTTTATTTCCCTCTT
25	TCTACCTACTGCTTTATTTCT
26	TTGTGTTCTACCTACTGCTT
27	TATCTTTTGTGTTCTACCTA
28	GAGTTTATCTTTTGTGTTCT
29	GAGCTGGAGTTTATCTTTTG
30	AGATGAGAGCTGGAGTTTAT
31	GACTTAGATGAGAGCTGGAG
32	GGAAGGGACTTAGATGAGAG
33	GGTAGAGGAAGGGACTTAGA
34	GCCCAGGTAGAGGAAGGGAC

Example 2 Characterization of target regions (hot spots) for increasing inclusion of ATP7B exon 6

[00196] To explore the possible use of splice-switching oligonucleotides as a treatment for Wilson disease patients carrying the M645R variant, a hepatic cell line, HepG2, carrying this mutation, 2F3, was derived. The 2F3 line carries the M645R variant at one allele and a large plasmid insertion at the other. This insertion is assumed to render that allele unable to produce functional protein, thus creating a situation analogous to a compound heterozygote, consistent with what found in Wilson disease patients.

[00197] The oligonucleotides listed in Table 2 were transfected into 2F3 cells, and RT-PCR products were analyzed using capillary electrophoresis. These oligonucleotides were designed to hybridize to the above-identified hotspots and to expand the search for additional hotspots. Percent spliced in (PSI) for

exon 6 was then calculated as well as the change in percent spliced in compared to an inactive control antisense oligonucleotide (dPSI) (Table 2). As seen in Table 2, certain antisense oligonucleotides have a negative dPSI indicating an increase in exon 6 exclusion (e.g., SEQ ID NOs: 115-118) which is opposite of the intended effect. Antisense oligonucleotides which are targeted to intronic regions either side of exon 6 are effective showing positive dPSIs.

[00198] Certain oligonucleotides with high dPSIs were aligned to the reference genome chr13: 52535914-52536146[hg19/b37] (SEQ ID NOs: 28-33, 47-51, 60, 61, 63-67, 76-81, 93-95, 120, 121, 130, 132, 133-137, 147, and 150). As seen in FIG. 3, a third hotspot for antisense oligonucleotide binding to induce exon 6 inclusion in ATP7B M645R mutants was discovered spanning the area covered by antisense oligonucleotides having sequences set forth in SEQ ID NOs: 119-124 (genomic antisense: positions 54472-54516 in SEQ ID NO:1; chr13:52536115-52536159), in addition to the areas covered by antisense oligonucleotides having sequences set forth in SEQ ID NOs: 3-12 (genomic antisense: positions 54522-54593 in SEQ ID NO: 1; chr13:52536038-52536109) and SEQ ID NOs: 28-34 (genomic antisense: positions 54665-54718 in SEQ ID NO: 1; chr13:52535913-52535966). Three exemplary core sequences were also identified in these areas: SEQ ID NO: 160 (genomic sense: TTATCTTTT; genomic antisense: positions 54672-54680 in SEQ ID NO: 1), SEQ ID NO: 161 (genomic sense: GACTTAGATGA; genomic antisense: positions 54691-54701 in SEQ ID NO: 1), and SEQ ID NO: 162 (genomic sense: TTTCAGCTTTTGAAA; genomic antisense: positions 54492-54506).

[00199] Table 2

SEQ ID NO	PSI	sequence	Start Chr13 [hg19/b37]	End Chr13 [hg19/b37]	length	dPSI
29	0.94903506	GAGCTGGAGTTTATCTTTTG	52535941	52535960	20	0.491132
30	0.93643194	AGATGAGAGCTGGAGTTTAT	52535935	52535954	20	0.478529
35	0.70974518	AGATGAGAGCTGGAGT	52535935	52535950	16	0.251842
36	0.79611546	GATGAGAGCTGGAGTT	52535936	52535951	16	0.338213
37	0.80660668	ATGAGAGCTGGAGTTT	52535937	52535952	16	0.348704
38	0.62920358	TGAGAGCTGGAGTTTA	52535938	52535953	16	0.171301
39	1	GAGAGCTGGAGTTTAT	52535939	52535954	16	0.542097
40	0.82915652	AGAGCTGGAGTTTATC	52535940	52535955	16	0.371254
41	0.80242893	GAGCTGGAGTTTATCT	52535941	52535956	16	0.344526
42	0.93324554	AGCTGGAGTTTATCTT	52535942	52535957	16	0.475343
43	0.91806589	GCTGGAGTTTATCTTT	52535943	52535958	16	0.460163
44	0.92097701	CTGGAGTTTATCTTTT	52535944	52535959	16	0.463074
45	0.88136593	TGGAGTTTATCTTTTG	52535945	52535960	16	0.423463
46	0.92978378	GGAGTTTATCTTTTGT	52535946	52535961	16	0.471881
47	0.93878061	GAGTTTATCTTTTGTG	52535947	52535962	16	0.480878
48	0.95719947	AGTTTATCTTTTGTGT	52535948	52535963	16	0.499297
49	0.96989269	GTTTATCTTTTGTGTT	52535949	52535964	16	0.51199
50	0.95199346	TTTATCTTTTGTGTTT	52535950	52535965	16	0.494091
51	0.94531532	TTATCTTTTGTGTTCT	52535951	52535966	16	0.487412
52	0.57919401	AGATGAGAGCTGGAGTT	52535935	52535951	17	0.121291
53	0.42220616	GATGAGAGCTGGAGTTT	52535936	52535952	17	-0.0357

SEQ ID NO	PSI	sequence	Start Chr13 [hg19/b37]	End Chr13 [hg19/b37]	length	dPSI
54	0.54186296	ATGAGAGCTGGAGTTTA	52535937	52535953	17	0.08396
55	0.75146207	TGAGAGCTGGAGTTTAT	52535938	52535954	17	0.293559
56	0.83250872	GAGAGCTGGAGTTTATC	52535939	52535955	17	0.374606
57	0.84600061	AGAGCTGGAGTTTATCT	52535940	52535956	17	0.388098
58	0.90189881	GAGCTGGAGTTTATCTT	52535941	52535957	17	0.443996
59	0.85045831	AGCTGGAGTTTATCTTT	52535942	52535958	17	0.392555
60	0.94235165	GCTGGAGTTTATCTTTT	52535943	52535959	17	0.484449
61	0.96260732	CTGGAGTTTATCTTTTG	52535944	52535960	17	0.504704
62	0.93743667	TGGAGTTTATCTTTTGT	52535945	52535961	17	0.479534
63	0.96488622	GGAGTTTATCTTTTGTG	52535946	52535962	17	0.506983
64	0.95897848	GAGTTTATCTTTTGTGT	52535947	52535963	17	0.501076
65	0.97370995	AGTTTATCTTTTGTGTT	52535948	52535964	17	0.515807
66	0.96669576	GTTTATCTTTTGTGTTT	52535949	52535965	17	0.508793
67	0.94990072	TTTATCTTTTGTGTTCT	52535950	52535966	17	0.491998
68	0.41427269	AGATGAGAGCTGGAGTTT	52535935	52535952	18	-0.04363
69	0.29548406	GATGAGAGCTGGAGTTTA	52535936	52535953	18	-0.16242
70	0.65272521	ATGAGAGCTGGAGTTTAT	52535937	52535954	18	0.194822
71	0.82775724	TGAGAGCTGGAGTTTATC	52535938	52535955	18	0.369854
72	0.89482076	GAGAGCTGGAGTTTATCT	52535939	52535956	18	0.436918
73	0.89873087	AGAGCTGGAGTTTATCTT	52535940	52535957	18	0.440828
74	0.93420851	GAGCTGGAGTTTATCTTT	52535941	52535958	18	0.476306
75	0.88748876	AGCTGGAGTTTATCTTTT	52535942	52535959	18	0.429586
76	0.95705511	GCTGGAGTTTATCTTTTG	52535943	52535960	18	0.499152
77	0.95710864	CTGGAGTTTATCTTTTGT	52535944	52535961	18	0.499206
78	0.95718201	TGGAGTTTATCTTTTGTG	52535945	52535962	18	0.499279
79	0.95723276	GGAGTTTATCTTTTGTGT	52535946	52535963	18	0.49933
80	0.9682069	GAGTTTATCTTTTGTGTT	52535947	52535964	18	0.510304
81	0.96703305	AGTTTATCTTTTGTGTTT	52535948	52535965	18	0.50913
82	0.89944316	GTTTATCTTTTGTGTTCT	52535949	52535966	18	0.44154
83	0.35701069	AGATGAGAGCTGGAGTTTA	52535935	52535953	19	-0.10089
84	0.37085899	GATGAGAGCTGGAGTTTAT	52535936	52535954	19	-0.08704
85	0.81174544	ATGAGAGCTGGAGTTTATC	52535937	52535955	19	0.353843
86	0.86312479	TGAGAGCTGGAGTTTATCT	52535938	52535956	19	0.405222
87	0.92042588	GAGAGCTGGAGTTTATCTT	52535939	52535957	19	0.462523
88	0.94378406	AGAGCTGGAGTTTATCTTT	52535940	52535958	19	0.485881
89	0.90121758	GAGCTGGAGTTTATCTTTT	52535941	52535959	19	0.443315
90	0.97875323	AGCTGGAGTTTATCTTTTG	52535942	52535960	19	0.52085
91	0.97165722	GCTGGAGTTTATCTTTTGT	52535943	52535961	19	0.513754
92	0.974632	CTGGAGTTTATCTTTTGTG	52535944	52535962	19	0.516729
93	0.98143024	TGGAGTTTATCTTTTGTGT	52535945	52535963	19	0.523527
94	0.96598484	GGAGTTTATCTTTTGTGTT	52535946	52535964	19	0.508082

SEQ ID NO	PSI	sequence	Start Chr13 [hg19/b37]	End Chr13 [hg19/b37]	length	dPSI
95	0.96947187	GAGTTTATCTTTTGTGTTT	52535947	52535965	19	0.511569
96	0.51215069	AGTTTATCTTTTGTGTTCT	52535948	52535966	19	0.054248
97	0.43801448	GATGAGAGCTGGAGTTTATC	52535936	52535955	20	-0.01989
98	0.84294034	ATGAGAGCTGGAGTTTATCT	52535937	52535956	20	0.385037
99	0.91506349	TGAGAGCTGGAGTTTATCTT	52535938	52535957	20	0.457161
100	0.95571491	GAGAGCTGGAGTTTATCTTT	52535939	52535958	20	0.497812
101	0.9250529	AGAGCTGGAGTTTATCTTTT	52535940	52535959	20	0.46715
102	0.91454039	AGCTGGAGTTTATCTTTTGT	52535942	52535961	20	0.456638
103	0.95584009	GCTGGAGTTTATCTTTTGTG	52535943	52535962	20	0.497937
104	0.96866415	CTGGAGTTTATCTTTTGTGT	52535944	52535963	20	0.510761
105	0.50424102	GTTGGGCCAGGTAGAGGAA	52535908	52535927	20	0.046338
106	0.46400523	GCAGAGTTGGGCCAGGTAG	52535903	52535922	20	0.006102
107	0.55283098	AGCTGGCAGAGTTGGGCCCA	52535898	52535917	20	0.094928
108	0.70816205	AGACCAGCTGGCAGAGTTGG	52535893	52535912	20	0.250259
109	0.44154369	AGACAAGACCAGCTGGCAGA	52535888	52535907	20	-0.01636
110	0.44883783	TGGGAAGACAAGACCAGCTG	52535883	52535902	20	-0.00907
111	0.25330097	CACCATGGGAAGACAAGACC	52535878	52535897	20	-0.2046
112	0.1584088	GAAGGCACCATGGGAAGACA	52535873	52535892	20	-0.29949
113	0.050187	AGGAGGAAGGCACCATGGGA	52535868	52535887	20	-0.40772
114	0.05564216	AATCCAGGAGGAAGGCACCA	52535863	52535882	20	-0.40226
115	0.0583237	TGGTTAAATATGCATTGGC	52536095	52536114	20	-0.39958
116	0.0814945	AAAATATGCATTGGCAGAAA	52536100	52536119	20	-0.37641
117	0.18600916	ATGCATTGGCAGAAAGCACT	52536105	52536124	20	-0.27189
118	0.2612183	TTGGCAGAAAGCACTTTTCA	52536110	52536129	20	-0.19668
119	0.7320813	AGAAAGCACTTTTCAGCTTT	52536115	52536134	20	0.274178
120	0.92565362	GCACTTTTCAGCTTTTGAAA	52536120	52536139	20	0.467751
121	0.9502844	TTTCAGCTTTGGAAATTAGA	52536125	52536144	20	0.492382
122	0.96330809	GCTTTGGAAATTAGAAAGTG	52536130	52536149	20	0.505405
123	0.80993619	GGAAATTAGAAAGTGAATCT	52536135	52536154	20	0.352033
124	0.71516338	TTAGAAAGTGAATCTAAAG	52536140	52536159	20	0.257261
125	0.52494194	GGTAGAGGAAGGGACTTA	52535918	52535935	18	0.067039
126	0.6689587	GTAGAGGAAGGGACTTAG	52535919	52535936	18	0.211056
127	0.68622134	TAGAGGAAGGGACTTAGA	52535920	52535937	18	0.228319
128	0.78427472	AGAGGAAGGGACTTAGAT	52535921	52535938	18	0.326372
129	0.7981034	GAGGAAGGGACTTAGATG	52535922	52535939	18	0.340201
130	0.89063399	AGGAAGGGACTTAGATGA	52535923	52535940	18	0.432731
131	0.9141607	GGAAGGGACTTAGATGAG	52535924	52535941	18	0.456258
132	0.89012483	GAAGGGACTTAGATGAGA	52535925	52535942	18	0.432222
133	0.93330513	AAGGGACTTAGATGAGAG	52535926	52535943	18	0.475402
134	0.85689543	AGGGACTTAGATGAGAGC	52535927	52535944	18	0.398993
135	0.82924288	GGGACTTAGATGAGAGCT	52535928	52535945	18	0.37134

SEQ ID NO	PSI	sequence	Start Chr13 [hg19/b37]	End Chr13 [hg19/b37]	length	dPSI
136	0.89801786	GGACTTAGATGAGAGCTG	52535929	52535946	18	0.440115
137	0.92780754	GACTTAGATGAGAGCTGG	52535930	52535947	18	0.469905
138	0.78670155	ACTTAGATGAGAGCTGGA	52535931	52535948	18	0.328799
139	0.85607847	CTTAGATGAGAGCTGGAG	52535932	52535949	18	0.398176
140	0.8755492	TTAGATGAGAGCTGGAGT	52535933	52535950	18	0.417646
141	0.81911248	TAGATGAGAGCTGGAGTT	52535934	52535951	18	0.36121
142	0.65870755	GTAGAGGAAGGGACTTAGAT	52535919	52535938	20	0.200805
143	0.68963137	TAGAGGAAGGGACTTAGATG	52535920	52535939	20	0.231729
144	0.71236425	AGAGGAAGGGACTTAGATGA	52535921	52535940	20	0.254461
145	0.77181564	GAGGAAGGGACTTAGATGAG	52535922	52535941	20	0.313913
146	0.91591182	AGGAAGGGACTTAGATGAGA	52535923	52535942	20	0.458009
147	0.93671833	GAAGGGACTTAGATGAGAGC	52535925	52535944	20	0.478815
148	0.9264989	AAGGGACTTAGATGAGAGCT	52535926	52535945	20	0.468596
149	0.96644251	AGGGACTTAGATGAGAGCTG	52535927	52535946	20	0.50854
150	0.9477045	GGGACTTAGATGAGAGCTGG	52535928	52535947	20	0.489802
151	0.88244294	GGACTTAGATGAGAGCTGGA	52535929	52535948	20	0.42454
152	0.82302381	ACTTAGATGAGAGCTGGAGT	52535931	52535950	20	0.365121
153	0.68879408	CTTAGATGAGAGCTGGAGTT	52535932	52535951	20	0.230891
154	0.61269452	TTAGATGAGAGCTGGAGTTT	52535933	52535952	20	0.154792
155	0.50052141	TAGATGAGAGCTGGAGTTTA	52535934	52535953	20	0.042619

Example 4 Treatment of 2F3 cells with a splice modulating antisense oligonucleotide increases protein level and copper tolerance

[00200] To model the effectiveness of an exon 6 inclusion inducing antisense oligonucleotide as a potential treatment for Wilson disease an *in vitro* copper sensitivity assay has been used. As seen in FIG. 4A, the M645R mutation reduces copper tolerance in 2F3 cells, mirroring Wilson disease where cells are unable to process copper as a result of ATP7B mutations. Transfection with an antisense oligonucleotide having the sequence set forth in SEQ ID 29 increases copper tolerance in 2F3 cells demonstrating that the copper sensitive phenotype can be at least partially rescued by transfection with an oligonucleotide shown herein to increase the inclusion of exon 6. Western blots against the ATP7B protein are consistent with these results, as 2F3 cells produce markedly less ATP7B protein than wild-type HepG2 cells (FIG. 4B). Transfection with an antisense oligonucleotide having the sequence set forth in SEQ ID 29 partially rescues this phenotype and increases ATP7B protein levels in 2F3 cells. This increase in ATP7B protein levels and function demonstrates that treatment with an antisense oligonucleotide known to increase the inclusion of exon 6 partially rescues the Wilson disease phenotype in an *in vitro* model.

OTHER EMBODIMENTS

[00201] Various modifications and variations of the described invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the

described modes for carrying out the invention that are obvious to those skilled in the art are intended to be within the scope of the invention.

WHAT IS CLAIMED IS:

CLAIMS

1. An antisense oligonucleotide comprising a nucleobase sequence at least 70% complementary to an ATP7B target sequence in exon 6, a 5'-flanking intron, a 3'-flanking intron, or a combination of exon 6 and the 5'-flanking or 3'-flanking intron.
2. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence reduces the binding of a splicing factor to an intronic splicing silencer in the 5'-flanking or 3'-flanking intron.
3. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence comprises at least one nucleotide located among positions 54672-54680 in SEQ ID NO: 1.
4. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence comprises at least one nucleotide located among positions 54691-54701 in SEQ ID NO: 1.
5. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence comprises at least one nucleotide located among positions 54492-54506 in SEQ ID NO: 1.
6. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence comprises at least one nucleotide located among positions 54472-54516, 54522-54593, and 54665-54718 in SEQ ID NO: 1.
7. The antisense oligonucleotide of claim 1, wherein the nucleobase sequence is complementary to a sequence within the 5'-flanking intron.
8. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence is located within the 5'-flanking intron among positions up to 54517 in SEQ ID NO: 1.
9. The antisense oligonucleotide of claim 1, wherein the nucleobase sequence has at least 70% sequence identity to SEQ ID NO: 119, 120, 121, 122, 123, or 124.
10. The antisense oligonucleotide of claim 1, wherein the nucleobase sequence has at least 70% sequence identity to SEQ ID NO: 122.
11. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence is located within the 5'-flanking intron among positions 54522 to 54581 in SEQ ID NO: 1.
12. The antisense oligonucleotide of claim 1, wherein the nucleobase sequence has at least 70% sequence identity to SEQ ID NO: 3, 4, 5, 6, 7, 8, 9, or 10.
13. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence is located within the combination of the 5'-flanking intron and exon 6.

14. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence is located within the combination of the 5'-flanking intron and exon 6 among positions 54562 to 54593 in SEQ ID NO: 1.
15. The antisense oligonucleotide of claim 1, wherein the nucleobase sequence has at least 70% sequence identity to SEQ ID NO: 11.
16. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence is located within exon 6 or the combination of the 3'-flanking intron and exon 6.
17. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence is located among positions 54631 to 54677 in SEQ ID NO: 1.
18. The antisense oligonucleotide of claim 1, wherein the nucleobase sequence has at least 70% sequence identity to SEQ ID NO: 22, 23, 24, or 25.
19. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence is located within the 3'-flanking intron.
20. The antisense oligonucleotide of claim 1, wherein the ATP7B target sequence is located among positions 54655 to 54738 in SEQ ID NO: 1.
21. The antisense oligonucleotide of claim 5, 6, 19, or 20, wherein the 5'-terminal nucleotide of the oligonucleotide is complementary to neither position 54695 nor position 54696 of SEQ ID NO: 1.
22. The antisense oligonucleotide of claim 1, wherein the nucleobase sequence has at least 70% sequence identity to SEQ ID NO: 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, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, or 155.
23. The antisense oligonucleotide of claim 1, wherein the nucleobase sequence has at least 70% sequence identity to SEQ ID NOs: 39, 48, 49, 50, 61, 63, 64, 65, 66, 67, 76, 77, 78, 79, 80, 81, 90, 91, 92, 93, 94, 95, 100, 103, 104, 121, and 149.
24. The antisense oligonucleotide of claim 9, 10, 12, 15, 18, 22, or 23, wherein the sequence identity is at least 90%.
25. The antisense oligonucleotide of claim 24, wherein the sequence identity is 100%.
26. The antisense oligonucleotide of any one of claims 1 to 25, wherein the antisense oligonucleotide comprises at least one modified nucleobase.

27. The antisense oligonucleotide of any one of claims 1 to 26, wherein the antisense oligonucleotide comprises at least one modified internucleoside linkage.
28. The antisense oligonucleotide of claim 27, wherein the modified internucleoside linkage is a phosphorothioate linkage.
29. The antisense oligonucleotide of claim 28, wherein the phosphorothioate linkage is a stereochemically enriched phosphorothioate linkage.
30. The antisense oligonucleotide of any one of claims 27 to 29, wherein at least 50% of internucleoside linkages in the antisense oligonucleotide are independently the modified internucleoside linkage.
31. The antisense oligonucleotide of claim 30, wherein at least 70% of internucleoside linkages in the antisense oligonucleotide are independently the modified internucleoside linkage.
32. The antisense oligonucleotide of claim 31, wherein all internucleoside linkages in the antisense oligonucleotide are independently the modified internucleoside linkage.
33. The antisense oligonucleotide of any one of claims 1 to 32, wherein the antisense oligonucleotide comprises at least one modified sugar nucleoside.
34. The antisense oligonucleotide of claim 33, wherein at least one modified sugar nucleoside is a 2'-modified sugar nucleoside.
35. The antisense oligonucleotide of claim 34, wherein at least one 2'-modified sugar nucleoside comprises a 2'-modification selected from the group consisting of 2'-fluoro, 2'-methoxy, and 2'-methoxyethoxy.
36. The antisense oligonucleotide of claim 35, wherein the 2'-modified sugar nucleoside comprises the 2'-methoxyethoxy modification.
37. The antisense oligonucleotide of any one of claims 33 to 36, wherein at least one modified sugar nucleoside is a bridged nucleic acid.
38. The antisense oligonucleotide of claim 37, wherein the bridged nucleic acid is a locked nucleic acid (LNA), ethylene-bridged nucleic acid (ENA), or cEt nucleic acid.
39. The antisense oligonucleotide of any one of claims 33 to 38, wherein all nucleosides in the antisense oligonucleotide are independently the modified sugar nucleosides.

40. The antisense oligonucleotide of any one of claims 1 to 26, wherein the antisense oligonucleotide is a morpholino oligomer.
41. The antisense oligonucleotide of any one of claims 1 to 40, further comprising a targeting moiety.
42. The antisense oligonucleotide of claim 41, wherein the targeting moiety is covalently conjugated at the 5'-terminus of the antisense oligonucleotide.
43. The antisense oligonucleotide of claim 41, wherein the targeting moiety is covalently conjugated at the 3'-terminus of the antisense oligonucleotide.
44. The antisense oligonucleotide of claim 41, wherein the targeting moiety is covalently conjugated at an internucleoside linkage of the antisense oligonucleotide.
45. The antisense oligonucleotide of any one of claims 42 to 44, wherein the targeting moiety is covalently conjugated through a linker.
46. The antisense oligonucleotide of claim 45, wherein the linker is a cleavable linker.
47. The antisense oligonucleotide of any one of claims 42 to 46, wherein the targeting moiety comprises N-acetylgalactosamine.
48. The antisense oligonucleotide of claim 47, wherein the targeting moiety is an N-acetylgalactosamine cluster.
49. The antisense oligonucleotide of any one of claims 1 to 48, wherein the antisense oligonucleotide comprises at least 12 nucleosides.
50. The antisense oligonucleotide of claim 49, wherein the antisense oligonucleotide comprises at least 16 nucleosides.
51. The antisense oligonucleotide of any one of claims 1 to 50, wherein the antisense oligonucleotide comprises a total of 50 nucleosides or fewer.
52. The antisense oligonucleotide of any one of claims 1 to 50, wherein the antisense oligonucleotide comprises a total of 30 nucleosides or fewer.
53. The antisense oligonucleotide of any one of claims 1 to 50, wherein the antisense oligonucleotide comprises a total of 20 nucleosides or fewer.
54. The antisense oligonucleotide of any one of claims 1 to 48, wherein the antisense oligonucleotide comprises a total of 16 to 20 nucleosides.

55. The antisense oligonucleotide of any one of claims 1 to 48, wherein the antisense oligonucleotide comprises a total of 16 to 19 nucleosides.
56. A pharmaceutical composition comprising the antisense oligonucleotide of any one of claims 1 to 55 and a pharmaceutically acceptable excipient.
57. A method of increasing the level of exon 6-containing ATP7B mRNA molecules in a cell expressing an aberrant ATP7B gene, the method comprising contacting the cell with the antisense oligonucleotide of any one of claims 1 to 55.
58. The method of claim 57, wherein the cell is in a subject.
59. A method of treating Wilson disease in a subject having an aberrant ATP7B gene, the method comprising administering a therapeutically effective amount of the antisense oligonucleotide of any one of claims 1 to 55 or the pharmaceutical composition of claim 56 to the subject in need thereof.
60. The method of claim 59, wherein the administering step is performed parenterally.
61. The method of any one of claims 57 to 60, wherein the aberrant ATP7B gene is ATP7B having a g.54646T>G mutation in SEQ ID NO: 1.

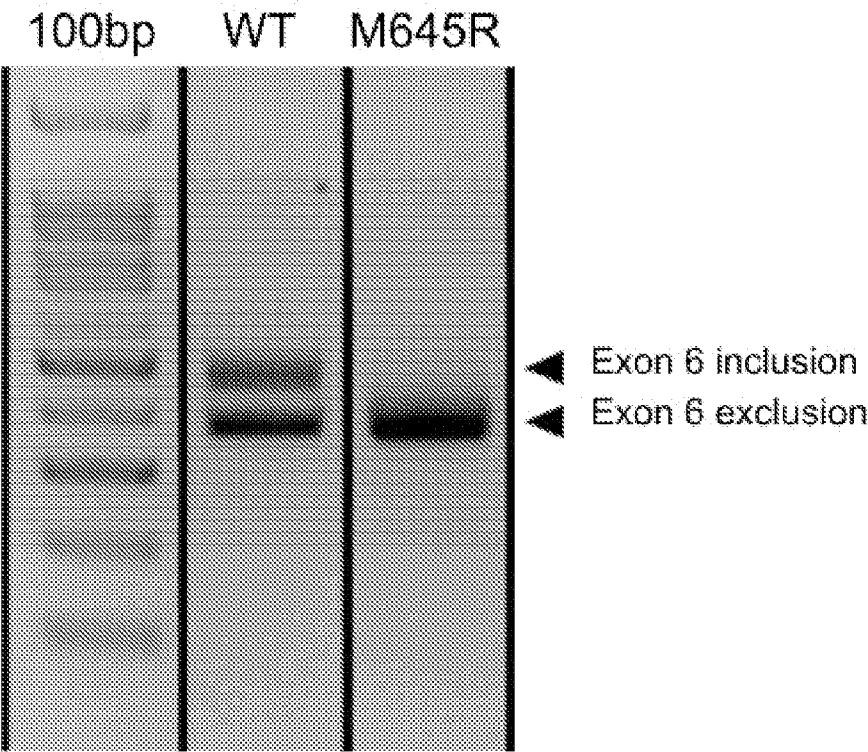


FIG. 1A

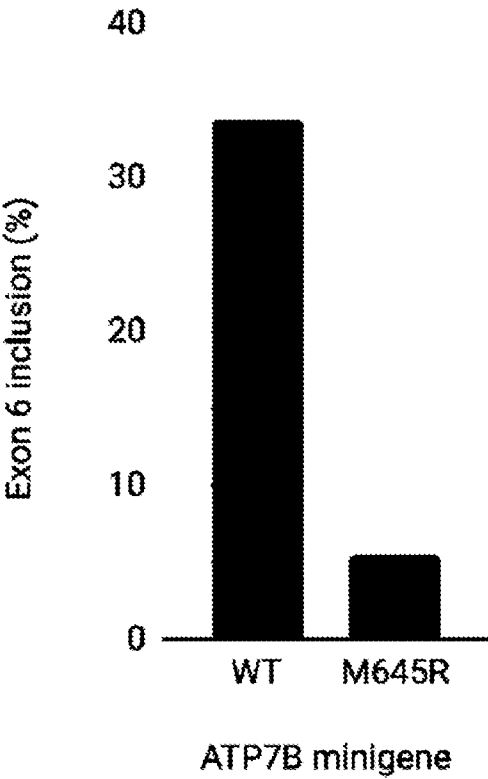


FIG. 1B

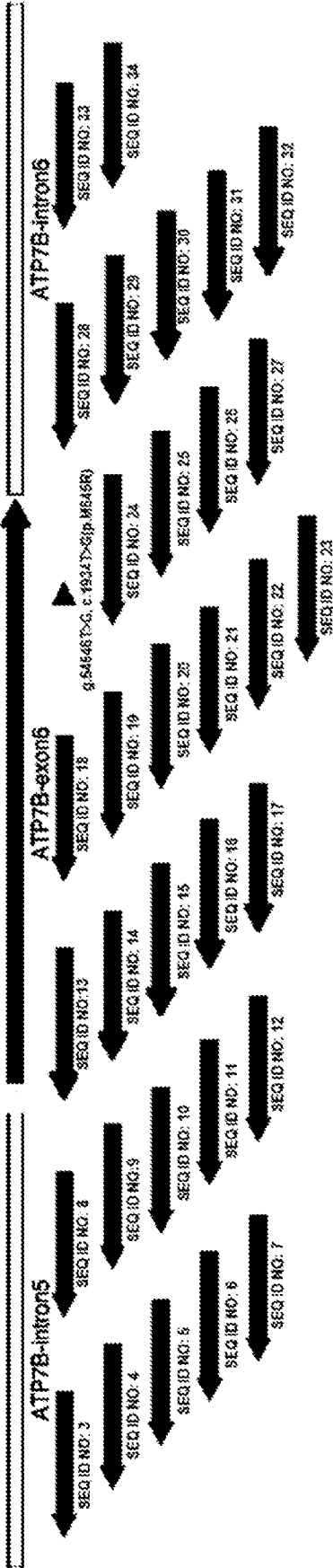


FIG. 2A

SEQ ID NO: 3	SEQ ID NO: 4	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 9	SEQ ID NO: 10	SEQ ID NO: 11	SEQ ID NO: 12	SEQ ID NO: 13	SEQ ID NO: 14	SEQ ID NO: 15	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 18	SEQ ID NO: 19	SEQ ID NO: 20	SEQ ID NO: 21	SEQ ID NO: 22	SEQ ID NO: 23	SEQ ID NO: 24	SEQ ID NO: 25	SEQ ID NO: 26	SEQ ID NO: 27	SEQ ID NO: 28	SEQ ID NO: 29	SEQ ID NO: 30	SEQ ID NO: 31	SEQ ID NO: 32	SEQ ID NO: 33	SEQ ID NO: 34
--------------	--------------	--------------	--------------	--------------	--------------	--------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------

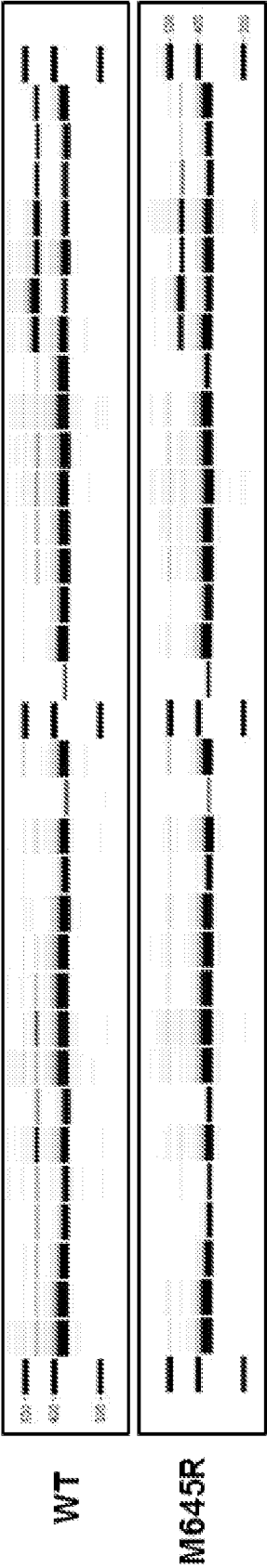


FIG. 2B

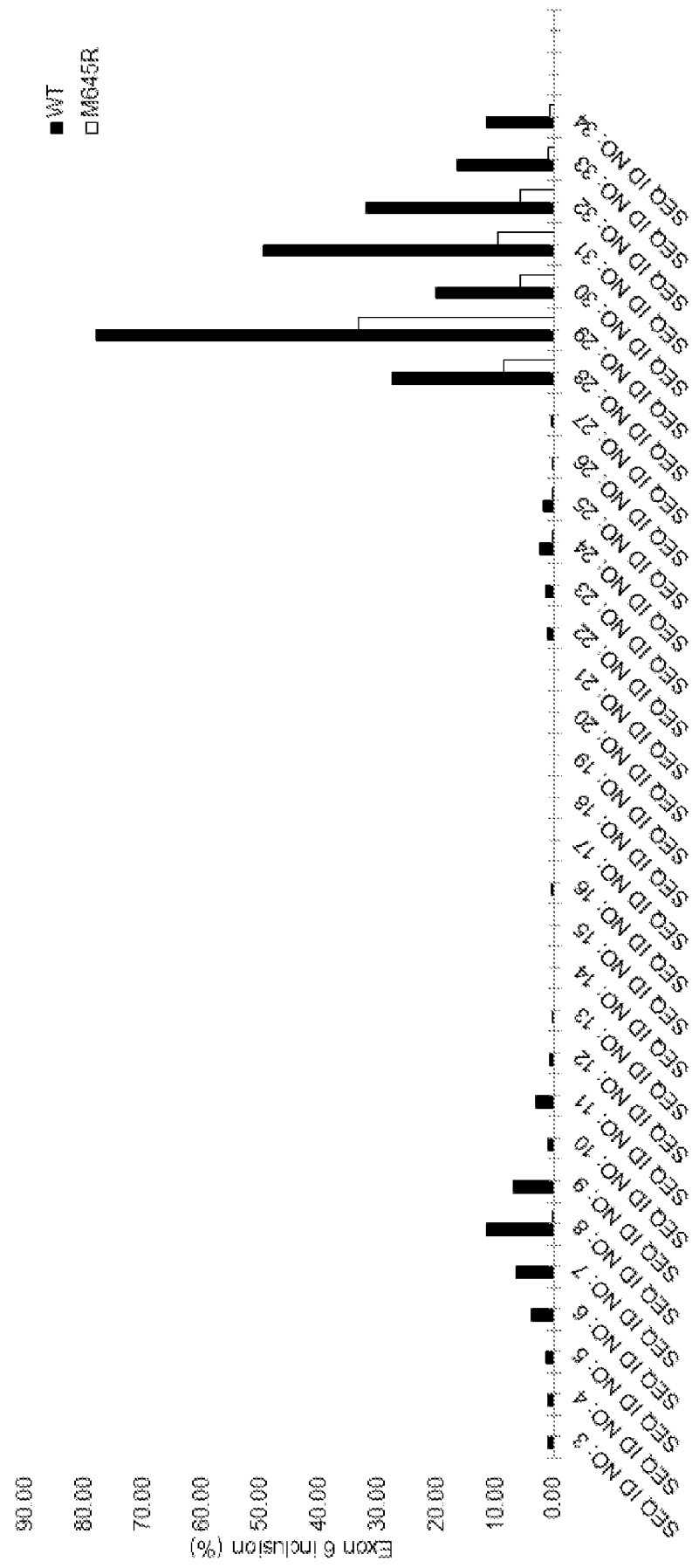


FIG. 2C

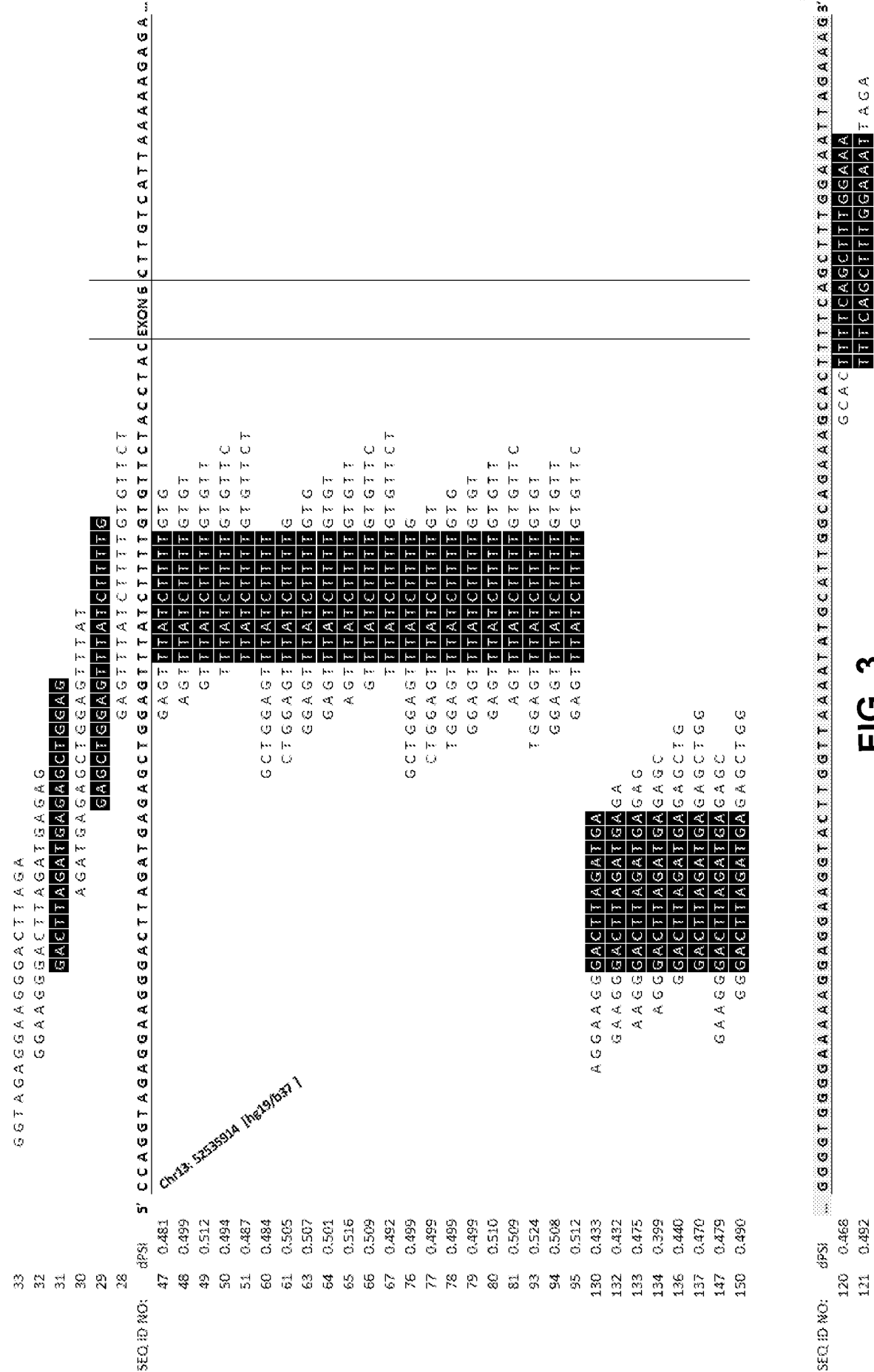


FIG. 3

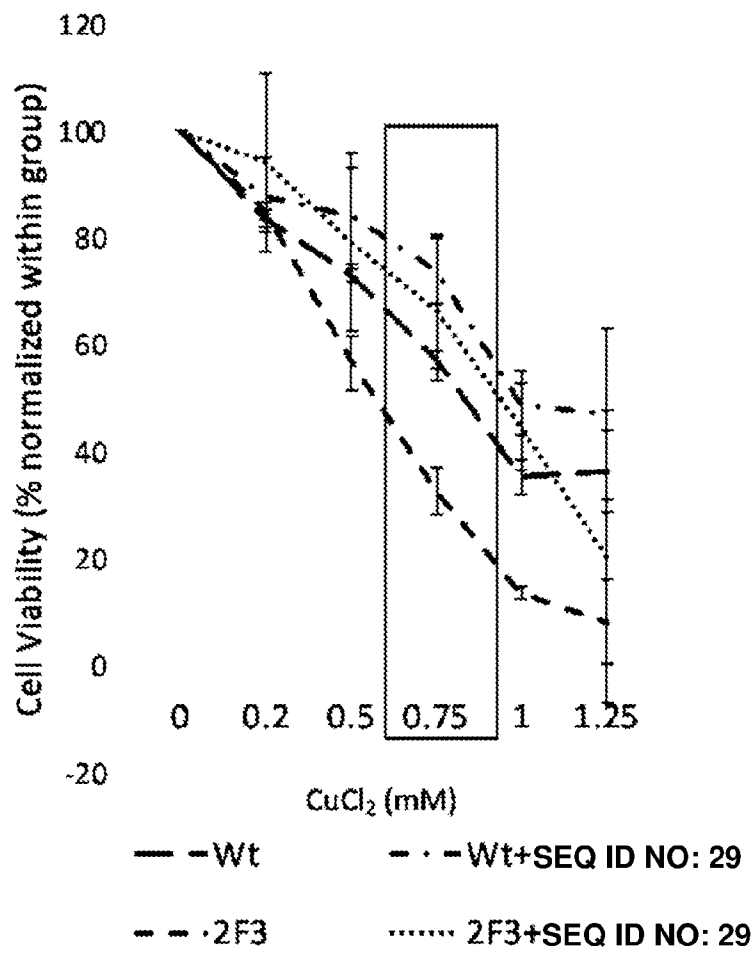


FIG. 4A

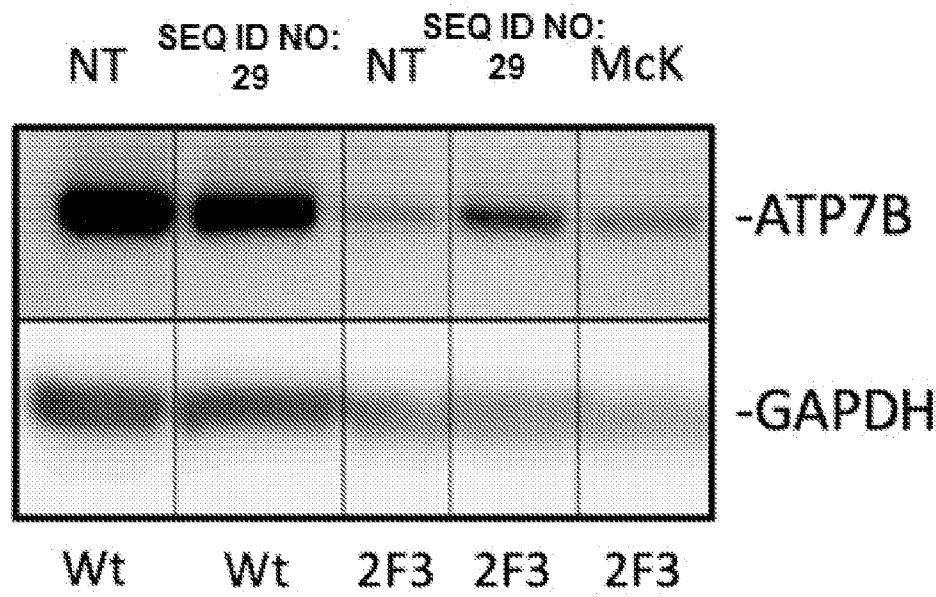


FIG. 4B

SEQUENCE LISTING

<110> Deep Genomics Incorporated

<120> OLIGONUCLEOTIDE THERAPY FOR WILSON DISEASE

<130> 51359-003W02

<150> US 62/630,565

<151> 2018-02-14

<160> 170

<170> PatentIn version 3.5

<210> 1

<211> 85826

<212> DNA

<213> Homo sapiens

<400> 1

cacaaaacca ctttttcctc ctgggcctcc aggcctgcc tggagagaggc tgccatgaag
60

gtctccgaca ttgcctggag acattttccc ccatgggtctt caggattaac attaggctcc
120

ttgctacttt tgaaaatttc tgcagccagc ttgaatttct cctaaaaaaaa tgggtctttc
180

ttttctactg cattgtcagg ctgcaaattt tctgaacttt tatgctctgt tccgctttta
240

aaatggaatg cttttaagag cacccaagtc acttcttgaa ttctttgctg cttagaaatt
300

tcttacacca gataccctaa atcatctctt tcaagttcaa agttccacaa atctctaggg
360

caggggcaaa atgcctctag tacctttgct aaaacataag aggagtcacc tttctccagt
420

tcccaacaag ttcttcatct ccatctgaga ccacctcagc ctggatttcg ttgtccatat
480

cactatccgc attttggttca aagccattca acaagtctct aggatgttcc aaactttccc
540

acattttcct ggcttcttct gagccctcca aactgttcca gcctctgcct gttgccaggt
600

tccaaagtcg cttccacatt tttgggtatc tcttcagcag caccacctc tactggtact
660

aatttatggg accaagttca ttttcacgct gctgataaag acatactcga aactgggaac
720

aaaaataggt ttaattggat ttacagttcc acatggctat gaggcctcag aatcatggcg
780

ggaggcataa ggcacttctt acatggtggc agcaagagaa aaatcagaaa gaagcaaaag
840

cagaaacccc tgataaaccc catcagatct cgtgaactta ctcactatca cgagaatagc
900

atgggaaaga ttggccccc a tgattcaatt acctccccct ggatccctcc cacaacatgt
960

gggcattctg ggagataaaa ttcaacttga ggtttgggtg gggtcacggc caaacatat
1020

cacatagttt tcaactgcca catgaaaaca tacaaattca tcagaaaagt aaattttgct
1080

tttttttagat acaggggtctc actctgtcac ccaggctgaa gtgcagtggt gcaatcatag
1140

caatcactgc agccccaaac tccttggctc aagcgatcct cctgcctcag actcctgaga
1200

agctggaact acaggtgcat gccaccacac ccagctaatt aaaaaaatgt tttttgtaga
1260

aatagggttt tgctatggtg cccaggctgg tctcaaactc ttggcctcaa gtaatcctcc
1320

tgcttggcc ttccaaagca ttgagattac aggcattgagc tgctgcacct ggccagaagt
1380

aaatatattta ttgatattta gttttgacat aaagttctca tgtgggattt cctattgtga
1440

acaaatcctc atttggttcct attgtgaaca aatctcatgt gggatttcct actgtgtact
1500

tgaagaggta tttgactcta tttgaaatcc cacatgagca ctttatttca aagctgaaat
1560

aattggtcta aaataatttg gtctttatat tctttataca cctttatact cttaaaaatt
1620

ttactgagga tacaatagag cttttgttta catcagagtt tctcatcctt gacactactg
1680

acattttgga ccaaataaatt atttgttgcg gaggttgtcc tatgtattat atttagtagc
1740

atctttgggc ttgaccact tagatgccag tagcaccctc ccaatcaaaa ttatcagttg
1800

ggcatggtgg ctcatgcctg taatcccaac attttgggag gctgaggtgg gaggattgct
1860

tgagcccagg agttcaaggc tgtagtaagc tatgatcatg ccactgcact ccagcctggg
1920

caacagagca agactcttat ctctaaaaaa aaaattaaaa aaaaaaagcc ctccagatat
1980

tgccaaatgc caaatgtctt gcgaagggca aaaatcaccc ctggttgaga accagtgggt
2040

tacactgctt atatctacca atatttagca tattagaaat gaaaactaaa gaactttaaa
2100

ctttttatga attcatttga aaataaaaat tctttcataa cacattaaca tatttttatt
2160

aaaaaactat cttttccaaa ccaaaaaaag aaatttagca agaaggggtgc caatgattta
2220

cactttttgc aaatcttttt aatgtctggc ttaacgtgaa accagatagt ggctcctcct
2280

ttctgtttct gcattcaatc tagtgatatg ttgtttttgt tgaagtatat gaagaaaatt
2340

caggctaact tagacatgta gtaggaaaag ggagtatatt ttaatatata ttgatcttat
2400

gtttaaaata tatttaactt ttcagctagt tacagatatt cttctctgat gctagaccac
2460

atttgacaag ttatgcagtt tcttaatgtc agttgcaata cagactctga aaccctatca
2520

atgacctttt cacactgtta attaaaatcc actatcactt agtttttatcc gctatcactt
2580

ttgcacttat caacatagta aaataggcaa atattacatt aatgttatta taaaaacaga
2640

tttgttgagg atccattgaa gggccttggg acctgcatgg tccatggagc acaatttgag
2700

aaccactgcc ttaaaggatg aacagagtct gacaaaagtg gtaagattta tcatttatct
2760

aacatttgga atatatgaat tcataaaagc aacactacag aggacagttt gctttctcag
2820

gacactcttt caaaagctta tttaaaatgc agctgaccca tataatgtaa aggtttttct
2880

ggcaattcct tcaaataatt taggtgggaa aattcatgtg acaagctttt caaactgaac
2940

tattaaacat ttaacttaaa agctgcataa taactaaagt taaacctgtc ctttaagtct
3000

ggttagtgag tttatttaaa cattataatg taaagcactg aggaatatta cgtatgacag
3060

ggaagaacca aaaggctctg ataacagact ttctcattca aaatgtttta agtaatgatt
3120

acagaatttg aagtagagtg gacaaaaacg aagaagaaat gcctgaactt ttgaaagaat
3180

taggtatata	ctgctattca	ggaagctatg	gtgttaatta	taaattttca	gcctgaacaa
3240					
aaatatgtag	gtggtaaaaa	aactaggtag	gtagaatata	ctaggccttc	tgttttatta
3300					
gtggagggca	aggcatttaa	aatacagtag	tactacaaca	acaaagggtgg	aaaaaagcct
3360					
gggagttttt	tcaccaaaat	gttaacagtg	gctatctccg	gatggtagca	ttcctggggg
3420					
tttttccttt	tacttatcta	ttttctaatt	tttctagaag	aaacagcata	tacacgtaat
3480					
acaaataaaa	taatatttta	tgtaacttct	tacatcataa	aaatttaaagt	aacagattat
3540					
caaaaagcag	aggattcaac	ataagcttta	tatttaagtg	acgtgttaac	aatggcagtt
3600					
ttggtgagtt	taagagtaga	aactaagatg	agaaggaaga	tccggtttct	tctttaccga
3660					
aaagagacat	gatagattgg	aaaatgtctc	gtggctcaat	cgtaaagaaa	aataaggaat
3720					
gaatatgaat	aatttatgac	agactaactg	gagtgtgaaa	gggaaaaaaaa	aagctgaagc
3780					
agggaaggga	actcgacccc	tcggccaaca	gtgaagcaga	agaaaacgcg	gtaaattgag
3840					
tctttaagac	ctagtctatt	ccagtctata	tagaagaccc	ccaagaaaaa	gaaagactgc
3900					
tcattggcct	gagaaattac	cttagattca	tgacaaggaa	ggccatttgc	ccgcaaaatt
3960					
tagctacact	ggacggggcaa	gtacccttac	agaagagaaa	acgtctgtga	gcccacacga
4020					
ccggctgctc	acctcaacaa	cttgcacagg	caccagctcc	tttcgccggc	cgccatcttc
4080					

cgccgacccc cgaactcagg aaacgcttca ctttcctttt ccctattggc tcctgagaaa
4140

gcaagccgtg ctcgccccgc cccacgggc caattgtgcg ttactattgg ttactggtag
4200

ccgcttccca cggccttcca gccaatagaa tatgccgagg cgtagactag tggtcggcgt
4260

ggcgcacacg gctcccgccc ccgtgggcgg gacagcagtg ggggggttggg ctgaggaggg
4320

cgtggcctgt gattgacagc cgtcgtccc tccctcggcc acctcccca ctagaagccc
4380

ccgcctgggc gcctgcgccc ccgttcccgg cccaaagccc gccgcccgtt ggaggccatt
4440

ggctggcctt tgcgcacagc ggatcgattt tccaggtgcg gagttcactc ttgccgcggt
4500

tgcttccttt gggacccacg gcgtccggca gccaggcgca gagtccgagg agggggcagc
4560

gcagagcgga cccgacgcgg cgccgccggg caccttcccc gcaggcggtg ggtgagccct
4620

gggagctgag tctgcggccc ggctctgcgc agctcacctg ccctcccgtt cccgcacacg
4680

cgtgagatcc cagtacagtg tcggagcgca ccagcgcgag gtggccgaga ccgcggagga
4740

ggacaggcct ccgccctgcg gcgccggcac ggcagaggac attgtggcac tggcacggca
4800

gagaacactg tggcaccggc ggggccggca gttccagggt gggcactccc agccacctgg
4860

ggagtgggcg agggtcagag gccactctc ccctcacgtt ctcacccccg tgccccagg
4920

tcgggaggac ggcggcgcgc aactttgaat catccgtgtg aagagggctg cggcttcccc
4980

ggtcccaaatt gaaggggagg ttcccggacc cctgttttgct ttagagccga gccgcgccga
5040

tgccctcaca ctctgcgcct cctctcccgg gactttaaca ccccgctctc ctccaccgac
5100

caggtgacct tttgctctga gccagatcag agaagaattc ggtgtccgtg cgggacgatg
5160

cctgagcagg agagacagat cacagccaga gaaggggcca gtcggaaagt gagttttggt
5220

ccccgcgtc cgcgagcgt gctcactccc accaggagga ttttctccc cgcgttcggc
5280

gcttactcgc cccagggggg tgcgcaggga aggccgagcc agaccccagc tccagggatg
5340

tcttggcgtg ggagaaaagg accctttaga aaaagtgtgc cgtcaactga tgtttgggag
5400

caatggaaaa cccgcacct caaaattcat ctttatgctt ttaaggggtc tgatgctgga
5460

gagataaagt gagcgtcgag ttgctgtgct agagccctaa tggctcaaaa cgatgaagtt
5520

agttacctgg agaagttact cagagcctgg tgcttggacc agatatggta tatcgaggtc
5580

ctggaacttt tataactggt ttaaattcat ttttaattgta caagtaatac attaagacag
5640

tcttttaaaa atggaaacag cacagaaata ttttgcacaa aaagcagatc tatttttcag
5700

tctttttcta ttcacgtaca tatgtcttca ttcggcaaatt atttattgag aactttctat
5760

gtttcaggca cacagtgttg agttgataca gtggttaaca agacatgtat ctatatgtaa
5820

tatgcacaca ttcagagaaa tctttagggt tcatttgtat gaggtgttaa aatttttgtt
5880

ttacctaaat gagatcactg tatgtgcatt ctcttttttc agtttaacga tatatcttag
5940

agattgtttc actaactgtc taattcagtg caacctctgc acagtttgga cgcatcatag
6000

tttattcatg gtttcctagt ggtggacggt ctgtttgcta ctttttttct ctctctcttg
6060

caagtgaaca tctttgtaca tgatttttaa tgctgacact atttcttttag gatagatact
6120

taggcctgga ttgctggaca gaacagtagg ctgtttcttt tccacagatt tctgaaacta
6180

attcactgac gttgctgggt gcccttttagt agtgaattca tctaggctcc ggtgatacag
6240

aaatgaaaac taacggatta gtggacagac agagagagggc tggttagtag gtggtttgta
6300

gggtatgatg actatgcatt tgataggagt atgaagaaat gttgggaggg gtggcttgta
6360

cagaggagat ttcaacttcc tacagtgtcc ccacctagag aaagaaatca ttatccttcc
6420

aaatcaagct caagacctgc cacctccttg aaacgttctg taactgctcc tgcccatggt
6480

gacctgtctt ctggaatcct ctgataatag cgttgaccag cgttattgag cttactgtgt
6540

gctgggcatt gtgttaggca ctcatatccc tctatgtgca ctctcgcggt taatcctcac
6600

aaaaaccctt tgatgtaggt attatatctt gttatcaaca tttattcaac aaattttttt
6660

gtgtgtacta tgtgctcaaa cacatagtac tcaaagtttt agacactggg gatttaggta
6720

ggaccgagac cttgccctca tgaatttaca tcctagttgg agataggaaa aaaaaagggg
6780

taaacaaatg 6840	tatagcgtag	gagttaaggg	aaaacttttg	tgccccgtga	agcgtcactg
aaaaatcaac 6900	tgacaaatgg	cagattaata	ggagagaagg	catacacatt	tactaacgtg
ccgtgagagt 6960	cttacaacat	aagatctcaa	aggaaggcaa	ggtggttgcc	acaattatac
tatcctgagg 7020	ttacagaaag	tgtgggggct	ctaagcttgg	ccaaaaacag	gttatggttg
taaatcaggt 7080	tataggaggg	agagaagagg	agacctggct	agcaaagggtg	gtcttatgta
gatgaaacct 7140	cacaggtagc	agccctcagg	gagaagaaat	gtttcagacc	ttgaagggtgc
caaattcagc 7200	taatctcacc	cagatccaga	caaggaaggg	ccctctgaga	aaaactggct
gcgtcaatgg 7260	aggttctcaa	cagatgctca	tctccctcac	agtaggacta	cttttgtctg
taggccctct 7320	gaacggccat	ctcaaaatat	ttccaataag	tctgtttttga	ggtgaaatgt
tttagtttcc 7380	ttcaatagct	ttatttcagg	tactagtaag	tactatgaag	aaaataaagc
agtttgggtc 7440	caggagttca	agaccagcct	ggcagcatat	caagacatcg	tttatacaaa
aaaaaaagaa 7500	aaattagctg	ggcataatgg	tgtgtgcctg	tcgaaccagc	tactcgggag
gctgaggcag 7560	gaggatggct	tgagcccagg	agtttgaggc	tacagtgagt	tgtgatccca
ccactgcact 7620	tcagcctgag	tgataaagca	agatcctgtc	tcaaaaaaaaa	agaaaagaaa
agaaagcaga 7680	ataagggaat	aagatagggg	ttgacatggg	tggagaggaa	gatgctgttt

taagtaggtg gtcaggggaat gatttctctt ggaaggtagg atttgagtag agacctgaat
7740

gattcagatt tcagagaagt agcctagggt gagaaacat aaagtgtgag agtcctacgg
7800

cagtcattgtg ataagaagaa atatatactt ggtcttgtcc caagttcctg acatagagct
7860

tcttaaaccg ttgtaatttc ctgagagatg gtgataagag catcttttgt tattcataac
7920

aagtctcttt caccctacc tgagtctctg ttaatgagag ggctcttgat aggcccctag
7980

atagcttcag gatggagact gattgccaga agaaccaacc atgtgattag gggattcgaa
8040

tggtcagccc cgtgctccaa cctctggaga ttgtgccaat caccagtga caatgatcta
8100

atcagtgatg ccggcagaat agagctttct ttaaattttt atttatttag ttttttgaga
8160

ggtagtctca ctctgttgct aggctggagt gcagtggcgg gatctcggtt cactgcaacc
8220

tgtgcctccc gggctcaagt gattctccgg cctcagcctc ctgagtagct gggattacag
8280

gcgcacgcca ccacaccag ctaatttttg tatttttagt ggagacgggg tttcaccatg
8340

ttggtcaggg tcatctctat ctcttgatct cgtgatccgc ccacctcggc ctcccaaagt
8400

gttggaatta caggcgtgag gccactgtgc ctggcctttt tttttttaac caatgcagtg
8460

gatcacaaat agagctttct taaaaaccct aaatgatggg ttttgagag cttcctgtta
8520

gtgaacacat caaggtgctg ggaaggtggt tcctggaaag gatatggaag ctcctggaaa
8580

cctccattcc catatcttgc cctatctttt ccatgtggtg gtttcttagc tgtatcttgt
8640

ataataaact ggtgatatta aataaagtgc tttcctaagt tctgcgagct gttctagcaa
8700

attaccaac ctgaggaggt aattgtagga acctcctgac tttacagttg gtcaggagta
8760

cagatggccc agaccagcg actggcatgt gaagtgtggg cattcttggtg gtcttggtggg
8820

cctgagccct tacctgtggg gtccgtgcta actccagatg gttggtgtcg gaattgaatt
8880

gaattgtagg gcaccagtt agtgtctgga gagagagcat tggttggtgt gaagaaaaac
8940

cccacacatt tgggtgtcaga agtgtcgtga gtggaaacag tacagaagac atgatgttag
9000

agtttgggga acaataggaa agccagagtg gctgtggcag agggaaggag tgggagtgat
9060

agaacaggag atcagagttc atcattgctt tacagatggg gaaactgagc cagacacaag
9120

ttaagtaact tgcccgatgt accttgctta taagaggcag atctaggatt gagtcctagc
9180

caggcctgcc tgactttctc agcgattttt tattgtcaaa acttacttga cacttggtca
9240

taccacctgg tatgattatt taatgaatat agcttttatt tgcaactctg tgcttcttga
9300

aatgggaacc gggtttttac cgtttaaaaa actatttccc acatgtaggt gcaccacagt
9360

tgttcttgaa tgaatggatt ctcgacatga tttctccaca ttgttttctc ttcagaggcc
9420

tctctgtggg gcagcagtga cctggtcatt cagccttcag tccacaaaca ttatggagca
9480

ctttctgttt gcctggtact gggccaagca gggggatgca tagatgaaga gaggtcagcc
9540

cctgtgctgg atgcatgctg ggacagagct gtctacatgg tacagcaccc acagaagggg
9600

cagaaaaatt ccttcagtgg gaggtggtgc tcgagtgggt ctttcaggaa gagcaggcat
9660

tgattgggca cacaggagca tggagagaag ggcattgaag gcagagacca gcatgtataa
9720

aggtccagag gactggaaat ggaaagaggc atttagggga ctgtgagggg cgtggcaagg
9780

ctggaagagt ggaaggctgg aaggctctgtg gaagaccctg tttatcaggc taggagtttg
9840

gactttaccc tacaggaaaa cagcgtaagt cactcctgca aggcagaggg tgaactgggc
9900

agggagctgc gggagaccag ctagagactg ttggaaatga tcgtcccat ccagaccaa
9960

ggtggctgaa aagcctgaaa aagaactgta agttgggtct tctaccctta ttctttatgg
10020

aattgtcttc tcagtagaga aactaggcta tccaaactca actgtcctga gtattagagg
10080

ccagtggctg ttttagcagc aacagagtgc agttgggtacc tacggagcag ggtaataaac
10140

taagcaagtg acagagatgg atttaagtgc ttctttacgt gacacacatt ttctttctcaa
10200

tgggagggat ggagatggcc cagtaacatg aacactgctc actgatgcca tagcatagat
10260

atcgaggaag tatccgtggt cagtgctgag agccccgaac aaagatttac ccacgaattt
10320

attaacagca agccagtcac tagcattggt tctatagata ttcattaact aaaagtatcc
10380

cttatgggaa aggaagggat gggccgaaat aaaggggtgg gtctggctag ttatctgcag
10440

caggaacatg cccttaaggc acagatcgct catgctgttt gtggtttaag gacaccttta
10500

agcagttttc tggcctgggc gggccaggtg ttccttgcct tcattccagt aaaccacaa
10560

ccttccagcg tgggcgttac ggccatcatg aacatttcac agtgctgcag agattttggt
10620

tatggtcagt tttggggcca gtttatggct agattttggg gggcctgttc ccaacagagg
10680

cacaaaggga ctaagataga aattggtacc aagaagtgga gtgttggtat aacagatccc
10740

tgaaaatgca gagttggctt tggaactggg taaagggtag aggctggaag agtttgagg
10800

tacatgctgg aaaaagccta gatagtgata aactaagcat taaggattat tttggtgagg
10860

gcttagaaga acagagctgt agacagaacc tgaatcttct tagagattac ttaagtggtc
10920

atgatcagaa tgttggtgga agtatggaca gcaaaggcca ttttgatgag gtcttagaca
10980

gaaatgagga atatttttatt gggaaactaga gaaaagccca ttctgttttt atttataaaa
11040

tggcagagaa cgtgggtcag ttgtgtctgt gccctaatgc tttgtggagg atagatgatg
11100

aactagatga agagtgatga actaggatat ttggtggaag aaatatctaa gcagagtgtg
11160

cagggtgctg catggcttct cttgattgct tataagaaaa tgcaagaaga aaacaaatta
11220

aagacagaaa ttataatcaa aagggaagca gaacttagaa agatttggaa aattttcagc
11280

ctggctagat cataaagaat aaaaaggctg ctgggaagaa tgaccttaa gggatttggg
11340

aggtcttttg gactgctcct cctgtcacag gcccagagtg cctggacat gagggcagga
11400

gggagaggac cataccccgc cagcaggctt atgggtctag ttgcccagg ccacctcgag
11460

cctctgcttc ctgcattctg cgggtgctcct tggccacccc aggtgcggct caggtgtggc
11520

ttgggctgct gctctggaag gtgcaggcta taaaccttg tgtccatgtg gtgctaactg
11580

cagctgagca gagttcaaga gctgtggagt gtggctactt tcatctagat ttccgaggat
11640

gtcttgagga acccggggcc caggcagaaa acgtctatgg ggaagagcca ccacagagtg
11700

ccccactag ggcagtacca tcagtcatga attgggcagt gccagactac atagttcagg
11760

gctccatcta gggggaatga gggaaagact ttgtaagtgt tcagggagc aaaacaaga
11820

gaataattga ttgggttaaag tagaaaggcc tagttataag ctgtcatgag attggagaaa
11880

aaaataaaaa taaggaataa aaaagaccct agttagagtt tagttggtgg ttcttgattg
11940

atgaagtctc tagttagaga tagttaatgg attgtgattg gttaagctaa gctttgtttt
12000

actgtttata ttgagttggg tttggtttgc ttacatagga acctgaaatg ttggagtcac
12060

ctcagcctaa ttgcttcctg attaattgatt tgtaacaca agttaggatg caactccatc
12120

ttcttgatac ctcattggcct aaaacagttg ttgggtgctgc agccgtgtat ttcaagttgc
12180

acatagctga aataagggaa aggcaaggat aatcctgctc tctctttttt tttgagacag
12240

agtcttgctc tgtcagccag gctggagtg c agtggcacga ttttggctca ctgcaacctc
12300

cctctgccgg gctaaggaaa ctcttctgcc tcagcctccc gagtagctgg gtttacaggc
12360

atgtgccacc atgcccggct aatttttgta atttttttta gtagagacgg ggtttcacca
12420

tgttggccag gctggtcttg aactcctgac ctcaggtaat ccgccccct cggcctccca
12480

aagtgctgga ttacagggtgt gagccaccgc gcccggccag tcctgctgtc ttttagggag
12540

cctgcttagg agccctgaca aacaactttg ccttatattc cattggctag aactttggcc
12600

agttttcaac atcttaaaat ggggaagattt tactatgtaa aactccagct atccaccttc
12660

tctcggaagc tctggtgaca ccaggccttt tgacaacagt tggctggagc tgaggaacag
12720

ctgcctgctc caggtggggc atgagctcct cctgtggccg caagccccgc atggcctggt
12780

ctgctcattt aagtcactgc ctacttgtga cagtacaggc aaacctttgc tttattgccc
12840

ttttgcagat gctgtatttt ttttttttac cagttgaagg tttgtggcaa ccctgcactg
12900

agcaagtcta ctgacgccat ttgtgtaaca gcatatactc acttcatgcc acattttggt
12960

aattctcaca atatttcaga tttttcatta ttattatatac tgttacgggtg atctgtgatc
13020

agtgctcttt gatgctactg ttataatcat cttggggccc gtgtaagatg gcaaagttaa
13080

tctataaatg tgtacattct gactgctcca ctgactccag cagttcatgt ctctgtctct
13140

ccctccctct caggctctccc tatttcctaa gacacaacaa tattaaaatt aggccaatta
13200

ataaccctac aatggattct aagtgttcaa gtgaaagagt ccactccctt taaatcaaaa
13260

gctagaaatg attaagctta gtgaggaagg cttgtcaaaa gctgagatag gctgaaacct
13320

aggcctctgg tgccaaacag ttagcaaagt ttttaatgca aaggtaaagt ccttgaagga
13380

aagtaagtac tactctagtg aacagataaa tgataagaaa gcaaaacagc cttactgctg
13440

atatggagaa agtttttagtg atttgaacat aagatcaaac cagacacaac attctcttaa
13500

gccaaagcct agtccagagc aaggctcctca gcattgaagc aaggccttcc actagcagaa
13560

agaccatgac ttactgacag ctcagatgat tgttagcatt tttttgccat aacatatttt
13620

aaaattaaag tatatgtatt tttttaggca taatgctatt gcacacttga tagactatag
13680

tgtaaacata actttttatat gcacaaggaa accaaaacct ttgtgtcact cacgttatta
13740

ggatattttgc tttattgtga tgatctagaa ccaaaccat ggatatctct gagagatgcc
13800

tgtattttgag cttagatccc tggcttgaag ggataaatga atgagtaagt gagggaggag
13860

ggagccagct tttgcatgag gcttgctccc ttcccctccc ctcccctccc ctctcccctc
13920

gcctccctgc ccctcttggt tcacagggtc tctgttgctc aggttggaat gcagtggcag
13980

gctcactgca gcctggatct cctcgggctc aggtgatacct atcacctcag cttctcgagt
14040

agttgggact gcaggcatgt gccaccatgc ccagataatt aaaaacattt ttttttttag
14100

agacacggtt tcaccatggt gccaagctg gttttaaact cctgggctca agtgatcctc
14160

ccatttcagc ctctcaaagt gctaggatta caggtgtag cccccacgcc tggccctgaa
14220

tgttgagtat tgtcggaaag tttagcagtc cttttacagc aagcacattt gtaagttggt
14280

tttgcttggc attctttgtc attcatttaa caaatgctta ttgagaaatt ctttaagtgt
14340

caagcactgt tttagacctt ggaataaagc agtgactgaa caaaatagat gtgaagaggc
14400

tgtgtgtggt ggttcagacc tctaatacca gcactttggg aggccaaagt aggcggattg
14460

cttgagctca ggagttcgag gccagcctgg ataacatagt gaaaccctg tctctactaa
14520

ataaataaat aaataagcaa gctgggcata gtggcacgca cctgtagccc caggtacttg
14580

ggaggctgag atgggaggat cacttgagcc agggaggcgg aggctgcagt gagccaagat
14640

catgccactg cactccagct tgggtgacag agtgaaaccc tgtccttaaa aaaaaaatg
14700

tgaaggattc ctgcacttgt gaaacttata tcctaactgg ggagatagtt gttaaaaaga
14760

gtaagtaaaa gatgtattat tgcatttggt ggtgattgag aattctcaag aaaaataaga
14820

caggggaagga tgatagggaa ttgggagtgt atgaaattta gatgggatca tttgggggaa
14880

atcattgaga agttgacgtt tgagcaaaga tttcatggag gtgaggaagt gagccctttg
14940

gatatgggat gttggagggga ggtattttcta agctagtaga cagttccgga ggcaggtgta
15000

tgctttgctt acctaaggaa cagcaagatc agaatgactg tttcagaatg agtaaggggc
15060

caggagtgtg aggataagtt ggaaggtaat gggagggctg caactgtata gtcttacagg
15120

tcatcattgt aaggatttta ggttttatta agcacatggg aaccccctag aggcatttta
15180

gaggaatgac atcttgtgac ttcctcttat aggtattctg acttttgtgt tgagaataga
15240

ctttaggggc taagagcaga agccaggagg tcgcttagga gattctaagc aatagtgcag
15300

gtgagatgag ctggtggcct ggaccagggt tcttgtggat gtgttgaagt cattagatgc
15360

tggatgcatt ttgttggtgg aactaacagg atttgctaaa catttacatg tgggtttgaa
15420

agaagatgat tcttttttac tttttaattt taaattcatt ttaattatta tttttgagac
15480

agaatctcac actgtcaccc agggtagagt gcagtgggtcc aatcacagat cgctgcaacc
15540

tccacctcct gggctcaagt gatcctccca cctcagcctt ctgagtagct gggactacaa
15600

gtgtgcacca ccatgcccag ctaatttttt attttttgaa gagatgggtgt ctactatgt
15660

tgcccaagct gttctgaaac atctgggctc aagtgggtcct cctgccttgg ccttccaaag
15720

tgctgggatt acaggcatga gcccctgcac ctggatgaga ttattcttat caaagaggga
15780

tccaggattt tttcctgggc aagtggaggg aaggagtgtt atttactaag gtggggaaga
15840

ctgtgggagc agcaggggaa gggcagaatg ttgagagctc tgttgtagac aagttaagtt
15900

tgcgatgcat atgaaacacc gaggagatgg tgatgaggca tttgggtctg taagtataga
15960

gttggaatt gtaggctata ggtggtgata tagtctgaat gttaatgtcc ctgccaaatt
16020

catatgttgg aacctaacc ccaaggtggg ggtattagga ggtgaggact ttaggaggtg
16080

atttggccat gagggctctg ccttctcgaa tgggattagt gcccttataa aagaggcttg
16140

agggagccac tttgctcttc tgccatttga ggatgcagca gcaaggcacc atctgtgaag
16200

caaagaacag ccttcaccgg atactgaatc tgctggctcc ttgatctggg acttcccagg
16260

ttcctgaact gtgagcaata aatTTTTgtt gcttataaat taccagctct aaggatTTTT
16320

gtgatagcag cctgagtgag ctaagaaagg tatacttctc ctggcctgga tgagtgtaca
16380

aggaatgagt agagatggaa gaggtccaag ggctgagctg ggggccactg cagcacttgg
16440

aggtttgggg acttggaggc tctggtgaaa gactctgagt aggagctgtt ggtgggaagg
16500

ataggaggaa aatcaggagg tggggcaact tggaggctga gcttggagac aaggactggt
16560

caccttgggc agtgtgtcac tggctgagag taagtatggg ggtgggcact ggggacctga
16620

ggagccagag aggggatagg accactgtct cggagagcga agggtgagcg gactaggggg
16680

tagttggact gccaggtggc agcacgggct ctcccgcagt tcgtggtgct gaatcagagt
16740

gtgggcagcc agcctggatg tgcgttctgt cacattcttg ttgagctgct agaggacagg
16800

cgtgaagcca agtggaactg agagtaaagc cgtgttggag ttttgcagga aggtgccaga
16860

tgatgagggc agccagcgaa gtcgaaggct catggaatga actgtgaaat ggaaacattg
16920

gcttttactt ttcgataatc ccattgtttc tctttttggc catcccatct cttcgtaatt
16980

tagaactatc ataatcgttt ttcactcaag actgttcagc caagtgcagt ttatttaaga
17040

gcactccttc agtgagctta cttaggtatg tgtgtgtttg tgcctgtctc ctgttaataa
17100

gctgatatga aagtcctgct gagctagttg tttcagtctt ggagatcagc agcgatatgg
17160

ggaggagaag ccaagcctct gttaaagccc accagtcctt ccctcccccg ctcccagcac
17220

ctagtgtttg aggaattcag actttgaata ctgctctttg ctactgttct tctctcctaa
17280

cttgccagta aacttccaag ggcagtatgt gtcttgtgct cctctcctgg tcttctgttg
17340

tgcatacact taaatacggc tctcaggctg tgaataatct gctttggtta acactcccct
17400

ctgttctgaa ggctgggctg ctagagccct gactgttggg ggctgtgctg tttgcagtgc
17460

tcatgttctc ttcttgggaa cccaacatat gcagctttaa tgtcaggctc tgccctcctg
17520

tcctgtcaca ggatccggga ctctctccc catcttggga gggatggcct ggggatttcg
17580

tctggaagc cctgcccagt gtgagcctct agctggactg agcaggtaat gtttgtgagt
17640

gctgaaatat tttaaccaga tgtccctcct ccaaagagcc ctttgtcctg gaggttcatg
17700

gaaacatgag ctgcccattct gcctttgggc ccaggcttct caggttgcaa caagattggt
17760

taggacacca aagagatatatt tattttttata caggaaaaat ccacagactg ggaaggcgaa
17820

taggaggtaa agaagatgaa ttttcttctg gaaaatggct cgtgaaattt gtacatatgg
17880

aagacagaat aatctagtgc aaggaaatcc ttgaagacct agaagaaata tgtgagggac
17940

tgatagagtt tgcccttgtg ttgtggtaag ggggacaaaa agtaccatga aattctaggt
18000

agaaagcaca ttctgagcag tatataattg aactcattta tttaaatatg gttaaataata
18060

cttattcaaa tgtaaatatg taaatatatta actgcttaaa tatacttatt taatgtctga
18120

tattgttcat caccatgtta ggtctgtggg cagctaagat atgagatcct tgtgttccag
18180

gatctgacca cgtttcccaa agtgagcaca tgcccatagg catcagtaac aagtgtggac
18240

tgagctgact gcagcctcgg tggacagaga agacttgtcc tccagcagtg ttacaggctg
18300

accacctgct ggcaccctga tgttgacatt agtgatgctg gaattcagat aaggagaga
18360

gctgctagca ggattccaca gagaggcagg acttaaagga taagaaggat tagatggtgc
18420

cggaagggg aggagcaggg aagccagggt cagggtgtgag gaaaggggaa ggtgctagaa
18480

agaggagttt ggacttaata ctgtagccta ggaagggtggc ccatgccttg gctgttccac
18540

acgtctctta gattctatgg gagcaaagga gaagccactg tgtttaccac agacttgcct
18600

tcttcccacc ctccccgctc atcttggtgt actcatcacc atgtgctcaa tctgcgaaac
18660

tggaaatggc attgtcatcc tcaggccccc actttttgtc acctgctgtg atccagcagt
18720

cactgagttg ggaggcctcc tcgtttggtt aatcgtattc tactattaaa ttttctgccc
18780

cttgtaatcc ctccctctcct gggctcttcat tcttgtttct gcccaaactg tttccagatc
18840

acaaatatga ttctgtgagt tctctcctca gaaactcctg gtgattctac accaccaca
18900

ggacaaagtc tagactctgc cctttctaac ctggccccag ccccttttc tggctgtctt
18960

ctgggtctga gccaccatct agaaatggtt tgaatggtga agtatattcaa tttcatggag
19020

ctaacagaat agattattta aaagcatgtg ggacatatga tcccaacata ttttccttat
19080

tagctattca taggcacagc acagaagaag catttatatt tataatgac tcattgtcacc
19140

aaggaaaagt taaaatatgt tcaataactc actttaaaag gcagcggaaa aatagatgtc
19200

acttatgggg aatattttca cccagaatga caaaattttt aaaaaatcta atgctcaaag
19260

cgatgagctt cagttacctg gaaaatctac ttctgatgcc ccataaatga aatgtcaggt
19320

ggtcagcagg gcatgtgagg ctccatgatag atcctttcag gaagggagag acatagataa
19380

gcaagacctg aagaatcact agaaagcaat cacagatgca cctgatgtat ttacatgggc
19440

tcccatagag gcagctgcac agccccacag tgctggcctc tggcagggag gctagggttg
19500

gggctcaagg ccttttagcat caaatatacc tttcctttga tgaaaatggg tttgaatgac
19560

atttcttgga aattgaggca ggttctttgt tttttggctg agcacttagg agataagaaa
19620

caaaccaga cacatttttg aggagttagt tgctcagcgt gtgctttaga agtgatagca
19680

accgccacaa gtgtggtaag aagcagcagg ataaagcagg cacgaagggg ctttgggtgca
19740

tggaccagcc ttggtgttgc ctttcctcaa agagaacaga tctctcagtc aggagtctt
19800

ggggcagggg tgacagtgtt aatgggggag gaagcgtcac tgtgggcaac ctatgaggag
19860

gcagccttca gtgaccttgg gcatgtgttc tgagtgccct cactgggtgc acaccttagt
19920

gacccctgc ttggtacattc tgctgtgaag aataacacgg agagacagag tgatcatttt
19980

accctgaatt caattccaca ctcacttgca gcaggaggta tggctggaat tagtggctgt
20040

ggtggagctg ttcctgccat gggggatgtg gtctctcctt ctgggtccccg agtgggtggg
20100

tgggaatgga ggcagcgctg cccgacagag caggctgctg ctggaattac cacagtgtcg
20160

tgtgtctcct cctattaatc cccatgtgga acagcaggta gcctgagaag cctgaattgt
20220

gccttagttc ctggccgtgg tgctgcttaa acagatgagc cttgttagcc agtttgtgca
20280

gtcagtttctt actgaagtca agccttttggc ccaagggaaa acaaagctac aggaaacatg
20340

gaacttttctt acacctttctt gaaaatacaa gtcattgttgc ctaggaaata cttctctcac
20400

caagtattctt gcacgctgcc aacattggag ctgtggaaga ggcagctgaa gcaattccaa
20460

gaaagtgcta gcagtttggc aataaaagga agtggcatat ctctcctgcc tctagactgg
20520

gtttctagga gtgcacatgg ccaaaaggaa actgctgcac tctggcaagt gattacatgt
20580

tgtgctcatg gacaccttgg ccgggtgtcc agtgtgcact gggtgagggg gtgcatgggc
20640

tgtgtgcaaa gatctgcagt gctgcacaca gtgtgttgag tgagaggggtg ccaagtgctg
20700

ggacggttgt agataaaca aagcaaatag gtgcagtact gtgcagcctg cccaagttcc
20760

caaagcttgc tgccaggctg aagactgctg gcagcctctc ctgggctggg tcaggagtag
20820

aaggggtgtct tccccttgct aattaggagg acacttctaa agcatttctc tctaacacaa
20880

ccaaatagtc tgctgaaggc gaactgcttt agtgaggagg gaagccgtgg ccgtgaactc
20940

agcgatagag aaagccaggg gtgatgtgcc ccagagcccc gagcagtcct tcctggcaac
21000

tctgagtgac caacatgtga ccctgctgtc tgtccacaag cagtggtcct ttaggaagag
21060

ccctgggggtt aggcattcag ccaggccagt gtctactgtg gaggaaggca ctccaccgtc
21120

agaggaagga gaatttccac agagagtttt gaatggaacc tgggaaatga gttccagtca
21180

ggtcagtgcattgcagatgggctaanaagcctggaagggcctggacccttttaaagat
21240

ggaggtttctctccatgtctcttggtgtgggatattggggagaggtctgggagcggaatg
21300

agagtgtgaatgttagcacctggcaggagtc aagctgattaattacaccatggcttggaag
21360

tttgggttttgttgctttttattttcagaa tattttgtaa gactctatga agcaaaagtt
21420

ttccacaatgcaagtcaaacaanaatgaatccttccttcttttactaatatgttttata
21480

gtgtataaccttattttgttaatctgtacattttaactaaacattttgaaatgctgcaaa
21540

atgaagttcttttcttaagcaagttgaagttttgttactccaaaggttaggcatgtcc
21600

aaaagtaagacagtataacgttttgtaaaaatgggtgcctttagcagcatttctgagcag
21660

actggtggtgtgtacagggcttcttctcattgacctagtacacgaaaataaaaacaact
21720

tttgaaggcatttactagttagatacaaggagggtttcataagttactttgcttcctcta
21780

agacttatttttgtgctgttcattgggttaaagtcctcaaggactaaaggaaattttgtgaa
21840

aagaactttcaaatgtgctatgcaagagatttacatacgttgaaatcacagggtacttgg
21900

atattgaagacagaaactcttttgagtatttaaagattctgcattttgtatcagggttt
21960

ttaaccaaggatataactatgctcggaggattttatttttaaatattaaaagtgtttttt
22020

ggtgttctgaatttgtggaaatttccaagagactcctttaagttttttcctaggcaaat
22080

ttaaaaatat gaaatgtaga attattcagt ggctgcccgt ttgggtagct tctacacaga
22140

ggttcgagtg gttttagggg agggaagtga ggtgatagat acaattttta ttactttggt
22200

ttctaaagta gttttataaa tttcatttta tgtctcttta aactgtatta agccagtagc
22260

taaactaatg tgttaaatca cttagaattt taaagagtga cagagcgttt aatcttttaa
22320

tcagcctttt catatgttgc cggttttatac aatatacggg gaacatttat acttttagcc
22380

atgtgtttta aatagtgaat tactaaatat gttgggggtga aaatgcttga ggcattttta
22440

aaaaatgtcc aaactgttac tttacaattt ccagcctata tctgaagttt tgagatatga
22500

cgtgtgctgg tgatttttta ggccagcacc aatctgtgac tagtggggcg acctgaagct
22560

gactagggga gcctcttcag aaaggagggc tggggaagag gccttggtgg aggggatctg
22620

agtgcagacc caacacaagt tgtgggacct gggcagggtga gtcttggaca ctctccttta
22680

gattcttcat tgcaaagtgg atgataaatc ctgcctcata tgattgtgaa agcactttgt
22740

agattatcat gcaaatgcaa aatagtgtta cttctgccaa aattatgctt tttcagaaga
22800

atttccaact ctgagatatg tttgatgtgt tgtctgaagg tctggacagg actgtgataa
22860

cccagtaatc tatctgagca cagttgcaac tctatttttc ctttgtttgt ttttttctaa
22920

gttttttatt ttgggacatg gtctcttgct gatgccgagg ctggagtgca gtggcatgat
22980

catggctcat agcaccctcg accttccggg ttcaaagtat tttcctgcct cagccccctg
23040

agtagctgcc accacaggtg catgccacca tttccagttt ctctttatatt gtagagatgt
23100

ggtctcccta tgttgcccag gctgggtctca aactcctggg ctcaaagtat cctcctgcct
23160

ccccctccca aagtgctggg attatagatg tgagccaccg tgcctggcag ctctattttt
23220

tctaaagctc tgattataaa agtaacatat actctatata aaaaacaaat ttacgatgga
23280

acgaggagaa tacaaaaaag tataaagaaa aagcagtcgt tttcccaaca aatatgaatg
23340

tgtatacgtg tatgtatggt tatgtttaca taatgtatat ttttaaaata aatatgggat
23400

caaattatac atactatctt atactttgca tttcttactg aatattttgt ggacatttct
23460

gtatgtcaat atatatgtaa atctgtatta tctttttttt tttttttttt tttttgagac
23520

agagtcttgc tctgtcatcc aggctataga gtgcagtggc gtaatcttgg ctcaactgcaa
23580

cctccatctc ctggggttcaa gtgattctcc tgcctcagcc tccagagtag ctgggactac
23640

aggcacctga caccatacct agctaatttt tgtgttttta gtagagatga ggtttcccca
23700

tgttgcccag gctgggtctca aactcctgag ctcaagtgat ccactcacct cagccatcca
23760

aagtgctggg attacaggcg tgaaccaccc tgcctggcct atataatcat tttttattgc
23820

agttccactt ttaaatttag gattaatttt tagaaagaat ttagttatcc ttaaatagaa
23880

gaaatttgca gtagctgggt tatgattata aaaaccaacc caaaagttaa tttcaattgt
23940

gactcaaagt ctgttataaa tcagctttgc accttgagaa ctaacattta gcacttggt
24000

tatggtagga accttactct tgtcacttaa tcattttaac aactttagta ggtatcatca
24060

gtccaccttt aacagacgag aaagctgagt gagagcttag gtttcctgaa aaatttaatc
24120

ccagggtccgt ctgatgacaa gtccttgctt ttccattgga aactgaagt attatttaat
24180

acactgggca aaatagagtt ttgaaagtat tggaaaactg ctaattcata gtgaatggaa
24240

agttagcttt taggtggcac gttgttcaga aaggccgctg ctatgaataa aacatttaaa
24300

aattgctggt tatgcctagc agatttggtt gcttatattt tacatttctt tttttttttt
24360

tttacatttc ttacacaatg cagagtacat tctgtacaat gtagaatgaa atcatgactt
24420

tggaaaacct gtttctttga ttcttctatg tttcagacag gcagagaggg gaaaaagatc
24480

tgtttttaaaa ggcataactg gtaaaagctc agagtctgga gggcacattt ccgcccgcctc
24540

ctctcccaca ttagaaagtt tctgtttgct gatgtgagca gggctaagcc tgggtgaggtt
24600

cctaggaatt aattatgaaa aggggggaaat agggaggagg gccttccccg caaatcccct
24660

ccataagagc acgctaacca gagtagaagt caggggctcc gtgctttggc ccagggtgttc
24720

ctatcaaagg ttccacagac aacctttcag gcccagaatg gtgagctgga gaggaatggg
24780

tgccctgccac tgaggctctc ccttctctag cccctcctta actctctcct ctacttgtct
24840

tggcaggtgg gaaatggttg gaggttacta gaaagttggt ccagctgtat acagaattac
24900

cctttgtttc cagaaacaac tgcttaagcc tccttactca tgactctggt gacttctttg
24960

gccaaagcttg gctgctgttt ttttttgttt gtttttttcc cgacagggtc tcaccgtgtc
25020

accaggttg gagtgcagtt gccgcagatc ttggctcact gcaacctcca cctcccgagc
25080

tcaagggatc ctcccatcag tgatcctccc accctagcct cctgagtagc tgggactaca
25140

ggtgcacact ggctgctggt ttaatttgtc atcgtatcct gtatctgaca gagttcagtc
25200

catggtatca agctattgac tgagtgaaca aatgaatgaa ctagaagtag atttatactg
25260

cagcagggag gattgatggt agacacattg aactgatggt gactataagg ttgttacgtg
25320

gtggacttag cattttttaga gaccagaata tgggaattttg ttcttaggta agtgggtggtt
25380

aggttctttt ttttttttct ttaaatatga gacagagtct cgctccgtcg ccaggctgga
25440

gtgcagtggc gcgatcttgg ctcagtgaac gctccgcctc ccaggctcaa gcgattcttc
25500

tgccctcaggc tcctgaatat ctgggactac aggtgcatgc caccacgccc aactaatatt
25560

tgtattttta gtagagatga ggtttcaccg tgttgtccag gatcgtctca atctcttgac
25620

cccgtgatcc acctgcctcg gcctcccgaa gtgctgggat tacaggtagt agccattgtg
25680

ccagggtggg ttcggtttta tttttgccta gttgaaatca ggatgacttg gctggcacag
25740

atatatccct tctaaaataa aggaaaaaca ataaatgaag tgggtgtgatc tagtcattta
25800

aagaagttat ctgtcattat ggaatcatca gcattcgcta ttattttatt tgcttttgta
25860

gagaccattt gaagtggatt cttgaattta gtttttccta aaagtagtat ctgctggata
25920

caaacattag tagcaaagta tagaagtgta gttcttttcc tctcttagat aaagaggact
25980

gtcttaatac agttctgcaa cttttttttc tttttaacat catggaagtc tttccacgtc
26040

aatacatgta aattcaccta tttttaaaaa accatgacat agattttctgt tcgatggact
26100

atattttatg tagacagttc tctatagaag aatattgaga tggtttccag tttttcacta
26160

tttgaatgta gcaatgaaca ttaatatgta attgtatatc tatatggtaa aattcccaga
26220

agtggaattt ctaggacaaa agcagtaata ctttttgaag agcaattagg caaaatttag
26280

ctttctacaa acagtgggta agaactgttt ttccttatac tcttgtttac tgtatcatca
26340

gactttttta attaataaaa taacatttat tttattaatg ttataataaa tgttatttta
26400

attttaaaaa agatttatta gataacttaat atttcttttt ctaaggatta ttttaagtag
26460

catttcccat catgcattta tagagtagtg ctaatttata aacattaggt ttgtacatgt
26520

cttaaaggaa tacatttggt ttatgtgtca tacctataat tctaggcaga aaccaatact
26580

ccttggagta cagagtaaaa aaaataaata aataaaacct ctttgttggg aaaggcttga
26640

aatacagagt gttattagag aagattctga agcaggaaaa tgctgttaat tgataaaagt
26700

tagatggaat aattgttaca atagacataa attaaaagca gcacattctc tgaaacggtt
26760

gattataaag ctgtgtagca cgagattgtg aatggggttt ctaagataat attgtgaagt
26820

gctccttgga gacgataccc tgtcttatag tttgtgaatt tctaaatttc attttgttgt
26880

atattgaaaa taacatattt gtatgtacat atgtggtgag gtttggttaat tttgttttta
26940

attttactca cttttttttt tttaggcaag tgtcagatat gctacatttt gtttctcagt
27000

aagggttaat attttccctt gctgtggttg agagaagcaa ataggcaatg aatagtcatt
27060

tgacacagaat taccagaaa gtttacagtc tgtcctcaga agggcagaga ttgggttaat
27120

gcacacatag ccctggcaca tagtaggccc tcctatgtgt ttgaatgaat gccttcattt
27180

ttaaagtaga taaaatttaa ataattgtca gtggaaagtc tttttggacc attgtgaaac
27240

aggattcaag gccagagtga agggcttggg ggaatatgta cactactttt aataattagt
27300

ttgccttatc agtaaaatgg agaaaactat gtagagtact cgaaccttcc tgggtcatgt
27360

ttgaagatca aatgaggtga tgtttgtctt catagtcatt tgttttcaaa catgtatgat
27420

gttaatcttt ttttaaaaaa ttatacttta agttctgggg tacatgtgca gaacgtgcag
27480

gtttgttaca tagatataca cgtgtcatgg tggtttgctg caccatcaa accgtcatct
27540

acattaggta tttctcctaa tggatatccct cccctagccc cccaccccc gacaggcccc
27600

agtgtgtgat gttccccctcc ctgtgtccat gtgttctcat tgtttagctc ccacttatga
27660

gtgagaacat gcactgtttg gttttctgtt cttgtgttag tttgctgaga atggtggttt
27720

ccagcttcat ccatgtccct gcaaagaaca tgaactcatc cttttttatg gctccatagt
27780

attccatggg gtatatgtgc cacattttct ttatccagtc tatcattaat gggcatttgg
27840

gttggttcca agtctttgct attgtgagta gtgccacaat aaacatatgt gtgcgtgtgt
27900

ctttatcgta gaatgattta caatcctctg ggtatatacc cagtaatggg attgctgggt
27960

caaatgggat ttctggttct agatccttga ggaatcgcca cactgtcttc cacaatgggt
28020

gaactaattt acactcccac caatagtgtg aaagtgttcc tatttctcta catcctctcc
28080

agcatctatt gtttcctgac tttttaatga ttgccatcct aactggcatg agatgggatac
28140

tcattgtggg tttgatttgc atgtctgtaa taaccagtga tgatgagctt tttttcatgt
28200

ttgttggtg cataaatgtc ttcttttgag aagtgtctgt tcataatctt caccacttt
28260

ttgatggggg tgtttttttt ttcttgtaaa tttgtttaag ttccttgtag atcctggata
28320

ttagcccttt gtcagatgga tagattgcaa aaattttctc ccattctgta ggttgccctgt
28380

tcactctgat gatagtttct tttgctgtgc agaagctctt tagtttaatt agatcccatt
28440

tgtcaatfff ggcttttggt gccattgctt ttgggtgttt agtcattaag ttgttgccca
28500

tttctgtgtc ctgaatggta ctgcctatgt tttcttctag ggtttttatg gttttaggtc
28560

ttacgtttaa gttttaatcc atcttgagtt aatffffgtg taagggtgtaa ggaaggggtc
28620

cagtttcagt tttctgcata tggctagcca gttttcccaa gaccatttat taaatagggg
28680

atcctttccc cattgctttt gtcagggttg tcaaagatcg gatgggtgta gatgtgtggt
28740

gttatfffctg aggcccatgt tctgttccat tggctctatat ctctgttttg gtaccagtac
28800

catgctgttt tggttactat agccttgtag tatagtttga agttaggtag tgtgatgcct
28860

ccggctttgt tctttttgct taggattgtc ttggctatct gggctctttt ttggttccgt
28920

atgaaattta aagtagtttt ttctaattct gtgaagaaag tcagtggtag tttgatgggg
28980

atagcattga atctataaat tactttgggc agtatggcca ttttcacgat attgattggt
29040

cctatgcatg agcatggaat gtttttccat ttgtttgtgt cctctcttat ttccttgagc
29100

agtggtttgt agttctcctt gaagagggtcc ttcacatcct ttgccaaatt gtattcctgg
29160

gtatfffatt ctctttgtag caattgtgaa tgggagttca ctcatgattt ggctttctgt
29220

tattgggtgta tagaaatgct tgtgattttt gcacattgat tttgtatcct gagactttgc
29280

tgaaattgct tatcagctta aggagatfff gggctgagat gatgggggttt tctaaatata
29340

caatcatgtc atctgcaaac agagacaatt tgacttcctc tcttcctatt tgaatatcct
29400

ttattttcttt ctcttgccctg attgccctgg ccagaacttc caatactgtg ttgaatagga
29460

gtggtgagag agggcatcct tgtcttgtgc cagttttcaa aaggagtgtc tccagttttt
29520

gccattcac tgtgatattg gctctgggtt tgtcataaat agctgttatt attttgagat
29580

acatcccatc agtacctagt ttattgagag tttttagcat gaaggggtgt tgaatTTTTT
29640

cgaaggctctt ctctgcgtct attgagataa tcatgtgggtt tttgtcattg gttctgttta
29700

tgtgatggat ttcgtttatt gatatgcatg tgttgaacca gccttgctgc atcctaggga
29760

tgaagccgac ttgatcatgg tggataagct ttttgatgtg ctgctggatt cggtttgcca
29820

gtattttatt gagaatTTTT gcatcgatgt tcatcaggga tattggcctg aaatTTTctt
29880

TTTTTgttct atctctgcaa ggttttggtg tcaggatgat gctggcctca taaaatgagt
29940

tagggaggag tccctctttt tctattgttt ggaatagttt cagaaggaaat ggtaccagct
30000

cctctttata cctctggtag aatttgactg tgtatccacc tgatcctaga ctttttttga
30060

ttggtaggct attaattact ccctcaattt cagaacttgt tatttgtcta ttcagggtt
30120

cgacttcctc ctagttttaga ctttggaggg tgtatgtgtc caggaattta tccatttctt
30180

ctagattttc tagttcattt gcttagaggt gtttgtagta ttctctgatg gtagtttgta
30240

tttctgtggg atcgggtggg atatctcctt tatcattttt tattgcatcg atttgattcc
30300

tctctctttt cttcttttatt agtctggcta gcggtctatc tattttgttg atcttttcaa
30360

aaaaccagct cctggattca ttgatttttt tgaaagggtt ttcattgtgc tatctccttc
30420

agttctgttt tgatcttagt tatttcttgt cttctgctag cgtttgaatt tctttgcgct
30480

tgctgctcta gttcttttaa ttgtgggtgt aggggtgttga ttttagagct ttccagcttt
30540

cttttgtggg catttagtcc tagaaatttc ccactaacca gtgctttggc tgtgtcccag
30600

agattctggg atgttgtgtc ttgtttctca ttggcttcaa agaacttatt tctgccttaa
30660

tttcattatt taccagtag tcattcagga gcaggttggt cagtttccat gtagttttac
30720

agttttgagt gagtttctta atcctgagtt ctaatttgat tgcactgtga tctgagagac
30780

tgtttgttat tatttctggt cttttgcatt tgctgaggag tgtttcactt ccaattatgt
30840

ggtcaatttt aggataactg caatgtgggt ctgagaagaa tgtatatatt gttgatttgg
30900

ggtggagagt tctgtagatg tctattaagt ctgtttggtc cagagctgat ttcaagccct
30960

gaatatcggt gttaattttc tgcctcattg atctgtctaa tattgacagt ggggtgttaa
31020

agtctccac tattactgtg tggaagtcta agtctttttg taggtctcta agaacttgct
31080

ttatgaatct ggggtgttcct gtattgggtg catatatatt taggatagtt agctcttctt
31140

gttgcat tga tccctttacc attatgtaag gcccttcttt gtctcttttg atctttttgg
31200

tttaaagtct gttttatcag agactaggat tccaaccccg actttttttt gctttccatt
31260

tgcttggtaa atcttcctcc atccctgtat ttcgagccta tgtgtgtgtt tgcacatgag
31320

gtgggtctcc tgaatacagc acatcaatga gtcttgactc tttatccaat tcgccagtct
31380

gtgtctttta attggggcat ttagcctgtt tacgtttaag gttagtattg ttatgtgtga
31440

atttgatcct gtcattatga tgctagctgg ttattttggc tattagt tga tgcagtttct
31500

tcatagtgtc tatgggtcttt acaatttggt atgtttttgc agtgactggt accagttggt
31560

cctttccatg tttagtgtt cccttcagga actcttgtaa ggcaggccta gtggtgacaa
31620

aatctctcag catttgcttg tctgtaaaga attttatttc tccttcactt atgaagctta
31680

gtttggctgg atatgaaatt ctggggttgaa aattattttc tttagaagt tcgaatattg
31740

gccctattc tcttctggct tgtaggggtt ctgcagagag atctgctgtt aagtctgatg
31800

agcttccctt tgtgggtaac ctgaccttct tctctggctg ccctcaacat ttttcttct
31860

atttcaagct tgggtgaatct gacgattatg tgtcttgggg ttgctcttct tgaggagtat
31920

ctttgttggt ttctctgtat ttctgaatt tgaatgttgg cctgtcttgc taggttgggg
31980

aagttctcct ggataaatatc ctgaagagtg ttttccaact tggttccatt ctccccgtca
32040

ttttcaggta caccagtcaa acttatgttt ggtcttttca catagtctta tatttcttgg
32100

aggctttgtt tgttcctttt cattcttttt tctctatctt gtcttcacgc tttatttcat
32160

taagttgatc ttcagtcgct gatattcttt cttctgcttg atcgatttgg ctattgatac
32220

ttgtgtatgc ttcacgaagt tcttgtgctg tgtttttcag ctccatcagg tcatttatgt
32280

tcctctctaa actgattatt ctagtttagca atttctctaa ccttttttca gggttcttag
32340

cttccttgcc ttgggttaga acatgctact ttagctcggt ggagtttttt tattaccac
32400

cttctaaaac ctacttctgt caatttgta aacgcattct ccatccagtt ttgttcctt
32460

gctggcgagg gagttgtgat cctttgaagg agaagaggca ttctggtttt tggaattttc
32520

aggcttgttg cgctggtttc tcccatctt catggattta tctaccttg ctttttgatg
32580

ttggtgacct tcggatggcg tttctgagtg gacgtccttt ttgttgatat tgatgctata
32640

cctttctggt tgttagtttt ccttctaaca gtcagggtccc tctgctgcag gtctgctaga
32700

gtttgctgga ggtccactcc agaccctggt tgcctgagta tcaccagcag aggctacaga
32760

acagcaaaga ttgctgcctg ttccttcctg gaagcattgt ccctgagggg cacctgccag
32820

atgccagcca gactctcttg tatggggtat cttttggccc ctgctggggg gtgtctcctg
32880

gtcaggagac acaggggtca cggacccact tgaggaggca gtctgaccct tagcagagct
32940

cgagcactgt gctaggagat ccactgctct cttcagagct ggcaggcagg gacctttgag
33000

actgctgaag ctgcacccac agccaccatt tgccccaggt gatctgtccc agggagatgg
33060

cttcacttat aagcccctga ctggggctgc tgcctttttt tcagagatgc cctgcccaga
33120

gaagaggaat ctagagaggc agtctagcta cagtagcttt gctgagctgc ggagggctcc
33180

agccagttgc catttccctg tggctttgtt tacactatga ggggaaaacc acctgctcaa
33240

gcctcagtaa tgggtggatgc ccctcccccg accaagcttg agcatcccag gtcgatttca
33300

gactgtttgt ctggcagtga gaatttcaaa ccagtggatc ttaacttgct gggctctgta
33360

ggggtgggat ccgctgagct agaccacttg gctccctggc ttcagccacc tttccaggcg
33420

agtgaacggg tctgtcttgc tggcattcca ggcaccactg gggatatgaaa aaaaaactcc
33480

tgcagctagc ttgggtgtctg cccaaacggc tgtccagttt tgtgcttgaa acccagggcc
33540

ctgggtgggtgt aggcacccga gggaatctcc tgggtctgtgg gttgtgaaga ccgtgggaaa
33600

agtgtagtgc ctggaccaga gtgcaccatt cctcatggca cagtccctcg cggcttccct
33660

tggctagtgg aggaggttca ctaaccactt gtgcttcctg catgaggcaa caccaccacc
33720

tgctttggct tgccctccat gggctgtacc cgctgtttta ccagatgagc tgggtacctc
33780

agctggaaat gcagaaatca ccctccttct gcgttgatct cgctgggagc tgcagactgg
33840

agctgttcct attcggccat cttgccacaa gctgatgtta gtcttttagg gccattagc
33900

tctcaagcct tcctgtacaa ctgtaagatg gccattatta ttactcccgc cagatgccag
33960

ccggagcttt cttgtatcta aacctgtgtt tgttttttaa aactcatcag gagagttttt
34020

tcctccaaag atcatcttaa ttgaagaact gaattaaaag catacatgtg ccctgtgtct
34080

acacttaaaa ctatccaaaa cacaaacctt tatacctaaa ggaaactgag gcctggattt
34140

gctcagataa tgctggaatt catacttaaa actctcagac ataccctcaa gaaattcaga
34200

agtaggaaag acagatattt aaacaaatat atgtgaaaat gctctttgaa ctataaagca
34260

atttacaaaa gtgtcatcaa ttcatcattg ttgttatctt tatcacagtc aataatattg
34320

ttgcccaaat tagaatggtt taacaaaaca ttttcttcct tgccttagag caggggtgtc
34380

caaacttttg actttcctgg gccacattag aagaagaatg tgtcttcggc cacacataaa
34440

atacgctaac actaataata gatgatgagc taaaaaaaaa aaaaaaaaaat tcacacccaa
34500

aaaaatctca ttatgtttta agaaagttta caaatTTgtg ttgggccgta ttcaaagcca
34560

tcctgggctg catgtggcctt gcaggccatg gggtggacag gcttgcctta gacaaacggg
34620

aaaatcacat ttatTTtctg cttgtgaaag taacacttgt ttttttaatt taagtttgaa
34680

aaattggaac tatgtacaga tattttttaaa aatcagtatt tcataactca ccaaaagata
34740

acctgaattt tctctactga tattttacgtt tatccccctg tatacttaat gttcatagag
34800

gttaatatgg gacattgatt cttcctcagt ggggggtttta gagaccaacc cccactggcc
34860

tggaatgcctg ggatgggtttg ggtgagatcc tttcttgagt caggcatcat ttttgacctg
34920

tgccctcccc tgtgaatttg tggcatccag agaattctta gataaggccg agcaaggcag
34980

tactcattcc agatattcag gatgtacgtg gtccagagct gttgtcttct ctgacttggtg
35040

tcttggtgtgc cttctcctgt aagtgtgggt gccttgatta gggatattga ggaaagtgga
35100

gcgagtgtgg atattttgct gtagttttcc tgaggtgaca gatactagtg ggtgaagcca
35160

gggctcaggc agctcagctc tcttggtgtgt tagcctggct tgctgcttgc aggtatggct
35220

gtttattttc atttcacttt tctgcattgt agtctcttcc tataattaac atggctgaaa
35280

agaaaagggtt actctaaaag ttgttttttaaaaatacaga gcctctcgga tttctcagca
35340

gcgtttaaat cccagctctg ttacttaatg ttgatttggtg ctcttggata agcttctgag
35400

tcagtttcaa cagtgaaatg gtgataatat ctccttctag aattgcccta agcatgaggg
35460

atataaaatg ccatcctcat aaaataacac tcagaatcat taattatttg ttattattgg
35520

atctgcgtgc atatgaaaga atccggactt ttttgatattt tgaatcccat gaaagttcag
35580

atgtgttaag ttgtatgggg ggaggtgata tattggaaaa gttgccttca tgtataacat
35640

gtttgcctca taagttgtta tgggtgttat acatccttca aagtgaaaga attcttgagg
35700

ataggcagca ggggactgac tggatccta ttttaagggtg atgggaatcc aagagtgatc
35760

aagagacctc caaagtggcc agtggtcctt cacttgtaca tgagttagta ttttttgatt
35820

tgtgatcaca gccactttct aacaaagcat gttgtttctg ctttttcact atcagattgg
35880

ctttgaaagc cttcatgaaa acatgccacc cagtttctga tttcttcata tcagaatggc
35940

ttatccagag attttgTTTT cctttctggc ttcattcttg tgaatcagta gttatTTTT
36000

taacctatat tttatTTTaa aagccctggc actctcagca aatttcgtgt gacttgggta
36060

cctagagagg tagcagagaa ttgaaattca cttttggact cccttctgtg gcctcctctg
36120

cttccagtcc tttctcatcc tgacgtctac catcatttat caaaaattca aaagctagaa
36180

aaagtcattc taaagcctgg tgttgggtct gcagaaaatg tcccaaattgt ctgtccgtaa
36240

ggagtgcaga atctagtggg aattagataa actagggcct tcacaaggct ttctgggata
36300

tccatgtgga accttggcct gctctccatt tctgtgactc tttgatgggc ccttaaattct
36360

aaacctgtgt ttgtTTTTta aaactcatca ggagagtttt ttccccaag gatcatctta
36420

attgaagaac tgaattaaaa gcatgcatat gccctgtgtc tacacttaaa actatccaaa
36480

cactaccagt tgtttaatga gggtgttcct agaccaacag atactttagt gttttgtaga
36540

gatatgttat ttgcaaggct ttgtcttctg taactgcttg gccttaaaaa cctgcctgtc
36600

tagaaagggg ggcctctgag tagtgtttga taataagaaa ttgctttcag tactactgca
36660

aacccttttg gcagaaacat agaatcaaca ttctagaaat taagggctgc tgttatggtt
36720

tggggtttgt aaagttagtag ttggagcgag gaccaaaaaa tgtaaactgg gttatcatga
36780

gggatttggt tttccaacaa actagatggg tgaggggttat gtctttctta ttatgaggaa
36840

gagggttttg cctaataagg tactatatga aaatgaatgt ctttgaatat cttcttagtt
36900

tttcttttat ccttggttct tttttaatta aagaagtatt taattaacac caactacaat
36960

caagtcactg ttctgagtct tctttgtgct catgttatatt taattttaaa ttaattttta
37020

gattttggta atgtgtgtat tgtacagtta aattatatag tgaaagtaat tctctgtatc
37080

ccagccacct ggtccccttc agctgcctaa tcttcctcaa attctgtact agcttttgca
37140

ctgtcccaga gataatccat gcctatacaa gcatccattt gtatgatatt ctcttgtttc
37200

acacaggtag tcctagcacc atgacagaat tgtagccctc tcagtgaact acacaccctt
37260

gtattcaagc tccgtgtaat tcctgcccac attgactgtg cttggccgtg tgactggttt
37320

tggccaatgg gtgttagcaa gtgtgatgca agcagagggt tgataagagc ttgcactttg
37380

aaatgtgttc tagaggagtg cttgctattg tgacactccc tctcagtacc cagaagcccc
37440

cagctgtgag gtccccacct agtcataaag agaagccaca tggaagagaa ccgggtccct
37500

gccaacagcc tcatctgagc ccctagccgg cagcaagcat ccacttggca gccatgtgat
37560

gaatgaacta tcttcttggga agtggaactg ccagcccccatt tcaggccacc cccagtgggtg
37620

ctatgtggag aagagacaag ctttccctgc caagccctgt ccgaattgaa aaattgtgag
37680

ctgttgctag taatgtacta aagttaggttt gtcattgtagc aacagataac aaacacatgc
37740

tatatatatg atatcttctg tgccttgctt ttgttgctga gcagtatatt agatatcagc
37800

ttatttcatt acatacagag tcacatcatt cattttaaga gctgaataat attccattta
37860

tggctgtaca gtaattaact agcctcttat tgatggacat tgatgtttct aatctttctc
37920

agcatttaat atctccatgg catgctttgt tttgcacatg aacaaatata tttattgggt
37980

aaattcttag aagtagagtt gctgagtaaa agagtacgta aatttgtaat tttaataggt
38040

attgcaaatt gccctccaga gaggttacac tcccactagc aatgtaggaa gactgtgcca
38100

gtgtgggttg aatgttacat ttcttttttaa tattttccca tgtatttttaa aaaccatatg
38160

catttctttt ttgatgaact gtcttctttt ttgtttggat tgttggcctt tttctttttt
38220

ggaggaactc tctgtattta aagaaattaa ctcttaaatgt atgctatgtg aagcaaattct
38280

tttttcccc agttttttat cttttgactt catttgaata attttttgtc cttattgtgt
38340

ttagttatgt caaatctttt atagcttttg gattttcttg gttttgttct ggtttgttgg
38400

tcatatttag gcttttctca ctttaagatt attgtagata ttctcctgtg ttttcttcta
38460

gtgtttttgt tatttctttt ttacatgtta atctttgatc tgcctgaaat gtcttttgac
38520

gtagacgttt ccaggtaggt gtccgagggtg gctagtcatt tgtctgttaa tattgttata
38580

ccatgtcatc cttactatac agtcatgatg ctttagtcaa tgatggactg cacatatgac
38640

actggtccca tgatattaca atgaaactga aaaattccta ttgcctagtg atctcgcagc
38700

tgttgtaatg tcatagtaat gcattactca tgtatttttg gtgagctggg gtaaaccaac
38760

ttactgtact gccagtcata taaaaataga gcacatacaa ttctgtacag tacataatac
38820

tcgataataa atgactatgt tactgggtta tgtattttact atactctatt gttatttttag
38880

agtgtactcc ttctacttat aaaaagaaaa agttaactgt aaagtagcct caggcaggtc
38940

ctaggaggta tccagaagaa ggcattgtta ttataggaga tgacagctcc atgcctgtta
39000

tcgcacctga agaccttcca gtgggagaag atgtagagat ggaagacagt tatactgatg
39060

atcctgacca tgtgtaggcc gaggctaatag tatgtgtttg tgtcttagtt ttttaacaaaa
39120

atgcttaaac agtttaaaaa taaaaaaatt ttaaaagctt ataggataag tatataaaga
39180

aaatatttga taaagctgtc cagtgtgttt gtgctttaag ctaagtgtta ttacaaaaga
39240

gtcgaaagta aaaaaaaaaa aaattaaacg tttatgaagt aaaaaagtta cagcaagctg
39300

aggttaatta ataaattatt attattgttt cagagacagg gtcttgTTTT gtctcccagg
39360

ctagagtaca gtggtgcaat catagctcac tgtaaccttg aactcctggg ctcaagcagt
39420

cctcctgcct cggcctccca aagcattgag attacaatga gccactgtct ggctatgttg
39480

atttattgtt gaagaaataa aactatTTTT tataaatTTa gtgtagccta agtttacagt
39540

gtttataaaa tcgacagtag tgcagtgctc taagtcttca catctactca acactcattc
39600

accgactcac ccagagcaac ttccagtcct gcaagctcca ttcattgataa gtgccccata
39660

cagggtgcacc atTTTTtaat cttttatatt gtgttttttac tgtacctttt ctatgttcag
39720

atatgttttag atacacaagt acttaccatt gtgtttacaat tgcctacagt attcagtaca
39780

gtaacatgct gtacagggtgt gtggcctagg agcaacaggc tacaccatac agctcagggtg
39840

tttggtaggc tacatcagct aggtgtgtgt aagtgactct atgatgggtca cacaacgatg
39900

aagtcacaat gatgactcat ttctcagaat gtatcaccat ccttaacatc attaaacatt
39960

catgtatttta ctggattcag caatatcttt tgaactgctg tcccccaaat tacttaatta
40020

ttccctTTTT tctcttagga aatgctacag tgaatatTTa aaattctatg taattattaa
40080

tcccaaacct ttgtggatat agaaatcaac atttatatag tctgcatttt acaagttatt
40140

taaacatttg tgttgttaag ctgctcaaaa gccggttggc aaggtctgca ttttatgcac
40200

aggaaaaagg agaagatcat tccatgtaaa ggagattcat ctgaagacaa gcaggattgt
40260

gtacaagtag agagaaatgt gcttttttat gatgtatcat gtataaaggc ctgttcccaa
40320

atttctgaag cacatggaag ttctggtgta aaaagtatat ctagttaaga cagctgtagg
40380

ctgggagcct gtgaaaggaa gattttgctt gttaatagaa tgactgagac tacatcatcc
40440

tgtgtagaaa ccagggaata acacttgtag tgtgtttgtgt taagaaaaca ggttctttaa
40500

ttaagtgatg tacctgaaat ttcaggtaat gactacagct ctccctgaaca cccaggaggg
40560

cggtgacact gagtcacagg cagaactatg aaattccttg tgaaactcaa ctttctgttc
40620

ttccagcctt tcccctcccg cccctcggct cagccacagg tgctcattaa acctcgccca
40680

gggagcatac ctccaaagct ttttcacagg catgtggccc acctttggat tccgacccta
40740

tggtcctggg aagtctgcat ttggagcgtg tacctgaagg gtgttaagca catttaacct
40800

ctgctaagat aatcagacag acaccctag aaagagatat agaagtaagg tgttgactat
40860

tctggggttt gtatttcagc aggaccccca aagtttgagc tgaaaggact aagttctgtg
40920

gggaggtatc agccagaggt ggggacctct tttgggatac ctgtgccttc ccaccgtgag
40980

ccgctgtgcc ctctgcctct gctgcaggcc agccgtgctg cattccatat gtaaagctgt
41040

gtcgtgcttt ctttctttcc ctcagaatcc tccacgtctg ttgtgtgaga acgacatttc
41100

actaaggagg acagacagtt agcttctgtc agagaagaat gtcagagctc tttcattttg
41160

tagatgctgc ctttagcttg ctttctgtgt atgttatgtt cttctaccct tgggatattt
41220

tgacaccagt ggcattgttt tccattttct cagtgccaga gaagctggga tgttgtagaa
41280

aatatttggg ttcaagggtta aaaaatgtta tttcttttct ttttagatct tatctaagct
41340

ttctttgcct acccgtgcct gggaaccagc aatgaagaag agttttgctt ttgacaatgt
41400

tggctatgaa ggtggtctgg atggcctggg cccttcttct caggtggcca ccagcacagt
41460

caggatcttg ggcattgactt gccagtcatt tgtgaagtcc attgaggaca ggattttcaa
41520

tttgaaaggc atcatcagca tgaagggttc cctggaacaa ggcagtgcca ctgtgaaata
41580

tgtgccatcg gttgtgtgcc tgcaacaggt ttgccatcaa attggggaca tgggcttcga
41640

ggccagcatt gcagaaggaa aggcagcctc ctggccctca aggtccttgc ctgcccagga
41700

ggctgtgggc aagctccggg tggagggcat gacctgccag tcctgtgtca gctccattga
41760

aggcaaggtc cggaactgc aaggagtagt gagagtcaaa gtctcactca gcaaccaaga
41820

ggccgtcatc acttatcagc cttatctcat tcagcccgaa gacctcaggg accatgtaaa
41880

tgacatggga tttgaagctg ccatcaagag caaagtggct cccttaagcc tgggaccaat
41940

tgatattgag cggttacaaa gcactaaccc aaagagacct ttatcttctg ctaaccagaa
42000

ttttaataat tctgagacct tggggcacca aggaagccat gtggtcaccc tccaactgag
42060

aatagatgga atgcattgta agtcttgctg cttgaatatt gaagaaaata ttggccagct
42120

cctagggggtt caaagtattc aagtgtcctt ggagaacaaa actgccaag taaagtatga
42180

cccttcttgt accagcccag tggctctgca gagggctatc gaggcacttc cacctgggaa
42240

ttttaagtt tctcttcctg atggagccga agggagtggg acagatcaca ggtcttcag
42300

ttctcattcc cctggctccc caccgagaaa ccagggtccag ggcacatgca gtaccactct
42360

gattgccatt gccggcatga cctgtgcatc ctgtgtccat tccattgaag gcatgatctc
42420

ccaactggaa ggggtgcagc aaatatcggg gtctttggcc gaagggactg caacagttct
42480

ttataatccc tctgtaatta gcccagaaga actcagagct gctatagaag acatgggatt
42540

tgaggcttca gtcgtttctg gtacgtagtg tgtttgaggc atgtcctgag cttgtctcct
42600

tttctctttg tgtcttatag ctccctggatg gtggatatagg tgagccctgc tccctgcctc
42660

ccatgtcaac agtgggaacaa aatcctgcaa actttccttg catgttaaata ttgttaagat
42720

agaggaaata taatctccaa gtgtctagat tttgcttttc atgtagttgg gattttgtat
42780

tagtctgttt tgtgttgctt ataacagaat ggctgaattt gggtaattta taaagatatg
42840

gaatttgttt cttatagtta ctgagaagtc taaggttgag ggccacatct ggtgagagcc
42900

ttctgccagt ggggactccg gtgccacagg gcatcacatg gcaggggtgc tgagcatgct
42960

cacgtggttg cttaggtggc tcttcctctt tttagaaagc ctcgatgcc aacctgtgat
43020

aaccattaa tccatgagtg gattccattc atgagggtag aaccctcatg atccagtcac
43080

ttcttaaagc cctcaccttt taacactgcc acattgggga ttatgtttcc aacacatgaa
43140

atttggggga cacagtcaaa ccatggctga cttgaatgcg aacctccttg gatttgggct
43200

tcttttctgg agagaattct gattctgtat cctgaaatag agcactgtgg gaccatccag
43260

gcccttttct taccaagaga actattatca ccttctcaaa agattttact gctaaacaac
43320

agaaccttag tctcatcacc tagagaccat gctggtaaaa atgagagtta aggaaaggat
43380

ttataaatga tttttttctc cctgtttgat gaatataaga gaagaggtaa aggtgagggga
43440

gccaccaggc ttttatatth ttttaattaaa aaatttttta attttttttag ggacacggtc
43500

ttgccctgtt gtccaagatg gaggatagtg acatgatcat aggtcactgc atccttgaac
43560

ctttgggctc aggtgatcct cgcacctcag cctcctgagt agctgggact acagggtgtgc
43620

accaccacac cgggctcatt ttttatthttt atttttttgt agagacggag tgttgctgtg
43680

tttcccaggc tgggtcttgaa ctcttggcct ctagcaatcc tcccgcctca gcctcccaaa
43740

gtgctgggat tgcaggcgtg agccactatg ctcagccacc atcatgggtt ttaatcagag
43800

gaagcacgaa gtcagattat attttagaag actggtgtgg tgtcacagta cagtgtccat
43860

ggaggcaatg agaccacgta ggaggccggg tgagaaagag gtgggtttga aagacattca
43920

ggagggcagc tccaccaagc ttgggggtcac caagaggggt gtgtgatggg gagggccatg
43980

gaggaaaggg gtaactggga tgagccttta gctgagatgc tggaggacag tgggatagaa
44040

gggccttgaa ttttgaggga gagtgtgggg ccggaaatgc agagtgtggg aagatgacag
44100

agacctagaa agagtgattg tgagatggga aagaattaga gggaggagga ggggtctcct
44160

agtcaccctg gacataatgc gagccgcaga ggccaaggtc tgaatcgtat cttcgcgggt
44220

taactagggg aggccattgg aggcctcggc agctgcttca ggaccctggg aaggtgggag
44280

cccagcttgg ctgctggtaa gggtgagagac cactagtagc ttagattagg gataaagagg
44340

atgatagggtg agtggagatg gttgagagaa atgctgaaga taaggaggga aggggcacat
44400

gttgtgaggg gcagctgtgt gggtggtgtg ttgtaaccga ttaggtccca gacatgggcc
44460

ctgcccgtgc gtgtctctgt ggactggacc aggacatggc tggcttcctc agccttttgg
44520

ctggagctcg gtactgtagg gagcagcggg cccaacacta agaagagctg ctttcagtgg
44580

**cgtgggcaag cccactggca gctcagtgct cagggatggc tgaactttgg tttcccacat
44640**

**catttcctgt ggtcactgag ataaagtttg tgtccttgaa aagttcactc agctgtgagt
44700**

**ctttgctgcc tcccctctgt gcccacagca cacacctact gctgaaggga gtgggcattc
44760**

**atgtggagtg gcctggcccc attttgtgct ttgtcttttg tatcaagttt caggaaggag
44820**

**tttctttaaa acaaaaatca ataaggaaaa cacaatagaa cagtaaatacc tgtacgcaat
44880**

**aggtaagaaa tatagtataa ggtaaacatc tgcctgtttt tgtgtgatgt ttaaagtctt
44940**

**tcaaccgtgc aatcaggtta ttagtggtaa aaagaaatta gcactgaact tacatattac
45000**

**agaataattg gatttgatta aataaaccta gaaagaaaag aatagtgatt tttttaattg
45060**

**ggcaggatta tgagaatttt atttagccga agaggtattt agggatatcag acttagcaaa
45120**

**taaaaatata ggacatctac ttaaatttta atttttcata aacaatgaat ggtttttttag
45180**

**tataaacatg tccagctcaa tattcatgac atatgtatag taaaaattat ttgttgtttg
45240**

**tctgagaatc aaatttaact ggatgtcctg tattttacct ggcaacttta atttggaata
45300**

**aaccattagt gtgccagttt gtatttgaag agcccttctt gattgattag gagctccttg
45360**

**aggacagcct tgcggccagc gtgatttggt tctccttgga ctcaatagca ccccgcatgc
45420**

**agtaagtatt cctgaaggaa atgaaagata aggtgaagtt acggggtagc tcagtggaag
45480**

gatttaatat agagaactat ttacaaaggt attagagagc tggaaatcca aataggcaat
45540

gggaagaacc ccagagatta acaacagcag ggagctctgt cctccaggat ggctgagggga
45600

caaggtagtt actggagccc aggggctgag gctgctcggg gggagccggg acaatgaacc
45660

ctcaccaaga gccctgaaac ctcttggttct gaaaaacata ttatttgctc ttctaatttg
45720

aatatttttct gacattttat cctagaaagc tgttctacta accctcttgg aaaccacagt
45780

gctgggaatt ccatgggtgca aactacagat ggtacaccta catctgtgca ggaagtggct
45840

ccccacactg ggaggctccc tgcaaaccat gccccggaca tcttgggcaaa gtccccacaa
45900

tcaaccagag cagtggcacc gcagaagtgc ttcttacaga tcaaaggcat gacctgtgca
45960

tcctgtgtgt ctaacataga aaggaatctg cagaaagaag ctggtaagag atgaacgccc
46020

atgttgaact taggaatgct gcgtatagac ctcgtattct cccttcagaa taccagcaa
46080

ctgtgtttat cagtagccct gataaccagg tagcaagttc attatatagc ataaatgttt
46140

gtctaattgtc cttgtataac tggcattgtg ttctttaaga agtatttagt ttgagataaa
46200

atatatctct gagtaagtta gccattgaaa atccagttta aaggtcatag taccttttgt
46260

ggatagactt ctcagttagg ttgtactagt attattccat attttgttcc ctcccatccc
46320

atcaaccctt tactataaaa tgaataaata acctgcattt tgtatagaag ttccaatttc
46380

aaatattcctt tcttggttggt cctgtaaacc agagagtgtt tgtaggaaca acaggacaga
46440

aatatgtgga aatgtcaaag ttttgtcctc cagaatgcat ttccaatact tttaaacctg
46500

aaagtacaat aaaattcggg tccattttatt tgtttggaag atctcttaac catgactcta
46560

ccttgccacc agaatcactt tctcttttct taccaccagtg atgtgtcttt ttacttgaac
46620

tccctggccc ttagaatcag ccctgaattc agatgacatg tttttccagc ctcttggtca
46680

ctatgcatta tagcttcctg cagctctaga gtgactgtta gcagatggag gtgttaatag
46740

gagatccctt gttatccatg gtctgactgc tcataaaatt agcctataat ctcttatagc
46800

ctatgtttct gttcctgact tgaatcctca tttttatcct atgtggcatc ctttctgaag
46860

cacttccttc ctttcattta tttgaacgag gctttttttc ccatgtaaaa ctaacacaat
46920

agtgttagca agattgtggg atttcttttt catacctgaa agtcattaga aactataaaa
46980

gagtaaggaa aaccactttt gaaaattcctt tattttttcc ttgtatgaaa gtatctggag
47040

aaaaggagaa aaaaaaaaaa ggatatacctt ccggtaacag cagtcaataa ctgcagacca
47100

catttccttat ttcattggtg tcttctaaag cctgagttag aatgtgtcca tgggcaggca
47160

gcatgagcat cacctgggag ctggatggac ctgcagactc tcaaacctca tcctagaccg
47220

actagtcaga agctctatct taacaaggct ctcgggtgat tcatatgtcc attaaagttt
47280

**gagaagcatt gttcttctag aatagtggtt gtcagacttc cctgtacatt aggatcacct
47340**

**ggggagcttt ttttttttaa aaaaaaatga tgccttggga ggccaaggcg agcggatcac
47400**

**aaggtcagga gttcaagacc agcttgacca acatggtgaa accccatctc tactaaaaat
47460**

**acaaaaatta gctgggcatg gtggcacgca cctgtaaccc cagctactca ggaggctgag
47520**

**gcaggagaat cacttgaacc ctggaggcag aggttgcagt gagccgagat catgctactg
47580**

**taccagcct gggcaacaga gtgagattcc atctcaaaaa aaaaaaaaaa agaaaaaaga
47640**

**aaaaaagatg cctgccacct accctgggat actctgattt ggttggctctg ggtaagagac
47700**

**cagacatcgt gattgtcgaa ggctttccaa gcaagcctca tgtgcagcct cattttcaaa
47760**

**cagtgtttga aaactgctgt tctagaggat tctgggaaga tgtgtttctt tgttcggtta
47820**

**tattgactgt gtcaacctag aggccctgcc caccagagt gttacagcca tgacctgatg
47880**

**gttccagggtg ttctctccgt gttggttgcc ttgatggcag gaaaggcaga gatcaagtat
47940**

**gaccagagg tcatccagcc cctcgagata gctcagttca tccaggacct gggttttgag
48000**

**gcagcagtca tggaggacta cgcaggctcc gatggcaaca ttgagctgac agtaagtact
48060**

**gtgggtgcgt tacgggggta caggcttctg acagtttgca ttttggacac atccatcttt
48120**

**gtgattagta aatttcccca tcttggacgt gtctggtttg ttgttgttgg gtattttggt
48180**

tgtttcaactt tcctttgaag ccgacaacaa tgacttggtc attttaagaa aggtcagtta
48240

cgttaaacag tcatggggag gatcccactg agagtattca ggggtgtagg aggaagagat
48300

ggtgagctgc aggagtatag aagaaacagg aattaaagac ggtggactgc cctgagctgg
48360

gtggtaggg ggagttgaac atgaacacag tccttcaagc ctaaaaagaa gctagcatgc
48420

ctgcaactat ctcaaggcag aagaggggaag tccatgtggt agaaagtgcc attctaagag
48480

aaaagggctt gtgatagggt catgggtgtc tagagggcca aagtgtcca caacactcag
48540

tctgagggga atgctccagg tagtgattgt tgccagtgtt ttgggtgggg aggtattatt
48600

tcccatctgt ctacgtgcca acacattatt gaaaacaatg gatgtgtctg atgcctagaa
48660

catgctatgt ttgaaagcac ttctgtactt tttgtacagc tctttttgtg taagcaaagg
48720

gagacttaga agaactcaca gagaattgaa tgatgctgct ttaccaggc cagagggttt
48780

ctacttcctt tgttgtgttt gtcagtgcca ttagtggttaa ctggatgttt gaggggaaag
48840

cactgaggac acagaggccc atgagctcct gataagccgt gttctttgtg gagcaaaccc
48900

tttgctgtgg tagggcctcc ctgtgtctgg agggcatagc tgatacttgc tgggacacag
48960

ctttggtcag cccacaaga ggtgcccaac aagaggggca gggctctctgc tgcaaacctc
49020

tggatctaca tgtggacatc tctgcctgct cagacttcta tgtggaaaat ctaaagcaag
49080

aggattattg gaaaagggtgt acattttcttt taaacaattt ttttattgta gtgtttctttt
49140

ttcttttagc acacatttat taaacaaagg gataatccta attaatacaa cacactttga
49200

aataactgca tgtaaaatgt ttgtgataaa gataattgaa cacagtaatg aaaaaaaaaa
49260

gaaagaaaca gtatggagat ttgctcattg aactgagctt ggtcattctc ttagttaact
49320

cctgtccaaa gtgatgatgg aattttttatt ctactttttc atagatccga gtacaggtga
49380

cattgttcat gacacactcc accactaatt tcccatcttt caattttctt gttattgtgc
49440

tttccttccc atcccactcc tgatgctgaa ccaatgcacc atctgtaaag ttgcagacag
49500

tctgtgtttt tctgccatca gctgtggttt cttcaaactc atctcccagg gtacaagaaa
49560

actgtgttgt tttcaaagtg ctctcgggtt ttgtggtgag gtttctgcca tcacaagtga
49620

tgatacaatc tggcttggcc attgcgccca ttttttgcaa agctattccc actcctagct
49680

ccttcatgta ttcataaaag cctttgctgt ccagcaggcg ccatcttcct tccagctgct
49740

gaactgtggc catggtgggt gcgggcgggc tggcgtgcag agcaggggtct acgtcggcat
49800

ggcagcatgc tgtgtctgta gtgtttctttt tttaaaaaaaa ccatgctgta tcaatagttg
49860

cattgtcatc acatattcag attttatgtg aatttctgca ttttaactagt cttctaggaa
49920

gtgctgacca ctgatactca gaagagtgaa ttgtctgttt gcttcatatt tgcctaatac
49980

atatcccagg aatagcacgt cacacttaag acttcccttt cccatgtcag gatacacaac
50040

atggtacggc acaagcgggtg taggctttaa aaaccgggtt tgagttttgg ccctgtcgct
50100

gcagggatac ttctgaacct ggtctgagaa gactcacctc atgctgcaga tggcagtggc
50160

tcagagcaga ttctggagcc agcccatctg gggacgtgct tcagctctac ccgtcactag
50220

ttgcgcgact ttggacaagt tcctaaacct ctcagtactg cggttttgag acttacctct
50280

taggtttggg gtgagtagca caatgcctag gacagcacca gccatgttgc aagtgcaaag
50340

taaatgctac ctgttattgt tgttgttata aaaacaaaag gagagcagag aagtaccatt
50400

gaaatatattg gatacttgat ttttaaaatt ctgttggcag tcacctgcca acagataagc
50460

tcctaagttt tcctgggtggg ccttgcagtc ccattttttc cttgggtctt gggccaagtt
50520

gaatgcattc tgattgcaga tgagctcccc tcttcttcct ccctaattcc agatgggacc
50580

cttggctctt tttgtattta ttctgggata actgatgcga atttcagcat gcagctccag
50640

tcacatcatt ccactgctca ttaccctca gtgacttcgc atcaggtgct agcttagcaa
50700

cacttttagct ggctgcttga gagccttcag gctgtggacc catccagccc tgcagccttc
50760

cctcctctcc ctcgagcatc ctgtgtccag agaagcctgg cttttgatgt ttgcctgaat
50820

acatcatgct atcctgcctg tcgtcttttt cgagcttctc cttctggtag tgtccttccc
50880

ttcacttgtc aatagcctat ttgtaccttg ggaccacctt tgattacatc tcctgcatca
50940

gccacttttc tcctgttaat aagaggtatt aacatcatta cccttggagc cagcagagca
51000

ctaaaatagg gaggatgcta gcagcagctg cagaccgcat gccagcaaca ctcaggatta
51060

cggagagtat ttcaggagtg cagcgagttt cacagaagtg cacaaccccg tacaggcatt
51120

agtatcccgg ctttgccagt gaggtaacag acagtggaga ggggcttata aaagcccatc
51180

tgtatcagag ccggaagtca agccttggtc tgttgccatc tgcttcacga ttggctgtct
51240

cttgtacctg ttttggggaa gagctactct ttagtccctg aagccgctgg aggaacgttg
51300

agaatctcac atgcgggggtt tggaggtgga gtccagggtc ttgagagcag tgctgaggag
51360

ggaaaggctc ttggctgcct gttacctaga ctccctggac tggctttcac aggctttcct
51420

tgatcctggg tctgtgggat tcttgccatc ctgtgttgca gatcacaggg atgacctgcg
51480

cgtcctgtgt ccacaacata gagtccaaac tcacgaggac aaatggcatc acttatgcct
51540

ccgttgccct tgccaccagc aaagcccttg ttaagtttga cccggaaatt atcgggtccac
51600

gggatattat caaaattatt gaggtaagta attcattaaa aaattgtagt caccttttta
51660

aaaacagtaa tatataatca gtgaagaaaa atgagaaaaat atagctaaga aaaaatgaaa
51720

ataaccaaaa ttcttcaact tttcccatgg aaataagctc agtcatttta aaataaactg
51780

caggtaagta aataataaat aatctgaggg aggatatcag tgaatgggta agagaaagat
51840

gaggaaaaga cattattatt gttttaattt tagatattcc ctaagaaaca tcaatattgt
51900

caagataccc tgtacatagg gttaaagtgc ctaattgtaa tttttaagtg tgatctttgc
51960

ccaaacgaag tatctgtaga gcaaaggaac gttggccttg gaaaaaatga tccacatttt
52020

gccactatgc caggccactc ttgttgtgga gagcttggcc gcttctgctt tggaaaatcc
52080

gtgagaattt tcatgcgtgc cctgcctctt actcccttcc cagtttgtcc acagacatgc
52140

tatttccttt ttacatttta tacaatacta ggctataaga gaatgagcat cagtggatgat
52200

ctgaatgggt taaggcagtc aaaactgaaa aagattaact tgtaacagct agtaaggata
52260

aattaagttt agtgggaact acatgttata aaatcgctta tgaaatgaag cctcttatag
52320

caggcttaat gtagtatacct gaatttcctc ttcatttgtgt gtgtgtgtgt gtgtgtgtgt
52380

gtgtgtttta agagagatgc tgtctgttat gtggcctggg aaagcttttt ttttctttct
52440

ttttttgaca gagtctcatt ctgtcgccca ggctggagtg cagtggcatg attttggctc
52500

actgcaacct ccgcctccca ggtcaagaga ttctcctacc ttagtctccc cagtagctgg
52560

gactacaggc gtgcaccgcc atgcctggct aatttttgta tttttacttg agatgggtgtt
52620

tcgccatgtt ggccaggccg gtctcaaact cctgacacct taggtgatcc acctgcctgg
52680

gcttcccaaa gtgctaggat tacagatgtg agccactgtg cccggccttg gagagttttt
52740

tttttttttt ttttttttga gacgtggtct cgctcttgtc ccccaggctg gagtgtgatg
52800

gcatgatctc ggctcactgc aacctccgcc tcctgggatc aagcgattct cctgcctcgg
52860

ttccccgagt agttgggatt acaggcacct gccaccatgc ctggctaatt tttgtatttt
52920

tggtagagac agggtttcat catgttttcc aggctggtct agaactcctg acgtcagggtg
52980

atccacctgc ctcagcctcc caaagtgggtg ggattacagg catgagccac cgtgcctagc
53040

cccgggaaag ctttttgaca tggaaatgat catgctgctg atgggtccagg gccccaccct
53100

tggaggcaga cagcatagct tttttatcag ctgaggaaag atgctcagag aggttaagta
53160

actttactca caatacccag ctcacggaca gcagaactgg tgttcaaacc caggtcttgt
53220

gcttccgagc tcagtggcct ttgtgagtct agtacaccac aaatcttaaa gaacagtagt
53280

gatatggcat tgatgaggct cagacatcag actgatgcag tcaacagcaa ggagagtccc
53340

atgtggtgct tttgaacaca acttagtcac catttctgat ttctttcttt cttttttttt
53400

ttgagttggt gttttgctct tgttgcccag gctggagtgc agtggcatgg tctcagctca
53460

ctgcaagctc tgcctcccgg gttcaagcaa ttctcctgcc tcagcctccc gagtagctgg
53520

gattacaggt atgcaccacc acgcctggct aatttttgta tttttagtag agatggggtt
53580

tcgccatggt tgtcaggctg gtctcgaact cctgacctca ggtgatccac ctgcctcggc
53640

ctcccaaagt gctgggatta cagggtgtgag ccactgtgcc cagccaaccc tttttgattt
53700

ctgatacctc acttgcttct atgctaccca gcctgctaag tgtggtgtgg catccatgct
53760

gaccttattt ggggactatt gctaattgtct gtcctccatt cttgcatctc tggggcattt
53820

tttgctaata taattaggct ggtattactt gtaaaaggat ttgtcttttag acagtctcag
53880

ctgctcttcc tcccacccca actttttaatt tttgtgtaaa agtcccagca ccagagttct
53940

ggggcccttc tcttaattgt gctaccagac aagagcattc gtggcagttt ttcctgtgtg
54000

tattcataaa gagtatcaga gatatagtgg tcagatgaac gtggtaacaa agcgcttgag
54060

ttaaactgtg agcaaagcac aagattgcgg tgtaggctct gttggggcac agtggaagct
54120

gtgcttccaa atgctgaagg agccgagttc tgccgcagtg gcaagaagcc tgggctggaa
54180

gctggaggct ggagctctga gtccagctct tctgagaagg gtggggcttt gggcaaggta
54240

tttatcctct ccgaccttct gtttcctcat acatggaatg aagatttctg accattaggt
54300

atctgggatt ccttcactca aaaatgctga gtttaaatta taagtcagca ctgctcagct
54360

ttggagtctg gggaagctgt cttcccagaa gtgctgaact tctgcacagg gtttgtaggg
54420

ctattgggta aagaagttgt aagcagaaaa cccacaaagt ctactgaggc acttttagat
54480

tcacttttcta atttccaaag ctgaaaagtg ctttctgcca atgcatatatt taaccaagta
54540

ccttcctcct ttttccccac ccctctcttt ttaatgacaa ggaaattggc tttcatgctt
54600

ccctggccca gagaaacccc aacgctcatc acttggacca caagatggaa ataaagcagt
54660

aggtagaaca caaaagataa actccagctc tcacttaagt cccttcctct acctgggccc
54720

aactctgcca gctgggtcttg tcttcccatg gtgccttcct cctggattag ctgcctttac
54780

ccttgtaatg tgaaccctct gaaacagtgc tgtccaaaag aagttctgga atgatggcag
54840

tgttctgtat ctgcactgcc cagtgtgata gccgctagtc acatgtagct agtgagcact
54900

tcaaatgcag ctagtgtgac tgtggaaatg aatttttaat ttttaaggaa tttcaattta
54960

gctaccttta gctaacagat tggacagcac ttaagttggg agggagggag agagtgtaga
55020

ttttttttcc tttgacaatg ttttcagact ttatttgcac gggttctaga actagaattt
55080

gaaggcaggc ataatggctg agtggaaga aaatggccca ggcgcgggag acctgagatc
55140

attcattatt tcagatggcc agttactcat tcagcaccat aaaggctgct gttgtgtgca
55200

gggcagtggt ctaggtgctg gggataaata ctttccctgt ctcgaggagg cccatagtct
55260

ggctgagaaa gagtgagtga gcaatgaaat caatggctct ggcagattct ggaaggtcta
55320

ggaaaactcc aagcagggga cataagatct gagcctcaaa gaatgagtgg aagtttgctt
55380

ggctgaagga attggtggcg gctacccttg gggactgctg tggcttggtt cagtaggcat
55440

gcctggggca caggaggagc ttacgggaga ggcattctgaa cagccagttg ggggtctggt
55500

tcttgggtag cagtccaaca gcctaagact tcagcctcaa aggccagtgg ttttcatgtc
55560

atgttttagg tgcatttgag gttccacaac cactcggttg gaatttcatg gatgctgcag
55620

ctaacatgag accacagctg tcagccatga cccccgattc caaactatgg tgttcattga
55680

gacagccagt gatcactggt ctcatggact ctattttaaa taattgttta aagatagaat
55740

gtataagttt atactaataa aatactgctt ttatctattt ttttctttgg gggctctgtat
55800

aagattgcat tttgaaaaga aagctatagg caaccatagg ccgaggatgg taggcacctt
55860

tggggctaga ttctcagctt tggggcctca cttgcctcac ccgtaaattg gaggtaacca
55920

cagtatgtga aggcaggggc tgggcctcaa catctttgga acctccttat agtgatgttt
55980

agacctctag atgctccctc agatggccag tgtcagagag agagagttct tactttcatt
56040

ttaacccttg gtggtctgtc ccagacatgt gacaaaggca ggtcttaaac tgtgtcctca
56100

gaaggggagt ggcttgtaat ccagggtgaca agcagcatct gatatatctg tgttgctgca
56160

tttgctttcc aggtggaaga agtctttcct gtgcagcctg gtgtttggca tccctgtcat
56220

ggccttaatg atctatatgc tgatacccag caacgagccc caccagtcca tggtcctgga
56280

ccacaacatc attccaggac tgtccattct aaatctcatc ttctttatct tgtgtacctt
56340

tgtccaggta tatatgagaa agtgggcaga cctctccctt ccatgctgtg tgtggccctc
56400

agatattctg cccgctaagc gcaaacatag tgcttttagct tggtttaaataa ggcactttat
56460

tgcagctttc ctgccccacc caggcagttg ctgctttttt tctctcttaa tgaccagtag
56520

gtggggcctt gggccccccc agtgtggggg actactaagg gcctggtttc actgtcagtt
56580

tcagttgatt gcaagtgatc agaactggcc ctggaatagc aacctccttg ctgcaagggtg
56640

tggtgaactc tgtgtctcca atgtgcatat tgtttagggt aatttacgtt tgttgctgga
56700

agcaacaact ggatcaccct tgagcccatg ccactcttga gtgttctatc ttaaaagggtg
56760

gagatatggg ccttgtcttt caggaagtta taatttaaaa gagaaaggca tgaaacgtct
56820

gctttggtgg tgatattaag gaactagctg tgtgtctggg gggaagtgtg ccctctatgt
56880

gtatatTTTTg gaatataaaa ctaaattccct gtcctaatat tggaatgttt aggggctatt
56940

tgggggcaag cattggtagg ctggattttt tttttttttt ttttttgaga caggggtctca
57000

ccctgtcacc caggctggag ttcagtgatg tgatcatagt tcaactgtagt cttgaactcc
57060

tgggctcaag cgatcctccc acctcagcct cctgagtagc taggaccaca agcgcctgcc
57120

accatgcctg gctagatagg gtcttgctat gtgcccaggt tgggtctcaaa ctcctggcct
57180

tcaacagtcc tccaagctca gtctcccaaa gctctgagat gacagggtgtg agccactgca
57240

cccagcctga atttgtgttt ttctatatatt tgtttgatta taagtgtaat gtacacttat
57300

tattaacaat tcaaaccgta ctgaagtata agaagaaaag aggaagatgc cctttcactg
57360

atcacattct gcctcccgcc caatctcatt ccccagtggt gctatgggtga tattaatagt
57420

tgggcattaa tctctccaaa tctttatata gagaaaatgc acacttgcgt gtatgtgtat
57480

gtatatgttc acgtaataca gatacagact tttttaactg aaatgggatc atacaagata
57540

gactgatgtg cagctggcct tttcacttaa ctatgtatta gacatctttt atgccagtgg
57600

gttatagata gagcccatc tctttaggga atggaatgcc ttgtgctgtc tacagtacaa
57660

atgagtcgtg atttttttcc ccaacttttc accattctct tattagagga catttgtgtt
57720

gtaagcactg ggtttttaaag gtctcacatg ctcttgggtca tccattcctg tggacagtag
57780

tcctctgaat gggaaagtat atttcataaa cgcccatcac agaggaagaa gtactgtcac
57840

gactgtgcac aaagctagag gctttgccat ccccagggcc cttggccctg tgtcgctcat
57900

tgaactctcc tccctacttg ctggcagcct tcaactgtcct tgtctttcag ctctcggtg
57960

ggtggtactt ctacgttcag gcctacaaat ctctgagaca caggtcagcc aacatggacg
58020

tgctcatcgt cctggccaca agcattgctt atgtttattc tctgggtcatc ctggtggttg
58080

ctgtggctga gaaggcggag aggagccctg tgacattctt cgacacgccc cccatgctct
58140

ttgtgttcat tgccctgggc cgggtggctgg aacacttggc aaaggtaaca gcagcttcag
58200

gttcagaaaa gagctgctcc ttcagtaaac aaatctcact tcctctgaac accatgttta
58260

gaattactaa ttatacacag catagagaca gacttaaaga aataggaaac ctccatataa
58320

ttaaggtgct ctagtcaact atctccaaat tggtcactac ttctgaaatc ccagctaata
58380

tggttaatta ttaaaacccat cattcagggtg tattgtgtaa actaaagaaa cctggccttc
58440

agggcagagc ctacattgtc tcttggtgca aaactgaata tgtagttggt ctagacatcc
58500

atacatcctc ctcaagggtca tgccattttt aacacccttc attgtcccca acactgaaat
58560

gtatcttaca agagaatgga actcaggccc taaaagaaat tagattgaac ctaaaagaaa
58620

atgggtcagtg ggaagactga tgtttgtttc cagtctgaaa tgtaattccc ggactttagt
58680

ggcacaggca ctgtgcatca ctgtggaagt gacatgtggc catgtgtggt ggatagcaag
58740

taacgcccac ctgcagagcc ttttatcgtg ccgtgcggct gtttctctcg caccagctgt
58800

ctctaacacc acgcttgtga ctctcaggct gggtttggac aggtctgctt tcgatagctc
58860

tcatttcaca ttctggttat ttcctagagc aaaacctcag aagccctggc taaactcatg
58920

tctctccaag ccacagaagc caccgttgtg acccttggtg aggacaattt aatcatcagg
58980

tgagttatgg ttatcaaatag tctttgtggg ttggtatctat caatctgtgt gagctgcatc
59040

agatgcccac gttgtattga cattgcaata gaccttgtga gtgtgggcag agacacagta
59100

agatcaccac tctcaatcca gctacgaaag caaggcattg aactataaaa ctagcaagat
59160

ttgtaggtgg tgtgtgtgtg cgaaatggca catggggagggt gtagatgatt agtattccag
59220

gaattaagaa gaggaaggcc agtgttggtc aagggaact tcacagagaa catagagtat
59280

gaattaggcc ttcctagtaa gatacacatt cagacaggca gagagataag gacttcaact
59340

tgctaaagaa taaaaattga aggagaaatt tatcttcttc ttaagggtgc ctgatattaa
59400

aaataataat tattattatt cctggaaact cttgttgcatt ttattgtttg ctagacacta
59460

ttctaattctc tttgtattta atctaaccac agatggaaac tttctgaggc aggactatt
59520

attactctta atttacagat gaggaagctg aggttcagag aggctaagca actggcccaa
59580

ggtcacacag gtattcagta atggaattgg gatgtaaccg aggctgggtg atccagagtc
59640

tattttctta acctttacac acactgactc ttaatgcata ttttaataata tggaaatttt
59700

tctaattaag gcttctgtat ttccttgggg ttagcttctt ttttatattc taggcaatgg
59760

ttatttcagc cctccttcatt attcctaaaa tctctccctg gtcccaccac attttctatt
59820

agcctgtggg ctgcgtttgg actactggaa ggagaagtgt tcagatggaa accctccagt
59880

tccctgtccc tcagcctaca cagctagcct cctccgtggc acctcccgt cctgccccgt
59940

tggtgggaga ggcgtctcct ccaatgcctc taccgggaa ctttctctg ttggttttct
60000

cagtggtttt tggtgtctct gattcttata ctccgatgtc atgcagactc taaccactcc
60060

cattttcaga ggcgttggtgc cttgcccctc ccttggtgtgc ccttggtccat gtatcacact
60120

gttgatggag ccagcctgcc ctagctgtgt cctggaggcc tgggcgtgaa ggctgagaca
60180

gacccttgc ttcccatgg ggtctgcagg gctgggtgcag ctggccctgg gcacgcccc
60240

ggtgagtaga ggaactgacc ggggccacac agatgcatga gtgttcatga ggtgcaggct
60300

ggctgctgcg cgggtggcact ggttttggtc tggatcctgg cttttaaatg atcatcaca
60360

ttgtaggact ggcaactgag atttgagtcc ttctgggtgtt ccaggctcac tccccagcac
60420

ttgccacaga tcgtctagtt ccactctgta gctgttccgt gaggcagcac tgtcatcatc
60480

ccagctttcc tcaggaaaag ccaagtgcag tgagggtggag taatatcccc agagtccatg
60540

gcaggcaggc agcagagctg ggattggaac cctggaggca cagagcccc ctcagccact
60600

gtgctgcgct gccccaaac aggaccctgt gtgacggcac ctggcatgct gggctctgct
60660

gtgccatggg tctgcagcag tcacatgcct tggtttcttt tttttttctt ctcttataaa
60720

ttaattgcat gaaagtttct gagaagtga aacgttcaga caggagacca agccgctggc
60780

agctgtgttg gttttgactc atgtttcctc caactcgcct tgctagggtc ttcggtagacag
60840

tttgaggcgc ctctgatggc actgactgca gttgcttgga actattccag ggatctgaac
60900

aggaggagaa ggcattctgcc ggccgaagtg cccacggggt gcgtgcagca cctgtcataa
60960

ctcagttctt ttccaagaat gccgattcat tttttggctg gggacgcacg gcctgactac
61020

taccatgcaa catgacagga cattattaag ctagggggcca gaagaccaga gttcattctg
61080

cctttggctc actgtttggg tgtgggtcatt cctacttgcc tcctgacagc acaccccgag
61140

cccaaccctc tgtcccaggc caaggctgca cccactgccg cctggattgc cgcagtctag
61200

gcctcccctg agctcaacca ttctgtgctt tgagggtgcac tggacactgg tttcccagtg
61260

actctgcctc ccacagggcc ctcccgcgtc tgggcctctg cttgtgcctt tgggtccatgg
61320

caccatctga aggccccag aggttctcgc ataactgccc tctttgactg aggcttgggc
61380

ttcttcctct gaccatccct ttgctgggga cacttggtgct ttctctttac tccaagtctt
61440

gggtgctggc agcgtttcta ggtgcgatcc gtctggcaca ctgtagtttc cttgctgctg
61500

ctgctgctta gtctccatgg gctttcttct ctcccttttc cttgtggatg gtccttgggg
61560

ttctgggtgg actgcctctg tctttgtgtg ttacacaccg tcactgggac gtctcactct
61620

cccacgtctt tggttaacag ccgtgagaag ctcatctctca aacctacgtc tagatcccag
61680

agactacctg ctgggttgag tgtttttctt ggttgactgg taaaggctca gactcggcct
61740

gtctgtattc ttcttcacct tccccactt tttccccaaa aatattcccc ttccttggtg
61800

cctgtctgtc tcaggcaaag gtaccctcat gccagctccc agtgtgtcct ctacaatgcc
61860

tgtctgccct cattcacaca tctgcagtcc atcctcatgg ttctagcacc tctgtctcta
61920

ggggcgatcat ccactggcac catgctggtc caggccactg tggccttgcc cctaaaacac
61980

cgcagctgtc tcttcatact gcttgccagcc atttatcaat ccagtatttt tccacattca
62040

aaccagtgat cttcctgaaa gcacaaattg atcatattac agtgctttct gaaagcacia
62100

attgatcata ttggaaaatc cttcaatggc tgcattgctgt ccttaggattc aaaaccacct
62160

tcctaactgt gcagtgaggc ccggcatccg cctgcatgcc tgttgagccc agcctgtctt
62220

gccatcccct gcgttgtgct ccagctgccc tgcaccactg gcagggtccca ggaagggact
62280

gcctctctct gcagacctct gcacacatcc ttccttccct ctgaactccc ttcttcctta
62340

acctttgcct ggctccttgc cttcagggtc cagcatggaa gttgcatcct ccagaagcct
62400

ttcctgattc cccaggcctg gctcagctgt ggctcctcag tgctctcaga acactctgtg
62460

cctcagccac cctgggatgt aaacatggcg ggtgctgaga tactcactgg tggaatgaaa
62520

ctgctgtcag atgttaaaga aacatactgc acacatgcat gtgtgtgctt gatattaaaa
62580

ggcttgcaca gccccatccc cagtctccca tgttgaaatt ttctggtttt tttttttttt
62640

tttacaattg atgggggaaac aaattacaaa tcatatctag accagtgacc ataatgctat
62700

aagttttttc tttgttttct taacttgtgt ttaatgagga aggggagtag aagagaggag
62760

acaagtgaag taaacagatt cctggcagaa actcttaaag atgaagtgct atgaaacaaa
62820

ccactgtctt ttaattaaga gggttcctat cgctcagcac ttggagagga atagtagctt
62880

ctgaagaata catgctattg agattcacca gtaggagcca cattaaccct cataaaggac
62940

tattccttgt aagctgctta gagttagaaa gatgtgggta atttagatca aaggagaaat
63000

tcaaagttag atttaacaga gattagcatt actaaagatt tcttagaaag acacccaact
63060

gtacattcgt gccaatTTta agtctcctta cagaagtgga cattttaaaa cttagtgatt
63120

aggttctgct ggaggagggt ctgaccttta gttcttgggg tctttcattt ctaaattgaa
63180

tgcagggagc tcctcctccc gacttgggct tctcaggcca tgtggagacc tctgtgggat
63240

gataaatcct ccaggttggt cagagggggc ctcaagggca atgttcagcc tggacttctc
63300

catcgctgga ggactttaaa attaagcaga gaatattctt gggttctcga gatccttccg
63360

ttctcctggc tcactcagcc accaaacaga ttacacagg tatactccca gcactaaaga
63420

cttcagtctt tgaaagtctc cttttcagta ctccatttat tttggttgcc atgattgaac
63480

tttctgtgtg tgtcttaact gtccatctgt gtgtctgctt ttagttatga ggctataaac
63540

ctcctgagag gaaggctcta ctctgtcctg tttgtaaccc cagcctaccc tgggtgcctgg
63600

tacacaagag gccaatcaat atgaaaatgt caattaggtg ataataaaaa tcaccagcat
63660

tcattgggca cttactgtgt acaggtgcc a tactaagcat ctcacgtaat cacaggaagt
63720

agccacacaa gccttgaatc ctcgtagcac agagactggt cctgaacaat ttttacaag
63780

gaaaataatg atgcaaatat ttaacataca tgccctgcata tggtagttac taagttttat
63840

atagttgtac atagcaacat ttaaataagt gggcatgtat gttcaatttt gtgcctttct
63900

aagacatctt caggttttgc tgaaatgggg gtaataagac tcaccttgta tgtgtctggg
63960

aagattcagt gaggtaaaac atgtaaaaat ctagcacagt tcctagcact cagtaaatgt
64020

ctatttctcc ttctcaagta agtaagatgg cttttgtgcc tgtgatagtt taaatgtagt
64080

ctttgtagaa tgaaagcaag gtgtaaaata acctcagagc cctgcttgct ttccatgcat
64140

gctcttctgg cttgttgacc attagctcat gagccagtat tacctataca attaattctg
64200

ttggggcccat gccatgcttt tagtgagctg ctgctgaagc tggccctggg ccgtaggtgc
64260

cctccctctg gctttgtctc tgttcctgcc gtcttctcct ctccccccgc agcatctctg
64320

tctctgcagg acatgtccca ctagcacaca gttcacatgt atgtacaaga cacaatcccc
64380

actttaatcc accctccctt ctactgctct gtttttctgt tccttcacaa aagccttctt
64440

taaaaacgtg actttcgctg ttgtctttgc cttctcactg cgtgctctct cagcccatc
64500

tagttgagcc tctgccttta ccagtttctg aacttggttt ggccagtgcc atcaagtaat
64560

ttccatatac cccaatccaa cattgcttgt gttaagtcct catttcagtc tctcaccagt
64620

atttccccct tgtctgtttc tttgattttt tttctttttt caatgtgtaa ctcacagtca
64680

ctgcaacaca caaaccctac gtgtagagtg tgatgaattt tacgtatatatt tgcccctgtg
64740

gaactacttg tgcagcgttt ggtgtggttg gatgggtgct ccctcctcct cctcctcatc
64800

cccactctca ttttcctcct gtgtcactgg ctggtccttc agcttctgct cagctttcag
64860

atagtggagt gtctcagggg tgagttctag gacccttct gtgttttctc cgtcactcct
64920

aggggcgctt atatgatctc atggctttta atgctatcta catgtgacag ctctcaggtt
64980

tgttctctac ctctgacctc ttctttccac ttcagactca tattttcagc tccctagtgt
65040

gcatttctac gtggaagtca aataggcatc tcaaaccgga catgtcaaaa tgaaccattt
65100

ctctccatcc cctccccctc tcttccctca ccctccctta gctgcttctc cataatcatc
65160

tccatctcat taaatgggtt ctctatccac ccagttgctc aaaatgaaaa tttggagctc
65220

atctttgatt cctttcttta ccctcatatt cagtgcattt tcacatcttg ttaacacatg
65280

atctaagaat atgctgtgta tttctttctcc ctctctctgt agtccagggtt acttccatct
65340

tccgcttggg ctgctgcaac agcctgactc tccccctcac tcctgtgatc atccctacag
65400

cccattctca tgcagcagcc aaagagatca cttaaagtgt aatcagattg tgtcactttc
65460

ctgctaaaag cttacatcc aaaacccatg ctctgctcca ggagactcgg catgatctgg
65520

actctgcctg cttctctgac ctcaacttgc acgtcactta cttcctgcca gccttgcttg
65580

ctttcttgct gtttctcaga cgcacccttc ctgcctgcct taaacctttg tgtcagttgt
65640

tttctctccg tgaacttcct ttccgtgtgt gttggacctt cttggcatgc aggtctcagc
65700

ttaaatgtta ccttcccaga gaagcccttc catgtctcct cctccctctg ccactctaag
65760

aagccctcca gtcacattat ccttggttttt ttcacacag cacttattaa tatttaatac
65820

ctgattctat cttgcttggt gtccatttct accacagaac ttgtcttcat gaggtagtag
65880

ggcctttgct gtgccattca cagctgtatc accagtgcct agaacagtgc ctggtattca
65940

gcagctgcac gataaatatc tgtaaatgaa cagatcaaata gaatactggt gctcagtata
66000

agcaaataca gtgtaactat tgtaacagct ggcctagaac ctgacccggt gaccgaatga
66060

gtggccatgt gagtgataag tggcgtttgt tgcagggagg agcaagtccc catggagctg
66120

gtgcagcggg gcgatatcgt caaggtggtc cctgggggaa agtttccagt ggatgggaaa
66180

gtcctggaag gcaataccat ggctgatgag tccctcatca caggtgagat ggcttgtttc
66240

atgttccctc aggaggatat catagcagct gtcaggtcac atgagtgctg gatggggctg
66300

agcaagtgac agttgtctct ttcctacgtc taggagaagc catgccagtc actaagaaac
66360

ccggaagcac tgtaattgcg gggctctataa atgcacatgg ctctgtgctc attaaagcta
66420

cccacgtggg caatgacacc actttggctc agattgtgaa actgggtggaa gaggctcaga
66480

tgtcaaaggt aatgaagaaa tttttaaaac taacttcac tttctcgttt tagaaattat
66540

gtgaagagtt ctgggaaatc agacagtttt attgagtaga gattgattag taaatgtggt
66600

taaatgaagg agattatccc aatctttatc catgcttggt gtgttttatt tcttcatagg
66660

ttgtaatttc ccatgggtctt ggtgttttat tttcataggc acccattcag cagctggctg
66720

accggtttag tggatatttt gtcccattta tcatcatcat gtcaactttg acgttggtgg
66780

tatggattgt aatcgggtttt atcgattttg gtgttggtca gagatacttt cctgtaagtt
66840

gaatgccttg ggctatatgg tggttgtggt tttaaataatc tactgacatt gatcctgttc
66900

tttcatatct tagattcact gggctttaat tattcattac attttatttg cttgcttctc
66960

ttattgacag caaaatctaa gccagagtag gataaacagt cactctcctt ctgcaggttt
67020

ctattaaaga tctcctttac attctactta gataattttc acagacacca attgggtgctc
67080

agttgtaatg cttgaagcat ttacaatatg tattattggt cctttttgag gttatgaggc
67140

atacatcata gaaaatagtt cactttaaat ctcaatggca tgataatatt atttaataca
67200

ttttaagaaa aatatctttt gatgtatata ttaccagta ccatctttga cctcccagct
67260

tccaagtaa tgtaaatagc tttttatggt tgtttgttta tttattttatt tttttgagat
67320

ggagtctccc tctgtttccc gggctggagt gcagtggctt gatctcggtt cactgcaacc
67380

tccacctccc aggttcaacc agttctcccg tctcagcctc ccaagtagct gggactacag
67440

gcacacgcca ccacacccag ctaatttttg tatttttagc agagatgggg tttcaccgta
67500

ttagtcaggc tggcttcgaa ctccctgacct caggtgatct gcctgtctcg gcctcctaaa
67560

gtgctgggat ttcaggtgtg agccactgca cccaggcgtg aatagctttt tataatagtt
67620

aaatattttac taaacattgt atgtgaggca atggaatata catttttgag tttaaaaaag
67680

tggttgaaga acactatctt aggaaatatg gttttattaa acctcagaaa aagaaaggaa
67740

aatacaggcc acccctcaaa aaaacatatc tgcttgagga ataaaagcta tattacatct
67800

gtgattctca gttggcagct gtgacatttt cgggtgtttgt gactgggggg atgctgctaa
67860

acatcctgca aggcacatca caggccctac aacaaagagt tatctggccc caaatgtcag
67920

tagcactgag gttgaaacct tgataatgta cttaaaaagc taatggcaca cttagaacag
67980

tatttggttaa tggcacaatg agcggaacaa tgtttttacc ttatcttatt cctagcccag
68040

tctagtgcctt gttataccat agagccatag aatgagaaac catactgggt gtggtgatgc
68100

acctgtagac tcagctactt gggaggctaa ggcagaggat catttgagtc caggagtttg
68160

aggctgtagt gtgctatgat cgtgcctgtg aatagccact gcactccaac ctgggcaacg
68220

tagcaaaaag tacaataaca gaatggataa tttttttctt ttaaaaaaaaa tccttcacat
68280

aattaaaagt atatcaaata ttctcaggac acatgtcatt tactaggatg tattttctcc
68340

ctcaaatttc ccatcttaaa ttgaacaatg ttcaagttaa acactttctg acatctaagg
68400

aacctcagtg ctcttaaaag taacttacag atatattggt agaatgtgcc tgtggctatc
68460

catggcacct gactgcctat aatttcggct ttttcgggaa agcagtgcggt tggcatatgg
68520

ccatgggaag actcagtcct aaacaagttg gaatctgtgg ctttctgctg ttccatgcct
68580

ttccatctta ctttggcttc ttttgttttt atctgcaata tcagggagga tgctggcaaa
68640

tgaggtagtg catgttcctt taagatactg tcactctctt gctgtacaaa agtcattgta
68700

gagaagggct catggtaact tccagaagtc caccagtc ttcagaaatg ccagggtgtga
68760

ttttgacaag aattagtgtg tttttaaaaa agattacata atctttatgt ctaaaaatct
68820

taatatatcc ctctacctaa atataactgt ttctatgaaa attaaaagac ccctttaata
68880

atacattatt ttagaatcct gcctttcaga gtgaaataat aaaatgtaga tggaagctaa
68940

agaggctgtg gtccatagcc ttaccttaga caactgtgct ttttactata agtaagtgaa
69000

catagtatct cttgattata ctccatagcc accagccatt tttaaaaata aaccaagata
69060

attcaagctc tcagaatgca aataatacca agtgtctgga aagtaggtca tattttgttt
69120

gttttgggta caaagaaaaa atagtgaaaa atacagagaa gaaaataaag atacctccaa
69180

atcatatccc tagaagtaat ttttttgcatt attggtggat acctttccag tattgttgaa
69240

atgctgtatt cagatacagg cggagatttt tgcacgaata aatcccagca catgcgattt
69300

tgtacttgct ttttaatttt gcagtgctgc tgggtacctt ccatatcagt atgttccatc
69360

atcattccaa tagttgcctc gtatcccatt atatacatca cacctacagt ttaattaacc
69420

aattccgtat tgggtgtttc cattttttat cattacaaaa acaaaacaaa acactgtggt
69480

aaatatcctt gcatcttttc ttattttatat gattatcacc ttagaatagc atttttaagc
69540

tttcctaata atgagaaaaa taacgataat agttactggt tcttaaatagc cgctcctatg
69600

ccagggtgta tgtactgtgt atattttatat gtttatatgt gtattgcctc attcaatcct
69660

tgccacaatt tatgagctat gcattatcat ctccattttg tgtatgagga aaccaggatc
69720

tggatctgtg tgggtccatct ccagagtctg ctttccccta tcaagggctt ttatttgact
69780

ctgctcctgt aatgcctctt tgagggaaac ctgtaggatg aagttggata gctgggatgt
69840

ggagagcagt aacgtgttct ctatgatggc agagcagtgt ggaataccat ctgtttccgg
69900

aacccaagtt cgtcacgttg tgtccagtgc cccctgaaa tgccttatg tgattagagt
69960

tctgggagct tccttattga actctcaacc tgcctctgac tctgtcctgt tttcagaacc
70020

ccaacaagca catctcccag acagaggtga tcatccggtt tgctttccag acgtccatca
70080

cggtgctgtg cattgcctgc ccctgctccc tggggctggc cacgcccacg gctgtcatgg
70140

tgggcaccgg ggtggccgcg cagaacggca tcctcatcaa gggaggcaag cccctggaga
70200

tggcgcacaa ggtcagcctg tagcacggct ttcccatcc tgagagatga aagtagtacc
70260

tgtttactat ttcacattga gagaaaagcc tgagagccac tcaagacagc agtggttaatt
70320

acatagaata ggaagtcaag tataactggg aataacaaca gtagcaacag agtagccacc
70380

agtcataata tggaacgctt tatcaacatt acttgattta attgcattaa caactctgtg
70440

tggtaggtgc ttatcattcc cattttacag atgaggaaac tgcagcagga ggagattagg
70500

taacttgacc aaacgcctgc agctaaccag aggcagagct gggattcaca cccaggtagc
70560

caggcttcag gacaaaagcc ccaccgtgac atctgtggga ttagtggtgg gtcctgtgac
70620

cctggggagt gaagagtgag tgacggtgtt tgtgaagagt cttccttttg cccaacttac
70680

taacctcttg ccttccacac tttcacctat ttaatcagta tttgctgggc actcactgcg
70740

tgtgagccct gcactggggtt ctggcgatgg tcatggtttc cagccgccaa ggagcccacc
70800

agggggaaaag acagacagtg aaaaagtaaa tagcaaagct gactgagtat tacgagatac
70860

tgagtattga ctggtggtgc ttgacgtcct ggggacactt gaacaggggc caatcctaga
70920

cttgggggtgc tcgactcagg gacaagggtg ctgcagggtg tcaggggaagc gaatacgaac
70980

gttcatctga gaccttactg ttgagtgaga gctggacaac ggagctgggg aagagtgttc
71040

caaaatgggg acagaagggtg gaaagcacct gtggaggtgg aacttcccct ccccgcggtta
71100

tccccagcat ctaacctgag acatcaagtg cattcagaag ttatctgttg aatgagcaaa
71160

tgaggaactg gaagaaaatc attgagagag aggaaccaga gatgaggccg ttgagggagg
71220

ggccagaatc caaaggctga agatttagta aagtgttctg aacgttagcc tagaagccac
71280

gtaggggatg gcatacccac gtgatgtgac tggatcgtgt gttgggaaga tcattccggc
71340

tgcagtgtta aggttgcgct gaaagctggc aagaatccct ggaggaagat gagttaagtg
71400

tgttgaagga aggaaggcca ccgctcagca tcagaaacaa aaagctgaat tattgtttgc
71460

ttgtaaggga gaaccctgct tgtgggacat ggtctgagaa gagcaacctg ctggccttaa
71520

atgggtaaaa ttatacactt caggaattac tggactttca gaagcagggc tcagggctcg
71580

actggccctg gtcattggaag tgcaggtacc tgaatgtagc tgctgcccatt atctgagccc
71640

agaggtgtgg ggactctgca gactgccatt gactaagtgt gaccaactta aaactggctc
71700

tggggggtttg tttgtcatta tggccataga atagtccttt cctttcttga actcagagtg
71760

tgaggactct attcccagta gcttttgcac ttaggcaggc aaaaggaagc tgaggttggg
71820

ctcggcggat ggcagtggac agggaggcag gtgtccctca gggtcctgac tcctgacctg
71880

atgtggggga gtatgtggat ggtgtcactg tctgggggac agtgctgtgc catctcctgt
71940

acagccgact gagggtttct ggctttgcct ggtcaaggtc aagcctcagt gtggctgggt
72000

tgactggaca catagggaaa cagtttcttt gcagtgggag ttgttgaacc ctgagattga
72060

acgacagagg atcacgttag gaagctgtgc aggtgtcttg tttcctgtct gaggcaggtt
72120

gggtgaagtt ctgcctcagg agtgtgacta tggaagcccc tccatctgta ttgtggtcag
72180

tgagttgtgg ttgttttttg cagataaaga ctgtgatgtt tgacaagact ggcaccatta
72240

cccatggcgt ccccagggtc atgcgggtgc tcctgctggg ggatgtggcc aactgcccc
72300

tcaggaaggt tctggctgtg gtggggactg cggaggccag cagtgaacac cccttgggcg
72360

tggcagtcac caaatactgt aaagaggtag gtggacttgg gcgtggccct gccctccccg
72420

ccaatgctct tttattcctc accatgtcct tctctcctag ctgccctcga ggagccttct
72480

ctgtgtgggc tggaaaacca cttagagggc ccttctgcag caggcggaga tatagggctc
72540

acagtcagcc cttgccacag ttccagggtta tcacaaaggg agttcagtgt cacaggcatt
72600

tgctgaacat caccaggcgc caggctcagg agaacaacag gagtgggata tgaactctac
72660

cctttgggag ctcacgggtct tggggagtag gggcagtcct gaaactcagt gctttgacag
72720

tgcttagtgc aaggggacac aaacatgaaa tgggagtact ggggtggaag ggagaaatct
72780

actcccctgg gagcacctgg gacagagagg aaccagcaga ggtgagccag agtgcccagc
72840

aggttcctgg cacttcgtgg agaagcattg ccttcttttg gtctctccac aggaaagggt
72900

ctattttggt ttctcttaat ttcagtcctga atataggggc agcagaacgc gtcacacaga
72960

atattgccct gacaactttt aaacagcctg agcaatactg ttagcagtaa atgtagaaag
73020

gtgacatcat tgccttgaaa aacaattgac aactcttcct gatcaggggc tcttccagaa
73080

acctcatttc tagttttgtc cttcatgtag ggtcagacag cccactcttt ctcacctcac
73140

ccttctgaca tttcaggggag aatcttgctc accctcagtc atcctttatt tgttgtgtgg
73200

tctggaaaac cgcttagaag ggtcctctgc agcaagtgga gatgaggggc tcacagtcag
73260

cccttgccac agttctaggt tgtcataaag tgaggtcggt gtcacagcca tttgctgaac
73320

atcaccaggc gtcaggctca ggagaacaac agaagtggga tgtgaactct accctttggg
73380

agctcacagc ctgggggagg acgggcagtc cagaaacccc agccattcct ttaccctgcc
73440

accttgtccc ctccctcctcc agggctctgat atacaggcag ttatctcttg atagtctcaa
73500

agagctttgt taacactact tattccattg tttttaatga aattatttgc aacataaaat
73560

tcacatctcc ctctcttttg catctgcagt cctctttaat ttgagggtgt tgatatgagt
73620

tgcctttttc ctttccagtc ggtaacctgt tcactgtcat gaggcaccca cccagggaag
73680

gatttttaggc tgtttaaaat agtctcattc aacttacagc tataatatct ttaccaaaagg
73740

agggagaaat aattacaagt tactagtcac tatcttaacc tttcctatct gttccacctc
73800

cctcccctcc tttctctcag ttcccgcctt ccgctgctct cttgccacct tcaccctgtg
73860

tccctgtcct gcctgcccc tcccttccact tcaccctctc tggcttacag tttcctcttc
73920

ctctctttcc accttcccag gaacttgga cagagacctt gggatactgc acggacttcc
73980

aggcagtgcc aggctgtgga attgggtgca aagtcagcaa cgtggaaggc atcctggccc
74040

acagtgagcg ccctttgagt gcaccggcca gtcacctgaa tgaggctggc agccttcccg
74100

cagaaaaagg tattgctggc ttttgtctct gcagctgggt aaaagtagag gtgggtcaaa
74160

ccacagagag caccacgccc agcagtgatt gcctctgctg tgcggcagac ggttcatggc
74220

taaggcaccc aagcctgcct cccacacccc aggaaagttt ctcttatgtt cttgggtgcg
74280

tctaaatfff gttgcttatg gctgccttaa attcatagag cattgttttag ctgacccatg
74340

atfffatgattc tcctctffftct tatcccatat cggctcttgct tctcataagc tacagagtag
74400

acagacggga attccactff cgttctffcct gtttccattc tgtatggtag aattctggta
74460

tttaciaaagc aggcagaaaa agtcagggaa agcagagggga aatgtctgcc caagccagga
74520

agggacactg cttggaggga cagctgctff tgaagaatgg cctffftffat tctcaaggta
74580

gcttgccagc caccagaata aagggtggaa aacactcaaa atctgtgact agtgffftffc
74640

ctaagtgttg gtagtgaaat tgtggcttag attgtgtgggt tagaagccat ccaccacctt
74700

ccagcctffa aaaccccagg atggcctffc catgagagct tccagagcct ggccgttcag
74760

tgacacctca gctagcagca cacgctgtta gaaagtgata ggtgagaggt gcattagcct
74820

cgggagttac agfffttaaac atffcatgta gacgtgtctg attatffgtt gattcacctg
74880

gggttgttga taccaagcag taatgaatgt agtcaacact tgctfftggt gctgggcaaa
74940

tacctgagtg cttctaattc ctgcagaaaa agaatacagtg gatacttcta agacaaagac
75000

ggacattgcc ttatgtfftaa tacgtffata ccatttccac ccatgtagca tgtactatta
75060

atgccccatc ttaattcata cccacatgtc ctacacatta atffftactff tffftffftff
75120

gtcctaaagg atgctgtcac aagaggtgct tacaaggtta cagffftffca gaatggffta
75180

aaggatattt tgctgttaaa aggattgcat ggtttttagt tcacagtga attggaccat
75240

ttagaaataa ccacagcctc ttttgaatag atgcagtccc ccagaccttc tctgtgctga
75300

ttggaaaccg tgagtggctg aggcgcaacg gtttaaccat ttctagcgat gtcagtgacg
75360

ctatgacaga ccacgagatg aaaggacaga cagccatcct ggtggctatt gacggtatct
75420

tctgcttctg ccttccttcc gctctctcag aaatacagtt ttctgcagat atcaggcaaa
75480

agagtcctcc ttataaaaag aaaagaagac aaacaaaagc cttcctctct taatttcagg
75540

cctgttttcc agaagattgt ctttttgtct tttctcttca ttctttccac tggctccatc
75600

ttggccccag agcagctctg gtctggcctc agtctccttg accttggaat tctaaactgt
75660

ctgtgtcatc atcattgctc ttgtccaggt gggcagttca gagtcatcct tctccttcat
75720

ccctgcagac agagcaggaa ttcacttccc agcagggtta actacaggat gcatttaa
75780

aatgagatcc atgccaccct tctctgtctg ctttatttat ttattattac tttctttttc
75840

tttttttagac tgagtttccc tcttgtcacc caggctgtag tgcagtggag tgggtgtgact
75900

tgtctcacta caagctccac cttccgggtt caagcaattc tcctgcctca gactcctgag
75960

tagctgggat tacaggcccc cgccaccacg cccagctaca ttttgtattt ttagtagaga
76020

cggggtttca ccaggctggg ctcgaactcc tgacctcagg tgatctgctt gccttggcct
76080

cccaaagtgc tgggattaca ggcgtgagcc accgcacctg cccttctatc tgctttatac
76140

agactgggtgt cgcacagagt acagggtctca agactgtgtg ttggccaggt gcggtggctc
76200

acgcctgtaa tcccagcact ttgggaggct gaggtgggtg gatcacctga ggtcgggagt
76260

tcgagaccag tctgaccaac atggagaaac cccgtctcta ctaaaaataa aaaaattagc
76320

ggggtgtggt ggcgcatgcc tctaattcca gctacttggg aggctgaggc aggagaattg
76380

cttgaacccg ggagttggag gttgtggtga gccaagatca caccactgcc ctccagcctg
76440

ggcaacaaga gcgtaactct gtctcaaaaa aaaaaaaaaa aaaaaaaaaa aaaagactgc
76500

atgtgtttta gtgctcagat tctatcctgg gctttacatc ccagagagat tagagtcttt
76560

actttgttaa tgttagtttc tgctaataga agtaacagta tatgcaagca agaaataaaa
76620

tacatgacac cttttgaatg ctccagccaa atatgtcttt ctctcacatt ctctctcttt
76680

tttccctttt tctaagtctt tgtaaattcca cccgcccctc ggctatcaaa ccagaattgg
76740

cacaactaaa ggagtcgtgc cctctcgtat gtcatccctt actctgttct cattcacacc
76800

tcagccctaa aagccctcct gcttctatag ctaagattca ctcaaagaga aattctgccg
76860

aaatgttttc tgcaggaaaa gacgaaggat tgattttcta gtttggacat ttatcacttt
76920

aagtgatatc tgttatgtct cttgaatttt agtagctaata gtgtattgaa ttcttattac
76980

aagccaggca cttcgcaaag tgtcatatcg gctccttgca gaaagtttgg aaaattcagc
77040

tataaaggag aaaatgaaat cactcgtaat cctattcctt ggggagccac tgcgaagatt
77100

tcaattatat tgcttccaga cttttgtgta catccgtaaa tgcattgtact ttttaaaaag
77160

caaaaaaacc acattgtgtt atataatttg atgttttgtt ttttgcacat aatattatat
77220

gatgagaaca ttgcaagtgt ggtatcttgg tgcggggtgc cccaacttgt gtagctgctg
77280

atgctggctg gttcgctcca ggtgtgctct gtgggatgat cgcaatcgca gacgctgtca
77340

agcaggaggc tgccctggct gtgcacacgc tgcagagcat ggggtgtggac gtggttctga
77400

tcacggggga caaccggaag acagccagag ctattgccac ccaggtacag ccctttaatg
77460

ttgcaaactg tgtaaagctc ataaaagcag ttagagacaa aagtaagcac cagttggccc
77520

agcactgagc tgtactcctg ctctttctcc cttgctggat gcttttctct ccacgtgttt
77580

cattctgtgg ccatgtagtg ttgcacgtga gcactgtggg gctgtgcgaa gatctcttcc
77640

ctggagcttg ggcgcatgca tcctggaaaa aaaggagttc cagggttctc taaccgtttc
77700

ccagtctttg caatggaaaa aaaagtatgt gacctaaagg ctgagataga cattttctaa
77760

cctttatgtt aagggaagat tttgccagct cagctcttct cactgggggtg ctcagataag
77820

tgactcctgg gcagttctcc agcagcccct gatgctgagt ttggtctgtg aggacagcac
77880

cactggatgg ctgcctgact tctcacacag cagcggctcc agtttggatt cctagaaatt
77940

atgcacctcc aaatgtgtgc agagctctag gcttctgaaa gcttaattgt gttccttgag
78000

ggaagcagca gtaccctgt tttccatggt cctccttgct aattatggac ctgttttagcg
78060

tccttcccc cttttctggc tagcacttgt tcctctattc agcaggcact tacgaagcac
78120

cacgaatgtg acaggcaggt gtggggccag gctctgtgag gagtgaaggg gggctctgccc
78180

ttgacgggct catggtcttg caggaagggt gaccagtgc tgtgtagatc tgagaagcgg
78240

gaaagggact aggaaaggaa aactgcacct tcccaggaca gagcaggcag caggagcatt
78300

tcccaggggc acagagtagc acggcggttt ggagatgccg cggggagctg aatggtatgg
78360

tcaagtgtgg caccagctgc cagaccctgg gaccctggga ctctgggacc ctcggaagca
78420

gaccaagag ttgagctgcc ctctctctggg agatgcatgg ggtctactgg atgtgaagca
78480

ggagagtagg gagggacatg gccggacgga tgcttttgaa agaccatggg ggcagcacat
78540

ggagaagaca gttggaggac tgagccgtgg acagggagac cagcaagggg gttgtttgca
78600

catccagaaa aggacagggc ctaaaccagt gcagggtgtt ggggcaggag ccagggataa
78660

actggccctg tgacagcaaa cctgcagggt gtggttgacc aacatcactg actggacca
78720

gaggccgagg gcagaggggg caagggtaac ttgaggtttc tgctgctatc tgataccttt
78780

tgccaacact aggcattgcc ttccttttgt cttagggttg catcaacaaa gtctttgcag
78840

aggtgctgcc ttcgcacaag gtggccaagg tccaggagct ccagaataaa gggaagaaag
78900

tcgccatggt ggggggatggg gtcaatgact ccccggcctt ggcccaggca gacatgggtg
78960

tggccattgg caccggcacg gatgtggcca tcgaggcagc cgacgtcgtc cttatcagag
79020

tgagcgtggc tgcagccagg ctgtgggtgc tgggagggca atgggcagac cccttcctca
79080

ctgtgtgctc ctctccatca gaatgatttg ctggatgtgg tggctagcat tcacctttcc
79140

aagaggactg tccgaaggat acgcatcaac ctggtcctgg cactgattta taacctggtt
79200

gggataccca ttgcagcagg taggcagctc ttaccactg tgctccagct gcgcccagaa
79260

aggcttctgt ctcccagggt cctgctgggg ttagtgagtg gctcactcac tggctggcta
79320

gaggcgtttt agaaaggctg tttttttttt tttttcatgt cccctgcttt atatgtcttt
79380

aaatgagaaa ctgtggagag gacctgaaac cttcattcac tctggtcctt tctaaaaagc
79440

tagtggagat gcttggctgt attttctgat ccctatttta ttgtggcctt attttctgag
79500

ccatttgтта agtcatatag tctgagaaat aatcattagt aagaattatg gggcatctgc
79560

tacatgctag gcacaattat tgggtgtttca gaaccttata ctgaacatta atcctgcca
79620

agtttcactc tgagttacac aaatgatcat ggaaagaaat ggctctttga ccaggcatgg
79680

tggctcatgc ctgtaattcc agcacttttg gaggctgagg caggaggggtc atttgagctc
79740

aggggttcaa gactcacctg ggcaacatgg tgagactcca tctctacaaa aaattttaaa
79800

gttagctggg cgtggtagcg tacaccttag tcccatctgg gagtacagag gctggggcag
79860

aaggatttat tgagcccagg agttggaggt ggcagtgacc catgatcgca ccactgtact
79920

gtagcctggg caagagtgaa accctatctc agaaaaaaag aaaaagaaaa gaaatagatt
79980

tttctgcttt gtaattcccc ctcattcccc ctccttaaga accatctctc ccccgcccc
80040

tgctccctgc cccctgcccc caggaacatg gcaatgaaaa gttccagagt gctcctccat
80100

gggctgagga gagatcaagt ttgtaatcac agcattcctg cctcagtctc ccttgtgtct
80160

gagtggttag tggtgacttg tgtaacaggg ccttggcttc ctgagcccag tgagtcaggc
80220

aggtttactt cttgtcccca gagccggctc aggagcttaa ttccccaggt gtatgtgcct
80280

ctgccacaca aagttgacca tacctgtctc ctgaccctgt cctgggacct ggctctcatt
80340

cctcttctaa agatgtcatg aagccctgga ggcgagatca tctgctagag aaggactgtg
80400

gccctcagcc gtgtttgtcc tctgtgggac tgtggctaca ggatgggcaa ctggctccct
80460

cagctcctgc cgcttttctc ctcccgtcc tcccactcct tgtgggtttgt cgctgccctg
80520

gagctcgccc tgaaatggga tgccacctgt gggtggttct gggaacatca gggcgagtgg
80580

aagagagcca ggcccactca acagcatccc acaccagct ttctaggaag cctcactttg
80640

gggggggcct gtgggcaaga tccattggca ggggcttctg agcacgccag gtggagacct
80700

cactgggggc cgctgagggg ggggcagccc ggtgcctgaa gccctctcct ggctcctctc
80760

cccagacctt ggtgtgagtg cgagttcttt cttccccagg tgtcttcatg cccatcggca
80820

ttgtgctgca gccctggatg ggctcagcgg ccatggcagc ctcctctgtg tctgtggtgc
80880

tctcatccct gcagctcaag tgggtgagtcc cctcaggtgg agcatctgtg ggaggctccc
80940

aggaggcctt gtcattctgc atgcttagca cagtgggaact tgcctctcat ttcccattca
81000

ttcacctggg acaaatgctg cagtggcagc ccacatggag cctgggtgtgc atgcatgtaa
81060

cttatcagca ccatgagcag gcaattcact gctgcccatt cacgtgctgg gctgccctta
81120

ctctgctcct caagggctgt ttatgtgaac atcaagcaga catgggtccc tgtgtatgat
81180

ctatcacttt tttattttaa aaagatgttc acttatttat gagtgcattc atttaacaaa
81240

agttagtggc agcctacaat gtgcaaggca tttgcttagg aagtttctaa atcaggtgag
81300

ctttcctgga accatagggg gtttaatgtc cttttccttg gaaactcttg aggttttgat
81360

actgaatgtt attaaaagat ggatgagagg ccttcaccag gcttagaaaa aaaaaaagcc
81420

ttgtttctag aatggctcag atgctgttgc gttcctgctt tccagctata agaagcctga
81480

cctggagagg tatgaggcac aggcgcatgg ccacatgaag cccctgacgg catcccaggt
81540

cagtgtgcac ataggcatgg atgacaggtg gcgggactcc cccagggcca caccatggga
81600

ccaggtcagc tatgtcagcc aggtgtcgct gtcctccctg acgtccgaca agccatctcg
81660

gcacagcgct gcagcagacg atgatgggga caagtggctt ctgctcctga atggcagggga
81720

tgaggagcag tacatctgat gacttcaggc aggcgggccg gggcagggac ttgcctccac
81780

tcaccacaag ctgagcagga cagccagcag caggatgggc tgagctagcc tccagctttg
81840

gggacttccg ctccctggat atgtccagtc atcctgccct gcagcacgcg gccttgtctg
81900

ggtgcagctg ggcttggcct ggagaggacg gccctgcctg cctcttggcc tcacgggacc
81960

gtcagcatgg gctttgtctt ggactctagt ccttggctgg actgtagaag gtgagaggcg
82020

agtcaccctc ctcacagacc tctgcttggga gtatttagga tgactgctgt gaaatggaga
82080

acagtttcat caggaccaa aaacctcact gggcctttcc agagaactgc agacctcact
82140

gtcaggggtct ttctgatgac gcctgtctgt gtgcatcatg tttctgagac cacagtttac
82200

ctcaggtgtg cctgttgctt tcttcctgca tagtctgttc ctttcttcgt acatagtctg
82260

ttccttttct ctccctgtgtg cttgtcagtg gggaccctc gcaaccctgc ctgtcacctg
82320

ggaggggtggg accaatgtcc ttgtgggtctt tgctgctgct ctcaggcgct tctccaatgc
82380

tctggagtgt gcatttcagc ttgaacctgc ttcctggctc acacatcccc agccagggag
82440

cttgccacac tcttcttcaa gttgaggaga gttctttttt gcttaaagcc cccttctcca
82500

tggagtgttg gcttctcaat agagtgttgt tgctgaccag ctggagtgag ggcctcagag
82560

cctgacctga gagtccgtac tcggcttcct gtgggggtgta ggttctcgcg attcaggacg
82620

tccttccata tccctgcccc gcctgtggtg cttgaaacgt ttgccccatg ggaaacgtat
82680

gtgtgcagga gcctccctgc acggcccaag gggcttcggtt ttcagtcttc tgactgtcac
82740

ctcgtgggggt tcagtagaga attcatgtga ctagcgcttg gccttgtgtg gcttggagga
82800

aatggtactg cccaaatagg aggaaaacac agcctccctg agcctgcatt ctgcacgctg
82860

cccaggggct tcagaaaagg agtggccaca gcaccccgaa gggagcatct gtttacctgg
82920

cagtggctct cagagcagca gaacgggttc agtttttagac tctgaagttg gttgtgattg
82980

acagaaccct ttgggagcaa actagtagag ttggattaaa ttctgggtga aacccttttc
83040

tcccacacaa aatagtttta gtgatttttt tcattgtcca ttacttgcca ggggcagttt
83100

tagcagcact tttgatagat tacgtctaata cctcccaacc aaccagcagg gtagctatta
83160

ctgtccacat tttaacaggca aggaaacagg ctccaagagg ctgaggactt tgcccaggat
83220

gacatagcca atggacaagc agtgtctgtc agctgtgaag gcttcactct tattgtcctt
83280

ctaccttgaa tagaagtttt cctgataaga ataaacgagg aaaaggtcct tgcctcctgg
83340

aagaacaaat ctaccagggtg atctattcat tgtttcaact cagaatgcac ttgattcagg
83400

aggtcattctg accttcacct tggatgggtta gtttcacttt ttacatatag tttttgcagg
83460

gtttttatfff ataaaatcca agcgcgctgt tgattgtgtt ttccttgfff tcagcccccc
83520

cactccagcc cgcagcacat ttccgctgtc cgtcagtaat tgtgtcctct ctttatgctt
83580

gcttgggggaa tgttgtttttc tgactaggct gatcattatc taaagaatct aattctgttg
83640

atfttttaaaa cttttaggac cataaacgft gtgttcatat atggacatgg aaatatttat
83700

ataattttat agaaaataac cttttagatg gtcaaagtgt aaggagtttt tttgtcagat
83760

aatcattttct acttcaaaaa catttcatgc aatattagaa taaagttcct gtcattcctc
83820

taaaaatctg cctgtgagtg agatgcgtgg tcagaggcca gtgccgtctt gtcttcacga
83880

gggcttctgt gaggcactgg cttggcttca tgtcttttta gggggcagag gaggatagga
83940

ttttaacagt aggggggaatg aggggtgggca gagccccctt gatctgccat ccccagcact
84000

ggtgtgagtc cttttctcta atgcaggcct gcctcagcgg cgtgtgggga gaacgctgaa
84060

aggcagctcc tgtcaaagtc cctgccccgc tggatcggac acaggcttaa attaccgcat
84120

cgttttgcacc cctctgtttct aggggtgagt tctgcaaaaag gctgtttctca aacacatctc
84180

aatcaggagg gatatgtttt gggttccttg gccttgaggg agtgagtcag gaggaggagc
84240

agaaaaatct tgactagtag gctctgctgg gaaagctctg tgggcatctt ggggtggccg
84300

gggggtgtaa gtgagccaca cctgaaggga gttgacagtg taatcttcac ccacaccagg
84360

acacttcaaa gagtggaagg aggcactatg aatagttatg ccaggacaac tggcctaaac
84420

cgggactatc caggcaaact gggacatata ctcaccttcc ttcaaaggga aatagacatt
84480

gtgacatctc cattgacaca gactacgtac tttagaacct tgcctgatca taggttccag
84540

agacatgctg gtggcatagt ccgttatcat atttcaaggt tctcccccat atccacgtct
84600

tggtagctac acatctctgt agaatcattg aggtgattat ctcattggctc attgtgactg
84660

catagctaga tggccagccc cgccctccag cataccctgg aggacttcat gccatgcgtg
84720

cttgactgtt tttccggatg cctgggtaca gaaatttcat gttgtgtgca tgtcccattg
84780

tgaaaagtga gggtgatacc accttaggaa gggatcccat ttcttttagaa gatagattct
84840

ttcttttttct tttttttggt atactttaag ttctaggtac atgtgcagag catgcagggt
84900

tgttacgtag gtacacatgt gccattgtgg tttgctgcac ccgtcaacct atcatctaca
84960

ttaggtatatt ctcgtaatgc tttccctccc ctaatcccc acccctcagc aggccccagt
85020

gtgtgatgtt cccctccctg cgtccatgtg ttctcattgt tcattctcca cttacgagtg
85080

agaatatgtg gtgttttggtt ttctgttctt gtgttagttt gctgagaatg atggtttcta
85140

gcttcatcca tgtccctgca aaggacatca actcatcctt tttttttttt tttttgagac
85200

gaactcatcc ttttttatgg ctacatagca ttccatgggtg tacatgtgcc acattttctt
85260

cagtctatca ttgatgggca ttgtcgttgg ttccaagcct ttactattgt gaacagtgct
85320

gcaataaaca tagtgtccat gtgtctttat agtagaatga tttataatcc tttggatact
85380

ttctaaaact tgatcacatt ttcctgatat attcctcagg cagaataaat cacagcattc
85440

caaataactt ttagtctttt aagaccttag tgaagaattg tgatgggatt accttttta
85500

taaacaaatg cagtaattct catTTTTTTT caaaaatgac atctggaaat tgacatgtta
85560

gaaaatagat tctcaaatta ttctgctaga ttaccaagga atcaagatta ataaataaaa
85620

cctTTTTTTT taaagattaa aaataaaaac ttattTTTTT taaatccagc atagaccaaa
85680

atcagccaat caaatagttt aattagaaag agattatttg aattgcaggg gttttttttt
85740

tttgctctaa aaaattatga cagtgtttta gtaggtaaatt ttcctaattgt gtatacaatt
85800

g a g t t a g a c a a t t t t t a t t t g c a t a c
85826

<210> 2

<211> 77

<212> DNA

<213> Homo sapiens

<400> 2
gaaattggct ttcattgcttc cctggcccag agaaacccca acgctcatca cttggaccac
60

a a g a t g g a a a t a a a g c a
77

<210> 3
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 3
g t a c t t g g t t a a a a t a t g c a
20

<210> 4
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 4
a g g a a g g t a c t t g g t t a a a a
20

<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 5
a a a a g g a g g a
20

a g g t a c t t g g

<210> 6
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 6
t g g g g a a a a a
20

g g a g g a a g g t

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 7

a g g g g t g g g g
20

a a a a a g g a g g

<210> 8

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 8

a a a g a g a g g g
20

g t g g g g a a a a

<210> 9

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 9
c a t t a a a a a g
20

a g a g g g g t g g

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

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 10
c t t g t c a t t a
20

a a a a g a g a g g

<210> 11
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 11
a a t t t c c t t g
20

t c a t t a a a a a

<210> 12

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

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 12
a a a g c c a a t t
20

t c c t t g t c a t

<210> 13
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 13
a g c a t g a a a g
20

c c a a t t t c c t

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

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 14
a g g g a a g c a t
20

g a a a g c c a a t

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

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 15
t g g g c c a g g g
20

a a g c a t g a a a

<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 16

t t t c t c t g g g
20

c c a g g g a a g c

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

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 17
t g g g g t t t c t
20

c t g g g c c a g g

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

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 18
g a g c g t t g g g
20

g t t t c t c t g g

<210> 19
<211> 20
<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 19

a g t g a t g a g c
20

g t t g g g g t t t

<210> 20

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 20

g g t c c a a g t g
20

a t g a g c g t t g

<210> 21

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 21
c t t g t g g t c c
20

a a g t g a t g a g

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

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 22
t t c c c t c t t g
20

t g g t c c a a g t

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

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 23
c t t t a t t t c c
20

c t c t t g t g g t

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

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 24
t a c t g c t t t a
20

t t t c c c t c t t

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

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 25
t c t a c c t a c t
20

g c t t t a t t t c

<210> 26
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 26
t t g t g t t c t a c c t a c t g c t t
20

<210> 27
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 27
t a t c t t t t g t g t t c t a c c t a
20

<210> 28
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 28

g a g t t t a t c t
20

t t t g t g t t c t

<210> 29

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 29

g a g c t g g a g t
20

t t a t c t t t t g

<210> 30

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 30

a g a t g a g a g c
20

t g g a g t t t a t

<210> 31
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 31
g a c t t a g a t g
20

a g a g c t g g a g

<210> 32
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 32
g g a a g g g a c t
20

t a g a t g a g a g

<210> 33
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 33
g g t a g a g g a a
20

g g g a c t t a g a

<210> 34
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 34
g c c c a g g t a g
20

a g g a a g g g a c

<210> 35
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 35

a g a t g a g a g c
16

t g g a g t

<210> 36

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(16)

<223> 2'-methoxyethoxy

<400> 36

g a t g a g a g c t
16

g g a g t t

<210> 37

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(16)

<223> 2'-methoxyethoxy

<400> 37

a t g a g a g c t g
16

g a g t t t

<210> 38

<211> 16

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 38
t g a g a g c t g g
16

a g t t t a

<210> 39
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 39
g a g a g c t g g a
16

g t t t a t

<210> 40
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 40
a g a g c t g g a g
16

t t t a t c

<210> 41
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 41
g a g c t g g a g t
16

t t a t c t

<210> 42
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 42
a g c t g g a g t t
16

t a t c t t

<210> 43
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 43
g c t g g a g t t t
16

a t c t t t

<210> 44
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 44
c t g g a g t t t a
16

t c t t t t

<210> 45
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 45
t g g a g t t t a t c t t t t g
16

<210> 46
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 46
g g a g t t t a t c t t t t g t
16

<210> 47
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)

<223> 2'-methoxyethoxy

<400> 47

g a g t t t a t c t
16

t t t g t g

<210> 48

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(16)

<223> 2'-methoxyethoxy

<400> 48

a g t t t a t c t t
16

t t g t g t

<210> 49

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(16)

<223> 2'-methoxyethoxy

<400> 49

g t t t a t c t t t
16

t g t g t t

<210> 50
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 50
t t t a t c t t t t
16

g t g t t c

<210> 51
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(16)
<223> 2'-methoxyethoxy

<400> 51
t t a t c t t t t g
16

t g t t c t

<210> 52
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 52
a g a t g a g a g c
17

t g g a g t t

<210> 53
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 53
g a t g a g a g c t
17

g g a g t t t

<210> 54
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 54

a t g a g a g c t g
17

g a g t t t a

<210> 55

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(17)

<223> 2'-methoxyethoxy

<400> 55

t g a g a g c t g g
17

a g t t t a t

<210> 56

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(17)

<223> 2'-methoxyethoxy

<400> 56

g a g a g c t g g a
17

g t t t a t c

<210> 57

<211> 17

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 57
a g a g c t g g a g t t t a t c t
17

<210> 58
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 58
g a g c t g g a g t t t a t c t t
17

<210> 59
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 59

a g c t g g a g t t
17

t a t c t t t

<210> 60
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 60

g c t g g a g t t t
17

a t c t t t t

<210> 61
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 61

c t g g a g t t t a
17

t c t t t t g

<210> 62
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 62
t g g a g t t t a t
17

c t t t t g t

<210> 63
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 63
g g a g t t t a t c
17

t t t t g t g

<210> 64
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 64
g a g t t t a t c t t t t g t g t
17

<210> 65
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)
<223> 2'-methoxyethoxy

<400> 65
a g t t t a t c t t t t g t g t t
17

<210> 66
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(17)

<223> 2'-methoxyethoxy

<400> 66

g t t t a t c t t t
17

t g t g t t c

<210> 67

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(17)

<223> 2'-methoxyethoxy

<400> 67

t t t a t c t t t t
17

g t g t t c t

<210> 68

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(18)

<223> 2'-methoxyethoxy

<400> 68

a g a t g a g a g c
18

t g g a g t t t

<210> 69
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 69
g a t g a g a g c t
18

g g a g t t t a

<210> 70
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 70
a t g a g a g c t g
18

g a g t t t a t

<210> 71
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 71
t g a g a g c t g g
18

a g t t t a t c

<210> 72
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 72
g a g a g c t g g a
18

g t t t a t c t

<210> 73
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 73
a g a g c t g g a g
18

t t t a t c t t

<210> 74
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 74
g a g c t g g a g t
18

t t a t c t t t

<210> 75
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 75
a g c t g g a g t t
18

t a t c t t t t

<210> 76
<211> 18

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 76
g c t g g a g t t t
18

a t c t t t t g

<210> 77
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 77
c t g g a g t t t a
18

t c t t t t g t

<210> 78
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 78
t g g a g t t t a t
18

c t t t t g t g

<210> 79
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 79
g g a g t t t a t c
18

t t t t g t g t

<210> 80
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 80
g a g t t t a t c t
18

t t t g t g t t

<210> 81
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 81
a g t t t a t c t t
18

t t g t g t t c

<210> 82
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 82
g t t t a t c t t t
18

t g t g t t c t

<210> 83
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)
<223> 2'-methoxyethoxy

<400> 83
a g a t g a g a g c t g g a g t t t a
19

<210> 84
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)
<223> 2'-methoxyethoxy

<400> 84
g a t g a g a g c t g g a g t t t a t
19

<210> 85
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)

<223> 2'-methoxyethoxy

<400> 85

a t g a g a g c t g
19

g a g t t t a t c

<210> 86

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(19)

<223> 2'-methoxyethoxy

<400> 86

t g a g a g c t g g
19

a g t t t a t c t

<210> 87

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(19)

<223> 2'-methoxyethoxy

<400> 87

g a g a g c t g g a
19

g t t t a t c t t

<210> 88
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)
<223> 2'-methoxyethoxy

<400> 88
a g a g c t g g a g t t t a t c t t t
19

<210> 89
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)
<223> 2'-methoxyethoxy

<400> 89
g a g c t g g a g t t t a t c t t t t
19

<210> 90
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)
<223> 2'-methoxyethoxy

<400> 90
a g c t g g a g t t
19

t a t c t t t t g

<210> 91
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)
<223> 2'-methoxyethoxy

<400> 91
g c t g g a g t t t
19

a t c t t t t g t

<210> 92
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)
<223> 2'-methoxyethoxy

<400> 92
c t g g a g t t t a
19

t c t t t t g t g

<210> 93
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)
<223> 2'-methoxyethoxy

<400> 93
t g g a g t t t a t
19

c t t t t g t g t

<210> 94
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)
<223> 2'-methoxyethoxy

<400> 94
g g a g t t t a t c
19

t t t t g t g t t

<210> 95
<211> 19

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)
<223> 2'-methoxyethoxy

<400> 95
g a g t t t a t c t t t t g t g t t c
19

<210> 96
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(19)
<223> 2'-methoxyethoxy

<400> 96
a g t t t a t c t t t t g t g t t c t
19

<210> 97
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 97
g a t g a g a g c t
20

g g a g t t t a t c

<210> 98
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 98
a t g a g a g c t g
20

g a g t t t a t c t

<210> 99
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 99
t g a g a g c t g g
20

a g t t t a t c t t

<210> 100
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 100
g a g a g c t g g a
20

g t t t a t c t t t

<210> 101
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 101
a g a g c t g g a g
20

t t t a t c t t t t

<210> 102
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 102
a g c t g g a g t t t a t c t t t t g t
20

<210> 103
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 103
g c t g g a g t t t a t c t t t t g t g
20

<210> 104
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 104

c t g g a g t t t a
20

t c t t t t g t g t

<210> 105

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 105

g t t g g g c c c a
20

g g t a g a g g a a

<210> 106

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 106

g c a g a g t t g g
20

g c c c a g g t a g

<210> 107
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 107
a g c t g g c a g a
20

g t t g g g c c c a

<210> 108
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 108
a g a c c a g c t g
20

g c a g a g t t g g

<210> 109
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 109
a g a c a a g a c c
20

a g c t g g c a g a

<210> 110
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 110
t g g g a a g a c a
20

a g a c c a g c t g

<210> 111
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 111
c a c c a t g g g a
20

a g a c a a g a c c

<210> 112
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 112
g a a g g c a c c a
20

t g g g a a g a c a

<210> 113
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 113
a g g a g g a a g g
20

c a c c a t g g g a

<210> 114
<211> 20

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 114
a a t c c a g g a g
20

g a a g g c a c c a

<210> 115
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 115
t g g t t a a a a t
20

a t g c a t t g g c

<210> 116
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 116
a a a a t a t g c a
20

t t g g c a g a a a

<210> 117
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 117
a t g c a t t g g c
20

a g a a a g c a c t

<210> 118
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 118
t t g g c a g a a a
20

g c a c t t t t c a

<210> 119
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 119
a g a a a g c a c t
20

t t t c a g c t t t

<210> 120
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 120
g c a c t t t t c a
20

g c t t t g g a a a

<210> 121
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 121
t t t c a g c t t t g g a a a t t a g a
20

<210> 122
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 122
g c t t t g g a a a t t a g a a a g t g
20

<210> 123
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 123

g g a a a t t a g a
20

a a g t g a a t c t

<210> 124

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 124

t t a g a a a g t g
20

a a t c t a a a a g

<210> 125

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(18)

<223> 2'-methoxyethoxy

<400> 125

g g t a g a g g a a
18

g g g a c t t a

<210> 126
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 126
g t a g a g g a a g
18

g g a c t t a g

<210> 127
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 127
t a g a g g a a g g
18

g a c t t a g a

<210> 128
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 128
a g a g g a a g g g
18

a c t t a g a t

<210> 129
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 129
g a g g a a g g g a
18

c t t a g a t g

<210> 130
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 130
a g g a a g g g a c
18

t t a g a t g a

<210> 131
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 131
g g a a g g g a c t
18

t a g a t g a g

<210> 132
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 132
g a a g g g a c t t
18

a g a t g a g a

<210> 133
<211> 18

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 133
a a g g g a c t t a
18

g a t g a g a g

<210> 134
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 134
a g g g a c t t a g
18

a t g a g a g c

<210> 135
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 135
g g g a c t t a g a
18

t g a g a g c t

<210> 136
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 136
g g a c t t a g a t
18

g a g a g c t g

<210> 137
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 137
g a c t t a g a t g
18

a g a g c t g g

<210> 138
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 138
a c t t a g a t g a
18

g a g c t g g a

<210> 139
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 139
c t t a g a t g a g
18

a g c t g g a g

<210> 140
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 140
t t a g a t g a g a
18

g c t g g a g t

<210> 141
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(18)
<223> 2'-methoxyethoxy

<400> 141
t a g a t g a g a g
18

c t g g a g t t

<210> 142
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 142

g t a g a g g a a g
20

g g a c t t a g a t

<210> 143

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 143

t a g a g g a a g g
20

g a c t t a g a t g

<210> 144

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> misc_feature

<222> (1)..(20)

<223> 2'-methoxyethoxy

<400> 144

a g a g g a a g g g
20

a c t t a g a t g a

<210> 145
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 145
g a g g a a g g g a
20

c t t a g a t g a g

<210> 146
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 146
a g g a a g g g a c
20

t t a g a t g a g a

<210> 147
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 147
g a a g g g a c t t
20

a g a t g a g a g c

<210> 148
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 148
a a g g g a c t t a
20

g a t g a g a g c t

<210> 149
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 149
a g g g a c t t a g
20

a t g a g a g c t g

<210> 150
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 150
g g g a c t t a g a
20

t g a g a g c t g g

<210> 151
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 151
g g a c t t a g a t
20

g a g a g c t g g a

<210> 152
<211> 20

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 152
a c t t a g a t g a
20

g a g c t g g a g t

<210> 153
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 153
c t t a g a t g a g
20

a g c t g g a g t t

<210> 154
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 154
t t a g a t g a g a
20

g c t g g a g t t t

<210> 155
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(20)
<223> 2'-methoxyethoxy

<400> 155
t a g a t g a g a g
20

c t g g a g t t t a

<210> 156
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 156
c c a g c a a a g c
20

c c t t g t t a a g

<210> 157
<211> 19
<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 157

g c t c g t t g c t
19

g g g t a t c a g

<210> 158

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 158

g a t c a c a g g g
19

a t g a c c t g c

<210> 159

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 159

g t t t a c g t c g
18

c c g t c c a g

<210> 160

<211> 9

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 160

t	t	a	t	c	t	t	t	t
9								

<210> 161
 <211> 11
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 161
 g a c t t a g a t g a

11

<210> 162
 <211> 15
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 162
 t t t c a g c t t t g g a a a

15

<210> 163
 <211> 58
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 163
 ccaggtagag gaagggactt agatgagagc tggagtttat cttttgtggt ctacctac

58

<210> 164
 <211> 98

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 164
cttgtcatta aaaagagagg ggtggggaaa aaggaggaag gtacttgggtt aaaatatgca
60

ttggcagaaa gcacttttca gctttggaaa ttagaaag
98

<210> 165
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 165
gatccaccgg tcgccaccat gacctgcgcg tcctgtgtc
39

<210> 166
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 166
tcctcgccct tgctcaccat ggacagtcct ggaatgatgt tgtgg
45

<210> 167
<211> 36
<212> DNA
<213> Artificial Sequence

