(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 762 326**

(51) Int. CI.:

***C12N 15/113*** (2010.01)
***A61P 25/28*** (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA MODIFICADA TRAS OPOSICIÓN T5

(86) Fecha de presentación y número de la solicitud internacional: **15.10.2013 PCT/US2013/065067**

(87) Fecha y número de publicación internacional: **24.04.2014 WO14062686**

(96) Fecha de presentación y número de la solicitud europea: **15.10.2013 E 13846313 (8)**

(97) Fecha y número de publicación de la concesión europea modificada tras oposición: **14.12.2022 EP 2906696**

(54) Título: **Métodos para modular la expresión de C9ORF72**

(30) Prioridad:

**15.10.2012 US 201261714139 P**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente modificada:
**27.04.2023**

(73) Titular/es:

**IONIS PHARMACEUTICALS, INC. (33.3%)**
**2855 Gazelle Court**
**Carlsbad, CA 92010, US;**
**LUDWIG INSTITUTE FOR CANCER RESEARCH (33.3%) y**
**THE REGENTS OF THE UNIVERSITY OF CALIFORNIA (33.3%)**

(72) Inventor/es:

**RIGO, FRANK;**
**CLEVELAND, DON W.;**
**LAGIER-TOURENNE, CLOTILDE;**
**RAVITS, JOHN M.;**
**BENNETT, C. FRANK;**
**FREIER, SUSAN M. y**
**BAUGHN, MICHAEL W.**

(74) Agente/Representante:

**IZQUIERDO BLANCO, María Alicia**

**ES 2 762 326 T5**

## DESCRIPCIÓN

Métodos para modular la expresión de C9ORF72

**Campo**

**[0001]** En la presente se proporcionan compuestos para su uso en el tratamiento de la esclerosis lateral amiotrófica (ELA) o la demencia frontotemporal (FTD) asociada a la expansión de la repetición del hexanucleótido C9ORF72.

**Antecedentes**

**[0002]** La esclerosis lateral amiotrófica (ELA) es una enfermedad neurodegenerativa fatal caracterizada clínicamente por parálisis progresiva que conduce a la muerte por insuficiencia respiratoria, típicamente en el plazo de los dos a tres años posteriores a la aparición de los síntomas (Rowland y Shneider, N. Engl. J. Med., 2001, 344, 1688-1700). La ELA es la tercera enfermedad neurodegenerativa más común en el mundo occidental (Hirtz et al., Neurology, 2007, 68, 326-337), y actualmente no existen terapias eficaces. Aproximadamente el 10% de los casos son de naturaleza familiar, mientras que la mayoría de los pacientes diagnosticados con la enfermedad se clasifican como esporádicos, ya que parecen ocurrir al azar en toda la población (Chio et al., Neurology, 2008, 70, 533-537) Hay un reconocimiento creciente, en base a datos clínicos, genéticos y epidemiológicos, de que la ELA y la demencia frontotemporal (FTD) representan un espectro superpuesto de la enfermedad, caracterizado patológicamente por la presencia de inclusiones positivas para TDP-43 en todo el sistema nervioso central (Lillo y Hodges)., J. Clin. Neurosci., 2009, 16, 1131-1135; Neumann et al., Science, 2006, 314, 130-133).

**[0003]** Hasta la fecha, se han descubierto una serie de genes causantes de la ELA familiar clásica, por ejemplo, SOD1, TARDBP, FUS, OPTN y VCP (Johnson et al., Neuron, 2010, 68, 857-864; Kwiatkowski et al., Science, 2009, 323, 1205-1208; Maruyama et al., Nature, 2010, 465, 223-226; Rosen et al., Nature, 1993, 362, 59-62; Sreedharan et al., Science, 2008, 319, 1668-1672; Vance et al., Brain, 2009, 129, 868-876). Recientemente, el análisis de enlace de afines que involucran múltiples casos de ELA, FTD y ELA-FTD había sugerido que había un locues importante para la enfermedad en el brazo corto del cromosoma 9 (Boxer et al., J. Neurol. Neurosurg. Psychiatry, 2011, 82, 196-203; Morita et al., Neurology, 2006, 66, 839-844; Pearson et al. J. Nerol., 2011, 258, 647-655; Vance et al., Brain, 2006, 129, 868-876). El locus del cromosoma 9p21ALS-FTD en el último gen autosómico dominante principal cuya mutación es causante de la ELA. La mutación causante de ELA-FTD es una expansión de repetición de hexanucleótidos grande (GGGGCC) en el primer intrón del gen C9ORF72 (Renton et al., Neuron, 2011, 72, 257-268; DeJesus-Hemandez et al., Neuron, 2011, 72, 245-256). Un haplotipo fundador, que cubre el gen C9ORF72, está presente en la mayoría de los casos vinculados a esta región (Renton et al., Neuron, 2011, 72, 257-268) Este locus en el cromosoma 9p21 representa casi la mitad de la ELA familiar y casi una cuarta parte de todos los casos de ELA en una cohorte de 405 pacientes finlandeses (Laaksovirta et al, Lancet Neurol., 2010, 9, 978-985).

**[0004]** Un haplotipo fundador, que cubre el gen C9ORF72, está presente en la mayoría de los casos vinculados a esta región.

**[0005]** Actualmente no hay terapias eficaces para tratar tales enfermedades neurodegenerativas. Por lo tanto, es un objeto proporcionar composiciones y métodos para el tratamiento de tales enfermedades neurodegenerativas.

**[0006]** La WO 2013/030588 describe métodos para diagnosticar una enfermedad neurodegenerativa. La US 2012/214865 describe métodos para ralentizar la progresión de la enfermedad de ELA familiar. Donnelly et al., 2012, Annals of Neurology., 72(16), S67-S68 informa sobre el desarrollo de biomarcadores y agentes terapéuticos de ELA C9orf72. Madsen, 2012, revista de noticias MDA/ALS, 1-2 informa sobre antisentido contra C9ORF72. La WO 2012/092367 describe compuestos de unión a ácidos nucleicos, métodos de elaboración y uso de los mismos. Renton, 2011, Neuron, 72(2), 257-268 informa que una expansión de repetición de hexanucleótidos en C9OF72 es la causa del ELA-FTD ligada al cromosoma 9p21. Dejesus-Hernández, 2011, Neuron, 72(2), 245-256 informa que la repetición del hexanucleótidos GGGGCC expandida en la región no codificante de C9ORF72 causa FTD y ELA ligadas al cromosoma 9p. La WO 2012/012467 describe la modulación del ARN retenido nuclear. Sha et al., 2012, Alzheimers Res Ther., 4(6), 46 informa sobre las implicaciones del tratamiento de C9ORF72.

**Sumario de la invención**

**[0007]** La invención proporciona un compuesto que comprende un oligonucleótido antisentido de cadena sencilla modificado que consiste de 12 a 30 nucleósidos enlazados, en donde el oligonucleótido antisentido de cadena sencilla modificado es por lo menos un 90% complementario a un ácido nucleico de C9ORF72, para su uso en el tratamiento de la expansión de repetición del hexanucleótido C9ORF72 asociada a esclerosis lateral amiotrófica (ELA) o demencia frontotemporal (FTD) reduciendo el porcentaje de células con focos nucleares y/o reduciendo el número de focos nucleares por célula en un animal que se ha identificado que tiene focos, en donde el oligonucleótido antisentido de cadena sencilla modificado comprende por lo menos un enlace internucleosídico de fosforotioato, por lo menos una 5-metilcitosina, por lo menos un azúcar bicíclico o por lo menos un Grupo 2'-O-

metoxietilo.

**Sumario de la divulgación**

**[0008]** La invención se expone en el conjunto de reivindicaciones adjunto.

**[0009]** En la invención, la modulación tiene lugar en una célula o tejido en un animal. En ciertas realizaciones, el animal es un humano. En ciertas realizaciones, se reducen los niveles de ARNm de C9ORF72. En ciertas realizaciones, se reducen los niveles de proteína de C9ORF72. En ciertas realizaciones, se reducen preferentemente ciertas variantes de ARNm de C9ORF72. En ciertas realizaciones, se reducen preferentemente variantes de ARNm de C9ORF72 que contienen el intrón 1. En ciertas realizaciones, el intrón 1 contiene una expansión de repetición de hexanucleótidos. En ciertas realizaciones, la expansión de repetición de hexanucleótidos está asociada con una enfermedad asociada a C9ORF72. En ciertas realizaciones, la expansión de repetición de hexanucleótidos está asociada con una enfermedad asociada a la expansión de repetición de hexanucleótidos de C9ORF72. En ciertas realizaciones, la expansión de repetición de hexanucleótidos comprende por lo menos 30 repeticiones de GGGGCC. En ciertas realizaciones, la expansión de repetición de hexanucleótidos está asociada con focos nucleares. En ciertos casos, los métodos descritos en la presente son útiles para reducir el ARNm de C9ORF72, los niveles de proteína de C9ORF72, y los focos nucleares. Tal reducción puede producirse de una manera dependiente del tiempo o de una manera dependiente de la dosis.

**[0010]** También se describen métodos útiles para prevenir, tratar y mejorar enfermedades, trastornos y afecciones asociadas con C9ORF72. Tales enfermedades, trastornos y afecciones asociadas con C9ORF72 pueden ser enfermedades neurodegenerativas. La enfermedad neurodegenerativa puede ser ELA, FTD, síndrome de degeneración corticobasal (CBD), síndrome de Parkinson atípico y degeneración olivopontocerebelosa (OPCD).

**[0011]** Tales enfermedades, trastornos y afecciones pueden tener uno o más factores de riesgo, causas o resultados en común. Ciertos factores de riesgo y causas para el desarrollo de una enfermedad neurodegenerativa, y, en particular, ELA y FTD, incluyen la predisposición genética y la edad avanzada.

**[0012]** En ciertos casos, los métodos de tratamiento incluyen administrar un inhibidor específico de C9ORF72 a un individuo con necesidad de ello. En estos casos, el inhibidor específico de C9ORF72 puede ser un ácido nucleico. En estos casos, el ácido nucleico puede ser un compuesto antisentido. En estos casos, el compuesto antisentido puede ser un oligonucleótido antisentido de cadena sencilla. En la invención, el oligonucleótido antisentido de cadena sencilla puede ser complementario a un ácido nucleico de C9ORF72.

**Descripción detallada**

**[0013]** Debe entenderse que tanto la descripción general anterior como la siguiente descripción detallada son ejemplares y explicativas únicamente y no son restrictivas de la invención, como se reivindica. En la presente, el uso del singular incluye el plural a menos que se indique específicamente lo contrario. Como se usa en la presente, el uso de "o" significa "y/o" a menos que se indique lo contrario. Adicionalmente, como se usa en la presente, el uso de "y" significa "y/o" a menos que se indique lo contrario. Además, el uso del término "incluyendo", así como otras formas, como "incluye" e "incluido", no es limitativo. Además, términos como "elemento" o "componente" abarcan tanto elementos como componentes que comprenden una unidad y elementos y componentes que comprenden más de una subunidad, a menos que se indique específicamente lo contrario.

**[0014]** Los encabezados de sección usados en la presente son solo con propósitos organizativos y no deben interpretarse como limitantes del tema descrito.

*Definiciones*

**[0015]** A menos que se proporcionen definiciones específicas, la nomenclatura utilizada en relación con, y los procedimientos y técnicas de química analítica, química orgánica sintética y química medicinal y farmacéutica descritas en la presente son las bien conocidas y comúnmente usadas en la técnica. Pueden usarse técnicas estándar para la síntesis química y el análisis químico.

**[0016]** A menos que se indique lo contrario, los siguientes términos tienen los siguientes significados:

"Grupo 2'-O-metoxietilo" (también 2'-MOE y 2'-$OCH_2CH_2$-$OCH_3$ y MOE) se refiere a una modificación de O-metoxi-etilo de la posición 2' de un anillo de furanosilo. Un azúcar modificado con 2'-O-metoxietilo es un azúcar modificado.

"Nucleósido 2'-MOE" (también nucleósido 2'-O-metoxietilo) significa un nucleósido que comprende un grupo 2'-O-metoxietilo.

"5-metilcitosina" significa una citosina modificada con un grupo metilo unido a la posición 5'. Una 5-metilcitosina es una nucleobase modificada.

**[0017]** "Aproximadamente" significa dentro del ±7% de un valor. Por ejemplo, si se afirma que "los compuestos afectaron por lo menos aproximadamente un 70% de inhibición de C9ORF72", está implícito que los niveles de C9ORF72 se inhiben dentro de un intervalo del 63% y el 77%.

**[0018]** "Administrado concomitantemente" se refiere a la administración conjunta de dos agentes farmacéuticos de cualquier manera en la que los efectos farmacológicos de ambos se manifiesten en el paciente al mismo tiempo. La administración concomitante no requiere que ambos agentes farmacéuticos se administren en una única composición farmacéutica, en la misma forma de dosificación, o por la misma vía de administración. Los efectos de ambos agentes farmacéuticos no necesitan manifestarse al mismo tiempo. Los efectos solo necesitan superponerse durante un período de tiempo y no necesitan ser coextensivos.

**[0019]** "Administrar" significa proporcionar un agente farmacéutico a un animal e incluye, pero no está limitado a, la administración por un profesional médico y la autoadministración.

**[0020]** "Mejora" o "mejorar" o "mejorando" se refiere a una disminución de por lo menos un indicador, signo o síntoma de una enfermedad, trastorno o afección. La gravedad de los indicadores puede determinarse mediante medidas subjetivas u objetivas, que son conocidas por los expertos en la técnica.

**[0021]** "Animal" se refiere a un animal humano o no humano, incluyendo, pero no limitados a, ratones, ratas, conejos, perros, gatos, cerdos y primates no humanos incluyendo, pero no limitados a, monos y chimpancés.

**[0022]** "Anticuerpo" se refiere a una molécula caracterizada por reaccionar específicamente con un antígeno de alguna manera, donde el anticuerpo y el antígeno se definen cada uno en términos del otro. El anticuerpo puede referirse a una molécula de anticuerpo completa o cualquier fragmento o región de la misma, como la cadena pesada, la cadena ligera, la región Fab y la región Fc.

**[0023]** "Actividad antisentido" significa cualquier actividad detectable o medible atribuible a la hibridación de un compuesto antisentido con su ácido nucleico objetivo. En ciertas realizaciones, la actividad antisentido es una disminución en la cantidad o expresión de un ácido nucleico o proteína objetivo codificado por dicho ácido nucleico objetivo.

**[0024]** "Compuesto antisentido" significa un compuesto oligomérico que es capaz de experimentar hibridación con un ácido nucleico objetivo a través de enlaces de hidrógeno. Los ejemplos de compuestos antisentido incluyen compuestos de cadena sencilla y de cadena doble como oligonucleótidos antisentido, ARNip, ARNhc, ARNmc, y compuestos basados en la ocupación. Los mecanismos antisentido incluyen, sin limitación, antisentido mediado por RNasa H; Mecanismos de ARNi, que utilizan la vía RISC e incluyen, sin limitación, mecanismos de ARNip, ARNmc y microARN; y mecanismos basados en la ocupación que incluyen, sin limitación, oligonucleótidos modificados uniformes. Ciertos compuestos antisentido pueden actuar a través de más de uno de tales mecanismos y/o mediante mecanismos adicionales.

**[0025]** "Inhibición antisentido" significa la reducción de los niveles de ácido nucleico objetivo o los niveles de proteína objetivo en presencia de un compuesto antisentido complementario a un ácido nucleico objetivo en comparación con los niveles de ácido nucleico objetivo o los niveles de proteína objetivo en ausencia del compuesto antisentido. La inhibición puede ser por cualquier medio, incluyendo la degradación de RNasa H, como con un gapmer, y mecanismos basados en el bloqueo/ocupación estérica, como con un oligonucleótido modificado uniformemente.

**[0026]** "Oligonucleótido antisentido" significa un oligonucleótido de cadena sencilla que tiene una secuencia de nucleobases que permite la hibridación con un segmento correspondiente de un ácido nucleico objetivo.

**[0027]** "Azúcar bicíclico" significa un anillo de furanosilo modificado por el puente de dos átomos. Un azúcar bicíclico es un azúcar modificado.

**[0028]** "Nucleósido bicíclico" (también BNA) significa un nucleósido que tiene una fracción de azúcar que comprende un puente que conecta dos átomos de carbono del anillo de azúcar, formando de este modo un sistema de anillo bicíclico. En ciertas realizaciones, el puente conecta el carbono 4' y el carbono 2' del anillo de azúcar.

**[0029]** "Enfermedad asociada a C9ORF72" significa cualquier enfermedad asociada con cualquier ácido nucleico de C9ORF72 o producto de expresión del mismo. Tales enfermedades pueden incluir una enfermedad neurodegenerativa. Tales enfermedades neurodegenerativas pueden incluir ELA y FTD.

**[0030]** "Enfermedad asociada a la expansión de repetición de hexanucleótidos de C9ORF72" significa cualquier enfermedad asociada con un ácido nucleico de C9ORF72 que contiene una expansión de repetición de hexanucleótidos. En ciertas realizaciones, la expansión de repetición de hexanucleótidos puede comprender GGGGCC, GGGGGG, GGGGGC o GGGGCC repetido por lo menos 30 veces. Tales enfermedades pueden incluir

una enfermedad neurodegenerativa. Tales enfermedades neurodegenerativas pueden incluir ELA y FTD.

**[0031]** "Ácido nucleico de C9ORF72" significa cualquier ácido nucleico que codifica C9ORF72. Por ejemplo, en ciertas realizaciones, un ácido nucleico de C9ORF72 incluye una secuencia de ADN que codifica C9ORF72, una secuencia de ARN transcrita a partir de ADN que codifica C9ORF72 (incluyendo ADN genómico que comprende intrones y exones), y una secuencia de ARNm que codifica C9ORF72. "ARNm de C9ORF72" significa un ARNm que codifica una proteína de C9ORF72.

**[0032]** "Inhibidor específico de C9ORF72" se refiere a cualquier agente capaz de inhibir específicamente la expresión de ARNm de C9ORF72 y/o proteína de C9ORF72 a nivel molecular. Por ejemplo, los inhibidores específicos de C9ORF72 incluyen ácidos nucleicos (incluyendo compuestos antisentido), ARNip, aptámeros, anticuerpos, péptidos, moléculas pequeñas y otros agentes capaces de inhibir la expresión del ARNm de C9ORF72 y/o proteína de C9ORF72. De manera similar, los inhibidores específicos de C9ORF72 pueden afectar a otros procesos moleculares en un animal.

**[0033]** "Estructura de tapa" o "fracción de tapa terminal" significa modificaciones químicas, que se han incorporado en cualquier extremo terminal de un compuesto antisentido.

**[0034]** "cEt" o "etilo restringido" significa un nucleósido bicíclico que tiene una fracción de azúcar que comprende un puente que conecta el carbono 4' y el carbono 2', en donde el puente tiene la fórmula: 4'-CH ($CH_3$)-O-2'.

**[0035]** "Nucleósido de etilo restringido " (también nucleósido de cET) significa un nucleósido que comprende una fracción de azúcar bicíclico que comprende un puente 4'-CH'($CH_3$)-O-2'.

**[0036]** "Región químicamente distinta" se refiere a una región de un compuesto antisentido que es de alguna manera químicamente diferente que otra región del mismo compuesto antisentido. Por ejemplo, una región que tiene nucleósidos de 2'-O-metoxietilo es químicamente distinta de una región que tiene nucleósidos sin modificaciones de 2'-O-metoxietilo.

**[0037]** "Compuesto antisentido quimérico" significa un compuesto antisentido que tiene por lo menos dos regiones químicamente distintas.

**[0038]** "Administración conjunta" significa la administración de dos o más agentes farmacéuticos a un individuo. Los dos o más agentes farmacéuticos pueden estar en una única composición farmacéutica, o pueden estar en composiciones farmacéuticas separadas. Cada uno de los dos o más agentes farmacéuticos puede administrarse a través de la misma vía o a través de diferentes vías de administración. La administración conjunta abarca la administración paralela o secuencial.

**[0039]** "Complementariedad" significa la capacidad de emparejamiento entre nucleobases de un primer ácido nucleico y un segundo ácido nucleico.

**[0040]** "Nucleobases contiguas" significa nucleobases inmediatamente adyacentes entre sí.

**[0041]** "Diluyente" significa un ingrediente en una composición que carece de actividad farmacológica, pero es farmacéuticamente necesario o deseable. Por ejemplo, el diluyente en una composición inyectada puede ser un líquido, por ejemplo, solución salina.

**[0042]** "Dosis" significa una cantidad especificada de un agente farmacéutico proporcionado en una única administración, o en un período de tiempo especificado. En ciertos casos, una dosis puede administrarse en uno, dos o más bolos, comprimidos o inyecciones. Por ejemplo, en ciertos casos en los que se desea la administración subcutánea, la dosis deseada requiere un volumen que no se acomoda fácilmente con una sola inyección, por lo tanto, pueden usarse dos o más inyecciones para lograr la dosis deseada. En ciertos casos, el agente farmacéutico se administra por infusión durante un período prolongado de tiempo o de forma continua. Las dosis pueden indicarse como la cantidad de agente farmacéutico por hora, día, semana o mes.

**[0043]** "Cantidad eficaz" significa la cantidad de agente farmacéutico suficiente para efectuar un resultado fisiológico deseado en un individuo que necesita el agente farmacéutico. La cantidad eficaz puede variar entre individuos dependiendo de la salud y la condición física del individuo a tratar, el grupo taxonómico de los individuos a tratar, la formulación de la composición, la evaluación de la condición médica del individuo y otros factores relevantes.

**[0044]** "Expresión" significa la conversión de la información de un gen de C9ORF72 en ARNm mediante transcripción y luego a proteína mediante traducción. La expresión puede dar como resultado una manifestación fenotípica del gen de C9ORF72.

**[0045]** "Completamente complementario" o "100% complementario" significa que cada nucleobase de un primer

ácido nucleico tiene una nucleobase complementaria en un segundo ácido nucleico. En ciertas realizaciones, un primer ácido nucleico es un compuesto antisentido y un ácido nucleico objetivo es un segundo ácido nucleico.

**[0046]** "Gapmer" significa un compuesto antisentido quimérico en el que una región interna que tiene una pluralidad de nucleósidos que soportan la escisión de RNasa H está colocada entre regiones externas que tienen uno o más nucleósidos, en donde los nucleósidos que comprenden la región interna son químicamente distintos del nucleósido o nucleósidos que comprenden las regiones externas. La región interna puede denominarse "brecha" y las regiones externas pueden denominarse "alas".

**[0047]** "Hueco reducido" significa un compuesto antisentido quimérico que tiene un segmento de hueco de 9 o menos 2'-desoxirribonucleósidos contiguos colocados entre e inmediatamente adyacentes a los segmentos de ala 5' y 3' que tienen de 1 a 6 nucleósidos.

**[0048]** "Hueco ampliado" significa un compuesto antisentido quimérico que tiene un segmento de hueco de 12 o más 2'-desoxirribonucleósidos contiguos colocados entre e inmediatamente adyacentes a segmentos de ala 5' y 3' que tienen de 1 a 6 nucleósidos.

**[0049]** "Expansión de repetición de hexanucleótidos" significa una serie de seis bases (por ejemplo, GGGGCC, GGGGGG, GGGGCG o GGGGGC) repetidas por lo menos dos veces. En ciertas realizaciones, la expansión de repetición de hexanucleótidos puede localizarse en el intrón 1 de un ácido nucleico de C9ORF72. En ciertas realizaciones, una expansión de repetición de hexanucleótidos patogénica incluye por lo menos 30 repeticiones de GGGGCC, GGGGGG, GGGGCG o GGGGGC en un ácido nucleico de C9ORF72 y está asociada con la enfermedad. En ciertas realizaciones, las repeticiones son consecutivas. En ciertas realizaciones, las repeticiones están interrumpidas por 1 o más nucleobases. En ciertas realizaciones, una expansión de repetición de hexanucleótidos de tipo salvaje incluye 23 repeticiones o menos de GGGGCC, GGGGGG, GGGGCG o GGGGGC en un ácido nucleico de C9ORF72. En ciertas realizaciones, las repeticiones son consecutivas. En ciertas realizaciones, las repeticiones están interrumpidas por 1 o más nucleobases.

**[0050]** "Hibridación" significa el apareamiento de moléculas de ácido nucleico complementarias. En la invención, las moléculas de ácido nucleico complementarias incluyen un compuesto antisentido y un ácido nucleico objetivo.

**[0051]** "Identificar un animal que tiene una enfermedad asociada a C9ORF72" significa identificar un animal al que se le ha diagnosticado una enfermedad asociada a C9ORF72 o que está predispuesto a desarrollar una enfermedad asociada a C9ORF72. Los individuos predispuestos a desarrollar una enfermedad asociada a C9ORF72 incluyen aquellos que tienen uno o más factores de riesgo para desarrollar una enfermedad asociada a C9ORF72, incluyendo que tienen un historial personal o familiar o predisposición genética de una o más enfermedades asociadas a C9ORF72. Dicha identificación puede lograrse mediante cualquier método, incluyendo la evaluación del historial médico de un individuo y pruebas o evaluaciones clínicas estándar, como las pruebas genéticas.

**[0052]** "Inmediatamente adyacente" significa que no hay elementos intermedios entre los elementos inmediatamente adyacentes.

**[0053]** "Individuo" significa un animal humano o no humano seleccionado para tratamiento o terapia.

**[0054]** "Inhibir C9ORF72" significa reducir la expresión del ARNm y/o los niveles de proteína de C9ORF72en presencia de un inhibidor específico de C9ORF72, que incluye un oligonucleótido antisentido de C9ORF72, en comparación con la expresión del ARNm y/o niveles de proteína de C9ORF72 en ausencia de un inhibidor específico de C9ORF72, como un oligonucleótido antisentido de C9ORF72.

**[0055]** "Enlace internucleosídico" se refiere al enlace químico entre nucleósidos.

**[0056]** “Nucleósidos enlazados” significa nucleósidos adyacentes que están unidos entre sí.

**[0057]** "Malapareamiento" o "nucleobase no complementaria" se refiere al caso en el que una nucleobase de un primer ácido nucleico no es capaz de emparejarse con la nucleobase correspondiente de un segundo ácido nucleico u objetivo.

**[0058]** "Enlace internucleosídico modificado" se refiere a una sustitución o cualquier cambio de un enlace internucleósido de origen natural (es decir, un enlace internucleósido de fosfodiéster).

**[0059]** “Nucleobase modificada” se refiere a cualquier nucleobase distinta de adenina, citosina, guanina, timidina o uracilo. Una "nucleobase no modificada" significa las bases de purina adenina (A) y guanina (G), y las bases de pirimidina timina (T), citosina (C) y uracilo (U).

**[0060]** “Nucleótido modificado” significa un nucleótido que tiene, independientemente, una fracción de azúcar

modificado, enlace internucleósido modificado o nucleobase modificada. Un "nucleósido modificado" significa un nucleósido que tiene, independientemente, una fracción de azúcar modificado o una nucleobase modificada.

**[0061]** "Oligonucleótido modificado" significa un oligonucleótido que comprende un enlace internucleosídico modificado, un azúcar modificado o una nucleobase modificada.

**[0062]** "Azúcar modificado" se refiere a una sustitución o cambio de un azúcar natural.

**[0063]** "Motivo" significa el patrón de regiones químicamente distintas en un compuesto antisentido.

**[0064]** "Enlace internucleosídico natural" significa un enlace fosfodiéster de 3' a 5'.

**[0065]** "Fracción de azúcar natural" significa un azúcar que se encuentra en el ADN (2'-H) o ARN (2'-OH).

**[0066]** "Ácido nucleico" se refiere a moléculas compuestas de nucleótidos monoméricos. Un ácido nucleico incluye ácidos ribonucleicos (ARN), ácidos desoxirribonucleicos (ADN), ácidos nucleicos de cadena sencilla, ácidos nucleicos de cadena doble, ácidos ribonucleicos interferentes pequeños (ARNip) y microARN (ARNmi).

**[0067]** “Nucleobase” significa una fracción heterocíclica capaz de emparejarse con una base de otro ácido nucleico.

**[0068]** "Secuencia de nucleobase" significa el orden de nucleobases contiguas independientes de cualquier azúcar, enlace o modificación de nucleobase.

**[0069]** "Nucleósido" significa una nucleobase unida a un azúcar.

**[0070]** "Mimético de nucleósido" incluye aquellas estructuras usadas para reemplazar el azúcar o el azúcar y la base y no necesariamente el enlace en una o más posiciones de un compuesto oligomérico como, por ejemplo, miméticos de nucleósidos que tienen morfolino, ciclohexenilo, ciclohexilo, tetrahidropiranilo, biciclo o miméticos de azúcar triciclo, por ejemplo, unidades de azúcar no de furanosa. El mimético de nucleótidos incluye aquellas estructuras usadas para reemplazar el nucleósido y el enlace en una o más posiciones de un compuesto oligomérico como, por ejemplo, ácidos nucleicos peptídicos o morfolinos (morfolinos enlazados por -N(H)-C(=O)-O- u otro enlace no fosfodiéster). El sustituto del azúcar se superpone con el término nucleósido mimético ligeramente más amplio, pero se pretende que indique el reemplazo de la unidad de azúcar (anillo de furanosa) solamente. Los anillos de tetrahidropiranilo proporcionados en la presente son ilustrativos de un ejemplo de un sustituto de azúcar en el que el grupo azúcar de furanosa se ha reemplazado con un sistema de anillo de tetrahidropiranilo.

**[0071]** “Nucleótido” significa un nucleósido que tiene un grupo fosfato enlazado covalentemente a la porción de azúcar del nucleósido.

**[0072]** "Compuesto oligomérico" u "oligómero" significa un polímero de subunidades monoméricas enlazadas que es capaz de hibridar con por lo menos una región de una molécula de ácido nucleico.

**[0073]** "Oligonucleótido" significa un polímero de nucleósidos enlazados, cada uno de los cuales puede estar modificado o no modificado, uno independientemente del otro.

**[0074]** "Administración parenteral" significa administración por inyección o infusión. La administración parenteral incluye administración subcutánea, administración intravenosa, administración intramuscular, administración intraarterial, administración intraperitoneal o administración intracraneal, por ejemplo, administración intratecal o intracerebroventricular.

**[0075]** "Péptido" significa una molécula formada enlazando por lo menos dos aminoácidos por enlaces amida. Péptido se refiere a polipéptidos y proteínas.

**[0076]** "Agente farmacéutico" significa la sustancia o sustancias en una composición farmacéutica que proporcionan un beneficio terapéutico cuando se administran a un individuo. Por ejemplo, en ciertos casos, un oligonucleótido antisentido dirigido a C9ORF72 es un agente farmacéutico.

**[0077]** "Composición farmacéutica" significa una mezcla de sustancias adecuadas para administrar a un individuo. Por ejemplo, una composición farmacéutica puede comprender uno o más agentes farmacéuticos y una solución acuosa estéril.

**[0078]** "Derivado farmacéuticamente aceptable" abarca sales, conjugados, profármacos o isómeros farmacéuticamente aceptables de los compuestos descritos en la presente.

**[0079]** "Sales farmacéuticamente aceptables" significa sales fisiológica y farmacéuticamente aceptables de

compuestos antisentido, es decir, sales que retienen la actividad biológica deseada del oligonucleótido original y no imparten efectos toxicológicos no deseados a los mismos.

**[0080]** "Enlace de fosforotioato" significa un enlace entre nucleósidos en donde el enlace de fosfodiéster se modifica reemplazando uno de los átomos de oxígeno no puente con un átomo de azufre. Un enlace de fosforotioato (P=S) es un enlace internucleosídico modificado.

**[0081]** Porción significa un número definido de nucleobases contiguas (es decir, enlazadas) de un ácido nucleico. En ciertas realizaciones, una porción es un número definido de nucleobases contiguas de un ácido nucleico objetivo. En ciertas realizaciones, una porción es un número definido de nucleobases contiguas de un compuesto antisentido.

**[0082]** "Prevenir" o "prevención" se refiere a retrasar o impedir el inicio o desarrollo de una enfermedad, trastorno o afección durante un período de tiempo de minutos a indefinidamente. Prevenir también significa reducir el riesgo de desarrollar una enfermedad, trastorno o afección.

**[0083]** "Profármaco" significa un agente terapéutico que se prepara en una forma inactiva que se convierte en una forma activa dentro del cuerpo o las células del mismo mediante la acción de enzimas endógenas u otras sustancias químicas o afecciones.

**[0084]** "Efectos secundarios" significa respuestas fisiológicas atribuibles a un tratamiento distintas de los efectos deseados. En ciertas realizaciones, los efectos secundarios incluyen reacciones en el sitio de inyección, anormalidades de la prueba de función hepática, anormalidades de la función renal, toxicidad hepática, toxicidad renal, anormalidades del sistema nervioso central, miopatías y malestar general.

**[0085]** “Oligonucleótido de cadena sencilla” significa un oligonucleótido que no hibrida con una cadena complementaria.

**[0086]** "Específicamente hibridable" se refiere a un compuesto antisentido que tiene un grado suficiente de complementariedad entre un oligonucleótido antisentido y un ácido nucleico objetivo para inducir un efecto deseado, a la vez que muestra efectos mínimos o ningún efecto sobre los ácidos nucleicos no objetivo bajo condiciones en las que se desea una unión específica, es decir, bajo condiciones fisiológicas en el caso de ensayos *in vivo* y tratamientos terapéuticos.

**[0087]** "Dirigir" o "dirigido" significa el proceso de diseño y selección de un compuesto antisentido que se hibridará específicamente con un ácido nucleico objetivo e inducirá un efecto deseado.

**[0088]** "Ácido nucleico objetivo", "ARN objetivo" y "transcrito de ARN objetivo" se refieren todos a un ácido nucleico capaz de ser el objetivo de compuestos antisentido.

**[0089]** "Segmento objetivo" significa la secuencia de nucleótidos de un ácido nucleico objetivo al que se dirige un compuesto antisentido. "Sitio objetivo 5' " se refiere al nucleótido más 5' de un segmento objetivo. "Sitio objetivo 3’ " se refiere al nucleótido más 3' de un segmento objetivo.

**[0090]** "Cantidad terapéuticamente eficaz" significa una cantidad de un agente farmacéutico que proporciona un beneficio terapéutico a un individuo.

**[0091]** "Tratar" o "tratamiento" se refiere a administrar una composición farmacéutica para efectuar una alteración o mejora de una enfermedad, trastorno o afección.

**[0092]** "Nucleótido no modificado" significa un nucleótido compuesto de nucleobases, fracciones de azúcar y enlaces internucleosídicos de origen natural. En ciertas realizaciones, un nucleótido no modificado es un nucleótido de ARN (es decir, β-D-ribonucleósidos) o un nucleótido de ADN (es decir, β-D-desoxirribonucleósido).

*Ciertas realizaciones*

**[0093]** En la presente se divulgan métodos para disminuir la expresión de ARNm y proteína de C9ORF72.

**[0094]** La ELA o FTD asociada con la expansión de la repetición de hexanucleótidos C9ORF72 puede ser familiar o esporádica.

**[0095]** La expansión de repetición de hexanucleótidos puede comprender GGGGCC, GGGGGG, GGGGGC o GGGGCG. En ciertos casos, los inhibidores específicos de C9ORF72 son ácidos nucleicos (incluyendo compuestos antisentido), péptidos, anticuerpos, moléculas pequeñas y otros agentes capaces de inhibir la expresión del ARNm de C9ORF72 y/o proteína de C9ORF72.

**[0096]** En la presente se divulgan métodos que comprenden administrar un inhibidor específico de C9ORF72 u homólogo de C9ORF72 a un animal.

**[0097]** En la presente se divulgan métodos que comprenden:

identificar un animal que tiene una enfermedad asociada a C9ORF72; y
administrar un inhibidor específico de C9ORF72.

**[0098]** En ciertas realizaciones, el animal es un humano.

**[0099]** En ciertos casos, el inhibidor específico de C9ORF72 es cualquiera de un compuesto antisentido, un aptámero, un anticuerpo, un péptido o una molécula pequeña.

**[0100]** En ciertas realizaciones, el compuesto antisentido comprende un oligonucleótido antisentido de cadena sencilla complementario a un ácido nucleico de C9ORF72.

**[0101]** En ciertas realizaciones, el ácido nucleico de C9ORF72 es un ácido nucleico de C9ORF72 humano.

**[0102]** En ciertas realizaciones, el animal tiene una enfermedad asociada a C9ORF72.

**[0103]** En ciertas realizaciones, la enfermedad asociada a C9ORF72 es una enfermedad asociada a la expansión de repetición de hexanucleótidos de C9ORF72.

**[0104]** En ciertas realizaciones, la enfermedad asociada a C9ORF72 es esclerosis lateral amiotrófica (ELA) o demencia frontotemporal (FTD).

**[0105]** En ciertas realizaciones, la enfermedad asociada a la expansión de repetición de hexanucleótidos de C9ORF72 es esclerosis lateral amiotrófica (ELA) o demencia frontotemporal (FTD).

**[0106]** En ciertas realizaciones, la esclerosis lateral amiotrófica (ELA) es ELA familiar o ELA esporádica.

**[0107]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla comprende por lo menos una modificación.

**[0108]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla es específicamente hibridable con un ácido nucleico de C9ORF72 humano.

**[0109]** En ciertos casos, el oligonucleótido antisentido de cadena sencilla es por lo menos un 75%, por lo menos un 80%, por lo menos un 85%, por lo menos un 90% o por lo menos un 95% complementario a una porción de igual longitud de un ácido nucleico de C9ORF72 humano.

**[0110]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla es 100% complementario a una porción de igual longitud de un ácido nucleico de C9ORF72 humano.

**[0111]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla comprende por lo menos un enlace internucleósido modificado.

**[0112]** En ciertas realizaciones, cada enlace internucleosídico del oligonucleótido antisentido de cadena sencilla es un enlace internucleósido modificado.

**[0113]** En ciertas realizaciones, el enlace internucleosídico modificado es un enlace internucleósido de fosforotioato.

**[0114]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla comprende por lo menos un nucleósido modificado.

**[0115]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla comprende por lo menos un nucleósido modificado que tiene un azúcar modificado.

**[0116]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla comprende por lo menos un nucleósido modificado que comprende un azúcar bicíclico.

**[0117]** En ciertas realizaciones, el azúcar bicíclico comprende un puente de 4' a 2' seleccionado entre: puente 4'-$(CH_2)_n$-O-2', en donde n es 1 o 2; y 4'$CH_2$-O-$CH_2$-2'.

**[0118]** En ciertas realizaciones, el azúcar bicíclico comprende un puente 4'-CH'($CH_3$)-O-2'.

**[0119]** En ciertas realizaciones, por lo menos un nucleósido modificado que tiene un azúcar modificado comprende una fracción de azúcar modificado 2'-modificado no bicíclico.

**[0120]** En ciertas realizaciones, la fracción de azúcar 2'-modificado comprende un grupo 2'-O-metoxietilo.

**[0121]** En ciertas realizaciones, la fracción de azúcar 2'-modificado comprende un grupo 2'-O-metilo.

**[0122]** En ciertas realizaciones, el por lo menos un nucleósido modificado que tiene un azúcar modificado comprende un sustituto de azúcar.

**[0123]** En ciertas realizaciones, el sustituto del azúcar es un morfolino.

**[0124]** En ciertas realizaciones, el sustituto del azúcar es un ácido nucleico peptídico.

**[0125]** En ciertas realizaciones, cada nucleósido está modificado.

**[0126]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla comprende por lo menos una nucleobase modificada.

**[0127]** En ciertas realizaciones, la nucleobase modificada es una 5'-metilcitosina.

**[0128]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla comprende:

un segmento de hueco que consiste de desoxinucleósidos enlazados;
un segmento de ala 5' que consiste de nucleósidos enlazados;
un segmento de ala 3' que consiste de nucleósidos enlazados;
en donde el segmento de hueco está colocado inmediatamente adyacente a y entre el segmento de ala 5' y el segmento de ala 3' y en donde cada nucleósido de cada segmento de ala comprende un azúcar modificado.

**[0129]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla comprende:

un segmento de hueco que consiste de diez desoxinucleósidos enlazados;
un segmento de ala 5' que consiste de cinco nucleósidos enlazados;
un segmento de ala 3' que consiste de cinco nucleósidos enlazados;
en donde el segmento de hueco está colocado inmediatamente adyacente y entre el segmento de ala 5' y el segmento de ala 3', en donde cada nucleósido de cada segmento de ala comprende un azúcar 2'-O-metoxietilo; y en donde cada enlace internucleosídico es un enlace fosforotioato.

**[0130]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla consiste de 15 nucleósidos unidos.

**[0131]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla consiste de 16 nucleósidos enlazados.

**[0132]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla consiste de 17 nucleósidos enlazados.

**[0133]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla consiste de 18 nucleósidos enlazados.

**[0134]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla consiste de 19 nucleósidos enlazados.

**[0135]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla consiste de 20 nucleósidos enlazados.

**[0136]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla consiste de 21 nucleósidos enlazados.

**[0137]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla consiste de 22 nucleósidos enlazados.

**[0138]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla consiste de 23 nucleósidos enlazados.

**[0139]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla consiste de 24 nucleósidos enlazados.

**[0140]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla consiste de 25 nucleósidos enlazados.

**[0141]** En ciertas realizaciones, el oligonucleótido antisentido de cadena sencilla complementario a un ácido nucleico de C9ORF72 es complementario a cualquiera de un intrón, un exón, una unión de corte y empalme, una unión de exón:exón, una repetición, la UTR 3' o la UTR 5' del ácido nucleico deC9ORF72.

**[0142]** En ciertas realizaciones, el intrón es cualquiera de los intrones 1, intrón 2, intrón 3, intrón 4, intrón 5, intrón 6, intrón 7, intrón 8, intrón 9 o intrón 10.

**[0143]** En ciertas realizaciones, el exón es cualquiera de exón 1a, exón 1b, exón 1c, exón 1d, exón 1d, exón 2, exón 3, exón 4, exón 5, exón 6, exón 7, exón 8, exón 9, exón 10 o exón 11.

**[0144]** En ciertas realizaciones, la repetición es una repetición de hexanucleótido.

**[0145]** En ciertas realizaciones, la repetición del hexanucleótido es cualquiera de GGGGCC, GGGGGG, GGGGCG o GGGGGC.

**[0146]** En ciertas realizaciones, se reduce la expresión de ARN de C9ORF72.

**[0147]** En ciertas realizaciones, la expresión de ARN de C9ORF72 se reduce en un núcleo.

**[0148]** En ciertas realizaciones, se reduce la expresión de proteína de C9ORF72.

**[0149]** En ciertas realizaciones, la expresión de proteína de C9ORF72 se reduce en un núcleo.

**[0150]** En ciertas realizaciones, la expresión se reduce en cualquiera de fibroblastos, linfoblastos inmortalizados, neuronas motoras espinales, células de purkinje y granulares en el cerebelo, neuronas corticales, astrocitos o microglia.

**[0151]** En ciertas realizaciones, se reduce la expresión de ARN de la isoforma V1 de C9ORF72.

**[0152]** En ciertas realizaciones, se reduce la expresión de proteína de la isoforma V1 de C9ORF72.

**[0153]** En ciertas realizaciones, se reduce la expresión de ARN de la isoforma V2 de C9ORF72.

**[0154]** En ciertas realizaciones, se reduce la expresión de proteína de la isoforma V2 de C9ORF72.

**[0155]** En ciertas realizaciones, se reduce la expresión de ARN de la isoforma V3 de C9ORF72.

**[0156]** En ciertas realizaciones, se reduce la expresión de proteína de la isoforma V3 de C9ORF72.

**[0157]** En ciertas realizaciones, se reduce la expresión del ARN de la isoforma V2 de C9ORF72 y la expresión del ARN de la isoforma V3 de C9ORF72 y no se reduce el ARN de la isoforma V1 de C9ORF72.

**[0158]** En ciertas realizaciones, se reduce la expresión de proteína de la isoforma V2 de C9ORF72 y la expresión de proteína de la isoforma V3 de C9ORF72 y no se reduce la proteína de la isoforma V1 de C9ORF72.

**[0159]** En ciertos casos, se reducen los focos de ARN de C9ORF72.

**[0160]** En la invención, los focos de ARN de C9ORF72 se reducen en un núcleo.

**[0161]** En ciertos casos, las composiciones descritas en la presente se administran por administración parenteral.

**[0162]** En ciertos casos, la administración parenteral es cualquiera de inyección o infusión.

**[0163]** En ciertos casos, la administración parenteral es cualquiera de administración intratecal o administración intracerebroventricular.

**[0164]** En ciertas realizaciones, se mejora o previene por lo menos un síntoma de una enfermedad asociada a C9ORF72.

**[0165]** En ciertas realizaciones, se mejora o previene por lo menos un síntoma de ELA o FTD asociada a la expansión de repetición de hexanucleótidos de C9ORF72.

**[0166]** En ciertas realizaciones, se ralentiza la progresión de por lo menos un síntoma de una enfermedad asociada a C9ORF72.

**[0167]** En ciertas realizaciones, se ralentiza la progresión de por lo menos un síntoma de ELA o FDT asociada a la expansión de repetición de hexanucleótidos de C9ORF72.

**[0168]** En ciertas realizaciones, el por lo menos un síntoma es cualquiera de la función motora, respiración, debilidad muscular, fasciculación y calambres musculares, dificultad para proyectar la voz, falta de aire, dificultad para respirar y tragar, comportamiento social inapropiado, falta de empatía, distracción, cambios en las preferencias alimentarias, agitación, emociones embotadas, descuido de la higiene personal, comportamiento repetitivo o compulsivo, y energía y motivación disminuidas.

*Compuestos antisentido*

**[0169]** Los compuestos oligoméricos incluyen, pero no están limitados a, oligonucleótidos, oligonucleósidos, análogos de oligonucleótidos, miméticos de oligonucleótidos, compuestos antisentido, oligonucleótidos antisentido y ARNip. Un compuesto oligomérico puede ser "antisentido" para un ácido nucleico objetivo, lo que significa que es capaz de experimentar hibridación con un ácido nucleico objetivo a través de enlaces de hidrógeno.

**[0170]** En ciertas realizaciones, un compuesto antisentido tiene una secuencia de nucleobases que, cuando se escribe en la dirección 5' a 3', comprende el complemento inverso del segmento objetivo de un ácido nucleico objetivo al que está dirigido. En ciertas de tales realizaciones, un oligonucleótido antisentido tiene una secuencia de nucleobases que, cuando se escribe en la dirección 5' a 3', comprende el complemento inverso del segmento objetivo de un ácido nucleico objetivo al que está dirigido.

**[0171]** En la invención, un compuesto antisentido dirigido a un ácido nucleico de C9ORF72 tiene una longitud de 12 a 30 subunidades. En ciertos casos, el compuesto antisentido es de 8 a 80, de 12 a 50, de 15 a 30, de 18 a 24, de 19 a 22, o de 20 subunidades enlazadas. En ciertos casos, los compuestos antisentido tienen 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79 u 80 subunidades enlazadas de longitud, o un intervalo definido por dos cualquiera de los valores anteriores. En la invención, el compuesto antisentido es un oligonucleótido antisentido modificado, y las subunidades enlazadas son nucleósidos.

**[0172]** En ciertas realizaciones, los oligonucleótidos antisentido dirigidos a un ácido nucleico de C9ORF72 pueden acortarse o truncarse. Por ejemplo, puede eliminarse una sola subunidad del extremo 5' (truncamiento 5'), o alternativamente del extremo 3' (truncamiento 3'). Un compuesto antisentido acortado o truncado dirigido a un ácido nucleico de C9ORF72 puede tener dos subunidades eliminadas del extremo 5', o alternativamente puede tener dos subunidades eliminadas del extremo 3' del compuesto antisentido. Alternativamente, los nucleósidos eliminados pueden dispersarse por todo el compuesto antisentido, por ejemplo, en un compuesto antisentido que tiene un nucleósido eliminado del extremo 5' y un nucleósido eliminado del extremo 3'.

**[0173]** Cuando hay una única subunidad adicional en un compuesto antisentido alargado, la subunidad adicional puede localizarse en el extremo 5' o 3' del compuesto antisentido. Cuando hay dos o más subunidades adicionales, las subunidades añadidas pueden estar adyacentes entre sí, por ejemplo, en un compuesto antisentido que tiene dos subunidades añadidas al extremo 5' (adición 5'), o alternativamente al extremo 3' (adición 3'), del compuesto antisentido. Alternativamente, las subunidades añadidas pueden estar dispersas por todo el compuesto antisentido, por ejemplo, en un compuesto antisentido que tiene una subunidad añadida al extremo 5' y una subunidad añadida al extremo 3'.

**[0174]** Es posible aumentar o disminuir la longitud de un compuesto antisentido, como un oligonucleótido antisentido, y/o introducir bases malapareadas sin eliminar la actividad. Por ejemplo, en Woolf et al. (Proc. Natl. Acad. Sci. USA 89:7305-7309, 1992), se probó una serie de oligonucleótidos antisentido de 13-25 nucleobases de longitud para determinar su capacidad de inducir la escisión de un ARN objetivo en un modelo de inyección de ovocitos. Los oligonucleótidos antisentido de 25 nucleobases de longitud con 8 u 11 bases malapareadas cerca de los extremos de los oligonucleótidos antisentido fueron capaces de dirigir la escisión específica del ARNm objetivo, aunque en menor medida que los oligonucleótidos antisentido que no contenían malapareamientos. De manera similar, se logró la escisión específica del objetivo usando oligonucleótidos antisentido de 13 nucleobases, incluyendo aquellos con 1 o 3 malapareamientos.

**[0175]** Gautschi et al (J. Natl. Cancer Inst. 93 463-471, marzo de 2001) demostraron la capacidad de un oligonucleótido que tiene un 100% de complementariedad con el ARNm de bcl-2 y que tiene 3 malapareamientos

con el ARNm de bcl-xL para reducir la expresión tanto de bcl-2 como de bcl-xL *in vitro* e *in vivo.* Además, este oligonucleótido demostró una potente actividad antitumoral *in vivo.*

**[0176]** Maher y Dolnick (Nuc. Acid. Res. 16: 3341-3358,1988) probaron una serie de oligonucleótidos antisentido en tándem de 14 nucleobases, y un oligonucleótido antisentido de 28 y 42 nucleobases compuesto por la secuencia de dos o tres de los oligonucleótidos antisentido en tándem, respectivamente, para su capacidad para detener la traducción de DHFR humano en un ensayo de reticulocitos de conejo. Cada uno de los tres oligonucleótidos antisentido de 14 nucleobases solo fue capaz de inhibir la traducción, aunque en menor nivel que los oligonucleótidos antisentido de 28 o 42 nucleobases.

*Motivos de compuestos antisentido*

**[0177]** En ciertas realizaciones, los compuestos antisentido dirigidos a un ácido nucleico de C9ORF72 tienen subunidades modificadas químicamente dispuestas en patrones, o motivos, para conferir a los compuestos antisentido propiedades como actividad inhibidora mejorada, afinidad de unión aumentada para un ácido nucleico objetivo o resistencia a la degradación por nucleasas *in vivo*.

**[0178]** Los compuestos antisentido quiméricos contienen típicamente por lo menos una región modificada para conferir una resistencia aumentada a la degradación de la nucleasa, una captación celular aumentada, una afinidad de unión aumentada por el ácido nucleico objetivo y/o una actividad inhibidora aumentada. Una segunda región de un compuesto antisentido quimérico puede servir opcionalmente como sustrato para la endonucleasa celular RNasa H, que escinde la cadena de ARN de un dúplex de ARN:ADN.

**[0179]** Los compuestos antisentido que tienen un motivo gapmer se consideran compuestos antisentido quiméricos. En un gapmer, una región interna que tiene una pluralidad de nucleótidos que soporta la escisión de RNasaH se coloca entre regiones externas que tienen una pluralidad de nucleótidos que son químicamente distintos de los nucleósidos de la región interna. En el caso de un oligonucleótido antisentido que tiene un motivo gapmer, el segmento de hueco sirve generalmente como sustrato para la escisión de la endonucleasa, mientras que los segmentos de ala comprenden nucleósidos modificados. En ciertas realizaciones, las regiones de un gapmer se diferencian por los tipos de fracciones de azúcar que comprenden cada región distinta. Los tipos de fracciones de azúcar que se usan para diferenciar las regiones de un gapmer pueden incluir en algunas realizaciones β-D-ribonucleósidos, β-D-desoxirribonucleósidos, nucleósidos 2'-modificados (tales nucleósidos 2'-modificados pueden incluir 2'-MOE, y 2'-O-$CH_3$, entre otros), y nucleósidos modificados con azúcar bicíclico (tales nucleósidos modificados con azúcar bicíclico pueden incluir aquellos que tienen un puente 4'-$(CH_2)$nO-2', donde n=1 o n=2 y 4'-$CH_2$-O-$CH_2$-2'). Preferiblemente, cada región distinta comprende fracciones de azúcar uniformes. El motivo ala-hueco-ala se describe con frecuencia como "X-Y-Z", donde "X" representa la longitud de la región de ala 5', "Y" representa la longitud de la región del hueco, y "Z" representa la longitud de la región de ala 3'. Como se usa en la presente, un gapmer descrito como "X-Y-Z" tiene una configuración tal que el segmento de hueco está colocado inmediatamente adyacente a cada uno de los segmentos de ala 5' y el segmento de ala 3'. Por tanto, no existen nucleótidos intermedios entre el segmento de ala 5' y el segmento de hueco, o el segmento de hueco y el segmento de ala 3'. Cualquiera de los compuestos antisentido descritos en la presente puede tener un motivo gapmer. En algunas realizaciones, X y Z son iguales, en otras realizaciones son diferentes. En una realización preferida, Y tiene entre 8 y 15 nucleótidos. X, Y o Z pueden ser cualquiera de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30 o más nucleótidos. Por tanto, los gapmers descritos en la presente incluyen, pero no están limitados a, por ejemplo 5-10-5, 5-10-4, 4-10-4, 4-10-3, 3-10-3, 2-10-2, 5-9-5, 5-9-4, 4-9-5, 5-8-5, 5-8-4, 4-8-5, 5-7-5, 4-7-5, 5-7-4, o 4-7-4.

**[0180]** En ciertas realizaciones, el compuesto antisentido tiene un motivo "wingmer", que tiene una configuración ala-hueco o hueco-ala, es decir, una configuración X-Y o Y-Z como se ha descrito anteriormente para la configuración gapmer. Por tanto, las configuraciones wingmer descritas en la presente incluyen, pero no están limitadas a, por ejemplo 5-10, 8-4, 4-12, 12-4, 3-14, 16-2, 18-1, 10-3, 2- 10, 1-10, 8-2, 2-13, 5-13, 5-8 o 6-8.

**[0181]** En ciertas realizaciones, los compuestos antisentido dirigidos a un ácido nucleico de C9ORF72 poseen un motivo gapmer 5-10-5.

**[0182]** En ciertas realizaciones, los compuestos antisentido dirigidos a un ácido nucleico de C9ORF72 poseen un motivo gapmer 5-10-4.

**[0183]** En ciertas realizaciones, los compuestos antisentido dirigidos a un ácido nucleico de C9ORF72 poseen un motivo gapmer 4-10-4.

**[0184]** En ciertas realizaciones, los compuestos antisentido dirigidos a un ácido nucleico de C9ORF72 poseen un motivo gapmer 4-10-3.

**[0185]** En ciertas realizaciones, los compuestos antisentido dirigidos a un ácido nucleico de C9ORF72 poseen un

motivo gapmer 5-9-5.

**[0186]** En ciertas realizaciones, un compuesto antisentido dirigido a un ácido nucleico de C9ORF72 tiene un motivo de hueco reducido. En ciertas realizaciones, un oligonucleótido antisentido de hueco reducido dirigido a un ácido nucleico de C9ORF72 tiene un segmento de hueco de 9, 8, 7 o 6 2'-desoxinucleótidos colocados inmediatamente adyacentes y entre los segmentos de ala de 5, 4, 3, 2, o 1 nucleósidos modificados químicamente. En ciertas realizaciones, la modificación química comprende un azúcar bicíclico. En ciertas realizaciones, el azúcar bicíclico comprende un puente de 4' a 2' seleccionado de entre: puente 4'-$(CH_2)_n$-O-2', en donde n es 1 o 2; y 4'-$CH_2$-O-$CH_2$-2'. En ciertas realizaciones, el azúcar bicíclico está comprende un puente 4'-$CH(CH_3)$-O-2'. En ciertas realizaciones, la modificación química comprende una fracción de azúcar 2'-modificado no bicíclico. En ciertas realizaciones, la fracción de azúcar 2'-modificado no bicíclico comprende un grupo 2'-O-metiletilo o un grupo 2'-O-metilo.

**[0187]** En ciertas realizaciones, un compuesto antisentido dirigido a un ácido nucleico de C9ORF72 se modifica uniformemente. En ciertas realizaciones, el compuesto antisentido comprende 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 nucleósidos. En ciertas realizaciones, cada nucleósido se modifica químicamente. En ciertas realizaciones, la modificación química comprende una fracción de azúcar 2'-modificado bicíclico. En ciertas realizaciones, la fracción de azúcar 2'-modificado comprende un grupo 2'-O-metoxietilo. En ciertas realizaciones, la fracción de azúcar 2'-modificado comprende un grupo 2'-O-metilo. En ciertas realizaciones, el compuesto antisentido modificado uniformemente puede dirigirse a C9ORF72, o cualquier porción del mismo, como una expansión de repetición de hexanucleótidos. En ciertas realizaciones, dirigirse a la expansión de repetición de hexanucleótidos con un compuesto antisentido modificado uniformemente redujo la repetición de ARN bloqueando la interacción con proteínas de unión a ARN. En ciertas realizaciones, esto dio como resultado que el ARN tóxico esté ausente de los focos y en su lugar se degrade.

*Ácidos nucleicos objetivo, regiones objetivo y secuencias de nucleótidos*

**[0188]** Las secuencias de nucleótidos que codifican C9ORF72 incluyen, sin limitación, las siguientes: el complemento de GENBANK Nº de registro NM_001256054.1 (incorporada en la presente como SEQ ID NO: 1), GENBANK Nº de registro NT_008413.18 truncado de la nucleobase 27535000 a 27565000 (incorporada en la presente como SEQ ID NO: 2), GENBANK Nº de registro BQ068108.1 (incorporada en la presente como SEQ ID NO: 3), GENBANK Nº de registro NM_018325.3 (incorporada en la presente como SEQ ID NO: 4), GENBANK Nº de registro DN993522. 1 (incorporada en la presente como SEQ ID NO: 5), GENBANK Nº de registro NM_145005.5 (incorporada en la presente como SEQ ID NO: 6), GENBANK Nº de registro DB079375.1 (incorporada en la presente como SEQ ID NO: 7), GENBANK Nº de registro BU194591.1 (incorporada en la presente como SEQ ID NO: 8), Identificador de secuencia 4141_014_A (incorporada en la presente como SEQ ID NO: 9), e identificador de secuencia 4008_73_A (incorporada en la presente como SEQ ID NO: 10).

**[0189]** Se entiende que la secuencia expuesta en cada SEQ ID NO en los Ejemplos contenidos en la presente es independiente de cualquier modificación a una fracción de azúcar, un enlace internucleosídico o una nucleobase. Como tal, los compuestos antisentido definidos por una SEQ ID NO pueden comprender, independientemente, una o más modificaciones a una fracción de azúcar, un enlace internucleosídico o una nucleobase. Los compuestos antisentido descritos por el Número Isis (Isis Nº) indican una combinación de secuencia de nucleobase y motivo.

**[0190]** En ciertas realizaciones, una región objetivo es una región estructuralmente definida del ácido nucleico objetivo. Por ejemplo, una región objetivo puede abarcar una UTR 3', una UTR 5', un exón, un intrón, una unión exón/intrón, una región de codificación, una región de inicio de traducción, una región de terminación de traducción, u otra región de ácido nucleico definida. Las regiones estructuralmente definidas para C9ORF72 pueden obtenerse mediante el número de registro de bases de datos de secuencias como NCBI. En ciertas realizaciones, una región objetivo puede abarcar la secuencia desde un sitio objetivo 5' de un segmento objetivo dentro de la región objetivo hasta un sitio objetivo 3' de otro segmento objetivo dentro de la misma región objetivo.

**[0191]** Dirigir incluye la determinación de por lo menos un segmento objetivo con el que hibrida un compuesto antisentido, de tal manera que se produce un efecto deseado. En ciertas realizaciones, el efecto deseado es una reducción en los niveles de ácido nucleico objetivo del ARNm. En ciertas realizaciones, el efecto deseado es la reducción de los niveles de proteína codificada por el ácido nucleico objetivo o un cambio fenotípico asociado con el ácido nucleico objetivo.

**[0192]** Una región objetivo puede contener uno o más segmentos objetivo. Pueden superponerse múltiples segmentos objetivo dentro de una región objetivo. Alternativamente, pueden no superponerse. En ciertas realizaciones, los segmentos objetivo dentro de una región objetivo están separados por no más de aproximadamente 300 nucleótidos. En ciertas realizaciones, los segmentos objetivo dentro de una región objetivo están separados por una cantidad de nucleótidos que es, es aproximadamente, no es más que, no es más que aproximadamente, 250, 200, 150, 100, 90, 80, 70, 60, 50, 40, 30, 20 o 10 nucleótidos en el ácido nucleico objetivo, o es un intervalo definido por dos cualquiera de los valores anteriores. En ciertas realizaciones, los segmentos objetivo dentro de una región objetivo están separados por no más de, o no más de aproximadamente, 5 nucleótidos en el

ácido nucleico objetivo. En ciertas realizaciones, los segmentos objetivo son contiguos. Se contemplan regiones objetivo definidas por un intervalo que tiene un ácido nucleico de partida que es cualquiera de los sitios objetivo 5' o sitios objetivo 3' enumerados en la presente.

**[0193]** Segmentos objetivo adecuados pueden encontrarse dentro de una UTR 5', una región de codificación, una UTR 3', un intrón, un exón o una unión exón/intrón. Los segmentos objetivo que contienen un codón de inicio o un codón de parada también son segmentos objetivo adecuados. Un segmento objetivo adecuado puede excluir específicamente una determinada región estructuralmente definida, como el codón de inicio o el codón de parada.

**[0194]** La determinación de segmentos objetivo adecuados puede incluir una comparación de la secuencia de un ácido nucleico objetivo con otras secuencias en todo el genoma. Por ejemplo, puede usarse el algoritmo BLAST para identificar regiones de similitud entre diferentes ácidos nucleicos. Esta comparación puede evitar la selección de secuencias de compuestos antisentido que pueden hibridar de manera no específica con secuencias distintas de un ácido nucleico objetivo seleccionado (es decir, secuencias no objetivo o fuera de objetivo).

**[0195]** Puede haber variación en la actividad (por ejemplo, como se define por el porcentaje de reducción de los niveles de ácido nucleico objetivo) de los compuestos antisentido dentro de una región objetivo. En ciertas realizaciones, las reducciones en los niveles de ARNm de C9ORF72 son indicativas de la inhibición de la expresión de C9ORF72. Las reducciones en los niveles de una proteína de C9ORF72 también son indicativas de inhibición de la expresión del ARNm objetivo. La reducción en presencia de focos de ARN de C9ORF72 expandidos es indicativa de la inhibición de la expresión de C9ORF72. Además, los cambios fenotípicos son indicativos de la inhibición de la expresión de C9ORF72. Por ejemplo, la función motora y la respiración mejoradas pueden ser indicativas de inhibición de la expresión de C9ORF72.

*Hibridación*

**[0196]** La hibridación se produce entre un compuesto antisentido divulgado en la presente y un ácido nucleico de C9ORF72. El mecanismo más común de hibridación implica la unión de hidrógeno (por ejemplo, unión de hidrógeno de Watson-Crick, Hoogsteen o Hoogsteen invertida) entre nucleobases complementarias de las moléculas de ácido nucleico.

**[0197]** La hibridación puede producirse bajo varias condiciones. Las condiciones estrictas dependen de la secuencia y están determinadas por la naturaleza y la composición de las moléculas de ácido nucleico que se hibridarán.

**[0198]** Los métodos para determinar si una secuencia es específicamente hibridable con un ácido nucleico objetivo son bien conocidos en la técnica. Los compuestos antisentido proporcionados en la presente son hibridables específicamente con un ácido nucleico de C9ORF72.

*Complementariedad*

**[0199]** Un compuesto antisentido y un ácido nucleico objetivo son complementarios entre sí cuando un número suficiente de nucleobases del compuesto antisentido puede unirse por hidrógeno con las nucleobases correspondientes del ácido nucleico objetivo, de tal manera que se produzca un efecto deseado (por ejemplo, inhibición antisentido de un ácido nucleico objetivo, tal como un ácido nucleico de C9ORF72).

**[0200]** Las nucleobases no complementarias entre un compuesto antisentido y un ácido nucleico de C9ORF72 pueden tolerarse siempre que el compuesto antisentido siga siendo capaz de hibridar específicamente con un ácido nucleico objetivo. Además, un compuesto antisentido puede hibridar sobre uno o más segmentos de un ácido nucleico de C9ORF72 de tal manera que los segmentos intermedios o adyacentes no estén involucrados en el evento de hibridación (por ejemplo, una estructura de giro, malapareamiento o estructura de horquilla).

**[0201]** En ciertos casos, los compuestos antisentido divulgados en la presente, o una porción especificada de los mismos, son, o son por lo menos, un 70%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% complementarios a un ácido nucleico de C9ORF72, una región objetivo, un segmento objetivo o una porción específica del mismo. El porcentaje de complementariedad de un compuesto antisentido con un ácido nucleico objetivo puede determinarse usando métodos rutinarios.

**[0202]** Por ejemplo, un compuesto antisentido en el que 18 de 20 nucleobases del compuesto antisentido son complementarias con una región objetivo y, por lo tanto, se hibridarían específicamente, representarían un 90 por ciento de complementariedad. En este ejemplo, las nucleobases no complementarias restantes pueden estar agrupadas o intercaladas con nucleobases complementarias y no necesitan ser contiguas entre sí o con nucleobases complementarias. Como tal, un compuesto antisentido que tiene 18 nucleobases de longitud que tiene 4 (cuatro) nucleobases no complementarias que están flanqueadas por dos regiones de complementariedad completa con el ácido nucleico objetivo tendría un 77,8% de complementariedad total con el ácido nucleico objetivo. El porcentaje de complementariedad de un compuesto antisentido con una región de un ácido nucleico objetivo

puede determinarse rutinariamente usando programas BLAST (herramientas de búsqueda de alineación local básica) y los programas PowerBLAST conocidos en la técnica (Altschul et al., J. Mol. Biol., 1990, 215, 403 410; Zhang y Madden, Genome Res., 1997, 7, 649 656). El porcentaje de homología, identidad de secuencia o complementariedad puede determinarse, por ejemplo, mediante el programa Gap (Wisconsin Sequence Analysis Package, Versión 8 para Unix, Genetics Computer Group, University Research Park, Madison Wis.), usando los ajustes por defecto, que usa el algoritmo de Smith y Waterman (Adv. Appl. Math., 1981, 2, 482 489).

**[0203]** En ciertas realizaciones, los compuestos antisentido proporcionados en la presente, o porciones especificadas de los mismos, son completamente complementarios (es decir, 100% complementarias) con un ácido nucleico objetivo, o una porción especificada del mismo. Por ejemplo, un compuesto antisentido puede ser completamente complementario con un ácido nucleico de C9ORF72, o una región objetivo, o un segmento objetivo o secuencia objetivo del mismo. Como se usa en la presente, "completamente complementario" significa que cada nucleobase de un compuesto antisentido es capaz de emparejar bases precisas con las nucleobases correspondientes de un ácido nucleico objetivo. Por ejemplo, un compuesto antisentido de 20 nucleobases es completamente complementario con una secuencia objetivo que tiene una longitud de 400 nucleobases, siempre que haya una porción correspondiente de 20 nucleobases del ácido nucleico objetivo que sea completamente complementaria con el compuesto antisentido. Completamente complementario también puede usarse en referencia a una porción especificada del primer y/o el segundo ácido nucleico. Por ejemplo, una porción de 20 nucleobases de un compuesto antisentido de 30 nucleobases puede ser "completamente complementaria" con una secuencia objetivo que tiene una longitud de 400 nucleobases. La porción de 20 nucleobases del oligonucleótido de 30 nucleobases es completamente complementaria con la secuencia objetivo si la secuencia objetivo tiene una porción correspondiente de 20 nucleobases en donde cada nucleobase es complementaria con la porción de 20 nucleobases del compuesto antisentido. Al mismo tiempo, el compuesto antisentido de 30 nucleobases entero puede o no ser completamente complementario con la secuencia objetivo, dependiendo de si las 10 nucleobases restantes del compuesto antisentido también son complementarias con la secuencia objetivo.

**[0204]** La localización de una nucleobase no complementaria puede estar en el extremo 5' o en el extremo 3' del compuesto antisentido. Alternativamente, la nucleobase o nucleobases no complementarias pueden estar en una posición interna del compuesto antisentido. Cuando están presentes dos o más nucleobases no complementarias, pueden ser contiguas (es decir, enlazadas) o no contiguas. En una realización, una nucleobase no complementaria está localizada en el segmento de ala de un oligonucleótido antisentido gapmer.

**[0205]** En ciertos casos, los compuestos antisentido que tienen, o tienen hasta 12, 13, 14, 15, 16, 17, 18, 19 o 20 nucleobases de longitud no comprenden más de 4, no más de 3, no más de 2, o no más de 1 nucleobases no complementarias con respecto a un ácido nucleico objetivo, como un ácido nucleico de C9ORF72, o una porción especificada del mismo.

**[0206]** En ciertos casos, los compuestos antisentido que tienen, o tienen hasta 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 o 30 nucleobases de longitud comprenden no más de 6, no más de 5, no más de 4, no más de 3, no más de 2, o no más de 1 nucleobases no complementarias con respecto a un ácido nucleico objetivo, como como un ácido nucleico de C9ORF72, o una porción especificada del mismo.

**[0207]** Los compuestos antisentido proporcionados en la presente también incluyen aquellos que son complementarios con una porción de un ácido nucleico objetivo. Como se usa en la presente, "porción" se refiere a un número definido de nucleobases contiguas (es decir, enlazadas) dentro de una región o segmento de un ácido nucleico objetivo. Una "porción" también puede referirse a un número definido de nucleobases contiguas de un compuesto antisentido. En ciertos casos, los compuestos antisentido son complementarios con por lo menos una porción de 8 nucleobases de un segmento objetivo. En ciertos casos, los compuestos antisentido son complementarios con por lo menos una porción de 9 nucleobases de un segmento objetivo. En ciertos casos, los compuestos antisentido son complementarios con por lo menos una porción de 10 nucleobases de un segmento objetivo. En ciertas realizaciones, los compuestos antisentido, son complementarios con por lo menos una porción de 11 nucleobases de un segmento objetivo. En ciertas realizaciones, los compuestos antisentido son complementarios con por lo menos una porción de 12 nucleobases de un segmento objetivo. En ciertas realizaciones, los compuestos antisentido son complementarios con por lo menos una porción de 13 nucleobases de un segmento objetivo. En ciertas realizaciones, los compuestos antisentido son complementarios con por lo menos una porción de 14 nucleobases de un segmento objetivo. En ciertas realizaciones, los compuestos antisentido son complementarios con por lo menos una porción de 15 nucleobases de un segmento objetivo. También se contemplan compuestos antisentido que son complementarios con por lo menos una porción de 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más nucleobases de un segmento objetivo, o un intervalo definido por dos cualquiera de estos valores.

*Identidad*

**[0208]** Los compuestos antisentido proporcionados en la presente también pueden tener un porcentaje de identidad definido para una secuencia de nucleótidos particular, SEQ ID NO, o compuesto representado por un número Isis

específico, o una porción del mismo. Como se usa en la presente, un compuesto antisentido es idéntico a la secuencia divulgada en la presente si tiene la misma capacidad de emparejamiento de nucleobases. Por ejemplo, un ARN que contiene uracilo en lugar de timidina en una secuencia de ADN divulgada se consideraría idéntico a la secuencia de ADN ya que tanto el uracilo como la timidina se emparejan con adenina. También se contemplan versiones acortadas y alargadas de los compuestos antisentido descritos en la presente, así como compuestos que tienen bases no idénticas con respecto a los compuestos antisentido proporcionados en la presente. Las bases no idénticas pueden ser adyacentes entre sí o estar dispersas por todo el compuesto antisentido. El porcentaje de identidad de un compuesto antisentido se calcula de acuerdo con el número de bases que tienen un emparejamiento de bases idéntico con respecto a la secuencia con la que se está comparando.

**[0209]** En ciertas realizaciones, los compuestos antisentido, o porciones de los mismos, son por lo menos un 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o 100% idénticos a uno o más de los compuestos antisentido o SEQ ID NO, o una porción de los mismos, divulgados en la presente.

**[0210]** En ciertas realizaciones, una porción del compuesto antisentido se compara con una porción de igual longitud del ácido nucleico objetivo. En ciertas realizaciones, una porción de 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 nucleobases se compara con una porción de igual longitud del ácido nucleico objetivo.

**[0211]** En ciertas realizaciones, una porción del oligonucleótido antisentido se compara con una porción de igual longitud del ácido nucleico objetivo. En ciertas realizaciones, una porción de 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 nucleobases se compara con una porción de igual longitud del ácido nucleico objetivo.

*Modificaciones*

**[0212]** Un nucleósido es una combinación de azúcar-base. La porción de nucleobases (también conocida como base) del nucleósido es normalmente una fracción de base heterocíclica. Los nucleótidos son nucleósidos que incluyen además un grupo fosfato enlazado covalentemente a la porción de azúcar del nucleósido. Para aquellos nucleósidos que incluyen un azúcar pentofuranosilo, el grupo fosfato puede unirse a la fracción de hidroxilo 2', 3' o 5' del azúcar. Los oligonucleótidos se forman a través del enlace covalente de nucleósidos adyacentes entre sí, para formar un oligonucleótido polimérico lineal. Dentro de la estructura oligonucleotídica, los grupos fosfato son referidos comúnmente como formadores de los enlaces internucleosídicos del oligonucleótido.

**[0213]** Las modificaciones a los compuestos antisentido abarcan sustituciones o cambios en los enlaces internucleosídicos, fracciones de azúcar o nucleobases. Los compuestos antisentido modificados se prefieren a menudo a las formas nativas debido a propiedades deseables como, por ejemplo, captación celular mejorada, afinidad mejorada por el objetivo de ácido nucleico, estabilidad aumentada en presencia de nucleasas o actividad inhibidora mejorada.

**[0214]** Los nucleósidos modificados químicamente también pueden emplearse para aumentar la afinidad de unión de un oligonucleótido antisentido acortado o truncado para su ácido nucleico objetivo. En consecuencia, a menudo pueden obtenerse resultados comparables con compuestos antisentido más cortos que tienen tales nucleósidos modificados químicamente.

*Enlaces internucleosídicos modificados*

**[0215]** El enlace internucleósido de origen natural del ARN y el ADN es un enlace fosfodiéster de 3' a 5'. Los compuestos antisentido que tienen uno o más enlaces internucleosídicos modificados, es decir, de origen no natural, se seleccionan a menudo sobre los compuestos antisentido que tienen enlaces internucleósidos de origen natural debido a propiedades deseables como, por ejemplo, captación celular mejorada, afinidad mejorada por los ácidos nucleicos objetivo y estabilidad aumentada en presencia de nucleasas.

**[0216]** Los oligonucleótidos que tienen enlaces internucleosídicos modificados incluyen enlaces internucleosídicos que retienen un átomo de fósforo, así como enlaces internucleósidos que no tienen un átomo de fósforo. Los enlaces internucleosídicos que contienen fósforo representativos incluyen, pero no están limitados a, fosfodiésteres, fosfotriésteres, metilfosfonatos, fosforamidato y fosforotioatos. Los métodos de preparación de enlaces que contienen fósforo y que no contienen fósforo son bien conocidos.

**[0217]** En ciertas realizaciones, los compuestos antisentido dirigidos a un ácido nucleico de C9ORF72 comprenden uno o más enlaces internucleosídicos modificados. En ciertas realizaciones, los enlaces internucleosídicos modificados se intercalan en todo el compuesto antisentido. En ciertas realizaciones, los enlaces internucleosídicos modificados son enlaces de fosforotioato. En ciertas realizaciones, cada enlace internucleosídico de un compuesto antisentido es un enlace internucleósido de fosforotioato.

*Fracciones de azúcar modificado*

**[0218]** Los compuestos antisentido pueden contener opcionalmente uno o más nucleósidos en los que se ha modificado el grupo azúcar. Tales nucleósidos modificados en el azúcar pueden impartir una estabilidad de nucleasas mejorada, afinidad de unión mejorada, alguna otra propiedad biológica beneficiosa a los compuestos antisentido. En ciertas realizaciones, los nucleósidos comprenden fracciones del anillo de ribofuranosa modificado químicamente. Los ejemplos de anillos de ribofuranosa modificados químicamente incluyen, sin limitación, la adición de grupos sustituyentes (incluyendo grupos sustituyentes 5' y 2', formación de puentes de átomos del anillo no geminales para formar ácidos nucleicos bicíclicos (BNA), la sustitución del átomo de oxígeno del anillo ribosilo con S, N(R) o $C(R_1)(R_2)$ (R, $R_1$ y R son cada uno independientemente H, alquilo $C_1$-$C_{12}$ o un grupo protector) y combinaciones de los mismos. Los ejemplos de azúcares modificados químicamente incluyen nucleósido 2'-F-5'-metilo sustituido (ver Solicitud Internacional de PCT WO 2008/101157 publicada el 21/08/08 para otros nucleósidos 5',2'-bis sustituidos divulgados) o reemplazo del átomo de oxígeno del anillo de ribosilo con S con sustitución adicional en la posición 2' (ver la Solicitud de Patente de Estados Unidos publicada US2005-0130923, publicada el 16 de junio de 2005) o alternativamente la sustitución 5' de un BNA (ver Solicitud internacional de PCT WO 2007/134181 publicada el 22/11/07 en donde el LNA está sustituido con, por ejemplo, un grupo 5'-metilo o uno 5'-vinilo).

**[0219]** Los ejemplos de nucleósidos que tienen fracciones de azúcar modificado incluyen, sin limitación, nucleósidos que comprenden grupos sustituyentes 5'-vinilo, 5'-metilo (*R* o *S*), 4'-S, 2'-F, 2'-$OCH_3$, 2'-$OCH_2CH_3$, 2'-$OCH_2CH_2F$ y 2'O-$(CH_2)_2OCH_3$. El sustituyente en la posición 2' también puede seleccionase de alilo, amino, azido, tio, O-alilo, O-alquilo $C_1$-$C_{10}$, $OCF_3$, $OCH_2$ F, $O(CH_2)_2SCH_3$, $O(CH_2)_2$-O-$N(R_m)(R_n)$, O-$CH_2$-C(=O)-$N(R_m)(R_n)$ y O-$CH_2$-C(=O)-$N(R_l)$-$(CH_2)_2$-$N(R_m)(R_n)$, donde cada $R_1$, $R_m$ y $R_n$ es, independientemente, H o alquilo $C_1$-$C_{10}$ sustituido o no sustituido.

**[0220]** Como se usa en la presente, "nucleósidos bicíclicos" se refiere a nucleósidos modificados que comprenden una fracción de azúcar bicíclico. Los ejemplos de nucleósidos bicíclicos incluyen, sin limitación, nucleósidos que comprenden un puente entre los átomos del anillo ribosilo 4' y 2'. En ciertas realizaciones, los compuestos antisentido proporcionados en la presente incluyen uno o más nucleósidos bicíclicos que comprenden un puente de 4' a 2'. Los ejemplos de tales nucleósidos bicíclicos con puente de 4' a 2' incluyen, pero no están limitados a, una de las fórmulas: 4'-$(CH_2)$-O-2' (LNA); 4'-$(CH_2)$-S-2'; 4'-$(CH_2)_2$-O-2' (ENA); 4'-$CH(CH_3)$-O-2' y 4'-$CH(CH_2OCH_3)$-O-2' (y análogos de los mismos ver Patente de Estados Unidos 7.399.845, concedida el 15 de julio de 2008); 4'-$C(CH_3)(CH_3)$-O-2' (y análogos de la misma, ver la Solicitud internacional publicada WO/2009/006478, publicada el 8 de enero de 2009); 4'-$CH_2$-$N(OCH_3)$-2' (y análogos de la misma, ver la Solicitud internacional publicada WO/2008/150729, publicada el 11 de diciembre de 2008); 4'-$CH_2$-$ON(CH_3)$-2' (ver la solicitud de patente publicada de Estados Unidos US2004-0171570, publicada el 2 de septiembre de 2004); 4'-$CH_2$-N(R)-O-2', en donde R es H, alquilo $C_1$-$C_{12}$, o un grupo protector (ver la Patente de Estados Unidos 7.427.672, concedida el 23 de septiembre de 2008); 4'-$CH_2$-C(H)$(CH_3)$-2' (ver Chattopadhyaya et al, J. Org Chem, 2009, 74, 118-134); y 4'-$CH_2$-C-(=$CH_2$)-2' (y análogos de la misma, ver la Solicitud internacional publicada WO 2008/154401, publicada el 8 de diciembre de 2008)

**[0221]** También pueden encontrarse informes adicionales relacionados con los nucleósidos bicíclicos en la bibliografía publicada (ver, por ejemplo: Singh et al., Chem. Commun., 1998, 4, 455-456; Koshkin et al., Tetrahedron, 1998, 54, 3607-3630; Wahlestedt et al., Proc. Natl. Acad. Sci. U. S. A., 2000, 97, 5633-5638; Kumar et al., Bioorg. Med. Chem. Lett., 1998, 8, 2219-2222; Singh et al., J. Org. Chem., 1998, 63, 10035-10039; Srivastava et al., J. Am. Chem. Soc., 2007, 129(26) 8362-8379; Elayadi et al., Curr. Opinion Invest. Drugs, 2001, 2, 558-561; Braasch et al., Chem. Biol., 2001, 8, 1-7; and Orum et al., Curr. Opinion Mol. Ther., 2001, 3, 239-243; Patentes de Estados Unidos Nº 6.268.490; 6.525.191; 6.670.461; 6.770.748; 6.794.499; 7.034.133; 7.053.207; 7.399.845; 7.547.684; y 7.696.345; Publicaciones de Patente de Estados Unidos Nº US2008-0039618; US2009-0012281; Nº de Serie de Patente de Estados Unidos 60/989,574; 61/026,995; 61/026,998; 61/056,564; 61/086,231; 61/097,787; y 61/099,844; Solicitudes internacionales de PCT publicadas WO 1994/014226; WO 2004/106356; WO 2005/021570; WO 2007/134181; WO 2008/150729; WO 2008/154401; y WO 2009/006478. Cada uno de los nucleósidos bicíclicos anteriores puede prepararse con una o más configuraciones de azúcar estereoquímicas que incluyen, por ejemplo, α-L-ribofuranosa y β-D-ribofuranosa (ver la solicitud internacional de PCT PCT/DK98/00393, publicada el 25 de marzo de 1999 como WO 99/14226)

**[0222]** En ciertas realizaciones, las fracciones de azúcar bicíclicos de nucleósidos de BNA incluyen, pero no están limitadas a, compuestos que tienen por lo menos un puente entre la posición 4' y 2' de la fracción de azúcar pentofuranosilo en donde dichos puentes comprenden independientemente 1 o de 2 a 4 grupos enlazados seleccionados independientemente de -$[C(R_a)(R_b)]_n$-, -$C(R_a)$=$C(R_b)$-, -$C(R_a)$=N-, -C(=O)-, -C(=$NR_a$)-, -C(=S)-, -O-, -$Si(R_a)_2$-, -$S(=O)_x$-, y -$N(R_a)$-;
en donde:

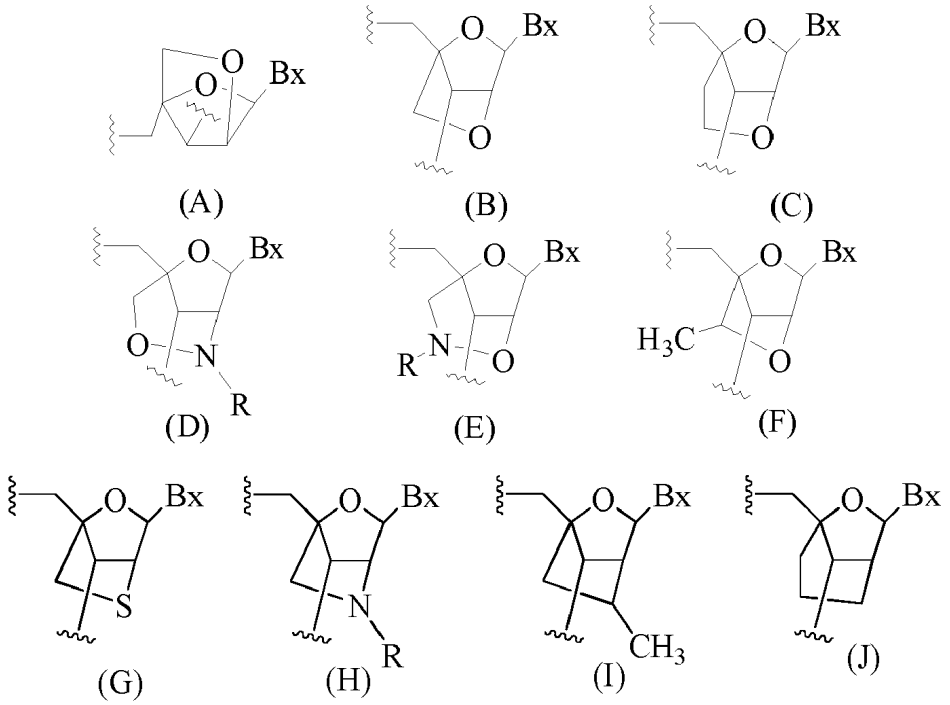
x es 0, 1 o 2;
n es 1, 2, 3 o 4;
cada $R_a$ y $R_b$ es, independientemente, H, un grupo protector, hidroxilo, alquilo $C_1$-$C_{12}$, alquilo $C_1$-$C_{12}$ sustituido, alquenilo $C_2$-$C_{12}$, alquenilo $C_2$-$C_{12}$ sustituido, alquinilo $C_2$-$C_{12}$, alquinilo $C_2$-$C_{12}$ sustituido, arilo $C_5$-$C_{20}$, arilo $C_5$-$C_{20}$ sustituido, radical heterociclo, radical heterociclo sustituido, heteroarilo, heteroarilo sustituido,

radical alicíclico $C_5$-$C_7$, radical alicíclico $C_5$-$C_7$ sustituido, halógeno, $OJ_1$, $NJ_1J_2$, $SJ_1$, $N_3$, $COOJ_1$, acilo (C(=O)-H), acilo sustituido, CN, sulfonilo ($S(=O)_2$-$J_1$), o sulfoxilo (S(=O)-$J_1$); y
cada $J_1$ y $J_2$ es, independientemente, H, alquilo $C_1$-$C_{12}$, alquilo $C_1$-$C_{12}$ sustituido, alquenilo $C_2$-$C_{12}$, alquenilo $C_2$-$C_{12}$ sustituido, alquinilo $C_2$-$C_{12}$, alquinilo $C_2$-$C_{12}$ sustituido, arilo $C_5$-$C_{20}$, arilo $C_5$-$C_{20}$ sustituido, acilo (C(=O)-H), acilo sustituido, un radical heterociclo, un radical heterociclo sustituido, aminoalquilo $C_1$-$C_{12}$, aminoalquilo $C_1$-$C_{12}$ sustituido o un grupo protector.

**[0223]** En ciertas realizaciones, el puente de una fracción de azúcar bicíclico es -$[C(R_a)(R_b)]_n$-, -$[C(R_a)(R_b)]_n$-O-, -$C(R_aR_b)$-N(R)-O- o -$C(R_aR_b)$-O-N(R)-. En ciertas realizaciones, el puente es 4'-$CH_2$-2', 4'-$(CH_2)_2$-2', 4'-$(CH_2)_3$-2', 4'-$CH_2$-O-2', 4'-$(CH_2)_2$-O-2', 4'-$CH_2$-ON(R)-2' y 4'-$CH_2$-N(R)-O-2'-, en donde cada R es, independientemente, H, un grupo protector o alquilo $C_1$-$C_{12}$.

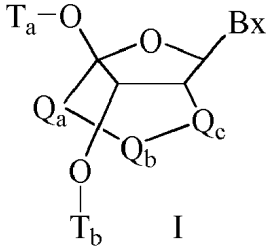
**[0224]** En ciertas realizaciones, los nucleósidos bicíclicos se definen adicionalmente por configuración isomérica. Por ejemplo, un nucleósido que comprende un puente 4'-2' metileno-oxi, puede estar en la configuración α-L o en la configuración β-D. Anteriormente, se han incorporado α-L-metilenoxi (4'-$CH_2$-O-2') BNA en oligonucleótidos antisentido que mostraban actividad antisentido (Frieden et al., Nucleic Acids Research, 2003, 21, 6365-6372).

**[0225]** En ciertas realizaciones, los nucleósidos bicíclicos incluyen, pero no están limitados a, (A) α-L-metileneoxi (4'-$CH_2$-O-2') BNA, (B) β-D-metileneoxi (4'-$CH_2$-O-2') BNA, (C) etilenoxi (4'-$(CH_2)_2$-O-2') BNA, (D) aminooxi (4'-$CH_2$-O-N(R)-2') BNA, (E) oxiamino (4'-$CH_2$-N(R)-O-2') BNA y (F) metil(metileneoxi) (4'-$CH(CH_3)$-O-2') BNA, (G) metilen-tio (4'-$CH_2$-S-2') BNA, (H) metilen-amino (4'-$CH_2$-N(R)-2') BNA, (I) metil carbocíclico (4'-$CH_2$-$CH(CH_3)$-2') BNA, y (J) propileno carbocíclico (4'-$(CH_2)_3$-2') BNA como se representa a continuación.




en donde Bx es la fracción de base y R es independientemente H, un grupo protector o alquilo $C_1$-$C_{12}$.

**[0226]** En ciertas realizaciones, se proporcionan nucleósidos bicíclicos que tienen la Fórmula I:

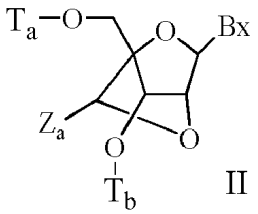


I

en donde:

Bx es una fracción de base heterocíclica;
$-Q_a-Q_b-Q_c-$ es $-CH_2-N(R_c)-CH_2-$, $-C(=O)-N(R_c)-CH_2-$, $-CH_2-ON(R_c)-$, $-CH_2-N(R_c)-O-$ o $-N(R_c)-O-CH_2$;
$R_c$ es alquilo $C_1$-$C_{12}$ o un grupo protector de amino; y
$T_a$ y $T_b$ son cada uno, independientemente H, un grupo protector de hidroxilo, un grupo conjugado, un grupo de fósforo reactivo, una fracción de fósforo o una unión covalente a un medio de soporte.

**[0227]** En ciertas realizaciones, se proporcionan nucleósidos bicíclicos que tienen la Fórmula II:



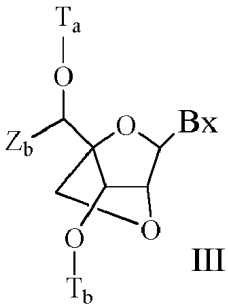
II

en donde:

Bx es una fracción de base heterocíclica;
$T_a$ y $T_b$ son cada uno, independientemente H, un grupo protector de hidroxilo, un grupo conjugado, un grupo de fósforo reactivo, una fracción de fósforo o una unión covalente a un medio de soporte;
$Z_a$ es alquilo $C_1$-$C_6$, alquenilo $C_2$-$C_6$, alquinilo $C_2$-$C_6$, alquilo $C_1$-$C_6$ sustituido, alquenilo $C_2$-$C_6$ sustituido, alquinilo $C_2$-$C_6$ sustituido, acilo, acilo sustituido, amida sustituida, tiol o tio sustituido.

**[0228]** En una realización, cada uno de los grupos sustituidos está, independientemente, mono o poli sustituido con grupos sustituyentes seleccionados independientemente de halógeno, oxo, hidroxilo, $OJ_c$, $NJ_cJ_d$, $SJ_c$, $N_3$, $OC(=X)J_c$, y $NJ_eC(=X)NJ_cJ_d$, en donde cada $J_c$, $J_d$ y $J_e$ es, independientemente, H, alquilo $C_1$-$C_6$, o alquilo $C_1$-$C_6$ sustituido y X es O $NJ_c$.

**[0229]** En ciertas realizaciones, se proporcionan nucleósidos bicíclicos que tienen la Fórmula III:

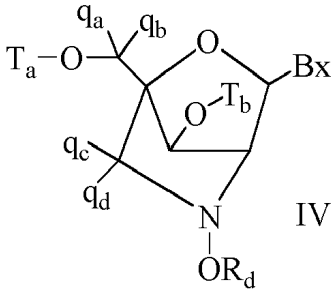


III

en donde:

Bx es una fracción de base heterocíclica;
$T_a$ y $T_b$ son cada uno, independientemente H, un grupo protector de hidroxilo, un grupo conjugado, un grupo de fósforo reactivo, una fracción de fósforo o una unión covalente a un medio de soporte;
$Z_b$ es alquilo $C_1$-$C_6$, alquenilo $C_2$-$C_6$, alquinilo $C_2$-$C_6$, alquilo $C_1$-$C_6$ sustituido, alquenilo $C_2$-$C_6$ sustituido, alquinilo $C_2$-$C_6$ sustituido o acilo sustituido (C(=O)-).

**[0230]** En ciertas realizaciones, se proporcionan nucleósidos bicíclicos que tienen la Fórmula IV:

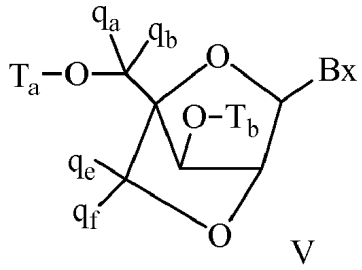


IV

en donde:

Bx es una fracción de base heterocíclica;
$T_a$ y $T_b$ son cada uno, independientemente H, un grupo protector de hidroxilo, un grupo conjugado, un grupo de fósforo reactivo, una fracción de fósforo o una unión covalente a un medio de soporte;
$R_d$ es alquilo $C_1$-$C_6$, alquilo $C_1$-$C_6$ sustituido, alquenilo $C_2$-$C_6$, alquenilo $C_2$-$C_6$ sustituido, alquinilo $C_2$-$C_6$, o alquinilo $C_2$-$C_6$ sustituido;
cada $q_a$, $q_b$, $q_c$ y $q_d$ es, independientemente, H, halógeno, alquilo $C_1$-$C_6$, alquilo $C_1$-$C_6$ sustituido, alquenilo $C_2$-$C_6$, alquenilo $C_2$-$C_6$ sustituido, alquinilo $C_2$-$C_6$, o alquinilo $C_2$-$C_6$ sustituido, alcoxilo $C_1$-$C_6$, alcoxilo $C_1$-$C_6$ sustituido, acilo, acilo sustituido, aminoalquilo $C_1$-$C_6$ o aminoalquilo $C_1$-$C_6$ sustituido$_6$;

**[0231]** En ciertas realizaciones, se proporcionan nucleósidos bicíclicos que tienen la Fórmula V:




V

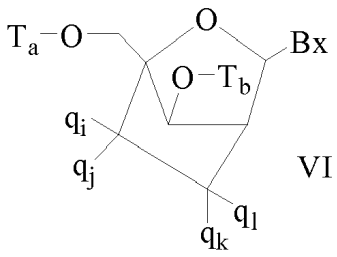
en donde:

Bx es una fracción de base heterocíclica;
$T_a$ y $T_b$ son cada uno, independientemente H, un grupo protector de hidroxilo, un grupo conjugado, un grupo de fósforo reactivo, una fracción de fósforo o una unión covalente a un medio de soporte;
$q_a$, $q_b$, $q_e$ q y $q_f$ son cada uno, independientemente, hidrógeno, halógeno, alquilo $C_1$-$C_{12}$, alquilo $C_1$-$C_{12}$ sustituido, alquenilo $C_2$-$C_{12}$, alquenilo $C_2$-$C_{12}$ sustituido, alquinilo $C_2$-$C_{12}$, alquinilo $C_2$-$C_{12}$ sustituido, alcoxi $C_1$-$C_{12}$, alcoxi $C_1$-$C_{12}$ sustituido, $OJ_j$, $SJ_j$, $SOJ_j$, $SO_2J_j$, $NJ_jJ_k$, $N_3$, CN, $C(=O)OJ_j$, $C(=O)NJ_jJ_k$, $C(=O)J_j$, $O$-$C(=O)NJ_jJ_k$, $N(H)C(=NH)NJ_jJ_k$, $N(H)C(=O)NJ_jJ_k$ o $N(H)C(=S)NJ_jJ_k$;
o $q_e$ y $q_f$ juntos son = $C(q_g)(q_h)$;
$q_g$ y $q_h$ son cada uno, independientemente, H, halógeno, alquilo $C_1$-$C_{12}$ o alquilo $C_1$-$C_{12}$ sustituido.

**[0232]** Se han descrito la síntesis y preparación de los monómeros de metileneoxi (4'-$CH_2$-O-2') BNA adenina, citosina, guanina, 5-metil-citosina, timina y uracilo, junto con su oligomerización y propiedades de reconocimiento de ácidos nucleicos. (Koshkin et al., Tetrahedron, 1998, 54, 3607-3630). Los BNA y la preparación de los mismos también se describen en la WO 98/39352 y la WO 99/14226.

**[0233]** También se han preparado análogos de metilenoxi (4'-$CH_2$-O-2') BNA y 2'-tio-BNA, (Kumar et al., Bioorg. Med. Chem. Lett., 1998, 8, 2219-2222) También se ha descrito la preparación de análogos de nucleósidos bloqueados que comprenden dúplex de oligodesoxirribonucleótidos como sustratos para polimerasas de ácidos nucleicos (Wengel et al., WO 99/14226) Además, en la técnica se ha descrito la síntesis de 2'-amino-BNA, un nuevo análogo de oligonucleótido de alta afinidad conformacionalmente restringido (Singh et al., J. Org. Chem., 1998, 63, 10035-10039). Además, se han preparado 2'-amino- y 2'-metilamino-BNA y se ha informado anteriormente sobre la estabilidad térmica de sus dúplex con cadenas de ARN y ADN complementarias.

**[0234]** En ciertas realizaciones, se proporcionan nucleósidos bicíclicos que tienen la Fórmula VI:




VI

en donde:

Bx es una fracción de base heterocíclica;
$T_a$ y $T_b$ son cada uno, independientemente H, un grupo protector de hidroxilo, un grupo conjugado, un grupo de fósforo reactivo, una fracción de fósforo o una unión covalente a un medio de soporte;
cada $q_i$, $q_j$, $q_k$ y $q_l$ es, independientemente, H, halógeno, alquilo $C_1$-$C_{12}$, alquilo $C_1$-$C_{12}$ sustituido, alquenilo $C_2$-$C_{12}$, alquenilo $C_2$-$C_{12}$ sustituido, alquinilo $C_2$-$C_{12}$, alquinilo $C_2$-$C_{12}$ sustituido, alcoxilo $C_1$-$C_{12}$, alcoxilo $C_1$-$C_{12}$ sustituido, $OJ_j$, $SJ_j$, $SOJ_j$, $SO_2J_j$, $NJ_jJ_k$, $N_3$, CN, $C(=O)OJ_j$, $C(=O)NJ_jJ_k$, $C(=O)J_j$, $O$-$C(=O)NJ_jJ_k$, $N(H)C(=NH)NJ_jJ_k$, $N(H)C(=O)NJ_jJ_k$ or $N(H)C(=S)NJ_jJ_k$; y
$q_i$ y $q_j$ o $q_l$ y q juntos son $=C(q_g)(q_h)$, en donde $q_g$ y $q_h$ son cada uno, independientemente, H, halógeno, alquilo $C_1$-$C_{12}$ o alquilo $C_1$-$C_{12}$ sustituido.

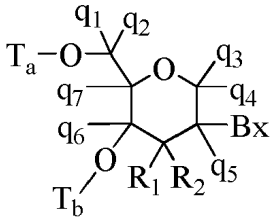
**[0235]** Se ha descrito un nucleósido bicíclico carbocíclico que tiene un puente 4'-$(CH_2)_3$-2' y el puente análogo de alquenilo 4'-CH=CH-$CH_2$-2' (Freier et al., Nucleic Acids Research, 1997, 25(22), 4429-4443 y Albaek et al., J. Org. Chem., 2006, 71, 7731-7740). También se ha descrito la síntesis y preparación de nucleósidos bicíclicos carbocíclicos junto con su oligomerización y estudios bioquímicos (Srivastava et al., J. Am. Chem. Soc., 2007, 129(26), 8362-8379).

**[0236]** Como se usa en la presente, "nucleósido bicíclico 4'-2' " o "nucleósido bicíclico de 4' a 2' " se refiere a un nucleósido bicíclico que comprende un anillo de furanosa que comprende un puente que conecta dos átomos de carbono del anillo de furanosa que conecta el átomo de carbono 2' y el átomo de carbono 4' del anillo de azúcar.

**[0237]** Como se usa en la presente, "nucleósidos monocíclicos" se refiere a nucleósidos que comprenden fracciones de azúcar modificado que no son fracciones de azúcar bicíclico. En ciertas realizaciones, la fracción de azúcar, o análogo de fracción de azúcar, de un nucleósido puede modificarse o sustituirse en cualquier posición.

**[0238]** Como se usa en la presente, "azúcar 2'-modificado" significa un azúcar de furanosilo modificado en la posición 2'. En ciertas realizaciones, tales modificaciones incluyen sustituyentes seleccionados de: un haluro, que incluye, pero no está limitado a, alcoxi sustituido y no sustituido, tioalquilo sustituido y no sustituido, amino alquilo sustituido y no sustituido, alquilo sustituido y no sustituido, alilo sustituido y no sustituido, y alquinilo sustituido y no sustituido. En ciertas realizaciones, las modificaciones 2' se seleccionan de sustituyentes que incluyen, pero no están limitados a: O $[(CH_2)_nO]_mCH_3$, $O(CH_2)_nNH_2$, $O(CH_2)_nCH_3$, $O(CH_2)_nF$, $O(CH_2)_nONH_2$, $OCH_2C(=O)N(H)CH_3$ y $O(CH_2)_nON[(CH_2)_nCH_3]_2$, donde n y m son de 1 a aproximadamente 10. Otros grupos 2'-sustituyentes también pueden seleccionarse de: alquilo $C_1$-$C_{12}$, alquilo sustituido, alquenilo, alquinilo, alcarilo, aralquilo, O-alcarilo o O-aralquilo, SH, $SCH_3$, OCN, Cl, Br, CN, F, $CF_3$, $OCF_3$, $SOCH_3$, $SO_2CH_3$, $ONO_2$, $NO_2$, $N_3$, $NH_2$, heterocicloalquilo, heterocicloalcarilo, aminoalquilamino, polialquilamino, sililo sustituido, un grupo de escisión de ARN, un grupo informador, un intercalador, un grupo para mejorar las propiedades farmacocinéticas, o un grupo para mejorar las propiedades farmacodinámicas de un compuesto antisentido, y otros sustituyentes que tienen propiedades similares. En ciertas realizaciones, los nucleósidos modificados comprenden una cadena lateral 2'-MOE (Baker et al., J. Biol. Chem., 1997, 272, 11944-12000). Se ha descrito que tal sustitución de 2'-MOE tiene una afinidad de unión mejorada en comparación con nucleósidos no modificados y con otros nucleósidos modificados como 2'-*O*-metilo, *O*-propilo y *O*-aminopropilo. También se ha demostrado que los oligonucleótidos que tienen el sustituyente 2'-MOE son inhibidores antisentido de la expresión génica con características prometedoras para uso *in vivo* (Martin, Helv. Chim. Acta, 1995, 78, 486-504; Altmann et al., Chimia, 1996, 50, 168-176; Altmann et al., Biochem. Soc. Trans., 1996, 24, 630-637; y Altmann et al., Nucleosides Nucleotides, 1997, 16, 917-926).

**[0239]** Como se usa en la presente, un "nucleósido de tetrahidropirano modificado" o "nucleósido de THP modificado" significa un nucleósido que tiene un "azúcar" de tetrahidropirano de seis miembros sustituido por el residuo de pentofuranosilo en nucleósidos normales (un sustituto del azúcar). Los nucleósidos de THP modificados incluyen, pero no están limitados a, lo que es referido en la técnica como ácido nucleico de hexitol (HNA), ácido nucleico de anitol (ANA), ácido nucleico de manitol (MNA) (ver Leumann, Bioorg. Med. Chem., 2002, 10, 841-854), fluoro HNA (F-HNA) o aquellos compuestos que tienen la fórmula VII:




VII

en donde independientemente para cada uno de dichos por lo menos un análogo de nucleósido de tetrahidropirano de Fórmula VII:

Bx es una fracción de base heterocíclica;
$T_a$ y $T_b$ son cada uno, independientemente, un grupo de enlace internucleosídico que enlaza el análogo de nucleósido de tetrahidropirano con el compuesto antisentido o uno de $T_a$ y $T_b$ es un grupo de enlace internucleósido que enlaza el análogo de nucleósido de tetrahidropirano con el compuesto antisentido y el otro de $T_a$ y $T_b$ es H, un grupo protector de hidroxilo, un grupo conjugado enlazado o un grupo 5' o 3' terminal;
$q_1$, $q_2$, $q_3$, $q_4$, $q_5$, $q_6$ y $q_7$ son cada uno independientemente, H, alquilo $C_1$-$C_6$, alquilo $C_1$-$C_6$ sustituido, alquenilo $C_2$-$C_6$, alquenilo $C_2$-$C_6$ sustituido, alquinilo $C_2$-$C_6$, o alquinilo $C_2$-$C_6$ sustituido; y cada uno de $R_1$ y $R_2$ se selecciona de hidrógeno, hidroxilo, halógeno, alcoxi sustituido o no sustituido, $NJ_1J_2$, $SJ_1$, $N_3$, $OC(=X)J_1$, $OC(=X)NJ_1J_2$, $NJ_3C(=X)NJ_1J_2$ y CN, en donde X es O, S o $NJ_1$ y cada $J_1$, $J_2$ y $J_3$ es, independientemente, H o alquilo $C_1$-$C_6$.

**[0240]** En ciertas realizaciones, se proporcionan nucleósidos de THP modificados de Fórmula VII en donde $q_1$, $q_2$, $q_3$, $q_4$, $q_5$, $q_6$ y $q_7$ son cada uno H. En ciertas realizaciones, por lo menos uno de $q_1$, $q_2$, $q_3$, $q_4$, $q_5$, $q_6$ y $q_7$ es diferente de H. En ciertas realizaciones, por lo menos uno de $q_1$, $q_2$, $q_3$, $q_4$, $q_5$, $q_6$ y $q_7$ es metilo. En ciertas realizaciones, se proporcionan nucleósidos THP de fórmula VII en donde uno de $R_1$ y $R_2$ es fluoro. En ciertas realizaciones, $R_1$ es fluoro y $R_2$ es H; $R_1$ es metoxi y $R_2$ es H, y $R_1$ es H y $R_2$ es metoxietoxi.

**[0241]** Como se usa en la presente, "2'-modificado" o "2'-sustituido" se refiere a un nucleósido que comprende un azúcar que comprende un sustituyente en la posición 2' diferente de H u OH. Los nucleósidos 2'-modificados incluyen, pero no están limitados a, nucleósidos bicíclicos en los que el puente que conecta dos átomos de carbono del anillo de azúcar conecta el carbono 2' y otro carbono del anillo de azúcar; y nucleósidos con sustituyentes 2' que no forman puentes como alilo, amino, azido, tio, O-alilo, O- alquilo $C_1$-$C_{10}$, $-OCF_3$, $O-(CH_2)_2-O-CH_3$, $2'-O(CH_2)_2SCH_3$, $O-(CH_2)_2-O-N(R_m)(R_n)$, u $O-CH_2-C(=O)-N(R_m)(R_n)$, donde cada $R_m$ y $R_n$ es, independientemente, H o alquilo $C_1$-$C_{10}$ sustituido o no sustituido. Los nucleósidos 2'-modificados pueden comprender además otras modificaciones, por ejemplo en otras posiciones del azúcar y/o en la nucleobase.

**[0242]** Como se usa en la presente, "2'-F" se refiere a un nucleósido que comprende un azúcar que comprende un grupo fluoro en la posición 2'.

**[0243]** Como se usa en la presente, "2'-OMe" o "2'-$OCH_3$" o "2'-O-metilo" se refiere cada uno a un nucleósido que comprende un azúcar que comprende un grupo $-OCH_3$ en la posición 2' del anillo de azúcar.

**[0244]** Como se usa en la presente, "MOE" o "2'-MOE" o "2'-$OCH_2CH_2OCH_3$ " o "2'-O-metoxietilo" se refieren cada uno a un nucleósido que comprende un azúcar que comprende un grupo $-OCH_2CH_2OCH_3$ en la posición 2' del anillo de azúcar.

**[0245]** Como se usa en la presente, "oligonucleótido" se refiere a un compuesto que comprende una pluralidad de nucleósidos enlazados. En la invención, se modifica uno o más de la pluralidad de nucleósidos. En ciertas realizaciones, un oligonucleótido comprende uno o más ribonucleósidos (ARN) y/o desoxirribonucleósidos (ADN).

**[0246]** También se conocen muchos otros sistemas de anillos sustitutos de azúcar biciclo y triciclo en la técnica que pueden usarse para modificar nucleósidos para su incorporación en compuestos antisentido (ver, por ejemplo, el artículo de revisión: Leumann, Bioorg. Med. Chem., 2002, 10, 841-854) Dichos sistemas de anillo pueden experimentar varias sustituciones adicionales para mejorar la actividad.

**[0247]** Los métodos para la preparación de azúcares modificados son bien conocidos por los expertos en la técnica.

**[0248]** En los nucleótidos que tienen fracciones de azúcar modificado, las fracciones de nucleobases (naturales, modificados o una combinación de las mismas) se mantienen para la hibridación con un objetivo de ácido nucleico apropiado.

**[0249]** En ciertas realizaciones, los compuestos antisentido comprenden uno o más nucleósidos que tienen fracciones de azúcar modificado. En ciertas realizaciones, la fracción de azúcar modificado es 2'-MOE. En ciertas realizaciones, los nucleósidos modificados con 2'-MOE están dispuestos en un motivo gapmer. En ciertas realizaciones, la fracción de azúcar modificado es un nucleósido bicíclico que tiene un grupo de puente (4'-$CH(CH_3)$-O-2'). En ciertas realizaciones, los nucleósidos modificados con (4'-$CH(CH_3)$-O-2') están dispuestos a lo largo de las alas de un motivo gapmer.

*Composiciones y métodos para formular composiciones farmacéuticas*

**[0250]** Los oligonucleótidos antisentido pueden mezclarse con sustancias activas o inertes farmacéuticamente aceptables para la preparación de composiciones o formulaciones farmacéuticas. Las composiciones y los métodos para la formulación de composiciones farmacéuticas dependen de una serie de criterios, incluyendo, pero no limitados a, la vía de administración, el alcance de la enfermedad, o la dosis a ser administrada.

**[0251]** Un compuesto antisentido dirigido a un ácido nucleico de C9ORF72 puede utilizarse en composiciones farmacéuticas combinando el compuesto antisentido con un diluyente o portador farmacéuticamente aceptable adecuado. Un diluyente farmacéuticamente aceptable incluye solución salina tamponada con fosfato (PBS). El PBS es un diluyente adecuado para su uso en composiciones que se administran por vía parenteral. Por consiguiente, en un caso, en los métodos descritos en la presente se emplea una composición farmacéutica que comprende un compuesto antisentido dirigido a un ácido nucleico de C9ORF72 y un diluyente farmacéuticamente aceptable. En ciertos casos, el diluyente farmacéuticamente aceptable es PBS. En la invención, el compuesto antisentido es un oligonucleótido antisentido modificado.

**[0252]** Las composiciones farmacéuticas que comprenden compuestos antisentido abarcan cualquier sal, éster farmacéuticamente aceptable, o sales de tales ésteres, o cualquier otro oligonucleótido que, tras la administración a un animal, incluido un humano, sea capaz de proporcionar (directa o indirectamente) el metabolito biológicamente activo o residuo del mismo. Por consiguiente, por ejemplo, la divulgación también se refiere a sales farmacéuticamente aceptables de compuestos antisentido, profármacos, sales farmacéuticamente aceptables de tales profármacos y otros bioequivalentes. Las sales farmacéuticamente aceptables adecuadas incluyen, pero no están limitadas a, sales de sodio y potasio.

**[0253]** Un profármaco puede incluir la incorporación de nucleósidos adicionales en uno o ambos extremos de un compuesto antisentido que se escinden por nucleasas endógenas dentro del cuerpo, para formar el compuesto antisentido activo.

*Compuestos antisentido conjugados*

**[0254]** Los compuestos antisentido pueden estar enlazados covalentemente a una o más fracciones o conjugados que mejoran la actividad, distribución celular o captación celular de los oligonucleótidos antisentido resultantes. Los grupos conjugados típicos incluyen fracciones de colesterol y fracciones de lípidos. Los grupos conjugados adicionales incluyen carbohidratos, fosfolípidos, biotina, fenazina, ácido fólico, fenantridina, antraquinona, acridina, fluoresceínas, rodaminas, cumarinas y colorantes.

**[0255]** Los compuestos antisentido también pueden modificarse para tener uno o más grupos estabilizadores que generalmente están unidos a uno o ambos extremos terminales de los compuestos antisentido para mejorar propiedades como, por ejemplo, la estabilidad de la nucleasa. En los grupos estabilizadores se incluyen las estructuras de tapa. Estas modificaciones terminales protegen el compuesto antisentido que tiene ácido nucleico terminal de la degradación de exonucleasas, y pueden ayudar en la administración y/o localización dentro de una célula. La tapa puede estar presente en el extremo terminal 5'(tapa 5'), o en el extremo terminal 3' (tapa 3'), o puede estar presente en ambos extremos terminales. Las estructuras de tapa son bien conocidas en la técnica e incluyen, por ejemplo, tapas invertidas de desoxi abásicas. Grupos estabilizadores 3' y 5' adicionales que pueden usarse para tapar uno o ambos extremos de un compuesto antisentido para impartir estabilidad de nucleasas incluyen los divulgados en la WO 03/004602 publicada el 16 de enero de 2003.

*Cultivo celular y tratamiento de compuestos antisentido*

**[0256]** Los efectos de los compuestos antisentido sobre el nivel, la actividad o la expresión de los ácidos nucleicos de C9ORF72 pueden probarse *in vitro* en una variedad de tipos de células. Los tipos de células usados para tales análisis están disponibles de proveedores comerciales (por ejemplo, American Type Culture Collection, Manassus, VA; Zen-Bio, Inc., Research Triangle Park, NC; Clonetics Corporation, Walkersville, MD) y se cultivan de acuerdo con las instrucciones del vendedor usando reactivos disponibles comercialmente (por ejemplo, Invitrogen Life Technologies, Carlsbad, CA). Los tipos de células ilustrativas incluyen, pero no están limitadas a, células HepG2, células Hep3B y hepatocitos primarios.

*Pruebas in vitro de oligonucleótidos antisentido*

**[0257]** En la presente se describen métodos para el tratamiento de células con oligonucleótidos antisentido, que pueden modificarse adecuadamente para el tratamiento con otros compuestos antisentido.

**[0258]** En general, las células se tratan con oligonucleótidos antisentido cuando las células alcanzan aproximadamente un 60-80% de confluencia en cultivo.

**[0259]** Un reactivo usado comúnmente para introducir oligonucleótidos antisentido en células cultivadas incluye el reactivo de transfección de lípidos catiónicos LIPOFECTINA (Invitrogen, Carlsbad, CA). Los oligonucleótidos antisentido se mezclan con LIPOFECTINA en OPTI-MEM 1 (Invitrogen, Carlsbad, CA) para lograr la concentración final deseada de oligonucleótido antisentido y una concentración de LIPOFECTINA que varía típicamente de 2 a 12 ug/ml por 100 nM de oligonucleótido antisentido.

**[0260]** Otro reactivo usado para introducir oligonucleótidos antisentido en células cultivadas incluye la

LIPOFECTAMINA (Invitrogen, Carlsbad, CA). El oligonucleótido antisentido se mezcla con LIPOFECTAMINA en medio sérico reducido OPTI-MEM 1 (Invitrogen, Carlsbad, CA) para lograr la concentración deseada de oligonucleótido antisentido y una concentración de LIPOFECTAMINA que generalmente varía de 2 a 12 ug/ml por 100nM de oligonucleótido antisentido.

**[0261]** Otra técnica usada para introducir oligonucleótidos antisentido en células cultivadas incluye la electroporación.

**[0262]** Las células se tratan con oligonucleótidos antisentido mediante métodos rutinarios. Las células se recogen típicamente 16-24 horas después del tratamiento con oligonucleótidos antisentido, momento en el cual se miden los niveles de ARN o proteína de los ácidos nucleicos objetivo mediante métodos conocidos en la técnica y descritos en la presente. En general, cuando los tratamientos se realizan en múltiples replicados, los datos se presentan como la media de los tratamientos replicados.

**[0263]** La concentración de oligonucleótido antisentido usada varía de una línea celular a otra. Los métodos para determinar la concentración óptima de oligonucleótidos antisentido para una línea celular particular son bien conocidos en la técnica. Los oligonucleótidos antisentido se usan típicamente a concentraciones que varían de 1 nM a 300 nM cuando se transfectan con LIPOFECTAMINA. Los oligonucleótidos antisentido se usan a concentraciones más altas que varían de 625 a 20.000 nM cuando se transfectan usando electroporación.

*Aislamiento de ARN*

**[0264]** El análisis de ARN puede realizarse en ARN celular total o poli(A)+ARNm. Los métodos de aislamiento de ARN son bien conocidos en la técnica. El ARN se prepara usando métodos bien conocidos en la técnica, por ejemplo, usando el reactivo TRIZOL (Invitrogen, Carlsbad, CA) de acuerdo con los protocolos recomendados por el fabricante.

*Análisis de inhibición de niveles objetivo o expresión*

**[0265]** La inhibición de los niveles o la expresión de un ácido nucleico de C9ORF72 puede ensayarse de varias maneras conocidas en la técnica. Por ejemplo, los niveles de ácido nucleico objetivo pueden cuantificarse mediante, por ejemplo, análisis de transferencia Northern, reacción en cadena de polimerasa competitiva (PCR), o PCR cuantitativa en tiempo real. El análisis de ARN puede realizarse en ARN celular total o poli(A)+ARNm. Los métodos de aislamiento de ARN son bien conocidos en la técnica. El análisis de transferencia Northern también es rutinario en la técnica. La PCR cuantitativa en tiempo real puede realizarse convenientemente usando el sistema de detección de secuencias ABI PRISM 7600, 7700 o 7900 disponible comercialmente, disponible en PE-Applied Biosystems, Foster City, CA y usado de acuerdo con las instrucciones del fabricante.

*Análisis cuantitativo por PCR en tiempo real de los niveles de ARN objetivo*

**[0266]** La cuantificación de los niveles de ARN objetivo puede lograrse mediante PCR cuantitativa en tiempo real usando el sistema de detección de secuencias ABI PRISM 7600, 7700 o 7900 (PE-Applied Biosystems, Foster City, CA) de acuerdo con las instrucciones del fabricante. Los métodos de PCR cuantitativa en tiempo real son bien conocidos en la técnica.

**[0267]** Antes de la PCR en tiempo real, el ARN aislado se somete a una reacción de transcriptasa inversa (RT), que produce ADN complementario (ADNc) que luego se usa como sustrato para la amplificación por PCR en tiempo real. Las reacciones de RT y PCR en tiempo real se realizan secuencialmente en el mismo pocillo de muestra. Los reactivos de RT y PCR en tiempo real se obtienen de Invitrogen (Carlsbad, CA). Las reacciones de PCR en tiempo real RT se llevan a cabo mediante métodos bien conocidos por los expertos en la técnica.

**[0268]** Las cantidades objetivo del gen (o ARN) obtenidas por PCR en tiempo real se normalizan usando o el nivel de expresión de un gen cuya expresión es constante, como la ciclofilina A, o cuantificando el ARN total usando RIBOGREEN (Invitrogen, Inc. Carlsbad, CA). La expresión de ciclofilina A se cuantifica por PCR en tiempo real, ejecutándose simultáneamente con el objetivo, multiplexando o por separado. El ARN total se cuantifica usando el reactivo de cuantificación de ARN RIBOGREEN (Invetrogen, Inc. Eugene, OR). Los métodos de cuantificación de ARN por RIBOGREEN se enseñan en Jones, L.J. y et al., (Analytical Biochemistry, 1998, 265, 368-374). Se usa un instrumento CYTOFLUOR 4000 (PE Applied Biosystems) para medir la fluorescencia de RIBOGREEN.

**[0269]** Las sondas y los cebadores están diseñados para hibridar con un ácido nucleico de C9ORF72. Los métodos para diseñar sondas y cebadores de PCR en tiempo real son bien conocidos en la técnica, y pueden incluir el uso de software como el software PRIMER EXPRESS (Applied Biosystems, Foster City, CA).

*Análisis de niveles de proteínas*

**[0270]** La inhibición antisentido de los ácidos nucleicos de C9ORF72 puede evaluarse midiendo los niveles de proteína de C9ORF72. Los niveles de proteína de C9ORF72 pueden evaluarse o cuantificarse de una variedad de maneras bien conocidas en la técnica como inmunoprecipitación, análisis de transferencia Western (inmunotransferencia), análisis de inmunoabsorción ligado a enzimas (ELISA), análisis cuantitativos de proteínas, análisis de actividad de proteínas (por ejemplo, ensayos de actividad de caspasa), inmunohistoquímica, inmunocitoquímica o clasificación de células activadas por fluorescencia (FACS). Los anticuerpos dirigidos a un objetivo pueden identificarse y obtenerse de una variedad de fuentes, como el catálogo de anticuerpos MSRS (Aerie Corporation, Birmingham, MI), o pueden prepararse mediante métodos convencionales de generación de anticuerpos monoclonales o policlonales bien conocidos en la técnica. Hay disponibles comercialmente anticuerpos útiles para la detección de C9ORF72 de ratón, rata, mono y humano.

*Pruebas in vivo de compuestos antisentido.*

**[0271]** Los compuestos antisentido, por ejemplo, los oligonucleótidos antisentido, se prueban en animales para evaluar su capacidad de inhibir la expresión de C9ORF72 y producir cambios fenotípicos, como la función motora y la respiración mejoradas. En ciertas realizaciones, la función motora se mide mediante rotarod, fuerza de agarre, ascenso de poste, rendimiento en campo abierto, haz de equilibrio, prueba de huella de pata trasera en el animal. En ciertas realizaciones, la respiración se mide mediante pletismógrafo de cuerpo entero, resistencia invasiva y medidas de cumplimiento en el animal. Las pruebas pueden realizarse en animales normales o en modelos experimentales de enfermedades. Para la administración a animales, los oligonucleótidos antisentido se formulan en un diluyente farmacéuticamente aceptable como solución salina tamponada con fosfato. La administración incluye vías de administración parenteral como intraperitoneal, intravenosa y subcutánea. El cálculo de la dosificación de oligonucleótidos antisentido y la frecuencia de dosificación está dentro de las capacidades de los expertos en la técnica, y depende de factores tales como la vía de administración y el peso corporal del animal. Después de un período de tratamiento con oligonucleótidos antisentido, se aísla el ARN del tejido del SNC o el CSF y se miden los cambios en la expresión del ácido nucleico de C9ORF72.

*Dirigiendo a C9ORF72*

**[0272]** Los oligonucleótidos antisentido descritos en la presente pueden hibridar con un ácido nucleico de C9ORF72 en cualquier etapa del procesamiento de ARN. Por ejemplo, en la presente se describen oligonucleótidos antisentido que son complementarios a un pre-ARNm o un ARNm maduro. Adicionalmente, los oligonucleótidos antisentido descritos en la presente pueden hibridar con cualquier elemento de un ácido nucleico de C9ORF72. Por ejemplo, en la presente se describen oligonucleótidos antisentido que son complementarios a un exón, un intrón, la UTR 5', la UTR 3', una región de repetición, una expansión de repetición de hexanucleótidos, una unión de corte y empalme, un exón: unión de corte y empalme de exón, un silenciador de corte y empalme exónico (ESS), un potenciador de corte y empalme exónico (ESE), exón 1a, exón 1b, exón 1c, exón Id, exón 1e, exón 2, exón 3, exón 4, exón 5, exón 6, exón 7, exón 8, exón 9, exón 10, exón 11, intrón 1, intrón 2, intrón 3, intrón 4, intrón 5, intrón 6, intrón 7, intrón 8, intrón 9, o intrón 10 de un ácido nucleico de C9ORF72.

**[0273]** En ciertas realizaciones, los oligonucleótidos antisentido descritos en la presente hibridan con todas las variantes de C9ORF72. En ciertas realizaciones, los oligonucleótidos antisentido descritos en la presente hibridan selectivamente con ciertas variantes de C9ORF72. En ciertas realizaciones, los oligonucleótidos antisentido descritos en la presente hibridan selectivamente con variantes de C9ORF72 que contienen una expansión de repetición de hexanucleótidos. En ciertas realizaciones, tales variantes de C9ORF72 que contienen una expansión de repetición de hexanucleótidos incluyen las SEQ ID NO: 1-3 y 6-10. En ciertas realizaciones, dicha expansión de repetición de hexanucleótidos comprende por lo menos 30 repeticiones de cualquiera de GGGGCC, GGGGGG, GGGGGC o GGGGCG.

**[0274]** En ciertas realizaciones, los oligonucleótidos antisentido descritos en la presente inhiben la expresión de todas las variantes de C9ORF72. En ciertas realizaciones, los oligonucleótidos antisentido descritos en la presente inhiben la expresión de todas las variantes de C9ORF72 por igual. En ciertas realizaciones, los oligonucleótidos antisentido descritos en la presente inhiben preferiblemente la expresión de ciertas variantes de C9ORF72. En ciertas realizaciones, los oligonucleótidos antisentido descritos en la presente inhiben preferiblemente la expresión de variantes de C9ORF72 que contienen una expansión de repetición de hexanucleótidos. En ciertas realizaciones, tales variantes de C9ORF72 que contienen una expansión de repetición de hexanucleótidos incluyen las SEQ ID NO: 1-3 y 6-10. En ciertas realizaciones, dicha expansión de repetición de hexanucleótidos comprende por lo menos 30 repeticiones de cualquiera de GGGGCC, GGGGGG, GGGGGC o GGGGCG. En ciertas realizaciones, la expansión de repetición de hexanucleótidos forma focos nucleares. Los oligonucleótidos antisentido descritos en la presente son útiles para reducir los focos nucleares. Los focos nucleares pueden reducirse en términos de porcentaje de células con focos así como en número de focos por célula.

**[0275]** En base a estudios anteriores dirigidos a expansiones de repetición, no es posible predecir si los oligonucleótidos antisentido dirigidos a C9ORF72 fuera de la expansión de repetición de hexanucleótidos inhibirían con éxito la expresión de C9ORF72 por dos razones. Primero, la expansión de repetición de C9ORF72 está

localizada en un intrón y no se sabe si el ARN en los focos contiene solo las repeticiones o también la secuencia intrónica flanqueante. Por ejemplo, un estudio anterior sobre la distrofia miotónica tipo 2 (DM2), que es una enfermedad provocada por una mutación de expansión de CCTG en el intrón 1 del gen ZNF9, determinó que las expansiones de DM2 grandes no previenen el corte y empalme pre-ARNm específico de alelos, la exportación nuclear de los transcritos, o niveles de ARNm o proteína en estado estacionario. El estudio demostró además que las inclusiones ribonucleares encontradas asociadas con la enfermedad están enriquecidas para la expansión de CCUG, pero no las secuencias intrónicas flanqueantes. Estos datos sugieren que los efectos moleculares en sentido descendente de la mutación DM2 pueden desencadenarse por la acumulación del tracto de repetición de CCUG solo. Por lo tanto, este estudio implica que dirigir la expansión de repeticionesn de CCUG solamente llevaría a una mejoría de la enfermedad, ya que dirigir las secuencias flanqueantes, especialmente la región en sentido descendente de la expansión de repeticiones, no afectaría a la formación de inclusiones ribonucleares (Margolis et al. Tararear. Mol. Genet., 2006, 15:1808-1815). En segundo lugar, no se sabe cómo de rápido se corta y acumulo en los focos el intrón 1 de C9ORF72 que contiene las repeticiones. Por lo tanto, no es posible predecir si dirigir el pre-ARNm daría como resultado la eliminación de la repetición de ARN y los focos.

*Características de C9OFF72*

**[0276]** Los oligonucleótidos antisentido descritos en la presente pueden hibridar con cualquier variante de C9ORF72 en cualquier estado de procesamiento dentro de cualquier elemento del gen de C9ORF72. Por ejemplo, los oligonucleótidos antisentido descritos en la presente pueden hibridar con un exón, un intrón, la UTR 5', la UTR 3' una región de repetición, una expansión de repetición hexanucleotídica, una unión de corte y empalme, un exón: unión de corte y empalme de exón, un silenciador de empalme exónico (ESS), un potenciador de empalme exónico (ESE), exón 1a, exón 1b, exón 1c, exón 1d, exón 1e, exón 2, exón 3, exón 4, exón 5, exón 6, exón 7, exón 8, exón 9, exón 10, exón 11, intrón 1, intrón 2, intrón 3, intrón 4, intrón 5, intrón 6, intrón 7, intrón 8, intrón 9 o intrón 10. Por ejemplo, los oligonucleótidos antisentido pueden dirigirse a cualquiera de los exones caracterizados a continuación en las Tablas 1-5 para las varias variantes de C9ORF72 descritas a continuación. Los oligonucleótidos antisentido descritos en la presente también pueden dirigirse a variantes no caracterizadas a continuación y tales variantes se caracterizan en GENBANK. Además, los oligonucleótidos antisentido descritos en la presente también pueden dirigirse a elementos distintos de los exones y tales elementos se caracterizan en GENBANK.

**Tabla 1**

| Segmentos Funcionales para NM_001256054.1 (SEQ ID NO: 1) | | | | |
|---|---|---|---|---|
| **Número de Exón** | **Sitio de inicio de ARNm** | **Sitio de parada de ARNm** | **Sitio de inicio en referencia a la SEQ ID NO: 2** | **Sitio de parada en referencia a la SEQ ID NO: 2** |
| exón 1C | 1 | 158 | 1137 | 1294 |
| exón 2 | 159 | 646 | 7839 | 8326 |
| exón 3 | 647 | 706 | 9413 | 9472 |
| exón 4 | 707 | 802 | 12527 | 12622 |
| exón 5 | 803 | 867 | 13354 | 13418 |
| exón 6 | 868 | 940 | 14704 | 14776 |

(continuación)

| Segmentos Funcionales para NM_001256054.1 (SEQ ID NO: 1) | | | | |
|---|---|---|---|---|
| **Número de Exón** | **Sitio de inicio de ARNm** | **Sitio de parada de ARNm** | **Sitio de inicio en referencia a la SEQ ID NO: 2** | **Sitio de parada en referencia a la SEQ ID NO: 2** |
| exón 7 | 941 | 1057 | 16396 | 16512 |
| exón 8 | 1058 | 1293 | 18207 | 18442 |
| exón 9 | 1294 | 1351 | 24296 | 24353 |
| exón 10 | 1352 | 1461 | 26337 | 26446 |
| exón 11 | 1462 | 3339 | 26581 | 28458 |

**Tabla 2**

| Segmentos Funcionales para NM_018325.3 (SEQ ID NO: 4) | | | | |
|---|---|---|---|---|
| **Número de Exón** | **Sitio de inicio de ARNm** | **Sitio de parada de ARNm** | **Sitio de inicio en referencia a la SEQ ID NO: 2** | **Sitio de parada en referencia a la SEQ ID NO: 2** |
| exón 1B | 1 | 63 | 1510 | 1572 |
| exón 2 | 64 | 551 | 7839 | 8326 |
| exón 3 | 552 | 611 | 9413 | 9472 |
| exón 4 | 612 | 707 | 12527 | 12622 |
| exón 5 | 708 | 772 | 13354 | 13418 |
| exón 6 | 773 | 845 | 14704 | 14776 |
| exón 7 | 846 | 962 | 16396 | 16512 |
| exón 8 | 963 | 1198 | 18207 | 18442 |
| exón 9 | 1199 | 1256 | 24296 | 24353 |
| exón 10 | 1257 | 1366 | 26337 | 26446 |
| exón 11 | 1367 | 3244 | 26581 | 28458 |

**Tabla 3**

| Segmentos Funcionales para NM_145005.5 (SEQ ID NO: 6) | | | | |
|---|---|---|---|---|
| **Número de Exón** | **Sitio de inicio de ARNm** | **Sitio de parada de ARNm** | **Sitio de inicio en referencia a la SEQ ID NO: 2** | **Sitio de parada en referencia a la SEQ ID NO: 2** |
| exón 1A | 1 | 80 | 1137 | 1216 |
| exón 2 | 81 | 568 | 7839 | 8326 |
| exón 3 | 569 | 628 | 9413 | 9472 |
| exón 4 | 629 | 724 | 12527 | 12622 |
| exón 5B (exón 5 en intrón 5) | 725 | 1871 | 13354 | 14500 |

**Tabla 4**

| Segmentos Funcionales para DB079375.1 (SEQ ID NO: 7) | | | | |
|---|---|---|---|---|
| **Número de Exón** | **Sitio de inicio de ARNm** | **Sitio de parada de ARNm** | **Sitio de inicio en referencia a la SEQ ID NO: 2** | **Sitio de parada en referencia a la SEQ ID NO: 2** |
| exón 1E | 1 | 35 | 1135 | 1169 |
| exón 2 | 36 | 524 | 7839 | 8326 |
| exón 3 (EST finaliza antes del final del exón completo) | 525 | 562 | 9413 | 9450 |

**Tabla 5**

| Segmentos Funcionales para BU194591.1 (SEQ ID NO: 8) | | | | |
|---|---|---|---|---|
| **Número de Exón** | **Sitio de inicio de ARNm** | **Sitio de parada de ARNm** | **Sitio de inicio en referencia a la SEQ ID NO: 2** | **Sitio de parada en referencia a la SEQ ID NO: 2** |
| exón 1D | 1 | 36 | 1241 | 1279 |
| exón 2 | 37 | 524 | 7839 | 8326 |
| exón 3 | 525 | 584 | 9413 | 9472 |
| exón 4 | 585 | 680 | 12527 | 12622 |
| exón 5B (exón 5 en intrón 5) | 681 | 798 | 13354 | 13465 |

*Ciertas indicaciones*

**[0277]** En ciertos casos, en la presente se divulgan métodos para tratar a un individuo que comprenden administrar una o más composiciones farmacéuticas descritas en la presente. En ciertos casos, el individuo tiene una enfermedad neurodegenerativa. En ciertos casos, el individuo puede estar en riesgo de desarrollar una enfermedad neurodegenerativa incluyendo, pero no limitada a, ELA o FTD. Puede haberse identificado que el individuo tiene una enfermedad asociada a C9ORF72. Puede haberse identificado que el individuo tiene una enfermedad asociada a la expansión de repetición de hexanucleótidos de C9ORF72. En la presente se divulgan métodos para reducir profilácticamente la expresión de C9ORF72 en un individuo. Ciertos casos incluyen tratar a un individuo con necesidad de ello mediante la administración a un individuo de una cantidad terapéuticamente eficaz de un compuesto antisentido dirigido a un ácido nucleico de C9ORF72.

**[0278]** En un caso, la administración de una cantidad terapéuticamente eficaz de un compuesto antisentido dirigido a un ácido nucleico de C9ORF72 se acompaña de la monitorización de los niveles de C9ORF72 en un individuo, para determinar la respuesta de un individuo a la administración del compuesto antisentido. La respuesta de un individuo a la administración del compuesto antisentido puede usarse por un médico para determinar la cantidad y la duración de la intervención terapéutica.

**[0279]** En ciertos casos, la administración de un compuesto antisentido dirigido a un ácido nucleico de C9ORF72 da como resultado la reducción de la expresión de C9ORF72 en por lo menos un 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 o 99%, o un intervalo definido por dos cualquiera de estos valores. En ciertos casos, la administración de un compuesto antisentido dirigido a un ácido nucleico de C9ORF72 da como resultado una función motora y respiración mejoradas en un animal. En ciertos casos, la administración de un compuesto antisentido de C9ORF72 mejora la función motora y la respiración en por lo menos un 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 o 99%, o un intervalo definido por dos cualquiera de estos valores.

**[0280]** En ciertos casos, las composiciones farmacéuticas que comprenden un compuesto antisentido dirigido a C9ORF72 se usan para la preparación de un medicamento para tratar a un paciente que padece o es susceptible a una enfermedad neurodegenerativa, incluyendo ELA y FTD.

*Ciertas terapias de combinación*

**[0281]** En ciertos casos, una o más composiciones farmacéuticas descritas en la presente se administran

conjuntamente con uno o más de otros agentes farmacéuticos. En ciertos casos, uno o más de estos otros agentes farmacéuticos están diseñados para tratar la misma enfermedad, trastorno o afección que la una o más composiciones farmacéuticas descritas en la presente. En ciertos casos, uno o más de estos otros agentes farmacéuticos están diseñados para tratar una enfermedad, trastorno o afección diferente a la una o más composiciones farmacéuticas descritas en la presente. En ciertos casos, el uno o más de estos otros agentes farmacéuticos están diseñados para tratar un efecto secundario no deseado de una o más composiciones farmacéuticas descritas en la presente. En ciertos casos, una o más composiciones farmacéuticas descritas en la presente se administran conjuntamente con otro agente farmacéutico para tratar un efecto no deseado de ese otro agente farmacéutico. En ciertos casos, una o más composiciones farmacéuticas descritas en la presente se administran conjuntamente con otro agente farmacéutico para producir un efecto combinatorio. En ciertos casos, una o más composiciones farmacéuticas descritas en la presente se administran conjuntamente con otro agente farmacéutico para producir un efecto sinérgico.

**[0282]** En ciertos casos, una o más composiciones farmacéuticas descritas en la presente y uno o más agentes farmacéuticos de otros se administran al mismo tiempo. En ciertos casos, una o más composiciones farmacéuticas descritas en la presente y uno o más de otros agentes farmacéuticos se administran en momentos diferentes. En ciertos casos, una o más composiciones farmacéuticas descritas en la presente y uno o más de otros agentes farmacéuticos se preparan juntos en una única formulación. En ciertos casos, una o más composiciones farmacéuticas descritas en la presente y uno o más de otros agentes farmacéuticos se preparan por separado.

**[0283]** En ciertos casos, los agentes farmacéuticos que pueden administrarse conjuntamente con una composición farmacéutica descrita en la presente incluyen Riluzol (Rilutek), Lioresal (Lioresal) y Dexpramipexol.

**[0284]** En ciertos casos, los agentes farmacéuticos que pueden administrarse conjuntamente con un inhibidor específico de C9ORF72 descrito en la presente incluyen, pero no están limitados a, un inhibidor adicional de C9ORF72. En ciertos casos, el agente farmacéutico administrado conjuntamente se administra antes de la administración de una composición farmacéutica descrita en la presente. En ciertos casos, el agente farmacéutico administrado conjuntamente se administra después de la administración de una composición farmacéutica descrita en la presente. En ciertos casos, el agente farmacéutico administrado conjuntamente se administra al mismo tiempo que una composición farmacéutica descrita en la presente. En ciertos casos, la dosis de un agente farmacéutico administrado conjuntamente es la misma que la dosis que se administraría si el agente farmacéutico administrado conjuntamente se administrara solo. En ciertos casos, la dosis de un agente farmacéutico administrado conjuntamente es menor que la dosis que se administraría si el agente farmacéutico administrado conjuntamente se administrara solo. En ciertos casos, la dosis de un agente farmacéutico administrado conjuntamente es mayor que la dosis que se administraría si el agente farmacéutico administrado conjuntamente se administrara solo.

**[0285]** En ciertos casos, la administración conjunta de un segundo compuesto aumenta el efecto de un primer compuesto, de tal manera que la administración conjunta de los compuestos da como resultado un efecto que es mayor que el efecto de administrar el primer compuesto solo. En otros casos, la administración conjunta da como resultado efectos que son aditivos a los efectos de los compuestos cuando se administran solos. En ciertos casos, la administración conjunta produce efectos que son supra-aditivos de los efectos de los compuestos cuando se administran solos. En ciertos casos, el primer compuesto es un compuesto antisentido. En ciertos casos, el segundo compuesto es un compuesto antisentido.

## EJEMPLOS

### *Divulgación no limitativa*

**[0286]** Aunque se han descrito ciertos compuestos, composiciones y métodos descritos en la presente con especificidad de acuerdo con ciertas realizaciones, los siguientes ejemplos sirven solo para ilustrar los compuestos descritos en la presente y no se pretende que limiten los mismos.

### Ejemplo 1: inhibición antisentido de C9ORF72 humano en células HepG2

**[0287]** Se diseñaron oligonucleótidos antisentido dirigidos a un ácido nucleico de C9ORF72 y se probaron para sus efectos sobre el ARNm de C9ORF72 *in vitro.* Los oligonucleótidos antisentido se probaron en una serie de experimentos que tenían condiciones de cultivo similares. Los resultados para cada experimento se presentan en tablas separadas mostradas a continuación. Las células HepG2 cultivadas a una densidad de 20.000 células por pocillo se transfectaron usando electroporación con oligonucleótido antisentido 7.000 nM. Después de un período de tratamiento de aproximadamente 24 horas, se aisló el ARN de las células y se midieron los niveles de ARNm de C9ORF72 por PCR cuantitativa en tiempo real. Se usó el conjunto de cebador sonda humano RTS3750 (secuencia directa TGTGACAGTTGGAATGCAGTGA, designada en la presente como SEQ ID NO: 15; secuencia inversa GCCACTTAAAGCAATCTCTGTCTTG, designada en la presente como SEQ ID NO: 16; secuencia de sonda TCGACTCTTTGCCCACCGCCA, designada en la presente como SEQ ID NO: 17) para medir los niveles de ARNm. Los niveles de ARNm de C9ORF72 se ajustaron de acuerdo con el contenido de ARN total, medido por

RIBOGREEN®. Los resultados se presentan como porcentaje de inhibición de C9ORF72, con respecto a las células de control no tratadas.

**[0288]** Los oligonucleótidos antisentido en las Tablas 6-8 se diseñaron como gapmers MOE 5-10-5. Los gapmers tienen una longitud de 20 nucleósidos, en donde el segmento de hueco central comprende diez 2'-desoxinucleósidos y está flanqueado por segmentos de ala tanto en el extremo 5' como en el extremo 3' que comprenden cinco nucleósidos cada uno. Cada nucleósido en el segmento de ala 5' y cada nucleósido en el segmento de ala 3' tiene una modificación MOE. Los enlaces internucleosídicos en cada gapmer son enlaces de fosforotioato. Todos los residuos de citosina en cada gapmer son 5-metilcitosinas. El "sitio de inicio" indica el nucleósido más 5' al que está dirigido el oligonucleótido antisentido en la secuencia del gen humano. El "sitio de parada" indica el nucleósido más 3' al que está dirigido el oligonucleótido antisentido en la secuencia del gen humano. Cada oligonucleótido antisentido enumerado en las Tablas 6-7 está dirigido o a la secuencia de ARNm de C9ORF72 humano, designada en la presente como SEQ ID NO: 1 (Nº de registro GENBANK NM_001256054.1) o la secuencia genómica de C9ORF72 humano, designada en la presente como SEQ ID NO: 2 (el complemento del Nº de registro GENBANK NT_008413.18 truncado de los nucleósidos 27535000 a 27565000), o ambos. 'n/a' indica que el oligonucleótido antisentido no se dirigió a esa secuencia génica particular. Los oligonucleótidos antisentido de la Tabla 8 están dirigidos o a la SEQ ID NO: 3 (Nº de registro de GENBANK BQ068108.1) o a la SEQ ID NO: 4 (Nº de registro de GENBANK NM_018325.3).

**Tabla 6**

| Nº ISIS | % de inhibición |
|---|---|
| 576816 | 78 |
| 576860 | 79 |

**Tabla 7**

| Nº ISIS | % de inhibición |
|---|---|
| 576974 | 56 |

**Tabla 8**

| Nº ISIS | % de inhibición |
|---|---|
| 576816 | 90 |
| 577061 | 28 |
| 577065 | 0 |
| 577083 | 70 |

**Ejemplo 2: inhibición antisentido dependiente de la dosis de C9ORF72 humano en células HepG2**

**[0289]** Se seleccionaron los oligonucleótidos antisentido del estudio descrito anteriormente que mostraban una inhibición *in vitro* significativa del ARNm de C9ORF72 y se probaron a varias dosis en células HepG2. Los oligonucleótidos antisentido se probaron en una serie de experimentos que tenían condiciones de cultivo similares. Los resultados para cada experimento se presentan en tablas separadas que se muestran a continuación. Las células se colocaron en placas a una densidad de 20.000 células por pocillo y se transfectaron usando electroporación con concentraciones de oligonucleótido antisentido de 82,3 nM, 246,9 nM, 740,7 nM, 2.222,2 nM, 6.666,7 nM o 20.000 nM. Después de un período de tratamiento de aproximadamente 16 horas, se aisló el ARN de las células y se midieron los niveles del ARNm de C9ORF72 por PCR cuantitativa en tiempo real. Se usó el conjunto de cebador sonda de C9ORF72 humano RTS3750 para medir los niveles de ARNm. Los niveles de ARNm de C9ORF72 se ajustaron de acuerdo con el contenido de ARN total, medido por RIBOGREEN®. Los resultados se presentan como porcentaje de inhibición de C9ORF72,

**[0290]** En la Tablas 9-10 también se presenta la concentración inhibidora máxima media ($IC_{50}$) de cada oligonucleótido. Como se ilustra, los niveles de ARNm de C9ORF72 se redujeron de una manera dependiente de la dosis en las células tratadas con oligonucleótidos antisentido.

**Tabla 9**

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

**Tabla 10**

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

**Ejemplo 3: inhibición antisentido dependiente de la dosis de C9ORF72 humano en células HepG2**

**[0291]** Se seleccionaron oligonucleótidos antisentido del estudio descrito anteriormente que mostraban una inhibición *in vitro* significativa del ARNm de C9ORF72 y se probaron a varias dosis en células HepG2. Los oligonucleótidos antisentido se probaron en una serie de experimentos que tenían condiciones de cultivo similares. Los resultados para cada experimento se presentan en tablas separadas mostradas a continuación. Las células se colocaron en placas a una densidad de 20.000 células por pocillo y se transfectaron usando electroporación con concentraciones de oligonucleótido antisentido de 246,9 nM, 740,7 nM, 2.222,2 nM, 6.666,7 nM o 20.000 nM. Después de un período de tratamiento de aproximadamente 16 horas, se aisló el ARN de las células y se midieron los niveles de ARNm total de C9ORF72, así como los niveles de ARNm de la transcripción del exón 1, mediante PCR cuantitativa en tiempo real. Se usó el conjunto de cebador sonda de C9ORF72 humano RTS3750 para medir los niveles totales del ARNm de C9ORF72. Se uso el conjunto cebador sonda RTS3905 (secuencia directa GGGTCTAGCAAGAGCAGGTG, designada en la presente como SEQ ID NO: 18; secuencia inversa GTCTTGGCAACAGCTGGAGAT, designada en la presente como SEQ ID NO: 19; secuencia de sonda TGATGTCGACTCTTTGCCCACCGC, designada en la presente como SEQ ID NO: 20) para medir la transcripción del mensaje del exón 1. Los niveles de ARNm de C9ORF72 se ajustaron de acuerdo con el contenido de ARN total, medido por RIBOGREEN®. Los resultados se presentan como porcentaje de inhibición de C9ORF72, con respecto a las células de control no tratadas.

**[0292]** En las Tablas 11 y 12 también se presenta la concentración inhibidora máxima media ($IC_{50}$) de cada oligonucleótido. Como se ilustra, los niveles del ARNm de C9ORF72 se redujeron de una manera dependiente de la dosis en las células tratadas con oligonucleótidos antisentido. 'n.d.' indica que no hay datos para esa dosis particular.

**Tabla 11**

| % de inhibición de niveles de ARNm de C9ORF72 totales | | | | | | |
|---|---|---|---|---|---|---|
| Nº ISIS | 246.9 nM | 740.7 nM | 2222.2 nM | 6666.7 nM | 20000.0 nM | $IC_{50}$ (μM) |
| 576816 | 29 | 53 | 84 | 90 | 92 | 0.60 |
| 576860 | 25 | 53 | 72 | 86 | 85 | 0.80 |
| 576974 | 36 | 49 | 64 | 65 | 68 | 0.95 |
| 577061 | 0 | 3 | 0 | 4 | 0 | >20.00 |
| 577065 | 7 | 0 | 1 | 6 | 0 | >20.00 |
| 577083 | 0 | 19 | 55 | 71 | 75 | 3.35 |

**Tabla 12**

| % de inhibición de niveles de ARNm de exón 1 de C9ORF72 | | | | | | |
|---|---|---|---|---|---|---|
| **Nº ISIS** | **246.9 nM** | **740.7 nM** | **2222.2 nM** | **6666.7 nM** | **20000.0 nM** | **$IC_{50}$ (μM)** |
| 576816 | 45 | 78 | 93 | n.d. | n.d. | 0.26 |
| 576860 | 43 | 36 | 71 | 95 | 91 | 0.66 |
| 577061 | 8 | 36 | 70 | 67 | 76 | 2.03 |
| 577065 | 20 | 55 | 67 | 82 | 62 | 1.06 |
| 577083 | 30 | 43 | 85 | 88 | 92 | 0.71 |

**Ejemplo 4: Efecto de la inhibición antisentido de C9ORF72 humano en fibroblastos primarios de pacientes con ELA**

**[0293]** Los oligonucleótidos antisentido, de los estudios descritos anteriormente y dirigidos a varias regiones del gen de C9ORF72 (SEQ ID NO: 2), se probaron en fibroblastos primarios de pacientes con ELA. La localización objetivo de cada oligonucleótido antisentido se especifica en la Tabla 13 a continuación. Como se observa en la Tabla, varios de los oligonucleótidos antisentido se dirigen al intrón 1 y, por lo tanto, a la secuencia pre-ARNm de C9ORF72. El intrón 1 contiene una repetición de hexanucleótidos (GGGGCC), que se expande por lo menos a 30 repeticiones en una célula patógena (Renton et al., Neuron, 2011, 72, 257-268).

**Tabla 13**

| Localización de oligonucleotidos antisentido dirgidos a SEQ ID NO: 1 | |
|---|---|
| **Nº ISIS** | **Localización** |
| 577061 | Intrón 1 |
| 577065 | Intrón 1 |
| 577083 | Intrón 1 |
| 576816 | Exón 2 |
| 576860 | Intrón 5 |
| 576974 | Exón 11 |

**[0294]** El efecto del tratamiento del oligonucleótido antisentido dirigido a C9ORF72 se evaluó en fibroblastos primarios derivados de pacientes con ELA. Los fibroblastos primarios derivados de biopsia de piel de un paciente humano con ELA (laboratorio del Dr. John Ravits) se colocaron en placas en medio DMEM suplementado con suero bovino fetal al 20%. Las células se trataron con los oligonucleótidos antisentido dirigidos a C9ORF72 a una concentración de 25 nM durante 4 h. Se trató un conjunto de células con un oligonucleótido de control (control ASO), sin objetivo humano conocido, y sirvió como control negativo. Después de 24 horas de transfección, las células se recogieron en reactivo TriZol para cuantificación del ARN, o se fijaron con paraformaldehído al 4% para hibridación *in* situ.

*Análisis de ARN*

**[0295]** Se recogieron las células y se analizaron los niveles de ciertas variantes de ARNm de C9ORF72, así como del ARNm de C9ORF72 total. Las variantes de ARNm medidas fueron SEQ ID NO: 1 (variante 3) y SEQ ID NO: 6 (variante 2). Como se presenta en las Tablas 1 y 3 enumeradas anteriormente, las SEQ ID NO 1 y 6 codifican porciones de la secuencia de C9ORF72 que incluyen el intrón 1. Debido a que el intrón 1 incluye la repetición de hexanucleótidos, la reducción de los niveles de ARNm de estas variantes es terapéutica en las células patógenas (células que contienen expansiones de repetición de hexanucleótidos de más de 30 repeticiones). Ciertas otras variantes de ARNm no incluyen la repetición de hexanucleótidos; por ejemplo, la SEQ ID NO: 4. Por lo tanto, es posible inhibir preferencialmente ciertas variantes de C9ORF72 diseñando oligonucleótidos antisentido para dirigirse a ciertas regiones del ácido nucleico de C9ORF72 objetivo, como, por ejemplo, el intrón 1.

**[0296]** Se midieron los niveles de la SEQ ID NO: 1 y 6 (variantes 3 y 2, respectivamente) en combinación mediante el conjunto de cebador sonda RTS3905, que se dirige al exón 1. El nivel de ARNm de C9orf72 total se midió

mediante el conjunto de cebador sonda RTS3750, que se dirige al exón 2. Los datos se presentan en la Tabla 14. En comparación con el control PBS, Isis 577061 e Isis 577065 (cada uno dirigido a intrón 1) redujeron la expresión de la variante de ARN 2/3 en más del 90% y redujeron el ARN total en un 18% y un 8%, respectivamente. Isis 577083 (dirigido al intrón 1), Isis 576816 (dirigido al exón 2), Isis 576860 (dirigido al exón 5) e Isis 576974 (dirigido al exón 11) lograron entre un 75% y un 96% de inhibición de la expresión de la variante de ARN 2/3 y la expresión del ARN total. El control ASO no inhibió significativamente los niveles de ARN, como se esperaba.

**Tabla 14**

| Porcentaje de inhibición, en comparación con el control de PBS control de niveles de ARN de C90RF72 en fibroblastos primarios | | | |
|---|---|---|---|
| **ID de oligo** | **Localización** | **Variante 2/3** | **ARN total** |
| Control ASO | n/a | 13 | 3 |
| 577061 | Intrón 1 | 95 | 18 |
| 577065 | Intrón 1 | 96 | 8 |
| 577083 | Intrón 1 | 89 | 81 |
| 576816 | Exón 2 | 96 | 95 |
| 576860 | Intrón 5 | 81 | 75 |
| 576974 | Exón 11 | 85 | 77 |

*Análisis de los focos de ARN de C9ORF72*

**[0297]** En años recientes, la acumulación intracelular de repeticiones de nucleótidos expandida como focos de ARN en el núcleo de las células afectadas ha aparecido como un mecanismo de enfermedad importante en pacientes con demencia frontotemporal y ELA (DeJesus-Hernandez, M. et al., Neuron. 2011. 72: 245 -56). Los focos nucleares de C9ORF72 se detectaron usando FISH con (5'TYE563-CCCCGGCCCCGGCCCC) (SEQ ID NO: 23) sonda de LNA (Exiqon). Los datos para el porcentaje de células con focos se presentan en la Tabla 15. Los resultados indican que los oligonucleótidos antisentido dirigidos al ARN de C9ORF72 disminuyeron el porcentaje de células que contenían focos de ARN de C9ORF72.

**[0298]** Se detectan focos de ARN individuales y múltiples en fibroblastos de pacientes. Se contaron los números de células con focos individuales o múltiples de fibroblastos que fueron tratados con ISIS 576816 o el control ASO. Los datos se presentan en la Tabla 16 e indican que el oligonucleótido antisentido dirigido a C9ORF72 redujo el número de focos por célula en comparación con el control ASO.

**Tabla 15**

| Porcentaje de fibroblastos primarios con focos | |
|---|---|
| | **% de células** |
| PBS | 37 |
| control ASO | 38 |
| 577061 | 9 |
| 577065 | 6 |
| 577083 | 6 |
| 576816 | 9 |
| 576860 | 12 |
| 576974 | 12 |

**Tabla 16**

| Número de células con focos individuales o múltiples | | |
|---|---|---|
| **Nº de focos** | **tratado con 576816** | **Control ASO** |
| 0 | 82 | 66 |
| 1 | 14 | 10 |
| 2 | 2 | 9 |
| 3 | 2 | 5 |
| 4 | 0 | 5 |
| 5 | 0 | 2 |
| 6 | 0 | 0 |
| 7 | 0 | 0 |
| 8 | 0 | 2 |
| 9 | 0 | 0 |
| 10+ | 0 | 1 |

**Ejemplo 5: Efecto de la inhibición de ARNip de C9ORF72 humano en fibroblastos primarios de pacientes con ELA**

**[0299]** Se obtuvieron ARNip dirigido a C9ORF72 y ARNip de control dirigido a GAPDH se obtuvieron de Thermo Scientific Dharmacon (On-Target más Smartpool). Se evaluó el efecto del tratamiento de ARNip en fibroblastos de pacientes con ELA.

**[0300]** Los fibroblastos primarios derivados de una biopsia de piel de pacientes humanos con ELA se colocaron en placas en medio DMEM suplementado con suero bovino fetal al 20%. Las células se trataron con el ARNip dirigido a C9ORF72 a una concentración de 40 nM durante 4 horas. También se trató un conjunto de células con ARNip dirigido a GAPDH a una concentración de 40 nM durante 4 horas y sirvió como control. Después de la transfección durante 24 horas, las células se recogieron en reactivo TriZol para cuantificación de ARN, o se fijaron con paraformaldehído al 4% para hibridación *in situ*.

**[0301]** Los niveles de ARN de C9ORF72 totales se midieron usando un conjunto de cebador (Directo: TCCTGTAATGGAACTGCTTTCA, designado en la presente como SEQ ID NO: 21; Inverso: GGTATCTGCTTCATCCAGCTTT, designado en la presente como SEQ ID NO: 22). El tratamiento con ARNip resultó en una inhibición del 73% de los niveles totales de ARN de C9ORF72 en comparación con el control.

**[0302]** Los focos nucleares de C9ORF72 se detectaron usando FISH con (5'TYE563-CCCCGGCCCCGGCCCC) (SEQ ID NO: 23) sonda de LNA (Exiqon). Los datos se presentan en la Tabla 17. Los resultados indican que el tratamiento con ARNip contra C9ORF72 no reduce los focos nucleares en los fibroblastos en comparación con el control.

**Tabla 17**

| Porcentaje de células con focos | |
|---|---|
| | **% de células** |
| ARNip-GAPDH | 28 |
| ARNip-C9ORF72 | 31 |

**Ejemplo 6: Inhibición y tolerabilidad in vivo de roedores con tratamiento de oligonucleótidos antisentido de C9ORF72**

**[0303]** Para evaluar la tolerabilidad de la inhibición de la expresión de C9orf72 *in vivo,* se diseñaron y evaluaron oligonucleótidos antisentido dirigidos a un ácido nucleico de C9orf72 murino en modelos de ratón y rata.

**[0304]** El ISIS 571883 se diseñó como un gapmer 5-10-5 MOE, de 20 nucleósidos de longitud, en el que el segmento de hueco central comprende diez 2'-desoxinucleósidos y está flanqueado por segmentos de ala en tanto el extremo 5' como en el extremo 3' que comprenden cinco nucleósidos cada uno. Cada nucleósido en el segmento de ala 5' y cada nucleósido en el segmento de ala 3' tiene una modificación MOE. Los enlaces internucleosídicos son enlaces de fosforotioato. Todos los residuos de citosina en todo el gapmer son 5-metilcitosinas. El ISIS 571883 tiene un sitio de inicio objetivo del nucleósido 33704 en la secuencia genómica murina de C9ORF72, designada en la presente como SEQ ID NO: 11 (el complemento del Nº de registro de GENBANK NT_166289.1 truncado de los nucleósidos 3587000 a 3625000).

**[0305]** El ISIS 603538 se diseñó como un gapmer 5-10-5 MOE, de 20 nucleósidos de longitud, en el que el segmento de hueco central comprende diez 2'-desoxinucleósidos y está flanqueado por segmentos de ala en tanto el extremo 5' como en el extremo 3' que comprenden cinco nucleósidos cada uno. Cada nucleósido en el segmento de ala 5' y cada nucleósido en el segmento de ala 3' tiene una modificación MOE. Los enlaces internucleosídicos son o enlaces de fosforotioato o enlaces de éster de fosfato (Gs Ao Co Co Gs Cs Ts Ts Gs As Gs Ts Ts Ts Gs Co Co Ao Cs A; en donde 's' indica un enlace internucleósido de fosforotioato, 'o' denota un enlace de éster de fosfato; y A, G, C, T denotan los nucleósidos relevantes). Todos los residuos de citosina en todo el gapmer son 5-metilcitosinas. EL ISIS 603538 tiene un sitio de inicio objetivo del nucleósido 2872 en la secuencia de ARNm de C9ORF72 de rata,

Experimento en ratones 1

**[0306]** Se inyectaron a grupos de 4 ratones C57BL/6 50 µg, 100 µg, 300 µg, 500 µg o 700 µg de ISIS 571883 administrado mediante una inyección intracerebroventricular en bolo. Un grupo de control de cuatro ratones C57/BL6 se trató de manera similar con PBS. Los animales se anestesiaron con isofluorano al 3% y se colocaron en un marco estereotáctico. Después de esterilizar el sitio quirúrgico, se inyectaron a cada ratón -0,2 mm anteroposterior desde el bregma na d 3 mm dorsoventral al bregma las dosis mencionadas anteriormente de ISIS 571883 usando una jeringuilla Hamilton. La incisión se cerró con suturas. Se permitió que los ratones se recuperaran durante 14 días, después de lo cual los animales se sacrificaron de acuerdo con un protocolo humano aprobado por el Comité Institucional de Cuidado y Uso de Animales. Se recogieron el cerebro y el tejido de la médula espinal y se congelaron al instante en nitrógeno líquido. Antes de congelar, el tejido cerebral se cortó transversalmente en cinco secciones usando una matriz de cerebro de ratón.

*Análisis de ARN*

**[0307]** El ARN se extrajo de una sección del cerebro de 2-3 mm posterior al sitio de inyección, de la corteza frontal del cerebro y de la sección lumbar del tejido de la médula espinal para el análisis de la expresión del ARNm de C9orf72. La expresión del ARNm de C9orf72 se midió por RT-PCR. Los datos se presentan en la Tabla 18. Los resultados indican que el tratamiento con dosis crecientes de ISIS 571883 dio como resultado una inhibición dependiente de la dosis de la expresión del ARNm de C9orf72.

**[0308]** También se evaluó la inducción del marcador microglial AIF-1 como medida de la toxicidad del SNC. Los datos se presentan en la Tabla 19. Los resultados indican que el tratamiento con dosis crecientes de ISIS 571883 no dio como resultado aumentos significativos en la expresión de ARNm de AIF-1. Por lo tanto, la inyección de ISIS 571883 se consideró tolerable en este modelo.

**Tabla 18**

| Porcentaje de inhibición de expresión de ARNm de C9orf72 en comparación con el control de PBS | | | |
|---|---|---|---|
| Dosis (µg) | Cerebro posterior | Corteza | Médula espinal |
| 50 | 22 | 8 | 46 |
| 100 | 22 | 12 | 47 |
| 300 | 55 | 47 | 67 |
| 500 | 61 | 56 | 78 |
| 700 | 65 | 65 | 79 |

**Tabla 19**

| Porcentaje de expresión de expresión de ARNm de AIF-1 en comparación con el control de PBS | | |
|---|---|---|
| Dosis (µg) | Cerebro posterior | Médula espinal |
| 50 | 102 | 89 |
| 100 | 105 | 111 |
| 300 | 107 | 98 |
| 500 | 131 | 124 |
| 700 | 122 | 116 |

Experimento en ratones 2

**[0309]** Se inyectaron a grupos de 4 ratones C57BL/6 cada uno 500 µg de ISIS 571883 administrados mediante una inyección intracerebroventricular en bolo en un procedimiento similar al descrito anteriormente. Un grupo de control de cuatro ratones C57/BL6 fueron tratados de manera similar con PBS. Los ratones se probaron en puntos temporales regulares después de la administración de ICV.

*Análisis de comportamiento.*

**[0310]** Se emplearon dos ensayos estándar para evaluar el comportamiento motor; el ensayo rotarod y el ensayo de fuerza de agarre. En el caso de los ensayos de rotarod, se midió el tiempo de latencia para caer. Los datos para los ensayos se presentan en las Tablas 20 y 21. Los resultados indican que no hubo cambios significativos en el comportamiento motor de los ratones como resultado de la inhibición antisentido de ISIS 571883 o debido a la inyección de ICV. Por lo tanto, la inhibición antisentido de C9orf72 se consideró tolerable en este modelo.

**Tabla 20**

| Latencia a la caída (seg) en el ensayo rotarod | | |
|---|---|---|
| Semanas después de inyección | PBS | ISIS 571883 |
| 0 | 66 | 66 |
| 4 | 91 | 70 |
| 8 | 94 | 84 |

**Tabla 21**

| Fuerza de agarre de la extremidad trasera media (g) en ensayo de fuerza de agarre | | |
|---|---|---|
| Semanas después de inyección | PBS | ISIS 571883 |
| 0 | 57 | 63 |
| 1 | 65 | 51 |
| 2 | 51 | 52 |
| 3 | 51 | 51 |
| 4 | 59 | 72 |
| 5 | 60 | 64 |
| 6 | 61 | 72 |
| 7 | 67 | 68 |
| 8 | 66 | 70 |
| 9 | 63 | 61 |
| 10 | 48 | 46 |

Experimento en ratas

**[0311]** A grupos de 4 ratas Sprague-Dawley cada uno se les inyectaron 700 µg, 1.000 µg o 3.000 µg de ISIS 603538 administrados mediante una inyección de bolo intratecal. Un grupo de control de cuatro ratas Sprague-Dawley se trató de manera similar con PBS. Los animales se anestesiaron con isofluorano al 3% y se colocaron en una estructura estereotáctica. Después de esterilizar el sitio quirúrgico, se inyectaron a cada rata 30 µl de solución de ASO administrada a través de un catéter intratecal de 8 cm a 2 cm en el canal espinal con un lavado de 50 µl. Se permitió que las ratas se recuperaran durante 4 semanas, después de lo cual los animales se sacrificaron de acuerdo con un protocolo humano aprobado por el Comité Institucional de Cuidado y Uso de Animales.

*Análisis de ARN*

**[0312]** El ARN se extrajo de una sección del cerebro de 2-3 mm posterior al sitio de inyección, de la corteza frontal del cerebro y de las secciones cervicales y lumbares del tejido de la médula espinal para el análisis de la expresión del ARNm de C9orf72. La expresión del ARNm de C9orf72 se midió por RT-PCR. Los datos se presentan en la Tabla 22. Los resultados indican que el tratamiento con dosis crecientes de ISIS 603538 dio como resultado una inhibición dependiente de la dosis de la expresión del ARNm de C9orf72.

**[0313]** También se evaluó la inducción del marcador microglial AIF-1 como medida de la toxicidad en el SNC. Los datos se presentan en la Tabla 23. Los resultados indican que el tratamiento con dosis crecientes de ISIS 603538 no produjo aumentos significativos en la expresión del ARNm de AIF-1. Por lo tanto, la inyección de ISIS 603538 se consideró tolerable en este modelo.

**Tabla 22**

| Porcentaje de inhibición de expresión del ARNm de C9orf72 en comparación con el control de PBS | | | | |
|---|---|---|---|---|
| Dosis (µg) | Cerebro (sección de 1 mm) | Corteza | Médula espinal (lumbar) | Médula espinal (cervical) |
| 700 | 21 | 4 | 86 | 74 |
| 1000 | 53 | 49 | 88 | 82 |
| 3000 | 64 | 62 | 88 | 80 |

**Tabla 23**

| Porcentaje de expresión de expresión del ARNm de AIF-1 mRNA en comparación con el control de PBS | | | | |
|---|---|---|---|---|
| Dosis (µg) | Cerebro (sección de1 mm) | Corteza | Médula espinal (lumbar) | Médula espinal (cervical) |
| 700 | 97 | 119 | 98 | 89 |
| 1000 | 105 | 113 | 122 | 96 |
| 3000 | 109 | 141 | 156 | 115 |

*Análisis de peso corporal.*

**[0314]** El peso corporal de las ratas se midió a intervalos temporales regulares. Los datos se presentan en la Tabla 24. Los resultados indican que el tratamiento con dosis crecientes de ISIS 603538 no tuvo cambios significativos sobre el peso corporal de las ratas.

**Tabla 24**

| Pesos corporales de las ratas (% del peso corporal inicial) | | | | | | |
|---|---|---|---|---|---|---|
| | Dosis (µg) | Semana 1 | Semana 2 | Semana 3 | Semana 4 | Semana 5 |
| PBS | | 100 | 94 | 103 | 105 | 109 |
| ISIS 603538 | 700 | 100 | 94 | 98 | 103 | 107 |
| | 1000 | 100 | 95 | 97 | 101 | 103 |
| | 3000 | 100 | 92 | 98 | 102 | 105 |

LISTADO DE SECUENCIAS

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

<120> MÉTODOS PARA MODULAR LA EXPRESIÓN DE G9ORF72

<130> BIOLC1214WO

<150> 61/714,139
<151> 2012-10-15

<160> 35

<170> FastSEQ para Windows Versión 4.0

<210> 1
<211> 3339
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (203)...(1648)

<400> 1

```
acgtaaccta cggtgtcccg ctaggaaaga gaggtgcgtc aaacagcgac aagttccgcc 60
cacgtaaaag atgacgcttg gtgtgtcagc cgtccctgct gcccggttgc ttctcttttg 120
ggggcggggt ctagcaagag caggtgtggg tttaggagat atctccggag catttggata 180
atgtgacagt tggaatgcag tg atg tcg act ctt tgc cca ccg cca tct cca  232
                          Met Ser Thr Leu Cys Pro Pro Pro Ser Pro
                           1               5                  10

gct gtt gcc aag aca gag att gct tta agt ggc aaa tca cct tta tta   280
Ala Val Ala Lys Thr Glu Ile Ala Leu Ser Gly Lys Ser Pro Leu Leu
                 15                  20                  25

gca gct act ttt gct tac tgg gac aat att ctt ggt cct aga gta agg   328
Ala Ala Thr Phe Ala Tyr Trp Asp Asn Ile Leu Gly Pro Arg Val Arg
             30                  35                  40

cac att tgg gct cca aag aca gaa cag gta ctt ctc agt gat gga gaa   376
His Ile Trp Ala Pro Lys Thr Glu Gln Val Leu Leu Ser Asp Gly Glu
         45                  50                  55

ata act ttt ctt gcc aac cac act cta aat gga gaa atc ctt cga aat   424
Ile Thr Phe Leu Ala Asn His Thr Leu Asn Gly Glu Ile Leu Arg Asn
     60                  65                  70

gca gag agt ggt gct ata gat gta aag ttt ttt gtc ttg tct gaa aag   472
Ala Glu Ser Gly Ala Ile Asp Val Lys Phe Phe Val Leu Ser Glu Lys
 75                  80                  85                  90

gga gtg att att gtt tca tta atc ttt gat gga aac tgg aat ggg gat   520
Gly Val Ile Ile Val Ser Leu Ile Phe Asp Gly Asn Trp Asn Gly Asp
```

```
                    95                  100                 105

cgc agc aca tat gga cta tca att ata ctt cca cag aca gaa ctt agt   568
Arg Ser Thr Tyr Gly Leu Ser Ile Ile Leu Pro Gln Thr Glu Leu Ser
            110                 115                 120

ttc tac ctc cca ctt cat aga gtg tgt gtt gat aga tta aca cat ata   616
Phe Tyr Leu Pro Leu His Arg Val Cys Val Asp Arg Leu Thr His Ile
        125                 130                 135

atc cgg aaa gga aga ata tgg atg cat aag gaa aga caa gaa aat gtc   664
Ile Arg Lys Gly Arg Ile Trp Met His Lys Glu Arg Gln Glu Asn Val
    140                 145                 150

cag aag att atc tta gaa ggc aca gag aga atg gaa gat cag ggt cag   712
Gln Lys Ile Ile Leu Glu Gly Thr Glu Arg Met Glu Asp Gln Gly Gln
155                 160                 165                 170

agt att att cca atg ctt act gga gaa gtg att cct gta atg gaa ctg   760
Ser Ile Ile Pro Met Leu Thr Gly Glu Val Ile Pro Val Met Glu Leu
                175                 180                 185

ctt tca tct atg aaa tca cac agt gtt cct gaa gaa ata gat ata gct   808
Leu Ser Ser Met Lys Ser His Ser Val Pro Glu Glu Ile Asp Ile Ala
            190                 195                 200

gat aca gta ctc aat gat gat gat att ggt gac agc tgt cat gaa ggc   856
Asp Thr Val Leu Asn Asp Asp Asp Ile Gly Asp Ser Cys His Glu Gly
        205                 210                 215

ttt ctt ctc aat gcc atc agc tca cac ttg caa acc tgt ggc tgt tcc   904
Phe Leu Leu Asn Ala Ile Ser Ser His Leu Gln Thr Cys Gly Cys Ser
    220                 225                 230

gtt gta gta ggt agc agt gca gag aaa gta aat aag ata gtc aga aca   952
Val Val Val Gly Ser Ser Ala Glu Lys Val Asn Lys Ile Val Arg Thr
235                 240                 245                 250

tta tgc ctt ttt ctg act cca gca gag aga aaa tgc tcc agg tta tgt   1000
Leu Cys Leu Phe Leu Thr Pro Ala Glu Arg Lys Cys Ser Arg Leu Cys
                255                 260                 265

gaa gca gaa tca tca ttt aaa tat gag tca ggg ctc ttt gta caa ggc   1048
Glu Ala Glu Ser Ser Phe Lys Tyr Glu Ser Gly Leu Phe Val Gln Gly
            270                 275                 280

ctg cta aag gat tca act gga agc ttt gtg ctg cct ttc cgg caa gtc   1096
Leu Leu Lys Asp Ser Thr Gly Ser Phe Val Leu Pro Phe Arg Gln Val
        285                 290                 295

atg tat gct cca tat ccc acc aca cac ata gat gtg gat gtc aat act   1144
Met Tyr Ala Pro Tyr Pro Thr Thr His Ile Asp Val Asp Val Asn Thr
    300                 305                 310

gtg aag cag atg cca ccc tgt cat gaa cat att tat aat cag cgt aga   1192
Val Lys Gln Met Pro Pro Cys His Glu His Ile Tyr Asn Gln Arg Arg
315                 320                 325                 330
```

```
tac atg aga tcc gag ctg aca gcc ttc tgg aga gcc act tca gaa gaa   1240
Tyr Met Arg Ser Glu Leu Thr Ala Phe Trp Arg Ala Thr Ser Glu Glu
                335                 340                 345

gac atg gct cag gat acg atc atc tac act gac gaa agc ttt act cct   1288
Asp Met Ala Gln Asp Thr Ile Ile Tyr Thr Asp Glu Ser Phe Thr Pro
            350                 355                 360

gat ttg aat att ttt caa gat gtc tta cac aga gac act cta gtg aaa   1336
Asp Leu Asn Ile Phe Gln Asp Val Leu His Arg Asp Thr Leu Val Lys
        365                 370                 375

gcc ttc ctg gat cag gtc ttt cag ctg aaa cct ggc tta tct ctc aga   1384
Ala Phe Leu Asp Gln Val Phe Gln Leu Lys Pro Gly Leu Ser Leu Arg
    380                 385                 390

agt act ttc ctt gca cag ttt cta ctt gtc ctt cac aga aaa gcc ttg   1432
Ser Thr Phe Leu Ala Gln Phe Leu Leu Val Leu His Arg Lys Ala Leu
395                 400                 405                 410

aca cta ata aaa tat ata gaa gac gat acg cag aag gga aaa aag ccc   1480
Thr Leu Ile Lys Tyr Ile Glu Asp Asp Thr Gln Lys Gly Lys Lys Pro
                415                 420                 425

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

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

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

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

aatcatgatc gctggtaaag tagctcagtg gtgtggggaa acgttcccct ggatcatact 1738
ccagaattct gctctcagca attgcagtta agtaagttac actacagttc tcacaagagc 1798
ctgtgagggg atgtcaggtg catcattaca ttgggtgtct cttttcctag atttatgctt 1858
ttgggataca gacctatgtt tacaatataa taaatattat tgctatcttt taaagatata 1918
ataataggat gtaaacttga ccacaactac tgtttttttg aaatacatga ttcatggttt 1978
acatgtgtca aggtgaaatc tgagttggct tttacagata gttgactttc tatcttttgg 2038
cattctttgg tgtgtagaat tactgtaata cttctgcaat caactgaaaa ctagagcctt 2098
taaatgattt caattccaca gaaagaaagt gagcttgaac ataggatgag ctttagaaag 2158
aaaattgatc aagcagatgt ttaattggaa ttgattatta gatcctactt tgtggattta 2218
gtccctggga ttcagtctgt agaaatgtct aatagttctc tatagtcctt gttcctggtg 2278
aaccacagtt agggtgtttt gtttatttta ttgttcttgc tattgttgat attctatgta 2338
gttgagctct gtaaaaggaa attgtatttt atgttttagt aattgttgcc aacttttaa 2398
attaattttc attatttttg agccaaattg aaatgtgcac ctcctgtgcc ttttttctcc 2458
ttagaaaatc taattacttg gaacaagttc agatttcact ggtcagtcat tttcatcttg 2518
ttttcttctt gctaagtctt accatgtacc tgctttggca atcattgcaa ctctgagatt 2578
ataaaatgcc ttagagaata tactaactaa taagatcttt ttttcagaaa cagaaaatag 2638
ttccttgagt acttccttct tgcatttctg cctatgtttt tgaagttgtt gctgtttgcc 2698
tgcaataggc tataaggaat agcaggagaa attttactga agtgctgttt tcctaggtgc 2758
```

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

<210> 2
<211> 30001
<212> ADN
<213> Homo sapiens

<400> 2

caaagaaaag ggggaggttt tgttaaaaaa gagaaatgtt acatagtgct ctttgagaaa 60
attcattggc actattaagg atctgaggag ctggtgagtt tcaactggtg agtgatggtg 120
gtagataaaa ttagagctgc agcaggtcat tttagcaact attagataaa actggtctca 180
ggtcacaacg ggcagttgca gcagctggac ttggagagaa ttacactgtg ggagcagtgt 240
catttgtcct aagtgctttt ctacccccta cccccactat tttagttggg tataaaaaga 300
atgacccaat ttgtatgatc aactttcaca aagcatagaa cagtaggaaa agggtctgtt 360
tctgcagaag gtgtagacgt tgagagccat tttgtgtatt tattcctccc tttcttcctc 420
ggtgaatgat taaaacgttc tgtgtgattt ttagtgatga aaaagattaa atgctactca 480
ctgtagtaag tgccatctca cacttgcaga tcaaaaggca cacagtttaa aaaacctttg 540
ttttttttaca catctgagtg gtgtaaatgc tactcatctg tagtaagtgg aatctataca 600
cctgcagacc aaaagacgca aggtttcaaa aatctttgtg ttttttacac atcaaacaga 660
atggtacgtt tttcaaaagt taaaaaaaaa caactcatcc acatattgca actagcaaaa 720
atgacattcc ccagtgtgaa aatcatgctt gagagaattc ttacatgtaa aggcaaaatt 780
gcgatgactt tgcaggggac cgtgggattc ccgcccgcag tgccggagct gtcccctacc 840
agggtttgca gtggagtttt gaatgcactt aacagtgtct tacggtaaaa acaaaatttc 900
atccaccaat tatgtgttga gcgcccactg cctaccaagc acaaacaaaa ccattcaaaa 960
ccacgaaatc gtcttcactt tctccagatc cagcagcctc ccctattaag gttcgcacac 1020
gctattgcgc caacgctcct ccagagcggg tcttaagata aaagaacagg acaagttgcc 1080
ccgccccatt tcgctagcct cgtgagaaaa cgtcatcgca catagaaaac agacagacgt 1140
aacctacggt gtcccgctag gaaagagagg tgcgtcaaac agcgacaagt tccgcccacg 1200
taaaagatga cgcttggtgt gtcagccgtc cctgctgccc ggttgcttct cttttggggg 1260
cggggtctag caagagcagg tgtgggttta ggaggtgtgt gtttttgttt ttcccaccct 1320
ctctccccac tacttgctct cacagtactc gctgagggtg aacaagaaaa gacctgataa 1380
agattaacca gaagaaaaca aggagggaaa caaccgcagc ctgtagcaag ctctggaact 1440
caggagtcgc gcgctagggg ccggggccgg ggccggggcg tggtcggggc gggccccggg 1500
gcgggcccgg ggcggggctg cggttgcggt gcctgcgccc gcggcggcgg aggcgcaggc 1560
ggtggcgagt gggtgagtga ggaggcggca tcctggcggg tggctgtttg gggttcggct 1620
gccgggaaga ggcgcgggta gaagcggggg ctctcctcag agctcgacgc atttttactt 1680
tccctctcat ttctctgacc gaagctgggt gtcgggcttt cgcctctagc gactggtgga 1740
attgcctgca tccgggcccc gggcttcccg gcggcggcgg cggcggcggc ggcgcaggga 1800
caagggatgg ggatctggcc tcttccttgc tttcccgccc tcagtacccg agctgtctcc 1860
ttcccgggga cccgctggga gcgctgccgc tgcgggctcg agaaaaggga gcctcgggta 1920
ctgagaggcc tcgcctgggg gaaggccgga gggtgggcgg cgcgcggctt ctgcggacca 1980
agtcggggtt cgctaggaac ccgagacggt ccctgccggc gaggagatca tgcgggatga 2040
gatgggggtg tggagacgcc tgcacaattt cagcccaagc ttctagagag tggtgatgac 2100
ttgcatatga gggcagcaat gcaagtcggt gtgctcccca ttctgtggga catgacctgg 2160
ttgcttcaca gctccgagat gacacagact tgcttaaagg aagtgactat tgtgacttgg 2220
gcatcacttg actgatggta atcagttgtc taaagaagtg cacagattac atgtccgtgt 2280
gctcattggg tctatctggc cgcgttgaac accaccaggc tttgtattca gaaacaggag 2340
ggaggtcctg cactttccca ggaggggtgg ccctttcaga tgcaatcgag attgttaggc 2400
tctgggagag tagttgcctg gttgtggcag ttggtaaatt tctattcaaa cagttgccat 2460

gcaccagttg ttcacaacaa gggtacgtaa tctgtctggc attacttcta cttttgtaca 2520
aaggatcaaa aaaaaaaaag atactgttaa gatatgattt ttctcagact ttgggaaact 2580
tttaacataa tctgtgaata tcacagaaac aagactatca tataggggat attaataacc 2640
tggagtcaga atacttgaaa tacggtgtca tttgacacgg gcattgttgt caccacctct 2700
gccaaggcct gccactttag gaaaaccctg aatcagttgg aaactgctac atgctgatag 2760
tacatctgaa acaagaacga gagtaattac cacattccag attgttcact aagccagcat 2820
ttacctgctc caggaaaaaa ttacaagcac cttatgaagt tgataaaata ttttgtttgg 2880
ctatgttggc actccacaat ttgctttcag agaaacaaag taaaccaagg aggacttctg 2940
tttttcaagt ctgccctcgg gttctattct acgttaatta gatagttccc aggaggacta 3000
ggttagccta cctattgtct gagaaacttg gaactgtgag aaatggccag atagtgatat 3060
gaacttcacc ttccagtctt ccctgatgtt gaagattgag aaagtgttgt gaactttctg 3120
gtactgtaaa cagttcactg tccttgaagt ggtcctgggc agctcctgtt gtggaaagtg 3180
gacggtttag gatcctgctt ctctttgggc tgggagaaaa taaacagcat ggttacaagt 3240
attgagagcc aggttggaga aggtggctta cacctgtaat gccagagctt tgggaggcgg 3300
aggcaagagg atcacttgaa gccaggagtt caagctcaac ctgggcaacg tagaccctgt 3360
ctctacaaaa aattaaaaac ttagccgggc gtggtgatgt gcacctgtag tcctagctac 3420
ttgggaggct gaggcaggag ggtcatttga gcccaagagt ttgaagttac cgagagctat 3480
gatcctgcca gtgcattcca gcctggatga caaaacgaga ccctgtctct aaaaaacaag 3540
aagtgagggc tttatgattg tagaattttc actacaatag cagtggacca accacctttc 3600
taaataccaa tcagggaaga gatggttgat tttttaacag acgtttaaag aaaaagcaaa 3660
acctcaaact tagcactcta ctaacagttt tagcagatgt taattaatgt aatcatgtct 3720
gcatgtatgg gattatttcc agaaagtgta ttgggaaacc tctcatgaac cctgtgagca 3780
agccaccgtc tcactcaatt tgaatcttgg cttccctcaa aagactggct aatgtttggt 3840
aactctctgg agtagacagc actacatgta cgtaagatag gtacataaac aactattggt 3900
tttgagctga ttttttttcag ctgcatttgc atgtatggat ttttctcacc aaagacgatg 3960
acttcaagta ttagtaaaat aattgtacag ctctcctgat tatacttctc tgtgacattt 4020
catttcccag gctatttctt ttggtaggat ttaaaactaa gcaattcagt atgatctttg 4080
tccttcattt tctttcttat tctttttgtt tgtttgtttg tttgtttttt tcttgaggca 4140
gagtctctct ctgtcgccca ggctggagtg cagtggcgcc atctcagctc attgcaacct 4200
ctgccacctc cgggttcaag agattctcct gcctcagcct cccgagtagc tgggattaca 4260
ggtgtccacc accacacccg gctaattttt tgtatttta gtagaggtgg ggtttcacca 4320
tgttggccag gctggtcttg agctcctgac ctcaggtgat ccacctgcct cggcctacca 4380
aagagctggg ataacaggtg tgacccacca tgcccggccc attttttttt tcttattctg 4440
ttaggagtga gagtgtaact agcagtataa tagttcaatt ttcacaacgt ggtaaaagtt 4500
tccctataat tcaatcagat tttgctccag ggttcagttc tgttttagga aatactttta 4560
ttttcagttt aatgatgaaa tattagagtt gtaatattgc ctttatgatt atccaccttt 4620
ttaacctaaa agaatgaaag aaaaatatgt ttgcaatata attttatggt tgtatgttaa 4680
cttaattcat tatgttggcc tccagtttgc tgttgttagt tatgacagca gtagtgtcat 4740
taccatttca attcagatta cattcctata tttgatcatt gtaaactgac tgcttacatt 4800
gtattaaaaa cagtggatat tttaaagaag ctgtacggct tatatctagt gctgtctctt 4860
aagactatta aattgataca acatatttaa aagtaaatat tacctaaatg aatttttgaa 4920
attacaaata cacgtgttaa aactgtcgtt gtgttcaacc atttctgtac atacttagag 4980
ttaactgttt tgccaggctc tgtatgccta ctcataatat gataaaagca ctcatctaat 5040
gctctgtaaa tagaagtcag tgctttccat cagactgaac tctcttgaca agatgtggat 5100
gaaattcttt aagtaaaatt gtttactttg tcatacattt acagatcaaa tgttagctcc 5160
caaagcaatc atatggcaaa gataggtata tcatagtttg cctattagct gctttgtatt 5220
gctattatta taaatagact tcacagtttt agacttgctt aggtgaaatt gcaattcttt 5280
ttactttcag tcttagataa caagtcttca attatagtac aatcacacat tgcttaggaa 5340
tgcatcatta ggcgattttg tcattatgca aacatcatag agtgtactta cacaaaccta 5400
gatagtatag cctttatgta cctaggccgt atggtatagt ctgttgctcc taggccacaa 5460
acctgtacaa ctgttactgt actgaatact atagacagtt gtaacacagt ggtaaatatt 5520
tatctaaata tatgcaaaca gagaaaaggt acagtaaaag tatggtataa aagataatgg 5580
tatacctgtg taggccactt accacgaatg gagcttgcag gactagaagt tgctctgggt 5640
gagtcagtga gtgagtggtg aattaatgtg aaggcctaga acactgtaca ccactgtaga 5700
ctataaacac agtacgctga agctacacca aatttatctt aacagttttt cttcaataaa 5760
aaattataac tttttaactt tgtaaacttt ttaatttttt aacttttaaa atacttagct 5820
tgaaacacaa atacattgta tagctataca aaaatatttt ttctttgtat ccttattcta 5880
gaagcttttt tctattttct attttaaatt ttttttttta cttgttagtc gtttttgtta 5940

aaaactaaaa cacacacact ttcacctagg catagacagg attaggatca tcagtatcac 6000
tcccttccac ctcactgcct tccacctcca catcttgtcc cactggaagg tttttagggg 6060
caataacaca catgtagctg tcacctatga taacagtgct ttctgttgaa tacctcctga 6120
aggacttgcc tgaggctgtt ttacatttaa cttaaaaaaa aaaaaagtag aaggagtgca 6180
ctctaaaata acaataaaag gcatagtata gtgaatacat aaaccagcaa tgtagtagtt 6240
tattatcaag tgttgtacac tgtaataatt gtatgtgcta tactttaaat aacttgcaaa 6300
atagtactaa gaccttatga tggttacagt gtcactaagg caatagcata ttttcaggtc 6360
cattgtaatc taatgggact accatcatat atgcagtcta ccattgactg aaacgttaca 6420
tggcacataa ctgtatttgc aagaatgatt tgttttacat taatatcaca taggatgtac 6480
ctttttagag tggtatgttt atgtggatta agatgtacaa gttgagcaag gggaccaaga 6540
gccctgggtt ctgtcttgga tgtgagcgtt tatgttcttc tcctcatgtc tgttttctca 6600
ttaaattcaa aggcttgaac gggccctatt tagcccttct gttttctacg tgttctaaat 6660
aactaaagct tttaaattct agccatttag tgtagaactc tctttgcagt gatgaaatgc 6720
tgtattggtt tcttggctag catattaaat atttttatct ttgtcttgat acttcaatgt 6780
cgttttaaac atcaggatcg ggcttcagta ttctcataac cagagagttc actgaggata 6840
caggactgtt tgcccatttt ttgttatggc tccagacttg tggtatttcc atgtcttttt 6900
ttttttttct ttttttgacc ttttagcggc tttaaagtat ttctgttgtt aggtgttgta 6960
ttacttttct aagattactt aacaaagcac cacaaactga gtggctttaa acaacagcaa 7020
tttattctct cacaattcta gaagctagaa gtccgaaatc aaagtgttga caggggcatg 7080
atcttcaaga gagaagactc tttccttgcc tcttcctggc ttctggtggt taccagcaat 7140
cctgagtgtt cctttcttgc cttgtagttt caacaatcca gtatctgcct tttgtcttca 7200
catggctgtc taccatttgt ctctgtgtct ccaaatctct ctccttataa acacagcagt 7260
tattggatta ggccccactc taatccagta tgaccccatt ttaacatgat tacacttatt 7320
tctagataag gtcacattca cgtacaccaa gggttaggaa ttgaacatat ctttttgggg 7380
gacacaattc aacccacaag tgtcagtctc tagctgagcc tttcccttcc tgtttttctc 7440
ctttttagtt gctatgggtt aggggccaaa tctccagtca tactagaatt gcacatggac 7500
tggatatttg ggaatactgc gggtctattc tatgagcttt agtatgtaac atttaatatc 7560
agtgtaaaga agccctttt taagttattt ctttgaattt ctaaatgtat gccctgaata 7620
taagtaacaa gttaccatgt cttgtaaaat gatcatatca acaaacattt aatgtgcacc 7680
tactgtgcta gttgaatgtc tttatcctga taggagataa caggattcca catctttgac 7740
ttaagaggac aaaccaaata tgtctaaatc atttggggtt ttgatggata tctttaaatt 7800
gctgaaccta atcattggtt tcatatgtca ttgtttagat atctccggag catttggata 7860
atgtgacagt tggaatgcag tgatgtcgac tctttgccca ccgccatctc cagctgttgc 7920
caagacagag attgctttaa gtggcaaatc acctttatta gcagctactt ttgcttactg 7980
ggacaatatt cttggtccta gagtaaggca catttgggct ccaaagacag aacaggtact 8040
tctcagtgat ggagaaataa cttttcttgc caaccacact ctaaatggag aaatccttcg 8100
aaatgcagag agtggtgcta tagatgtaaa gttttttgtc ttgtctgaaa agggagtgat 8160
tattgtttca ttaatctttg atggaaactg gaatggggat cgcagcacat atggactatc 8220
aattatactt ccacagacag aacttagttt ctacctccca cttcatagag tgtgtgttga 8280
tagattaaca catataatcc ggaaaggaag aatatggatg cataaggtaa gtgatttttc 8340
agcttattaa tcatgttaac ctatctgttg aaagcttatt ttctggtaca tataaatctt 8400
atttttttaa ttatatgcag tgaacatcaa acaataaatg ttatttattt tgcatttacc 8460
ctattagata caaatacatc tggtctgata cctgtcatct tcatattaac tgtggaaggt 8520
acgaaatggt agctccacat tatagatgaa aagctaaagc ttagacaaat aaagaaactt 8580
ttagaccctg gattcttctt gggagccttt gactctaata ccttttgttt ccctttcatt 8640
gcacaattct gtcttttgct tactactatg tgtaagtata acagttcaaa gtaatagttt 8700
cataagctgt tggtcatgta gcctttggtc tctttaacct ctttgccaag ttcccaggtt 8760
cataaaatga ggaggttgaa tggaatggtt cccaagagaa ttccttttaa tcttacagaa 8820
attattgttt tcctaaatcc tgtagttgaa tatataatgc tatttacatt tcagtatagt 8880
tttgatgtat ctaaagaaca cattgaattc tccttcctgt gttccagttt gatactaacc 8940
tgaaagtcca ttaagcatta ccagttttaa aaggcttttg cccaatagta aggaaaaata 9000
atatctttta aaagaataat tttttactat gtttgcaggc ttacttcctt ttttctcaca 9060
ttatgaaact cttaaaatca ggagaatctt ttaaacaaca tcataatgtt taatttgaaa 9120
agtgcaagtc attcttttcc tttttgaaac tatgcagatg ttacattgac tgttttctgt 9180
gaagttatct ttttttcact gcagaataaa ggttgttttg attttatttt gtattgttta 9240
tgagaacatg catttgttgg gttaatttcc tacccctgcc cccatttttt ccctaaagta 9300
gaaagtattt ttcttgtgaa ctaaattact acacaagaac atgtctattg aaaaataagc 9360
aagtatcaaa atgttgtggg ttgttttttt aaataaattt tctcttgctc aggaaagaca 9420

agaaaatgtc cagaagatta tcttagaagg cacagagaga atggaagatc aggtatatgc 9480
aaattgcata ctgtcaaatg tttttctcac agcatgtatc tgtataaggt tgatggctac 9540
atttgtcaag gccttggaga catacgaata agcctttaat ggagctttta tggaggtgta 9600
cagaataaac tggaggaaga tttccatatc ttaaacccaa agagttaaat cagtaaacaa 9660
aggaaaatag taattgcatc tacaaattaa tatttgctcc ctttttttct ctgtttgccc 9720
agaataaatt ttggataact tgttcatagt aaaaataaaa aaaattgtct ctgatatgtt 9780
ctttaaggta ctacttctcg aacctttccc tagaagtagc tgtaacagaa ggagagcata 9840
tgtacccctg aggtatctgt ctggggtgta ggcccaggtc cacacaatat ttcttctaag 9900
tcttatgttg tatcgttaag actcatgcaa tttacatttt attccataac tattttagta 9960
ttaaaatttg tcagtgatat ttcttaccct ctcctctagg aaaatgtgcc atgtttatcc 10020
cttggctttg aatgcccctc aggaacagac actaagagtt tgagaagcat ggttacaagg 10080
gtgtggcttc ccctgcggaa actaagtaca gactatttca ctgtaaagca gagaagttct 10140
tttgaaggag aatctccagt gaagaaagag ttcttcactt ttacttccat ttcctcttgt 10200
gggtgaccct caatgctcct tgtaaaactc caatatttta aacatggctg ttttgccttt 10260
ctttgcttct ttttagcatg aatgagacag atgatacttt aaaaaagtaa ttaaaaaaaa 10320
aaacttgtga aaatacatgg ccataataca gaacccaata caatgatctc ctttaccaaa 10380
ttgttatgtt tgtacttttg tagatagctt tccaattcag agacagttat tctgtgtaaa 10440
ggtctgactt aacaagaaaa gatttccctt tacccaaaga atcccagtcc ttatttgctg 10500
gtcaataagc agggtcccca ggaatggggt aactttcagc accctctaac ccactagtta 10560
ttagtagact aattaagtaa acttatcgca agttgaggaa acttagaacc aactaaaatt 10620
ctgcttttac tgggattttg ttttttcaaa ccagaaacct ttacttaagt tgactactat 10680
taatgaattt tggtctctct tttaagtgct cttcttaaaa atgttatctt actgctgaga 10740
agttcaagtt tgggaagtac aaggaggaat agaaacttaa gagattttct tttagagcct 10800
cttctgtatt tagccctgta ggatttttct ttttttttct ttttttggtg ttgttgagct 10860
tcagtgaggc tattcattca cttatactga taatgtctga gatactgtga atgaaatact 10920
atgtatgctt aaacctaaga ggaaatattt tcccaaaatt attcttcccg aaaaggagga 10980
gttgcctttt gattgagttc ttgcaaatct cacaacgact ttattttgaa caatactgtt 11040
tggggatgat gcattagttt gaaacaactt cagttgtagc tgtcatctga taaaattgct 11100
tcacagggaa ggaaatttaa cacggatcta gtcattattc ttgttagatt gaatgtgtga 11160
attgtaattg taaacaggca tgataattat tactttaaaa actaaaaaca gtgaatagtt 11220
agttgtggag gttactaaag gatggttttt ttttaaataa aactttcagc attatgcaaa 11280
tgggcatatg gcttaggata aaacttccag aagtagcatc acatttaaat tctcaagcaa 11340
cttaataata tggggctctg aaaaactggt taaggttact ccaaaaatgg ccctgggtct 11400
gacaaagatt ctaacttaaa gatgcttatg aagactttga gtaaaatcat ttcataaaat 11460
aagtgaggaa aaacaactag tattaaattc atcttaaata atgtatgatt taaaaaatat 11520
gtttagctaa aaatgcatag tcatttgaca atttcattta tatctcaaaa aatttactta 11580
accaagttgg tcacaaaact gatgagactg gtggtggtag tgaataaatg agggaccatc 11640
catatttgag acactttaca tttgtgatgt gttatactga attttcagtt tgattctata 11700
gactacaaat ttcaaaatta caatttcaag atgtaataag tagtaatatc ttgaaatagc 11760
tctaaaggga atttttctgt tttattgatt cttaaaatat atgtgctgat tttgatttgc 11820
atttgggtag attatacttt tatgagtatg gaggttaggt attgattcaa gttttcctta 11880
cctatttggt aaggatttca aagtcttttt gtgcttggtt ttcctcattt ttaaatatga 11940
aatatattga tgacctttaa caaatttttt ttatctcaaa ttttaaagga gatcttttct 12000
aaaagaggca tgatgactta atcattgcat gtaacagtaa acgataaacc aatgattcca 12060
tactctctaa agaataaaag tgagctttag ggccgggcat ggtcagaaat ttgacaccaa 12120
cctggccaac atggcgaaac cccgtctcta ctaaaaatac aaaaatcagc cgggcatggt 12180
ggcggcacct atagtcccag ctacttggga ggatgagaca ggagagtcac ttgaacctgg 12240
gaggagaggt tgcagtgagc tgagatcacg ccattgcact ccagcctgag caatgaaagc 12300
aaaactccat ctcaaaaaaa aaaaaagaaa agaaagaata aaagtgagct ttggattgca 12360
tataaatcct ttagacatgt agtagacttg tttgatactg tgtttgaaca aattacgaag 12420
tattttcatc aaagaatgtt attgtttgat gttatttta ttttttattg cccagcttct 12480
ctcatattac gtgatttttct tcacttcatg tcactttatt gtgcagggtc agagtattat 12540
tccaatgctt actggagaag tgattcctgt aatggaactg ctttcatcta tgaaatcaca 12600
cagtgttcct gaagaaatag atgtaagttt aaatgagagc aattatacac tttatgagtt 12660
ttttggggtt atagtattat tatgtatatt attaatattc taatttaat agtaaggact 12720
ttgtcataca tactattcac atacagtatt agccacttta gcaaataagc acacacaaaa 12780
tcctggattt tatggcaaaa cagaggcatt tttgatcagt gatgacaaaa ttaaattcat 12840
tttgtttatt tcattacttt tataattcct aaaagtggga ggatcccagc tcttatagga 12900

```
gcaattaata tttaatgtag tgtcttttga aacaaaactg tgtgccaaag tagtaaccat 12960
taatggaagt ttacttgtag tcacaaattt agtttcctta atcatttgtt gaggacgttt 13020
tgaatcacac actatgagtg ttaagagata cctttaggaa actattcttg ttgttttctg 13080
attttgtcat ttaggttagt ctcctgattc tgacagctca gaagaggaag ttgttcttgt 13140
aaaaattgtt taacctgctt gaccagcttt cacatttgtt cttctgaagt ttatggtagt 13200
gcacagagat tgtttttgg ggagtcttga ttctcggaaa tgaaggcagt gtgttatatt 13260
gaatccagac ttccgaaaac ttgtatatta aaagtgttat ttcaacacta tgttacagcc 13320
agactaattt ttttatttt tgatgcattt tagatagctg atacagtact caatgatgat 13380
gatattggtg acagctgtca tgaaggcttt cttctcaagt aagaattttt cttttcataa 13440
aagctggatg aagcagatac catcttatgc tcacctatga caagatttgg aagaaagaaa 13500
ataacagact gtctacttag attgttctag ggacattacg tatttgaact gttgcttaaa 13560
tttgtgttat ttttcactca ttatatttct atatatattt ggtgttattc catttgctat 13620
ttaaagaaac cgagtttcca tcccagacaa gaaatcatgg ccccttgctt gattctggtt 13680
tcttgtttta cttctcatta aagctaacag aatcctttca tattaagttg tactgtagat 13740
gaacttaagt tatttaggcg tagaacaaaa ttattcatat ttatactgat ctttttccat 13800
ccagcagtgg agtttagtac ttaagagttt gtgcccttaa accagactcc ctggattaat 13860
gctgtgtacc cgtgggcaag gtgcctgaat tctctataca cctatttcct catctgtaaa 13920
atggcaataa tagtaatagt acctaatgtg tagggttgtt ataagcattg agtaagataa 13980
ataatataaa gcacttagaa cagtgcctgg aacataaaaa cacttaataa tagctcatag 14040
ctaacatttc ctatttacat ttcttctaga aatagccagt atttgttgag tgcctacatg 14100
ttagttcctt tactagttgc tttacatgta ttatcttata ttctgtttta aagtttcttc 14160
acagttacag attttcatga aattttactt ttaataaaag agaagtaaaa gtataaagta 14220
ttcactttta tgttcacagt cttttccttt aggctcatga tggagtatca gaggcatgag 14280
tgtgtttaac ctaagagcct taatggcttg aatcagaagc actttagtcc tgtatctgtt 14340
cagtgtcagc ctttcataca tcattttaaa tcccatttga ctttaagtaa gtcacttaat 14400
ctctctacat gtcaatttct tcagctataa aatgatggta tttcaataaa taaatacatt 14460
aattaaatga tattatactg actaattggg ctgttttaag gctcaataag aaaatttctg 14520
tgaaaggtct ctagaaaatg taggttccta tacaaataaa agataacatt gtgcttatag 14580
cttcggtgtt tatcatataa agctattctg agttatttga agagctcacc tacttttttt 14640
tgtttttagt ttgttaaatt gttttatagg caatgttttt aatctgtttt ctttaactta 14700
cagtgccatc agctcacact tgcaaacctg tggctgttcc gttgtagtag gtagcagtgc 14760
agagaaagta aataaggtag tttatttttat aatctagcaa atgatttgac tctttaagac 14820
tgatgatata tcatggattg tcatttaaat ggtaggttgc aattaaaatg atctagtagt 14880
ataaggaggc aatgtaatct catcaaattg ctaagacacc ttgtggcaac agtgagtttg 14940
aaataaactg agtaagaatc atttatcagt ttattttgat agctcggaaa taccagtgtc 15000
agtagtgtat aaatggtttt gagaatatat taaaatcaga tatataaaaa aaattactct 15060
tctatttccc aatgttatct ttaacaaatc tgaagatagt catgtacttt tggtagtagt 15120
tccaaagaaa tgttatttgt ttattcatct tgatttcatt gtcttcgctt tccttctaaa 15180
tctgtccctt ctagggagct attgggatta agtggtcatt gattattata ctttattcag 15240
taatgtttct gacccttcc ttcagtgcta cttgagttaa ttaaggatta atgaacagtt 15300
acatttccaa gcattagcta ataaactaaa ggattttgca cttttcttca ctgaccatta 15360
gttagaaaga gttcagagat aagtatgtgt atctttcaat ttcagcaaac ctaatttttt 15420
aaaaaaagtt ttacatagga aatatgttgg aaatgatact ttacaaagat attcataatt 15480
ttttttgta atcagctact ttgtatattt acatgagcct taatttatat ttctcatata 15540
accatttatg agagcttagt atacctgtgt cattatattg catctacgaa ctagtgacct 15600
tattccttct gttacctcaa acaggtggct ttccatctgt gatctccaaa gccttaggtt 15660
gcacagagtg actgccgagc tgctttatga agggagaaag gctccatagt tggagtgttt 15720
ttttttttt ttttaaacat ttttcccatc ctccatcctc ttgagggaga atagcttacc 15780
ttttatcttg ttttaatttg agaaagaagt tgccaccact ctaggttgaa aaccactcct 15840
ttaacataat aactgtggat atggtttgaa tttcaagata gttacatgcc tttttatttt 15900
tcctaataga gctgtaggtc aaatattatt agaatcagat ttctaaatcc cacccaatga 15960
cctgcttatt ttaaatcaaa ttcaataatt aattctcttc tttttggagg atctggacat 16020
tctttgatat ttcttacaac gaatttcatg tgtagaccca ctaaacagaa gctataaaag 16080
ttgcatggtc aaataagtct gagaaagtct gcagatgata taattcacct gaagagtcac 16140
agtatgtagc caaatgttaa aggttttgag atgccataca gtaaatttac caagcatttt 16200
ctaaatttat ttgaccacag aatccctatt ttaagcaaca actgttacat cccatggatt 16260
ccaggtgact aaagaatact tatttcttag gatatgtttt attgataata acaattaaaa 16320
tttcagatat ctttcataag caaatcagtg gtctttttac ttcatgtttt aatgctaaaa 16380
```

tattttcttt tatagatagt cagaacatta tgccttttc tgactccagc agagagaaaa 16440
tgctccaggt tatgtgaagc agaatcatca tttaaatatg agtcagggct ctttgtacaa 16500
ggcctgctaa aggtatagtt tctagttatc acaagtgaaa ccacttttct aaaatcattt 16560
ttgagactct ttatagacaa atcttaaata ttagcattta atgtatctca tattgacatg 16620
cccagagact gacttccttt acacagttct gcacatagac tatatgtctt atggatttat 16680
agttagtatc atcagtgaaa caccatagaa taccctttgt gttccaggtg ggtccctgtt 16740
cctacatgtc tagcctcagg actttttttt ttttaacaca tgcttaaatc aggttgcaca 16800
tcaaaaataa gatcatttct ttttaactaa atagatttga attttattga aaaaaaattt 16860
taaacatctt taagaagctt ataggattta agcaattcct atgtatgtgt actaaaatat 16920
atatatttct atatataata tatattagaa aaaaattgta tttttctttt atttgagtct 16980
actgtcaagg agcaaaacag agaaatgtaa attagcaatt atttataata cttaaaggga 17040
agaaagttgt tcaccttgtt gaatctatta ttgttatttc aattatagtc ccaagacgtg 17100
aagaaatagc tttcctaatg gttatgtgat tgtctcatag tgactacttt cttgaggatg 17160
tagccacggc aaaatgaaat aaaaaaattt aaaaattgtt gcaaatacaa gttatattag 17220
gcttttgtgc attttcaata atgtgctgct atgaactcag aatgatagta tttaaatata 17280
gaaactagtt aaaggaaacg tagtttctat ttgagttata catatctgta aattagaact 17340
tctcctgtta aaggcataat aaagtgctta atacttttgt ttcctcagca ccctctcatt 17400
taattatata attttagttc tgaaagggac ctataccaga tgcctagagg aaatttcaaa 17460
actatgatct aatgaaaaaa tatttaatag ttctccatgc aaatacaaat catatagttt 17520
tccagaaaat acctttgaca ttatacaaag atgattatca cagcattata atagtaaaaa 17580
aatggaaata gcctctttct tctgttctgt tcatagcaca gtgcctcata cgcagtaggt 17640
tattattaca tggtaactgg ctaccccaac tgattaggaa agaagtaaat ttgttttata 17700
aaaatacata ctcattgagg tgcatagaat aattaagaaa ttaaaagaca cttgtaattt 17760
tgaatccagt gaatacccac tgttaatatt tggtatatct ctttctagtc ttttttccc 17820
ttttgcatgt attttcttta agactcccac ccccactgga tcatctctgc atgttctaat 17880
ctgctttttt cacagcagat tctaagcctc tttgaatatc aacacaaact tcaacaactt 17940
catctataga tgccaaataa taaattcatt tttatttact taaccacttc ctttggatgc 18000
ttaggtcatt ctgatgtttt gctattgaaa ccaatgctat actgaacact tctgtcacta 18060
aaactttgca cacactcatg aatagcttct taggataaat ttttagagat ggatttgcta 18120
aatcagagac catttttaa aattaaaaaa caattattca tatcgtttgg catgtaagac 18180
agtaaatttt ccttttattt tgacaggatt caactggaag ctttgtgctg cctttccggc 18240
aagtcatgta tgctccatat cccaccacac acatagatgt ggatgtcaat actgtgaagc 18300
agatgccacc ctgtcatgaa catatttata atcagcgtag atacatgaga tccgagctga 18360
cagccttctg gagagccact tcagaagaag acatggctca ggatacgatc atctacactg 18420
acgaaagctt tactcctgat ttgtacgtaa tgctctgcct gctggtactg tagtcaagca 18480
atatgaaatt gtgtctttta cgaataaaaa caaaacagaa gttgcattta aaaagaaaga 18540
aatattacca gcagaattat gcttgaagaa acatttaatc aagcattttt ttcttaaatg 18600
ttcttctttt tccatacaat tgtgtttacc ctaaaatagg taagattaac ccttaaagta 18660
aatatttaac tatttgttta ataaatatat attgagctcc taggcactgt tctaggtacc 18720
gggcttaata gtggccaacc agacagcccc agccccagcc cctacattgt gtatagtcta 18780
ttatgtaaca gttattgaat ggacttatta acaaaaccaa agaagtaatt ctaagtcttt 18840
tttttcttga catatgaata taaaatacag caaaactgtt aaaatatatt aatggaacat 18900
ttttttactt tgcattttat attgttattc acttcttatt ttttttaaa aaaaaaagcc 18960
tgaacagtaa attcaaaagg aaaagtaatg ataattaatt gttgagcatg gacccaactt 19020
gaaaaaaaaa atgatgatga taaatctata atcctaaaac cctaagtaaa cacttaaaag 19080
atgttctgaa atcaggaaaa gaattatagt atacttttgt gtttctcttt tatcagttga 19140
aaaaaggcac agtagctcat gcctgtaaga acagagcttt gggagtgcaa ggcaggcgga 19200
tcacttgagg ccaggagttc cagaccagcc tgggcaacat agtgaaaccc catctctaca 19260
aaaaataaaa aagaattatt ggaatgtgtt tctgtgtgcc tgtaatccta gctattccga 19320
aagctgaggc aggaggatct tttgagccca ggagtttgag gttacaggga gttatgatgt 19380
gccagtgtac tccagcctgg ggaacaccga gactctgtct tatttaaaaa aaaaaaaaaa 19440
aaaatgcttg caataatgcc tggcacatag aaggtaacag taagtgttaa ctgtaataac 19500
ccaggtctaa gtgtgtaagg caatagaaaa attggggcaa ataagcctga cctatgtatc 19560
tacagaatca gtttgagctt aggtaacaga cctgtggagc accagtaatt acacagtaag 19620
tgttaaccaa aagcatagaa taggaatatc ttgttcaagg gacccccagc cttatacatc 19680
tcaaggtgca gaaagatgac ttaatatagg acccatttt tcctagttct ccagagtttt 19740
tattggttct tgagaaagta gtaggggaat gttttagaaa atgaattggt ccaactgaaa 19800
ttacatgtca gtaagttttt atatattggt aaatttagt agacatgtag aagttttcta 19860

attaatctgt gccttgaaac attttctttt ttcctaaagt gcttagtatt ttttccgttt 19920
tttgattggt tacttgggag ctttttgag gaaatttagt gaactgcaga atgggtttgc 19980
aaccatttgg tatttttgtt ttgtttttta gaggatgtat gtgtatttta acatttctta 20040
atcattttta gccagctatg tttgttttgc tgatttgaca aactacagtt agacagctat 20100
tctcattttg ctgatcatga caaaataata tcctgaattt ttaaattttg catccagctc 20160
taaattttct aaacataaaa ttgtccaaaa aatagtattt tcagccacta gattgtgtgt 20220
taagtctatt gtcacagagt cattttactt ttaagtatat gtttttacat gttaattatg 20280
tttgttattt ttaattttaa ctttttaaaa taattccagt cactgccaat acatgaaaaa 20340
ttggtcactg gaatttttt tttgactttt attttaggtt catgtgtaca tgtgcaggtg 20400
tgttatacag gtaaattgcg tgtcatgagg gtttggtgta caggtgattt cattacccag 20460
gtaataagca tagtacccaa taggtagttt tttgatcctc acccttctcc caccctcaag 20520
taggccctgg tgttgctgtt tccttctttg tgtccatgta tactcagtgt ttagctccca 20580
cttagaagtg agaacatgcg gtagttggtt ttctgttcct ggattagttc acttaggata 20640
atgacctcta gctccatctg gtttttatgg ctgcatagta ttccatggtg tatatgtatc 20700
acattttctt tatccagtct accattgata ggcatttagg ttgattccct gtctttgtta 20760
tcatgaatag tgctgtgatg aacatacaca tgcatgtgtc tttatggtag aaaaatttgt 20820
attcctttag gtacatatag aataatgggg ttgctagggt gaatggtagt tctattttca 20880
gttatttgag aaatcttcaa actgcttttc ataatagcta aactaattta cagtcccgcc 20940
agcagtgtat aagtgttccc ttttctccac aaccttgcca acatctgtga ttttttgact 21000
ttttaataat agccattcct agagaattga tttgcaattc tctattagtg atattaagca 21060
ttttttcata tgctttttag ctgtctgtat atattcttct gaaaaatttt catgtccttt 21120
gcccagtttg tagtggggtg ggttgttttt tgcttgttaa ttagttttaa gttccttcca 21180
gattctgcat atccctttgt tggatacatg gtttgcagat atttttctcc cattgtgtag 21240
gttgtctttt actctgttga tagtttcttt tgccatgcag gagctcgtta ggtcccattt 21300
gtgtttgttt ttgttgcagt tgcttttggc gtcttcatca taaaatctgt gccagggcct 21360
atgtccagaa tggtatttcc taggttgtct tccagggttt ttacaatttt agattttacg 21420
tttatgtctt taatccatct tgagttgatt tttgtatatg gcacaaggaa ggggtccagt 21480
ttcactccaa ttcctatggc tagcaattat cccagcacca tttattgaat acggagtcct 21540
ttccccattg cttgtttttt gtcaactttg ttgaagatca gatggttgta agtgtgtggc 21600
tttatttctt ggctctctat tctccattgg tctatgtgtc tgtttttata acagtaccct 21660
gctgttcagg ttcctatagc cttttagtat aaaatcggct aatgtgatgc ctccagcttt 21720
gttctttttg cttaggattg ctttggctat ttgggctcct ttttgggtcc atattaattt 21780
taaaacagtt ttttctggtt ttgtgaagga tatcattggt agtttatagg aatagcattg 21840
aatctgtaga ttgctttggg cagtatggcc attttaacaa tattaattct tcctatctat 21900
gaatatggaa tgtttttcca tgtgtttgtg tcatctcttt atacctgatg tataaagaaa 21960
agctggtatt attcctactc aatctgttcc aaaaaattga ggaggaggaa ctcttcccta 22020
atgaggccag catcattctg ataccaaaac ctggcagaga cacaacagaa aaaagaaaac 22080
ttcaggccaa tatccttgat gaatatagat gcaaaaatcc tcaacaaaat actagcaaac 22140
caaatccagc agcacatcaa aaagctgatc tactttgatc aagtaggctt tatccctggg 22200
atgcaaggtt ggttcaacat acacaaatca ataagtgtga ttcatcacat aaacagagct 22260
aaaaacaaaa accacaagat tatctcaata ggtagagaaa aggttgtcaa taaaatttaa 22320
catcctccat gttaaaaacc ttcagtaggt caggtgtagt gactcacacc tgtaatccca 22380
gcactttggg aggccaaggc gggcatatct cttaagccca ggagttcaag acgagcctag 22440
gcagcatggt gaaaccccat ctctacaaaa aaaaaaaaaa aaaaaaatta gcttggtatg 22500
gtgacatgca cctatagtcc cagctattca ggaggttgag gtgggaggat tgtttgagcc 22560
cgggaggcag aggttggcag cgagctgaga tcatgccacc gcactccagc ctgggcaacg 22620
gagtgagacc ctgtctcaaa aaagaaaaat cacaaacaat cctaaacaaa ctaggcattg 22680
aaggaacatg cctcaaaaaa ataagaacca tctatgacag acccatagcc aatatcttac 22740
caaatgggca aaagctggaa gtattctcct tgagaaccgt aacaagacaa ggatgtccac 22800
tctcaccact ccttttcagc atagttctgg aagtcctagc cagagcaatc aggaaagaga 22860
aagaaagaaa gacattcaga taggaagaga agaagtcaaa ctatttctgt ttgcaggcag 22920
tataattctg tacctagaaa atctcatagt ctctgcccag aaactcctaa atctgttaaa 22980
aatttcagca aagttttggc attctctata ctccaacacc ttccaaagtg agagcaaaat 23040
caagaacaca gtcccattca caatagccgc aaaacgaata aaatacctag gaatccagct 23100
aaccagggag gtgaaagatc tctatgagaa ttacaaaaca ctgctgaaag aaatcagaga 23160
tgacacaaac aaatggaaat gttctttttt aacaccttgc tttatctaat tcacttatga 23220
tgaagatact cattcagtgg aacaggtata ataagtccac tcgattaaat ataagcctta 23280
ttctctttcc agagcccaag aaggggcact atcagtgccc agtcaataat gacgaaatgc 23340

taatattttt ccccttacg gtttctttct tctgtagtgt ggtacactcg tttcttaaga 23400
taaggaaact tgaactacct tcctgtttgc ttctacacat acccattctc ttttttgcc 23460
actctggtca ggtataggat gatccctacc actttcagtt aaaaactcct cctcttacta 23520
aatgttctct taccctctgg cctgagtaga acctagggaa aatggaagag aaaaagatga 23580
aagggaggtg gggcctggga agggaataag tagtcctgtt tgtttgtgtg tttgctttag 23640
cacctgctat atcctaggtg ctgtgttagg cacacattat tttaagtggc cattatatta 23700
ctactactca ctctggtcgt tgccaaggta ggtagtactt tcttggatag ttggttcatg 23760
ttacttacag atggtgggct tgttgaggca aacccagtgg ataatcatcg gagtgtgttc 23820
tctaatctca ctcaaatttt tcttcacatt ttttggtttg ttttggtttt tgatggtagt 23880
ggcttatttt tgttgctggt ttgttttttg ttttttttttg agatggcaag aattggtagt 23940
tttatttatt aattgcctaa gggtctctac ttttttttaaa agatgagagt agtaaaatag 24000
attgatagat acatacatac ccttactggg gactgcttat attctttaga gaaaaaatta 24060
catattagcc tgacaaacac cagtaaaatg taaatatatc cttgagtaaa taaatgaatg 24120
tatattttgt gtctccaaat atatatatct atattcttac aaatgtgttt atatgtaata 24180
tcaatttata agaacttaaa atgttggctc aagtgaggga ttgtggaagg tagcattata 24240
tggccatttc aacatttgaa ctttttttctt ttcttcattt tcttcttttc ttcaggaata 24300
tttttcaaga tgtcttacac agagacactc tagtgaaagc cttcctggat caggtaaatg 24360
ttgaacttga gattgtcaga gtgaatgata tgacatgttt tctttttttaa tatatcctac 24420
aatgcctgtt ctatatattt atattcccct ggatcatgcc ccagagttct gctcagcaat 24480
tgcagttaag ttagttacac tacagttctc agaagagtct gtgagggcat gtcaagtgca 24540
tcattacatt ggttgcctct tgtcctagat ttatgcttcg ggaattcaga cctttgttta 24600
caatataata aatattattg ctatctttta aagatataat aataagatat aaagttgacc 24660
acaactactg ttttttgaaa catagaattc ctggtttaca tgtatcaaag tgaaatctga 24720
cttagctttt acagatataa tatatacata tatatatcct gcaatgcttg tactatatat 24780
gtagtacaag tatatatata tgtttgtgtg tgtatatata tatagtacga gcatatatac 24840
atattaccag cattgtagga tatatatatg tttatatatt aaaaaaaagt tataaactta 24900
aaaccctatt atgttatgta gagtatatgt tatatatgat atgtaaaata tataacatat 24960
actctatgat agagtgtaat atatttttta tatatatttt aacatttata aaatgataga 25020
attaagaatt gagtcctaat ctgttttatt aggtgctttt tgtagtgtct ggtctttcta 25080
aagtgtctaa atgatttttc cttttgactt attaatgggg aagagcctgt atattaacaa 25140
ttaagagtgc agcattccat acgtcaaaca acaaacattt taattcaagc attaacctat 25200
aacaagtaag ttttttttttt ttttttgaga aagggaggtt gtttatttgc ctgaaatgac 25260
tcaaaaatat ttttgaaaca tagtgtactt atttaaataa catctttatt gtttcattct 25320
tttaaaaaat atctacttaa ttacacagtt gaaggaaatc gtagattata tggaacttat 25380
ttcttaatat attacagttt gttataataa cattctgggg atcaggccag gaaactgtgt 25440
catagataaa gctttgaaat aatgagatcc ttatgtttac tagaaatttt ggattgagat 25500
ctatgaggtc tgtgacatat tgcgaagttc aaggaaaatt cgtaggcctg gaatttcatg 25560
cttctcaagc tgacataaaa tccctcccac tctccacctc atcatatgca cacattctac 25620
tcctacccac ccactccacc ccctgcaaaa gtacaggtat atgaatgtct caaaaccata 25680
ggctcatctt ctaggagctt caatgttatt tgaagatttg ggcagaaaaa attaagtaat 25740
acgaaataac ttatgtatga gttttaaaag tgaagtaaac atggatgtat tctgaagtag 25800
aatgcaaaat ttgaatgcat ttttaaagat aaattagaaa acttctaaaa actgtcagat 25860
tgtctgggcc tggtggctta tgcctgtaat cccagcactt tgggagtccg aggtgggtgg 25920
atcacaaggt caggagatcg agaccatcct gccaacatgg tgaaaccccg tctctactaa 25980
gtatacaaaa attagctggg cgtggcagcg tgtgcctgta atcccagcta cctgggaggc 26040
tgaggcagga gaatcgcttg aacccaggag gtgtaggttg cagtgagtca agatcgcgcc 26100
actgcacttt agcctggtga cagagctaga ctccgtctca aaaaaaaaaa aaaatatcag 26160
attgttccta cacctagtgc ttctatacca cactcctgtt agggggcatc agtggaaatg 26220
gttaaggaga tgtttagtgt gtattgtctg ccaagcactg tcaacactgt catagaaact 26280
tctgtacgag tagaatgtga gcaaattatg tgttgaaatg gttcctctcc ctgcaggtct 26340
ttcagctgaa acctggctta tctctcagaa gtactttcct tgcacagttt ctacttgtcc 26400
ttcacagaaa agccttgaca ctaataaaat atatagaaga cgatacgtga gtaaaactcc 26460
tacacggaag aaaaaccttt gtacattgtt tttttgtttt gtttcctttg tacattttct 26520
atatcataat ttttgcgctt ctttttttttt ttttttttttt ttttttttca ttatttttag 26580
gcagaaggga aaaaagccct ttaaatctct tcggaacctg aagatagacc ttgatttaac 26640
agcagagggc gatcttaaca taataatggc tctggctgag aaaattaaac caggcctaca 26700
ctcttttatc tttggaagac ctttctacac tagtgtgcaa gaacgagatg ttctaatgac 26760
tttttaaatg tgtaacttaa taagcctatt ccatcacaat catgatcgct ggtaaagtag 26820

ctcagtggtg tggggaaacg ttcccctgga tcatactcca gaattctgct ctcagcaatt 26880
gcagttaagt aagttacact acagttctca caagagcctg tgaggggatg tcaggtgcat 26940
cattacattg ggtgtctctt ttcctagatt tatgcttttg ggatacagac ctatgtttac 27000
aatataataa atattattgc tatcttttaa agatataata ataggatgta aacttgacca 27060
caactactgt ttttttgaaa tacatgattc atggtttaca tgtgtcaagg tgaaatctga 27120
gttggctttt acagatagtt gactttctat cttttggcat tctttggtgt gtagaattac 27180
tgtaatactt ctgcaatcaa ctgaaaacta gagcctttaa atgatttcaa ttccacagaa 27240
agaaagtgag cttgaacata ggatgagctt tagaaagaaa attgatcaag cagatgttta 27300
attggaattg attattagat cctactttgt ggatttagtc cctgggattc agtctgtaga 27360
aatgtctaat agttctctat agtccttgtt cctggtgaac cacagttagg gtgttttgtt 27420
tattttattg ttcttgctat tgttgatatt ctatgtagtt gagctctgta aaaggaaatt 27480
gtattttatg ttttagtaat tgttgccaac tttttaaatt aattttcatt atttttgagc 27540
caaattgaaa tgtgcacctc ctgtgccttt tttctcctta gaaaatctaa ttacttggaa 27600
caagttcaga tttcactggt cagtcatttt catcttgttt tcttcttgct aagtcttacc 27660
atgtacctgc tttggcaatc attgcaactc tgagattata aaatgcctta gagaatatac 27720
taactaataa gatcttttt tcagaaacag aaaatagttc cttgagtact tccttcttgc 27780
atttctgcct atgtttttga agttgttgct gtttgcctgc aataggctat aaggaatagc 27840
aggagaaatt ttactgaagt gctgttttcc taggtgctac tttggcagag ctaagttatc 27900
ttttgttttc ttaatgcgtt tggaccattt tgctggctat aaaataactg attaatataa 27960
ttctaacaca atgttgacat tgtagttaca caaacacaaa taaatatttt atttaaaatt 28020
ctggaagtaa tataaaaggg aaaatatatt tataagaaag ggataaaggt aatagagccc 28080
ttctgccccc cacccaccaa atttacacaa caaaatgaca tgttcgaatg tgaaaggtca 28140
taatagcttt cccatcatga atcagaaaga tgtggacagc ttgatgtttt agacaaccac 28200
tgaactagat gactgttgta ctgtagctca gtcatttaaa aaatatataa atactacctt 28260
gtagtgtccc atactgtgtt ttttacatgg tagattctta tttaagtgct aactggttat 28320
tttctttggc tggtttattg tactgttata cagaatgtaa gttgtacagt gaaataagtt 28380
attaaagcat gtgtaaacat tgttatatat cttttctcct aaatggagaa ttttgaataa 28440
aatatatttg aaattttgcc tctttcagtt gttcattcag aaaaaaatac tatgatattt 28500
gaagactgat cagcttctgt tcagctgaca gtcatgctgg atctaaactt tttttaaaat 28560
taattttgtc ttttcaaaga aaaaatattt aaagaagctt tataatataa tcttatgtta 28620
aaaaaacttt ctgcttaact ctctggattt cattttgatt tttcaaatta tatattaata 28680
tttcaaatgt aaaatactat ttagataaat tgtttttaaa cattcttatt attataatat 28740
taatataacc taaactgaag ttattcatcc caggtatcta atacatgtat ccaaagtaaa 28800
aatccaagga atctgaacac tttcatctgc aaagctagga ataggtttga cattttcact 28860
ccaagaaaaa gttttttttt gaaaatagaa tagttgggat gagaggtttc tttaaaagaa 28920
gactaactga tcacattact atgattctca aagaagaaac caaaacttca tataatacta 28980
taaagtaaat ataaaatagt tccttctata gtatatttct ataatgctac agtttaaaca 29040
gatcactctt atataatact attttgattt tgatgtagaa ttgcacaaat tgatatttct 29100
cctatgatct gcagggtata gcttaaagta acaaaaacag tcaaccacct ccatttaaca 29160
cacagtaaca ctatgggact agttttatta cttccatttt acaaatgagg aaactaaagc 29220
ttaaagatgt gtaatacacc gcccaaggtc acacagctgg taaaggtgga tttcatccca 29280
gacagttaca gtcattgcca tgggcacagc tcctaactta gtaactccat gtaactggta 29340
ctcagtgtag ctgaattgaa aggagagtaa ggaagcaggt tttacaggtc tacttgcact 29400
attcagagcc cgagtgtgaa tccctgctgt gctgcttgga gaagttactt aacctatgca 29460
aggttcattt tgtaaatatt ggaaatggag tgataatacg tacttcacca gaggatttaa 29520
tgagacctta tacgatcctt agttcagtac ctgactagtg cttcataaat gctttttcat 29580
ccaatctgac aatctccagc ttgtaattgg ggcatttaga acatttaata tgattattgg 29640
catggtaggt taaagctgtc atcttgctgt tttctatttg ttctttttgt tttctcctta 29700
cttttggatt tttttattct actatgtctt ttctattgtc ttattaacta tactctttga 29760
tttattttag tggttgtttt agggttatac ctctttctaa tttaccagtt tataaccagt 29820
ttatatacta cttgacatat agcttaagaa acttactgtt gttgtctttt tgctgttatg 29880
gtcttaacgt ttttatttct acaaacatta taaactccac actttattgt tttttaattt 29940
tacttataca gtcaattatc ttttaaagat atttaaatat aaacattcaa aacaccccaa 30000
t 30001

<210> 3

<211> 1031
<212> ADN
<213> Homo sapiens

<400> 3

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

<210> 4
<211> 3244
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (108) ... (1553)

<400> 4

```
gggcggggct gcggttgcgg tgcctgcgcc cgcggcggcg gaggcgcagg cggtggcgag 60
tggatatctc cggagcattt ggataatgtg acagttggaa tgcagtg atg tcg act   116
                                                    Met Ser Thr
                                                     1

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

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

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

gta ctt ctc agt gat gga gaa ata act ttt ctt gcc aac cac act cta   308
Val Leu Leu Ser Asp Gly Glu Ile Thr Phe Leu Ala Asn His Thr Leu
             55                  60                  65

aat gga gaa atc ctt cga aat gca gag agt ggt gct ata gat gta aag   356
Asn Gly Glu Ile Leu Arg Asn Ala Glu Ser Gly Ala Ile Asp Val Lys
         70                  75                  80
```

```
ttt ttt gtc ttg tct gaa aag gga gtg att att gtt tca tta atc ttt   404
Phe Phe Val Leu Ser Glu Lys Gly Val Ile Ile Val Ser Leu Ile Phe
     85                  90                  95

gat gga aac tgg aat ggg gat cgc agc aca tat gga cta tca att ata   452
Asp Gly Asn Trp Asn Gly Asp Arg Ser Thr Tyr Gly Leu Ser Ile Ile
100                 105                 110                 115

ctt cca cag aca gaa ctt agt ttc tac ctc cca ctt cat aga gtg tgt   500
Leu Pro Gln Thr Glu Leu Ser Phe Tyr Leu Pro Leu His Arg Val Cys
                120                 125                 130

gtt gat aga tta aca cat ata atc cgg aaa gga aga ata tgg atg cat   548
Val Asp Arg Leu Thr His Ile Ile Arg Lys Gly Arg Ile Trp Met His
            135                 140                 145

aag gaa aga caa gaa aat gtc cag aag att atc tta gaa ggc aca gag   596
Lys Glu Arg Gln Glu Asn Val Gln Lys Ile Ile Leu Glu Gly Thr Glu
        150                 155                 160

aga atg gaa gat cag ggt cag agt att att cca atg ctt act gga gaa   644
Arg Met Glu Asp Gln Gly Gln Ser Ile Ile Pro Met Leu Thr Gly Glu
    165                 170                 175

gtg att cct gta atg gaa ctg ctt tca tct atg aaa tca cac agt gtt   692
Val Ile Pro Val Met Glu Leu Leu Ser Ser Met Lys Ser His Ser Val
180                 185                 190                 195

cct gaa gaa ata gat ata gct gat aca gta ctc aat gat gat gat att   740
Pro Glu Glu Ile Asp Ile Ala Asp Thr Val Leu Asn Asp Asp Asp Ile
                200                 205                 210

ggt gac agc tgt cat gaa ggc ttt ctt ctc aat gcc atc agc tca cac   788
Gly Asp Ser Cys His Glu Gly Phe Leu Leu Asn Ala Ile Ser Ser His
            215                 220                 225

ttg caa acc tgt ggc tgt tcc gtt gta gta ggt agc agt gca gag aaa   836
Leu Gln Thr Cys Gly Cys Ser Val Val Val Gly Ser Ser Ala Glu Lys
        230                 235                 240

gta aat aag ata gtc aga aca tta tgc ctt ttt ctg act cca gca gag   884
Val Asn Lys Ile Val Arg Thr Leu Cys Leu Phe Leu Thr Pro Ala Glu
    245                 250                 255

aga aaa tgc tcc agg tta tgt gaa gca gaa tca tca ttt aaa tat gag   932
Arg Lys Cys Ser Arg Leu Cys Glu Ala Glu Ser Ser Phe Lys Tyr Glu
260                 265                 270                 275

tca ggg ctc ttt gta caa ggc ctg cta aag gat tca act gga agc ttt   980
Ser Gly Leu Phe Val Gln Gly Leu Leu Lys Asp Ser Thr Gly Ser Phe
                280                 285                 290

gtg ctg cct ttc cgg caa gtc atg tat gct cca tat ccc acc aca cac   1028
Val Leu Pro Phe Arg Gln Val Met Tyr Ala Pro Tyr Pro Thr Thr His
            295                 300                 305

ata gat gtg gat gtc aat act gtg aag cag atg cca ccc tgt cat gaa   1076
```

```
Ile Asp Val Asp Val Asn Thr Val Lys Gln Met Pro Pro Cys His Glu
        310                 315                 320

cat att tat aat cag cgt aga tac atg aga tcc gag ctg aca gcc ttc   1124
His Ile Tyr Asn Gln Arg Arg Tyr Met Arg Ser Glu Leu Thr Ala Phe
    325                 330                 335

tgg aga gcc act tca gaa gaa gac atg gct cag gat acg atc atc tac   1172
Trp Arg Ala Thr Ser Glu Glu Asp Met Ala Gln Asp Thr Ile Ile Tyr
340                 345                 350                 355

act gac gaa agc ttt act cct gat ttg aat att ttt caa gat gtc tta   1220
Thr Asp Glu Ser Phe Thr Pro Asp Leu Asn Ile Phe Gln Asp Val Leu
                360                 365                 370

cac aga gac act cta gtg aaa gcc ttc ctg gat cag gtc ttt cag ctg   1268
His Arg Asp Thr Leu Val Lys Ala Phe Leu Asp Gln Val Phe Gln Leu
            375                 380                 385

aaa cct ggc tta tct ctc aga agt act ttc ctt gca cag ttt cta ctt   1316
Lys Pro Gly Leu Ser Leu Arg Ser Thr Phe Leu Ala Gln Phe Leu Leu
        390                 395                 400

gtc ctt cac aga aaa gcc ttg aca cta ata aaa tat ata gaa gac gat   1364
Val Leu His Arg Lys Ala Leu Thr Leu Ile Lys Tyr Ile Glu Asp Asp
    405                 410                 415

acg cag aag gga aaa aag ccc ttt aaa tct ctt cgg aac ctg aag ata   1412
Thr Gln Lys Gly Lys Lys Pro Phe Lys Ser Leu Arg Asn Leu Lys Ile
420                 425                 430                 435

gac ctt gat tta aca gca gag ggc gat ctt aac ata ata atg gct ctg   1460
Asp Leu Asp Leu Thr Ala Glu Gly Asp Leu Asn Ile Ile Met Ala Leu
                440                 445                 450

gct gag aaa att aaa cca ggc cta cac tct ttt atc ttt gga aga cct   1508
Ala Glu Lys Ile Lys Pro Gly Leu His Ser Phe Ile Phe Gly Arg Pro
            455                 460                 465

ttc tac act agt gtg caa gaa cga gat gtt cta atg act ttt taa       1553
Phe Tyr Thr Ser Val Gln Glu Arg Asp Val Leu Met Thr Phe
        470                 475                 480

atgtgtaact taataagcct attccatcac aatcatgatc gctggtaaag tagctcagtg 1613
gtgtggggaa acgttcccct ggatcatact ccagaattct gctctcagca attgcagtta 1673
agtaagttac actacagttc tcacaagagc ctgtgagggg atgtcaggtg catcattaca 1733
ttgggtgtct cttttcctag atttatgctt ttgggataca gacctatgtt tacaatataa 1793
taaatattat tgctatcttt taaagatata ataataggat gtaaacttga ccacaactac 1853
tgtttttttg aaatacatga ttcatggttt acatgtgtca aggtgaaatc tgagttggct 1913
tttacagata gttgactttc tatcttttgg cattctttgg tgtgtagaat tactgtaata 1973
cttctgcaat caactgaaaa ctagagcctt taaatgattt caattccaca gaaagaaagt 2033
gagcttgaac ataggatgag ctttagaaag aaaattgatc aagcagatgt ttaattggaa 2093
ttgattatta gatcctactt tgtggattta gtccctggga ttcagtctgt agaaatgtct 2153
aatagttctc tatagtcctt gttcctggtg aaccacagtt agggtgtttt gtttatttta 2213
ttgttcttgc tattgttgat attctatgta gttgagctct gtaaaaggaa attgtatttt 2273
atgttttagt aattgttgcc aacttttaa attaattttc attatttttg agccaaattg 2333
aaatgtgcac ctcctgtgcc ttttttctcc ttagaaaatc taattacttg gaacaagttc 2393
agatttcact ggtcagtcat tttcatcttg ttttcttctt gctaagtctt accatgtacc 2453
```

tgctttggca atcattgcaa ctctgagatt ataaaatgcc ttagagaata tactaactaa 2513
taagatcttt ttttcagaaa cagaaaatag ttccttgagt acttccttct tgcatttctg 2573
cctatgtttt tgaagttgtt gctgtttgcc tgcaataggc tataaggaat agcaggagaa 2633
attttactga agtgctgttt tcctaggtgc tactttggca gagctaagtt atcttttgtt 2693
ttcttaatgc gtttggacca ttttgctggc tataaaataa ctgattaata taattctaac 2753
acaatgttga cattgtagtt acacaaacac aaataaatat tttatttaaa attctggaag 2813
taatataaaa gggaaaatat atttataaga aagggataaa ggtaatagag cccttctgcc 2873
ccccacccac caaatttaca caacaaaatg acatgttcga atgtgaaagg tcataatagc 2933
tttcccatca tgaatcagaa agatgtggac agcttgatgt tttagacaac cactgaacta 2993
gatgactgtt gtactgtagc tcagtcattt aaaaaatata taaatactac cttgtagtgt 3053
cccatactgt gtttttttaca tggtagattc ttatttaagt gctaactggt tattttcttt 3113
ggctggttta ttgtactgtt atacagaatg taagttgtac agtgaaataa gttattaaag 3173
catgtgtaaa cattgttata tatcttttct cctaaatgga gaattttgaa taaaatatat 3233
ttgaaatttt g 3244

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

<220>

<221> característica miscelánea
<222> 693, 722
<223> n = A,T,C o G

<400> 5

cacgaggctt tgatatttct tacaacgaat ttcatgtgta gacccactaa acagaagcta 60
taaaagttgc atggtcaaat aagtctgaga aagtctgcag atgatataat tcacctgaag 120
agtcacagta tgtagccaaa tgttaaaggt tttgagatgc catacagtaa atttaccaag 180
cattttctaa atttatttga ccacagaatc cctattttaa gcaacaactg ttacatccca 240
tggattccag gtgactaaag aatacttatt tcttaggata tgttttattg ataataacaa 300
ttaaaatttc agatatcttt cataagcaaa tcagtggtct ttttacttca tgttttaatg 360
ctaaaatatt ttcttttata gatagtcaga acattatgcc tttttctgac tccagcagag 420
agaaaatgct ccaggttatg tgaagcagaa tcatcattta aatatgagtc agggctcttt 480
gtacaaggcc tgctaaagga ttcaactgga agctttgtgc tgcctttccg gcaagtcatg 540
tatgctccat atcccaccac acacatagat gtggatgtca atactgtgaa gcagatgcca 600
ccctgtcatg aacatattta taatcagcgt agatacatga gatccgagct gacagccttc 660
tggagagcca cttcagaaga agacatggct cangatacga tcatctacac tgacgaaagc 720
tntactcctg atttgaatat ttttcaagat gtcttacaca g 761

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

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

<400> 6

acgtaaccta cggtgtcccg ctaggaaaga gaggtgcgtc aaacagcgac aagttccgcc 60
cacgtaaaag atgacgcttg atatctccgg agcatttgga taatgtgaca gttggaatgc 120
agtg atg tcg act ctt tgc cca ccg cca tct cca gct gtt gcc aag aca 169
Met Ser Thr Leu Cys Pro Pro Pro Ser Pro Ala Val Ala Lys Thr
1 5 10 15

gag att gct tta agt ggc aaa tca cct tta tta gca gct act ttt gct 217

```
Glu Ile Ala Leu Ser Gly Lys Ser Pro Leu Leu Ala Ala Thr Phe Ala
                 20                  25                  30

tac tgg gac aat att ctt ggt cct aga gta agg cac att tgg gct cca   265
Tyr Trp Asp Asn Ile Leu Gly Pro Arg Val Arg His Ile Trp Ala Pro
             35                  40                  45

aag aca gaa cag gta ctt ctc agt gat gga gaa ata act ttt ctt gcc   313
Lys Thr Glu Gln Val Leu Leu Ser Asp Gly Glu Ile Thr Phe Leu Ala
         50                  55                  60

aac cac act cta aat gga gaa atc ctt cga aat gca gag agt ggt gct   361
Asn His Thr Leu Asn Gly Glu Ile Leu Arg Asn Ala Glu Ser Gly Ala
     65                  70                  75

ata gat gta aag ttt ttt gtc ttg tct gaa aag gga gtg att att gtt   409
Ile Asp Val Lys Phe Phe Val Leu Ser Glu Lys Gly Val Ile Ile Val
 80                  85                  90                  95

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

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

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

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

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

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

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

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

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

```
cagactccct ggattaatgc tgtgtacccg tgggcaaggt gcctgaattc tctatacacc 1273
tatttcctca tctgtaaaat ggcaataata gtaatagtac ctaatgtgta gggttgttat 1333
aagcattgag taagataaat aatataaagc acttagaaca gtgcctggaa cataaaaaca 1393
cttaataata gctcatagct aacatttcct atttacattt cttctagaaa tagccagtat 1453
ttgttgagtg cctacatgtt agttccttta ctagttgctt tacatgtatt atcttatatt 1513
ctgttttaaa gtttcttcac agttacagat tttcatgaaa ttttacttt aataaaagag 1573
aagtaaaagt ataaagtatt cacttttatg ttcacagtct tttcctttag gctcatgatg 1633
gagtatcaga ggcatgagtg tgtttaacct aagagcctta atggcttgaa tcagaagcac 1693
tttagtcctg tatctgttca gtgtcagcct ttcatacatc attttaaatc ccatttgact 1753
ttaagtaagt cacttaatct ctctacatgt caatttcttc agctataaaa tgatggtatt 1813
tcaataaata aatacattaa ttaaatgata ttatactgac taattgggct gttttaaggc 1873
aaaaaaaaaa aaaaaaaaaa aaaaaaaa                                    1901
```

<210> 7
<211> 562
<212> ADN
<213> Homo sapiens

<220>
<221> característica miscelánea

<222> 166
<223> n = A,T,C o G

<400> 7

```
agacgtaacc tacggtgtcc cgctaggaaa gagagatatc tccggagcat ttggataatg 60
tgacagttgg aatgcagtga tgtcgactct ttgcccaccg ccatctccag ctgttgccaa 120
gacagagatt gctttaagtg gcaaatcacc tttattagca gctacntttt gcttactggg 180
acaatattct tggtcctaga gtaaggcaca tttgggctcc aaagacagaa caggtacttc 240
tcagtgatgg agaaataact tttcttgcca accacactct aaatggagaa atccttcgaa 300
atgcagagag tggtgctata gatgtaaagt tttttgtctt gtctgaaaag ggagtgatta 360
ttgtttcatt aatctttgat ggaaactgga atggggatcg cagcacatat ggactatcaa 420
ttatacttcc acagacagaa cttagtttct acctcccact tcatagagtg tgtgttgata 480
gattaacaca tataatccgg aaaggaagaa tatggatgca taaggaaaga caagaaaatg 540
tccagaagat tatcttagaa gg                                          562
```

<210> 8
<211> 798
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (81) ... (590)

<400> 8

```
gggctctctt ttgggggcgg ggtctagcaa gagcagatat ctccggagca tttggataat 60
gtgacagttg gaatgcagtg atg tcg act ctt tgc cca ccg cca tct cca gct 113
                        Met Ser Thr Leu Cys Pro Pro Pro Ser Pro Ala
                         1               5                  10

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

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

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

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

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

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

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

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

cgg aaa gga aga ata tgg atg cat aag gaa aga caa gaa aat gtc cag   545
Arg Lys Gly Arg Ile Trp Met His Lys Glu Arg Gln Glu Asn Val Gln
140                 145                 150                 155

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

agtattattc caatgcttac tggagaagtg attcctgtaa tgggactgct ttcatctatg 650
aaatcacaca gtgttcctga agaaatagat atagctgata cagtactcca tgatgatgat 710
atttggtgac agctgtcatg aaaggctttc ttctcaagta ggaatttttt cttttcataa 770
aagctgggat gaagccagat tcccatct                                   798
```

<210> 9
<211> 169
<212> ADN
<213> Homo sapiens

<400> 9

```
aaacagcgac aagttccgcc cacgtaaaag atgatgcttg gtgtgtcagc cgtccctgct 60
gcccggttgc ttctcttttg ggggcggggt ctagcaagag cagatatctc cggagcattt 120
ggataatgtg acagttggaa tgcggtgatg tcgactcttt gcccaccgc             169
```

<210> 10
<211> 176
<212> ADN
<213> Homo sapiens

<400> 10

```
aaaacgtcat cgcacataga aaacagacag acgtaaccta cggtgtcccg ctaggaaaga 60
gaggtgcgtc aaacagcgac aagttccgcc cacgtaaaag atgacgcttg atatctccgg 120
agcatttgga taatgtgaca gttggaatgc agtgatgtcg actctttgcc caccgc     176
```

<210> 11
<211> 38001
<212> ADN
<213> Mus musculus

<400> 11

tgtctctagg taaaattttg aaggaaaaaa aaaacactaa gaaggtatat tccttcaaag 60
ttccagtctt attctgaagt gtaatgttat gttagtttga ctcacagaca ggttttaaag 120
aagggcttac ttcaagagga caccaaacaa ataccttcta ttcctagtgg gctctggaat 180
cacagaaaac tgacccaatc aattacattg atagctctgg cttactacag acaagcaaat 240
tatcttaagt gtgcatgcat gcgcgtgtat gtgtgttagt acctaacacc cacctgggaa 300
cttttcagct tttcagtgtg ggatatagta taaacgtcta ttcctcgtgt tgtggattag 360
ctgactggcc tcactcagct gccttcctta cctgcaaact cacccacttt gactacagca 420
tcgcactctt aaccctagcc ttccaaacat ggtcctatgc tatttctgtg tgtctggatg 480
tatttttaac tctcagatgt atacttcatt tatgagatat acatctgaag accacggtac 540
aaaacactgt aagaacttga tagaatgaca actgctaggt aaaaaaaaaa aaaaaaaaaa 600
aaaaaaaaaa aaaaaaaagc atacaatacc tggtgagagt tctattttta ccgaaggtgg 660
tattgatagg tattctgtta ttaatgcctt tcttttccct ataaatgatg aaaagttgct 720
ggaaaataat aaacactact catctgtagt gaaaagccac aatacagtta caaaccaatc 780
aatcaatcaa taaatcagac gtcatggtgt tcttttccca aaggttaaaa aacaaagtgc 840
actgtgctat ttggcaaaaa tgacgtttag aagaaaacac ggtgactacg cacagagggt 900
gggggaatca ttgtgcttgt tgcggagtga acacgtacag tgtgcacgca gacttacggc 960
atttaaccgt gtcataggga ccaaaggaaa tccactcact cactaaatat ttgttgagca 1020
cccactacct gccaactccc aaacaaaaca aagcaaaact acttacaacc acaaactacg 1080
cttcgtaacc tagatagata acgcaggtga cactatctat ctaggttgag ctcagctctg 1140
cccatgcttt tcctgagcgg ctcttggaag aaaagctaca aagcccatga cagcctccgc 1200
ctggccagct gccactggca tctcaaggct ggcaaagcaa agtgaaagcg ccaacccgga 1260
acttacggag tcccacgagg gaaccgcggc gcgtcaagca gagacgagtt ccgcccacgt 1320
gaaagatggc gtttgtagtg acagccatcc caattgccct ttccttctag gtggaaagtg 1380
gtgtctagac agtccaggga gggtgtgcga gggaggtgcg ttttggttgc ctcagctcgc 1440
aacttaactc cacaacggtg accaaggaca aaagaaggaa acaagactgc agagatccgc 1500
accggggagc cctgcagatt ctgggtctgc tgtggactgg gggcgggact gcgactgggc 1560
gggcctgggg gcgtgtccgg ggcggggcgg tcccggggcg gggcccggag cgggctgcgg 1620
ttgcggtccc tgcgccggcg gtgaaggcgc agcagcggcg agtgggtgag tgagacgcgc 1680
gggcggaggg gggctgctgc cacggtcggc tcgcgggccg gccggctccg ggtaccagcg 1740
gggtttttt ctccttcgag gtgaactcct ccctgtcccc cgggcgaaag agcccttggc 1800
cttgcaggag ttgcgggggc cgcggcggtg cggaggggat ggggatgggc ctcatctttg 1860
ctgtccgccc gcgctccccg atcccgaccc ggagcgtctc ccgggcccct gagggaaccc 1920
tccgggagta cggcgagcgc ggcccccacc gccacaagcc tgggccccag gggcctggcc 1980
cggcgacagc tggtgggtcc tgcgacccag tcaggtctcc cgagggtccc cgcccgggag 2040
gagaaagcgc cggtgggatg gagtaaggac ggacagaaca acacgcaggc aggatttcgc 2100
agaagtttgc aaggagtgcg gatgcccact tacatgggct gctactctta ccaggttgtt 2160
ccccagttct gtgggacgtg acctggttgc ctcacagctc cgcggttgta cagacttatt 2220
aaaggaagtg accattgtga cttgggcatc acttgactga tggtaatcag ttgcagagag 2280
agaagtgcac tgattaagtc tgtccacaca gggtctgtct ggccaggagt gcatttgcct 2340
gggagggatt ggttgcgctt tctggtgtgg ggactattag gctcttgtag agttttgtcc 2400
cggcagatgg ataaatttct tgttacactg ttcccgttcg tcaccagttg agaaaaacgg 2460
gtacacagtc tgtctcagta gtacttttac tttatattaa gggcccaaaa gggactggaa 2520
aatactttaa gatagaatcg ttagtccact tggaaaactt aaaatatgag agagagaggg 2580
ggggggggaga gagagagaga gagagagaga gaaaggaagg aagaaggagg aagaggagga 2640
ggaaagagat tgagattatg ttaataatat ggaatcagaa tatttgaaat atagtaagcg 2700
tcccctcagt taaagaggac attccaggag gcccccagta tagcctgaaa tctcaggaaa 2760
cgcctacata cacccatcgt gtggatatag gtgttttccc ttcattacat ttcatacaca 2820
gatgttaaag tttagaaagt aggcacaata agagattaca aataactgat aataaagtcg 2880
agccattgca gctgctctgt aaaagtcctg tgaatgtgat cgctttgtgt ttcaaagtaa 2940
cttactgtac ttcacccctg ttaagcaaaa caagattcac ctgaacgcag gcaccttggt 3000
accttggcag acaccagatc tgataaccaa gaggatggag aagtagtggc agacagtgtg 3060
gagagcatga atatgctaga caaaagggtg aatcataacc taggagcaga aagcaggtat 3120

ttcatcatcc tccacagtaa aaacctatgt cacgtaaaaa acctacaagt agtttttctt 3180
ttactctttt tgaatgaaag cttgctacag gcactgaaag ttaaaataat ctgtggatca 3240
ggaggaacag gggttttctg tctgagtcac tgctgactag cacctcagtg accattggca 3300
ctgtgggaaa ccccagagtc agttggaaac ttcgaaacta aaggtgacgg tgttcttatt 3360
tcatagaaca caaaaaataa gaggggttac agcctgcgct gcagactgga cattcaacaa 3420
gcatttaaat ttctgggaga caaatgtaaa tataacttta aaagttggta aaatactctg 3480
tttggctatg ttggccatcc aatgtttgct tttagaaaat gactgaatgg ataaaacgtc 3540
tatcttttga gcctgcccta gacccccatg ttgagtgaat actgtccaag tgttaggtta 3600
gccggcctga gaaacttgga tctaggcaag atggcacagt cctggtgtca tgagtatgca 3660
tgtgagtttt ggctgaaatt gaacatttgt agagaatgac aaaggctggt ctggcaagta 3720
gtccactgtc tttacagtgg tcttggttag ttcctgtttg gctgagaggg ctggttgatg 3780
gctgtcctgc ccctcttccc acaagtggaa gccttatggt ataattcttg atcacagtag 3840
cagtaggcaa atgaacttcc tcaaagcagc ctggaaagct gatttttttt tctttctttc 3900
tcttttttt ttttttttca caaggttaaa gaaaaaacaa agggcttcaa atgtgccagt 3960
ctgctaacag tgttaacatg tttattaaca taaataaact ttattagttt ttggaagtat 4020
tggttaagcc ctcgtgaccc ctgaactcgg tttatagagt gatgagtcgt agcctcactc 4080
tggtttggac tctggcttct ctcagaagac tctgtggcta atgttaacct tctgaagtag 4140
ccagaaaaca tataagcaaa agtctgtgag gttgaaatga atttttggc cacatttgta 4200
tatgggttcc caccaatgct aacttcaggt gttagtaata tcagactcac agcttccctg 4260
attacacttc gctataagac tttatttttt aggtcatagg aatttcccct ttttcatgat 4320
tcctaaatca tgaaataaca tagtctaaaa atacggtatt cctgaaataa acaatttcta 4380
agttttaagc tgcgtgctat tctgaacagt ctgatgccct cttgtagctt ttactgtgtc 4440
ctaccccggg catggttgat tcctttgtcc aaacatctgt ctgttgtatc cacactggat 4500
tgcaccacct gcgtgctagt cagtcactca gacattttag ttataaggta gcttatattt 4560
actccttatt ttatttaata atggcctcat agcaaggcgg taatgatact ggtaatttgg 4620
gtttgcttaa gaggagccat gaagtagttt taaatgaaaa ggtgaaaatt cccactatag 4680
tttggagggg gaggctatac tggtactact acgattcacg gtaagactaa atcttctgtg 4740
aaattatgaa ggagaaaaag ttacactggt ctggtcttgc tgttggatta attttatagt 4800
tataaccact gtacatgata aataacccta aaacaatgaa tttgtaggtg gatggcataa 4860
tctgaaaacc atgttctgag cagttgatgg cagcaggctg tgctggaagt gttaggcata 4920
tttatagatt tcagcccaag ttctgaagag gctggagaga tggctcagtg gttaagagtg 4980
cttgctattg cagaggacct aggttcctct acaggcacca ggcaagcgtg ggacacactg 5040
agatacatac agacaaaaca taaaattaaa taaattgtgc ataataatac tagtaatata 5100
tgagtaaaat aaggataaat acacatcata attaaataaa taaattgtaa agttccctag 5160
aagtgagggt caccaagcca ttcacaagat ggctgcgctg atgcagggat atatgtgaac 5220
tagaaaaagg tcaaacttaa cagagaagtt ccaaggcatg ctactgcagg cttggctagc 5280
atgcttgacc tgcagaaatg ctgacggcca ctgggaggtt ttcacaaatg aggaattaga 5340
agaacttttt ttactaatct ccagaaaaaa aaaagggaag aagaaactga agcagcctgt 5400
gatgtggacc agaaacgcag tgacagtaac atgtgtgaca ttgcaaaggc atgaaaggac 5460
agagctgtgg aatacagacc tcaggtggag ctcagcatag agtcattcgg ggattatgcc 5520
tgctgcagca acaaaaggat gagctcaaaa gagacaccga cttctgaatg cagtgggtgt 5580
ttgttttgtt ttgtttcaaa tgaattgggc agaaaacttt ccagctgtgg aagcttctga 5640
accgtccctt gctgctgaca tctaagcgtc cgctgtgtcc cagctcagtg atctagggtc 5700
ttccaaacag atggtccggt gctgagcact ttgaatctca atcctgagtt tctaccacgc 5760
ctttggccat ttaattccca gataaaagac acatacaacc tttatattta taataaacct 5820
tagtcagcac aagagctgag caaatatctg tcctctatgc tattatatct attacccagc 5880
caataacccc attctataat ttgctgtgct tcatctgggc tgctcttaac ttcagtcagc 5940
cagcccacgt ggccattatt ttaagatttt tttaccccat agtgtcttct cactttactt 6000
tacatttttc tctctctcct catggttctc ctctgacccc aagcctagga accctaaacc 6060
ccacccatgt ctcttctgcc catctattgg ctgtaggcat ctttattcac caatcaggat 6120
aacttggagg caaggttaag tagtctcctg ggtctaggtg ctgtctctgg gagcaaccag 6180
tatttagcat agcaaaagac cagacctcca caatgatcac tctgaccatc ggggcagaag 6240
gcacctacta gcctgtgcca ctcacctcac tttgttgaat cacatcttat cctgtagtgt 6300
gtatcactgc ctgttatcac aggaaaaagt gagtcccatc aaataagatg tttcagaaag 6360
agaccatgtt catataatta tcattctggt aagcttttaa tggttatatt ttgttattaa 6420
tctctttgtt cctattttgc aaattatacc ttacagtaaa tatatatgca tccaatgggg 6480
tctttgaatt cctccccggg gagtaggagg actctttgag gatgggctgc atttaaagct 6540
aaacaacgca acatgacctt tagtccttat agatagccta gagatgagac taaataaaag 6600

aaatggtata taatgcttta agtttcccaa tcagcttaaa agcttttcct ataaatcttt 6660
aagattatgc tctggggctc aatactgctt caagaagggc ttttcttttg tatttagaat 6720
tattcacctt tttaaacaaa aggagaaaat ggaatagaaa tatgtttgca acataatttt 6780
atgactatgt gtttatttcg cgtgttctgt gggcctgcag tttgctgctg ttaatgagga 6840
caacagtggc accaatacag tttccactca gattacattc tctgttccct ttctgaaagc 6900
tgccctctcc actgggccca aaagagtcag tatcttaaac aagctgtaca acttagataa 6960
ccatggtctc ttcagactag ttaattgaca tatattaaaa agtaaatagt accaaagtga 7020
atttctgaaa ttaaaaatga acatttaaaa actctaggta aactattcct tagagttaag 7080
tgttttgcca agttctgtaa tcataatatg atagaaacgc tcactcagca ttctaaatat 7140
agaagttact ccttcgcatg acactctaat tcttgataag gtggagaaag agagagagag 7200
agggggagag acagaaaata tggtggttca aggaccattt gagggaatta gttatgttct 7260
tccgtcctct gtggatctta ggggttgaat acagtcattg agctcggtgg atggctgtcc 7320
tgttgaaagg tctgcccagc agagcaaata gacttttta tttacatgga catccgtttg 7380
tgactaatct aatgttcact cccaaagtaa tcacacagac agagaggtag cttccttcag 7440
tactcttacc ttacatgaat cctaccattt tgttatttt tttccacttt aaatctttga 7500
ttatgtgttt ttaattagaa aatttgcata caaatttcca tacagtatgt agaattgact 7560
gtgtttgaat gggtgaagat ccacatgtgt aaccctagct ctggactggc tctgagcttg 7620
tttgctcttc tcttttgtgt tctgagtaac tgaaactctt tcattttagc agcttagtat 7680
gcgcccttca cattgctgtg ctgcctgctg cactaacatt actcctttgc ttatgttccc 7740
cttcctgatt cagtgtcatt ttaagcagta gtactggacc tcagtacctt agccggagct 7800
cactgaggtg acagggctga ggctctgctg ctgtcttttg agcttacctc tttttaatgt 7860
tttatggtat ttctgctgcc aggtttgggg gttttgtttt gttttgtttt ttgttttttg 7920
ttttttttaa ttttctagga acacctagaa aacacaaact aggaaactta aaagagcagc 7980
gtcttgttcc ctgcgttcta gaaagtccaa gcctaatgcc agtgtcatgg ttgtcaggaa 8040
catgagcctc tgaaggcttc ttgggaaacc tttcttgtct caacacctct ggtggcaagc 8100
agtagtccat ggtactctct ctgtccacgg tcagcatccc agtccctgcc ctttatcttt 8160
gtgcagccga ccagctttgc tttagtctgt ctccttctca ggtctccttc cccgctcctc 8220
ttaagcacag cagtcattgg attagagccc atccttccct cggatggccc atttgaccta 8280
attttacgta tttgtaacta aggtcccatt tacttacaca gggccctccc cttcctgttt 8340
tgttctttag ctgaaatggt ttggagacca aatatccaat cattacaatt gtgcacaagc 8400
tatgttcatt tggaggtaat aaaggctcat tctttgcttc tattggtatg tgacattttt 8460
ctaagtcact tggggtttga tagatatctt taaatggctg aacctgatca ctgttctttt 8520
gtatgtccct gtttagctat tgcaagcgtt cggataatgt gagacctgga atgcagtgag 8580
acctgggatg cagggatgtc gactatctgc cccccaccat ctcctgctgt tgccaagaca 8640
gagattgctt taagtggtga atcacccttg ttggcggcta cctttgctta ctgggataat 8700
attcttggtc ctagagtaag gcatatttgg gctccaaaga cagaccaagt gcttctcagt 8760
gatggagaaa taacttttct tgccaaccac actctaaatg gagaaattct tcgaaatgca 8820
gagagtgggg ctatagatgt aaaatttttt gtcttatctg aaaaaggggt aattattgtt 8880
tcattaatct tcgacggaaa ctggaatgga gatcggagca cttatggact atcaattata 8940
ctgccgcaga cagagctgag cttctacctc ccacttcaca gagtgtgtgt tgacaggcta 9000
acacacatta ttcgaaaagg aagaatatgg atgcataagg taaggggctt ttgagcttga 9060
tcatggtagc ctggccaatg aaagtttttt tctggtacag ttacacttaa gttttggaaa 9120
ttatatgctg ctaacaccag acagctgtta tgttgtgtct cctgggcaca gaaagccctg 9180
ctctcatgcc tggggtcttc acagtcctaa tggaaagtaa gatcttataa acattgtgtc 9240
tgagtttgtt ctggaagctg tgactctacc ttcttgtttt cctttccctg tgtgactttg 9300
tcctttgctt acaacagtgc aaaagtataa atattctcag attttgataa gctgtcagcc 9360
acacagcctt agtaactaag ctgctgtccc acgctcccag ttctgtataa cgaggatgga 9420
ccaattagat tctaaggagt tattcctttc aatttgcaaa tttagctaaa ggaaatattg 9480
ttttctcctg atatttacat tgcttttcat tttcagcata tctaaagaac aaacctaatt 9540
ctccttccta ctttctagtt taatataatc ctaaaaatcc attaaaacat gactaattct 9600
ataaggcctc taacctacaa agggaagtag cattttgaaa agaatagttt tctctattat 9660
acctattcat gcagacttcc ttccttattt ctgacatact taacaaaaat catttagatt 9720
caaacagttt agctgcaggt gatattacag acaagtaatc ccagtgctct atctagtctg 9780
aggcaaaagg atttgagctc agtgccagcc tgttctatct acctggtgag ttccagtccc 9840
ataaataaac aaactaaaac aaccgttcct ctgttcctca gatgcgagtc gatcttgttt 9900
gatttaaata gtgtgtaatt attttctttt gaagctgcag gtgttatgtg ggctgtttta 9960
gactaaattc tctctttact gtggagtaaa gggtgctgtg attgtatttc atgttctctg 10020
cgagagcttg aacttgttgg gctaatcgct tgtctccatc ctgtctcccc acctgcgtaa 10080

aaagtatttt cctgtgagct gtacatgata gagcatatct acattgaaaa atgaacgagc 10140
atcaaaatgg atttgttaaa gtaaattttc tttttcttag gaaagacaag aaaatgtcca 10200
gaaaattgtc ttggaaggca cagagaggat ggaagatcag gtacagtgca tatcacatgc 10260
tgcctgtggc aggtcctctt tgcttatgtc ggtataaagt tggtgggtac ttctggtaag 10320
gacctgagga tacattcatt tgacggaagg agcctgaaaa tgagtattct tgttaagctg 10380
tatagaatga actgaataaa aatttctgca gcctaagttt gaattttaaa aaaatttaat 10440
tacatctaca aattagtatt tggccaccct ttttcaatca gcaagaatat gtttgaggtc 10500
atttatttgt agtaaaattg catgcagttt atttatttta ttgaaaatag gtttttaaa 10560
ctatattttc tgattatggt tttccctcct ctgaatcctc ctagaacctc cacctaccca 10620
aatctatatc tgttctttct ctctctcatt aggatacaat caggcatgta aaataatagt 10680
agtagtagta gtaataataa tgtaaaataa gttaaagtaa aaacaaacca gagtaggaca 10740
acataaatag aagtagaaaa gagccaaata agaaattcaa gaaacacata tagacacaga 10800
cacaatattt gcatacacag aaattgcata aaaccgcaag actggaaacc ataatatgta 10860
tgtaaggtgg agtgggaagc cctgacagca cagtgagtaa agcactttca aaaacaccac 10920
tgactttgtg ttgtgttgcc tgtctgctgg gcatgaggcc tggccttaga gagtggtgtg 10980
tatacccagg aagacttaca taaacactta gcttttcatt tgtgacctga tagcaattgg 11040
aaatagtgtc tgggctaggc attccggctt attgccactt cccctcagca ctgaggcccc 11100
atctgaatcg gatccgtgca acccttgtgc atatgcagtt ttaaaagtta tcccttctgc 11160
aactatgctc acaggagttg ccgtcttaag ggagtgagca caccctgag gcatggctcc 11220
aggggtgcag agccagccat aggcacagtt ttttttaaaa ggtttatgtt gtagttttga 11280
aactcaaatt tatgtgtatt tgtggcagat tgtttgaatg ttgaaatttg ccagtaacat 11340
cttttatctt cttcccttta gcctggcatg ccacccaccc tcatttgtcc ttgtcaaact 11400
ccagtaatta aacatggcta tgtggccttt tctctcattt tccttagcat ggctaaggag 11460
aatgggactt aaaaaataat aatcatcatt ttaagtatgt ctgagggttt gaggatatag 11520
tggtagaata tctgcctagc ttccatagct tgatcctaca tttgatccct ggcaaaacac 11580
acacacacac acatatacac acacataaaa tgacttttat aaagttagtg tgctgtgctg 11640
tgatgaacag tgccatagga aatattcttg gaaaagacct gaaactaaat gctctaaaag 11700
gtctaatctt tacttgcttg ctgatcgtta agcagagtct ccaagtataa agtcactttc 11760
accaacctct gcactggatt tctggagtaa ttagggagag tcatttcaat ataagaaaat 11820
ttagtaccaa ataaaatttt cattcagtga aattttgttt ttgaaagtaa gagcccactg 11880
tggtggtttg aatatgcttg gcccagggag tgtcctgtaa gatttttgtt gttgttgaac 11940
tccattgaga cttatgttga caataaatgc ctgagagtcc atgtctaaaa tgctgtacct 12000
gtctgaaccc aacggagata aaacttacca tttctgaaaa ggatgaggtg ttttatttac 12060
atagctgatg taatgtgctt gcaacagctc tattatgaat cttaatacta cttcagtata 12120
tcacagcact tcaggaaatt taacatacat tgtttaattc catgtcttaa ttgtatttgt 12180
aaacagacat ttcagcagtt actctaaaaa gtagaaataa tgagtggttg cttctggtca 12240
ttaggatgaa atattgaaat gataaaattt tctgggctgg agagatggct cagaggttaa 12300
gagcactgac tgctcttcca gagatcctga gttcaattcc cagcaaccac atggtagctc 12360
acaaccatct gtaatgggga tctgatgccc tcttctggtg tgtctgaaga caactacagt 12420
gaactcatac aaataaaaat aaataaatct ttttttaaaa atctatatct gcataggcat 12480
ttctagatta ggataaattt tccaaaggaa ataagcacct ccatgataag ggcattggaa 12540
atgaagcccc cgcccccacc cccggtctgc acgtgtgttg aggatgagat ctagggcctc 12600
cttatacatg ccaggcagct gttctgtcac caagtggaat ataatcctca acccttaatt 12660
tgaggttcta actttaaaat agatgtgagg ggtttaaata atcatttcat gaaacttaaa 12720
tgagcaagtt tattactgag gtgagtataa gtaattgata attttaaata tatttagctg 12780
agattgatag acacttggca atgtcagcat cttatttagg tgatcataaa ctgatgggag 12840
aaatggtaaa tgttaggggg tgtcgctcat gtcacacacc gcagttatgc tgcaaacaag 12900
atgccgggaa atagaaattc aaggtcttgt tttgcgggtg cagactcttc tgtctcactg 12960
attctatgtg gtaacttcag tatgcatttg gatagattat gtcccatttt gaatgtggaa 13020
gctggctgtt gagaggagac ttcctggtga attccttttt ctaagcatta ccatctgtct 13080
tagtcagggt ttctattcct gcacaaacat tatgaccaag aagcacttgg ggaggaaagg 13140
gtttattcag cttacacttc cacactgctg ttcatcacca aggaagtcag gactggaact 13200
taagcaggtc aggaagcagg agctgatgca gaggccacgg agggatgttc tttactggct 13260
tgcttccctg gcttgctcag cctgctgtct tatagaaccc aagactacta gcctagggat 13320
ggcaccaccc acaatgggcc ctccccccct gatcactaat tgagaaaatg ccccacagct 13380
ggatctcatg gaggcatttc ctcaactgaa actcctttct ctgtgataac tccagcctgt 13440
gtcaagttga cacacaaaac cagccagtac aacatctttt cacatttaat ttttctcact 13500
ttaaacgtgg cctttaacaa gcgcttataa aaatgcttaa gcttaaatgt tatttaagct 13560

taatatactt aatatacagc actgtagctt aaatgttgca tgtgagagta tatgataagc 13620
catgctcacc aaggaaaaga agcttaaaga gcataaaaac cctgacagcg gtttctgagt 13680
gggaggctcg gggactgtgc tgagcaattc caaccaaggg tgttttactc tctgcctcca 13740
tttgaaatgt ttttcctgca caacctaccc accctgtgat ttcgttcact cgattatgtt 13800
tgatctaggg tcagagtatc attcccatgc ttactgggga agtcattcct gtaatggagc 13860
tgcttgcatc tatgaaatcc cacagtgttc ctgaagacat tgatgtaagt gtcatgtatc 13920
ttttatgggt tcccttgagt ggtgagtggg tggatgtgtg gtgcatgtgc gtgtgtgtgc 13980
ttgcatactg ggaattgaac ccaagtcctc aggaagagca gccggtgctc ttaagcactg 14040
agccatctct tcagaacctc ttccaccagt ttctttgacc atttgttgag aatattccag 14100
tcacacattt tccgtgagta aatctctcta atgctgattt gtcattaagc tcagtctcct 14160
aattctgata gctaagaagg gtaaattatt aaaaagtgcc ctttactctt cctggccaat 14220
tccccttttgt tcttctgaaa agtgcataga cagcatcact ttatagatca ccttgatgct 14280
cgtgagaggg ctggctcgtg ctggctctag acttcggcac acttattaag agttctccca 14340
acactgtaaa cagactaatt tttatattgt gcattttaga tagctgatac agtgctcaat 14400
gatgatgaca ttggtgacag ctgtcacgaa ggctttcttc tcaagtaaga attttacttc 14460
tttttctgaa tgctaagtaa agcagattaa aaatcttaat gctcacccat gacaagattt 14520
acagggaaaa gatggtagaa aacctacttc ctccaattat ttagggtcaa catggcacat 14580
ttgagcttac acgtgttgtt ctcacccata caacagtggc atatctgaca ttactcttcc 14640
cacagtctaa aaaggcagag tttccgtagt acccagggaa gttctggtct gtgtttgggt 14700
ctggtttctt ctttcaattc tcactaagta taacccttag gaatctatca agttgagttg 14760
cattttaaat tcctgtgaat tcttcaggtc tagaaatgga aatcattcat attttagact 14820
gacatttttc atcttcttgt gtaatttaac atttaagaac ttgagctcta atatcagact 14880
gtctaggtta caactgggaa aacttggtga agctacccaa agctgaacct ccattttctt 14940
acctgtgaaa tgtgaacagt gataacagct agtttcttgg gtccttgtag gcaccaaatg 15000
acaggataat ataaagcacc taggacagtg gagccaatga gccaggagcc agtgtgccct 15060
attatatctg ctctaagaaa gacagtaagt ggaatagcca atactgactg tcttagtcag 15120
gctttctatt cctgaacaaa aaacatcatg accaagaagc aagctgggga ggaaagggtt 15180
tattcagctt acacttccac gttgctgttc ctcaccaaag gaagtcagga ctggaactca 15240
gatcaggaaa caggagcaga tgcagaggcc atggaggaat gttacttact agcttgcttt 15300
cttatagacc ccaagactac cagcccagag atggtcccac ccacaaggga ccctgccccc 15360
ttgatcacta attgagaaaa tgccccacag ctggatctca tggaggcatt tccccaactg 15420
aaactccttt ctctatgata actccagcct gtttcaagtt gacacaaaac cagccagtac 15480
gctgaccgag cagctgtgtg ttcctctgca gggctgtgtt ctctgtttgt ccctcatctc 15540
ctgttgtagt ctcctttaca gttacagact gtcatcagta acgagagaga agtgaatagg 15600
attttgttaa agtgtttact tctatgtcac attcccttcc tataataagc tcacagtgaa 15660
ataccaggtg accgtgctta acggcatcta ttacctaact ggggtatctt tttccttaaa 15720
atggatttaa ttttatgtgt gtttgaatac ctgcatatgt gtatgtacac catatttatg 15780
tatgcctggt acctgaaaaa gggaaaagag ggctttggct ttcttgaaac tagatggttg 15840
tgagtctcca tgtgggttct ggattgtctc tgcaagagcg gcaggcacac tttagcagtg 15900
agccgctcct gtcccgagtt gtcttaagac ctgtgaaagg tccctaaaaa atgcagggtt 15960
ttacccgaat aaaagatgac atcatgcaga tggctttggt gttcatcaag ctcttgtgtg 16020
ttgtcctaac cttgctgggc tttgtcgttg tgaagctgta actccgtcaa tgttttcctt 16080
acctacagtg ccatcagctc acacctgcag acctgtggct gttccgttgt agttggcagc 16140
agtgcagaga aagtaaataa ggtaattcgt tctacagttg aacatgatct gacttttatc 16200
atcactagca tatcatacat tatcatctaa acagtaggct gcaattgaaa taaccccata 16260
gtataaggaa gcaatgtaat tttaccaaat ttctctgaca ccctctagca gaactgactc 16320
taatagaatg agtaagaatt caattaccaa attaattttg atactctttt ttatttttgt 16380
tattactttt ttatttttatt ttaattaggt attttcttca tttacatttc caatgctatc 16440
ccaaaagttt cccataccct cccacccact cccactcccc tatccaccca ctccccttttg 16500
gccttggcgt tcacctgtac tgagacatat aaaatttgca agaccaatgg gcctctcttt 16560
ccaatgatgg ccaactagac catcttctga tacatatgca gctagagaca cgagctccag 16620
ggggtactgg ttagttcata ttgttgttcc acctaaaggg ttgcagaccc ctttagctcc 16680
ttaggtactt tctctagctc ctccattggg ggccctgtga tccatccaat agctgactgt 16740
gagcatccac ttctctgttt gctaggcccc agcatagcct cacaagagac agctatatca 16800
gggtcctttt agcaaaatct tgctagtgtg tgcaatggtg tcagcgtttg gaagctgatt 16860
atgagatgga tccccaggat ggcagtatct agatcgtcca tcctttcgtc tcagttccaa 16920
actttgtctc tgtaactcct tccatgggtg ttttgttccc aattctaaga agggacaaag 16980
tgtccacact ttggttttca ttcttcttga atttcatgtg ttttgcaaat tgtatcttat 17040

atcttgggta tcctaagttt ctgggctaat atccacttat cagtgagtac atattgtgtg 17100
agttcctttg tgattgggtt acctcactca ggatgatgcc ctccaagtcc atccatttgc 17160
ctaggaattt cataaattca ttctttttaa tagctgagta gtactccatt gtataaatgt 17220
accacatttt ctgtatccat tcctctgttg aaggacatct gggttctttc cagcttctgg 17280
ctattataaa taaggctgct atgaacatag tggagcatgt gaccttctta ccggttggaa 17340
catcttctgg atatatgccc aggagaggta ttgtgggatc ctccggtagt actatgtcca 17400
attttctgag gaacggccag actgatttcc agagtggttg tacaagcttg caattccacg 17460
aacaatggag gagtattcct atttctccac atcctcgcca gcatctgctg tcacctgaat 17520
ttttcatcgt agccattctg actggtgtga ggtggaatct cagggttgtt ttgatttgca 17580
tttacctgat gattaaggat gctgagtttt tttttcaggt gcttctctgc cattcggtat 17640
tcctcaggtg agaattcttg gtttagctct gagccccatt tttaatgggg ttatttgatt 17700
ttctggagtc caccttcttg agttctttat atatattgga tattagtccc ctatctgatt 17760
taggataggt aaagatcctt tccaaatctg ttggtgacct ttttgtctta ttgatggtgt 17820
cttttgcctt acagaagctt tgcaatttta tgaggtacca tttgtcgatt ctcgctctta 17880
cagcacaagc cattgatgtt ctattcagga atttttcccc tgagccaata tcttcgaggc 17940
tgttccccac tctctcctct ataagcttca ctgtctctgg ttttatgtgg agttccttga 18000
tccacatgga tttgacatta gtacaaggaa ataggaatgg attaatttgc attcttctac 18060
atgatatccg ccagttgtgc tagcaccatt tgttgaaaat gctttttcc actggatggt 18120
tttagctccc ttgtcaaaga tcaagtgacc ataggtgtgt gggttcattt ctgggtcttc 18180
aattctattc cattggtcta cttgtctgta tataccacta ccatgcagtt tttatcacaa 18240
ttgccctgta gtacagcttt aggtcaggca tggtgattcc accagaggat cttttatcct 18300
tgagaagagt ttttgctatc ctaggttttt tgttattcca gatgaatttg catattgccc 18360
tttctaattc gttgaagaat tgagttggaa ttttgatggg gattgcattg aatctgtaga 18420
ttgcttttgg caagatagcc atttttacaa tgttgatcct gccaatccat gagcatggga 18480
gatctttcca tcttctgaga tcttctttaa tttctttctt cagagacttt aagttcttgt 18540
catacagatc tttcacttcc ttagagtcac gccaaggtat tttatattat ttgtgactat 18600
tgagaagggt gttgttttcc taatttcttt ctcagcctgt ttatcctttg tatagagaaa 18660
ggccattact tgtttgagtt aattttatat ccagctactt cattgaagct gtttatcaga 18720
tttaggagtt ctctggtgga attcttaggg tcacttatat atactaccat atcatctgca 18780
aaaagtgata ttttgacttc ttcctttcca atttgtatcc ccttgatctc ctcttgttat 18840
cgaattgctc tggctaagac ttcaagtaca gtgttgaata gggaggaaga aagtggacag 18900
ccttgtctag tccctgattt tagtggggtt gcttccagct tctcaccatt tactttgatg 18960
ttggctactg gtttgctgta gattgctttt atcatgttta ggtatgggcc ttgaattcct 19020
gatctttcca agacttttat catgaatggg tgttggattt tgacaaatgc tttctcctca 19080
tctaacgaga tgatcatgtg gtttttgtct ttgagtttat ataatggatt acattgatgg 19140
atttccgtat attgaaccat ctctgcatcc ctggaataaa acctacttgg tcaggatgga 19200
tgattgtttt gatgagttct tggattcagt tagtgagaat tttactgagt atttttgcat 19260
caatattcat aagggaaatt ggtctgaagt tctctatctt tgttggttct ttctgtggtt 19320
taggtatcag agtaattgtg gcttcataga atgagttggg tagagtacct tctgcttctg 19380
ttttgtggaa tagtttgtga agaactggaa ttagatcttc tttgaaggtc tgatagaact 19440
ctgcactaaa cccatctggt cctgggattt tttttggttg ggagactatt aatgactgct 19500
tctatttctt taggggatat aggactgttt agatcattaa cctgatcttg atttaacttt 19560
ggtacctggt atctgtctag aaacttgtcc atttcatcca ggttctccag ttttgttgag 19620
tatagccttt tgtagaagga tctgatggtg ttttggattt cttcaggatc tgttgttatg 19680
tctccctttt catttctgat tttgttaatt agaatacttt ccctgtggcc tctagtgagt 19740
ctggctaagg gtttatctat cttgttgatt ttctctaaga accagctcct tgattggttg 19800
attctttgaa tagttcttct tgtttccact tggttgattt cacccctgag tttgattgtt 19860
tcctgccgtc tactcctctt gggtgaattt gcttcttttt gttctagagc ttttaggtgt 19920
gttgtcaagc tgctaatgtg tgctctctct agtttccttt tggaggcact cagagctatg 19980
agttttcccc ttagaaatgc tatcattgtg tcccataagt ttgggtatgt agtggcttca 20040
ttttcattaa actccaaaaa gtccttaatt tctttcttca ttccttcctt gaccaaggta 20100
tcattgagaa gactgttgtt cagtttccac gtgaatgttg gctttctatt atttattttg 20160
ttattgaaga tcagccttag tccatggtga tctgatagga tgcatgggac aatttcaata 20220
tttttgtata tgttgaggct tgtttttctg accaattatg tggtcaattt tggagaaggt 20280
accatgaggt gctgagaaga aggtatatcc ttttgtttta ggataaaatg ttctgtagat 20340
atctgtcaga tccatttgtt tcattacttc tgttagtttc actgtgtccc tgtttagttt 20400
ctgtttccac gatctgtcca ttggtgaaag tggccatctt tatagtcact gaagacatac 20460
aaatacatat tcatatcaac tggaacaaac ctaatttctt tttaaatgtt ttacatggaa 20520

ataagttagg ggttgttatt tgcattacaa agttactcat ccctttcctt cttttctttt 20580
ttttttttt ttttttttg agaacaagcc tgtgtactta tatgaacttt aatttgccaa 20640
attcataatt cttattcaat catttatgac agaatgctaa aactctcatt atattttagc 20700
taggcattta gagctgttat gtgtaacccc aaaaagtagc tttccacttg agatgctgaa 20760
ggccttgggt tccgtgggct gtcatcatgg ttggctgtat gaaaagagaa aggctccatt 20820
gtttgggcat cacttaaata ttttttcacc tttcatcttc ttttaggtta agtagcttgt 20880
ccttgatcat ttcatttttg agagacaact tgccactact ctagttgaaa agtgctgtct 20940
tgacgctgtc tctggctgtg gtcagagtcc agcagagctg cacagctggt tacctttctc 21000
tgtacagctc taggccaact cttcttactg gcgaccattt ctaaatccac cattcacttg 21060
ttccccatga aagtgagtag ggtttttttct gtggaagatt ttgggcagtc ctgttgccac 21120
tttgcatcag acaatagttc cctcattgaa acacgcagtt tattctccag agcggtctgc 21180
ccactccaaa ggcagtaggt gctgggtaga gatatgccaa gtatcacact aggctatgac 21240
tgctcactca gatcactcgg atgaagcttt catggccaaa tacagttgag aaagaacaaa 21300
tattcttcac ttagagagca acaagagtta ttcaagtgta acaagttctg agattccatg 21360
cagttgattt accagctact tcctaaactt aactggccac aaaatccctt tgtaagcagt 21420
atgttgtttt gacccatgcc ctgtcaaagg atactcctta cttgggaact gttttaatga 21480
tggcaacaaa aatttctatt taaatttatt tcataagcaa gcaaagatct ttttacttca 21540
cattccaatg ttgactcttt tcctctagat agtaagaacg ctgtgccttt ttctgacacc 21600
agcagagagg aaatgctcca ggctgtgtga agcagaatcg tcctttaagt acgaatcggg 21660
actctttgtg caaggcttgc taaaggtaca cttgccgatc atttatcatg tgtgacgcaa 21720
caagtagaga tggagggtac aaataatcac tgagaggctt tggaaagtat attgttagca 21780
tttaatgtct catagtttta gttgtctggg tactggtttg ttttcatcat tctgagcatg 21840
aagtgtatgt cttagggatt tatagttcgt atcatgtatg aaacaccatg gggtaatatt 21900
tatatttcac ttggttccct ctagctatgt gtctggcccc agtgctttcc ttgtaaatgc 21960
atgcttgaat cagactgagc tgatatgata atgttgatgc tccttttgct tactgagtgg 22020
ctatgaatat gcaccatact tactcattgt aagaaattaa aatgtctctt aaggatgtaa 22080
acatagcaaa atgaagcaaa acaaaagcga tgctgtttta ggtaccctaa ctgaccttgt 22140
gtattcaagg agcattccta cttctgtgat gcaaaagctg tctacactgg gcagatctac 22200
aaccagcatt aaaccaaata gggaatcact gaaatcacgt tatcaaagat gagaaacaag 22260
ataataatgt ctactttcac ggcttttatt caggtctagt gctataagtt tttgccaaaa 22320
caaaaatgaa aacatagact ctgggctgag gctttccctt agcagaaaag tgcttacttg 22380
ttgtgtccgg ccagcagatc acagcctggg ttctagcctg gaaaggcatt ttggaaacct 22440
ggaagagaag aggggctagg taacgagaga aagaacggag ccaagtcaaa agcaactctg 22500
atcaaagctc aattttacta tatcagcacg cagttataaa ggaggggaag ggggggccaa 22560
tagcaaggcg gcaggttcca gcagtgggcg tggcagaccg attgagccgg caagctcctt 22620
ccaggtgtaa acagtggagc cctaaggctg ggggagggga ggctacactt agcatgcctg 22680
atgccctaga tgccacctaa atgacaaatc cagtccagta caggatgtag agcacccccc 22740
cccaaaaaat tatttttttt gtataccaga aatgaaattg ctgagaaaaa aaaatgaaga 22800
ccataattat actcccagta gctacaaact aaacagcccc atagatgaag tgagtgatgt 22860
ctgctgtgac aattatgaaa tgaaagaagt aaagatgaac aaatgaaggg aagacatcca 22920
gtactcagga ctgaaagact gctgctaaaa tgcctatcca acccagagct ctctgcagac 22980
tctggacaga tccgctctag atgtgaagat ggtctttttt ttttttttt ttttttggtt 23040
tttcgagaca gggtttctct gtgtagccct ggctgtccag gaactcactc tgtagaccag 23100
gctggcctcg aactcagaaa tccgcctgcc tctgcctccc gagtgctggg attaaaggcg 23160
tgcaccacca tacctggctt tttgtgaaga tgttcttaac agaactagaa agaagtaccc 23220
cttggtttgc tgcccttctg atgcagtatc cccaaaggct cgcatgcact gaacatttca 23280
tcttacctgg tgccactgtt gggaagtgat ggaaatgcga ggaattgtag cctcgttgag 23340
atgtttctca ttaaggcact gggggcatac ctatggagca tacagtagga acctggtttg 23400
caacctctcc cctctccatc caggctctcc cctgtgcacc tggccttggt gttctgccac 23460
tccatgaacc caaagtaaag tggactatgc ccttagactg taacagtgag tcagaagaaa 23520
catttcctct ttaaagctga gttttctggg tgctttgtca tgttaatgga gtctgattag 23580
tacagaccct gagtaggcag ggcaatctta tgcagaaaca tcaaagctgg tagcatagac 23640
atacctaatt tcacaataga cactgatgga ctcagtctgg agtacttaca gtaagaatat 23700
acagcagaga tacggagctc tcttacagtg gtgctctggg agaactggcc gtcctgtgaa 23760
gaaaagccag agtggctcat tctcaccaga cacaaactga gctcataaga cgcttgaacc 23820
tgagatcctg gtcagcagcc actagaagaa aacttaggag aaaccattca acacgtcagt 23880
ctggggaaaa gggtggtttt ggttttggtt ttggtttttt agtatattcc ccaaatcaaa 23940
aacaacaaaa cccaaacttg acagatgaca tcacactgca aagcttttgc acaaccaaga 24000

aagcaacctg cagagtgcag taataaccca cagaaggaga ggagatactt gtgggcagtt 24060
catcacacag gtcaatataa gcaagtactg atagtgtggc catctccaaa gaagatatga 24120
aaataactgg tatatatgaa gtagtactta gcattgctgc gtatatggta aattcaaaac 24180
catgatgaga tattgcccca cttagatgga tattatcaaa acaacatcaa aaagtgacaa 24240
atgctttcaa ggatatgggg aaagtgtact tgcaggaatt taaattatta atttgccatt 24300
caagaggata ggatggcagt ttaaattaaa aaactagaag tggtagagca gtcgcctaga 24360
acatacaagg ttcagcacta taataaatga gcaattagac atttgaagca acaatctcac 24420
cactaggcaa gtcctaaaag aaatggactc gcttcttctt cttcgggaaa acaccaaatg 24480
gcagatgacg ccggtgcagc gggagggccc agaggacctg ggggctcagg attaggaggc 24540
cgcggcggct tccacggagg attcggcagc ggtcttaggg gccgtggtcg tggccgaggc 24600
cgtggccgtg gtcgaggccg cggggctcgt ggaggtaaag ctgaagacaa ggagtggatc 24660
cccgtcacca agctgggccg cctggttaag gacatgaaga tcaagtcctt ggaggagatc 24720
tacctgttct ccctgcgcat taaggagtct gagatcattg atttcttcct gggtgcgtcc 24780
ctaaaggatg aggttctgaa aatcatgcca gtgcagaagc agactcgggc tggccagcgg 24840
accaggttca aggctttcgt cgctattggg gactacaatg gtcacgttgg tcttggtgtt 24900
aagtgctcca aggaggttgc tactgccatc cgaggggcca tcatcttggc caagctttcc 24960
atcgtccctg tgcggagagg ctactggggg aacaagattg gcaagcccca cactgttcca 25020
tgcaaggtga caggccgctg tggctctgtg ctggtgcgtc tcatccctgc ccccagaggc 25080
actggcattg tctctgctcc tgaagctcct gatgatggcc ggtatagatg actgctacac 25140
ttcagccaga ggctgcactg ccaccctggg caactttgct aaggccacct ttgatgccat 25200
ctccaagact tacagctacc tgacccccga cctctggaaa gagactgtct tcaccaagtc 25260
tccttatcag gaattcacgg atcatcttgt gaaaacccac accagagtct ctgttcagag 25320
gacccaggct ccagctgtgg ctaccacata agggttttta tatgagaaaa ataaaagaat 25380
taagtctgct gaaaaaaaaa aaaaaagaaa gaaagaaaga aaagaaatgg actcggtatg 25440
tggatgaagc ccaggcacct tcatctgtgt tgcagcacga gtcaccatgc aggatcagtc 25500
taaacgccca tgcacaaatg aatggtacat agccacagtg aagtgtttga ccacaaaaag 25560
gaaagtcagt tgtgataagt gaaacaagcc aggcacagaa agataaatgc tgcatgttat 25620
cattatgtgt aaaggctaaa acgtttatct catacaagta gaaggtaaat acggagacta 25680
ccagaactta taaagagttc taggaaaaag ctatagagag gctcagggtt gaataactaa 25740
aattatacct aaaataacta aaaggatagc ttacaatatt ctgtagcact gtagaataat 25800
tgtgacagtt tgttgtattt ttctggtttg tgtatgtggg agagaaagta tgtggacaga 25860
ggttgatatc aagtgtctga ctctgcactg cattatttta ggcagggtct ctctctaacc 25920
attgaatgga ctggctaggc agtggtgccc taacatctac ctgtccgtac atctcccaat 25980
actaggttat aagtacactg ggttttaagt acaggctata ggtatagata taggctacag 26040
gtatagatat aggctgctgc aactgattac atgggtgctg ggaacctaac ataggttggg 26100
tcctcatgtt tacacagaaa tcagtactgt gcctactgag tcatttcccc agttctagta 26160
tttgtttttt aaatagctag taattggaat tgtgaatgtt cctaacaaaa gaaaatgata 26220
actatctgag atgctagtta tgataccctg agtgaatcac actttgtgtg catgtactga 26280
aattcattgt accctgaaaa tacaaaaatt gctctgtgtt gattggctag atgcatgtgt 26340
attagtcagc aatctctaga gtaataaaac ttagatatat gggatgtatt agacttttgg 26400
ccttacaggc caagatccag ctaatccatc agtggcaggc tgtgaacagt aagtctaaga 26460
atccaatagt tgttcagtcc acaaggccgg gtggctcagc tgccttctgt atacagtgga 26520
atcccaaaga aataggcgcc aaagctagtg aggaatggtc ttgctagcaa agcgaaggtg 26580
aaggtaatca ggcagaagac aagaccttcc tttttccgtg tccttatata ggctcctagc 26640
agaacaagtg gcccagacta gatgtggatt aaatgttttg ggtttggttt ggtttgattt 26700
ggtttggttt ggtttggttt ggtttggttt ggtttggttt ggtttggctt ttcgagacag 26760
ggtttctctg tatagccctg gctgtcctgg gttgtagacc aggctggcct caaactcaga 26820
aatctcttgc ctctgcttcc caagtgctgg gattaaaggc gtgcacacca ctacgcccgg 26880
ctcaatagca ttaaatggca tgtcttttcc tatctcaaat gatctggatt aaaagagtgt 26940
cttcctacct caaaggtctg gattagaagt ggatctttct acttcagatt aagttaaact 27000
ctctcacagg tgtgccctct acttttggat ttttggttct agatggagtc aacatgacaa 27060
ccaaaagtaa ctattacaag tccacccaat atcaacttga tacacaatca tatctcctta 27120
tgtcataatt aatttccaaa tgaaaacaat aaccatgtca taaaaacacc taaacatgaa 27180
taactattcc acatacaatc agaaatgcat tcattatata tttaaccaag tcctaattat 27240
gcctaacgtg atataactat tcttcataca acagcaaaca tgataaattt acaataggtg 27300
gcaatgtctt attcttttaa tatctcaaac ttaaatatga taaccattga tgttatctta 27360
attgatgtta tatcatatga taaagaaatt gatgaaagaa agcacaaatg tctgtataaa 27420
tgctttctta agaaaatagg acagaaactc tgtcaattat aatcatcttt tctgcaacta 27480

gtcatgtggc cttagtattt ataactacct tcctctgcta aaccattttg tattttctcc 27540
acccttggca agaacctcag caggtcttgg ctcttttcct ggaggagtga cccatacctt 27600
cattccttac atgtatgtgc cctttgtcat cctgcctgga ccaggttgtt gtaacattga 27660
ctttaatcac aggacatcgt agcaccaaca catgccccaa aggatctcct gccctataga 27720
cataaccttt cttacctcca tagtggggag gcagtcccag tcctccttgg tagtctgcat 27780
cagtcacgcc tcctaacact gttattcctt tcttagccgg ttgacttaag ggcatcagaa 27840
ggccaaagtt gccagaggaa aatctgagct tccagttcaa tgaatgtaat gttgttctag 27900
gcaagcagaa ctgaaggtct caggaatagg aagcaaacac ttcccatgga tcactacagg 27960
gtgagagtga gtagaattat tctcttttct accacttgac tcctggacct atggatcctg 28020
gtatcaaaga aaatgtctca tatattgtac actgattcag agcatgcctt ctggaaaacc 28080
ctgccccagc ccttcatact gctgccatca aattgtcacc tgtgtcttcc tggtaccaac 28140
ttttgtcctg gttagggtta ctattgctgt gaggaaacac catgagcacc aaagcaactt 28200
ggggagaaat gggtttattc agcttatgcg tctacatcac agctcatcat caaaggaagt 28260
cagaacagga gctcaagcag ggcaggaatc tggaggccgt ggaggaaagc tgctgactgg 28320
ctcgctccct aggcttgctc agactgctta tagaactcag gaccaccagc tccagggtgg 28380
ccccaccccg caatggattg ggccctccct caggaatcac aattgcccca cagacttacc 28440
tacagcctag gcattttgga ggctttgagt ctgcctcctc tctgatgatt ctagcttttg 28500
tcaagttgaa gcaaaagtag acaggcctta aactcacaac aacccacctg cctcaatttt 28560
ctgagtgcta atattatatc aatttaaaat ttaaatataa catataaagg gcaatagaaa 28620
ggactagatt catgtaatgg atacaagtta tggaagatgt gtgtgtgtgt gtctgtctgt 28680
ctgtgtgtgt gtttctagtt taattctgtc atgatttttt tcttgtaggt ggtaggtgag 28740
tgcatggaat acatttgata ctgaaagggt aaattgaatg tggagcctca cagcttctgt 28800
tccacatgcc tatgataacc gtagaaattc atggattagt atagacgttg agtctggtta 28860
attttggtgt gtgatattta tatatatatg tatatatata tgtgtgtgta tgtatgtatg 28920
tatgtatata tatatatgta tatgtgtata tatatatata tatatatata tatgcaagat 28980
ttcttataat taagtttaca aaattaaaaa ctatcttaaa aattgaattc ttgcaaataa 29040
aaatttagct tttggtgatt ggattcttaa tatggttgat gtttacctag aaagttaaaa 29100
gccctgagtt cagtctccac tttcaccccc aaaatgaaaa tcagcttttg ggtttcagat 29160
catgagctca gaattaaaga aaacacattt ctaactttgc ttttacaaat cttaatttta 29220
ccaatttcct ttaaagtcac aatgagatac acagtacttc ctagcacccc ttgttcaatt 29280
agataatgtg atttctgaaa gagctccctc tacacagggc acagggcagg tgcaaaactg 29340
tgattgggtg aaatacctgc gagctctcca agcaaagcca ggcctatttg ctttagctgc 29400
cacatcgggt tcttagaccc gacatccctt cccacctgta tcctccctaa ttccttccaa 29460
ccccacaaca ctaggtagga gagaaagaag gttagtggtg gaagtttgca cacatctttt 29520
tagactattt cctactgatt aggggtgtta ggtccttgag acaagtccag tcttcattgt 29580
caggatatct ccaacttctt cttctcatct ctttgctcac aaagtttatc acaagttgat 29640
aaactacaac aacaggaacc agcagtagca aggacatcag agttgtatag ctttccagaa 29700
aatactttga tatacagtaa ttatcctagc ctttaagagt gaaagatttg gcagcctctg 29760
tgttctacac tcagcataat accttgtata ctgcaggtat ttgctgcatg gtaagtggct 29820
gcccagctac ctagaaagag gtaaatactt ttctattaac atacatattc atttagatat 29880
aggaagaaga taaaacaatg gagaaaggca gtcataattt tacagaccag caagtaaacg 29940
cattaacttg gcataggtct ttgtagtctt tttctgcagt gcgtatttcc tgcagtgccc 30000
acaccctaca gttggattgc acgtggcatg ttctgaccca ctttttatgg tatactgtgt 30060
actgtcactg tcaacacaaa tggtagtggc tggattttta tacagtatca gcttgaaggt 30120
tatttctgaa caagccctgt accagattca caggaatatg catctcttat cattactata 30180
ttcttttaac aattgcttct ctcagttggc atgtggtcag tgagttctct cttccttctg 30240
acaggatgca acaggcagtt ttgtcctacc cttccggcaa gttatgtatg ccccgtaccc 30300
caccacgcac attgatgtgg atgtcaacac tgtcaagcag atgccaccgt gtcatgaaca 30360
tatttataat caacgcagat acatgaggtc agagctgaca gccttctgga gggcaacttc 30420
agaagaggac atggcgcagg acaccatcat ctacacagat gagagcttca ctcctgattt 30480
gtatgtgacg cttggcctta ggtgtcattg ttaaacaaca taaaacttct catttatgag 30540
taaaaacagt gcaagttgta tttaaaagaa aagaaatatg acaagcacat actcaggcac 30600
tttttcttta ttttcttaac tttaaggttt ttttttttt aagatttatt tattattata 30660
tctaagtaca ctgtagctgt cttcagacac accagaagag ggcgtcagat ctcattacaa 30720
atggttgtga gccaccatgt ggttgctggg atttgaactc aggacctttg gaagagcagt 30780
cagtgctctt acctgctgag ccatcttgcc acccccaact ttaaattttt tatactatta 30840
tttttagaca gtctcactgg gcctaatgac ttacataggt ggcctggaac tcactatata 30900
gatcaggcta gccttcaact cccagatatc cacctgcctc tgccacccaa atacttggat 30960

taaaggcgtg tgcctccata cctagcctaa atcttcattt cttaaaatac tgttttgcta 31020
agataggtaa agatttcctc ttaaaaataa atacttagca aatatatacc gatctcctaa 31080
ttacttaatg aagggccagc ttaatagtta tcagtcagtt atcagtgcca gcccctactg 31140
ctgggaattt agtgtataac gttcattgta tggtagactg aagtaattct aagtattttt 31200
ttcttgggtg tgactatcaa acacagaaaa gtatttgaaa tttataaaga gaacaggttt 31260
tttctttgca ttttatattt tgctatttat ttcttaccag aagatgcgag cagcaaagta 31320
aaaggcagta agtgctgatg ggtttggagg aacttgggat tttaattata aaacttcaag 31380
aaagcatttc aatggtgttc tagagtctaa aaaagaatag tgagacccta ttcctgttct 31440
ctccgatcaa ccaagagctt gaaatggtgc tagtccttag tatacactga aaagacgcta 31500
agtgtggtca tcccggttgg agggctttag gaagcagtga ccctggacca atgggtgtca 31560
ccgtgtgtct gaagaagaaa gcagagctga aacaagaggc gcatggtagg gacaccagca 31620
gccacagtaa actgctgccc agaggtccct gtgtggggct gcagaattaa aagaacccat 31680
tctacacagc tctgctgtgc tctgttagtg ctgagaaagg ttgagaggaa ttgtttcaga 31740
agaggaatcg ttcaaattga actcttatgt cactagttca catactggca atcttggaaa 31800
acatagaaat tttctcactg agtctgcgtg cctgcgtctt cctcgtgact aatatacttg 31860
aagtcctgtt tatttttta gttgattgtt tagaatctct tctcaggaaa tgaggtaaac 31920
ttgaatggat ttgcaccatg ttagtgtttt tgttttgaat atgtttgttt ggaagatttg 31980
aagaaaaagc aattgttcag ctattctggc atgacaaaat catgtcatga attttagaat 32040
tttatttcca gttctaagta aatgttttga atataaaatt gtcagaaata ttttcagcca 32100
caagattata tcttctatta ttgtgggctc atgatagtat cagtgtggtt taaataatat 32160
tcacttttga gtctgggagg tttgaggttt cagattcagg gactcacaca ctgggcaatt 32220
actgtaccac tatgcagttg cttattagta ccacagagta attcccagtt aagttacttt 32280
taattttaac ctttttaaga taaaagcagt ctgatgatac attaaagtcg gacatttcct 32340
tgaagatagt ctttcctttt ccagcttttg tgatccagat ctcattcagt aaagcagaaa 32400
ttgggaaata gtggacttaa gttctaaggg acccacaaac cccgtgactg tgctgtccgt 32460
tttcagccag taaccatgaa gtgctggcgt cccttccagc gccccttttct ccatttggtg 32520
cactcatccc tcaaggctga gaggcgtgct gctctcctgt ctatttccct cttccccatg 32580
gttcctgggc agtgatgttg tgatctctac catctgagtc ttgctttgca tttatcttac 32640
tgtgaaaaat gttatatttt ccctctgaca tgaatataat agcctaggga aagacagaag 32700
taaaacactg aaagggaatg ggggctgaga aaaaaacagt cattagcttc tgtctggcca 32760
gcatgctgaa gtgggtcacc tcagttggcc attttgtctg aacgttacat gccagccaac 32820
cttagctgcg gtagtaataa gttatgctgc tggctcatac ttacagatgg taagtctctt 32880
gacctgaggc aaacgtgtaa ggtgacggtt ctaaacacac tgatggacag gcacatgccc 32940
tgcctggata gcctcaaaac acaaacagtg tacaaatgta cccttgcgtt aaagtggatc 33000
tatgtgcgtt tgtgtttatt ttctgtgcat taagtatgta tatgtatgtg tgtttatatt 33060
gtgcacattg agtatatgca tgtgtgttta cactgaatac tgaacccacg gcctcctgca 33120
aactaagtat gcattccaaa tgcacacatc tgtcttctta cacatctgtt tataaaactt 33180
caacttttt actagagcaa gaagttgtgg aatgtaactc tgtaaaaccg tttaatatct 33240
gaacctttt cttcttagga atattttcca agatgtctta cacagagaca ctctagtgaa 33300
agccttcctg gatcaggtaa atatgatgcc acccattgcc agacaaaaga acatcatata 33360
ttttctttta aaatatgtcc cacagtgcct acagaatata taaaaagcac caaagaatta 33420
aagtgctaga ggccttttcta aagtctgtaa acggattcct ctttgaatta ttaatgggaa 33480
atagcctgta tattaaccgt taaagcagca ttctccatcc tagtggctgc ttcaggtcca 33540
accctctgcc tttagaattt ttgtggttgg tgaagacagg ggtgtgcttt catttgtgtt 33600
aattgaattg aaaatattct taaaacttag gttgcttctg cttaaatggt agcatcctta 33660
ttgtctctgt ttttaaaagt atctgatgag taaacatctg gagatggtac tggattctat 33720
gcgacttgtt tctatacgta agcagagctt tgtcataata gcatgctggg aatcaggcca 33780
agatcctgtg ccatagacat agagttgaga tgaggagaac ctcgtgttca ctgggacttg 33840
tgggtctggg tctgtgtgag gtgaggacag cctgtaatcc caagtctctg aagctgaaaa 33900
gtcccctcct ctactccaca caacctgaag tcattgactt agttatttcc ataataaaat 33960
aaggagatat tttaaggtag aatacaagat ctaagtgcat taaactaggg aatctgaaaa 34020
ggggacagtg ggtttccaga catttgccgc taccagagtc ttgccctttg gaaatcggaa 34080
gaaatggctg taatgggtgt tgtgtgtcag atcctgtcaa caatgtcgcg gaagctgcac 34140
tgtcttgtgt ccctgcaggt cttccatttg aagcctggcc tgtctctcag gagtactttc 34200
cttgcacagt tcctcctcat tcttcacaga aaagccttga cactaatcaa gtacatcgag 34260
gatgatacgt gagtcctgct cctctagagg aaagccttta tgcattgaca gttgctgttc 34320
gttccctttg aacattgtct gtattataat gcgggggttt ttgtctcttt tgttttgttt 34380
ataggcagaa ggggaaaaag ccctttaagt ctcttcggaa cctgaagata gatcttgatt 34440

taacagcaga gggcgatctt aacataataa tggctctagc tgagaaaatt aagccaggcc 34500
tacactcttt catctttggg agacctttct acactagtgt acaagaacgt gatgttctaa 34560
tgacctttg accgtgtggt ttgctgtgtc tgtctcttca cagtcacacc tgctgttaca 34620
gtgtctcagc agtgtgtggg cacatccttc ctcccgagtc ctgctgcagg acagggtaca 34680
ctacacttgt cagtagaagt ctgtacctga tgtcaggtgc atcgttacag tgaatgactc 34740
ttcctagaat agatgtactc ttttagggcc ttatgtttac aattatccta agtactattg 34800
ctgtctttta aagatatgaa tgatggaata tacacttgac cataactgct gattggtttt 34860
ttgttttgtt ttgtttgttt tcttggaaac ttatgattcc tggtttacat gtaccacact 34920
gaaaccctcg ttagctttac agataaagtg tgagttgact tcctgcccct ctgtgttctg 34980
tggtatgtcc gattacttct gccacagcta aacattagag catttaaagt ttgcagttcc 35040
tcagaaagga acttagtctg actacagatt agttcttgag agaagacact gatagggcag 35100
agctgtaggt gaaatcagtt gttagccctt cctttataga cgtagtcctt cagattcggt 35160
ctgtacagaa atgccgaggg gtcatgcatg ggccctgagt atcgtgacct gtgacaagtt 35220
ttttgttggt ttattgtagt tctgtcaaag aaagtggcat ttgtttttat aattgttgcc 35280
aacttttaag gttaattttc attatttttg agccgaatta aaatgcgcac ctcctgtgcc 35340
tttcccaatc ttggaaaata taatttcttg gcagagggtc agatttcagg gcccagtcac 35400
tttcatctga ccacccttg cacggctgcc gtgtgcctgg cttagattag aagtccttgt 35460
taagtatgtc agagtacatt cgctgataag atctttgaag agcagggaag cgtcttgcct 35520
ctttcctttg gtttctgcct gtactctggt gtttcccgtg tcacctgcat cataggaaca 35580
gcagagaaat ctgacccagt gctatttttc taggtgctac tatggcaaac tcaagtggtc 35640
tgtttctgtt cctgtaacgt tcgactatct cgctagctgt gaagtactga ttagtggagt 35700
tctgtgcaac agcagtgtag gagtatacac aaacacaaat atgtgtttct atttaaaact 35760
gtggacttag cataaaaagg gagaatatat ttatttttta caaaagggat aaaaatgggc 35820
cccgttcctc acccaccaga tttagcgaga aaaagctttc tattctgaaa ggtcacggtg 35880
gctttggcat tacaaatcag aacaacacac actgaccatg atggcttgtg aactaactgc 35940
aaggcactcc gtcatggtaa gcgagtaggt cccacctcct agtgtgccgc tcattgcttt 36000
acacagtaga atcttatttg agtgctaatt gttgtctttg ctgctttact gtgttgttat 36060
agaaaatgta agctgtacag tgaataagtt attgaagcat gtgtaaacac tgttatatat 36120
cttttctcct agatggggaa ttttgaataa aataccttg aaattctgtg tatgttttag 36180
ttcattattt agggaaaacg ctgctgtgaa agggggcgtg atcagcttcc tattctgcga 36240
cagtcgtgtt gaacggaacc cattggtttt catcttcgct ccccccccct tggtttttcg 36300
agacagggtt tctctgtata gccctggctg tcctggacct cactctgtag accaggctgg 36360
cctcgaactc agaaatctac ctgcctctgc ctcccaagtg ctgggaggca gttgccccac 36420
caactagtct tctttttca aagaagatat ttaaagctaa cgaataatgc tagactctta 36480
catcttaaaa aaaaaagaag agaaaagaaa agaaaaggta atcacactgc ccagtgtgta 36540
gtgcatgctt ctacttccgg tccttgggag atgggggcag gatgagacgc tccagaccgg 36600
cttccaatac agagttcaag acccactgag ctacgtgagg ctacacgagc ctgcctttaa 36660
aaacataaag ctaaagcttt cttcttaact tccagtattg caccttgatt ccccttcaa 36720
atttcacata caaaataatt cttaaattct cttttgaaaa atgttctact gaggccagag 36780
agacagttcg cttggtaaag gtgcctgttg ccaaacgtga taacctgagt taaatcatag 36840
ccccacatgg gggaggaaga aacccccgca gcttgccctc tgatgccatg tatgcactaa 36900
aacacgcacg tgtgtgcgca cacatttttt aagttcctat tacattgata gtaatataat 36960
ttaaactgat ttattctccc caagtcattg atacgggtgt ccaacgtaaa atccagcggc 37020
tgaacaaagc acttttaggc gctttaagtt ggaaagcaag aaacggagat tgacactgtc 37080
actccaagag aaaactcttc gtagtagcga gatcggctgt ggagtgaaga tgctcagagg 37140
ctgggaacgc acacagctca ggagtggata gcatccccca gcctcaactc ctaacactgg 37200
gaaagcgtag ggctctcaga tgaggaaaca aaaccataca aagctgctgc aagctaaaca 37260
gaaaaatagt ggcattacac taactgttgt ggaattgtac agaccgattc tcctcccaat 37320
ctgccgagtg tgggcggctt gagagaatga agagagctac tggcctcagg taacagtgct 37380
tcccacagga ctgtctcagg ctgccaccac cataaatagc attttagacg tgacagagct 37440
aaggcttgac acacagccaa aagctactca cattccattt catccccagc tgttctgtca 37500
tcgctaagca cagagcattc agcacagctc ttccctgtgg tgggtactca gcactgttga 37560
gttgaaagga ttgaaaaaac tcaagactat gttctcaaac attttttaa gctcttttta 37620
aaaccacctt agaatgaaag cttttgactt cttattaaca tgcactaact tcatatacac 37680
atttagtgtt attgtacagg cacgaagcat actctggtca gaacctgtct cctttggtcc 37740
accctcccca ccgttttcag cttctattcc accttccata cgtctcaaga tccacatgtg 37800
agagggaaca ctcagagcct tgtctttctg tatctgggat atctcactta acatgatatt 37860
ctccagttct gttccatcca tttcattgca aagagcaaga tttcactcta cagccaaata 37920

```
acacatttgt ccatgtatat ccgtattttt ccttattcat ctgttgaatg gcacaagact 37980
gatatcatgg gtaatatcta t                                           38001
```

<210> 12
<211> 3435
<212> ADN
<213> Rattus norvegicus

<220>
<221> CDS
<222> (370) ... (1815)

<400> 12

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

gcc atc agc tca cat ctg cag acc tgc ggc tgt tct gtg gtg gta ggc   1083
Ala Ile Ser Ser His Leu Gln Thr Cys Gly Cys Ser Val Val Val Gly
        225                 230                 235

agc agt gca gag aaa gta aat aag ata gta aga aca ctg tgc ctt ttt   1131
Ser Ser Ala Glu Lys Val Asn Lys Ile Val Arg Thr Leu Cys Leu Phe
    240                 245                 250

ctg aca cca gca gag agg aag tgc tcc agg ctg tgt gaa gcc gaa tcg   1179
Leu Thr Pro Ala Glu Arg Lys Cys Ser Arg Leu Cys Glu Ala Glu Ser
255                 260                 265                 270

tcc ttt aaa tac gaa tct gga ctc ttt gta caa ggc ttg cta aag gat   1227
Ser Phe Lys Tyr Glu Ser Gly Leu Phe Val Gln Gly Leu Leu Lys Asp
                275                 280                 285

gcg act ggc agt ttt gta cta cct ttc cgg caa gtt atg tat gcc cct   1275
Ala Thr Gly Ser Phe Val Leu Pro Phe Arg Gln Val Met Tyr Ala Pro
            290                 295                 300

tat ccc acc aca cac atc gat gtg gat gtc aac act gtc aag cag atg   1323
Tyr Pro Thr Thr His Ile Asp Val Asp Val Asn Thr Val Lys Gln Met
        305                 310                 315

cca ccg tgt cat gaa cat att tat aat caa cgc aga tac atg agg tca   1371
Pro Pro Cys His Glu His Ile Tyr Asn Gln Arg Arg Tyr Met Arg Ser
    320                 325                 330

gag ctg aca gcc ttc tgg agg gca act tca gaa gag gac atg gct cag   1419
Glu Leu Thr Ala Phe Trp Arg Ala Thr Ser Glu Glu Asp Met Ala Gln
335                 340                 345                 350

gac acc atc atc tac aca gat gag agc ttc act cct gat ttg aat att   1467
Asp Thr Ile Ile Tyr Thr Asp Glu Ser Phe Thr Pro Asp Leu Asn Ile
                355                 360                 365

ttc caa gat gtc tta cac aga gac act cta gtg aaa gcc ttt ctg gat   1515
Phe Gln Asp Val Leu His Arg Asp Thr Leu Val Lys Ala Phe Leu Asp
            370                 375                 380
```

```
cag gtc ttc cat ttg aag cct ggc ctg tct ctc agg agt act ttc ctt   1563
Gln Val Phe His Leu Lys Pro Gly Leu Ser Leu Arg Ser Thr Phe Leu
        385                 390                 395

gca cag ttc ctc ctc att ctt cac aga aaa gcc ttg aca cta atc aag   1611
Ala Gln Phe Leu Leu Ile Leu His Arg Lys Ala Leu Thr Leu Ile Lys
    400                 405                 410

tac ata gag gat gac acg cag aag ggg aaa aag ccc ttt aag tct ctt   1659
Tyr Ile Glu Asp Asp Thr Gln Lys Gly Lys Lys Pro Phe Lys Ser Leu
415                 420                 425                 430

cgg aac ctg aag ata gat ctt gat tta aca gca gag ggc gac ctt aac   1707
Arg Asn Leu Lys Ile Asp Leu Asp Leu Thr Ala Glu Gly Asp Leu Asn
                435                 440                 445

ata ata atg gct cta gct gag aaa att aag cca ggc cta cac tct ttc   1755
Ile Ile Met Ala Leu Ala Glu Lys Ile Lys Pro Gly Leu His Ser Phe
            450                 455                 460

atc ttc ggg aga cct ttc tac act agt gtc caa gaa cgt gat gtt cta   1803
Ile Phe Gly Arg Pro Phe Tyr Thr Ser Val Gln Glu Arg Asp Val Leu
        465                 470                 475

atg act ttt taa acatgtggtt tgctccgtgt gtctcatgac agtcacactt       1855
Met Thr Phe
    480

gctgttacag tgtctcagcg ctttggacac atccttcctc cagggtcctg ccgcaggaca 1915
cgttacacta cacttgtcag tagaggtctg taccagatgt caggtacatc gttgtagtga 1975
atgtctcttt tcctagacta gatgtaccct cgtagggact tatgtttaca accctcctaa 2035
gtactagtgc tgtcttgtaa ggatacgaat gaagggatgt aaacttcacc acaactgctg 2095
gttggttttg ttgtttttgt tttttgaaac ttataattca tggtttacat gcatcacact 2155
gaaaccctag ttagcttttt acaggtaagc tgtgagttga ctgcctgtcc ctgtgttctc 2215
tggcctgtac gatctgtggc gtgtaggatc acttttgcaa caactaaaaa ctaaagcact 2275
ttgtttgcag ttctacagaa agcaacttag tctgtctgca gattcgtttt tgaaagaaga 2335
catgagaaag cggagtttta ggtgaagtca gttgttggat cttcctttat agacttagtc 2395
ctttagatgt ggtctgtata gacatgccca accatcatgc atgggcactg aatatcgtga 2455
actgtggtat gctttttgtt ggtttattgt acttctgtca aagaaagtgg cattggtttt 2515
tataattgtt gccaagtttt aaggttaatt ttcattattt ttgagccaaa ttaaaatgtg 2575
cacctcctgt gcctttccca atcttggaaa atataatttc ttggcagaag gtcagatttc 2635
agggcccagt cactttcgtc tgacttccct ttgcacagtc cgccatgggc ctggcttaga 2695
agttcttgta aactatgcca gagagtacat tcgctgataa aatcttcttt gcagagcagg 2755
agagcttctt gcctctttcc tttcatttct gcctggactt tggtgttctc cacgttccct 2815
gcatcctaag gacagcagga gaactctgac cccagtgcta tttctctagg tgctattgtg 2875
gcaaactcaa gcggtccgtc tctgtccctg taacgttcgt accttgctgg ctgtgaagta 2935
ctgactggta aagctccgtg ctacagcagt gtagggtata cacaaacaca agtaagtgtt 2995
ttatttaaaa ctgtggactt agcataaaaa gggagactat atttattttt tacaaaaggg 3055
ataaaaatgg aacccttttcc tcacccacca gatttagtca gaaaaaaaca ttctattctg 3115
aaaggtcaca gtggttttga catgacacat cagaacaacg cacactgtcc atgatggctt 3175
atgaactcca agtcactcca tcatggtaaa tgggtagatc cctccttcta gtgtgccaca 3235
ccattgcttc ccacagtaga atcttattta agtgctaagt gttgtctctg ctggtttact 3295
ctgttgtttt agagaatgta agttgtatag tgaataagtt attgaagcat gtgtaaacac 3355
tgttatacat cttttctcct agatggggaa tttggaataa aataccttta aaattcaaaa 3415
aaaaaaaaaa aaaaaaaaaa                                           3435
```

<210> 13
<400> 13
000

<210> 14

<400> 14
000

<210> 15
<211> 22
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 15
tgtgacagtt ggaatgcagt ga 22

<210> 16
<211> 25
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 16
gccacttaaa gcaatctctg tcttg 25

<210> 17
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Sonda

<400> 17
tcgactcttt gcccaccgcc a 21

<210> 18
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 18
gggtctagca agagcaggtg 20

<210> 19
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 19
gtcttggcaa cagctggaga t 21

<210> 20

<211> 24
<212> ADN
<213> Secuencia artificial

<220>
<223> Sonda

<400> 20
tgatgtcgac tctttgccca ccgc 24

<210> 21
<211> 22
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 21
tcctgtaatg gaactgcttt ca 22

<210> 22
<211> 22
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 22
ggtatctgct tcatccagct tt 22

<210> 23
<211> 16
<212> ADN
<213> Secuencia artificial

<220>
<223> Sonda

<400> 23
ccccggcccc ggcccc 16

**REIVINDICACIONES**

1. Un compuesto que comprende un oligonucleótido antisentido modificado de cadena sencilla que consiste de 12 a 30 nucleósidos enlazados, en el que el oligonucleótido antisentido modificado de cadena sencilla es por lo menos un 90% complementario a un ácido nucleico de C9ORF72, para su uso en el tratamiento de esclerosis lateral amiotrófica (ELA) o demencia frontotemporal (FTD) reduciendo el porcentaje de células con focos nucleares y/o reduciendo el número de focos nucleares por célula en un animal al que se ha identificado que tiene focos, en donde el oligonucleótido antisentido modificado de cadena sencilla comprende por lo menos un enlace internucleósido de fosforotioato, por lo menos una 5-metilcitosina, por lo menos un azúcar bicíclico o por lo menos un grupo 2'-O-metoxietilo.

2. El compuesto para el uso de acuerdo con la reivindicación 1, en el que la secuencia de nucleobases del oligonucleótido antisentido modificado de cadena sencilla es 100% complementaria a una porción de igual longitud de un ácido nucleico de C9ORF72.

3. El compuesto para el uso de acuerdo con la reivindicación 1 o la reivindicación 2, en el que la secuencia de nucleobases del oligonucleótido antisentido modificado de cadena sencilla es complementaria a una expansión de repetición de hexanucleótidos, opcionalmente en donde la expansión de repetición de hexanucleótidos es cualquiera de GGGGCC, GGGGGG, GGGGGC o GGGGCG.

4. El compuesto para el uso de acuerdo con cualquiera de las reivindicaciones 1-3, en el que el ácido nucleico de C9ORF72 tiene la secuencia de nucleobases de cualquiera de las SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, o SEQ ID NO: 10.

5. El compuesto para el uso de acuerdo con cualquiera de las reivindicaciones 1-4, en el que el oligonucleótido antisentido modificado de cadena sencilla comprende:

un segmento de hueco que consiste de desoxinucleósidos enlazados;
un segmento de ala 5' que consiste de nucleósidos enlazados;
un segmento de ala 3' que consiste de nucleósidos enlazados;

en donde el segmento de hueco está colocado inmediatamente adyacente a y entre el segmento de ala 5' y el segmento de ala 3' y en donde cada nucleósido de cada segmento de ala comprende un azúcar modificado.

6. El compuesto para el uso de acuerdo con cualquiera de las reivindicaciones 1-5, en el que el oligonucleótido antisentido modificado de cadena sencilla comprende por lo menos un enlace internucleosídico de fosforotioato.

7. El compuesto para el uso de acuerdo con la reivindicación 6, en el que cada enlace internucleosídico es un enlace internucleósido de fosforotioato.

8. El compuesto para el uso de acuerdo con cualquiera de las reivindicaciones 1-7, en el que el oligonucleótido antisentido modificado de cadena sencilla comprende por lo menos una 5-metilcitosina.

9. El compuesto para el uso de acuerdo con cualquiera de las reivindicaciones 1-8, en el que el oligonucleótido antisentido modificado de cadena sencilla comprende por lo menos un nucleósido modificado que comprende un azúcar bicíclico.

10. El compuesto para el uso de acuerdo con la reivindicación 9, en el que el azúcar bicíclico comprende un puente químico entre las posiciones 4' y 2' del azúcar.

11. El compuesto para el uso de acuerdo con la reivindicación 10, en el que el puente químico es a) 4'-CH(R)-O-2' y en el que R es metilo; b) 4'-CH(R)-O-2' y en el que R es H; o c) 4'-CH(R)-O-2' y en el que R es $-CH_2-O-CH_3$.

12. El compuesto para el uso de acuerdo con cualquiera de las reivindicaciones 1-11, en el que el oligonucleótido antisentido modificado de cadena sencilla comprende por lo menos un nucleósido modificado que comprende un grupo 2'-O-metoxietilo.

13. El compuesto para el uso de acuerdo con cualquiera de las reivindicaciones 1-12, en donde el oligonucleótido antisentido modificado de cadena sencilla está unido covalentemente a un grupo conjugado