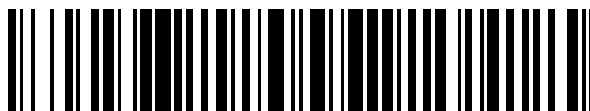


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 485 848**

51 Int. Cl.:

A61K 48/00 (2006.01)

C07H 21/04 (2006.01)

C12N 15/11 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.09.2004 E 04783980 (8)**

97 Fecha y número de publicación de la concesión europea: **21.05.2014 EP 1670518**

54 Título: **ARN de interferencia para el tratamiento de trastornos relacionados con la ganancia de función**

30 Prioridad:

12.09.2003 US 502678 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

14.08.2014

73 Titular/es:

**UNIVERSITY OF MASSACHUSETTS (100.0%)
365 PLANTATION STREET
WORCESTER, MA 01605, US**

72 Inventor/es:

**ARONIN, NEIL y
ZAMORE, PHILLIP D.**

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 485 848 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

ARN de interferencia para el tratamiento de trastornos relacionados con la ganancia de función

5 Antecedentes de la invención

ARN de interferencia (ARNi) es el mecanismo de silenciamiento génico post-transcripcional específico de una secuencia, iniciado por los ARN bicatenarios (ARNbc) homólogos al gen que se suprime. Los ARNbc son procesados por Dicer, una ribonucleasa celular III, para generar dúplex de aproximadamente 21 nt con salientes 3' (ARN pequeño de interferencia, ARNpi) que median la degradación del ARNm específico de una secuencia. En células de mamíferos, las moléculas de ARNpi son capaces de silenciar específicamente la expresión génica sin inducción de la vía de respuesta inespecífica de interferón. Por lo tanto, los ARNpi se han convertido en una nueva y poderosa alternativa para otras herramientas genéticas tales como oligonucleótidos antisentido y ribozimas para analizar la función de los genes. Además, los ARNpi se están desarrollando con fines terapéuticos, con el fin de silenciar los genes causantes de enfermedades en los seres humanos.

Las enfermedades de repetición de trinucleótidos comprenden un grupo recientemente reconocido de trastornos hereditarios. La mutación genética común es un aumento en una serie de una repetición de trinucleótidos en particular. Hasta la fecha, la repetición de trinucleótidos más frecuente es CAG, que codifica para el aminoácido glutamina. Se conocen al menos 9 enfermedades de repetición de CAG y existen más de 20 variedades de estas enfermedades, incluyendo la enfermedad de Huntington, la enfermedad de Kennedy y muchas enfermedades espinocerebelosas. Estos trastornos comparten un componente neurodegenerativo en el cerebro y / o la médula espinal. Cada enfermedad tiene un patrón específico de neurodegeneración en el cerebro y la mayoría tiene una herencia autosómica dominante.

El inicio de las enfermedades en general se produce alrededor de los 30 a 40 años de edad, pero en la enfermedad de Huntington repeticiones de CAG en el gen de la Huntingtina de > 60 presagiar un inicio juvenil.

La investigación reciente por parte de los presentes inventores ha demostrado que la mutación genética (aumento de la longitud de repeticiones de CAG desde normal <36 en el gen de la Huntingtina hasta > 36 en el caso de la enfermedad) se asocia con la síntesis de una proteína Huntingtina mutante, que tiene > 36 poliglutaminas (Aronin et al., 1995). También se ha demostrado que la proteína forma agregados citoplasmáticos e inclusiones nucleares (Difiglia et al., 1997) y se asocia con vesículas (Aronin et al., 1999). Las vías patogénicas precisas no se conocen.

Se cree que la enfermedad de Huntington (y por implicación otras enfermedades de repetición de trinucleótidos) es causada, al menos en parte, por interacciones aberrantes de proteínas, que causan deterioro de procesos neuronales críticos, disfunción neuronal y en última instancia la muerte neuronal (neurodegeneración en áreas del cerebro llamadas el cuerpo estriado y corteza). En la búsqueda de un tratamiento efectivo para estas enfermedades, los investigadores en este campo destacaron la comprensión de la patogénesis de la enfermedad y en un principio trataron de interceder al nivel de las presuntas interacciones de proteínas aberrantes. Sin embargo, no existe un tratamiento efectivo para la enfermedad de Huntington o de otras enfermedades causadas por repetición de trinucleótidos. Además, hoy en día se observa que múltiples procesos anormales pueden ser activos en estos tipos de enfermedad.

45 Resumen de la invención

La presente invención se define en y por medio de las reivindicaciones anexas.

Se han propuesto otros métodos basados en ARNi para la destrucción de genes mutantes en los que los ARNpi están dirigidos, por ejemplo, a una mutación puntual que se presenta en un solo alelo en el gen mutante (por ejemplo, la mutación puntual en el gen para la superóxido dismutasa (SOD) asociada con la esclerosis lateral amiotrófica (ELA)). Sin embargo, hay una diferencia clave entre la ELA y las enfermedades de repetición de trinucleótidos, tales como la enfermedad de Huntington. ELA tiene una mutación puntual en un alelo como el cambio genético, mientras que las enfermedades de repetición de trinucleótidos tienen una región de repetición de CAG expandida en un alelo como el cambio genético. El uso de ARNi contra la región de repetición expandida de CAG tiene complicaciones potenciales. Se sabe que existen más de 80 genes normales con regiones de repetición de CAG en las células. Por lo tanto, los ARNpi dirigidos a estas repeticiones de CAG no se pueden utilizar sin correr el riesgo de la destrucción generalizada de los ARNm que contiene la repetición de CAG normal. Del mismo modo, el enfoque en sitios no específicos de alelo daría lugar a la pérdida tanto de Huntingtina normal como mutante lo que provoca disfunción neuronal.

La invención utilizan la tecnología de ARN de interferencia (ARNi) contra regiones polimórficas seleccionadas (es decir, regiones que contienen polimorfismos específicos alélicos o específicos del alelo) que son distintos del sitio de mutación en los genes que codifican proteínas mutantes. La presente invención proporciona tratamientos efectivos para enfermedades por ganancia de función resultantes de mutaciones por supresión, mutaciones por inserción, mutaciones puntuales, y similares, siempre que el gen mutante codifique una proteína que tenga una función que

normalmente no esté asociada con una proteína de tipo silvestre.

La presente invención proporciona un tratamiento efectivo para la enfermedad de Huntington (HD). La invención también puede proporcionar tratamientos efectivos para otros trastornos por poliglutamina y / o una enfermedad por repetición de trinucleótidos, como se describe en detalle en la presente invención.

En una realización, la proteína mutante contiene una región de poliglutamina expandida. En otra realización, el gen que codifica la proteína mutante contiene una región de repetición de trinucleótidos expandida.

La invención puede utilizar agentes de ARNi homólogos a un polimorfismo alélico en el gen que codifica, por ejemplo, una proteína Huntingtina mutante para el tratamiento de la enfermedad de Huntington. En una forma de realización preferida, el agente de ARNi puede dirigir un polimorfismo alélico seleccionado del grupo que consiste de P1-P5. En una forma de realización preferida adicional, el agente de ARNi puede dirigir un polimorfismo alélico seleccionado del grupo que consiste de P6-P43.

En una realización adicional, la invención proporciona agentes de ARNi que constan de una primera y una segunda cadena que contienen cada una 16 - 25 nucleótidos. La primera cadena de la presente invención puede ser homóloga a una región de un gen que codifica una proteína mutante de ganancia de función, en la que la secuencia de nucleótidos de la proteína mutante de ganancia de función comprende un polimorfismo alélico. La segunda cadena incluye 16 - 25 nucleótidos complementarios a la primera cadena. El agente de ARNi también puede tener una porción de bucle que comprende 4 - 11, por ejemplo, 4, 5, 6, 7, 8, 9, 10, 11, nucleótidos que conecta las dos secuencias de nucleótidos. La región objetivo de la secuencia de ARNm se puede localizar en una región 5' no traducida (UTR) o una UTR 3' del ARNm de una proteína mutante.

La invención proporcionar un constructo de expresión que comprende un ácido nucleico aislado que codifica una molécula de ácido nucleico con una primera secuencia de 16 - 25 nucleótidos homólogos a un polimorfismo alélico dentro de, por ejemplo, el gen que codifica una proteína Huntingtina mutante. El constructo de expresión puede ser, por ejemplo, un vector viral, un vector retroviral, un casete de expresión o un plásmido. El constructo de expresión puede tener también una secuencia del promotor de la ARN polimerasa II o una secuencia del promotor de la ARN Polimerasa II, tal como, el promotor de ARN_p U6 del promotor H1.

En aún otras formas de realización, la presente invención proporciona células huésped (por ejemplo, células de mamífero) que comprenden moléculas de ácido nucleico y constructos de expresión de la presente invención.

En incluso otras formas de realización, la presente invención proporciona composiciones terapéuticas que comprenden las moléculas de ácido nucleico de la invención y un vehículo farmacéuticamente aceptable.

Otras características y ventajas de la invención serán evidentes a partir de la siguiente descripción detallada y las reivindicaciones.

Breve descripción de los dibujos

Figuras 1a - k: gen de la Huntingtina humana, secuencia de nucleótidos (SEQ ID NO: 1)

Figuras 2a - b: proteína Huntingtina humana, secuencia de aminoácidos (SEQ ID NO: 2)

Figura 3: Sentido (SEQ ID NO: 3) y antisentido (SEQ ID NO: 4) de la secuencia de ARN objetivo de la Huntingtina (htt)

Figuras 4: Análisis termodinámico de los extremos 5' de la cadena de ARN_{pi} para el dúplex de ARN_{pi}

Figuras 5a - c: Reacciones de ARNi *in vitro* programadas con ARN_{pi} que dirige un polimorfismo dentro del ARNm de la Huntingtina (htt). (a) ARN_{pi} estándar. (b) ARN_{pi} mejorado mediante la reducción de la fuerza de apareamiento de bases del extremo 5' de la cadena antisentido del dúplex de ARN_{pi}. (c) ARN_{pi} mejorado mediante la reducción del extremo 5' no emparejado de la cadena antisentido del dúplex de ARN_{pi}.

Figura 6a - b. ARNi de la proteína Htt endógena en células HeLa. (a) inmunotransferencia de la proteína Htt humana. (b) Cuantificación de la misma.

Descripción detallada de la invención

La presente invención utiliza tecnología de interferencia de ARN (ARNi) contra los polimorfismos alélicos situados dentro de un gen que codifica una proteína mutante de ganancia de función. El ARNi destruye el correspondiente ARNm mutante con especificidad y selectividad de nucleótidos. Agentes de ARN de la presente invención son dirigidos a las regiones polimórficas de un gen mutante, lo que resulta en la escisión del ARNm mutante. Estos agentes de ARN, actúan a través de una serie de interacciones proteína-nucleótido, para escindir los ARNm mutantes. Las células destruyen el ARNm escindido, impidiendo así la síntesis de la proteína mutante correspondiente por ejemplo, la proteína Huntingtina.

La invención puede utilizar agentes de ARNi homólogos a un polimorfismo alélico en el gen que codifica, por ejemplo, una proteína Huntingtina mutante para el tratamiento de la enfermedad de Huntington. El agente de ARNi

puede dirigir el polimorfismo alélico seleccionado del grupo que consiste de P1-P5. El agente de ARNi puede dirigir un polimorfismo alélico seleccionado del grupo que consiste de P6-P43.

5 En una realización adicional, la invención proporciona agentes de ARNi que comprenden de una primera y una segunda cadenas cada una conteniendo 16-25 nucleótidos. La primera cadena de la presente invención es homóloga a una región de un gen que codifica una proteína mutante de ganancia de función, en donde la secuencia de nucleótidos de la proteína mutante de ganancia de función comprende un polimorfismo alélico. La segunda cadena incluye 16-25 nucleótidos complementarios a la primera cadena. El agente de ARNi también puede tener una porción de bucle que comprende 4-11, por ejemplo, 4, 5, 6, 7, 8, 9, 10, 11, nucleótidos que conectan las dos
10 secuencias de nucleótidos. La región objetivo de la secuencia de ARNm se puede localizar en una región 5' no traducida (UTR) o una UTR 3' del ARNm de una proteína mutante.

15 La invención puede proporcionar un constructo de expresión que comprende un ácido nucleico aislado que codifica una molécula de ácido nucleico con una primera secuencia de 16-25 nucleótidos homólogos a un polimorfismo alélico dentro de, por ejemplo, el gen que codifica una proteína Huntingtina mutante. El constructo de expresión puede ser, por ejemplo, un vector viral, un vector retroviral, un casete de expresión o un plásmido. El constructo de expresión también puede tener una secuencia del promotor de la ARN polimerasa II o una secuencia del promotor de la ARN Polimerasa II, tal como, el promotor de ARN_p U6 del promotor H1.

20 En aún otras formas de realización, la presente invención proporciona células huésped (por ejemplo, células de mamífero) que comprenden moléculas de ácido nucleico y constructos de expresión de la presente invención.

25 En incluso otras formas de realización, la presente invención proporciona composiciones terapéuticas que comprenden las moléculas de ácido nucleico de la invención y un vehículo farmacéuticamente aceptable.

Para que la invención pueda entenderse más fácilmente, se definen primero ciertos términos.

30 El término "nucleósido" se refiere a una molécula que tiene una base de purina o de pirimidina enlazada covalentemente a un azúcar de ribosa o de desoxirribosa. Los nucleósidos de ejemplo incluyen adenosina, guanosina, citidina, uridina y timidina. Ejemplos adicionales de nucleósidos incluyen inosina, 1-metil inosina, seudouridina, 5,6-dihidrouridina, ribotimidina, ²N-metilguanosina y ^{2,2}N,N-dimetilguanosina (también denominados como nucleósidos "raros"). El término "nucleótido" se refiere a un nucleósido que tiene uno o más grupos fosfato unidos en enlaces éster a la fracción de azúcar. Los ejemplos de nucleótidos incluyen monofosfatos, difosfatos y trifosfatos de nucleósidos. Los términos "polinucleótido" y "molécula de ácido nucleico" se utilizan aquí de forma
35 intercambiable y se refieren a un polímero de nucleótidos unidos entre sí por un enlace fosfodiéster entre los átomos de carbono 5' y 3'.

40 El término "ARN" o "molécula de ARN" o "molécula de ácido ribonucleico" se refiere a un polímero de ribonucleótidos. El término "ADN" o "molécula de ADN" o "molécula de ácido desoxirribonucleico" se refiere a un polímero de desoxirribonucleótidos. El ADN y el ARN se pueden sintetizar de forma natural (por ejemplo, por replicación del ADN o transcripción del ADN, respectivamente). El ARN puede ser post-transcripcionalmente modificado. El ADN y el ARN también se pueden sintetizar químicamente. El ADN y el ARN pueden ser monocatenarios (es decir, ADN_{mc} y ARN_{mc}, respectivamente) o multicatenarios (por ejemplo, bicatenarios, es decir, ADN_{bc} y ARN_{bc}, respectivamente). "ARNm" o "ARN mensajero" es ARN monocatenario que especifica la secuencia de aminoácidos de una o más cadenas polipeptídicas. Esta información se traduce durante la síntesis de proteína cuando los ribosomas se enlazan al ARNm.
45

50 Como se usa en la presente invención, el término "ARN pequeño de interferencia" ("ARNpi") (también denominado en la técnica como "ARN corto de interferencia") se refiere a un ARN (o análogo de ARN) que comprende entre aproximadamente 10-50 nucleótidos (o análogos de nucleótido) que es capaz de dirigir o mediar la interferencia de ARN. Preferiblemente, un ARNpi comprende entre aproximadamente 15-30 nucleótidos o análogos de nucleótido, más preferiblemente entre aproximadamente 16-25 nucleótidos (o análogos de nucleótido), aún más preferiblemente entre aproximadamente 18-23 nucleótidos (o análogos de nucleótido), y aún más preferiblemente entre aproximadamente 19-22 nucleótidos (o análogos de nucleótido) (por ejemplo, 19, 20, 21 o 22 nucleótidos o análogos de nucleótido). El término ARNpi "corto" se refiere a un ARNpi que comprende ~21 nucleótidos (o análogos de nucleótido), por ejemplo, 19, 20, 21 o 22 nucleótidos. El término ARNpi "largo" se refiere a un ARNpi que comprende ~24-25 nucleótidos, por ejemplo, 23, 24, 25 o 26 nucleótidos. Los ARNpi cortos pueden, en algunos casos, incluir menos de 19 nucleótidos, por ejemplo, 16, 17 o 18 nucleótidos, siempre que el ARNpi más corto retenga la capacidad para mediar el ARNi. Del mismo modo, los ARNpi largos pueden, en algunos casos, incluir más de 26
60 nucleótidos, siempre que el ARNpi más largo retenga la capacidad para mediar el ARNi sin un procesamiento adicional, por ejemplo, un procesamiento enzimático, para un ARNpi corto.

65 El término "análogo de nucleótido" o "nucleótido alterado" o "nucleótido modificado" se refiere a un nucleótido no estándar, que incluye ribonucleótidos o desoxirribonucleótidos que no son de origen natural. Los análogos de nucleótido preferidos se modifican en cualquier posición con el fin de alterar ciertas propiedades químicas del

- nucleótido reteniendo aún la capacidad del análogo de nucleótido para llevar a cabo su función prevista. Ejemplos de las posiciones del nucleótido que puede ser sometidas a derivación incluyen la posición 5, por ejemplo, 5-(2-amino) propil uridina, 5-bromo uridina, 5-propino uridina, 5-propenil uridina, etc.; la posición 6, por ejemplo, 6-(2-amino) propil uridina; la posición 8 para adenosina y / o guanosinas, por ejemplo, 8-bromo guanosina, 8-cloro guanosina, 8-fluoroguanosina, etc. Los análogos de nucleótido también incluyen nucleótidos desaza, por ejemplo, 7-desaza-adenosina; nucleótidos O- y N-modificados (por ejemplo, alquilados, por ejemplo, la N6-metil adenosina, o como también se conocen en el arte); y otros análogos de nucleótido modificados en forma heterocíclica tales como aquellos descritos en Herdewijn, *Antisense Nucleic Acid Drug Dev.*, 2000 agosto 10 (4): 297 - 310.
- 10 Los análogos de nucleótido también pueden comprender modificaciones a la porción de azúcar de los nucleótidos. Por ejemplo, el grupo 2' OH puede ser reemplazado por un grupo seleccionado entre H, OR, R, F, Cl, Br, I, SH, SR, NH₂, NHR, NR₂, COOR, u OR, en donde R es alquilo, alqueno, alquino, arilo C₁-C₆ sustituido o no sustituido, etc. Otras posibles modificaciones incluyen aquellas descritas en las patentes de los Estados Unidos Nos. 5.858.988, y 6.291.438.
- 15 El grupo fosfato del nucleótido también puede modificarse, por ejemplo, mediante la sustitución de uno o más de los oxígenos del grupo fosfato con azufre (por ejemplo, fosforotioatos), o haciendo otras sustituciones que permitan que el nucleótido lleve a cabo su función prevista, como se describe, por ejemplo, en Eckstein, *Antisense Nucleic Acid Drug Dev.* 2000 abril 10 (2): 117 - 21, Rusckowski et al. *Antisense Nucleic Acid Drug Dev.* 2000 octubre 10 (5): 333 - 45, Stein, *Antisense Nucleic Acid Drug Dev.* 2001 octubre 11 (5): 317 - 25, Vorobjev et al. *Antisense Nucleic Acid Drug Dev.* 2001 abril 11 (2): 77 - 85, y la patente de los Estados Unidos No. 5.684.143. Algunas de las modificaciones anteriormente referenciadas (por ejemplo, modificaciones del grupo fosfato) disminuyen preferiblemente la velocidad de hidrólisis, por ejemplo, de polinucleótidos que comprenden dichos análogos *in vivo* o *in vitro*.
- 20 El término "oligonucleótido" se refiere a un polímero corto de nucleótidos y / o de análogos de nucleótido. El término "análogo de ARN" se refiere a un polinucleótido (por ejemplo, un polinucleótido sintetizado químicamente) que tiene al menos un nucleótido alterado o modificado en comparación con un ARN correspondiente no alterado o no modificado, pero que retiene la misma naturaleza o función o una similar que el correspondiente ARN inalterado o no modificado. Como se discutió anteriormente, los oligonucleótidos se pueden enlazar a través de los enlaces que dan como resultado una menor tasa de hidrólisis del análogo de ARN en comparación con una molécula de ARN con enlaces fosfodiéster. Por ejemplo, los nucleótidos del análogo pueden comprender enlaces metileno diol, etileno diol, oximetiltio, oxieltio, oxicarbonilo, fosforodiamidato, fosforamidato, y / o fosforotioato. Los análogos de ARN preferidos incluyen ribonucleótidos y / o desoxirribonucleótidos modificados en su cadena principal y/o el azúcar.
- 25 Tales alteraciones o modificaciones pueden incluir además de material no nucleotídico, tal como el(los) extremo(s) del ARN o internamente (en uno o más nucleótidos del ARN). Un análogo de ARN sólo necesita ser suficientemente similar al ARN natural que tenga la capacidad para mediar la interferencia de ARN.
- 30 Como se usa aquí, el término "ARN de interferencia" ("ARNi") se refiere a una degradación intracelular selectiva del ARN. El ARNi se produce en las células de forma natural para remover los ARN foráneos (por ejemplo, ARN virales). El ARNi natural procede a través de fragmentos escindidos del ARNbc libre que dirige el mecanismo de degradación a otras secuencias de ARN similares. Alternativamente, el ARNi puede ser iniciado por la mano del hombre, por ejemplo, para silenciar la expresión de los genes objetivo.
- 35 Un agente de ARNi que tiene una cadena que es una "secuencia suficientemente complementaria con una secuencia de ARNm objetivo para dirigir la interferencia de ARN específica del objetivo (ARNi)" significa que la cadena tiene una secuencia suficiente para activar la destrucción del ARNm objetivo mediante la maquinaria o el proceso del ARNi.
- 40 Como se usa aquí, el término "ARN aislado" (por ejemplo, "ARNpi aislado" o "precursor de ARNpi aislado") se refiere a moléculas de ARN que están sustancialmente libres de otro material celular, o medio de cultivo cuando se produce por técnicas recombinantes, o sustancialmente libre de precursores químicos u otros productos químicos cuando se sintetiza químicamente.
- 45 El término "in vitro" tiene su significado reconocido en la técnica, por ejemplo, la participación de reactivos o extractos purificados, por ejemplo, extractos de células. El término "in vivo" también tiene su significado reconocido en la técnica, por ejemplo, la participación de células vivas, por ejemplo, células inmortalizadas, células primarias, líneas celulares, y / o células en un organismo.
- 50 Tal como se usa en el presente documento, el término "transgén" se refiere a cualquier molécula de ácido nucleico, que se inserta por un medio artificial en una célula, y se convierte en parte del genoma del organismo que se desarrolla a partir de la célula. Tal transgén puede incluir un gen que es parcial o totalmente heterólogo (es decir, foráneo) para el organismo transgénico, o puede representar un gen homólogo a un gen endógeno del organismo. El término "transgén" también significa una molécula de ácido nucleico que incluye una o más secuencias de ácido nucleico seleccionadas, por ejemplo, los ADN que codifican uno o más precursores de ARN modificados por

ingeniería genética, para ser expresados en un organismo transgénico, por ejemplo, animal, que es en parte o enteramente heterólogo, es decir, foráneo, al animal transgénico, u homólogo a un gen endógeno del animal transgénico, pero que está diseñado para ser insertado en el genoma del animal en una ubicación que difiere de aquella del gen natural. Un transgén incluye uno o más promotores y cualquier otro ADN, tal como intrones, necesarios para la expresión de la secuencia de ácido nucleico seleccionado, todos operativamente enlazados a la secuencia seleccionada, y pueden incluir una secuencia potenciadora.

Un gen "involucrado" en una enfermedad o trastorno incluye un gen, la expresión normal o aberrante o la función que afecta o causa la enfermedad o trastorno o al menos un síntoma de dicha enfermedad o trastorno.

El término "mutación de ganancia de función" como se usa aquí, se refiere a cualquier mutación en un gen en el que la proteína codificada por dicho gen (i.e., La proteína mutante) adquiere una función que normalmente no está asociada con la proteína (es decir, la proteína de tipo silvestre) lo que causa o contribuye a una enfermedad o trastorno. La mutación de ganancia de función puede ser una supresión, adición o sustitución de un nucleótido o nucleótidos en el gen que da lugar al cambio en la función de la proteína codificada. La mutación de ganancia de función cambia la función de la proteína mutante o causa interacciones con otras proteínas. La mutación de ganancia de función puede provocar una disminución o la remoción de la proteína de tipo silvestre normal, por ejemplo, por la interacción de la proteína mutante alterada, con dicha proteína de tipo silvestre normal.

El término "polimorfismo" como se usa aquí, se refiere a una variación (por ejemplo, una supresión, inserción, o sustitución) en una secuencia génica que se identifica o se detecta cuando se comparan la misma secuencia génica de diferentes fuentes de sujetos (pero del mismo organismo). Por ejemplo, un polimorfismo puede ser identificado cuando se comparan la misma secuencia del gen de diferentes sujetos (pero del mismo organismo). La identificación de tales polimorfismos es de rutina en la técnica, siendo similares las metodologías a aquellas utilizadas para detectar, por ejemplo, mutaciones puntuales de cáncer de mama. La identificación puede ser hecha, por ejemplo, a partir del ADN extraído de linfocitos de un sujeto, seguido de la amplificación de regiones polimórficas utilizando cebadores específicos para dicha región polimórfica. Alternativamente, el polimorfismo se puede identificar cuando se comparan dos alelos del mismo gen. Una variación en la secuencia entre dos alelos del mismo gen dentro de un organismo se denomina aquí como un "polimorfismo alélico". El polimorfismo puede estar en un nucleótido dentro de una región de codificación, pero, debido a la degeneración del código genético, no se codifica ningún cambio en la secuencia de aminoácidos. Alternativamente, las secuencias polimórficas pueden codificar un aminoácido diferente en una posición particular, pero el cambio en el aminoácido no afecta la función de la proteína. También se pueden encontrar regiones polimórficas en las regiones no codificantes del gen.

El término "dominio de poliglutamina," tal como se utiliza aquí, se refiere a un segmento o dominio de una proteína que consiste de residuos de glutamina consecutivos enlazados a enlaces peptídicos. En una realización, la región consecutiva incluye al menos 5 residuos de glutamina.

El término "dominio expandido de poliglutamina" o "segmento expandido de poliglutamina", como se usa aquí, se refiere a un segmento o dominio de una proteína que incluye al menos 35 residuos de glutamina consecutivos unidos por enlaces peptídicos. Tales segmentos expandidos se encuentran en sujetos afectados con un trastorno de poliglutamina, tal como se describe en el presente documento, ya sea que el sujeto haya demostrado o no que manifiesta síntomas.

El término "repetición de trinucleótidos" o "región de repetición de trinucleótidos" como se usa aquí, se refiere a un segmento de una secuencia de ácido nucleico por ejemplo, que consiste de repeticiones consecutivas de una secuencia particular de trinucleótidos. Las repeticiones de trinucleótidos pueden incluir al menos 5 secuencias consecutivas de trinucleótidos. Los ejemplos de secuencias de trinucleótidos incluyen, pero no se limitan a, CAG, CGG, GCC, GAA, CTG, y / o CGG.

El término "enfermedades de repetición de trinucleótidos" como se utiliza aquí, se refiere a cualquier enfermedad o trastorno caracterizado por una región de repetición expandida de trinucleótidos situada dentro de un gen, siendo la región de repetición expandida de trinucleótidos la causante de la enfermedad o el trastorno. Los ejemplos de enfermedades de repetición de trinucleótidos incluyen, pero no se limitan a ataxia espinocerebelosa tipo 12, ataxia espinocerebelosa tipo 8, síndrome de X frágil, retardo mental de XE frágil, ataxia de Friedreich y la distrofia miotónica. Las enfermedades de repetición de trinucleótidos preferidas para el tratamiento de acuerdo con la presente invención son aquellas caracterizadas o causadas por una región de repetición expandida de trinucleótidos en el extremo 5' de la región de codificación de un gen, el gen que codifica una proteína mutante que causa o es la causante de la enfermedad o el trastorno. Ciertas enfermedades de trinucleótidos, por ejemplo, el síndrome de X frágil, donde la mutación no está asociada con una región de codificación pueden no ser adecuadas para el tratamiento de acuerdo con la presente invención, ya que no hay un ARNm adecuado para ser dirigido por el ARNi. En contraste, una enfermedad tal como la ataxia de Friedreich puede ser adecuada para el tratamiento de acuerdo con las metodologías de la invención porque, a pesar de que la mutación causante no está dentro de una región de codificación (es decir, se encuentra dentro de un intrón), la mutación puede estar dentro, por ejemplo, de un precursor de ARNm (por ejemplo, un precursor de ARNm empalme previamente).

El término "trastorno de poliglutaminas" como se usa aquí, se refiere a cualquier enfermedad o trastorno caracterizado por una expansión de las repeticiones (CAG)_n en el extremo 5' de la región de codificación (que codifica por lo tanto una región expandida de poliglutamina en la proteína codificada). En una forma de realización, los trastornos de poliglutamina se caracterizan por una degeneración progresiva de las células nerviosas. Los ejemplos de trastornos de poliglutamina incluyen, pero no se limitan a: la enfermedad de Huntington, ataxia espinocerebelosa tipo 1, ataxia espinocerebelosa tipo 2, ataxia espinocerebelosa tipo 3 (también conocida como enfermedad de Machado-Joseph), y ataxia espinocerebelosa tipo 6, ataxia espinocerebelosa tipo 7 y atrofia dentato-rubro-pálido-Luisiana.

La frase "examen de la función de un gen en una célula u organismo" se refiere al examen o estudio de la expresión, actividad, función o fenotipo que surge de los mismos.

Diversas metodologías de la presente invención incluyen una etapa que implica la comparación de un valor, nivel, rasgo, característica, propiedad, etc. para un "control adecuado", a que se hace referencia indistintamente en este documento como un "control adecuado". Un "control adecuado" o "control apropiado" es cualquier control o estándar familiar para alguien normalmente capacitado en la técnica útil para propósitos de comparación. Un "control adecuado" o "control apropiado" puede ser un valor, nivel, rasgo, característica, propiedad, etc. determinado antes de la realización de una metodología de ARNi, como se describe en el presente documento. Por ejemplo, una tasa de transcripción, nivel de ARNm, tasa de traducción, nivel de proteína, actividad biológica, característica o propiedad celular, genotipo, fenotipo, etc. se pueden determinar antes de la introducción de un agente de ARNi de la invención en una célula u organismo. Un "control adecuado" o "control apropiado" puede ser un valor, nivel, rasgo, característica, propiedad, etc. determinado en una célula u organismo, por ejemplo, un control o una célula normal u organismo, que exhiba, por ejemplo, rasgos normales. Un "control adecuado" o "control apropiado" es un valor, nivel, función, característica, propiedad predefinida, etc.

Varios aspectos de la invención se describen con más detalle en las siguientes subsecciones.

I. Trastornos de poliglutamina

Los trastornos de poliglutamina son una clase de enfermedades o trastornos caracterizados por una mutación genética común. En particular, la enfermedad o los trastornos se caracterizan por una repetición expandida del trinucleótido CAG que da lugar, en la proteína codificada, a un tramo expandido de residuos de glutamina. Los trastornos de poliglutamina son similares en que las enfermedades se caracterizan por una degeneración progresiva de las células nerviosas. A pesar de sus similitudes, los trastornos de poliglutamina se presentan en diferentes cromosomas y por lo tanto se presentan en segmentos totalmente diferentes del ADN. Los ejemplos de trastornos de poliglutamina incluyen la enfermedad de Huntington, la atrofia dentato-rubro-pálido-Luisiana, atrofia espinobulbar muscular, ataxia espinocerebelosa tipo 1, ataxia espinocerebelosa tipo 2, ataxia espinocerebelosa tipo 3, ataxia espinocerebelosa tipo 6 y ataxia espinocerebelosa tipo 7 (Tabla 3).

Tabla 1. Trastornos poliglutamina

Enfermedad	Gen	Locus	Proteína	Tamaño de la repetición CAG	
				Normal	Enfermedad
Atrofia espinobulbar muscular (enfermedad de Kennedy)	<i>AR</i>	Xq13-21	Receptor andrógeno (AR)	9-36	38-62
Enfermedad de Huntington	<i>HD</i>	4p16.3	Huntingtina	6-35	36-121
Atrofia dentato-rubro-pálido-Luisiana (síndrome de Haw-River)	<i>DRPLA</i>	12p13.31	Atrofina-1	6-35	49-88
Ataxia espinocerebelosa tipo 1	<i>SCA1</i>	6p23	Ataxina-1	6-44 ^a	39-82
Ataxia espinocerebelosa tipo 2	<i>SCA2</i>	12q24.1	Ataxina-2	15-31	36-63
Ataxia espinocerebelosa tipo 3 (enfermedad de Machado-Joseph)	<i>SCA3</i> (<i>MJD1</i>)	14q32.1	Ataxina-3	12-40	55-84
Ataxia espinocerebelosa tipo 6	<i>SCA6</i>	19p13	Subunidad del canal de calcio que depende del voltaje α_{1A}	4-18	21-33
Ataxia espinocerebelosa tipo 7	<i>SCA7</i>	13p12-13	Ataxina-7	4-35	37-306

^a Los alelos con 21 o más repeticiones son interrumpidos por 1-3 unidades CAT; los alelos de enfermedad contienen tramos de CAG.

Los trastornos de poliglutamina de la invención se caracterizan por (por ejemplo, los dominios que tienen aproximadamente entre 30 a 35 residuos de glutamina, aproximadamente entre 35 a 40 residuos de glutamina, aproximadamente entre 40 a 45 residuos de glutamina y tienen alrededor de 45 o más residuos de glutamina. El dominio de poliglutamina contiene típicamente residuos de glutamina consecutivos (Q n > 36).

II. Enfermedad de Huntington

La enfermedad de Huntington, hereda como una enfermedad autosómica dominante, provoca deterioro de la cognición y enfermedad motora. Los pacientes pueden vivir más de una década con una debilidad severa, antes de la muerte prematura por hambre o infección. La enfermedad comienza en la cuarta o quinta década de la mayoría de los casos, pero un subgrupo de pacientes manifiesta la enfermedad en la adolescencia. La mutación genética para la enfermedad de Huntington es una repetición de CAG alargada en el gen de la Huntingtina. La repetición de CAG varía en número de 8 a 35 en los individuos normales (Kremer et al., 1994). La mutación genética por ejemplo, un aumento de la longitud de la repeticiones de CAG desde normal hasta menos de 36 en el gen de la Huntingtina hasta más de 36 en la enfermedad se asocia con la síntesis de una proteína Huntingtina mutante, que tiene más de 36 poliglutamatos (Aronin et al., 1995). En general, los individuos con 36 o más repeticiones de CAG tendrán la enfermedad de Huntington. Prototípico de hasta veinte otras enfermedades con un CAG alargado como la mutación subyacente, la enfermedad de Huntington todavía no tiene ninguna terapia efectiva. Una variedad de intervenciones - tales como la interrupción de las rutas apoptóticas, la adición de reactivos para reforzar la eficiencia mitocondrial, y el bloqueo de los receptores NMDA - han demostrado ser prometedores en cultivos celulares y en el modelo de ratón de la enfermedad de Huntington. Sin embargo, en el mejor de estos enfoques se pone de manifiesto una corta prolongación de la supervivencia celular o animal.

La enfermedad de Huntington cumple con el dogma central de la genética: un gen mutante sirve como una plantilla para la producción de un ARNm mutante; el ARNm mutante entonces dirige la síntesis de una proteína mutante (Aronin et al., 1995; DiFiglia et al., 1997). La Huntingtina mutante (proteína) probablemente se acumula en neuronas selectivas en el cuerpo estriado y la corteza, interrumpe hasta aquí actividades celulares determinadas, y causa disfunción neuronal y la muerte (Aronin et al., 1999; Laforet et al., 2001). Debido a que una sola copia de un gen mutante es suficiente para causar la enfermedad de Huntington, el tratamiento más parsimonioso volvería al gen mutante ineficaz. Enfoques teóricos podrían incluir la detención de la transcripción del gen de la Huntingtina mutada, la destrucción del ARNm mutante, y el bloqueo de la traducción. Cada uno tiene el mismo resultado - la pérdida de la Huntingtina mutante.

III. Gen de la Huntingtina

El gen de la enfermedad relacionada con la enfermedad de Huntington se llama Huntington o (htt). El locus de la Huntingtina es grande, abarcando 180 kb y que consiste de 67 exones. El gen de la Huntingtina se expresa ampliamente y es necesario para el desarrollo normal. Se expresa como 2 formas alternativamente poliadeniladas que muestran diferente abundancia relativa en diversos tejidos fetales y adultos. El transcripto más grande es de aproximadamente 13,7 kb y se expresa predominantemente en el cerebro fetal y de adultos, mientras que el transcripto más pequeño de aproximadamente 10,3 kb es más ampliamente expresado. Los dos transcriptos difieren con respecto a sus regiones 3' no traducidas (Lin et al., 1993). Se predice que ambos mensajes codifican una proteína de 348 kilodaltons que contiene 3144 aminoácidos. Se cree que el defecto genético que conduce a la enfermedad de Huntington confiere una nueva propiedad en el ARNm o altera la función de la proteína. La secuencia de aminoácidos de la proteína Huntingtina humana se expone en la Figura 2 (SEQ ID NO: 2).

Una secuencia de nucleótidos consenso del gen de Huntingtina humana (ADNc) se expone en la Figura 1 (SEQ ID NO: 1). La región de codificación consiste de los nucleótidos 316 - 9750 de la SEQ ID NO: 1. Las dos señales de poliadenilación alternativas se encuentran en los nucleótidos 10.326 a 10.331 y los nucleótidos 13.644 a 13649, respectivamente. Los dos sitios de poliadenilación correspondientes se encuentran en los nucleótidos 10348 y 13672, respectivamente. La primera señal / sitio de poliadenilación es aquella del transcripto de 10,3 kb. La segunda señal / sitio de poliadenilación es aquella del transcripto de 13,7 kb, el transcripto predominante en el cerebro.

Se identificaron cinco (5) polimorfismos en el gen htt humano como se describe en el Ejemplo I. Se han identificado 38 polimorfismos adicionales en la secuencia del gen de la Huntingtina a través del análisis de SNP (polimorfismo de nucleótido único) (véase la Tabla 3). Los polimorfismos expuestos en las Tablas 2 y 3 representan los sitios preferidos para dirigir a través de un solo nucleótido específico de ARNi, como se describe en este documento.

Tabla 2. Sitios polimórficos (P) en el gen htt de líneas celulares humanas.

Línea celular	P1 (2886)	P2 (4034)	P3 (6912)	P4 (7222)	P5 (7246)
GFP-Htt (construido de 9 kb)	C	G	A	T	C
HeLa	t	a	A	g	C
HEK 293T	t	a	G	g	t
HepG2	t	a	G	g	t
FP-4	t	a	g, A	g	t, C

Tabla 3. Sitios polimórficos (P) en el gen htt humano identificados por el análisis de SNP.

	consenso	polimorfismo	db xref
--	----------	--------------	---------

complemento 103	G	A	P6	dbSNP:396875
complemento 432	T	C	P7	dbSNP:473915
complemento 474	C	A	P8	dbSNP:603765
1509	T	C	P9	dbSNP:1065745
complemento 1857	T	C	P10	dbSNP:2301367
3565	G	C, A	P11,P12	dbSNP:1065746
3594	T	G	P13	dbSNP:1143646
3665	G	C	P14	dbSNP:1065747
complemento 4122	G	A	P15	dbSNP:363099
complemento 4985	G	A	P16	dbSNP:363129
complemento 5480	T	G	P17	dbSNP:363125
6658	T	G	P18	dbSNP:1143648
complemento 6912	T	C	P19	dbSNP:362336
complemento 7753	G	A	P20	dbSNP:3025816
complemento 7849	G	C	P21	dbSNP:3025814
complemento 8478	T	C	P22	dbSNP:2276881
8574	T	C	P23	dbSNP:2229985
complemento 9154	C	A	P24	dbSNP:3025807
9498	T	C	P25	dbSNP:2229987
complemento 9699	G	A	P26	dbSNP:362308
complemento 9809	G	A	P27	dbSNP:362307
complemento 10064	T	C	P28	dbSNP:362306
complemento 10112	G	C	P29	dbSNP:362268
complemento 10124	G	C	P30	dbSNP:362305
complemento 10236	T	G	P31	dbSNP:362304
complemento 10271	G	A	P32	dbSNP:362303
complemento 10879	G	A	P33	dbSNP:1557210
complemento 10883	G	A	P34	dbSNP:362302
complemento 10971	C	A	P35	dbSNP:3025805
complemento 11181	G	A	P36	dbSNP:362267
complemento 11400	C	A	P37	dbSNP:362301
11756..11757	G	-	P38	dbSNP:5855774
12658	G	A	P39	dbSNP:2237008
complemento 12911	T	C	P40	dbSNP:362300
complemento 13040	G	A	P41	dbSNP:2530595
13482	G	A	P42	dbSNP:1803770
13563	G	A	P43	dbSNP:1803771

La presente invención está dirigida a la Huntingtina mutada utilizando ARN de interferencia (Hutvagner et al., 2002). Una cadena de ARN bicatenario (ARNbi) complementa una región polimórfica en el ARNm de la Huntingtina mutada. Después de la introducción del ARNpi en las neuronas, el ARNpi se desenrolla parcialmente, se enlaza a la región polimórfica dentro del ARNm de la Huntingtina de una manera específica del sitio, y activa una ARNm nucleasa. Esta nucleasa escinde el ARNm de la Huntingtina, deteniendo de esta forma la traducción de la Huntingtina mutante. Las células mismas se deshacen del ARNm parcialmente digerido, lo que impide la traducción, o las células digieren proteínas parcialmente traducidas. Las neuronas sobreviven en la Huntingtina de tipo silvestre (a partir del alelo normal); este enfoque evita los estragos de la Huntingtina mutante mediante la eliminación de su producción.

IV. Diseño del ARNpi

Los ARNpi se diseñan de la siguiente manera. En primer lugar, se selecciona una porción del gen objetivo (por ejemplo, el gen htt) que incluye el polimorfismo. Los ejemplos de polimorfismos se seleccionan de la región 5' no traducida de un gen objetivo. La escisión del ARNm en estos sitios debería eliminar la traducción de la proteína mutante correspondiente. Los polimorfismos de otras regiones del gen mutante también son adecuados para manipulación dirigida. Se diseña una cadena sentido con base en la secuencia de la porción seleccionada. Preferiblemente, la porción (y la cadena sentido correspondiente) incluye cerca de 19 a 25 nucleótidos, por ejemplo, 19, 20, 21, 22, 23, 24 o 25 nucleótidos. Más preferiblemente, la porción (y la correspondiente cadena sentido) incluye 21, 22 o 23 nucleótidos. La persona capacitada en la técnica apreciará, sin embargo, que los ARNpi que tienen una longitud de menos de 19 nucleótidos o mayor a 25 nucleótidos también pueden funcionar para mediar ARNi. En consecuencia, los ARNpi de tal longitud están también dentro del alcance de la presente invención siempre que conserven la capacidad de mediar el ARNi. Se ha demostrado que agentes de ARNi más largos provocan una respuesta de interferón o de PKR en ciertas células de mamífero que pueden ser indeseables. Preferiblemente, los agentes de ARNi de la invención no provocan una respuesta de PKR (es decir, son de una longitud suficientemente corta). Sin embargo, agentes de ARNi más largos pueden ser útiles, por ejemplo, en tipos

de células incapaces de generar una respuesta de PRK o en situaciones en las que la respuesta de PKR ha sido subregulada o moderada por medios alternativos.

5 La secuencia de cadena sentido se diseña de tal manera que el polimorfismo esté esencialmente en el medio de la cadena. Por ejemplo, si se elige un ARNpi de 21 nucleótidos, el polimorfismo se encuentra, por ejemplo, en los nucleótidos 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 o 16 (es decir, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 o 16 nucleótidos del extremo 5' de la cadena sentido. Para un ARNpi de 22 nucleótidos, el polimorfismo se encuentra, por ejemplo, en los nucleótidos 7, 8, 9, 10, 11, 12, 13, 14, 15 o 16. Para un ARNpi de 23 nucleótidos, el polimorfismo se encuentra, por ejemplo, en 7, 8, 9, 10, 11, 12, 13, 14, 15 o 16. Para un ARNpi de 24 nucleótidos, el polimorfismo se encuentra, por ejemplo, en 9, 10, 11, 12, 13, 14 o 16. Para un ARNpi de 25 nucleótidos, el polimorfismo se encuentra, por ejemplo, en 9, 10, 11, 12, 13, 14, 15, 16 o 17. Moviendo el polimorfismo a una posición fuera del centro puede, en algunos casos, reducir la eficiencia de la escisión por el ARNpi. Tales composiciones, es decir, las composiciones menos eficientes, pueden ser deseables para ser usadas si se detecta apagado del silenciamiento del ARNm de tipo silvestre.

15 La cadena antisentido es habitualmente de la misma longitud que la cadena sentido e incluye nucleótidos complementarios. Las cadenas pueden ser totalmente complementarias, es decir, las cadenas son de extremos romos cuando se alinean o se hibridan. Las cadenas se pueden alinear o hibridar de tal manera que se generen salientes de 1, 2 o 3 nucleótidos, es decir, el extremo 3' de la cadena sentido se extiende 1, 2 o 3 nucleótidos más lejos que el extremo 5' de la cadena antisentido y / o el extremo 3' de la cadena antisentido se extiende 1, 2 o 3 nucleótidos más lejos que el extremo 5' de la cadena sentido. Los salientes pueden comprender (o consistir de) los nucleótidos correspondientes a la secuencia del gen objetivo (o complemento del mismo). Alternativamente, los salientes pueden comprender (o consistir de) desoxirribonucleótidos, por ejemplo los dT, o análogos de nucleótidos, u otro material no nucleótido adecuado.

20 Para facilitar la entrada de la cadena antisentido en RISC (y por lo tanto aumentar o mejorar la eficiencia de la escisión y el silenciamiento objetivo), la resistencia del par de bases entre el extremo 5' de la cadena sentido y el extremo 3' de la cadena antisentido puede ser alterada, por ejemplo, disminuida o reducida, tal como se describe en detalle en las solicitudes de patente provisionales Nos. 60/475.386, titulada "Métodos y composiciones para controlar la eficacia del silenciamiento del ARN" (presentada el 2 de junio de 2003) y 60/475.331, titulada "Métodos y composiciones para mejorar la eficacia y especificidad del ARNi" (presentada el 2 de junio de 2003). La resistencia del par de bases puede ser menor debido a menos pares de bases G:C entre el extremo 5' de la primera cadena o la cadena antisentido y el extremo 3' de la segunda cadena o la cadena sentido que entre el extremo 3' de la primera cadena o la cadena antisentido y el extremo 5' de la segunda cadena o la cadena sentido. La resistencia del par de bases puede ser menor debido a al menos un par de bases no coincidentes entre el extremo 5' de la primera cadena o la cadena antisentido y el extremo 3' de la segunda cadena o la cadena sentido. Preferiblemente, se selecciona el par de bases no coincidentes del grupo que consiste de G:A, C:A, C:U, G:G, A:A, C:C y U:U. La resistencia del par de bases puede ser menor debido a por lo menos un par de bases que oscila, por ejemplo, G:U, entre el extremo 5' de la primera cadena o la cadena antisentido y el extremo 3' de la segunda cadena o la cadena sentido. La resistencia del par de bases es menor debido a por lo menos un par de bases que comprende un nucleótido raro, por ejemplo, inosina (I). Preferiblemente, el par de bases se selecciona del grupo que consiste de un I:A, I:U y I:C. La resistencia del par de bases es menor debido a por lo menos un par de bases que comprende un nucleótido modificado. El nucleótido modificado se puede seleccionar del grupo que consiste de 2-amino-G, 2-amino-A, 2,6-diamino-G, y 2, 6-diamino-A.

45 El diseño de los ARNpi adecuados para dirigir los polimorfismos htt expuestos en la Tabla 2 se describe en detalle a continuación

50	P1 ADN sentido antisentido (salientes de 2 nt) véase la Figura 5	TGTGCTGACTCTGAGGAACAG UGUGCUGACUCUGAGGAACAG ACACGACUGAGACUCCUUGUC	(SEQ ID NO: 5) (SEQ ID NO: 6) (extremos romos, 21 mer)	(SEQ ID NO: 7)
55	P2 ADN sentido antisentido	CATACCTCAAACCTGCATGATG CAUACCUCAAACUGCAUGAUG GUAUGGAGUUUGACGUACUAC	(SEQ ID NO: 8) (SEQ ID NO: 9) (extremos romos, 21 mer)	(SEQ ID NO: 10)
60	P3 ADN sentido antisentido	GCCTGCAGAGCCGGCGGCCTA GCCUGCAGAGCCGGCGGCCUA CGGACGUCUCGGCCGCGGAU	(SEQ ID NO: 11) (SEQ ID NO: 12) (extremos romos, 21 mer)	(SEQ ID NO: 13)
65	P4 ADN sentido antisentido	ACAGAGTTTGTGACCCACGCC ACAGAGUUUGUGACCCACGCC UGUCUCAAACACUGGGUGCGG	(SEQ ID NO: 14) (SEQ ID NO: 15) (extremos romos, 21 mer)	(SEQ ID NO: 16)

P5 ADN	TCCCTCATCTACTGTGTGCAC	(SEQ ID NO: 17)
sentido	UCCCUCAUCUACUGUGUGCAC	(SEQ ID NO: 18)
antisentido	AGGGAGUAGAUGACACACGUG	(extremos romos, 21 mer) (SEQ ID NO: 19)

5 Los ARNpi se pueden diseñar de acuerdo con las enseñanzas de los ejemplos anteriores para cualquier otro de los polimorfismos encontrados en el gen htt. Además, la tecnología es aplicable para dirigir cualquier otro gen de la enfermedad que tenga polimorfismos asociados, es decir, polimorfismos que no causan enfermedad.

10 Para validar la efectividad con que los ARNpi destruyen los ARNm mutantes (por ejemplo, el ARNm de la Huntingtina mutante), el ARNpi se incubaba con ADNc mutante (por ejemplo, ADNc de Huntingtina mutante) en un sistema de expresión de ARNm *in vitro* con base en *Drosophila*. Los ARNm mutantes sintetizados recientemente, marcados en forma radioactiva con ³²P, (por ejemplo, ARNm de Huntingtina mutante) se detectan autorradiográficamente en un gel de agarosa. La presencia de ARNm mutante escindido indica la actividad de ARNm nucleasa. Los controles adecuados incluyen la omisión de ARNpi y el uso de ADNc de la Huntingtina de tipo silvestre. Alternativamente, se seleccionan los ARNpi de control que tengan la misma composición de nucleótidos que el ARNpi seleccionado, pero sin complementariedad de secuencia significativa con el gen objetivo apropiado. Dichos controles negativos pueden diseñarse mezclando en forma aleatoria la secuencia de nucleótidos del ARNpi seleccionado; puede llevarse a cabo una búsqueda de homología para garantizar que el control negativo carezca de homología con cualquier otro gen en el genoma apropiado. Además, los ARNpi de control negativo pueden ser diseñados mediante la introducción de uno o más faltas de correspondencia de bases en la secuencia.

Los sitios de complementación de ARNpi - ARNm se seleccionan de tal manera que den lugar a una especificidad óptima de ARNm y una escisión máxima del ARNm.

25 Aunque la presente invención ofrece principalmente el direccionamiento de regiones polimórficas en el gen mutante objetivo (por ejemplo, en htt mutante) distinto de la mutación de la región expandida de CAG, la persona capacitada en la técnica apreciará que el direccionamiento de la región mutante puede tener aplicabilidad como estrategia terapéutica en ciertas situaciones. El direccionamiento de la región mutante se puede lograr utilizando ARNpi que complemente CAG en serie. El ARNpi^{CAG} se enlazaría a los ARNm con complementación de CAG, pero puede esperarse que tenga una mayor oportunidad de enlazarse a una serie extendida de CAG. Múltiples ARNpi^{CAG} se enlazarían al ARNm de Huntingtina mutante (en contraposición con una menor cantidad para el ARNm de Huntingtina de tipo silvestre); por lo tanto, es más probable que se escinda el ARNm de la Huntingtina mutante. La inactivación exitosa de ARNm utilizando este enfoque también eliminaría el ARNm de la Huntingtina normal o de tipo silvestre. Otros genes normales también podrían ser inactivados, al menos en cierta medida (aproximadamente 70) que también tienen repeticiones de CAG, donde sus ARNm podrían interactuar con el ARNpi. Este enfoque contaría por lo tanto con una estrategia de desgaste - más del ARNm de la Huntingtina mutante sería destruido que del ARNm de la Huntingtina de tipo silvestre o los otros aproximadamente 69 ARNm que codifican para poliglutaminas.

40 V. Agentes de ARNi

La presente invención incluye moléculas de ARNpi diseñadas, por ejemplo, como se describió anteriormente. Las moléculas de ARNpi de la invención se pueden sintetizar químicamente, o pueden transcribirse *in vitro* a partir de una plantilla de ADN, o *in vivo*, por ejemplo, a partir de ARNph, o, mediante el uso de la enzima DICER humana recombinante, para escindir *in vitro* las plantillas de ARNbc transcrito en grupos de ARNi que median el ARN dúplex de 20, 21 o 23 pb. Las moléculas de ARNpi se pueden diseñar usando cualquier método conocido en la técnica.

50 En lugar del agente de ARNi que es un ácido ribonucleico de interferencia, por ejemplo, un ARNpi o ARNph como se describió anteriormente, el agente de ARNi puede codificar un ácido ribonucleico de interferencia, por ejemplo, un ARNph, como se describió anteriormente. En otras palabras, el agente de ARNi puede ser una plantilla transcripcional del ácido ribonucleico de interferencia. Por lo tanto, los agentes de ARNi de la presente invención también pueden incluir los ARN pequeños de horquilla (ARNph), y constructos de expresión modificados por ingeniería genética para expresar los ARNpi. La transcripción de los ARNph se inicia en el promotor de una polimerasa III (pol III), y se cree que se termina en la posición 2 de un sitio de terminación de la transcripción de 4-5-timina. Tras la expresión, se cree que los ARNph se pliegan en una estructura de tallo-bucle con salientes de 3'UU; posteriormente, se procesan los extremos de estos ARNph, convirtiendo los ARNph en moléculas del tipo ARNpi de aproximadamente 21 - 23 nucleótidos (Brummelkamp et al., 2002;... Lee et al., 2002, ver más arriba; Miyagishi et al., 2002; Paddison et al., 2002, ver más arriba; Paul et al., 2002, ver más arriba; Sui et al., 2002, ver más arriba; Yu et al., 2002, ver más arriba. Más información sobre el diseño y el uso de ARNph se puede encontrar en Internet en las siguientes direcciones: Katahdin.cshl.org:9331/RNAi/docs/BseRIBamHI_Strategy.pdf y Katahdin.cshl.org:9331/RNAi/docs/Web_version_of_PCR_strategy1.pdf.

65 Los constructos de expresión de la presente invención incluyen cualquier constructo adecuado para su uso en el sistema de expresión apropiado e incluyen, pero no se limitan a, vectores retrovirales, casetes de expresión lineal, plásmidos y vectores virales o derivados de virus, tal como se conoce en la técnica. Tales constructos de expresión pueden incluir uno o más promotores inducibles, sistemas promotores de ARN Pol III tales como promotores de

ARNpn U6 o promotores de ARN polimerasa III H1, u otros promotores conocidos en la técnica. Los constructos pueden incluir una o ambas cadenas del ARNpi. Los constructos de expresión que expresan ambas cadenas también pueden incluir estructuras de bucle que unen las dos cadenas, o cada cadena puede ser transcrita por separado a partir de promotores separados dentro del mismo constructo. Cada cadena también puede transcribirse a partir de un constructo de expresión separado (Tuschl, T., 2002, ver más arriba).

Los ARNpi sintéticos pueden ser suministrados a células mediante métodos conocidos en la técnica, incluyendo la transfección de liposomas catiónicos y electroporación. Sin embargo, estos ARNpi exógeno en general, muestran la persistencia de corto plazo del efecto de silenciamiento (4-5 días en células cultivadas), que puede ser beneficioso en sólo ciertas formas de realización. Para obtener una supresión a largo plazo de los genes objetivo (es decir, genes mutantes) y para facilitar el suministro bajo ciertas circunstancias, se puede expresar uno o más ARNpi dentro de las células de constructos de ADN recombinante. Tales métodos para expresar dúplex de ARNpi dentro de las células de constructos de ADN recombinante para permitir la supresión del gen objetivo a más largo plazo en las células son conocidos en la técnica, incluyendo sistemas promotores Pol III de mamíferos (por ejemplo, sistemas promotores de U6/ARNpn (Tuschl, T., 2002, ver más arriba) capaces de expresar los ARNpi bicatenarios funcionales; (Bagella et al., 1998; Lee et al., 2002, ver más arriba; Miyagishi et al., 2002, ver más arriba; Paul et al., 2002, ver más arriba; Yu et al., 2002), ver más arriba; Sui et al., 2002, ver más arriba). La terminación de la transcripción por la ARN Pol III se produce en corridas de cuatro residuos T consecutivos en la plantilla de ADN, proporcionando un mecanismo para poner fin a la transcripción de ARNpi en una secuencia específica. El ARNpi es complementario a la secuencia del gen objetivo en orientaciones 5' - 3' y 3' - 5', y las dos cadenas del ARNpi se puede expresar en el mismo constructo o en constructos separados. Los ARNpi de horquilla, dirigidos por el promotor de ARNpn H1 o U6 y expresado en células, puede inhibir la expresión del gen objetivo (Bagella et al., 1998; Lee et al., 2002, ver más arriba; Miyagishi et al., 2002, ver más arriba; Paul et al., 2002, ver más arriba; Yu et al., 2002), ver más arriba; Sui et al., 2002, ver más arriba). Los constructos que contienen la secuencia de ARNpi bajo el control del promotor de T7 también vuelven los ARNpi funcionales cuando se cotransfectan en las células con un vector que expresa ARN polimerasa de T7 (Jacque et al., 2002, ver más arriba). Un solo constructo puede contener múltiples secuencias que codifican para los ARNpi, tal como múltiples regiones del gen que codifica htt mutante, dirigiendo el mismo gen o genes múltiples, y puede ser dirigidos, por ejemplo, por sitios del promotor PolIII separado.

Las células animales expresan una serie de ARN no codificante de aproximadamente 22 nucleótidos llamado micro ARN (miARN) que puede regular la expresión génica a nivel post transcripcional o translacional durante el desarrollo animal. Una característica común de los miARN es que todos ellos son cortados de un tallo-bucle de ARN precursor de 70 nucleótidos, probablemente por Dicer, una enzima de tipo ARNasa III, o un homólogo de la misma. Mediante la sustitución de las secuencias madre del precursor de miARN con una secuencia complementaria al ARNm objetivo, se puede utilizar un constructo de vector que expresa el precursor modificado por ingeniería genética para producir los ARNpi para iniciar el ARNi contra objetivos de ARNm específicos en células de mamífero (Zeng et al., 2002, ver más arriba). Cuando se expresa por medio de vectores de ADN que contienen promotores de polimerasa III, los bucles diseñados de micro ARN pueden silenciar la expresión génica (McManus et al., 2002, ver más arriba). Los micro ARN que dirigen polimorfismos también pueden ser útiles para bloquear la traducción de proteínas mutantes, en ausencia de silenciamiento de genes mediado por ARNpi. Tales aplicaciones pueden ser útiles en situaciones, por ejemplo, donde un ARNpi diseñado causó silenciamiento fuera del objetivo de proteína de tipo silvestre.

Los mecanismos de suministro mediados por virus también se pueden utilizar para inducir el silenciamiento específico de genes objetivo a través de la expresión de ARNpi, por ejemplo, mediante la generación de adenovirus recombinantes que albergan ARNpi bajo el control de transcripción del promotor de ARN Pol II bajo (Xia et al., 2002, ver más arriba). La infección de células HeLa por estos adenovirus recombinantes permite una expresión disminuida del gen objetivo endógeno. La inyección de los vectores de adenovirus recombinantes en ratones transgénicos que expresan los genes objetivo de los ARNpi da como resultado la reducción *in vivo* de la expresión del gen objetivo. Id. En un modelo animal, la electroporación de todo el embrión puede suministrar eficientemente ARNpi sintético en embriones de ratón posteriores a la implantación (Calegari et al., 2002). En ratones adultos, el suministro eficiente de ARNpi se puede lograr mediante la técnica de suministro a "alta presión", una inyección rápida (en un lapso de 5 segundos) de un gran volumen de ARNpi que contiene solución en animales a través de la vena de la cola (Liu et al., 1999, ver más arriba; McCaffrey et al., 2002, ver más arriba; Lewis et al., 2002. También se pueden utilizar nanopartículas y liposomas para suministrar ARNpi en animales.

Las composiciones de ácido nucleico de la invención incluyen tanto los ARNpi modificados como los ARNpi no modificados como se conoce en la técnica, tales como los derivados de ARNpi entrecruzados o los derivados que tienen fracciones que no son de nucleótidos enlazadas, por ejemplo a sus extremos 5' o 3'. La modificación de los derivados de ARNpi de esta forma puede mejorar la captación celular o mejorar las actividades de direccionamiento celular de los derivados de ARNpi resultantes en comparación con el ARNpi correspondiente, son útiles para rastrear el derivado de ARNpi en la célula, o mejorar la estabilidad del derivado de ARNpi en comparación con el correspondiente ARNpi.

Los precursores de ARN modificados por ingeniería genética, introducidos en las células u organismos enteros tal como se describe en el presente documento, dará lugar a la producción de una molécula de ARNpi deseada. Tal molécula de ARNpi se asociará entonces con los componentes de proteína endógena de la ruta del ARNi para unirse y dirigir una secuencia de ARNm específica para la escisión y destrucción. De esta manera, el ARNm que es dirigido por el ARNpi generado a partir del precursor de ARN modificado por ingeniería genética se agotará en la célula u organismo, llevando a una disminución en la concentración de la proteína codificada por ese ARNm en la célula u organismo. Los precursores de ARN son típicamente moléculas de ácido nucleico que codifican de forma individual o bien una cadena de un ARNbc o codifican toda la secuencia de nucleótidos de una estructura de bucle en horquilla de ARN.

Las composiciones de ácido nucleico de la invención pueden ser no conjugadas o pueden conjugarse con otra fracción, tal como una nanopartícula, para mejorar una propiedad de las composiciones, por ejemplo, un parámetro farmacocinético tal como la absorción, la eficacia, la biodisponibilidad, y / o la vida media. La conjugación puede realizarse por métodos conocidos en la técnica, por ejemplo, utilizando los métodos de Lambert et al., *Drug Deliv. Rev.*: 47 (1), 99 - 112 (2001) (describe ácidos nucleicos cargados en nanopartículas de polialquiloacrilato (PACA)); . Fattal et al., *J. Control Release* 53 (1-3): 137 - 43 (1998) (describe ácidos nucleicos unidos a nanopartículas); Schwab et al., *Ann. Oncol. 5 Suppl. 4*: 55 - 8 (1994) (describe ácidos nucleicos ligados a agentes de intercalación, grupos hidrófobos, policationes o nanopartículas de PACA); y Godard et al., *Eur. J. Biochem.* 232 (2): 404 - 10 (1995) (describe ácidos nucleicos ligados a nanopartículas).

Las moléculas de ácido nucleico de la presente invención también se pueden marcar usando cualquier método conocido en la técnica; por ejemplo, las composiciones de ácido nucleico pueden marcarse con un fluoróforo, por ejemplo, Cy3, fluoresceína, o rodamina. La marcación puede llevarse a cabo utilizando un kit, por ejemplo, el kit de marcación de ARNpi SILENCER^{MR} (Ambion). Además, se puede marcar en forma radioactiva el ARNpi, por ejemplo, utilizando ³H, ³²P, u otro isótopo apropiado.

Además, debido a que se cree que el ARNi progresa a través de al menos un producto intermedio de ARN monocatenario, la persona capacitada en la técnica apreciará que también se puede diseñar ARNpi-mc (por ejemplo, la cadena antisentido de un ARNpi-bc) (por ejemplo, mediante síntesis química), generar (por ejemplo, generar enzimáticamente) o expresar (por ejemplo, a partir de un vector o un plásmido) como se describe en el presente documento y utilizado de acuerdo a las metodologías reivindicadas. Además, en invertebrados, el ARNi puede ser activado de manera efectiva mediante los ARNbc largos (por ejemplo, los ARNbc de aproximadamente 100 - 1000 nucleótidos de longitud, preferiblemente de aproximadamente 200 - 500, por ejemplo, de aproximadamente 250, 300, 350, 400 o 450 nucleótidos de longitud) que actúan como efectores de ARNi. (Brondani et al., *Proc Natl Acad Sci, EE.UU.*, 2001, 4 de diciembre; 98 (25): 14428 - 33. Epub 2001, 27 de noviembre).

VI. Métodos para introducir los ARN, vectores y células huésped

Los métodos físicos para la introducción de ácidos nucleicos incluyen la inyección de una solución que contiene el ARN, el bombardeo por partículas cubiertas por el ARN, remojar la célula o el organismo en una solución del ARN, o la electroporación de las membranas celulares en presencia del ARN. Un constructo viral empaquetado en una partícula viral conseguiría tanto introducción eficiente de un constructo de expresión en la célula como la transcripción del ARN codificado por el constructo de expresión. Se pueden utilizar otros métodos conocidos en la técnica para introducir ácidos nucleicos en las células, tales como el transporte por el portador mediado por lípidos, el transporte mediado por productos químicos, tal como fosfato de calcio, y similares. Por lo tanto, se puede introducir el ARN junto con componentes que realizan una o más de las siguientes actividades: mejorar la absorción de ARN por la célula, inhibir la hibridación de cadenas sencillas, estabilizar las cadenas sencillas, o bien aumentar la inhibición del gen objetivo.

El ARN puede ser introducido directamente en la célula (es decir, intracelularmente); o introducido extracelularmente en una cavidad, espacio intersticial, en la circulación de un organismo, introducido oralmente, o puede ser introducido mediante el baño de una célula u organismo en una solución que contiene el ARN. La circulación vascular o extravascular, la sangre o el sistema linfático, y el líquido cefalorraquídeo son sitios en los que se puede introducir el ARN.

La célula que tiene el gen objetivo puede ser de la línea germinal o somática, totipotente o pluripotente, de división o de no división, parénquima o epitelio, inmortalizada o transformada, o similares. La célula puede ser una célula madre o una célula diferenciada. Los tipos de células que se diferencian incluyen adipocitos, fibroblastos, miocitos, cardiomiocitos, endotelio, neuronas, glía, células sanguíneas, megacariocitos, linfocitos, macrófagos, neutrófilos, eosinófilos, basófilos, mastocitos, leucocitos, granulocitos, queratinocitos, condrocitos, osteoblastos, osteoclastos, hepatocitos, y células de las glándulas endocrinas o exocrinas.

Dependiendo del gen objetivo particular y de la dosis de material de ARN bicatenario suministrado, este proceso puede proporcionar una pérdida parcial o completa de la función para el gen objetivo. Una reducción o pérdida de la expresión génica en al menos 50%, 60%, 70%, 80%, 90%, 95% o 99% o más de las células objetivo es ejemplar. La

inhibición de la expresión génica se refiere a la ausencia (o disminución observable) en el nivel de proteína y / o del producto de ARNm de un gen objetivo. La especificidad se refiere a la capacidad de inhibir el gen objetivo sin efectos evidentes sobre otros genes de la célula. Las consecuencias de la inhibición se pueden confirmar mediante el examen de las propiedades exteriores de la célula u organismo (como se presenta a continuación en los ejemplos) o por técnicas bioquímicas tales como hibridación de la solución de ARN, la protección de la nucleasa, hibridación tipo Northern, transcripción inversa, el seguimiento de la expresión génica con un microarreglo, enlazamiento de anticuerpos, ensayo inmunoabsorbente ligado a enzima (ELISA), transferencia tipo Western, radioinmunoensayo (RIA), otros inmunoensayos, y análisis de células activadas por fluorescencia (FACS).

Para la inhibición mediada por ARN en una línea celular o de todo el organismo, se ensaya convenientemente la expresión del gen mediante el uso de un gen reportero o de resistencia a fármacos cuyo producto proteico se ensaya fácilmente. Tales genes reporteros incluyen la acetohidroxiácido sintasa (AHAS), fosfatasa alcalina (AP), beta galactosidasa (LacZ), beta glucuronidasa (GUS), cloranfenicol acetiltransferasa (CAT), proteína fluorescente verde (GFP), peroxidasa de rábano picante (HRP), luciferasa (Luc), nopalina sintasa (NOS), octopina sintasa (OCS), y derivados de los mismos. Múltiples marcadores seleccionables se encuentran disponibles que confieren resistencia a ampicilina, bleomicina, cloranfenicol, gentamicina, higromicina, kanamicina, lincomicina, metotrexato, fosfinotricina, puromicina, y tetraciclina. Dependiendo del ensayo, la cuantificación de la cantidad de expresión génica permite determinar un grado de inhibición que es mayor al 10%, 33%, 50%, 90%, 95% o 99% en comparación con una célula no tratada de acuerdo con la presente invención. Las dosis más bajas de material inyectado y los tiempos más largos después de la administración del agente de ARNi pueden resultar en la inhibición en una fracción más pequeña de células (por ejemplo, al menos 10%, 20%, 50%, 75%, 90%, o 95% de las células objetivo). La cuantificación de la expresión génica en una célula puede mostrar cantidades similares de inhibición a nivel de la acumulación de ARNm objetivo o traducción de la proteína objetivo. Como ejemplo, la eficiencia de la inhibición puede determinarse mediante la evaluación de la cantidad de producto génico en la célula; el ARNm puede ser detectado con una sonda de hibridación que tenga una secuencia de nucleótidos fuera de la región usada para el ARN bicatenario inhibidor, o el polipéptido traducido puede ser detectado con un anticuerpo producido contra la secuencia de polipéptido de esa región.

El ARN se puede introducir en una cantidad que permite el suministro de al menos una copia por célula. Las dosis más altas (por ejemplo, por lo menos 5, 10, 100, 500 o 1.000 copias por célula) de material pueden producir una inhibición más efectiva; dosis más bajas también pueden ser útiles para aplicaciones específicas.

La eficacia de un agente de ARNi de la invención (por ejemplo, un ARNpi dirigido a un polimorfismo en un gen mutante) puede ser analizado por su capacidad para degradar específicamente ARNm mutante (por ejemplo, ARNm de htt mutante y / o la producción de proteína Huntingtina mutante) en las células, en particular, en las neuronas (por ejemplo, líneas clonales neuronales corticales o de cuerpo estriado y / o de neuronas primarias). También son adecuados para ensayos de validación con base en células otras células fácilmente transfectables, por ejemplo, células HeLa o células COS. Las células se transfectan con los ADNc mutantes o de tipo silvestre humanos (por ejemplo, ADNc de Huntingtina mutante o de tipo silvestre humano). Se cotransfectan ARNpi estándar, ARNpi modificado o vectores capaces de producir ARNpi modificadas a partir de ARNm de bucle en forma de U. Se mide la reducción selectiva en ARNm mutante (por ejemplo, ARNm de Huntingtina mutante) y / o la proteína mutante (por ejemplo, la Huntingtina mutante). La reducción de ARNm mutante o de proteína se puede ser comparado con los niveles de ARNm normal o de proteína. El ARNm normal o la proteína introducidos exógenamente (o el ARNm normal exógeno o la proteína) se pueden analizar para propósitos de comparación. Cuando se utilizan células neuronales, que se sabe que son algo resistentes a las técnicas de transfección estándar, puede ser deseable introducir agentes de ARNi (por ejemplo, los ARNpi) mediante absorción pasiva.

VII. Métodos de tratamiento:

La presente invención permite métodos tanto profilácticos como terapéuticos para tratar a un sujeto en riesgo de (o susceptible a) una enfermedad o trastorno causado, en su totalidad o en parte, por una proteína mutante de ganancia de función. La enfermedad o el trastorno pueden ser una enfermedad o trastorno causados por repetición de trinucleótidos. La enfermedad o el trastorno puede ser un trastorno por poliglutaminas. La enfermedad o el trastorno puede ser un trastorno asociado con la expresión de Huntingtina y en el que la alteración de la Huntingtina, especialmente la amplificación del número de copias de repetición de CAG, conduce a un defecto en el gen de la Huntingtina (estructura o función) o proteína Huntingtina (estructura o función o expresión), de tal manera que las manifestaciones clínicas incluyen aquellas observadas en pacientes con enfermedad de Huntington.

"Tratamiento", o "tratar" como se usa en el presente documento, se definen como la aplicación o administración de un agente terapéutico (por ejemplo, un agente de ARN o vector o transgén que lo codifica) a un paciente, o la aplicación o administración de un agente terapéutico a un tejido aislado o línea celular de un paciente, que padece la enfermedad o el trastorno, un síntoma de enfermedad o del trastorno o una predisposición hacia una enfermedad o trastorno, con el propósito de curar, sanar, aliviar, mitigar, alterar, remediar, mejorar, medrar o afectar la enfermedad o el trastorno, los síntomas de la enfermedad o el trastorno, o la predisposición hacia la enfermedad.

La invención proporciona un método para prevenir en un sujeto, una enfermedad o trastorno como se describió anteriormente, mediante la administración al sujeto de un agente terapéutico (por ejemplo, un agente de ARNi o vector o transgén que lo codifique). Los sujetos con riesgo de padecer la enfermedad pueden ser identificados, por ejemplo, mediante cualquiera o una combinación de ensayos de diagnóstico o pronóstico como se describe aquí. La administración de un agente profiláctico puede presentarse antes de la manifestación de los síntomas característicos de la enfermedad o el trastorno, de manera que se evitan la enfermedad o el trastorno o, alternativamente, se retrasa su progreso.

La invención se refiere a métodos de tratamiento de sujetos terapéuticamente, es decir, la alteración del inicio de los síntomas de la enfermedad o el trastorno. La invención puede implicar poner en contacto una célula que expresa un mutante de ganancia de función con un agente terapéutico (por ejemplo, un agente de ARNi o un vector o transgén que codifica al mismo) que es específico para un polimorfismo en el gen, de tal manera que se consigue la interferencia específica de la secuencia con el gen. Estos métodos pueden realizarse *in vitro* (por ejemplo, cultivando la célula con el agente) o, alternativamente, *in vivo* (por ejemplo, mediante la administración del agente a un sujeto).

Con respecto tanto a los métodos profilácticos como a los terapéuticos de tratamiento, tales tratamientos pueden ser adaptados o modificados específicamente, con base en el conocimiento obtenido del campo de la farmacogenómica. "Farmacogenómica", como se usa aquí, se refiere a la aplicación de tecnologías genómicas tales como la secuenciación de genes, genética estadística, y análisis de expresión génica a fármacos en desarrollo clínico y en el mercado. Más específicamente, el término se refiere al estudio de cómo los genes de un paciente determinan su respuesta a un medicamento (por ejemplo, el "fenotipo de respuesta a los fármacos", o el "genotipo de respuesta al fármaco" de un paciente). Por lo tanto, la invención proporciona métodos para adaptar el tratamiento profiláctico o el tratamiento terapéutico de un individuo, ya sea con las moléculas del gen objetivo de la presente invención o los moduladores de genes objetivo de acuerdo con el genotipo de respuesta al fármaco de ese individuo. La farmacogenómica permite que un médico o especialista clínico dirija tratamientos profilácticos o terapéuticos a los pacientes que más se beneficiarán del tratamiento y para evitar el tratamiento de los pacientes que experimentan efectos secundarios relacionados con fármacos tóxicos.

Los agentes terapéuticos pueden ser probados en un modelo animal apropiado. Por ejemplo, se puede utilizar un agente de ARNi (o vector de expresión o transgén que lo codifica) como se describe en la presente memoria en un modelo animal para determinar la eficacia, toxicidad, o efectos secundarios del tratamiento con dicho agente. Alternativamente, se puede utilizar un agente terapéutico en un modelo animal para determinar el mecanismo de acción de tal agente. Por ejemplo, se puede utilizar un agente en un modelo animal para determinar la eficacia, toxicidad, o efectos secundarios del tratamiento con tal agente. Alternativamente, se puede utilizar un agente en un modelo animal para determinar el mecanismo de acción de tal agente.

VIII. Composiciones farmacéuticas

La invención se refiere a los usos de los agentes descritos anteriormente para tratamientos profilácticos y / o terapéuticos como se describe más adelante. En consecuencia, los moduladores (por ejemplo, agentes de ARNi) de la presente invención se pueden incorporar en composiciones farmacéuticas adecuadas para administración. Tales composiciones típicamente comprenden la molécula de ácido nucleico, proteína, anticuerpo, o compuesto modulador y un vehículo farmacéuticamente aceptable. Como se usa aquí, la expresión "vehículo farmacéuticamente aceptable" pretende incluir cualquiera y todos los disolventes, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y retardantes de la absorción, y similares, compatibles con la administración farmacéutica. El uso de tales medios y agentes para sustancias farmacéuticamente activas es bien conocido en la técnica. Excepto en la medida en que cualquier medio o agente convencional sea incompatible con el compuesto activo, su uso en las composiciones está contemplado. Compuestos activos suplementarios también se pueden incorporar en las composiciones.

Una composición farmacéutica de la invención se formula para que sea compatible con su vía de administración prevista. Los ejemplos de rutas de administración incluyen parenteral, por ejemplo, intravenosa, intradérmica, subcutánea, intraperitoneal, intramuscular, oral (por ejemplo, inhalación), transdérmica (tópica), y la administración transmucosa. Las soluciones o suspensiones usadas para aplicación parenteral, intradérmica, o subcutánea pueden incluir los siguientes componentes: un diluyente estéril tal como agua para inyección, solución salina, aceites fijos, polietilenglicoles, glicerina, propilenglicol u otros disolventes sintéticos; agentes antibacterianos tales como alcohol bencílico o metil parabenos; antioxidantes tales como ácido ascórbico o bisulfito de sodio; agentes quelantes tales como ácido etilendiaminotetraacético; amortiguadores tales como acetatos, citratos o fosfatos y agentes para el ajuste de la tonicidad tales como cloruro de sodio o dextrosa. El pH puede ajustarse con ácidos o bases, tales como ácido clorhídrico o hidróxido de sodio. La preparación parenteral puede tenerse en ampollas, jeringas desechables o viales de múltiples dosis hechos de vidrio o plástico.

Las composiciones farmacéuticas adecuadas para uso inyectable incluyen soluciones acuosas estériles (solubles en agua) o dispersiones y polvos estériles para la preparación extemporánea de soluciones o dispersiones inyectables estériles. Para la administración intravenosa, los vehículos adecuados incluyen solución salina fisiológica, agua

5 bacteriostática, Cremóforo EL^{MR} (BASF, Parsippany, NJ) o amortiguador salino de fosfato (PBS). En todos los casos, la composición debe ser estéril y debe ser fluida de tal manera que permita ser inyectada fácilmente. Debe ser estable bajo las condiciones de fabricación y almacenamiento y debe conservarse frente a la acción contaminante de microorganismos tales como bacterias y hongos. El vehículo puede ser un disolvente o medio de dispersión que contiene, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol, y polietilenglicol líquido, y similares), y mezclas adecuadas de los mismos. La fluidez apropiada puede mantenerse, por ejemplo, mediante el uso de un recubrimiento tal como lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de una dispersión y mediante el uso de tensoactivos. La prevención de la acción de microorganismos puede lograrse mediante diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido ascórbico, timerosal, y similares. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares, polialcoholes tales como manitol, sorbitol, cloruro de sodio en la composición. La absorción prolongada de las composiciones inyectables puede lograrse incluyendo en la composición un agente que retrase la absorción, por ejemplo, monoestearato de aluminio y gelatina.

15 Las soluciones inyectables estériles pueden prepararse mediante la incorporación del compuesto activo en la cantidad requerida en un disolvente apropiado con uno o una combinación de ingredientes enumerados anteriormente, según se requiera, seguido de esterilización por filtración. Generalmente, las dispersiones se preparan incorporando el compuesto activo en un vehículo estéril que contiene un medio de dispersión básico y los otros ingredientes requeridos de aquellos enumerados anteriormente. En el caso de polvos estériles para la preparación de soluciones inyectables estériles, los métodos preferidos de preparación son secado al vacío y secado por congelación que produce un polvo del ingrediente activo más cualquier ingrediente adicional deseado a partir de una solución previamente filtrada estéril del mismo.

25 Las composiciones orales generalmente incluyen un diluyente inerte o un vehículo comestible. Se pueden colocar en cápsulas de gelatina o comprimirse en tabletas. Para el propósito de la administración terapéutica oral, el compuesto activo puede incorporarse con excipientes y usarse en forma de comprimidos, pastillas, o cápsulas. Las composiciones orales también se pueden preparar usando un vehículo fluido para uso como un enjuague bucal, en el que el compuesto en el vehículo fluido se aplica por vía oral y se hacen buches y se expectora o se traga. Agentes de enlazamiento farmacéuticamente compatibles, y / o se pueden incluir materiales adyuvantes como parte de la composición. Los comprimidos, píldoras, cápsulas, pastillas y similares pueden contener cualquiera de los siguientes ingredientes, o compuestos de una naturaleza similar: un aglutinante tal como celulosa microcristalina, goma de tragacanto o gelatina; un excipiente tal como almidón o lactosa, un agente disgregante tal como ácido algínico, Primogel, o almidón de maíz; un lubricante tal como estearato de magnesio o Sterotes; un deslizante tal como dióxido de silicio coloidal; un agente edulcorante tal como sacarosa o sacarina; o un agente aromatizante tal como menta, salicilato de metilo, o aromatizante de naranja.

40 Para la administración por inhalación, los compuestos se suministran en forma de un atomizador de aerosol desde un recipiente o dispensador presurizado que contiene un propulsor adecuado, por ejemplo, un gas tal como dióxido de carbono, o un nebulizador.

45 La administración sistémica también puede ser por vía transmucosa o transdérmica. Para administración transmucosa o transdérmica, se usan penetrantes apropiados para la barrera que va a ser permeada en la formulación. Tales penetrantes son generalmente conocidos en la técnica, e incluyen, por ejemplo, para administración transmucosa, detergentes, sales biliares, y derivados de ácido fusídico. La administración transmucosa se puede lograr a través del uso de atomizadores nasales o supositorios. Para administración transdérmica, los compuestos activos se formulan en pomadas, ungüentos, geles, o cremas como se conoce generalmente en la técnica.

50 Los compuestos también se pueden preparar en forma de supositorios (por ejemplo, con bases de supositorio convencionales tales como manteca de cacao y otros glicéridos) o enemas de retención para administración rectal.

55 Los compuestos activos se pueden preparar con vehículos que protegerán el compuesto contra la eliminación rápida del cuerpo, tales como una formulación de liberación controlada, incluyendo implantes y sistemas de suministro microencapsulados. Se pueden usar polímeros biodegradables, biocompatibles, tales como acetato de vinil etileno, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres, y ácido poliláctico. Los métodos para la preparación de tales formulaciones serán evidentes para aquellos capacitados en la técnica. Los materiales también se pueden obtener comercialmente a través de Alza Corporation y Nova Pharmaceuticals, Inc. Las suspensiones liposomales (incluyendo liposomas dirigidos a células infectadas con anticuerpos monoclonales para antígenos virales) también se pueden utilizar como portadores farmacéuticamente aceptables. Estos se pueden preparar de acuerdo con métodos conocidos por aquellos capacitados en la técnica, por ejemplo, como se describe en la patente de los Estados Unidos No. 4.522.811.

65 Es especialmente ventajoso formular composiciones orales o parenterales en forma de unidad de dosificación para facilidad de administración y uniformidad de dosificación. La forma de unidad de dosificación como se usa en el presente documento se refiere a unidades físicamente discretas adecuadas como dosificaciones unitarias para el

5 sujeto a tratar; conteniendo cada unidad una cantidad predeterminada del compuesto activo calculada para producir el efecto terapéutico deseado junto con el vehículo farmacéutico requerido. La especificación para las formas unitarias de dosificación de la invención están dictadas por y dependen directamente de las características únicas del compuesto activo y el efecto terapéutico particular a conseguir, y las limitaciones inherentes en la técnica de formación de composiciones de tal compuesto activo para el tratamiento de individuos.

10 La toxicidad y la eficacia terapéutica de tales compuestos se puede determinar mediante procedimientos farmacéuticos estándar en cultivos celulares o animales experimentales, por ejemplo, para determinar la LD50 (la dosis letal para el 50% de la población) y la ED50 (la dosis terapéuticamente efectiva en el 50% de la población). La relación de la dosis entre los efectos tóxicos y terapéuticos es el índice terapéutico y puede expresarse como la relación LD50/ED50. Se prefieren los compuestos que exhiben índices terapéuticos grandes. Aunque los compuestos que exhiben efectos tóxicos secundarios pueden utilizarse, se debe tener cuidado en diseñar un sistema de administración que dirija dichos compuestos al sitio del tejido afectado con el fin de minimizar el daño potencial a células no infectadas y, de ese modo, reducir los efectos secundarios.

15 Los datos obtenidos a partir de los ensayos de cultivo celular y estudios en animales pueden usarse para formular un intervalo de dosificación para uso en seres humanos. La dosificación de tales compuestos se encuentra preferiblemente dentro de un intervalo de concentraciones circulantes que incluyen la ED50 con poca o ninguna toxicidad. La dosificación puede variar dentro de este intervalo dependiendo de la forma de dosificación empleada y la vía de administración utilizada. Para cualquier compuesto usado en el método de la invención, la dosis terapéuticamente efectiva puede estimarse inicialmente a partir de ensayos de cultivo celular. Se puede formular una dosis en modelos animales para conseguir un intervalo de concentración en plasma circulante que incluya la EC50 (es decir, la concentración del compuesto de ensayo que logra la mitad de una respuesta máxima) tal como se determina en cultivo celular. Tal información se puede utilizar para determinar con mayor precisión las dosis útiles en seres humanos. Los niveles en plasma se pueden medir, por ejemplo, por cromatografía líquida de alto rendimiento.

Las composiciones farmacéuticas pueden incluirse en un recipiente, envase, o dispensador junto con instrucciones para la administración.

30 Esta invención se ilustra adicionalmente mediante los siguientes ejemplos que no deben interpretarse como limitantes.

Ejemplos

35 A diferencia de otros tipos de enfermedades autosómicas dominantes, la enfermedad de Huntington no contiene una mutación puntual, por ejemplo) el cambio de un solo nucleótido. Por lo tanto, la estrategia para diseñar un ARNpi dirigido contra una mutación puntual en el alelo de la enfermedad no se puede implementar. En lugar de ello, la presente invención dirige los ARNpi diseñados contra los polimorfismos en el gen de la Huntingtina, de los cuales hay unos 30 disponibles en el GenBank. La presente invención también identifica el polimorfismo en el alelo de la enfermedad de Huntington que difiere del alelo de tipo silvestre, de manera que el ARNpi destruye sólo el ARNm de la enfermedad y deja intacto el ARNm del alelo de tipo silvestre (normal). Por lo tanto, sólo se destruye la proteína Huntingtina mutante y la proteína normal queda intacta.

Ejemplo I: Pruebas de agentes de ARNi (por ejemplo, los ARNpi) contra htt mutante en lisados de *Drosophila*

45 Se diseñó un ARNpi dirigido a la posición 2886 en el ARNm de htt como se describió anteriormente. La secuencia del ARNpi se representa en la Figura 5a (SEQ ID NO: 24 sentido; 25 antisentido). Se desprotegió el ARN sintético (Dharmacon) de acuerdo con el protocolo del fabricante. Se hibridaron cadenas de ARNpi (Elbashir et al., 2001a).

50 Se prepararon los ARN objetivo como sigue. Se transcribieron los ARN objetivo con ARN polimerasa T7 recombinante, marcada con histidina, a partir de productos de PCR como se describe (Nykänen et al., 2001; Hutvagner et al., 2002). Las plantillas de PCR para htt sentido y antisentido se generaron por amplificación de 0,1 ng / ml (concentración final) de la plantilla de plásmido que codifica el ADNc de htt usando los siguientes pares de cebadores: objetivo sentido de htt, 5'-GCG TAA TAC GAC TCA CTA TAG GAA CAG TAT GTCTCA GAC ATC-3' (SEQ ID NO: 30) y 5'-AAG UUCG UAU UCC GCG UAC GU-3' (SEQ ID NO: 31); objetivo anti-sentido de htt, 5'-GCG TAA TAC GAG TCA CTA TAG GAC AAG CCT AAT TAG TGA TGC-3' (SEQ ID NO: 32) y 5'-GAA CAG TAT GTC TCA GAC ATC-3' (SEQ ID NO: 33).

60 Se analizó el ARNpi usando un ensayo *in vitro* de ARNi, con lisados de embriones de *Drosophila*. Las reacciones *in vitro* y el análisis de ARNi se llevaron a cabo como se describió anteriormente (Tuschl et al., 1999; Zamore et al., 2000; Haley et al., 2003). Se usaron los ARN objetivo en concentraciones de ~ 5 nM de modo que las reacciones se encuentran principalmente bajo condiciones de una sola rotación. La escisión del objetivo bajo estas condiciones es proporcional a la concentración de ARNpi.

65 La figura 5a muestra la eficiencia del ARNpi dirigido contra la posición 2886 en el htt mutante. Los datos demuestran claramente que el ARNpi dirige la escisión del objetivo sentido hasta un mayor grado que el observado para el

objetivo antisentido. Sin embargo, se observa que este ARNpi diseñado primero no produjo una molécula muy activa, al menos en este ensayo *in vitro*. El análisis termodinámico de la resistencia del par de bases en los dos extremos del dúplex de ARNpi indicaron resistencias de las pares de bases más o menos equivalentes. La Figura 4 representa el análisis termodinámico de ARNpi sentido (SEQ ID NO: 20; 22, respectivamente) y antisentido (SEQ ID NO: 21; 23, respectivamente) de los extremos 5' de la cadena para el dúplex de ARNpi en 5a. Se calculó ΔG (kcal / mol) en NaCl 1 M a 37 °C.

Para mejorar la eficacia del dúplex de ARNpi diseñado, se modificó el extremo 5' de la cadena sentido o la posición 19 de la cadena antisentido del ARNpi de htt probado en la Figura 5a para producir dúplex de ARNpi en los que el extremo 5' de la cadena sentido estaba ya sea totalmente desapareada (Figura 5c; SEQ ID NO: 28 sentido, SEQ ID NO: 29 antisentido) o en un par de bases A:U (Figura 5b; SEQ ID NO: 26 sentido, SEQ ID NO: 27 antisentido). El desapareamiento del extremo 5' de una cadena de ARNpi - la cadena sentido - en este caso hace que la cadena funcione para la exclusión de la otra cadena. Cuando el extremo 5' de la cadena sentido de htt estaba presente en un par de bases A:U y el extremo 5' de la cadena antisentido de htt estaba en un par G:C, la cadena sentido dominó la reacción (Figura 5b-c), pero la cadena antisentido de htt retuvo una actividad similar a la observada para el ARNpi originalmente diseñado.

Ejemplo II: ARNi desactivado de la proteína htt en células cultivadas

En un primer experimento, se ensayaron los ARNpi que dirigen un polimorfismo en el ARNm de htt (es decir, el polimorfismo en la posición 2886 en el ARNm de htt) para determinar su capacidad para subregular la proteína htt endógena en células HeLa. Las células HeLa se cultivaron y transfectaron de la siguiente manera. Se mantuvieron células HeLa a 37 °C en medio de Eagle modificado de Dulbecco (DMEM, Invitrogen) suplementado con 10% de suero bovino fetal (FBS), 100 unidades / ml de penicilina y 100 μ g / ml de estreptomycin (Invitrogen). Se pasaron regularmente las células en subconfluencia y se sembraron hasta un 70% de confluencia 16 horas antes de la transfección. La transfección transitoria mediada por Lipofectamina^{MR} (Invitrogen) de los ARNpi se realizó en placas de 6 pozos por duplicado (Falcon) como se describe para líneas de células adherentes por el fabricante. Se añadió una mezcla de transfección estándar que contenía ARNpi 100-150 nM y 9-10 μ l de Lipofectamina^{MR} en 1 ml de OPTI-MEM® (Invitrogen) reducido en suero a cada pozo. Se incubaron las células en mezcla de transfección a 37 °C durante 6 horas y después se cultivaron en DMEM libre de antibióticos. Para el análisis de transferencias tipo Western en diferentes intervalos de tiempo, se recogieron las células transfectadas, se lavaron dos veces con solución salina amortiguada con fosfato (PBS, Invitrogen), se congelaron rápidamente en nitrógeno líquido, y se almacenaron a -80 °C para el análisis.

Se ensayaron tres ARNpi contra a una secuencia objetivo común en el exón 1 y cuatro ARNpi se ensayaron para el polimorfismo de la posición 2886. Se realizó el análisis de transferencias tipo Western como sigue. Se recogieron las células tratadas con ARNpi como se describió anteriormente y se lisaron en amortiguador de lisis del reportero enfriado en hielo (Promega) que contenía inhibidor de proteasa (completo, libre de EDTA, 1 tableta/ 10 ml de amortiguador, Roche Molecular Biochemicals). Después del aclaramiento de los lisados resultantes por centrifugación, se cuantificó la proteína en los lisados claros mediante el kit de ensayo de proteína Dc (Bio-Rad). Se resolvieron las proteínas en 60 μ g de lisado celular total mediante 10% de SDS - PAGE, se transfirieron a una membrana de difluoruro de polivinilideno (PVDF, Bio-Rad), y se inmunotransfirieron con anticuerpos contra CD80 (Santa Cruz). Se visualizó el contenido de proteína con un kit de transferencia de quimioluminiscencia BM (Roche Molecular Biochemicals). Se expusieron las transferencias a una película de rayos X (Kodak MR-1) durante diferentes tiempos (30 s hasta 5 min). La Figura 6a presenta los resultados del análisis Western. La tubulina sirvió como control de carga. Los datos se cuantificaron y normalizaron en la Figura 6b. De los ARNpi analizados, 2886-4, mostraron en forma reproducible una mayor eficacia en las células HeLa cultivadas (Figura 6). Este ARNpi también mostró en forma reproducible una eficacia mejorada *in vitro* (no se muestra). El ARNpi de GFP es un ARNpi de control que no comparte homología de secuencia con el ARNm de htt.

Se pueden analizar igualmente los ARNpi contra regiones polimórficas en el ARNm de htt en células transfectadas con ADNc de htt humano o en células transfectadas con constructos del reportero de htt. Cotransfecciones transitorias mediadas por Lipofectamina^{MR} (Invitrogen) de los ADNc o plásmidos reporteros y los ARNpi se llevan a cabo como se describió anteriormente. Para analizar la capacidad de los ARNpi para dirigir los constructos reportados de htt, se utilizó ARNi para inhibir la expresión de GFP-htt en líneas de células HeLa humanas cultivadas. En resumen, se transfectaron las células HeLa con dúplex de ARNpi de GFP-htt, dirigiendo la secuencia de ARNm de GFP-htt. Para analizar los efectos de ARNi contra GFP-htt, se prepararon lisados de células tratadas con dúplex de ARNpi dúplex en diversos momentos después de la transfección. Se llevaron a cabo experimentos de transferencias tipo Western como se describió anteriormente. En resumen, se recogieron las células HeLa en diversos momentos después de la transfección, se resolvió su contenido de proteína en 10% de SDS-PAGE, se transfirieron a membranas de PVDF, y se hicieron inmunotransferencias con anticuerpos apropiados. Los resultados de este estudio indicaron que el ARNpi contra GFP puede eliminar la expresión de la expresión de GFP-htt en las células HeLa transfectadas con el gen GFP-htt. Para los estudios dirigidos se introduce exógenamente htt, los procedimientos son como se describió, excepto que los anticuerpos anti-Htt se utilizan para inmunotransferencia.

El ARNi se puede utilizar para inhibir la expresión de htt en las células neuronales cultivadas también. Los ejemplos de células incluyen líneas de células PC12 (Scheitzer et al., Thompson et al.) y NT3293 (Tagle et al.) como se describió anteriormente. Los ejemplos de células adicionales incluyen células transfectadas de forma estable, por ejemplo, células neuronales o células derivadas neuronalmente. Las líneas de células PC12 que expresan el exón 1 del gen de la Huntingtina humana (Htt) se pueden usar aunque la expresión del exón 1 reduce la supervivencia de las células. Las células PC12 con GFP-Htt que tienen un gen GFP-Htt inducible también se pueden usar para analizar o validar la eficacia de ARNpi.

Ejemplo III: suministro de ARNpi de Htt en un entorno *in vivo*

Los modelos de ratones R6/2 (que expresan el producto de ADNc de htt humano htt de R6/2) son un modelo animal aceptado para estudiar la eficacia del suministro de ARNpi en un entorno *in vivo*. Se utilizaron ratones R6/2 modificados por ingeniería genética para probar la eficacia de ARNpi en el extremo 5' del ARNm de Huntingtina. Se inyectó ARNpi de Htt en el cuerpo estriado de ratones R6/2 a través de una bomba Alzet. Los ratones fueron tratados durante 14 días con el sistema de suministro de la bomba Alzet / ARNpi.

Los resultados de este estudio indicaron que dos ratones que recibieron el ARNpi con Trans-IT TKO (Mirus), ya sea como una solución de 20 o 200 nM a razón de 0,25 µl/hora no mostraron deterioro de la discapacidad motora desde el día 67 al día 74. En general, se espera que estos R6/2 tengan una reducción continua en una barra en rotación más allá del día 60.

Referencias

Aronin et al., *Neuron*, Nov; 15(5): 1193 - 201 (1995)
 Aronin et al., *Phil Trans Royal Society*, June 29; 354 (1386): 995 - 1003 (1999)
 Bagella et al., *J. Cell. Physiol.* 177: 206 - 213 (1998)
 Brummelkamp et al., *Science* 296: 550 - 553 (2002)
 Calegari et al., *Proc. Natl. Acad. Sci. USA* 99(22): 14236 - 40 (2002)
 Difiglia et al., *Science*, Sep 26;277(5334): 1990 - 3 (1997)
 Elbashir et al., *Genes Dev* 15, 188 - 200 (2001a)
 Haley et al., *Methods* 30, 330 - 336 (2003)
 Hutvagner and Zamore, *Science* 297, 2056 - 2060 (2002)
 Jacque et al., (2002)
 Kremer et al., (1994)
 Laforet et al., *J. Neurosci.*, Dec 1; 21(23): 9112 - 23 (2001)
 Lee et al., *EMBO J.* 21: 4663 - 4670 (2002)
 Lewis et al., *Nature Genetics* 32: 107 - 108 (2002)
 Lin et al., (1993)
 Liu et al., (1999)
 McCaffrey et al., *Gene Ther.* 2002 Dec; 9(23): 1563 (2002)
 McManus et al., *RNA* 8, 842 - 850 (2002)
 Miyagishi et al., *Nature Biotechnol.* 20: 497 - 500 (2002)
 Nykänen et al., *Cell* 107, 309 - 321 (2001)
 Paddison et al., *Genes Dev* 16, 948 - 958 (2002)
 Paul et al., *Nat Biotechnol* 20, 505 - 508 (2002)
 Scheitzer et al.
 Sui et al., *Proc Natl Acad Sci USA* 99, 5515 - 5520 (2002)
 Tagle et al.
 Thompson et al.
 Tuschl, T., *Nat Biotechnol.* 2002 May; 20(5): 446 - 8 (2002)
 Tuschl et al., *Genes Dev* 13, 3191 - 3197 (1999)
 Xia et al., (2002)
 Yohrling G.J. et al., *Mol Cell Neurosci.* May; 23(1): 28 - 38 (2003)
 Yu et al., *Proc Natl Acad Sci USA* 99, 6047 - 6052 (2002)
 Zamore et al., *Cell* 101, 25 - 33 (2000)
 Zamore et al., *Nature Medicine*, volumen 9 Número 3 páginas 266 - 267 (2003)
 Zeng et al., (2002)

Listado de secuencias

<110> University of Massachusetts ARONIN, Neil ZAMORE, Phillip D.

<120> ARN DE INTERFERENCIA PARA EL TRATAMIENTO DE TRASTORNOS RELACIONADOS CON LA GANANCIA DE FUNCIÓN

<130> UMY-083PC

<150> 60/502678
<151> 2003-09-12

5 <160> 33

<170> FastSEQ para Windows Versión 4.0

10 <210> 1
<211> 13672
<212> ADN
<213> Homo sapiens

15 <400> 1

```

ttgctgtgtg aggcagaacc tgcgggggca ggggcgggct ggttccctgg ccagccattg 60
gcagagtccg caggctaggg ctgtcaatca tgctggcccg cgtggccccg cctccgcccg 120
cgcgcccccg cctccgcccg cgcaogtctg ggacgcaagg cgccgtgggg gctgccggga 180
cgggtccaag atggacggcc gctcaggttc tgcttttacc tgcggcccag agccccattc 240
attgccccgg tgctgagcgg cgccgcgagt cggcccgagg cctccgggga ctgccgtgcc 300
gggcgggaga ccgccatggc gaccctggaa aagctgatga aggccctcga gtccctcaag 360
tccttccagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 420
cagcagcagc aacagcccgc accgcccgcg ccgcccgcgc cgccctcctca gcttccctcag 480
ccgcccgcgc aggcacagcc gctgctgcct cagccgcagc cgcccccgcc gccgcccccg 540
ccgccacccg gcccggtgt ggctgaggag ccgctgcacc gaccaaagaa agaactttca 600
gctaccaaga aagaccgtgt gaatcattgt ctgacaatat gtgaaaacat agtggcacag 660
tctgtcagaa attctccaga atttcagaaa cttctgggca tcgctatgga actttttctg 720
ctgtgcagtg atgacgcaga gtcagatgtc aggatggtgg ctgacgaatg cctcaacaaa 780
gttatcaaag ctttgatgga ttctaattctt ccaaggttac agctcgagct ctataaggaa 840
atataaaaaga atgggtgcccc tcggagtttg cgtgctgccc tgtggagggt tgctgagctg 900
gctcacctgg ttcggcctca gaaatgcagg ccttacctgg tgaaccttct gccgtgectg 960
actcgaacaa gcaagagacc cgaagaatca gtccaggaga ccttggctgc agctgttccc 1020
aaaattatgg cttctttttgg caatttttgca aatgacaatg aaattaaggt tttgttaaag 1080
gccttcatag cgaacctgaa gtcaagctcc cccaccattc ggccggacagc ggctggatca 1140
gcagtgagca tctgccagca ctcaagaagg acacaatatt tctatagttg gctactaaat 1200
gtgctcttag gcttactcgt tcctgtcgag gatgaacact ccactctgct gattcttggc 1260
gtgctgctca ccctgaggta tttggtgccc ttgctgcagc agcagggtcaa ggacacaagc 1320
ctgaaaggca gcttcggagt gacaaggaaa gaaatggaag tctctccttc tgcagagcag 1380
cttgtccagg tttatgaact gacgttacat catacacagc accaagacca caatgttgtg 1440
accggagccc tggagctggt gcagcagctc ttcagaacgc ctccaccgga gcttctgcaa 1500
accctgaccg cagtcggggg cattgggagc ctcaccgctg ctaaggagga gtctggtggc 1560
cgaagccgta gtgggagtat tgtggaactt atagctggag ggggttccctc atgcagccct 1620
gtcctttcaa gaaaacaaaa aggcaaagtg ctcttaggag aagaagaagc cttggaggat 1680
gactctgaat cgagatcgga tgtcagcagc tctgccttaa cagcctcagt gaaggatgag 1740
atcagtggag agctggctgc ttcttcaggg gtttccactc cagggtcagc aggtcatgac 1800
atcatcacag aacagcccag gtcacagcac acactgcagg cggactcagt ggatctggcc 1860
agctgtgact tgacaagctc tgccactgat ggggatgagg aggatatctt gagccacagc 1920
tccagccagg tcagcgcctg cccatctgac cctgccatgg acctgaatga tgggaccagc 1980
gcctcgtcgc ccatcagcga cagctcccag accaccaccg aagggcctga ttcagctggt 2040
acccttccag acagttctga aattgtgtta gacggtaccg acaaccagta tttgggcctg 2100

```

cagattggac agccccagga tgaagatgag gaagccacag gtattcttcc tgatgaagcc 2160
 tgggaggcct tcaggaactc ttccatggcc cttcaacagg cacatttatt gaaaaacatg 2220
 agtcactgca ggcagccttc tgacagcagt gttgataaat ttgtgttgag agatgaagct 2280
 actgaaccgg gtgatcaaga aaacaagcct tgccgcatca aagggtgacat tggacagtcc 2340
 actgatgatg actctgcacc tcttgtccat tgtgtccgcc ttttatctgc ttcgtttttg 2400
 ctaacagggg gaaaaaatgt gctggttccg gacagggatg tgagggtcag cgtgaaggcc 2460
 ctggccctca gctgtgtggg agcagctgtg gccctccacc cggaatcttt cttcagcaaa 2520
 ctctataaag ttccctcttga caccacggaa taccctgagg aacagtatgt ctcagacatc 2580
 ttgaactaca tcgatcatgg agaccacag gttcgaggag cactgccat tctctgtggg 2640
 accctcatct gctccatcct cagcaggctc cgcttccacg tgggagattg gatgggcacc 2700
 attagaacct tcacaggaaa tacattttct ttggcggatt gcattccttt gctgocgaaa 2760
 aactgaagg atgagtcttc tgttacttgc aagttagctt gtacagctgt gaggaactgt 2820
 gtcattgagtc tctgcagcag cagctacagt gattaggac cagctgatg catcgatgtg 2880
 ctgactctga ggaacagttc ctattggctg gtgaggacag agcttctgga aacccttgca 2940
 gagattgact tcaggctggg gagctttttg gaggcaaaaag cagaaaactt acacagaggg 3000
 gctcatcatt atacagggct tttaaaactg caagaacgag tgctcaataa tgttgtcatc 3060
 ctttggcttg gagatgaaga ccccagggtg cgacatgttg ccgcagcatc actaattagg 3120
 cttgtcccaa agctgtttta taaatgtgac caaggacaag ctgatccagt agtggccgtg 3180
 gcaagagatc aaagcagtgt ttacctgaaa cttctcatgc atgagacgca gcctccatct 3240
 ctttctccg tcagcacaat aaccagaata tatagaggct ataacctact accaagcata 3300
 acagacgtca ctatggaaaa taacctttca agagttattg cagcagttt ccatgaacta 3360
 atcacatcaa ccaccagagc actcacattt ggatgctgtg aagctttgtg tcttctttcc 3420
 actgcttcc cagtttgcatt ttggagtta ggttggcact gtggagtgcc tccactgagt 3480
 gcctcagatg agtctaggaa gagctgtacc gttgggtagg ccacaatgat tctgaccctg 3540
 ctctcgtcag cttggttccc attggatctc tcagcccatc aagatgcttt gattttggcc 3600
 ggaaacttgc ttgcagccag tgctcccaa tctctgagaa gttcatgggc ctctgaagaa 3660
 gaagccaacc cagcagccac caagcaagag gaggtctggc cagccctggg ggaccgggcc 3720
 ctggtgccca tgggtggagca gctcttctct cacctgctga aggtgattaa catttgtgcc 3780
 cacgtcctgg atgacgtggc tcctggacc gcaataaagg cagccttgcc ttctctaaca 3840
 aacccccctt ctctaagtcc catccgacga aaggggaagg agaaagaacc aggagaaca 3900
 gcatctgtac cgttgagtc caagaaaggc agtgaggcca gtgcagcttc tagacaatct 3960
 gatacctcag gtccgtttac aacaagtaaa tctctcatcag tggggagttt ctatcatctt 4020
 cttcatacc tcaactgca tgatgtcctg aaagctacac acgctaacta caaggtcacg 4080
 ctggatcttc agaacagcac ggaaaagttt ggagggtttc tccgctcagc cttggatgtt 4140
 cttctcaga tactagagct ggccacactg caggacattg ggaagtgtgt tgaagagatc 4200
 ctaggatacc tgaatcctg ctttagtca gaaccaatga tggcaactgt ttgtgttcaa 4260
 caattggtga agactctctt tggcacaaac ttggcctccc agtttgatgg ctatcttcc 4320
 aaccccagca agtcacaagg ccgagcacag cgccttggct cctccagtgt gaggccaggc 4380
 ttgtaccact actgcttcat ggcctcgtag acccaactca cccaggccct cgctgacgcc 4440
 agcctgagga acatgggtgca ggcggagcag gagaacgaca cctcgggatg gtttgatgtc 4500
 ctccagaaag tgtctaccca gttgaagaca aacctcacga gtgtcaciaa gaaccgtgca 4560
 gataagaatg ctattcataa tcacattcgt ttgtttgaac ctcttgttat aaaagcttta 4620
 aacagtgaca cgactacaac atgtgtgcag ttacagaagc aggtttttaga ttgtctggcg 4680
 cagctggttc agttacgggt taattactgt cttctggatt cagatcaggt gtttattggc 4740
 tttgtattga aacagtttga atacattgaa gtgggcccagt tcaggggaatc agaggcaatc 4800
 attccaaca tctttttctt cttgggatta ctatcttatg aacgctatca ttcaaaacag 4860
 atcattggaa ttccataaat cattcagctc tgtgatggca tcatggccag tggaggaag 4920
 gctgtgacac atgccatacc ggctctgcag cccatagtcc acgacctctt tgtattaaga 4980
 ggaacaaata aagctgatgc aggaaaagag cttgaaacc aaaagagggt ggtgggtgca 5040
 atgttactga gactcatcca gtaccatcag gtgttggaga tgttcattct tgtctgcag 5100
 cagtgcaca aggagaatga agacaagtgg aagcgactgt ctcgacagat agctgacatc 5160
 atcctcccaa tggtagccaa acagcagatg cacattgact ctcatgaagc ccttggagtg 5220
 ttaaatacat tatttgagat tttggccctc tctccctcc gtccgtaga catgctttta 5280
 cggagtatgt tcgtcactcc aacacaaatg gcgtccgtga gcaactgttca actgtggata 5340
 tcgggaattc tggccatttt gagggttctg atttcccagt caactgaaga tattgttctt 5400
 tctcgtatc aggagctctc cttctctccg tatttaaatc cctgtacagt aattaatagg 5460
 ttaagagatg gggacagtac ttcaacgcta gaagaacaca gtgaagggaa acaataaag 5520
 aattgcccag aagaaacatt ttcaaggttt ctattacaac tggttgggtat tcttttagaa 5580
 gacattgtta caaaacagct gaaggtggaa atgagtgagc agcaacatac tttctattgc 5640
 caggaactag gcacactgct aatgtgtctg atccacatct tcaagtctgg aatgttccgg 5700
 agaatcacag cagctgccac taggctgttc cgcagtgatg gctgtggcgg cagtttctac 5760

accctggaca gcttgaactt gggggctcgt tccatgatca ccaccacccc ggccctgggtg 5820
 ctgctctggg gtcagatact gctgcttgtc aaccacacccg actaccgctg gtgggcagaa 5880
 gtgcagcaga ccccgaaaag acacagtctg tccagcacia agttacttag tccccagatg 5940
 tctggagaag aggaggattc tgacttggca gccaaacttg gaatgtgcaa tagagaaata 6000
 gtacgaagag gggctctcat tctcttctgt gattatgtct gtcagaacct ccatgactcc 6060
 gagcacttaa cgtggctcat tghtaaatcac attcaagatc tgatcagcct tccccacgag 6120
 cctccagtac aggacttcat cagtgcctgt catcggaaact ctgctgccag cggcctgttc 6180
 atccaggcaa ttcagtctcg ttgtgaaaac ctttcaactc caaccatgct gaagaaaact 6240
 cttcagtgct tggaggggat ccatctcagc cagtccggag ctgtgctcac gctgtatgtg 6300
 gacaggcttc tgtgcacccc tttccgtgtg ctggctcgca tggctcgacat ccttgcttgt 6360
 cgccgggtag aaatgcttct ggctgcaaat ttacagagca gcatggcca gttgccaatg 6420
 gaagaactca acagaatcca ggaatacctt cacagcagcg ggctcgctca gagacaccaa 6480
 aggctctatt cctgctgga caggtttctg ctctccacca tgcaagactc acttagtccc 6540
 tctcctccag tctcttccca cccgctggac ggggatgggc acgtgtcact ggaaacagtg 6600
 agtccggaca aagactggta cgttcatctt gtcaaatccc agtggtggac caggtcagat 6660
 tctgcaactg tggaggggag agagctggg aatcggattc ctgctgaaga tatgaatgcc 6720
 ttcagatgta actcggagtt caacctaacg ctgctagctc catgcttaag cctagggatg 6780
 agtgaatatt ctgggtggca gaagagtgcc ctttttgaag cagcccgtga ggtgactctg 6840
 gccctgtgta gcggcaccgt gcagcagctc cctgctgtcc atcatgtctt ccagcccag 6900
 ctgectgcag agccggcggc ctactggagc aagttgaatg atctgtttg ggatgctgca 6960
 ctgtatcagt ccctgcccac tctggcccgg gccctggcac agtacctggg ggtggtctcc 7020
 aaactgccc gtcatcttga ccttctctct gagaaagaga aggacattgt gaaattcgtg 7080
 gtggcaacc ttgaggccct gtctctggat ttgatccatg agcagatccc gctgagtctg 7140
 gatctccagg cagggtgga ctgctgtgc ctggcctgc agctgctgg cctctggagc 7200
 gtggctctct ccacagagtt tgtgaoccac gectgctccc tcatctactg tgtgcaactc 7260
 atcctggagg ccgttgcaag gcagcctgga gagcagcttc ttagtccaga aagaaggaca 7320
 aatacccaa aagccatcag cgaggaggag gaggaagtat atccaaacac acagaatcct 7380
 aagtatatca ctgcagcctg tgagatggg gcagaaatgg tggagtctct gcagtcggg 7440
 ttggccttgg gtcataaaag gaatagcggc gtgcccggct ttctcacgcc attgctcagg 7500
 aacatcatca tcagcctggc ccgctgccc cttgtcaaca gctacacacg tbtgccccca 7560
 ctggtgtgga agcttggatg gtcacccaaa ccgggagggg attttggcac agcattccct 7620
 gagatcccc tggagttcct ccaggaaaag gaagtcttta aggagttcat ctaccgcatc 7680
 aacacactag gctggaccag tctgactcag tttgaagaaa cttgggccac cctccttggg 7740
 gtccctggta cgcagccct cgtgatggag caggaggaga gccaccaga agaagacaca 7800
 gagaggacc agatcaacgt cctggccgtg caggccatca cctcactggg gctcagtga 7860
 atgactgtgc ctgtggccgg caaccagct gtaagtgtct tggagcagca gcccgggac 7920
 aagcctctga aagctctga caccagttt gggaggaagc tgagcattat cagagggatt 7980
 gtggagcaag agattcaagc aatggtttca aagagagaga atattgccac ccatattta 8040
 tatcaggcat gggatcctgt ccttctctct tctccggcta ctacaggtg cctcatcagc 8100
 cacgagaagc tgctgtaca gatcaacccc gagcgggagc tggggagcat gagctacaaa 8160
 ctccggcagg tgtccataca ctccgtgtgg ctggggaaca gcatcacacc cctgaggggag 8220
 gaggaatgg acgaggaaga ggaggaggag gccgacgcc ctgcaacctc gtcaccacc 8280
 acgtctccag tcaactccag gaaacaccgg gctggagttg acatccactc ctgttccgag 8340
 tttttgctt agttgtacag ccgctggatc ctgcccgtca gctcagccag gaggacccc 8400
 gccatcctga tcagtgagg ggtcagatcc cttctagtgg tctcagactt gttcaccgag 8460
 cgcaaccagt ttgagctgat gtatgtgacg ctgacagaac tgcgaagggt gcacccttca 8520
 gaagacgaga tccctcgctca gtacctggtg cctgcccact gcaaggcagc tgccgtcctt 8580
 gggatggaca agcccggtgg ggagcctgtc agccgctgc tggagagcag gctcaggagc 8640
 agccacctgc ccagcagggt tggagccctg cacggcgtcc tctatgtgct ggagtccgac 8700
 ctgctggacg acactgccc cagctcatc ccggtcatca gcgactatct cctctccaac 8760
 ctgaaaggg tgcgccactg cgtgaacatt cacagccagc agcacgtact ggtcatgtgt 8820
 gccactgctg tttacctcat tgagaactat cctctggaag tagggccgga attttcagca 8880
 tcaataatac agatgtgtgg ggtgatgctg tctggaagtg aggagtccac cccctccatc 8940
 atttaccact gtgccctcag aggcctggag cgcctctctg tctctgagca gctctcccgc 9000
 ctggatgcag aatcgctgg caagctgagt gtggacagag tgaacgtgca cagcccgcac 9060
 cgggccatgg cggctctgg cctgatgctc acctgcatgt acacaggaaa ggagaaagt 9120
 agtccgggta gaacttcaga ccctaactct gcagcccccg acagcgagtc agtgattgtt 9180
 gctatggagc gggatctctg tttttttgat aggatcagga aaggctttcc ttgtgaagcc 9240
 agagtgggtg ccaggatcct gccccagttt ctgacagact tcttcccacc ccaggacatc 9300
 atgaacaaag tcatcggaga gtttctgtcc aaccagcagc cataccccca gttcatggcc 9360
 accgtgggtg ataagggtgt tcagactctg cacagcaccg ggcagtcgtc catggctccg 9420

gactgggtca tgctgtccct ctccaacttc acgcagaggg ccccggtcgc catggccacg 9480
tggagcctct cctgcttctt tgtcagecgc tccaccagcc cgtgggtcgc ggcgatcctc 9540
ccacatgtca tcagcaggat gggcaagctg gagcagggtg acgtgaacct tttctgcctg 9600
gtcgcacacag acttctacag acaccagata gaggaggagc tcgaccgcag ggccttccag 9660
tctgtgcttg aggtggttg agccccagga agcccatatc accggctgct gacttgttta 9720
cgaaatgtcc acaaggtcac cacctgctga gcgccatggt gggagagact gtgaggcggc 9780
agctggggcc ggagcctttg gaagtctgtg cccttgtgcc ctgcctccac cgagccagct 9840
tggtccttat gggcttccgc acatgccgcg gggggccagg caacgtgctg gtctctgcca 9900
tgtggcagaa gtgctctttg tggcagtgcc caggcagggg gtgtctgcag tccctgggtgg 9960
gctgagcctg aggccttcca gaaagcagga gcagctgtgc tgcaccccat gtgggtgacc 10020
aggtcctttc tcctgatagt cacctgctgg ttgttgccag gttgcagctg ctcttgcatac 10080
tgggcccagaa gtccctccctc ctgcaaggctg gctgttggcc cctctgctgt cctgcagtag 10140
aagtgcccg gagcaggctt tgggaacct ggcctgggtc tccctgggtg ggtgtgcatg 10200
ccacgccccg tgtctggatg cacagatgcc atggcctgtg ctgggcccag ggctgggggt 10260
gctagacacc cggcaccatt ctcccttctc tcttttcttc tcaggattta aaatttaatt 10320
atatcagtaa agagattaat ttaacgaac tctttctatg cccgtgtaa gtatgtgaat 10380
cgcaaggcct gtgctgcatg cgacagcgtc cgggggtggg gacagggccc cgggccacgc 10440
tccctctcct gtagccactg gcatagccct cctgagcacc cgctgacatt tccgttgtac 10500
atgttctctg ttatgcattc acaaggtgac tgggatgtag agaggcgtta gtgggcaggt 10560
ggccacagca ggactgagga caggccccca ttatcctagg ggtgcgctca actgcagccc 10620
ctcctcctcg ggcacagacg actgtcgttc tccaccacc agtcagggac agcagcctcc 10680
ctgtcactca gctgagaagg ccagccctcc ctggctgtga gcagcctcca ctgtgtccag 10740
agacatgggc ctccactcc tgttccttgc tagccctggg gtggcgtctg cctaggagct 10800
ggctggcagg tgttgggacc tgcctgctca tggatgcatg ccctaagagt gtcactgagc 10860
tgtgttttgt ctgagcctct ctcggtcaac agcaaagctt ggtgtccttg cactgtagt 10920
gacagagccc agcatccctt ctgccccgt tccagctgac atcttgcacg gtgaccctt 10980
ttagttagga gagtgcagat ctgtgctcat cggagactgc cccacggccc tgtcagagcc 11040
gccactccta tcccaggac aggtccctgg accagcctcc tgtttgcagg cccagaggag 11100
ccaagtcatt aaaatggaag tggattctgg atggccgggc tgcctgctgat gtaggagctg 11160
gatttgggag ctctgcttgc cgactggctg tgagacgagg caggggctct gcttccctcag 11220
ccctagaggc gagccaggca aggttggcga ctgtcatgtg gcttggttt gtcatgccc 11280
tcgatgtttt gggatattgaa tgtggtaatg ggaggaaatg ttggaactct gtgcaggtgc 11340
tgccttgaga ccccaagct tccacctgtc cctctcctat gtggcagctg gggagcagct 11400
gagatgtgga cttgtatgct gccacatac gtgaggggga gctgaaagg agcccctgct 11460
caaaggagc cctcctctg agcagcctct gccaggcctg tatgaggctt tcccaccag 11520
ctcccaacag aggcctcccc cagccaggac cacctgctcc tgcgtggcgg gacagaggag 11580
cggtagaaaag gggccgatg ttgaggagg cccttaagg aagctaactg attataacac 11640
gtaagaaaat caccattctt ccgtattggt tggggctcc tgtttctcat cctagctttt 11700
tcttgaaaaa gcccgtaga aggtttggga acgaggggaa agttctcaga actgttgctg 11760
ctccccacc gcctcccgcc tccccgcag gttatgtcag cagctctgag acagcagtat 11820
cacaggccag atgttgttcc tggctagatg tttacatttg taagaaataa cactgtgaat 11880
gtaaaacaga gccattccct tggaatgcat atcgctgggc tcaacataga gtttgtctc 11940
ctcttgttta cgacgtgatc taaaccagtc cttagcaagg ggctcagaac accccgctct 12000
ggcagtaggt gtccccacc cccaaagacc tgcctgtgtg ctccggagat gaatatgagc 12060
tcattagtaa aatgacttc acccagcat atacataaag tatccatgca tgtgcatata 12120
gacacatcta taattttaca cacacacctc tcaagacgga gatgcatggc ctctaagagt 12180
gcccgtgtcg gttcttctg gaagtgact ttccttagac ccgccaggtc aagttagccg 12240
cgtgacggac atccaggcgt gggacgtggg cagggcaggg ctattcatt gccactagg 12300
atccactgg ogaagatgg ctccatatca gctctctgca gaagggagga agactttatc 12360
atgttcctaa aaatctgtg caagcaccca tcgtattatc caaattttgt tgcaaatgtg 12420
attaatttgg ttgtcaagtt ttgggggtgg gctgtgggga gattgctttt gtttctctgc 12480
tggtaatatc gggaaagatt ttaatgaaac cagggtagaa ttgtttggca atgactgaa 12540
gcgtgtttct tccccaaaat gtgcctccct tccgctgcgg gccagctga gtctatgtag 12600
gtgatgtttc cagctgcaa gtgctctttg ttaactgtcca ccctcattc tgccagcgca 12660
tgtgtccttt caaggggaaa atgtgaagct gaacccccctc cagacacca gaatgtagca 12720
tctgagaagg ccctgtgccc taaaggacac ccctcgcccc catcttcatg gaggggtca 12780
tttcagagcc ctccggagcca atgaacagct cctcctcttg gagctgagat gagccccacg 12840
tggagctcgg gacggatagt agacagcaat aactcgggtg gtggccgcct ggcaggtgga 12900
acttccctcc gttgcgggggt ggagtgaggt tagttctgtg tgtctgggtg gtggagtcag 12960
gcttctcttg ctacctgtga gcatccttcc cagcagacat cctcatcggg ctttgcctc 13020
ccccgcttc ctccctctgc ggggaggacc cgggaccaca gctgtggcc agggtagact 13080

ES 2 485 848 T3

```
tggagctgtc ctccagaggg gtcacgtgta ggagtgagaa gaaggaagat cttgagagct 13140
gctgagggac cttggagagc tcaggatggc tcagacgagg aactcgcctt gccgggcctg 13200
gccctcctgg gaaggagggg gctgctcaga atgccgcatg acaactgaag gcaacctgga 13260
aggttcaggg cccgctcttc ccccatgtgc ctgtcacgct ctgggtgcagt caaaggaacg 13320
ccttcccctc agttgtttct aagagcagag tctcccgcctg caatctgggt ggtaactgcc 13380
agccttggag gatcgtggcc aacgtggacc tgcctacgga ggggtgggctc tgaccaagt 13440
ggggcctcct tgcccaggtc tcaactgctt gcaccgtggc cagagggact gtcagctgag 13500
cttgagctcc cctggagcca gcagggctgt gatgggcgag tcccggagcc ccaccagac 13560
ctgaatgctt ctgagagcaa agggaaggac tgacgagaga tgtatattta atttttaac 13620
tgctgcaaac attgtacatc caaattaaag ggaaaaaatg gaaacatca at 13672
```

<210> 2

<211> 3144

5 <212> PRT

<213> Homo sapiens

<400> 2

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

Lys Gly Ser Phe Gly Val Thr Arg Lys Glu Met Glu Val Ser Pro Ser
 340 345 350
 Ala Glu Gln Leu Val Gln Val Tyr Glu Leu Thr Leu His His Thr Gln
 355 360 365
 His Gln Asp His Asn Val Val Thr Gly Ala Leu Glu Leu Leu Gln Gln
 370 375 380
 Leu Phe Arg Thr Pro Pro Pro Glu Leu Leu Gln Thr Leu Thr Ala Val
 385 390 395 400
 Gly Gly Ile Gly Gln Leu Thr Ala Ala Lys Glu Glu Ser Gly Gly Arg
 405 410 415
 Ser Arg Ser Gly Ser Ile Val Glu Leu Ile Ala Gly Gly Gly Ser Ser
 420 425 430
 Cys Ser Pro Val Leu Ser Arg Lys Gln Lys Gly Lys Val Leu Leu Gly
 435 440 445
 Glu Glu Glu Ala Leu Glu Asp Asp Ser Glu Ser Arg Ser Asp Val Ser
 450 455 460
 Ser Ser Ala Leu Thr Ala Ser Val Lys Asp Glu Ile Ser Gly Glu Leu
 465 470 475 480
 Ala Ala Ser Ser Gly Val Ser Thr Pro Gly Ser Ala Gly His Asp Ile
 485 490 495
 Ile Thr Glu Gln Pro Arg Ser Gln His Thr Leu Gln Ala Asp Ser Val
 500 505 510
 Asp Leu Ala Ser Cys Asp Leu Thr Ser Ser Ala Thr Asp Gly Asp Glu
 515 520 525
 Glu Asp Ile Leu Ser His Ser Ser Ser Gln Val Ser Ala Val Pro Ser
 530 535 540
 Asp Pro Ala Met Asp Leu Asn Asp Gly Thr Gln Ala Ser Ser Pro Ile
 545 550 555 560
 Ser Asp Ser Ser Gln Thr Thr Thr Glu Gly Pro Asp Ser Ala Val Thr
 565 570 575
 Pro Ser Asp Ser Ser Glu Ile Val Leu Asp Gly Thr Asp Asn Gln Tyr
 580 585 590
 Leu Gly Leu Gln Ile Gly Gln Pro Gln Asp Glu Asp Glu Glu Ala Thr
 595 600 605
 Gly Ile Leu Pro Asp Glu Ala Ser Glu Ala Phe Arg Asn Ser Ser Met
 610 615 620
 Ala Leu Gln Gln Ala His Leu Leu Lys Asn Met Ser His Cys Arg Gln
 625 630 635 640
 Pro Ser Asp Ser Ser Val Asp Lys Phe Val Leu Arg Asp Glu Ala Thr
 645 650 655
 Glu Pro Gly Asp Gln Glu Asn Lys Pro Cys Arg Ile Lys Gly Asp Ile
 660 665 670
 Gly Gln Ser Thr Asp Asp Asp Ser Ala Pro Leu Val His Cys Val Arg
 675 680 685
 Leu Leu Ser Ala Ser Phe Leu Leu Thr Gly Gly Lys Asn Val Leu Val
 690 695 700
 Pro Asp Arg Asp Val Arg Val Ser Val Lys Ala Leu Ala Leu Ser Cys
 705 710 715 720
 Val Gly Ala Ala Val Ala Leu His Pro Glu Ser Phe Phe Ser Lys Leu
 725 730 735
 Tyr Lys Val Pro Leu Asp Thr Thr Glu Tyr Pro Glu Glu Gln Tyr Val
 740 745 750
 Ser Asp Ile Leu Asn Tyr Ile Asp His Gly Asp Pro Gln Val Arg Gly
 755 760 765
 Ala Thr Ala Ile Leu Cys Gly Thr Leu Ile Cys Ser Ile Leu Ser Arg
 770 775 780
 Ser Arg Phe His Val Gly Asp Trp Met Gly Thr Ile Arg Thr Leu Thr
 785 790 795 800
 Gly Asn Thr Phe Ser Leu Ala Asp Cys Ile Pro Leu Leu Arg Lys Thr
 805 810 815
 Leu Lys Asp Glu Ser Ser Val Thr Cys Lys Leu Ala Cys Thr Ala Val

			820						825				830			
Arg	Asn	Cys	Val	Met	Ser	Leu	Cys	Ser	Ser	Ser	Tyr	Ser	Glu	Leu	Gly	
		835					840					845				
Leu	Gln	Leu	Ile	Ile	Asp	Val	Leu	Thr	Leu	Arg	Asn	Ser	Ser	Tyr	Trp	
	850					855					860					
Leu	Val	Arg	Thr	Glu	Leu	Leu	Glu	Thr	Leu	Ala	Glu	Ile	Asp	Phe	Arg	
865					870					875					880	
Leu	Val	Ser	Phe	Leu	Glu	Ala	Lys	Ala	Glu	Asn	Leu	His	Arg	Gly	Ala	
			885						890					895		
His	His	Tyr	Thr	Gly	Leu	Leu	Lys	Leu	Gln	Glu	Arg	Val	Leu	Asn	Asn	
		900						905				910				
Val	Val	Ile	His	Leu	Leu	Gly	Asp	Glu	Asp	Pro	Arg	Val	Arg	His	Val	
		915					920					925				
Ala	Ala	Ala	Ser	Leu	Ile	Arg	Leu	Val	Pro	Lys	Leu	Phe	Tyr	Lys	Cys	
	930					935					940					
Asp	Gln	Gly	Gln	Ala	Asp	Pro	Val	Val	Ala	Val	Ala	Arg	Asp	Gln	Ser	
945					950					955					960	
Ser	Val	Tyr	Leu	Lys	Leu	Leu	Met	His	Glu	Thr	Gln	Pro	Pro	Ser	His	
			965						970					975		
Phe	Ser	Val	Ser	Thr	Ile	Thr	Arg	Ile	Tyr	Arg	Gly	Tyr	Asn	Leu	Leu	
		980						985					990			
Pro	Ser	Ile	Thr	Asp	Val	Thr	Met	Glu	Asn	Asn	Leu	Ser	Arg	Val	Ile	
		995					1000					1005				
Ala	Ala	Val	Ser	His	Glu	Leu	Ile	Thr	Ser	Thr	Thr	Arg	Ala	Leu	Thr	
	1010					1015					1020					
Phe	Gly	Cys	Cys	Glu	Ala	Leu	Cys	Leu	Leu	Ser	Thr	Ala	Phe	Pro	Val	
1025					1030					1035					1040	
Cys	Ile	Trp	Ser	Leu	Gly	Trp	His	Cys	Gly	Val	Pro	Pro	Leu	Ser	Ala	
			1045						1050					1055		
Ser	Asp	Glu	Ser	Arg	Lys	Ser	Cys	Thr	Val	Gly	Met	Ala	Thr	Met	Ile	
		1060						1065					1070			
Leu	Thr	Leu	Leu	Ser	Ser	Ala	Trp	Phe	Pro	Leu	Asp	Leu	Ser	Ala	His	
	1075					1080						1085				
Gln	Asp	Ala	Leu	Ile	Leu	Ala	Gly	Asn	Leu	Leu	Ala	Ala	Ser	Ala	Pro	
	1090					1095					1100					
Lys	Ser	Leu	Arg	Ser	Ser	Trp	Ala	Ser	Glu	Glu	Glu	Ala	Asn	Pro	Ala	
1105					1110					1115					1120	
Ala	Thr	Lys	Gln	Glu	Glu	Val	Trp	Pro	Ala	Leu	Gly	Asp	Arg	Ala	Leu	
			1125						1130					1135		
Val	Pro	Met	Val	Glu	Gln	Leu	Phe	Ser	His	Leu	Leu	Lys	Val	Ile	Asn	
		1140						1145					1150			
Ile	Cys	Ala	His	Val	Leu	Asp	Asp	Val	Ala	Pro	Gly	Pro	Ala	Ile	Lys	
	1155					1160						1165				
Ala	Ala	Leu	Pro	Ser	Leu	Thr	Asn	Pro	Pro	Ser	Leu	Ser	Pro	Ile	Arg	
	1170					1175					1180					
Arg	Lys	Gly	Lys	Glu	Lys	Glu	Pro	Gly	Glu	Gln	Ala	Ser	Val	Pro	Leu	
1185					1190					1195					1200	
Ser	Pro	Lys	Lys	Gly	Ser	Glu	Ala	Ser	Ala	Ala	Ser	Arg	Gln	Ser	Asp	
			1205						1210					1215		
Thr	Ser	Gly	Pro	Val	Thr	Thr	Ser	Lys	Ser	Ser	Ser	Leu	Gly	Ser	Phe	
			1220					1225					1230			
Tyr	His	Leu	Pro	Ser	Tyr	Leu	Lys	Leu	His	Asp	Val	Leu	Lys	Ala	Thr	
	1235					1240						1245				
His	Ala	Asn	Tyr	Lys	Val	Thr	Leu	Asp	Leu	Gln	Asn	Ser	Thr	Glu	Lys	
	1250					1255					1260					
Phe	Gly	Gly	Phe	Leu	Arg	Ser	Ala	Leu	Asp	Val	Leu	Ser	Gln	Ile	Leu	
1265					1270					1275					1280	
Glu	Leu	Ala	Thr	Leu	Gln	Asp	Ile	Gly	Lys	Cys	Val	Glu	Glu	Ile	Leu	
			1285						1290					1295		
Gly	Tyr	Leu	Lys	Ser	Cys	Phe	Ser	Arg	Glu	Pro	Met	Met	Ala	Thr	Val	
		1300						1305					1310			

Cys Val Gln Gln Leu Leu Lys Thr Leu Phe Gly Thr Asn Leu Ala Ser
 1315 1320 1325
 Gln Phe Asp Gly Leu Ser Ser Asn Pro Ser Lys Ser Gln Gly Arg Ala
 1330 1335 1340
 Gln Arg Leu Gly Ser Ser Ser Val Arg Pro Gly Leu Tyr His Tyr Cys
 1345 1350 1355 1360
 Phe Met Ala Pro Tyr Thr His Phe Thr Gln Ala Leu Ala Asp Ala Ser
 1365 1370 1375
 Leu Arg Asn Met Val Gln Ala Glu Gln Glu Asn Asp Thr Ser Gly Trp
 1380 1385 1390
 Phe Asp Val Leu Gln Lys Val Ser Thr Gln Leu Lys Thr Asn Leu Thr
 1395 1400 1405
 Ser Val Thr Lys Asn Arg Ala Asp Lys Asn Ala Ile His Asn His Ile
 1410 1415 1420
 Arg Leu Phe Glu Pro Leu Val Ile Lys Ala Leu Lys Gln Tyr Thr Thr
 1425 1430 1435 1440
 Thr Thr Cys Val Gln Leu Gln Lys Gln Val Leu Asp Leu Leu Ala Gln
 1445 1450 1455
 Leu Val Gln Leu Arg Val Asn Tyr Cys Leu Leu Asp Ser Asp Gln Val
 1460 1465 1470
 Phe Ile Gly Phe Val Leu Lys Gln Phe Glu Tyr Ile Glu Val Gly Gln
 1475 1480 1485
 Phe Arg Glu Ser Glu Ala Ile Ile Pro Asn Ile Phe Phe Phe Leu Val
 1490 1495 1500
 Leu Leu Ser Tyr Glu Arg Tyr His Ser Lys Gln Ile Ile Gly Ile Pro
 1505 1510 1515 1520
 Lys Ile Ile Gln Leu Cys Asp Gly Ile Met Ala Ser Gly Arg Lys Ala
 1525 1530 1535
 Val Thr His Ala Ile Pro Ala Leu Gln Pro Ile Val His Asp Leu Phe
 1540 1545 1550
 Val Leu Arg Gly Thr Asn Lys Ala Asp Ala Gly Lys Glu Leu Glu Thr
 1555 1560 1565
 Gln Lys Glu Val Val Val Ser Met Leu Leu Arg Leu Ile Gln Tyr His
 1570 1575 1580
 Gln Val Leu Glu Met Phe Ile Leu Val Leu Gln Gln Cys His Lys Glu
 1585 1590 1595 1600
 Asn Glu Asp Lys Trp Lys Arg Leu Ser Arg Gln Ile Ala Asp Ile Ile
 1605 1610 1615
 Leu Pro Met Leu Ala Lys Gln Gln Met His Ile Asp Ser His Glu Ala
 1620 1625 1630
 Leu Gly Val Leu Asn Thr Leu Phe Glu Ile Leu Ala Pro Ser Ser Leu
 1635 1640 1645
 Arg Pro Val Asp Met Leu Leu Arg Ser Met Phe Val Thr Pro Asn Thr
 1650 1655 1660
 Met Ala Ser Val Ser Thr Val Gln Leu Trp Ile Ser Gly Ile Leu Ala
 1665 1670 1675 1680
 Ile Leu Arg Val Leu Ile Ser Gln Ser Thr Glu Asp Ile Val Leu Ser
 1685 1690 1695
 Arg Ile Gln Glu Leu Ser Phe Ser Pro Tyr Leu Ile Ser Cys Thr Val
 1700 1705 1710
 Ile Asn Arg Leu Arg Asp Gly Asp Ser Thr Ser Thr Leu Glu Glu His
 1715 1720 1725
 Ser Glu Gly Lys Gln Ile Lys Asn Leu Pro Glu Glu Thr Phe Ser Arg
 1730 1735 1740
 Phe Leu Leu Gln Leu Val Gly Ile Leu Leu Glu Asp Ile Val Thr Lys
 1745 1750 1755 1760
 Gln Leu Lys Val Glu Met Ser Glu Gln Gln His Thr Phe Tyr Cys Gln
 1765 1770 1775
 Glu Leu Gly Thr Leu Leu Met Cys Leu Ile His Ile Phe Lys Ser Gly
 1780 1785 1790
 Met Phe Arg Arg Ile Thr Ala Ala Ala Thr Arg Leu Phe Arg Ser Asp

	1795						1800					1805					
Gly	Cys	Gly	Gly	Ser	Phe	Tyr	Thr	Leu	Asp	Ser	Leu	Asn	Leu	Arg	Ala		
	1810						1815					1820					
Arg	Ser	Met	Ile	Thr	Thr	His	Pro	Ala	Leu	Val	Leu	Leu	Trp	Cys	Gln		
	1825						1830					1835			1840		
Ile	Leu	Leu	Leu	Val	Asn	His	Thr	Asp	Tyr	Arg	Trp	Trp	Ala	Glu	Val		
							1845					1850			1855		
Gln	Gln	Thr	Pro	Lys	Arg	His	Ser	Leu	Ser	Ser	Thr	Lys	Leu	Leu	Ser		
							1860					1865			1870		
Pro	Gln	Met	Ser	Gly	Glu	Glu	Glu	Asp	Ser	Asp	Leu	Ala	Ala	Lys	Leu		
							1875					1880			1885		
Gly	Met	Cys	Asn	Arg	Glu	Ile	Val	Arg	Arg	Gly	Ala	Leu	Ile	Leu	Phe		
	1890						1895					1900					
Cys	Asp	Tyr	Val	Cys	Gln	Asn	Leu	His	Asp	Ser	Glu	His	Leu	Thr	Trp		
	1905						1910					1915			1920		
Leu	Ile	Val	Asn	His	Ile	Gln	Asp	Leu	Ile	Ser	Leu	Ser	His	Glu	Pro		
							1925					1930			1935		
Pro	Val	Gln	Asp	Phe	Ile	Ser	Ala	Val	His	Arg	Asn	Ser	Ala	Ala	Ser		
							1940					1945			1950		
Gly	Leu	Phe	Ile	Gln	Ala	Ile	Gln	Ser	Arg	Cys	Glu	Asn	Leu	Ser	Thr		
							1955					1960			1965		
Pro	Thr	Met	Leu	Lys	Lys	Thr	Leu	Gln	Cys	Leu	Glu	Gly	Ile	His	Leu		
							1970					1975			1980		
Ser	Gln	Ser	Gly	Ala	Val	Leu	Thr	Leu	Tyr	Val	Asp	Arg	Leu	Leu	Cys		
	1985						1990					1995			2000		
Thr	Pro	Phe	Arg	Val	Leu	Ala	Arg	Met	Val	Asp	Ile	Leu	Ala	Cys	Arg		
							2005					2010			2015		
Arg	Val	Glu	Met	Leu	Leu	Ala	Ala	Asn	Leu	Gln	Ser	Ser	Met	Ala	Gln		
							2020								2030		
Leu	Pro	Met	Glu	Glu	Leu	Asn	Arg	Ile	Gln	Glu	Tyr	Leu	Gln	Ser	Ser		
							2035								2040		
Gly	Leu	Ala	Gln	Arg	His	Gln	Arg	Leu	Tyr	Ser	Leu	Leu	Asp	Arg	Phe		
							2050								2055		
Arg	Leu	Ser	Thr	Met	Gln	Asp	Ser	Leu	Ser	Pro	Ser	Pro	Pro	Val	Ser		
							2065					2070			2075		
Ser	His	Pro	Leu	Asp	Gly	Asp	Gly	His	Val	Ser	Leu	Glu	Thr	Val	Ser		
							2085					2090			2095		
Pro	Asp	Lys	Asp	Trp	Tyr	Val	His	Leu	Val	Lys	Ser	Gln	Cys	Trp	Thr		
							2100					2105			2110		
Arg	Ser	Asp	Ser	Ala	Leu	Leu	Glu	Gly	Ala	Glu	Leu	Val	Asn	Arg	Ile		
							2115								2120		
Pro	Ala	Glu	Asp	Met	Asn	Ala	Phe	Met	Met	Asn	Ser	Glu	Phe	Asn	Leu		
							2130								2135		
Ser	Leu	Leu	Ala	Pro	Cys	Leu	Ser	Leu	Gly	Met	Ser	Glu	Ile	Ser	Gly		
							2145					2150			2155		
Gly	Gln	Lys	Ser	Ala	Leu	Phe	Glu	Ala	Ala	Arg	Glu	Val	Thr	Leu	Ala		
							2165					2170			2175		
Arg	Val	Ser	Gly	Thr	Val	Gln	Gln	Leu	Pro	Ala	Val	His	His	Val	Phe		
							2180					2185			2190		
Gln	Pro	Glu	Leu	Pro	Ala	Glu	Pro	Ala	Ala	Tyr	Trp	Ser	Lys	Leu	Asn		
							2195								2200		
Asp	Leu	Phe	Gly	Asp	Ala	Ala	Leu	Tyr	Gln	Ser	Leu	Pro	Thr	Leu	Ala		
							2210								2215		
Arg	Ala	Leu	Ala	Gln	Tyr	Leu	Val	Val	Val	Ser	Lys	Leu	Pro	Ser	His		
							2225					2230			2235		
Leu	His	Leu	Pro	Pro	Glu	Lys	Glu	Lys	Asp	Ile	Val	Lys	Phe	Val	Val		
							2245					2250			2255		
Ala	Thr	Leu	Glu	Ala	Leu	Ser	Trp	His	Leu	Ile	His	Glu	Gln	Ile	Pro		
							2260					2265			2270		
Leu	Ser	Leu	Asp	Leu	Gln	Ala	Gly	Leu	Asp	Cys	Cys	Cys	Leu	Ala	Leu		
							2275								2280		

Gln Leu Pro Gly Leu Trp Ser Val Val Ser Ser Thr Glu phe Val Thr
 2290 2295 2300
 His Ala Cys Ser Leu Ile Tyr Cys Val His Phe Ile Leu Glu Ala Val
 2305 2310 2315 2320
 Ala Val Gln Pro Gly Glu Gln Leu Leu Ser Pro Glu Arg Arg Thr Asn
 2325 2330 2335
 Thr Pro Lys Ala Ile Ser Glu Glu Glu Glu Glu Val Asp Pro Asn Thr
 2340 2345 2350
 Gln Asn Pro Lys Tyr Ile Thr Ala Ala Cys Glu Met Val Ala Glu Met
 2355 2360 2365
 Val Glu Ser Leu Gln Ser Val Leu Ala Leu Gly His Lys Arg Asn Ser
 2370 2375 2380
 Gly Val Pro Ala Phe Leu Thr Pro Leu Leu Arg Asn Ile Ile Ile Ser
 2385 2390 2395 2400
 Leu Ala Arg Leu Pro Leu Val Asn Ser Tyr Thr Arg Val Pro Pro Leu
 2405 2410 2415
 Val Trp Lys Leu Gly Trp Ser Pro Lys Pro Gly Gly Asp Phe Gly Thr
 2420 2425 2430
 Ala Phe Pro Glu Ile Pro Val Glu Phe Leu Gln Glu Lys Glu Val Phe
 2435 2440 2445
 Lys Glu Phe Ile Tyr Arg Ile Asn Thr Leu Gly Trp Thr Ser Arg Thr
 2450 2455 2460
 Gln Phe Glu Glu Thr Trp Ala Thr Leu Leu Gly Val Leu Val Thr Gln
 2465 2470 2475 2480
 Pro Leu Val Met Glu Gln Glu Glu Ser Pro Pro Glu Glu Asp Thr Glu
 2485 2490 2495
 Arg Thr Gln Ile Asn Val Leu Ala Val Gln Ala Ile Thr Ser Leu Val
 2500 2505 2510
 Leu Ser Ala Met Thr Val Pro Val Ala Gly Asn Pro Ala Val Ser Cys
 2515 2520 2525
 Leu Glu Gln Gln Pro Arg Asn Lys Pro Leu Lys Ala Leu Asp Thr Arg
 2530 2535 2540
 Phe Gly Arg Lys Leu Ser Ile Ile Arg Gly Ile Val Glu Gln Glu Ile
 2545 2550 2555 2560
 Gln Ala Met Val Ser Lys Arg Glu Asn Ile Ala Thr His His Leu Tyr
 2565 2570 2575
 Gln Ala Trp Asp Pro Val Pro Ser Leu Ser Pro Ala Thr Thr Gly Ala
 2580 2585 2590
 Leu Ile Ser His Glu Lys Leu Leu Leu Gln Ile Asn Pro Glu Arg Glu
 2595 2600 2605
 Leu Gly Ser Met Ser Tyr Lys Leu Gly Gln Val Ser Ile His Ser Val
 2610 2615 2620
 Trp Leu Gly Asn Ser Ile Thr Pro Leu Arg Glu Glu Glu Trp Asp Glu
 2625 2630 2635 2640
 Glu Glu Glu Glu Glu Ala Asp Ala Pro Ala Pro Ser Ser Pro Pro Thr
 2645 2650 2655
 Ser Pro Val Asn Ser Arg Lys His Arg Ala Gly Val Asp Ile His Ser
 2660 2665 2670
 Cys Ser Gln Phe Leu Leu Glu Leu Tyr Ser Arg Trp Ile Leu Pro Ser
 2675 2680 2685
 Ser Ser Ala Arg Arg Thr Pro Ala Ile Leu Ile Ser Glu Val Val Arg
 2690 2695 2700
 Ser Leu Leu Val Val Ser Asp Leu Phe Thr Glu Arg Asn Gln Phe Glu
 2705 2710 2715 2720
 Leu Met Tyr Val Thr Leu Thr Glu Leu Arg Arg Val His Pro Ser Glu
 2725 2730 2735
 Asp Glu Ile Leu Ala Gln Tyr Leu Val Pro Ala Thr Cys Lys Ala Ala
 2740 2745 2750
 Ala Val Leu Gly Met Asp Lys Ala Val Ala Glu Pro Val Ser Arg Leu
 2755 2760 2765
 Leu Glu Ser Thr Leu Arg Ser Ser His Leu Pro Ser Arg Val Gly Ala

2770		2775		2780
Leu His Gly Val	Leu Tyr Val	Leu Glu Cys Asp	Leu Leu Asp Asp	Thr
2785	2790	2795		2800
Ala Lys Gln Leu	Ile Pro Val	Ile Ser Asp Tyr	Leu Leu Ser Asn	Leu
	2805	2810		2815
Lys Gly Ile Ala	His Cys Val	Asn Ile His Ser	Gln Gln His Val	Leu
	2820	2825		2830
Val Met Cys Ala	Thr Ala Phe Tyr	Leu Ile Glu Asn Tyr	Pro Leu Asp	
	2835	2840		2845
Val Gly Pro Glu	Phe Ser Ala Ser	Ile Ile Gln Met Cys	Gly Val Met	
	2850	2855		2860
Leu Ser Gly Ser	Glu Glu Ser Thr	Pro Ser Ile Ile Tyr	His Cys Ala	
2865	2870	2875		2880
Leu Arg Gly Leu	Glu Arg Leu Leu	Leu Ser Glu Gln	Leu Ser Arg Leu	
	2885	2890		2895
Asp Ala Glu Ser	Leu Val Lys Leu	Ser Val Asp Arg	Val Asn Val His	
	2900	2905		2910
Ser Pro His Arg	Ala Met Ala Ala	Leu Gly Leu Met	Leu Thr Cys Met	
	2915	2920		2925
Tyr Thr Gly Lys	Glu Lys Val Ser	Pro Gly Arg Thr	Ser Asp Pro Asn	
	2930	2935		2940
Pro Ala Ala Pro	Asp Ser Glu Ser	Val Ile Val Ala	Met Glu Arg Val	
2945	2950	2955		2960
Ser Val Leu Phe	Asp Arg Ile Arg	Lys Gly Phe Pro	Cys Glu Ala Arg	
	2965	2970		2975
Val Val Ala Arg	Ile Leu Pro Gln	Phe Leu Asp Asp	Phe Phe Pro Pro	
	2980	2985		2990
Gln Asp Ile Met	Asn Lys Val Ile	Gly Glu Phe Leu	Ser Asn Gln Gln	
	2995	3000		3005
Pro Tyr Pro Gln	Phe Met Ala Thr	Val Val Tyr Lys	Val Phe Gln Thr	
	3010	3015		3020
Leu His Ser Thr	Gly Gln Ser Ser	Met Val Arg Asp	Trp Val Met Leu	
3025	3030	3035		3040
Ser Leu Ser Asn	Phe Thr Gln Arg	Ala Pro Val Ala	Met Ala Thr Trp	
	3045	3050		3055
Ser Leu Ser Cys	Phe Phe Val Ser	Ala Ser Thr Ser	Pro Trp Val Ala	
	3060	3065		3070
Ala Ile Leu Pro	His Val Ile Ser	Arg Met Gly Lys	Leu Glu Gln Val	
	3075	3080		3085
Asp Val Asn Leu	Phe Cys Leu Val	Ala Thr Asp Phe Tyr	Arg His Gln	
	3090	3095		3100
Ile Glu Glu Glu	Leu Asp Arg Arg	Ala Phe Gln Ser	Val Leu Glu Val	
3105	3110	3115		3120
Val Ala Ala Pro	Gly Ser Pro Tyr	His Arg Leu Leu	Thr Cys Leu Arg	
	3125	3130		3135
Asn Val His Lys	Val Thr Thr Cys			
	3140			

<210> 3
 <211> 40
 5 <212> ADN
 <213> Secuencia artificial

<220>
 <223> constructo sintético

10 <400> 3
 ugcagcugau caucgaugug cugaccuga ggaacaguuc

40

<210> 4
 <211> 40
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> constructo sintético
 <400> 4
 10 gaacuguucc ucagggucag cacaucgaug aucagcugca 40
 <210> 5
 <211> 21
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> constructo sintético
 20 <400> 5
 tgtgctgact ctgaggaaca g 21
 <210> 6
 <211> 21
 25 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> constructo sintético
 30 <400> 6
 ugugcugacu cugaggaaca g 21
 <210> 7
 <211> 21
 35 <212> ADN
 <213> Secuencia artificial
 <220>
 40 <223> constructo sintético
 <400> 7
 cuguuccuca gagucagcac a 21
 45 <210> 8
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 50 <220>
 <223> constructo sintético
 <400> 8
 55 catacctcaa actgcatgat g 21
 <210> 9
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 60 <220>
 <223> constructo sintético
 <400> 9
 65 cauaccucaa acugcaugau g 21

<210> 10
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> constructo sintético
 <400> 10
 10 caucaugcag uuugaggau g 21
 <210> 11
 <211> 21
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> constructo sintético
 20 <400> 11
 gcctgcagag ccggcggcct a 21
 <210> 12
 <211> 21
 25 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> constructo sintético
 30 <400> 12
 gccugcagag ccggcggccu a 21
 <210> 13
 <211> 21
 35 <212> ADN
 <213> Secuencia artificial
 <220>
 40 <223> constructo sintético
 <400> 13
 uaggccgccg gcucgcagg c 21
 45 <210> 14
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 50 <220>
 <223> constructo sintético
 <400> 14
 55 acagagtttg tgaccacgc c 21
 <210> 15
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 60 <220>
 <223> constructo sintético
 <400> 15
 65 acagaguug ugaccacgc c 21

<210> 16
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> constructo sintético
 <400> 16
 10 ggcguggguc acaaacucug u 21
 <210> 17
 <211> 21
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> constructo sintético
 20 <400> 17
 tcctcatct actgtgtgca c 21
 <210> 18
 <211> 21
 25 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> constructo sintético
 30 <400> 18
 uccucaucu acugugugca c 21
 <210> 19
 <211> 21
 35 <212> ADN
 <213> Secuencia artificial
 <220>
 40 <223> constructo sintético
 <400> 19
 gugcacacag uaugagagg a 21
 45 <210> 20
 <211> 5
 <212> ADN
 <213> Secuencia artificial
 50 <220>
 <223> constructo sintético
 <400> 20
 55 ugugc 5
 <210> 21
 <211> 7
 <212> ADN
 <213> Secuencia artificial
 60 <220>
 <223> constructo sintético
 <400> 21
 65 gcacauc 7

<210> 22
 <211> 5
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> constructo sintético

 <400> 22
 10 guugc 5

 <210> 23
 <211> 7
 <212> ADN
 15 <213> Secuencia artificial

 <220>
 <223> constructo sintético

 <400> 23
 20 ggaagag 7

 <210> 24
 <211> 21
 <212> ADN
 25 <213> Secuencia artificial

 <220>
 <223> constructo sintético
 30
 <400> 24
 ugugcugacc cugaggaaca g 21

 <210> 25
 <211> 21
 <212> ADN
 35 <213> Secuencia artificial

 <220>
 <223> constructo sintético
 40
 <400> 25
 guuccucagg gucagcaau c 21

 <210> 26
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 45
 <220>
 <223> constructo sintético
 50
 <400> 26
 55 ugugcugacc cugaggaaaa g 21

 <210> 27
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 60
 <220>
 <223> constructo sintético

 <400> 27
 65 uuuccucagg gucagcaau c 21

<210> 28
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> constructo sintético
 <400> 28
 10 ugugcugacc cugaggaaaa g 21
 <210> 29
 <211> 21
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> constructo sintético
 20 <400> 29
 guuccucagg gucagacau c 21
 <210> 30
 <211> 42
 25 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> constructo sintético
 30 <400> 30
 gcgtaatacg actcactata ggaacagtat gtctcagaca tc 42
 <210> 31
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 35 <220>
 <223> constructo sintético
 40 <400> 31
 uucgaaguau uccgcguacg u 21
 45 <210> 32
 <211> 42
 <212> ADN
 <213> Secuencia artificial
 50 <220>
 <223> constructo sintético
 <400> 32
 55 gcgtaatacg actcactata ggacaagcct aattagtgat gc 42
 <210> 33
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 60 <220>
 <223> constructo sintético
 <400> 33
 65 gaacagtatg tctcagacat c 21

REIVINDICACIONES

- 5 1. Una cantidad efectiva de un agente de ARNi seleccionado de agentes de ARNpi y ARNph para uso en el tratamiento de la enfermedad de Huntington causada por una mutación en el gen htt que codifica una proteína Huntingtina mutante de ganancia de función, dirigiendo un solo polimorfismo de nucleótido que no causa enfermedad localizado en un sitio distinto de una mutación que causa enfermedad dentro del gen htt que codifica la proteína Huntingtina mutante, de tal manera que se presenta la interferencia específica de dicho gen, en donde el agente de ARNi se dirige a una región polimórfica en el gen que es distinta de una región expandida de CAG dentro del gen mutante.
- 10 2. El agente de ARNi de la reivindicación 1 para uso de acuerdo con la reivindicación 1, en donde dicho gen comprende una región expandida de repetición de trinucleótidos.
- 15 3. El agente de ARNi de la reivindicación 1 para uso de acuerdo con la reivindicación 1, en donde dicha proteína mutante comprende un dominio expandido de poliglutamina.
- 20 4. El agente de ARNi de la reivindicación 1 para uso de acuerdo con cualquiera de las reivindicaciones 1 a 3, en donde el agente de ARNi comprende una primera cadena que comprende aproximadamente 16-25 nucleótidos homóloga a una región del gen que comprende el polimorfismo y una segunda cadena que comprende aproximadamente 16-25 nucleótidos complementarios a la primera cadena.
- 25 5. El agente de ARNi de la reivindicación 1 para uso de acuerdo con la reivindicación 1, en donde la cantidad efectiva es una cantidad efectiva para inhibir la expresión o actividad de la proteína mutante.
- 30 6. Un agente de ARNi que comprende una primera cadena que comprende aproximadamente 16-25 nucleótidos homólogos a una región de un gen que codifica una proteína Huntingtina mutante de ganancia de función, comprendiendo dicha región un solo polimorfismo de nucleótido que no causa enfermedad localizado en un sitio distinto de una mutación que causa enfermedad, y una segunda cadena que comprende aproximadamente 16-25 nucleótidos complementarios a la primera cadena, en donde el agente de ARNi dirige la escisión específica objetivo de un ARNm transcrito del gen que codifica la proteína Huntingtina mutante.
- 35 7. El agente de ARNi de cualquiera de las reivindicaciones precedentes, en donde dicho polimorfismo se selecciona ya sea de entre el grupo que comprende PI-P5, o seleccionado de entre el grupo que comprende P6-P43.
- 40 8. El agente de ARNi de la reivindicación 6 o 7, en donde la primera cadena comprende una secuencia de nucleótidos idéntica a la secuencia de la región que comprende el polimorfismo.
- 45 9. El agente de ARNi de la reivindicación 6, 7 u 8, que comprende además una porción de bucle que comprende 4 - 11 nucleótidos que conectan las dos cadenas.
- 50 10. Una molécula de ácido nucleico aislada que codifica el agente de ARNi de cualquiera de las reivindicaciones 6 a 9.
- 55 11. Un vector que comprende la molécula de ácido nucleico de la reivindicación 10.
- 60 12. El vector de la reivindicación 11, que es un vector viral, vector retroviral, casete de expresión, o plásmido; y en cuyo caso comprende además opcionalmente un promotor de ARN polimerasa III o ARN polimerasa II; y en cuyo caso opcionalmente además en donde el promotor de la ARN polimerasa III es el promotor U6 o HI.
- 65 13. Una célula huésped que comprende el agente de ARNi, la molécula de ácido nucleico, o el vector de cualquiera de las reivindicaciones 6 a 12.
14. La célula huésped de la reivindicación 13, que es una célula huésped de mamífero, en cuyo caso opcionalmente, en donde la célula huésped es una célula de mamífero no humano o una célula humana.
15. Una composición que comprende el agente de ARNi, la molécula de ácido nucleico, el vector o la célula huésped de cualquiera de las reivindicaciones 6 a 14 y un vehículo farmacéuticamente aceptable.
- 60 16. Un agente de ARNi seleccionado de ARNpi y ARNph para uso en el tratamiento de la enfermedad de Huntington mediante la identificación de un único polimorfismo de nucleótido en el gen htt que codifica una proteína Huntingtina mutante de ganancia de función y la administración del agente de ARNi que dirige dicho polimorfismo de tal manera que se reduce la proteína mutante, en donde el polimorfismo es un polimorfismo de un solo nucleótido que no causa enfermedad situado en un sitio distinto de una mutación dentro de un región polimórfica que es distinta de una región expandida de CAG dentro del gen mutante.

1 TTGCTGTGTG AGGCAGAACC TCGGGGGGCA GGGGGGGGCT GGTTCCTCTGG CCAGCCATTG
61 GCAGAGTCCG CAGGETAGGG CTGTCAATCA TGCTGGCCCG CGTGGCCCCG CCTCCGCCGG
121 CGGGCCCCG CCTCCGCCGG CGCACGCTG GACGCAAGG CGCCGTGGG GTCGCCGGGA
181 CGGGTCCAAG ATGGACGGCC GCTCAGGTC TGTTTTACC TCGGGCCAG AGCCCCATT
241 ATTGCCCCGG TGTGAGCGG CGCCGGAGT CGGCCGAGG CCTCCGGGA CTGCCGTGCC
301 GGGCGGAGA CCGCCGCG GACCTGGAA AAGCTGATGA AGGCCTTGA GTCCCTCAAG
361 TCCTTCCAGC AGCAGCAGCA GCAGCAGCAG CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG
421 CAGCAGCAGC AACAGCCGCC ACCCGCCCG CCGCCGCCG CGCCTCCTCA GCTTCCTCAG
481 CCGCCGCCCG AGGCACAGCC GCTGTGCCT CAGCCGCAGC CGCCCCCGCC GCCGCCCCCG
541 CCGCCACCCG GCCCGGCTGT GGCTGAGGAG CCGCTGCACC GACCAAGAA AGAACTTTCA
601 GCTACCAAGA AAGACCGTGT GAATCATTGT CTGACAATAT GTGAAAACAT AGTGGCACAG
661 TCTGTCAGAA ATTCTCCAGA AHTTCAGAAA CTTCTGGCA TCGCTATGGA ACTTTTCTG
721 CTGTGCAGTG ATGACGCAGA GTCAGATGTC AGGATGGTGG CTGACGAAATG CCTCAACAAA
781 GTTATCAAAG CTTTGATGGA TTCTAATCTT CCAAGGTTAC AGCTCGAGCT CTATRAAGAA
841 ATTAATAAGA ATGGTCCCC TCGGAGTTG CGTGTGCCC TGTGGAGGTT TGCTGAGCTG
901 GCTCACCTGG TTCGGCCTCA GAAATGCAGG CCTTACCTGG TGAACCTTCT GCCGTGCCCTG
961 ACTCGRACAA GCAAGAGACC CGAAGAATCA GTCCAGGAGA CCTTGGCTGC AGCTGTTCCC
1021 AAAATTATGG CTTCTTTTGG CAATTTTCCA AATGACAATG AATTAAGGT TTTGTTAAAG
1081 GCCTTCATAG CGAACCTGAA GTCAGCTCC CCCACCATC GCGGACAGC GGCTGGATCA
1141 GCAGTGAGCA TCTGCCAGCA CTCAGAAGG ACACAATATT TCTATAGTTG GCTACTAAAT
1201 GTGCTCTTAG GCTTACTCGT TCCTGTCCGAG GATGAACACT CCACTCTGCT GATTCTTGGC

FIG.1A

1261 GTGCTGCTCA CCCTGAGGTA TTTGGTGCCC TTGCTGCAGC AGCAGGTCAA GGACACAAGC
1321 CTGAARAGCA GCTTCGGAGT GACAAGGAAA GAATGGAAG TCTCTCCTTC TGCAGAGCAG
1381 CTTGTCCAGG TTTATGAACT GACGTTACAT CATAACACAGC ACCAAGACCA CAATGTTGTG
1441 ACCGGAGCCC TGGAGCTGTT GCAGCAGCTC TTCAGAAGCC CTCACCCCGA GCTTCIGCAA
1501 ACCCTGACCG CAGTCGGGGG CATTGGGCAG CTCACGGCTG CTAAGGAGGA GTCGTGGTGGC
1561 CGAAGCCGTA GTGGGAGTAT TGTGGAACCT ATAGCTGGAG GGGGTTCCCTC ATGCAGCCCT
1621 GTCCTTTCAA GAAAACAAAA AGGCAAAAGTG CTCTTAGGAG AAGAAGAAGC CTTGGAGGAT
1681 GACTCTGAAT CGAGATCGGA TGTCAGCAGC TCTGCCCTAA CAGCCTCAGT GAAGGATGAG
1741 ATCAGTGGAG AGCTGGCTGC TTCTTCAGGG GTTTCACCTC CAGGGTCAGC AGGTCATGAC
1801 ATCATCACAG AACAGCCACG GTCACAGCAC ACACTGCAGG CGGACTCAGT GGATCTGGCC
1861 AGCTGTGACT TGACAAGCTC TGCCACTGAT GGGGATGAGG AGGATATCTT GAGCCACACAG
1921 FCCAGCCAGG TCAGGCGCGT CCCATCTGAC CCTGCCATGG ACCTGAATGA TGGGACCCAG
1981 GCCTCGTCCG CCATCAGCGA CAGCTCCCAG ACCACCACCG AAGGGCCTGA TTCAGCTGTT
2041 ACCCCTTCAG ACAGTTCTGA AATTGTGTTA GACGGTACCG ACAACCAGTA TTTGGGCCCTG
2101 CAGATTGGAC AGCCCCAGGA TGAAGATGAG GAAGCCACAG GTATTCTTCC TGATGAAGCC
2161 TCGGAGGCCT TCAGGAACTC TTCCATGGCC CTTCAACAGG CACATTTATT GAAAAACATG
2221 AGTCACTGCA GGCAGCCTTC TGACAGCAGT GTTGATAAAT TTGTGTTGAG AGATGAAGCT
2281 ACTGAACCGG GTGATCAAGA AAACAAGCCT TGCCGCATCA AAGGTGACAT TGGACAGTCC
2341 ACTGATGATG ACTCTGCACC TCTTGTCCAT TGTTGTCGCC TTTTATCTGC TTCGTTTTTG
2401 CTAACAGGGG GAAAAAATGT GCTGGTTCCG GACAGGGATG TGAGGGTCAG CGTGAAGGCC
2461 CTGGCCCTCA GCTGTGTGGG AGCAGCTGTG GCCCTCCACC CGGAATCTTT CTTACAGCAA

FIG.1B

2521 CTCATATAAG TTCCTCTTGA CACCACGGAA TACCCTGAGG AACAGTATGT CTCAGACATC
 2581 TTGAACTACA TCGATCAATGG AGACCCACAG GTTCGAGGAG CCACTGCCAT TCTCTGTGGG
 2641 ACCCTCATCT GCTCCATCCT CAGCAGTCC CGCTTCCACG TGGGAGATTG GATGGGCACC
 2701 ATTAGAACCC TCACAGGAAA TACATTTTCT TTGGCGGATT GCATTCCTTT GCTGCGGAAA
 2761 ACACTGAAGG ATGAGTCTTC TGTTACTTGC AAGTTAGCTT GTACAGCTGT GAGGAACTGT
 2821 GTCATGAGTC TCTGCAGCAG CAGCTACAGT GAGTTAGGAC TGCAGCTGAT CATCGATGTG
 2881 CTGACTCTGA GGAACAGTTC CTATTGGCTG GTGAGGACAG AGCTTCTGGA AACCCTTGCA
 2941 GAGATTGACT TCAGGCTGGT GAGCTTTTTG GAGCAAAAG CAGAAAACCTT ACACAGAGGG
 3001 GTCATCATT ATACAGGGCT TTTAAAACCTG CAAGAACGAG TGCTCAATAA TGTTGTCAATC
 3061 CATTGTCTTG GAGATGAAGA CCCCAGGGTG CGACATGTTG CCGCAGCATC ACTAATTAGG
 3121 CTTGTCCCAA AGCTGTTTTA TAAATGTGAC CAAGGACAAG CTGATCCAGT AGTGGCCCGTG
 3181 GCAAGAGATC AAAGCAGTGT TTACCTGAAA CTTCTCATGC ATGAGACGCA GCCTCCATCT
 3241 CATTCTCCG TCAGCACAAAT AACCAAGATA TATAGAGGCT ATAACCTACT ACCAAGCATA
 3301 ACAGACGTCA CTATGGAAA TAACCTTTCA AGAGTTATTG CAGCAGTTTC TCATGAACTA
 3361 ATCACATCAA CCACCAGAGC ACTCACATTT GGATGCTGTG AAGCTTTGTG TCTTCTTTCC
 3421 ACTGCCTTCC CAGTTTGCAT TTGGAGTTTA GGTGGCAGT GTGGAGTGCC TCCACTGAGT
 3481 GCCTCAGATG AGTCTAGGAA GAGCTGTACC GTTGGGATGG CCACAATGAT TCTGACCCCTG
 3541 CTCTCGTCAG CTTGGTTCCC ATTGGATCTC TCAGCCCCATC AAGATGCTTT GATTTTGGCC
 3601 GGAAACTTGC TTGCAGCCAG TGCTCCCAA TCTCTGAGAA GTTCATGGGC CTCTGAAGAA
 3661 GAAGCCAACC CAGCAGCCAC CAAGCAAGAG GAGGTCTGGC CAGCCCTGGG GGACCCGGCC
 3721 CTGGTGCCCA TGGTGGAGCA GCTCTTCTCT CACCTGCTGA AGGTGATTAA CATTGTGTCC

FIG.1C

3781 CACGTCCTGG ATGACGTTGG TCCTGGACCC GCAATAAAGG CAGCCTTGCC TTCTCTAACA
3841 AACCCCCCTT CTCTAAGTCC CATCCGACGA AAGGGGAAGG AGAAGAACC AGGAGAACA
3901 GCATCTGTAC CGTTGAGTCC CAAGAAAGGC AGTGAGGCCA GTGCAGCTTC TAGACAAATCT
3961 GATACCTCAG GTCCGTGTAC AACAAGTAAA TCCTCATCAC TGGGGAGTTT CTATCATCTT
4021 CCTTCATACC TCMAACTGCA TGATGTCTG AAAGCTACAC ACGCTAACTA CAAGGTCACG
4081 CTGGATCTTC AGAACAGCAC GGAAGAAGTTT GGAGGGTTTC TCCGCTCAGC CTTGGATGTT
4141 CTTTCTCAGA TACTAGAGCT GGCCACACTG CAGGACATTG GGAAGTGTGT TGAAGAGATC
4201 CTAGGATACC TGAATCCTG CTTTAGTCGA GAACCAATGA TGGCAACTGT TTGTGTTCAA
4261 CAATTGTTGA AGACTCTCTT TGGCACAAAC TTGGCCTCCC AGTTTGATGG CTTATCTTCC
4321 AACCCACGCA AGTCACAAGG CCGAGCACAG CGCCTTGGCT CCTCCAGTGT GAGGCCAGGC
4381 TTGTACCACT ACTGCTTCAT GSCCCOGTAC ACCCACTTCA CCCAGGCCCT CGCTGACGCC
4441 AGCCTGAGGA ACATGGTGCA GCGGGAGCAG GAGAACGACA CCTCGGGATG GTTTGATGTC
4501 CTCCAGAAAG TGTCTACCCA GTTGAAGACA AACCTCACGA GTGTCACAAA GAACCGTGCA
4561 GATAAGAATG CTATTCATAA TCACATTCGT TTGTTTGAAC CTCTTGTTAT ARAAGCTTTA
4621 AACAGTACA CGACTACAAC ATGTGTGCAG TTACAGAAGC AGGTTTAGA TTTGCTGGCG
4681 CAGCTGGTTC AGTTACGGGT TAATTACTGT CTTCTGGATT CAGATCAGGT GTTTATTGGC
4741 TTTGTATTGA AACAGTTTGA ATACATTGAA GTGGGCCAGT TCAGGGAATC AGAGGCAATC
4801 ATTCCAAAACA TCTTTTCTT CTTGGTATTA CTAICTTATG AAGCTATCA TTCAAAACAG
4861 ATCATTGGAA TTCCTAAAT CATTACGCTC TGTGATGGCA TCATGGCCAG TGAAGGAAG
4921 GCTGTGACAC ATGCCATACC GGCTCTGCAG CCCATAGTCC ACGACCTCTT TGTATTAAGA
4981 GGAACAAATA AAGCTGATGC AGGAAAAGAG CTTGAAACCC AAAAAGAGGT GGTGGTGTCA

FIG.1D

5041 ATGTTACTGA GACTCATCCA GTACCATCAG GTGTTGGAGA TGTTCACTCT TGTCCTGCAG
5101 CAGTGCCACA AGGAGAATGA AGACAAGTGG AAGCGACTGT CTCGACAGAT AGCTGACATC
5161 ATCCTCCCAA TGTTAGCCAA ACAGCAGATG CACATTGACT CTCATGAAGC CCTTGGAGTG
5221 TTAATAACAT TATTTGAGAT TTTGGCCCTC TCCTCCCTCC GTCCGGTAGA CATGCTTTTA
5281 CGGAGTATGT TCGTCACTCC AAACACAATG GCGTCCGIGA GCACTGTTCA ACTGTGGATA
5341 TCGGGAATTC TGGCCATTTT GAGGGTTCTG ATTTCCCCAGT CAACTGAAGA TATTGTTCTT
5401 TCTCGTATTC AGGAGCTCTC CTTCTCTCCG TATTTAATCT CCTGTACAGT AATTAATAGG
5461 TTAAGAGATG GGGACAGTAC TTCAACGCTA GAAGAACACA GTGAAGGGAA ACAATAAAG
5521 AATTTGCCAG AAGAAACATF TTCAAGGTTT CTATTACAAC TGGTTGGTAT TCTTTTAGAA
5581 GACATTGTTA CAAAACAGCT GAAGGTGGAA ATGAGTGAGC AGCAACATAC TTTCTATTGC
5641 CAGGAACTAG GCACACTGCT AATGTGTCTG ATCCACATCT TCAAGTCTGG AATGTTCCGG
5701 AGAATCACAG CAGCTGCCAC TAGGCTGTTT CCGAGTGATG GCTGTGGCGG CAGTTTCTAC
5761 ACCCTGGACA GCTTGAACIT GCGGGCTCGT TCCATGATCA CCACCCACCC GGCCCTGGTG
5821 CTGCTCTGGT GTCAGATACT GCTGCTTGTC AACCCACCCG ACTACCCGCTG GTGGGCAGAA
5881 GTGCAGCAGA CCCCAGAAAAG ACACAGTCTG TCCAGCACAA AGTTACTTAG TCCCCAGATG
5941 TCTGGAGAAG AGGAGGATTC TGACTTGGCA GCCAAACTTG GAATGTGCAA TAGAGAAATA
6001 GTACGAAGAG GGGCTCTCAT TCTCTTCTGT GATTATGICT GTCAGAACCCT CCATGACTCC
6061 GAGCACTTAA CGTGGCTCAT TGTAATCAC ATTCAAGATC TGATCAGCCT TTCCCCACGAG
6121 CCTCCAGTAC AGGACTTCAT CAGTGCCGPT CATCGGAACT CTGCTGCCAG CGGCCTGTTC
6181 ATCCAGGCAA TTCAGTCTCG TTGTGAAAAC CTTTCAACTC CAACCATGCT GAAGAAAAC
6241 CTTCAGTGCT TGGAGGGGAT CCATCTCAGC CAGTCGGGAG CTGTGCTCAC GCTGTATGTG

FIG.1E

6301 GACAGGCTTC TGTGCACCCC TTTCCGTGTG CTGGCTCGCA TGGTCGACAT CCTTGCTTGT
 6361 CGCCGGGTAG AAATGCTTCT GGCTGCAAAAT TTACAGAGCA GCAITGGCCCA GTTGCCAATG
 6421 GAAGAACTCA ACAGAAATCCA GGAATACCTT CAGAGCAGCG GGCTCGCTCA GAGACACCAA
 6481 AGGCTCTATT CCCTGCTGGA CAGGTTTCGT CTCTCCACCA TGCAAGACTC ACTTAGTCCC
 6541 TCTCCTCCAG TCTCTTCCCA CCCGCTGGAC GGGGATGGC ACGTGTCACT GGAACACAGT
 6601 AGTCCGGACA AAGACTGGTA CGTTCATCTT GTCAAATCCC AGTGTGGAC CAGGTCAGAT
 6661 TCTGCACTGC TGGAAGGTGC AGAGCTGGTG AATCGGATT C TGCTGAAGA TATGAATGCC
 6721 TTCATGATGA ACTCGGAGTT CAACCTAAGC CTGCTAGCTC CATGCTTAAG CCTAGGGATG
 6781 AGTGAATTT CTGGTGGCCA GAAGAGTGCC CTTTTGAAG CAGCCCGTGA GGTGACTCTG
 6841 GCCCGTGTGA GCGGCACCGT GCAGCAGCTC CCTGCTGTCC ATCATGTCTT CCAGCCCGAG
 6901 CTGCCTGCAG AGCCGGCGGC CTACTGGAGC AAGTTGAATG ATCTGTTGG GGATGCTGCA
 6961 CTGTATCAGT CCCTGCCCCAC TCTGGCCCCG GCCCTGGCAC AGTACCTGGT GGTGGICTCC
 7021 AAAC TGCCCCA GTCATTTGCA CCTTCCTCCT GAGAAAGAGA AGGACATTGT GAAATTCGTG
 7081 GTGGCAACCC TTGAGGCCCT GTCCTGGCAT TTGATCCATG AGCAGATCCC GCTGAGTCTG
 7141 GATCTCCAGG CAGGGCTGGA CTGCTGCTGC CTGGCCCTGC AGCTGCCTGG CCTCTGGAGC
 7201 GTGGTCTCCT CCACAGAGTT TGTGACCCAC GCCTGCTCCC TCATCTACTG TGTGCACTTC
 7261 ATCCTGGAGG CCGTTGCAGT GCAGCCTGGA GAGCAGCTTC TTAGTCCAGA AAGAAGGACA
 7321 AATAACCCAA AAGCCATCAG CGAGGAGGAG GAGGAAGTAG ATCCAACAC ACAGAACTCT
 7381 AAGTATATCA CTGCAGCCTG TGAGATGGTG GCAGAAATGG TGGAGTCTCT GCAGTCGGTG
 7441 TTGGCCTTGG GTCATAAAAG GAATAGCGGC GTGCCGGCGT TTCACAGCC ATTGCTCAGG
 7501 AACATCATCA TCAGCCTGGC CCGCCTGCC CTTGTCAACA GCTACACACG TGTGCCCCCA

7561 CTGGTGTGGA AGCTTGGATG GTCACCCAAA CCGGGAGGGG ATTTTGGCAC AGCATTCCCT
 7621 GAGATCCCCG TGGAGTTCCT CCAGGAAAAG GAAGTCTTTA AGGAGTTCAT CTACCCGCATC
 7681 AACACACTAG GCTGGACCCAG TCGTACTCAG TTTGAAGAAA CTTGGGCCAC CCTCCTTGGT
 7741 GTCCTGGTGA CGCAGCCCCT CGTGATGGAG CAGGAGGAGA GCCCACCAGA AGAAGACACA
 7801 GAGAGGACCC AGATCAACGT CCTGGCCGTG CAGGCCATCA CCTCACTGGT GCTCAGTGCA
 7861 ATGACTGTGC CTGTGGCCGG CAACCCAGCT GTAAGCTGCT TGGAGCAGCA GCCCCGGAAC
 7921 AAGCCTCTGA AAGCTCTCGA CACCAGGTTT GGGAGGAAGC TGAGCATTAT CAGAGGGATT
 7981 GTGGAGCAAG AGATTCAAGC AATGGTTTCA AAGAGAGAGA ATATTGCCAC CCATCATTTA
 8041 TATCAGGCAT GGGATCCTGT CCCTTCTCTG TCTCCGGCTA CTACAGGTGC CCTCATCAGC
 8101 CACGAGAAGC TGCTGCTACA GATCAACCCC GAGCGGGAGC TGGGGAGCAT GAGCTACAAA
 8161 CTCGGCCAGG TGTCATACA CTCCGTGTGG CTGGGGAACA GCATCACACC CCTGAGGGAG
 8221 GAGGAATGGG ACGAGGAAGA GGAGGAGGAG GCCGACGCC CTTGCACCTTC GTCACCACCC
 8281 ACGTCTCCAG TCAACTCCAG GAAACACCGG GCTGGAGTTG ACATCCACTC CTGTTCCGAG
 8341 TTTTGTCTG AGTTGTACAG CCGCTGGATC CTGCCGTCCA GCTCAGCCAG GAGGACCCCG
 8401 GCCATCCTGA TCAGTGAGGT GGTCAGATCC CTTCTAGTGG TCTCAGACTT GTTCACCGAG
 8461 CGCAACCAGT TTGAGCTGAT GTATGTGACG CTGACAGAAC TGGGAAGGGT GCACCCCTCA
 8521 GAAGACGAGA TCCTCGCTCA GTACCTGGTG CCTGCCACCT GCAAGGCAGC TGCCGTCCCTT
 8581 GGGATGGACA AGGCCGTGGC GGAGCCTGTC AGCCGCCCTGC TGGAGAGCAC GCTCAGGAGC
 8641 AGCCACCTGC CCAGCAGGGT TGGAGCCCTG CACGGCGTCC TCTATGTGCT GGAGTGGAC
 8701 CTGCTGGACG ACACTGCCAA GCAGCTCATC CCGGTCATCA GCGACTATCT CCTCTCCAAC
 8761 CTGAAAAGGA TCGCCCACTG CGTGAACATT CACAGCCAGC AGCACGTACT GGTCATGTGT

FIG.1G

8821 GCCACTGCCT TTTACCTCAT TGAGAACTAT CCTCTGGACG TAGGGCCGGA ATTTTCAGCA
 8881 TCAATAATAC AGATGTGIGG GGTGATGCTG TCTGGAAGTG AGGAGTCCAC CCCCTCCATC
 8941 ATTTACCACT GTGCCCTCAG AGGCCTGGAG CGCCTCCTGC TCTCTGAGCA GCTCTCCCGC
 9001 CTGGATGCAG AATCGCTGGT CAAGCTGAGT GTGGACAGAG TGAACGTGCA CAGCCCGCAC
 9061 CGGGCCATGG CGGCTCTGGG CCTGATGCTC ACCTGCATGT ACACAGGAAA GGAGAAAGTC
 9121 AGTCCGGGTA GAACTTCAGA CCTAATCCT GCAGCCCCCG ACAGCGAGTC AGTGATGTGT
 9181 GCTATGGAGC GGGTATCTGT TCTTTTGTAT AGGATCAGGA AAGGCTTTCC TTGTGAAGCC
 9241 AGAGTGGTGG CCAGGATCCT GCCCCAGTTT CTAGACGACT TCTTCCCACC CCAGGACATC
 9301 ATGAACAAG TCATCGGAGA GTTCTGTCC AACCAGCAGC CATAACCCCA GTTCATGGCC
 9361 ACCGTGGTGT ATAAGGTGTT TCAGACTCTG CACAGCACCG GGCAGTCTG CATGGTCCGG
 9421 GACTGGGTCA TGCTGTCCCT CTCCAATTC ACGCAGAGGG CCCC GGTCGC CATGGCCACG
 9481 TGGAGCCTCT CCTGCTTCTT TGTCAGCGCG TCCACCAGCC CGTGGGTCGC GCGGATCCTC
 9541 CCACATGTCA TCAGCAGGAT GGGCAAGCTG GAGCAGGTGG ACGTGAACCT TTTCTGCCCTG
 9601 GTCGCCACAG ACTTCTACAG ACACCAGATA GAGGAGGAGC TCGACCCGAG GGCCTTCCAG
 9661 TCTGTGCTTG AGGTGGTTGC AGCCCCAGGA AGCCCATATC ACCGGCTGCT GACTTGTTTA
 9721 CGAAATGTCC ACAAGGTAC CACCTGC XXXXXXXXXX GCGCCATGGT GGGAGAGACT GTGAGGCCGGC
 9781 AGCTGGGGCC GGAGCCTTTG GAAGTCTGTG CCCTTGTGCC CTGCCCTCCAC CGAGCCAGCT
 9841 TGGTCCCTAT GGGCTTCCGC ACATGCCCGG GCGGCCCAGG CAACGTGCGT GTCTCTGCCA
 9901 TGTGGCAGAA GTGCTCTTTG TGGCAGTGGC CAGGCAGGGA GTGTCTGCAG TCCTGGTGGG
 9961 GCTGAGCCTG AGGCCTTCCA GAAAGCAGGA GCAGCTGTGC TGCACCCCAT GTGGGTGACC
 10021 AGGTCCCTTC TCCTGATAGT CACCTGCTGG TTGTTGCCAG GTTGCAGCTG CTCTTGCATC

FIG.1H

10081 TGGCCAGAA GTCCCTCCCTC CTGCAGGCTG GCTGTGGCC CCTCTGCTGT CCTGCAGTAG
 10141 AAGGTGCCGT GAGCAGGCTT TGGGAACACT GGCCTGGGTC TCCCTGGTGG GGTGTGCATG
 10201 CCACGCCCCG TGTCTGGATG CACAGATGCC ATGGCCTGTG CTGGCCAGT GGCTGGGGGT
 10261 GCTAGACACC CGGCACCATT CTCCCTTCTC TCTTTTCTTC TCAGGATTTA AAATTTAATT
 10321 ATATCAGTAA AGAGATTAAT TTTAACGAAC TCTTTCTATG CCCGTGTAAA GTATGTGAAT
 10381 CGCAAGGCCCT GTGCTGCATG CGACAGCGTC CGGGGTGGTG GACAGGGCCC CCGGCCACGC
 10441 TCCCTCTCCT GTAGCCACTG GCATAGCCCT CCTGAGCACC CGCTGACATT TCCGTTGTAC
 10501 ATGTTCTCTGT TTATGCATT CACAAGGTGAC TGGGATGTAG AGAGCCGTTA GTGGGCAGGT
 10561 GGCCACAGCA GGAATGAGGA CAGGCCCCCA TTATCCTAGG GGTGGCTCA ACTGCAGCCC
 10621 CTCCTCCTCG GGCACAGACG ACTGTGTTT TCCACCCACC AGTCAGGGAC AGCAGCCCTCC
 10681 CTGTCACTCA GCTGAGAAGG CCAGCCCTCC CTGGCTGTGA GCAGCCCTCCA CTGTGTCCAG
 10741 AGACATGGGC CTCCCCTCC TGTTCCTTGC TAGCCCTGGG GTGGCGTCTG CCTAGGAGCT
 10801 GGCTGGCAGG TGTGGGACC TGCTGCTCCA TGGATGCATG CCCTAAGAGT GTCACCTGAGC
 10861 TGTGTTTTGT CTGAGCCTCT CTCGGTCAAC AGCAAAGCTT GGTGTCTTGG CACTGTTAGT
 10921 GACAGAGCCC AGCATCCCTT CTGCCCCCGT TCCAGCTGAC ATCTTGCACG GTGACCCCTT
 10981 TTAGTCAGGA GAGTGCAGAT CTGTGCTCAT CGGAGACTGC CCCACGGCCC TGTCAGAGCC
 11041 GCCACTCCTA TCCCACGGAC AGTCCCCTGG ACCAGCCTCC TGTITGCAGG CCCAGAGGAG
 11101 CCAAGTCATT AAAATGGAAG TGGATTCTGG ATGGCCGGGC TGCTGCTGAT GTAGGAGCTG
 11161 GATTGGGAG CTCTGCTTGC CGACTGGCTG TGAGACGAGG CAGGGGCTCT GCTTCCTCAG
 11221 CCCTAGAGGC GAGCCAGGCA AGTTGGCGA CTGTCAATG GCTTGGTTG GTCATGCCCG
 11281 TCGATGTTTT GGGTATTGAA TGTGGTAAGT GGAGGAAATG TTGGAACCTCT GTGCAGGTGC

FIG.1I

11341 TGCCTTGAGA CCCCCAAGCT TCCACCTGTC CCTCTCCTAT GTGGCAGCTG GGGAGCAGCT
11401 GAGATGTGGA CTTGTATGCT GCCCACATAC GTGAGGGGGA GCTGAAGAAGG AGCCCCCTGCT
11461 CAAAGGGAGC CCCTCCTCTG AGCAGCCTCT GCCAGGCCCTG TATGAGGCTT TTTCCACCAG
11521 CTCCCACAG AGGCCTCCCC CAGCCAGGAC CACCTCGTCC TCGTGGCGGG GCAGCAGGAG
11581 CGGTAGAAAG GGTCCGATG TTTGAGGAG CCCTTAAGGG AAGCTACTGA ATTATAACAC
11641 GTAAGAAAAT CACCATTCCT CCGTATTGGT TGGGGGCTCC TGTTCCTCAT CCTAGCTTTT
11701 TCCTGGAAAA GCCCGCTAGA AGGTTTGGGA ACGAGGGGAA AGTTCACAGA ACTGTTGCTG
11761 CTCCCACCC GCCTCCCGCC TCCCCCGCAG GTTATGTCAG CAGCTCTGAG ACAGCAGTAT
11821 CACAGGCCAG ATGTTGTTCC TGGCTAGATG TTTACATTTG TAAGAAATAA CACTGTGAAT
11881 GTAAAACAGA GCCATTCCCT TGGAAATGCAT ATCGCTGGGC TCAACATAGA GTTTGTCTTC
11941 CTCTTGTTTA CGACGTGATC TAAACCAGTC CTTAGCAAGG GGCTCAGAAC ACCCCGCTCT
12001 GGCAGTAGGT GTCCCCCACC CCCAAGACC TGCCTGTGTG CTCCGGAGAT GAATATGAGC
12061 TCATTAGTAA AATGACTTC ACCCAGGCAT ATACATAAAG TATCCATGCA TGTGCATATA
12121 GACACATCTA TAATTTTACA CACACACCTC TCAAGACGGA GATGCATGGC CTCTAAGAGT
12181 GCCCCGTGCG GTTCTTCC TGAAAGTTGACT TTCCTTAGAC CCGCCAGGTC AAGTTAGCCG
12241 CGTGACGGAC ATCCAGGCCG GGGACGTGGT CAGGGCAGGG CTCATTTCAT GCCCCACTAGG
12301 ATCCCACCTGG CGAAGATGGT CTCCATATCA GCTCTCTGCA GAAGGGAGGA AGACTTTATC
12361 ATGTTCCCTAA AAATCTGTGG CAAGCACCCA TCGTATTATC CAAATTTTGT TGCAAAATGTG
12421 ATTAATTTGG TTGTCAAGTT TTGGGGGTGG GCTGTGGGGA GATTGCTTTT GTTTTCCTGC
12481 TGGTAATATC GGGAAAGATT TTAATGAAC CAGGGTAGAA TTGTTTGGCA ATGCACTGAA
12541 GCGTGTCTCT TTCCCAAAAT GTGCCTCCCT TCCGCTGCGG GCCCAGCTGA GTCTATGTAG

FIG.1J

12601 GTGATGTTTC CAGCTGCCAA GTGCTCTTTG TTA CTGTCCA CCCTCATTTT TGCCAGCGCA
 12661 TGTGTCCTTT CAAGGGGAAA ATGTGAAGCT GAACCCCTC CAGACACCCA GAATGTAGCA
 12721 TCTGAGAAGG CCTGTGCCC TAAAGGACAC CCTCGCCCC CATCTTCATG GAGGGGGTCA
 12781 TTTCAGAGCC CTCGGAGCCA ATGAACAGCT CCTCCTCTTG GAGCTGAGAT GAGCCCCACG
 12841 TGGAGCTCGG GACGGATAGT AGACAGCAAT AACTCGGTGT GTGGCCGCCT GGCAGGTGGA
 12901 ACTTCCCTCC GTTGGGGGGT GGAGTGAGGT TAGTCTGTG TGTCTGGTGG GTGGAGTCAG
 12961 GCTTCTCTTG CTACCTGTGA GCATCCTTCC CAGCAGACAT CCTCATCGGG CTTTGTCCCT
 13021 CCCCCTCTTC CTCCTCTGC GGGGAGGACC CGGGACCACA GCTGCTGGCC AGGGTAGACT
 13081 TGGAGCTGTC CTCAGAGGG GTCACGTGTA GGAGTGAGAA GAAGGAAGAT CTTGAGAGCT
 13141 GCTGAGGGAC CTTGGAGAGC TCAGGATGGC TCAGACGAGG AACTCGCTT GCCGGGCCCTG
 13201 GCCCTCCTGG GAAGGAGGGA GCTGCTCAGA ATGCCGCATG ACAACTGAAG GCAACCTGGA
 13261 AGGTTCAGGG CCCGCTCTC CCCCATGTGC CTGTCACGCT CTGGTGCAGT CAAAGGAACG
 13321 CCTTCCCCTC AGTTGTTTCT AAGAGCAGAG TCTCCCCTG CAATCTGGGT GGTAAC TGCC
 13381 AGCCTTGGAG GATCGTGGCC AACGTGGACC TGCCTACGGA GGGTGGGCTC TGACCCAAAGT
 13441 GGGGCTCCT TGCCCAGGTC TCACTGCTTT GCACCGTGGT CAGAGGGACT GTCAGCTGAG
 13501 CTTGAGCTCC CCTGGAGCCA GCAGGGCTGT GATGGGGCAG TCCC GGAGCC CCACCCAGAC
 13561 CTGAATGCTT CTGAGAGCAA AGGGAAGGAC TGACGAGAGA TGTATATTTA ATTTTTTAAC
 13621 TGCTGCAAAC ATTGTACATC CAAATTAAG GGAATAAATG GAAACCATCA AT

FIG.1K

1 matleklmka feslksfqqq qqqqqqqqqq qqqqqqqqqq pppppppppp pqlpqqppqa
 61 qpllpqppp pppppppppp avaeplhrp kkelstakkd rvnhcltice nivaqsvrns
 121 pefqkllgia melflcsdd aedsvrmvad eclnkvikal mdsnlprlql elykeikkng
 181 aprslraalw rfaelahlv pdkcrpylvn llpcltrtsk rpeesvgetl aaavpkimas
 241 fgnfandnei kvllkafian lkssptirr taagsavsic qhsrrtqfyf swllnvlgl
 301 lvpvedehst llilgvlltl rylvplllqq vktstlkgf gvtirkemevs psaeqlvqv
 361 eltlhhtqh dhnvvtgale llqqlfrtpp pellqtltav ggigqltaak eesggrsrsg
 421 siveliaagg sscspvlrsk qkgkvllgee ealeddesr sdvssalta svkdeisgel
 481 aassgvstpg saghditeq prsqhtlqad svdlascldt ssatdgdced ilshsssqvs
 541 avpsdpamd ndgtqasspi sdssqttteg pdsavtppsds seivldgtdn qylglqigqp
 601 qdedeeatgi lpdeaseafr digqstddd aplvhcvrll sasflltggk nvlvpdrdrv vsvkalalasc
 661 qenkpcrik digqstddd aplvhcvrll sasflltggk nvlvpdrdrv vsvkalalasc
 721 vgaavalhpe sffsklykvp ldtteypee qvdsilnyid hgdppvzgat ailcgtlics
 781 ilsrsrfhvq dwmgtirtlt gntfsladci plllrktlkde ssvtcklact avrncvmslc
 841 sssyseglq liidvltlrn sswlvrtel letlaeidfr lvsfleakae nlhrghahyt
 901 gllkgervl nnvihilgd edprvrhvae aslirlvpkl fykcdqggad pvvavardqs
 961 svylkllmhe tqppshfsvs titriyrgyn llpsitdvtm ennlsvviaa vshelitstt
 1021 raltfgccea lcllstafov ciwslgwhcg vpplsasdes rkscvtvgmt miltllssaw
 1081 fpldlsahqd alilagnlla asapkslrsv waseeeanpa atkqeevwpal gdralvpmv
 1141 eqlfshllkv inicahvldd vapgpaikaa lpsltnppsl spirrkqkek epgeqasvpl
 1201 spkkgseasa asrqsdtsqp vttskssslg sfyhlpsylk lhdvltkatha nykvtldlqn
 1261 stekfggflr saldvlsqil elatldqdig cveeilgylk scfsrepmma tvcvqqlkt
 1321 lfgtnlasqf dglssnpsks qgraqrlgss svrpglyhyc fmaythftq aladaslrnm
 1381 vqaecqndts gwfdvlgkvs tqkntltsv tknraknai nhhirlfep1 vikalkqytt
 1441 ttcvqlqkv ldllaqlvql rvnycldsd qvfigfvlkq feyievqgfr eseaiipnif
 1501 fflvllsyer yhsqigigip kiiqlcdgim asgrkavtha ipalqpivhd lfvlrgrtnka
 1561 dagkeletqk evvsmllrl iqyhqvlemf ilvlgqchke nedkwrler qiadiilpml
 1621 akqgmhidsh ealgvlnltlf eilapsslrp vdmllrsmfv tpntmasvst vqlwisgila
 1681 ilrvlisqst edivlsriqe lsfspylisc tvinlrldgd ststleehse gkqiknlpee

FIG.2A

1741 tfsrlllqlv gilledivtk qlkvemsegg htfcyqelgt llmclihfk sgmfrritaa
1801 atrlfrsdgc ggsfytldsl nlrarsmitt hpalvllwcq illlvnhtdy rrwaevqgtp
1861 krhslsstk1 lspqmsgeee dsdlaaklgm cnreivrrga lilfcdyvcq nlhdsehltw
1921 livnhiqdli slsheppvqd fisavhrnsa asglfiqaiq srcenlstpt mlkktlqcle
1981 gihlsqsgav ltlyvdrllc tpfvrarmv dilacrrvem llaanlqssm aqlpmeelnr
2041 iqeylqssgl aqrhqrlysl ldrfirlstmq dsllspppvs shpldgdghv sletvspdkd
2101 wyvhlvksqc wtrdsalle gaelvnrifa edmnafmms efnlslapc lslgmseisg
2161 gqksalfeaa revtlarvsg tvqqlpavhh vfqelpaep aaywsklndl fgdaalyqsl
2221 ptlaralagy lvvsklpsh lhlppeked ivkfvvatle alswhliheq iplsldlqag
2281 ldccclalql pglwsvsst efvthacsl iycvhfileav avqpeqlls pertrntpka
2341 iseeeeevdp ntqnpkyita acemvaemve slqsvlalgh krnsgvpaf1 tp1lrniis
2401 larlplvnsy trvpplvwkl gwsppkpggdf gtafpeipve flqekevfk fiyrntlgw
2461 tsrtqfeetw atllgvlvtq plvmeqeesp peedertqi nvlavqaits lvlsamtvpv
2521 agnpavscle qqprnkplka ldtrfgrkls iirgiveqei qamvskreni athhlygawd
2581 pvpslspatt galishekl1 lqinperelg smsyklgqvs ihsvwlgnsi tplreeewde
2641 eeeeeadapa psspptspvn srkhragvdi hscsqflllel ysrwilpsss arrtpailis
2701 evvrsllvvs dlfterngfe lmyvtltelr rvhpsedeil agylvpatck aaavlgmdka
2761 vaepvsrille stlrsshlpv rvgalhgvlv vlecldlddt akqlipvisd yllsnlkgia
2821 hcvnihsqqh vlvmcatafy lienypldvg pefsasiiqm cgvmlsgsee stpsiiyhca
2881 lrglerlls eqlsrldaes lvklsvdrvn vhsphramaa lglmltcmtyt gkekvspgrrt
2941 sdpnpaapds esvivamerv svlfdrirkv fpcearvvar ilpqflddff ppqdimnkvi
3001 geflnqqpy pqfmatvvyk vftqlhstgg ssmvrdwvml slsftqrap vamatswslsc
3061 ffsastspw vaailphvis rmgklegvdv nlfclvatdf yrhgieeeld rrafqsvlev
3121 vaapgsphyh lltclrnvhk vtcc

FIG.2B

FIG. 3

htt sentido objetivo: 5'...ugcagcugaucaucgauugcugacccugaggaacaguuc...-3'
htt antisentido objetivo: 3'...acgucgacuaguagcuacacgacugggacuccuugucaag...-5'

FIG. 4

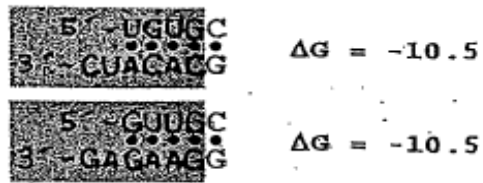


FIG. 5A

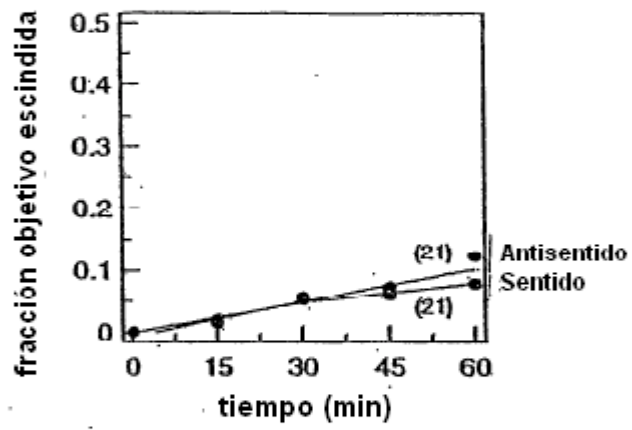
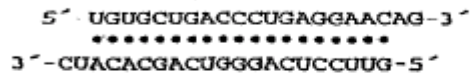


FIG. 5B

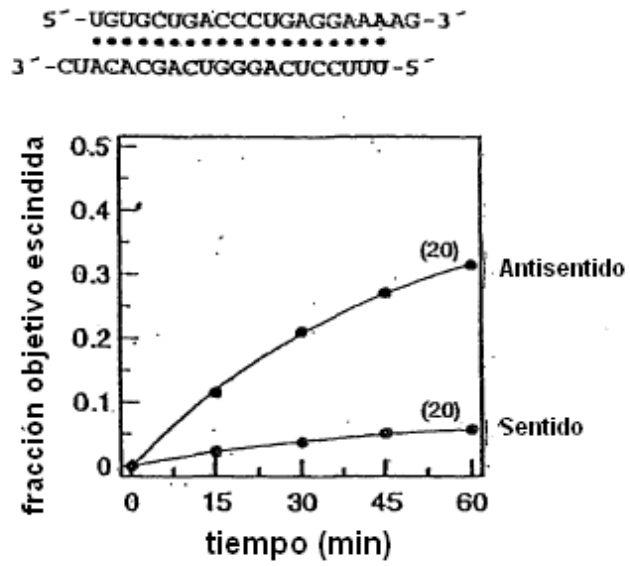


FIG. 5C

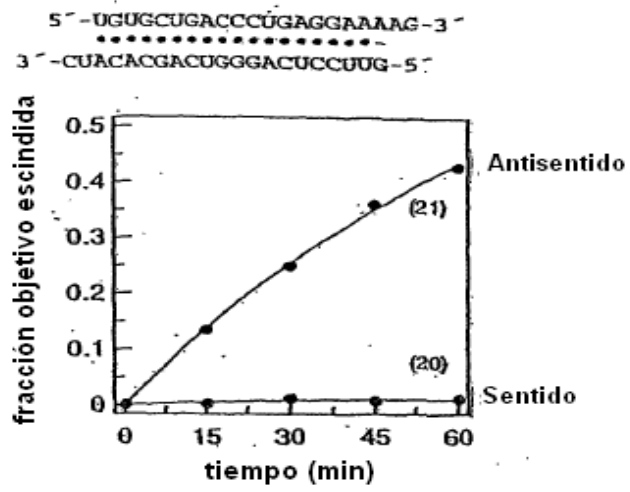
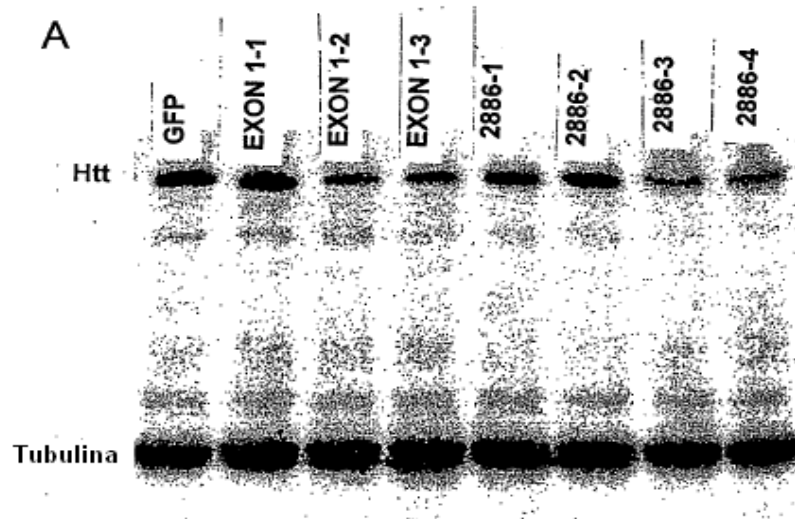


FIG. 6



B

