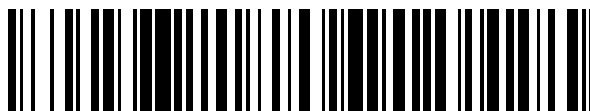


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 808 561**

51 Int. Cl.:

C07H 21/04 (2006.01)

A61K 48/00 (2006.01)

C12N 15/113 (2010.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.09.2004** **E 14168602 (2)**

97 Fecha y número de publicación de la concesión europea: **29.04.2020** **EP 2821085**

54 Título: **Interferencia por ARN para el tratamiento de trastornos de 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:

01.03.2021

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:

LEHMANN NOVO, María Isabel

ES 2 808 561 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

Interferencia por ARN para el tratamiento de trastornos de ganancia de función

Solicitudes relacionadas

La presente solicitud de patente reivindica el beneficio de la solicitud de patente provisional de EE. UU. N° de serie 60/502.678, titulada "RNA Interference for the Treatment of Gain-of-Function Disorders", presentada el 12 de septiembre de 2003.

Antecedentes de la invención

La interferencia por ARN (iARN) es el mecanismo de silenciamiento génico postranscripcional específico de secuencia por ARN bicatenarios (ARNbc) homólogos al gen que se expresa. Los ARNbc se procesan por Dicer, una ribonucleasa III celular, para generar dúplex de aproximadamente 21 nt con nucleótidos protuberantes en 3' (ARN interferente pequeño, ARNip) que median en la degradación del ARNm específico de secuencia. En células de mamífero, las moléculas de ARNip son capaces de silenciar específicamente la expresión génica sin inducción de la vía de respuesta del interferón no específica. Así, los ARNip se han convertido en una alternativa nueva y poderosa a otras herramientas genéticas, tales como oligonucleótidos no codificantes y ribozimas para analizar la función génica. Además, se están desarrollando ARNip para fines terapéuticos con el objetivo de silenciar los genes de enfermedad en seres humanos.

Las enfermedades de repetición de trinucleótidos comprenden un grupo recientemente reconocido de trastornos heredados. La mutación genética común es un aumento en una serie de una repetición de trinucleótidos particular. Hasta la fecha, la repetición de trinucleótidos más frecuente es CAG, que codifica el aminoácido glutamina. Se conocen al menos 9 enfermedades de repetición de CAG y existen más de 20 variedades de estas enfermedades, que incluyen enfermedad de Huntington, 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 dominante autosómica.

La aparición de las enfermedades ocurre, en general, de los 30 a 40 años, pero en enfermedad de Huntington las repeticiones de CAG en el gen huntingtina de >60 predicen una aparición juvenil.

La investigación reciente por los presentes inventores ha mostrado que la mutación genética (aumento en la longitud de repeticiones de CAG de normal <36 en el gen huntingtina a >36 en enfermedad) está asociada con la síntesis de una proteína huntingtina mutante, que tiene >36 poliglutaminas (Aronin et al., 1995). También se ha mostrado que la proteína forma agregados citoplásmicos e inclusiones nucleares (Difiglia et al., 1997) y se asocia con vesículas (Aronin et al., 1999). No se conocen vías patógenas precisas.

Se cree que la enfermedad de Huntington (y por implicación otras enfermedades de repetición de trinucleótidos) se provoca, al menos en parte, por interacciones aberrantes de proteína, que provocan la alteración de los procesos neuronales críticos, disfunción neuronal y por último lugar muerte neuronal (neurodegeneración en áreas del cerebro denominadas el cuerpo estriado y la corteza). En la búsqueda de un tratamiento eficaz para estas enfermedades, los investigadores en este campo enfatizaron el entendimiento de la patogénesis de la enfermedad e inicialmente buscaron mediar al nivel de las supuestas interacciones aberrantes de proteína. Sin embargo, no existe tratamiento eficaz para la enfermedad de Huntington u otras enfermedades de repetición de trinucleótidos. Además, se ha apreciado ahora que múltiples procesos anormales podrían ser activos en estos tipos de enfermedad.

Sumario de la divulgación

La invención presentemente reivindicada se define en y por las reivindicaciones adjuntas.

La presente divulgación se refiere a los métodos de tratamiento de una variedad de enfermedades de ganancia de función. En particular, la divulgación proporciona métodos para la destrucción selectiva de ARNm mutantes transcritos a partir de genes mutantes de ganancia de función, previniendo así la producción de proteínas mutantes codificadas por dichos genes. Se han propuesto otros métodos basados en iARN para la destrucción de genes mutantes en los que se silencian los ARNip, por ejemplo, una mutación puntual que ocurre en un único alelo en el gen mutante (por ejemplo, la mutación puntual en el gen superóxido dismutasa (SOD) asociado a esclerosis lateral amiotrófica (ELA)). Sin embargo, existe una diferencia clave entre ELA y enfermedades de repetición de trinucleótidos, tales como la enfermedad de Huntington. La ELA tiene una mutación puntual en un alelo como cambio genético, mientras que las enfermedades de repetición de trinucleótidos tienen una región ampliada de repeticiones de CAG en un alelo como cambio genético. El uso de iARN contra la región ampliada de repeticiones de CAG tiene posibles complicaciones. Se sabe que existen más de 80 genes normales con regiones de repeticiones CAG en células. Así, no se pueden usar los ARNip que silencian estas repeticiones de CAG sin arriesgar la destrucción generalizada de ARNm que contienen repeticiones de CAG normales. Asimismo, el silenciamiento de sitios específicos no alélicos daría como resultado la pérdida de tanto huntingtina normal como mutante que provoca la disfunción neuronal.

Los métodos de la divulgación utilizan tecnología de interferencia por ARN (iARN) contra regiones polimórficas seleccionadas (es decir, regiones que contienen polimorfismos específicos de alelo o alélicos) que son distintos del

sitio de mutación en los genes que codifican proteínas mutantes. Las metodologías de la presente divulgación son tratamientos eficaces para enfermedades de ganancia de función resultantes de mutaciones de delección, mutaciones de inserción, mutaciones puntuales y similares, a condición de que el gen mutante codifique una proteína que tiene una función normalmente no asociada a la proteína natural.

5 En un aspecto preferido, las metodologías de la presente divulgación proporcionan un tratamiento eficaz para la enfermedad de Huntington (EH). Las metodologías también proporcionan tratamientos eficaces para otros trastornos de poliglutamina y/o enfermedad de repetición de trinucleótidos, como se describe en detalle en el presente documento.

10 Por consiguiente, en un aspecto, la presente divulgación proporciona un método de tratamiento de un sujeto que tiene o está en riesgo de tener una enfermedad caracterizada o provocada por una proteína mutante de ganancia de función administrando al sujeto una cantidad eficaz de un agente de iARN que silencia un polimorfismo alélico dentro de un gen que codifica una proteína mutante, por ejemplo, la proteína huntingtina, de forma que la interferencia específica de secuencia de un gen ocurre dando como resultado un tratamiento eficaz para la enfermedad. En una realización, la proteína mutante contiene una región ampliada de poliglutamina. En otra realización, el gen que codifica la proteína mutante contiene una región ampliada de repeticiones de trinucleótidos.

15 En otra realización más, el método de la divulgación se puede usar para tratar enfermedad de Huntington y una variedad de otras enfermedades seleccionadas del grupo que consiste en ataxia espinocerebelosa de tipo 1, ataxia espinocerebelosa de tipo 2, ataxia espinocerebelosa de tipo 3, ataxia espinocerebelosa de tipo 6, ataxia espinocerebelosa de tipo 7, ataxia espinocerebelosa de tipo 8, ataxia espinocerebelosa de tipo 12, distrofia miotónica, enfermedad muscular espinobulbar y atrofia dentato-rubro-pálido-luisiana.

20 El método de la divulgación usa agentes de iARN homólogos a un polimorfismo alélico dentro del gen que codifica, por ejemplo, una proteína huntingtina mutante para el tratamiento de enfermedad de Huntington. En una realización preferida, el agente de iARN silencia el polimorfismo alélico seleccionado del grupo que consiste en P1-P5. En una realización preferida adicional, el agente de iARN silencia un polimorfismo alélico seleccionado del grupo que consiste en P6-P43.

25 En una realización adicional, la divulgación proporciona agentes de iARN que comprenden una primera y una segunda cadena que contienen cada una 16-25 nucleótidos. La primera cadena de la presente divulgación 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 iARN 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. En otras realizaciones más, la región diana de la secuencia de ARNm se localiza en una región no traducida (UTR) 5' o una UTR 3' del ARNm de una proteína mutante.

30 En otra realización, la divulgación proporciona una construcción de expresión que comprende un ácido nucleico aislado que codifica una molécula de ácido nucleico con una primera secuencia de 16-25 nucleótidos homóloga a un polimorfismo alélico dentro de, por ejemplo, el gen que codifica una proteína huntingtina mutante. La construcción de expresión puede ser, por ejemplo, un vector viral, vector retroviral, casete de expresión o plásmido. La construcción de expresión también puede tener una secuencia promotora de ARN polimerasa II o secuencia promotora de ARN polimerasa II, tal como el promotor de ARNnp U6 o promotor H1.

35 En aún otras realizaciones, la presente divulgación proporciona células hospedadoras, por ejemplo, células de mamífero, que comprenden moléculas de ácidos nucleicos y construcciones de expresión de la presente divulgación.

En otras realizaciones más, la presente divulgación proporciona composiciones terapéuticas que comprenden las moléculas de ácidos nucleicos de la divulgación y un vehículo farmacéuticamente aceptable.

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

Breve descripción de los dibujos

Figura 1a-k: Gen huntingtina humana, secuencia de nucleótidos (SEQ ID NO: 1)

Figura 2a-b: Proteína huntingtina humana, secuencia de aminoácidos (SEQ ID NO: 2)

50 Figura 3: Codificante (SEQ ID NO: 3) y no codificante (SEQ ID NO: 4) de la secuencia de ARN diana de huntingtina (htt)

Figura 4: Análisis termodinámico de los extremos 5' de la cadena de ARNip para el dúplex de ARNip

Figura 5a-c: Reacciones *in vitro* de iARN programadas con ARNip que silencia un polimorfismo dentro del ARNm de huntingtina (htt). (a) ARNip estándar. (b) ARNip mejorado por la reducción de la intensidad del apareamiento

de bases del extremo 5' de la cadena no codificante del dúplex de ARNip. (c) ARNip mejorado por la reducción del desapareamiento del extremo 5' de la cadena no codificante del dúplex de ARNip.

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

5 Descripción detallada de la divulgación

La presente divulgación se refiere a métodos y reactivos para tratar una variedad de enfermedades de ganancia de función. En un aspecto, la divulgación se refiere a métodos y reactivos para tratar una variedad de enfermedades caracterizadas por una mutación en un alelo o copia de un gen, codificando la mutación una proteína que es suficiente para contribuir a o provocar la enfermedad. Preferentemente, los métodos y reactivos se usan para tratar enfermedades causadas o caracterizadas por una mutación que se hereda en un modo dominante autosómico. En una realización, los métodos y reactivos se usan para tratar una variedad de enfermedades neurodegenerativas provocadas por una mutación de ganancia de función, por ejemplo, trastornos de poliglutamina y/o enfermedades de repetición de trinucleótidos, por ejemplo, enfermedad de Huntington. En otra realización, los métodos y reactivos se usan para tratar enfermedades provocadas por una ganancia de función en un oncogén, siendo el producto génico mutado una ganancia de función mutante, por ejemplo, cánceres provocados por una mutación en el oncogén *ret* (por ejemplo, *ret-1*), por ejemplo, tumores endocrinos, tumores medulares de tiroides, tumores de la hormona paratiroidea, neoplasia endocrina múltiple de tipo 2 y similares. En otra realización, los métodos y reactivos de la divulgación se pueden usar para tratar una variedad de cánceres gastrointestinales conocidos por ser provocados por mutaciones de ganancia de función autosómicamente heredadas.

La presente divulgación utiliza tecnología de interferencia por ARN (iARN) contra polimorfismos alélicos localizados dentro de un gen que codifica una proteína mutante de ganancia de función. La iARN destruye el ARNm mutante correspondiente con especificidad y selectividad por nucleótidos. Los agentes de ARN de la presente divulgación se dirigen a regiones polimórficas de un gen mutante, dando como resultado la escisión de ARNm mutante. Estos agentes de ARN, mediante una serie de interacciones proteína-nucleótido, funcionan escindiendo los ARNm mutantes. Las células destruyen el ARNm escindido, previniendo así la síntesis de proteína mutante correspondiente, por ejemplo, la proteína huntingtina.

Por consiguiente, en un aspecto, la presente divulgación proporciona un método de tratamiento de un sujeto que tiene o está en riesgo de tener una enfermedad caracterizada o provocada por una proteína mutante de ganancia de función administrando al sujeto una cantidad eficaz de un agente de iARN que silencia un polimorfismo alélico dentro de un gen que codifica una proteína mutante, por ejemplo, la proteína huntingtina, de forma que la interferencia específica de secuencia de un gen ocurra dando como resultado un tratamiento eficaz para la enfermedad. En una realización, la proteína mutante contiene una región ampliada de poliglutamina. En otra realización, el gen que codifica la proteína mutante contiene una región ampliada de repeticiones de trinucleótidos.

En otra realización más, el método de la divulgación se puede usar para tratar enfermedad de Huntington y una variedad de otras enfermedades seleccionadas del grupo que consiste en ataxia espinocerebelosa de tipo 1, ataxia espinocerebelosa de tipo 2, ataxia espinocerebelosa de tipo 3, ataxia espinocerebelosa de tipo 6, ataxia espinocerebelosa de tipo 7, ataxia espinocerebelosa de tipo 8, ataxia espinocerebelosa de tipo 12, distrofia miotónica, enfermedad muscular espinobulbar y atrofia dentato-rubro-pálido-luisiana.

El método de la divulgación usa agentes de iARN homólogos a un polimorfismo alélico dentro del gen que codifica, por ejemplo, una proteína huntingtina mutante para el tratamiento de enfermedad de Huntington. En una realización preferida, el agente de iARN silencia un polimorfismo alélico seleccionado del grupo que consiste en P1-P5. En una realización preferida adicional, el agente de iARN silencia un polimorfismo alélico seleccionado del grupo que consiste en P6-P43.

En una realización adicional, la divulgación proporciona agentes de iARN que comprenden una primera y una segunda cadena que contiene cada una 16-25 nucleótidos. La primera cadena de la presente divulgación 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 iARN 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 secuencias de nucleótidos. En otras realizaciones más, la región diana de la secuencia de ARNm se localiza en una región no traducida (UTR) 5' o una UTR 3' del ARNm de una proteína mutante.

En otra realización, la divulgación proporciona una construcción de expresión que comprende un ácido nucleico aislado que codifica una molécula de ácido nucleico con una primera secuencia de 16-25 nucleótidos homóloga a un polimorfismo alélico dentro de, por ejemplo, el gen que codifica una proteína huntingtina mutante. La construcción de expresión puede ser, por ejemplo, un vector viral, vector retroviral, casete de expresión o plásmido. La construcción de expresión también puede tener una secuencia promotora de ARN polimerasa II o secuencia promotora de ARN polimerasa II, tal como el promotor de ARNnp U6 o promotor H1.

En aún otras realizaciones, la presente divulgación proporciona células hospedadoras, por ejemplo, células de mamífero, que comprenden moléculas de ácidos nucleicos y construcciones de expresión de la presente divulgación.

En otras realizaciones más, la presente divulgación proporciona composiciones terapéuticas que comprenden las moléculas de ácidos nucleicos de la divulgación y un vehículo farmacéuticamente aceptable.

- 5 De manera que la divulgación pueda ser más fácilmente entendida, se definen primero ciertos términos.

El término "nucleósido" se refiere a una molécula que tiene una base de purina o pirimidina ligada covalentemente a un azúcar de ribosa o desoxirribosa. Los nucleósidos a modo de ejemplo incluyen adenosina, guanosina, citidina, uridina y timidina. Los nucleósidos a modo de ejemplo adicionales incluyen inosina, 1-metilinosina, pseudouridina, 5,6-dihidouridina, ribotimidina, ²N-metilguanosina y ^{2,2}NN-dimetilguanosina (también denominados nucleósidos "raros").

- 10 El término "nucleótido" se refiere a un nucleósido que tiene uno o más grupos fosfato unidos con enlaces éster al resto de azúcar. Los nucleótidos a modo de ejemplo incluyen monofosfatos, difosfatos y trifosfatos de nucleósido. Los términos "polinucleótido" y "molécula de ácido nucleico" se usan indistintamente en el presente documento y se refieren a un polímero de nucleótidos unidos juntos por un enlace fosfodiéster entre los átomos de carbono 5' y 3'.

El término "ARN" o "molécula de ARN" o "molécula de ácido ribonucleico" se refiere a un polímero de ribonucleótidos.

- 15 El término "ADN" o "molécula de ADN" o "molécula de ácido desoxirribonucleico" se refiere a un polímero de desoxirribonucleótidos. Se pueden sintetizar ADN y ARN naturalmente (por ejemplo, por replicación de ADN o transcripción de ADN, respectivamente). El ARN se puede modificar postranscripcionalmente. El ADN y el ARN también se pueden sintetizar químicamente. El ADN y el ARN pueden ser monocatenarios (es decir, ARNmc y ADNmc, respectivamente) o de múltiples cadenas (por ejemplo, bicatenario, es decir, ARNbc y ADNbc, respectivamente).
20 "ARNm" o "ARN mensajero" es ARN monocatenario que especifica la secuencia de aminoácidos de una o más cadenas de polipéptidos. Esta información se traduce durante la síntesis de proteínas cuando los ribosomas se unen al ARNm.

Como se usa en el presente documento, el término "ARN interferente pequeño" ("ARNip") (también denominado en la técnica "ARN interferentes pequeños") se refiere a un ARN (o análogo de ARN) que comprende entre
25 aproximadamente 10-50 nucleótidos (o análogos de nucleótidos) que es capaz de dirigir o mediar en la interferencia por ARN. Preferentemente, un ARNip comprende entre aproximadamente 15-30 nucleótidos o análogos de nucleótidos, más preferentemente entre aproximadamente 16-25 nucleótidos (o análogos de nucleótidos), incluso más preferentemente entre aproximadamente 18-23 nucleótidos (o análogos de nucleótidos), e incluso más preferentemente entre aproximadamente 19-22 nucleótidos (o análogos de nucleótidos) (por ejemplo, 19, 20, 21 o 22
30 nucleótidos o análogos de nucleótidos). El término ARNip "corto" se refiere a un ARNip que comprende ~21 nucleótidos (o análogos de nucleótidos), por ejemplo, 19, 20, 21 o 22 nucleótidos. El término ARNip "largo" se refiere a un ARNip que comprende ~24-25 nucleótidos, por ejemplo, 23, 24, 25 o 26 nucleótidos. Los ARNip cortos pueden incluir, en algunos casos, menos de 19 nucleótidos, por ejemplo, 16, 17 o 18 nucleótidos, a condición de que el ARNip más corto retenga la capacidad para mediar en iARN. Asimismo, los ARNip largos pueden incluir, en algunos casos, más de 26
35 nucleótidos, a condición de que el ARNip más largo retenga la capacidad para mediar en el procesamiento adicional ausente de iARN, por ejemplo, procesamiento enzimático, a un ARNip corto.

El término "análogo de nucleótidos" o "nucleótido alterado" o "nucleótido modificado" se refiere a un nucleótido no estándar, que incluye ribonucleótidos o desoxirribonucleótidos que no existen de forma natural. Los análogos
40 preferidos de nucleótidos se modifican en cualquier posición para alterar ciertas propiedades químicas del nucleótido, aún retienen la capacidad del análogo de nucleótidos para realizar su función prevista. Los ejemplos de posiciones del nucleótido que pueden ser derivatizadas incluyen la posición 5, por ejemplo, 5-(2-amino)propiluridina, 5-bromouridina, 5-propinouridina, 5-propeniluridina, etc.; la posición 6, por ejemplo, 6-(2-amino)propiluridina; la posición 8 para adenosina y/o guanosinas, por ejemplo, 8-bromoguanosina, 8-cloroguanosina, 8-fluoroguanosina, etc. Los análogos de nucleótidos también incluyen nucleótidos deaza, por ejemplo, nucleótidos 7-deaza-adenosina; modificados en O y
45 N (por ejemplo, alquilados, por ejemplo, N6-metiladenosina, o como se conoce de otro modo en la técnica); y otros análogos de nucleótidos heterocíclicamente modificados tales como los descritos en Herdewijn, Antisense Nucleic Acid Drug Dev., 2000 Aug. 10(4):297-310.

Los análogos de nucleótidos también pueden comprender modificaciones en la porción de azúcar de los nucleótidos. Por ejemplo, el grupo 2' OH se puede sustituir por un grupo seleccionado de H, OR, R, F, Cl, Br, I, SH, SR, NH₂, NHR,
50 NR₂, COOR u OR, en donde R es alquilo C₁-C₆ sustituido o sin sustituir, alqueno, alquino, arilo, etc. Otras posibles modificaciones incluyen las descritas en las patentes de EE. UU. N° 5.858.988 y 6.291.438.

El grupo fosfato del nucleótido también se puede modificar, por ejemplo, sustituyendo uno o más de los oxígenos del grupo fosfato con azufre (por ejemplo, fosforotioatos), o haciendo otras sustituciones que permitan al nucleótido
55 realizar su función prevista, tal como se describe en, por ejemplo, Eckstein, Antisense Nucleic Acid Drug Dev. 2000 Apr. 10(2):117-21, Ruszkowski et al. Antisense Nucleic Acid Drug Dev. 2000 Oct. 10(5):333-45, Stein, Antisense Nucleic Acid Drug Dev. 2001 Oct. 11(5): 317-25, Vorobjev et al. Antisense Nucleic Acid Drug Dev. 2001 Apr. 11(2):77-85, y la patente de EE. UU. N° 5.684.143. Ciertas de las modificaciones anteriormente referenciadas (por ejemplo, modificaciones de grupo fosfato) disminuyen preferentemente la tasa de hidrólisis de, por ejemplo, polinucleótidos que comprenden dichos análogos *in vivo* o *in vitro*.

El término "oligonucleótido" se refiere a un polímero corto de nucleótidos y/o análogos de nucleótidos. El término "análogo de ARN" se refiere a un polinucleótido (por ejemplo, un polinucleótido químicamente sintetizado) que tiene al menos un nucleótido alterado o modificado en comparación con un ARN correspondiente sin alterar o sin modificar, pero que retiene la misma naturaleza o función o similar que el ARN sin alterar o sin modificar correspondiente. Como se trata anteriormente, los oligonucleótidos se pueden unir con enlaces que dan como resultado una tasa de hidrólisis más baja del análogo de ARN en comparación con una molécula de ARN con enlaces fosfodiéster. Por ejemplo, los nucleótidos del análogo pueden comprender enlaces metilenodiol, etilenodiol, oximetiltio, oxietiltio, oxicarbonilo, fosforodiamidato, fosforoamidato y/o fosforotioato. Los análogos de ARN preferidos incluyen ribonucleótidos y/o desoxirribonucleótidos modificados en el azúcar y/o esqueleto. Dichas alteraciones o modificaciones pueden incluir además la adición de material no de nucleótido, tal como al (a los) extremo(s) del ARN o internamente (en uno o más nucleótidos del ARN). Un análogo de ARN solo necesita ser suficientemente similar al ARN natural que tiene la capacidad para mediar (media) en la interferencia por ARN.

Como se usa en el presente documento, el término "interferencia por ARN" ("iARN") se refiere a una degradación intracelular selectiva de ARN. La iARN ocurre en las células naturalmente para retirar ARN extraños (por ejemplo, ARN virales). La iARN natural avanza por fragmentos escindidos del ARNbc libre que dirigen el mecanismo degradativo a otras secuencias de ARN similares. Alternativamente, la iARN se puede iniciar por la mano del hombre, por ejemplo, para silenciar la expresión de genes diana.

Un agente de iARN que tiene una cadena que tiene "secuencia suficientemente complementaria a una secuencia de ARNm diana para dirigir la interferencia por ARN (iARN) específica de diana" significa que la cadena tiene una secuencia suficiente para desencadenar la destrucción del ARNm diana por la maquinaria o proceso de iARN.

Como se usa en el presente documento, el término "ARN aislado" (por ejemplo, "ARNip aislado" o "precursor de ARNip 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 sintetizan químicamente.

El término "*in vitro*" tiene su significado reconocido en la técnica, por ejemplo, que implica reactivos o extractos purificados, por ejemplo, extractos de células. El término "*in vivo*" también tiene su significado reconocido en la técnica, por ejemplo, que implica células vivas, por ejemplo, células inmortalizadas, células primarias, líneas celulares y/o células en un organismo.

Como se usa en el presente documento, el término "transgén" se refiere a cualquier molécula de ácido nucleico que se inserta por arteificio en una célula, y llega a ser parte del genoma del organismo que se desarrolla de la célula. Dicho transgén puede incluir un gen que es parcial o completamente heterólogo (es decir, extraño) al 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 ácidos nucleicos seleccionadas, por ejemplo, ADN, que codifican uno o más precursores de ARN manipulados, para expresión en un organismo transgénico, por ejemplo, animal, que es parcialmente o completamente heterólogo, es decir, extraño, al animal transgénico, u homólogo a un gen endógeno del animal transgénico, pero que se diseña para ser insertado en el genoma del animal en una localización que se diferencia de la del gen natural. Un transgén incluye uno o más promotores y cualquier otro ADN, tales como intrones, necesarios para la expresión de la secuencia de ácidos nucleicos seleccionada, todos operativamente unidos a la secuencia seleccionada, y puede incluir una secuencia potenciadora.

Un gen "implicado" en una enfermedad o trastorno incluye un gen, cuya expresión o función normal o aberrante afecta o provoca la enfermedad o trastorno o al menos un síntoma de dicha enfermedad o trastorno

El término "mutación de ganancia de función", como se usa en el presente documento, se refiere a cualquier mutación en un gen en que la proteína codificada por dicho gen (es decir, la proteína mutante) adquiere una función normalmente no asociada a la proteína (es decir, la proteína natural) que provoca o contribuye a una enfermedad o trastorno. La mutación de ganancia de función puede ser una delección, adición o sustitución de un nucleótido o nucleótidos en el gen que da lugar al cambio en la función de la proteína codificada. En una realización, la mutación de ganancia de función cambia la función de la proteína mutante o provoca interacciones con otras proteínas. En otra realización, la mutación de ganancia de función provoca una disminución o retirada de la proteína natural normal, por ejemplo, por interacción de la proteína mutante alterada con dicha proteína natural normal.

El término "polimorfismo", como se usa en el presente documento, se refiere a una variación (por ejemplo, una delección, inserción o sustitución) en una secuencia de gen que se identifica o detecta cuando se compara la misma secuencia génica de sujetos de fuentes diferentes (pero del mismo organismo). Por ejemplo, se puede identificar un polimorfismo cuando se compara la misma secuencia de genes de diferentes sujetos (pero del mismo organismo). La identificación de dichos polimorfismos es rutinaria en la técnica, siendo las metodologías similares a las usadas para detectar, por ejemplo, mutaciones puntuales de cáncer de mama. La identificación se puede hacer, por ejemplo, de ADN extraído de linfocitos de un sujeto, seguido por amplificación de regiones polimórficas usando cebadores específicos para dicha región polimórfica. Alternativamente, el polimorfismo se puede identificar cuando se comparan dos alelos del mismo gen. Una variación en secuencia entre dos alelos del mismo gen dentro de un organismo se denomina en el presente documento un "polimorfismo alélico". El polimorfismo puede ser un nucleótido dentro de una

región codificante, pero, debido a la degeneración del código genético, no codifica 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 regiones no codificantes del gen.

- 5 El término "dominio de poliglutamina", como se usa en el presente documento, se refiere a un segmento o dominio de una proteína que consiste en restos consecutivos de glutamina unidos por enlaces peptídicos. En una realización, la región consecutiva incluye al menos 5 restos de glutamina.

- 10 El término "dominio ampliado de poliglutamina" o "segmento ampliado de poliglutamina", como se usa en el presente documento, se refiere a un segmento o dominio de una proteína que incluye al menos 35 restos consecutivos de glutamina unidos por enlaces peptídicos. Dichos segmentos ampliados se encuentran en sujetos afectados por un trastorno por poliglutaminas, como se describe en el presente documento, tanto si se ha mostrado como si no que el sujeto manifiesta síntomas.

- 15 El término "repetición de trinucleótidos" o "región de repetición de trinucleótidos", como se usa en el presente documento, se refiere a un segmento de una secuencia de ácidos nucleicos, por ejemplo, que consiste en repeticiones consecutivas de una secuencia de trinucleótidos particular. En una realización, la repetición de trinucleótidos incluye al menos 5 secuencias de trinucleótidos consecutivas. Las secuencias de trinucleótidos a modo de ejemplo incluyen, pero no se limitan a, CAG, CGG, GCC, GAA, CTG y/o CGG.

- 20 El término "enfermedades de repetición de trinucleótidos", como se usa en el presente documento, se refiere a cualquier enfermedad o trastorno caracterizado por una región ampliada de repeticiones de trinucleótidos localizada dentro de un gen, siendo la región ampliada de repeticiones de trinucleótidos la causante de la enfermedad o trastorno. Los ejemplos de enfermedades de repetición de trinucleótidos incluyen, pero no se limitan a, ataxia espinocerebelosa de tipo 12, ataxia espinocerebelosa de tipo 8, síndrome del cromosoma X frágil, retraso mental del cromosoma XE frágil, ataxia de Friedreich y distrofia miotónica. Las enfermedades de repetición de trinucleótidos preferidas para el tratamiento según la presente divulgación son las caracterizadas o provocadas por una región ampliada de repeticiones de trinucleótidos en el extremo 5' de la región codificante de un gen, el gen que codifica una proteína mutante que provoca o es causante de la enfermedad o trastorno. Ciertas enfermedades de trinucleótidos, por ejemplo, síndrome del cromosoma X frágil, donde la mutación no está asociada a una región codificante, pueden no ser adecuadas para el tratamiento según las metodologías de la presente divulgación, no hay ARNm adecuado para ser silenciado por iARN. Por el contrario, la enfermedad tal como ataxia de Friedreich puede ser adecuada para el tratamiento según las metodologías de la divulgación debido a que, aunque la mutación causante no está dentro de una región codificante (es decir, se encuentra dentro de un intrón), la mutación puede estar dentro de, por ejemplo, un precursor de ARNm (por ejemplo, un precursor de ARNm previamente cortado y empalmado).

- 35 El término "trastorno de las poliglutaminas", como se usa en el presente documento, se refiere a cualquier enfermedad o trastorno caracterizado por una expansión de repeticiones de (CAG)_n en el extremo 5' de la región codificante (codificando así una región ampliada de poliglutamina en la proteína codificada). En una 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: enfermedad de Huntington, ataxia espinocerebelosa de tipo 1, ataxia espinocerebelosa de tipo 2, ataxia espinocerebelosa de tipo 3 (también conocida como enfermedad de Machado-Joseph) y ataxia espinocerebelosa de tipo 6, ataxia espinocerebelosa de tipo 7 y atrofia dentato-rubro-pálido-luisiana.

- 40 La expresión "examinar la función de un gen en una célula u organismo" se refiere a examinar o estudiar la expresión, actividad, función o fenotipo que surge del mismo.

- Diversas metodologías de la presente divulgación incluyen la etapa que implica comparar un valor, nivel, rasgo, característica, propiedad, etc., con un "control adecuado", denominado indistintamente en el presente documento "control apropiado". Un "control adecuado" o "control apropiado" es cualquier control o patrón conocido por un experto habitual en la técnica útil para fines de comparación. En una realización, un "control adecuado" o "control apropiado" es un valor, nivel, rasgo, característica, propiedad, etc., determinado antes de la realización de una metodología de iARN, como se describe en el presente documento. Por ejemplo, se pueden determinar una tasa de transcripción, nivel de ARNm, tasa de traducción, nivel de proteína, actividad biológica, característica o propiedad celular, genotipo, fenotipo, etc., antes de introducir un agente de iARN de la divulgación en una célula u organismo. En otra realización, un "control adecuado" o "control apropiado" es un valor, nivel, rasgo, característica, propiedad, etc., determinado en una célula u organismo, por ejemplo, un control o célula normal u organismo, que presenta, por ejemplo, rasgos normales. En otra realización más, un "control adecuado" o "control apropiado" es un valor predefinido, nivel, rasgo, característica, propiedad, etc.

Diversos aspectos de la divulgación se describen con más detalle en las siguientes subsecciones.

55 I. Trastornos de poliglutamina

Los trastornos de poliglutamina son una clase de enfermedad o trastornos caracterizados por una mutación genética común. En particular, la enfermedad o los trastornos se caracterizan por una repetición ampliada del trinucleótido CAG que da lugar, en la proteína codificada, a un tramo ampliado de restos de glutamina. Los trastornos de poliglutamina

son similares en que las enfermedades se caracterizan por una degeneración progresiva de células nerviosas. A pesar de sus similitudes, los trastornos de poliglutamina ocurren en diferentes cromosomas y así ocurren en segmentos de ADN completamente diferentes. Los ejemplos de trastornos de poliglutamina incluyen enfermedad de Huntington, atrofia dentato-rubro-pálido-luisiana, atrofia muscular espinobulbar, ataxia espinocerebelosa de tipo 1, ataxia espinocerebelosa de tipo 2, ataxia espinocerebelosa de tipo 3, ataxia espinocerebelosa de tipo 6 y ataxia espinocerebelosa de tipo 7 (Tabla 3).

Tabla 1. Trastornos de poliglutamina

Enfermedad	Gen	Locus	Proteína	Tamaño de repetición de CAG	
				Normal	Enfermedad
Atrofia muscular espinobulbar (enfermedad de Kennedy)	<i>AR</i>	Xq13-21	Receptor de andrógenos (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	Atropina-1	6-35	49-88
Ataxia espinocerebelosa de tipo 1	<i>SCA1</i>	6p23	Ataxina-1	6-44 ^a	39-82
Ataxia espinocerebelosa de tipo 2	<i>SCA2</i>	12q24.1	Ataxina-2	15-31	36-63
Ataxia espinocerebelosa de tipo 3 (enfermedad de Machado-Joseph)	<i>SCA3 (MJD1)</i>	14q32.1	Ataxina-3	12-40	55-84
Ataxia espinocerebelosa de tipo 6	<i>SCA6</i>	19p13	Subunidad α_{1A} de los canales de calcio dependientes de la tensión	4-18	21-33
Ataxia espinocerebelosa de tipo 7	<i>SCA7</i>	13p12-13	Ataxina-7	4-35	37-306
^a Alelos con 21 o más repeticiones se interrumpen por 1-3 unidades de CAT; los alelos de enfermedad contienen tramos de CAG puros.					

Los trastornos de poliglutamina de la divulgación se caracterizan por (por ejemplo, dominios que tienen entre aproximadamente 30 y 35 restos de glutamina, entre aproximadamente 35 y 40 restos de glutamina, entre aproximadamente 40 y 45 restos de glutamina y que tienen aproximadamente 45 o más restos de glutamina. El dominio de poliglutamina contiene normalmente restos de glutamina consecutivos (Q n>36).

II. Enfermedad de Huntington

La enfermedad de Huntington, heredada como una enfermedad dominante autosómica, provoca alteración de la cognición y enfermedad motora. Los pacientes pueden vivir más de una década con una intensa debilitación, antes de la muerte prematura por inanición o infección. La enfermedad empieza en la cuarta o quinta década en la mayoría de los casos, pero un subconjunto de pacientes manifiesta la enfermedad en la adolescencia. La mutación genética para la enfermedad de Huntington es una repetición de CAG ampliada en el gen huntingtina. La repetición de CAG varía en número desde 8 hasta 35 en individuos normales (Kremer et al., 1994). La mutación genética, por ejemplo, un aumento en la longitud de las repeticiones de CAG desde normal inferior a 36 en el gen huntingtina hasta más de 36 en la enfermedad, está asociada 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 enfermedad de Huntington. Prototípico de nada menos que veinte otras enfermedades con una CAG ampliada como mutación subyacente, la enfermedad de Huntington todavía no tiene terapia eficaz. Se han mostrado prometedoras una variedad de intervenciones -- tales como interrupción de las vías apoptóticas, adición de reactivos para reforzar la eficiencia de las mitocondrias y bloqueo de receptores de NMDA -- en cultivos celulares y modelo de ratón de enfermedad de Huntington. Sin embargo, en el mejor de los casos, estos enfoques revelan una prolongación corta de la supervivencia celular o animal.

La enfermedad de Huntington obedece el dogma principal de la genética: un gen mutante sirve de molde para la producción de un ARNm mutante; el ARNm mutante dirige entonces la síntesis de una proteína mutante (Aronin et al., 1995; DiFiglia et al., 1997). La huntingtina mutante (proteína) se acumula probablemente en neuronas selectivas en el cuerpo estriado y la corteza, altera actividades celulares determinadas hasta ahora y provoca disfunción neuronal y muerte (Aronin et al., 1999; Laforet et al., 2001). Debido a que una única copia de un gen mutante es suficiente para provocar la enfermedad de Huntington, el tratamiento más parco sería volver ineficaz el gen mutante. Los enfoques teóricos podrían incluir detener la transcripción génica de la huntingtina mutante, destruyendo el ARNm mutante y bloqueando la traducción. Cada uno tiene el mismo resultado -- pérdida de huntingtina mutante.

III. Gen huntingtina

El gen de enfermedad ligado a la enfermedad de Huntington se llama Huntington o (htt). El sitio huntingtina es grande, que abarca 180 kb y que consiste en 67 exones. El gen huntingtina se expresa ampliamente y se requiere para el desarrollo normal. Se expresa como 2 formas alternativamente poliadeniladas que presentan diferente abundancia relativa en diversos tejidos fetales y adultos. El transcrito más grande tiene aproximadamente 13,7 kb y se expresa predominantemente en el cerebro adulto y fetal mientras que el transcrito más pequeño de aproximadamente 10,3 kb se expresa más ampliamente. Los dos transcritos se diferencian con respecto a sus regiones no traducidas 3' (Lin et al., 1993). Se predice que ambos mensajeros codifican una proteína de 348 kilodálton que contiene 3144 aminoácidos. Se cree que el defecto genético que conduce a la enfermedad de Huntington confiere una nueva propiedad al 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).

En la Figura 1 (SEQ ID NO: 1) se expone una secuencia de nucleótidos consenso del gen huntingtina humano (ADNc). La región codificante consiste en los nucleótidos 316 a 9750 de SEQ ID NO: 1. Las dos señales de poliadenilación alternativas se encuentran en los nucleótidos 10326 a 10331 y los nucleótidos 13644 a 13649, respectivamente. Los dos sitios de poliadenilación correspondientes se encuentran en los nucleótidos 10348 y 13672, respectivamente. La primera señal de poliadenilación/sitio es la del transcrito de 10,3 kb. La segunda señal de poliadenilación/sitio es la del transcrito de 13,7 kb, el transcrito 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 huntingtina por análisis de SNP (polimorfismo de un solo nucleótido) (véase la Tabla 3). Los polimorfismos expuestos en las Tablas 2 y 3 representan sitios preferidos para silenciar mediante iARN específica de un único nucleótido, como se describe en el presente 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 (construcción 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 identificado por 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

	consenso	polimorfismo	db xref
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 divulgación silencia huntingtina mutante usando interferencia por ARN (Hutvagner et al., 2002). Una cadena de ARN bicatenario (ARNip) complementa una región polimórfica dentro del ARNm de huntingtina mutante. Después de la introducción de ARNip en neuronas, el ARNip se desenrolla parcialmente, se une a una región polimórfica dentro del ARNm de huntingtina en un modo específico de sitio, y activa una ARNm nucleasa. Esta nucleasa escinde el ARNm de huntingtina, deteniendo así la traducción de la huntingtina mutante. Las células se libran ellas mismas del ARNm parcialmente digerido, impidiendo así la traducción, o las células digieren proteínas

parcialmente traducidas. Las neuronas sobreviven sobre la huntingtina no mutante (del alelo normal); este enfoque previene los deterioros de huntingtina mutante eliminando su producción.

IV. Diseño de ARNip

Se diseñan ARNip del siguiente modo. Primero, se selecciona una porción del gen diana (por ejemplo, el gen htt) que incluye el polimorfismo. Los polimorfismos a modo de ejemplo se seleccionan de la región no traducida 5' de un gen diana. La escisión de ARNm en estos sitios debe eliminar la traducción de proteína mutante correspondiente. Los polimorfismos de otras regiones del gen mutante también son adecuados para su direccionamiento. Se diseña una cadena codificante basándose en la secuencia de la porción seleccionada. Preferentemente, la porción (y la cadena codificante correspondiente) incluye aproximadamente 19 a 25 nucleótidos, por ejemplo, 19, 20, 21, 22, 23, 24 o 25 nucleótidos. Más preferentemente, la porción (y cadena codificante correspondiente) incluye 21, 22 o 23 nucleótidos. El experto apreciará, sin embargo, que los ARNip que tienen una longitud inferior a 19 nucleótidos o superior a 25 nucleótidos también pueden funcionar para mediar en la iARN. Por consiguiente, los ARNip de dicha longitud también están dentro del alcance de la presente divulgación, a condición de que retengan la capacidad de mediar en iARN. Se ha demostrado que agentes de iARN más largos provocan una respuesta de interferón o PKR en ciertas células de mamífero que pueden ser no deseables. Preferentemente, los agentes de iARN de la divulgación no provocan una respuesta de PKR (es decir, son de una longitud suficientemente corta). Sin embargo, agentes de iARN más largos pueden ser útiles, por ejemplo, en tipos de células incapaces de generar una respuesta de PKR o en situaciones donde la respuesta de PKR se ha regulado por disminución o disminuido por medios alternativos.

La secuencia de la cadena codificante se diseña de forma que el polimorfismo esté esencialmente en el centro de la cadena. Por ejemplo, si se elige un ARNip de 21 nucleótidos, el polimorfismo está en, por ejemplo, el nucleótido 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 desde el extremo 5' de la cadena codificante. Para un ARNip de 22 nucleótidos, el polimorfismo está en, por ejemplo, el nucleótido 7, 8, 9, 10, 11, 12, 13, 14, 15 o 16. Para un ARNip de 23 nucleótidos, el polimorfismo está en, por ejemplo, 7, 8, 9, 10, 11, 12, 13, 14, 15 o 16. Para un ARNip de 24 nucleótidos, el polimorfismo está en, por ejemplo, 9, 10, 11, 12, 13, 14 o 16. Para un ARNip de 25 nucleótidos, el polimorfismo está en, por ejemplo, 9, 10, 11, 12, 13, 14, 15, 16 o 17. El mover el polimorfismo hacia una posición descentrada puede reducir, en algunos casos, la eficiencia de escisión por el ARNip. Dichas composiciones, es decir, composiciones menos eficientes, pueden ser deseables para su uso si se detecta el silenciamiento del ARNm no mutante.

La cadena no codificante es rutinariamente de la misma longitud que la cadena codificante e incluye nucleótidos complementarios. En una realización, las cadenas son completamente complementarias, es decir, las cadenas son de extremos romos cuando se alinean o hibridan. En otra realización, las cadenas comprenden alinear o hibridar tal que se generen nucleótidos protuberantes de 1, 2 o 3 nucleótidos, es decir, el extremo 3' de la cadena codificante se extiende 1, 2 o 3 nucleótidos más que el extremo 5' de la cadena no codificante y/o el extremo 3' de la cadena no codificante se extiende 1, 2 o 3 nucleótidos más que el extremo 5' de la cadena codificante. Los nucleótidos protuberantes pueden comprender (o consistir en) nucleótidos correspondientes a la secuencia del gen diana (o complemento de la misma). Alternativamente, los nucleótidos protuberantes pueden comprender (o consistir en) desoxirribonucleótidos, por ejemplo, dTs, o análogos de nucleótidos, u otro material no de nucleótido adecuado.

Para facilitar la entrada de la cadena no codificante en RISC (y así aumentar o mejorar la eficiencia de la escisión y el silenciamiento diana), se puede alterar la intensidad del par de bases entre el extremo 5' de la cadena codificante y el extremo 3' de la cadena no codificante, por ejemplo, disminuir o reducir, como se describe en detalle en la solicitud de patente provisional de EE. UU. N° 60/475.386 titulada "*Methods and Compositions for Controlling Efficacy of RNA Silencing*" (presentada el 2 de junio de 2003) y 60/475.331 titulada "*Methods and Compositions for Enhancing the Efficacy and Specificity of RNAi*" (presentada el 2 de junio de 2003).

En una realización de estos aspectos de la divulgación, la intensidad del par de bases es menos debido a pares de bases G:C menores entre el extremo 5' de la primera cadena o no codificante y el extremo 3' de la segunda cadena o codificante que entre el extremo 3' de la primera cadena o no codificante y el extremo 5' de la segunda cadena o codificante. En otra realización, la intensidad del par de bases es menos debido a al menos un par de bases desapareado entre el extremo 5' de la primera cadena o no codificante y el extremo 3' de la segunda cadena o codificante. Preferentemente, el par de bases desapareado se selecciona del grupo que consiste en G:A, C:A, C:U, G:G, A:A, C:C y U:U. En otra realización, la intensidad del par de bases es menos debido a al menos un par de bases de titubeo, por ejemplo, G:U, entre el extremo 5' de la primera cadena o no codificante y el extremo 3' de la segunda cadena o codificante. En otra realización, la intensidad del par de bases es menos debido a al menos un par de bases que comprende un nucleótido raro, por ejemplo, inosina (I). Preferentemente, el par de bases se selecciona del grupo que consiste en I:A, I:U y I:C. En otra realización más, la intensidad del par de bases es menos debido a al menos un par de bases que comprende un nucleótido modificado. En realizaciones preferidas, el nucleótido modificado se selecciona del grupo que consiste en 2-amino-G, 2-amino-A, 2,6-diamino-G y 2,6-diamino-A.

El diseño de ARNip adecuados para silenciar los polimorfismos de htt expuestos en la Tabla 2 se describe con detalle a continuación

ADN de P1	TGTGCTGACTC T GAGGAACAG	(SEQ ID NO: 5)	
codificante	UGUGCUGACUC U GAGGAACAG	(SEQ ID NO: 6)	
no codificante	ACACGACUGAGACUCCUUGUC	(extremos romos, 21-mero)	(SEQ ID NO: 7)
(2 nucleótidos protuberantes) véase la Figura 5			
ADN de P2	CATACCTCA A ACTGCATGATG	(SEQ ID NO: 8)	
codificante	CAUACCUCA A ACUGCAUGAUG	(SEQ ID NO: 9)	
no codificante	GUAUGGAGUUUGACGUACUAC	(extremos romos, 21-mero)	(SEQ ID NO: 10)
ADN de P3	GCCTGCAGA G CCGGCGGCCTA	(SEQ ID NO: 11)	
codificante	GCCUGCAGA G CCGGCGGCCUA	(SEQ ID NO: 12)	
no codificante	CGGACGUCUCGGCCGCCGAU	(extremos romos, 21-mero)	(SEQ ID NO: 13)
ADN de P4	ACAGAGTTT G TGACCCACGCC	(SEQ ID NO: 14)	
codificante	ACAGAGUUU G UGACCCACGCC	(SEQ ID NO: 15)	
no codificante	UGUCUCA A ACACUGGGUGCGG	(extremos romos, 21-mero)	(SEQ ID NO: 16)
ADN de P5	TCCCTCATC T ACTGTGTGCAC	(SEQ ID NO: 17)	
codificante	UCCCUCAUC U ACUGUGUGCAC	(SEQ ID NO: 18)	
no codificante	AGGGAGUAGAUGACACACGUG	(extremos romos, 21-mero)	(SEQ ID NO: 19)

Se pueden diseñar ARNip según las enseñanzas anteriores a modo de ejemplo para cualquier otro polimorfo encontrado en el gen htt. Además, la tecnología es aplicable a silenciar cualquier otro gen de enfermedad que tenga asociados polimorfismos, es decir, polimorfismos que no causan enfermedad.

5 Para validar la eficacia por la que los ARNip destruyen ARNm mutantes (por ejemplo, ARNm de huntingtina mutante), el ARNip se incuba con ADNc mutante (por ejemplo, ADNc de huntingtina mutante) en un sistema de expresión de ARNm *in vitro* basado en *Drosophila*. Radiomarcados con ³²P, los ARNm mutantes recién sintetizados (por ejemplo, ARNm de huntingtina mutante) se detectan autorradiográficamente sobre un gel de agarosa. La presencia de ARNm mutante escindido indica actividad de ARNm nucleasa. Los controles adecuados incluyen la omisión de ARNip y el uso de ADNc de huntingtina no mutante. Alternativamente, se seleccionan ARNip de control que tienen la misma composición de nucleótidos que el ARNip seleccionado, pero sin complementariedad de secuencias significativa con el gen diana apropiado. Dichos controles negativos se pueden diseñar reordenando aleatoriamente la secuencia de nucleótidos del ARNip seleccionado; se puede realizar una búsqueda de homología para garantizar que el control negativo carezca de homología con cualquier otro gen en el genoma apropiado. Además, se pueden diseñar ARNip de control negativo introduciendo uno o más desapareamientos de bases en la secuencia.

15 Se seleccionan sitios de complementación ARNip-ARNm que dan como resultado la especificidad óptima de ARNm y la máxima escisión de ARNm.

Mientras que la presente divulgación caracteriza principalmente el silenciamiento de regiones polimórficas en el gen diana mutante (por ejemplo, en htt mutante) distinto de la mutación de la región ampliada de CAG, el experto apreciará que silenciar la región mutante puede tener aplicabilidad como estrategia terapéutica en ciertas situaciones. El silenciamiento de la región mutante se puede llevar a cabo usando ARNip que complementa CAG en serie. El ARNip^{cag} se uniría a ARNm con complementación de CAG, pero cabría esperar tener mayor oportunidad de unirse a una serie ampliada de CAG. Se unirían múltiples ARNip^{cag} al ARNm de huntingtina mutante (a diferencia de menos para el ARNm de huntingtina no mutante); así, es más probable que se escinda el ARNm de huntingtina mutante. La satisfactoria inactivación de ARNm usando este enfoque también eliminaría ARNm de huntingtina normal o no mutante. También inactivados, al menos de algún modo, podrían estar otros genes normales (aproximadamente 70) que también tienen repeticiones de CAG, donde sus ARNm podrían interactuar con el ARNip. Este enfoque se

basaría así en una de estrategia de desgaste -- se destruiría más del ARNm de huntingtina mutante que del ARNm de huntingtina no mutante o los otros aproximadamente 69 ARNm que codifican poliglutaminas.

V. Agentes de iARN

La presente divulgación incluye moléculas de ARNip diseñadas, por ejemplo, como se ha descrito anteriormente. Las moléculas de ARNip de la divulgación pueden ser químicamente sintetizadas, o se pueden transcribir *in vitro* de un molde de ADN, o *in vivo*, de, por ejemplo, ARNhp, o, usando enzima DICER humana recombinante, para escindir moldes de ARNbc transcritos *in vitro* en conjuntos de ARN dúplex de 20, 21 o 23 pb que median en la iARN. Las moléculas de ARNip se pueden diseñar usando cualquier método conocido en la técnica.

En un aspecto, en lugar del agente de iARN que es un ácido ribonucleico interferente, por ejemplo, un ARNip o ARNhp como se ha descrito anteriormente, el agente de iARN puede codificar un ácido ribonucleico interferente, por ejemplo, un ARNhp, como se ha descrito anteriormente. En otras palabras, el agente de iARN puede ser un molde transcripcional del ácido ribonucleico interferente. Así, los agentes de iARN de la presente divulgación también pueden incluir ARN de horquilla pequeña (ARNhp), y construcciones de expresión manipuladas para expresar ARNhp. La transcripción de ARNhp se inicia en un promotor de la polimerasa III (pol III), y se cree que termina en la posición 2 de un sitio de terminación de 4-5-timina de la transcripción. Tras la expresión, se cree que los ARNhp se pliegan en una estructura de tallo-bucle con nucleótidos protuberantes UU de 3'; posteriormente, los extremos de estos ARNhp se procesan, convirtiendo los ARNhp en moléculas de tipo ARNip de aproximadamente 21-23 nucleótidos (Brummelkamp et al., 2002; Lee et al., 2002, arriba; Miyagishi et al., 2002; Paddison et al., 2002, arriba; Paul et al., 2002, arriba; Sui et al., 2002 arriba; Yu et al., 2002, arriba. Se puede encontrar en internet más información sobre el diseño y uso de ARNhp en las siguientes direcciones: katahdin.cshl.org:9331/RNAi/docs/BseRI-BamHI_Strategy.pdf and katahdin.cshl.org:9331/RNAi/docs/Web_version_of_PCR-strategy1.pdf.

Las construcciones de expresión de la presente divulgación incluyen cualquier construcción adecuada 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, como se conoce en la técnica. Dichas construcciones de expresión pueden incluir uno o más promotores inducibles, sistemas de promotores de ARN Pol III tales como los promotores de ARNnp U6 o promotores de la ARN polimerasa III HI, u otros promotores conocidos en la técnica. Las construcciones pueden incluir una o ambas cadenas del ARNip. Las construcciones de expresión que expresan ambas cadenas también pueden incluir estructuras de bucle que unen ambas cadenas, o cada cadena se puede transcribir por separado de promotores separados dentro de la misma construcción. Cada cadena también se puede transcribir de una construcción de expresión separada (Tuschl, T., 2002, arriba).

Se pueden administrar ARNip sintéticos en células por métodos conocidos en la técnica, que incluyen transfección de liposomas catiónicos y electroporación. Sin embargo, estos ARNip exógenos muestran, en general, persistencia a corto plazo del efecto de silenciamiento (4~5 días en células cultivadas), que puede ser beneficioso en solo ciertas realizaciones. Para obtener la supresión a largo plazo de los genes diana (es decir, genes mutantes) y para facilitar la administración en ciertas circunstancias, se pueden expresar uno o más ARNip dentro de células de construcciones de ADN recombinante. Se conocen en la técnica dichos métodos para expresar dúplex de ARNip dentro de células de construcciones de ADN recombinante para permitir la supresión de genes de diana a largo plazo en células, que incluyen sistemas de promotores de mamífero Pol III (por ejemplo, sistemas de promotores HI o U6/ARNnp (Tuschl, T., 2002, arriba) capaces de expresar ARNip bicatenarios funcionales; (Bagella et al., 1998; Lee et al., 2002, arriba; Miyagishi et al., 2002, arriba; Paul et al., 2002, arriba; Yu et al., 2002), arriba; Sui et al., 2002, arriba). La terminación transcripcional por ARN Pol III ocurre en series de cuatro restos T consecutivos en el molde de ADN, que proporciona un mecanismo para terminar el transcrito de ARNip en una secuencia específica. El ARNip es complementario a la secuencia del gen diana en las orientaciones 5'-3' y 3'-5', y las dos cadenas del ARNip se pueden expresar en la misma construcción o en construcciones separadas. El ARNip de horquilla, conducido por el promotor HI o ARNnp U6 y expresado en células, puede inhibir la expresión de genes diana (Bagella et al., 1998; Lee et al., 2002, arriba; Miyagishi et al., 2002, arriba; Paul et al., 2002, arriba; Yu et al., 2002), arriba; Sui et al., 2002, arriba). Las construcciones que contienen secuencia de ARNip bajo el control del promotor T7 también hacen funcional el ARNip cuando se cotransfecta en las células con un vector que expresa ARN polimerasa T7 (Jacque et al., 2002, arriba). Una única construcción puede contener múltiples secuencias que codifican ARNip, tales como múltiples regiones del gen que codifican htt mutante, que silencia el mismo gen o múltiples genes, y se puede accionar, por ejemplo, por sitios de promotor Pol III separados.

Las células de animales expresan una variedad de ARN no codificantes de aproximadamente 22 nucleótidos denominados microARN (miARN) que pueden regular la expresión génica al nivel postranscripcional o traduccional durante el desarrollo animal. Una característica común de los miARN es que todos se escinden de un tallo-bucle de ARN de precursor de aproximadamente 70 nucleótidos, probablemente por Dicer, una enzima de tipo RNasa III, o un homólogo de la misma. Se puede usar la sustitución de las secuencias de tallo del precursor de miARN con secuencia complementaria al ARNm diana, una construcción de vector que expresa el precursor manipulado, para producir ARNip para iniciar la iARN contra dianas de ARNm específicas en células de mamífero (Zeng et al., 2002, arriba). Cuando se expresa por vectores de ADN que contienen promotores de polimerasa III, las horquillas diseñadas como micro-ARN pueden silenciar la expresión génica (McManus et al., 2002, arriba). Los micro-ARN que silencian polimorfismos también pueden ser útiles para bloquear la traducción de proteínas mutantes, en ausencia de

silenciamiento génico mediado por ARNip. Dichas aplicaciones pueden ser útiles en situaciones, por ejemplo, donde un ARNip diseñado causó el silenciamiento inespecífico de la proteína natural.

También se pueden usar mecanismos de administración mediados por virus para inducir el silenciamiento específico de genes diana mediante la expresión de ARNip, por ejemplo, generando adenovirus recombinantes que alojan ARNip bajo el control de la transcripción del promotor de ARN Pol II (Xia et al., 2002, arriba). La infección de células HeLa por estos adenovirus recombinantes permite la reducida expresión de genes diana endógenos. La inyección de los vectores de adenovirus recombinante en ratones transgénicos que expresan los genes diana del ARNip da como resultado la reducción *in vivo* de la expresión del gen diana. Ídem. En un modelo animal, la electroporación de embrión completo puede suministrar eficientemente ARNip sintético en embriones de ratón después de la implantación (Calegari et al., 2002). En ratones adultos, se puede llevar a cabo la eficiente administración de ARNip por la técnica de administración de "alta presión", una inyección rápida (en el plazo de 5 segundos) de un gran volumen de ARNip que contiene disolución en el animal por la vena de la cola (Liu et al., 1999, arriba; McCaffrey et al., 2002, arriba; Lewis et al., 2002). También se pueden usar nanopartículas y liposomas para administrar ARNip en animales.

Las composiciones de ácido nucleico de la divulgación incluyen tanto ARNip sin modificar como ARNip modificados como se conoce en la técnica, tales como los derivados de ARNip reticulados o derivados que tienen restos no de nucleótido unidos, por ejemplo, a sus extremos 3' o 5'. La modificación de derivados de ARNip de esta forma puede mejorar la captación celular o potenciar las actividades de silenciamiento celular del derivado de ARNip resultante en comparación con el ARNip correspondiente, son útiles para rastrear el derivado de ARNip en la célula, o mejorar la estabilidad del derivado de ARNip en comparación con el ARNip correspondiente.

Los precursores de ARN manipulados, introducidos en células u organismos completos como se describe en el presente documento, conducirán a la producción de una molécula de ARNip deseada. Dicha molécula de ARNip se asociará entonces con componentes de proteína endógena de la vía de iARN para unirse a y silenciar una secuencia de ARNm específica para la escisión y destrucción. En este modo, el ARNm a ser silenciado por el ARNip generado a partir del precursor de ARN manipulado se agotará de la célula u organismo, que conduce a una disminución en la concentración de la proteína codificada por ese ARNm en la célula u organismo. Los precursores de ARN normalmente son moléculas de ácidos nucleicos que codifican individualmente una cadena de un ARNbc o codifican toda la secuencia de nucleótidos de una estructura de bucle de horquilla de ARN.

Las composiciones de ácido nucleico de la divulgación pueden estar sin conjugar o se pueden conjugar con otro resto, tal como una nanopartícula, para potenciar una propiedad de las composiciones, por ejemplo, un parámetro farmacocinético tal como absorción, eficacia, biodisponibilidad y/o semivida. La conjugación se puede llevar a cabo por métodos conocidos en la técnica, por ejemplo, usando los métodos de Lambert et al., Drug Deliv. Rev. 47(1), 99-112 (2001) (describe ácidos nucleicos cargados en nanopartículas de polialquilcianoacrilato (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 unidos a agentes intercalantes, grupos hidrófobos, policationes o nanopartículas de PACA); y Godard et al., Eur. J. Biochem. 232(2):404-10 (1995) (describe ácidos nucleicos unidos a nanopartículas).

Las moléculas de ácidos nucleicos de la presente divulgación también se pueden marcar usando cualquier método conocido en la técnica; por ejemplo, las composiciones de ácido nucleico se pueden marcar con un fluoróforo, por ejemplo, Cy3, fluoresceína o rodamina. El marcado se puede llevar a cabo usando un kit, por ejemplo, el kit de marcado de ARNip SILENCER™ (Ambion). Además, el ARNip se puede radiomarcarse, por ejemplo, usando ³H, ³²P, u otro isótopo apropiado.

Además, debido a que se cree que la iARN progresa por al menos un producto intermedio de ARN monocatenario, el experto apreciará que también se pueden diseñar ARNip-mc (por ejemplo, la cadena no codificante de un ARNip-bc) (por ejemplo, para síntesis química) generados (por ejemplo, generados enzimáticamente) o expresados (por ejemplo, a partir de un vector o plásmido) como se describe en el presente documento y utilizados según las metodologías reivindicadas. Además, en invertebrados, la iARN puede ser eficazmente desencadenada por ARNbc largos (por ejemplo, ARNbc de aproximadamente 100 - 1000 nucleótidos de longitud, preferentemente aproximadamente 200 - 500, por ejemplo, aproximadamente 250, 300, 350, 400 o 450 nucleótidos en longitud) que actúan como efectores de iARN (Brondani et al., Proc Natl Acad Sci USA. 2001 Dec 4;98(25):14428-33. Epub 2001 Nov 27).

VI. Métodos de introducción de ARN, vectores y células hospedadoras

Los métodos físicos de introducción de ácidos nucleicos incluyen la inyección de una disolución que contiene el ARN, bombardeo por partículas cubiertas por el ARN, impregnación de la célula u organismo en una disolución del ARN, o electroporación de membranas celulares en presencia del ARN. Una construcción viral encapsidada en una partícula viral lograría tanto la eficiente introducción de una construcción de expresión en la célula como la transcripción de ARN codificado por la construcción de expresión. Se pueden usar otros métodos conocidos en la técnica para introducir ácidos nucleicos en células, tales como el transporte de vehículos mediado por lípidos, transporte mediado por productos químicos, tal como fosfato de calcio, y similares. Así, el ARN se puede introducir junto con componentes que desempeñan una o más de las siguientes actividades: potenciar la captación de ARN por la célula, inhibir la

hibridación de cadenas individuales, estabilizar las cadenas individuales o aumentar de otro modo la inhibición del gen diana.

El ARN puede ser directamente introducido en la célula (es decir, intracelularmente); o introducido extracelularmente en una cavidad, espacio intersticial, en la circulación de un organismo, introducido por vía oral, o se puede introducir sumergiendo una célula u organismo en una disolución que contiene el ARN. La circulación vascular o extravascular, el sistema circulatorio o linfático y el líquido cefalorraquídeo son sitios donde se puede introducir el ARN.

La célula que tiene el gen diana puede ser de la línea germinal o somática, totipotente o pluripotente, divisora o no divisora, parénquima o epitelio, immortalizada o transformada, o similares. La célula puede ser una célula madre o una célula diferenciada. Los tipos de células que están diferenciados incluyen adipocitos, fibroblastos, miocitos, cardiomiocitos, endotelio, neuronas, glía, glóbulos sanguíneos, 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 diana particular y la dosis de material de ARN bicatenario suministrado, este proceso puede proporcionar pérdida de función parcial o completa para el gen diana. Una reducción o pérdida de la expresión génica en al menos 50 %, 60 %, 70 %, 80 %, 90 %, 95 % o 99 % o más de células silenciadas es a modo de ejemplo. Inhibición de expresión génica se refiere a la ausencia (o disminución observable) al nivel de proteína y/o producto de ARNm de un gen diana. La especificidad se refiere a la capacidad para inhibir el gen diana sin manifestar efectos sobre otros genes de la célula. Las consecuencias de la inhibición se pueden confirmar por examen de las propiedades exteriores de la célula u organismo (como se presenta más adelante en los ejemplos) o por técnicas bioquímicas tales como hibridación en disolución de ARN, protección con nucleasa, hibridación Northern, transcripción inversa, monitorización de la expresión génica con una micromatriz, unión de anticuerpos, ensayo inmunoanalítico de adsorción (ELISA), transferencia Western, radioinmunoensayo (RIA), otros inmunoensayos y análisis por citometría de flujo (FACS).

Para la inhibición mediada por ARN en una línea celular u organismo completo, la expresión génica se ensaya convenientemente usando un gen indicador o de resistencia a fármaco cuyo producto de proteína se ensaya fácilmente. Dichos genes indicadores incluyen acetohidroxiácido sintasa (AHAS), fosfatasa alcalina (AP), beta-galactosidasa (LacZ), beta-glucoronidasa (GUS), cloranfenicol acetiltransferasa (CAT), proteína verde fluorescente (GFP), peroxidasa de rábano picante (HRP), luciferasa (Luc), nopalina sintasa (NOS), octopina sintasa (OCS) y derivados de las mismas. Están disponibles múltiples marcadores de selección que confieren resistencia a ampicilina, bleomicina, cloranfenicol, gentamicina, higromicina, kanamicina, lincomicina, metotrexato, fosfotricina, puromicina y tetraciclina. Dependiendo del ensayo, la cuantificación de la cantidad de expresión génica permite determinar un grado de inhibición que es mayor que 10 %, 33 %, 50 %, 90 %, 95 % o 99 % en comparación con una célula no tratada según la presente divulgación. Dosis más bajas de material inyectado y tiempos más largos después de la administración de agente de iARN pueden dar como resultado la inhibición en una fracción más pequeña de células (por ejemplo, al menos 10 %, 20 %, 50 %, 75 %, 90 % o 95 % de células silenciadas). La cuantificación de la expresión génica en una célula puede mostrar cantidades similares de inhibición al nivel de acumulación de ARNm diana o la traducción de proteína diana. Como un ejemplo, la eficiencia de inhibición se puede determinar evaluando la cantidad de producto génico en la célula; se puede detectar ARNm con una sonda de hibridación que tiene una secuencia de nucleótidos fuera de la región usada para el ARN bicatenario inhibidor, o se puede detectar polipéptido traducido con un anticuerpo producido contra la secuencia de polipéptidos de esa región.

El ARN se puede introducir en una cantidad que permite la administración de al menos una copia por célula. Dosis más altas (por ejemplo, al menos 5, 10, 100, 500 o 1000 copias por célula) de material pueden dar inhibición más eficaz; también puede ser útiles dosis más bajas para aplicaciones específicas.

En un aspecto preferido, se prueba la eficacia de un agente de iARN de la divulgación (por ejemplo, un ARNip que silencia un polimorfismo en un gen mutante) para su capacidad para degradar específicamente ARNm mutante (por ejemplo, ARNm de htt mutante y/o la producción de la proteína huntingtina mutante) en células, en particular, en neuronas (por ejemplo, cuerpo estriado o líneas clonales neuronales corticales y/o neuronas primarias). También son adecuados para los ensayos de validación basados en células otras células fácilmente transfectables, por ejemplo, células HeLa o células COS. Las células se transfectan con ADNc humanos no mutantes o mutantes (por ejemplo, ADNc no mutante humano o de huntingtina mutante). Se co-transfectan los ARNip convencionales, ARNip modificados o vectores capaces de producir ARNip de ARNm en bucle en U. Se mide la reducción selectiva en ARNm mutante (por ejemplo, ARNm de huntingtina mutante) y/o proteína mutante (por ejemplo, huntingtina mutante). Se puede comparar la reducción de ARNm o proteína mutante con niveles de ARNm o proteína normal. Se puede ensayar ARNm o proteína normal exógenamente introducida (o ARNm o proteína normal endógena) para fines de comparación. Cuando se utilizan células neuronales, que se conocen por ser algo resistentes a las técnicas de transfección convencionales, se puede desear introducir agentes de iARN (por ejemplo, ARNip) por captación pasiva.

VII. Métodos de tratamiento:

La presente divulgación proporciona métodos profilácticos y terapéuticos para tratar un sujeto en riesgo de (o susceptible a) una enfermedad o trastorno causado, por completo o en parte, por una proteína mutante de ganancia de función. En una realización, la enfermedad o trastorno es una enfermedad o trastorno de repetición de

trinucleótidos. En otra realización, la enfermedad o trastorno es un trastorno de las poliglutaminas. En una realización preferida, la enfermedad o trastorno es un trastorno asociado a la expresión de huntingtina y en la que la alteración de huntingtina, especialmente la amplificación del número de copias de la repetición de CAG, conduce a un defecto en el gen huntingtina (estructura o función) o la proteína huntingtina (estructura o función o expresión), de forma que las manifestaciones clínicas incluyen las observadas en pacientes con enfermedad de Huntington.

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

En un aspecto, la divulgación proporciona un método de prevención en un sujeto de una enfermedad o trastorno como se ha descrito anteriormente, administrando al sujeto un agente terapéutico (por ejemplo, un agente o vector o transgén de iARN que codifica el mismo). Los sujetos en riesgo de la enfermedad se pueden identificar, por ejemplo, por cualquiera o una combinación de ensayos de diagnóstico o de pronóstico como se describen en el presente documento. La administración de un agente profiláctico puede ocurrir antes de la manifestación de los síntomas característicos de la enfermedad o trastorno, de forma que la enfermedad o trastorno se prevenga o, alternativamente, se retrase en su progresión.

Otro aspecto de la divulgación se refiere a métodos para tratar sujetos terapéuticamente, es decir, alterar la aparición de síntomas de la enfermedad o trastorno. En una realización a modo de ejemplo, el método modulador de la divulgación implica poner en contacto una célula que expresa una ganancia de función mutante con un agente terapéutico (por ejemplo, un agente de iARN o vector o transgén que codifica el mismo) que es específico para un polimorfismo dentro del gen, de forma que se logre la interferencia específica de secuencia con el gen. Estos métodos se pueden realizar *in vitro* (por ejemplo, cultivando la célula con el agente) o, alternativamente, *in vivo* (por ejemplo, administrando el agente a un sujeto).

Con respecto a tanto métodos profilácticos como terapéuticos de tratamiento, dichos tratamientos se pueden adaptar o modificar específicamente, basándose en el conocimiento obtenido del campo de la farmacogenómica. "Farmacogenómica", como se usa en el presente documento, se refiere a la aplicación de tecnologías de genómica, tales como secuenciación génica, genética estadística y análisis de expresión génica a fármacos en desarrollo clínico y a la venta. Más específicamente, el término se refiere al estudio de cómo los genes de un paciente determinan su respuesta a un fármaco (por ejemplo, "fenotipo de respuesta del fármaco" o "genotipo de respuesta del fármaco" de un paciente). Así, otro aspecto de la divulgación proporciona métodos para adaptar el tratamiento profiláctico o terapéutico de un individuo con cualquiera de las moléculas de gen diana de la presente divulgación o moduladores de gen diana según el genotipo de respuesta del fármaco individual. La farmacogenómica permite que un profesional clínico o médico dirija los tratamientos profilácticos o terapéuticos a los pacientes que se beneficiarán más del tratamiento y para evitar el tratamiento de pacientes que experimentarán efectos secundarios relacionados con fármacos tóxicos.

Se pueden probar agentes terapéuticos en un modelo animal apropiado. Por ejemplo, se puede usar un agente de iARN (o vector de expresión o transgén que codifica el mismo) como se describe en el presente documento en un modelo animal para determinar la eficacia, la toxicidad o los efectos secundarios del tratamiento con dicho agente. Alternativamente, se puede usar un agente terapéutico en un modelo animal para determinar el mecanismo de acción de dicho agente. Por ejemplo, se puede usar un agente en un modelo animal para determinar la eficacia, la toxicidad o los efectos secundarios del tratamiento con dicho agente. Alternativamente, se puede usar un agente en un modelo animal para determinar el mecanismo de acción de dicho agente.

VIII. Composiciones farmacéuticas

La divulgación se refiere al uso de los agentes anteriormente descritos para tratamientos profilácticos y/o terapéuticos como se describe más adelante. Por consiguiente, los moduladores (por ejemplo, agentes de iARN) de la presente divulgación se pueden incorporar en composiciones farmacéuticas adecuadas para su administración. Dichas composiciones normalmente comprenden la molécula de ácido nucleico, proteína, anticuerpo o compuesto modulador y un vehículo farmacéuticamente aceptable. Como se usa en el presente documento, la expresión "vehículo farmacéuticamente aceptable" pretende incluir todos y cada uno de los disolventes, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y de retraso de la absorción, y similares, compatibles con la administración farmacéutica. Se conoce bien en la técnica el uso de dichos medios y agentes para sustancias farmacéuticamente activas. Excepto en la medida de que cualquier medio o agente convencional sea incompatible con el compuesto activo, se contempla el uso del mismo en las composiciones. También se pueden incorporar compuestos activos suplementarios en las composiciones.

Se formula una composición farmacéutica de la divulgación para ser compatible con su vía de administración prevista. Los ejemplos de las vías de administración incluyen parenteral, por ejemplo, administración intravenosa, intradérmica, subcutánea, intraperitoneal, intramuscular, oral (por ejemplo, inhalación), transdérmica (tópica) y transmucosa. Las

disoluciones o suspensiones usadas para administración parenteral, intradérmica o subcutánea pueden incluir los siguientes componentes: un diluyente estéril tal como agua para inyección, solución salina, aceites no volátiles, polietilenglicoles, glicerina, propilenglicol u otros disolventes sintéticos; agentes antibacterianos tales como alcohol bencílico o metilparabenos; antioxidantes tales como ácido ascórbico o bisulfito de sodio; agentes quelantes tales como ácido etilendiaminetetraacético; tampones tales como acetatos, citratos o fosfatos y agentes para el ajuste de la tonicidad tales como cloruro sódico o dextrosa. El pH se puede ajustar con ácidos o bases, tales como ácido clorhídrico o hidróxido sódico. La preparación parenteral se puede encerrar en ampollas, jeringas desechables o viales de dosis múltiple fabricados de vidrio o plástico.

Las composiciones farmacéuticas adecuadas para el uso inyectable incluyen disoluciones o dispersiones acuosas estériles (solubles en agua) y polvos estériles para la preparación extemporánea de disoluciones inyectables estériles o dispersión. Para administración intravenosa, los vehículos adecuados incluyen solución salina fisiológica, agua bacteriostática, Cremophor EL™ (BASF, Parsippany, NJ) o solución salina tamponada con fosfato (PBS). En todos los casos, la composición debe ser estéril y debe ser líquida hasta el punto de que exista fácil inyectabilidad. Debe ser estable en las condiciones de fabricación y almacenamiento y se debe conservar contra 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 se puede mantener, por ejemplo, usando un recubrimiento tal como lecitina, por el mantenimiento del tamaño de partículas requerido en el caso de dispersión y usando tensioactivos. La prevención de la acción de microorganismos se puede lograr por 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 sódico en la composición. La absorción prolongada de las composiciones inyectables se puede provocar incluyendo en la composición un agente que retrasa la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Se pueden preparar disoluciones inyectables estériles incorporando el compuesto activo en la cantidad requerida en un disolvente apropiado con uno o una combinación de los ingredientes enumerados anteriormente, según se requiera, seguido por esterilización por filtración. En general, se preparan dispersiones incorporando el compuesto activo en un vehículo estéril que contiene un medio de dispersión básico y los otros ingredientes requeridos de los enumerados anteriormente. En el caso de polvos estériles para la preparación de disoluciones inyectables estériles, los métodos preferidos de preparación son secado a vacío y liofilización que da un polvo del principio activo más cualquier ingrediente deseado adicional de una disolución previamente esterilizada por filtración de la misma.

Las composiciones orales incluyen, en general, un diluyente inerte o un vehículo comestible. Se pueden encerrar en cápsulas de gelatina o comprimir en comprimidos. Con el fin de la administración terapéutica oral, el compuesto activo se puede incorporar con excipientes y usar en forma de comprimidos, trociscos o cápsulas. Las composiciones orales también se pueden preparar usando un vehículo líquido para su uso como un enjuague bucal, en donde el compuesto en el vehículo líquido se aplica por vía oral y se enjuaga y expectora o traga. Se pueden incluir aglutinantes y/o materiales de adyuvante farmacéuticamente compatibles como parte de la composición. Los comprimidos, píldoras, cápsulas, trociscos y similares pueden contener cualquiera de los siguientes ingredientes, o compuestos de una naturaleza similar: un aglutinante tal como celulosa microcristalina, goma 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 edulcorante tal como sacarosa o sacarina; o un aromatizante tal como menta, salicilato de metilo o aromatizante de naranja.

Para administración por inhalación, los compuestos se administran en forma de un espray en aerosol de un recipiente o dispensador a presión que contiene un propulsor apropiado, por ejemplo, un gas tal como dióxido de carbono, o un nebulizador.

La administración sistémica también puede ser por medios transmucosos o transdérmicos. Para administración transmucosa o transdérmica, se usan en la formulación penetrantes apropiados para que atraviesen la barrera. Dichos penetrantes se conocen, en general, en la técnica, e incluyen, por ejemplo, para administración transmucosa, detergentes, sales biliares y derivados de ácido fusídico. La administración transmucosa se puede llevar a cabo mediante el uso de esprays nasales o supositorios. Para administración transdérmica, los compuestos activos se formulan en pomadas, bálsamos, geles o cremas, como se conoce, en general, en la técnica.

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.

En una realización, los compuestos activos se preparan con vehículos que protegerán el compuesto de la rápida eliminación del cuerpo, tal como una formulación de liberación controlada, que incluye implantes y sistemas de administración microencapsulada. Se pueden usar polímeros biodegradables biocompatibles, tales como etilenoacetato de vinilo, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres y ácido poliláctico. Los métodos para la preparación de dichas formulaciones serán evidentes para los expertos en la técnica. Los materiales también se pueden obtener comercialmente de Alza Corporation y Nova Pharmaceuticals, Inc. También se pueden usar suspensiones liposomales (que incluyen liposomas dirigidos a células infectadas con anticuerpos monoclonales a

antígenos virales) como vehículos farmacéuticamente aceptables. Estos se pueden preparar según métodos conocidos para los expertos en la técnica, por ejemplo, como se describe en la patente de EE. UU. N° 4.522.811.

Es especialmente ventajoso formular composiciones orales o parentales en forma unitaria de dosificación para facilitar la administración y la uniformidad de la dosificación. Forma unitaria de dosificación, como se usa en el presente documento, se refiere a unidades físicamente discretas aptas como administraciones unitarias para el sujeto que se va a tratar; conteniendo cada unidad una cantidad predeterminada de compuesto activo calculada para producir el efecto terapéutico deseado en asociación con el vehículo farmacéutico requerido. La memoria descriptiva para las formas unitarias de dosificación de la divulgación viene dictada por y es directamente dependiente de las características únicas del compuesto activo y el efecto terapéutico particular a lograr, y las limitaciones inherentes en la técnica de combinar dicho compuesto activo para el tratamiento de individuos.

La toxicidad y eficacia terapéutica de dichos compuestos se puede determinar por procedimientos farmacéuticos convencionales en cultivos celulares o animales experimentales, por ejemplo, para determinar la DL50 (la dosis letal para el 50 % de la población) y la DE50 (la dosis terapéuticamente eficaz en el 50 % de la población). La relación de dosis entre efectos tóxicos y terapéuticos es el índice terapéutico y se puede expresar como la relación DL50/DE50. Se prefieren los compuestos que presentan grandes índices terapéuticos. Aunque se pueden usar compuestos que presentan efectos secundarios tóxicos, se debe tener cuidado en el diseño de un sistema de administración que dirija dichos compuestos al sitio de tejido afectado para minimizar el posible daño a células no infectadas y, así, reducir los efectos secundarios.

Se pueden usar los datos obtenidos de los ensayos de cultivo celular y se pueden usar estudios en animales en la formulación de un intervalo de administración para su uso en seres humanos. La administración de dichos compuestos radica preferentemente dentro de un intervalo de concentraciones circulantes que incluyen la DE50 con poca o ninguna toxicidad. La administración puede variar dentro de este intervalo dependiendo de la forma farmacéutica empleada y la vía de administración utilizada. Para cualquier compuesto usado en el método de la divulgación, la dosis terapéuticamente eficaz se puede estimar inicialmente de ensayos de cultivo celular. Se puede formular una dosis en modelos animales para lograr un intervalo de concentración plasmática circulante que incluye la CE50 (es decir, la concentración del compuesto de prueba que alcanza una respuesta) como se determina en cultivo celular. Dicha información se puede usar para determinar con más exactitud dosis útiles en los seres humanos. Se pueden medir niveles en plasma, por ejemplo, por cromatografía líquida de alta resolución.

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

La presente divulgación se ilustra además por los siguientes ejemplos que no se debe interpretar como limitantes.

EJEMPLOS

A diferencia de otros tipos de enfermedades autosómicas dominantes, la enfermedad de Huntington no contiene una mutación puntual, por ejemplo, cambio de un único nucleótido. Por tanto, no se puede implementar la estrategia para diseñar ARNip dirigidos contra una mutación puntual en el alelo de enfermedad. En su lugar, la presente divulgación se dirige a ARNip diseñado contra polimorfismos en el gen huntingtina, de los que hay aproximadamente 30 disponibles en GenBank. La presente divulgación también identifica el polimorfismo en el alelo de la enfermedad de Huntington que se diferencia del alelo no mutante, de manera que el ARNip destruye solo el ARNm de la enfermedad y deja intacto el ARNm del alelo no mutante (normal). Así, solo se destruye la proteína huntingtina mutante y la proteína normal está intacta.

Ejemplo I: Prueba de agentes de iARN (por ejemplo, ARNip) contra htt mutante en lisados de *Drosophila*

Se diseñó una posición 2886 que silencia ARNip en el ARNm de htt como se describe arriba. La secuencia del ARNip se representa en la Figura 5a (SEQ ID NO: 24 codificante; 25 no codificante). Se desprotegió ARN sintético (Dharmacon) según el protocolo del fabricante. Se hibridaron cadenas de ARNip (Elbashir et al., 2001a).

Se prepararon ARN diana del siguiente modo. Se transcribieron ARN diana con T7 ARN polimerasa recombinante, marcada con histidina, de productos de PCR como se ha descrito (Nykänen et al., 2001; Hutvagner et al., 2002). Se generaron moldes de PCR para codificante y no codificante de htt amplificando 0,1 ng/mL (concentración final) de molde de plásmido que codifica ADNc de htt usando los siguientes pares de cebadores: diana codificante de htt, 5'-GCG TAA TAC GAC TCA CTA TAG GAA CAG TAT GTCTCA GAC ATC-3' (SEQ ID NO: 30) y 5'-UUCG AAG UAU UCC GCG UAC GU-3' (SEQ ID NO: 31); diana no codificante de htt, 5'-GCG TAA TAC GAC 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).

Se probó el ARNip usando un ensayo de iARN *in vitro*, que caracteriza los lisados de embrión de *Drosophila*. Se llevaron a cabo las reacciones y los análisis de iARN *in vitro* como se ha descrito previamente (Tuschl et al., 1999; Zamore et al., 2000; Haley et al., 2003). Se usaron dianas de ARN a concentración de ~ 5 nM de manera que las reacciones estuvieran principalmente en condiciones de reemplazo único. La escisión diana en estas condiciones es proporcional a la concentración de ARNip.

La Figura 5a muestra la eficacia del ARNip dirigido contra la posición 2886 en el htt mutante. Los datos demuestran claramente que el ARNip dirige la escisión de la diana codificante a un mayor grado que la observada para la diana no codificante. Sin embargo, se reconoce que este ARNip diseñado en primer lugar no produjo una molécula muy activa, al menos en este ensayo *in vitro*. El análisis termodinámico de la intensidad del par de bases en los dos extremos del dúplex de ARNip indicó intensidades de pares de bases aproximadamente equivalentes. La Figura 4 representa el análisis termodinámico de extremos 5' de cadenas codificantes (SEQ ID NO: 20; 22 respectivamente) y no codificantes (SEQ ID NO: 21; 23 respectivamente) de ARNip para el dúplex de ARNip en 5a. Se calculó ΔG (kcal/mol) en NaCl 1 M a 37 °C.

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

Ejemplo II: Inactivación de iARN de la proteína Htt en células cultivadas

En un primer experimento, se probaron ARNip que silencian un polimorfismo en el ARNm de htt (es decir, el polimorfismo en la posición 2886 en el ARNm de htt) para su capacidad para regular por disminución la proteína Htt endógena en células HeLa. Se cultivaron y transfectaron células HeLa del siguiente modo. Se mantuvieron células HeLa a 37 °C en medio Eagle modificado por Dulbecco (DMEM, Invitrogen) complementado con 10 % de suero bovino fetal (FBS), 100 unidades/mL de penicilina y 100 µg/mL de estreptomina (Invitrogen). Las células se sometieron a pases regulares a sub-confluencia y se sembraron a 70 % de confluencia 16 horas antes de la transfección. Se realizó transfección transitoria mediada por Lipofectamine™ (Invitrogen) de ARNip en placas de 6 pocillos por duplicado (Falcon) como se describe para las líneas de células adherentes por el fabricante. Se añadió a cada pocillo una mezcla de transfección estándar que contenía ARNip 100-150 nM y 9-10 µL de Lipofectamine™ en 1 mL de OPTI-MEM® reducido en suero (Invitrogen). Las células se incubaron en mezcla de transfección a 37 °C durante 6 horas y se cultivaron adicionalmente en DMEM libre de antibiótico. Para el análisis de transferencia Western en diversos intervalos de tiempo, se recogieron las células transfectadas, se lavaron dos veces con solución salina tamponada con fosfato (PBS, Invitrogen), se ultracongelaron en nitrógeno líquido y se conservaron a -80 °C para el análisis.

Se probaron tres ARNip contra una secuencia diana común en el exón 1 y se probaron cuatro ARNip para el polimorfismo de la posición 2886. Se realizó el análisis de transferencia Western del siguiente modo. Se recogieron células tratadas con ARNip como se ha descrito anteriormente y se lisaron en tampón de lisis indicador frío en hielo (Promega) que contenía inhibidor de la proteasa (completo, libre de EDTA, 1 comprimido/10 mL de tampón, Roche Molecular Biochemicals). Después de aclarar los lisados resultantes por centrifugación, se cuantificó la proteína en los lisados claros por el kit del ensayo de proteínas Dc (Bio-Rad). Se resolvieron proteínas en 60 µg de lisado celular total por 10 % de SDS-PAGE, se transfirieron sobre una membrana de poli(difluoruro de vinilideno) (PVDF, Bio-Rad) y se inmunotransfirieron con anticuerpos contra CD80 (Santa Cruz). Se visualizó el contenido de proteína con un kit de inmunotransferencia BM Chemiluminescence (Roche Molecular Biochemicals). Las transferencias se expusieron a película de rayos X (Kodak MR-1) durante diversos tiempos (30 s a 5 min). La Figura 6a representa los resultados del análisis Western. La tubulina sirvió de control de carga. Se cuantifican los datos y se normalizaron en la Figura 6b. Del ARNip probado, 2886-4, mostró reproduciblemente eficacia potenciada en células HeLa cultivadas (Figura 6). Este ARNip también mostró reproduciblemente eficacia *in vitro* potenciada (no mostrada). ARNip de GFP es un ARNip de control que no comparte homología de secuencias con ARNm de htt.

Asimismo, se puede probar ARNip contra regiones polimórficas en el ARNm de htt en células transfectadas con ADNc de htt humano o en células transfectadas con construcciones indicadoras de htt. Las cotransfecciones transitorias mediadas por Lipofectamine™ (Invitrogen) de ADNcs o plásmidos indicadores y ARNip se realizan como se describe arriba. Para probar la capacidad de ARNip para silenciar las construcciones de htt informadas, se usó iARN para inhibir la expresión de GFP-htt en líneas celulares HeLa humanas cultivadas. Brevemente, se transfectaron células HeLa con dúplex de ARNip de GFP-htt, que silencia la secuencia de ARNm de GFP-htt. Para analizar los efectos de iARN contra GFP-htt, se prepararon lisados a partir de células tratadas con dúplex de ARNip en diversos momentos después de la transfección. Se llevaron a cabo experimentos de transferencia Western como se describe arriba. Brevemente, se recogieron las células HeLa en diversos momentos después de la transfección, se resolvió su contenido de proteína sobre 10 % de SDS-PAGE, se transfirió sobre membranas de PVDF y se inmunotransfirió con anticuerpos apropiados. Los resultados de este estudio indicaron que el ARNip contra GFP puede eliminar la expresión de la expresión de GFP-htt en células HeLa transfectadas con el gen GFP-htt. Para estudios que silencian htt exógenamente introducida, los procedimientos son como se describen, excepto que se usan anticuerpos anti-Htt para la inmunotransferencia.

También se puede usar iARN para inhibir la expresión de htt en células neuronales cultivadas. Las células a modo de ejemplo incluyen las líneas celulares PC12 (Scheitzer et al., Thompson et al.) y NT3293 (Tagle et al.) como se ha descrito previamente. Células adicionales a modo de ejemplo incluyen células establemente transfectadas, por

ejemplo, células neuronales o células neuronalmente derivadas. Se pueden usar líneas celulares PC12 que expresan el exón 1 del gen huntingtina humana (Htt), aunque la expresión del exón 1 reduce la supervivencia celular. También se pueden usar células GFP-Htt PC12 que tienen un gen GFP-htt inducible para probar o validar la eficacia de ARNip.

Ejemplo III: Administración de ARNip de Htt en un entorno *in vivo*

- 5 Los modelos de ratones R6/2 (que expresan el producto de ADNc de htt de R6/2 humano) son un modelo animal aceptado para estudiar la eficacia de la administración de ARNip en un entorno *in vivo*. Se usaron ratones R6/2 genéticamente manipulados para probar la eficacia de ARNip en el extremo 5' de ARNm de huntingtina. Se inyectó ARNip de Htt en el cuerpo estriado de ratones R6/2 mediante una bomba Alzet. Los ratones se trataron durante 14 días con el sistema de administración de bomba ARNip/Alzet.
- 10 Los resultados de este estudio indicaron que los dos ratones que recibieron el ARNip con Trans-IT TKO (Mirus) como una disolución 20 o 200 nM a 0,25 µL/hora no mostraron deterioro de la alteración motora desde el día 67 hasta el día 74. En general, se espera que estos R6/2 tengan una reducción continuada en la varilla giratoria más allá del día 60.

Referencias

- Aronin et al., Neuron, Nov; 15(5):1193-201 (1995)
- 15 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)
- 20 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)
- 25 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)
- 30 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)
- 35 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.
- 40 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 U S A 99,6047-6052 (2002)

5 Zamore et al., Cell 101, 25-33 (2000)

Zamore et al., Nature Medicine, volume 9 Number 3 pp 266 - 267 (2003)

Zeng et al., (2002)

Equivalentes

10 Los expertos en la técnica reconocerán, o serán capaces de determinar usando no más de experimentación rutinaria, muchos equivalentes a las realizaciones específicas de la invención descritas en el presente documento. Dichos equivalentes pretenden estar englobados por las siguientes reivindicaciones.

LISTADO DE SECUENCIAS

<110> Universidad de Massachusetts

ARONIN, Neil

15 ZAMORE, Phillip D.

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

<130> UMY-083PC

20

<150> 60/502678

<151> 12-09-2003

<160> 33

25

<170> FastSEQ para Windows Versión 4.0

<210> 1

<211> 13672

30

<212> ADN

<213> Homo sapiens

<400> 1

ES 2 808 561 T3

ttgctgtgtg	aggcagaacc	tgcgggggca	ggggcgggct	ggttccctgg	ccagccattg	60
gcagagtccg	caggctaggg	ctgtcaatca	tgttgcccg	cgtggccccg	cctccgcccg	120
cgcgcccccg	cctccgcccg	cgcacgtctg	ggacgcaagg	cgccgtgggg	gctgccggga	180
cgggtccaag	atggacggcc	gctcaggttc	tgtttttacc	tgcggcccag	agccccattc	240
attgcccccg	tgttgagcgg	cgccgcgagt	cgcccgagg	cctccgggga	ctgccgtgcc	300
gggggggaga	ccgcatggc	gacctggaa	aagctgatga	aggccttcga	gtccctcaag	360
tccttcagc	agcagcagca	gcagcagcag	cagcagcagc	agcagcagca	gcagcagcag	420
cagcagcagc	aacagccgcc	accgccgccg	ccgccgccgc	cgctccctca	gcttcctcag	480
ccgcgcgcgc	aggcacagcc	gctgctgcct	cagccgcagc	cgcccccgcc	gccgcccccg	540
ccgccacccg	gcccggctgt	ggctgaggag	ccgctgcacc	gaccaaagaa	agaactttca	600
gctaccaaga	aagaccgtgt	gaatcattgt	ctgacaatat	gtgaaaacat	agtggcacag	660
tctgtcagaa	attctccaga	atttcagaaa	cttctgggca	tgcctatgga	actttttctg	720
ctgtgcagtg	atgacgcaga	gtcagatgtc	aggatgggtg	ctgacgaatg	cctcaacaaa	780
gttatcaaa	ctttgatgga	ttctaattct	ccaaggttac	agctcgagct	ctataaggaa	840
attaaaaaga	atggtgccc	tgcgagtttg	cgtgctgccc	tgtggaggtt	tgtcgagctg	900
gtcacctgg	ttcggcctca	gaaatgcagg	ccttacctgg	tgaaccttct	gccgtgcctg	960
actcgaacaa	gcaagagacc	cgaagaatca	gtccaggaga	ccttggtctg	agctgttccc	1020
aaaattatgg	cttcttttgg	caattttgca	aatgacaatg	aaattaaggt	tttgtaaag	1080
gccttcatag	cgaacctgaa	gtcaagctcc	cccaccattc	ggcgacagc	ggctggatca	1140
gcagtgagca	tctgccagca	ctcaagaagg	acacaatatt	tctatagttg	gctactaaat	1200
gtgctcttag	gcttactcgt	tcctgtcgag	gatgaacact	ccactctgct	gattcttggc	1260
gtgctgctca	ccctgaggtg	tttggtgccc	ttgctgcagc	agcaggtcaa	ggacacaagc	1320
ctgaaaggca	gcttcggagt	gacaaggaaa	gaaatggaag	tctctccttc	tgcagagcag	1380
cttgtccagg	tttatgaact	gacgttacat	catacacagc	accaagacca	caatggtgtg	1440
accggagccc	tggagctgtt	gcagcagctc	ttcagaacgc	ctccaccoga	gcttctgcaa	1500
accctgaccg	cagtcggggg	cattggggcag	ctcaccgctg	ctaaggagga	gtctggtggc	1560
cgaagccgta	gtgggagtat	tgtggaactt	atagctggag	ggggttcctc	atgcagccct	1620
gtcctttcaa	gaaaacaaaa	aggcaaagtg	ctcttaggag	aagaagaagc	cttggaggat	1680
gactctgaat	cgagatcgga	tgtcagcagc	tctgccttaa	cagcctcagt	gaaggatgag	1740
atcagtgagg	agctggctgc	ttcttcaggg	gtttccactc	cagggtcagc	aggctcatgac	1800
atcatcacag	aacagccacg	gtcacagcac	acactgcagg	cggactcagt	ggatctggcc	1860
agctgtgact	tgacaagctc	tgccactgat	ggggatgagg	aggatatctt	gagccacagc	1920
tcagccagc	tcagcgccgt	cccatctgac	cctgccatgg	acctgaatga	tgggaccag	1980
gcctcgtcgc	ccatcagcga	cagctcccag	accaccaccg	aagggcctga	ttcagctgtt	2040
accccttcag	acagttctga	aattgtgtta	gacggtaccg	acaaccagta	tttgggcctg	2100
cagattggac	agccccagga	tgaagatgag	gaagccacag	gtattcttcc	tgatgaagcc	2160
tcggaggcct	tcaggaactc	ttccatggcc	cttcaacagg	cacattttatt	gaaaaacatg	2220

agtcactgca	ggcagccttc	tgacagcagt	gttgataaat	ttgtgttgag	agatgaagct	2280
actgaaccgg	gtgatcaaga	aaacaagcct	tgccgcacat	aagggtgacat	tggacagctcc	2340
actgatgatg	actctgcacc	tcttgtccat	tgtgtccgcc	ttttatctgc	ttcgtttttg	2400
ctaaccagggg	gaaaaaatgt	gctgggtccg	gacagggatg	tgagggtcag	cgtgaaggcc	2460
ctggccctca	gctgtgtggg	agcagctgtg	gccctccacc	cggaaatcttt	cttcagcaaa	2520
ctctataaag	ttcctcttga	caccacggaa	taccctgagg	aacagtatgt	ctcagacatc	2580
ttgaactaca	tcgatcatgg	agaccacag	gttcgaggag	ccactgccat	tctctgtggg	2640
accctcatct	gctccatcct	cagcaggtcc	cgcttccacg	tgggagattg	gatgggcacc	2700
attagaaccc	tcacagggaa	tacattttct	ttggcggatt	gcatttccttt	gctgcgga	2760
acactgaagg	atgagtcctc	tgttactttg	aagtttagctt	gtacagctgt	gaggaactgt	2820
gtcatgagtc	tctgcagcag	cagctacagt	gagttaggac	tgcagctgat	catcgatgtg	2880
ctgactctga	ggaacagttc	ctattggctg	gtgaggacag	agcttctgga	aacctttgca	2940
gagattgact	tcaggctggg	gagctttttg	gaggcaaaag	cagaaaaactt	acacagaggg	3000
gctcatcatt	atacagggct	tttaaaactg	caagaacgag	tgctcaataa	tgttgtcatc	3060
catttgcttg	gagatgaaga	ccccagggtg	cgacatgttg	ccgcagcatc	actaattagg	3120
cttggaccaa	agctgtttta	taaatgtgac	caaggacaag	ctgatccagt	agtggccgtg	3180
gcaagagatc	aaagcagtgt	ttacctgaaa	cttctcatgc	atgagacgca	gcctccatct	3240
catttctccg	tcagcacaat	aaccagaata	tatagaggct	ataacctact	accaagcata	3300
acagacgtca	ctatggaaaa	taacctttca	agagtatttg	cagcagtttc	tcatgaacta	3360
atcacatcaa	ccaccagagc	actcacattt	ggatgctgtg	aagctttgtg	tcttctttcc	3420
actgccttcc	cagtttgcac	ttggagttaa	ggttggcact	gtggagtgcc	tccactgagt	3480
gcctcagatg	agtctaggaa	gagctgtacc	gttgggatgg	ccacaatgat	tctgacctcg	3540
ctctcgtcag	cttgggtccc	attggatctc	tcagcccatc	aagatgcttt	gattttggcc	3600
ggaaacttgc	ttgcagccac	tgctcccaaa	tctctgagaa	gttcatgggc	ctctgaaaga	3660
gaagccaacc	cagcagccac	caagcaagag	gaggtctggc	cagccctggg	ggaccggggc	3720
ctggtgccca	tggtggagca	gctcttctct	cacctgctga	aggtgatata	catttgtgcc	3780
cacgtcctgg	atgacgtggc	tcctggacc	gcaataaagg	cagccttgcc	ttctctaaca	3840
aacccccctt	ctctaagttc	catccgacga	aagggaagg	agaaagaacc	aggagaacaa	3900
gcatctgtac	cgttgagtcc	caagaaaggc	agtgaaggcca	gtgcagcttc	tagacaatct	3960
gatacctcag	gtcctgttac	aacaagtaaa	tcctcatcac	tggggagttt	ctatcatctt	4020
ccttcatacc	tcaaatgca	tgatgtcctg	aaagctacac	acgctaacta	caaggctacg	4080
ctggtacttc	agaacagcac	ggaaaagtgt	ggagggtttc	tccgctcagc	cttggtattg	4140
ctttctcaga	tactagagct	ggccacactg	caggacattg	ggaagtgtgt	tgaagagatc	4200
ctaggatacc	tgaaatcctg	ctttagtoga	gaaccaatga	tggcaactgt	ttgtgttcaa	4260
caattgttga	agactctctt	tggcacaaa	ttggcctccc	agtttgatgg	cttatcttcc	4320
aacccccctt	agtcacaagg	ccgagcacag	cgcttgggtc	cctccagtgt	gaggcaggcc	4380
ttgtaccact	actgcttcat	ggccccgtac	accacttca	cccaggccct	cgtgacgcc	4440
agcctgagga	acatggtgca	ggcggagcag	gagaacgaca	cctcgggatg	gtttgatgtc	4500
ctccagaaag	tgtctaccca	gttgaagaca	aacctcacga	gtgtcacaaa	gaaccgtgca	4560
gataagaatg	ctattcataa	tcacattcgt	ttgtttgaac	ctcttgttat	aaaagcttta	4620
aaacagtaca	cgactacaac	atgtgtgcag	ttacagaagc	aggttttaga	tttgctggcg	4680
cagctgggtc	agttacgggt	taattactgt	cttctggatt	cagatcaggt	gtttattggc	4740
tttgtattga	aacagtttga	atacattgaa	gtgggccagt	tcagggaatc	agaggcaatc	4800
attccaaaca	tcttttctct	cttggattta	ctatcttatg	aacgctatca	ttcaaaacag	4860
atcattggaa	ttcctaaaat	cattcagctc	tgtgatggca	tcattggccag	tggaaaggag	4920
gctgtgacac	atgccatacc	ggctctgcag	cccatagtcc	acgacctctt	tgtattaaga	4980
ggaacaaaata	aagctgatgc	aggaaaagag	cttgaaaacc	aaaaagaggt	ggtgggtgtca	5040
atgttactga	gactcatcca	gtaccatcag	gtgttgagga	tgttcattct	tgtcctgcag	5100
cagtgccaca	aggagaatga	agacaagtgg	aagcgactgt	ctcgacagat	agctgacatc	5160
atcctcccaa	tgttagccaa	acagcagatg	cacattgact	ctcatgaagc	ccttggagtg	5220
ttaaatacat	tatttgagat	tttggccctc	tcctccctcc	gtccggtaga	catgctttta	5280
cggagtatgt	tcgtcactcc	aaacacaatg	gcgtccgtga	gcactgttca	actgtggata	5340
tcgggaattc	tggtccattt	gagggttctg	atttcccagt	caactgaaga	tattgttctt	5400
tctcgtattc	aggagctctc	cttctctccg	tatttaactc	cctgtacagt	aattaatagg	5460
ttaagagcat	gggacagtac	ttcaacgcta	gaagaacaca	gtgaagggaa	acaaataaag	5520
aatttgccag	aagaaacatt	ttcaaggttt	ctattacaac	tggttggtat	tcttttagaa	5580
gacattgtta	caaaacagct	gaagggtgaa	atgagtgagc	agcaacatac	tttctattgc	5640
caggaaactag	gcacactgct	aatgtgtctg	atccacatct	tcaagtctgg	aatgttccgg	5700
agaatcacag	cagctgccac	taggctgttc	cgcagtgatg	gctgtggcgg	cagtttctac	5760
accctggaca	gcttgaactt	gcgggctcgt	tccatgatca	ccaccacccc	ggccttgggt	5820
ctgctctggg	gtcagatact	gctgctgtgc	aaccacaccg	actaccgctg	gtgggcagaa	5880
gtgcagcaga	ccccgaaaag	acacagctctg	tccagcacaa	agttacttag	tccccagatg	5940
tctggagaag	aggaggattc	tgacttggca	gccaaaacttg	gaatgtgcaa	tagagaaata	6000

gtacgaagag gggctctcat tctcttctgt gattatgtct gtcagaacct ccatgactcc 6060
gagcacttaa cgtggctcat tgtaaatcac attcaagatc tgatcagcct ttcccacgag 6120
cctccagtac aggacttcat cagtgcggtt catcggaact ctgctgccag cggcctgttc 6180
atccaggcaa ttcagtctcg ttgtgaaaac ctttcaactc caaccatgct gaagaaaact 6240
cttcagtgtc tggaggggat ccatctcagc cagtccggag ctgtgctcac gctgtatgtg 6300
gacaggtctc tgtgcacccc tttccgtgtg ctggctcgca tggctgacat ccttgcttgt 6360
cgccgggtag aaatgcttct ggctgcaaat ttacagagca gcatggccca gttgccaatg 6420
gaagaactca acagaatcca ggaatacctt cagagcagcg ggctcgctca gagacaccaa 6480
aggctctatt ccctgtgga cagggttctg ctctccacca tgcaagactc acttagtccc 6540
tctcctccag tctcttccca cccgctggac ggggatgggc acgtgtcact ggaacagtg 6600
agtccggaca aagactggta cgttcatctt gtcaaatccc agtggtggac cagggtcagat 6660
tctgcaactg tggaaagtgc agagctgggt aatcggattc ctgctgaaga tatgaatgcc 6720
ttcatgatga actcggagtt caacctaaag ctgctagctc catgcttaag cctagggatg 6780
agtgaatttt ctggtggcca gaagagtgc ctttttgaag cagcccgtga ggtgactctg 6840
gcccgtgtga gcggaacggt gcagcagctc cctgctgtcc atcatgtctt ccagcccag 6900
ctgctcagca agcggcgagg ctactggagc aagttgaatg atctgtttgg ggtgctgca 6960
ctgtatcagt ccctgcccac tctggcccg gcccgtggac agtacctggt ggtggtctcc 7020
aaactgcccc gtcatttgca ccttcctcct gagaaagaga aggacattgt gaaattctgt 7080
gtggaacccc ttgaggccct gtcctggcat ttgatccatg agcagatccc gctgagctgt 7140
gatctccagc cagggtgga ctgctgctgc ctggccctgc agctgcctgg cctctggag 7200
gtggtctcct ccacagagtt tgtgaccac gectgctccc tcatctactg tgtgcacttc 7260
atcctggagg ccgttgcaagt gcagcctgga gagcagcttc ttagtccaga aagaaggaca 7320
aataccccc aaagccatcag cgaggaggag gaggaagtag atccaaacac acagaatcct 7380
aagtatatca ctgcagccga tgagatggtg gcagaaatgg tggagtctct gcagtcgtg 7440
ttggccttgg gtcataaaag gaatagcggc gtgccggcgt ttctcacgcc attgctcag 7500
aacatcatca tcagcctggc ccgctgccc cttgtcaaca gctacacacg tgtgccccca 7560
ctggtgtgga agcttggatg gtcacccaaa ccgggagggg attttggcac agcattccct 7620
gagatccccg tggagtctct ccaggaaaag gaagtcttta aggagttcat ctaccgcatc 7680
aacacactag gctggaccag tctgactcag tttgaagaaa cttgggccac cctccttgt 7740
gtcctgggtga cgcagcccct cgtgatggag caggaggaga gccaccaga agaagacaca 7800
gagaggaccc agatcaacgt cctggcctg caggccatca cctcactggt gctcagtcca 7860
atgactgttc ctgtggccg caaccagct gtaagctgct tggagcagca gcccggaaac 7920
aagcctctga aagctctcga caccaggtt gggaggaaag tgagcattat cagagggatt 7980
gtggaagca agattcaagc aatggtttca aagagagaga atattgccac ccatcattta 8040
tatcaggcat gggatcctgt cccttctctg tctccggcta ctacagtgct cctcatcagc 8100
cacgagaaag tctgtctaca gatcaacccc gagcgggagc tggggagcat tagctacaaa 8160
ctcgccagg tgtccataca ctccgtgtgg ctggggaaca gcatcacacc cctgaggagg 8220
gaggaatggg acgaggaaga ggaggaggag gccagcggcc ctgcaccttc gtcaccaccc 8280
acgtctccag tcaactccag gaaacaccgg gctggagttg acatccactc ctgttcgag 8340
tttttgcctt agttgtacag ccgctggatc ctgcccctca gctcagccag gaggaccgg 8400
gccatcctga tcagtgggt ggtcagatcc cttctagtgg tctcagactt gttcacccag 8460
cgcaaccagt ttgagctgat gtatgtgacg ctgacagaaac tgcgaagggt gcacccttca 8520
gaagacgaga tcctcgctca gtacctggtg cctgccacct gcaaggcagc tggcgtctt 8580
gggatggaca aggcctggc ggagcctgtc agccgcctgc tggagagcac gctcaggagc 8640
agccacctgc ccagcagggt tggagccctg cacggcgtcc tctatgtgct ggagtgcgac 8700
ctgctggagc acactgccaa gcagctcatc ccggtcatca gcgactatct cctctccaac 8760
ctgaaaggga tcgcccactg cgtgaacatt cacagccagc agcacgtact ggtcatgtgt 8820
gccactgcgt tttacctcat tgagaactat cctctggagc tagggccgga attttcagca 8880
tcaataatac agatgtgtgg ggtgatgctg tctggaagtg aggagtcac ccctccatc 8940
atttaccact gtgcccctcag aggcctggag cgcctcctgc tctctgagca gctctccgc 9000
ctggtatgca aatcgtgtg caagctgagt gtggacagag tgaacgtgca cagcccgcac 9060
cgggccatgg cggctctggg cctgatgctc acctgcatgt acacaggaaa ggagaaagt 9120
agtccgggta gaacttcaga ccctaactct gcagccccc agacgagtc agtgattgtt 9180
gctatggagc gggatctgt tctttttgat aggatcagga aaggctttcc ttgtgaagcc 9240
agagtgggtg ccaggatcct gcccagttt ctgacgact tcttccacc ccaggacatc 9300
atgaacaaa tcatcgagga gtttctgtcc aaccagcagc cataccccc gttcatggcc 9360
accgtggtgt ataaggtgtt tcagactctg cacagcaccg ggcagtcgtc catggtccgg 9420
gactgggtca tgctgtccct ctccaacttc acgcagaggg ccccggctgc catggccacg 9480
tggagcctct cctgcttctt tgtcagcgcg tccaccagcc cgtgggtcgc ggcgactctc 9540
ccacatgtca tcagcaggat gggcaagctg gagcaggtgg acgtgaacct tttctgcctg 9600
gtcggccacag acttctacag acaccagata gaggaggagc tcgaccgcag ggccttccag 9660
tctgtgcttg aggtggttgc agcccagga agcccatatc accggctgct gacttgttta 9720
cgaaatgtcc acaaggtcac cacctgctga gcgccatggt gggagagact gtgaggcgcc 9780

ES 2 808 561 T3

agctggggcc	ggagcctttg	gaagtctgtg	cccttggtgc	ctgcctccac	cgagccagct	9840
tgttccctat	gggcttccgc	acatgccgcg	ggcggccagg	caacgtgcgt	gtctctgcca	9900
tgtggcagaa	gtgctctttg	tggcagtgcc	caggcaggga	gtgtctgcag	tcctggtggg	9960
gctgagcctg	aggccttcca	gaaagcagga	gcagctgtgc	tgcaccccat	gtgggtgacc	10020
aggtcctttc	tcctgatagt	cacctgctgg	ttgttgccag	gttgccagctg	ctcttgcatc	10080
tgggccagaa	gtcctccctc	ctgcaggctg	gctgttggcc	cctctgctgt	cctgcagtag	10140
aaggtgccgt	gagcaggctt	tgggaacact	ggcctgggtc	tccttggtgg	ggtgtgcatg	10200
ccacgccccg	tgtctggatg	cacagatgcc	atggcctgtg	ctgggcccagt	ggctgggggt	10260
gctagacacc	cggcaccatt	ctcccttctc	tcttttcttc	tcaggattta	aaattttaatt	10320
atatcagtaa	agagattaat	tttaacgaac	tctttctatg	cccgtgtaaa	gtatgtgaat	10380
cgcaaggcct	gtgctgcatg	cgacagcgtc	cggggtgggtg	gacagggccc	ccggccacgc	10440
tcctctcctt	gtagccactg	cgatagccct	cctgagcacc	cgctgacatt	tccgttgtag	10500
atgttctctg	ttatgcattc	acaagggtgac	tgggatgtag	agaggcggtta	gtggggcagg	10560
ggccacagca	ggactgagga	caggccccca	ttatcctagg	ggtgcgctca	actgcagccc	10620
ctcctcctcg	ggcacagacg	actgtcgttc	tccacccacc	agtcaggggac	agcagcctcc	10680
ctgtcactca	gctgagaagg	ccagccctcc	ctggctgtga	gcagccctcca	ctgtgtccag	10740
agacatgggc	ctccactcc	tgttccttgc	tagccctggg	gtggcgctctg	cctaggagct	10800
ggctggcagg	tgttgggacc	tgtgctcca	tggatgcatg	ccctaagagt	gtcactgagc	10860
tgtgttttgt	ctgagcctct	ctcggtaaac	agcaaaagctt	ggtgtccttg	caactgttagt	10920
gctgacccc	agcatccctt	ctgccccctg	tccagctgac	atcttgacag	gtgacccttc	10980
ttagttagga	gagtgcagat	ctgtgctcat	cggagactgc	cccacggccc	tgtcagagcc	11040
gccactccta	tccccaggac	aggtccctgg	accagcctcc	tgtttgcagg	cccagaggag	11100
ccaagtccatt	aaaaatggaag	tggattcttg	atggccgggc	tgtgctgat	gtaggagctg	11160
gatttgggag	ctctgcttgc	gcactggctg	tgagacgagg	caggggctct	gcttctcttc	11220
ccctagaggc	gagccaggca	aggttggcga	ctgtcatgtg	gcttggtttg	gtcatgcccg	11280
tcgatgtttt	gggtattgaa	tgtggttaagt	ggaggaaatg	ttggaactct	gtgcaggtgc	11340
tgccttgaga	cccccaagct	tccacctgtc	cctctcctat	gtggcagctg	gggagcagct	11400
gagatgtgga	cttgtatgct	gcccacatac	gtgaggggga	gctgaaaggg	agcccccttc	11460
caaaggggag	ccctcctctg	agcagcctct	gccaggcctg	tatgaggctt	ttcccaccag	11520
ctcccaacag	aggcctcccc	cagccaggac	cacctcgtcc	tcgtggcggg	gcagcaggag	11580
cggtagaaa	gggtccgatg	tttgaggagg	cccttaaggg	aagctactga	attataaac	11640
gacacatcta	caccattctt	ccgtatttgt	tgggggctcc	tgtttctcat	cctagctttt	11700
tcctggaaaa	gcccgttaga	aggtttggga	acgaggggaa	agttctcaga	actgttgctg	11760
ctccccaccc	gcctcccggc	tcccccgcat	gttatgtcag	cagctctgag	acagcagtat	11820
cacaggccag	atgttgttcc	tggctagatg	tttacatttg	taagaataaa	caactgtgaat	11880
gtaaaacaga	gccattccct	tggaaatgcat	atcgctgggc	tcaacataga	gtttgtcttc	11940
ctcttggtta	cgacgtgatc	taaaaccagtc	cctagcaagg	ggctcagaac	accccgtctt	12000
ggcagtaggt	gtccccacc	cccaaagacc	tgcctgtgtg	ctccggagat	gaatatgagc	12060
tcattagtaa	aaatgacttc	acccaacgat	atacataaag	tatccatgca	tgtgcataa	12120
gacacatcta	taattttaca	cacacacctc	tcaagacgga	gatgcattgg	ctctaagagt	12180
gcccgtgtcg	gttcttctctg	gaagttgact	ttccttagac	cgcgcaggtc	aagttagccg	12240
cgtgacggac	atccaggcgt	gggacgttgt	cagggcaggg	ctcattcatt	gcccactagg	12300
atcccactgg	cgaagatggt	ctccatatca	gctctctgca	gaaggaggga	agactttatc	12360
atgttcccta	aaatctgtgg	caagcaccca	tcgtattatc	caaattttgt	tgcaaatgtg	12420
attaatttgg	ttgtcaagtt	ttgggggtgg	gctgtgggga	gattgctttt	gttttctctg	12480
tggtaatatc	gggaaagatt	ttaatgaaac	cagggttagaa	ttgtttggca	atgcactgaa	12540
gcgtgtttct	ttcccaaaat	gtgcctccct	tccgctgcgg	gcccagctga	gtctatgtag	12600
gtgatgtttc	cagctgcca	gtgctctttg	ttactgtcca	ccctcatttc	tgccagcgca	12660
tgtgtccttt	caaggggaaa	atgtgaagct	gaacccctc	cagacaccca	gaatgtagca	12720
tctgagaagg	ccctgtgccc	taaaggacac	ccctcgcccc	catcttcatt	gagggggtca	12780
tttcagagcc	ctcggagcca	atgaacagct	cctcctcttg	gagctgagat	gagccccacg	12840
tggagctcgg	gacggatagt	agacagcaat	aactcgggtg	gtggccgcct	ggcaggtgga	12900
acttctctcc	gttgccgggt	ggagtggagt	tagttctgtg	tgtctggtgg	gtggagtcag	12960
gcttctcttg	ctacctgtga	gcacctcttc	cagcagacat	cctcatcggt	ctttgtccct	13020
cccccgcttc	ctccctctgc	ggggaggacc	cgggaccaca	gctgctggcc	agggtagact	13080
tggagctgtc	ctccagaggg	gtcacgtgta	ggagtggaga	gaagggaagt	cctgagagct	13140
gctgaggggc	cttgagagac	tcaggatggc	tcagacgagg	acactcgctt	gcccggcctg	13200
gccctcctgg	gaaggaggga	gctgctcaga	atgccgcatg	acaactgaag	gcaacctgga	13260
aggttcaggg	cccgtctctc	ccccatgtgc	ctgtcacgct	ctggtgcagt	caaagggaac	13320
ccttccccctc	agttgtttct	aagagcagag	tctcccgctg	caatctgggt	ggtaactgcc	13380
agccttgagg	gategtggcc	aacgtggacc	tgcctacgga	gggtgggctc	tgacccaagt	13440
ggggcctcct	tgcccaggtc	tactgctttt	gcaccgtggg	cagaggggact	gtcagctgag	13500
cttgagctcc	cctggagcca	gcagggctgt	gatgggcgag	tcccggagcc	ccaccagac	13560
ctgaatgctt	ctgagagcaa	aggggaaggac	tgacgagaga	tgtatatatta	attttttaac	13620
tgctgcaaac	attgtacatc	caaattaaag	ggaaaaaatg	gaaacatca	at	13672

<210> 2

5

<211> 3144

<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
 20      25      30
Gln Gln Gln Gln Gln Gln Gln Gln 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 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 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 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	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	

ES 2 808 561 T3

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	

ES 2 808 561 T3

				1925										1930									1935								
Pro	Val	Gln	Asp	Phe	Ile	Ser	Ala	Val	His	Arg	Asn	Ser	Ala	Ala	Ser																
			1940											1945																	
Gly	Leu	Phe	Ile	Gln	Ala	Ile	Gln	Ser	Arg	Cys	Glu	Asn	Leu	Ser	Thr																
		1955												1960																	
Pro	Thr	Met	Leu	Lys	Lys	Thr	Leu	Gln	Cys	Leu	Glu	Gly	Ile	His	Leu																
		1970												1975																	
Ser	Gln	Ser	Gly	Ala	Val	Leu	Thr	Leu	Tyr	Val	Asp	Arg	Leu	Leu	Cys																
1985					1990										1995																
Thr	Pro	Phe	Arg	Val	Leu	Ala	Arg	Met	Val	Asp	Ile	Leu	Ala	Cys	Arg																
		2005												2010																	
Arg	Val	Glu	Met	Leu	Leu	Ala	Ala	Asn	Leu	Gln	Ser	Ser	Met	Ala	Gln																
		2020												2025																	
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																	
Pro	Asp	Lys	Asp	Trp	Tyr	Val	His	Leu	Val	Lys	Ser	Gln	Cys	Trp	Thr																
		2100												2105																	
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	Arg	Glu	Val	Thr	Leu	Ala			</														

ES 2 808 561 T3

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

ES 2 808 561 T3

2930	2935	2940
Pro Ala Ala Pro Asp Ser Glu Ser Val Ile Val Ala Met Glu Arg Val		
2945	2950	2955
Ser Val Leu Phe Asp Arg Ile Arg Lys Gly Phe Pro Cys Glu Ala Arg		2960
	2965	2970
Val Val Ala Arg Ile Leu Pro Gln Phe Leu Asp Asp Phe Phe Pro Pro		2975
	2980	2985
Gln Asp Ile Met Asn Lys Val Ile Gly Glu Phe Leu Ser Asn Gln Gln		2990
	2995	3000
Pro Tyr Pro Gln Phe Met Ala Thr Val Val Tyr Lys Val Phe Gln Thr		3005
	3010	3015
Leu His Ser Thr Gly Gln Ser Ser Met Val Arg Asp Trp Val Met Leu		3020
3025	3030	3035
Ser Leu Ser Asn Phe Thr Gln Arg Ala Pro Val Ala Met Ala Thr Trp		3040
	3045	3050
Ser Leu Ser Cys Phe Phe Val Ser Ala Ser Thr Ser Pro Trp Val Ala		3055
	3060	3065
Ala Ile Leu Pro His Val Ile Ser Arg Met Gly Lys Leu Glu Gln Val		3070
	3075	3080
Asp Val Asn Leu Phe Cys Leu Val Ala Thr Asp Phe Tyr Arg His Gln		3085
	3090	3095
Ile Glu Glu Glu Leu Asp Arg Arg Ala Phe Gln Ser Val Leu Glu Val		3100
3105	3110	3115
Val Ala Ala Pro Gly Ser Pro Tyr His Arg Leu Leu Thr Cys Leu Arg		3120
	3125	3130
Asn Val His Lys Val Thr Thr Cys		3135
	3140	

<210> 3

<211> 40

5 <212> ADN

<213> Secuencia artificial

<220>

<223> construcción sintética

10 <400> 3

ugcagcugau caucgaugug cugacccuga ggaacaguuc 40

<210> 4

<211> 40

15 <212> ADN

<213> Secuencia artificial

<220>

<223> construcción sintética

20 <400> 4

gaacuguucc ucagggucag cacaucgaug aucagcugca 40

	<210> 5	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
5		
	<220>	
	<223> construcción sintética	
	<400> 5	
10	tgtgctgact ctgaggaaca g	21
	<210> 6	
	<211> 21	
	<212> ADN	
15	<213> Secuencia artificial	
	<220>	
	<223> construcción sintética	
20	<400> 6	
	ugugcugacu cugaggaaca g	21
	<210> 7	
	<211> 21	
25	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> construcción sintética	
30	<400> 7	
	cuguuccuca gagucagcac a	21
	<210> 8	
35	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	

<220>
 <223> construcción sintética

 <400> 8
 5 catacctcaa actgcatgat g 21

 <210> 9
 <211> 21
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> construcción sintética

 15 <400> 9
 cauaccucaaa acugcaugau g 21

 <210> 10
 <211> 21
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> construcción sintética
 25
 <400> 10
 caucaugcag uuugaggau g 21

 <210> 11
 30 <211> 21
 <212> ADN
 <213> Secuencia artificial

 <220>
 35 <223> construcción sintética

 <400> 11

	gcctgcagag ccggcggcct a	21
	<210> 12	
	<211> 21	
5	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> construcción sintética	
10	<400> 12	
	gccugcagag ccggcggccu a	21
	<210> 13	
15	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
20	<223> construcción sintética	
	<400> 13	
	uaggccgccg gcucugcagg c	21
25	<210> 14	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> construcción sintética	
	<400> 14	
35	acagagtttg tgaccacgc c	21
	<210> 15	
	<211> 21	

<212> ADN
 <213> Secuencia artificial

 <220>
 5 <223> construcción sintética

 <400> 15
 acagaguug ugacccacgc c 21

 10 <210> 16
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> construcción sintética

 <400> 16
 ggcguggguc acaaacucug u 21
 20
 <210> 17
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> construcción sintética

 <400> 17
 30 tccctcatct actgtgtgca c 21

 <210> 18
 <211> 21
 <212> ADN
 35 <213> Secuencia artificial

 <220>

	<223> construcción sintética	
	<400> 18	
5	uccucaucu acugugugca c	21
	<210> 19	
	<211> 21	
	<212> ADN	
10	<213> Secuencia artificial	
	<220>	
	<223> construcción sintética	
	<400> 19	
15	gugcacacag uagaugaggg a	21
	<210> 20	
	<211> 5	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> construcción sintética	
25	<400> 20	
	ugugc	5
	<210> 21	
	<211> 7	
30	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> construcción sintética	
35	<400> 21	
	gcacauc	7

	<210> 22	
	<211> 5	
	<212> ADN	
	<213> Secuencia artificial	
5		
	<220>	
	<223> construcción sintética	
	<400> 22	
10	guugc	5
	<210> 23	
	<211> 7	
	<212> ADN	
15	<213> Secuencia artificial	
	<220>	
	<223> construcción sintética	
20	<400> 23	
	ggaagag	7
	<210> 24	
	<211> 21	
25	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> construcción sintética	
30		
	<400> 24	
	ugugcugacc cugaggaaca g	21
	<210> 25	
35	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	

<220>
 <223> construcción sintética

 <400> 25
 5 guuccucagg gucagcacau c 21

 <210> 26
 <211> 21
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> construcción sintética

 15 <400> 26
 ugugcugacc cugaggaaaa g 21

 <210> 27
 <211> 21
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> construcción sintética
 25
 <400> 27
 uuuccucagg gucagcacau c 21

 <210> 28
 30 <211> 21
 <212> ADN
 <213> Secuencia artificial

 <220>
 35 <223> construcción sintética

 <400> 28

	ugugcugacc cugaggaaaa g	21
	<210> 29	
	<211> 21	
5	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> construcción sintética	
10	<400> 29	
	guuccucagg gucagcacau c	21
	<210> 30	
15	<211> 42	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
20	<223> construcción sintética	
	<400> 30	
	gcgtaatacg actcactata ggaacagtat gtctcagaca tc	42
25	<210> 31	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> construcción sintética	
	<400> 31	
35	uucgaaguau uccgcguacg u	21
	<210> 32	
	<211> 42	

<212> ADN

<213> Secuencia artificial

<220>

5 <223> construcción sintética

<400> 32

gcgtaatacg actcactata ggacaagcct aattagtgat gc 42

10 <210> 33

<211> 21

<212> ADN

<213> Secuencia artificial

15 <220>

<223> construcción sintética

<400> 33

gaacagtatg tctcagacat c 21

20

REIVINDICACIONES

1. Un agente de iARN para su uso en un método de tratamiento de un sujeto que tiene o está en riesgo de tener una enfermedad caracterizada o provocada por una proteína mutante de ganancia de función, comprendiendo el método administrar al sujeto una cantidad eficaz del agente de iARN que silencia un polimorfismo alélico dentro de un gen que codifica una proteína mutante de forma que ocurra la interferencia específica de secuencia de un gen dando como resultado un tratamiento eficaz de la enfermedad, en donde el agente de iARN silencia una región polimórfica que es distinta de la mutación de la región ampliada de CAG en el gen que codifica la proteína mutante, en donde la enfermedad es enfermedad de Huntington y la proteína mutante es la proteína huntingtina.
2. El agente de iARN para el uso de la reivindicación 1, en donde el agente de iARN silencia un polimorfismo de un solo nucleótido.
3. El agente de iARN para el uso de la reivindicación 1 o 2, en donde el agente de iARN silencia un polimorfismo alélico seleccionado del grupo que consiste en P1-P5 o del grupo que consiste en P6-P43.
4. El agente de iARN para el uso de una cualquiera de las reivindicaciones 1 a 3, en donde el agente de iARN silencia el polimorfismo alélico P27.
5. El agente de iARN para el uso de una cualquiera de las reivindicaciones 1 a 3, en donde el agente de iARN silencia el polimorfismo alélico P5.
6. El agente de iARN para el uso de una cualquiera de las reivindicaciones 1 a 3, en donde el agente de iARN silencia el polimorfismo alélico P17.
7. El agente de iARN para el uso de cualquier reivindicación precedente, en donde el agente de iARN comprende una primera y una segunda cadena que contiene cada una 16-25 nucleótidos, en donde la primera cadena es homóloga a una región del gen que codifica la proteína mutante de ganancia de función y la secuencia de nucleótidos de la proteína mutante de ganancia de función comprende un polimorfismo alélico, y en donde la segunda cadena es complementaria a la primera cadena.
8. El agente de iARN para el uso de cualquier reivindicación precedente, en donde el agente de iARN es una construcción de expresión manipulada para expresar ARNhp.
9. El agente de iARN para el uso de cualquier reivindicación precedente, en donde el agente de iARN es una construcción de expresión seleccionada de vectores retrovirales, casetes de expresión lineal, plásmidos, vectores virales y vectores derivados de virus.
10. El agente de iARN para el uso de la reivindicación 8 o 9, en donde la construcción de expresión comprende un promotor de la ARN polimerasa II o un promotor de la ARN polimerasa III.
11. El agente de iARN para el uso de la reivindicación 10, en donde la secuencia promotora de ARN III es el promotor de ARNnp U6 o promotor H1.
12. El agente de iARN para el uso de cualquiera de las reivindicaciones 8 a 11, en donde la construcción de expresión comprende un ácido nucleico aislado que codifica una molécula de ácido nucleico con una primera secuencia de 16-25 nucleótidos homóloga a un polimorfismo alélico dentro de la proteína huntingtina mutante.
13. El agente de iARN para el uso de cualquier reivindicación precedente, en donde el agente de iARN comprende una célula hospedadora.
14. El agente de iARN para el uso de la reivindicación 13, en donde la célula hospedadora es una célula de mamífero.
15. El agente de iARN para el uso de la reivindicación 13 o 14, en donde la célula hospedadora es una célula de mamífero no humana.
16. El agente de iARN para el uso de la reivindicación 13 o 14, en donde la célula hospedadora es una célula humana.
17. Una construcción de expresión que comprende un ácido nucleico aislado que codifica una molécula de ácido nucleico con una primera secuencia de 16-25 nucleótidos homóloga a un polimorfismo alélico dentro del gen que codifica una proteína huntingtina mutante, en donde el polimorfismo alélico se selecciona del grupo que consiste en P1-P5 o se selecciona del grupo que consiste en P6-P43.
18. La construcción de expresión de la reivindicación 17, en donde el polimorfismo alélico es P27.
19. La construcción de expresión de la reivindicación 17, en donde el polimorfismo alélico es P5.
20. La construcción de expresión de la reivindicación 17, en donde el polimorfismo alélico es P17.

21. La construcción de expresión de cualquiera de las reivindicaciones 17 a 20, en donde la construcción de expresión se selecciona de vectores retrovirales, casetes de expresión lineal, plásmidos, vectores virales y vectores derivados de virus.

1 TTGCTGTGTG AGGCAGAAACC TGGGGGGGGCA GGGGGGGGGCT GGTCCCTGG CCAGCCATTG
 61 GCAGAGTCCG CAGGETAGGG CTGTCAATCA TGTGGCCCG CGTGGCCCGG CCTCCGCCGG
 121 CGCGGCCCCG CCTCCGCCGG CGCAGTCTG GGACGCAAG CGCCGTGGG GCTGCCGGGA
 181 CGGGTCCAAG ATGGACGGCC GCTCAGGTTT TGTCTTACC TCGGGCCCG AGCCCCATT
 241 ATTGCCCCGG TGTGAGCGG CGCCGCGAGT CGGCCCGAGG CCTCCGGGG CTGCCGTGCC
 301 GGGCGGGAGA CCGCCCTGCG GACCCCTGGAA AAGCTGATGA AGGCTTCGA GTCCCTCAAG
 361 TCCTTCCAGC AGCAGCAGCA GCAGCAGCAG CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG
 421 CAGCAGCAGC AACAGCCGCC ACCGCCCGG CGCGCGCGC CGCCTCCTCA GCTTCCTCAG
 481 CGCCCGCCG AGGCACAGCC GCTGCTGCCT CAGCGCAGC CGCCCGCGC GCCGCCCGCG
 541 CGCCACCCG GCCCGGCTGT GGCTGAGGAG CCGCTGCACC GACCAAGAA AGAATTTCA
 601 GCTACCAAGA AAGACCGTGT GAATCATTT GTGAAAACAT AGTGGCACAG
 661 TCTGTCAGAA ATTCTCCAGA ATTTAGAAA CTTCTGGGA TCGCTATGGA ACTTTTCTG
 721 CTGTGCAGTG ATGACGCAGA GTCAGATGTC AGGATGGTGG CTGACGAATG CCTCAACAAA
 781 GTTATCAAAG CTTTGATGGA TTCTAATCTT CCAAGGTAC AGCTCGAGCT CTATAAGGAA
 841 ATTAAAAAGA ATGGTGCCCC TCGGAGTTT CGTGCTGCCC TGTGGAGGTT TGCTGAGCTG
 901 GCTCACCTGG TTCGGCCTCA GAAATGCAGG CCTTACCTGG TGAACCTTCT GCCGTGCCGT
 961 ACTCGAACA GCAAGAGACC CGAAGAATCA GTCCAGGAGA CCTTGGCTGC AGCTGTTCCC
 1021 AAAATTATGG CTTCTTTTGG CAATTTTGCA AATGACAATG AAATTAAGGT TTTGTTAAAG
 1081 GCCTTCATAG CGAACCTGAA GTCAAGCTCC CCCACATTC GCGGGACAGC GGCTGGATCA
 1141 GCAGTGAGCA TCTGCCAGCA CTCAGAAGG ACACAATATT TCTATAGTTG GCTACTAAAT
 1201 GTGCTCTTAG GCTTACTCGT TCCTGTCGAG GATGAACACT CCACTCTGCT GATTCTTGGC

FIG.1A

FIG.1B

1261 GTGCTGCTCA CCTGAGGTA TTGTGTGCC TTGCTGCAGC AGCAGGTCAA GGACACAAGC
 1321 CTGAAAGGCA GCTTCGGAGT GACAAGGAAA GAAATGGAAG TCTCTCCTTC TGCAGAGCAG
 1381 CTTGTCCAGG TTTATGAACT GACGTTACAT CATAACACAGC ACCAAGACCA CAATGTGTGTG
 1441 ACCGGAGCCC TGGAGCTGTT GCAGCAGCTC TTCAGAACGC CTCACACCCGA GCTTCTGCAA
 1501 ACCCTGACCG CAGTCGGGGG CATTGGGCAG CTCACGGCTG CTAAGGAGGA GTCTGGTGGC
 1561 CGAAGCCGTA GTGGGAGTAT TGTGGAACCTT ATAGCTGGAG GGGGTTCTTC ATGCAGCCCT
 1621 GTCCTTTCAA GAAACAAAA AGGCAAAAGTG CTCTTAGGAG AAGAAGAAGC CTTGGAGGAT
 1681 GACTCTGAAT CGAGATCGGA TGTACGACAG TCTGCCCTTAA CAGCCTCAGT GAAGGATGAG
 1741 ATCAGTGGAG AGCTGGCTGC TTCTTCAGGG GTTTCACATC CAGGGTCAGC AGGTCATGAC
 1801 ATCATCACAG AACAGCCACG GTCACAGCAC ACACTGCAGG CGGACTCAGT GGATCTGGCC
 1861 AGCTGTGACT TGACAAAGCTC TGCCACTGAT GGGGATGAGG AGGATATCTT GAGCCACAGC
 1921 TCCAGCCAGG TCAGCGCCGT CCCATCTGAC CCTGCCATGG ACCTGAATGA TGGGACCCAG
 1981 GCCTCGTCGC CCATCAGCGA CAGCTCCCAG ACCACCACCG AAGGGCCTGA TTCAGCTGTT
 2041 ACCCCTTCAG ACAGTTCTGA AATTGTGTTA GACGGTACCG ACAACCAGTA TTTGGGCCCTG
 2101 CAGATTGGAC AGCCCCAGGA TGAAGATGAG GAAGCCACAG GTATTCTTCC TGATGAAGCC
 2161 TCGGAGGCCT TCAGGAACTC TTCCATGGCC CTTCAACACAG CACATTTATT GAAAAACATG
 2221 AGTCACTGCA GGCAGCCTTC TGACAGCAGT GTTGATAAAT TTGTGTTGAG AGATGAAGCT
 2281 ACTGAACCGG GTGATCAAGA AAACAAGCCT TGCCGCATCA AAGGTGACAT TGGACAGTCC
 2341 ACTGATGATG ACTCTGCACC TCTTGTCCAT TGTGTCGCC TTTTATCTGC TTCGTTTTTG
 2401 CTAACAGGGG GAAAAAATGT GCTGTTCCG GACAGGGATG TGAGGGTCAG CGTGAAGGCC
 2461 CTGGCCCTCA GCTGTGTGG AGCAGCTGTG GCCCTCCACC CGGAATCTTT CTCAGCAAA

2521 CTCTATAAAG TTCCCTCTTGA CACCACGGAA TACCCTGAGG AACAGTATGT CTCAGACATC
 2581 TTGAACCTACA TCGATCATGG AGACCCACAG GTTCGAGGAG CCACTGCCAT TCTCTGTGG
 2641 ACCCTCATCT GCTCCATCCT CAGCAGGTCC CGCTTCCACG TGGGAGATTG GATGGGCACC
 2701 ATTAGAACC TCACAGGAAA TACATTTTCT TTGGCGGATT GCATTCCCTT GCTGCGGAAA
 2761 AACTGAAGG ATGAGTCTTC TGTACTTGC AAGTTAGCTT GTACAGCTGT GAGAACTGT
 2821 GTCATGAGTC TCTGCAGCAG CAGCTACAGT GAGTTAGGAC TGCAGCTGAT CATCGATGTG
 2881 CTGACTTCTGA GGAACAGTTC CTATTGGCTG GTGAGGACAG AGCTTCTGGA AACCTTGCA
 2941 GAGATTGACT TCAGGCTGGT GAGCTTTTTG GAGGCCAAAAG CAGAAAACTT ACACAGAGG
 3001 GCTCATCATT ATACAGGGCT TTTAAAACTG CAAGAACGAG TGCTCAATAA TGTGTGATC
 3061 CATTGTCTTG GAGATGAAGA CCCCAGGGTG CGACATGTTG CCGCAGCATC ACTAATTAGG
 3121 CTTGTCCCAA AGCTGTTTTA TAAATGTGAC CAAGGACAAG CTGATCCAGT AGTGGCCGTG
 3181 GCAAGAGATC AAAGCAGTGT TTACCTGAAA CTTCATGTC ATGAGACGCA GCCTCCATCT
 3241 CATTTCTCG TCAGCACAAAT AACCAGAATA TATAGAGGCT ATAACTTACT ACCAAGCATA
 3301 ACAGACGTCA CTATGGAAAA TAACCTTTCA AGAGTTATTG CAGCAGTTTC TCATGAACTA
 3361 ATCACATCAA CCACCAGAGC ACTCACATTT GGATGCTGTG AAGCTTTGTG TCTTCTTTCC
 3421 ACTGCCCTCC CAGTTTGCAT TTGGAGTTTA GGTGGGCACT GTGGAGTGCC TCCACTGAGT
 3481 GCCTCAGATG AGTCTAGGAA GAGCTGTACC GTTGGGATGG CCACAATGAT TCTGACCCTG
 3541 CTCTCGTCAG CTTGGTTCCC ATTGGATCTC TCAGCCCCATC AAGATGCTTT GATTTTGGCC
 3601 GAAAACTTGC TTGCAGCCAG TGCTCCCCAA TCTCTGAGAA GTTCATGGGC CTCTGAAGAA
 3661 GAAGCCAACC CAGCAGCCAC CAAGCAAGAG GAGGTCTGGC CAGCCCTGGG GGACCGGGCC
 3721 CTGGTGCCCA TGGTGGAGCA GCTCTTCTCT CACCTGCTGA AGGTGATTAA CATTTGTGCC

FIG.1C

FIG.1D

3781 CACGTCCTGG ATGACGTGGC TCCTGGACCC GCAATAAAGG CAGCCTTGCC TTCTCTAACA
 3841 AACCCCCCTT CTCTAAGTCC CATCCGACGA AAGGGHAGG AGAAAGAACC AGGAGAACAA
 3901 GCATCTGTAC CGTTGAGTCC CAAGAAAGGC AGTGAGGCCA GTGCAGCTTC TAGACAATCT
 3961 GATACCTCAG GTCCTGTTAC AACAAAGTAA TCCTCATCAC TGGGAGTTT CTATCATCTT
 4021 CCTTCATACC TCAAACTGCA TGATGCTCTG AAAGCTACAC ACGCTAACTA CAAGGTCACG
 4081 CTGGATCTTC AGAACAGCAC GGAAGAGTTT GGAGGGTTT TCCGCTCAGC CTTGGATGTT
 4141 CTTTCTCAGA TACTAGAGCT GGCCACACTG CAGGACATTG GGAAGTGTGT TGAAGAGATC
 4201 CTAGGATACC TGAATCCCTG CTTTAGTCTG GAACCAATGA TGGCAACTGT TTGTGTTCAA
 4261 CAATTGTTGA AGACTCTCTT TGGCACAAAC TTGGCCTCCC AGTTTGATGG CTTATCTTCC
 4321 AACCCACAGCA AGTCACAAGG CCGAGCACAG CGCCTTGGCT CCTCCAGTGT GAGGCCAGGC
 4381 TTGTACCACT ACTGCTTCAT GGCCCCGTAC ACCCACTTCA CCCAGGCCCT CGCTGACGCC
 4441 AGCCTGAGGA ACATGGTGCA GGCGGAGCAG GAGAACGACA CCTCGGGATG GTTTGATGTC
 4501 CTCCAGAAAAG TGCTTACCCA GTTGAAGACA AACCTCACGA GTGTCACAAA GAACCGTGCA
 4561 GATAAGAATG CTATTCATAA TCACATTCTG TTGTTTGAAC CTCTTGTTAT AAAAGCTTTA
 4621 AACACAGTACA CGACTACAAC ATGTGTGCAG TTACAGAAGC AGGTTTTAGA TTTGCTGGCG
 4681 CAGCTGGTTC AGTTACGGGT TAATTACTGT CTTCTGGATT CAGATCAGGT GTTATTGGC
 4741 TTTGTATTGA AACAGTTTGA ATACATTGAA GTGGGCCAGT TCAGGGAATC AGAGGCAATC
 4801 ATTCCAAAACA TCTTTTCTT CTTGGTATTA CTATCTTATG AACGCTATCA TTCAAAAACAG
 4861 ATCATTTGGA TTCCATAAAT CATTCAGCTC TGTGATGGCA TCATGGCCAG TGAAGGAAG
 4921 GCTGTGACAC ATGCCATACC GGCTCTGCAG CCCATAGTCC ACGACCTCTT TGTATTAGA
 4981 GGAACAAATA AAGCTGATGC AGGAAAAGAG CTTGAAAACC AAAAAGAGGT GGTGGTGTCA

FIG.1E

5041 ATGTTACTGA GACTCATCCA GTACCATCAG GTGTTGGAGA TGTTCATTCT TGTCTGCAG
 5101 CAGTGCCACA AGGAGAATGA AGACAAGTGG AAGCAGCTGT CTCGACAGAT AGCTGACATC
 5161 ATCCTCCCAA TGTTAGCCAA ACAGCAGATG CACATTGACT CTCATGAAGC CCTTGGAGTG
 5221 TTAATAACAT TATTGAGAT TTTGGCCCCCT TCCTCCCTCC GTCCGGTAGA CATGCTTTTA
 5281 CGGAGTATGT TCGTCACTCC AAACACAATG GCGTCCGTGA GCACTGTTCA ACTGTGGATA
 5341 TCGGGAAATC TGGCCATTTT GAGGGTTCTG ATTTCCAGT CAACTGAAGA TATTGTTCTT
 5401 TCTCGTATTC AGGAGCTCTC CTTCTCTCCG TATTAAATCT CCTGTACAGT AATTAATAGG
 5461 TTAAGAGATG GGGACAGTAC TTCAACGCTA GAAGAACACA GTGAAGGGAA ACAATAAAG
 5521 AATTGGCCAG AAGAAACATT TTCAAGGTTT CTATTACAAC TGGTTGGTAT TCTTTTAGAA
 5581 GACATTGTTA CAAAACAGCT GAAGGTGGAA ATGAGTGAGC AGCAACATAC TTTCTATTGC
 5641 CAGGAAGTAG GCACACTGCT AATGTGCTG ATCCACATCT TCAAGTCTGG AATGTTCCGG
 5701 AGAATCACAG CAGCTGCCAC TAGGCTGTTT CGCAGTGATG GCTGTGGCGG CAGTTTCTAC
 5761 ACCCTGGACA GCTTGAACCT GCGGGCTCGT TCCATGATCA CCACCCACCC GGCCCTGGTG
 5821 CTGCTCTGGT GTCAGATACT GCTGCTTGTC AACCACACCG ACTACCGCTG GTGGGCAGAA
 5881 GTGCAGCAGA CCCCAGAAAAG ACACAGTCTG TCCAGCACAA AGTTACTTAG TCCCCAGATG
 5941 TCTGGAGAAG AGGAGGATTC TGACTTGGCA GCCAAACTTG GAATGTGCAA TAGAGAAAATA
 6001 GTACGAAGAG GGGCTCTCAT TCTCTTCTGT GATTATGTCT GTCAGAACCT CCATGACTCC
 6061 GAGCACTTAA CGTGGCTCAT TGTAATCAC ATTCAGATC TGATCAGCCT TTCCCACGAG
 6121 CCTCCAGTAC AGGACTTCAT CAGTGCCGTT CATCGGAAC CTGCTGCCAG CGGCCCTGTT
 6181 ATCCAGGCAA TTCAGTCTCG TTGTGAAAAC CTTTCAACTC CAACCATGCT GAAGAAAAC
 6241 CTTCAGTGCT TGGAGGGGAT CCATCTCAGC CAGTCGGGAG CTGTGCTCAC GCTGTATGTG

FIG.1F

6301 GACAGGCTTC TGTGCACCCC TTTCGCTGTG CTGGCTCGCA TGTCGACAT CCTTGCTTGT
 6361 CGCCGGGTAG AAATGCTTCT GGCTGCAAAAT TTACAGAGCA GCATGGCCCA GTTGCCAATG
 6421 GAAGAACTCA ACAGAAATCCA GGAATACCTT CAGAGCAGCG GGCTCGCTCA GAGACACCAA
 6481 AGGCTCTATT CCCTGCTGGA CAGGTTTCGT CTCTCCACCA TGCAAGACTC ACTTAGTCCC
 6541 TCTCCTCCAG TCTCTTCCCA CCCGCTGGAC GGGGATGGG ACGTGTCACT GGAACACAGTG
 6601 AGTCCGGACA AAGACTGGTA CGTTTCATCTT GTCAAAATCCC AGTGTGGAC CAGGTCAGAT
 6661 TCTGCACATG TGAAGGTGC AGAGCTGGTG AATCGGATTC CTGCTGAAGA TATGAATGCC
 6721 TTCATGATGA ACTCGGAGTT CAACCTAAGC CTGCTAGCTC CATGCTTAAG CCTAGGGATG
 6781 AGTGAAATTT CTGGTGGCCA GAAGAGTGCC CTTTTTGAAG CAGCCCGTGA GGTGACTCTG
 6841 GCCCGTGTGA GCGGCACCGT GCAGCAGCTC CCTGCTGTCC ATCATGTCTT CCAGCCCGAG
 6901 CTGCCTGCAG AGCCGGCGGC CTA~~CT~~GGAGC AAGTTGAATG ATCTGTTTGG GGATGCTGCA
 6961 CTGTATCAGT CCTTGCCAC TCTGGCCCGG GCCCTGGCAC AGTACCTGGT GGTGGTCTCC
 7021 AAATGCCCC GTCAATTTGCA CCTTCCTCCT GTCTGGCAT TTGATCCATG AGCAGATCCC GCTGAGTCTG
 7081 GTGGCAACCC TTGAGGCCCT GTCTGGCAT GTCTGGCATG AGCAGATCCC GCTGAGTCTG
 7141 GATCTCCAGG CAGGGCTGGA CTGCTGCTGC CTGGCCCTGC AGCTGCCCTGG CCTCTGGAGC
 7201 GTGGTCTCCT CCACAGAGTT TGTGACCCAC GCCTGCTCCC TCATCTACTG TGTCGACTTC
 7261 ATCCTGGAGG CCGTTGCAGT GCAGCCTGGA GAGCAGCTTC TTAGTCCAGA AAGAAGGACA
 7321 AATACCCCAA AAGCCATCAG CGAGGAGGAG GAGGAAGTAG ATCCAAACAC ACAGAATCCT
 7381 AAGTATATCA CTGCAGCCTG TGAGATGGTG GCAGAAATGG TGGAGTCTCT GCAGTCGGTG
 7441 TTGGCCTTGG GTCATAAAAG GAATAGCGGC GTGCCGGCGT TTCTCAGGCC ATTGCTCAGG
 7501 AACATCATCA TCAGCCTGGC CCGCCTGCC CTTGTCAACA GCTACACACG TGTGCCCCCA

FIG.1G

7561 CTGGTGTGGA AGCTTGGATG GTCACCCAAA CCGGAGGGG ATTTTGGCAC AGCATTCCT
7621 GAGATCCCCG TGGAGTTCCT CCAGGAAAAG GAAGTCTTTA AGGAGTTCAT CTACCGCATC
7681 AACACACTAG GCTGGACCAG TCGTACTICAG TTTGAAGAAA CTTGGGCCAC CCTCCTTGGT
7741 GTCCTGGTGA CGCAGCCCCCT 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 CCTTCTCTG TCTCCGGCTA CTACAGGTGC CCTCATCAGC
8101 CACGAGAAGC TGCTGCTACA GATCAACCCC GAGCGGGAGC TGGGAGCAT GAGCTACAAA
8161 CTCGGCCAGG TGTCATACA CTCCGTGTGG CTGGGGAACA GCATCACACC CCTGAGGGAG
8221 GAGGAATGGG ACGAGGAAGA GGAGGAGGAG GCCGACGCC CTGCACCTTC GTCACCACCC
8281 ACGTCTCCAG TCAACTCCAG GAAACACCGG GCTGGAGTTG ACATCCACTC CTGTTCCGAG
8341 TTTTGTCTTG AGTTGTACAG CCGCTGGATC CTGCCGTCCA GCTCAGCCAG GAGGACCCCG
8401 GCCATCCTGA TCAGTGAGGT GGTCAAGTCC CTTCTAGTGG TCTCAGACTT GTTCACCCGAG
8461 CGCAACCAGT TTGAGCTGAT GTATGTGACG CTGACAGAAC TGCGAAGGGT GCACCCCTCA
8521 GAAGACGAGA TCCTCGCTCA GTACCTGGTG CCTGCCACCT GCAAGGCAGC TGCCGTCTCT
8581 GGGATGGACA AGGCCGTGGC GGAGCTGTC AGCCGCCCTGC TGGAGAGCAC GCTCAGGAGC
8641 AGCCACCTGC CCAGCAGGGT TGGAGCCCTG CACGGCGTCC TCTATGTGCT GGAGTCCGAC
8701 CTGCTGGACG ACACTGCCAA GCAGCTCATC CCGGTCAATCA GCGACTATCT CCTCTCCAAC
8761 CTGAAAGGGA TCGCCCACTG CGTGAACATT CACAGCCAGC AGCACGTACT GGTCAATGTGT

FIG.1H

8821 GCCACTGCGT TTTACCTCAT TGAGAACTAT CCTCTGGACG TAGGGCCGGA ATTTTCAGCA
 8881 TCAATAATAC AGATGTGTGG 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 AGTGATTGTT
 9181 GCTATGGAGC GGGTATCTGT TCTTTTGTAT AGGATCAGGA AAGGTTTCC TTGTGAAGCC
 9241 AGAGTGGTGG CCAGGATCCT GCCCCAGTTT CTAGACGACT TCTTCCCACC CCAGGACATC
 9301 ATGAACAAAG TCATCGGAGA GTTCTGTGCC AACCAGCAGC CATACCCCCA GTTCATGGCC
 9361 ACCGTGGTGT ATAAGGTGTT TCAGACTCTG CACAGCACCG GGCAGTCGTC CATGGTCCGG
 9421 GACTGGGTCA TGCTGTCCCT CTCCAACTTC ACGCAGAGGG CCCCAGTCGC CATGGCCACG
 9481 TGGAGCCTCT CCTGCTTCTT TGTCAGCGCG TCCACCAGCC CGTGGGTCGC GGCATCCTC
 9541 CCACATGTCA TCAGCAGGAT GGGCAAGCTG GAGCAGGTGG ACGTGAACCT TTTCTGCCCTG
 9601 GTCGCCACAG ACTTCTACAG ACACCAGATA GAGGAGGAGC TCGACCGCAG GGCCTTCCAG
 9661 TCTGTGCTTG AGGTGGTTGC AGCCCCAGGA AGCCCATATC ACCGGCTGCT GACTTGTTTA
 9721 CGAAATGTCC ACAAGGTCAC CACCTGC  GCGCCATGTT GGGAGAGACT GTGAGGGCGG
 9781 AGCTGGGGCC GGAGCCTTTG GAAGTCTGTG CCCTTGTGCC CTGCCCTCCAC CGAGCCAGCT
 9841 TGGTCCCTAT GGGCTTCCGC ACATGCCGCG GCGGCCCAGG CAACGTGCGT GTCTCTGCCA
 9901 TGTGGCAGAA GTGCTCTTTG TGGCAGTGGC CAGGCAGGGA GTGTCTGCAG TCCTGTGGG
 9961 GCTGAGCCTG AGGCCTTCCA GAAAGCAGGA GCAGCTGTGC TGCACCCCAT GTGGGTGACC
 10021 AGGTCCTTTC TCCTGATAGT CACCTGCTGG TTGTTGCCAG GTTGCAGCTG CTCCTTGCATC

FIG.1I

10081 TGGGCCAGAA GTCCTCCCTC CTGCAGGCTG GCTGTTGGCC CCTCTGCTGT CCTGCAGTAG
 10141 AAGGTGCCGT GAGCAGGCTT TGGGAACACT GGCCTGGGTC TCCCTGGTGG GGTGTGCATG
 10201 CCACGCCCCG TGTCTGGATG CACAGATGCC ATGGCCTGTG CTGGGCCAGT GGCTGGGGGT
 10261 GCTAGACACC CGGCACCAT TCTCCCTTCTC TCTTTTCTTC TCAGGATTTA AAATTTAATT
 10321 ATATCAGTAA AGAGATTAAT TTTAAACGAAC TCTTTCTATG CCGGTGTAAA GTATGTGAAT
 10381 CGCAAGGCCT GTGCTGCAIG CGACAGCGTC CGGGGTGGTG GACAGGGCCC CCGGCCACGC
 10441 TCCCTCTCCT GTAGCCACTG GCATAGCCCT CCTGAGCACC CGCTGACATT TCCGTTGTAC
 10501 ATGTTCTGT TTATGCATC ACAAGGTGAC TGGGATGTAG AGAGGCGTTA GTGGGCAGGT
 10561 GGCCACAGCA GGA CTGAGGA CAGGCCCCCA TTATCCTAGG GGTGCGCTCA ACTGCAGCCC
 10621 CTCCTCCTCG GGCACAGACG ACTGTCGTTT TCCACCCACC AGTCAGGGAC AGCAGCCTCC
 10681 CTGTCACTCA CTTGAGAAG CCAGCCCTCC CTGGCTGTGA GCAGCCTCCA CTGTGTCCAG
 10741 AGACATGGGC CTCCCCTCC TGTTCCTTGC TAGCCCTGGG GTGGCGTCTG CCTAGGAGCT
 10801 GGCTGGCAGG TGTGGGACC TGCTGCTCCA TGGATGCATG CCTAAGAGT GTCAC TGAGC
 10861 TGTGTTTGT CTGAGCCTCT CTCGGTCAAC AGCAAAGCTT GGTGTCCTTG CACTGTTAGT
 10921 GACAGAGCCC AGCATCCCTT CTGCCCCCGT TCCAGCTGAC ATCTTGACAG GTGACCCCTT
 10981 TTAGTCAGGA GAGTGAGAT CTGTGCTCAT CGGAGACTGC CCCACGGCCC TGTCAGAGCC
 11041 GCCACTCCTA TCCCAGGAC AGGTCCCTGG ACCAGCCTCC TGTTCGAGG CCCAGAGGAG
 11101 CCAAGTCATT AAAATGGAAG TGGATTCTGG ATGGCCGGGC TGCTGCTGAT GTAGGAGCTG
 11161 GATTGGGAG CTCTGCTTGC CGACTGGCTG TGAGACGAGG CAGGGGCTCT GCTTCCTCAG
 11221 CCCTAGAGGC GAGCCAGGCA AGGTTGGCGA CTGTCAATGT GCTTGGTTTG GTCATGCCCG
 11281 TCGATGTTTT GGGTATTGAA TGTGGTAAAGT GGAGGAAATG TTGGAAC TCT GTGCAGGTGC

FIG.1J

11341 TGCCTTGAGA CCCCCAAGCT TCCACCTGTC CCTCTCCTAT GTGGCAGCTG GGGAGCAGCT
 11401 GAGATGTGGA CTTGTATGCT GCCACATAC GTGAGGGGA GCTGAAAGG AGCCCTTGCT
 11461 CAAAGGGAGC CCTCCTCTG AGCAGCCTCT GCCAGGCTG TATGAGGCTT TTCCCACCAG
 11521 CTCCCAACAG AGGCCTCCCC CAGCCAGGAC CACCTCGTCC TCGTGGCGGG GCAGCAGGAG
 11581 CGGTAGAAAG GGGTCCGATG TTTGAGGAG CCCTTAAGG AAGCTACTGA ATTATAACAC
 11641 GTAAGAAAAAT CACCATTTCTT CCGTATTGGT TGGGGGCTCC TGTTCCTCAT CCTAGCTTTT
 11701 TCCTGGAAAA GCGCGCTAGA AGGTTTGGGA ACGAGGGGAA AGTTCTCAGA ACTGTTGCTG
 11761 CTCCCACCC GCCTCCCGCC TCCCCCGCAG GTTATGTCAG CAGCTCTGAG ACAGCAGTAT
 11821 CACAGGCCAG ATGTTGTTC TGGCTAGATG TTTACATTTG TAAGAAATAA CACTGTGAAT
 11881 GTAAAAACAGA GCCATTCCCT TGGAAATGCAT ATCGCTGGG TCAACATAGA GTTTGTCTTC
 11941 CTC TTGTTTA CGACGTGATC TAAACCAAGT CTTAGCAAG GGCTCAGAAC ACCCGCTCT
 12001 GGCAGTAGGT GTCCCCCACC CCCAAAGACC TGCCTGTGTG CTCCGGAGAT GAAATATGAGC
 12061 TCATTAGTAA AAATGACTTC ACCCACGCAT ATACATAAAG TATCCATGCA TGTGCATATA
 12121 GACACATCTA TAATTTTACA CACACACCTC TCAAGACGGA GATGCATGGC CTC TAAGAGT
 12181 GCCCGTGTG GTTCTTCCTG GAAGTTGACT TTCCTTAGAC CCGCCAGGTC AAGTTAGCCG
 12241 CGTGACGGAC ATCCAGGCGT GGGACGTGGT CAGGGCAGGG CTCATTCAIT GCCCACTAGG
 12301 ATCCCACTGG CGAAGATGGT CTCCATATCA GCTCTCTGCA GAAGGGAGGA AGACTTTATC
 12361 ATGTTCCCTAA AAATCTGTGG CAAGCACCCA TCGTATTATC CAAATTTTGT TGCAAAATGTG
 12421 ATTAATTTGG TTGTCAAGTT TTGGGGGTGG GCTGTGGGA GATTGCTTTT GTTTTCCTGC
 12481 TGGTAATATC GGGAAAGATT TTAATGAAC CAGGGTAGAA TTGTTTGGCA ATGCACTGAA
 12541 GCGTGTCTTCT TTCCCCAAAT GTGCCTCCCT TCCGCTGCGG GCCCAGCTGA GTCTATGTAG

FIG.1K

12601 GTGATGTTTC CAGCTGCCAA GTGCTCTTTG TTACTGTCCA CCTCATTTTC TGCCAGCGCA
 12661 TGTGTCCTTT CAAGGGGAAA ATGTGAAGCT GAACCCCTTC CAGACACCCA GAATGTAGCA
 12721 TCTGAGAAAG CCTGTGCCC TAAAGGACAC CCTCGCCCC CATCTTCATG GAGGGGTCA
 12781 TTTCAGAGCC CTCGGAGCCA ATGAACAGCT CCTCCTCTTG GAGCTGAGAT GAGCCCCACG
 12841 TGGAGCTCGG GACGGATAGT AGACAGCAAT AACTCGGTGT GTGGCCGCCCT GGCAGGTGGA
 12901 ACTTCCTCCC GTTGCCTGGT GGAGTGAGGT TAGTTCTGTG TGTCTGGTGG GTGGAGTCAG
 12961 GCTTCTCTTG CTACCTGTGA GCATCCTTCC CAGCAGACAT CCTCATCGGG CTTTGTCCCT
 13021 CCCCCGCTTC CTCCTCTGC GGGGAGGACC CGGGACCACA GCTGCTGGCC AGGGTAGACT
 13081 TGGAGCTGTC CTCCAGAGGG GTCACGTGTA GGAGTGAGAA GAAGGAAGAT CTTGAGAGCT
 13141 GCTGAGGGAC CTTGGAGAGC TCAGGATGGC TCAGACGAGG ACACTCGCTT GCCGGGCCCTG
 13201 GCCCTCCTGG GAAGGAGGGA GCTGCTCAGA ATGCCGCAAG ACAACTGAAG GCAACCTGGA
 13261 AGGTTCAAGG CCGCTCTTC CCCCATGTGC CTGTCACGCT CTGGTGCAGT CAAAGGAACG
 13321 CCTTCCCCTC AGTTGTTTCT AAGAGCAGAG TCTCCCCTG CAATCTGGGT GGTAAC TGCC
 13381 AGCCTTGGAG GATCGTGGC AACGTGGACC TGCCTACGGA GGTGGGCTC TGACCCAAGT
 13441 GGGGCTCCT TGCCAGGTC TCACTGCTTT GCACCGTGGT CAGAGGGACT GTCAGCTGAG
 13501 CTTGAGCTCC CCTGGAGCCA GCAGGGCTGT GATGGGCGAG TCCCGGAGCC CCACCCAGAC
 13561 CTGAATGCTT CTGAGAGCAA AGGGAAGGAC TGACGAGAGA TGTATATTTA ATTTTAAAC
 13621 TGC TGCAAAC ATTGTACATC CAAATTAAAG GGAAAAAATG GAAACCATCA AT

FIG.2A

1 matleklmka feslksfqqq qqqqqqqqq qqqqqqqqq pppppppppp pqlpqpqpqa
 61 qpllpqpqp pppppppppp avaeephrp kkeksatkdd rvnhclice nivaqsvrns
 121 pefqkllgia melfllcsdd aesdvrmvad eclnkvikal mdsnlprlql elykeikkng
 181 aprslraalw rfaelahlvr pdkcrpylvr llpcltrtsk rpeesvqetl aaavpkimas
 241 fgnfandnei kvllkafian lkssptirr taagsavsic qhsrrtqfyf swllnvlgl
 301 lvpvedehst llilgvlltl rylvllqqg vkdtslkgsf gvtrkemevs psaeqlvqv
 361 eltlhhtghq dhnvvtgale llqqlfrtpp pellqtltaf ggigqltaak eesggrsrsg
 421 siveliaagg sscspvlsrc qkgkvllgee ealeddesr sdvssalta svkdeisgel
 481 aassgvstpg saghdiiteq prsqhtlqad svdlascldt ssatdgdced ilshsssqvs
 541 avpsdpamd1 ndgtqasspi sdssqtteq pdsavtspds seivldgtdn qylglqigqp
 601 qdedeeatgi lpdeaseafr nssmalqgah llknmshcrq psdssvdkfv lrdeatepgd
 661 genkpcrirk digqstddds aplvhcrrll sasflitggk nvlvpdrdrvr vsvkalalsc
 721 vgaavalhpe sffsklykvp ldtteypee qvdsilnyid hgdpgvrgat ailcgtlics
 781 ilsrfrfhvg dwmgirtlt gntfsladci pllrrktlkde ssvtcklact avrncvmslc
 841 sssyselglq liidvltlrn ssywlrvtel letlaeidfr lvsfleakae nlhrgahhyt
 901 glklqervl nnvvihlldg edprvrhva aslirlvpkl fykcdqgqad pvvavardqs
 961 svylklmhe tqppshfsvs titriyrgyn llpsitdvtm ennlsrviaa vshelittstt
 1021 raltfgccea lcllstafrp ciwslgwhcg vpplsasdes rkscvtgmat miltllssaw
 1081 fpdlisahqd alilagnlla asapkslrns waseeanpa atkqeevwa lgdralvpmv
 1141 eqlfshllkv inicahvldd vaggpaikaa lpsltnppsl sprrrkgekep egeqasvpl
 1201 spkkgseasa asrqsdtsqp vttskssslg sfyhlpsylk lhdvllkatha nykvtldlqn
 1261 stekfggflr saldvlsqil elatlqdigk cveeilgylk scfsrepmma tvcvqgllkt
 1321 lfgtnlasqf dglssnpsks qgradrqgss svrpglyhyc fnapythftq aladaslrnm
 1381 vqaeqendts gwfdvldkvs tqlktnltsv tknradkna1 hnhirlfepl vikalkqytt
 1441 ttcvqlqkv ldllaqlvql rvnyclldsd qvfigfvlkq feyievgqfr eseaiipnif
 1501 fflvllsyer yhsqigligip kiiqlcdgim asgrkavtha ipalqpivhd lfvrlgtuka
 1561 dagkeletqk evvsmllrl iqyqhvlmf ilvlqgchke nedkwkrlsr qiadiilpml
 1621 akqgmhidsh ealgvnlrlf eilapsslrp vdmllrsmfv tpntmasvst vqlwisgila
 1681 ilrvlisqst edivlsriqe lsfspylisc tvinrlrdgd ststleehse gkqiknlpee

FIG.2B

1741 tfsrfillqlv gilledivtk qlkvemseqq htfcyqelgt llmclihifk sgmfrritaa
 1801 atrlfrsdgc ggsfytlsl nlrarsmitt hpalvllwcq illllvnhtdy rwaevqgtp
 1861 krhslsstkl lspqmsgeee dsdlaaklgm cnreivrrga lilfcdyvcq nlhdsehlw
 1921 livnhiqqli slsheppvqd fisavhrnsa asglfiqaiq srcenlstpt mlkktlqcle
 1981 gihlsqsgav ltlyvdrllc tpfrrvlarmv dilacrrvem llaanlqssm aqlpmeelnr
 2041 iqeylqssgl aqrhqrlysl ldrfrlstmq dsllspspvvs shpldgdghv sletvspdkd
 2101 wyvhlvksqc wtrdsalle gaelvnripa edmnafmms efnlsllapc lslgmseisg
 2161 gqksalfeaa revtlarvsg tvqqlpavhh vfqelpaep aaywsklndl fgdaalyqsl
 2221 ptlaralagy lvvsklpsh lhlppekekd ivkfvtatle alswhliheq ipslldlqag
 2281 ldccclalql pglwsvvssst efvthacsli ycvhfileav avqpgedlls pertrntpka
 2341 iseeeeevdp ntqnpkyita acemvaemve slqsvlalgh krnsgvpafll tpllrniis
 2401 larlplvnsy trvpplvwkl gwspkpggdf gtafpeipve flqekvfke fiyrntlgw
 2461 tsrtqfeetw atllgvlvtq plvmeqeesp peedtertqi nvlavqaits lvlsamtpvp
 2521 agnpavscle qqprnkplka ldtrfgrkls iirgivegei qamvskreni athhlyqawd
 2581 pvpslspatt galishekll lqinperelg smsyklgqvs ihsvwlgnsl tplreewde
 2641 eeeeeadapa pspptspvn srkhragvdi hscsqfllel ysrrwilpss arrrtpailis
 2701 evvrsllvvs dlfternqfe lmyvtltelr rvhpsedeil aqylvpatck aaavlgmdka
 2761 vaepvsrllle stlrsshlpv rvgalhgvlv vlecldlddt akqlipvisd yllsnlkgia
 2821 hcvnihsqgh vlvmcatafy lienypldvg pefsasiiqm cgvmlsgsee stpsliyhca
 2881 lrglerllls eqslrldaes lvklsvdrvn vhsphramaa lglmltcmtyt gkekvspgrrt
 2941 sdnpaapds esvivamerv svlfdrirkq fpcearvvar ilpqflddff ppqdimnkvi
 3001 geflenqgpy pqfmatvvyk vfqtlhstgq smmvrdrwml slsnftgrap vamatwsisc
 3061 fivsastspw vaailphvis rmqgleqgdv nlfclvatdf yrhgieeeld rrafqsvlev
 3121 vaapgsphyrl lltclrnvhk vtcc

FIG. 3

diana codificante de *htt*: 5'-...ugcagcugaucaucgaugugcugacccugaggaacaguuc...-3'
diana no codificante de *htt*: 3'-...acgucgacuaguagcuacacgacugggacuccuugucaag...-5'

FIG. 4

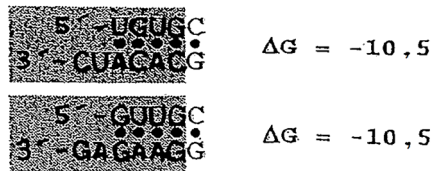


FIG. 5A

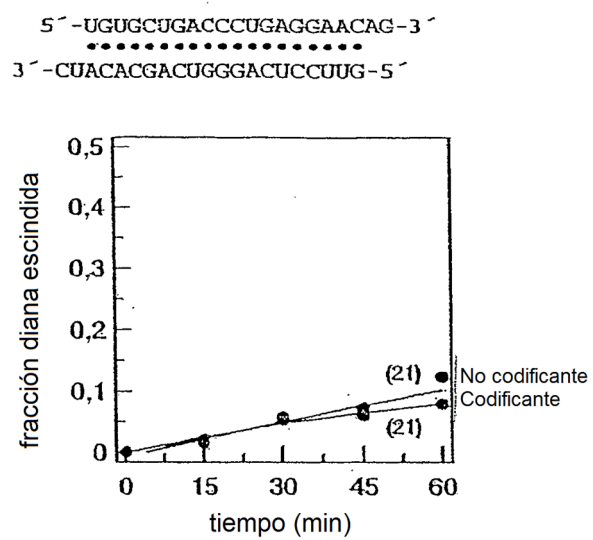


FIG. 5B

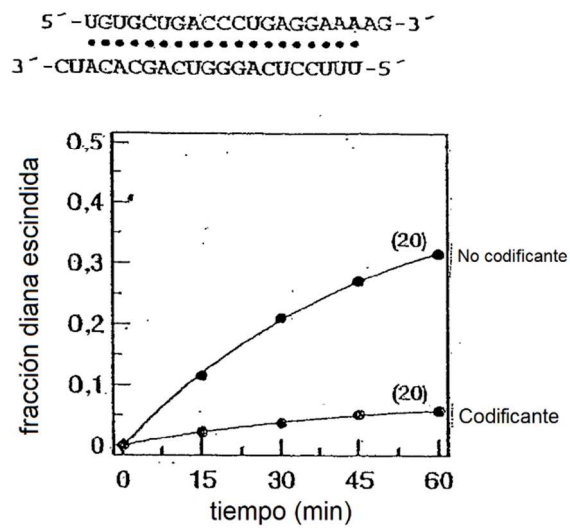


FIG. 5C

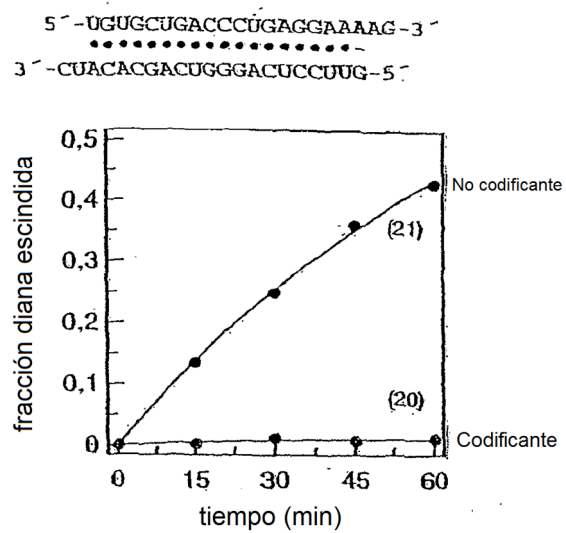
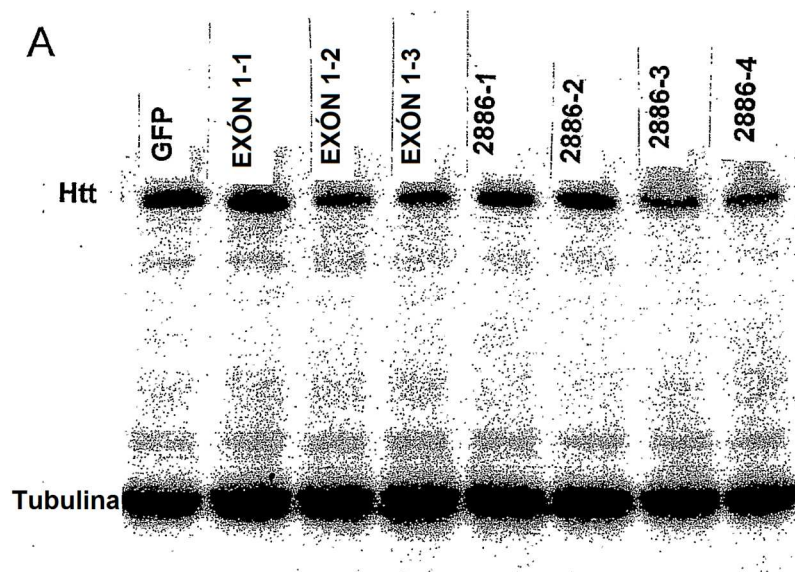


FIG. 6



B

