

SUPRESIÓN DEL GEN DE LA HUNTINGTINA INDUCIDA POR LA ARNI

ANTECEDENTES DE LA INVENCIÓN

5 El gen de la huntingtina, referido también como el gen de la HTT o HD (enfermedad de Huntington), codifica para la proteína huntingtina. El gen de la huntingtina es un gran gen de aproximadamente 13.5 kb (la proteína huntingtina es de aproximadamente 350 kDa). La enfermedad de Huntington es un trastorno neurodegenerativo genético causado por una mutación genética en el gen de la huntingtina. La mutación genética implica un segmento de ADN del gen de la huntingtina
10 conocido como la repetición del trinucleótido CAG. Normalmente, el segmento de CAG en el gen de la huntingtina de los humanos se repite múltiples veces, es decir, aproximadamente 10 a 35 veces. Las personas que desarrollan la enfermedad de Huntington tienen una expansión del número de repeticiones de CAG en por lo menos un alelo. Una persona afectada hereda usualmente el alelo mutado de un progenitor afectado. En casos raros, un individuo con la enfermedad de Huntington no
15 tiene un progenitor con el trastorno (HD esporádica). Las personas con 36 a 39 repeticiones de CAG pueden desarrollar signos y síntomas de la enfermedad de Huntington, mientras que las personas con 40 o más repeticiones casi siempre desarrollan el trastorno. El incremento en el tamaño de la repetición de CAG lleva a la producción de una proteína huntingtina alargada (mutada). Esta proteína es procesada en la célula en fragmentos más pequeños que son citotóxicos y que se
20 acumulan y agregan en las neuronas. Esto resulta en la alteración de la función normal y muerte final de las neuronas. Este es el proceso que ocurre en el cerebro que subyace a los signos y síntomas de la enfermedad de Huntington.

La interferencia del ARN (ARNi) es un mecanismo de ocurrencia natural que implica la subregulación del ARNm específica de la secuencia. La subregulación del ARNm resulta en una
25 reducción de la cantidad de proteína que es expresada. La interferencia del ARN es desencadenada por el ARN de doble cadena. Una de las cadenas de ARN de doble cadena es sustancialmente o completamente complementaria a su objetivo, el ARNm. Esta cadena se denomina la cadena guía. El mecanismo de interferencia del ARN implica la incorporación de la cadena guía en el complejo de silenciamiento inducido por el ARN (RISC). Este complejo es un complejo de cambios múltiples que
30 por medio del apareamiento de bases complementarias se une a su ARNm objetivo. Una vez unido a su ARNm objetivo, puede escindir el ARNm o reducir la eficacia de la traducción. Desde su descubrimiento, la interferencia del ARN se ha usado ampliamente para suprimir la expresión de genes objetivo específicos. Los factores desencadenantes para inducir la interferencia del ARN que se han usado implican el uso de moléculas de ARNsi o moléculas de ARNsh. Además, moléculas que
35 pueden desencadenar naturalmente la ARNi, las denominadas moléculas de ARNmi, se han usado

para obtener moléculas de ARNm artificiales que imitan a sus contrapartes de ocurrencia natural. Estas estrategias tienen en común que proveen sustancialmente moléculas de ARN de doble cadena que están diseñadas para focalizar un gen de elección. Los procedimientos terapéuticos basados en la ARNi que usan la modalidad de ARNi específica de la secuencia están bajo desarrollo, y varios se
 5 usan actualmente en pruebas clínicas (véase entre otros, a Davidson y McCray, *Nature Reviews - Genetics*, 2011; Vol. 12; 329-340).

Puesto que la enfermedad de Huntington implica la expresión de una proteína huntingtina mutada, la acumulación de la misma llevando a la enfermedad, la interferencia del ARN provee una oportunidad para tratar la enfermedad, ya que puede reducir la expresión del gen de la
 10 huntingtina. El paradigma que subyace a este procedimiento implica una reducción de la proteína Htt mutante para reducir de esta manera los efectos tóxicos que resultan de la proteína Htt mutante para lograr una reducción y/o el retraso de los síntomas de la enfermedad de Huntington, o incluso para prevenir los síntomas de la enfermedad de Huntington en conjunto. Se ha planteado la hipótesis de la focalización de la supresión del gen de la huntingtina en la técnica anterior,
 15 incluyendo el listado de aproximadamente dos mil secuencias objetivo de ARNs hipotéticas (véase el documento WO2005105995). Estrategias para reducir la expresión del gen de la huntingtina se conocen en la técnica, e implican la focalización específica de genes de la huntingtina mutantes (véase, por ejemplo, los documentos US20090186410 y US20110172291). En forma alternativa, la interferencia del ARN se ha usado también para focalizar genes mutantes y no mutantes (véase, por
 20 ejemplo, Rodríguez-Lebron *et al.*, 2005, *Mol Ther.* Vol. 12 No. 4: 618-633; Franich *et al.*, 2008, *Mol Ther.* Vol. 16 No. 5; 947-956; Drouet *et al.*, 2009, *Annals of Neurology*; Vol. 65 No. 3; 276-285, y McBride *et al.*, *Mol Ther.* Diciembre de 2011; 19 (12): 2152-62; documentos US20080015158 y WO2008134646). En el último caso, se mostró que la supresión de la expresión de la proteína huntingtina de tipo silvestre no tiene efecto perjudicial aparente alguno.

25

BREVE DESCRIPCIÓN DE LA INVENCION

La presente invención provee un ARN de doble cadena que comprende una primera secuencia de ARN y una segunda secuencia de ARN en donde la primera y segunda secuencias de
 30 ARN son sustancialmente complementarias, en donde la primera secuencia de ARN tiene una longitud de secuencia de por lo menos 19 nucleótidos y es sustancialmente complementaria a SEQ ID NO. 1. Un gran número de secuencias objetivo se puso a prueba para la supresión efectiva de la expresión del gen de la huntingtina. Se encontró que el ARN de doble cadena seleccionado de la presente invención es efectivo en la reducción de la expresión del gen de la huntingtina. Dicho ARN
 35 de doble cadena cuando se provee en una célula, ya sea directamente por medio de transfección o

indirectamente por medio del suministro de ADN (por ejemplo, transfección) o por medio de la expresión mediada por el vector después de lo cual dicho ARN de doble cadena puede ser expresado, es capaz de reducir la expresión de un gen de la huntingtina mutado y un gen de la huntingtina normal. Además, se mostró que el ARN de doble cadena de la invención fue capaz de reducir la expresión del gen objetivo cuando se proveyó como un ARNsi o en un andamio de ARNmi. Cuando se puso a prueba en un modelo animal, se mostró que un ARN de doble cadena de conformidad con la invención fue capaz de reducir la muerte de las células neuronales y los agregados de huntingtina. El ARN de doble cadena como se proveyó en la presente invención provee una mejora en comparación con las moléculas de ARN de doble cadena en la técnica que focalizan el gen de la huntingtina, o provee por lo menos una alternativa a las mismas.

El ARN de doble cadena de conformidad con la invención puede proveerse como un ARNsi, un ARNsh, un pre-ARNmi o pri-ARNmi. Dichas moléculas de ARN de doble cadena pueden ser suministradas a las células objetivo directamente, por ejemplo, por medio de la captación celular usando, por ejemplo, métodos de transfección. De preferencia, dicho suministro se logra usando un vector de terapia génica, en donde un cassette de expresión para el ARNsi, ARNsh, pre-ARNmi o pri-ARNmi se incluye en un vector. De esta manera, las células pueden ser provistas con un suministro constante de ARN de doble cadena para lograr la supresión durable del gen de la huntingtina sin que se requiera una administración repetida. De preferencia, el vector viral de elección es AAV5. La presente invención provee también de esta manera el uso médico de un ARN de doble cadena de conformidad con la invención, tal como en el tratamiento de la enfermedad de Huntington, en donde dicho uso médico puede comprender también un cassette de expresión o un vector viral, tal como AAV5, capaz de expresar dicho ARN de doble cadena de la invención.

BREVE DESCRIPCIÓN DE LAS FIGURAS

Figura 1. Secuencias objetivo y del gen de la huntingtina humana (HTT). (A) Representación esquemática del gen de la HTT humana (L27350.1) con expansiones de CAG (negro) y secuencias objetivo para miH1-H21 (gris claro). (B) Secuencia de ARN del exón 1 del gen de la HTT (SEQ ID NO. 2). La secuencia de repetición de CAG abarca los nucleótidos 367-429. (C) Representación esquemática de las secuencias objetivo puestas a prueba para el exón 1 (H1, 185-205; H2, 186-206; H3, 189-209; H4, 191-211; H5, 194-214; H6, 196-216; H7, 250-270; H8, 261-281; H9, 310-330; H10, 311-331; H11, 339-359; H12, 345-365; H13, 454-474; H14, 459-479; H15, 477-497; H16, 486-506; H17, 492-512; H18, 498-518; H19, 549-569; H20, 557-577; H21, 558-578; en donde H1-H21 corresponde a SEQ ID NOs. 23-43). Las secuencias representadas son secuencias de ADN. Los números se refieren a las secuencias de nucleótidos de ARN correspondientes en SEQ

ID NO. 2. Las secuencias objetivo de ARN correspondientes de SEQ ID NO. 2 tienen la secuencia que aparece en C), salvo que en donde el ADN codifica para una "T", el ARN codifica para una "U".

Figura 2. Ejemplos de moléculas de ARN de doble cadena y cassettes de expresión.

(A) Ejemplos del andamio de pri-/pre-ARNmi para miH12 pre-miH12-155 (SEQ ID NO. 44) y el andamio de pre-miH12-451 (SEQ ID NO. 45) con la guía de miH12 (gris) indicada. (B) Representación esquemática que esboza las moléculas de ARN de doble cadena de conformidad con la invención y la forma en que pueden ser procesadas por la maquinaria de ARNi. Un ARN de doble cadena puede ser un ARN de horquilla corto (1) o un ARNsi extendido (2). El ARN de horquilla o el ARNsi extendido tienen la primera secuencia de ARN en el extremo proximal, como se indica (indicado con 1 y paréntesis). Un ARN de horquilla corto o un ARNsi extendido pueden ser procesados por la maquinaria de ARNi en la célula para producir un ARNsi (3), el cual puede ser también un ARN de doble cadena de conformidad con la invención, de los cuales una cadena que comprende la primera secuencia de ARN puede ser incorporada en el RISC (4). Un ARN de doble cadena puede estar comprendido en una secuencia de pri-ARNmi (5) o una secuencia de pre-ARNmi (6). El pri-ARNmi puede ser procesado por la maquinaria de ARNi para producir un pre-ARNmi y posteriormente un dúplex de ARNmi maduro (7), del cual una cadena que comprende la primera secuencia de ARN puede ser incorporada en el RISC (4). La posición de la primera secuencia de ARN en el pre-ARNmi, pri-ARNmi y dúplex de ARNmi se indica como 1 y paréntesis. (C) Secuencia de ADN del cassette de expresión pVD-CMV-miH12-155 (**promotor del CMV** (1-588), secuencia intermedia, secuencia de la proteína fluorescente verde = secuencia de GFP (713-1432), flanco de pri-ARNmi 5' (1433-1514), pre-ARNmi 5', **cadena guía (primera secuencia de ARN)** (1520-1540), secuencia de bucle (espira), **cadena viajera (segunda secuencia de ARN)** (1560-1578), pre-ARNmi 3'), flanco de pri-ARNmi 3' (1584-1704), **señal de poliA de TK del HSV** (1705-1976). (D) Secuencia de ADN del cassette de expresión pVD-CAG-miH12-451 (**promotor de CAG** (43-1715), flanco de pri-ARNmi 5' (1716-2017), pre-ARNmi 5', **cadena guía (primera secuencia de ARN)** (2034-2054), segunda secuencia de ARN y pre-ARNmi 3', flanco de pri-ARNmi 3' (2090-2320), **señal de poliA de hGH** (2321-2417). (E) Secuencia de ADN del cassette de expresión pVD-PGK-miH12-451 (**promotor de PGK** (23-277), flanco de pri-ARNmi 5' (278-794), pre-ARNmi 5', **cadena guía (primera secuencia de ARN)** (811-831), segunda secuencia de ARN y pre-ARNmi 3', flanco de pri-ARNmi 3' (867-1097), **señal de poliA de hGH** (1098-1194). (F) Secuencia de pri-miH12-155 que es codificada por el cassette de expresión pVD-CMV-miH12-155. (G) Secuencia de pri-miH12-451 que es codificada por el cassette de expresión pVD-CAG-miH12-451. **(G) Secuencia de pri-miH12-451** que es codificada por el cassette de expresión pVD-PGK-miH12-451. Para las figuras (E), (F) y (G), el tipo de fuente es el mismo que se usó anteriormente para el ADN correspondiente. Las secuencias de promotor están en negritas. La secuencia de la proteína fluorescente verde está en cursivas y subrayada (sólo C), las

secuencias de pri-ARNmi tienen un tipo de fuente normal, la cadena guía (primera secuencia de ARN) está en negritas y en cursivas, y la cadena viajera o segunda secuencia de ARN está en cursivas, las secuencias de pre-ARNmi están subrayadas, y la señal de poliA está en negritas y subrayada.

5 Figura 3. Eficacia de supresión de la expresión de miH1-21 *in vitro*. (A) Supresión de la expresión de HTT total por la focalización del exón 1. LucHTT fue co-transfectada en células Hek293T con miH1-miH21. Se midió la fluorescencia de renilla y de la luciferasa de luciérnaga 48 h post-transfección. Se usó a miScr (ctrl = control) como un control negativo, y se ajustó a 100%. miH12 mostró la eficiencia de supresión de la expresión más fuerte. (B) La supresión de la expresión
10 de LucHTT fue por siH12 sintético con 19-23 nucleótidos de longitud.

 Figura 4. Eficacia de supresión de la expresión de miH12-451 *in vitro* con diferentes promotores. (A) El reportero de LucHTT ha sido co-transfectado con las variantes CAG-miH12 o PGK-miH12, y se determinó la eficacia de supresión de la expresión como se describió anteriormente para la figura 3. (B) La actividad de la cadena viajera (*) de miH12-451* expresada de los promotores de
15 CAG o PGK se midió en reporteros de LucHTT* específicos. No se detectó actividad de la cadena viajera.

 Figura 5. Eficacia *in vivo* de miH12 suministrado por AAV5. (A) Plan experimental. Los ratones fueron co-inyectados con AAV5-Luc73QHTT y AAV5-CMV-miScr-155 o AAV5-CMV-miH12-155 a una relación de 1:5. Los puntos de medición se indican con flechas. (B) La supresión de la
20 expresión de AAV5-Luc73QHTT en los animales 6 semanas p.i. se midió por IVIS. (C) Tendencia de la supresión de la expresión de AAV5-Luc73QHTT por AAV5-miH12 hasta 6 semanas p.i.

 Figura 6. Prueba del concepto de la supresión de la expresión de la HTT humana en el modelo mecanicista de HD de la rata. (A) Plan experimental. (B) Histología del cerebro que muestra menos neurodegeneración (DARP32) y menos agregados de Htt mutantes (EM48) en el
25 grupo de AAV5-CMV-miH12-155. (C) Histología de la GFP en el cerebro. (D) Histología del cerebro con el marcador de activación inmune Iba1.

 Figura 7. Supresión de la expresión de la HTT humana en el modelo de ratón HD Hu97/18 humanizado. (A) Eficiencia de la transducción en el cerebro de murino tras la inyección intraestriatal lenta, la inyección intraestriatal por difusión mejorada por convección (CED) o la
30 inyección intracerebral ventricular (ICV) de AAV5-CMV-miH12-155. La fluorescencia de la GFP se observó 5 semanas post-inyección. (B) Medición por Western blot de la supresión de la expresión de la HTT humana en el cerebro de murino tras el suministro de AAV5-miHTT. (C) Cuantificación de la HTT por Western blot.

 Figura 8. Comparación del objetivo H12 seleccionado con secuencias objetivo de la
35 técnica anterior. LucHTT fue co-transfectada en células Hek293T con las moléculas de ARNs

indicadas (A) y las construcciones de ARNmi (B y C). Se midió la fluorescencia de renilla y de la luciferasa de luciérnaga 48 h post-transfección. miH12 y siH12 mostraron la eficiencia de supresión de la expresión más fuerte.

5

DESCRIPCIÓN DETALLADA DE LA INVENCION

La presente invención provee un ARN de doble cadena que comprende una primera secuencia de ARN y una segunda secuencia de ARN, en donde la primera y segunda secuencias de ARN son sustancialmente complementarias, en donde la primera secuencia de ARN tiene una longitud de secuencia de por lo menos 19 nucleótidos y es sustancialmente complementaria a SEQ ID NO. 1.

SEQ ID NO. 1 (5'-CUUCGAGUCCCUCAAGUCCUU-3') corresponde a una secuencia objetivo del gen de la huntingtina del exón 1 (SEQ ID NO. 2). El exón 1, como se representa en la figura 1B, tiene 21 secuencias de repetición de CAG de los nucleótidos (nt.) 367-429. La secuencia del exón 1, como se representa en la figura 1B, corresponde a un gen de la huntingtina normal que no está asociado con la enfermedad. Los genes de la huntingtina mutantes correspondientes asociados con la enfermedad de Huntington comprenden mucho más que 21 secuencias de repetición de CAG. Como se dijo, con 36 a 39 repeticiones de CAG pueden desarrollarse signos y síntomas de la enfermedad de Huntington, mientras que con 40 o más repeticiones casi siempre se desarrolla el trastorno. La secuencia SEQ ID NO. 1 objetivo está comprendida en sustancialmente todas las secuencias del exón 1, independientemente del número de repeticiones de CAG.

SEQ ID NO. 1 corresponde a los números de nucleótido 345-365 de SEQ ID NO. 2. 18 diferentes secuencias objetivo en el exón 1 se pusieron a prueba para focalización usando moléculas de ARN de doble cadena que fueron diseñadas para inducir una inhibición específica de la secuencia de SEQ ID NO. 2 (véase la figura 1 y la figura 3A), y se encontró que, en particular, la focalización de esta secuencia del exón 1 fue útil en la reducción de la expresión del gen de la huntingtina. Se encontró que moléculas de ARNsi que varían en longitud, es decir, consistiendo de 19, 20, 21, 22 y 23 pares de bases consecutivos con 2 salientes de nucleótidos adicionalmente, son efectivas contra esta secuencia, así como dos andamios de ARNmi separados que poseen una secuencia de 21 nucleótidos complementaria a SEQ ID NO. 1 en la posición de la secuencia guía (véase las figuras 3A, 3B y 4). Por lo tanto, la primera secuencia de ARN que es sustancialmente complementaria a la secuencia objetivo de la huntingtina SEQ ID NO. 1 tiene una longitud de secuencia de por lo menos 19 nucleótidos.

La primera secuencia de ARN de conformidad con la invención está comprendida en la cadena guía del ARN de doble cadena, referida también como cadena antisentido, ya que es

35

complementaria ("anti") a la secuencia objetivo sentido. La segunda secuencia de ARN está comprendida en la cadena viajera, referida también como "cadena sentido", ya que puede tener una identidad de secuencia sustancial con, o puede ser idéntica a, la secuencia objetivo. La primera y segunda secuencias de ARN están comprendidas en un ARN de doble cadena, y son sustancialmente complementarias. Dicho ARN de doble cadena de conformidad con la invención es para inducir la interferencia del ARN para reducir de esta manera la expresión del gen de tipo silvestre y mutante de la huntingtina. Por lo tanto, se entiende que sustancialmente complementarias significa que no se requiere que tengan todos los nucleótidos de la primera y segunda secuencias de ARN con bases apareadas, es decir, que sean completamente complementarias, o todos los nucleótidos de la primera secuencia de ARN y SEQ ID NO. 1 con bases apareadas. En tanto el ARN de doble cadena sea capaz de inducir la interferencia del ARN para de esta manera focalizar específicamente una secuencia que comprende SEQ ID NO. 1, dicha complementariedad sustancial se contempla en la invención.

Por lo tanto, en una modalidad, el ARN de doble cadena de conformidad con la invención comprende una primera secuencia de ARN y una segunda secuencia de ARN en donde la primera y segunda secuencias de ARN son sustancialmente complementarias, y en donde la primera secuencia de ARN tiene una longitud de secuencia de por lo menos 19 nucleótidos y es sustancialmente complementaria a SEQ ID NO. 1, es capaz de inducir la interferencia del ARN en la secuencia para reducir específicamente la expresión de un transcrito de ARN que comprende SEQ ID NO. 1. En una modalidad adicional, dicha inducción de la interferencia del ARN para reducir la expresión de un transcrito de ARN que comprende SEQ ID NO. 1, significa que es para reducir la expresión del gen de la huntingtina humana.

Puede determinarse fácilmente si este es el caso usando pruebas de reportero de luciferasa estándar y controles apropiados, tales como se describe en los ejemplos y como es sabido en la técnica (Zhuang *et al.*, 2006 Methods Mol Biol. 2006; 342: 181-7). Por ejemplo, un reportero de luciferasa que comprende SEQ ID No. 1 puede usarse para mostrar que el ARN de doble cadena de conformidad con la invención es capaz de suprimir la expresión específica de la secuencia. Además, como se muestra en la sección de ejemplos, la expresión de la huntingtina puede determinarse con anticuerpos específicos para determinar la cantidad de expresión en un análisis Western blot, como es el caso del análisis Northern blot que detecta la cantidad de transcritos de ARN.

Por lo tanto, el ARN de doble cadena de conformidad con la invención es para su uso en la inducción de la interferencia del ARN. El ARN de doble cadena de conformidad con la invención es para su uso en la reducción de la expresión de transcritos que comprenden SEQ ID NO. 1 tales como, por ejemplo, SEQ ID NO. 2 o similares, con un número variable de repeticiones de CAG.

Como se dijo, el ARN de doble cadena es capaz de inducir la interferencia del ARN. Estructuras de ARN de doble cadena son bien conocidas en la técnica, y son adecuadas para la inducción de la ARNi. Por ejemplo, un pequeño ARN de interferencia (ARNsi) comprende dos cadenas de ARN separadas, una cadena comprendiendo la primera secuencia de ARN y la otra

5 cadena comprendiendo la segunda secuencia de ARN. Un diseño de ARNsi que se usa con frecuencia implica 19 pares de bases consecutivos con salientes 3' de dos nucleótidos (véase la figura 2A). Este diseño se basa en el procesamiento de Dicer (enzima ribonucleasa III) observado de las más grandes moléculas de ARN de doble cadena que resulta en moléculas de ARNsi que tienen estas características. La saliente 3' puede estar comprendida en la primera secuencia de ARN. La saliente

10 3' puede estar además de la primera secuencia de ARN. La longitud de las dos cadenas de las cuales un ARNsi está compuesto puede ser de 19, 20, 21, 22, 23, 24, 25, 26 o 27 nucleótidos o más. Cada una de las dos cadenas comprende la primera y segunda secuencias de ARN. La cadena que comprende la primera secuencia de ARN también puede consistir de la misma. La cadena que comprende la primera secuencia de ARN también puede consistir de la primera secuencia de ARN y

15 la secuencia saliente.

Las moléculas de ARNsi pueden servir también como sustratos de Dicer. Por ejemplo, un sustrato de Dicer puede ser un componente de 27 partes que consiste de dos cadenas de ARN que tienen 27 pares de bases consecutivos. La primera secuencia de ARN está posicionada en el extremo 3' del dúplex de 27 partes. En el extremo 3', como con las moléculas de ARNsi, es una

20 saliente de dos nucleótidos. La saliente 3' puede estar comprendida en la primera secuencia de ARN. La saliente 3' puede estar además de la primera secuencia de ARN. 5' de la primera secuencia de ARN, pueden incluirse secuencias adicionales que son complementarias a la secuencia objetivo adyacente a SEQ ID NO. 1 o no. El otro extremo del sustrato de Dicer de ARNsi es de extremos rasurados. Este diseño de sustrato de Dicer da como resultado una preferencia en el procesamiento

25 por Dicer tal que un ARNsi se forma como el diseño de ARNsi como se describió anteriormente, teniendo 19 pares de bases consecutivos y 2 salientes de nucleótidos en ambos extremos 3'. En cualquier caso, las moléculas de ARNsi, o similares, se componen de dos cadenas de ARN separadas (Fire *et al.*, 1998, Nature. Febrero 19 de 1998; 391 (6669): 806-11), cada cadena de ARN comprendiendo o consistiendo de la primera y segunda secuencias de ARN de conformidad con la

30 invención.

El ARN de doble cadena de conformidad con la invención no requiere que la primera y segunda secuencias de ARN estén comprendidas en dos cadenas separadas. La primera y segunda secuencias de ARN pueden estar comprendidas también en una sola cadena de ARN tal como, por ejemplo, un ARNsh. Un ARNsh puede comprender de 5' - segunda secuencia de ARN - secuencia de

35 bucle - primera secuencia de ARN - secuencia saliente opcional de 2 nt - 3'. En forma alternativa, un

ARNsh puede comprender de 5' - primera secuencia de ARN - secuencia de bucle - segunda secuencia de ARN - secuencia saliente opcional de 2 nt - 3'. Dicha molécula de ARN forma pares de bases intramoleculares por medio de la primera y segunda secuencias de ARN sustancialmente complementarias. Secuencias de bucle adecuadas son bien conocidas en la técnica (entre otros, como se muestra en Dallas *et al.*, 2012 Nucleic Acids Res. Octubre de 2012; 40 (18): 9255-71, y en Schopman *et al.*, Antiviral Res. Mayo de 2010; 86 (2): 204-11).

La secuencia de bucle puede ser también una secuencia de tallo-bucle, por la cual la región de doble cadena del ARNsh es extendida. Sin que sea limitado por la teoría, al igual que el substrato de Dicer de ARNsi como se describió anteriormente, un ARNsh es usualmente procesado por Dicer para obtener, por ejemplo, un ARNsi que tiene un diseño de ARNsi tal como se describió anteriormente que tiene, por ejemplo, 19 pares de bases consecutivos y 2 salientes de nucleótidos en ambos extremos 3'. En caso de que el ARN de doble cadena sea procesado por Dicer, se prefiere tener la primera y segunda secuencias de ARN en el extremo.

Un ARN de doble cadena de conformidad con la invención puede ser incorporado también en un pre-ARNmi o andamio de pri-ARNmi. Las moléculas de micro ARN, es decir, ARNmi, son cadenas guía que se originan de moléculas de ARN de doble cadena que se expresan, por ejemplo, en células de mamífero. Un ARNmi es procesado de una molécula precursora de pre-ARNmi, similar al procesamiento de un ARNsh o un ARNsi extendido como se describió anteriormente, por la maquinaria de ARNi, e incorporado en un complejo de silenciamiento activado inducido por el ARN (RISC) (Tijsterman M, Plasterk RH. Dicers at RISC; the mechanism of RNAi Cell. Abril 2 de 2004; 117 (1): 1-3). Sin que sea limitado por la teoría, un pre-ARNmi es una molécula de horquilla que puede ser parte de una molécula de ARN más grande (pri-ARNmi), por ejemplo, comprendida en un intrón, que es procesada primero por Drosha para formar una molécula de horquilla de pre-ARNmi. La molécula de pre-ARNmi es una molécula tipo ARNsh que posteriormente puede ser procesada por Dicer para dar como resultado un dúplex de doble cadena tipo ARNsi. El ARNmi, es decir, la cadena guía, que es parte del dúplex de ARN de doble cadena, es incorporado posteriormente en el RISC. Una molécula de ARN tal como está presente en la naturaleza, es decir, un pri-ARNmi, un pre-ARNmi o un dúplex de ARNmi, puede usarse como un andamio para producir un ARNmi artificial que focaliza específicamente un gen de elección. Con base en la estructura de ARN predicha, por ejemplo, como se predijo usando, por ejemplo, el software m-fold, la secuencia de ARNmi natural como está presente en la estructura de ARN (es decir, dúplex, pre-ARNmi o pri-ARNmi), y la secuencia presente en la estructura que es complementaria con la misma, son removidas y reemplazadas con una primera secuencia de ARN y una segunda secuencia de ARN de conformidad con la invención. La primera secuencia de ARN y la segunda secuencia de ARN pueden seleccionarse de modo que las estructuras de ARN que se forman, es decir, pre-ARNmi, pri-ARNmi

y/o dúplex de ARNmi, se asemejan a las secuencias originales predichas correspondientes. El pre-ARNmi, pri-ARNmi y dúplex de ARNmi (que consisten de dos cadenas de ARN separadas que son hibridadas por medio de apareamiento de bases complementarias), como se encuentran en la naturaleza, con frecuencia no son completamente apareadas en sus bases, es decir, no todos los nucleótidos que se corresponden con la primera y segunda cadenas como se definió anteriormente son apareados en sus bases, y la primera y segunda cadenas con frecuencia no son de la misma longitud. Cómo usar moléculas precursoras de ARNmi como andamios para cualquier secuencia objetivo seleccionada y primera secuencia de ARN sustancialmente complementaria se describe, por ejemplo, en Liu YP *Nucleic Acids Res.* Mayo de 2008; 36(9): 2811-24.

En cualquier caso, como se desprende de lo anterior, el ARN de doble cadena que comprende la primera y segunda secuencias de ARN puede comprender nucleótidos y/o secuencias de nucleótidos adicionales. El ARN de doble cadena puede estar comprendido en una secuencia de ARN individual o comprendido en dos cadenas de ARN separadas. Sin que sea limitado por la teoría, cualquier diseño que se use para el ARN de doble cadena, es diseñado de modo que una secuencia antisentido que comprenda la primera secuencia de ARN de la invención puede ser procesada por la maquinaria de ARNi de modo que pueda ser incorporada en el RISC para que tenga su acción. Dicha secuencia que consista o que comprenda la primera secuencia de ARN de la invención, es capaz de focalizar específicamente la secuencia SEQ ID NO. 1. Por lo tanto, en tanto el ARN de doble cadena sea capaz de inducir la ARNi, dicho ARN de doble cadena se contempla en la invención. Por lo tanto, en una modalidad, el ARN de doble cadena de conformidad con la invención está comprendido en un andamio de pre-ARNmi, un andamio de pri-ARNmi, un ARNsh o un ARNsi.

El término complementario se define en la presente como los nucleótidos de una secuencia de ácido nucleico que pueden unirse a otra secuencia de ácido nucleico a través de enlaces de hidrógeno, es decir, los nucleótidos que son capaces de exhibir apareamiento de bases. Los ribonucleótidos, los bloques de construcción del ARN, están compuestos de monómeros (nucleótidos) que contienen un azúcar, fosfato y una base que es una purina (guanina, adenina) o pirimidina (uracilo, citosina). Las cadenas de ARN complementarias forman el ARN de doble cadena. Un ARN de doble cadena puede formarse de dos cadenas de ARN complementarias separadas, o las dos cadenas de ARN complementarias pueden estar comprendidas en una cadena de ARN. En las cadenas de ARN complementarias, los nucleótidos citosina y guanina (C y G) pueden formar un par de bases, guanina y uracilo (G y U), y uracilo y adenina (U y A). El término complementariedad sustancial significa que no se requiere que la primera y segunda secuencias de ARN sean completamente complementarias, o que la primera secuencia de ARN y SEQ ID NO. 1 sean completamente complementarias. Por ejemplo, el primer y segundo nucleótidos como se muestran en la figura 2A son sustancialmente complementarios y no completamente complementarios.

En una modalidad, la complementariedad sustancial entre la primera secuencia de ARN y SEQ ID NO. 1 consiste de no tener no apareamientos, un nucleótido no apareado o dos nucleótidos no apareados. Se entiende que un nucleótido no apareado significa que sobre la longitud entera de la primera secuencia de ARN que forma pares de bases con SEQ ID NO. 1, un nucleótido no forma pares de bases con SEQ ID NO. 1. El no tener no apareamientos significa que todos los nucleótidos forman pares de bases con SEQ ID NO. 1, y el tener 2 no apareamientos significa que dos nucleótidos no forman pares de bases con SEQ ID NO. 1. La primera secuencia de ARN puede ser también mayor de 21 nucleótidos, y en este escenario la complementariedad sustancial se determina sobre la longitud entera de SEQ ID NO. 1. Esto significa que SEQ ID NO. 1 en esta modalidad no tiene uno o dos no apareamientos sobre su longitud entera cuando forma pares de bases con la primera secuencia de ARN. Por ejemplo, como se muestra en la figura 3B y en los ejemplos, se pusieron a prueba moléculas de ARNs que tienen una primera secuencia de nucleótidos de 22 y 23 nucleótidos de longitud. Estas primeras secuencias de nucleótidos no tuvieron no apareamientos y fueron completamente complementarias a SEQ ID NO. 1. Puede permitirse el tener unos cuantos no apareamientos entre la primera secuencia de nucleótidos y SEQ ID NO. 1 de conformidad con la invención, en tanto el ARN de doble cadena de conformidad con la invención sea capaz de reducir la expresión de los transcritos que comprenden SEQ ID NO. 1, tal como un reportero de luciferasa o, por ejemplo, un transcrito que comprenda SEQ ID NO. 1. En esta modalidad, la complementariedad sustancial entre la primera secuencia de ARN y SEQ ID NO. 1 consiste de no tener uno o dos no apareamientos sobre la longitud entera de la primera secuencia de ARN o SEQ ID NO. 1, cualquiera sea la más corta.

En una modalidad, la primera secuencia de ARN y SEQ ID NO. 1 tienen por lo menos 15, 16, 17, 18 o 19 nucleótidos que forman pares de bases. De preferencia, la primera secuencia de ARN y SEQ ID NO. 1 son sustancialmente complementarias, dicha complementariedad comprendiendo por lo menos 19 pares de bases. En otra modalidad, la primera secuencia de ARN tiene por lo menos 8, 9, 10, 11, 12, 13 o 14 nucleótidos consecutivos que forman pares de bases con nucleótidos consecutivos de SEQ ID NO. 1. En otra modalidad, la primera secuencia de ARN tiene por lo menos 19 nucleótidos consecutivos que forman pares de bases con nucleótidos consecutivos de SEQ ID NO. 1. En otra modalidad, la primera secuencia de ARN comprende por lo menos 19 nucleótidos consecutivos que forman pares de bases con 19 nucleótidos consecutivos de SEQ ID NO. 1. En todavía otra modalidad, la primera secuencia de ARN tiene por lo menos 17 nucleótidos que forman pares de bases con SEQ ID NO. 1, y tiene por lo menos 15 nucleótidos consecutivos que forman pares de bases con nucleótidos consecutivos de SEQ ID NO. 1. La longitud de la secuencia del primer nucleótido es cuando mucho de 21, 22, 23, 24, 25, 26 o 27 nucleótidos.

Como se dijo, un no apareamiento de conformidad con la invención significa que un nucleótido de la primera secuencia de ARN no forma pares de bases con SEQ ID NO. 1. Los nucleótidos que no forman pares de bases son A y C, C y U o A y G. Un no apareamiento puede resultar también de una delección de un nucleótido, o una inserción de un nucleótido. Cuando el no

5 apareamiento es una delección en la primera secuencia de ARN, esto significa que un nucleótido de SEQ ID NO. 1 no forma pares de bases con la primera secuencia de ARN en comparación con la longitud entera de la primera secuencia de ARN. Los nucleótidos que pueden formar pares de bases son A-U, G-C y G-U. Un par de bases de G-U es referido también como un bamboleo de G-U, o un par de bases de bamboleo. En una modalidad, el número de pares de bases de G-U entre la primera

10 secuencia de ARN y SEQ ID NO. 1 es 0, 1 ó 2. En una modalidad, no existen no apareamientos entre la primera secuencia de ARN y SEQ ID NO. 1, y se permite un par de bases de G-U o pares de G-U. De preferencia, puede no haber pares de bases de G-U entre la primera secuencia de ARN y SEQ ID NO. 1, o la primera secuencia de ARN y SEQ ID NO. 1 tienen sólo pares de bases que son A-U o G-C. De preferencia, no existen pares de bases de G-U y no existen no apareamientos entre la primera

15 secuencia de ARN y SEQ ID NO. 1. Por lo tanto, la primera secuencia de ARN del ARN de doble cadena de conformidad con invención es de preferencia completamente complementaria a SEQ ID NO. 1, dicha complementariedad consistiendo de pares de bases G-C y A-U.

Puede no requerirse que exista complementariedad completa (es decir, apareamiento de bases completo (sin no apareamientos) y ningún par de bases G-U) entre la

20 primera secuencia de nucleótidos y SEQ ID NO. 1, y dicha primera secuencia de nucleótidos puede permitir aún supresión suficiente de la expresión génica. Asimismo, el no tener complementariedad completa puede contemplarse, por ejemplo, para evitar o reducir la supresión génica específica de la secuencia fuera del objetivo, mientras se mantiene la inhibición específica de la secuencia de los transcritos que comprenden SEQ ID NO. 1. Sin embargo, puede preferirse tener complementariedad

25 completa, puesto que puede resultar en inhibición más potente. Sin que sea limitado por la teoría, el tener complementariedad completa entre la primera secuencia de ARN y SEQ ID NO. 1 puede permitir que el RISC activado que comprende dicha primera secuencia de ARN escinda su secuencia objetivo, mientras que el tener no apareamientos puede permitir sólo la inhibición de la traducción, lo último resultando en inhibición menos potente.

30 En una modalidad, la primera secuencia de ARN tiene una longitud de secuencia de por lo menos 19 nucleótidos, de preferencia 20 nucleótidos, más preferiblemente de por lo menos 21 nucleótidos. La longitud de la secuencia puede ser también de por lo menos 22 nucleótidos, o por lo menos 23 nucleótidos. La primera secuencia de ARN de conformidad con la invención puede seleccionarse de SEQ ID NOs. 3-7.

CUADRO 1**Primeras secuencias de ARN**

SEQ ID NO.	Primera secuencia de ARN	Longitud
3	5'-AAGGACUUGAGGGACUCGA-3'	19
4	5'-AAGGACUUGAGGGACUCGAA-3'	20
5	5'-AAGGACUUGAGGGACUCGAAG-3'	21
6	5'-AAGGACUUGAGGGACUCGAAGG-3'	22
7	5'-AAGGACUUGAGGGACUCGAAGGC-3'	23

Se ha mostrado que las primeras secuencias de ARN del cuadro 1 inhiben específicamente a los transcritos que comprenden SEQ ID NO. 1 como se describe en la sección de ejemplos.

En una modalidad, la primera secuencia de nucleótidos del ARN de doble cadena de conformidad con la invención es completamente complementaria con SEQ ID NO. 1. Esto significa que no existen no apareamientos entre la primera secuencia de ARN y SEQ ID NO. 1 sobre la longitud entera de la primera secuencia de ARN o SEQ ID NO. 1, cualquiera que sea la más corta. De preferencia, la primera secuencia de nucleótidos y SEQ ID NO. 1 son completamente complementarias, comprendiendo sólo los pares de bases de G-C y A-U. De preferencia, la primera secuencia de ARN se selecciona del grupo que consiste de SEQ ID NOs. 3-7, que son completamente complementarias con SEQ ID NO. 1. Muy preferiblemente, la primera secuencia de ARN es SEQ ID NO. 5. Cuando la primera secuencia de nucleótidos es de 21 nucleótidos o menos (SEQ ID NOs. 3, 4 y 5), todos los nucleótidos de la primera secuencia de nucleótidos forman pares de bases con SEQ ID NO. 1. Cuando la primera secuencia de nucleótidos es mayor de 21 nucleótidos de longitud (SEQ ID NOs. 6 y 7), todos los nucleótidos de SEQ ID NO. 1 forman pares de bases con la primera secuencia de nucleótidos. Los nucleótidos adicionales que están comprendidos en la primera secuencia de ARN no forman pares de bases con SEQ ID NO. 1. Cuando la primera secuencia de nucleótidos es mayor de 21 nucleótidos y los nucleótidos adicionales serán parte de la cadena guía, de preferencia los nucleótidos adicionales son complementarios a la secuencia que flanquea la secuencia de SEQ ID NO. 1 como está presente en SEQ ID NO. 2.

Con respecto a la segunda secuencia de ARN, la segunda secuencia de ARN es sustancialmente complementaria con la primera secuencia de ARN. La segunda secuencia de ARN combinada con la primera secuencia de ARN forma un ARN de doble cadena. Como se dijo, este es para formar un substrato adecuado para la maquinaria de interferencia del ARN, de modo que una secuencia guía derivada de la primera secuencia de ARN está comprendida en el RISC a fin de inhibir específicamente la expresión de la secuencia de su objetivo, es decir, la expresión del gen de la

huntingtina. Como se dijo, dicho ARN de doble cadena está comprendido de preferencia en un andamio de pre-ARNmi, un andamio de pri-ARNmi, un ARNsh o un ARNsi.

La secuencia de la segunda secuencia de ARN tiene similitudes con la secuencia objetivo. Sin embargo, puede seleccionarse la complementariedad sustancial con la primera secuencia de ARN para que tenga menos complementariedad sustancial en comparación con la complementariedad sustancial entre la primera secuencia de ARN y SEQ ID NO. 1. Por lo tanto, la segunda secuencia de ARN puede comprender 0, 1, 2, 3, 4 o más no apareamientos, 0, 1, 2, 3 o más pares de bases de bamboleo de G-U, y puede comprender inserciones de 0, 1, 2, 3 o 4 nucleótidos y/o deleciones de 0, 1, 2, 3 o 4 nucleótidos. De preferencia, la primera secuencia de ARN y la segunda secuencia de ARN son sustancialmente complementarias, dicha complementariedad comprendiendo 0, 1, 2 o 3 pares de bases de G-U, y/o en donde dicha complementariedad comprende por lo menos 17 pares de bases.

Estos no apareamientos, pares de bases de bamboleo de G-U, inserciones y deleciones, son con respecto a la primera secuencia de ARN, es decir, la región de doble cadena que se forma entre la primera y segunda secuencias de ARN. En tanto la primera y segunda secuencias de ARN puedan sustancialmente formar pares de bases, y sean capaces de inducir la inhibición específica de la secuencia de SEQ ID NO. 1, se permite dicha complementariedad sustancial de conformidad con la invención. Se entiende también que sustancialmente la complementariedad entre la primera secuencia de ARN y la segunda secuencia de ARN puede depender del diseño del ARN de doble cadena de elección. Puede depender, por ejemplo, del andamio de ARNmi que se elige en el cual el ARN de doble cadena será incorporado.

CUADRO 2

Segundas secuencias de ARN

SEQ ID NO.	Segunda secuencia de ARN	Longitud
8	5'-UCGAGUCCCUCAAGUCCUU-3'	19
9	5'-UUCGAGUCCCUCAAGUCCUU-3'	20
10	5'-CUUCGAGUCCCUCAAGUCCUU-3'	21
11	5'-CCUUCGAGUCCCUCAAGUCCUU-3'	22
12	5'-GCCUUCGAGUCCCUCAAGUCCUU-3'	23
13	5'-CUUCGAGUCUCAAGUCCUU-3'	19
14	5'-ACGAGUCCCUCAAGUCCUC-3'	19

En una modalidad, una segunda secuencia de ARN se selecciona del grupo que consiste de SEQ ID NOS. 8-14. En el cuadro 2, se enlistan ejemplos de dichas segundas secuencias de ARN de conformidad con la invención. SEQ ID NOS. 8, 9, 10, 11 y 12 son completamente

complementarias con SEQ ID NO. 1 en toda su longitud. SEQ ID NOs. 13 y 14 pueden combinarse con un primer nucleótido que tenga una secuencia que corresponda a SEQ ID NO. 5 de 21 nucleótidos, que sea complementaria con SEQ ID NO. 1 en toda su longitud. SEQ ID NO. 13 es complementaria con SEQ ID NO. 5 que tiene una delección de dos nucleótidos (resultando en los 2
5 nucleótidos correspondientes de SEQ ID NO. 5 que no forma pares de bases) y 19 nucleótidos que forman pares de bases. SEQ ID NO. 14 es complementaria con SEQ ID NO. 5 que tiene una delección de dos nucleótidos, dos no apareamientos y 17 nucleótidos que forman pares de bases. La complementariedad puede verse también en la figura 2B, ya que la combinación de SEQ ID NO. 5 y SEQ ID NO. 13 está presente en miH12_155, y la combinación de SEQ ID NO. 5 y SEQ ID NO. 14
10 está presente en miH12_451a. Por lo tanto, como se desprende de lo anterior, la segunda secuencia de ARN no requiere complementariedad con la primera secuencia de ARN, pero puede comprender delecciones, inserciones y mutaciones que resultan en no apareamientos, en comparación con SEQ ID NO. 1.

Como se desprende de lo anterior, la complementariedad sustancial entre la primera
15 secuencia de ARN y la segunda secuencia de ARN puede comprender no apareamientos, delecciones y/o inserciones respecto a una primera y segunda secuencias de ARN que son completamente complementarias (es decir, completamente con pares de bases). En una modalidad, la primera y segunda secuencias de ARN tienen por lo menos 11 pares de bases consecutivos. Por lo tanto, por lo menos 11 nucleótidos consecutivos de la primera secuencia de ARN y por lo menos 11 nucleótidos
20 consecutivos de la segunda secuencia de ARN son completamente complementarios. En otra modalidad, la primera y segunda secuencias de ARN tienen por lo menos 15 nucleótidos que forman pares de bases. Dicho apareamiento de bases entre por lo menos 15 nucleótidos de la primera secuencia de ARN y por lo menos 15 nucleótidos de la segunda secuencia de ARN puede consistir de los pares de bases de G-U, G-C y A-U, o puede consistir de los pares de bases de G-C y A-U. En
25 todavía otra modalidad, la primera y segunda secuencias de ARN tienen por lo menos 15 nucleótidos que forman pares de bases, y tienen por lo menos 11 pares de bases consecutivos. En todavía otra modalidad, la primera secuencia de ARN y la segunda secuencia de ARN son sustancialmente complementarias, en donde dicha complementariedad comprende por lo menos 17 pares de bases. Dichos 17 pares de bases pueden ser de preferencia 17 pares de bases consecutivos, dicho
30 apareamiento de bases consistiendo de los pares de bases de G-U, G-C y A-U, o consistiendo de los pares de bases de G-C y A-U.

En una modalidad, la primera y segunda secuencias de nucleótidos se seleccionan del grupo de SEQ ID NOs. 3 y 8; 4 y 9; 5 y 10; 5 y 13; 5 y 14; 6 y 11; y 7 y 12. Se mostró que estas combinaciones de la primera y segunda secuencias de nucleótidos son efectivas cuando están
35 comprendidas en moléculas de ARNsi o andamios de ARNmi.

La primera y segunda secuencias de nucleótidos que son sustancialmente complementarias de preferencia no forman un ARN de doble cadena de 30 pares de bases consecutivos o más, ya que pueden desencadenar una respuesta inmune innata a través de la vía de la proteína cinasa activada por el ARN de doble cadena (ARNds). Por lo tanto, el ARN de doble

5 cadena es de preferencia de menos de 30 pares de bases consecutivos. De preferencia, un andamio de pre-ARNmi, un andamio de pri-ARNmi, un ARNsh, o un ARNsi que comprende el ARN de doble cadena de conformidad con la invención, no comprenden 30 pares de bases consecutivos.

De preferencia, el ARN de doble cadena de conformidad con la invención está comprendido en un andamio de pre-ARNmi o pri-ARNmi. Un andamio de pri-ARNmi comprende un

10 andamio de pre-ARNmi. El andamio de pre-ARNmi comprende el ARN de doble cadena de la invención, es decir, la primera secuencia de ARN y la segunda secuencia de ARN. De preferencia, el ARN de doble cadena de conformidad con la invención está comprendido en un andamio de pri-ARNmi derivado de miR-451a (referido también como miR-451) o miR-155. Ejemplos de moléculas de ARN de doble cadena de conformidad con la invención comprendidas en un andamio de pre-

15 ARNmi se representan en la figura 2A. Las secuencias de estas moléculas de pre-ARNmi se enlistan en el cuadro 3 a continuación.

CUADRO 3

Andamios de pre-ARNmi con SEQ ID NO. 5

20

SEQ ID NO.	Nombre	Secuencia
15	pre-miR451a	5'-CUUGGGAUUGGCAAGGAAGGACUUGAGGGACUCG AAGACGAGUCCCUCAAGUCCUCUCUUGCUAUACCCAGA-3'
16	pre-miR155	5'-UGCUGAAGGACUUGAGGGACUCGAAGGUUUUGGCCA CUGACUGACCUUCGAGUCUCAAGUCCUUCAGGA-3'

25

La secuencia de pre-ARNm de SEQ ID NO. 15 consiste de un brazo 5' que corresponde a los nucleótidos 1-16 de SEQ ID NO. 15, una primera secuencia de ARN que corresponde a SEQ ID NO. 5 de los nucleótidos 17-37 de SEQ ID NO. 15, una segunda secuencia de ARN que corresponde a SEQ ID NO. 14 de los nucleótidos 38-56 de SEQ ID NO. 15, y un brazo 3'

30 que corresponde a los nucleótidos 57-72 de SEQ ID NO. 15. El andamio de pre-miR451a de conformidad con la invención comprende el brazo 5' que corresponde a los nucleótidos 1-16 de SEQ ID NO. 15, seguido de la primera secuencia de nucleótidos de conformidad con la invención, la segunda secuencia de nucleótidos de conformidad con la invención, y el brazo 3' que corresponde a los nucleótidos 57-72 de SEQ ID NO. 15. De preferencia, la primera y segunda secuencias de ARN se

35 seleccionan de modo que cuando están comprendidas en el andamio de pri-miR451a, se obtiene una

estructura predicha altamente similar o como se muestra en la figura 2B. De preferencia, los pares de bases que se forman entre la primera y segunda secuencias de ARN son los pares de bases de G-C, G-U y A-U, y de preferencia la longitud de la secuencia de la primera secuencia de ARN es de 21 nucleótidos y la longitud de la segunda secuencia de ARN es de preferencia de 19 nucleótidos. De preferencia, los pares de bases que se forman entre la primera y segunda secuencias de ARN son los pares de bases de G-C y A-U, y de preferencia, la longitud de la secuencia de la primera secuencia de ARN es de 21 nucleótidos y la longitud de la segunda secuencia de ARN es de preferencia de 19 nucleótidos. Sin que sea limitado por la teoría, este andamio de pri-R451a puede preferirse, ya que no resulta en una cadena viajera que será procesada por la maquinaria de la ARNi para ser incorporada en el RISC (Cheloufi *et al.*, Junio 3 de 2010; 465 (7298): 584-9). A partir de un dúplex de ARNsi o ARNmi, en principio ambas cadenas pueden ser incorporadas en el RISC. Puesto que la cadena viajera (que corresponde a la segunda secuencia) puede resultar en la focalización de transcritos diferentes de un transcrito de la huntingtina, el uso de pri-miR451a o el andamio de pre-miR451a puede permitir evitar dicha focalización no deseada. Cuando se puso a prueba para la actividad potencial de la "cadena viajera", no se detectó actividad con un andamio de pri-451a (véase las figuras 4A y 4B). Se entiende que el procesamiento de la horquilla de pre-ARNmi es independiente de Dicer, y que será escindida por Ago2 (Yang *et al.*, Proc Natl Acad Sci U S A. Agosto 24 de 2010; 107 (34): 15163-8).

La secuencia de pre-ARNm de SEQ ID NO. 16 consiste de un brazo 5' que corresponde a los nucleótidos 1-5 de SEQ ID NO. 16, una primera secuencia de ARN que corresponde a SEQ ID NO. 5 de los nucleótidos 6-26 de SEQ ID NO. 16, una secuencia de bucle de los nucleótidos 27-45 de SEQ ID NO. 16, una segunda secuencia de ARN que corresponde a SEQ ID NO. 13 de los nucleótidos 46-64 de SEQ ID NO. 16, y un brazo 3' que corresponde a los nucleótidos 65-69 de SEQ ID NO. 16. El andamio de pre-155 que comprende la primera y segunda secuencias de ARN de conformidad con la invención, comprende el brazo 5' que corresponde a los nucleótidos 1-5 de SEQ ID NO. 16, la primera secuencia de ARN, una secuencia de bucle que corresponde a los nucleótidos 27-45 de SEQ ID NO. 16, la segunda secuencia de ARN, y un brazo 3' que corresponde a los nucleótidos 65-69 de SEQ ID NO. 16. De preferencia, la primera y segunda secuencias de ARN se seleccionan de modo que cuando están comprendidas en el andamio de pre-miR155 (o andamio de pri-ARNmi) se obtiene una estructura predicha altamente similar, como se muestra en la figura 2B. De preferencia, los pares de bases que se forman entre la primera y segunda secuencias de ARN son los pares de bases de G-C o A-U, y de preferencia la longitud de la secuencia de la primera secuencia de ARN es de 21 nucleótidos y la longitud de la segunda secuencia de ARN es de preferencia de 19 nucleótidos. Dichos 19 nucleótidos son de preferencia completamente

complementarios con los nucleótidos 2-18 de una primera secuencia de nucleótidos de 21 nucleótidos de longitud.

Una secuencia de pre-ARNmi cuando está comprendida en una secuencia de ARN más grande, requiere secuencias de flaqueo de cadena sencilla 5'- y 3'- que permitan el reconocimiento y la escisión por Drosha. Dichas secuencias que son adecuadas para permitir el reconocimiento y la escisión por Drosha son la secuencia 5'-pri-ARNmi y la secuencia 3'-pri-ARNmi. Por ejemplo, la secuencia pre-miH12_155 expresada como se representa en la figura 5A es flanqueada por 5'-pri-ARNmi y las secuencias 3'-pri-ARNmi de miR155 en el vector de expresión CMV-miH12-155 (figura 2C). La secuencia de pri-ARNmi que comprende miRH12_155 se enlista en la figura 2F, el 5'-pri-ARNmi que corresponde a los nucleótidos 1-87 y el 3'-pri-ARNmi que corresponde a 147-272. Del mismo modo, las secuencias de pri-ARNmi CAG-miH12-451 y PGK-miH12-451 que son expresadas por sus vectores respectivos en las figuras 2D y 2E se muestran en las figuras 2G y 2H. El 5'-pri-ARNmi y el 3'-pri-ARNmi del ARN de CAG-miH12-451 expresado corresponden a los nucleótidos 1-302 y 375-605, y el 5'-pri-ARNmi y el 3'-pri-ARNmi del ARN de PGK-miH12-451 expresado corresponden respectivamente a los nucleótidos 1-516 y 589-819. La longitud de los flancos de cadena sencilla puede variar, pero es típicamente de aproximadamente 80 nt (Zeng y Cullen, J Biol Chem. Julio 29 de 2005; 280 (30): 27595-603; Cullen, Mol Cell. Diciembre 22 de 2004; 16 (6): 861-5). La longitud mínima de los flancos de cadena sencilla puede determinarse fácilmente como cuando se vuelve demasiado corta. El procesamiento por Drosha puede fallar, y la inhibición específica de la secuencia será reducida o incluso estará ausente. En una modalidad, el andamio de pri-ARNmi que posee la primera y segunda secuencias de ARN de conformidad con la invención tiene un flanco de secuencia 5' y un flanco de secuencia 3' respecto a la estructura de re-ARNmi predicha de por lo menos 50 nucleótidos. El pre-ARNmi y las secuencias derivadas de pri-ARNmi se derivan todos de preferencia de la misma secuencia de pri-ARNmi de ocurrencia natural.

La secuencia de pre-ARNmi de SEQ ID NO. 15 y SEQ ID NO. 16 es codificada por las secuencias de ADN como se representan en las figuras 2C (SEQ ID NO. 16), 2D (SEQ ID NO. 15) y 2E (SEQ ID NO. 15). Las secuencias de pri-ARNmi que comprenden dichas secuencias de pre-ARNmi se representan en las figuras 2F (SEQ ID NO. 16), 2G (SEQ ID NO. 15) y 2H (SEQ ID NO. 15). El pri-ARNmi codificado por CMV-miH12-155 corresponde a los nucleótidos 1433-1704 de la figura 2C, de CAG-miH12-451 a los nucleótidos 1716-2320 de la figura 2D y de PGK-miH12-451 a los nucleótidos 278-1097 de la figura 2E. Del mismo modo, la primera y segunda secuencias de ARN serán incorporadas como se describió anteriormente para las moléculas de pre-ARNmi.

Las moléculas de ARN de doble cadena de conformidad con la invención, incorporadas en un ARNsi, ARNsh, andamio de pri-ARNm o andamio de pre-ARNmi pueden proveerse en una célula usando métodos conocidos en la técnica, tales como lipofección,

transfección o usando cualquier otro medio adecuado para lo mismo. Las moléculas de ARN de doble cadena de conformidad con la invención pueden ser moléculas de ARN de doble cadena sintéticas o moléculas de ARN de doble cadena naturales. Las moléculas de ARN de doble cadena sintéticas pueden comprender ácidos nucleicos que contienen análogos conocidos de nucleótidos naturales.

5 Dicho ARN de doble cadena tiene propiedades similares en comparación con sus contrapartes naturales, y una actividad de interferencia del ARN similar a o mejorada sobre las moléculas de ARN de doble cadena que consisten enteramente de ARN de doble cadena no sintético. Por ejemplo, las moléculas de ARNsi sintéticas pueden incluir en su diseño el uso de ácido nucleico bloqueado (un anillo de ribosa unido por un puente de metileno (naranja) entre los átomos de 2'-O y 4'-C),

10 nucleótidos modificados tales como nucleótidos que comprenden fosforotioatos, 2'-O-Me, 2'-O-alilo y 2'-desoxi-fluorouridina. Es bien sabido que las moléculas de ARN de doble cadena, por ejemplo, las moléculas de ARNsi, pueden acomodar un gran número de modificaciones en las posiciones con apareamiento de bases y sin apareamiento de bases sin pérdida significativa de actividad. De preferencia, el ARN de doble cadena de la invención es un ARN de doble cadena que consiste de

15 nucleótidos naturales, tales como los obtenidos de la expresión de un ARN de doble cadena.

Por lo tanto, en una modalidad, dichas moléculas de ARN de doble cadena de la invención son codificadas por una secuencia de ADN. Dicha secuencia de ADN que codifica para dicho ARN de doble cadena, por ejemplo, como el comprendido en un ARNsi, ARNsh, andamio de pri-ARNm o andamio de pre-ARNmi, está comprendida en un cassette de expresión. Se entiende que

20 cuando el ARN de doble cadena será, por ejemplo, un ARNsi, que consiste de dos cadenas de ARN, existen dos cassettes de expresión requeridos: Uno que codifica para una cadena de ARN que comprende la primera secuencia de ARN, y el otro cassette que codifica para una cadena de ARN que comprende la primera cadena de ARN. Cuando el ARN de doble cadena está comprendido en una sola molécula de ARN, por ejemplo, que codifica para un ARNsh, pre-ARNmi o pri-ARNmi, un

25 cassette de expresión puede ser suficiente. Un cassette de expresión de pol II puede comprender una secuencia de promotor y una secuencia que codifica para el ARN que será expresado, seguido de una secuencia de poliadenilación. En caso de que el ARN de doble cadena que es expresado comprenda un andamio de pri-ARNmi, la secuencia de ARN codificada puede codificar para las secuencias de intrón y las secuencias de exón y de UTR 3' y UTRs 3'. Un cassette de expresión de

30 pol III comprende en general una secuencia de promotor, seguida de la secuencia de ADN que codifica para el ARN (por ejemplo, secuencia de ARNsh, pre-ARNmi, o una cadena de las moléculas de ARN de doble cadena estará comprendida, por ejemplo, en un ARNsi o ARNsi extendido). Un cassette de expresión de pol I puede comprender un promotor de pol I, seguido de la secuencia codificante de ARN y una caja 3'. Los cassettes de expresión para las moléculas de ARN de doble

35 cadena son bien conocidos en la técnica, y cualquier tipo de cassette de expresión pueden ser

suficiente, por ejemplo, puede usarse un promotor de pol III, un promotor de pol II o un promotor de pol I (véase, entre otros, a ter Brake *et al.*, Mol Ther. Marzo de 2008; 16 (3): 557-64; y Maczuga *et al.*, BMC Biotechnol. Julio 24 de 2012; 12:42). Ejemplos de cassettes de expresión que expresan un ARN de doble cadena de conformidad con la invención, se representan en las figuras 2C-2E.

5 De preferencia, se usa un promotor de pol II, tal como el promotor de PGK, un promotor de CBA o un promotor de CMV (véase las figuras 2C-2D). Puesto que la enfermedad de Huntington afecta a las neuronas, puede ser particularmente útil usar un promotor neuroespecífico. Ejemplos de promotores neuroespecíficos adecuados son la enolasa específica de neuronas (NSE), la sinapsina 1 humana, la caMK cinasa y la tubulina. Otros promotores adecuados que pueden
10 contemplarse son los promotores inducibles, es decir, un promotor que inicia la transcripción sólo cuando la célula hospedera es expuesta a algún estímulo particular.

Dichos cassettes de expresión de conformidad con la invención pueden ser transferidos a una célula, por ejemplo, usando métodos de transfección. Cualquier medio adecuado puede ser suficiente para transferir un cassette de expresión de conformidad con la invención. De
15 preferencia, se usan vectores de terapia génica que transfieren establemente el cassette de expresión a las células de modo que pueda lograrse la expresión estable de las moléculas de ARN de doble cadena que inducen la inhibición específica de la secuencia del gen de la huntingtina como se describió anteriormente. Vectores adecuados pueden ser vectores lentivirales, sistemas de vectores basados en retrotransposones, o vectores de AAV. Se entiende que, por ejemplo, los vectores
20 lentivirales poseen un genoma de ARN y el genoma de ARN codificará para dicho cassette de expresión de modo que, después de la transducción de una célula, dicha secuencia de ADN y dicho cassette de expresión se forman. De preferencia, se usa un vector viral tal como un AAV. De preferencia, el vector de AAV que se usa es un vector de AAV de serotipo 5. El AAV de serotipo 5 puede ser particularmente útil para la transducción de las neuronas, como se muestra en los
25 ejemplos. La producción de vectores de AAV que comprenden cualquier cassette de expresión de interés se describe, por ejemplo, en los documentos WO2007/046703, WO2007/148971, WO2009/014445, WO2009/104964, WO2011/122950 y WO2013/036118, los cuales se incorporan en la presente en su totalidad.

Las secuencias de AAV que pueden usarse en la presente invención para la
30 producción de vectores de AAV, por ejemplo, producidos en líneas de células de insecto o de mamífero, pueden derivarse del genoma de cualquier serotipo de AAV. En general, los serotipos de AAV tienen secuencias genómicas de homología significativa a los niveles de aminoácidos y ácidos nucleicos, proveen una serie idéntica de funciones genéticas, producen viriones que son en esencia físicamente y funcionalmente equivalentes, y se replican y ensamblan por mecanismos
35 prácticamente idénticos. Para la secuencia genómica de los varios serotipos de AAV y una visión

general de las similitudes genómicas véase, por ejemplo, el número de acceso del GenBank U89790; número de acceso del GenBank J01901; número de acceso del GenBank AF043303; número de acceso del GenBank AF085716; Chlorini *et al.* (1997, J. Vir. 71: 6823-33); Srivastava *et al.* (1983, J. Vir. 45: 555-64); Chlorini *et al.* (1999, J. Vir. 73: 1309-1319); Rutledge *et al.* (1998, J. Vir. 72: 309-319); y Wu *et al.* (2000, J. Vir. 74: 8635-47). Los serotipos de AAV 1, 2, 3, 4 y 5 son la fuente preferida de las secuencias de nucleótidos de AAV para su uso en el contexto de la presente invención. De preferencia, las secuencias de ITR de AAV para su uso en el contexto de la presente invención se derivan de AAV1, AAV2 y/o AAV5. Asimismo, las secuencias que codifican para Rep52, Rep40, Rep78 y/o Rep68 se derivan de preferencia de AAV1, AAV2 y AAV5. Las secuencias que codifican para las proteínas de la cápside VP1, VP2 y VP3 para su uso en el contexto de la presente invención pueden tomarse sin embargo de cualquiera de los 42 serotipos conocidos, más preferiblemente de AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 o AAV9 o las partículas tipo AAV recién desarrolladas obtenidas, por ejemplo, por técnicas de entremezclado de la cápside y colecciones de cápsides de AAV.

En otra modalidad, se provee una célula hospedera que comprende dicha secuencia de ADN, o dicho cassette de expresión de conformidad con la invención. Por ejemplo, dicho cassette de expresión o secuencia de ADN puede estar comprendido en un plásmido contenido en bacterias. Dicho cassette de expresión o secuencia de ADN puede estar comprendido también en una célula de producción que produce, por ejemplo, un vector viral.

Como se muestra en la sección de ejemplos, y como se explicó anteriormente, el ARN de doble cadena de conformidad con la invención, la secuencia de ADN de conformidad con la invención, el cassette de expresión de conformidad con la invención y el vector de terapia génica de conformidad con la invención, son para su uso en un tratamiento médico, en particular para su uso en el tratamiento de la enfermedad de Huntington. Dicho tratamiento médico cuando se usa un vector de AAV (o igualmente para un vector de terapia génica) comprende una infusión directa del vector de AAV de la invención en el cerebro. Dicha infusión directa en otras modalidades comprende una infusión intratecal del vector en el fluido cerebroespinal. Dicha infusión intratecal representa una forma eficiente para suministrar el vector de terapia génica hacia el SNC y para focalizar las neuronas. De preferencia, las estructuras corticales y del cuerpo estriado son focalizadas por medio del suministro de vectores de AAV por difusión mejorada por convección (CED) intraestriatal a través de inyecciones en el cuerpo estriado. Más preferiblemente, para una mayor cobertura del SNC, las inyecciones son en el cuerpo estriado y también en el tálamo. Por lo tanto, los vectores de AAV se suministran por vía intraestriatal, o se suministran por vía intraestriatal e intratálamica a través de inyecciones por difusión mejorada por convección (CED) en el cuerpo estriado, o el cuerpo estriado y el tálamo. Dichas inyecciones se llevan a cabo de preferencia a través de inyecciones guiadas por

MRI. Dichos métodos de tratamiento son útiles en particular para sujetos humanos que tienen la enfermedad de Huntington. Se entiende que el tratamiento de la enfermedad de Huntington implica sujetos humanos que tienen la enfermedad de Huntington, incluyendo sujetos humanos que tienen una predisposición genética de desarrollar la enfermedad de Huntington que aún no muestran los
 5 signos de la enfermedad. Por lo tanto, el tratamiento de sujetos humanos que padecen la enfermedad de Huntington incluye el tratamiento de cualquier sujeto humano que porta un alelo de la huntingtina con más de 35 repeticiones de CAG.

Modalidades

- 10 1. Un ARN de doble cadena que comprende una primera secuencia de ARN y una segunda secuencia de ARN en donde la primera y segunda secuencias de ARN son sustancialmente complementarias, en donde la primera secuencia de ARN tiene una longitud de secuencia de por lo menos 19 nucleótidos y es sustancialmente complementaria con SEQ ID NO. 1.
- 15 2. Un ARN de doble cadena de conformidad con la modalidad 1, en donde dicho ARN de doble cadena es capaz de reducir la expresión del gen de la huntingtina.
3. Un ARN de doble cadena de conformidad con la modalidad 1 o la modalidad 2, en donde el ARN de doble cadena está comprendido en un andamio de pre-ARNmi, un andamio de pri-ARNmi, un ARNsh o un ARNsi, de preferencia un andamio de pre-ARNmi.
- 20 4. Un ARN de doble cadena de conformidad con cualquiera de las modalidades 1 a 3, en donde la primera secuencia de ARN tiene una longitud de secuencia de por lo menos 20 nucleótidos, de preferencia de por lo menos 21 nucleótidos.
5. Un ARN de doble cadena de conformidad con cualquiera de las modalidades 1 a 4, en donde la primera secuencia de ARN es completamente complementaria a SEQ ID NO. 1.
- 25 6. Un ARN de doble cadena de conformidad con cualquiera de las modalidades 1 a 5, en donde la primera cadena se selecciona del grupo que consiste de SEQ ID NO. 3, SEQ ID NO. 4, SEQ ID NO. 5, SEQ ID NO. 6 y SEQ ID NO. 7.
7. Un ARN de doble cadena de conformidad con cualquiera de las modalidades 1 a 6, en donde la primera cadena y la segunda cadena se seleccionan del grupo que consiste de las combinaciones de SEQ ID NO. 3 y 8; SEQ ID NO. 4 y 9; SEQ ID NO. 5 y 10; SEQ ID NO. 5 y 13; SEQ
 30 ID NO. 5 y 14; SEQ ID NO. 6 y 11; y SEQ ID NO. 7 y 12.
8. Un ARN de doble cadena de conformidad con cualquiera de las modalidades 1 a 7, en donde el ARN de doble cadena está comprendido en un andamio de pre-ARNmi derivado de miR-451a o miR-155.
- 35 9. Una secuencia de ADN que codifica para el ARN de doble cadena de conformidad con cualquiera de las modalidades 1 a 8.

10. Un cassette de expresión que codifica para un ARN de doble cadena de conformidad con cualquiera de las modalidades 1 a 9.

11. Un cassette de expresión de conformidad con la modalidad 10, en donde el cassette de expresión comprende el promotor de PGK, un promotor de CMV, un promotor neuroespecífico o un promotor de CBA.

12. Un vector de terapia génica que comprende el cassette de expresión de conformidad con la modalidad 10 u 11.

13. Un vector de terapia génica de conformidad con la modalidad 12, en donde el vector de terapia génica es un vector de AAV, de preferencia un vector de AAV de serotipo 5.

14. Una célula hospedera que comprende la secuencia de ADN de conformidad con la modalidad 9 o el cassette de expresión de conformidad con la modalidad 10 o la modalidad 11.

15. Un ARN de doble cadena de conformidad con cualquiera de las modalidades 1 a 8, una secuencia de ADN de conformidad con la modalidad 9, un cassette de expresión de conformidad con la modalidad 10 o la modalidad 11, y un vector de terapia génica de conformidad con la modalidad 12 o la modalidad 13, para su uso en un tratamiento médico.

16. Un ARN de doble cadena de conformidad con cualquiera de las modalidades 1 a 8, una secuencia de ADN de conformidad con la modalidad 9, un cassette de expresión de conformidad con la modalidad 10 o la modalidad 11, y un vector de terapia génica de conformidad con la modalidad 12 o la modalidad 13, para su uso en el tratamiento de la enfermedad de Huntington.

EJEMPLOS

Construcciones de expresión de andamios de ARNmi y moléculas de ARNsi

Para crear vectores de andamios de ARNmi basados en miR155, 21 secuencias de nucleótidos (pares de bases) completamente complementarias con las secuencias objetivo seleccionadas en HTT como se indica en la figura 1, fueron incluidas en la estructura de base de pri-mir-155 de pcDNA6.2-GW/EmGFP-miR (Invitrogen, Carlsbad, CA), resultando en pVD-CMV-miHTT-155 (un ejemplo de una secuencia de cassette de expresión se representa en la figura 2C, las secuencias de pre-ARNmi y pri-ARNmi como están comprendidas en el ARN expresado se representan en las figuras 2B y 2F, respectivamente). Las construcciones de pri-mir-155 se diseñaron con base en las instrucciones provistas por Invitrogen (BLOCK-iT, kits de vectores de expresión de ARNi de Pol II de miR, versión E, Junio 22 de 2007, 25-0857) uniendo oligonucleótidos de doble cadena sintéticos en el sitio Bsal de pcDNA6.2-GW/emGFP-mir155. La estructura de todos los pre-ARNmi artificiales codificados por las construcciones de miHtt se verificó usando el software

Mfold (Nucleic Acids Res. 31 (13), 3406-15, (2003), usando la versión 3.5 de Mfold, disponible en línea en <http://mfold.rna.albany.edu/?q=mfold>). La estructura predicha del andamio de pre-ARNmi-155 que posee la secuencia correspondiente con SEQ ID NO. 5 se muestra en la figura 2B. La secuencia de ADN que codifica para la construcción de expresión para el andamio de pri-ARNmi-155 con SEQ ID NO. 5 como una primera secuencia seleccionada, se enlista en la figura 2C (SEQ ID NO. 17), y esta construcción es referida también como miH12 y su objetivo H12. Para las otras 20 secuencias objetivo seleccionadas, las secuencias fueron diseñadas para ser completamente complementarias con las secuencias objetivo como se representa en la figura 1, y el andamio de ARNmi que tiene las mismas características estructurales como se representa en la figura 2B, es decir, la segunda secuencia de ARN embebida en el andamio, corresponde a la secuencia para la cual la primera secuencia de ARN seleccionada es completamente complementaria, pero que tiene una delección de 2 nucleótidos en el centro. Como control de manera similar, una secuencia de ARN mezclada se usó como una primera secuencia de ARN, y como anteriormente, una segunda secuencia de ARN fue diseñada para crear un vector pVD-CMV-miScr-155. Las construcciones contenían GFP para permitir la expresión del ARNmi y la visualización de la transducción *in vitro* e *in vivo*.

Se creó un vector de andamio de ARNmi basado en miR451a. La secuencia de ADN que codifica para el andamio de pri-miR-451 se sintetizó con base en la secuencia predicha madura de mir-451 que es reemplazada por la secuencia de focalización de H12, es decir, SEQ ID NO. 5 como primera secuencia de ARN. La segunda secuencia de ARN fue diseñada para ser completamente complementaria a los nucleótidos 2-18 de la primera secuencia de ARN. La segunda secuencia de ARN se seleccionó de modo que la estructura de ARN predicha de la secuencia de pre-ARNmi artificial adoptara una estructura similar a la estructura de tipo silvestre original. La estructura del pre-ARNmi codificado por las construcciones se verificó usando el software Mfold (Nucleic Acids Res. 31 (13), 3406-15, (2003), usando la versión 3.5 de Mfold, disponible en línea en <http://mfold.ma.albany.edu/?q=mfold>). La estructura predicha usando a SEQ ID NO. 5 se muestra en la figura 2B. Se obtuvieron dos vectores que expresan andamios de miR451 diferentes. Las secuencias de ADN de las construcciones de expresión se representan en las figuras 2D y 2E, y las secuencias de andamios de pri-ARNmi respectivas correspondientes comprendidas en el ARN expresado se enlistan en las figuras 2G y 2H. La figura 2D muestra la secuencia de ADN del cassette de expresión de pVD-CAG-miH12-451, que expresa el andamio de ARNmi con un promotor de CAG, y en la figura 2E se muestra la secuencia de ADN del cassette de expresión de pVD-PGK-miH12-451, que usa un promotor de PGK.

Se diseñó ARNsi sintético que focaliza a HTT en el objetivo miH12, es decir, SEQ ID NO. 1, con longitudes de 19-23 pb (cuadro 1). Las moléculas de ARNsi comprendían la primera y

segunda secuencias de nucleótidos que corresponden a SEQ ID NOs. 3 y 8; 4 y 9; 5 y 10; 6 y 11; y 7 y 12. Las moléculas de ARNsi fueron diseñadas para tener salientes de 3'-UU en ambas cadenas.

Construcciones de reportero

5 Las construcciones de LucHTT psiCheck-2 que contienen la secuencia completa del exón 1 de HTT fueron diseñadas y clonadas siguiendo las instrucciones provistas por Promega (vectores siCHECK™, C8011, Promega Benelux b.v., Leiden, Los Países Bajos). LucHTT abarca a SEQ ID NO. 1 y las secuencias de flanco de la misma como están presentes en SEQ ID NO. 2 (véase la figura 1B). El reportero de LucH12_451a comprende la secuencia complementaria a la segunda
10 secuencia de ARN diseñada para ser expresada por pVD-CAG-miH12-451 y pVD-PGK-miH12-451. Todas las construcciones han sido secuenciadas, y la secuencia correcta ha sido verificada. La eficacia de supresión de la expresión de todos los andamios de ARNmi y moléculas de ARNsi se determinó en reporteros específicos de luciferasa *in vitro*. Células Hek293T fueron co-transfectadas con la miHtt y el reportero de luciferasa a una relación 1:1 (miR-155, PGK-miH12-451) o relación
15 1:10 (CAG-miH12-451, es decir, el promotor de CAG es muy fuerte). Se midió la supresión de la expresión de la luciferasa de Renilla 48 horas post-transfección (p.t.), y se midió la luciferasa de luciérnaga como un control interno. Se usó a miScr como control negativo, y se ajustó a 100%.

Resultados *in vitro*

20 Entre las construcciones de miH1-miH21 que focalizan el exón 1, miH12 indujo la más fuerte supresión de la expresión del reportero de luciferasa con una reducción de 75-80% (figura 3A). Se mostró que todas las moléculas de ARNsi que focalizan a H12, es decir, SEQ ID NO. 1, tienen una eficiencia de supresión de la expresión similar, mostrando tras un incremento en la dosis una supresión de la expresión más fuerte. Las moléculas de ARNsi de 19 y 21 pares de bases
25 mostraron un poco más de inhibición en comparación con las otras moléculas de ARNsi puestas a prueba. Se realizó un análisis de la secuenciación de la siguiente generación (NGS) del ARN expresado de pVD-CMV-miHTT-155, y mostró una preferencia en las cadenas guía y viajera (véase, por ejemplo, la figura 2B (7)) que corresponde a la primera secuencia de ARN y la segunda secuencia de ARN diseñadas para ser parte del andamio de ARNmi. Se pusieron a prueba pVD-CAG-
30 miH12-451 y pVD-PGK-miH12-451 para la focalización de LucHTT y LucH12_451a. Las construcciones de CAG y PGK mostraron inhibición específica de la secuencia (figura 4A), mientras que un reportero de Luc que estaba comprendiendo una secuencia completamente complementaria a la segunda secuencia de ARN de las construcciones no fue reducido.

Supresión de la expresión *in vivo* usando vectores de AAV en ratones

La construcción de expresión CMV-miH12-155 de pVD-CMV-miHTT-155 fue clonada en una estructura de base de vector de AAV5 y AAV5 producido usando el sistema de producción de baculovirus. Como control, CMV-miScr-155 fue incorporado también en una estructura de base de AAV5 para servir como un control negativo. Para el monitoreo de la eficiencia de transducción en el cerebro, los cassettes de expresión contenían GFP (figura 5A). Un vector de AAV-5 posee una construcción de reportero de luciferasa, Luc73QHtt, que comprende la secuencia objetivo SEQ ID NO. 1 (es decir, la secuencia completa del exón 1 de HTT con 73 repeticiones de CAG) y las secuencias de flanqueo del mismo como están presentes en SEQ ID NO. 2 (véase la figura 1B). Ratones Balb/6 (N5) fueron co-inyectados por la vía intraestriatal con 2 μ l de AAV5-Luc73QHtt/SNP (3.6×10^{12} gc/ml), y AAV5-miScr (1.8×10^{13} gc/ml) o AAV5-miH12 (1.8×10^{13} gc/ml) a una relación 1:5. Un grupo separado fue inyectado con AAV5-Luc73QHtt/SNP y PBS. La expresión de luciferasa se monitoreó a 2, 4 y 6 días p.i. por MS. Ya a 1 semana p.i., había una clara supresión de la expresión de Luc19QHtt/wt por miH12 en comparación con sólo miScr y Luc19QHtt/wt (figuras 7B y 7C). Se mostró una tendencia que indica una disminución significativa en la expresión del reportero de luciferasa en el tiempo, siendo casi indetectable (miH12 # 1 y # 2) en el cerebro en comparación con los grupos control, indicando una fuerte supresión de la expresión de la HTT objetivo por CMV-miH12-155. Al final del experimento, la fluorescencia de Luc73QHtt/SNP en el grupo de CMV-miH12-155 fue aproximadamente 1 log menor en comparación con miScr.

20

Supresión de la expresión *in vivo* usando vectores de AAV en modelos animales de HD

Se puso a prueba AAV5-CMV-miH12-155 en el modelo de rata de LV-171-82Q HD (Drouet *et al.*, 2009, Ann Neurol. Marzo de 2009; 65 (3): 276-85). El modelo se basa en la sobreexpresión, en el cuerpo estriado, de los primeros 171 aminoácidos del fragmento mutante de la HTT con 82 repeticiones de CAG enlazado a un fragmento del exón 67 que contiene el SNP C/T. Se inyectó a ratas con HD con AAV5-CMV-miH12-155 o un control de AAV5-CMV-miScr-155 (figura 6A). Las ratas fueron inyectadas por la vía intraestriatal con LV-171-82Q y una semana después con AAV5-CMV-miH12-155 o AAV5-CMV-miScr-155. El efecto neuroprotector de AAV5-CMV-miH12-155 se determinó con base en la tinción histológica (DARP32, EM48, GFP e Iba1) de secciones de cerebro de rata con HD en los puntos de tiempo temprano y tardío (figuras 6B, 6C y 6D, 2 semanas p.i.). Las lesiones indicadas con DARP32 (figura 6B, panel superior) indican neurodegeneración, y pueden observarse como manchas blancas en las secciones del cerebro. El panel muestra claramente menos muerte neuronal y por ende la ausencia de manchas blancas en las secciones del cerebro de ratas AAV5-CMV-miH12-155. Las secciones de cerebro fueron teñidas para agregados de HTT mutantes

35

(EM48, figura 6B) vistos como pequeños puntos marrones en los portaobjetos. Hubo claramente menos neurodegeneración y menos agregados de HTT mutantes en ratas con HD inyectadas con AAV5-CMV-miH12-155 en comparación con el grupo control (figura 6B, paneles intermedio y de más abajo). Se obtuvieron resultados similares en el punto de tiempo tardío de 8 semanas post-inyección (datos no mostrados). Los resultados de la histología con la GFP (tinción marrón) indicaron transducción estriatal eficiente con todos los vectores usados en el estudio actual (figura 6C). Además, casi no se detectó respuesta inmune a las 8 semanas p.i. con base en la tinción con Iba1 (figura 6D).

AAV5-CMV-miH12-155 fue posteriormente puesto a prueba adicionalmente en el modelo de ratón con HD Hu97/18 humanizado (Southwell *et al.*, 2013, Hum Mol Genet. Enero 1 de 2013; 22 (1): 18-34). Este modelo tiene el gen Hdh de murino reemplazado por dos copias de la HTT humana, una que posee 97 repeticiones de CAG y la otra 18. Se ha realizado la caracterización detallada de los cambios motores, psiquiátricos, cognitivos, electrofisiológicos, neuropatológicos y bioquímicos en el modelo de ratón Hu97/18 como resultado de progresión de la enfermedad. Se inyectó AAV5-CMV-miH12-155 en el modelo de ratón con HD Hu97/18 humanizado de 2 meses por suministro a través del suministro intracerebral por medio de difusión mejorada por convección (IC-CED) en el cuerpo estriado, o suministro lento (IC-suministro lento) en el cuerpo estriado, o suministro intracerebroventricular (ICV) en los ventrículos del cerebro. La fluorescencia de la GFP indicó la transducción completa del cuerpo estriado del ratón tras el suministro lento y por CED (figura 7A). El análisis Western blot de la HTT humana mostró supresión de la expresión por AAV5-CMV-miH12-155 en el cuerpo estriado cuando se aplicó suministro por CED (figura 7B).

Comparación de H12 con secuencias objetivo de la técnica anterior

Se compararon las secuencias objetivo de la técnica anterior en la proximidad de la secuencia de H12 con las secuencias de H12. Se construyeron moléculas de ARNsi y andamios de ARNmi, y se llevó a cabo una comparación directa usando el sistema de reportero de luciferasa como se describió anteriormente. Bajas concentraciones de moléculas de ARNsi fueron transfectadas (0.25 nM) por triplicado a fin de evitar los efectos fuera de objetivo que sesgan los resultados. Las moléculas de ARNsi se obtuvieron con cadenas guía y viajera completamente complementarias (pares de bases de G-C y A-U) y que tienen una saliente UU-3' en ambas cadenas. Se usó un ARNsi revuelto como control, y los valores se midieron respecto al control. El ARNsi H12 mostró una inhibición más fuerte (véase la figura 8A). Del mismo modo, se construyeron andamios de ARNmi a base de miR-155 como se describió anteriormente de acuerdo con las instrucciones de Invitrogen, y se hizo también una comparación directa. Los andamios R6.1 y R6.2 se hicieron reemplazando 19 y 18 nucleótidos que son perfectamente complementarios a las secuencias objetivo de R6.1 y R6.2 en

la secuencia guía del andamio de miR-155 diseñado de Invitrogen. Por lo tanto, dependiendo del procesamiento de pre-ARNmi por Dicer, la cadena guía R6 procesada puede contener nucleótidos del andamio de miR-155 en los extremos de la secuencia. Los andamios de ARNmi fueron transfectados como se describió anteriormente usando diferentes relaciones entre la construcción de ARNmi y el reportero (construcción de andamio de ARNmi: reportero de luciferasa). Una construcción de ARNmi revuelta se usó como control, y los valores se midieron respecto al control. La figura 8B muestra una relación 1:1, mientras que la figura 8C muestra una relación 1:10. El andamio de ARNmi H12 mostró una inhibición más fuerte. H12 mostró fuerte inhibición pronunciada para las moléculas de ARNsi y las moléculas de ARNmi, en particular, a bajas concentraciones que pueden considerarse más relevantes para aplicación *in vivo*.

CUADRO 4

Objetivos de la técnica anterior en comparación con H12. Se muestran las secuencias objetivo

Objetivo	Secuencia objetivo	SI.	L.	Id.
H12	5'-CUUCGAGUCCCUCAAGUCCUU-3'	34	21	21
H11	5'-GAAGGCCUUCGAGUCCCUCAA-3'	33	21	15
R1 (siHUNT-2)	5'-GGCCUUCGAGUCCCUCAAGUCC-3'	46	21	18
R2 (shD2)	5'-GGCCUUCGAGUCCCUCAAGUC----3'	47	21	18
R3 (ARNsi-DExón1)	5'-AGGCCUUCGAGUCCCUCAAGU-3'	48	21	17
R4 (HDAS 07)	5'-AUGAAGGCCUUCGAGUCCCU-3'	49	21	13
R6.1 (54)	5'-GCCUUCGAGUCCCUCAAGU-3'	50	19	17
R6.2 (55)	5'-CCUUCGAGUCCCUCAAGU-3'	51	18	17

Los nucleótidos objetivo que tienen identidad con la secuencia objetivo H12, están subrayados. (SI. indica SEQ ID NO.; L. indica la longitud de nucleótidos, Id. indica el número de nucleótidos que tienen identidad con H12). (R1 se deriva del ARNsi siHUNT-2 de Rodríguez-Lebron *et al.*, 2005, Mol Ther. Vol. 12 No. 4: 618-633, R2 se deriva de un ARNsh expresado shD2 de Franich *et al.*, 2008, Mol Ther. Vol. 16 No. 5; 947-956), R3 se deriva del ARNsi-DExón1 del documento US20080015158, R4 se deriva de HDAS 07 del documento WO2008134646, R6.1 y R6.2 se derivan de una lista de aproximadamente 1750 moléculas de ARNsi hipotéticas diseñadas para focalizar el gen de la huntingtina (véase el documento WO2005105995).