

MOLÉCULAS DE TRAN-SEMPALME DE HTT

REFERENCIA CRUZADA

La presente solicitud reivindica el beneficio de la solicitud de patente provisional estadounidense n.º 63/485.142, presentada el 15 de febrero de 2023 y la solicitud de patente provisional estadounidense n.º 63/485.146, presentada el 15 de febrero de 2023, cada una de las cuales se incorpora a la presente en su totalidad mediante esta referencia.

LISTADO DE SECUENCIAS

La presente solicitud contiene un Listado de secuencias que ha sido presentado de forma electrónica en formato XML y se incorpora a la presente en su totalidad mediante esta referencia. Dicha copia XML, creada el 14 de febrero de 2024, se denomina "61313-709_601_SL.xml" y tiene un tamaño de 521.958 bytes.

ANTECEDENTES

La enfermedad de Huntington (HD, por su sigla en inglés) es un trastorno neurodegenerativo progresivo asociado al movimiento. La enfermedad se asocia con la pérdida de neuronas piramidales en la corteza, pérdida de neuronas espinosas medianas en el cuerpo estriado y pérdida de neuronas hipotalámicas. La causa genética de la HD es la herencia autosómica dominante de una repetición expandida del trinucleótido CAG en el exón 1 del gen *HTT*, en donde la presencia de más de 40 repeticiones de CAG en esta región es causante de la enfermedad.

El locus *HTT* es grande, abarca 180 kb y consta de 67 exones, y la expresión del gen *HTT* es necesaria para el desarrollo normal. Aunque la proteína HTT se expresa ampliamente, el cerebro es el que se ve más gravemente afectado por la expansión patológica de repeticiones de trinucleótidos CAG, con efectos tempranos observados en el cuerpo estriado y la corteza motora. En general, se cree que el mecanismo subyacente de la patogénesis de la HD es la expansión de repeticiones CAG somáticas en *la HTT* que ocurre en las regiones cerebrales afectadas (por ejemplo, el cuerpo estriado) de los pacientes con HD.

Los pacientes con HD tienen una neurodegeneración progresiva que conduce a la muerte, generalmente entre 10 y 20 años después del inicio de la enfermedad. Actualmente no existen tratamientos modificadores de la HD. Los tratamientos actuales se limitan a proporcionar alivio sintomático.

COMPENDIO

Se describen en este documento moléculas de trans-empalme de ácido nucleico que incluyen un dominio codificante que comprende uno o más exones *HTT*, un sitio de empalme y un dominio de unión que se une a un intrón objetivo de un pre-ARNm *HTT*. En algunas

modalidades, las construcciones que comprenden moléculas de trans-empalme de ácido nucleico pueden incluir además secuencias adicionales que codifican ARN antisentido que se une al pre-ARNm objetivo y al hacerlo bloquea el cis-empalme, promoviendo así el trans-empalme de la molécula de trans-empalme de ácido nucleico al pre-ARNm objetivo (por ejemplo, pre-ARNm *HTT*). En algunas modalidades, las moléculas de trans-empalme de ácidos nucleicos pueden comprender además un segundo dominio de unión que se une a un intrón objetivo en un segundo pre-ARNm objetivo (por ejemplo, pre-ARNm *MSH3*), con lo que el trans-empalme en el intrón objetivo del segundo pre-ARNm objetivo genera un pre-ARNm híbrido que comprende uno o más exones *HTT* y exones del segundo pre-ARNm objetivo (por ejemplo, pre-ARNm *MSH3*), en donde el pre-ARNm híbrido se procesa a ARNm híbrido que está destinado a la degradación mediante, por ejemplo, desintegración mediada por sin sentido. En algunas modalidades, las construcciones que comprenden moléculas de trans-empalme de ácidos nucleicos pueden incluir además secuencias adicionales que codifican ARN nuclear pequeño (ARNnp) que bloquea el procesamiento normal de un segundo pre-ARNm objetivo (por ejemplo, pre-ARNm *MSH3*) a través de, por ejemplo, la introducción de un codón de parada prematuro en el segundo pre-ARNm objetivo (por ejemplo, pre-ARNm *MSH3*) que se dirige al ARNm del segundo objetivo (por ejemplo, ARNm *MSH3*) para la desintegración mediada por sin sentido. En algunas modalidades, las construcciones que comprenden moléculas de trans-empalme de ácidos nucleicos pueden comprender además elementos que codifican ARN antisentido que bloquea el procesamiento normal de un segundo pre-ARNm objetivo (por ejemplo, pre-ARNm de *MSH3*) a través de, por ejemplo, el bloqueo de las uniones de empalme de *MSH3* o la hibridación con el 5' UTR o la secuencia de codificación inicial de *MSH3* para bloquear la traducción de *MSH3*. En algunas modalidades, las construcciones que comprenden moléculas de trans-empalme de ácidos nucleicos pueden incluir además secuencias adicionales que codifican microARN (miARN) específicos para el ARNm de *HTT* endógeno o específicos para el ARNm de *MSH3* que promueven la escisión del ARNm de *HTT* endógeno o el ARNm de *MSH3*, respectivamente, reduciendo así los niveles de ARNm de *HTT* endógeno o ARNm de *MSH3*. También se incluyen las composiciones que comprenden dichas moléculas de trans-empalme de ácidos nucleicos, así como las composiciones que comprenden combinaciones de moléculas de trans-empalme de ácidos nucleicos con agentes terapéuticos adicionales (por ejemplo, oligonucleótidos antisentido o construcciones codificantes de ARN antisentido). Las moléculas de trans-empalme de ácidos nucleicos y las composiciones que las comprenden se pueden usar solas o en combinación con agentes terapéuticos adicionales en métodos para tratar la enfermedad de Huntington (HD). Las moléculas de trans-empalme de ácidos nucleicos y las composiciones que las comprenden también se usan solas o en combinación con agentes terapéuticos adicionales para tratar la HD o en la preparación de medicamentos para tratar la HD.

Se describe una molécula de trans-empalme de ácido nucleico *HTT* que comprende: (a) un dominio codificante que comprende el exón 1 de *HTT* y el exón 2 de *HTT*; (b) un dominio de empalme; y (c) un dominio de unión que se une a un intrón objetivo de un pre-ARNm de *HTT*,

en donde el intrón objetivo comprende el intrón 2. En algunas modalidades, el dominio de unión comprende, consiste esencialmente en, o consiste en cualquiera de las SEQ ID NO: 60-81, o una secuencia que tiene al menos un 90 % de identidad con cualquiera de las SEQ ID NO: 60-81.

5 También se describe en la presente una molécula de trans-empalme de ácido nucleico HTT que comprende: (a) un dominio de codificación que comprende los exones 1-3 de HTT; (b) un dominio de empalme; y (c) un dominio de unión que se une a un intrón objetivo de un pre-ARNm de HTT, en donde el intrón objetivo comprende el intrón 3. En algunas modalidades, el dominio de unión comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO: 158-174, o una secuencia que tiene al menos 90% de identidad con cualquiera de las SEQ ID NO: 158-174.

15 También se describe en la presente una molécula de trans-empalme de ácido nucleico HTT que comprende: (a) un dominio de codificación que comprende el exón 1 de HTT; (b) un dominio de empalme; y (c) un dominio de unión que se une a un intrón objetivo de un pre-ARNm de HTT, en donde el intrón objetivo comprende el intrón 1, y en donde el dominio de unión comprende cualquiera de las SEQ ID NO: 8-21.

20 En algunas modalidades, en cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente, el dominio codificante comprende, consiste esencialmente en, o consiste en el exón 1 de HTT; el exón 1 de HTT y el exón 2 de HTT; o los exones 1-3 de HTT. En algunas modalidades, el dominio codificante comprende, consiste esencialmente en, o consiste en cualquiera de las SEQ ID NO: 3, 59, 157 o 349-353 o una secuencia con al menos un 90 % de identidad con las SEQ ID NO: 3, 59, 157 o 349-353. En algunas modalidades, el dominio codificante, el dominio de empalme y el dominio de unión están unidos operativamente en dirección 5' a 3'.

30 En algunas modalidades, las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente comprenden además un enlazador, en donde el dominio de codificación, el dominio de empalme, el enlazador y el dominio de unión están unidos operativamente en una dirección de 5' a 3'. En algunas modalidades, el enlazador comprende, consiste esencialmente en, o consiste en: una secuencia que varía de 20 a 50 nucleótidos de longitud, en donde el enlazador comprende 60-80% de guaninas intercaladas con timidina/uridina; 65-75% de guaninas intercaladas con timidina/uridina; o 66-74% de guaninas intercaladas con timidina/uridina; una secuencia que varía de 20 a 45 nucleótidos de longitud, en donde el enlazador comprende 60-80% de guaninas intercaladas con timidina/uridina; 65-75% de guaninas intercaladas con timidina/uridina; o 66-74% de guaninas intercaladas con timidina/uridina; o una secuencia que va de 22 a 42 nucleótidos de longitud, en donde el enlazador comprende 60-80% de guaninas intercaladas con timidina/uridina; 65-75% de guaninas intercaladas con timidina/uridina; o 66-74% de guaninas intercaladas con

- timidina/uridina. En algunas modalidades, el enlazador comprende, consiste esencialmente en, o consiste en cualquiera de los siguientes: SEQ ID NO: 38 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 38; SEQ ID NO: 39 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 39; SEQ ID NO: 40 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 40; o SEQ ID NO: 41 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 41. En algunas modalidades, el enlazador comprende, consiste esencialmente en, o consiste en cualquiera de: SEQ ID NO: 37 o una secuencia que tiene al menos 90% de identidad con SEQ ID NO: 37; SEQ ID NO: 42 o una secuencia que tiene al menos 90% de identidad con SEQ ID NO: 42; SEQ ID NO: 43 o una secuencia que tiene al menos 90% de identidad con SEQ ID NO: 43; SEQ ID NO: 44 o una secuencia que tiene al menos 90% de identidad con SEQ ID NO: 44; SEQ ID NO: 45 o una secuencia que tiene al menos 90% de identidad con SEQ ID NO: 45; SEQ ID NO: 46 o una secuencia que tiene al menos 90% de identidad con SEQ ID NO: 46; SEQ ID NO: 106 o una secuencia que tiene al menos 90% de identidad con SEQ ID NO: 106; SEQ ID NO: 107 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 107; SEQ ID NO: 108 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 108; SEQ ID NO: 109 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 109; SEQ ID NO: 110 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 110; SEQ ID NO: 111 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 111; SEQ ID NO: 112 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 112; SEQ ID NO: 197 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 197; o SEQ ID NO: 198 o una secuencia que tiene al menos un 90% de identidad con SEQ ID NO: 198.
- En algunas modalidades, las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente comprenden además un terminador de triple hélice, en donde el dominio de codificación, el dominio de empalme, el enlazador, cuando están presentes, el dominio de unión, y el terminador de triple hélice están unidos operativamente en una dirección de 5' a 3'. En algunas modalidades, el terminador de triple hélice comprende, consiste esencialmente en, o consiste en SEQ ID NO: 5 o una secuencia que tiene al menos un 90 % de identidad con SEQ ID NO: 5. En algunas modalidades, el terminador de triple hélice comprende, consiste esencialmente en, o consiste en SEQ ID NO: 6.
- En algunas modalidades, las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente comprenden además una región no traducida 5' (5' UTR), en donde el 5' UTR, el terminador de hélice triple, en donde el dominio de codificación, el dominio de empalme, el enlazador, cuando está presente, el dominio de unión, y el terminador de hélice triple, cuando están presentes, están unidos operativamente en una dirección de 5' a 3'. En algunas modalidades, el 5' UTR es un *HTT* 5' UTR. En algunas modalidades, el *HTT* 5' UTR comprende, consiste esencialmente en, o consiste en cualquiera de las SEQ ID NO: 136 o 192 o una secuencia que tiene al menos un 90 % de identidad con cualquiera de las SEQ ID NO:

136 o 192.

En algunas modalidades, las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente comprenden además una secuencia que codifica una etiqueta de epítipo, en donde la 5' UTR, cuando está presente, la etiqueta de epítipo, el dominio codificante, el dominio de empalme, el enlazador, cuando está presente, el dominio de unión, y el terminador de hélice triple, cuando están presentes, están unidos operativamente en una dirección de 5' a 3'. En algunas modalidades, la secuencia que codifica la etiqueta del epítipo comprende, consiste esencialmente en, o consiste en una SEQ ID NO: 4.

También se describe en la presente una molécula de trans-empalme de ácido nucleico HTT que comprende: (a) un dominio de codificación que comprende el exón 1 de HTT y el exón 2 de HTT; (b) un dominio de empalme; (c) un enlazador; y (d) un dominio de unión que se une a un intrón objetivo de un pre-ARNm de HTT, en donde el intrón objetivo comprende el intrón 2. En algunas modalidades, el dominio de unión comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO: 60-81, o una secuencia que tiene al menos 90% de identidad con cualquiera de las SEQ ID NO: 60-81.

También se describe en la presente una molécula de trans-empalme de ácido nucleico HTT que comprende: (a) un dominio de codificación que comprende los exones 1-3 de HTT; (b) un dominio de empalme; (c) un enlazador; y (d) un dominio de unión que se une a un intrón objetivo de un pre-ARNm de HTT, en donde el intrón objetivo comprende el intrón 3. En algunas modalidades, el dominio de unión comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO:158-174, o una secuencia que tiene al menos 90% de identidad con cualquiera de las SEQ ID NO:158-174.

También se describe en la presente una molécula de trans-empalme de ácido nucleico HTT que comprende: (a) un dominio de codificación que comprende el exón 1 de HTT; (b) un dominio de empalme; (c) un enlazador; y (d) un dominio de unión que se une a un intrón objetivo de un pre-ARNm de HTT, en donde el intrón objetivo comprende el intrón 1. En algunas modalidades, el dominio de unión comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO: 8-21, o una secuencia que tiene al menos 90% de identidad con cualquiera de las SEQ ID NO: 8-21.

En algunas modalidades, el enlazador comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO: SEQ ID NO: 37 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO:37; SEQ ID NO: 38 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 38; SEQ ID NO: 39 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 39; SEQ ID NO: 40 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 40; SEQ ID NO: 41 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 41; SEQ ID NO: 42 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 42; SEQ ID NO: 43 o una secuencia que tiene al menos 90 % de

identidad con SEQ ID NO: 43; SEQ ID NO: 44 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 44; SEQ ID NO: 45 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 45; SEQ ID NO: 46 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 46; SEQ ID NO: 106 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 106; SEQ ID NO: 107 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 107; SEQ ID NO: 108 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 108; SEQ ID NO: 109 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 109; SEQ ID NO: 110 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 110; SEQ ID NO: 111 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 111; SEQ ID NO: 112 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 112; SEQ ID NO: 197 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 197; o SEQ ID NO: 198 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 198. En algunas modalidades, la molécula de trans-empalme de ácido nucleico HTT comprende además un terminador de triple hélice, en donde el dominio de codificación, el dominio de empalme, el enlazador, el dominio de unión y el terminador de triple hélice están unidos operativamente en una dirección de 5' a 3'; y opcionalmente, comprende además un 5' UTR, en donde el 5' UTR, cuando está presente, el dominio de codificación, el dominio de empalme, el enlazador, el dominio de unión y el terminador de triple hélice, cuando está presente, están unidos operativamente en una dirección de 5' a 3'.

También se describe en este documento una molécula de trans-empalme de ácido nucleico que comprende un enlazador, en donde el enlazador comprende, consiste esencialmente en, o consiste en una secuencia que varía de 20 a 50 nucleótidos de longitud, en donde el enlazador comprende 60-80 % de guaninas intercaladas con timidina/uridina; 65-75 % de guaninas intercaladas con timidina/uridina; o 66-74 % de guaninas intercaladas con timidina/uridina. En algunas modalidades, el enlazador comprende, consiste esencialmente en, o consiste en una secuencia que varía de 20 a 45 nucleótidos de longitud, en donde el enlazador comprende 60-80 % de guaninas intercaladas con timidina/uridina; 65-75 % de guaninas intercaladas con timidina/uridina; o 66-74 % de guaninas intercaladas con timidina/uridina. En algunas modalidades, el enlazador comprende, consiste esencialmente en, o consiste en una secuencia que varía de 22 a 42 nucleótidos de longitud, en donde el enlazador comprende 60-80 % de guaninas intercaladas con timidina/uridina; 65-75 % de guaninas intercaladas con timidina/uridina; o 66-74 % de guaninas intercaladas con timidina/uridina. En algunas modalidades, el enlazador comprende, consiste esencialmente en, o consiste en: SEQ ID NO: 38 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 38; SEQ ID NO: 39 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 39; SEQ ID NO: 40 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 40; o SEQ ID NO: 41 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 41.

También se describe en la presente una molécula de trans-empalme de ácido nucleico que

comprende un enlazador, en donde el enlazador comprende, consiste esencialmente en, o consiste en: SEQ ID NO: 37 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO:37; SEQ ID NO: 42 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 42; SEQ ID NO: 43 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 43; SEQ ID NO: 44 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 44; SEQ ID NO: 45 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 45; SEQ ID NO: 46 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 46; SEQ ID NO: 106 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 106; SEQ ID NO: 107 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 107; SEQ ID NO: 108 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 108; SEQ ID NO: 109 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 109; SEQ ID NO: 110 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 110; SEQ ID NO: 111 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 111; SEQ ID NO: 112 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 112; SEQ ID NO:197 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 197; o SEQ ID NO: 198 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 198.

En algunas modalidades de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente, la molécula de trans-empalme de ácido nucleico HTT comprende además un dominio de unión que se une a un intrón objetivo de un pre-ARNm *MSH3*. En algunas modalidades, el intrón objetivo de *MSH3* comprende cualquiera del intrón 5 o el intrón 15 de *MSH3*. En algunas modalidades, el dominio de unión que se une a un intrón objetivo de un pre-ARNm *MSH3* comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO: 140, 142, 144, 146, 209, o 210 o una secuencia que tiene al menos 90% de identidad con cualquiera de las SEQ ID NO: 140, 142, 144, 146, 209 o 210. En algunas modalidades, la molécula de trans-empalme de ácido nucleico comprende cualquiera de las SEQ ID NO: 149 – 154 o las SEQ ID NO: 212-223, o una secuencia que tiene al menos un 90 % de identidad con cualquiera de las SEQ ID NO: 149 – 154 o las SEQ ID NO: 212-223.

En algunas modalidades, cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente comprende además una secuencia de ácido nucleico que codifica un pri-miARN que comprende una secuencia de microARN (miARN) específica para el exón 1 del ARNm HTT endógeno, en donde el exón 1 de la molécula de trans-empalme de ácido nucleico comprende un cambio en la secuencia de nucleótidos que altera la unión del miARN al ARNm codificado al menos en parte por la molécula de trans-empalme de ácido nucleico. En algunas modalidades, la secuencia de miARN comprende cualquiera de las SEQ ID NO: 339 o 342 o una secuencia de ácido nucleico que tiene al menos un 90 % de identidad con cualquiera de las SEQ ID NO: 339 o 342. En algunas modalidades, la secuencia de ácido nucleico que codifica el pri-miARN comprende cualquiera de las SEQ ID NO: 341 o 344. En algunas modalidades, el pri-miARN comprende una secuencia de andamiaje mir-33. En

algunas modalidades, el pri-miARN comprende una secuencia de andamiaje mir-30a, una secuencia de bucle mir-30a, una secuencia de andamiaje mir-155, una secuencia de bucle mir-155, una secuencia de andamiaje mir-33 o una secuencia de bucle mir-33. En algunas modalidades, la secuencia de andamiaje mir-30a comprende una secuencia de andamiaje 5' establecida en la SEQ ID NO: 227 o una secuencia de andamiaje 3' establecida en SEQ ID NO: 228; en donde la secuencia de bucle mir-30a comprende SEQ ID NO: 229; en donde la secuencia de andamiaje mir-155 comprende una secuencia de andamiaje 5' establecida en SEQ ID NO: 230 o una secuencia de andamiaje 3' establecida en SEQ ID NO: 231; en donde la secuencia de bucle mir-155 comprende SEQ ID NO: 232; en donde la secuencia de andamiaje mir-33 comprende una secuencia de andamiaje 5' establecida en SEQ ID NO: 259 o una secuencia de andamiaje 3' establecida en SEQ ID NO: 260; o en donde la secuencia de bucle mir-33 comprende SEQ ID NO: 261.

También se describe en este documento una construcción de ácido nucleico de omisión de exón MSH3 que comprende, operativamente unido: (a) una secuencia que codifica un ARN antisentido que promueve la omisión de exón de un exón objetivo del pre-ARNm de MSH3, en donde el exón objetivo es cualquiera de los exones 2-4, 6-8 o 15 de MSH3, en donde el exón objetivo comprende una unión exón-intrón 5' y una secuencia de unión exón-intrón 3'; y (b) una secuencia que codifica una secuencia de ARN nuclear pequeño (ARNnp). En algunas modalidades, la construcción de ácido nucleico de omisión de exón MSH3 comprende además un promotor U1 y un terminador U1 unidos operativamente a (a) y (b). En algunas modalidades, el ARNnp es un ARNnp modificado. En algunas modalidades, el ARNnp modificado comprende una secuencia OPT U7 Sm o una secuencia ARNnp U2. En algunas modalidades, el ARN antisentido se dirige a la unión exón-intrón 5' o a la unión exón-intrón 3' del exón objetivo. En algunas modalidades, el ARN antisentido comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO: 274, 275, 276, 277, 278, 279, 280, 300, 302, 301, 303, 281, 282, 306, 308, 305, 307, 311, 313, 310, 312, 316, 318, 315, 317, 321, 323, 320 o 322 o una secuencia que tiene al menos un 90 % de identidad con cualquiera de las SEQ ID NO: 274, 275, 276, 277, 278, 279, 280, 300, 302, 301, 303, 281, 282, 306, 308, 305, 307, 311, 313, 310, 312, 316, 318, 315, 317, 321, 323, 320 o 322. En algunas modalidades, una construcción de ácido nucleico de omisión de exón MSH3 descrita anteriormente comprende cualquiera de las SEQ ID NO: 284, 285, 286, 287, 288, 289, 290, 325, 326, 291, 292, 328, 329, 331, 332, 334, 335, 337 y 338.

En algunas modalidades, el ARN antisentido se dirige a la unión exón-intrón 5' y a la unión exón-intrón 3'. En algunas modalidades, el ARN antisentido comprende una secuencia que es al menos 80% complementaria a la secuencia completa del exón objetivo.

En algunas modalidades, el ARN antisentido comprende, además: (a) una secuencia que es al menos 80% complementaria a una secuencia de 5 nucleótidos anterior a la unión exón-intrón 5'; y (b) una secuencia que es al menos 80% complementaria a una secuencia de 5

nucleótidos posterior la unión exón-intrón 3'. En algunas modalidades, el ARN antisentido comprende cualquiera de las SEQ ID NO: 299, 304, 309, 314, o 319 o una secuencia que tiene al menos 90% de identidad con cualquiera de las SEQ ID NO: 299, 304, 309, 314 o 319. En algunas modalidades, una construcción de ácido nucleico de omisión de exón MSH3 descrita anteriormente comprende cualquiera de las SEQ ID NO: 324, 327, 330, 333 o 336.

En algunas modalidades, el ARN antisentido comprende, unido operativamente en una dirección 5' a 3': (a) una secuencia que se dirige a la unión exón-intrón 3'; (b) una secuencia de enlace de al menos 15 nucleótidos que no se une al exón objetivo; y (c) una secuencia que se dirige a la unión exón-intrón 5'. En algunas modalidades, la secuencia de enlace es menos del 50 % complementaria a todas las secuencias del exón objetivo de la misma longitud que el enlazador. En algunas modalidades, el ARN antisentido comprende cualquiera de las SEQ ID NO: 300, 301, 302, 303, 305, 306, 307, 308, 310, 311, 312, 313, 315, 316, 317, 318, 320, 321, 322 o 323, o cualquier combinación de estos. En algunas modalidades, la construcción de ácido nucleico de omisión de exón MSH3 comprende cualquiera de las SEQ ID NO: 325, 326, 328, 329, 331, 332, 334, 335, 337, o 338.

En algunas modalidades, el ARN antisentido se dirige al exón 7 de MSH3. En algunas modalidades, la construcción de ácido nucleico de omisión de exón MSH3 comprende la SEQ ID NO: 309. En algunas modalidades, la construcción de ácido nucleico de omisión de exón MSH3 comprende de 5' a 3': (a) SEQ ID NO: 310 (In7/Ex7 ARNas), SEQ ID NO: 298 (enlazador) y SEQ ID NO: 311 (In7/Ex7 ARNas); o (b) SEQ ID NO: 312 (In7/Ex7 ARNas), SEQ ID NO: 298 (enlazador) y SEQ ID NO: 313 (In7/Ex7 ARNas). En algunas modalidades, el ARN antisentido comprende cualquiera de las SEQ ID NO: 309, 310, 311, 312, o 313 o cualquier combinación de estos o una secuencia que tiene al menos 90% de identidad con cualquiera de las SEQ ID NO: 309, 310, 311, 312 o 313. En algunas modalidades, la construcción de ácido nucleico de omisión de exón MSH3 comprende al menos una de las SEQ ID NO: 330-332, o cualquier combinación de estas.

También se describe en este documento una construcción de ácido nucleico de miARN MSH3 que comprende una secuencia que codifica un pri-miARN que comprende una secuencia de andamiaje, una secuencia de bucle y una secuencia de miARN que se dirige al ARNm MSH3 endógeno, en donde: (a) la secuencia de andamiaje se deriva de mir-30a, mir-33 o mir-155; (b) la secuencia de bucle se deriva de mir-22, mir-30a, mir-33 o mir-155; y (c) la secuencia de miARN comprende cualquiera de las SEQ ID NO: 224, 244, 246, 248, 250, 252, 254, 256 o 257 o una secuencia que tiene al menos un 90 % de identidad con cualquiera de las SEQ ID NO: 224, 244, 246, 248, 250, 252, 254, 256 o 257. En algunas modalidades, la secuencia de andamiaje comprende cualquiera de las SEQ ID NO: 227, 228, 230, 231, 259 o 260. En algunas modalidades, la secuencia de bucle comprende cualquiera de las SEQ ID NO: 229, 232 o 261. En algunas modalidades, la secuencia pri-miARN comprende cualquiera de las SEQ ID NO: 234, 235, 238-241 o 262-269. En algunas modalidades, la secuencia que codifica el pri-

miARN está unida operativamente a un promotor U6 o un promotor CMV.

También se describe en este documento una molécula de trans-empalme de ácido nucleico MSH3 que comprende: (a) una secuencia de dominio codificante; (b) un dominio de empalme; y (c) un dominio de unión que se une a un intrón objetivo de un pre-ARNm de MSH3; en donde la secuencia de dominio codificante no es una secuencia de dominio codificante de MSH3. En algunas modalidades, la secuencia del dominio codificante comprende una secuencia que da como resultado un cambio de marco en un ARNm MSH3 maduro cuando se transempalma en el pre-ARNm MSH3. En algunas modalidades, la secuencia del dominio codificante comprende uno o más de los exones 1, 2 y 3 de HTT. En algunas modalidades, el intrón objetivo del pre-ARNm MSH3 es el intrón 5 o el intrón 15. En algunas modalidades, el dominio de unión comprende cualquiera de las SEQ ID NO: 140, 142, 144, 146, 209, o 210 o una secuencia que tiene al menos 90% de identidad con cualquiera de las SEQ ID NO: 140, 142, 144, 146, 209 o 210.

También se describe en este documento una construcción de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 que comprende: (a) cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente; y (b) cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente. En algunas modalidades, (a) y (b) están comprendidos en un solo vector. En algunas modalidades, el vector único es un vector AAV. En algunas modalidades, la construcción de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 comprende cualquiera de las SEQ ID NO: 356, 357, 363, o 364. En algunas modalidades, el vector AAV es un vector scAAV o ssAAV. En algunas modalidades, la construcción de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 comprende cualquiera de las SEQ ID NO: 369, 370, y 371.

También se divulga en la presente una construcción de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 que comprende: (a) cualquiera de las moléculas de ácido nucleico de trans-empalme HTT descritas anteriormente; y (b) cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente. En algunas modalidades, (a) y (b) están comprendidos en un solo vector. En algunas modalidades, la construcción de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 comprende cualquiera de las SEQ ID NO: 358 o 359. En algunas modalidades, el vector único es un vector AAV.

También se divulga en la presente una construcción de ácido nucleico de trans-empalme HTT y miARN MSH3 que comprende: (a) cualquiera de las moléculas de ácido nucleico de trans-empalme HTT descritas anteriormente; y (b) cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente. En algunas modalidades, (a) y (b) están comprendidos en un solo vector. En algunas modalidades, la construcción de ácido nucleico de trans-empalme HTT y miARN MSH3 comprende cualquiera de las SEQ ID NO: 354 o 355.

En algunas modalidades, el vector es un vector de AAV.

También se describe un vector AAV que comprende cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente. En algunas modalidades, el vector AAV comprende cualquiera de las SEQ ID NO: 356, 357, 363 o 364.

También se describe un vector AAV que comprende cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente o cualquiera de las moléculas de trans-empalme de ácido nucleico descritas anteriormente.

También se describe una molécula de trans-empalme de ácido ribonucleico que comprende cualquiera de las SEQ ID NO: 23-36, 47-56, 83-105, 113-125, 175-191 o 199-206.

También se describe una molécula de trans-empalme de ácido ribonucleico transcrita de cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente o de cualquiera de las moléculas de trans-empalme de ácido nucleico descritas anteriormente.

En algunas modalidades de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente, el pre-ARNm de HTT comprende al menos una mutación asociada con la enfermedad de Huntington (HD). En algunas modalidades, al menos una mutación asociada con HD comprende una expansión de repeticiones CAG en un alelo del gen *HTT*. En algunas modalidades, la expansión de repeticiones CAG en un alelo del gen *HTT* comprende más de 35 repeticiones CAG. En algunas modalidades, la al menos una mutación asociada con HD es autosómica dominante. En algunas modalidades, la al menos una mutación asociada con HD se expresa en al menos una de las neuronas piramidales corticales, las neuronas espinosas medianas estriatales o las neuronas hipotalámicas.

También se describe un vector que comprende cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico descritas anteriormente; cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 descritas anteriormente; o cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y miARN MSH3 descritas anteriormente.

También se describe un vector que comprende cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente. En algunas modalidades, el vector comprende un dominio regulador 5' unido de forma operativa 5' al dominio codificante. En algunas

modalidades, el dominio regulador 5' está unido de forma operativa a una región no traducida 5'. En algunas modalidades, el dominio regulador 5' comprende un promotor constitutivo o un promotor específico del tejido. En algunas modalidades, el promotor constitutivo es un promotor CMV o un promotor CAGGS.

5

También se describe un plásmido proviral que comprende cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico descritas anteriormente; cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 descritas anteriormente; o cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y miARN MSH3 descritas anteriormente.

También se describe un virus adenoasociado (AAV) que comprende cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico descritas anteriormente; cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 descritas anteriormente; o cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y miARN MSH3 descritas anteriormente.

También se describe un virus adenoasociado (AAV) que comprende cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente, en donde el AAV comprende opcionalmente un dominio regulador 5' unido operativamente 5' a la molécula de trans-empalme de ácido nucleico. En algunas modalidades, el AAV comprende un dominio regulador 5' unido de forma operativa 5' al dominio codificante. En algunas modalidades, el dominio regulador 5' está unido de forma operativa a una región no traducida 5'. En algunas modalidades, el dominio regulador 5' comprende un promotor constitutivo. En algunas modalidades, el promotor constitutivo es un promotor CMV o un promotor CAGGS. En algunas modalidades, el AAV exhibe tropismo neuronal. En algunas modalidades, el AAV es AAV9, AAV8, AAV5, AAV2, AAV7 o AAV2.7m8, AAV-retro, AAV1, AAV4 o AAV-PHP.eB.

40

También se describe una composición que comprende cualquiera de las moléculas de trans-

empalme de ácido nucleico HTT descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico descritas anteriormente; cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 y trans-empalme HTT descritas anteriormente; cualquiera de los vectores descritos anteriormente; cualquiera de los plásmidos provirales descritos anteriormente, o cualquiera de los AAV descritos anteriormente. En algunas modalidades, la composición comprende además un excipiente farmacéuticamente aceptable. En algunas modalidades, la composición comprende además al menos un oligonucleótido antisentido o una construcción que codifica al menos un ARN antisentido que inhibe el cis-empalme del pre-ARNm de HTT. En algunas modalidades, el al menos un oligonucleótido antisentido comprende cualquiera de las SEQ ID NO: 126-135 o la construcción que codifica el al menos un ARN antisentido se une a una secuencia objetivo unida por cualquiera de las SEQ ID NO: 126-135. En algunas modalidades, el al menos un oligonucleótido antisentido comprende la SEQ ID NO: 131 o la construcción que codifica el al menos un ARN antisentido se une a una secuencia objetivo unida por la SEQ ID NO: 131.

También se describe un método de expresar HTT biológicamente activo en una célula objetivo para restaurar los niveles funcionales de la proteína HTT en la célula objetivo, el método comprende transducir la célula objetivo con cualquiera de las moléculas de trans-empalme de ácido nucleico de HTT descritas anteriormente; cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 y trans-empalme HTT descritas anteriormente; cualquiera de los vectores descritos anteriormente; cualquiera de los plásmidos provirales descritos anteriormente; cualquiera de los AAV descritos anteriormente; o cualquiera de las composiciones descritas anteriormente. En algunas modalidades, se reemplaza al menos el 5%, al menos el 10%, al menos el 15%, al menos el 20%, al menos el 25%, al menos el 30%, al menos el 35%, al menos el 40% o al menos el 45% del pre-ARNm de HTT que comprende al menos una mutación asociada con HD en la célula objetivo. En algunas modalidades, se reemplaza al menos 50%, al menos el 55%, al menos el 60%, al menos el 65%, al menos el 70%, al menos el 75%, al menos el 80 %, al menos el 85% del pre-ARNm de HTT que comprende al menos una mutación asociada con HD en la célula objetivo. En

algunas modalidades, se reemplaza al menos el 90%, al menos el 95%, al menos el 96%, al menos el 97%, al menos el 98%, al menos el 99% o el 100% del pre-ARNm de HTT que comprende al menos una mutación asociada con HD en la célula objetivo. En algunas modalidades, los niveles funcionales de HTT se restauran en la célula objetivo mediante la expresión de la proteína HTT biológicamente funcional y/o el ARN *HTT* mutante y se reducen las transcripciones relacionadas (por ejemplo, *HTT1a*).

También se describe un método para reducir la expresión de HTT que comprende una repetición poliglutamina que excede 35 residuos de glutamina consecutivos en un sujeto, el método que comprende transfectar o transducir una célula objetivo, más particularmente una neurona, en el sujeto con cualquiera de las moléculas de trans-empalme de ácido nucleico de HTT descritas anteriormente; cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 y trans-empalme HTT descritas anteriormente; cualquiera de los vectores descritos anteriormente; cualquiera de los plásmidos provirales descritos anteriormente; cualquiera de los AAV descritos anteriormente; o cualquiera de las composiciones descritas anteriormente.

También se describe un método para corregir al menos una mutación en una secuencia de exón HTT en un pre-ARNm HTT en una célula objetivo de un sujeto, el método que comprende administrarle al sujeto cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico descritas anteriormente; cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 y trans-empalme HTT descritas anteriormente; cualquiera de los vectores descritos anteriormente; cualquiera de los plásmidos provirales descritos anteriormente; cualquiera de los AAV descritos anteriormente; o cualquiera de las composiciones descritas anteriormente.

También se describe un método para tratar la enfermedad de Huntington (HD) en un sujeto que lo necesita, el método que comprende administrar al sujeto cualquiera de las moléculas de trans-empalme de ácido nucleico de HTT descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico descritas anteriormente; cualquiera de las

construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 y trans-empalme HTT descritas anteriormente; cualquiera de los vectores descritos anteriormente; cualquiera de los plásmidos provirales descritos anteriormente; cualquiera de los AAV descritos anteriormente; o cualquiera de las composiciones descritas anteriormente en una cantidad terapéuticamente eficaz.

En algunas modalidades de los métodos descritos anteriormente, el método que comprende la administración de cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico descritas anteriormente; cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 y trans-empalme HTT descritas anteriormente; cualquiera de los vectores descritos anteriormente; cualquiera de los plásmidos provirales descritos anteriormente; cualquiera de los AAV descritos anteriormente; o cualquiera de las composiciones descritas anteriormente al cerebro del sujeto. En algunas modalidades, el sujeto es un mamífero, preferentemente un roedor, un primate no humano o un ser humano. En algunas modalidades, el sujeto está genéticamente predispuesto a tener HD o ha sido diagnosticado con HD.

También se describe cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico descritas anteriormente; cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 y trans-empalme HTT descritas anteriormente; cualquiera de los vectores descritos anteriormente; cualquiera de los plásmidos provirales descritos anteriormente; cualquiera de los AAV descritos anteriormente; o cualquiera de las composiciones descritas anteriormente

para usar en la prevención o para tratar la HD en un sujeto que lo necesita.

También se describe en la presente cualquiera de las moléculas de trans-empalme de ácido nucleico HTT descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico descritas anteriormente; cualquiera de las construcciones de ácido nucleico de omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 descritas anteriormente; cualquiera de las moléculas de trans-empalme de ácido nucleico MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de trans-empalme HTT, miARN HTT y omisión de exón MSH3 descritas anteriormente; cualquiera de las construcciones de ácido nucleico de miARN MSH3 y trans-empalme HTT descritas anteriormente; cualquiera de los vectores descritos anteriormente; cualquiera de los plásmidos provirales descritos anteriormente; cualquiera de los AAV descritos anteriormente; o cualquiera de las composiciones descritas anteriormente para usar en la preparación de un medicamento para el tratamiento o la prevención de la HD en un sujeto que lo necesita.

También se describe un método que comprende introducir en una célula una molécula de trans-empalme de ácido nucleico configurada para empalmarse tanto a un primer pre-ARNm objetivo como a un segundo pre-ARNm objetivo, en donde el empalme al primer pre-ARNm objetivo corrige un defecto en el primer pre-ARNm objetivo, y en donde el empalme al segundo pre-ARNm objetivo introduce un defecto en el segundo pre-ARNm objetivo. En algunas modalidades, la molécula de trans-empalme de ácido nucleico comprende un primer dominio de unión configurado para dirigirse a un intrón del primer pre-ARNm objetivo y un segundo dominio de unión configurado para dirigirse a un intrón del segundo pre-ARNm objetivo. En algunas modalidades, la molécula de trans-empalme de ácido nucleico comprende además una secuencia de dominio codificante que comprende una secuencia funcional de uno o más exones del primer pre-ARNm objetivo que corrige el defecto en el primer pre-ARNm objetivo. En algunas modalidades, el defecto en el segundo pre-ARNm objetivo comprende un cambio de marco en la secuencia de codificación del segundo pre-ARNm objetivo. En algunas modalidades, el cambio de marco crea un codón de terminación prematura en el segundo pre-ARNm objetivo. En algunas modalidades, el defecto comprende el codón de inicio endógeno del segundo pre-ARNm objetivo que se está eliminando. En algunas modalidades, el defecto comprende un UTR 5' insertado que impide la traducción de la proteína codificada por el segundo pre-ARNm objetivo. En algunas modalidades, el defecto comprende un UTR 3' insertado que desestabiliza el pre-ARNm o impide la exportación del segundo pre-ARNm objetivo desde el núcleo. En algunas modalidades, el defecto comprende la eliminación de una tapa 5' o una cola de poliA 3' del segundo pre-ARNm objetivo. En algunas modalidades, el defecto provoca la desintegración mediada por sin sentido del segundo pre-ARNm objetivo. En algunas modalidades, la introducción hace que la abundancia del producto génico del segundo pre-ARNm objetivo en la célula se reduzca en comparación con la abundancia del

producto génico antes de la introducción.

También se describe un método para reducir la abundancia de una proteína en una célula, comprendiendo el método la introducción en la célula de una molécula de trans-empalme de ácido nucleico que introduce un defecto en un pre-ARNm que codifica la proteína. En algunas modalidades, el defecto comprende uno o más de los siguientes: (a) un desplazamiento del marco introducido en la secuencia codificante del pre-ARNm; (b) eliminación del codón de inicio endógeno del pre-ARNm; (c) introducción de un codón de parada prematuro en la secuencia codificante del pre-ARNm; (d) reemplazo de la secuencia codificante endógena del pre-ARNm con una secuencia codificante alternativa; (e) inserción de un UTR 5' que impide la traducción de la secuencia codificante endógena del pre-ARNm; (f) inserción de un UTR 3' que desestabiliza el pre-ARNm; (g) inserción de un UTR 3' que impide que el pre-ARNm se exporte desde el núcleo; (h) eliminación de una tapa 5' del pre-ARNm; o (i) eliminación de una cola de poliA 3' del pre-ARNm. En algunas modalidades, la proteína es MSH3. En algunas modalidades, la molécula de trans-empalme de ácido nucleico comprende un dominio de unión que se une a un intrón del pre-ARNm. En algunas modalidades, la molécula de trans-empalme del ácido nucleico comprende una secuencia heteróloga de dominio de codificación.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La Figura 1 representa el rango de repeticiones de trinucleótidos CAG encontradas en el exón 1 del gen *HTT*. Como se muestra allí, las repeticiones de trinucleótidos CAG 8-35 reflejan el rango fenotípicamente normal (tipo salvaje) encontrado en humanos que no presentan signos de enfermedad asociada con el gen *HTT*. Las repeticiones de trinucleótidos CAG 35-39 están asociadas con una penetrancia incompleta de HD. La presencia de más de 40 repeticiones del trinucleótido CAG en esta región del exón 1 del gen *HTT* es causa de enfermedad de HD.

La Figura 2 representa la expansión somática de la repetición CAG del trinucleótido CAG expandido en el exón 1 del gen *HTT*, que es el mecanismo subyacente de la patogénesis de la HD.

La Figura 3 presenta la edición de exones mediante trans-empalme de pre-ARNm de *HTT* como un enfoque terapéutico para la HD. Los editores de exones dirigidos al intrón 1, al intrón 2 o al intrón 3 fueron diseñados y probados para verificar la eficiencia del reemplazo del exón 1 del *HTT* mutante.

La Figura 4 muestra una estructura general ejemplar de construcciones que codifican editores de exones de ARN dirigidos al intrón 1 de *HTT*. El editor de exones ejemplar producido por la construcción representada comprende un UTR 5', una secuencia codificante del exón 1, un sitio donante de empalme, un enlazador, un dominio de unión y una secuencia terminadora. Se varió el dominio de unión para apuntar a diferentes posiciones a lo largo del intrón 1 de

HTT. En algunas modalidades, el promotor es un promotor CMV; en algunas modalidades, el 5' UTR es el *HTT* 5' UTR; en algunas modalidades, el enlazador es el enlazador de 40 mer; en algunas modalidades, el *HTT* 5' UTR se combina con el enlazador de 40 mer; en algunas modalidades, se incluye una etiqueta de epítipo, un ejemplo de la cual es una etiqueta 3X FLAG de extremo N para la detección de proteínas en el objetivo. Las modalidades mencionadas anteriormente se pueden combinar, en donde al menos una de estas modalidades se incluye en un Editor de Exones de ARN y cualquiera y todas sus combinaciones, incluida una combinación de todas estas modalidades en un Editor de Exones de ARN.

La Figura 5 presenta un gráfico que representa los editores de exones de ARN dirigidos al intrón 1 de *HTT*, que exhiben diferentes niveles de eficiencia de trans-empalme (% de reemplazo) dependiendo de dónde se dirige el dominio de unión en el intrón. Las células HEK293 se transfectaron con editores de exones de ARN dirigidos al intrón 1 de *HTT* que se dirigen a varias regiones del intrón 1. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR. La nomenclatura para el dominio de unión es: (posición de la base del nucleótido en el intrón de la posición de inicio del extremo 5' del dominio de unión, numeración de acuerdo con la SEQ ID NO: 1)_(longitud (nt) del dominio de unión). Por ejemplo, 701_150 indica que el dominio de unión es la secuencia de complemento inverso de las bases 701-850 en el intrón 1, numeración de acuerdo con SEQ ID NO: 1. Los resultados mostrados se relacionan con editores de exones de ARN dirigidos al intrón 1 de *HTT* a modo de ejemplo que comprenden los objetivos del dominio de unión indicados, en donde el 5' UTR comprende el *HTT* 5' UTR (SEQ ID NO: 136) y el enlazador comprende el enlazador de 40 mer (SEQ ID NO: 37). NBD_150 es un editor de control en el que el dominio de enlace que apunta a *HTT* se reemplaza con un dominio de enlace que no apunta a *HTT*.

La Figura 6 representa gráficamente las posiciones y las eficiencias de trans-empalme (% de reemplazo) de los editores de exones de ARN dirigidos al intrón 1 de *HTT*. Se determinaron diferentes niveles de eficiencia de trans-empalme (% de reemplazo) que reflejaban dónde se dirigían los dominios de unión en el intrón. Las células HEK293 se transfectaron con editores de exones de ARN dirigidos al intrón 1 de *HTT* que se dirigen a varias regiones del intrón 1. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR. La nomenclatura para el dominio de unión es: (posición de la base del nucleótido en el intrón de la posición de inicio del extremo 5' del dominio de unión, numeración de acuerdo con la SEQ ID NO: 1)_(longitud (nt) del dominio de unión). Por ejemplo, 701_150 indica que el dominio de unión es la secuencia de complemento inverso de las bases 701-850 en el intrón 1, numeración de acuerdo con SEQ ID NO: 1. Los resultados mostrados se relacionan con editores de exones de ARN dirigidos al intrón 1 de *HTT* de ejemplo que comprenden los objetivos del dominio de unión indicados, en donde el 5' UTR comprende el *HTT* 5' UTR y el enlazador comprende el enlazador de 40 mer.

La Figura 7 representa una estructura general de construcciones ejemplares que codifican editores de exones de ARN dirigidos al intrón 1 de *HTT*. Se muestra un editor de exones ejemplar en donde la expresión está impulsada por un promotor CMV. El editor de exones ejemplar representado comprende el UTR 5' de *HTT*, la etiqueta FLAG 3X de extremo N, la secuencia codificante del exón 1, un sitio donante de empalme, un enlazador, un dominio de unión (*HTT_intron1_11704_100*) y una secuencia terminadora.

La Figura 8 muestra la actividad de editores de exones de ARN dirigidos al intrón 1 de *HTT* de ejemplo que incluían varios enlazadores. Como se muestra allí, algunos enlazadores aumentaron la eficiencia del trans-empalme en relación con el enlazador de 40 mer en los editores de exones del intrón 1 de *HTT* (*HTT_intron1_11704_100*). Las células HEK293 fueron transfectadas con editores de exones de ARN dirigidos al intrón 1 de *HTT* que comprendían los enlazadores indicados. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR.

La Figura 9 describe una estructura general de construcciones ejemplares que codifican editores de exones de ARN que se dirigen al intrón 2 de *HTT*. El editor de exones ejemplar representado comprende un UTR 5', las secuencias codificantes de los exones 1-2, un sitio donante de empalme, un enlazador, un dominio de unión y una secuencia terminadora. Se varió el dominio de unión para apuntar a diferentes posiciones a lo largo del intrón 2 de *HTT*. En algunas modalidades, el 5' UTR es el *HTT* 5' UTR; en algunas modalidades, el enlazador es el enlazador de 40 mer; en algunas modalidades, el *HTT* 5' UTR se combina con el enlazador de 40 mer. En algunas modalidades, el promotor es un promotor CMV; en algunas modalidades, el 5' UTR es el *HTT* 5' UTR; en algunas modalidades, el enlazador es el enlazador de 40 mer; en algunas modalidades, el promotor CMV se combina con el 5' UTR de *HTT*; en algunas modalidades, el promotor CMV se combina con el 5' UTR de *HTT* y el enlazador de 40 mer; en algunas modalidades, se incluye una etiqueta de epítipo, un ejemplo de la cual es una etiqueta 3X FLAG de extremo N para la detección de proteínas en el objetivo. Las modalidades mencionadas anteriormente se pueden combinar, en donde al menos una de estas modalidades se incluye en un Editor de Exones de ARN y cualquiera y todas sus combinaciones, incluida una combinación de todas estas modalidades en un Editor de Exones de ARN.

La Figura 10 muestra la actividad de varios editores de exones de ARN dirigidos al intrón 2 de *HTT* ilustrativos, que exhiben diferentes niveles de eficiencia de trans-empalme (% de reemplazo) dependiendo de dónde se dirige el dominio de unión en el intrón. Las células HEK293 fueron transfectadas con editores de exones de ARN dirigidos al intrón 2 de *HTT* ilustrativos que se dirigen a varias regiones del intrón 2. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR. La nomenclatura para el dominio de unión es: (posición de la base del

nucleótido en el intrón de la posición de inicio del extremo 5' del dominio de unión, numeración de acuerdo con la SEQ ID NO: 57)_(longitud (nt) del dominio de unión). Los resultados mostrados se relacionan con editores de exones de ARN dirigidos al intrón 2 de *HTT* de ejemplo que comprenden los objetivos del dominio de unión indicados, en donde el 5' UTR comprende el *HTT* 5' UTR y el enlazador comprende el enlazador de 40 mer. NBD, editor de control en el que el dominio de enlace que apunta a *HTT* se reemplaza con un dominio de enlace que no apunta a *HTT*. Mutante de empalme, editor de control que carece de un sitio donante de empalme funcional.

La Figura 11 representa gráficamente las posiciones y las eficiencias de trans-empalme (% de reemplazo) de los editores de exones de ARN dirigidos al intrón 2 de *HTT*. Se determinaron diferentes niveles de eficiencia de trans-empalme (% de reemplazo) que reflejaban dónde se dirigían los dominios de unión en el intrón. Las células HEK293 fueron transfectadas con editores de exones de ARN dirigidos al intrón 2 de *HTT* que se dirigen a varias regiones del intrón 2. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR. La nomenclatura para el dominio de unión es: (posición de la base del nucleótido en el intrón de la posición de inicio del extremo 5' del dominio de unión, numeración de acuerdo con la SEQ ID NO: 57)_(longitud (nt) del dominio de unión). Los resultados mostrados se relacionan con editores de exones de ARN dirigidos al intrón 2 de *HTT* de ejemplo que comprenden los objetivos del dominio de unión indicados, en donde el 5' UTR comprende el *HTT* 5' UTR y el enlazador comprende el enlazador de 40 mer.

Las FIGS. 12A y 12B representan editores de exones de ARN dirigidos al intrón 2 de *HTT* ilustrativos que se dirigen a la región anterior del punto de ramificación. A) Los editores de exones de ARN fueron diseñados para dirigirse a la parte anterior del punto de ramificación del intrón 2 y varían en la longitud del dominio de unión. La expresión del editor de exones está impulsada por un promotor CMV. Los editores de exones comprenden el UTR 5' de *HTT*, la etiqueta FLAG 3X de extremo N para la detección de proteínas en el objetivo, la secuencia codificante de los exones 1-2, un sitio donante de empalme, el enlazador 41mer_2, el dominio de unión indicado, y una secuencia terminadora. La nomenclatura para el dominio de unión es: (posición de la base del nucleótido en el intrón de la posición de inicio del extremo 5' del dominio de unión, numeración de acuerdo con la SEQ ID NO: 57)_(longitud (nt) del dominio de unión). Se varió el dominio de unión para apuntar a diferentes longitudes en la parte anterior del punto de ramificación en el intrón 2. B) La longitud del dominio de unión influye en la eficiencia de la orientación, según los cálculos del % de reemplazo. Las células HEK293 fueron transfectadas con editores de exones de ARN dirigidos al intrón 2 de *HTT* que variaban en sus longitudes de dominio de unión. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR.

La Figura 13 representa una estructura general de construcciones ejemplares que codifican editores de exones de ARN dirigidos al intrón 2 de HTT. La expresión del editor de exones está impulsada por un promotor CMV. Los editores de exones ejemplares comprenden el UTR 5' de HTT, la etiqueta FLAG 3X de extremo N, las secuencias codificantes de los exones 1-2, un sitio donante de empalme, un enlazador, un dominio de unión (HTT_intron2_12061_150) y una secuencia terminadora.

La Figura 14 muestra que los editores de exones que comprenden los enlazadores indicados no exhiben eficiencias de trans-empalme significativamente diferentes en relación con el enlazador de 40 mer en el intrón 2 de HTT (HTT_intron2_12061_150). Las células HEK293 fueron transfectadas con editores de exones de ARN dirigidos al intrón 2 de HTT que variaban con respecto a sus enlazadores incluidos en estos. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR.

La Figura 15 presenta una caricatura que representa una reacción de trans-empalme y su competencia con respecto al cis-empalme. Abreviaturas utilizadas en este: terminador MALAT1 (term); dominio de unión (BD); enlazador (L); sitio de empalme (SS). La molécula empalmada en cis resultante del cis-empalme comprende la expansión de repeticiones CAG. La molécula trans-empalmada quimérica resultante del trans-empalme 5' mediado por el RTM 5' de HTT comprende un exón 1 de HTT que comprende un número normal de repeticiones de CAG (8-35). Abreviaturas utilizadas en este: MALAT1 terminador (term); dominio de unión (BD); enlazador (L); sitio de empalme (SS); codón optimizado (C/O).

La Figura 16 presenta una caricatura que muestra oligonucleótidos antisentido (ASO) diseñados para bloquear sitios de cis-empalme competitivos (ASO8-10), así como sitios de cis-empalme para el exón en la parte anterior (ASO2-7). Cada uno de estos ASO se cotransfectó con el editor de exones dirigido al intrón 2 de HTT indicado (HTT_intron2_12061_150) y se analizó la eficiencia de trans-empalme *in vitro* en células HEK293. Abreviaturas utilizadas en este: terminador MALAT1 (term); dominio de unión (BD); enlazador (L); sitio de empalme (SS).

La Figura 17 muestra la actividad de reemplazo porcentual de un editor de exones dirigido al intrón 2 de HTT ejemplar (HTT_intron2_12061_150) en combinación con los ASO indicados. ASO6, que fue diseñado para bloquear el cis-empalme del intrón en la parte anterior, condujo a una mejora en las eficiencias del trans-empalme *in vitro*. Las células HEK293 se cotransfectaron con una construcción de editor de exones de ARN (REEC) dirigido al intrón 2 de HTT ejemplar y ASO diseñados para bloquear el sitio de cis-empalme competitivo o ASO diseñados para bloquear el empalme del intrón en la parte anterior. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR.

La Figura 18 muestra imágenes representativas de RT-qPCR y análisis de transferencia Western (sondeadas para el epítipo FLAG de extremo N) de lisados de células completas de células HEK293 transfectadas con editores de exones de ARN *HTT* que comprenden los elementos indicados. Cabe destacar que los niveles de detección de proteínas en el objetivo se correlacionan con la eficiencia del trans-empalme, representada por el % de reemplazo de ARN *HTT*. Las células HEK293 fueron transfectadas con editores de exones *HTT* etiquetados con FLAG en el extremo N que tienen un rango de actividad basado en ensayos de RT-qPCR (panel superior). Se utilizó el anticuerpo anti-FLAG para detectar la proteína generada luego de un trans-empalme exitoso en lisados de células completas (panel inferior). Se utilizó el anticuerpo anti- proteína de Huntington para detectar la proteína nativa y ONT generada luego de un trans-empalme exitoso en lisados de células completas (panel inferior). La intensidad de la banda FLAG ONT escala con el rendimiento relativo del editor de exones basado en qPCR.

La Figura 19 presenta un esquema que representa un mecanismo potencial de un enfoque terapéutico híbrido diseñado para tratar la HD. El enfoque terapéutico híbrido combina agentes que inhiben la expansión somática de CAG (por ejemplo, mediante la reducción de MSH3) con editores de exones de ARN dirigidos al pre-ARNm de *HTT*. Los editores de exones de ARN dirigidos al pre-ARNm *HTT* sirven para reemplazar cualquier ARN *HTT* mutante que pudiera producirse a partir del ADN que “escapa” a la inhibición del proceso de expansión somática.

La Figura 20 presenta un esquema que representa un mecanismo de acción potencial de un editor de exones de ARN con dominio de unión en tándem dirigido al pre-ARNm *HTT* y *MSH3*. En una modalidad ejemplar, la expresión de un editor de exones es dirigido por un promotor de CMV. Tales editores de exón ejemplares pueden comprender la UTR 5' de *HTT*, la etiqueta de extremo N 3X FLAG, la secuencia codificante del exón 1 de *HTT*, el sitio donante de empalme, un enlazador, un dominio de unión a MSH3 (dirigido, por ejemplo, al intrón 5 o al intrón 15 del pre-ARNm *MSH3*), un dominio de unión a *HTT* (por ejemplo, *HTT_intron1_11704_100*) y una secuencia terminadora. El dominio de unión de *HTT* apuntará al editor de exones para producir el ARN *HTT* corregido después de un trans-empalme exitoso, mientras que el dominio de unión de *MSH3* apuntará al editor de exones para producir una molécula de ARN quimérica *HTT* exón 1-*MSH3* con un codón de parada prematuro que estará sujeto a una desintegración mediada por sin sentido (NMD) y conducirá a la reducción posterior de la expresión de *MSH3*.

Las Figuras 21A-21C presentan resultados que muestran que los editores de exones de ARN del dominio de unión en tándem dirigidos a *HTT* y *MSH3* exhiben un trans-empalme exitoso a ambos pre-ARNm. RT-qPCR de A) eficiencia de trans-empalme en el objetivo (ONT) de *HTT* (a través del dominio de unión de *HTT*), B) eficiencia de trans-empalme quimérico *HTT-MSH3* (a través del dominio de unión de *MSH3*) y C) nivel de expresión de transcripción de ARN *MSH3*, en células HEK293 transfectadas con el editor de exones de ARN de dominio de unión

en tándem dirigido al intrón 1 de *HTT* y al intrón 5 de *MSH3*.

Las Figuras 22A-22C presentan resultados que muestran que los editores de exones de ARN del dominio de unión en tándem dirigidos a *HTT* y *MSH3* exhiben un trans-empalme exitoso a ambos pre-ARNm. RT-qPCR de A) eficiencia de trans-empalme en el objetivo (ONT) de *HTT* (a través del dominio de unión de *HTT*), B) eficiencia de trans-empalme quimérico *HTT-MSH3* (a través del dominio de unión de *MSH3*) y C) nivel de expresión de transcripción de ARN *MSH3*, en células HEK293 transfectadas con el editor de exones de ARN de dominio de unión en tándem dirigido al intrón 1 de *HTT* y al intrón 15 de *MSH3*.

La Figura 23 representa una estructura general de construcciones ejemplares que codifican editores de exones de ARN dirigidos al intrón 3 de *HTT*. El editor de exones ejemplar representado comprende un UTR 5', las secuencias codificantes de los exones 1-3, un sitio donante de empalme, un enlazador, un dominio de unión y una secuencia terminadora. Se varió el dominio de unión para apuntar a diferentes posiciones a lo largo del intrón 3 de *HTT*. En algunas modalidades, el 5' UTR es el 5' UTR *HTT*; en algunas modalidades, el enlazador es el enlazador de 40mer. En algunas modalidades, el promotor es un promotor CMV; en algunas modalidades, el 5' UTR es el *HTT* 5' UTR; en algunas modalidades, el enlazador es el enlazador de 40 mer; en algunas modalidades, el promotor CMV se combina con el 5' UTR de *HTT*; en algunas modalidades, el promotor CMV se combina con el 5' UTR de *HTT* y el enlazador de 40 mer; en algunas modalidades, se incluye una etiqueta de epítipo, un ejemplo de la cual es una etiqueta 3X FLAG de extremo N para la detección de proteínas en el objetivo. Las modalidades mencionadas anteriormente se pueden combinar, en donde al menos una de estas modalidades se incluye en un Editor de Exones de ARN y cualquiera y todas sus combinaciones, incluida una combinación de todas estas modalidades en un Editor de Exones de ARN.

La Figura 24 muestra la actividad de varios editores de exones de ARN dirigidos al intrón 3 de *HTT* de ejemplo que muestran diferentes niveles de eficiencia de trans-empalme (% de reemplazo) dependiendo del sitio de unión dentro del intrón al que se dirige el dominio de unión. Las células HEK293 fueron transfectadas con editores de exones de ARN dirigidos al intrón 3 de *HTT* que se dirigen a varias regiones del intrón 3. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR. La nomenclatura para el dominio de unión es: (posición de la base del nucleótido en el intrón 3 de la posición de inicio del extremo 5' del dominio de unión, numeración de acuerdo con la SEQ ID NO: 155)_(longitud (nt) del dominio de unión).

La Figura 25 muestra una comparación directa de los editores de exones dirigidos al intrón 2 de *HTT* con los editores de exones dirigidos al intrón 3 de *HTT*. Las células HEK293 se transfectaron con editores de exones de ARN dirigidos al intrón 2 y al intrón 3 de *HTT* en paralelo. Las células se recolectaron 48 horas después de la transfección y se analizaron para

determinar la eficiencia del trans-empalme mediante RT-qPCR. La nomenclatura para el dominio de unión es: (posición de la base del nucleótido en el intrón 3 de la posición de inicio del extremo 5' del dominio de unión)_(longitud (nt) del dominio de unión).

5 La Figura 26 muestra que la mitigación del auto empalme no afecta la eficiencia del trans-empalme de los editores de exones HTT. En la Tabla 1 se presentan los sitios de empalme críptico identificados en el editor de exones original y los cambios de secuencia realizados en los editores de exones mitigados con auto empalme. Las células HEK293 se transfectaron con editores de exones de ARN dirigidos al intrón 2 de HTT (HTT_intron2_12061_150) con o sin
10 mitigación de auto empalme. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR.

La Figura 27 muestra una imagen representativa de RT-qPCR y análisis de transferencia Western de lisados de células HEK293 transfectadas con editores de exones de ARN HTT que
15 prueban 2 promotores y diferentes combinaciones de 5' UTR. Las células HEK293 se transfectaron con editores de exones HTT etiquetados con FLAG en el extremo N impulsados por el promotor CMV o CAGGS, con o sin el 5' UTR HTT, y se probó el sitio de empalme de tipo salvaje (GTAAGT) dirigido al intrón 2 (HTT_intron2_12061_150), el mutante de empalme dirigido al intrón 2 (HTT_intron2_12061_150) o el sitio de empalme de tipo salvaje con un
20 dominio de unión no dirigido (NBD). El ARN de estas células se sometió a RT-qPCR (panel superior). Se utilizó el anticuerpo α -FLAG para detectar la proteína generada luego de un trans-empalme exitoso (que comprende un epítipo FLAG de extremo N) en lisados de células completas (panel inferior). ONT: proteína HTT trans-empalmada con éxito en el objetivo. NSP: proteína no empalmada.

25 La Figura 28 muestra que la proteína no empalmada (NSP) se reduce de manera combinatoria mediante la inclusión de tres repeticiones en tándem del sitio de unión del ARNnp U1 (3X UBS; SEQ ID NO: 345) y un elemento rico en AU (ARE; SEQ ID NO: 346) en un editor de exones dirigido al intrón 1 de 5' HTT ejemplar. Las células HEK293 fueron transfectadas con editores
30 de exones de ARN dirigidos al intrón 1 de HTT (HTT_intron1_11704_100) que fueron variados en su región de enlace para incluir los elementos de reducción de NSP indicados. Las células se recolectaron 48 horas después de la transfección y se sometieron a análisis de transferencia Western. ONT: proteína HTT trans-empalmada con éxito en el objetivo. NSP: proteína no empalmada.

35 La Figura 29 muestra que la proteína no empalmada (NSP) se reduce de manera combinatoria mediante la inclusión de tres repeticiones en tándem del sitio de unión del ARNnp U1 (3X UBS; SEQ ID NO: 345) y un elemento rico en AU (ARE; SEQ ID NO: 346) en el editor de exones dirigido al intrón 2 de 5' HTT. Las células HEK293 fueron transfectadas con editores de exones
40 de ARN dirigidos al intrón 2 de HTT (HTT_intron2_12061_150) que fueron variados en su región de enlace para incluir los elementos de reducción de NSP indicados. Las células se

recolectaron 48 horas después de la transfección, se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR (panel superior) o se sometieron a análisis de transferencia Western (panel inferior).

5 La Figura 30 muestra un mecanismo de acción de un editor de exones de ARN con dominio de unión en tándem dirigido al pre-ARNm de intrón 2 *HTT* y *MSH3*. La expresión del editor de exones está dirigida por un promotor CMV y contiene la UTR 5' de *HTT*, la etiqueta de extremo N 3X FLAG, la secuencia codificante de los exones 1 y 2 de *HTT*, el sitio donante de empalme, el enlazador, el dominio de unión a *MSH3* (dirigido al intrón 5 o al intrón 15 del pre-ARNm
10 *MSH3*), el dominio de unión a *HTT* (*HTT_intron2_12061_150*) y una secuencia terminadora. El dominio de unión de *HTT* apuntará al editor de exones para producir el ARN *HTT* corregido después de un trans-empalme exitoso, mientras que el dominio de unión de *MSH3* apuntará al editor de exones para producir una molécula de ARN quimérica *HTT* de exón 1+2 -*MSH3* con un codón de parada prematuro que estará sujeto a una desintegración mediada por sin sentido (NMD) y conducirá a la reducción posterior de la expresión de *MSH3*.
15

Las FIGS. 31A y 31B presentan perfiles de RT-qPCR de trans-empalme de *HTT* y producción de quimera *HTT-MSH3* (a través de trans-empalme de *MSH3*) en editores de exones del dominio de unión en tándem. RT-qPCR de A) eficiencia de trans-empalme en el objetivo (ONT) de *HTT* (a través del dominio de unión dirigido al intrón 2 de *HTT*) y B) eficiencia de trans-empalme quimérico *HTT-MSH3* (a través del dominio de unión dirigido al intrón 5 de *MSH3*), en células HEK293 transfectadas con el editor de exones de ARN de dominio de unión en tándem dirigido al intrón 2 de *HTT* y al intrón 5 de *MSH3*. El dominio de unión dirigido al intrón 2 de *HTT* fue *HTT_intron2_12061_150* para todos los editores de exones probados aquí,
20 mientras que *MSH3_intron5_213_100* y *MSH3_intron5_188_150* se probaron para los dominios de unión dirigidos al intrón 5 de *MSH3*. Los dominios de unión se posicionaron en tándem y el orden de los dominios de unión fue el indicado. También se probó un editor de exones para cada dominio de unión de *MSH3* con la triple hélice MALAT1 colocada entre los dos dominios de unión en tándem.
25

30 Las Figuras 32A y 32B presentan perfiles de RT-qPCR de trans-empalme de *HTT* y producción de quimera *HTT-MSH3* (a través de trans-empalme de *MSH3*) en editores de exones del dominio de unión en tándem. RT-qPCR de A) eficiencia de trans-empalme en el objetivo (ONT) de *HTT* (a través del dominio de unión dirigido al intrón 2 de *HTT*), y B) eficiencia de trans-empalme quimérico *HTT-MSH3* (a través del dominio de unión dirigido al intrón 15 de *MSH3*), en células HEK293 transfectadas con el editor de exones de ARN de dominio de unión en tándem dirigido al intrón 2 de *HTT* y al intrón 15 de *MSH3*. El dominio de unión dirigido al intrón 2 de *HTT* fue *HTT_intron2_12061_150* para todos los editores de exones probados aquí,
35 mientras que *MSH3_intron15_6523_120* y *MSH3_intron15_6498_150* se probaron para los dominios de unión dirigidos al intrón 15 de *MSH3*. Los dominios de unión se posicionaron en tándem y el orden de los dominios de unión fue el indicado. También se probó un editor de
40

exones para cada dominio de unión de MSH3 con la triple hélice MALAT1 colocada entre los dos dominios de unión en tándem.

La Figura 33 presenta una caricatura que representa el silenciamiento de MSH3 por un miARN dirigido al ARNm de *MSH3*. MSH3 puede ser silenciado por miARN que atacan al ARNm de *MSH3* y degradan la transcripción.

La Figura 34 muestra el análisis de transferencia Western de construcciones de ARNi dirigidas al exón 23 de MSH3. Se analizaron las construcciones (SEQ ID NO: 234, 235, 238-241) que comprenden la secuencia activa del miARN dirigido al exón 23 de MSH3 TTAATCCATAACTCCTTGC (SEQ ID NO: 224), así como las construcciones de control (SEQ ID NO: 236, 237, 242 y 243). Se realizó el análisis ImageJ en los análisis de transferencia Western para analizar el silenciamiento de la proteína MSH3 (panel superior). Se diseñaron y probaron ARNh_c impulsados por el promotor U6 y miméticos de pri-miARN impulsados por el promotor CMV. Las variaciones incluyen la ubicación de la hebra (brazo 5' o brazo 3') de la hebra guía, incluyendo un bulto en la estructura del tallo y variando el andamiaje de miARN. Los controles negativos incluyen construcciones que contienen una secuencia no dirigida o un control sin bucle de horquilla.

La Figura 35 muestra análisis de transferencia Western y RT-qPCR de las construcciones ARNi dirigidas a MSH3. Se analizaron construcciones que codifican miARN dirigidos a diferentes regiones de la transcripción *MSH3*. Se realizó el análisis ImageJ en los análisis de transferencia Western para analizar el silenciamiento de la proteína MSH3. Se diseñaron y probaron imitadores de pri-miARN impulsados por el promotor CMV dirigidos a diferentes secuencias exónicas de MSH3.

La Figura 36 presenta una caricatura que representa el silenciamiento de MSH3 por ARN antisentido basado en ARN nuclear pequeño (ARNnp) (ARNas). MSH3 puede ser inactivado por ARN antisentido codificados en un andamiaje de ARNnp que se unen a las uniones de empalme de *MSH3*, evitando la inclusión de exones. Esto conduce a la omisión de exones y a la generación de un codón de parada prematuro, lo que en última instancia causa NMD de la transcripción MSH3. La ilustración aquí representa un ejemplo de un modulador de empalme de MSH3 que apunta a la omisión del exón 2.

La Figura 37 representa los niveles relativos de expresión de ARN de MSH3 de las uniones exón 1- exón 2 y exón 2- exón 3 en moduladores de empalme de MSH3 que apuntan a la omisión del exón 2 (abajo) y una caricatura que ilustra la región objetivo del ARNas en el pre-ARNm MSH3 (arriba). Las transcripciones del modulador de empalme son las SEQ ID NO: 284-287. SEQ ID NO: 284: U7SmOPT que contiene MSH3 In1/Ex2 ARNas (SEQ ID NO: 274); SEQ ID NO: 285 (U7SmOPT que contiene MSH3 Ex2/In2 ARNas (SEQ ID NO: 275); SEQ ID NO: 286: U7SmOPT que contiene MSH3 In1/Ex2 + Ex2/In2 ARNas (SEQ ID NO: 276); SEQ

ID NO: 287: MSH3 In1/Ex2 + Ex2/In2 ARNas largo (160 nt) (SEQ ID NO: 277).

La Figura 38 representa los niveles relativos de expresión de ARN de MSH3 de las uniones exón 2 - exón 3 y exón 3 - exón 4 en moduladores de empalme de MSH3 que apuntan a la omisión del exón 3 (abajo) y una caricatura que ilustra la región objetivo del ARNas en el pre-ARNm de MSH3 (arriba). Las transcripciones del modulador de empalme son las SEQ ID NO: 288-290. SEQ ID NO: 288: U7SmOPT que contiene MSH3 In2/Ex3 ARNas (SEQ ID NO: 278); SEQ ID NO: 289: U7SmOPT que contiene MSH3 Ex3/In3 ARNas (SEQ ID NO: 279); SEQ ID NO: 290: U7SmOPT que contiene MSH3 In2/Ex3 + Ex3/In3 ARNas (SEQ ID NO: 280).

La Figura 39 representa los niveles relativos de expresión de ARN de MSH3 de las uniones exón 3 - exón 4 y exón 4 - exón 5 en moduladores de empalme de MSH3 que apuntan a la omisión del exón 4 (abajo) y una caricatura que ilustra la región objetivo del ARNas en el pre-ARNm de MSH3 (arriba). Las transcripciones del modulador de empalme son las SEQ ID NO: 291-293. SEQ ID NO: 291: U7SmOPT que contiene MSH3 In3/Ex4 ARNas (SEQ ID NO: 281); SEQ ID NO: 292: U7SmOPT que contiene MSH3 Ex4/In4 ARNas (SEQ ID NO: 282); SEQ ID NO: 293: U7SmOPT que contiene MSH3 In3/Ex4 + Ex4/In4 ARNas (SEQ ID NO: 283).

La Figura 40 representa los niveles relativos de expresión de ARN de MSH3 de las uniones exón 2 - exón 3 y exón 3 - exón 4 en moduladores de empalme de MSH3 que apuntan a la omisión del exón 3 (abajo) y una caricatura que ilustra la región objetivo del ARNas en el pre-ARNm de MSH3 (arriba). Las transcripciones del modulador de empalme son las SEQ ID NO: 290, 324-326. SEQ ID NO: 290: U7SmOPT que contiene MSH3 In2/Ex3 + Ex3/In3 ARNas (SEQ ID NO: 280); SEQ ID NO: 324: Modulador de empalme MSH3 U7 SmOPT que contiene In3/Ex3/In2 ARNas (SEQ ID NO: 299); SEQ ID NO: 325: Modulador de empalme MSH3 U7 SmOPT que contiene In3/Ex3-1 + enlazador + Ex3/In2-1 ARNas (SEQ ID NO: 300 + 298 + 301); SEQ ID NO: 326: Modulador de empalme MSH3 U2 que contiene In3/Ex3-2 + enlazador + Ex3/In2-2 ARNas (SEQ ID NO: 302 + 298 + 303).

La Figura 41 representa los niveles relativos de expresión de ARN de MSH3 de las uniones exón 5 - exón 6 y exón 6 - exón 7 en moduladores de empalme de MSH3 que apuntan a la omisión del exón 6 (abajo) y una caricatura que ilustra la región objetivo del ARNas en el pre-ARNm de MSH3 (arriba). Las transcripciones del modulador de empalme son las SEQ ID NO: 327-329. SEQ ID NO: 327: Modulador de empalme MSH3 U7 SmOPT que contiene In6/Ex6/In5 ARNas (SEQ ID NO: 304); SEQ ID NO: 328: Modulador de empalme MSH3 U7 SmOPT que contiene In6/Ex6-1 + enlazador + Ex6/In5-1 ARNas (SEQ ID NO: 305 + 298 + 306); SEQ ID NO: 329: Modulador de empalme MSH3 U2 que contiene In6/Ex6-2 + enlazador + Ex6/In5-2 ARNas (SEQ ID NO: 307 + 298 + 308).

La Figura 42 representa los niveles relativos de expresión de ARN de MSH3 de las uniones exón 6 - exón 7 y exón 7 - exón 8 en moduladores de empalme de MSH3 que apuntan a la

- omisión del exón 7 (abajo) y una caricatura que ilustra la región objetivo del ARNas en el pre-ARNm de *MSH3* (arriba). Las transcripciones del modulador de empalme son las SEQ ID NO: 330-332. El ARN antisentido (ARNas) comprendido en SEQ ID NO: 330 es *MSH3* In7/Ex7/In6 ARNas (región ARNas SEQ ID NO: 309); los ARNas comprendidos en SEQ ID NO: 331 son In7/Ex7-1 (región ARNas SEQ ID NO: 310) + enlazador (SEQ ID NO: 298) + Ex7/In6-1 (región ARNas SEQ ID NO: 311); los ARNas comprendidos en SEQ ID NO: 332 son In7/Ex7-2 (región ARNas SEQ ID NO: 312) + enlazador (SEQ ID NO: 298) + Ex7/In6-2 (región ARNas SEQ ID NO: 313).
- La Figura 43 representa los niveles relativos de expresión de ARN de *MSH3* de las uniones exón 7 - exón 8 y exón 8 - exón 9 en moduladores de empalme de *MSH3* que apuntan a la omisión del exón 8 (abajo) y una caricatura que ilustra la región objetivo del ARNas en el pre-ARNm de *MSH3* (arriba). Las transcripciones del modulador de empalme son las SEQ ID NO: 333-335. SEQ ID NO: 333: Modulador de empalme *MSH3* U7 SmOPT que contiene In8/Ex8/In7 ARNas (SEQ ID NO: 314); SEQ ID NO: 334: Modulador de empalme *MSH3* U7 SmOPT que contiene In8/Ex8-1 + enlazador + Ex8/In7-1 ARNas (SEQ ID NO: 315 + 298 + 316); SEQ ID NO: 335: Modulador de empalme *MSH3* U2 que contiene In8/Ex8-2 + enlazador + Ex8/In7-2 ARNas (SEQ ID NO: 317 + 298 + 318).
- La Figura 44 representa los niveles relativos de expresión de ARN de *MSH3* de las uniones exón 14 - exón 15 y exón 15 - exón 16 en moduladores de empalme de *MSH3* que apuntan a la omisión del exón 15 (abajo) y una caricatura que ilustra la región objetivo del ARNas en el pre-ARNm de *MSH3* (arriba). Las transcripciones del modulador de empalme son las SEQ ID NO: 336-338. SEQ ID NO: 336: Modulador de empalme *MSH3* U7 SmOPT que contiene In15/Ex15/In14 ARNas (SEQ ID NO: 319); SEQ ID NO: 337: Modulador de empalme *MSH3* U7 SmOPT que contiene In15/Ex15-1 + enlazador + Ex15/In14-1 ARNas (SEQ ID NO: 320 + 298 + 321); SEQ ID NO: 338: Modulador de empalme *MSH3* U2 que contiene In15/Ex15-2 + enlazador + Ex15/In14-2 ARNas (SEQ ID NO: 322 + 298 + 323).
- Las FIGS. 45A, 45B y 45C muestran que los moduladores de empalme del exón 7 de *MSH3* muestran una reducción de los niveles de proteína y ARN de *MSH3*. Las células HEK293 se transfectaron con moduladores de empalme basados en ARNnp diseñados para omitir el exón 7 de *MSH3*. Las células se recolectaron 48 horas después de la transfección, se analizaron para detectar el silenciamiento de *MSH3* mediante RT-qPCR (Figura 45B) o se sometieron a análisis de transferencia Western (Figura 45C). Las transcripciones del modulador de empalme son las SEQ ID NO: 330-332. SEQ ID NO: 330: Modulador de empalme *MSH3* U7 SmOPT que contiene In7/Ex7/In6 ARNas (SEQ ID NO: 309); SEQ ID NO: 331: Modulador de empalme *MSH3* U7 SmOPT que contiene In7/Ex7-1 + enlazador + Ex7/In6-1 ARNas (SEQ ID NO: 310 + 298 + 311); SEQ ID NO: 332: Modulador de empalme *MSH3* U2 que contiene In7/Ex7-2 + enlazador + Ex7/In6-2 ARNas (SEQ ID NO: 312 + 298 + 313).

La Figura 46 presenta una estrategia de combinación para corregir la *HTT* mutante mediante el trans-empalme y el silenciamiento de especies de *HTT* no editadas (incluyendo *HTT1a*) con un microARN (miARN).

5 La Figura 47 muestra una molécula híbrida vectorizada que combina un editor de exones *HTT* con un miARN dirigido al ARNm *HTT* sin editar. Se puede agregar un ARN de horquilla corta (ARNhc) o un microARN (miARN) diseñado para reducir la expresión del gen *HTT* a un editor de exones dentro del mismo cistrón (por ejemplo, en un intrón del editor de exones) o como
10 un cistrón separado con sus propias secuencias reguladoras. El ARNi puede reducir la expresión del objetivo no editado (por ejemplo, *HTT* mutante). Con el propósito de reducir selectivamente *HTT* no editado con ARNi el Editor de Exones comprende un CDS de *HTT* que comprende una porción de secuencia alterada que hace que la *HTT* editada sea resistente al ARNhc o miARN.

15 Las FIGS. 48A y 48B presentan perfiles de trans-empalme y de silenciamiento de *HTT* del editor de exones *HTT*, miARN-1 *HTT*, y una molécula híbrida dual del editor de exones *HTT* y miARN-1 *HTT*. A) % de transcripciones *HTT* trans-empalmadas (editadas) en todas las transcripciones *HTT*, y B) número de copias de transcripciones *HTT* no editadas y trans-empalmadas (editadas) en cada uno de los tratamientos. Las células HEK293 se transfectaron
20 con el editor de exones de ARN dirigido al intrón 2 de *HTT*, miARN-1 *HTT* y una molécula híbrida dual de editor de exones dirigido al intrón 2 de *HTT* (SEQ ID NO: 204) + miARN-1 *HTT* (secuencia codificante de miARN-1 *HTT* SEQ ID NO: 341 que contiene la secuencia activa de miARN-1 *HTT* SEQ ID NO: 339). Las células se recolectaron 48 horas después de la transfección y el ARN se sometió a RT- qPCR. Molécula híbrida dual del editor de exones *HTT*
25 y miARN-1 *HTT* (SEQ ID NO: 354), cuyo híbrido dual comprende SEQ ID NO: 341 y SEQ ID NO: 204.

Las FIGS. 49A-49C presentan perfiles de trans-empalme y de silenciamiento de *HTT* de moléculas que contienen miARN-1 *HTT* y miARN-2 *HTT* in vitro. Las células HEK293 se
30 transfectaron con el editor de exones de ARN dirigido al intrón 2 de *HTT* y moléculas híbridas duales del editor de exones dirigido al intrón 2 de *HTT* + miARN-1 *HTT* o miARN-2 *HTT*. Las células se recolectaron 48 horas después de la transfección y el ARN se sometió a análisis de RT-qPCR para A) perfil de trans-empalme, B) perfil de silenciamiento de *HTT* y C) análisis del número de copias de *HTT*. Molécula híbrida dual del editor de exones *HTT* y miARN-1 *HTT*
35 (SEQ ID NO: 354), cuyo híbrido dual comprende SEQ ID NO: 341 y SEQ ID NO: 204. Molécula híbrida dual del editor de exones *HTT* y miARN-2 *HTT* (SEQ ID NO: 355), cuyo híbrido dual comprende SEQ ID NO: 344 y SEQ ID NO: 204.

La Figura 50 muestra que el miARN-1 *HTT* silencia con éxito las transcripciones *HTT* no editadas con una interacción mínima con el editor de exones y la transcripción *HTT* editada.
40 Las células HEK293 fueron transfectadas con un editor de exones *HTT* +/- moléculas híbridas

duales miARN-1 HTT. Las células se recolectaron 48 horas después de la transfección, se sometieron a ensayo para determinar las eficiencias del silenciamiento de HTT y en el trans-empalme por RT-qPCR (paneles superiores) o se sometieron al análisis de transferencia Western (paneles inferiores). Molécula híbrida dual del editor de exones HTT y miARN-1 HTT (SEQ ID NO: 354), cuyo híbrido dual comprende SEQ ID NO: 341 y SEQ ID NO: 204.

La Figura 51 presenta una caricatura que representa la reducción de MSH3 por modulación de empalme en combinación con trans-empalme HTT.

Las Figuras 52A y 52B presentan resultados que muestran el desempeño de las moléculas híbridas duales del editor de exones HTT + modulador de empalme MSH3. Las células HEK293 fueron transfectadas con un editor de exones HTT +/- modulador de empalme MSH3. Las células se recolectaron 48 horas después de la transfección y se analizaron A) los perfiles de trans-empalme y B) los perfiles de silenciamiento de MSH3 por RT-qPCR. Moléculas híbridas duales Modulador de empalme MSH3 + Editor de exones HTT: Modulador de empalme MSH3 + Editor de exones HTT (SEQ ID NO: 356), cuyo híbrido dual comprende SEQ ID NO: 331 y SEQ ID NO: 204.

Las Figuras 53A-D presentan una comparación de AAV autocomplementario (scAAV) con AAV de cadena simple (ssAAV) utilizando una molécula híbrida dual de editor de exones HTT + modulador de empalme MSH3. Las células HEK293 fueron transducidas con scAAV o ssAAV que expresan una molécula híbrida dual: editor de exones HTT + modulador de empalme con omisión del exón 7 MSH3. Se utilizó el serotipo AAV2. Las células se recolectaron 48 horas después de la transducción y se sometieron a análisis de transferencia RT-qPCR y Western. Las secuencias indicadas se insertaron entre los ITR de AAV2, ya sea ssAAV o scAAV. Molécula híbrida dual Modulador de empalme MSH3 + Editor de exones HTT (SEQ ID NO: 357), cuyo híbrido dual comprende SEQ ID NO: 331 y SEQ ID NO: 204, en una orientación de cabeza a cabeza.

La Figura 54 presenta resultados que muestran el rendimiento de las moléculas híbridas triples editor de exones HTT + miARN HTT + modulador de empalme MSH3 en comparación con sus controles. Se transfectaron células HEK293 con editor de exones HTT +/- miARN-1 HTT o miARN-2 HTT +/- modulador de empalme MSH3. Las células se cosecharon 48 horas después de la transfección y se analizaron para el panel superior) % de transcripciones trans-empalmadas HTT, panel central) perfiles de silenciamiento de HTT, y el panel inferior) perfiles de silenciamiento de MSH3 por RT-qPCR. También se analizaron moléculas híbridas de control que comprendían un editor de exones con una mutación donante de empalme, un modulador de empalme con una secuencia de ARNas codificada o un miARN con una secuencia de ARNas codificada. El valor "1" para el miARN HTT indica el miARN-1 HTT; el valor "2" para el miARN HTT indica que se utilizó el miARN-2 HTT. SM, mutante de empalme. Scr, control desordenado. Se utilizó miR-33 para el andamiaje de miARN. Editor de exones

HTT (SEQ ID NO: 204); molécula híbrida dual de Editor de exones HTT y miARN-1 HTT (SEQ ID NO: 354), cuyo híbrido dual comprende SEQ ID NO: 341 y SEQ ID NO: 204; moléculas híbridas duales Editor de exones HTT + Modulador de empalme MSH3: Modulador de empalme MSH3 + Editor de exones HTT (SEQ ID NO: 356), cuyo híbrido dual comprende SEQ ID NO: 331 y SEQ ID NO: 204; molécula híbrida triple de editor de exones HTT + miARN-1 HTT + Modulador de empalme MSH3 (SEQ ID NO: 358), cuyo híbrido triple comprende SEQ ID NO: 331, SEQ ID NO: 341 y SEQ ID NO: 204; Molécula híbrida triple de editor de exones HTT + miARN-2 HTT + modulador de empalme MSH3 (SEQ ID NO: 359), cuyo híbrido triple comprende SEQ ID NO: 331, SEQ ID NO: 344 y SEQ ID NO: 204; molécula híbrida dual de editor de exones HTT y miARN-2 HTT (SEQ ID NO: 355), cuyo híbrido dual comprende SEQ ID NO: 344 y SEQ ID NO: 204.

La Figura 55 muestra los perfiles de trans-empalme de moléculas HD en iCell GlutaNeurons medidos por RT-ddPCR y transferencia Western. Las iCell GlutaNeurons se transdujeron con una molécula híbrida dual de editor de exones HTT + miARN HTT o una molécula híbrida dual de editor de exones HTT + modulador de empalme MSH3, ambas empaquetadas en AAV2.7m8. Las células se recolectaron para RT-ddPCR y análisis de transferencia Western después de 18-21 días. Las secuencias indicadas se insertaron entre los ITR de AAV2.7m8. Molécula híbrida dual de editor de exones HTT y miARN-1 HTT (SEQ ID NO: 354), cuyo híbrido dual comprende SEQ ID NO: 341 y SEQ ID NO: 204; molécula híbrida dual de editor de exones HTT + modulador de empalme MSH3 (SEQ ID NO: 357), cuyo híbrido dual comprende SEQ ID NO: 331 y SEQ ID NO: 204, en una orientación cabeza a cabeza; molécula híbrida triple de editor de exones HTT + miARN-1 HTT + modulador de empalme MSH3 (SEQ ID NO: 358), cuyo híbrido triple comprende SEQ ID NO: 331, SEQ ID NO: 341 y SEQ ID NO: 204.

La Figura 56 muestra los perfiles de silenciamiento de miARN de HTT en iCell GlutaNeurons medidos por RT-ddPCR y análisis de transferencia Western. Las iCell GlutaNeurons se transdujeron con una molécula híbrida dual de editor de exones HTT + miARN HTT o una molécula híbrida dual de editor de exones HTT + modulador de empalme MSH3, ambas empaquetadas en AAV2.7m8. Las células se recolectaron para RT-ddPCR y análisis de transferencia Western después de 18 días. Las secuencias indicadas se insertaron entre los ITR de AAV2.7m8. Molécula híbrida dual de editor de exones HTT y miARN-1 HTT (SEQ ID NO: 354), cuyo híbrido dual comprende SEQ ID NO: 341 y SEQ ID NO: 204; molécula híbrida dual de editor de exones HTT + modulador de empalme MSH3 (SEQ ID NO: 357), cuyo híbrido dual comprende SEQ ID NO: 331 y SEQ ID NO: 204, en una orientación cabeza a cabeza.

La Figura 57 muestra un perfil de silenciamiento de MSH3 del modulador de empalme de MSH3 en iCell GlutaNeurons medidos por RT-ddPCR y análisis de transferencia Western. Las iCell GlutaNeurons se transdujeron con una molécula híbrida dual de editor de exones HTT + miARN-1 HTT o una molécula híbrida dual de editor de exones HTT + modulador de empalme MSH3, ambas empaquetadas en AAV2.7m8. Las células se recolectaron para RT-ddPCR y

análisis de transferencia Western después de 18 días. Las secuencias indicadas se insertaron entre los ITR de AAV2.7m8. Molécula híbrida dual de editor de exones HTT y miARN-1 HTT (SEQ ID NO: 354), cuyo híbrido dual comprende SEQ ID NO: 341 y SEQ ID NO: 204; molécula híbrida dual de editor de exones HTT + modulador de empalme MSH3 (SEQ ID NO: 357), cuyo

5 híbrido dual comprende SEQ ID NO: 331 y SEQ ID NO: 204, en una orientación cabeza a cabeza.

La Figura 58 muestra los perfiles de trans-empalme de HTT en el cerebro del ratón BAC-CAG. Se realizaron inyecciones neonatales de ICV (a 1E+11 o 3E+11 vg/animal) en ratones BAC-CAG con las moléculas HD indicadas empaquetadas en AAV9. La corteza y el cuerpo estriado del ratón se recolectaron 4 semanas después de la inyección y se perfilaron las eficiencias del reemplazo de exones HTT por trans-empalme mediante RT-ddPCR y análisis de transferencia Western. SEQ ID NO: 369: scAAV, modulador de empalme Msh3 de ratón + CMVp:Editor de exones (híbrido dual de SEQ ID NO: 362 y SEQ ID NO: 204, orientación cabeza a cabeza);

10 SEQ ID NO: 370: ssAAV, modulador de empalme Msh3 de ratón + CMVp:Editor de exones (híbrido dual de SEQ ID NO: 362 y SEQ ID NO: 204, orientación cabeza a cabeza); SEQ ID NO: 371: ssAAV, modulador de empalme Msh3 de ratón + CAGGSp:Editor de exones (híbrido dual de SEQ ID NO: 362 y SEQ ID NO: 204, orientación cabeza a cabeza).

La Figura 59 representa una relación observada entre el número de copias de ARN del editor de exones y la eficiencia de trans-empalme (% de reemplazo de HTT) en el cerebro del ratón BAC-CAG. SEQ ID NO: 369: scAAV, modulador de empalme Msh3 de ratón + CMVp:Editor de exones (híbrido dual de SEQ ID NO: 362 y SEQ ID NO: 204, orientación cabeza a cabeza);

15 SEQ ID NO: 370: ssAAV, modulador de empalme Msh3 de ratón + CMVp:Editor de exones (híbrido dual de SEQ ID NO: 362 y SEQ ID NO: 204, orientación cabeza a cabeza); SEQ ID NO: 371: ssAAV, modulador de empalme Msh3 de ratón + CAGGSp:Editor de exones (híbrido dual de SEQ ID NO: 362 y SEQ ID NO: 204, orientación cabeza a cabeza).

DESCRIPCIÓN DETALLADA

Los siguientes ejemplos se proporcionan para ilustrar algunas modalidades de la presente divulgación, pero no pretenden limitar el alcance de la divulgación; se entenderá por su naturaleza ejemplar que se pueden utilizar alternativamente otros procedimientos, metodologías o técnicas conocidos por los expertos en la técnica.

Las composiciones y métodos que se describen en la presente memoria involucran moléculas de trans-empalme (por ejemplo, moléculas de trans-empalme de pre-ARNm) para tratar enfermedades o trastornos causados por una mutación en el gen *HTT*. Tales mutaciones comprenden una repetición de trinucleótido de CAG expandido en el exón 1 del gen *HTT*. Las composiciones y métodos que se describen en la presente memoria emplean moléculas de

35 trans-empalme de pre-ARNm para terapia génica (por ejemplo, terapia génica *in vivo*, por ejemplo, administrada por virus adenoasociado) para tratar enfermedades causadas por una

repetición trinucleotídica CAG expandida en *HTT*, tal como HD. Las composiciones y métodos descritos en este documento también emplean moléculas de trans-empalme de pre-ARNm para terapia génica (por ejemplo, terapia génica *in vivo*) en combinación con otros agentes terapéuticos descritos en este documento para tratar enfermedades causadas por al menos una mutación en *HTT*, tal como HD.

Las composiciones y métodos descritos en este documento también involucran moléculas terapéuticas que inhiben la expresión de MSH3. Las construcciones de silenciamiento de MSH3 se pueden utilizar solas en métodos terapéuticos o se pueden utilizar en combinación con agentes terapéuticos correctores de *HTT* y/o agentes terapéuticos de silenciamiento de *HTT*. Las construcciones de silenciamiento de MSH3 también se pueden utilizar en combinación con otros agentes terapéuticos distintos de los agentes terapéuticos correctores de *HTT*, como otros agentes terapéuticos que están diseñados para corregir trastornos de expansión de repeticiones de trinucleótidos.

Como se describe aquí, la HD es un trastorno neurodegenerativo progresivo hereditario, para el cual solo hay disponible terapia paliativa. Esta terapia incluye medicamentos, fisioterapia y asesoramiento, intervenciones que proporcionan cierto alivio sintomático. Aunque la presentación de la HD varía, la enfermedad se caracteriza por una pérdida progresiva de la capacidad de controlar el movimiento, regular las emociones y mantener los atributos cognitivos. La HD generalmente se presenta en personas de entre 30 y 40 años. La enfermedad se asocia con la pérdida de neuronas piramidales en la corteza, pérdida de neuronas espinosas medianas en el cuerpo estriado y pérdida de neuronas hipotalámicas. La causa genética de la HD es la herencia autosómica dominante de una repetición expandida del trinucleótido CAG en el exón 1 del gen *HTT*, en donde la presencia de más de 40 repeticiones de CAG en esta región es causante de la enfermedad. Ver la Figura 1.

El locus *HTT* es grande, abarca 180 kb y consta de 67 exones. La expresión del gen *HTT* es necesaria para el desarrollo normal. Aunque la proteína HTT se expresa ampliamente, el cerebro es el que se ve más gravemente afectado por la expansión patológica de repeticiones de trinucleótidos CAG, con efectos patológicos tempranos observados en el cuerpo estriado y la corteza motora. El mecanismo subyacente de la patogénesis de la HD es la expansión de repeticiones CAG somáticas en la *HTT* que ocurre en las regiones cerebrales afectadas (por ejemplo, el cuerpo estriado) de los pacientes con HD. Ver la Figura 2. La evidencia genética humana implica genes en la vía de reparación de desajustes de ADN (ej. *MSH2*, *MSH3*, *FAN1*, *MLH1*) en el control de este proceso y la modificación del resultado clínico de la HD.

A pesar de los considerables esfuerzos realizados por muchas compañías biotecnológicas y farmacéuticas para desarrollar terapias para la HD, actualmente no existen tratamientos que modifiquen la enfermedad para la HD. Los presentes inventores buscaron abordar esta necesidad utilizando una variedad de modalidades diferentes, cada una de las cuales puede

implementarse sola o en combinación para brindar una intervención terapéutica para la HD.

Reemplazo de exones por trans-empalme de pre-ARNm HTT

El reemplazo de exones mediante trans-empalme de pre-ARNm es un enfoque terapéutico muy adecuado para la HD debido a su capacidad de reemplazar el exón 1 de *HTT* mutante y al mismo tiempo preservar la copia de tipo salvaje de *HTT*. Además, este enfoque puede, en teoría, tratar al 100% de la población con HD, ya que puede abordar todo el espectro de variabilidad genética en los pacientes con HD. En otras palabras, no se limita a abordar variaciones específicas de pacientes con HD (por ejemplo, SNP), sino que sirve como un agente terapéutico pan-específico capaz de corregir defectos genéticos en la totalidad de uno o más exones.

Como se describe en este documento, los presentes inventores diseñaron y probaron editores de exones de ARN que se dirigen al intrón 1, intrón 2 o intrón 3 del pre-ARNm *HTT*. En consecuencia, estos editores de exones de ARN pueden reemplazar todo el exón 1, los exones 1 y 2, o los exones 1-3 del ARNm de *HTT*, corrigiendo así todas y cada una de las mutaciones patógenas presentes en el exón 1, los exones 1 y 2, o los exones 1-3 (por ejemplo, corrigiendo la repetición del trinucleótido CAG expandida en el exón 1 a un número de tipo salvaje), respectivamente. Tal como se muestra en la Figura 3, la edición de exones de ARN a través de trans-empalme permite el reemplazo del exón 1 de *HTT* mutante con el exón 1 de *HTT* de tipo salvaje utilizando cualquiera de los siguientes editores de exones dirigidos al intrón 1, un editor de exones dirigido al intrón 2 o un editor de exones dirigido al intrón 3.

En algunas modalidades, un editor de exones de ARN que se dirige al intrón 1 del pre-ARNm de *HTT* comprende un promotor CMV (SEQ ID NO: 137) (usado aquí para experimentación *in vitro* en células HEK293), un 5' UTR (por ejemplo, 5' UTR de *HTT* (SEQ ID NO: 136 o 192)), una etiqueta de epítipo (por ejemplo, una etiqueta 3X FLAG (SEQ ID NO: 4)) para la detección de la proteína HTT en el objetivo (ONT) generada después de un trans-empalme exitoso, un exón 1 de *HTT* (SEQ ID NO: 3) (que puede modificarse por codones seguido de la secuencia nativa), una secuencia donante de empalme (GTAAGT), un enlazador (por ejemplo, el enlazador de 40 mer (SEQ ID NO: 37)), un dominio de unión que se dirige a cualquiera de varias regiones en el pre-ARNm del intrón 1 de *HTT* y un terminador de triple hélice (p. ej. Terminador MALAT1 (SEQ ID NO: 5 o una versión modificada de este, por ejemplo, SEQ ID NO: 6)). Ver, por ejemplo, la Figura 4. En algunas modalidades, el editor de exones de ARN que se dirige al intrón 1 del pre-ARNm de *HTT* no comprende uno o más de un promotor de CMV (SEQ ID NO: 137), un 5' UTR (por ejemplo, 5' UTR de *HTT* (SEQ ID NO: 136 o 192)), una etiqueta de epítipo (por ejemplo, una etiqueta 3X FLAG (SEQ ID NO: 4)), la secuencia donante de empalme GTAAGT, el enlazador de 40 mer (SEQ ID NO: 37), un terminador de triple hélice (por ejemplo, terminador MALAT1 (SEQ ID NO: 5 o una versión modificada de este, por ejemplo, SEQ ID NO: 6)). En algunas modalidades, el promotor CMV puede reemplazarse por un promotor diferente. Se pueden elegir promotores que tengan propiedades

adecuadas para estudios *in vivo* y para un agente terapéutico que comprenda un editor de exones de ARN. En algunas modalidades, un editor de exones de ARN ejemplar no comprende una etiqueta FLAG ni ninguna etiqueta de epítipo. En tales modalidades, un agente terapéutico que comprende un editor de exones de ARN puede no comprender ninguna etiqueta de epítipo, lo que puede reducir el potencial de inmunogenicidad. La Figura 5 muestra que el nivel de reemplazo del exón *HTT* (eficiencia de trans-empalme) varía según el objetivo del dominio de unión en el intrón 1. Los resultados aquí presentados demuestran que el objetivo en el extremo 3' del intrón 1 cerca del punto de ramificación se correlaciona con una mayor eficiencia de trans-empalme en el pre-ARNm *HTT*. Ver la Figura 6.

Los presentes inventores buscaron luego mejorar aún más la eficiencia del trans-empalme de los editores de exones dirigidos al intrón 1 de *HTT* seleccionando un dominio de unión ejemplar, *HTT_intron1_11704_100*, y determinando los efectos de diferentes enlaces vinculados operativamente a este dominio de unión en el contexto de los editores de exones de ARN. Ver, por ejemplo, la Figura 7, que presenta un esquema de la disposición de un editor de exones de ARN ejemplar que comprende un dominio de unión, *HTT_intron1_11704_100*. Los resultados presentados en este documento demuestran que varios enlazadores identificados mediante procedimientos exhaustivos de prueba y error (enlazador GU de 23 mer, enlazador GU de 33 mer, enlazador GU de 34 mer, enlazador GU de 41 mer, 40 mer_2, 41 mer_2, 68 mer, 84 mer y 60 mer) confieren niveles más altos de eficiencia de trans-empalme en relación con los conferidos por el de 40 mer (SEQ ID NO: 37). Ver, por ejemplo, la Figura 8. Las secuencias del enlazador GU de 23 mer (contenido de guanina de ~69-70 %), enlazador GU de 33 mer (contenido de guanina de ~66-67 %), enlazador GU de 34 mer (contenido de guanina de ~67-68 %), enlazador GU de 41 mer (contenido de guanina de ~73-74 %), 40 mer_2, 41 mer_2, 68 mer, 84 mer y 60 mer corresponden a las SEQ ID NO: 38-46, respectivamente.

Los presentes inventores probaron a continuación editores de exones de ARN dirigidos al intrón 2 de *HTT*. Ver la Figura 9 para un esquema general de editores de exones de ARN dirigidos al intrón 2 de *HTT*. De manera similar a los editores de exones orientados al intrón 1, inicialmente se realizó un escaneo del dominio de unión en el intrón 2 para identificar los dominios de unión que se unen a las regiones objetivo en el intrón 2 que están asociadas con y promueven altos niveles de eficiencia de trans-empalme. Ver, por ejemplo, la Figura 10. Los presentes inventores determinaron que el extremo 3' del intrón, 10-20 nucleótidos (nt) en la parte anterior del punto de ramificación A, es una región particularmente eficaz para el trans-empalme en el intrón 2 de *HTT*. Ver, por ejemplo, la Figura 11. Análisis adicionales que examinaron la asociación de la longitud del dominio de unión en la región del intrón en la parte anterior del punto de ramificación con la funcionalidad identificaron longitudes de dominio de unión de entre 125-200 nt como las que tenían las mayores eficiencias relativas de trans-empalme. Ver, por ejemplo, la Figura 12.

Los presentes inventores investigaron más a fondo la funcionalidad combinatoria de los dominios de unión unidos operativamente a diferentes enlazadores como se muestra gráficamente en la Figura 13. Los presentes inventores examinaron los efectos de una variedad de enlaces ejemplares sobre la actividad de trans-empalme de un dominio de unión de alto rendimiento para el intrón 2 de *HTT* (*HTT_intron2_12061_150*). Tal como se muestra en la Figura 14, la actividad de trans-empalme de los editores de exones de ARN que comprenden *HTT_intron2_12061_150* unidos operativamente a los enlazadores indicados no varió significativamente. De hecho, solo se observó un efecto mínimo cuando *HTT_intron2_12061_150* se vinculó operativamente a diferentes enlazadores, lo que sugiere que el trans-empalme usando *HTT_intron2_12061_150* está altamente optimizado.

Los resultados presentados en este documento muestran que ninguno de los enlazadores probados mejoró significativamente la eficiencia del trans-empalme en comparación con la conferida por el enlazador de 40 meros, lo que sugiere que los editores de exones *HTT_intron2_12061_150* han logrado una alta eficiencia a través de la afinidad de unión y el reclutamiento de la maquinaria del espliceosoma relevante. Los resultados también indican que se pueden incluir diferentes secuencias de enlace en un editor de exones de ARN *HTT* que sea eficaz para inducir el trans-empalme. Los presentes inventores probaron experimentalmente si la inhibición del cis-empalme podría contribuir a un trans-empalme mejorado mediante el diseño de oligonucleótidos antisentido (ASO) que bloquean específicamente el sitio de cis-empalme competidor (ASO8 (SEQ ID NO: 133), ASO9 (SEQ ID NO: 134), ASO10 (SEQ ID NO: 135)), así como los sitios de cis-empalme involucrados en el empalme del intrón en la parte anterior (ASO2-7 (SEQ ID NO: 127-132, respectivamente)), y coexpresando cada uno de estos ASO con un editor de exones de ARN dirigido al intrón 2. Ver la Figura 16. La Figura 17 muestra que el editor de exones dirigido al intrón 2 de *HTT* (*HTT_intron2_12061_150*) en combinación con ASO6 (SEQ ID NO: 131), que bloquea el cis-empalme del intrón en la parte anterior, condujo a una mejora en las eficiencias del trans-empalme *in vitro*. Algunas modalidades descritas en este documento implican estrategias para bloquear estos eventos de cis-empalme utilizando ARN antisentido expresado en el mismo plásmido que el editor de exones de ARN *HTT*. Ver también la sección Modalidades Combinatorias a continuación.

Para confirmar la presencia de una proteína *HTT* quimérica *genuina* generada después del trans-empalme, se realizó un análisis de transferencia Western en lisados celulares hechos de células HEK293 transfectadas con editores de exones *HTT* seleccionados, en donde las transferencias se sondearon utilizando anticuerpos que se unen específicamente al epítipo FLAG de extremo N. Tal como se muestra en la Figura 18, el análisis de transferencia Western confirmó el trans-empalme en el objetivo (ONT) y la generación resultante de proteína en el objetivo generada después del trans-empalme al detectar la proteína *HTT* etiquetada con FLAG que tenía el peso molecular previsto. En particular, la intensidad de la banda de proteína anti-FLAG ONT se correlacionó con la eficiencia del trans-empalme (% de reemplazo de ARN).

Ver la Figura 18.

Los presentes inventores también diseñaron y probaron los editores de exones de ARN que se dirigen al intrón 3 *HTT*. Ver la Figura 23 para un esquema general de editores de exones de ARN dirigidos al intrón 3 de *HTT*. De manera similar a los editores de exones dirigidos al intrón 1 y al intrón 2, inicialmente se realizó un escaneo del dominio de unión para identificar los dominios de unión que se unen a las regiones objetivo en el intrón 3 que están asociadas con y promueven altos niveles de eficiencia de trans-empalme. Ver, por ejemplo, la Figura 24. Los presentes inventores determinaron que el extremo 3' del intrón, 5-15 nt en la parte anterior del punto de ramificación A, es una región particularmente eficaz para el trans-empalme en el intrón 3 de *HTT*.

Para comparar el rendimiento de los editores de exones dirigidos al intrón 2 con respecto a los editores de exones dirigidos al intrón 3, los presentes inventores realizaron transfecciones de editores de exones que contenían los dominios de unión de mayor rendimiento para cada intrón objetivo en paralelo y analizaron el ARN mediante RT-ddPCR. Los resultados presentados en este documento demuestran que los editores de exones dirigidos al intrón 3 de *HTT* exhiben una eficiencia de trans-empalme ligeramente mayor en comparación con los editores de exones dirigidos al intrón 2. Ver, por ejemplo, la Figura 25.

Para investigar el auto empalme críptico (ya sea por cis-empalme de concatémeros de AAV o por trans-empalme intermolecular) en los editores de exones *HTT* y las construcciones que los comprenden, se identificaron sitios de empalme críptico utilizando predicción in silico de sitios de auto empalme (ver Tabla 1) para mitigar dichos eventos no deseados. Se generó un editor de exones dirigido al intrón 2 de *HTT* mitigado por un sitio de empalme críptico ejemplar (SEQ ID NO: 203) (en el que se alteraron 4 de los sitios de auto empalme identificados) y se probó su eficiencia de trans-empalme y se comparó con la del editor de exones original (inalterado) (SEQ ID NO: 95). Los resultados presentados aquí confirman que los cambios menores de secuencia aplicados no influyen en la actividad de trans-empalme del editor de exones mitigado por el sitio de empalme críptico ejemplar. Ver, por ejemplo, la Figura 26.

Las rondas iniciales de moléculas del editor de exones *HTT* se expresaron a partir de un promotor CMV (SEQ ID NO: 137) y contenían el UTR 5' *HTT* nativo (SEQ ID NO: 136). A continuación, los presentes inventores exploraron cómo funcionan los editores de exones cuando se expresan desde un promotor CAGGS (SEQ ID NO: 196). Dado que el promotor CAGGS contiene un 5' UTR (exón CBA + exón de beta-globina de conejo), se incluyeron en este estudio pruebas de editores de exones con o sin el 5' UTR *HTT* nativo en combinación con el promotor CAGGS. Cabe destacar que, si bien los resultados se presentan en el contexto de editores de exones dirigidos al intrón 2 de *HTT*, los efectos de estas modificaciones son aplicables a editores de exones dirigidos a otros intrones.

En experimentos de transfección de células HEK293, los presentes inventores encontraron que el editor de exones impulsado por el promotor CAGGS que contiene tanto el UTR 5' de CAGGS como el UTR 5' de HTT poseía un porcentaje de reemplazo de ARNm de HTT similar al impulsado por un promotor CMV. Ver, por ejemplo, la Figura 27. Cuando se eliminó el UTR 5' de HTT del editor de exones impulsado por el promotor CAGGS, se observó una disminución en el % de reemplazo de ARN de HTT y un aumento significativo en la expresión de la proteína ONT mediante análisis de transferencia Western, lo que sugiere la presencia de un elemento activador de la traducción de proteínas en el UTR 5' de CAGGS. Al mismo tiempo, los presentes inventores observaron un aumento de las señales de proteína no empalmada (NSP) y de proteína de fondo (posible OFT) con la eliminación del UTR 5' de HTT.

Para explorar las estrategias de reducción de NSP en el contexto de los editores de exones 5' HTT, se incorporaron diversos elementos a los editores de exones dirigidos al intrón 1 de HTT y a los editores de exones dirigidos al intrón 2 de HTT y se probaron para evaluar su impacto, individualmente y en combinación, en el rendimiento del editor de exones. En el contexto de un contexto de NSP mínimo (es decir, codón de parada en o inmediatamente después de la secuencia donante de empalme en el Editor de exones para minimizar los neoantígenos), estas estrategias de reducción de NSP incluyen: 1) tres repeticiones en tándem del sitio de unión de ARNnp U1 (3X UBS) (SEQ ID NO: 345) y 2) un elemento rico en AU (ARE) (SEQ ID NO: 346).

En los editores de exones dirigidos al intrón 1 de HTT, los presentes inventores demostraron que la inclusión de 3X UBS en el dominio enlazador redujo los niveles de NSP en ~75 % en relación con el control de enlazador de solo 40 mer de referencia. La inclusión de ARE en el enlazador redujo los niveles de NSP en ~ 40 % en relación con el control del enlazador de referencia de solo 40 mer. La combinación de 3X UBS y ARE en el editor de exones redujo el nivel de NSP en ~ 88% en relación con el control del enlazador de referencia de solo 40 mer. Ver, por ejemplo, la Figura 28.

De manera similar, en los editores de exones dirigidos al intrón 2 de HTT, los presentes inventores demostraron que la inclusión de 3X UBS en el dominio de enlace redujo los niveles de NSP en ~ 38% en relación con el control de enlace de solo 40 mer de referencia. La inclusión de ARE en el enlazador redujo los niveles de NSP en ~ 33% en relación con el control del enlazador de referencia de solo 40 mer. La combinación de 3X UBS y ARE en el editor de exones trabajó combinatoriamente en la reducción del nivel de NSP en aproximadamente 66 % en comparación con el control del enlazador de referencia de solo 40 mer. Ver, por ejemplo, la Figura 29.

Los editores de exones de ARN dirigidos a un intrón *HTT* (por ejemplo, intrón 1, intrón 2 o intrón 3) pueden incluir elementos específicos del objetivo tales como, por ejemplo, un dominio de unión específico para un intrón objetivo *HTT* y una secuencia de dominio de codificación que codifica una secuencia de codificación *HTT* (por ejemplo, una secuencia que codifica todo

o parte del exón 1, exón 2 o exón 3, o cualquier combinación de estos (por ejemplo, exones 1 y 2 o exones 1-3)). Los editores de exones de ARN dirigidos a un intrón *HTT* pueden incluir uno o más elementos independientes del objetivo que pueden mejorar el funcionamiento de un editor de exones. Los elementos independientes del objetivo pueden incluir, por ejemplo, una etiqueta de epítipo, un enlazador, una secuencia donante de empalme, una o más repeticiones de un sitio de unión de ARNnp U1, un elemento rico en AU o un terminador. Las modalidades descritas en este documento pueden incluir uno o más de dichos elementos independientes del objetivo. Las modalidades descritas en este documento pueden excluir uno o más de dichos elementos independientes del objetivo. Los editores de exones de ARN que carecen de uno o más de los elementos independientes del objetivo descritos en este documento pueden ser capaces de efectuar la edición de exones. Por ejemplo, en algunas modalidades, los editores de exones de ARN descritos en este documento no incluyen un terminador de triple hélice, un terminador MALAT-1 o cualquier secuencia terminadora. En algunas modalidades, los editores de exones de ARN descritos en este documento no incluyen un promotor CMV o un promotor CAGGS. En algunas modalidades, los editores de exones de ARN descritos en este documento no incluyen ninguna secuencia 5' UTR, o no incluyen una secuencia *HTT* 5' UTR. En algunas modalidades, los editores de exones de ARN descritos en este documento no incluyen una secuencia de donante de empalme GTAAGT. En algunas modalidades, los editores de exones de ARN descritos en este documento no incluyen ninguna etiqueta de epítipo, o no incluyen una etiqueta FLAG 3X. En algunas modalidades, los editores de exones de ARN descritos en este documento no incluyen ninguno de los enlazadores descritos en este documento, incluido cualquiera de las SEQ ID NO: 37-46. En algunas modalidades, los editores de exones de ARN descritos en este documento no incluyen tres repeticiones en tándem del sitio de unión de ARNnp U1 (3X UBS) (SEQ ID NO: 345). En algunas modalidades, los editores de exones de ARN descritos en este documento no incluyen un elemento rico en AU (ARE) (SEQ ID NO: 346).

Inhibición de la expansión de repeticiones CAG somáticas

Cada vez hay más pruebas de que el mecanismo subyacente de la patogénesis de la HD es la inestabilidad somática del tracto de repetición CAG. Esto respalda la búsqueda de factores que modifiquen la expansión somática como objetivos terapéuticos viables. Los presentes inventores diseñaron y probaron estrategias para inhibir la expansión somática de CAG. Ver, por ejemplo, la Figura 19. Estudios recientes de asociación del genoma completo (GWAS) realizados en pacientes con HD identificaron muchos componentes de la vía de reparación de desajustes del ADN como modificadores genéticos clave asociados con la tasa de progresión de la enfermedad. Por ejemplo, en pacientes con HD, las variantes de un solo nucleótido (SNV) en el gen MSH3 que reducen su expresión están relacionadas con una edad tardía de aparición de la enfermedad. Se especula que MSH3 es un buen objetivo para una terapia HD y la evidencia sugiere que el silenciamiento de MSH3 conduce a la inhibición de la expansión de repeticiones CAG somáticas tanto in vitro como in vivo. En este trabajo, los presentes inventores aplicaron varios enfoques para reducir los niveles de MSH3.

Los presentes inventores exploraron el uso del trans-empalme de una manera novedosa para reducir los niveles de expresión de un gen objetivo, mediante el trans-empalme en el pre-ARNm objetivo para producir un ARN que se degrada rápidamente o que no puede producir una proteína funcional. Los presentes inventores imaginaron que un editor de exones 5' o un editor de exones 3' pueden lograr este resultado. Un editor de exones 5' puede reemplazar uno o más de los exones 5' del ARNm objetivo para eliminar el codón de inicio y parte o toda la secuencia codificante, reemplazándola con una secuencia codificante alternativa o una secuencia no codificante. Un editor de exones 3' puede reemplazar uno o más de los exones 3' del objetivo, reemplazándolos con una secuencia codificante alternativa o una secuencia no codificante. En el ARNm quimérico resultante, puede que no haya traducción de ningún aminoácido del ARNm objetivo. Los editores de exones 5' también pueden insertar regiones no traducidas (UTR) 5' que impiden la traducción. Los editores de exones 3' pueden insertar UTR 3' que desestabilizan la transcripción o impiden que se exporte desde el núcleo o se traduzca. Se supone que la porción del ARNm objetivo que queda después de la reacción de trans-empalme se degrada rápidamente, ya que carecerá de una tapa 5' o de una cola de poliA 3'.

Los presentes inventores razonaron que los editores de exones que editan un objetivo, como *HTT*, pueden al mismo tiempo inhibir un segundo gen, como *MSH3*, a través de la inclusión de dos dominios de unión diferentes, uno dirigido a *HTT* y el otro a *MSH3*. Una modalidad de cómo se puede lograr este doble efecto es que un editor de exones *HTT* 5' que contiene una porción 5' de *HTT* puede trans-empalmarse a *HTT* para reconstituir *HTT* de tipo salvaje funcional y al mismo tiempo trans-empalmarse a *MSH3* para producir un ARNm *MSH3* no funcional. En una modalidad de la cual, el ARNm quimérico no funcional *HTT-MSH3* codifica un polipéptido *HTT* que termina en el primer codón de terminación en marco en la porción *MSH3* del ARNm. De este modo, un pre-ARNm *MSH3* normal y de tipo salvaje se transforma en un ARNm *HTT-MSH3* quimérico no funcional que codifica pocos o ningún aminoácido de la secuencia codificante *MSH3* y es probable que sufra una desintegración mediada por sin sentido. Como se muestra aquí, este enfoque funciona con dominios de unión dirigidos a varios intrones *MSH3* diferentes y, de hecho, funcionaría con dominios de unión dirigidos a cualquier intrón *MSH3*. Además, aunque los ejemplos presentados aquí se relacionan con un editor de exones 5', se podría implementar un enfoque similar con un editor de exones 3' con una expectativa razonable de éxito.

En consecuencia, los presentes inventores han diseñado y probado estrategias para tratar la HD en una etapa más temprana de la patogénesis mediante la inhibición de la expansión somática de CAG. Las modalidades descritas en este documento, como aquellas descritas para inhibir la expansión somática de CAG, se pueden utilizar solas o junto con el trans-empalme mediado por el editor de exones de ARN para corregir mutaciones patógenas. Aunque la inhibición de la expansión somática por sí sola es un enfoque terapéutico viable

para la HD, los presentes inventores han probado enfoques terapéuticos híbridos y moléculas de acción dual/híbridas diseñadas para inhibir la expansión somática de CAG y corregir el ARN del exón 1 de *HTT*. Ver, por ejemplo, las Figuras 19 y 20. Por consiguiente, las modalidades descritas en este documento, como aquellas diseñadas para inhibir la expansión somática de CAG, se pueden utilizar solas o junto con el trans-empalme mediado por el editor de exones de ARN para corregir mutaciones patógenas.

Además de lo anterior, para inhibir un mecanismo subyacente de la patogénesis de HD (es decir, la expansión somática de CAG) y al mismo tiempo reemplazar el exón 1 de *HTT* mutante, los presentes inventores diseñaron un conjunto de editores de exones de ARN de acción dual/híbridos que contienen secuencias de orientación para pre-ARNm tanto de *HTT* como de *MSH3*. En algunas modalidades, la expresión del editor de exones está impulsada por un promotor CMV. En algunas modalidades, un editor de exón ejemplar comprende el UTR 5' de *HTT*, la etiqueta FLAG 3X de extremo N, una secuencia codificante del exón 1 de *HTT*, un dominio de empalme (sitio donante de empalme), un enlazador, un dominio de unión a *MSH3*, un dominio de unión a *HTT* y una secuencia terminadora. El dominio de unión de *HTT* dirige el editor de exones al pre-ARNm de *HTT* y conduce a la producción de ARN de *HTT* corregido después de un trans-empalme exitoso. El dominio de unión de *MSH3* dirige el editor de exones al pre-ARNm de *MSH3*, lo que conduce a la producción de una molécula de ARN quimérica del exón 1 de *HTT* - *MSH3* con un codón de parada prematuro (debido a un cambio de marco resultante del reemplazo de los exones *MSH3* en la parte anterior con el exón 1 de *HTT*), que está sujeto a una desintegración mediada por sin sentido (NMD) y, por lo tanto, conduce a la posterior reducción de la expresión de *MSH3*. Ver, por ejemplo, la Figura 20.

Mediante el uso de editores de exones que contienen dominios de unión en tándem, los experimentos de los presentes inventores mostraron un trans-empalme exitoso en pre-ARNm tanto de *HTT* como de *MSH3*, medido mediante RT-qPCR diseñada para detectar moléculas de ARN de *HTT* " corregido" quimérico y ARN de exón 1 de *HTT* quimérico-*MSH3*, respectivamente. Ver, por ejemplo, las Figuras 21 y 22. Las secuencias de los dominios de unión de *MSH3* utilizados en las construcciones probadas se establecen en SEQ ID NO: 140 y 142 (Figura 21) y en SEQ ID NO: 144 y 146 (Figura 22), las secuencias de los dominios de unión de *HTT* utilizados en las construcciones probadas se establecen en SEQ ID NO: 20, 141 y 145 (FIGs. 21, 22), y la secuencia del CDS de *HTT* en las construcciones probadas se establece en SEQ ID NO: 3. Cabe destacar que los niveles de ARN quimérico del exón 1 de *HTT* - *MSH3* podrían subestimarse debido a NMD. Para abordar esto, los presentes inventores también midieron los niveles de transcripción de ARN de *MSH3*.

En otra modalidad, los presentes inventores también diseñaron y probaron editores de exones de dominio de unión en tándem que se dirigen al intrón 2 de *HTT* (utilizando el dominio de unión *HTT_intron2_12061_150* (SEQ ID NO: 72)) junto con el pre-ARNm *MSH3*. Ver, por ejemplo, las Figuras 30-32. Las secuencias de los dominios de unión de *MSH3* utilizados en

las construcciones probadas se establecen en SEQ ID NO: 140, 209, 144 y 210 y la secuencia del CDS de *HTT* en las construcciones probadas se establece en SEQ ID NO: 59.

- El análisis RT-qPCR de los editores de exones del dominio de unión en tándem dirigidos al intrón 2 de *HTT* y al intrón 5 de *MSH3* mostró un trans-empalme exitoso tanto en *HTT* como en *MSH3*, medido mediante RT-qPCR diseñada para detectar moléculas de ARN de *HTT* corregido quimérico y ARN quimérico del exón 2 de *HTT* al exón 6 de *MSH3*. Ver, por ejemplo, las Figuras 31A y 31B. *MSH3_intron5_188_150* (SEQ ID NO: 209) tuvo una mayor eficiencia de trans-empalme de *MSH3* en comparación con *MSH3_intron5_213_100* (SEQ ID NO: 140). De manera similar, los editores de exones del dominio de unión en tándem dirigidos al intrón 2 de *HTT* y al intrón 15 de *MSH3* mostraron un trans-empalme exitoso tanto en *HTT* como en *MSH3*. Ver, por ejemplo, las Figuras 32A y 32B. Cabe destacar que los niveles de ARN del exón 2- *MSH3* del *HTT* quimérico podrían estar subestimados debido a la NMD.
- Los resultados presentados aquí demuestran que los editores de exones del dominio de unión en tándem son eficaces y es probable que tengan beneficios terapéuticos. Además, estos experimentos identificaron dominios de unión eficientes para la orientación del trans-empalme del pre-ARNm *MSH3*. Más allá de su uso en un contexto de dominio de unión en tándem como se describe aquí, estos dominios de unión dirigidos a *MSH3* también se pueden usar en el contexto de un editor de exones con una secuenciación codificante inductora de NMD; dichos editores de exones se pueden usar solos para apuntar a el silenciamiento de *MSH3* o empaquetarse en el mismo vector AAV con un editor de exones dirigido a *HTT* y co-administrarse al tejido objetivo para un enfoque combinatorio.
- En algunas modalidades, se puede utilizar un editor de exones *MSH3* para inhibir la expresión de *MSH3*, en donde el editor de exones *MSH3* tiene un dominio de unión dirigido a un intrón del pre-ARNm de *MSH3*, como se describió anteriormente, pero no incluye un dominio de unión dirigido al pre-ARNm de *HTT* o cualquier otro pre-ARNm. Los editores de exones de ARN con dominio de unión dual descritos anteriormente se pueden modificar eliminando un dominio de unión que se dirige a un intrón *HTT*. En algunas modalidades, los editores de exones *MSH3* pueden administrarse sin administrar también un tratamiento que se dirija a otra secuencia genética o la corrija. En algunas modalidades, dichos editores de exones *MSH3* pueden usarse en un método para tratar o prevenir un trastorno de expansión de repeticiones de trinucleótidos.
- Silenciamiento de *MSH3* mediante ARNi**
- La interferencia de ARN (ARNi) es un mecanismo natural mediante el cual el ARN de doble cadena (ARNdc) induce el silenciamiento genético de una manera específica de secuencia al apuntar al ARNm para su degradación. En un diseño inicial para identificar construcciones de ARNi vectorizadas que pudieran silenciar eficazmente *MSH3*, se diseñaron varias construcciones basadas en una secuencia de miARN (dirigida a *MSH3* Ex23, TTAATCCATAACTCCTTGC; SEQ ID NO: 224). Se construyeron ARNhc impulsados por el

promotor Pol III como posibles controles positivos. Estos ARNhc impulsados por el promotor U6 se probaron junto con construcciones que expresan transcripciones similares a miARN primarios (pri-miARN) impulsados por el promotor Pol II (promotor CMV) dirigidos a *MSH3*. Ver, por ejemplo, la Figura 33.

5 Tal como se muestra, por ejemplo, en la Figura 34, el análisis de transferencia Western reveló que para este miARN dirigido a Ex23 de *MSH3* (TTAATCCATAACTCCTTGC; SEQ ID NO: 224), se observó una reducción de alrededor del 50 % de la proteína MSH3 cuando se expresó en un contexto de andamiaje de miR-30a. Además, la eficiencia del silenciamiento no mejoró al colocar el ARN guía en el brazo 5' o al agregar una protuberancia a la estructura del tallo (37% y 23% de silenciamiento, respectivamente).

15 Los diseños adicionales centrados en variar la secuencia activa de miARN mostraron que además de la secuencia dirigida al exón 23 probada en la Figura 34, las secuencias de miARN dirigidas al exón 22 (SEQ ID NO: 246), dirigidas al exón 9 (SEQ ID NO: 248), dirigidas al exón 12 (SEQ ID NO: 256) y dirigidas al exón 15 (SEQ ID NO: 257) funcionaron igualmente bien, observándose un silenciamiento de la proteína MSH3 de alrededor del 50%. Ver, por ejemplo, la Figura 35. Esta modalidad de ARNi MSH3 por sí sola puede ser terapéuticamente relevante para muchos trastornos de expansión de repeticiones y/o podría combinarse con un editor de exones para un mayor potencial terapéutico.

25 En algunas modalidades, las construcciones de ARNi dirigidas a MSH3 divulgadas en este documento se pueden usar junto con editores de exones de ARN que se dirigen a HTT, como se describió anteriormente, o que se dirigen a otros genes asociados con enfermedades de expansión de repeticiones de trinucleótidos. En otras modalidades, las construcciones de ARNi dirigidas a MSH3 divulgadas en este documento se pueden utilizar independientemente de los enfoques terapéuticos dirigidos a HTT. En algunas modalidades, las construcciones de ARNi dirigidas a MSH3 divulgadas en este documento se pueden administrar para tratar o prevenir una enfermedad de expansión de repeticiones de trinucleótidos.

30 Modulación de empalme de MSH3

MSH3 se puede inactivar incorporando secuencias que sean complementarias a las uniones de empalme de MSH3 en secuencias de ARNnp, como el ARNnp U7. Los presentes inventores diseñaron y probaron estrategias para bloquear eventos de cis-empalme mediante ARN antisentido expresado a partir del mismo plásmido que el editor de exones de ARN dirigido a HTT. Los ARNnp modificados como el U7 SmOPT se diseñaron 1) cambiando la secuencia objetivo (por ejemplo, la secuencia de unión de histonas en la región 5' del U7 ARNnp) a la secuencia complementaria del gen que se va a modificar, y 2) cambiando el sitio de unión (AAUUUGUCUAG; SEQ ID NO: 367; U7 Sm WT) para las proteínas específicas de U7 snRNP (Lsm10 y Lsm11) a la secuencia de consenso derivada de las principales ribonucleoproteínas nucleares pequeñas ricas en uridina (U snRNP) espliceosómicas (AAUUUUUGGAG; SEQ ID

NO: 368; U7 Sm OPT), lo que conduce a la formación de un núcleo proteico heptamérico de tipo espliceosómico envuelto alrededor de U7 Sm OPT. En algunas modalidades, las moléculas de ARNas descritas en este documento consisten en el promotor U1, ARNnp (con una secuencia de ARNas dirigida a los límites intrón-exón de *MSH3* y al sitio de unión de consenso Sm) y un terminador U1. Ver, por ejemplo, la Figura 36.

Los presentes inventores diseñaron y probaron moduladores de empalme de *MSH3* dirigidos a la omisión de los exones 2, 3 y 4 de *MSH3*. Para cada conjunto, se probaron secuencias de ARNas contra el límite intrón-exón en la parte anterior (por ejemplo, límite intrón 1/exón 2 para el modulador de empalme del exón 2 (SEQ ID NO: 274)), límite exón-intrón posterior (por ejemplo, límite exón 2/intrón 2 para el modulador de empalme del exón 2 (SEQ ID NO: 275)), y secuencias en tándem (SEQ ID NO: 276 y 277) que consistían en los dos ARNas. Las secuencias de moduladores de empalme *MSH3* ejemplares que median la omisión del exón 2 se muestran en las SEQ ID NO: 284-286, que comprenden elementos moduladores de empalme *MSH3* SEQ ID NO: 274-276, respectivamente. Para las moléculas de omisión del exón 2 de *MSH3*, también se probó una molécula de ARNas dirigida a los límites intrón 1/exón 2 y exón 2/intrón 2 sin el andamiaje de ARNnp U7SmOPT. Los ensayos de RT-qPCR que miden los niveles de productos de ARNm *MSH3* empalmados mostraron que los moduladores de empalme de U7SmOPT (pero no la secuencia de ARNas que no contiene el andamiaje U7SmOPT) exhibieron una omisión intencionada de los exones objetivo. Las secuencias de moduladores de empalme *MSH3* ejemplares que median la omisión del exón 3 se muestran en las SEQ ID NO: 288-290, que comprenden elementos moduladores de empalme *MSH3* SEQ ID NO: 278-280, respectivamente. Ver, por ejemplo, las Figuras 37-39. Cabe destacar que la omisión del exón 4 de *MSH3* no provocará NMD, sino que provocará un producto *MSH3* más corto debido a que el exón 4 contiene 213 nucleótidos, lo que dará como resultado un producto de ARNm con el exón 4 omitido dentro del marco. Las secuencias de moduladores de empalme *MSH3* ejemplares que median la omisión del exón 4 se muestran en las SEQ ID NO: 291-293, que comprenden elementos moduladores de empalme *MSH3* SEQ ID NO: 281-283, respectivamente. Los presentes inventores observaron productos proteicos truncados para los moduladores de empalme dirigidos al exón 4 en análisis de transferencia Western evaluados con anticuerpos para visualizar *MSH3*, lo que proporcionó evidencia de que los moduladores de empalme estaban omitiendo los exones objetivo como estaba previsto.

Las rondas posteriores de diseños probaron moduladores de empalme *MSH3* dirigidos a la omisión del exón 3 (las secuencias de moduladores de empalme *MSH3* ejemplares se muestran en SEQ ID NO: 324-326, que comprenden los elementos moduladores de empalme *MSH3* SEQ ID NO: 299, SEQ ID NO: 300 + 298 + 301 y SEQ ID NO: 302 + 298 + 303, respectivamente), el exón 6 (las secuencias de moduladores de empalme *MSH3* ejemplares se muestran en SEQ ID NO: 327-329, que comprenden los elementos moduladores de empalme *MSH3* SEQ ID NO: 304, SEQ ID NO: 305 + 298 + 306 y SEQ ID NO: 307 + 298 + 308, respectivamente), el exón 7 (secuencias de moduladores de empalme *MSH3* ejemplares

se muestran en las SEQ ID NO: 330-332, que comprenden los elementos moduladores de empalme de MSH3 SEQ ID NO: 309, SEQ ID NO: 310 + 298 + 311 y SEQ ID NO: 312 + 298 + 313, respectivamente), el exón 8 (las secuencias de moduladores de empalme de MSH3 de ejemplo se muestran en las SEQ ID NO: 333-335, que comprenden los elementos moduladores de empalme de MSH3 SEQ ID NO: 314, SEQ ID NO: 315 + 298 + 316 y SEQ ID NO: 317 + 298 + 318, respectivamente), y el exón 15 (las secuencias de moduladores de empalme de MSH3 de ejemplo se muestran en las SEQ ID NO: 336-338, que comprenden los elementos moduladores de empalme de MSH3 SEQ ID NO: 319, SEQ ID NO: 320 + 298 + 321 y SEQ ID NO: 322 + 298 + 323, respectivamente). Estos exones fueron seleccionados por su idoneidad para la modulación de empalme, teniendo en cuenta factores como la longitud del neoepítipo si se produce una proteína en lugar de transcripciones que experimentan NMD, características de empalme como empalme mediado por espliceosoma mayor o menor, empalme alternativo, etc.

Los presentes inventores diseñaron y probaron moduladores de empalme MSH3 dirigidos a la omisión de los exones 3, 6, 7, 8 y 15 de *MSH3*. Para cada conjunto, se probaron 1) una molécula de U7SmOPT que contiene una secuencia antisentido a lo largo de toda la longitud del exón objetivo y 12-13 nucleótidos en los intrones flanqueantes, 2) una molécula de U7SmOPT que contiene una secuencia antisentido al límite exón/intrón corriente abajo y al límite intrón/exón en la parte anterior, separados por un enlazador no estructurado, y 3) una molécula de ARNnp U2 que contiene una secuencia antisentido al límite exón/intrón corriente abajo y al límite intrón/exón en la parte anterior, separados por un enlazador no estructurado. Las secuencias de los moduladores de empalme *MSH3* probados se establecen en las SEQ ID NO: 290 y 324-326 (Figura 40); SEQ ID NO: 327-329 (Figura 41); SEQ ID NO: 330-332 (Figura 42); SEQ ID NO: 333-335 (Figura 43); y SEQ ID NO: 336-338 (Figura 44). SEQ ID NO: indicando la región ARNas correspondiente SEQ ID NO: y la región objetivo se indican en cada una de las respectivas Figura 40-44. Los presentes inventores identificaron moduladores de empalme eficaces para todos los exones objetivo probados. En general, la molécula U7SmOPT que contiene una secuencia antisentido al límite exón/intrón corriente abajo y al límite intrón/exón en la parte anterior, separados por un enlazador no estructurado (por ejemplo, SEQ ID NO: 328, 331, 334 y 337), mostró un grado particularmente alto de reducción en los niveles de expresión de MSH3, con algunas excepciones. Ver, por ejemplo, las Figuras 40-44. Los análisis de transferencia Western evaluados con anticuerpos para visualizar MSH3 indicaron que los niveles de silenciamiento de la proteína MSH3 se corresponden con los perfiles de eficiencia de omisión de exones del ARN. Las secuencias de los moduladores de empalme *MSH3* probados se establecen en las SEQ ID NO: 330-332. Ver, por ejemplo, la Figura 45.

En algunas modalidades, los moduladores de empalme a MSH3 divulgados en este documento se pueden usar junto con editores de exones de ARN que se dirigen a HTT, como se describió anteriormente, o que se dirigen a otros genes asociados con enfermedades de

expansión de repeticiones de trinucleótidos. En algunas modalidades, los moduladores de empalme de MSH3 divulgados en este documento se pueden utilizar independientemente de los enfoques terapéuticos dirigidos a HTT. En algunas modalidades, los moduladores de empalme de MSH3 divulgados en este documento se pueden administrar para tratar o prevenir una enfermedad de expansión de repeticiones de trinucleótidos.

Silenciamiento de especies tóxicas de HTT

Para mejorar la eficacia de los editores de exones HTT descritos en este documento, los presentes inventores imaginaron enfoques combinatorios que abordan además especies tóxicas como HTT1a. Para este fin, se diseñaron y probaron moléculas híbridas vectorizadas que comprenden el editor de exones HTT, así como un miARN que se dirige al exón 1 del *HTT* sin editar, elementos que se pretendía empaquetar en un único AAV. Ver, por ejemplo, la Figura 46. Las secuencias del miARN (miARN-1 HTT y miARN-2 HTT) dirigidas a HTT sin editar se establecen en las SEQ ID NO: 339 y 342. El CDS del editor de exones HTT se modificó en codones para evitar la interacción entre el miARN y el editor de exones, evitando así la degradación del propio editor de exones, así como de los productos HTT corregidos. Ver, por ejemplo, la Figura 47.

Los presentes inventores evaluaron primero la eficacia del editor de exones solo, miARN HTT solo y las moléculas del editor de exones + miARN HTT cuando se combinan las dos modalidades. En este experimento se probaron un editor de exones dirigido al intrón 2 de HTT (SEQ ID NO: 204) y un miARN-1 de HTT (secuencia activa SEQ ID NO: 339). SEQ ID NO: 341 corresponde a la secuencia pri-miARN codificante del miARN-1 HTT (SEQ ID: 339) en el andamio mir-33 (SEQ ID: 259, 260, 261). Como se muestra, por ejemplo, en la Figura 48, la proporción de transcripciones de *HTT* transempalmadas (editadas) entre todas las transcripciones de *HTT* aumentó en presencia de miARN de HTT debido a la reducción del número total de copias de la transcripción de *HTT*. Sin embargo, el número de copias de las transcripciones editadas y transempalmadas de *HTT* no varió con o sin miARN-1 HTT, lo que indica que hubo una interacción mínima o nula entre miARN-1 HTT y el editor de exones.

Se realizó una comparación adicional de dos miARN independientes dirigidos al exón 1 de HTT, miARN-1 HTT (secuencia activa SEQ ID NO: 339; secuencia codificante de pri-miARN SEQ ID NO: 341) y miARN-2 HTT (secuencia activa SEQ ID NO: 342). SEQ ID NO: 344 corresponde a la secuencia pri-miARN codificante del miARN-2 HTT (SEQ ID: 342) en el andamio mir-33 (SEQ ID: 259, 260, 261). La secuencia codificante del editor de exones que apunta al intrón 2 de HTT se optimizó aún más en cuanto a codones para conferir y garantizar resistencia a ambos miARN (las secuencias de dominio codificante variantes de ejemplo se establecen en las SEQ ID NO: 59 y 349-351). No se observó ningún impacto negativo del miARN-1 de HTT ni del miARN-2 de HTT en el trans-empalme, como lo demuestra el número de copias de *HTT* trans-empalmadas sin cambios (o potencialmente aumentadas) y un mayor porcentaje de ARNm de *HTT* trans-empalmado por el editor de exones en presencia de

miARN. Ver, por ejemplo, la Figura 49. Los presentes inventores demostraron eficiencias similares de silenciamiento de HTT con miARN-1 HTT y miARN-2 HTT, sin que ninguno de los miARN afectara negativamente el rendimiento del editor de exones, como lo evidencian los números de copias sin cambios de las transcripciones de *HTT* transempalmadas con o sin cada miARN. Ver, por ejemplo, la Figura 49.

A continuación, los presentes inventores evaluaron las moléculas duales editores de exones HTT + miARN a nivel de proteína. Las pruebas de la molécula híbrida dual del editor de exones dirigido al intrón 2 de HTT + miARN de HTT indicaron que en presencia del miARN de HTT (miARN-1 HTT; SEQ ID NO: 339), se logró una reducción de ~50 % de las transcripciones de *HTT* sin editar (Figura 50, panel izquierdo). El análisis de la actividad de trans-empalme del editor de exones mostró que el porcentaje de ARN *HTT* editado aumentó en presencia de miARN HTT debido a la reducción de los números de copias de *HTT* no editadas y, por lo tanto, totales por el miARN (Figura 50, panel derecho).

Construcción de moléculas híbridas

Como se describe en este documento, los presentes inventores exploraron un enfoque híbrido en donde se apuntaron a múltiples mecanismos y/o especies patógenas. Como se detalla en este documento, estos enfoques pueden incluir el reemplazo de HTT mutante a través de un editor de exones, el silenciamiento de HTT mutante y transcripciones relacionadas (por ejemplo, HTT1a) a través de ARNi y la reducción de MSH3 a través de: trans-empalme o modulación de empalme vectorizado. Estos enfoques, cuando se combinan, representan un mecanismo de acción multimodal que, no obstante, puede administrarse en una única AAV. Un subconjunto de los diseños destinados a silenciar a MSH3 son por diseño una molécula híbrida (por ejemplo, el editor de exones de dominio de unión en tándem). De manera similar, los presentes inventores diseñaron el miARN dirigido a HTT con la idea de que se combinaría con el Editor de exones HTT; por lo tanto, se implementó la consideración adicional de modificar el codón del sitio objetivo del miARN en el Editor de exones para minimizar la interacción entre el miARN dirigido a HTT y el Editor de exones. Estos enfoques combinatorios se pueden utilizar para lograr un alto grado de eficacia terapéutica en algunos pacientes con HD, si no en todos. De hecho, estos enfoques pueden tener efectos drásticos en la modificación de la enfermedad.

Editor de exones HTT + Modulador de empalme MSH3

Aunque la inhibición de la expansión somática por sí sola puede tener un efecto terapéutico positivo en la HD, una molécula híbrida que inhibe la expansión somática de CAG y corrige el ARN del exón 1 de *HTT* podría tener potencialmente una promesa terapéutica significativa debido a la capacidad de este enfoque para combatir la enfermedad simultáneamente en múltiples frentes. Cualquiera de los editores de exones de ARN dirigidos a HTT descritos en este documento se puede utilizar en combinación con uno o más de los moduladores de empalme MSH3 descritos en este documento. Cualquiera de los editores de exones de ARN

dirigidos a HTT descritos en este documento se puede utilizar sin utilizar también ninguno de los moduladores MSH3 descritos en este documento, y cualquiera de los moduladores de empalme MSH3 descritos en este documento también se puede utilizar sin utilizar también ninguno de los editores de exones de ARN dirigidos a HTT.

5 Para lograr este efecto combinatorio, se diseñaron vectores híbridos que codifican el modulador de empalme MSH3 + el editor de exones HTT. Ver, por ejemplo, la Figura 51. Los exámenes iniciales de estas moléculas híbridas duales utilizaron un editor de exones dirigido al intrón 2 de *HTT* (HTT_intron2_12061_150) (SEQ ID NO: 204) y un modulador de empalme de omisión del exón 7 *MSH3* ejemplar (SEQ ID NO: 331). Sin embargo, el efecto combinatorio sería igualmente aplicable a cualquier combinación de editor de exones HTT y modulador de empalme MSH3 (las secuencias de moduladores de empalme MSH3 ejemplares se muestran, por ejemplo, en SEQ ID NO: 330-332, que comprenden elementos moduladores de empalme MSH3 SEQ ID NO: 309, SEQ ID NO: 310 + 298 + 311 y SEQ ID NO: 312 + 298 + 313, 10 respectivamente y SEQ ID NO: 284-293, 324-329 y 333-338. Se entiende además que el efecto combinatorio observado con el editor de exones de orientación del intrón 2 de HTT (HTT_intron2_12061_150) ejemplar (SEQ ID NO: 204) descrito aquí sería aplicable de manera similar a otros editores de exones de orientación del intrón 2 de HTT (HTT_intron2_12061_150), incluidos aquellos que comprenden cualquiera de las SEQ ID NO: 20 59, 350 o 351 correspondientes a la secuencia del dominio de codificación modificada. En consecuencia, un editor de exones dirigido al intrón 2 de HTT podría comprender cualquiera de las secuencias de dominio de codificación variante establecidas en cualquiera de las SEQ ID NO: 59 y 349-351). Se entiende además que el efecto combinatorio observado con el Editor de Exones (SEQ ID NO: 204) dirigido al intrón 2 de HTT (HTT_intron2_12061_150) descrito aquí sería igualmente aplicable a los Editores de Exones dirigidos al intrón 3 de HTT (que comprenden, por ejemplo, HTT_intron3_4223_150 o HTT_intron3_4233_150) o a los Editores de Exones dirigidos al intrón 1 de HTT (que comprenden, por ejemplo, HTT_intron1_11704_100 o HTT_intron1_11724_100). Como se muestra aquí, los presentes inventores examinaron el Editor de Exones solo, el Modulador de Empalme MSH3 solo y la 25 molécula híbrida dual Editor de Exones + Modulador de Empalme MSH3 para ver sus perfiles de trans-empalme, así como las eficiencias de silenciamiento de MSH3. El análisis RT-qPCR no mostró cambios en el rendimiento del editor de exones con o sin el modulador de empalme MSH3, como lo indican los niveles similares de % de reemplazo de HTT. Ver, por ejemplo, la Figura 52A. De manera similar, el rendimiento del modulador de empalme MSH3 no cambia con o sin el editor de exones. Ver, por ejemplo, la Figura 52B. Este experimento demostró que no hay interacción entre el editor de exones y el modulador de empalme MSH3 cuando se administran conjuntamente en el mismo fragmento de ADN.

30 Para un trans-empalme eficiente se requieren niveles suficientes de administración de AAV al tejido objetivo y altos niveles de expresión del editor de exones. Para lograr altos niveles de expresión del vector, se probó el AAV autocomplementario (scAAV). Para comparar el

rendimiento del AAV de cadena simple convencional (ssAAV) con el scAAV, se insertó una molécula híbrida dual ejemplar de editor de exones dirigido al intrón 2 *HTT* + modulador de empalme de omisión del exón 7 *MSH3* en un plásmido compatible con ssAAV (que contiene ITR de tipo salvaje) y un plásmido compatible con scAAV (ITR/ITR- Δ trs), se empaquetó en AAV2 y se usó para la transducción de células HEK293. La caracterización de la expresión de *MSH3*, tanto a nivel de ARN como de proteína, sugirió que las moléculas híbridas duales tenían un mejor desempeño para ambas modalidades cuando se empaquetaban en scAAV (ver, por ejemplo, Figura 53, panel izquierdo). La caracterización del trans-empalme también respaldó el mayor rendimiento de las moléculas empaquetadas en scAAV en comparación con ssAAV, como lo demuestra el mayor % de reemplazo de la transcripción *HTT* y la correspondiente proteína *HTT* trans-empalmada de longitud completa elevada detectada por análisis de transferencia Western (ver, por ejemplo, Figura 53, panel derecho). Los presentes inventores también analizaron el nivel de expresión de los editores de exones y observaron una mayor expresión de transgenes en las células transducidas con scAAV. Una consideración al usar vectores scAAV es el espacio de carga limitado (aproximadamente la mitad del tamaño en comparación con ssAAV), lo que limita su uso cuando se impulsa la expresión del editor de exones desde un promotor más grande (por ejemplo, el promotor CAGGS).

Editor de exones *HTT* + miARN *HTT* + Modulador de empalme *MSH3*

Los presentes inventores también diseñaron, construyeron y probaron moléculas híbridas triples que contienen múltiples modalidades que se dirigen a múltiples mecanismos y/o especies patógenas. Estos enfoques incluyen el reemplazo de *HTT* mutante a través de un editor de exones, el silenciamiento de *HTT* mutante y transcripciones relacionadas (por ejemplo, *HTT1a*) a través de ARNi y la reducción de *MSH3* a través de modulación de empalme. Estos enfoques, cuando se combinan, representan un mecanismo de acción multimodal que, no obstante, puede administrarse en una única AAV. Estos enfoques combinatorios se pueden utilizar para lograr un alto grado de eficacia terapéutica en algunos pacientes con HD, si no en todos. De hecho, estos enfoques pueden tener efectos drásticos en la modificación de la enfermedad. Como se describió anteriormente, cada una de estas modalidades está diseñada para minimizar la interacción entre sí (por ejemplo, la modificación del codón del editor de exones que tiene resistencia al miARN *HTT*).

Para explorar este enfoque, se realizó un análisis de moléculas híbridas triples que contenían el editor de exones dirigido al intrón 2 de *HTT* impulsado por el promotor CAGGS (SEQ ID NO: 196) (SEQ ID NO: 204), un modulador de empalme de omisión del exón 7 de *MSH3* ejemplar (SEQ ID NO: 331) y SEQ ID NO: 341 [miARN-1 *HTT* (SEQ ID: 339) que codifica la secuencia de pri-miARN en el andamio mir-33 (SEQ ID: 259, 260, 261)] o SEQ ID NO: 344 [miARN-2 *HTT* (SEQ ID: 342) que codifica la secuencia de pri-miARN en el andamio mir-33 (SEQ ID: 259, 260, 261)]. Se comparó la actividad del triple híbrido con una molécula híbrida de control que contenía un editor de exones mutante de empalme, un modulador de empalme que contenía una secuencia de control desordenada y un miARN que contenía una secuencia de

control desordenada. El efecto combinatorio sería igualmente aplicable a cualquier combinación de editor de exones HTT, modulador de empalme MSH3 (las secuencias de moduladores de empalme MSH3 ejemplares se muestran, por ejemplo, en SEQ ID NO: 330-332, que comprenden elementos moduladores de empalme MSH3 SEQ ID NO: 309, SEQ ID NO: 310 + 298 + 311 y SEQ ID NO: 312 + 298 + 313, respectivamente y SEQ ID NO: 284-293, 324-329 y 333-338) y miARN-1 HTT o miARN-2 HTT. Se entiende además que el efecto combinatorio observado con el editor de exones de orientación del intrón 2 de HTT (HTT_intron2_12061_150) ejemplar (SEQ ID NO: 204) descrito aquí sería aplicable de manera similar a otros editores de exones de orientación del intrón 2 de HTT (HTT_intron2_12061_150), incluidos aquellos que comprenden cualquiera de las SEQ ID NO: 59, 350 o 351 correspondientes a la secuencia del dominio de codificación modificada. En consecuencia, un editor de exones dirigido al intrón 2 de HTT podría comprender cualquiera de las secuencias de dominio de codificación variante establecidas en cualquiera de las SEQ ID NO: 59 y 349-351). Se entiende además que el efecto combinatorio observado con el editor de exones de orientación del intrón 2 de HTT (HTT_intron2_12061_150) ejemplar (SEQ ID NO: 204) descrito aquí sería aplicable de manera similar a los editores de exón dirigidos al intrón 3 de HTT (que comprenden, por ejemplo, HTT_intron3_4223_150 o HTT_intron3_4233_150) o a los editores de exón dirigidos al intrón 1 de HTT (que comprenden, por ejemplo, HTT_intron1_11704_100 o HTT_intron1_11724_100).

La eficiencia del trans-empalme para el editor de exones HTT fue similar con o sin el modulador de empalme MSH3 (ver, por ejemplo, Figura 54; panel superior), como se vio anteriormente en la Figura 52. En presencia de HTT-miARN 1 o HTT-miARN 2, el % de transcripciones de *HTT* editadas aumentó debido a la reducción de las transcripciones de HTT no editadas (y, por lo tanto, de las transcripciones de *HTT* totales) por el miARN de HTT (ver, por ejemplo, Figura 54; Panel superior). El rendimiento de los miARN HTT se evaluó en la Figura 54; panel central, y se observaron eficiencias similares con miARN-1 HTT y miARN-2 HTT, ya sea en un contexto híbrido dual (Editor de exones + miARN HTT) o híbrido triple (Editor de exones + miARN HTT + Modulador de empalme MSH3). La evaluación de la eficiencia de silenciamiento de MSH3 respaldó la inalteración de la eficiencia de silenciamiento por parte del modulador de empalme, ya sea en presencia del editor de exones o del miARN HTT. Ver, por ejemplo, la Figura 54; Panel inferior. Estos resultados demostraron que no hay interacción entre el editor de exones, el miARN HTT y el modulador de empalme de MSH3 cuando se administran conjuntamente en el mismo fragmento de ADN. Cabe destacar que el efecto combinatorio sería igualmente aplicable a cualquier combinación de editor de exones HTT, miARN HTT y modulador de empalme MSH3.

Prueba del mecanismo en las iCell GlutaNeurons

Para confirmar y corroborar la eficacia de las moléculas mencionadas anteriormente en sistemas celulares adicionales in vitro, los presentes inventores utilizaron iCell GlutaNeurons (FujiFilm), neuronas corticales humanas enriquecidas con glutamato derivadas de células

madre pluripotentes inducidas (iPS). Las moléculas de ejemplo se empaquetaron en AAV2.7m8 y se utilizaron para la transducción para expresar moléculas de interés en las iCell GlutaNeurons.

5 Los presentes inventores probaron dos AAV2.7m8 diferentes para un experimento de prueba de mecanismo para tres modalidades separadas: 1) Reemplazo de exón HTT por trans-empalme, 2) silenciamiento de HTT por un miARN dirigido al exón 1 de *HTT*, y 3) reducción de MSH3 por modulación de empalme del exón 7 de *MSH3*. Cabe destacar que un AAV probado (SEQ ID NO: 354)) empaquetó una molécula híbrida dual que expresaba tanto el editor de exones dirigido al intrón 2 de *HTT* (SEQ ID NO: 204) como el miARN-1 de HTT (secuencia activa SEQ ID NO: 339; secuencia de pri-miARN codificante SEQ ID NO: 341), mientras que el otro AAV probado (SEQ ID NO: 357) empaquetó una molécula híbrida dual que expresaba el editor de exones dirigido al intrón 2 de HTT (SEQ ID NO: 204) y el empalme de omisión del exón 7 de *MSH3*. Modulador (SEQ ID NO: 331), donde cada uno fue empaquetado en AAV2.7m8.

El análisis RT-ddPCR que mide los niveles de % de ARN *HTT* editado (trans-empalmado) demostró que sin el miARN HTT, la molécula híbrida dual editor de exones HTT + modulador de empalme MSH3 resultó en un reemplazo de ~37% por trans-empalme en iCell GlutaNeurons. En la molécula híbrida dual Editor de exón + miARN HTT, los presentes inventores observaron un % de ARN editado de ~50 % debido en parte al número reducido de copias de *HTT* nativas como resultado del miARN HTT. El análisis de transferencia Western de la etiqueta FLAG de extremo N confirmó la presencia de la proteína HTT de longitud completa trans-empalmada. Ver, por ejemplo, la Figura 55.

Un análisis adicional de los niveles de transcripción *HTT* sin editar sugirió que sin el miARN HTT, la molécula híbrida dual editor de exones HTT + modulador de empalme MSH3 demostró una reducción del 47 % del ARN *HTT* sin editar como consecuencia del trans-empalme (reemplazo de ARN nativo con corregido) en iCell GlutaNeurons. En presencia de la modalidad de miARN HTT, la molécula híbrida dual Editor de exón + miARN HTT silenció el 50% de las transcripciones de *HTT* no editadas por el miARN, lo que dio como resultado una relación de transcripciones de *HTT* editadas:no editadas aún más favorable. El análisis de transferencia Western para visualizar la HTT total (proteínas HTT transempalmadas y sin editar) confirmó el silenciamiento de la *HTT* sin editar con la modalidad de miARN HTT adicional. Ver, por ejemplo, la Figura 56.

Los presentes inventores también probaron la modalidad de modulación de empalme MSH3 en iCell GlutaNeurons. El análisis de transferencia Western de la proteína MSH3 mostró que la molécula híbrida dual Editor de exones + modulador de empalme de exón 7 de MSH3 (SEQ ID NO: 331) resultó en un silenciamiento de >40% de la proteína MSH3 en estas células. Ver, por ejemplo, la Figura 57. En general, los presentes inventores confirmaron una actividad

similar para cada una de las tres modalidades de HD (editor de exones *HTT*, miARN *HTT* y modulador de empalme *MSH3*) en células HEK293 y iCell GlutaNeurons.

Modalidades combinatorias adicionales y detalles relacionados con estas:

- 5 Mejorar el reemplazo de exones por trans-empalme de pre-ARNm *HTT*
Los resultados presentados aquí (por ejemplo, Figura 17) muestran que el bloqueo de eventos de cis-empalme de *HTT* puede conducir a una mejora en la eficiencia del trans-empalme y al reemplazo del exón 1 de *HTT*. Por consiguiente, las modalidades de la presente divulgación incluyen estrategias para bloquear estos eventos de cis-empalme mediante ARN antisentido (ARNas). Para expresar el bloqueador de cis-empalme del mismo plásmido que un editor de exones de ARN dirigido a *HTT*, se pueden utilizar ARNnp modificados, como el ARNnp U7 modificado (U7 Sm OPT), que previamente ha demostrado actuar como modulador de empalme. Como se detalla en la presente, los ARNnp tales como el U7 SmOPT se diseñaron 1) cambiando la secuencia objetivo (por ejemplo, la secuencia de unión de histonas en la 10 región 5' del U7 ARNnp) a la secuencia complementaria del gen que se va a modificar, y 2) cambiando el sitio de unión (AAUUUGUCUAG; SEQ ID NO: 367; U7 Sm WT) para las proteínas específicas de U7 snRNP (Lsm10 y Lsm11) a la secuencia de unión de consenso derivada de las principales U snRNP espliceosómicas (AAUUUUUGGAG; SEQ ID NO: 368; U7 Sm OPT), lo que conduce a la unión de proteínas Sm encontradas en snRNP 15 espliceosómicas a la U7 Sm OPT. En algunas modalidades, las moléculas de ARNas pueden comprender el promotor U1, ARNnp (con una secuencia de ARNas dirigida a los límites intrón-exón de *HTT* y al sitio de unión de consenso Sm) y un terminador U1. Las secuencias de ARNas ejemplares incluyen (SEQ ID NO: 127-135). De estos, una combinación del editor de exones dirigido al intrón 2 de *HTT* (*HTT_intron2_12061_150*) con ASO6 (SEQ ID NO: 131), 20 que bloquea el cis-empalme del intrón en la parte anterior, condujo a una mejora particularmente significativa en las eficiencias del trans-empalme *in vitro*. Ver, por ejemplo, las Figuras 16 y 17.

- Las modalidades dirigidas a la inhibición de la expansión de repeticiones CAG somáticas
30 Sin querer limitarse a la teoría, en el caso de la HD y muchos otros trastornos de expansión de repeticiones, se cree que el mecanismo subyacente de la patogénesis es la expansión somática del tracto de repetición. Como se detalla aquí, la evidencia genética humana apoya firmemente la idea de que el resultado clínico de la HD puede explicarse, al menos en parte, por la inestabilidad somática debido a la expansión somática de los tractos de repetición CAG.
35 Por lo tanto, inhibir el mecanismo subyacente de la patogénesis que surge de la inestabilidad somática debería alterar el curso de la enfermedad y podría ser una terapia potencial para la HD y muchos otros trastornos recurrentes.

- Con el objetivo de abordar este aspecto de la patología de la HD, los presentes inventores han diseñado y probado estrategias para tratar la HD en una etapa más temprana de la patogénesis 40 mediante la inhibición de la expansión somática de CAG. Las modalidades descritas en este

- documento, como aquellas descritas para inhibir la expansión somática de CAG, se pueden utilizar solas o junto con el trans-empalme mediado por el editor de exones de ARN para corregir mutaciones patógenas. Aunque la inhibición de la expansión somática por sí sola puede tener un efecto terapéutico positivo en la HD (ver la sección de modalidades combinatorias anterior), los presentes inventores han probado enfoques terapéuticos híbridos y moléculas de acción dual/híbridas diseñadas para inhibir la expansión somática de CAG y corregir el ARN del exón 1 de *HTT*. Ver, por ejemplo, las Figuras 19 y 20. Los presentes inventores también han demostrado que este enfoque terapéutico híbrido es eficaz utilizando moléculas de acción dual/híbridas diseñadas para inhibir la expansión somática de CAG y corregir el ARN del exón 2 de *HTT*. Ver, por ejemplo, las Figuras 30-32, 52 y 53. Por consiguiente, las modalidades descritas en este documento, como aquellas diseñadas para inhibir la expansión somática de CAG, se pueden utilizar solas o junto con el trans-empalme mediado por el editor de exones de ARN para corregir mutaciones patógenas.
- Los mecanismos que controlan este proceso involucran maquinaria celular, por ejemplo, enzimas de reparación de desajustes de ADN, incluidas *MSH2* y *MSH3*. Se especula que *MSH3* es un buen objetivo para inhibir la expansión de repeticiones CAG somáticas. En este trabajo, los presentes inventores describen varios enfoques actuales diseñados para reducir los niveles de *MSH3* e inhibir la expansión de repeticiones CAG somáticas.
- Silenciamiento de *MSH3* y otros genes mediante trans-empalme: Como se describe en este documento, el trans-empalme se puede utilizar para reducir la expresión de *MSH3* u otros genes, ya sea solo o en combinación con otras aplicaciones de trans-empalme.
- Modalidad A1.1. Trans-empalme para el silenciamiento de *MSH3* y otros genes: Como se describe en este documento, el trans-empalme puede utilizarse para la edición de ARNm, pero también se implementa aquí para reducir los niveles de expresión de un gen objetivo. Como se describe en la presente, se puede lograr reducir los niveles de expresión de un gen objetivo mediante el trans-empalme en el pre-ARNm objetivo de ARNm maduro para producir un ARN que se degrada rápidamente o que no puede producir una proteína funcional. Ver, por ejemplo, las Figuras 20-22 y 30-32. Los editores de exones tanto 5' como 3' pueden lograr este resultado. Un editor de exones 5', por ejemplo, puede reemplazar uno o más de los exones 5' del ARNm objetivo para eliminar el codón de inicio y parte o toda la secuencia codificante, reemplazándola con una secuencia codificante alternativa o una secuencia no codificante. Un editor de exones 3', por ejemplo, puede reemplazar uno o más de los exones 3' del objetivo, reemplazándolos con una secuencia codificante alternativa o una secuencia no codificante. En el ARNm quimérico resultante, puede no haber traducción de ningún aminoácido del ARNm objetivo. Los editores de exón 5' también pueden, por ejemplo, insertar regiones no traducidas (UTR) 5' que impidan la traducción. Los editores de exón 3' también pueden, por ejemplo, insertar UTR 3' que desestabilicen la transcripción o impidan que se exporte desde el núcleo o se traduzca. Se supone que la porción del ARNm objetivo que queda después de la reacción

de trans-empalme se degrada rápidamente, ya que carecerá de una tapa 5' (editor de exones 5') o de una cola de poliA 3' (editor de exones 3').

Modalidad A1.2. Editores de exones de dominio de unión en tándem: Los editores de exones que editan un objetivo, tal como *HTT*, pueden al mismo tiempo inhibir un segundo gen, como *MSH3*, a través de la inclusión de dos dominios de unión diferentes, uno dirigido a *HTT* y el otro a *MSH3*. En las Figuras 19-22 y 30-32 se presentan ejemplos de cómo se puede lograr este doble efecto. Brevemente, un editor de exones *HTT* 5' que contiene una porción 5' de *HTT* puede trans-empalmarse a *HTT* para reconstituir un ARNm de *HTT* de tipo salvaje funcional y al mismo tiempo trans-empalmarse a *MSH3* para producir un ARNm *MSH3* no funcional. En esta modalidad, el ARNm quimérico no funcional *HTT-MSH3* codificaría un polipéptido *HTT* que termina en el primer codón de terminación en marco en la porción *MSH3* del ARNm. Como consecuencia, el pre-ARNm de *MSH3* que experimenta dicho proceso de trans-empalme se alteraría para convertirse en un ARNm que codifica pocos o ningún aminoácido de la secuencia codificante de *MSH3* y probablemente experimentaría una desintegración mediada por sin sentido. Este enfoque es aplicable a dominios de unión dirigidos a cualquier intrón *MSH3*. Aunque los ejemplos presentados aquí (Figura 19-22 y 30-32) utilizan editores de exones 5', se podría adoptar un enfoque similar con un editor de exones 3'.

Modalidad A.1.3. Editores de exones de ARN en tándem: Los editores de exones en tándem son dos editores de exones diferentes insertados en el mismo plásmido, AAV u otro vehículo de administración. Pueden disponerse en configuraciones de cabeza a cabeza o de cabeza a cola. Cada editor de exones puede tener sus propias secuencias reguladoras, incluido un promotor, o puede generarse a través de un evento de escisión a partir de una única transcripción bicistrónica. Un editor de exones puede apuntar a un gen o intrón y el segundo editor de exones puede apuntar a un gen o intrón diferente. En algunas modalidades, un editor de exones puede reemplazar *HTT* mutante con *HTT* de tipo salvaje y un segundo editor de exones puede reducir la expresión de *MSH3*. En algunas modalidades, un editor de exones puede apuntar a un intrón de *HTT* y el segundo editor puede apuntar a un segundo intrón de *HTT*, lo que aumenta la eficiencia general de la edición de exones de *HTT*.

Modalidad A.2. Reducción de *MSH3* u otros genes mediante ARNi en combinación con trans-empalme de *HTT*: Se puede agregar un ARN de horquilla corta (ARNhc) o un microARN (miARN) diseñado para reducir la expresión del gen a un editor de exones dentro del mismo cistrón (por ejemplo, en un intrón del editor de exones) o como un cistrón separado con sus propias secuencias reguladoras. El ARNi puede reducir la expresión del objetivo no editado (por ejemplo, *HTT* mutante) u otro gen (por ejemplo, *MSH3*). Con el fin de reducir selectivamente la *HTT* mutante con ARNip, el editor de exones reemplazaría una porción de la transcripción de la *HTT* mutante con una secuencia resistente al ARNhc o miARN. Ver, por ejemplo, las Figuras 33-35, 46-50.

Modalidad A.3. Reducción de *MSH3* mediante modulación de empalme vectorizado o bloqueo de la traducción: En algunas modalidades, *MSH3* se inactiva bloqueando la inclusión de uno o más exones de *MSH3* cuya ausencia impide que se traduzca una proteína funcional, lo que
 5 provoca una terminación prematura de la traducción. En algunas modalidades, *MSH3* se inactiva mediante la incorporación de secuencias complementarias a las uniones de empalme de *MSH3* en secuencias de ARN nuclear pequeño (ARNnp), como, por ejemplo, el ARNnp U7, lo que impide la inclusión de exones. Aquí, la modulación de empalme puede ser vectorizada, es decir, entregada por un AAV que también puede incluir un editor de exones. Ver, por
 10 ejemplo, las Figuras 36-45, 51-54.

Modalidad A.3.1. ARNas basada en ARN nuclear pequeña (por ejemplo, U7SmOPT) contra *MSH3*: En algunas modalidades, *MSH3* se inactiva mediante la incorporación de secuencias complementarias a las uniones de empalme de *MSH3* en secuencias de ARNnp, como, por
 15 ejemplo, el ARNnp U7. Ver, por ejemplo, las Figuras 36-45.

Modalidad A.3.2. Reducción combinada de *MSH3* con trans-empalme de *HTT*: En algunas modalidades, las estrategias A.3.1 y otras descritas en este documento se combinan en el mismo fragmento de ADN, o el mismo fragmento que un editor de exones, o se empaquetan
 20 conjuntamente en un único AAV para su administración conjunta en los tejidos del paciente. Ver, por ejemplo, las Figuras 33-54.

Reducción de posibles especies tóxicas de *HTT*: Se puede agregar un ARN de horquilla corta (ARNhc) o un microARN (miARN) diseñado para reducir la expresión del gen a un editor de
 25 exones como un cistrón separado con sus propias secuencias reguladoras. El ARNi puede reducir la expresión del objetivo no editado (por ejemplo, *HTT* mutante). Con el fin de reducir selectivamente la *HTT* mutante con ARNi, el editor de exones reemplazaría una porción de la transcripción de la *HTT* mutante con una secuencia resistente al ARNhc o miARN. Ver las
 30 Figura 46-50 y 54-56.

Con el fin de abordar la posibilidad de que el tratamiento de la HD se pueda mejorar mediante el uso de un enfoque terapéutico combinatorio, los presentes inventores han diseñado un enfoque híbrido de múltiples frentes en el que se apuntaron a múltiples mecanismos y/o
 35 especies patógenas. En algunas modalidades, estos enfoques incluyeron la sustitución de la *HTT* mutante mediante un editor de exones, el silenciamiento de la *HTT* mutante y sus transcritos relacionados (por ejemplo, *HTT1a*) mediante ARNi, y la reducción de *MSH3* mediante, por ejemplo, trans-empalme, ARN antisentido, ARNi, modulación del empalme vectorizado o bloqueo de la traducción, o cualquier combinación de estos. Al combinarse en el mismo fragmento de ADN, o en el mismo fragmento que un editor de exones, estos enfoques
 40 representan un mecanismo de acción multimodal que, no obstante, puede administrarse en un único AAV para su coadministración en los tejidos del paciente. Las modalidades que utilizan

tal enfoque combinatorio se pueden usar para lograr eficacia terapéutica mejorada en algunos, si no todos, los pacientes con HD. Este enfoque podría conferir incluso más efectos modificadores de la enfermedad.

5 Modelos animales

Modelo de ratón BACHD: Este modelo animal es un modelo de ratón transgénico mediado por cromosoma artificial bacteriano (BAC) (BACHD), en el que se expresa huntingtina mutante humana de longitud completa (fl-mhtt). Se desarrolló expresando fl-mhtt que comprende 97 repeticiones de glutamina bajo el control de la maquinaria reguladora HTT endógena en el BAC. Las repeticiones de glutamina están codificadas por 97 repeticiones CAG-CAA mixtas. Los ratones BACHD recapitulan la enfermedad de HD y la progresión de la enfermedad en humanos, ya que presentan déficits motores progresivos, disfunción sináptica neuronal y neuropatología selectiva de aparición tardía, que incluye atrofia cortical y estriatal significativa y degeneración de neuronas oscuras estriatales. Basándose en la robustez bien documentada de los fenotipos conductuales y neuropatológicos, los ratones BACHD son reconocidos como un modelo de ratón fl-mhtt adecuado para estudios preclínicos. Los ratones BACHD se describen en detalle en Gray et al. (2008, J Neuroscience 28:6182; cuyo contenido se incorpora aquí en su totalidad) y están disponibles comercialmente.

En algunas modalidades, la eficacia terapéutica de las moléculas de trans-empalme de ácidos nucleicos/editores de exones de ARN descritos en este documento en el contexto del modelo de ratón BACHD se puede medir por al menos uno de los siguientes: un aumento en el porcentaje de reemplazo de ARN de *HTT* patógeno, una reducción en *HTT* patógeno, una reducción en agregados de *HTT* patógenos, una mejora en la coordinación motora y el equilibrio medidos, por ejemplo, mediante una prueba de rotarod, una disminución en la atrofia del prosencéfalo o cualquier combinación de estos.

Modelo de ratón BAC-CAG: Este modelo animal es un modelo de ratón transgénico BAC genómico humano de HD que expresa huntingtina mutante humana (mHTT) que comprende repeticiones CAG largas ininterrumpidas y somáticamente inestables (120-130 repeticiones CAG puras) y exhibe fenotipos progresivos relacionados con la enfermedad. A diferencia de otros modelos transgénicos mHTT que tienen repeticiones codificantes de poliglutamina estables, interrumpidas por CAA, los ratones BAC-CAG presentan inclusiones nucleares robustas selectivas del estriado y una desregulación transcripcional que también se observa en pacientes con HD y en modelos de huntingtina knockin. BAC-CAG se describen en detalle en Gu et al. (2022, Neuron 110:1173; cuyo contenido se incorpora aquí en su totalidad) y están disponibles comercialmente. Como se describe allí, la transcriptopatía estriatal en los modelos HD está correlacionada con su longitud de repetición CAG ininterrumpida en lugar de la longitud de la poliglutamina. Como también se describe allí, la inestabilidad de la repetición CAG somática y la agregación nuclear de mHTT se correlacionan mejor con la patogénesis molecular selectiva del cuerpo estriado de inicio temprano y los déficits locomotores y del

sueño, mientras que las patologías asociadas a la repetición de ARN y la traducción no AUG (RAN) asociada a la repetición pueden tener un impacto en funciones patogénicas menos selectivas o tardías, respectivamente.

5 En algunas modalidades, la eficacia terapéutica de las moléculas de trans-empalme de ácidos nucleicos/editores de exones de ARN descritos en este documento en el contexto del modelo de ratón BAC-CAG se puede medir por al menos uno de los siguientes: un aumento en el porcentaje de reemplazo de ARN de *HTT* patógeno, una reducción en *HTT* patógeno, una reducción en agregados de *HTT* patógenos, una mejora en la coordinación motora y el equilibrio medidos, por ejemplo, mediante una prueba de rotarod, una disminución en la atrofia del prosencéfalo o una disminución de la transcriptopatía específica del estriado, o cualquier combinación de estos.

15 Para demostrar in vivo la prueba del mecanismo en el cerebro, los presentes inventores utilizaron el modelo de ratón BAC-CAG (Gu et al. 2022). Como se describió anteriormente, esta cepa de ratón transgénico fue diseñada para contener el locus genómico *HTT* mutante humano (>120 repeticiones CAG ininterrumpidas en el exón 1 de *HTT*) y mostrar características anormales que se asemejan al fenotipo de la enfermedad, incluidos agregados nucleares de *HTT* y desregulación transcripcional a los ≥ 12 meses de edad. Los presentes inventores realizaron estudios de prueba del mecanismo de trans-empalme de *HTT* en el cerebro del ratón BAC-CAG con moléculas ejemplares empaquetadas en AAV9. En este estudio se probaron tres artículos de prueba de AAV diferentes: editor de exones dirigido al intrón 2 de *HTT* impulsado por el promotor CMV (SEQ ID NO: 204) empaquetado en AAV autocomplementario (scAAV) (SEQ ID NO: 363, que comprende SEQ ID NO: 204), editor de exones dirigido al intrón 2 de *HTT* impulsado por el promotor CMV (SEQ ID NO: 204) empaquetado en AAV de cadena doble (ssAAV) (SEQ ID NO: 363, que comprende SEQ ID NO: 204) y editor de exones dirigido al intrón 2 de *HTT* impulsado por el promotor CAGGS (SEQ ID NO: 204) empaquetado en ssAAV (SEQ ID NO: 364, que comprende SEQ ID NO: 204). Los tres artículos de prueba también contenían un modulador de empalme del exón 7 MSH3 sustituto dirigido al ratón (SEQ ID NO: 362), empaquetado conjuntamente como una molécula híbrida. La secuencia completa de artículos de prueba es la siguiente: editor de exones dirigido al intrón 2 de *HTT* impulsado por el promotor CMV empaquetado en scAAV (SEQ ID NO: 369), editor de exones dirigido al intrón 2 de *HTT* impulsado por el promotor CMV empaquetado en ssAAV (SEQ ID NO: 370) y editor de exones dirigido al intrón 2 de *HTT* impulsado por el promotor CAGGS empaquetado en ssAAV (SEQ ID NO: 371). Para probar la respuesta a la dosis del editor de exones, se seleccionaron 2 dosis diferentes para inyecciones: $1E+11$ genomas vectoriales (vgs) y $3E+11$ vgs por animal, inyectados bilateralmente. Se seleccionaron dos vías de administración independientes: inyección intracerebroventricular (ICV) en el P0 neonatal (Tabla 2) e inyección intraestriatal a las 8 semanas de edad. La corteza y el cuerpo estriado del ratón se recolectaron 4 semanas después de la inyección y se perfilaron las eficiencias del reemplazo de exones *HTT* por trans-empalme y silenciamiento de

MSH3 por modulación de empalme mediante RT-ddPCR y análisis de transferencia Western.

Tabla 2. Diseño del estudio para el estudio de prueba del mecanismo in vivo en el cerebro del ratón BAC-CAG mediante inyección ICV. Para cada tratamiento, la edad en el momento de la recolección terminal fue de 4 semanas y la región del cerebro de la cual se realizó la recolección fue la corteza estriada.

Genotipo	Sexo	Tratamiento	ssAAV frente a scAAV	Serotipo	Edad del tratamiento	ROA	Dosis
HEMI	emparejado	CMVp:HTT/MSH3 (herramienta) híbrido	scAAV	AAV9	P0	ICV	1,00E+11
		CMVp:HTT/MSH3 (herramienta) híbrido					3,00E+11
		CMVp:HTT/MSH3 (herramienta) híbrido	ssAAV				1,00E+11
		CMVp:HTT/MSH3 (herramienta) híbrido					3,00E+11
		CAGGSp:HTT/MSH3 (herramienta) híbrido					3,00E+11
		Vehículo					

Como se muestra en la Figura 58, el análisis de los perfiles de trans-empalme de HTT demostró que se logró un reemplazo de HTT de más del 30 % en el cerebro del ratón mediante inyección de ICV neonatal. El editor de exones impulsado por el promotor CAGGS superó a su equivalente promotor CMV, como lo indica un mayor porcentaje de reemplazo de HTT tanto en la corteza como en el estriado. Entre los editores de exones impulsados por el promotor CMV, se observó una clara respuesta a la dosis, en la que los animales que recibieron 3E+11 vg tuvieron un mayor porcentaje de reemplazo de HTT que los que recibieron 1E+11 vg. Significativamente, la proteína HTT de longitud completa trans-empalmada se detectó mediante transferencia Western contra la etiqueta FLAG de extremo N. Este experimento también indicó que, al aumentar los niveles de transcripciones del Editor de exones, se logró un mayor nivel de trans-empalme (Figura 59), lo que sugiere que identificar un promotor más fuerte o administrar una dosis más alta podría resultar en una actividad del Editor de exones aún mayor.

I. Definiciones

Como se utiliza en la presente, "trans-empalme" significa unir una primera molécula de ARN que contiene uno o más exones (por ejemplo, exones exógenos o exones que son parte de un

dominio de codificación de una molécula de trans-empalme) a una segunda molécula de ARN (por ejemplo, una molécula pre-ARNm, por ejemplo, una molécula de pre-ARNm endógena) y reemplazar una parte de la segunda molécula de ARN con una parte de la primera molécula de ARN a través de un mecanismo mediado por espliceosoma. El mecanismo general para una reacción de trans-empalme del ARN se ilustra en la Figura 3.

Una "molécula de trans-empalme de ácido nucleico" o "molécula de trans-empalme" tiene tres elementos principales: (a) un dominio de unión que confiere especificidad al unir la molécula de trans-empalme a su gen objetivo (por ejemplo, pre-ARNm); (b) un dominio de empalme (por ejemplo, un dominio de empalme que tiene un sitio de empalme 3' o 5'); y (c) un dominio de codificación configurado para ser trans-empalmado en el ácido nucleico objetivo, que puede reemplazar uno o más exones en el ácido nucleico objetivo (por ejemplo, uno o más exones mutados). Una "molécula de trans-empalme de pre-ARNm" o "RTM" hace referencia a una molécula de trans-empalme de ácido nucleico que se dirige al pre-ARNm. Los términos "molécula de trans-empalme de ácido nucleico" y "molécula de trans-empalme" hacen referencia tanto (1) al ADN que codifica el ARN, donde la transcripción del ARN es la molécula efectora que se une físicamente al pre-ARNm objetivo; y (2) la transcripción del ARN en sí. Para mayor claridad, la expresión "secuencia de codificación" (por ejemplo, secuencia de codificación de molécula de trans-empalme) se utiliza en la presente memoria para especificar que el sujeto codifica el efector (por ejemplo, la secuencia de codificación es ADN y el efector es ARN). En algunas modalidades, una secuencia de codificación de una molécula de trans-empalme puede incluir ADNc, por ejemplo, como parte de un exón funcional (por ejemplo, un exón *HTT* funcional) para el reemplazo de un exón *HTT* mutado.

Como se utiliza en la presente, la expresión "editor de exones" puede usarse para hacer referencia a una molécula de trans-empalme o un vector que la comprende (por ejemplo, un vector AAV que comprende ADN que codifica una transcripción de ARN que es una molécula de trans-empalme).

Como se utiliza en la presente, "eficiencia de trans-empalme" hace referencia a una relación entre el nivel de expresión detectado del producto de ARN trans-empalmado deseado (es decir, una molécula de ARN quimérico que incluye los exones funcionales de la molécula de trans-empalme enlazados de forma operativa al pre-ARNm objetivo endógeno generado por una reacción de empalme de ARN) y la cantidad de ADN o ARN introducido para la molécula de trans-empalme (o molécula de referencia). En algunos casos, el nivel de expresión de un producto de ARN cortado y trans-empalmado se detecta a partir del ARN que se aísla de células o tejidos mediante secuenciación de ARN.

Tal como se utiliza en la presente, "% de reemplazo de ARN" hace referencia a la parte de la población total de ARNm objetivo que ha experimentado un trans-empalme (TS) exitoso, y se calcula a través de la siguiente ecuación: % de TS en el objetivo (ONT) = 100*(número de

copias de ONT/(número de copias de ONT + número de copias nativas)).

Como se utiliza en la presente, "eficiencia de trans-empalme relativa" hace referencia a una relación entre una eficiencia de trans-empalme de prueba y una eficiencia de trans-empalme de referencia, en donde la eficiencia de trans-empalme de prueba es la eficiencia de trans-empalme de una molécula de trans-empalme (por ejemplo, una molécula de trans-empalme de ácido nucleico descrita en la presente), y la eficiencia de trans-empalme de referencia es la eficiencia de trans-empalme de una molécula de referencia (por ejemplo, una molécula de referencia que tiene los mismos elementos que la molécula de trans-empalme de ácido nucleico excepto que el dominio de unión se reemplaza con un dominio de unión desordenado o un dominio de unión no dirigido (por ejemplo, un dominio de unión que comprende, o consiste en, la SEQ ID NO: 7). La eficiencia relativa de trans-empalme de una molécula de trans-empalme se puede dar como una relación (también conocida como aumento de veces) de la eficiencia de trans-empalme del ARN de prueba sobre la eficiencia de trans-empalme de referencia probada en condiciones similares.

Como se utiliza en la presente memoria, el término «unido de forma operativa» o «de forma operativa unido» hace referencia a una disposición de elementos, donde los componentes descritos de esta manera están configurados para realizar su función habitual. Un ácido nucleico está "unido de forma operativa" a otra secuencia del ácido nucleico cuando se coloca en una relación funcional con la otra secuencia del ácido nucleico. Los elementos no necesitan ser contiguos para estar unidos de forma operativa. Así, por ejemplo, pueden estar presentes secuencias intermedias entre secuencias unidas de forma operativa (por ejemplo, un dominio de unión y una secuencia de codificación pueden estar separados por secuencias intermedias y el dominio de unión todavía se considera como "unido de forma operativa" a la secuencia de codificación).

Como se utiliza en la presente memoria, el término "dominio de codificación" hace referencia a una secuencia del ácido nucleico (por ejemplo, una secuencia de ARN, una secuencia de ADN o una combinación de ARN y ADN) que codifica una parte de una proteína (por ejemplo, una proteína objetivo en la que se está corrigiendo una mutación). Por lo tanto, un dominio de codificación puede incluir uno o más exones funcionales (por ejemplo, una secuencia de exones funcionales). En algunos casos, uno o más exones funcionales de un dominio de codificación no están separados por intrones (por ejemplo, como en el pre-ARNm endógeno) sino adyacentes entre sí (por ejemplo, tal como el ADNc). En algunos casos, un dominio de codificación puede incluir uno o más intrones (por ejemplo, intrones nativos) o regiones no traducidas (UTR, por ejemplo, UTR nativas) entre o adyacentes a (por ejemplo, por delante o detrás de) los exones.

Como se utiliza en la presente memoria, una "región no traducida *HTT* 5' nativo" o "UTR 5' *HTT* nativo" hace referencia a una secuencia de más de 20 nucleótidos de longitud que tiene

al menos un 90 % de identidad de secuencia con una región de un gen *HTT* nativo (por ejemplo, un gen *HTT* humano) que está 5' al codón de inicio ATG. Un ejemplo de una región 5' *HTT* nativa no traducida se da mediante la secuencia de ADN de la SEQ ID NO: 136. Un ejemplo de una versión modificada de unas regiones no traducidas 5' *HTT* se da mediante la

Como se utiliza en la presente memoria, una "secuencia funcional de exones 5' *HTT*" hace referencia a una secuencia del ácido nucleico que comprende uno o más de los exones *HTT* 1-3 (por ejemplo, exón 1; o exón 1 y exón 2; o exones y exón 2 y exón 3) que codifican una parte funcional (biológicamente activa) de la proteína *HTT*. En algunas modalidades, una "secuencia funcional de exones 5' *HTT*" hace referencia a una secuencia del ácido nucleico que comprende el exón 1 de *HTT* o los exones 1-2 de *HTT* o los exones 1-3 de *HTT* que codifica una parte funcional (biológicamente activa) de la proteína *HTT*. Cuando se trans-empalma a un exón *HTT* endógeno 3' al sitio de unión, una secuencia funcional de exones *HTT* 5' proporciona la expresión de la proteína *HTT* funcional (por ejemplo, proteína *HTT* no mutada). En algunos casos, la secuencia funcional de los exones 5' *HTT* incluye una secuencia de exones adyacentes al exón al que se va a trans-empalmar la molécula de trans-empalme (por ejemplo, una molécula de trans-empalme que se une al intrón 2 de *HTT* y se trans-empalma con el exón 3 de *HTT* endógeno puede incluir una secuencia funcional de exones 5' *HTT* que incluye los exones 1 y 2, o una molécula de trans-empalme que se une al intrón 3 de *HTT* y se trans-empalma con el exón 4 de *HTT* endógeno puede incluir una secuencia funcional de exones 5' *HTT* que incluye los exones 1-3).

Como se utiliza en la presente memoria, el término "funcional", cuando se utiliza en el contexto de una proteína, hace referencia a una proteína biológicamente activa. El término "funcional" también puede usarse para hacer referencia a la cantidad de actividad de una proteína que es necesaria para sustentar las funciones celulares normales. Con respecto a *HTT*, el término "funcional" puede usarse para hacer referencia a la cantidad de actividad de la proteína *HTT* que es necesaria para restaurar los niveles de actividad de *HTT* para sustentar las funciones celulares normales dentro del contexto de, por ejemplo, las neuronas piramidales del córtex, neuronas espinosas medias del cuerpo estriado y/o neuronas hipotalámicas. Además, reducir la cantidad de *HTT* mutante y/o sesgar la proporción mutante:salvaje contribuirá significativamente a la reducción de la neurodegeneración de neuronas piramidales en la corteza, neuronas espinosas medianas en el cuerpo estriado y/o neuronas hipotalámicas. En el contexto del tratamiento de una afección asociada con la actividad patógena de *HTT* (por ejemplo, HD) o del uso de un agente terapéutico que comprende una molécula de trans-empalme de ácidos nucleicos descrita en este documento, "funcional" se refiere a la reducción de la cantidad, por ejemplo, de proteína *HTT* defectuosa (no funcional) que comprende trectos de poliglutamina en exceso de 35 o 40 repeticiones de glutamina para eliminar uno o más síntomas de una afección asociada con la actividad patógena de *HTT* (por ejemplo, HD). En algunas modalidades, dichos métodos o usos conducen a una disminución de la actividad

patógena de HTT y a un aumento de la actividad de la proteína HTT de tipo salvaje. En algunas modalidades, dicha disminución en la actividad de la proteína HTT patógena disminuye los niveles de actividad de HTT patógena en al menos un 10%, al menos un 20%, al menos un 30%, al menos un 40%, al menos un 50%, al menos un 60%, al menos un 70%, al menos un 75%, al menos un 80%, al menos un 85%, al menos un 90% o al menos un 95% de la actividad de HTT (por ejemplo, un 96%, un 97%, un 98%, un 99% o un 100%) en relación con la de una célula de control (sin tratar) en la que se expresa HTT patógena mutada. En algunas modalidades, dicho aumento en la actividad de la proteína HTT restaura los niveles de actividad de HTT a al menos un 10 %, al menos un 20 %, al menos un 30 %, al menos un 40 %, al menos un 50 %, al menos un 60%, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, o al menos un 95 % de la actividad de HTT (por ejemplo, 96 %, 97 %, 98 %, 99 %, o un 100 %) en relación con la de una célula en la que HTT está presente en niveles normales, de tipo salvaje, tales como los presentes en células en las que se expresa HTT no mutado.

Como se utiliza en la presente memoria, "sitio de empalme críptico corregido", "sitio de empalme críptico mitigado" o "sitio de empalme críptico resistente" hace referencia a una molécula de trans-empalme de ácido nucleico o una parte de esta (por ejemplo, una secuencia de dominio de codificación en esta) que ha sido modificada para cambiar nucleótidos individuales en esta para reducir la frecuencia de empalme que ocurre en un sitio de empalme críptico identificado en el contexto de la molécula de trans-empalme de ácido nucleico. En algunas modalidades, las modificaciones no dan lugar a ningún cambio en las secuencias de aminoácidos codificadas por estas. En algunas modalidades, la secuencia de ácido nucleico resistente al sitio de empalme críptico dentro de una molécula de trans-empalme de ácido nucleico está en una secuencia de dominio codificante (CDS). En algunas modalidades, el CDS de *HTT* resistente al sitio de empalme críptico comprende, consiste esencialmente en, o consiste en el exón 1 o los exones 1 y 2 o los exones 1, 2 y 3 del gen *HTT*, en donde se han identificado sitios de empalme crípticos en el contexto de la molécula de trans-empalme de ácido nucleico y en donde al menos uno de los sitios de empalme crípticos se ha modificado para reducir la frecuencia de empalme en el al menos un sitio, sin alterar los aminoácidos codificados por este.

Un "dominio de empalme", como se utiliza en la presente memoria, hace referencia a una secuencia del ácido nucleico que tiene motivos que son reconocidos por el espliceosoma y median el trans-empalme. Un dominio de empalme incluye un sitio de empalme (por ejemplo, un único sitio de empalme, es decir, uno y solo un sitio de empalme), que puede ser un sitio de empalme 3' o un sitio de empalme 5'. Un dominio de empalme puede incluir otros elementos reguladores. En algunas modalidades, el dominio de empalme comprende GUAAGT o GTAAGT. En algunas modalidades, el sitio de empalme consiste esencialmente en GUAAGT o GTAAGT. En algunas modalidades, el sitio de empalme consiste en GUAAGT o GTAAGT.

Como se utiliza en la presente memoria, el "dominio de unión" de una molécula de trans-empalme es una secuencia de polinucleótidos que se une a un gen objetivo en un sitio de unión a través de hibridación (es decir, complementariedad total o parcial con el sitio de unión).

- 5 Como se utiliza en la presente memoria, la expresión "sitio de unión" hace referencia a una secuencia de polinucleótido endógena en el pre-ARNm objetivo (por ejemplo, un pre-ARNm de un gen endógeno, por ejemplo, *HTT*) que está unido por el dominio de unión de una molécula de trans-empalme de ácido nucleico. El sitio de unión se extiende desde el nucleótido más 5' unido por el dominio de unión hasta el nucleótido más 3' unido por el dominio de unión.
- 10 En algunas modalidades, el sitio de unión tiene la misma longitud que el dominio de unión. En otras modalidades, el sitio de unión se encuentra entre 1-10 nucleótidos más largo o más corto que el dominio de unión (es decir, algunos de los nucleótidos del sitio de unión o del dominio de unión no están hibridados). En modalidades que involucran dominios de unión que tienen al menos dos secuencias no superpuestas con al menos un 80 % de complementariedad con
- 15 el sitio de unión, el sitio de unión puede ser sustancialmente más corto que el dominio de unión.

Como se utiliza en la presente memoria, "complementariedad" y sus variaciones gramaticales hacen referencia al porcentaje de bases de nucleótidos de una secuencia dada que se empareja a través de enlaces de hidrógeno con una secuencia de referencia.

- 20 Como se utiliza en la presente memoria, una secuencia dada (por ejemplo, una secuencia de dominio de unión) es "100 % complementaria a" o tiene "100 % de complementariedad" con una secuencia de referencia (por ejemplo, un sitio de unión de pre-ARNm endógeno) si cada una de las bases de nucleótidos de la secuencia dada se empareja a través de enlaces de hidrógeno con la secuencia de referencia, hibridándose así para formar una secuencia de
- 25 doble cadena (por ejemplo, a través del apareamiento de bases de Watson-Crick, por ejemplo, cada A se empareja con una T o U, y cada C se empareja con una G). Por ejemplo, un dominio de unión que está en una orientación antisentido a un sitio de unión es complementario al sitio de unión. El emparejamiento de ARN incluye el emparejamiento de G con U; por lo tanto, un
- 30 dominio unión de ARN que tiene emparejamiento de G-U con su sitio de unión puede ser un 100 % complementario con el sitio de unión. Por consiguiente, un dominio de unión que es exactamente el complemento inverso de su sitio de unión (es decir, las A del dominio de unión están emparejadas con las U del sitio de unión) se puede modificar para reemplazar una o más de las A con las G o las C con las T sin afectar sustancialmente la unión.

- 35 Como se usa en la presente memoria, una secuencia dada (por ejemplo, una secuencia de dominio de unión) es "al menos X% complementaria a" o tiene "X% de complementariedad" con una secuencia de referencia (por ejemplo, un sitio de unión al pre-ARNm endógeno) si X% de las bases nucleótidas de los pares de secuencias dados a través de la unión al hidrógeno con la secuencia de referencia, por ejemplo, hibridar para formar una secuencia de cadena
- 40 doble (por ejemplo, a través del apareamiento de base de Watson-Crick, por ejemplo, A se

empareja con T o U, y C se empareja con G). Por ejemplo, una secuencia de dominio de unión que tiene una longitud de 150 bases es al menos un 90 % complementaria a un sitio de unión que tiene una longitud de 150 bases si al menos 135 de sus 150 residuos se aparean a través de enlaces de hidrógeno con el sitio de unión a través del apareamiento de bases Watson-Crick, dejando 15 o menos nucleótidos no coincidentes.

Con respecto a las secuencias presentadas en este documento y en el listado de secuencias adjunto, se entiende que las transcripciones de ARN codificadas por secuencias de ADN comprenden una uridina (U) en posiciones correspondientes a la timidina (T) como se indica en la secuencia de ADN correspondiente. En algunos casos, las secuencias de componentes del editor de exones de ARN se divulgan aquí como secuencias de ADN. Para cualquier secuencia divulgada en este documento como una secuencia de ADN, también se contempla una secuencia de ARN con U sustituida por cada T en la secuencia. Por lo tanto, si se identifica que una SEQ ID NO determinada tiene una secuencia que puede incluirse en un editor de exones de ARN, también se contempla una versión del SEQ ID NO con U sustituida por cada T.

"Unión" entre un dominio de unión y un intrón, como se utiliza en la presente memoria, hace referencia a la unión de hidrógeno (por ejemplo, formación de doble hélice o emparejamiento de Watson Crick) entre el dominio de unión y el intrón objetivo en un grado suficiente para mediar el trans-empalme al poner la molécula de trans-empalme junto con el objetivo (por ejemplo, pre-ARNm). En algunas modalidades, los enlaces de hidrógeno entre el dominio de unión y el intrón objetivo se encuentran entre bases de nucleótidos que son complementarias y están en orientación antisentido entre sí (por ejemplo, hibridadas entre sí).

Como se utiliza en la presente memoria, un "intrón artificial" hace referencia a una secuencia del ácido nucleico no de codificación que vincula (directa o indirectamente) un dominio de unión a un dominio de codificación. Un intrón artificial incluye un dominio de empalme y puede incluir además una o más secuencias espaciadoras y/u otros elementos reguladores.

Como se utiliza en la presente memoria, el término "mutación" puede usarse para hacer referencia a cualquier secuencia del ácido nucleico aberrante que cause un producto proteico o ARN defectuoso (por ejemplo, un producto proteico no funcional, una proteína no biológicamente activa, un producto proteico que tiene una función reducida, un producto proteico que tiene una función aberrante y/o un producto proteico que se produce en cantidades menores a las normales o mayores a las normales). Las mutaciones incluyen mutaciones de pares de bases (por ejemplo, polimorfismos de un solo nucleótido), duplicaciones, mutaciones sin sentido, mutaciones por desplazamiento del marco de lectura, deleciones, inserciones y mutaciones de empalme. En algunas modalidades, una mutación hace referencia a una secuencia del ácido nucleico que es diferente en una o más porciones de su secuencia que una secuencia del ácido nucleico de tipo salvaje correspondiente o una

variante funcional de esta. En algunas modalidades, una mutación hace referencia a una secuencia del ácido nucleico que codifica una proteína que tiene una secuencia de aminoácidos que es diferente de una proteína de tipo salvaje correspondiente o una variante funcional de esta. Un "exón mutado" (por ejemplo, un exón *HTT* mutado) hace referencia a un exón que contiene una mutación o una secuencia de exones que refleja una mutación en una región diferente, como un exón críptico resultante de una mutación en un intrón.

El término "*HTT*" (*Huntingtina*) se refiere a cualquier *HTT* nativa de cualquier fuente vertebrada, incluyendo mamíferos tales como primates (por ejemplo, humanos, monos verdes africanos y monos cynomolgus) y roedores (por ejemplo, ratones y ratas), a menos que se indique lo contrario, así como variantes funcionalmente equivalentes o mejoradas (por ejemplo, variantes naturales o sintéticas), mutantes, muteínas, análogos, subunidades, complejos receptores, isotipos, variantes de empalme y fragmentos de estas. Se pueden determinar variantes funcionalmente equivalentes y mejoradas basándose en la señalización *HTT* conocida. *HTT* abarca *HTT* sin procesar de longitud completa, así como cualquier forma de *HTT* que es el resultado del procesamiento nativo en la célula. Se proporciona una secuencia *HTT* humana de ejemplo como secuencia de referencia del Centro Nacional de Información Biotecnológica (NCBI): NG_009378. En algunos casos, un fragmento *HTT* está codificado por un agente terapéutico que comprende una secuencia que tiene al menos un 95 % de identidad de secuencia con cualquiera de las SEQ ID NO: 3, 59, 157, 349, 350, 351, 352 o 353 (por ejemplo, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % o un 100 % de identidad de secuencia con cualquiera de las SEQ ID NO: 3, 59, 157, 349, 350, 351, 352 o 353 una porción funcional de este y/o una variantemodificada por codones de este.

Como se utiliza en la presente memoria, una "variante" hace referencia a un polinucleótido que difiere en al menos un residuo de ácido nucleico de la secuencia de polinucleótido de referencia, como una secuencia de polinucleótido de origen natural, o un polipéptido (por ejemplo, una secuencia de cápside de AAV) que difiere en al menos un residuo de aminoácido de la secuencia de polipéptido de referencia, tal como una secuencia de polipéptido de origen natural o, por ejemplo, cualquiera de las secuencias de rAAV descritas en la presente memoria. En este contexto, la diferencia en al menos un residuo puede incluir, por ejemplo, una sustitución de un residuo de ácido nucleico por otro ácido nucleico, una eliminación o una inserción, o una sustitución de un residuo de aminoácido por otro aminoácido. Una variante puede ser un homólogo, una isoforma o una variante de transcripción de un polinucleótido como se define en la presente memoria, donde el homólogo, la isoforma o la variante de transcripción se caracteriza por un grado de identidad u homología, respectivamente, como se define en la presente memoria.

En algunos casos, una variante de un polinucleótido o polipéptido incluye al menos una sustitución de ácido nucleico (por ejemplo, 1-100 sustituciones de ácido nucleico o aminoácido, 1-50 sustituciones de ácido nucleico o aminoácido, 1-20 sustituciones de ácido nucleico o

aminoácido, 1-10 sustituciones de ácido nucleico o aminoácido, por ejemplo, 1 sustitución de ácido nucleico o aminoácido, 2 sustituciones de ácido nucleico o aminoácido, 3 sustituciones de ácido nucleico o aminoácido, 4 sustituciones de ácido nucleico o aminoácido, 5 sustituciones de ácido nucleico o aminoácido, 6 sustituciones de ácido nucleico o aminoácido, 7 sustituciones de ácido nucleico o aminoácido, 8 sustituciones de ácido nucleico o aminoácido, 9 sustituciones de ácido nucleico o aminoácido o 10 sustituciones de ácido nucleico o aminoácido). Las sustituciones de ácidos nucleicos que dan como resultado que el polipéptido expresado tenga un aminoácido intercambiado de la misma clase se denominan en la presente memoria sustituciones conservadoras. En particular, se trata de aminoácidos que tienen cadenas laterales alifáticas, cadenas laterales cargadas positiva o negativamente, grupos aromáticos en las cadenas laterales o aminoácidos cuyas cadenas laterales pueden formar puentes de hidrógeno, por ejemplo, cadenas laterales que tienen una función hidroxilo. Por sustitución conservadora, por ejemplo, un aminoácido que tiene una cadena lateral polar puede ser reemplazado por otro aminoácido que tenga una cadena lateral polar correspondiente, o, por ejemplo, un aminoácido caracterizado por una cadena lateral hidrófoba puede ser sustituido por otro aminoácido que tenga una cadena lateral hidrófoba correspondiente (por ejemplo, serina (treonina) por treonina (serina) o leucina (isoleucina) por isoleucina (leucina)).

En algunos casos, las inserciones, las eliminaciones y/o las sustituciones no conservadoras también están incluidas en el término variante, por ejemplo, en aquellas posiciones que no causan una modificación sustancial de la estructura tridimensional de la proteína. Las modificaciones a una estructura tridimensional mediante inserción/es o eliminación/es pueden ser determinadas fácilmente por Un experto en la técnica, por ejemplo, utilizando espectros CD (espectros de dicroísmo circular).

El término «homólogo» hace referencia al grado de identidad entre secuencias de dos secuencias de ácidos nucleicos. La homología de secuencias se determina comparando dos secuencias alineadas en condiciones estándar a lo largo de la longitud de la secuencia a comparar. Las secuencias que se van a comparar en la presente memoria pueden tener una adición o eliminación (por ejemplo, un espacio y similares) en la alineación óptima de las dos secuencias. En algunas modalidades, la homología de secuencia se calcula creando una alineación utilizando, por ejemplo, el algoritmo ClustalW (Nucleic Acid Res., 1994, 22(22): 4673-4680). También se pueden utilizar programas informáticos de análisis de secuencias comúnmente disponibles, como Vector NTI, GENETYX, BLAST, o herramientas de análisis proporcionadas por bases de datos públicas.

El término "AAV" o "serotipo AAV" como se utiliza en la presente memoria hace referencia a las docenas de virus adenoasociados naturales y disponibles, así como a los AAV artificiales. Entre los AAV aislados o diseñados a partir de primates humanos o no humanos (NHP, por su sigla en inglés) y bien caracterizados, el AAV2 humano es el primer AAV que se desarrolló

como vector de transferencia genética; se ha utilizado ampliamente para experimentos eficientes de transferencia genética en diferentes tejidos objetivo y modelos animales. AAV9, AAV-retro, AAV1, AAV4, AAV8, AAV5, AAV-PHP.eB, por ejemplo, se encuentran entre los serotipos de AAV que son neurotrópicos por naturaleza.

5 Como se utiliza en la presente memoria, en relación con AAV, el término variante significa cualquier secuencia de AAV que se deriva de una secuencia de AAV conocida, incluidas aquellas que comparten al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 95 %, al menos un 97 %, al menos un 99 % o más
10 de identidad de secuencia en la secuencia de aminoácidos o de ácidos nucleicos. En otra modalidad, la cápside de AAV incluye variantes que pueden incluir hasta aproximadamente un 10 % de variación de cualquier secuencia de cápside de AAV descrita o conocida. Es decir, la cápside de AAV comparte aproximadamente un 90 % de identidad con aproximadamente un 99,9 % de identidad, aproximadamente un 95 % con aproximadamente un 99 % de identidad
15 o aproximadamente un 97 % con aproximadamente un 98 % de identidad con una cápside de AAV proporcionada en la presente memoria y/o conocida en la técnica. En una modalidad, la cápside de AAV comparte al menos un 95 % de identidad con una cápside de AAV. Al determinar el porcentaje de identidad de una cápside de AAV, la comparación puede realizarse sobre cualquiera de las proteínas variables (por ejemplo, vp1, vp2 o vp3).

20 Las ITR u otros componentes de AAV se pueden aislar o diseñar fácilmente utilizando técnicas disponibles para aquellos expertos en la técnica a partir de un AAV. Dichos AAV pueden aislarse, diseñarse u obtenerse de fuentes académicas, comerciales o públicas (por ejemplo, la American Type Culture Collection, Manassas, Virginia). Como alternativa, las secuencias de
25 AAV pueden diseñarse mediante medios sintéticos u otros medios adecuados haciendo referencia a secuencias publicadas tales como las que están disponibles en la literatura o en bases de datos como, por ejemplo, GenBank, PubMed o similares. Los virus AAV pueden diseñarse mediante técnicas de biología molecular convencionales, lo que hace posible optimizar estas partículas para la administración específica de secuencias de ácidos nucleicos
30 a las células, para minimizar la inmunogenicidad, para ajustar la estabilidad y la vida útil de las partículas, para una degradación eficiente, para una administración precisa al núcleo, etc.

Como se utiliza en la presente memoria, el término "sujeto", "individuo" o "paciente" incluye a cualquier mamífero que necesite estos métodos de tratamiento o profilaxis, incluidos los
35 primates, como los humanos. Otros mamíferos que necesitan dicho tratamiento o profilaxis incluyen primates no humanos (NHP; por ejemplo, monos cynomolgus y monos verdes africanos), perros, gatos u otros animales domésticos, caballos, ganado, animales de laboratorio (por ejemplo, ratones, ratas o conejos), etc. El individuo puede ser macho o hembra. En una modalidad, el individuo tiene una enfermedad o trastorno causado por una mutación
40 en el gen *HTT* (por ejemplo, HD). En otra modalidad, el individuo corre el riesgo de desarrollar una enfermedad o trastorno causado por una mutación en el gen *HTT*. En otra modalidad, el

individuo mostró signos clínicos de una enfermedad o trastorno causado por una mutación en el gen *HTT*, tal como HD. El individuo puede tener cualquier edad durante la cual el tratamiento o la terapia profiláctica pueden ser beneficiosos. Por ejemplo, en algunas modalidades, el individuo tiene entre 0 y 5 años, entre 5 y 10 años, entre 10 y 20 años, entre 20 y 30 años, entre 30 y 40 años, entre 30 y 50 años, entre 40 y 50 años, entre 50 y 60 años, entre 60 y 70 años o más de 70 años.

Como se utilizan en la presente memoria, las expresiones "trastorno asociado con una mutación" o "mutación asociada con un trastorno" hacen referencia a una correlación entre un trastorno y una mutación. En algunas modalidades, se sabe o se sospecha que un trastorno asociado con una mutación es causado total o parcialmente, o directa o indirectamente, por la mutación. Por ejemplo, un individuo que tenga la mutación puede correr el riesgo de desarrollar el trastorno, y el riesgo puede depender además de otros factores, como otras mutaciones (por ejemplo, independientes) (por ejemplo, en el mismo gen o en uno diferente) o factores ambientales.

Como se utiliza en la presente memoria, el término «tratamiento», o una derivación gramatical de este, se define como reducir la progresión de una enfermedad, reducir la gravedad de un síntoma de una enfermedad, retardar la progresión de un síntoma de una enfermedad, eliminar un síntoma de una enfermedad o retrasar la aparición de una enfermedad. En algunas modalidades, el término "tratamiento" se utiliza para hacer referencia a un efecto persistente o duradero de un agente terapéutico tal como un editor de exones de ARN descrito en la presente memoria.

Como se utiliza en la presente memoria, el término "prevención" de un trastorno, o una derivación gramatical de este, se define como la reducción del riesgo de aparición de una enfermedad, por ejemplo, como terapia profiláctica para un individuo que corre el riesgo de desarrollar un trastorno asociado con una mutación. Un individuo puede ser caracterizado como "en riesgo" de desarrollar un trastorno mediante la identificación de una mutación asociada con el trastorno, de acuerdo con cualquier método adecuado conocido en la técnica o descrito en la presente memoria. En algunas modalidades, un individuo que está en riesgo de desarrollar un trastorno tiene una o más mutaciones *HTT* asociadas con el trastorno. De manera adicional o alternativa, un individuo puede ser caracterizado como "en riesgo" de desarrollar un trastorno si tiene antecedentes familiares del trastorno.

El tratamiento o la prevención de un trastorno en un individuo se puede realizar administrando directamente la molécula de trans-empalme o el editor de exones de ARN (por ejemplo, dentro de un vector, por ejemplo, un vector AAV o una partícula AAV) al individuo. De manera alternativa, se pueden administrar al individuo células hospedadoras que contienen la molécula de trans-empalme.

El término "administrar" o una derivación gramatical de este, como se utiliza en los métodos que se describen en la presente memoria, hace referencia a administrar una molécula de trans-empalme o un editor de exones de ARN (por ejemplo, dentro de un vector, por ejemplo, un vector de AAV o una partícula de AAV) o una composición de estos, o una célula tratada ex vivo, al individuo que lo necesita, por ejemplo, un individuo que tiene una mutación o defecto en *HTT*. Por ejemplo, en una modalidad en la que se apuntan a células estriatales (por ejemplo, neuronas espinosas medianas) o células en la corteza (por ejemplo, neuronas piramidales), el método implica administrar una molécula de trans-empalme o un editor de exones de ARN (por ejemplo, dentro de un vector, por ejemplo, un vector AAV o partícula AAV) o una composición de estos al individuo mediante administración intracerebral (IC) (por ejemplo, inyección de administración lenta o inyección de difusión mejorada por convección), administración intracerebroventricular (ICV) o administración intratecal. En algunas modalidades, las inyecciones IC implican la implantación estereotáxica de manguitos guía de microinyección para mejorar la administración a un locus específico en el cerebro. En otra modalidad, la composición se administra por vía sistémica (por ejemplo, por vía intravenosa). Un experto en la técnica podrá seleccionar otros métodos de administración en vista de esta descripción.

En la presente memoria, la expresión "modular la expresión de *HTT*" hace referencia a disminuir la expresión de *HTT* mutado endógeno y/o aumentar la expresión de *HTT* trans-empalmado. La modulación de la expresión de *HTT* puede usarse para referirse, por ejemplo, a la disminución de la expresión de *HTT* endógeno (por ejemplo, mutado) y/o al aumento de la expresión de *HTT* trans-empalmado (por ejemplo, transcripción *HTT* o producto proteico que tiene un sitio de mutación corregida mediada por una molécula de trans-empalme en relación con su transcripción o producto proteico mutado endógeno). Tras la sustitución del exón *HTT* endógeno que comprende el sitio de mutación mediante trans-empalme, se expresa una proteína *HTT* funcional.

Como se utiliza en la presente memoria, "optimización de codones" hace referencia a modificar una secuencia del ácido nucleico para cambiar ácidos nucleicos individuales sin ningún cambio resultante en el aminoácido codificado. Las secuencias modificadas de esta manera se denominan en la presente memoria "optimizadas por codones". Este proceso se puede realizar en cualquiera de las secuencias descritas en esta memoria descriptiva para mejorar el rendimiento o la estabilidad. La optimización por codones se puede realizar de una manera como la descrita en, por ejemplo, las patentes estadounidenses números 7.561.972, 7.561.973 o 7.888.112, cada una de las cuales se incorpora a la presente mediante referencia en su totalidad. La secuencia que rodea el sitio de inicio de la traducción se puede convertir en una secuencia Kozak consenso de acuerdo con métodos conocidos. Véase, por ejemplo, Kozak y cols., 1987. *Nucleic Acids Res.* 15 (20): 8125-8148, que se incorpora a la presente memoria a modo de referencia en su totalidad.

La expresión "aceptable desde un punto de vista farmacéutico" significa seguro para su

administración a un mamífero, como un ser humano. En algunas modalidades, una composición aceptable desde un punto de vista farmacéutico está aprobada por un organismo regulador del gobierno federal o estatal o figura en la Farmacopea de los EE. UU. u otra farmacopea generalmente reconocida para su uso en animales, y más particularmente en seres humanos.

El término "portador" hace referencia a un diluyente, adyuvante, excipiente o vehículo con el que se administra una molécula terapéutica (por ejemplo, una molécula de trans-empalme o una molécula de trans-empalme que incluye un vector o célula de la presente invención). Se describen ejemplos de portadores farmacéuticos adecuados en «Remington's Pharmaceutical Sciences», Mack Publishing Co., Easton, PA., 18.^a edición.

Los términos «un» y «una» significan «uno o más de». Por ejemplo, se entiende que «un gen» representa uno o más de dichos genes. Como tal, los términos "un" y "una", "uno o más de un (o una)" y "al menos uno de un (o una)" se usan indistintamente en la presente memoria.

Como se utiliza en la presente memoria, el término "aproximadamente" hace referencia a un valor dentro de un margen de variabilidad de $\pm 10\%$ respecto del valor de referencia, a menos que se especifique lo contrario.

II. Moléculas de trans-empalme

Se proporcionan en este documento moléculas de trans-empalme de ácidos nucleicos útiles para corregir mutaciones en *HTT* mediante la sustitución de al menos un exón de *HTT* mutado por un exón de *HTT* funcional [por ejemplo, un exón de *HTT* 5' al sitio de unión, por ejemplo, exón 1 (SEQ ID NO: 348; *HTT* 5' UTR + exón 1, que incluye 21 repeticiones de CAG (dentro del rango normal de número de repeticiones), exón 2 (SEQ ID NO: 365), exón 3 (SEQ ID NO: 366 de *HTT*]. Tener en cuenta que si bien la secuencia del exón 1 de *HTT* establecida en la SEQ ID NO: 348 incluye 21 repeticiones de CAG, la cantidad de repeticiones de CAG en secuencias genómicas individuales puede ser diferente (es decir, mayor o menor). Las posiciones de numeración de nucleótidos correspondientes en secuencias que tienen diferentes cantidades de repeticiones de CAG se pueden determinar fácilmente teniendo en cuenta cualquier variación en la cantidad de repeticiones de CAG. En algunas modalidades, la molécula de trans-empalme de ácido nucleico es una molécula de trans-empalme de ARN (RTM, por su sigla en inglés). El diseño de la molécula de trans-empalme permite el reemplazo de la parte defectuosa o mutada del o los exones del pre-ARNm con una secuencia del ácido nucleico, por ejemplo, el o los exones que tienen una secuencia funcional (por ejemplo, normal) sin la mutación. La secuencia funcional puede ser una secuencia de tipo salvaje, de origen natural, o una secuencia corregida con alguna otra modificación, por ejemplo, optimización de codones.

Las moléculas de trans-empalme comprenden un dominio de unión, un dominio de empalme y un dominio de codificación. En algunas modalidades, la molécula de trans-empalme de ácido

nucleico tiene un dominio regulador 5' que tiene una región *HTT* 5' nativa no traducida (por ejemplo, una secuencia que tiene al menos un 80 % de identidad de secuencia con cualquiera de las SEQ ID NO: 136 o 192). En algunas modalidades, la molécula de trans-empalme de ácido nucleico tiene un sitio de empalme de GTAAGT. En algunas modalidades, la molécula de trans-empalme de ácido nucleico tiene un dominio enlazador que tiene una longitud mayor a 25 nucleótidos. En algunas modalidades, la molécula de trans-empalme de ácido nucleico tiene un dominio enlazador que comprende, consiste esencialmente o consiste en cualquiera de la SEQ ID NO: 37-46 y 106-112 o una secuencia que tiene al menos un 91 % de identidad, al menos un 92 % de identidad, al menos un 93 % de identidad, al menos un 94 % de identidad, al menos un 95 % de identidad, al menos un 96 % de identidad, al menos un 97 % de identidad, al menos un 98 % de identidad, o al menos un 99 % de identidad con cualquiera de las SEQ ID NO: 37-46 y 106-112.

En algunas modalidades, una molécula de trans-empalme descrita en la presente memoria incluye, unida de forma operativa en una dirección 5' a 3': una región no traducida 5', una secuencia de dominio de codificación (por ejemplo, un dominio de codificación, por ejemplo, una secuencia que codifica una secuencia funcional de exones *HTT*, por ejemplo, una secuencia funcional de exones *HTT* 5' al sitio de unión), un dominio de empalme (por ejemplo, un sitio de empalme), un dominio de unión, un dominio de unión, una región descendente 3' y un dominio terminador.

En algunas modalidades, las moléculas de trans-empalme del ácido nucleico descritas en la presente memoria están configuradas para corregir al menos una mutación en un alelo del gen *HTT* en un sujeto ubicada en una región 5' del gen *HTT* (por ejemplo, una región 5' al intrón 1, 2 o 3) uniéndose al intrón objetivo 1, 2 o 3 y mediando el trans-empalme de un dominio de codificación que tiene una secuencia funcional de exones *HTT* 5' a un exón *HTT* endógeno 3' al intrón objetivo. Este trans-empalme repara el gen *HTT* defectuoso en la célula objetivo de un individuo reemplazando el o los exones defectuosos y eliminando la parte defectuosa del pre-ARNm objetivo, produciendo un ARNm de *HTT* de tipo salvaje capaz de transcribir una proteína *HTT* funcional en la célula.

HTT

Un gen *HTT* al que se dirige una molécula de trans-empalme descrita en este documento puede contener una o múltiples mutaciones asociadas con HD. Se proporciona una secuencia *HTT* humana de ejemplo como secuencia de referencia del Centro Nacional de Información Biotecnológica (NCBI): NG_009378. Además de las secuencias publicadas, también se incluyen todas las correcciones obtenidas posteriormente o las secuencias variantes conservadoras de origen natural y no causantes de enfermedades que se encuentran en la población humana o de otros mamíferos. También se incluyen reemplazos de nucleótidos conservadores adicionales o aquellos que causan optimizaciones de codones. Las secuencias proporcionadas por los números de acceso a la base de datos también pueden utilizarse para

buscar secuencias homólogas en el mismo organismo mamífero o en otro.

Se prevé que las secuencias de ácido nucleico *HTT* y las proteínas resultantes expresadas puedan tolerar ciertas modificaciones menores a nivel de ácido nucleico, que incluyen, por ejemplo, modificaciones en las bases de nucleótidos que son silenciosas con respecto al aminoácido codificado. En otras modalidades, se prevén modificaciones de bases de ácidos nucleicos que cambian los aminoácidos, por ejemplo, para mejorar la expresión del péptido/proteína resultante. En algunas modalidades, se prevé la modificación de las variaciones alélicas, causadas por la degeneración natural del código genético.

También se incluyen como modificaciones de los genes *HTT* los análogos o versiones modificadas de las secuencias de aminoácidos codificadas. Normalmente, dichos análogos difieren de las proteínas identificadas específicamente en solo uno a cuatro cambios de codón. Los reemplazos conservadores son aquellos que tienen lugar dentro de una familia de aminoácidos que están relacionados en sus cadenas laterales y propiedades químicas.

La secuencia del ácido nucleico de un gen *HTT* funcional puede derivarse de cualquier mamífero que exprese de forma nativa *HTT* funcional o un homólogo de este. En otras modalidades, se realizan ciertas modificaciones a la secuencia del gen *HTT* para mejorar la expresión en la célula objetivo. Estas modificaciones pueden incluir la optimización por codones.

Como se describe en la presente anteriormente, la HD es causada por la expansión de repeticiones de trinucleótidos CAG en el exón 1 en exceso de 35 repeticiones CAG (penetración incompleta) o en exceso de 40 repeticiones CAG (aparición juvenil o adulta) en el gen *HTT*, que se heredan de manera autosómica dominante. Ver la Figura 1. Las composiciones que comprenden moléculas de trans-empalme descritas en este documento pueden corregir las repeticiones CAG expandidas en el exón 1, independientemente de cuántas repeticiones estén presentes, porque las moléculas de trans-empalme reemplazan la totalidad del exón 1 patógeno del gen *HTT*.

Dominios de codificación

En algunas modalidades, el dominio codificante de una molécula de trans-empalme 5' incluye todos los exones *HTT* (por ejemplo, exones *HTT* funcionales) que están en 5' con respecto al intrón *HTT* objetivo (por ejemplo, intrón *HTT* 1 [(SEQ ID NO: 348; *HTT* 5' UTR + exón 1, que incluye 21 repeticiones CAG (dentro del rango normal de número de repeticiones) o puede comprender 35-39 repeticiones CAG (penetración incompleta) o 40+ repeticiones CAG (HD de inicio en la edad adulta o de inicio en la juventud)], intrón 2 (SEQ ID NO: 365) y/o intrón 3 (SEQ ID NO: 366)). Por ejemplo, en modalidades en las que una molécula de trans-empalme 5' se dirige al intrón 2 de *HTT*, el dominio de codificación puede incluir los exones *HTT* funcionales 1-2. En algunas modalidades, los exones *HTT* funcionales 1-2 están codificados por una

secuencia que comprende cualquiera de las SEQ ID NO: 59, 349, 350 o 351. En algunas modalidades, los exones *HTT* funcionales 1-2 están codificados por una secuencia que comprende cualquiera de las SEQ ID NO: 59, 349, 350 o 351, que comprende además un codón de iniciación ATG en el extremo terminal 5'. En algunas modalidades, el dominio de unión se une al intrón 2 y el dominio de codificación incluye los exones *HTT* funcionales 1-2.

En algunas modalidades en las que una molécula de trans-empalme 5' se dirige al intrón 3 de *HTT*, el dominio de codificación puede incluir los exones *HTT* funcionales 1-3. En algunas modalidades, los exones *HTT* funcionales 1-3 están codificados por una secuencia que comprende cualquiera de las SEQ ID NO: 157, 352 o 353. En algunas modalidades, los exones *HTT* funcionales 1-3 están codificados por una secuencia que comprende las SEQ ID NO: 157, 352 o 353, que comprende además un codón de iniciación ATG en el extremo terminal 5'. En algunas modalidades, el dominio de unión se une al intrón 3 y el dominio de codificación incluye los exones *HTT* funcionales 1-3.

En algunas modalidades, donde una molécula de trans-empalme 5' se dirige al intrón 1 de *HTT*, el dominio codificante puede incluir el exón 1 de *HTT* funcional. En algunas modalidades, el exón 1 de *HTT* funcional está codificado por una secuencia que comprende SEQ ID NO: 3. En algunas modalidades, el exón 1 de *HTT* funcional está codificado por una secuencia que comprende SEQ ID NO: 3, que además comprende un codón de inicio ATG en el extremo terminal 5'. En algunas modalidades, el dominio de unión se une al intrón 1 y el dominio de codificación incluye el exón 1 *HTT* funcional.

En algunas modalidades, una secuencia de codificación de dominio codificante (por ejemplo, de un transgén que codifica un RTM) incluye ADNc de exones *HTT* (por ejemplo, exones *HTT*) para el reemplazo de exones *HTT* mutados. Por ejemplo, uno o más exones *HTT* funcionales dentro del dominio de codificación pueden ser una secuencia de ADNc. En algunas modalidades, todo el dominio de codificación es una secuencia de ADNc. De manera adicional o alternativa, todo o una parte del dominio de codificación, o uno o más exones *HTT* funcionales de este, puede ser una secuencia de origen natural (por ejemplo, una secuencia que tiene un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 %, un 99 % o un 100 % de identidad de secuencia con un exón *HTT* endógeno).

En algunas modalidades, la totalidad o una parte del dominio de codificación, o la secuencia de codificación de dominio codificante, o uno o más exones *HTT* funcionales de este, es una secuencia optimizada para codones en la que se ha modificado una secuencia del ácido nucleico, por ejemplo, para mejorar la expresión o la estabilidad, sin dar como resultado un cambio en el aminoácido codificado. La optimización por codones se puede realizar de una manera como la descrita en, por ejemplo, las patentes estadounidenses números 7.561.972, 7.561.973 o 7.888.112, cada una de las cuales se incorpora a la presente mediante referencia en su totalidad. Para la administración a través de un AAV recombinante, como se describe en

este documento, en una modalidad, el dominio codificante puede ser una secuencia de ácido nucleico de hasta 4000 bases de nucleótidos de longitud.

5 En algunas modalidades, se describe en la presente memoria una molécula de ARN que comprende, en una dirección de 5' a 3': (a) una secuencia de dominio de codificación de ADNc; (b) una secuencia donante de empalme; y (c) una secuencia de dominio de unión configurada para unirse a un intrón de una molécula de ARN endógena; donde la secuencia de dominio de codificación comprende una mutación de nucleótidos en relación con la secuencia de la molécula de ARN endógena, donde la mutación de nucleótidos altera un sitio de empalme
10 críptico dentro de la secuencia de dominio de codificación. En algunas modalidades, la mutación de nucleótido es una mutación de nucleótido sinónima. En algunas modalidades, el sitio de empalme críptico se identifica experimentalmente. En algunas modalidades, el sitio de empalme críptico se predice basándose en un análisis informático.

15 También se incluye en la presente memoria un método para modificar una molécula de ARN en una célula, dicho método comprende proporcionar a la célula una molécula de ARN exógena que comprende, en una dirección de 5' a 3': (a) una secuencia de dominio de codificación de ADNc que comprende una mutación de nucleótido que difiere de la de una molécula de ARN objetivo endógena en la célula; (b) una secuencia donante de empalme
20 configurada para empalmarse con una secuencia aceptora de empalme de la molécula de ARN objetivo endógena; y (c) una secuencia de dominio de unión configurada para unirse a un intrón de la molécula de ARN objetivo endógena; donde la mutación de nucleótido interrumpe un sitio de empalme críptico dentro de la secuencia de dominio de codificación de la molécula de ARN exógena. En algunas modalidades, la mutación de nucleótido es una
25 mutación de nucleótido sinónima. En algunas modalidades, el sitio de empalme críptico se identifica experimentalmente. En algunas modalidades, el sitio de empalme críptico se predice basándose en un análisis in silico.

30 También se incluye en la presente memoria un método para aumentar la eficiencia del trans-empalme de un editor de exones de ARN que comprende la introducción de una mutación en una secuencia de dominio de codificación del editor de exones de ARN, donde la mutación interrumpe un sitio de empalme críptico en la secuencia de dominio de codificación del editor de exones de ARN. En algunas modalidades, la mutación de nucleótido es una mutación de nucleótido sinónima. En algunas modalidades, el sitio de empalme críptico se identifica
35 experimentalmente. En algunas modalidades, el sitio de empalme críptico se predice basándose en un análisis in silico.

Los cambios de nucleótidos que mitigan el sitio de empalme críptico pueden incluir cambios que eliminan o reducen la capacidad de un sitio de empalme críptico para ser utilizado en una
40 reacción de empalme. Por ejemplo, un sitio de empalme críptico identificado en el contexto de un editor de exones de ARN generalmente comprende un sitio de empalme, un tracto de

polipirimidina y un punto de ramificación. En algunas modalidades, se pueden introducir uno o más cambios de nucleótidos en al menos uno de un sitio de empalme, un tracto de polipirimidina o un punto de ramificación, o cualquier combinación de estos de un sitio de empalme críptico identificado en el contexto de un editor de exones de ARN. En algunas modalidades, el cambio de nucleótido se determina de modo de minimizar el impacto potencial en una proteína codificada por el mismo. Un experto en la técnica observará que, si un cambio de nucleótido realizado para reducir la frecuencia de uso del sitio de empalme críptico también altera el aminoácido codificado por un ARN empalmado en trans, se preferirían cambios de aminoácidos conservadores a cambios de aminoácidos no conservadores. Además, una persona experta podría analizar fácilmente la secuencia y la estructura de la proteína con la mirada puesta en los dominios funcionales y las secuencias significativas en ellos para evaluar si podría esperarse razonablemente que dichos cambios alteren la función de una proteína codificada por una proteína empalmada en trans. Un experto en la técnica también podría probar una proteína que comprenda dicho cambio de aminoácido para determinar si se altera la actividad biológica utilizando ensayos conocidos en la técnica. En algunas modalidades, se cambia más de un nucleótido dentro de un sitio de empalme críptico identificado. En algunas circunstancias, la determinación de cuántos nucleótidos se deben cambiar se realiza empíricamente basándose en predicciones *in silico* y/o resultados experimentales. En algunas modalidades, se pueden introducir una o más mutaciones sinónimas (también denominadas en la presente memoria como al menos una) en al menos uno de un sitio de empalme, un tracto de polipirimidina o un punto de ramificación, o cualquier combinación de estos de un sitio de empalme críptico identificado en el contexto de un editor de exones de ARN. Las mutaciones sinónimas no alteran la secuencia de aminoácidos de una proteína codificada por un ARN cortado y empalmado en trans. En algunas modalidades, se pueden introducir más de una mutación sinónima en al menos uno de un sitio de empalme, una polipirimidina o un punto de ramificación, o cualquier combinación de estos de un sitio de corte y empalma identificado en el contexto de un editor de exones de ARN.

Además de lo anterior, se analizaron los resultados experimentales y la información de la secuencia en general de la siguiente manera. Los cambios para eliminar los sitios de empalme crípticos identificados experimentalmente se realizaron buscando y reemplazando ciertos elementos de los sitios aceptores de empalme. Los sitios AG (y más fuertemente los sitios CAG) al final de un sitio aceptor de empalme fueron priorizados para la introducción de cambios de nucleótidos. Si no se encontró un sitio AG o no se pudo cambiar sin introducir una mutación no sinónima, entonces se escaneó la secuencia 42 a 4 pares de bases delante del sitio de empalme en busca de puntos de ramificación (secuencias que coincidían con YNAH). Cualquier secuencia de puntos de ramificación identificada se analizan y se consideran para la introducción de una o más mutaciones de nucleótidos para reducir el uso del sitio de empalme críptico en el sitio de empalme críptico identificado experimentalmente. Además, también se escanean las secuencias para detectar la presencia de tractos de polipirimidina (múltiples Y inmediatamente antes del extremo AG). Normalmente, dichos tractos de

polipirimidina comprenden al menos 5 pirimidinas dentro de 10 pares de bases delante del sitio de empalme. Una vez identificados, dichos trectos de polipiridimidina se analizan luego y se consideran para la introducción de una o más mutaciones de nucleótidos para reducir el uso del sitio de empalme críptico en el sitio de empalme críptico identificado de forma experimental.

5 En algunas modalidades, el sitio de empalme críptico, o sitio de empalme fuera del objetivo, que se cambia para mitigar el empalme fuera del objetivo es un sitio que se ha identificado empíricamente como un sitio de empalme fuera del objetivo. Dichos sitios pueden identificarse, por ejemplo, utilizando técnicas descritas en el Ejemplo 6 de WO 2023/220742, que se
10 incorpora al presente documento como referencia en su totalidad. En algunas modalidades, todos los sitios de empalme críptico que tienen una frecuencia de uso superior a un umbral predeterminado se modifican mediante cambios de nucleótidos que mitigan el sitio de empalme críptico.

15 En algunas modalidades, el sitio de empalme críptico, o el sitio de empalme fuera del objetivo, que se modifica para mitigar el corte y empalme fuera del objetivo es un sitio que se predicho que es un sitio de empalme fuera del sitio. Tales predicciones se pueden realizar basándose en el análisis de secuencia para identificar un sitio de empalme canónico, un tracto de polipirimidina y/o un punto de ramificación de un supuesto sitio de empalme críptico en este.

20 En algunas modalidades, un cambio de nucleótido que mitiga el sitio de empalme críptico hace que una secuencia de nucleótidos que coincide con una secuencia de consenso del sitio de empalme canónico ya no coincida con la secuencia canónica. En algunas modalidades, un cambio de nucleótido que mitiga el sitio de empalme críptico elimina un posible nucleótido del
25 sitio de empalme. En algunas modalidades, un cambio de nucleótido que mitiga el sitio de empalme críptico elimina un posible nucleótido de tracto de polipirimidina. En algunas modalidades, un cambio de nucleótido que mitiga el sitio de empalme críptico elimina un posible nucleótido punto de ramificación. En algunas modalidades, un cambio de nucleótido que mitiga el sitio de empalme críptico es un cambio de nucleótido sinónimo. En algunas
30 modalidades, un cambio de nucleótido que mitiga el sitio de empalme críptico provoca un cambio en un aminoácido codificado por el editor de exones. En algunas modalidades, el cambio de aminoácido es una sustitución de aminoácido conservadora.

35 Para abordar el potencial de auto empalme críptico (ya sea por cis-empalme de concatémoros de AAV o por trans-empalme intermolecular), los presentes inventores utilizaron una predicción *in silico* de sitios de auto empalme para mitigar dichos eventos no deseados. Ver, por ejemplo, la Tabla 1.

Tabla 1. Mitigación de sitios de auto empalme crípticos identificados dentro del Editor de exones original.

Región	Aceptor de empalme o donante	Cambio de mitigación de auto empalme	SEQ ID NO de la secuencia original	SEQ ID NO de la secuencia mitigada de auto empalme
5' UTR	Aceptor	G>T en base 20, GA>TT en base 60-61	136	192
3X FLAG	Aceptor	G>A en la base 51	4	193
En la parte anterior del terminador MALAT1	Aceptor	AG>CT en la base 2-3	195	194

Dominios de unión

Las moléculas de trans-empalme HTT descritas en la presente memoria presentan un dominio de unión (BD, por su sigla en inglés) configurado para ligarse/unirse a un intrón y/o exón HTT objetivo. En algunos casos, el intrón HTT objetivo es el intrón 2 HTT. En una modalidad, el dominio de unión es una secuencia del ácido nucleico que es al menos un 80 % complementaria a (por ejemplo, al menos un 85% complementaria a, al menos un 86% complementaria a, al menos un 87% complementaria a, al menos un 88% complementaria a, al menos un 89% complementaria a, al menos un 90% complementaria a, al menos un 91% complementaria a, al menos un 92% complementaria a, al menos un 93% complementaria a, al menos un 94% complementaria a, al menos un 95% complementaria a, al menos un 96% complementaria a, al menos un 97% complementaria a, al menos un 98% complementaria a, al menos un 99% complementaria a, o 100% complementaria a) una secuencia del pre-ARNm del intrón HTT objetivo (por ejemplo, un intrón HTT objetivo), que puede suprimir el cis-empalme objetivo endógeno mientras mejora el trans-empalme entre la molécula de trans-empalme y el pre-ARNm de HTT objetivo (por ejemplo, mediante la creación de una molécula quimérica que tiene una parte de ARNm de HTT endógeno y un dominio de codificación que tiene uno o más exones HTT funcionales que codifican las secuencias de aminoácidos HTT de tipo salvaje). En una modalidad que implica secuencias de codificación de moléculas de trans-empalme (por ejemplo, vectores que codifican moléculas de trans-empalme), la secuencia de codificación del dominio de unión codifica una secuencia del ácido nucleico que es al menos un 80 % complementaria a (por ejemplo, al menos un 85 % complementaria, al menos un 86 % complementaria, al menos un 87 % complementaria, al menos un 88 % complementaria, al menos un 89 % complementaria, al menos un 90 % complementaria, al menos un 91 % complementaria, al menos un 92 % complementaria, al menos un 93 % complementaria, al menos un 94 % complementaria, al menos un 95 % complementaria, al menos un 96 % complementaria, al menos un 97 % complementaria, al menos un 98 % complementaria, al menos un 99 % complementaria, o un 100 % complementaria a) una secuencia del pre-ARNm del intrón HTT objetivo.

En algunos casos, la presente invención proporciona moléculas de trans-empalme (o vectores de las mismas) que se unen a *HTT* en el intrón 2, por ejemplo, en donde la molécula de trans-empalme de ácido nucleico está configurada para trans-empalmar un dominio codificante al exón 3 de *HTT* endógeno. En particular, las moléculas de trans-empalme descritas en este documento incluyen aquellas en las que el dominio de unión se une a un sitio de unión que tiene uno o más (por ejemplo, seis o más, ocho o más, diez o más, o doce o más) de los nucleótidos 1 a 200, 1500 a 2500 o 10 500 a 12 251 de SEQ ID NO: 57.

En algunos casos, el sitio de unión incluye cualesquiera seis o más nucleótidos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT* (por ejemplo, ocho o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT*, diez o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT*, 12 o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT*, 20 o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT*, 30 o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT*, 40 o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT*, 50 o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT*, 100 o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT*, 150 o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT*, 200 o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT*, o 250 o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 200, 1500 a 2500, o 10 500 a 12 251 del intrón 2 de *HTT*).

En algunos casos, el sitio de unión incluye cualesquiera seis o más nucleótidos consecutivos dentro de los nucleótidos 10 000 a 12 251 del intrón 2 de *HTT* (por ejemplo, cualesquiera ocho o más ácidos nucleicos consecutivos dentro de los nucleótidos 10 000 a 12 251 del intrón 2 de *HTT*, cualesquiera diez o más ácidos nucleicos consecutivos dentro de los nucleótidos 10 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 12 o más ácidos nucleicos consecutivos dentro de los nucleótidos 10 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 20 o más ácidos nucleicos consecutivos dentro de los nucleótidos 10 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 30 o más ácidos nucleicos consecutivos dentro de los nucleótidos 10 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 40 o más ácidos nucleicos consecutivos dentro de los nucleótidos 10 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 50 o más ácidos nucleicos consecutivos dentro de los nucleótidos 10 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 100 o más ácidos nucleicos consecutivos dentro de los nucleótidos 10 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 150 o más ácidos nucleicos consecutivos dentro de los nucleótidos 10 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 200 o más ácidos nucleicos consecutivos dentro de los nucleótidos 10 000

a 12 251 del intrón 2 de *HTT*, o cualesquiera 250 o más ácidos nucleicos consecutivos dentro de los nucleótidos 10 000 a 12 251 del intrón 2 de *HTT*).

5 En algunos casos, el sitio de unión incluye cualesquiera seis o más nucleótidos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT* (por ejemplo, cualesquiera ocho o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT*, cualesquiera diez o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 12 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 20 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 30 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 40 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 50 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 100 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 150 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 200 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT*, cualesquiera 250 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 000 a 12 251 del intrón 2 de *HTT*).

20 En algunos casos, el sitio de unión incluye cualesquiera seis o más nucleótidos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT* (por ejemplo, cualesquiera ocho o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT*, diez o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT*, 12 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT*, 20 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT*, 30 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT*, 40 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT*, 50 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT*, 100 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT*, 150 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT*, 200 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT*, o 250 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 950 a 12 251 del intrón 2 de *HTT*).

En algunos casos, un dominio de unión tiene al menos dos secuencias no superpuestas con al menos un 80 % de complementariedad con el sitio de unión.

40 En algunos casos, el dominio de unión incluye una secuencia del ácido nucleico que tiene al menos un 80 % de identidad (por ejemplo, al menos un 81 %, al menos un 82 %, al menos un

83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 60-81.

En algunas modalidades, el dominio de unión es una secuencia de ADN que tiene al menos un 80 % de identidad (por ejemplo, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 60, 62, o 67-81.

En algunos casos, el dominio de unión incluye una secuencia del ácido nucleico que tiene al menos un 80 % de identidad (al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 60 o 67-81.

En algunas modalidades, el dominio de unión es una secuencia de ADN que tiene al menos un 80 % de identidad (por ejemplo, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 67 o 70-81.

En algunos casos, el dominio de unión incluye una secuencia del ácido nucleico que tiene al menos un 80 % de identidad (al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera

de las SEQ ID NO: 71-74 o 77-79.

En algunos casos, el intrón *HTT* *objetivo* es el intrón 3 *HTT*. En una modalidad, el dominio de unión es una secuencia del ácido nucleico que es al menos un 80 % complementaria a (por ejemplo, al menos un 85% complementaria a, al menos un 86% complementaria a, al menos un 87% complementaria a, al menos un 88% complementaria a, al menos un 89% complementaria a, al menos un 90% complementaria a, al menos un 91% complementaria a, al menos un 92% complementaria a, al menos un 93% complementaria a, al menos un 94% complementaria a, al menos un 95% complementaria a, al menos un 96% complementaria a, al menos un 97% complementaria a, al menos un 98% complementaria a, al menos un 99% complementaria a, o 100% complementaria a) una secuencia del pre-ARNm del intrón *HTT* objetivo (por ejemplo, un intrón *HTT* objetivo), que puede suprimir el cis-empalme objetivo endógeno mientras mejora el trans-empalme entre la molécula de trans-empalme y el pre-ARNm de *HTT* objetivo (por ejemplo, mediante la creación de una molécula quimérica que tiene una parte de ARNm de *HTT* endógeno y un dominio de codificación que tiene uno o más exones *HTT* funcionales que codifican las secuencias de aminoácidos *HTT* de tipo salvaje). En una modalidad que implica secuencias de codificación de moléculas de trans-empalme (por ejemplo, vectores que codifican moléculas de trans-empalme), la secuencia de codificación del dominio de unión codifica una secuencia del ácido nucleico que es al menos un 80 % complementaria a (por ejemplo, al menos un 85 % complementaria, al menos un 86 % complementaria, al menos un 87 % complementaria, al menos un 88 % complementaria, al menos un 89 % complementaria, al menos un 90 % complementaria, al menos un 91 % complementaria, al menos un 92 % complementaria, al menos un 93 % complementaria, al menos un 94 % complementaria, al menos un 95 % complementaria, al menos un 96 % complementaria, al menos un 97 % complementaria, al menos un 98 % complementaria, al menos un 99 % complementaria, o un 100 % complementaria a) una secuencia del pre-ARNm del intrón *HTT* objetivo.

En algunos casos, la presente invención proporciona moléculas de trans-empalme (o vectores de estas) que se unen a *HTT* en el intrón 3, por ejemplo, en donde la molécula de trans-empalme de ácido nucleico está configurada para transempalmar un dominio codificante al exón 4 endógeno de *HTT*. En particular, las moléculas de trans-empalme descritas en el presente documento incluyen aquellas en las que el dominio de unión se une a un sitio de unión que tiene uno o más (por ejemplo, seis o más, ocho o más, diez o más, o doce o más) de los nucleótidos 3100 a 4429 de SEQ ID NO: 155.

En algunos casos, el sitio de unión incluye cualesquiera seis o más nucleótidos consecutivos dentro de los nucleótidos 3100 a 4429 del intrón 3 de *HTT* (por ejemplo, cualesquiera ocho o más ácidos nucleicos consecutivos dentro de los nucleótidos 3100 a 4429 del intrón 3 de *HTT*, diez o más ácidos nucleicos consecutivos dentro de los nucleótidos 3100 a 4429 del intrón 3 de *HTT*, 12 o más ácidos nucleicos consecutivos dentro de los nucleótidos 3100 a 4429 del

intrón 3 de *HTT*, 20 o más ácidos nucleicos consecutivos dentro de los nucleótidos 3100 a 4429 del intrón 3 de *HTT*, 30 o más ácidos nucleicos consecutivos dentro de los nucleótidos 3100 a 4429 del intrón 3 de *HTT*, 40 o más ácidos nucleicos consecutivos dentro de los nucleótidos 3100 a 4429 del intrón 3 de *HTT*, 50 o más ácidos nucleicos consecutivos dentro de los nucleótidos 3100 a 4429 del intrón 3 de *HTT*, 100 o más ácidos nucleicos consecutivos dentro de los nucleótidos 3100 a 4429 del intrón 3 de *HTT*, 150 o más ácidos nucleicos consecutivos dentro de los nucleótidos 3100 a 4429 del intrón 3 de *HTT*, 200 o más ácidos nucleicos consecutivos dentro de los nucleótidos 3100 a 4429 del intrón 3 de *HTT*, o 250 o más ácidos nucleicos consecutivos dentro de los nucleótidos 3100 a 4429 del intrón 3 de *HTT*).

En algunos casos, el sitio de unión incluye cualesquiera seis o más nucleótidos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT* (por ejemplo, cualesquiera ocho o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT*, diez o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT*, 12 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT*, 20 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT*, 30 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT*, 40 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT*, 50 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT*, 100 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT*, 150 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT*, 200 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT*, o 250 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4429 del intrón 3 de *HTT*).

En algunos casos, el sitio de unión incluye cualesquiera seis o más nucleótidos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT* (por ejemplo, cualesquiera ocho o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT*, diez o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT*, 12 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT*, 20 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT*, 30 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT*, 40 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT*, 50 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT*, 100 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT*, 150 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT*, 200 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT*, o 250 o más ácidos nucleicos consecutivos dentro de los nucleótidos 4100 a 4388 del intrón 3 de *HTT*).

En algunos casos, un dominio de unión tiene al menos dos secuencias no superpuestas con

al menos un 80 % de complementariedad con el sitio de unión.

En algunas modalidades, el dominio de unión es una secuencia de ADN que tiene al menos un 80 % de identidad (por ejemplo, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 158-174.

En algunos casos, el dominio de unión incluye una secuencia del ácido nucleico que tiene al menos un 80 % de identidad (por ejemplo, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 164-174.

En algunas modalidades, el dominio de unión es una secuencia de ADN que tiene al menos un 80 % de identidad (por ejemplo, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 170-174.

En algunas modalidades, el dominio de unión es una secuencia de ADN que tiene al menos un 80 % de identidad (por ejemplo, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 171-172.

Un dominio de unión puede estar unido de forma operativa 3' a un dominio de empalme (por ejemplo, conectado directamente a un dominio de empalme o tener secuencias intermedias que conecten el extremo 3' del dominio de empalme y el extremo 5' del dominio de unión).

En algunos casos, el intrón *HTT* *objetivo* es el intrón 1 *HTT*. En una modalidad, el dominio de unión es una secuencia del ácido nucleico que es al menos un 80 % complementaria a (por ejemplo, al menos un 85% complementaria a, al menos un 86% complementaria a, al menos un 87% complementaria a, al menos un 88% complementaria a, al menos un 89% complementaria a, al menos un 90% complementaria a, al menos un 91% complementaria a, al menos un 92% complementaria a, al menos un 93% complementaria a, al menos un 94% complementaria a, al menos un 95% complementaria a, al menos un 96% complementaria a, al menos un 97% complementaria a, al menos un 98% complementaria a, al menos un 99% complementaria a, o 100% complementaria a) una secuencia del pre-ARNm del intrón *HTT* objetivo (por ejemplo, un intrón *HTT* objetivo), que puede suprimir el cis-empalme objetivo endógeno mientras mejora el trans-empalme entre la molécula de trans-empalme y el pre-ARNm de *HTT* objetivo (por ejemplo, mediante la creación de una molécula quimérica que tiene una parte de ARNm de *HTT* endógeno y un dominio de codificación que tiene uno o más exones *HTT* funcionales que codifican las secuencias de aminoácidos *HTT* de tipo salvaje). En una modalidad que implica secuencias de codificación de moléculas de trans-empalme (por ejemplo, vectores que codifican moléculas de trans-empalme), la secuencia de codificación del dominio de unión codifica una secuencia del ácido nucleico que es al menos un 80 % complementaria a (por ejemplo, al menos un 85 % complementaria, al menos un 86 % complementaria, al menos un 87 % complementaria, al menos un 88 % complementaria, al menos un 89 % complementaria, al menos un 90 % complementaria, al menos un 91 % complementaria, al menos un 92 % complementaria, al menos un 93 % complementaria, al menos un 94 % complementaria, al menos un 95 % complementaria, al menos un 96 % complementaria, al menos un 97 % complementaria, al menos un 98 % complementaria, al menos un 99 % complementaria, o un 100 % complementaria a) una secuencia del pre-ARNm del intrón *HTT* objetivo.

En algunos casos, la presente invención proporciona moléculas de trans-empalme (o vectores de estas) que se unen a *HTT* en el intrón 1, por ejemplo, en donde la molécula de trans-empalme de ácido nucleico está configurada para transempalmar un dominio codificante al exón 2 endógeno de *HTT*. En particular, las moléculas de trans-empalme descritas en el presente documento incluyen aquellas en las que el dominio de unión se une a un sitio de unión que tiene uno o más (por ejemplo, seis o más, ocho o más, diez o más, o doce o más) de los nucleótidos 1 a 1000 o 11 500 a 11 850 de SEQ ID NO: 1.

En algunos casos, el sitio de unión incluye cualesquiera seis o más nucleótidos consecutivos dentro de los nucleótidos 1 a 1000 o 11 500 a 11 850 del intrón 1 de *HTT* (por ejemplo, ocho o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 1000 o 11 500 a 11 850 del intrón 1 de *HTT*, diez o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 1000 o 11 500 a 11 850 del intrón 1 de *HTT*, 12 o más ácidos nucleicos consecutivos dentro de los nucleótidos 1 a 1000 o 11 500 a 11 850 del intrón 1 de *HTT*, 20 o más ácidos nucleicos

a 11 850 del intrón 1 de *HTT*, o cualesquiera 250 o más ácidos nucleicos consecutivos dentro de los nucleótidos 11 650 a 11 850 del intrón 1 de *HTT*).

5 En algunos casos, un dominio de unión tiene al menos dos secuencias no superpuestas con al menos un 80 % de complementariedad con el sitio de unión.

10 En algunas modalidades, el dominio de unión es una secuencia de ADN que tiene al menos un 80 % de identidad (por ejemplo, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 8 o 14-21.

15 En algunos casos, el dominio de unión incluye una secuencia del ácido nucleico que tiene al menos un 80 % de identidad (por ejemplo, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 16-21.

25 En algunas modalidades, el dominio de unión es una secuencia de ADN que tiene al menos un 80 % de identidad (por ejemplo, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 17, 18, 20, o 21.

35 En algunas modalidades, el dominio de unión es una secuencia de ADN que tiene al menos un 80 % de identidad (por ejemplo, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 86 %, al menos un 87 %, al menos un 88 %, al menos un 89 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad; por ejemplo, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % de identidad) con cualquiera de las SEQ ID NO: 17 o 18.

Un dominio de unión puede estar unido de forma operativa 3' a un dominio de empalme (por ejemplo, conectado directamente a un dominio de empalme o tener secuencias intermedias que conecten el extremo 3' del dominio de empalme y el extremo 5' del dominio de unión).

5

Como se detalla en la presente memoria, la primera etapa del diseño del editor de exones *HTT* comenzó con la evaluación y selección de una secuencia BD altamente eficiente, que es complementaria al intrón pre-ARNm objetivo. Como se muestra en este documento, los BD que se dirigen al intrón 2 de *HTT* y al intrón 3 de *HTT* exhibieron los niveles más altos de trans-empalme y, por lo tanto, se han identificado como elementos efectivos de editores de exones dirigidos a *HTT* ejemplares para el tratamiento de la población de pacientes con HD. Ver las Figuras 5, 10, 12, 24 y 25.

10

Dominios de empalme

15

Para los editores de exones 5', el dominio de empalme puede incluir un sitio donante de empalme (sitio de empalme 5') para mediar el trans-empalme.

20

Para los editores de exones 3', el dominio de empalme puede incluir un sitio de empalme, un punto de ramificación y/o un tracto de polipirimidina (PPT) para mediar el trans-empalme. En algunas modalidades, un dominio de empalme tiene un único sitio de empalme, lo que denota que el sitio de empalme está diseñado para un trans-empalme preferencial, pero no para un cis-empalme, debido a la falta de un sitio de empalme correspondiente.

25

Un experto en la técnica podrá seleccionar dominios de empalme alternativos de acuerdo con métodos y principios conocidos. En una modalidad, la secuencia de consenso del sitio de empalme 5' es la secuencia del ácido nucleico AG/GURAGU (donde / indica el sitio de empalme). En otra modalidad, los sitios de empalme endógenos que corresponden al exón y al intrón proximales al sitio de empalme se pueden emplear para mantener cualquier señal reguladora de empalme.

30

En una modalidad, un sitio de empalme 5' adecuado comprende GTAAGT o GUAAGT.

35

Un dominio de empalme puede estar unido de forma operativa 5' a un dominio terminador (por ejemplo, conectado directamente a un dominio terminador o tener secuencias intermedias que conecten el extremo 3' del dominio de empalme y el extremo 5' del dominio terminador, por ejemplo, un dominio enlazador y/o un dominio de unión y/o una secuencia 3' posterior).

Región no traducida 5'

40

En algunos casos, la molécula de trans-empalme de ácido nucleico incluye una región no traducida 5'. En algunas modalidades, la región no traducida 5' comprende, consiste esencialmente, o consiste en una región no traducida 5' de *HTT* nativa. En algunas

modalidades, la región no traducida 5' comprende una secuencia que tiene al menos un 80 % de identidad de secuencia (por ejemplo, al menos un 85 % de identidad de secuencia, al menos un 86 % de identidad de secuencia, al menos un 87 % de identidad de secuencia, al menos un 88 % de identidad de secuencia, al menos un 89 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 91 % de identidad de secuencia, al menos un 92 % de identidad de secuencia, al menos un 93 % de identidad de secuencia, al menos un 94 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 96 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de identidad de secuencia, al menos un 99 % de identidad de secuencia o un 100 % de identidad de secuencia) con la SEQ ID NO: 136.

En algunas modalidades, una región no traducida 5' puede estar unida de forma operativa 5' a un dominio de codificación (por ejemplo, conectada directamente a un dominio de codificación o tener secuencias intermedias que conecten el extremo 3' de la región no traducida 5' y el extremo 5' del dominio de codificación).

En algunas modalidades, la molécula de trans-empalme de ácido nucleico no incluye una región no traducida de 5' o no incluye una región no traducida de 5' *HTT* nativa.

20 *Dominios reguladores 5'*

En algunos casos, la molécula de trans-empalme de ácido nucleico está unida de forma operativa a un dominio regulador 5' unido de forma operativa 5' al dominio de codificación (por ejemplo, unido directamente al dominio de codificación, o unido a través de un dominio intermedio, por ejemplo, una región no traducida). Un dominio regulador 5' puede incluir un promotor (por ejemplo, un promotor constitutivo, por ejemplo, el promotor CMV o un promotor EF1-alfa). En algunos casos, el dominio regulador 5' incluye un promotor (por ejemplo, un promotor constitutivo, por ejemplo, el promotor CMV/CMV) unido de forma operativa a una región no traducida *HTT* 5' nativa. En algunas modalidades, el dominio regulador 5' unido de forma operativa a una región no traducida 5' de *HTT* nativa comprende una secuencia que tiene al menos un 80 % de identidad de secuencia (por ejemplo, al menos un 85 % de identidad de secuencia, al menos un 86 % de identidad de secuencia, al menos un 87 % de identidad de secuencia, al menos un 88 % de identidad de secuencia, al menos un 89 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 91 % de identidad de secuencia, al menos un 92 % de identidad de secuencia, al menos un 93 % de identidad de secuencia, al menos un 94 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 96 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de identidad de secuencia, al menos un 99 % de identidad de secuencia, o un 100 % de identidad de secuencia) con la SEQ ID NO: 137.

En algunas modalidades, el dominio regulador 5' unido de forma operativa a una región no traducida 5' de *HTT* nativa comprende una secuencia que tiene al menos un 80 % de identidad

de secuencia (por ejemplo, al menos un 85 % de identidad de secuencia, al menos un 86 % de identidad de secuencia, al menos un 87 % de identidad de secuencia, al menos un 88 % de identidad de secuencia, al menos un 89 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 91 % de identidad de secuencia, al menos un 92 % de identidad de secuencia, al menos un 93 % de identidad de secuencia, al menos un 94 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 96 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de identidad de secuencia, al menos un 99 % de identidad de secuencia, o un 100 % de identidad de secuencia) con la SEQ ID NO: 196.

En algunas modalidades, el promotor CMV se reemplaza por un promotor CAGGS, en donde el promotor CAGGS se utiliza para impulsar la expresión de un editor de exones de ARN descrito en este documento. Ver, por ejemplo, la Figura 27. Como se muestra allí, al comparar CAGGS 5' UTR con o sin HTT 5' UTR, la traducción de proteínas parece estar regulada a través de HTT 5' UTR cuando CAGGS 5' UTR + HTT 5' UTR están unidos de forma operativa. Cuando se elimina el HTT 5' UTR, dejando solo el CAGGS 5' UTR, los presentes inventores observaron una expresión de proteína más fuerte, actividad que puede deberse a algún elemento que regula positivamente la traducción en el CAGGS 5' UTR. Los experimentos de transducción presentados aquí, en donde se introdujeron editores de exones de ARN en las células a través de un AAV, también demostraron que el promotor CAGGs impulsa la expresión significativa de editores de exones de ARN, moduladores de empalme MSH3 y construcciones codificantes de miARN. Ver, por ejemplo, las Figuras 55, 56, 58 y 59.

En algunas modalidades, un dominio regulador 5' puede unirse de forma operativa 5' a un dominio de codificación (por ejemplo, directamente conectado a un dominio de codificación o tienen secuencias intermedias que conectan el extremo 3' del dominio regulador 5' y el extremo 5' del dominio de codificación).

En algunas modalidades, se pueden utilizar otros tipos de promotores. En algunas modalidades, la molécula de trans-empalme de ácido nucleico no está unida operativamente a un promotor CMV o a un promotor CAGGS.

Dominios enlazadores

Como se analiza en la presente memoria, aumentar la eficiencia del trans-empalme sigue siendo un objetivo importante para la implementación de moléculas de trans-empalme del ácido nucleico como agentes terapéuticos. En un esfuerzo por diseñar moléculas de trans-empalme de ácidos nucleicos que tengan eficiencias de trans-empalme que puedan satisfacer los requisitos para su uso en una intervención terapéutica, los presentes inventores probaron una pluralidad de secuencias enlazadoras entre el dominio de empalme (SD) y el dominio de unión (BD) e identificaron una serie de enlazadores ejemplares que mejoraron el rendimiento respecto al de un enlazador 40mer hasta un grado estadísticamente significativo. Ver, por

ejemplo, la Figura 8.

Además de lo anterior, las moléculas de trans-empalme del ácido nucleico pueden incluir un dominio enlazador en una o más posiciones con la molécula. En algunas modalidades, el dominio enlazador está unido de forma operativa 3' al dominio de empalme o al sitio de empalme (por ejemplo, conectado directamente al dominio de empalme o al sitio de empalme). El dominio enlazador puede tener cualquier tamaño adecuado. En algunas modalidades, el dominio enlazador tiene más de 20 nucleótidos de longitud (por ejemplo, entre 20 y 100 nucleótidos de longitud o entre 20 y 85 nucleótidos de longitud). En algunos casos, el dominio enlazador comprende, consiste esencialmente o consiste en una secuencia del ácido nucleico que tiene al menos un 80 % de identidad (por ejemplo, al menos un 85 % de identidad de secuencia, al menos un 86 % de identidad de secuencia, al menos un 87 % de identidad de secuencia, al menos un 88 % de identidad de secuencia, al menos un 89 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 91 % de identidad de secuencia, al menos un 92 % de identidad de secuencia, al menos un 93 % de identidad de secuencia, al menos un 94 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 96 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de identidad de secuencia, al menos un 99 % de identidad de secuencia, o un 100 % de identidad de secuencia) con cualquiera de las SEQ ID NO: 37-46 o 106-112.

Con frecuencia se incluye una secuencia de enlace en las moléculas de trans-empalme, en donde se ubica entre el donante de empalme y el dominio de unión para ofrecer flexibilidad y accesibilidad a cada elemento. Como se describe en la presente memoria, se evaluaron las contribuciones funcionales de diferentes elementos en diferentes combinaciones para determinar la actividad conferida por ellos en el contexto de moléculas de trans-empalme. En algunas modalidades, la molécula de trans-empalme de ácido nucleico no incluye ninguna de las SEQ ID NO: 37-46 o 106-112, o puede estar ausente una secuencia de enlace.

Los dominios enlazadores descritos en este documento también pueden incluirse en modalidades de editores de exones de ARN que no se dirigen a *HTT* o *MSH3*. Las modalidades divulgadas en este documento incluyen editores de exones de ARN que incluyen un dominio de unión que se dirige a un pre-ARNm endógeno, una secuencia de dominio codificante que codifica una secuencia de aminoácidos funcional y un dominio enlazador entre la secuencia del dominio codificante y el dominio de unión, en donde el dominio enlazador comprende uno o más de las SEQ ID NO: 37-46 o 106-112.

Dominios terminadores de transcripción 3'

En algunos casos, la molécula de trans-empalme incluye un dominio terminador de la transcripción 3'. En algunas modalidades, dichos dominios terminadores de transcripción 3' forman una estructura helicoidal triple que tapa de forma efectiva el extremo 3' de la molécula

- de trans-empalme. En algunos casos, el dominio terminador de la transcripción 3' proviene del ARN largo no de codificación humano MALAT1 (por ejemplo, MALAT1 de tipo salvaje). En algunas modalidades, el dominio terminador de la transcripción 3' incluye un dominio similar al ARNt. Los dominios terminadores de transcripción 3' útiles como parte de las presentes moléculas de trans-empalme *HTT* se describen en la publicación de patente internacional n.º WO 2020/214973, que se incorpora en la presente memoria a modo de referencia en su totalidad. Por ejemplo, en algunas modalidades, la región de un RTM operativamente unida al extremo 3' del dominio de unión incluye un dominio terminador que comprende, consiste esencialmente en o consiste en un dominio MALAT1+mascRNA de tipo salvaje, como SEQ ID NO: 5. En algunas modalidades, la región de un editor de exones de ARN operativamente unida al extremo 3' del dominio de unión incluye un dominio terminador que comprende, consiste esencialmente en o consiste en un dominio MALAT1+masc ARN mutado (ARN anti-masc Mut1), como SEQ ID NO: 6.
- En algunas modalidades, la molécula de trans-empalme de ácido nucleico no incluye un dominio terminador de transcripción 3' o no incluye un terminador de transcripción derivado de MALAT1.

Dominios de empalme

- En algunas modalidades, los editores de exones de ARN ejemplares descritos en este documento incluyen aquellos que comprenden dominios de unión que se unen al intrón 1 de *HTT*, en donde dichos editores de exones de ARN ejemplares de unión al intrón 1 pueden comprender cualquiera de las SEQ ID NO: 23-36 y 47-56, secuencias que comprenden la secuencia codificante (por ejemplo, SEQ ID NO: 3), dominio de empalme, dominio enlazador, dominio de unión y terminador de transcripción 3'). En algunas modalidades, la secuencia codificante comprende además un codón de inicio ATG en el extremo 5'.

- En algunas modalidades, los editores de exones de ARN ejemplares descritos en este documento incluyen aquellos que comprenden dominios de unión que se unen al intrón 2 de *HTT*, en donde dichos editores de exones de ARN ejemplares de unión al intrón 2 pueden comprender cualquiera de las SEQ ID NO: 83-105 y 113-125, secuencias que comprenden la secuencia codificante (por ejemplo, SEQ ID NO: 59, 349, 350, o 351), dominio de empalme, dominio enlazador, dominio de unión y terminador de transcripción 3'). En algunas modalidades, la secuencia codificante comprende además un codón de inicio ATG en el extremo 5'.

- En algunas modalidades, los editores de exones de ARN ejemplares descritos en este documento incluyen aquellos que comprenden dominios de unión que se unen al intrón 3 de *HTT*, en donde dichos editores de exones de ARN ejemplares de unión al intrón 3 pueden comprender cualquiera de las SEQ ID NO: 175-191, secuencias que comprenden la secuencia codificante (por ejemplo, cualquiera de las SEQ ID NO: 157, 352, o 353), dominio de empalme,

dominio enlazador, dominio de unión y terminador de transcripción 3'). En algunas modalidades, la secuencia codificante comprende además un codón de inicio ATG en el extremo 5'.

5 En algunas modalidades, los editores de exón de ARN ejemplares descritos en el presente documento incluyen aquellos que comprenden dominios de unión que se unen al intrón 1 de *HTT* y dominios de unión que se unen a un intrón de *MSH3*, en los que dichos editores de exón de ARN híbridos/duales *HTT/MSH3* ejemplares pueden comprender cualquiera de las SEQ ID NO: 149-154, cuyas secuencias comprenden una secuencia codificante (por ejemplo, SEQ ID
10 NO: 3), dominio de empalme, dominio enlazador, dominio de unión y terminador de transcripción 3'. En algunas modalidades, la secuencia codificante comprende además un codón de inicio ATG en el extremo 5'. En algunas modalidades, los editores de exones de ARN *HTT/MSH3* híbridos/duales ejemplares comprenden un dominio de unión capaz de unirse tanto a un intrón *HTT* como a un intrón *MSH3*.

15 En algunas modalidades, los editores de exón de ARN ejemplares descritos en el presente documento incluyen aquellos que comprenden dominios de unión que se unen al intrón 2 de *HTT* y dominios de unión que se unen a un intrón de *MSH3*, en los que dichos editores de exón de ARN híbridos/duales *HTT/MSH3* ejemplares pueden comprender cualquiera de las SEQ ID NO: 212-223, cuyas secuencias comprenden la secuencia codificante (por ejemplo, SEQ ID
20 NO: 59), dominio de empalme, dominio enlazador, dominio de unión y terminador de transcripción 3'). En algunas modalidades, la secuencia codificante comprende además un codón de inicio ATG en el extremo 5'. En algunas modalidades, los editores de exones de ARN *HTT/MSH3* híbridos/duales ejemplares comprenden un dominio de unión capaz de unirse tanto
25 a un intrón *HTT* como a un intrón *MSH3*.

En algunas modalidades, un editor de exones de ARN ejemplar incluye un dominio de unión que se une a un intrón *HTT* (por ejemplo, intrón 2) y un dominio de unión que se une a un intrón *MSH3* (por ejemplo, intrón 5 o intrón 15). En algunas modalidades, el dominio de unión
30 de *MSH3* comprende intron5_213_100 (SEQ ID NO: 140), intron5_188_150 (SEQ ID NO: 209), intron15_6523_120 (SEQ ID NO: 144) o intron15_6498_150 (SEQ ID NO: 210). En algunas modalidades, el dominio de unión *HTT* comprende *HTT_intron2_12061_150* (SEQ ID NO: 95). En algunas modalidades, el dominio de unión de *MSH3* está en 5' con respecto al dominio de unión de *HTT*. En algunas modalidades, el dominio de unión de *MSH3* está en 3' con respecto
35 al dominio de unión de *HTT*. En algunas modalidades, un terminador MALAT1 se encuentra entre el dominio de unión *MSH3* y el dominio de unión *HTT*. En algunas modalidades, un terminador MALAT1 no se encuentra entre el dominio de unión *MSH3* y el dominio de unión *HTT*. En algunas modalidades, un terminador MALAT1 se encuentra a 3' tanto del dominio de unión *MSH3* como del dominio de unión *HTT*.

40 En algunas modalidades, un editor de exones de ARN ejemplar incluye un dominio de unión

que se une a un intrón *MSH3* (por ejemplo, intrón 5 o intrón 15). En algunas modalidades, el dominio de unión de *MSH3* comprende intron5_213_100 (SEQ ID NO: 140), intron5_188_150 (SEQ ID NO: 209), intron15_6523_120 (SEQ ID NO: 144) o intron15_6498_150 (SEQ ID NO: 210). En algunas modalidades, un editor de exones de ARN dirigido a *MSH3* comprende además cualquier dominio de empalme divulgado en este documento, cualquier secuencia 3X UBS descrita en este documento, cualquier elemento rico en AU descrito en este documento, cualquier dominio enlazador descrito en este documento y/o cualquier secuencia terminadora divulgada en este documento. En algunas modalidades, se administra un editor de exones de ARN dirigido a *MSH3* junto con un editor de exones dirigido a *HTT*. En algunas modalidades, un editor de exones dirigido a *MSH3* no se usa junto con un editor de exones dirigido a *HTT*.

En algunas modalidades, la expresión de *MSH3* se reduce mediante un miARN dirigido al ARNm de *MSH3*. En algunas modalidades, el pri-miARN comprende uno o más de mir-30a [andamio 5' (SEQ ID NO: 227); andamio 3' (SEQ ID NO: 228); bucle (SEQ ID NO: 229)], mir155 [andamio 5' (SEQ ID NO: 230); andamio 3' (SEQ ID NO: 231); bucle (SEQ ID NO: 232)], o mir-33 [andamio 5' (SEQ ID NO: 259), andamio 3' (SEQ ID NO: 260); bucle (SEQ ID NO: 261)]. En algunas modalidades, la pri-miARN comprende una o más de las SEQ ID NO: 234, 235, 238-241, o 262-269. En algunas modalidades, la secuencia activa de miARN comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO: 224, 244, 246, 248, 250, 252, 254, 256 o 257 o una secuencia al menos 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% idéntica a cualquiera de las SEQ ID NO: 224, 244, 246, 248, 250, 252, 254, 256 o 257. El miARN dirigido al ARNm *MSH3* se puede utilizar en combinación con cualquiera de los editores de exones de ARN descritos en este documento, incluidos, por ejemplo, los editores de exones de ARN que se dirigen a *HTT*, *MSH3* o ambos. El miARN dirigido al ARNm *MSH3* se puede utilizar sin ningún editor de exones de ARN descrito en este documento, pero en su lugar se puede utilizar de forma independiente para reducir la expresión de *MSH3*. El miARN dirigido al ARNm *MSH3* pueden usarse en un método para tratar o prevenir un trastorno de expansión de repeticiones de trinucleótidos.

En algunas modalidades, la expresión de *MSH3* se reduce mediante ARN antisentido basado en ARNnp, que puede inducir la omisión de exones durante el procesamiento del pre-ARNm. En algunas modalidades, la construcción de ARNnp comprende uno o más de los siguientes: SEQ ID NO: 274, que se dirige a la unión entre el intrón 1 y el exón 2 de *MSH3*; SEQ ID NO: 275, que se dirige a la unión entre el exón 2 y el intrón 2 de *MSH3*; SEQ ID NO: 278, 301 o 303, que se dirigen a la unión entre el intrón 2 y el exón 3 de *MSH3*; SEQ ID NO: 279, 300 o 302, que se dirigen a la unión entre el exón 3 y el intrón 3 de *MSH3*; SEQ ID NO: 281, que se dirige a la unión entre el intrón 3 y el exón 4 de *MSH3*; SEQ ID NO: 282, que se dirige a la unión entre el exón 4 y el intrón 4 de *MSH3*; SEQ ID NO: 306 o 308, que se dirigen a la unión entre el intrón 5 y el exón 6 de *MSH3*; SEQ ID NO: 305 o 307, que se dirigen a la unión entre el exón 6 y el intrón 6 de *MSH3*; SEQ ID NO: 311 o 313, que se dirigen a la unión entre el intrón 6 y el exón 7 de *MSH3*; SEQ ID NO: 310 o 312, que se dirigen a la unión entre el exón

7 y el intrón 7 de *MSH3*; SEQ ID NO: 316 o 318, que se dirigen a la unión entre el intrón 7 y el exón 8 de *MSH3*; SEQ ID NO: 315 o 317, que se dirigen a la unión entre el exón 8 y el intrón 8 de *MSH3*; SEQ ID NO: 321 o 323, que se dirigen a la unión entre el intrón 14 y el exón 15 de *MSH3*; o SEQ ID NO: 320 o 322, que se dirigen a la unión entre el exón 15 y el intrón 15 de *MSH3*.

En algunas modalidades, se utilizan uno, dos, tres, cuatro, cinco o más de las construcciones de ARNas anteriores (por ejemplo, se administran a un paciente) en conjunto. Por ejemplo, en algunas modalidades, se utilizan conjuntamente una construcción dirigida a la unión intrón 1-exón 2 y una construcción dirigida a la unión exón 2-intrón 2; se utilizan conjuntamente una construcción dirigida a la unión intrón 2-exón 3 y una construcción dirigida a la unión exón 3-intrón 3; se utilizan conjuntamente una construcción dirigida a la unión intrón 3-exón 4 y una construcción dirigida a la unión exón 4-intrón 4; se utilizan conjuntamente una construcción dirigida a la unión intrón 5-exón 6 y una construcción dirigida a la unión exón 6-intrón 6; se utilizan conjuntamente una construcción dirigida a la unión intrón 6-exón 7 y una construcción dirigida a la unión exón 7-intrón 7; se utilizan conjuntamente una construcción dirigida a la unión intrón 7-exón 8 y una construcción dirigida a la unión exón 8-intrón 8; o se utilizan conjuntamente una construcción dirigida a la unión del intrón 14 con el exón 15 y/o una construcción dirigida a la unión del exón 15 con el intrón 15.

En algunas modalidades, se utiliza una única construcción de ARNas que se dirige a dos uniones intrón-exón. En algunas modalidades, la construcción de ARNas incluye una secuencia que es al menos parcialmente complementaria a una secuencia de exón completa más una porción de secuencia intrónica a cada lado del exón, tal como, por ejemplo, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 nucleótidos de secuencia intrónica a cada lado de la secuencia de exón. Estas construcciones pueden tener como objetivo, por ejemplo, las uniones intrón-exón en cada lado de cualquiera de los exones 2, 3, 4, 5, 6, 7, 8 o 15 de *MSH3*. En algunas modalidades, la construcción de ARNas es una construcción U7SmOPT que comprende SEQ ID NO: 324 (secuencia In3/Ex3/In2; comprende SEQ ID NO: 299), que se dirige a toda la longitud del exón 3 y las uniones flanqueantes a cada lado. En algunas modalidades, la construcción de ARNas es una construcción U7SmOPT que comprende SEQ ID NO: 327 (secuencia In6/Ex6/In5; comprende SEQ ID NO: 304), que se dirige a toda la longitud del exón 6 y las uniones flanqueantes a cada lado. En algunas modalidades, la construcción de ARNas es una construcción U7SmOPT que comprende SEQ ID NO: 330 (secuencia In7/Ex7/In6; comprende SEQ ID NO: 309), que se dirige a toda la longitud del exón 7 y las uniones flanqueantes a cada lado. En algunas modalidades, la construcción de ARNas es una construcción U7SmOPT que comprende SEQ ID NO: 333 (secuencia In8/Ex8/In7; comprende SEQ ID NO: 314), que se dirige a toda la longitud del exón 8 y las uniones flanqueantes a cada lado. En algunas modalidades, la construcción de ARNas es una construcción U7SmOPT que comprende SEQ ID NO: 336 (secuencia In15/Ex15/In14; comprende SEQ ID NO: 319), que se dirige a toda la longitud del exón 15 y las uniones

flanqueantes a cada lado.

En algunas modalidades, se utiliza una única construcción de ARNas que se dirige a dos uniones intrón-exón. En algunas modalidades, la construcción de ARNas incluye una
 5 secuencia que es al menos parcialmente complementaria a una secuencia de unión intrón-exón 5', una secuencia que es al menos parcialmente complementaria a una secuencia de unión intrón-exón 3' y un enlazador no estructurado que une estas dos secuencias. Estas construcciones pueden tener como objetivo, por ejemplo, las uniones intrón-exón en cada lado de los exones 2, 3, 4, 5, 6, 7, 8 o 15 de *MSH3*. En algunas modalidades, la construcción de
 10 ARNas es una construcción U7SmOPT que comprende la SEQ ID NO: 325 [U7SmOPT In3/Ex3 (SEQ ID NO: 300) + enlazador + Ex3/In2 (SEQ ID NO: 301)], que se dirige a las uniones a ambos lados del exón 3. En algunas modalidades, la construcción de ARNas es una construcción U2 ARNnp que comprende la SEQ ID NO: 326 [U2 In3/Ex3 (SEQ ID NO: 302) + enlazador + Ex3/In2 (SEQ ID NO: 303)], que se dirige a las uniones a ambos lados del exón
 15 3. En algunas modalidades, la construcción de ARNas es una construcción U7SmOPT que comprende la SEQ ID NO: 328 [U7SmOPT In6/Ex6 (SEQ ID NO: 305) + enlazador + Ex6/In5 (SEQ ID NO: 306)], que se dirige a las uniones a ambos lados del exón 6. En algunas modalidades, la construcción de ARNas es una construcción de ARNnp U2 que comprende SEQ ID NO: 329 [U2 In6/Ex6 (SEQ ID NO: 307) + enlazador + Ex6/In5 (SEQ ID NO: 308)], que
 20 se dirige a las uniones a ambos lados del exón 6. En algunas modalidades, la construcción de ARNas es una construcción U7SmOPT que comprende SEQ ID NO: 331 [U7SmOPT In7/Ex7 (SEQ ID NO: 310) + enlazador + Ex7/In6 (SEQ ID NO: 311)], que se dirige a las uniones a ambos lados del exón 7. En algunas modalidades, la construcción de ARNas es una construcción de ARNnp U2 que comprende SEQ ID NO: 332 [U2 In7/Ex7 (SEQ ID NO: 312) + enlazador + Ex7/In6 (SEQ ID NO: 313)], que se dirige a las uniones a ambos lados del exón
 25 7. En algunas modalidades, la construcción de ARNas es una construcción U7SmOPT que comprende SEQ ID NO: 334 [U7SmOPT In8/Ex8 (SEQ ID NO: 315) + enlazador + Ex8/In7 (SEQ ID NO: 316)], que se dirige a las uniones a ambos lados del exón 8. En algunas modalidades, la construcción de ARNas es una construcción de ARNnp U2 que comprende SEQ ID NO: 335 [U2 In8/Ex8 (SEQ ID NO: 317) + enlazador + Ex8/In7 (SEQ ID NO: 318)], que se dirige a las uniones a ambos lados del exón 8. En algunas modalidades, la construcción de ARNas es una construcción U7SmOPT que comprende SEQ ID NO: 337 [U7SmOPT In15/Ex15 (SEQ ID NO: 320) + enlazador + Ex15/In14 (SEQ ID NO: 321)], que se dirige a las uniones a ambos lados del exón 15. En algunas modalidades, la construcción de ARNas es
 30 una construcción de U2 ARNnp que comprende SEQ ID NO: 338 [U2 In15/Ex15 (SEQ ID NO: 322) + enlazador + Ex15/In14 (SEQ ID NO: 323)], que se dirige a las uniones a cada lado del exón 15.

Algunas modalidades de construcciones moduladoras de empalme *MSH3* incluyen, unidas
 40 operativamente, una secuencia que codifica una secuencia de ARN nuclear pequeño (ARNnp) (por ejemplo, una secuencia U7 Sm OPT o una secuencia U2 ARNnp) y una secuencia que

codifica un ARN antisentido que promueve la omisión de exón de un exón objetivo del pre-ARNm *MSH3*. La omisión de exones puede introducir cambios de marco y/o codones de parada prematuros, que pueden inducir una desintegración mediada por sin sentido o perjudicar de otro modo la producción de MSH3 funcional. El ARN antisentido que promueve la omisión de exón de un exón objetivo puede apuntar a una o ambas de las uniones exón-intrón 5' y la unión exón-intrón 3' del exón objetivo del pre-ARNm *MSH3*. Tal como se utiliza en el presente documento, se dice que un ARN antisentido "se dirige" a una unión exón-intrón particular si tiene suficiente complementariedad con una secuencia que rodea la unión exón-intrón que es capaz de promover la omisión del exón objetivo durante el procesamiento del pre-ARNm. En algunas modalidades, una secuencia de ARN antisentido incluye un tramo contiguo de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45 o 50 nucleótidos que es 100 % complementario a un tramo contiguo de nucleótidos de la misma longitud en el pre-ARNm que incluye una unión exón-intrón. En algunas modalidades, una secuencia de ARN antisentido incluye un tramo contiguo de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45 o 50 nucleótidos que es al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% complementario a un tramo contiguo de nucleótidos de la misma longitud en el pre-ARNm que incluye una unión exón-intrón. En algunas modalidades, una secuencia de ARN antisentido incluye un tramo contiguo de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45 o 50 nucleótidos que es al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% complementario a un tramo contiguo de nucleótidos de la misma longitud en el pre-ARNm que incluye una unión exón-intrón o está dentro de 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 nucleótidos de una unión exón-intrón. Una persona experta comprenderá que cuando existe una complementariedad menor a la completa, la funcionalidad de omisión de exones se puede mejorar incrementando la longitud del tramo contiguo de nucleótidos parcialmente complementarios. Una persona experta también comprenderá que el contenido de GC de la secuencia que rodea una unión exón-intrón puede influir en la capacidad de una secuencia de ARN antisentido para apuntar eficazmente a la unión exón-intrón e inducir la omisión. Un mayor contenido de GC aumenta la resistencia de la hibridación, lo que puede conducir a un menor estiramiento requerido de complementariedad.

En algunas modalidades, un modulador de empalme MSH3 comprende una secuencia de ARN antisentido que se dirige a dos uniones exón-intrón. En algunas modalidades, una secuencia de ARN antisentido que se dirige a dos uniones exón-intrón comprende una secuencia que se une a la mayor parte o a la totalidad de una secuencia de exón completa de un pre-ARNm. En algunas modalidades, el ARN antisentido es 100% complementario a toda la secuencia de exones del pre-ARNm. En algunas modalidades, el ARN antisentido es al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% complementario a la secuencia de exones del pre-ARNm. En algunas modalidades, el ARN antisentido comprende además secuencias de nucleótidos que son al

menos parcialmente complementarias a las secuencias de intrones en la parte anterior y/o en la parte posterior de la secuencia de exón de pre-ARNm objetivo. En algunas modalidades, dichas secuencias de intrones tienen al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 nucleótidos de longitud o más. En algunas modalidades, la secuencia de ARN antisentido comprende una secuencia que es al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% complementaria a tales secuencias de intrones.

En algunas modalidades, una secuencia de ARN antisentido que se dirige a dos uniones exón-intrón de un exón objetivo comprende (a) dos secuencias no adyacentes que se anillan cada una a una secuencia que rodea una unión exón-intrón del exón objetivo y (b) una secuencia de enlace entre las dos secuencias no adyacentes que tiene un bajo grado de complementariedad (por ejemplo, menos del 50 %) con el exón objetivo. En algunas modalidades, la secuencia de enlace comprende al menos 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45 o 50 nucleótidos o de 10 a 30, 10 a 50, 15 a 25 o 18 a 22 nucleótidos. En algunas modalidades, el enlazador tiene menos de 70, 60, 50, 40, o 30 % de complementariedad con todos los tramos contiguos de la secuencia del exón objetivo que tengan la misma longitud que el enlazador.

En algunas modalidades, cualquiera de los moduladores de empalme MSH3 descritos en este documento se puede utilizar en combinación con cualquier editor de exones de ARN HTT descrito en este documento. En algunas modalidades, los moduladores de empalme MSH3 descritos en este documento no se utilizan en combinación con cualquier editor de exones de ARN descrito en este documento. Cualquiera de los moduladores de empalme MSH3 descritos en este documento se puede utilizar de forma independiente en métodos de tratamiento de trastornos de expansión de repeticiones de trinucleótidos.

En algunas modalidades, se codifican múltiples construcciones en un solo vector, como, por ejemplo, un vector AAV. Cualquiera de los editores de exones, secuencias de miARN y ARNas descritos en este documento se pueden combinar en un solo vector. En algunas modalidades, dos o más de un editor de exones, un miARN y una construcción de ARNas se codifican en un solo vector. En algunas modalidades, el editor de exones, miARN y/o ARNas, según corresponda, se dirigen al mismo gen. Por ejemplo, un solo vector puede codificar un editor de exones *HTT* para generar un ARNm *HTT* corregido y un miARN dirigido a *HTT* y/o un asARN dirigido a *HTT* para reducir la cantidad de *HTT* defectuoso. En algunas modalidades, un solo vector codifica un editor de exones *HTT* (por ejemplo, un editor de exones que comprende cualquiera de las SEQ ID NO: 83-104, SEQ ID NO: 113-125, SEQ ID NO: 175-191, SEQ ID NO: 199-206, o SEQ ID NO: 23-36) y un miARN dirigido a *HTT* (por ejemplo, SEQ ID NO: 339 y/o 342). En algunas modalidades, el vector comprende la secuencia establecida en cualquiera de las SEQ ID NO: 354 o 355 o una secuencia que tiene al menos 80%, 85%, 90%, 95%, 96%, 97%, 98% o 99% de identidad con dichas secuencias. En algunas modalidades, el editor de

exones, miARN y/o ARNas, según corresponda, se dirigen a diferentes genes. En algunas modalidades, por ejemplo, un solo vector codifica un editor de exones *HTT* para generar ARNm *HTT* corregido y un editor de exones *MSH3*, miARN *MSH3* y/o un ARNas *MSH3* (por ejemplo, una construcción de ARNnp) para inhibir la expresión de *MSH3*. En algunas modalidades, un solo vector codifica un editor de exones *HTT* (por ejemplo, un editor de exones que comprende cualquiera de las SEQ ID NO: 83-104, SEQ ID NO: 113-125, SEQ ID NO: 175-191, SEQ ID NO: 199-206, o SEQ ID NO: 23-36) y un ARNas dirigido a *MSH3* (por ejemplo, una construcción que comprende cualquiera una o más de las SEQ ID NO: 284-293 o 324-338). En algunas modalidades, el vector comprende la secuencia establecida en las SEQ ID NO: 356 o 357 o una secuencia que tiene al menos 80%, 85%, 90%, 95%, 96%, 97%, 98% o 99% de identidad con dichas secuencias. En algunas modalidades, el vector codifica un editor de exones *HTT* (por ejemplo, un editor de exones que comprende cualquiera de las SEQ ID NO: 83-104, SEQ ID NO: 113-125, SEQ ID NO: 175-191, SEQ ID NO: 199-206, o SEQ ID NO: 23-36), un ARNas dirigido a *MSH3* (por ejemplo, una construcción que comprende cualquiera una o más de las SEQ ID NO: 284-293 o 324-338), y un miARN que se dirige a *HTT* (por ejemplo, SEQ ID NO: 339 y/o 342). En algunas modalidades, el vector comprende la secuencia establecida en las SEQ ID NO: 358 o 359 o una secuencia que tiene al menos 80, 85, 90, 95, 96, 97, 98 o 99% de identidad con dichas secuencias.

Las diferentes modalidades terapéuticas descritas en este documento pueden combinarse en varias combinaciones, ya sea que estén codificadas en el mismo vector o en vectores diferentes. Por ejemplo, cualquiera de los editores de exones *HTT* descritos en este documento con un dominio de unión que se dirige al intrón 1 se puede combinar con cualquiera de las construcciones de ARNas dirigidos a *MSH3* descritos en este documento y/o cualquiera de las construcciones de miARN dirigidos a *HTT* descritos en este documento. En algunas modalidades, el editor de exones *HTT* con un dominio de unión que se dirige al intrón 1, el ARNas dirigido a *MSH3* y el miARN dirigido a *HTT* se incluyen en el mismo vector. A modo de otro ejemplo, cualquiera de los editores de exones *HTT* descritos en este documento con un dominio de unión que se dirige al intrón 2 se puede combinar con cualquiera de las construcciones de ARNas dirigidos a *MSH3* descritos en este documento y/o cualquiera de las construcciones de miARN dirigidos a *HTT* descritos en este documento. En algunas modalidades, el editor de exones *HTT* con un dominio de unión que se dirige al intrón 2, el ARNas dirigido a *MSH3* y el miARN dirigido a *HTT* se incluyen en el mismo vector. A modo de otro ejemplo, cualquiera de los editores de exones *HTT* descritos en este documento con un dominio de unión que se dirige al intrón 3 se puede combinar con cualquiera de las construcciones de ARNas dirigidos a *MSH3* descritos en este documento y/o cualquiera de las construcciones de miARN dirigidos a *HTT* descritos en este documento. En algunas modalidades, el editor de exones *HTT* con un dominio de unión que se dirige al intrón 3, el ARNas dirigido a *MSH3* y el miARN dirigido a *HTT* se incluyen en el mismo vector.

En algunas modalidades, la unión de una molécula de trans-empalme al pre-ARNm objetivo

está mediada por el porcentaje de complementariedad (es decir, en función de las características de apareamiento de bases de los ácidos nucleicos), la formación de triple hélice o la interacción proteína-ácido nucleico (como se describe en los documentos citados en la presente memoria). En una modalidad, la molécula de trans-empalme de ácido nucleico incluye moléculas de ADN, ARN o híbridas de ADN/ARN, donde el ADN o ARN es de cadena simple o de cadena doble. También se incluyen en la presente memoria ARN o ADN, que pueden hibridarse con uno de los ARN o ADN mencionados anteriormente, preferiblemente en condiciones estrictas, por ejemplo, a 60 °C en tampón SSC 2,5x y varios lavados a 37 °C en una concentración de tampón menor, por ejemplo, tampón SSC 0,5x. Estos ácidos nucleicos pueden codificar proteínas que muestran actividad de fosfato lipídico fosfatasa y/o asociación con membranas plasmáticas. Cuando se sintetizan moléculas de trans-empalme *in vitro*, dichas moléculas de trans-empalme se pueden modificar en la fracción base, la fracción de azúcar o la cadena principal de fosfato, por ejemplo, para mejorar la estabilidad de la molécula, la hibridación con el ARNm objetivo, el transporte a la célula, la estabilidad en las células a la escisión enzimática, etc. Por ejemplo, la modificación de una molécula de trans-empalme para reducir la carga general puede mejorar la captación celular de la molécula. Además, se pueden realizar modificaciones para reducir la susceptibilidad a la degradación química o por nucleasas. Las moléculas de ácido nucleico pueden sintetizarse de tal manera que puedan conjugarse con otra molécula, por ejemplo, un péptido, un agente de reticulación desencadenado por hibridación, un agente de transporte, un agente de escisión desencadenado por hibridación, etc.

Se pueden introducir otras modificaciones bien conocidas en las moléculas de ácido nucleico como medio para aumentar la estabilidad intracelular y la vida media (véase también más arriba para oligonucleótidos). Se conocen posibles modificaciones en la técnica. Las modificaciones que se pueden realizar a la estructura de las moléculas de trans-empalme sintéticas incluyen modificaciones de la cadena principal.

Ensayos de líneas celulares

En algunos casos, las moléculas de trans-empalme descritas en la presente memoria se prueban en líneas celulares cultivadas. Para examinar, seleccionar y mejorar la funcionalidad de los editores de exones de ARN, se pueden adquirir o diseñar líneas celulares cultivadas para expresar el pre-ARNm de *HTT* objetivo en un nivel suficiente.

Plataforma de detección de editores de exones de ARN

Como se describe en la presente, los editores de exones de ARN 5' comprenden varios elementos de secuencia funcionales, como un dominio de unión (BD) para el direccionamiento del pre-ARNm y un enlazador que permite el acceso al sitio donante de empalme (SD). Al diseñar un editor de exones de ARN para un gen objetivo determinado, se prueba una variedad de opciones de secuencia para cada uno de estos elementos para determinar su capacidad de contribuir a una alta eficiencia de trans-empalme (TS). Estas pruebas se pueden realizar

mediante (A) clonación y transfección de variantes individuales del editor de exones de ARN y análisis de la eficiencia mediante RT-qPCR/ddPCR y transferencia Western o (B) clonación y agrupación de editores de exones de ARN en un enfoque basado en biblioteca de alto rendimiento (HT) que se basa en la secuenciación de próxima generación (NGS) y el análisis computacional para evaluar la eficiencia. A continuación, se describen ambos enfoques.

Cribado del editor de exones de ARN en formato individual:

Este enfoque se puede aplicar para probar una cantidad pequeña de elementos variables dentro de una secuencia de editor de exones de ARN antes de iniciar una prueba multiplexada basada en biblioteca, para validar el rendimiento de los editores de exones de ARN identificados en una prueba multiplexada y/o para mejorar el rendimiento de los candidatos principales. La evaluación de la eficiencia del TS se realiza a nivel de ARN y proteína.

A nivel de ARN, la actividad de TS se evalúa mediante el aislamiento del ARN total de las células seguido de transcripción inversa y PCR cuantitativa en tiempo real (RT-qPCR) que mide el número de copias de ARN de, por ejemplo, los siguientes objetivos: RNF20 (gen de mantenimiento para normalización); ARNm natural (*HTT*); ARN editor de exones; ARN editado por exones (ONT) en el objetivo, que es el producto de TS positivo; ONT + Editor de exones + OFT (fuera del objetivo): un solo ensayo que captura estos tres *objetivos*. OFT representa moléculas de ARN incorrectas a las cuales el editor de exones de ARN puede trans-empalmar.

La eficiencia de TS de ONT, también denominada porcentaje de reemplazo, representa la parte de la población total de ARNm de *HTT* que ha experimentado una TS exitosa y se calcula mediante la siguiente ecuación: % de TS de ONT = $100 \times (\text{número de copias de ONT} / (\text{número de copias de ONT} + \text{número de copias naturales}))$.

La eficiencia TS del editor de exones de ARN es la parte de la población de transcripción del editor de exones de ARN que se ha cortado y trans-empalmado correctamente en el ARN de *HTT* y se calcula mediante la siguiente ecuación: % TS del editor de exones = $100 \times (\text{número de copias de ONT} / (\text{número de copias de ONT} + \text{número de copias del editor de exones} + \text{número de copias de OFT}))$.

A nivel de proteína, la actividad de TS se mide mediante análisis transferencia Western aplicado a proteínas extraídas de muestras de células o tejidos. La beta-actina, que es una proteína citoesquelética, o tubulina, puede utilizarse como control de carga. Para aquellas construcciones que incluyen una etiqueta, por ejemplo, una etiqueta FLAG en su extremo N, se puede utilizar un anticuerpo (Ab) específico para la etiqueta (por ejemplo, un Ab específico de FLAG) para sondear transferencias Western y evaluar los niveles de proteína ONT.

III. Vectores

Las moléculas de trans-empalme se pueden administrar a las células objetivo de un individuo

utilizando varias técnicas, por ejemplo, utilizando vectores de virus adenoasociados (AAV) recombinantes u otras modalidades de vectores, como vectores no virales. Por lo tanto, en la presente memoria se proporcionan vectores que comprenden/codifican moléculas de trans-empalme (por ejemplo, vectores virales o no virales que comprenden/codifican moléculas de trans-empalme, por ejemplo, vectores de ADN que comprenden/codifican moléculas de trans-empalme). Se puede utilizar cualquier vector de ácido nucleico adecuado junto con las presentes composiciones y métodos para diseñar y ensamblar los componentes de la molécula de trans-empalme y un AAV recombinante. En una modalidad, el vector es un AAV recombinante que lleva la molécula de trans-empalme impulsada por un promotor que expresa una molécula de trans-empalme en células seleccionadas de un individuo. Los métodos para el ensamblaje de los vectores recombinantes son conocidos en la técnica. Ver, por ejemplo, Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Nueva York, 1989; Kay, M. A. et al., *Nat. Medic*, 2001, 7(l):33-40; y Walther W. y Stein U., *Drugs* 2000, 60(2):249-71.

En determinadas modalidades descritas en la presente memoria, la molécula de trans-empalme se administra a las células seleccionadas, por ejemplo, células fotorreceptoras, que necesitan tratamiento por medio de un vector AAV. Hay una variedad de serotipos de AAV disponibles de origen natural. Existen muchas variantes naturales en la cápside de AAV, lo que permite la identificación y el uso de un AAV con propiedades específicamente adecuadas para las células neuronales. Los vectores AAV artificiales pueden diseñarse mediante técnicas de biología molecular convencionales, lo que hace posible optimizar estas partículas para la administración específica de células de las secuencias de ácidos nucleicos de moléculas de trans-empalme, para minimizar la inmunogenicidad, para ajustar la estabilidad y la vida útil de las partículas, para una degradación eficiente, para una administración precisa al núcleo, etc. Por ejemplo, dichas cápsides artificiales pueden generarse mediante cualquier técnica adecuada, utilizando una secuencia de AAV seleccionada (por ejemplo, un fragmento de una proteína de cápside vp1) en combinación con secuencias heterólogas que pueden obtenerse de un AAV seleccionado diferente, porciones no contiguas del mismo AAV, de una fuente viral que no sea AAV o de una fuente no viral. Un AAV artificial puede ser, de modo no taxativo, un AAV pseudotipado, una cápside de AAV quimérico, una cápside de AAV recombinante o una cápside de AAV «humanizada». Los vectores pseudotipados, en los que la cápside de un AAV se reemplaza con una proteína de cápside heteróloga, son útiles para administrar las moléculas de trans-empalme descritas en la presente memoria.

La expresión de las moléculas de empalme en trans descritas en la presente memoria se puede lograr en las células seleccionadas a través de la administración por AAV diseñados de forma recombinante o AAV artificiales que contienen secuencias que comprenden/codifican la molécula de empalme en trans deseada. El uso de AAV es un modo común de administración exógena de ADN, ya que es relativamente no tóxico, proporciona una transferencia genética eficiente y se puede optimizar fácilmente para propósitos específicos. Entre los serotipos bien

caracterizados de AAV aislados de primates humanos o no humanos, el serotipo humano 2 ha sido ampliamente utilizado para experimentos eficientes de transferencia genética en diferentes tejidos objetivo y modelos animales.

5 En algunas modalidades, el AAV es AAV1 o una variante de este (por ejemplo, SEQ ID NO: 6 o 64 de US20030138772 o SEQ ID NO: 11 o 27 de US20150159173), AAV2 o una variante de este (por ejemplo, SEQ ID NO: 7 o 70 de US20030138772, SEQ ID NO: 7 o 23 de US20150159173, o SEQ ID NO: 7 de US20150159173), AAV2G9 o una variante de este, AAV3 o una variante de este (por ejemplo, SEQ ID NO: 8 o 71 de US20030138772), AAV3a o una
10 variante de este, AAV3b o una variante de este (por ejemplo, SEQ ID NO: 1 y 10 de la patente estadounidense n.º 6,156,303), AAV3-3 o una variante de este (por ejemplo, SEQ ID NO: 200 y 217 de WO2005033321), AAV4 o una variante de este (por ejemplo, SEQ ID NO: 63 de US20030138772), AAV4-4 o una variante de este (por ejemplo, SEQ ID NO: 201 o 218 de WO2005033321), AAV5 o una variante de este (por ejemplo, SEQ ID NO: 114 de
15 US20030138772), AAV6 o una variante de este (por ejemplo, SEQ ID NO: 65 de US20030138772), AAV6.1 o una variante de este (por ejemplo, SEQ ID NO: 29 de US20150159173), AAV6.2 o una variante de este, AAV6.1.2 o una variante de este, AAV7 o una variante de este (por ejemplo, SEQ ID NO: 1-3 de US20030138772), AAV7.2 o una variante de este, AAV8 o una variante de este (por ejemplo, SEQ ID NO: 4 y 95 de
20 US20030138772 o AAV8(b) (que tiene la secuencia de aminoácidos de Pro-Glu-Arg-Thr-Ala-Met-Ser-Leu-Pro en las posiciones de aminoácidos 587-595 en comparación con el AAV8 de tipo salvaje, tal como se describe en la patente estadounidense n.º 9,567,376, que se incorpora a la presente mediante referencia en su totalidad)), AAV9 o una variante de este (por ejemplo, SEQ ID NO: 5 y 100 de US20030138772), AAV9.9 o una variante de este, AAV9.11 o una
25 variante de este, AAV9.13 o una variante de este, AAV9.16 o una variante de este, AAV9.24 o una variante de este, AAV9.45 o una variante de este, AAV9.47 o una variante de este, AAV9.61 o una variante de este, AAV9.68 o una variante de este, AAV9.84 o una variante de este (ver, por ejemplo, N. Pulicherla et al. Molecular Therapy 19(6):1070-1078 (2011), que se incorpora a la presente mediante referencia en su totalidad), AAV10 o una variante de este
30 (por ejemplo, SEQ ID NO: 117 de US20030138772), AAV11 o una variante de este (por ejemplo, SEQ ID NO: 118 de US20030138772), AAV12 o una variante de este (por ejemplo, SEQ ID NO: 119 de US20030138772), AAV16.3 o una variante de este, AAV24.1 o una variante de este, AAV27.3 o una variante de este, AAV42.12 o una variante de este, AAV42-1b o una variante de este, AAV42-2 o una variante de este, AAV42-3a o una variante de este, AAV42-3b o una variante de este, AAV42-4 o una variante de este, AAV42-5a o una variante de este, AAV42-5b o una variante de este, AAV42-6b o una variante de este, AAV42-8 o una
35 variante de este, AAV42-10 o una variante de este, AAV42-11 o una variante de este, AAV42-12 o una variante de este, AAV42-13 o una variante de este, AAV42-15 o una variante de este, AAV42-aa o una variante de este, AAV43-1 o una variante de este, AAV43-12 o una variante de este, AAV43-20 o una variante de este, AAV43-21 o una variante de este, AAV43-23 o una
40 variante de este, AAV43-25 o una variante de este, AAV43-5 o una variante de este, AAV44.1

o una variante de este, AAV44.2 o una variante de este, AAV44.5 o una variante de este, AAV223.1 o una variante de este, AAV223.2 o una variante de este, AAV223.4 o una variante de este, AAV223.5 o una variante de este, AAV223.6 o una variante de este, AAV223.7 o una variante de este, AAV1-7/rh.48 o una variante de este, AAV1-8/rh.49 o una variante de este, AAV2-15/rh.62 o una variante de este, AAV2-3/rh.61 o una variante de este, AAV2-4/rh.50 o una variante de este, AAV2-5/rh.51 o una variante de este, AAV3.1/hu.6 o una variante de este, AAV3.1/hu.9 o una variante de este, AAV3-9/rh.52 o una variante de este, AAV3-11/rh.53 o una variante de este, AAV4-8/rh.64 o una variante de este, AAV4-9/rh.54 o una variante de este (por ejemplo, SEQ ID NO: 116 de WO2005033321), AAV4-19/rh.55 o una variante de este (por ejemplo, SEQ ID NO: 117 de WO2005033321), AAV5-3/rh.57 o una variante de este, AAV5-22/rh.58 o una variante de este, AAV7.3/hu.7 o una variante de este, AAV16.8/hu.10 o una variante de este, AAV16.12/hu.11 o una variante de este, AAV29.3/bb.1 o una variante de este, AAV29.5/bb.2 o una variante de este, AAV106.1/hu.37 o una variante de este, AAV114.3/hu.40 o una variante de este, AAV127.2/hu.41 o una variante de este, AAV127.5/hu.42 o una variante de este, AAV128.3/hu.44 o una variante de este, AAV130.4/hu.48 o una variante de este, AAV145.1/hu.53 o una variante de este, AAV145.5/hu.54 o una variante de este, AAV145.6/hu.55 o una variante de este, AAV161.10/hu.60 o una variante de este, AAV161.6/hu.61 o una variante de este, AAV33.12/hu.17 o una variante de este, AAV33.4/hu.15 o una variante de este, AAV33.8/hu.16 o una variante de este, AAV52/hu.19 o una variante de este, AAV52.1/hu.20 o una variante de este, AAV58.2/hu.25 o una variante de este, AAVA3.3 o una variante de este, AAVA3.4 o una variante de este, AAVA3.5 o una variante de este, AAVA3.7 o una variante de este, AAVC1 o una variante de este, AAVC2 o una variante de este, AAVC5 o una variante de este, AAV-DJ o una variante de este (por ejemplo, SEQ ID NO: 2 o 3 de US20140359799), AAV-DJ8 o una variante de este, AAVF3 o una variante de este, AAVF5 o una variante de este, AAVH2 o una variante de este, AAVH6 o una variante de este, AAVLK03 o una variante de este, AAVH-1/hu.1 o una variante de este, AAVH-5/hu.3 o una variante de este, AAVLG-10/rh.40 o una variante de este, AAVLG-4/rh.38 o una variante de este, AAVLG-9/hu.39 o una variante de este, AAVN721-8/rh.43 o una variante de este, AAVCh.5 o una variante de este (por ejemplo, SEQ ID NO 46 de US20150159173), AAVCh.5R1 o una variante de este, AAVcy.2 o una variante de este, AAVcy.3 o una variante de este, AAVcy.4 o una variante de este, AAVcy.5 o una variante de este (por ejemplo, SEQ ID NO: 8 y 24 de US20150159173), AAVCy.5R1 o una variante de este, AAVCy.5R2 o una variante de este, AAVCy.5R3 o una variante de este, AAVCy.5R4 o una variante de este, AAVcy.6 o una variante de este, AAVhu.1 o una variante de este (por ejemplo, SEQ ID NO: 144 de WO2005033321), AAVhu.2 o una variante de este (por ejemplo, SEQ ID NO: 143 de WO2005033321), AAVhu.3 o una variante de este (por ejemplo, SEQ ID NO: 145 de WO2005033321), AAVhu.4 o una variante de este (por ejemplo, SEQ ID NO: 141 de WO2005033321), AAVhu.5 o una variante de este, AAVhu.6 o una variante de este (por ejemplo, SEQ ID NO: 84 de WO2005033321), AAVhu.7 o una variante de este (por ejemplo, SEQ ID NO: 150 de WO2005033321), AAVhu.9 o una variante de este (por ejemplo, SEQ ID NO: 155 de WO2005033321), AAVhu.10 o una variante de este (por ejemplo,

SEQ ID NO: 156 de WO2005033321), AAVhu.11 o una variante de este (por ejemplo, SEQ ID NO: 153 de WO2005033321), AAVhu.13 o una variante de este (SEQ ID NO: 16 y 32 de US20150159173), AAVhu.15 o una variante de este (por ejemplo, SEQ ID NO: 147 de WO2005033321), AAVhu.16 o una variante de este (por ejemplo, SEQ ID NO: 148 de WO2005033321), AAVhu.17 o una variante de este (por ejemplo, SEQ ID NO: 83 de WO2005033321), AAVhu.18 o una variante de este (por ejemplo, SEQ ID NO: 149 de WO2005033321), AAVhu.19 o una variante de este (por ejemplo, SEQ ID NO: 133 de WO2005033321), AAVhu.20 o una variante de este (por ejemplo, SEQ ID NO: 134 de WO2005033321), AAVhu.21 o una variante de este (por ejemplo, SEQ ID NO: 135 de WO2005033321), AAVhu.22 o una variante de este (por ejemplo, SEQ ID NO: 138 de WO2005033321), AAVhu.23.2 o una variante de este (por ejemplo, SEQ ID NO: 137 de WO2005033321), AAVhu.24 o una variante de este (por ejemplo, SEQ ID NO: 136 de WO2005033321), AAVhu.25 o una variante de este (por ejemplo, SEQ ID NO: 146 de WO2005033321), AAVhu.26 o una variante de este (por ejemplo, SEQ ID NO: 17 y 33 de US20150159173), AAVhu.27 o una variante de este (por ejemplo, SEQ ID NO: 140 de WO2005033321), AAVhu.28 o una variante de este (por ejemplo, SEQ ID NO: 42 de US20150159173), AAVhu.29 o una variante de este (por ejemplo, SEQ ID NO: 132 de WO2005033321), AAVhu.29R o una variante de este, AAVhu.31 o una variante de este (por ejemplo, SEQ ID NO: 121 de WO2005033321), AAVhu.32 o una variante de este (SEQ ID NO: 122 de WO2005033321), AAVhu.34 o una variante de este (por ejemplo, SEQ ID NO: 125 de WO2005033321), AAVhu.35 o una variante de este (por ejemplo, SEQ ID NO: 164 de WO2005033321), AAVhu.37 o una variante de este (por ejemplo, SEQ ID NO: 18 y 34 de US20150159173), AAVhu.39 o una variante de este (por ejemplo, SEQ ID NO: 102 de WO2005033321), AAVhu.40 o una variante de este (por ejemplo, SEQ ID NO: 87 de WO2005033321), AAVhu.41 o una variante de este (por ejemplo, SEQ ID NO: 91 de WO2005033321), AAVhu.42 o una variante de este (por ejemplo, SEQ ID NO: 85 de WO2005033321), AAVhu.43 o una variante de este (por ejemplo, SEQ ID NO: 160 de WO2005033321), AAVhu.44 o una variante de este (por ejemplo, SEQ ID NO: 45 de US20150159173), AAVhu.44R1 o una variante de este, AAVhu.44R2 o una variante de este, AAVhu.44R3 o una variante de este, AAVhu.45 o una variante de este (por ejemplo, SEQ ID NO: 127 de WO2005033321), AAVhu.46 o una variante de este (por ejemplo, SEQ ID NO: 159 de WO2005033321), AAVhu.47 o una variante de este (por ejemplo, SEQ ID NO: 128 de WO2005033321), AAVhu.48 o una variante de este (por ejemplo, SEQ ID NO: 38 de US20150159173), AAVhu.48R1 o una variante de este, AAVhu.48R2 o una variante de este, AAVhu.48R3 o una variante de este, AAVhu.49 o una variante de este (por ejemplo, SEQ ID NO: 189 de WO2005033321), AAVhu.51 o una variante de este (por ejemplo, SEQ ID NO: 190 de WO2005033321), AAVhu.52 o una variante de este (por ejemplo, SEQ ID NO: 191 de WO2005033321), AAVhu.53 o una variante de este (por ejemplo, SEQ ID NO: 19 y 35 de US20150159173), AAVhu.54 o una variante de este (por ejemplo, SEQ ID NO: 188 de WO2005033321), AAVhu.55 o una variante de este (por ejemplo, SEQ ID NO: 187 de WO2005033321), AAVhu.56 o una variante de este (por ejemplo, SEQ ID NO: 192 de

WO2005033321), AAVhu.57 o una variante de este (por ejemplo, SEQ ID NO: 193 de WO2005033321), AAVhu.58 o una variante de este (por ejemplo, SEQ ID NO: 194 de WO2005033321), AAVhu.60 o una variante de este (por ejemplo, SEQ ID NO: 184 de WO2005033321), AAVhu.61 o una variante de este (por ejemplo, SEQ ID NO: 185 de WO2005033321), AAVhu.63 o una variante de este (por ejemplo, SEQ ID NO: 195 de WO2005033321), AAVhu.64 o una variante de este (por ejemplo, SEQ ID NO: 196 de WO2005033321), AAVhu.66 o una variante de este (por ejemplo, SEQ ID NO: 197 de WO2005033321), AAVhu.67 o una variante de este (por ejemplo, SEQ ID NO: 198 de WO2005033321), AAVhu.14/9 o una variante de este, AAVhu.t 19 o una variante de este, AAVrh.2 o una variante de este (por ejemplo, SEQ ID NO: 39 de US20150159173), AAVrh.2R o una variante de este, AAVrh.8 o una variante de este (por ejemplo, SEQ ID NO: 41 de US20150159173), AAVrh.8R o una variante de este, AAVrh.10 o una variante de este (por ejemplo, SEQ ID NO: 9 y 25 de US20150159173), AAVrh.12 o una variante de este, AAVrh.13 o una variante de este (por ejemplo, SEQ ID NO: 10 y 26 de US20150159173), AAVrh.13R o una variante de este, AAVrh.14 o una variante de este, AAVrh.17 o una variante de este, AAVrh.18 o una variante de este, AAVrh.19 o una variante de este, AAVrh.20 o una variante de este (por ejemplo, SEQ ID NO: 1 de US20150159173), AAVrh.21 o una variante de este, AAVrh.22 o una variante de este, AAVrh.23 o una variante de este, AAVrh.24 o una variante de este, AAVrh.25 o una variante de este, AAVrh.31 o una variante de este, AAVrh.32 o una variante de este, AAVrh.33 o una variante de este, AAVrh.34 o una variante de este, AAVrh.35 o una variante de este, AAVrh.36 o una variante de este, AAVrh.37 o una variante de este (por ejemplo, SEQ ID NO: 40 de US20150159173), AAVrh.37R2 o una variante de este, AAVrh.38 o una variante de este (por ejemplo, SEQ ID NO: 86 de WO2005033321), AAVrh.39 o una variante de este (por ejemplo, SEQ ID NO: 3, 20, o 36 de US20150159173), AAVrh.40 o una variante de este (por ejemplo, SEQ ID NO: 92 de WO2005033321), AAVrh.43 o una variante de este (por ejemplo, SEQ ID NO: 21 y 37 de US20150159173), AAVrh.46 o una variante de este (por ejemplo, SEQ ID NO: 4 y 22 de US20150159173), AAVrh.48 o una variante de este (por ejemplo, SEQ ID NO: 44 de US20150159173), AAVrh.48.1 o una variante de este (por ejemplo, SEQ ID NO: 44 de US20150159173), AAVrh.48.1.2 o una variante de este, AAVrh.48.2 o una variante de este, AAVrh.49 o una variante de este (por ejemplo, SEQ ID NO: 103 de WO2005033321), AAVrh.50 o una variante de este (por ejemplo, SEQ ID NO: 108 de WO2005033321), AAVrh.51 o una variante de este (por ejemplo, SEQ ID NO: 104 de WO2005033321), AAVrh.52 o una variante de este (por ejemplo, SEQ ID NO: 96 de WO2005033321), AAVrh.53 o una variante de este (por ejemplo, SEQ ID NO: 97 de WO2005033321), AAVrh.54 o una variante de este (por ejemplo, SEQ ID NO: 49 de US20150159173), AAVrh.56 o una variante de este (por ejemplo, SEQ ID NO: 152 de WO2005033321), AAVrh.57 o una variante de este (por ejemplo, SEQ ID NO: 105 de WO2005033321), AAVrh.58 o una variante de este (por ejemplo, SEQ ID NO: 48 de US20150159173), AAVrh.61 o una variante de este (por ejemplo, SEQ ID NO: 107 de WO2005033321), AAVrh.62 o una variante de este (por ejemplo, SEQ ID NO: 114 de WO2005033321), AAVrh.64 o una variante de este (por ejemplo, SEQ ID NO: 43 de

US20150159173), AAVrh.64R1 o una variante de este, AAVrh.64R2 o una variante de este, AAVrh.67 o una variante de este (por ejemplo, SEQ ID NO: 47 de US20150159173), AAVrh.73 o una variante de este (por ejemplo, SEQ ID NO: 5 de US20150159173), o AAVrh.74 o una variante de este (por ejemplo, SEQ ID NO: 6 de US2015015917). Los ejemplos no taxativos de variantes incluyen las SEQ ID NO: 9, 27-45, 47-62, 66-69, 73-81, 84-94, 96, 97, 99 y 101-113 de US20030138772, cuyo contenido se incorpora aquí como referencia en su totalidad, y las SEQ ID NO: 1, 2, 4-82, 89, 90, 93-95, 98, 100, 101, 109-113, 118-120, 124, 126, 131, 139, 142, 151, 154, 158, 161, 162, 165-183, 202, 204-212, 215, 219 y 224-236 de WO2005033321, cuyo contenido se incorpora aquí como referencia en su totalidad. En una modalidad, el serotipo de AAV es cualquiera de los descritos en U.S. 2021/0189430, cuyo contenido se incorpora a la presente memoria a modo de referencia en su totalidad. La secuencia de aminoácidos del AAV puede incluir una o más sustituciones de aminoácidos en una proteína de la cápside del AAV en una o más posiciones que interactúan con un proteoglicano de sulfato de heparán o en una o más posiciones correspondientes a los aminoácidos 484, 487, 527, 532, 585 o 588, numeración basada en la numeración VP1 del AAV2.

A menos que se especifique lo contrario, las ITR de AAV y otros componentes de AAV seleccionados que se describen en la presente memoria pueden seleccionarse fácilmente entre cualquier serotipo de AAV, incluidos, entre otros, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 u otros serotipos de AAV conocidos y desconocidos. En una modalidad, las ITR son de AAV2. Aquellas ITR u otros componentes de AAV se pueden aislar o diseñar fácilmente utilizando técnicas disponibles para aquellos expertos en la técnica a partir de un AAV. Dicho AAV puede aislarse u obtenerse de fuentes académicas, comerciales o públicas (por ejemplo, la American Type Culture Collection, Manassas, VA). De manera alternativa, las secuencias de AAV se pueden obtener a través de medios sintéticos u otros medios mediante referencia a las secuencias publicadas como las que están disponibles en la bibliografía o en bases de datos como, por ejemplo, GenBank, PubMed, o similares.

Los fragmentos de AAV deseables para ensamblar en vectores incluyen las proteínas cap, incluidas vp1, vp2, vp3 y las regiones hipervariables, las proteínas rep, incluidas rep 78, rep 68, rep 52 y rep 40, y las secuencias que codifican estas proteínas. Estos fragmentos pueden utilizarse fácilmente en una variedad de sistemas vectoriales y células hospedadoras. Dichos fragmentos pueden utilizarse solos, en combinación con otras secuencias o fragmentos de serotipos de AAV, o en combinación con elementos de otras secuencias virales de AAV o no AAV. Como se utiliza en la presente memoria, los serotipos de AAV artificiales incluyen, de modo no taxativo, AAV con una proteína de cápside que no se encuentra en la naturaleza. Una cápside artificial de este tipo se puede generar mediante cualquier técnica adecuada, utilizando una secuencia de AAV seleccionada (por ejemplo, un fragmento de una proteína de la cápside vp1) en combinación con secuencias heterólogas que se pueden obtener de un serotipo de AAV seleccionado diferente, porciones no contiguas del mismo serotipo de AAV, de una fuente vírica no AAV o de una fuente no vírica. Un serotipo de AAV artificial puede ser, de modo no

taxativo, un AAV pseudotipado, una cápside de AAV quimérico, una cápside de AAV recombinante o una cápside de AAV “humanizada”. Los vectores pseudotipados, en los que se utiliza la cápside de un AAV con las ITR de un AAV que tiene una proteína de cápside diferente, son útiles como se describe en la presente memoria. En una modalidad, el AAV es AAV2/5 (es decir, un AAV que tiene ITR de AAV2 y una cápside de AAV5). En una modalidad, el AAV es AAV2/8 (es decir, un AAV que tiene ITR de AAV2 y una cápside de AAV8). En una modalidad, el AAV incluye una cápside AAV8. Dicha cápside de AAV8 incluye la secuencia de aminoácidos que se encuentra en la secuencia de referencia de NCBI: YP_077180.1. En otra modalidad, la cápside de AAV8 incluye una cápside codificada por los nt 2121 a 4337 del número de acceso de GenBank: AF513852.1.

En una modalidad, los vectores útiles en las composiciones y métodos que se describen en la presente memoria contienen, como mínimo, secuencias que codifican una cápside de serotipo de AAV seleccionado, por ejemplo, una cápside de AAV2, o un fragmento de esta. En otra modalidad, los vectores útiles contienen, como mínimo, secuencias que codifican una proteína rep del serotipo de AAV seleccionado, por ejemplo, la proteína rep AAV2, o un fragmento de esta. Opcionalmente, dichos vectores pueden contener proteínas cap y rep de AAV. En los vectores en los que se proporcionan tanto la representación de AAV como la de cap, las secuencias de representación de AAV y de AAV cap pueden ser ambas de un mismo origen de serotipo, por ejemplo, un origen de AAV2.

Como alternativa, se pueden utilizar vectores en los que las secuencias rep sean de un serotipo de AAV que difiere del que proporciona las secuencias cap. En una modalidad, las secuencias rep y cap se expresan a partir de fuentes separadas (por ejemplo, vectores separados o una célula hospedadora y un vector). En otra modalidad, estas secuencias rep se fusionan en marco con secuencias cap de un serotipo de AAV diferente para formar un vector de AAV quimérico, como los descritos en la patente estadounidense n.º 7.282.199, que se incorpora a modo de referencia en la presente memoria.

Un AAV recombinante (rAAV, por sus siglas en inglés) adecuado se genera cultivando una célula hospedadora que contiene una secuencia del ácido nucleico que codifica una proteína de la cápside del serotipo de AAV, o un fragmento de esta, como se define en la presente memoria; un gen rep funcional; un minigén compuesto, por ejemplo, de ITR de AAV y una secuencia del ácido nucleico de una molécula de empalme en trans; y suficientes funciones auxiliares para permitir el empaquetamiento del minigén en la proteína de la cápside de AAV. Los componentes necesarios para cultivarse en la célula hospedadora para empaquetar un minigén de AAV en una cápside de AAV pueden proporcionarse a la célula hospedadora en trans. Como alternativa, uno o más de los componentes requeridos (por ejemplo, minigén, secuencias rep, secuencias cap y/o funciones auxiliares) pueden ser proporcionados por una célula hospedadora estable que ha sido diseñada para contener uno o más de los componentes requeridos utilizando métodos conocidos por aquellos expertos en la técnica.

En una modalidad, el AAV incluye un promotor (o un fragmento funcional de un promotor). La selección del promotor que se empleará en el rAAV puede realizarse entre un amplio número de promotores constitutivos o inducibles que pueden expresar el transgén seleccionado en la célula objetivo deseada. Véase, por ejemplo, la lista de promotores identificados en la publicación de patente internacional n.º WO 2014/012482, incorporada a modo de referencia en la presente memoria. En una modalidad, el promotor es específico de la célula. La expresión "específico de la célula" significa que el promotor particular seleccionado para el vector recombinante puede dirigir la expresión del transgén seleccionado en un tipo de célula particular. En algunas modalidades, el promotor es específico para la expresión del transgén en células neuronales. En algunas modalidades, el promotor es específico para la expresión en neuronas corticales (por ejemplo, neuronas piramidales de la corteza). En algunas modalidades, el promotor es específico para la expresión del transgén en neuronas estriatales (neuronas espinosas medias del cuerpo estriado). En algunas modalidades, el promotor es específico para la expresión del transgén en neuronas hipotalámicas. En algunas modalidades, el transgén se expresa en al menos uno de los tipos de células o células.

En otra modalidad, el promotor es el promotor natural del gen objetivo que se va a expresar. Los promotores útiles incluyen, sin limitación, el promotor CAGGS y promotores neuronales específicos, incluidos, sin limitación, un promotor del gen de sinapsina 1 humano, un promotor de enolasa específica de neurona (NSE), un promotor de sinapsina 1 humano, un promotor de quinasa CaMK o un promotor MeCP2. Otros promotores adecuados incluyen promotores inducibles, en donde dichos promotores inician la transcripción solo cuando la célula hospedadora está expuesta a un estímulo que actúa como un desencadenante para activar el promotor.

Otras secuencias reguladoras convencionales contenidas en el minigen o rAAV también se describen en documentos tales como WO 2014/124282 y otras citadas e incorporadas a modo de referencia en la presente memoria. Un experto en la técnica puede hacer una selección entre estas y otras secuencias de control de expresión sin alejarse del alcance descrito en la presente memoria.

El elemento genético seleccionado puede administrarse mediante cualquier método adecuado, incluidos los descritos en el presente documento. Los métodos utilizados para construir cualquier modalidad descrita en la presente memoria son conocidos por aquellos con experiencia en la manipulación de ácidos nucleicos e incluyen ingeniería genética, ingeniería recombinante y técnicas sintéticas. Ver, por ejemplo, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, Nueva York. De manera similar, los métodos para generar viriones rAAV son muy conocidos y la selección de un método adecuado no es una limitación de los métodos y construcciones que se describen en la presente memoria. Ver, por ejemplo, K. Fisher et al., *J. Virol.*, 1993 70: 520-532 y la patente

estadounidense n.º 5.478.745, cada una de las cuales se incorpora a modo de referencia en la presente memoria.

En otra modalidad, la molécula de trans-empalme está incluida en un plásmido provírico, tal como los descritos en la publicación de patente internacional n.º WO 2012/158757, incorporada a la presente memoria a modo de referencia. Un plásmido provírico de este tipo contiene un genoma de AAV recombinante modular que comprende en asociación operativa: una secuencia ITR de AAV2 de tipo salvaje 5' flanqueada por sitios de restricción únicos que permiten la eliminación o el reemplazo fácil de dicha ITR; un promotor que comprende una secuencia de citomegalovirus de 49 ácidos nucleicos delante de una secuencia de beta actina de pollo de citomegalovirus (CMV), o un promotor/potenciador específico de fotorreceptor, el promotor flanqueado por sitios de restricción únicos que permiten la eliminación o el reemplazo fácil de toda la secuencia promotora, y la secuencia delante flanqueada por sitios de restricción únicos que permiten la eliminación o el reemplazo fácil únicamente de la secuencia CMV o potenciadora delante, de la secuencia promotora. La molécula de trans-empalme descrita en la presente memoria se puede insertar en el sitio de un policonector de multiclonación, donde la molécula de trans-empalme está unida de forma operativa a, y bajo el control regulador del promotor. Una secuencia de poliadenilación de la hormona de crecimiento bovina flanqueada por sitios de restricción únicos que permiten la fácil eliminación o reemplazo de dicha secuencia de poli A; y una secuencia de ITR 3' de AAV2 de tipo salvaje flanqueada por sitios de restricción únicos que permiten la fácil eliminación o reemplazo de la ITR 3'; también son parte de dicho plásmido. La estructura principal del plásmido comprende los elementos necesarios para la replicación en células bacterianas, por ejemplo, un gen de resistencia a la kanamicina, y está flanqueada por secuencias terminadoras/aislantes de la transcripción.

En algunas modalidades, un plásmido provírico comprende: (a) un genoma de AAV recombinante modular que comprende en asociación operativa: (i) una secuencia ITR de AAV2 5' de tipo salvaje flanqueada por sitios de restricción únicos que permiten la fácil eliminación o reemplazo de dicha ITR; (ii) un promotor que comprende (A) una secuencia de CMV de 49 ácidos nucleicos delante de una secuencia de beta actina de CMV-pollo o (B) un promotor/potenciador específico de célula neuronal. El promotor está flanqueado por sitios de restricción únicos que permiten la fácil eliminación o reemplazo de toda la secuencia promotora, y la secuencia ascendente está flanqueada por sitios de restricción únicos que permiten la fácil eliminación o reemplazo únicamente de la secuencia potenciadora o CMV ascendente, de la secuencia promotora. También forma parte de este plásmido provírico una secuencia policonectora de clonación múltiple que permite la inserción de una secuencia de molécula de empalme en trans que incluye cualquiera de las descritas en la presente memoria, donde la molécula de empalme en trans está unida de forma operativa a, y bajo el control regulador del promotor; una secuencia de poliadenilación de hormona de crecimiento bovina flanqueada por sitios de restricción únicos que permiten la fácil eliminación o reemplazo de dicha secuencia de poli A; y una secuencia ITR 3' de AAV2 de tipo salvaje flanqueada por

sitios de restricción únicos que permiten la fácil eliminación o reemplazo de la ITR 3'. El plásmido provírico también contiene una estructura plasmídica que comprende los elementos necesarios para la replicación en células bacterianas y que comprende además un gen de resistencia a la kanamicina, estando dicha estructura plasmídica flanqueada por secuencias terminadoras/aislantes de la transcripción. El plásmido provírico descrito en la presente memoria también puede contener en la estructura principal del plásmido una secuencia de relleno de fago lambda no de codificación de 5,1 kb para aumentar la longitud de la estructura principal y evitar el empaquetamiento inverso de genomas de AAV no funcionales.

En otro aspecto adicional, se modifica el promotor del plásmido provírico para reducir el tamaño del promotor y permitir que se inserten secuencias de moléculas de empalme en trans más grandes en el rAAV. En una modalidad, el promotor híbrido CMV/CBA, que normalmente incluye un exón no de codificación y un intrón que suman aproximadamente 1000 pares de bases, se reemplaza con un intrón quimérico de 130 pares de bases, como se describe en la publicación de patente internacional n.º WO 2017/087900, que se incorpora a la presente memoria a modo de referencia en su totalidad.

En algunas modalidades, el promotor CMV se reemplaza por un promotor CAGGS, en donde el promotor CAGGS se utiliza para impulsar la expresión de un editor de exones de ARN descrito en este documento. Ver, por ejemplo, la Figura 27. Como se muestra allí, al comparar CAGGS 5' UTR con o sin HTT 5' UTR, la traducción de proteínas parece estar regulada a través de HTT 5' UTR cuando CAGGS 5' UTR + HTT 5' UTR están unidos de forma operativa. Cuando se elimina el HTT 5' UTR, dejando solo el CAGGS 5' UTR, los presentes inventores observaron una expresión de proteína más fuerte, actividad que puede deberse a algún elemento que regula positivamente la traducción en el CAGGS 5' UTR. Los experimentos de transducción presentados aquí, en donde se introdujeron editores de exones de ARN en las células a través de un AAV, también demostraron que el promotor CAGGS impulsa una expresión significativa de los editores de exones de ARN. Ver, por ejemplo, las Figuras 55, 56, 58 y 59.

Estos plásmidos províricos se emplean luego en metodologías de empaquetamiento actualmente convencionales para generar un virus recombinante que expresa el transgén de la molécula de trans-empalme transportado por los plásmidos províricos. Las líneas celulares de producción adecuadas son fácilmente seleccionadas por un experto en la técnica. Por ejemplo, una célula hospedadora adecuada puede seleccionarse de cualquier organismo biológico, incluidas células procariotas (por ejemplo, bacterianas) y células eucariotas, incluidas células de insectos, células de levadura y células de mamíferos. Brevemente, el plásmido provírico se transfecta en una célula de empaquetamiento seleccionada, donde puede existir transitoriamente. De manera alternativa, el minigén o casete de expresión genética con sus ITR flanqueantes se integra de forma estable en el genoma de la célula hospedadora, ya sea cromosómicamente o como un episoma. Se conocen técnicas de

transfección adecuadas y pueden utilizarse fácilmente para administrar el genoma de AAV recombinante a la célula hospedadora. Normalmente, los plásmidos províricos se cultivan en las células hospedadoras que expresan las proteínas cap y/o rep. En las células hospedadoras, el minigén que consiste en la molécula de empalme en trans con ITR de AAV
 5 flanqueantes se rescata y se empaqueta en la proteína de la cápside o la proteína de la envoltura para formar una partícula viral infecciosa. De este modo, se produce una partícula infecciosa de AAV recombinante cultivando una célula de empaquetamiento que lleva el plásmido provírico en presencia de suficientes secuencias virales para permitir el empaquetamiento del genoma viral del casete de expresión genética en una envoltura o
 10 cápside de AAV infecciosa.

Como alternativa, las moléculas de empalme en trans se pueden administrar utilizando un vector que no sea AAV, por ejemplo, un vector no viral. Se puede utilizar cualquier tecnología de vector no viral adecuada conocida en la técnica o descrita en la presente memoria. Dichos
 15 vectores no virales susceptibles de administración de moléculas de trans-empalme incluyen liposomas (por ejemplo, liposomas catiónicos, liposomas unilamelares o liposomas multilamelares), nanopartículas (por ejemplo, nanopartículas poliméricas, nanopartículas lipídicas (LNP), nanopartículas PEGiladas (por ejemplo, LNP PEGiladas), nanopartículas peptídicas, nanopartículas metálicas y similares), dendrímeros (por ejemplo, dendrímeros
 20 catiónicos, por ejemplo, dendrímeros de polipropilenoimina), exosomas (por ejemplo, exosomas inmunológicamente inertes y/o dirigidos, por ejemplo, elaborados utilizando técnicas descritas en Alvarez-Erviti, et al., 2011, *Nat. Biotechnol.* 29:341) y microvesículas. En algunos casos, las moléculas de trans-empalme descritas en la presente memoria pueden administrarse utilizando péptidos penetrantes en células (CPP), que pueden translocar la
 25 membrana plasmática de una célula objetivo y facilitar la administración de una molécula de trans-empalme al interior de la célula objetivo.

IV. Composiciones farmacéuticas y kits

Se proporcionan en la presente memoria composiciones farmacéuticas que incluyen una
 30 molécula de trans-empalme de ácido nucleico, un plásmido provírico o un rAAV que comprende cualquiera de las moléculas de trans-empalme del ácido nucleico *HTT* descritas en la presente memoria. En algunas modalidades, la composición farmacéutica incluye cualquiera de las moléculas de trans-empalme 5' descritas en la presente memoria.

Estas composiciones farmacéuticas pueden prepararse de manera que estén libres de contaminación y sean adecuadas para la administración *in vivo*. Las composiciones farmacéuticas descritas en la presente memoria pueden evaluarse para detectar contaminación mediante métodos convencionales y luego formularse en una composición farmacéutica destinada a una vía de administración adecuada. Se pueden formular de manera
 40 similar otras composiciones que contienen la molécula de trans-empalme, por ejemplo, ADN sin conjugar, con un portador adecuado. Dicha formulación implica el uso de un vehículo o

portador aceptable desde un punto de vista farmacéutico y/o fisiológico, particularmente dirigido para la administración a la célula objetivo (por ejemplo, una neurona). En una modalidad, los portadores adecuados para la administración a las células objetivo incluyen solución salina amortiguada, una solución isotónica de cloruro de sodio u otros tampones, por ejemplo, HEPES, para mantener el pH en niveles fisiológicos apropiados y, opcionalmente, otros agentes medicinales, agentes farmacéuticos, agentes estabilizadores, tampones, portadores, adyuvantes, diluyentes, etc.

En algunas modalidades, el portador es un líquido para inyección. Entre los portadores aceptables desde un punto de vista fisiológico se incluyen el agua estéril libre de pirógenos y la solución salina tamponada con fosfato estéril libre de pirógenos. Una variedad de dichos portadores conocidos se proporcionan en la patente estadounidense n.º 7,629,322, incorporada en la presente memoria a modo de referencia. En una modalidad, el portador es una solución isotónica de cloruro de sodio. En otra modalidad, el portador es una solución salina equilibrada. En una modalidad, el portador incluye Tween. Si el virus se va a almacenar a largo plazo, puede congelarse en presencia de glicerol o TWEEN®20.

En otras modalidades, las composiciones que contienen moléculas de trans-empalme descritas en la presente memoria incluyen un tensioactivo. Se pueden incluir tensioactivos útiles, tal como Pluronic F68 (Poloxamer 188, también conocido como LUTROL® F68), ya que evitan que el AAV se adhiera a superficies inertes y, de este modo, garantizan la administración de la dosis deseada. A modo de ejemplo, una composición ilustrativa diseñada para el tratamiento de HD descritas en la presente memoria comprende un vector adenoasociado recombinante que lleva una secuencia del ácido nucleico que codifica una molécula de trans-empalme 5' como se describe en la presente memoria, bajo el control de secuencias reguladoras que expresan la molécula de trans-empalme en una célula neuronal de un sujeto mamífero, y un portador farmacéuticamente aceptable. El portador es una solución isotónica de cloruro de sodio e incluye un tensioactivo Pluronic F68. En una modalidad, la molécula de trans-empalme es cualquiera de las descritas en la presente memoria.

En aún otra modalidad de ejemplo, la composición comprende un virus rAAV que comprende cualquiera de las moléculas de trans-empalme *HTT* descritas en la presente memoria para la corrección del gen *HTT*, la secuencia del ácido nucleico bajo el control de un promotor que dirige la expresión de la molécula de trans-empalme en células neuronales del cerebro (por ejemplo, neuronas en la corteza, cuerpo estriado y/o hipotálamo), en donde la composición se formula con un portador y componentes adicionales adecuados para la administración intracerebral (por ejemplo, mediante administración lenta o infusión por convección) o administración intracerebroventricular. En otra modalidad más, la composición o los componentes para la producción o el ensamblaje de esta composición, incluidos los portadores, las partículas de rAAV, los tensioactivos y/o los componentes para generar el rAAV, así como el hardware de laboratorio adecuado para preparar la composición, se pueden

incorporar en un kit. Dichos kits pueden incluir además instrucciones para administrar la composición a un individuo, por ejemplo, como un tratamiento para la HD.

Además, en la presente memoria se proporcionan kits que contienen una composición farmacéutica que comprende una molécula de trans-empalme 5' (por ejemplo, en donde la molécula de trans-empalme está empaquetada en cualquier vector AAV descrito en la presente memoria). En algunas modalidades, el kit incluye instrucciones para mezclar la composición farmacéutica antes de la administración.

V. Métodos y usos

Las moléculas de trans-empalme del ácido nucleico (por ejemplo, moléculas de trans-empalme del ácido nucleico y vectores de codificación de moléculas de trans-empalme del ácido nucleico) y las composiciones descritas anteriormente son útiles para expresar *HTT* funcional, y/o modular la expresión de *HTT*, en una célula objetivo (por ejemplo, una neurona, por ejemplo, neuronas piramidales de la corteza, neuronas espinosas medianas del cuerpo estriado y/o neuronas hipotalámicas) de un individuo en, por ejemplo, métodos para tratar enfermedades o trastornos asociados con mutaciones en el gen *HTT*, como la HD, incluyendo el retraso o la mejora de los síntomas asociados con la HD.

En algunas modalidades, los síntomas de HD incluyen, sin limitación, movimientos espasmódicos o retorcidos involuntarios (corea); problemas musculares, como rigidez o contractura muscular (distonía); movimientos oculares lentos o inusuales; alteración de la marcha, la postura y el equilibrio; dificultad para hablar o tragar. En algunas modalidades, los síntomas de HD incluyen las siguientes categorías: musculares (por ejemplo, marcha anormal, aumento de la actividad muscular, movimientos involuntarios, problemas de coordinación, pérdida de masa muscular y/o espasmos musculares); cognitivos (por ejemplo, amnesia, delirios, falta de concentración, confusión mental, lentitud en la actividad y/o dificultad para pensar y comprender); conductuales (por ejemplo, comportamiento compulsivo, inquietud, irritabilidad o falta de control); psicológicos (por ejemplo, delirio, depresión, alucinaciones y/o paranoia); y anímicos (por ejemplo, ansiedad, apatía y/o cambios de humor). Otros síntomas comúnmente observados en pacientes con HD incluyen pérdida de memoria, temblor y/o pérdida de peso.

Algunas modalidades de los métodos de tratamiento descritos en este documento, o de construcciones y/o moléculas para uso en métodos de tratamiento o en la preparación de medicamentos para el tratamiento, implican apuntar a MSH3 sin apuntar también a HTT. Algunas modalidades implican métodos para tratar o prevenir trastornos de expansión de repeticiones de trinucleótidos mediante la administración a un sujeto de un agente terapéutico que reduce la expresión de MSH3. En algunas modalidades, el trastorno de expansión de repeticiones de trinucleótidos es una enfermedad de poliglutamina. En algunas modalidades, la enfermedad de poliglutamina es atrofia dentatorubropalidoluisiana, enfermedad de

Huntington, atrofia muscular espinal y bulbar, ataxia espinocerebelosa tipo 1, ataxia espinocerebelosa tipo 2, ataxia espinocerebelosa tipo 3, ataxia espinocerebelosa tipo 6, ataxia espinocerebelosa tipo 7, ataxia espinocerebelosa tipo 17 o enfermedad similar a la de Huntington 2. En algunas modalidades, el trastorno de expansión de repeticiones de trinucleótidos es la enfermedad de Huntington. En algunas modalidades, el trastorno de expansión de repeticiones de trinucleótidos es una enfermedad no relacionada con poliglutamina. En algunas modalidades, la enfermedad no relacionada con poliglutamina es el síndrome del cromosoma X frágil, el síndrome de temblor/ataxia asociado al cromosoma XE frágil, el retraso mental asociado al cromosoma X frágil, la ataxia de Friedreich, la distrofia miotónica tipo 1, la ataxia espinocerebelosa tipo 8, la ataxia espinocerebelosa tipo 12, la distrofia muscular oculofaríngea, la insuficiencia ovárica prematura asociada al cromosoma X frágil, el síndrome FRA2A, el síndrome FRA7A o la encefalopatía epiléptica infantil temprana.

Las moléculas de trans-empalme del ácido nucleico (por ejemplo, moléculas de trans-empalme del ácido nucleico y vectores de codificación de moléculas de trans-empalme del ácido nucleico) y las composiciones descritas anteriormente son además útiles para expresar *HTT* funcional, y/o modular la expresión de *HTT*, en una célula objetivo (por ejemplo, una neurona, por ejemplo, neuronas piramidales de la corteza, neuronas medianas espinosas del cuerpo estriado y/o neuronas hipotalámicas) de un individuo para su aplicación, por ejemplo, a su uso para el tratamiento de enfermedades o trastornos asociados a mutaciones en el gen *HTT*, como la HD, incluyendo el retraso o la mejora de los síntomas asociados a la HD según se describe en el presente documento o a su uso en la preparación de un medicamento para el tratamiento de enfermedades o trastornos asociados a mutaciones en el gen *HTT*, como la HD, incluyendo el retraso o la mejora de los síntomas asociados a la HD según se describe en el presente documento. Dichos métodos y usos implican poner en contacto un gen *HTT* objetivo (por ejemplo, pre-ARNm de *HTT*) con una molécula de trans-empalme como se describe en la presente memoria [por ejemplo, una molécula de trans-empalme 5', o una mezcla de moléculas de trans-empalme como se describe en la presente memoria, una composición (por ejemplo, una composición farmacéutica) que la comprende o un medicamento que la comprende], en condiciones en las que un dominio de codificación de la molécula de trans-empalme se empalma al pre-ARNm de *HTT* objetivo para reemplazar una parte del pre-ARNm objetivo que porta uno o más defectos o mutaciones, con un ARNm funcional (es decir, sano), o normal o de tipo salvaje o corregido del gen objetivo, con el fin de corregir la expresión de *HTT* en la célula objetivo. Por tanto, los métodos y composiciones se utilizan para tratar las patologías de HD asociadas a mutaciones específicas como, por ejemplo, repeticiones de CAG de más de 40 repeticiones a nivel genómico y tramos de poliglutamina de más de 40 glutaminas en proteínas transcritas y traducidas a partir de repeticiones de CAG genómicas de más de 40 repeticiones (enfermedad que provoca la expansión de repeticiones de CAG genómicas).

En algunas modalidades, se proporcionan en la presente memoria métodos para expresar *HTT*

funcional en una célula objetivo, poniendo en contacto (por ejemplo, transduciendo) la célula objetivo con cualquiera de las moléculas de trans-empalme del ácido nucleico, vectores (por ejemplo, vectores AAV) o composiciones descritas en la presente memoria. En una modalidad, la puesta en contacto implica la administración directa de la composición (por ejemplo, una composición farmacéutica) al individuo afectado. En otra modalidad, la puesta en contacto puede ocurrir *ex vivo* con una célula cultivada (por ejemplo, una célula neuronal o su precursor) y la célula cultivada tratada reimplantada en el individuo. En otra modalidad, el método implica la administración de un rAAV que porta cualquiera de las moléculas de transsplicing 5' *HTT* (incluyendo moléculas de transsplicing duales 5' *HTT/MSH3* con dominios de unión en tándem y moléculas de transsplicing 5' *HTT* y 5' o 3' *MSH3* dispuestas en tándem); moléculas de transsplicing 5' o 3' *MSH3* que inducen NMD; moduladores de empalme de *MSH3*; pri-miARN que comprende miARN de *MSH3*; o pri-miARN que comprende miARN de *HTT* descritos en este documento, o cualquier combinación de estos. En otra modalidad más, el método implica administrar una mezcla de rAAV que lleva una molécula de trans-empalme 5' *HTT* y una molécula de trans-empalme 5' *MSH3* rAAV. Cada una de estas modalidades se puede combinar con ASO (por ejemplo, SEQ ID NO: 131) que reducen el cis-empalme para *HTT* o *MSH3*, promoviendo así el trans-empalme a las respectivas moléculas de trans-empalme. En algunas modalidades, los ASO que reducen el cis-empalme para uno o cada uno de *HTT* y *MSH3* se utilizan en combinación con moléculas de trans-empalme 5' *HTT* (incluidas moléculas de trans-empalme duales 5' *HTT/MSH3*) o combinaciones de moléculas de trans-empalme 5' *HTT* y moléculas de trans-empalme 5' *MSH3*. En algunas modalidades, los ARN ASO antisentido que reducen el cis-empalme para uno o cada uno de *HTT* y *MSH3* se utilizan en combinación con moléculas de trans-empalme 5' *HTT* (incluidas moléculas de trans-empalme duales 5' *HTT/MSH3*) o combinaciones de moléculas de trans-empalme 5' *HTT* y moléculas de trans-empalme 5' *MSH3*. Estos métodos comprenden administrar a un individuo que lo necesita una concentración eficaz de una composición de cualquiera de los descritos en la presente.

En algunas modalidades, los métodos incluyen seleccionar una o más moléculas de trans-empalme para tratar a un individuo que tiene un trastorno asociado con mutación/es en *HTT*. En algunas modalidades, se abarca en la presente memoria el uso de una o más moléculas de trans-empalme para tratar a un individuo que tiene un trastorno asociado con una o más mutaciones en *HTT* o el uso de las mismas en la preparación de un medicamento para el tratamiento de un individuo que tiene un trastorno asociado con una o más mutaciones en *HTT*. Dichos métodos y usos incluyen la selección de una o más moléculas de trans-empalme para tratar a un individuo que tiene un trastorno asociado con una mutación en *HTT* o para el uso de dichas una o más moléculas de trans-empalme seleccionadas en el tratamiento de un individuo que tiene un trastorno asociado con una mutación en *HTT* o para el uso de dichas una o más moléculas de trans-empalme seleccionadas en la preparación de un medicamento para el tratamiento de un individuo que tiene un trastorno asociado con una o más mutaciones en *HTT*. Dicha selección puede basarse en el genotipo del individuo. En algunas modalidades,

un trastorno asociado con *HTT* puede ser un trastorno autosómico recesivo. En algunos casos, el individuo es homocigoto o heterocigoto compuesto para las mutaciones en *HTT*. En la técnica se conocen métodos para detectar e identificar mutaciones particulares en *HTT*.

- 5 Los métodos de la invención incluyen seleccionar una única molécula de trans-empalme basándose en la ubicación de una única mutación en *HTT* (por ejemplo, una mutación de un alelo del individuo). Como se describe en este documento, las mutaciones causales asociadas con la HD comprenden la expansión patológica (más de 35 o más de 40 repeticiones de CAG) de las repeticiones de CAG en el exón 1 del gen *HTT*. Por lo tanto, en algunas modalidades, 10 los métodos de la invención incluyen administrar una única molécula de trans-empalme para corregir la expansión patológica (más de 35 o más de 40 repeticiones de CAG) de las repeticiones de CAG en el exón 1 del gen *HTT*, por ejemplo, sin tener en cuenta la ubicación de cualquier otra mutación que pueda existir en el otro alelo.
- 15 Además se presentan métodos que implican la selección de una única molécula de trans-empalme 5' para corregir la expansión patológica (más de 35 o más de 40 repeticiones de CAG) de las repeticiones de CAG en el exón 1 del pre-ARNm de *HTT* y reducir los niveles de *MSH3*, de modo que una única molécula de trans-empalme capaz de ser empaquetada en un vector AAV es capaz de corregir la expansión patológica de las repeticiones de CAG en *HTT* 20 a nivel de ARN y reducir los niveles de *MSH3* celular para retardar o inhibir la expansión de las repeticiones de CAG en el gen *HTT* en el genoma.

En otras modalidades, se proporcionan en la presente memoria métodos para corregir mutaciones múltiples dentro de un gen *HTT* utilizando dos moléculas de trans-empalme: una 25 molécula de trans-empalme 5' *HTT* y una molécula de trans-empalme 5' *MSH3*.

Las moléculas de trans-empalme de ácidos nucleicos descritas en este documento y los vectores, plásmidos provirales y AAV que los comprenden, así como las composiciones que comprenden dichas moléculas de trans-empalme de ácidos nucleicos y vectores, plásmidos 30 provirales y AAV que los comprenden son para uso en el tratamiento médico, en particular para uso en el tratamiento de la HD. En algunas modalidades, cuando se utiliza, por ejemplo, un vector AAV (u otro vector de terapia génica), el vector AAV puede administrarse mediante infusión directa en el cerebro. En algunas modalidades, la infusión directa comprende una infusión intratecal del vector AAV en el líquido cefalorraquídeo. La infusión intratecal ofrece un 35 modo de administración eficiente al SNC, desde donde se pueden dirigir las neuronas. En algunas modalidades, las estructuras corticales y estriatales pueden ser el objetivo a través de la administración de difusión mejorada por convección intraestriatal (CED) mediante inyecciones en el estriado. En algunas modalidades, las inyecciones pueden dirigirse al cuerpo estriado y al tálamo para proporcionar una mayor cobertura de las estructuras del cerebro 40 implicadas en la HD. En algunas modalidades, los vectores AAV pueden administrarse por vía intraestriatal o intraestriatalámicamente a través de inyecciones de CED en el estriado o en el

estriado y el tálamo. Estas inyecciones pueden realizarse mediante inyecciones guiadas por resonancia magnética. Estos métodos de tratamiento son particularmente útiles para sujetos humanos que padecen HD. Este tratamiento involucra a sujetos humanos que padecen HD, incluidos aquellos que tienen una predisposición genética para desarrollar HD y que no presentan síntomas de HD. En consecuencia, en algunas modalidades, el tratamiento de sujetos humanos con HD puede incluir el tratamiento de cualquier sujeto humano que sea portador de un alelo Huntingtina con más de 35 repeticiones CAG.

En algunas modalidades, una concentración eficaz de un virus adenoasociado recombinante que lleva una molécula de trans-empalme como se describe en la presente memoria varía entre aproximadamente 10^8 y 10^{13} genomas vectoriales por mililitro (vg/ml). Las unidades infecciosas de rAAV se miden como se describe en McLaughlin et al., *J. Virol.* 1988, 62: 1963. En otra modalidad, la concentración varía entre 10^9 y 10^{13} vg/ml. En otra modalidad, la concentración eficaz es de aproximadamente $1,5 \times 10^{11}$ vg/ml. En otra modalidad, la concentración efectiva es de aproximadamente 5×10^{11} vg/ml. En una modalidad, la concentración eficaz es de aproximadamente $1,5 \times 10^{10}$ vg/ml. En otra modalidad, la concentración efectiva es de aproximadamente $2,8 \times 10^{11}$ vg/ml. En incluso otra modalidad, la concentración efectiva es de aproximadamente $1,5 \times 10^{12}$ vg/ml. En otra modalidad, la concentración eficaz es de aproximadamente $1,5 \times 10^{13}$ vg/ml.

Es deseable que se utilice la dosis eficaz más baja (copias totales del genoma entregadas) del virus para reducir el riesgo de efectos indeseables, como toxicidad y otros problemas relacionados con la administración al cerebro. Una dosis eficaz de un virus adenoasociado recombinante que lleva una molécula de trans-empalme como se describe en la presente memoria varía entre aproximadamente 10^8 y 10^{13} genomas vectoriales (vg) por dosis (es decir, por inyección). En una modalidad, la dosis varía entre 10^9 y 10^{13} vg. En otra modalidad, la dosis eficaz es de aproximadamente $1,5 \times 10^{11}$ vg. En otra modalidad, la dosis eficaz es de aproximadamente 5×10^{11} vg. En una modalidad, la dosis eficaz es de aproximadamente $1,5 \times 10^{10}$ vg. En otra modalidad, la dosis eficaz es de aproximadamente $2,8 \times 10^{11}$ vg. En incluso otra modalidad, la dosis eficaz es de aproximadamente $1,5 \times 10^{12}$ vg. En otra modalidad, la concentración efectiva es de aproximadamente $1,5 \times 10^{13}$ vg. El médico tratante puede seleccionar otras dosis en estos rangos o en otras unidades, teniendo en cuenta el estado físico del individuo que se está tratando, incluida la edad del individuo, la composición que se está administrando y el trastorno particular, la célula objetivo y el grado en que se ha desarrollado el trastorno, si es progresivo.

En algunas modalidades, la composición puede administrarse en un volumen de desde aproximadamente 50 μ l hasta aproximadamente 1 ml, incluidos todos los números dentro del rango, dependiendo del tamaño del área a tratar, el título viral utilizado, la vía de administración y el efecto deseado del método. En una modalidad, el volumen es de aproximadamente 50 μ l. En otra modalidad, el volumen es de aproximadamente 70 μ l. En otra modalidad, el volumen

es de aproximadamente 100 µl. En otra modalidad, el volumen es de aproximadamente 125 µl. En otra modalidad, el volumen es de aproximadamente 150 µl. En otra modalidad, el volumen es de aproximadamente 175 µl. En incluso otra modalidad, el volumen es de aproximadamente 200 µl. En otra modalidad, el volumen es de aproximadamente 250 µl. En otra modalidad, el volumen es de aproximadamente 300 µl. En otra modalidad, el volumen es de aproximadamente 350 µl. En otra modalidad, el volumen es de aproximadamente 400 µl. En otra modalidad, el volumen es de aproximadamente 450 µl. En otra modalidad, el volumen es de aproximadamente 500 µl. En otra modalidad, el volumen es de aproximadamente 600 µl. En otra modalidad, el volumen es de aproximadamente 750 µl. En otra modalidad, el volumen es de aproximadamente 850 µl. En otra modalidad, el volumen es de aproximadamente 1000 µl.

En algunos casos, los tratamientos y usos que se describen en la presente memoria reemplazan el 10 % o más del ARNm de *HTT* objetivo en la célula objetivo (por ejemplo, el 11 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 12 % o más del ARNm de *HTT* objetivo en la/s célula/s objetivo, el 13 % o más del ARNm de *HTT* objetivo en la/s célula/s objetivo, el 14 % o más del ARNm de *HTT* objetivo en la/s célula/s objetivo, el 15 % o más del ARNm de *HTT* objetivo en la/s célula/s objetivo, el 16 % o más del ARNm de *HTT* objetivo en la/s célula/s objetivo, el 17 % o más del ARNm de *HTT* objetivo en la/s célula/s objetivo, el 18 % o más del ARNm de *HTT* objetivo en la/s célula/s objetivo, o el 19 % o más del ARNm de *HTT* objetivo en la o las células objetivo). En algunos casos, los tratamientos y usos descritos en este documento reemplazan el 20 % o más del ARNm de *HTT* objetivo en la célula objetivo (por ejemplo, el 21 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 22 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 23 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 24 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 25 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 26 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 27 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 28 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 29 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 30 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 31 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 32 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 33 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 34 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 35 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 36 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 37 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 38 % o más del ARNm de *HTT* objetivo en la o las células objetivo, el 39 % o más del ARNm de *HTT* objetivo en la o las células objetivo, 40 % o más del ARNm de *HTT* objetivo en la o las células objetivo, 41 % o más del ARNm de *HTT* objetivo en la o las células objetivo, 42 % o más del ARNm de *HTT* objetivo en la o las células objetivo, 43 % o más del ARNm de *HTT* objetivo en la o las células objetivo, 44 % o más del ARNm de *HTT* en la o las células objetivo, 45 % o más del ARNm de *HTT* objetivo en la o las células objetivo, 46 % o más del ARNm de

HTT objetivo en la o las células objetivo, 47% o más del ARNm de *HTT* objetivo en la o las células objetivo, 48% o más del ARNm de *HTT* objetivo en la o las células objetivo, 49% o más del ARNm de *HTT* objetivo en la o las células objetivo, o 50% o más del ARNm de *HTT* objetivo en la o las células objetivo). En algunos casos, los tratamientos y usos descritos en la presente memoria reemplazan un 50% o más del ARNm de *HTT* objetivo en la célula objetivo (por ejemplo, un 51 % o más del ARNm de *HTT* objetivo en la o las células objetivo, 52% o más del ARNm de *HTT* en la o las células objetivo, 53% o más del ARNm de *HTT* en la o las células objetivo, 54% o más del ARNm de *HTT* en la o las células objetivo, 55% o más del ARNm de *HTT* en la o las células objetivo, 56% o más del ARNm de *HTT* en la o las células objetivo, 57% o más del ARNm de *HTT* en la o las células objetivo, 58% o más del ARNm de *HTT* en la o las células objetivo, 59% o más del ARNm de *HTT* en la o las células objetivo, 60% o más del ARNm de *HTT* en la o las células objetivo, 61% o más del ARNm de *HTT* en la o las células objetivo, 62% o más del ARNm de *HTT* en la o las células objetivo, 63% o más del ARNm de *HTT* en la o las células objetivo, 64% o más del ARNm de *HTT* en la o las células objetivo, 65% o más del ARNm de *HTT* en la o las células objetivo, 66% o más del ARNm de *HTT* en la o las células objetivo, 67% o más del ARNm de *HTT* en la o las células objetivo, 68% o más del ARNm de *HTT* en la o las células objetivo, 69% o más del ARNm de *HTT* en la o las células objetivo, 70% o más del ARNm de *HTT* en la o las células objetivo, 71% o más del ARNm de *HTT* en la o las células objetivo, 72% o más del ARNm de *HTT* en la o las células objetivo, 73% o más del ARNm de *HTT* en la o las células objetivo, 74% o más del ARNm de *HTT* en la o las células objetivo, 75% o más del ARNm de *HTT* en la o las células objetivo, 76% o más del ARNm de *HTT* en la o las células objetivo, 77% o más del ARNm de *HTT* en la o las células objetivo, 78% o más del ARNm de *HTT* en la o las células objetivo, 79% o más del ARNm de *HTT* en la o las células objetivo, o 80% o más del ARNm de *HTT* en la o las células objetivo (por ejemplo, 81% o más del ARNm de *HTT* en la o las células objetivo, 82% o más del ARNm de *HTT* en la o las células objetivo, 83% o más del ARNm de *HTT* en la o las células objetivo, 84% o más del ARNm de *HTT* en la o las células objetivo, 85% o más del ARNm de *HTT* en la o las células objetivo, 86% o más del ARNm de *HTT* en la o las células objetivo, 87% o más del ARNm de *HTT* en la o las células objetivo, 88% o más del ARNm de *HTT* en la o las células objetivo, 89% o más del ARNm de *HTT* en la o las células objetivo, 90% o más del ARNm de *HTT* en la o las células objetivo, 91% o más del ARNm de *HTT* en la o las células objetivo, 92% o más del ARNm de *HTT* en la o las células objetivo, 93% o más del ARNm de *HTT* en la o las células objetivo, 94% o más del ARNm de *HTT* en la o las células objetivo, 95% o más del ARNm de *HTT* en la o las células objetivo, 96% o más del ARNm de *HTT* en la o las células objetivo, 97% o más del ARNm de *HTT* en la o las células objetivo, 98% o más del ARNm de *HTT* en la o las células objetivo, 99% o más del ARNm de *HTT* en la o las células objetivo, o 100% o más del ARNm de *HTT* en la o las células objetivo.

Para cada uno de los métodos y usos descritos, el tratamiento o uso puede utilizarse para prevenir la aparición de más daños o para rescatar tejido que tenga una enfermedad leve, moderada o avanzada. Como se utiliza en la presente memoria, el término «rescate» significa

prevenir la progresión de la enfermedad, prevenir la propagación del daño a las células no lesionadas y/o mejorar el daño en las células lesionadas.

Así, en una modalidad, la composición se administra antes de la aparición de la enfermedad.

5 En otra modalidad, la composición se administra antes del desarrollo de los síntomas. En otra modalidad, la composición se administra después del desarrollo de los síntomas. En otra modalidad más, la composición se administra cuando menos del 90 % de las células objetivo están funcionando o permanecen, por ejemplo, en comparación con un tejido de referencia. En incluso otra modalidad, la composición se administra cuando más del 10 % de las células objetivo funcionan o permanecen, por ejemplo, en comparación con un tejido de referencia. En 10 otra modalidad más, la composición se administra cuando más del 20 % de las células objetivo funcionan o permanecen. En incluso otra modalidad, la composición se administra cuando más del 30 % de las células objetivo funcionan o permanecen. En incluso otra modalidad, la composición se administra cuando más del 40 % de las células objetivo funcionan o permanecen. En incluso otra modalidad, la composición se administra cuando más del 50 % 15 de las células objetivo funcionan o permanecen. En incluso otra modalidad, la composición se administra cuando más del 60 % de las células objetivo funcionan o permanecen. En incluso otra modalidad, la composición se administra cuando más del 70 % de las células objetivo funcionan o permanecen. En incluso otra modalidad, la composición se administra cuando más del 80 % de las células objetivo funcionan o permanecen. En incluso otra modalidad, la composición se administra cuando más del 90 % de las células objetivo funcionan o permanecen. En incluso otra modalidad, la composición se administra cuando más del 95 % 20 de las células objetivo funcionan o permanecen.

25 En otra modalidad más, cualquiera de los métodos o usos descritos anteriormente se realiza en combinación con otra terapia, o terapia secundaria. La terapia puede ser cualquier terapia ahora conocida o aún desconocida que ayude a prevenir, detener o mejorar estas mutaciones o defectos o cualquiera de los efectos asociados con ellos. La terapia secundaria se puede administrar antes, simultáneamente con o después de la administración de una composición 30 farmacéutica descrita anteriormente. En una modalidad, una terapia secundaria implica enfoques no específicos para mantener la salud de las células neuronales, como la administración de factores neurotróficos, antioxidantes y/o agentes antiapoptóticos. Los enfoques no específicos se logran mediante la inyección de proteínas, ADN recombinante, vectores virales recombinantes, células madre, tejido fetal o células genéticamente 35 modificadas. Entre estos últimos podrían incluirse células genéticamente modificadas que estén encapsuladas.

Para su uso en estos métodos, el volumen y el título viral de cada inyección se determinan individualmente y pueden ser iguales o diferentes de otras inyecciones realizadas en, por 40 ejemplo, el cerebro. Las dosis, administraciones y pautas posológicas podrán ser determinados por el médico tratante teniendo en cuenta las enseñanzas de esta descripción.

VI. Ejemplos

Los ejemplos que siguen no limitan el alcance de las modalidades descritas en la presente memoria. Un experto en la técnica apreciará que se pueden realizar modificaciones en los siguientes ejemplos que pretenden estar comprendidas en el espíritu y alcance de la invención.

Ejemplo 1: Eficiencias de trans-empalme de HTT dirigidos a editores de exones de ARN

Los editores de exones dirigidos al intrón 1, al intrón 2 o al intrón 3 fueron diseñados y probados para verificar la eficiencia del reemplazo del exón 1 del *HTT* mutante (Figura 3). En resumen, las células HEK293 se transfectaron con editores de exones de ARN *HTT* dirigidos al intrón 1, al intrón 2 o al intrón 3 que se dirigen a varias regiones del intrón respectivo. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR. Los editores de exones producidos por la construcción representado en la Figura 4 comprendían un UTR 5', una secuencia codificante del exón 1, un sitio donante de empalme, un enlazador, un dominio de unión y una secuencia terminadora. Se varió el dominio de unión para apuntar a diferentes posiciones a lo largo del intrón 1 de *HTT*. Se evaluaron los editores de exones de ARN dirigidos al intrón 1 de *HTT* ilustrativos que exhiben diferentes niveles de eficiencia de trans-empalme (% de reemplazo) dependiendo de dónde se dirige el dominio de unión en el intrón testículos. Las Figuras 5 y 6 muestran el % de reemplazo para los dominios de unión probados. Los resultados mostrados se relacionan con editores de exones de ARN dirigidos al intrón 1 de *HTT* de ejemplo que comprenden los objetivos del dominio de unión indicados, en donde el 5' UTR comprende el *HTT* 5' UTR (SEQ ID NO: 136) y el enlazador comprende el enlazador de 40 mer (SEQ ID NO: 37). NBD_150 es un editor de control en el que el dominio de enlace que apunta a *HTT* se reemplaza con un dominio de enlace que no apunta a *HTT*. Los resultados demuestran que dirigirse al extremo 3' del intrón 1 cerca del punto de ramificación se correlaciona con una mayor eficiencia de trans-empalme en el pre-ARNm *HTT*.

Los editores de exones se muestran en la Figura 7 en donde la expresión fue impulsada por un promotor CMV. El editor de exones representado comprendía el UTR 5' de *HTT*, la etiqueta FLAG 3X de extremo N, la secuencia codificante del exón 1, un sitio donante de empalme, un enlazador, un dominio de unión (*HTT_intron1_11704_100*) y una secuencia terminadora. Se midió la actividad del ARN objetivo del intrón 1 de *HTT* ejemplar que incluía varios enlazadores (Figura 8). Como se muestra allí, algunos enlazadores aumentaron la eficiencia del trans-empalme en relación con el enlazador de 40 mer en los editores de exones del intrón 1 de *HTT* (*HTT_intron1_11704_100*). Los editores de exones representados en la Figura 9 comprendían un UTR 5', secuencias codificantes de los exones 1-2, un sitio donante de empalme, un enlazador, un dominio de unión y una secuencia terminadora. Se varió el dominio de unión para apuntar a diferentes posiciones a lo largo del intrón 2 de *HTT*. Se midió la actividad de varios editores de exones de ARN dirigidos al intrón 2 de *HTT* ilustrativos, que exhiben diferentes niveles de eficiencia de trans-empalme (% de reemplazo) dependiendo de dónde se

dirige el dominio de unión en el intrón y se mostró en las Figuras 10-11. Los resultados mostrados se relacionan con editores de exones de ARN dirigidos al intrón 2 de HTT de ejemplo que comprenden los objetivos del dominio de unión indicados, en donde el 5' UTR comprende el HTT 5' UTR y el enlazador comprende el enlazador de 40 mer. NBD, un editor de a control en el que el dominio de enlace que apunta a HTT se reemplazó con un dominio de enlace que no apunta a HTT. También se evaluó un mutante de empalme, un editor de control que carece de un sitio donante de empalme funcional.

Los editores de exones de ARN dirigidos al intrón 2 de HTT que se dirigen a la región anterior del punto de ramificación se diseñaron para dirigirse a la parte anterior del punto de ramificación del intrón 2 y variar en la longitud del dominio de unión (Figura 12A). La expresión del editor de exones fue impulsada por un promotor CMV. Los editores de exones comprendían el UTR 5' de HTT, la etiqueta FLAG 3X de extremo N para la detección de proteínas en el objetivo, la secuencia codificante de los exones 1-2, un sitio donante de empalme, el enlazador 41mer_2, el dominio de unión indicado, y una secuencia terminadora. El dominio de unión se varió para dirigirse a diferentes longitudes en la parte anterior del punto de ramificación en el intrón 2. La longitud del dominio de unión influyó en las eficiencias de selección, según los cálculos del % de sustitución (Figura 12B). Análisis que examinaron la asociación de la longitud del dominio de unión en la región del intrón en la parte anterior del punto de ramificación con la funcionalidad identificaron longitudes de dominio de unión de entre 125-200 nt como las que tenían las mayores eficiencias relativas de trans-empalme.

Se probaron construcciones que codifican editores de exones de ARN dirigidos al intrón 2 de HTT con diferentes enlazadores (Figura 13). La expresión del editor de exones fue impulsada por un promotor CMV. Los editores de exones ejemplares comprendían el UTR 5' de HTT, la etiqueta FLAG 3X de extremo N, las secuencias codificantes de los exones 1-2, un sitio donante de empalme, un enlazador, un dominio de unión (HTT_intron2_12061_150) y una secuencia terminadora. Los editores de exones que comprenden los enlazadores indicados no exhibieron eficiencias de trans-empalme significativamente diferentes en relación con el enlazador de 40 mer en el intrón 2 de HTT (HTT_intron2_12061_150) (Figura 14). Se probaron construcciones que codifican editores de exones de ARN dirigidos al intrón 3 de HTT (Figura 23). Se midió la actividad de varios editores de exones de ARN dirigidos al intrón 3 de HTT ilustrativos, que muestran diferentes niveles de eficiencia de trans-empalme (% de reemplazo) dependiendo del sitio de unión dentro del intrón al que se dirige por el dominio de unión (Figura 24). Se realizó una comparación directa de los editores de exones dirigidos al intrón 2 de HTT con los editores de exones dirigidos al intrón 3 de HTT (Figura 25).

La mitigación del auto empalme no afecta la eficiencia del trans-empalme de los editores de exones HTT. En la Tabla 1 se presentan los sitios de empalme críptico identificados en el editor de exones original y los cambios de secuencia realizados en los editores de exones mitigados con auto empalme (Figura 26). En breve, las células HEK293 se transfectaron con editores de

exones de ARN dirigidos al intrón 2 de HTT (HTT_intron2_12061_150) con o sin mitigación de auto empalme. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR (Figura 26).

5 Se produjeron imágenes de RT-qPCR y análisis de transferencia Western de lisados de células HEK293 transfectadas con editores de exones de ARN HTT que prueban 2 promotores y diferentes combinaciones de 5' UTR (Figura 27). En resumen, las células HEK293 se transfectaron con editores de exones HTT etiquetados con FLAG en el extremo N impulsados por el promotor CMV o CAGGS, con o sin el 5' UTR HTT, y se probó el sitio de empalme de tipo salvaje (GTAAGT) dirigido al intrón 2 (HTT_intron2_12061_150), el mutante de empalme
10 dirigido al intrón 2 (HTT_intron2_12061_150) o el sitio de empalme de tipo salvaje con un dominio de unión no dirigido (NBD). El ARN de estas células se sometió a RT-qPCR (Figura 27 panel superior). α Se utilizó el anticuerpo FLAG para detectar la proteína generada luego de un trans-empalme exitoso (que comprende un epítipo FLAG de extremo N) en lisados de
15 células completas (Figura 27, panel inferior). ONT se refiere a la proteína HTT trans-empalmada con éxito en el objetivo y NSP se refiere a la proteína no empalmada.

La proteína no empalmada (NSP) se reduce de manera combinatoria mediante la inclusión de tres repeticiones en tándem del sitio de unión del ARNnp U1 (3X UBS; SEQ ID NO: 345) y un
20 elemento rico en AU (ARE; SEQ ID NO: 346) en un editor de exones dirigido al intrón 1 de 5' HTT ejemplar (Figura 28). En resumen, las células HEK293 fueron transfectadas con editores de exones de ARN dirigidos al intrón 1 de HTT (HTT_intron1_11704_100) que fueron variados en su región de enlace para incluir los elementos de reducción de NSP indicados. Las células se recolectaron 48 horas después de la transfección y se sometieron a análisis de transferencia
25 Western. ONT se refiere a la proteína HTT trans-empalmada con éxito en el objetivo y NSP se refiere a la proteína no empalmada. En los editores de exones dirigidos al intrón 1 de HTT, se demostró que la inclusión de 3X UBS en el dominio enlazador redujo los niveles de NSP en ~75 % en relación con el control de enlazador de solo 40 mer de referencia. La inclusión de ARE en el enlazador redujo los niveles de NSP en ~ 40 % en relación con el control del
30 enlazador de referencia de solo 40 mer. La combinación de 3X UBS y ARE en el editor de exones redujo el nivel de NSP en ~ 88% en relación con el control del enlazador de referencia de solo 40 mer.

La proteína no empalmada (NSP) se reduce de manera combinatoria mediante la inclusión de tres repeticiones en tándem del sitio de unión del ARNnp U1 (3X UBS; SEQ ID NO: 345) y un
35 elemento rico en AU (ARE; SEQ ID NO: 346) en un editor de exones dirigido al intrón 2 de 5' HTT ejemplar (Figura 29). En resumen, las células HEK293 fueron transfectadas con editores de exones de ARN dirigidos al intrón 2 de HTT (HTT_intron2_12061_150) que fueron variados en su región de enlace para incluir los elementos de reducción de NSP indicados. Las células se recolectaron 48 horas después de la transfección, se analizaron para determinar la
40 eficiencia del trans-empalme mediante RT-qPCR (Figura 29, panel superior) o se sometieron

a análisis de transferencia Western (Figura 29, panel inferior). En los editores de exones dirigidos al intrón 2 de HTT, los presentes inventores demostraron que la inclusión de 3X UBS en el dominio enlazador redujo los niveles de NSP en ~38% en relación con el control de enlazador de solo 40 mer de referencia. La inclusión de ARE en el enlazador redujo los niveles de NSP en ~ 33% en relación con el control del enlazador de referencia de solo 40 mer. La combinación de 3X UBS y ARE en el editor de exones trabajó combinatoriamente en la reducción del nivel de NSP en aproximadamente 66 % en comparación con el control del enlazador de referencia de solo 40 mer.

Ejemplo 2: Editores de exones de ARN con oligonucleótidos antisentido (ASO)

Se probó una reacción de trans-empalme y la competencia de esta con respecto al cis-empalme (Figura 15). Los oligonucleótidos antisentido (ASO) se diseñaron para bloquear sitios de cis-empalme competitivos (ASO8-10), así como sitios de cis-empalme para el exón en la parte anterior (ASO2-7) (Figura 16). Cada uno de estos ASO se cotransfectó con el editor de exones dirigido al intrón 2 de HTT indicado (HTT_intron2_12061_150) y se analizó la eficiencia de trans-empalme in vitro en células HEK293. En breve, las células HEK293 se cotransfectaron con una construcción de editor de exones de ARN (REEC) dirigido al intrón 2 de HTT ejemplar y ASO diseñados para bloquear el sitio de cis-empalme competitivo o ASO diseñados para bloquear el empalme del intrón en la parte anterior. Las células se recolectaron 48 horas después de la transfección y se analizaron para determinar la eficiencia del trans-empalme mediante RT-qPCR.

Se midió la actividad de reemplazo porcentual de un editor de exones dirigido al intrón 2 de HTT ejemplar (HTT_intron2_12061_150) en combinación con los ASO indicados (Figura 17). ASO6, que fue diseñado para bloquear el cis-empalme del intrón en la parte anterior, condujo a una mejora en las eficiencias del trans-empalme in vitro. Se produjeron las imágenes representativas de RT-qPCR y análisis de transferencia Western (sondeadas para el epítipo FLAG de extremo N) de lisados de células completas de células HEK293 transfectadas con editores de exones de ARN HTT que comprenden los elementos indicados (Figura 18, panel superior). Cabe destacar que los niveles de detección de proteínas en el objetivo se correlacionan con la eficiencia del trans-empalme, representada por el % de reemplazo de ARN HTT. Las células HEK293 fueron transfectadas con editores de exones HTT etiquetados con FLAG en el extremo N que tienen un rango de actividad basado en ensayos de qPCR (Figura 18 panel superior). Se utilizó el anticuerpo anti-FLAG para detectar la proteína generada luego de un trans-empalme exitoso en lisados de células completas (Figura 18, panel inferior). La intensidad de la banda FLAG en el objetivo (ONT) escala con el rendimiento relativo del editor de exones basado en qPCR.

Se evaluó una estrategia de combinación para corregir la HTT mutante mediante el trans-empalme y el silenciamiento de especies de HTT no editadas (incluyendo HTT1a) con un microARN (miARN) (Figura 46). Se usó una molécula híbrida vectorizada que combina un

editor de exones HTT con un miARN dirigido al ARNm HTT sin editar (Figura 47). Se agregó un ARN de horquilla corta (ARNhc) o un microARN (miARN) diseñado para reducir la expresión del gen HTT a un editor de exones dentro del mismo cistón (por ejemplo, en un intrón del editor de exones) o como un cistón separado con sus propias secuencias reguladoras. El ARNi se diseñó para reducir la expresión del objetivo no editado (por ejemplo, HTT mutante). Con el propósito de reducir selectivamente HTT no editado con ARNi, el Editor de Exones comprende una secuencia de dominio de codificación de HTT que comprende una porción de secuencia alterada que hace que la HTT editada sea resistente al ARNhc o miARN.

Los perfiles de trans-empalme y de silenciamiento de HTT del editor de exones HTT, miARN-1 HTT, y una molécula híbrida dual del editor de exones HTT y miARN-1 HTT se presentan en la Figura 48. Tal como se muestra allí, la Figura 48A) presenta el % de transcripciones HTT trans-empalmadas (editadas) en todas las transcripciones HTT, y la Figura 48B) presenta el número de copias de transcripciones HTT no editadas y trans-empalmadas (editadas) en cada una de las afecciones evaluadas. En resumen, las células HEK293 se transfectaron con el editor de exones de ARN dirigido al intrón 2 de HTT, miARN-1 HTT, y una molécula híbrida dual del editor de exones dirigido al intrón 2 de HTT + miARN-1 HTT como se indica en la Figura 48. Las células se recolectaron 48 horas después de la transfección y el ARN se sometió a RT- qPCR. La molécula híbrida dual del editor de exones HTT y miARN-1 HTT (SEQ ID NO: 354), comprendía la SEQ ID NO: 341 y la SEQ ID NO: 204.

Se produjeron los perfiles de trans-empalme y de silenciamiento de HTT de moléculas que contienen miARN-1 HTT y miARN-2 HTT in vitro (Figura 49A-C). En resumen, las células HEK293 se transfectaron con el editor de exones de ARN dirigido al intrón 2 de HTT y moléculas híbridas duales del editor de exones dirigido al intrón 2 de HTT + miARN-1 HTT o miARN-2 HTT. Las células se recolectaron 48 horas después de la transfección y el ARN se sometió a análisis de RT-qPCR para un perfil de trans-empalme (Figura 49A), un perfil de silenciamiento de HTT (Figura 49B), y un análisis del número de copias de HTT (FIG 49C). La molécula híbrida dual del editor de exones HTT y miARN-1 HTT (SEQ ID NO: 354), comprendía la SEQ ID NO: 341 y la SEQ ID NO: 204. La molécula híbrida dual del editor de exones HTT y miARN-2 HTT (SEQ ID NO: 355), comprendía la SEQ ID NO: 344 y la SEQ ID NO: 204.

miARN-1 silencia con éxito las transcripciones *HTT* no editadas con una interacción mínima con el editor de exones y la transcripción HTT editada (Figura 50). En resumen, las células HEK293 fueron transfectadas con un editor de exones HTT +/- moléculas híbridas duales miARN-1 HTT. Las células se recolectaron 48 horas después de la transfección, se sometieron a ensayo para determinar las eficiencias del silenciamiento de *HTT* y en el trans-empalme por RT-qPCR (paneles superiores) o se sometieron al análisis de transferencia Western (paneles inferiores). La molécula híbrida dual del editor de exones HTT y miARN-1 HTT (SEQ ID NO: 354), comprendía la SEQ ID NO: 341 y la SEQ ID NO: 204.

Ejemplo 3: Editores de exones de ARN con inhibidores de la expansión de CAG somáticos

Se diseñó y probó un mecanismo potencial de un enfoque terapéutico híbrido para tratar la HD. El enfoque terapéutico híbrido combinó agentes que inhiben la expansión somática de CAG (por ejemplo, mediante la reducción de MSH3) con editores de exones de ARN dirigidos al pre-ARNm de *HTT* (Figura 19). Los editores de exones de ARN dirigidos al pre-ARNm *HTT* sirvieron para reemplazar cualquier ARN *HTT* mutante que pudo haberse producido a partir del ADN que escapó a la inhibición del proceso de expansión somática.

Se evaluó un mecanismo de acción potencial de un editor de exones de ARN con dominio de unión en tándem dirigido al pre-ARNm *HTT* y *MSH3*. (Figura 20). La expresión de un editor de exones fue dirigida por un promotor de CMV. Tales editores de exón ejemplares comprendieron la UTR 5' de *HTT*, la etiqueta de extremo N 3X FLAG, la secuencia codificante del exón 1 de *HTT*, el sitio donante de empalme, un enlazador, un dominio de unión a *MSH3* (dirigido, por ejemplo, al intrón 5 o al intrón 15 del pre-ARNm *MSH3*), un dominio de unión a *HTT* (por ejemplo, *HTT_intron1_11704_100*) y una secuencia terminadora. El dominio de unión de *HTT* apuntó al editor de exones para producir el ARN *HTT* corregido después de un trans-empalme exitoso, mientras que el dominio de unión de *MSH3* apuntó al editor de exones para producir una molécula de ARN quimérica *HTT* exón 1-*MSH3* con un codón de parada prematuro que estuvo sujeto a una desintegración mediada por sin sentido (NMD) y condujo a la reducción posterior de la expresión de *MSH3*.

Los editores de exones de ARN del dominio de unión en tándem dirigidos a *HTT* y *MSH3* exhiben un trans-empalme exitoso a ambos pre-ARNm (Figura 21A-21C). Se realizó RT-qPCR de eficiencia de trans-empalme en el objetivo (ONT) de *HTT* (a través del dominio de unión de *HTT*) (Figura 21A), eficiencia de trans-empalme quimérico *HTT-MSH3* (a través del dominio de unión de *MSH3*) (Figura 21B), y nivel de expresión de transcripción de ARN *MSH3* (Figura 21C), en células HEK293 transfectadas con el editor de exones de ARN de dominio de unión en tándem dirigido al intrón 1 de *HTT* y al intrón 5 de *MSH3*.

Los editores de exones de ARN del dominio de unión en tándem dirigidos a *HTT* y *MSH3* exhiben un trans-empalme exitoso a ambos pre-ARNm (FIG 22A-22C). Se realizó RT-qPCR de eficiencia de trans-empalme en el objetivo (ONT) de *HTT* (a través del dominio de unión de *HTT*) (Figura 22A), eficiencia de trans-empalme quimérico *HTT-MSH3* (a través del dominio de unión de *MSH3*) (Figura 22B), y nivel de expresión de transcripción de ARN *MSH3* (Figura 22C), en células HEK293 transfectadas con el editor de exones de ARN de dominio de unión en tándem dirigido al intrón 1 de *HTT* y al intrón 15 de *MSH3*.

Se evaluó un mecanismo de acción de un editor de exones de ARN con dominio de unión en tándem dirigido al intrón 2 de *HTT* y el pre-ARNm *MSH3* (Figura 30). La expresión del editor de exones estaba dirigida por un promotor CMV y contenía la UTR 5' de *HTT*, la etiqueta de extremo N 3X FLAG, la secuencia codificante de los exones 1 y 2 de *HTT*, el sitio donante de

empalme, el enlazador, el dominio de unión a MSH3 (dirigido al intrón 5 o al intrón 15 del pre-ARNm *MSH3*), el dominio de unión a HTT (HTT_intron2_12061_150) y una secuencia terminadora. El dominio de unión de HTT apuntó al editor de exones para producir el ARN *HTT* corregido después de un trans-empalme exitoso, mientras que el dominio de unión de MSH3
 5 apuntó al editor de exones para producir una molécula de ARN quimérica *HTT* exón 1+2 - *MSH3* con un codón de parada prematuro que estuvo sujeto a una desintegración mediada por sin sentido (NMD) y condujo a la reducción posterior de la expresión de MSH3.

Se produjeron los perfiles de RT-qPCR de trans-empalme de HTT y producción de quimera
 10 HTT-MSH3 (a través de trans-empalme de MSH3) en editores de exones del dominio de unión en tándem (Figura 31A-31B). Se realizó RT-qPCR de eficiencia de trans-empalme en el objetivo (ONT) de HTT (a través del dominio de unión dirigido al intrón 2 de *HTT*) (Figura 31A) y eficiencia de trans-empalme quimérico HTT-MSH3 (a través del dominio de unión dirigido al intrón 5 de *MSH3*) (Figura 31B), en células HEK293 transfectadas con el editor de exones de
 15 ARN de dominio de unión en tándem dirigido al intrón 2 de *HTT* y al intrón 5 de *MSH3*. El dominio de unión dirigido al intrón 2 de *HTT* fue HTT_intron2_12061_150 para todos los editores de exones probados aquí, mientras que MSH3_intron5_213_100 y MSH3_intron5_188_150 se probaron para los dominios de unión dirigidos al intrón 5 de MSH3. Los dominios de unión se posicionaron en tándem y el orden de los dominios de unión fue el
 20 indicado. También se probó un editor de exones para cada dominio de unión de MSH3 con el terminador de triple hélice MALAT1 colocado entre los dos dominios de unión en tándem.

Se produjeron los perfiles de RT-qPCR de trans-empalme de HTT y producción de quimera
 25 HTT-MSH3 (a través de trans-empalme de MSH3) en editores de exones del dominio de unión en tándem (Figura 32A-32B). Se realizó RT-qPCR de eficiencia de trans-empalme en el objetivo (ONT) de HTT (a través del dominio de unión dirigido al intrón 2 de *HTT*) (Figura 32A) y eficiencia de trans-empalme quimérico HTT-MSH3 (a través del dominio de unión dirigido al intrón 15 de *MSH3*) (Figura 32B), en células HEK293 transfectadas con el editor de exones de
 30 ARN de dominio de unión en tándem dirigido al intrón 2 de *HTT* y al intrón 15 de *MSH3*. El dominio de unión dirigido al intrón 2 de *HTT* fue HTT_intron2_12061_150 para todos los editores de exones probados aquí, mientras que MSH3_intron15_6523_120 y MSH3_intron15_6498_150 se probaron para los dominios de unión dirigidos al intrón 15 de *MSH3*. Los dominios de unión se posicionaron en tándem y el orden de los dominios de unión fue el indicado. También se probó un editor de exones para cada dominio de unión de MSH3
 35 con el terminador de triple hélice MALAT1 colocado entre los dos dominios de unión en tándem.

MSH3 puede ser silenciado por miARN que atacan al ARNm de *MSH3* y degradan la transcripción (Figura 33). Se realizó el análisis de transferencia Western de construcciones de
 40 ARNi dirigidas al exón 23 de *MSH3* (Figura 34). Se analizaron las construcciones que contenían la secuencia activa del miARN dirigido al exón 23 de *MSH3*

TTAATCCATAACTCCTTGC (SEQ ID NO: 224). Se realizó el análisis ImageJ en los análisis de transferencia Western para analizar el silenciamiento de la proteína MSH3 (panel superior). Se diseñaron y probaron ARNhc impulsados por el promotor U6 y miméticos de pri-miARN impulsados por el promotor CMV. Las variaciones incluyen la ubicación de la hebra (brazo 5' o brazo 3') de la hebra guía, incluyendo un bulto en la estructura del tallo y variando el andamiaje de miARN. Los controles negativos incluyen construcciones que contienen una secuencia no dirigida o un control sin bucle de horquilla.

Se realizó el análisis de transferencia Western y RT-qPCR de las construcciones ARNi dirigidas a *MSH3* (Figura 35). Se analizaron construcciones que contienen miARN dirigidos a diferentes regiones de la transcripción *MSH3*. Se realizó el análisis ImageJ en los análisis de transferencia Western para analizar el silenciamiento de la proteína MSH3. Se diseñaron y probaron imitadores de pri-miARN impulsados por el promotor CMV dirigidos a diferentes secuencias exónicas de *MSH3*.

Se probó el silenciamiento de *MSH3* mediante ARN antisentido (ARNas) basado en ARN nuclear pequeño (ARNnp) (Figura 36). *MSH3* puede ser inactivado por ARN antisentido codificados en un andamiaje de ARNnp que se unen a las uniones de empalme de *MSH3*, evitando la inclusión de exones. Esto conduce a la omisión de exones y a la generación de un codón de parada prematuro, lo que en última instancia causa NMD de la transcripción *MSH3*. Se midieron los niveles relativos de expresión de ARN de *MSH3* de las uniones exón 1 - exón 2 y exón 2 - exón 3 en moduladores de empalme de *MSH3* que apuntan a la omisión del exón 2 (Figura 37). Se midieron los niveles relativos de expresión de ARN de *MSH3* de las uniones exón 2 - exón 3 y exón 3 - exón 4 en moduladores de empalme de *MSH3* que apuntan a la omisión del exón 3 (Figura 38). Se midieron los niveles relativos de expresión de ARN de *MSH3* de las uniones exón 3 - exón 4 y exón 4 - exón 5 en moduladores de empalme de *MSH3* que apuntan a la omisión del exón 4 (Figura 39). Se midieron los niveles relativos de expresión de ARN de *MSH3* de las uniones exón 2 - exón 3 y exón 3 - exón 4 en moduladores de empalme de *MSH3* que apuntan a la omisión del exón 3 (Figura 40). Se midieron los niveles relativos de expresión de ARN de *MSH3* de las uniones exón 5 - exón 6 y exón 6 - exón 7 en moduladores de empalme de *MSH3* que apuntan a la omisión del exón 6 (Figura 41). Se midieron los niveles relativos de expresión de ARN de *MSH3* de las uniones exón 6 - exón 7 y exón 7 - exón 8 en moduladores de empalme de *MSH3* que apuntan a la omisión del exón 7 (Figura 42). La transcripción del modulador de empalme fue la SEQ ID NO: 331. El ARN antisentido (ARNas) comprendido en este fue In7/Ex7-1 (región ARNas SEQ ID NO: 310) + enlazador + Ex7/In6-1 (región ARNas SEQ ID NO: 311). Se midieron los niveles relativos de expresión de ARN de *MSH3* de las uniones exón 7 - exón 8 y exón 8 - exón 9 en moduladores de empalme de *MSH3* que apuntan a la omisión del exón 8 (Figura 43). Se midieron los niveles relativos de expresión de ARN de *MSH3* de las uniones exón 14 - exón 15 y exón 15 - exón 16 en moduladores de empalme de *MSH3* que apuntan a la omisión del exón 15 (Figura 44).

Los moduladores de empalme del exón 7 de *MSH3* muestran la reducción de ARN *MSH3* y los niveles de proteína (Figura 45A-C). En resumen, las células HEK293 se transfectaron con moduladores de empalme basados en ARNnp diseñados para omitir el exón 7 de *MSH3*. Las células se recolectaron 48 horas después de la transfección, se analizaron para determinar el silenciamiento de *MSH3* mediante RT-qPCR (Figura 45B) o se sometieron a análisis de transferencia Western (Figura 45C).

Reducción de *MSH3* por modulación de empalme en combinación con trans-empalme HTT probado (Figura 51). Los resultados mostraron el desempeño de las moléculas híbridas duales del editor de exones HTT + modulador de empalme *MSH3* (Figura 52A-B). En resumen, las células HEK293 fueron transfectadas con un editor de exones HTT +/- modulador de empalme *MSH3*. Las células se recolectaron 48 horas después de la transfección y se analizaron los perfiles de trans-empalme (Figura 52A) y los perfiles de silenciamiento de *MSH3* (Figura 52B) por RT-qPCR. Las moléculas híbridas duales del Modulador de empalme *MSH3* y el Editor de exones HTT fueron el Modulador de empalme *MSH3* + Editor de exones HTT (SEQ ID NO: 356), que comprendía SEQ ID NO: 331 y SEQ ID NO: 204.

Se realizó una comparación de AAV autocomplementario (scAAV) con AAV de cadena simple (ssAAV) utilizando una molécula híbrida dual de editor de exones HTT + modulador de empalme *MSH3* (Figura 53). En resumen, las células HEK293 fueron transducidas con scAAV o ssAAV que expresan una molécula híbrida dual: editor de exones HTT + modulador de empalme con omisión del exón 7 *MSH3*. Se utilizó el serotipo AAV2. Las células se recolectaron 48 horas después de la transducción y se sometieron a análisis de transferencia RT-qPCR y Western. La molécula híbrida dual de Modulador de empalme *MSH3* y Editor de exones HTT (SEQ ID NO: 357), comprendía SEQ ID NO: 331 y SEQ ID NO: 204, en una orientación de cabeza a cabeza.

Se evaluó el rendimiento de las moléculas híbridas triples editor de exones HTT + miARN HTT + modulador de empalme *MSH3* en comparación con sus controles (Figura 54; paneles superior, medio e inferior). En resumen, se transfectaron células HEK293 con editor de exones HTT +/- miARN-1 HTT o miARN-2 HTT +/- modulador de empalme *MSH3*. Las células se cosecharon 48 horas después de la transfección y se analizaron para el % de transcripciones trans-empalmadas HTT (Figura 54; panel superior), perfiles de silenciamiento de HTT (Figura 54; panel del medio), y perfiles de silenciamiento de *MSH3* por RT-qPCR (Figura 54; panel inferior). También se analizaron moléculas híbridas de control que comprendían un editor de exones con una mutación donante de empalme, un modulador de empalme con una secuencia de ARNas codificada o un miARN con una secuencia de ARNas codificada. El valor "1" para el miARN HTT indica el miARN HTT-1; el valor "2" para el miARN HTT indica que se utilizó el miARN HTT-2. SM hace referencia a mutante de empalme. Scr hace referencia a control desordenado. Se utilizó miR-33 para el andamiaje de miARN. La molécula híbrida dual del editor de exones HTT y miARN-1 HTT (SEQ ID NO: 354), comprendía la SEQ ID NO: 341 y la SEQ ID NO: 204. La molécula híbrida dual del editor de exones HTT y miARN-2 HTT (SEQ ID

- NO: 355), comprendía la SEQ ID NO: 344 y la SEQ ID NO: 204. Las moléculas híbridas duales del Modulador de empalme MSH3 + el Editor de exones HTT comprendieron el Modulador de empalme MSH3 + Editor de exones HTT (SEQ ID NO: 356), que comprendía SEQ ID NO: 331 y SEQ ID NO: 204. La molécula triple híbrida editor de exones HTT + miARN HTT + modulador de empalme MSH3 (SEQ ID NO: 358), comprendía SEQ ID NO: 331, SEQ ID NO: 341 y SEQ ID NO: 204. La molécula triple híbrida editor de exones HTT + miARN HTT + modulador de empalme MSH3 (SEQ ID NO: 359), comprendía SEQ ID NO: 331, SEQ ID NO: 344 y SEQ ID NO: 204.
- Los perfiles de silenciamiento de miARN HTT en iCell GlutaNeurons se midieron mediante RT-ddPCR y análisis de transferencia Western (Figura 56). En resumen, las iCell GlutaNeurons se transdujeron con una molécula híbrida dual de editor de exones HTT + miARN HTT o una molécula híbrida dual de editor de exones HTT + modulador de empalme MSH3, ambas empaquetadas en AAV2.7m8. Las células se recolectaron para RT-ddPCR y análisis de transferencia Western después de 18 días. La molécula híbrida dual del editor de exones HTT y miARN-1 HTT (SEQ ID NO: 354), comprendía la SEQ ID NO: 341 y la SEQ ID NO: 204. La molécula híbrida dual de Modulador de empalme MSH3 + Editor de exones HTT (SEQ ID NO: 357), comprendía SEQ ID NO: 331 y SEQ ID NO: 204, en una orientación de cabeza a cabeza.
- Se produjo un perfil del silenciamiento de MSH3 del modulador de empalme MSH3 en iCell GlutaNeurons medido por RT-ddPCR y análisis de transferencia Western (Figura 57). En resumen, las iCell GlutaNeurons se transdujeron con una molécula híbrida dual de editor de exones HTT + miARN-1 HTT o una molécula híbrida dual de editor de exones HTT + modulador de empalme MSH3, ambas empaquetadas en AAV2.7m8. Las células se recolectaron para RT-ddPCR y análisis de transferencia Western después de 18 días.
- Se produjeron los perfiles de trans-empalme de HTT en el cerebro del ratón BAC-CAG (Figura 58). En resumen, se realizaron inyecciones neonatales de ICV (a $1E+11$ o $3E+11$ vg/animal) en ratones BAC-CAG con las moléculas HD indicadas empaquetadas en AAV9. La corteza y el cuerpo estriado del ratón se recolectaron 4 semanas después de la inyección y se perfilaron las eficiencias del reemplazo de exones HTT por trans-empalme mediante RT-ddPCR y análisis de transferencia Western. Como se muestra en la Figura 58, el análisis de los perfiles de trans-empalme de HTT demostró que se logró un reemplazo de HTT de más del 30 % en el cerebro del ratón mediante inyección de ICV neonatal. El editor de exones impulsado por el promotor CAGGS superó a su equivalente promotor CMV, como lo indica un mayor porcentaje de reemplazo de HTT tanto en la corteza como en el estriado. Entre los editores de exones impulsados por el promotor CMV, se observó una clara respuesta a la dosis, en la que los animales que recibieron $3E+11$ vg tuvieron un mayor porcentaje de reemplazo de HTT que los que recibieron $1E+11$ vg. Significativamente, la proteína HTT de longitud completa trans-empalmada se detectó mediante transferencia Western contra la etiqueta FLAG de extremo N. Este experimento también indicó que, al aumentar los niveles de transcripciones del Editor de

exones, se logró un mayor nivel de trans-empalme (Figura 59), lo que sugiere que identificar un promotor más fuerte o administrar una dosis más alta podría resultar en una actividad del Editor de exones aún mayor. Además, se observó una relación entre el número de copias de ARN del editor de exones y la eficiencia de trans-empalme (% de reemplazo de HTT) en el cerebro del ratón BAC-CAG (FIG 59).

Si bien en la presente se han mostrado y descrito modalidades preferidas de la presente descripción, será evidente para los expertos en la técnica que dichas modalidades se proporcionan únicamente a modo de ejemplo. Los expertos en la técnica podrán realizar numerosas variaciones, cambios y sustituciones sin apartarse de la divulgación. Se entenderá que se pueden emplear varias alternativas a las modalidades de la presente descripción en la práctica de la presente descripción. Se pretende que las reivindicaciones siguientes definan el alcance de la descripción y que los métodos y estructuras dentro del alcance de estas reivindicaciones y sus equivalentes sean abarcados por estas.

VII. MODALIDADES ENUMERADAS

1. Una molécula de trans-empalme de ácido nucleico que comprende:

(a) un dominio codificante que comprende el exón 1 de HTT y el exón 2 de HTT;

(b) un dominio de empalme; y

(c) un dominio de unión que se une a un intrón objetivo de un pre-ARNm de HTT, en donde el intrón objetivo comprende el intrón 2.

2. La molécula de trans-empalme de ácido nucleico de la modalidad 1, en donde el dominio de unión comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO: 60-81.

3. Una molécula de trans-empalme de ácido nucleico que comprende:

(a) un dominio codificante que comprende el exón 1 de HTT;

(b) un dominio de empalme; y

(c) un dominio de unión que se une a un intrón objetivo de un pre-ARNm de HTT, en donde el intrón objetivo comprende el intrón 1, y en donde el dominio de unión comprende cualquiera de las SEQ ID NO: 8-19.

3. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades anteriores, en donde el dominio de codificación comprende, consiste esencialmente en o consiste en el exón 1 de HTT o el exón 1 de HTT y el exón 2 de HTT.

4. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades anteriores, en donde el dominio de codificación comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO: 3 o 59.

5. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades anteriores, en donde el dominio de codificación, el dominio de empalme, y el dominio de unión están unidos operativamente en una dirección de 5' a 3'.

5 6. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades anteriores, que comprende además un enlazador, en donde el dominio de codificación, el dominio de empalme, el enlazador y el dominio de unión están unidos operativamente en una dirección de 5' a 3'.

10 7. Una molécula de trans-empalme de ácido nucleico de la modalidad 6, en donde el enlazador comprende, consiste esencialmente en o consiste en:

una secuencia que varía de 20 a 50 nucleótidos de longitud, en donde el enlazador comprende 60-80 % de guaninas intercaladas con timidina/uridina; 65-75 % de guaninas intercaladas con timidina/uridina; o 66-74 % de guaninas intercaladas con timidina/uridina;

15 una secuencia que varía de 20 a 45 nucleótidos de longitud, en donde el enlazador comprende 60-80 % de guaninas intercaladas con timidina/uridina; 65-75 % de guaninas intercaladas con timidina/uridina; o 66-74 % de guaninas intercaladas con timidina/uridina; o

20 una secuencia que varía de 22 a 42 nucleótidos de longitud, en donde el enlazador comprende 60-80 % de guaninas intercaladas con timidina/uridina; 65-75 % de guaninas intercaladas con timidina/uridina; o 66-74 % de guaninas intercaladas con timidina/uridina.

8. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 6 o 7, en donde el enlazador comprende, consiste esencialmente en o consiste en cualquiera de:

SEQ ID NO: 38 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 38;

25 SEQ ID NO: 39 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 39;

SEQ ID NO: 40 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 40; o SEQ ID NO: 41 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 41.

30 9. La molécula de trans-empalme de ácido nucleico de la modalidad 6, en donde el enlazador comprende, consiste esencialmente en o consiste en cualquiera de:

SEQ ID NO: 42 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 42;

SEQ ID NO: 43 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 43;

SEQ ID NO: 44 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 44;

SEQ ID NO: 45 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 45;

35 SEQ ID NO: 46 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 46;

SEQ ID NO: 106 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 106;

SEQ ID NO: 107 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 107;

SEQ ID NO: 108 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 108;

SEQ ID NO: 109 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 109;

40 SEQ ID NO: 110 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 110;

SEQ ID NO: 111 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 111;

0

SEQ ID NO: 112 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 112.

5 10. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades anteriores, que comprenden además un terminador de triple hélice, en donde el dominio de codificación, el dominio de empalme, el enlazador, cuando están presentes, el dominio de unión, y el terminador de triple hélice están unidos operativamente en una dirección de 5' a 3'.

10 11. La molécula de trans-empalme de ácido nucleico de la modalidad 10, en donde el terminador de triple hélice comprende, consiste esencialmente o consiste en la SEQ ID NO: 5 o una secuencia que tiene al menos 90 % de identidad con la SEQ ID NO: 5.

15 12. La molécula de trans-empalme de ácido nucleico de la modalidad 11, en donde el terminador de triple hélice comprende, consiste esencialmente en o consiste en la SEQ ID NO: 6.

20 13. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades anteriores, que comprende además una región no traducida 5' (5' UTR), en donde el 5'UTR, el terminador de hélice triple, en donde el dominio de codificación, el dominio de empalme, el enlazador, cuando está presente, el dominio de unión, y el terminador de hélice triple, cuando está presente, están unidos operativamente en una dirección de 5' a 3'

25 14. La molécula de trans-empalme de ácido nucleico de la modalidad 13, en donde la 5' UTR es una *HTT* 5' UTR.

15. La molécula de trans-empalme de ácido nucleico de la modalidad 14, en donde la 5' UTR *HTT* comprende, consiste esencialmente o consiste en la SEQ ID NO: 136 o una secuencia que tiene al menos 90 % de identidad con la SEQ ID NO: 136.

30 16. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades anteriores, que comprende además una secuencia que codifica una etiqueta de epítipo, en donde la 5' UTR, cuando está presente, la etiqueta de epítipo, el dominio de codificación, el dominio de empalme, el enlazador, cuando está presente, el dominio de unión, y el terminador de hélice triple, cuando está presente, están unidos operativamente en una dirección de 5' a 3'.

40 17. La molécula de trans-empalme de ácido nucleico de la modalidad 16, en donde la secuencia que codifica la etiqueta de epítipo comprende, consiste esencialmente o conste en cualquiera de una SEQ ID NO: 4.

18. Una molécula de trans-empalme de ácido nucleico que comprende:

- (a) un dominio codificante que comprende el exón 1 de HTT y el exón 2 de HTT;
- (b) un dominio de empalme;
- (c) un enlazador; y
- (d) un dominio de unión que se une a un intrón objetivo de un pre-ARNm de HTT, en donde el intrón objetivo comprende el intrón 2.

19. La molécula de trans-empalme de ácido nucleico de la modalidad 18, en donde el dominio de unión comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO: 60-81.

20. Una molécula de trans-empalme de ácido nucleico que comprende:

- (a) un dominio codificante que comprende el exón 1 de HTT;
- (b) un dominio de empalme;
- (c) un enlazador; y
- (d) un dominio de unión que se une a un intrón objetivo de un pre-ARNm de HTT, en donde el intrón objetivo comprende el intrón 1.

21. La molécula de trans-empalme de ácido nucleico de la modalidad 20, en donde el dominio de unión comprende, consiste esencialmente en o consiste en cualquiera de las SEQ ID NO: 8-21.

22. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 18-21, en donde el enlazador comprende, consiste esencialmente o consiste en cualquiera de las SEQ ID NO:

- SEQ ID NO: 38 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 38;
- SEQ ID NO: 39 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 39;
- SEQ ID NO: 40 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 40;
- SEQ ID NO: 41 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 41;
- SEQ ID NO: 42 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 42;
- SEQ ID NO: 43 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 43;
- SEQ ID NO: 44 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 44;
- SEQ ID NO: 45 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 45;
- SEQ ID NO: 46 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 46;
- SEQ ID NO: 106 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 106;
- SEQ ID NO: 107 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 107;
- SEQ ID NO: 108 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 108;
- SEQ ID NO: 109 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 109;
- SEQ ID NO: 110 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 110;
- SEQ ID NO: 111 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 111;
- o
- SEQ ID NO: 112 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 112.

23. La molécula de trans-empalme de ácido nucleico de la modalidad 22, que comprende además un terminador de triple hélice, en donde el dominio de codificación, el dominio de empalme, el enlazador, el dominio de unión y el terminador de triple hélice están unidos operativamente en una dirección de 5' a 3'; y opcionalmente, comprende además un 5' UTR, en donde el 5' UTR, cuando está presente, el dominio de codificación, el dominio de empalme, el enlazador, el dominio de unión y el terminador de triple hélice, cuando está presente, están unidos operativamente en una dirección de 5' a 3'.
24. Una molécula de trans-empalme de ácido nucleico que comprende un enlazador, en donde el enlazador comprende, consiste esencialmente en, o consiste en una secuencia que varía de 20 a 50 nucleótidos de longitud, en donde el enlazador comprende 60-80 % de guaninas intercaladas con timidina/uridina; 65-75 % de guaninas intercaladas con timidina/uridina; o 66-74 % de guaninas intercaladas con timidina/uridina.
25. La molécula de trans-empalme de ácido nucleico de la modalidad 24, en donde el enlazador comprende, consiste esencialmente en, o consiste en una secuencia que varía de 20 a 45 nucleótidos de longitud, en donde el enlazador comprende 60-80 % de guaninas intercaladas con timidina/uridina; 65-75 % de guaninas intercaladas con timidina/uridina; o 66-74 % de guaninas intercaladas con timidina/uridina.
26. La molécula de trans-empalme de ácido nucleico de la modalidad 24 o la modalidad 25, en donde el enlazador comprende, consiste esencialmente en, o consiste en una secuencia que varía de 22 a 42 nucleótidos de longitud, en donde el enlazador comprende 60-80 % de guaninas intercaladas con timidina/uridina; 65-75 % de guaninas intercaladas con timidina/uridina; o 66-74 % de guaninas intercaladas con timidina/uridina.
27. La molécula de trans-empalme de ácido nucleico de las modalidades 24-26, en donde el enlazador comprende, consiste esencialmente en o consiste en:
- SEQ ID NO: 38 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 38; SEQ ID NO: 39 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 39; SEQ ID NO: 40 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 40; o SEQ ID NO: 41 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 41.
28. Una molécula de trans-empalme de ácido nucleico que comprende un enlazador, en donde el enlazador comprende, consiste esencialmente en, o consiste en:
- SEQ ID NO: 42 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 42; SEQ ID NO: 43 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 43; SEQ ID NO: 44 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 44; SEQ ID NO: 45 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 45; SEQ ID NO: 46 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 46;

- SEQ ID NO: 106 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 106;
 SEQ ID NO: 107 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 107;
 SEQ ID NO: 108 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 108;
 SEQ ID NO: 109 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 109;
 5 SEQ ID NO: 110 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 110;
 SEQ ID NO: 111 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 111;
 o
 SEQ ID NO: 112 o una secuencia que tiene al menos 90 % de identidad con SEQ ID NO: 112.
- 10 29. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-28, que comprende además un dominio de unión que se une a un intrón objetivo de un pre-ARNm *MSH3* o una segunda molécula de trans-empalme de ácido nucleico que comprende un dominio de unión que se une a un intrón objetivo de un pre-ARNm *MSH3*.
- 15 30. La molécula de trans-empalme de ácido nucleico de la modalidad 29, en donde el intrón objetivo de *MSH3* comprende cualquiera del intrón 5 o el intrón 15 de *MSH3*.
- 20 31. La molécula de trans-empalme de ácido nucleico o la segunda molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 29 o 30, en donde el dominio de unión que se une a un intrón objetivo de un pre-ARNm de *MSH3* comprende, consiste esencialmente en, o consiste en cualquiera de las SEQ ID NO: 140, 142, 143, 144, 146 o 147.
- 25 32. La molécula de trans-empalme de ácido nucleico o una segunda molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 29-31, en donde la molécula de trans-empalme de ácido nucleico comprende cualquiera de las SEQ ID NO: 149 - 154.
33. Una molécula de trans-empalme de ácido ribonucleico que comprende cualquiera de las SEQ ID NO: 23-36, 47-56, 83-105, o 113-125.
- 30 34. Una molécula de trans-empalme del ácido ribonucleico transcrita de la molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-33.
- 35 35. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-34, en donde el pre-ARNm de HTT endógeno comprende al menos una mutación asociada con una enfermedad de Huntington (HD).
- 40 36. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-35, en donde la al menos una mutación asociada con HD comprende una expansión de repeticiones de CAG en un alelo del gen *HTT*.
37. La molécula de trans-empalme de ácido nucleico de la modalidad 36, en donde la

expansión de repeticiones CAG en un alelo del gen *HTT* comprende más de 35 repeticiones CAG.

- 5 38. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 35-37, en donde la al menos una mutación asociada con HD es autosómica dominante.
- 10 39. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 36-38, en donde la al menos una mutación asociada con HD se expresa en al menos una de las neuronas piramidales corticales, las neuronas espinosas medianas estriatales o las neuronas hipotalámicas.
40. Un vector que comprende la molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-39.
- 15 41. El vector de la modalidad 40, en donde el vector comprende un dominio regulador 5' unido de forma operativa 5' al dominio de codificación.
- 20 42. El vector de la modalidad 41, en donde el dominio regulador 5' está unido de forma operativa a una región no traducida 5'.
43. El vector de cualquiera de las modalidades 41-42, en donde el dominio regulador 5' comprende un promotor constitutivo o un promotor específico de tejido.
- 25 44. El vector de la modalidad 43, en donde el promotor constitutivo es un promotor CMV.
- 30 45. Un plásmido provírico que comprende la molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-39.
46. Un virus adenoasociado (AAV) que comprende la molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-39, en donde el AAV comprende opcionalmente un dominio regulador 5' unido de forma operativa 5' a la molécula de trans-empalme de ácido nucleico.
- 35 47. El AAV de la modalidad 46, en donde el AAV comprende un dominio regulador 5' unido de forma operativa 5' al dominio de codificación.
48. El AAV de cualquiera de las modalidades 46 o 47, en donde el dominio regulador 5' está unido de forma operativa a una región no traducida 5'.
- 40 49. El AAV de cualquiera de las modalidades 46-48, en donde el dominio regulador 5' comprende un promotor constitutivo.

50. El AAV de la modalidad 49, en donde el promotor constitutivo es un promotor CMV.
51. El AAV de cualquiera de las modalidades 46 a 50, en donde el AAV exhibe tropismo neuronal.
52. El AAV de cualquiera de las modalidades 46-50, en donde el AAV es AAV9, AAV8, AAV5, o AAV2.
53. Una composición que comprende la molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-39, el vector de cualquiera de las modalidades 40-44, el plásmido provírico de la modalidad 45 o el AAV de cualquiera de las modalidades 46-52.
54. La composición de la modalidad 53, que comprende un excipiente farmacéuticamente aceptable.
55. La composición de cualquiera de las modalidades 53 o 54, que comprende además al menos un oligonucleótido antisentido o una construcción que codifica al menos un ARN antisentido que inhibe el cis-empalme del pre-ARNm de HTT.
56. La composición de la modalidad 55, en donde el al menos un oligonucleótido antisentido comprende cualquiera de las SEQ ID NO: 126-135 o la construcción que codifica el al menos un ARN antisentido se une a una secuencia objetivo unida por cualquiera de las SEQ ID NO: 126-135.
57. La composición de cualquiera de las modalidades 55-56, en donde el al menos un oligonucleótido antisentido comprende la SEQ ID NO: 131 o la construcción que codifica el al menos un ARN antisentido se une a una secuencia objetivo unida por la SEQ ID NO: 131.
58. Un método para expresar HTT biológicamente activo en una célula objetivo para restaurar los niveles funcionales de la proteína HTT en la célula objetivo, dicho método comprende la transducción de la célula objetivo con la molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-39, el vector de cualquiera de las modalidades 40-44, el plásmido provírico de la modalidad 45, el AAV de cualquiera de las modalidades 46-52, o la composición de cualquiera de las modalidades 53-57.
59. El método de la modalidad 58, en donde se reemplaza al menos el 5%, al menos el 10%, al menos el 15%, al menos el 20%, al menos el 25%, al menos el 30%, al menos el 35%, al menos el 40% o al menos el 45% del pre-ARNm de HTT que comprende al menos una mutación asociada con HD en la célula objetivo.

60. El método de la modalidad 59, en donde se reemplaza al menos 50%, al menos el 55%, al menos el 60%, al menos el 65%, al menos el 70%, al menos el 75%, al menos el 80 %, al menos el 85% del pre-ARNm de HTT que comprende al menos una mutación asociada con HD en la célula objetivo.

- 5
61. El método de la modalidad 60, en donde se reemplaza al menos el 90%, al menos el 95%, al menos el 96%, al menos el 97%, al menos el 98%, al menos el 99% o el 100% del pre-ARNm de HTT que comprende al menos una mutación asociada con HD en la célula objetivo.
- 10
62. El método de cualquiera de las modalidades 58-61, en donde los niveles funcionales de HTT se restauran en la célula objetivo mediante la expresión de la proteína HTT biológicamente funcional.
- 15
63. Un método para reducir la expresión de HTT que comprende una repetición de poliglutamina que excede 35 residuos de glutamina consecutivos en un sujeto, el método comprende transfectar o transducir una célula objetivo, más particularmente una neurona, en el sujeto con la molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-39, el vector de cualquiera de las modalidades 40-44, el plásmido provírico de la modalidad 45, el AAV de cualquiera de las modalidades 46-52, o la composición de cualquiera de las
- 20
- modalidades 53-57.
64. Un método para corregir al menos una mutación en una secuencia de exón HTT en un pre-ARNm de HTT en una célula objetivo de un sujeto, el método comprende administrar al sujeto la molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-39, el
- 25
- vector de cualquiera de las modalidades 40-44, el plásmido provírico de la modalidad 45, el AAV de cualquiera de las modalidades 46-52, o la composición de cualquiera de las modalidades 53-57.
65. Un método para tratar HD en un sujeto que lo necesita, dicho método comprende
- 30
- administrar al sujeto la molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-39, el vector de cualquiera de las modalidades 40-44, el plásmido provírico de la modalidad 45, el AAV de cualquiera de las modalidades 46-52, o la composición de las modalidades 53-57 en una cantidad terapéuticamente eficaz.
- 35
66. El método de cualquiera de las modalidades 58-65, dicho método comprende la administración de la molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-39, el vector de cualquiera de las modalidades 40-44, el plásmido provírico de la modalidad 45, el AAV de cualquiera de las modalidades 46-52, o la composición de cualquiera de las modalidades 53-57 al cerebro del sujeto.
- 40
67. El método de cualquiera de las modalidades 58-66, en donde el sujeto es un mamífero,

preferentemente un roedor, un primate no humano o un ser humano.

68. El método de cualquiera de las modalidades 63-67, en donde el sujeto está genéticamente predispuesto a tener HD o ha sido diagnosticado con HD.

5

69. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-39, el vector de cualquiera de las modalidades 40-44, el plásmido provírico de la modalidad 45, el AAV de cualquiera de las modalidades 46-52, o la composición de cualquiera de las modalidades 53-57 para usar en la prevención o el tratamiento HD en un sujeto que lo necesita.

10

70. La molécula de trans-empalme de ácido nucleico de cualquiera de las modalidades 1-39, el vector de cualquiera de las modalidades 40-44, el plásmido provírico de la modalidad 45, el AAV de cualquiera de las modalidades 46-52, o la composición de cualquiera de las modalidades 53-57 para usar en la preparación de un medicamento para el tratamiento o prevención de HD en un sujeto que lo necesita.

15

71. Un método que comprende introducir en una célula una molécula de trans-empalme de ácido nucleico configurada para empalmarse tanto a un primer pre-ARNm objetivo como a un segundo pre-ARNm objetivo, en donde el empalme al primer pre-ARNm objetivo corrige un defecto en el primer pre-ARNm objetivo, y en donde el empalme al segundo pre-ARNm objetivo introduce un defecto en el segundo pre-ARNm objetivo.

20

72. El método de la modalidad 71, en donde la molécula de trans-empalme de ácido nucleico comprende un primer dominio de unión configurado para dirigirse a un intrón del primer pre-ARNm objetivo y un segundo dominio de unión configurado para dirigirse a un intrón del segundo pre-ARNm objetivo.

25

73. El método de la modalidad 71 o 72, en donde la molécula de trans-empalme de ácido nucleico comprende además una secuencia de dominio codificante que comprende una secuencia funcional de uno o más exones del primer pre-ARNm objetivo que corrige el defecto en el primer pre-ARNm objetivo.

30

74. El método de cualquiera de las modalidades 71 a 73, en donde el defecto en el segundo pre-ARNm objetivo comprende un cambio de marco en la secuencia de codificación del segundo pre-ARNm objetivo.

35

75. El método de la modalidad 74, en donde el cambio de marco crea un codón de terminación prematura en el segundo pre-ARNm objetivo.

76. El método de cualquiera de las modalidades 71 a 75, en donde el defecto comprende el codón de inicio endógeno del segundo pre-ARNm objetivo que se está eliminando.

40

77. El método de cualquiera de las modalidades 71 a 76, en donde el defecto comprende un UTR 5' insertado que impide la traducción de la proteína codificada por el segundo pre-ARNm objetivo.

5

78. El método de cualquiera de las modalidades 71 a 77, en donde el defecto comprende un UTR 3' insertado que desestabiliza el pre-ARNm o impide la exportación del segundo pre-ARNm objetivo desde el núcleo.

10

79. El método de cualquiera de las modalidades 71 a 78, en donde el defecto comprende la eliminación de una tapa 5' o una cola de poliA 3' del segundo pre-ARNm objetivo.

80. El método de cualquiera de las modalidades 71 a 79, en donde el defecto causa una desintegración mediada por sin sentido del segundo pre-ARNm objetivo.

15

81. El método de cualquiera de las modalidades 71 a 80, en donde la introducción hace que la abundancia del producto génico del segundo pre-ARNm objetivo en la célula se reduzca en comparación con la abundancia del producto génico antes de la introducción.

20

82. Un método para reducir la abundancia de una proteína en una célula, comprendiendo el método la introducción en la célula de una molécula de trans-empalme de ácido nucleico que introduce un defecto en un pre-ARNm que codifica la proteína.

83. El método de la modalidad 82, en donde el defecto comprende uno o más de los siguientes:

25

(a) un cambio de marco introducido en la secuencia codificante del pre-ARNm;
 (b) eliminación del codón de inicio endógeno del pre-ARNm;
 (c) introducción de un codón de parada prematuro en la secuencia codificante del pre-ARNm;
 (d) sustitución de la secuencia codificante endógena del pre-ARNm por una secuencia codificante alternativa;

30

(e) inserción de un UTR 5' que impide la traducción de la secuencia codificante endógena del pre-ARNm;

(f) inserción de un UTR 3' que desestabiliza el pre-ARNm;

(g) inserción de 3' UTR que impide que el pre-ARNm se exporte desde el núcleo;

(h) eliminación de una tapa 5' del pre-ARNm; o

35

(i) eliminación de una cola de poliA 3' del pre-ARNm.

84. El método de la modalidad 82 o 83, en donde la proteína es MSH3.

85. El método de la modalidad 84, en donde la molécula de trans-empalme de ácido nucleico comprende un dominio de unión que se une a un intrón del pre-ARNm.

40

86. El método de la modalidad 85, en donde la molécula de trans-empalme de ácido nucleico comprende una secuencia heteróloga de dominio de codificación.