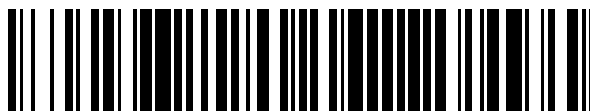


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 759 317**

51 Int. Cl.:

C12N 15/113 (2010.01)

A61K 31/7088 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.05.2015 PCT/US2015/031783**

87 Fecha y número de publicación internacional: **26.11.2015 WO15179525**

96 Fecha de presentación y número de la solicitud europea: **20.05.2015 E 15795373 (8)**

97 Fecha y número de publicación de la concesión europea: **04.09.2019 EP 3146051**

54 Título: **Compuestos terapéuticos para la enfermedad de Huntington**

30 Prioridad:
20.05.2014 US 201462000895 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
08.05.2020

73 Titular/es:
**UNIVERSITY OF IOWA RESEARCH FOUNDATION
(100.0%)
2660 University Capitol Center
Iowa City, IA 52242, US**

72 Inventor/es:
**DAVIDSON, BEVERLY, L. y
MAS MONTEYS, ALEJANDRO**

74 Agente/Representante:
SUGRAÑES MOLINÉ, Pedro

ES 2 759 317 T3

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

DESCRIPCIÓN

Compuestos terapéuticos para la enfermedad de Huntington

5 ANTECEDENTES DE LA INVENCION

[1] La iARN dirige el silenciamiento de genes específico de secuencia mediante un ARN de hebra doble (ARNhd) que es procesado a ARN inhibitorios pequeños funcionales (de aproximadamente 21nt). En la naturaleza, la iARN para la regulación de la expresión de genes se produce principalmente por medio de ARN conocidos como microARN (miARN). Los microARN maduros (aproximadamente entre 19 y 25 nucleótidos) son procesados a partir de transcritos primarios de miARN grandes (pri-miARN) que contienen regiones tallo-bucle. Mediante una serie de eventos de procesamiento catalizados por ribonucleasas, Drosha y Dicer, se libera la región dúplex del miARN y luego se incorpora una hebra individual (la hebra "guía" antisentido) en el complejo silenciador inducido por ARN (RISC), generando de esta manera un complejo funcional con capacidad de formar un apareamiento de bases y silenciar los transcritos diana. El modo de represión de diana depende principalmente del grado de complementariedad; la escisión del transcripto habitualmente requiere un grado alto de apareamiento de bases, mientras que la represión de la traducción y la desestabilización del ARNm se produce cuando los ARN pequeños se unen de manera imperfecta a los transcritos diana (frecuentemente en la UTR 3'). De hecho, para los últimos, pueden ser suficientes tramos cortos de complementariedad – tan cortos como de 6 pb – para producir el silenciamiento génico.

[2] En WO 2012/109667 A1 se describen vectores basados en microARN para la administración de moléculas de ARNi para la terapia de la enfermedad de Huntington.

25 SÍNTESIS DE LA INVENCION

[3] La presente invención se relaciona con un ácido nucleico que codifica para un transcripto primario de miARN artificial (pri-miARN) que consiste en, en orden de posición, una región flanqueante 5', una región no guía, una región bucle, una región guía, y una región flanqueante 3', en donde la región guía consiste en SEQ ID NO: 37 (miHDss3), SEQ ID NO:6 (miHDS1v5U) o SEQ ID NO:7 (miHDS1v6A), y la región no guía es por lo menos 80% complementaria con la región guía.

[4] La presente invención se relaciona además con un ARN codificado por el ácido nucleico de la invención.

[5] La presente invención se relaciona además con un cassette de expresión que codifica para el ácido nucleico aislado de la invención.

[6] La presente invención se relaciona además con un vector que comprende el cassette de expresión de la invención.

[7] La presente invención se relaciona además con un ácido nucleico aislado de entre 80 y 4000 nucleótidos de longitud, que comprende un ácido nucleico que codifica para un transcripto primario de miARN artificial (pri-miARN) que consiste en, en orden de posición, una región flanqueante 5', una región no guía, una región bucle, una región guía, y una región flanqueante 3', en donde la región guía consiste en SEQ ID NO: 37 (miHDss3), SEQ ID NO:6 (miHDS1v5U) o SEQ ID NO:7 (miHDS1v6A), y la región no guía es por lo menos 80% complementaria con la región guía.

[8] La presente invención se relaciona además con un ácido nucleico aislado que consiste en Pri-miHDS1v5U (SEQ ID NO:8), Pri-miHDS1v6A (SEQ ID NO:9), Pre-miHDS1v5U (SEQ ID NO:10), o Pre-miHDS1v6A (SEQ ID NO:11).

[9] La presente invención se relaciona además con el ácido nucleico de la invención, el cassette de expresión de la invención, o el vector de la invención para usar en el tratamiento de la enfermedad de Huntington.

[10] Se describe un vector lanzadera de miARN aislado que expresa un ARNip terapéutico con toxicidad fuera de diana limitada. En determinadas formas de realización, la incorporación de un ARNip que exhibe toxicidad fuera de diana en el contexto de un vector lanzadera de miARN de la presente divulgación limita la toxicidad fuera de diana del ARNip. En determinadas formas de realización, el vector lanzadera de miARN expresa un ARNip terapéutico en el cerebro con toxicidad fuera de diana limitada. En determinadas formas de realización, el vector lanzadera de miARN expresa un ARNip terapéutico en el cuerpo estriado con toxicidad fuera de diana limitada. En determinadas formas de realización, el vector lanzadera de miARN expresa un ARNip terapéutico en el cerebro con toxicidad fuera de diana limitada.

[11] Se describe un ácido nucleico aislado que codifica para un transcripto primario (pri-miARN) que incluye, en orden de posición, una región flanqueante 5', una región no guía (pasajero), una región bucle, una región guía, y una región flanqueante 3', en donde la región guía consiste en SEQ ID NO: 37 (miHDss3), SEQ ID NO:6

(miHDS1v5U) o SEQ ID NO:7 (miHDS1v6A), y la región no guía es por lo menos 80% complementaria con la región guía. En determinadas formas de realización, la región no guía es por lo menos 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% complementaria con la región guía. En determinadas formas de realización, la región flanqueante 5' está unida de manera contigua a la región no guía, la región bucle está posicionada entre la región no guía y la región guía, y la región guía está unida de manera contigua a la región flanqueante 3'. Como se usa en la presente, el término "región guía de ARNip" es una secuencia de hebra simple de ARN que es complementaria a la secuencia diana. Como se usa en la presente, el término "región no guía de ARNip" es una secuencia de hebra simple de ARN que es complementaria a la "región guía de ARNip". Por lo tanto, en las condiciones apropiadas, la región guía del ARNip y la región no guía del ARNip se asocian para formar un dúplex de ARN. Como se usa en la presente, todas las secuencias de ácido nucleico están indicadas, como es costumbre, en una dirección 5' a 3'.

[12] En determinadas formas de realización, la región flanqueante 5' contiene una secuencia de unión 5' unida de manera contigua a la región no guía. Como se usa en la presente, el término "sitio de unión" o una "secuencia de unión" es una secuencia de ácido nucleico corta menor de 60 nucleótidos que conecta dos secuencias de ácido nucleico diferentes. En determinadas formas de realización, el sitio de unión tiene una longitud de cualquier entero entre 4 y 50, inclusive. En determinadas formas de realización, la secuencia de unión 5' consiste en entre 5 y 8 nucleótidos (por ejemplo, consiste en 6 nucleótidos). En determinadas formas de realización, la secuencia de unión 5' codifica para GUGAGCGA (SEQ ID NO:13) o GUGAGCGC (SEQ ID NO:14).

[13] En determinadas formas de realización, la región flanqueante 5' comprende además una secuencia protuberante 5' posicionada corriente arriba a la secuencia de unión 5'. Como se usa en la presente, el término "secuencia protuberante" es una región del ácido nucleico que no es complementaria con el ácido nucleico opuesto en un dúplex. Por ejemplo, un dúplex contendrá una región de ácidos nucleicos complementarios, luego una región de ácidos nucleicos no complementarios, seguido por una segunda región de ácidos nucleicos complementarios. Las regiones de ácidos nucleicos complementarios se unirán entre sí, mientras que la región no complementaria central no se unirá, formando de esta manera una "protuberancia". En determinadas formas de realización las dos hebras del ácido nucleico posicionadas entre las dos regiones complementarias serán de diferente longitud, formando de esta manera una "protuberancia". En determinadas formas de realización, la secuencia protuberante 5' contendrá entre 2 y 15 nucleótidos. En determinadas formas de realización, la secuencia protuberante 5' consiste en aproximadamente entre 1 y 10 nucleótidos. En determinadas formas de realización, la secuencia protuberante 5' codifica para UAAACUCGA (SEQ ID NO:15). En determinadas formas de realización, la secuencia protuberante 5' tiene entre 0 y 50% de complementariedad con la secuencia protuberante 3'.

[14] En determinadas formas de realización, la región flanqueante 5' contiene además una secuencia espaciadora 5' posicionada corriente arriba a la secuencia protuberante 5'. En determinadas formas de realización, la secuencia espaciadora 5' consiste en entre 9 y 12 nucleótidos, tal como entre 10 y 12 nucleótidos. En determinadas formas de realización, la secuencia espaciadora 5' tiene entre 60 y 100% de complementariedad con una secuencia espaciadora 3'. En determinadas formas de realización, la secuencia protuberante 5' comprende un sitio de clonado, tal como un sitio XhoI. En determinadas formas de realización, la secuencia espaciadora 5' es UGGUACCGUU (SEQ ID NO:16).

[15] En determinadas formas de realización, la región flanqueante 5' contiene además una secuencia 5' corriente arriba posicionada corriente arriba a la secuencia espaciadora 5'. En determinadas formas de realización, la secuencia 5' corriente arriba tiene aproximadamente entre 5 y 5000 nucleótidos de longitud, tal como entre 30 y 2000 nucleótidos de longitud.

[16] En determinadas formas de realización, la región flanqueante 3' contiene una secuencia de unión 3' unida de manera contigua a la región guía. En determinadas formas de realización, el sitio de unión tiene una longitud de cualquier entero entre 4 y 50, inclusive. En determinadas formas de realización, la secuencia de unión 3' consiste en entre 5 y 8 nucleótidos, (por ejemplo, consiste en 6 nucleótidos). En determinadas formas de realización, la secuencia de unión 3' es por lo menos aproximadamente 85% complementaria con una secuencia de unión 5'. En determinadas formas de realización, la secuencia de unión 3' codifica para CGCYUAC (SEQ ID NO:17), en donde Y es C o U. En determinadas formas de realización, la secuencia de unión 3' codifica para CGCCUAC (SEQ ID NO:18).

[17] En determinadas formas de realización, la región flanqueante 3' comprende además una secuencia protuberante 3' posicionada corriente abajo de la secuencia de unión 3'. En determinadas formas de realización, la secuencia protuberante 3' comprende un sitio de clonado, tal como un sitio SpeI/XbaI o un sitio SpeI. En determinadas formas de realización, la secuencia protuberante 3' consiste en aproximadamente entre 1 y 15 nucleótidos (tal como entre 2 y 15 nucleótidos o entre 1 y 10 nucleótidos). En determinadas formas de realización, la secuencia protuberante 3' codifica para UAG (SEQ ID NO: 30). En determinadas formas de realización, la secuencia protuberante 5' es complementaria a la secuencia protuberante 3' en solo un nucleótido en cada extremo de la secuencia.

[18] En determinadas formas de realización, la región flanqueante 3' contiene además una secuencia

espaciadora 3' posicionada corriente abajo de la secuencia protuberante 3'. En determinadas formas de realización, la secuencia espaciadora 3' consiste en entre 9 y 12 nucleótidos, tal como entre 10 y 12 nucleótidos. En determinadas formas de realización, la secuencia espaciadora 3' es AGCGGCCGCCA (SEQ ID NO:19). En determinadas formas de realización, la secuencia espaciadora 3' es por lo menos aproximadamente 70% complementaria con una secuencia espaciadora 5'.

[19] En determinadas formas de realización, la región flanqueante 3' contiene además una secuencia 3' corriente abajo posicionada corriente abajo de la secuencia espaciadora 3'. En determinadas formas de realización, una secuencia 5' corriente arriba no se aparea significativamente con la secuencia 3' corriente abajo. Como se usa en la presente, el término "no se aparea significativamente con" se refiere a que las dos hebras tienen menos del 20% de homología. En determinadas formas de realización, la secuencia 3' corriente abajo tiene aproximadamente entre 5 y 5000 nucleótidos de longitud, tal como entre 30 y 2000 nucleótidos de longitud.

[20] En determinadas formas de realización, la región bucle tiene entre 4 y 20 nucleótidos de longitud, tal como entre 15 y 19 nucleótidos de longitud. Entre el 0 y 50% de la región bucle puede ser complementaria con la otra parte de la región bucle. Como se usa en la presente, el término "región bucle" es una secuencia que une dos hebras complementarias de ácido nucleico. En determinadas formas de realización, entre 1 y 3 nucleótidos de la región bucle que son inmediatamente contiguos a las hebras complementarias de ácido nucleico pueden ser complementarios con los últimos 1 a 3 nucleótidos de la región bucle. Por ejemplo, los dos primeros ácidos nucleicos en la región bucle pueden ser complementarios con los dos últimos nucleótidos de la región bucle. En determinadas formas de realización, la región bucle tiene 17 nucleótidos de longitud. En determinadas formas de realización, la región bucle codifica para CUNNNNNNNNNNNNNNGG (SEQ ID NO:20) o CC- NNNNNNNNNNNNNNGG (SEQ ID NO:21). En determinadas formas de realización, la región bucle codifica para CUGUGAAGCCACA- GAUGGG (SEQ ID NO:22) o CCGUGAAGCCACAGAUGGG (SEQ ID NO:23).

[21] Se describe un ARN codificado por el ácido nucleico descrito en la presente.

[22] Además, se describe un cassette de expresión que contiene un promotor unido de manera contigua a un ácido nucleico descrito en la presente. En determinadas formas de realización, el promotor es un promotor de polII o a polIII, tal como un promotor U6 (por ejemplo, un promotor U6 de ratón). En determinadas formas de realización, el cassette de expresión contiene además un gen marcador. En determinadas formas de realización, el promotor es un promotor de polII. En determinadas formas de realización, el promotor es un promotor específico de tejido. En determinadas formas de realización, el promotor es un promotor inducible. En determinadas formas de realización, el promotor es un promotor de polIII.

[23] En determinadas formas de realización, el cassette de expresión además comprende un gen marcador.

[24] Se describe un vector que contiene un cassette de expresión descrito en la presente. En determinadas formas de realización, el vector es un vector de virus adenoasociado (AAV). En determinadas formas de realización, el AAV es AAV1, AAV2, AAV5, AAV6 y/o AAV9. En determinadas formas de realización, el AAV es AAV2. En determinadas formas de realización, el AAV es AAV2/1. Los ejemplos de dichos AAV se encuentran en Davidson y col., PNAS (2000) 97:3428-3432. En determinadas formas de realización, el AAV es AAV2/1. En determinadas formas de realización, el AAV es AAV2/5. Como se usa en la presente, el término AAV2/1 se usa para referirse a una ITR de AAV2 y cápside de AAV1, el término AAV2/2 es una ITR de AAV2 y cápside de AAV2, el término AAV2/4 es una ITR de AAV2 y cápside de AAV4, etc. En determinadas formas de realización, el AAV es AAV1, AAV2, AAV5, AAV6 y/o AAV9. En determinadas formas de realización, el AAV es AAV1. En determinadas formas de realización, el AAV es AAV2. En determinadas formas de realización, el AAV es AAV5. En determinadas formas de realización, el AAV es un AAV6. En determinadas formas de realización, el AAV es un AAV8. En determinadas formas de realización, el AAV es un AAV9. En determinadas formas de realización, el AAV es un AAVrh10.

[25] En determinadas formas de realización, la cápside de AAV tiene por lo menos 80% de homología con cualquier proteína de cápside de serotipo de AAV de referencia VP1, VP2 y/o VP3, por ejemplo, con una proteína de cápside de AAV1 VP1, VP2 y/o VP3, o por ejemplo, con una proteína de cápside de AAV2 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV3 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV4 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV5 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV6 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV7 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV8 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV9 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAVrh10 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAVrh74 VP1, VP2 y/o VP3.

[26] En determinadas formas de realización, la cápside de AAV tiene 100% de homología con cualquier proteína de cápside de serotipo de AAV de referencia VP1, VP2 y/o VP3, por ejemplo, con una proteína de cápside de AAV1 VP1, VP2 y/o VP3, o por ejemplo, con una proteína de cápside de AAV2 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV3 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV4 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV5 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV6 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV7 VP1, VP2 y/o VP3, o por ejemplo, una

proteína de cápside de AAV8 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAV9 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAVrh10 VP1, VP2 y/o VP3, o por ejemplo, una proteína de cápside de AAVrh74 VP1, VP2 y/o VP3.

5 [27] Se describe un animal no humano que comprende el ácido nucleico, el cassette de expresión, el vector o dúplex descrito en la presente.

10 [28] La presente invención provee un ácido nucleico aislado de entre 80 y 4000 nucleótidos de longitud que comprende un ácido nucleico que codifica para un transcrito primario de miARN artificial (pri-miARN) que consiste en, en orden de posición, una región flanqueante 5', una región no guía, una región bucle, una región guía, y una región flanqueante 3', en donde la región guía consiste en SEQ ID NO: 37 (miHDss3), SEQ ID NO:6 (miHDS1v5U) o SEQ ID NO:7 (miHDS1v6A), y la región no guía es por lo menos 80% complementaria con la región guía.

15 [29] Se describe un ácido nucleico aislado que consiste en Pri-miHDS1v5U (SEQ ID NO:8), Pri-miHDS1v6A (SEQ ID NO:9), Pre-miHDS1v5U (SEQ ID NO:10), o Pre-miHDS1v6A (SEQ ID NO:11). En una forma de realización, un miHDS1 de longitud completa (SEQ ID NO:12) tiene la siguiente secuencia:

20 5'-GCGUUUAGUGAACCGUCAGAUGGUACCGUUUAAACUCGAGUGAGCGAUGCUGGCUCGCAUGGUC
GAUACUGUAAAGCCACAGAUGGGUGUCGACCAUGCGAGCCAGCACCGCCUACUAGAGCGGCCGC-3'
(SEQ ID NO:12)

25 [30] Se describe un dúplex de ARN aislado que comprende una región guía de ácido nucleico y una región no guía de ácido nucleico, en donde la región guía es SEQ ID NO: 37 (miHDss3), SEQ ID NO:6 (miHDS1v5U) o SEQ ID NO:7 (miHDS1v6A) y la región no guía es por lo menos 80% complementaria con la región guía. En determinadas formas de realización, el dúplex de ARN aislado tiene entre 19 y 30 pares de bases de longitud. Determinadas formas de realización incluyen un cassette de expresión que codifica para el ácido nucleico aislado descrito precedentemente. En determinadas formas de realización el cassette de expresión además comprende un gen marcador.

30 [31] Se describe un método para inducir interferencia de ARN mediante la administración a un sujeto de un ácido nucleico, un cassette de expresión, un vector, o una composición descrita en la presente.

35 [32] Se describe un vector que contiene un promotor U6 ligado operativamente a un ácido nucleico que codifica para un miARN. Los sitios de inicio de transcripción predichos de construcciones de la presente divulgación son diferentes de aquellos usados por los investigadores en el pasado. En determinadas formas de realización, el U6miARN tiene un extremo 5' extendido. Si el extremo 5' se trunca para que se parezca a la estrategia previa basada en CMV, la eficacia del silenciamiento se reduce de manera severa. También se describen secuencias flanqueantes mejoradas que presentan una eficacia mejorada respecto a las secuencias flanqueantes de miR-30 naturales. El uso de la presente estrategia de miARN parece aliviar la toxicidad asociada con las estrategias tradicionales de ARNhp. La estrategia de miARN en general no genera cantidades excesivas de ARNi como lo hacen las estrategias de U6ARNhp.

45 [33] Como se usa en la presente el término "secuencia de tallo" es una secuencia que es complementaria con otra secuencia en la misma molécula, en donde las dos hebras complementarias se alinean para formar un dúplex (por ejemplo, las regiones no guía y guía). El dúplex que se forma puede ser completamente complementario, o puede ser menos que completamente complementario, tal como 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81%, 80%, 75%, o 70% de complementariedad entre sí. Además, en determinadas formas de realización, una hebra puede contener más nucleótidos que la otra hebra, permitiendo la formación de un bucle lateral.

50 [34] Se describen vectores que contienen los cassettes de expresión descritos en la presente. Los ejemplos de vectores apropiados incluyen vectores adenovirales, lentivirales, virales adenoasociados (AAV), poliovirus, virus de herpes simple (HSV), o virales basados en Maloney murino. En una forma de realización, el vector es un vector de virus adenoasociado. Estos cassettes y vectores pueden estar contenidos en una célula, tal como una célula de mamífero. Un mamífero no humano puede contener el cassette o vector.

55 [35] Se describen células (tales como una célula de mamífero) que contienen moléculas de ácido nucleico, cassettes de expresión o vectores descritos en la presente. También se describe un mamífero no humano que contiene las moléculas de ácido nucleico, cassettes de expresión o vectores descritos en la presente.

60 [36] Se describe un ácido nucleico, un cassette de expresión, un vector, o una composición como se describe en la presente para usar en terapia, tal como para tratar una enfermedad neurodegenerativa.

65 [37] Se describe una molécula de ARNi aislado que tiene un microARN que tiene una sobreextensión en el extremo 3'. En determinadas formas de realización, la sobreextensión es una repetición de entre 2 y 5 nucleótidos. En determinadas formas de realización, la sobreextensión es una secuencia UU (SEQ ID NO:24), UUU (SEQ ID

NO:25), UUUU (SEQ ID NO:26), CUU (SEQ ID NO:27), CUUU (SEQ ID NO:28) o CUUUU (SEQ ID NO:29). En determinadas formas de realización, el microARN es un microARN de origen natural. En determinadas formas de realización, microARN es un microARN artificial. En determinadas formas de realización, la molécula de ARNi produce un nivel disminuido de toxicidad fuera de diana.

5 [38] Se describe un método para inducir una interferencia de ARN de baja toxicidad mediante la administración a un sujeto de un ácido nucleico, un cassette de expresión, un vector, o una composición como se describe en la presente. En determinadas formas de realización, el cassette de expresión contiene un promotor de polII.

10 [39] Se describe un método para inducir interferencia de ARN de baja toxicidad mediante la administración a un sujeto de un cassette de expresión que codifica para un promotor de polII ligado operativamente a un ácido nucleico que codifica para un miARN. En determinadas formas de realización, el miARN comprende una sobreextensión en 5' o 3' de 2 o 3 nucleótidos. En determinadas formas de realización, el miARN comprende una sobreextensión en 3' de 2 nucleótidos. En determinadas formas de realización, el miARN es un miARN artificial.

15 [40] Se describe un método para tratar un sujeto con enfermedad de Huntington mediante la administración al sujeto de un ácido nucleico, un cassette de expresión, un vector, o una composición como se describe en la presente para tratar la enfermedad de Huntington.

20 [41] Se describe un método para suprimir la acumulación de huntingtina en una célula mediante la introducción de moléculas de ácido nucleico (por ejemplo, un ácido ribonucleico (ARN)) descritas en la presente a la célula en una cantidad suficiente para suprimir la acumulación de huntingtina en la célula. En determinadas formas de realización, la acumulación de huntingtina es suprimida en por lo menos 10%. En determinadas formas de realización, la acumulación de huntingtina es suprimida en por lo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 95%, o 99%. En determinadas formas de realización, la supresión de la acumulación de la proteína es en una cantidad suficiente para provocar un efecto terapéutico, por ejemplo, para reducir la formación de ovillos.

25 [42] Se describe un método para prevenir los efectos citotóxicos de la huntingtina mutante en una célula mediante la introducción de moléculas de ácido nucleico (por ejemplo, un ácido ribonucleico (ARN)) descritas en la presente a la célula en una cantidad suficiente para suprimir la acumulación de huntingtina. En determinadas formas de realización, las moléculas de ácido nucleico previenen los efectos citotóxicos de huntingtina, por ejemplo, en una célula neuronal.

30 [43] Se describe un método para inhibir la expresión de un gen de huntingtina en una célula mediante la introducción de una molécula de ácido nucleico (por ejemplo, un ácido ribonucleico (ARN)) descrita en la presente a la célula en una cantidad suficiente para inhibir la expresión de la huntingtina, y en donde el ARN inhibe la expresión del gen de huntingtina. En determinadas formas de realización, la huntingtina es inhibida en por lo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 95%, o 99%.

35 [44] Se describe un método para inhibir la expresión de un gen de huntingtina en un mamífero (por ejemplo, un humano o un mamífero no humano) mediante (a) la provisión a un mamífero que contiene una célula neuronal, en donde la célula neuronal contiene el gen de huntingtina y la célula neuronal es susceptible a interferencia de ARN, y el gen de huntingtina se expresa en la célula neuronal; y (b) el contacto del mamífero con un ácido ribonucleico (ARN) o un vector descrito en la presente, inhibiendo de esta manera la expresión del gen de huntingtina. En determinadas formas de realización, la acumulación de huntingtina es suprimida en por lo menos 10%. En determinadas formas de realización, la huntingtina es inhibida en por lo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 95%, o 99%. En determinadas formas de realización, la célula está localizada *in vivo* en un mamífero.

40 [45] Se describe un vector viral que comprende un promotor y una lanzadera de microARN (miARN) que contiene un ARNi incorporado específico de una secuencia diana. En determinadas formas de realización, el promotor es un promotor inducible. En determinadas formas de realización, el vector es un vector adenoviral, lentiviral, viral adenoasociado (AAV), poliovirus, HSV, o viral basado en Maloney murino. En determinadas formas de realización, la secuencia diana es una secuencia asociada con la enfermedad de Huntington. La secuencia diana, en determinadas formas de realización, es una secuencia que codifica para huntingtina.

45 [46] Se describe un método para prevenir efectos citotóxicos de la enfermedad neurodegenerativa en un mamífero que lo requiere, mediante la introducción del vector que codifica para un miARN descrito en la presente en una célula en una cantidad suficiente para suprimir la acumulación de una proteína asociada con la enfermedad de Huntington, y en donde el ARN previene los efectos citotóxicos de la enfermedad de Huntington (también referida como HD, y la proteína involucrada es huntingtina, también denominada htt).

50 [47] Se describe un método para inhibir la expresión de una proteína asociada con la enfermedad de Huntington en un mamífero que lo requiere, mediante la introducción del vector que codifica para un miARN descrito en la presente en una célula en una cantidad suficiente para inhibir la expresión de la proteína huntingtina, en donde el ARN inhibe la expresión de la proteína huntingtina. La proteína huntingtina es inhibida en por lo menos 10%, 20%,

30%, 40%, 50%, 60%, 70%, 80%, 90% 95%, o 99%.

[48] Se describen compuestos, composiciones, y métodos útiles para modular la expresión del gen de la enfermedad de Huntington usando moléculas de ácido nucleico de interferencia cortas (ARNip). También se describen compuestos, composiciones, y métodos útiles para modular la expresión y actividad de otros genes involucrados en las vías de la expresión y/o actividad del gen de la HD mediante interferencia de ARN (ARNi) usando moléculas de ácido nucleico pequeño. En particular, la presente divulgación describe moléculas de ácido nucleico pequeño, tales como moléculas de ácido nucleico de interferencia pequeño (ANip), ARN de interferencia pequeño (ARNip), ARN de hebra doble (ARNhd), micro-ARN (miARN), y ARN de horquilla pequeño (ARNhp) y métodos usados para modular la expresión de genes de HD. Una molécula de ARNip de la presente divulgación puede ser, por ejemplo, sintetizada químicamente, expresada a partir de un vector o sintetizada enzimáticamente.

[49] Como se usa en la presente cuando una reivindicación indica un ARN "que corresponde a", se refiere al ARN que tiene la misma secuencia que el ADN, excepto en que el uracilo está sustituido por timina.

[50] Además, se describe un método para silenciar sustancialmente un gen diana de interés o alelo diana para el gen de interés con el objetivo de proveer un efecto terapéutico. Como se usa en la presente el término "silenciar sustancialmente" o "silenciado sustancialmente" se refiere a disminuir, reducir, o inhibir la expresión del gen diana o alelo diana en por lo menos aproximadamente 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85% y 100%. Como se usa en la presente el término "efecto terapéutico" se refiere a un cambio en las anomalías asociadas del estado de la enfermedad, que incluye déficits patológicos y del comportamiento; un cambio en el tiempo de progresión del estado de la enfermedad; una reducción, atenuación, o alteración de un síntoma de la enfermedad; o una mejora en la calidad de vida de la persona afectada con la enfermedad. Los efectos terapéuticos pueden ser medidos de manera cuantitativa por un médico o cualitativa por un paciente afectado con el estado de enfermedad blanco del ARNip. En determinadas formas de realización en donde se silencia sustancialmente el alelo mutante y el de tipo salvaje, el término efecto terapéutico define una condición en la cual el silenciamiento de la expresión del alelo de tipo salvaje no tiene un efecto deletéreo o dañino sobre las funciones normales tal que el paciente podría no tener un efecto terapéutico.

[51] En una forma de realización, se describe un método para tratar o prevenir la enfermedad de Huntington en un sujeto u organismo que comprende poner en contacto el sujeto u organismo con un ARNip de la divulgación en condiciones adecuadas para modular la expresión del gen de HD en el sujeto u organismo de manera que puede lograrse el tratamiento o prevención de la enfermedad de Huntington. En una forma de realización, el blanco del gen de HD comprende ambos alelos de HD (por ejemplo, un alelo que comprende una expansión de repetición de trinucleótido (CAG) y un alelo de tipo salvaje). La molécula de ARNip de la divulgación puede expresarse a partir de vectores como se describe en la presente o de cualquier otra manera conocida en el arte para dirigirla a tejidos o células apropiadas en el sujeto u organismo.

[52] En una forma de realización, se describe un método para tratar o prevenir la enfermedad de Huntington en un sujeto u organismo que comprende, poner en contacto al sujeto u organismo con una molécula de ARNip de la divulgación mediante administración local a tejidos o células relevantes, tales como células y tejidos del cerebro (por ejemplo, ganglios basales, cuerpo estriado, o corteza), por ejemplo, mediante la administración de vectores o cassettes de expresión de la divulgación que proveen moléculas de ARNip de la divulgación a células relevantes (por ejemplo, ganglios de la base, cuerpo estriado, o corteza). En una forma de realización, se administra el ARNip, vector, o cassette de expresión al sujeto u organismo mediante administración aumentada estereotáctica o por convección al cerebro. Por ejemplo, en la Patente de los EE.UU. No. 5.720.720 se proveen métodos y dispositivos útiles para la administración aumentada estereotáctica y por convección de reactivos al cerebro. Dichos métodos y dispositivos pueden ser usados fácilmente para la administración de los ARNip, vectores, o cassettes de expresión de la invención a un sujeto u organismo. En las Solicitudes de Patente de los EE.UU. Nos. 2002/0141980; 2002/0114780; y 2002/0187127 se proveen métodos y dispositivos útiles para la administración aumentada estereotáctica y por convección de reactivos que pueden adaptarse fácilmente para la administración de los ARNip, vectores, o cassettes de expresión de la invención a un sujeto u organismo. Los dispositivos particulares que pueden ser útiles para la administración de los ARNip, vectores, o cassettes de expresión de la invención a un sujeto u organismo se describen por ejemplo en la Solicitud de Patente de los EE.UU. No. 2004/0162255. La molécula de ARNip de la divulgación puede ser expresada a partir de vectores como se describe en la presente o de cualquier otra manera conocida en el arte para dirigirla a tejidos o células apropiadas en el sujeto u organismo.

[53] Los métodos de administración de vectores virales incluyen, a título enunciativo no taxativo, las rutas intraarterial, intramuscular, intravenosa, intranasal y oral. En general, los viriones de AAV pueden introducirse en las células del SNC usando técnicas de transducción *in vivo* o *in vitro*. Si se transducen *in vitro*, la célula receptora deseada será extraída del sujeto, transducida con los viriones de AAV y reintroducida en el sujeto. Como alternativa, pueden usarse células singénicas o xenogénicas en donde aquellas células no generarán una respuesta inmunológica inapropiada en el sujeto.

[54] Se han descrito métodos adecuados para la administración e introducción de las células transducidas en un sujeto. Por ejemplo, las células pueden ser transducidas *in vitro* por combinación de los viriones de AAV

recombinantes con las células del SNC por ejemplo, en medio apropiado, y el cribado de aquellas células que portan el ADN de interés puede realizarse usando técnicas convencionales tales como transferencia Southern y/o PCR, o mediante el uso de marcadores de selección. Las células transducidas pueden ser formuladas luego en composiciones farmacéuticas, descritas con mayor detalle más adelante, e introducción de la composición en el sujeto mediante diferentes técnicas, tales como injerto, inyección intramuscular, intravenosa, subcutánea e intraperitoneal.

[55] En una forma de realización, para la administración *in vivo*, los viriones de AAV se formulan en composiciones farmacéuticas y en general se administrarán por vía parenteral, por ejemplo, para inyección intramuscular directamente en el músculo esquelético o cardíaco o mediante inyección en el SNC.

[56] En una forma de realización, los vectores virales de la divulgación se administran al SNC mediante sistemas mejorados de convección (CED) que pueden administrar de manera eficaz vectores viral, por ejemplo, AAV, en regiones grandes del cerebro de un sujeto (por ejemplo, cuerpo estriado y/o corteza). Tal como se describe en detalle y se ejemplifica más adelante, estos métodos son adecuados para una variedad de vectores virales, por ejemplo vectores de AAV que portan genes terapéuticos (por ejemplo, ARNip).

[57] Cualquier dispositivo de administración por convección mejorada puede ser apropiado para la administración de vectores virales. En una forma de realización, el dispositivo es una bomba osmótica o una bomba de infusión. Las bombas osmóticas y de infusión se encuentran disponibles de una variedad de proveedores, por ejemplo Alzet Corporation, Hamilton Corporation, Aiza, Inc. (Palo Alto, Calif.). Habitualmente, un vector viral se administra por medio de dispositivos CED como a continuación. Se inserta un catéter, cánula y otro dispositivo de inyección en el tejido del SNC en el sujeto seleccionado. A la luz de las divulgaciones en la presente, una persona con experiencia en el arte podría determinar fácilmente cuál área general del SNC es un blanco apropiado. Por ejemplo, cuando se administra un vector AAV que codifica para un gen terapéutico para tratar la HD, el cuerpo estriado es un área adecuada del cerebro como blanco. Hay disponibles mapas estereotácticos y dispositivos de posicionamiento, por ejemplo de ASI Instruments, Warren, Mich. El posicionamiento también puede llevarse a cabo con el uso de mapas anatómicos obtenidos mediante CT y/o imágenes de MRI del cerebro del sujeto para ayudar a guiar el dispositivo de inyección al objetivo seleccionado. Además, debido a que los métodos descritos en la presente pueden ser realizados de manera que áreas relativamente grandes del cerebro incorporen los vectores virales, se requieren menos cánulas de infusión. Debido a que las complicaciones quirúrgicas están relacionadas con el número de penetraciones, los métodos descritos en la presente también son útiles para reducir los efectos secundarios observados con las técnicas de administración convencionales.

[58] En una forma de realización, las composiciones farmacéuticas comprenderán suficiente material genético para producir una cantidad terapéuticamente eficaz del ARNip de interés, es decir, una cantidad suficiente para reducir o aliviar los síntomas del estado de la enfermedad de interés o una cantidad suficiente para conferir el beneficio deseado. Las composiciones farmacéuticas también pueden contener un excipiente farmacéuticamente aceptable. Dichos excipientes incluyen cualquier agente farmacéutico que no induce el mismo la producción de anticuerpos dañinos para el individuo que recibe la composición, y el cual puede administrarse sin toxicidad indebida. Los excipientes farmacéuticamente aceptables incluyen, a título enunciativo no taxativo, sorbitol, Tween80, y líquidos tales como agua, solución salina, glicerol y etanol. También pueden incluir sales farmacéuticamente aceptables, por ejemplo, sales de ácidos minerales tales como clorhidratos, bromhidratos, fosfatos, sulfatos, y similares; y las sales de ácidos orgánicos tales como acetatos, propionatos, malonatos, benzoatos, y similares. Además, las sustancias auxiliares, tales como agentes humectantes o agentes emulsificantes, sustancias amortiguadoras de pH, y similares, pueden estar presentes en dichos vehículos. En REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Pub. Co., N.J. 1991) se encuentra disponible una descripción profunda de los excipientes farmacéuticamente aceptables.

[59] Tal como resulta evidente para aquellos con experiencia en el arte en vista de las divulgaciones de esta memoria descriptiva, una cantidad eficaz del vector viral que debe agregarse puede determinarse de manera empírica. La administración puede realizarse en una dosis, de manera continua o intermitente durante el curso del tratamiento. Los métodos para determinar los medios más eficaces y las dosificaciones de administración son bien conocidos para aquellos con experiencia en el arte y variarán con el vector viral, la composición de la terapia, las células diana, y el sujeto que está siendo tratado. Las administraciones individuales y múltiples pueden ser realizadas con el nivel de dosis y el patrón que es seleccionado por el médico a cargo.

[60] Se debe comprender que el vector viral administrado podría expresar más de un transgén. Como alternativa, también se pueden administrar vectores por separado, en donde cada uno expresa uno o más transgenes diferentes, al SNC como se describe en la presente. Además, también se pretende que los vectores virales administrados por los métodos de la presente invención se combinen con otras composiciones y terapias adecuadas.

[61] Además, se describen miARN o ARNhp, un cassette de expresión y/o un vector como se describe en la presente para usar en el tratamiento médico o diagnóstico.

[62] Se describe el uso de un miARN o ARNhp, un cassette de expresión y/o un vector como se describe en la presente para preparar un medicamento útil para tratar una condición susceptible al ARNi en un animal, por ejemplo, útil para tratar la enfermedad de Huntington.

5 [63] También se describe un ácido nucleico, cassette de expresión, vector, o composición de la divulgación para usar en terapia.

[64] También se describe un ácido nucleico, cassette de expresión, vector, o composición de la divulgación para tratar, por ejemplo, para usar en el tratamiento profiláctico o terapéutico de la enfermedad de Huntington.

10 [65] En determinadas formas de realización, el agente se administra al cerebro del sujeto. En determinadas formas de realización, el agente se administra directamente al cerebro o a través del torrente sanguíneo. En determinadas formas de realización, el agente terapéutico se administra por vía intracraneal. En determinadas formas de realización, el agente terapéutico se administra a la cisterna magna, cuerpo estriado, corteza o ventrículo, espacio subaracnoideo y/o espacio intratecal del sujeto. En determinadas formas de realización, el sujeto es humano. En determinadas formas de realización, el sujeto es un mamífero no humano. En determinadas formas de realización, el agente se inyecta en 1 a 5 ubicaciones en el cerebro, tal como en una, dos, o tres ubicaciones en el cerebro. En determinadas formas de realización, el método además comprende la administración adicional del rAAV en el ventrículo del cerebro del primate no humano, espacio subaracnoideo y/o espacio intratecal. Más específicamente, se describe un método para administrar un ácido nucleico a una célula en contacto con el LCR circulante, tal como una célula ependimaria, una célula pial, célula meníngea, una célula del endotelio cerebral, que comprende administrar a la célula una partícula de AAV que contiene un vector comprendido por el ácido nucleico insertado entre un par de repeticiones terminales invertidas de AAV, administrando de esta manera el ácido nucleico a la célula.

25 [66] También se describe un método para poner en contacto una célula con el ácido nucleico, el cassette de expresión, el vector, o el dúplex descrito en la presente, para tratar la enfermedad de Huntington, en donde la célula es una célula ependimaria, pial, endotelial, del ventrículo del cerebro, y/o meníngea.

30 **DESCRIPCIÓN BREVE DE LAS FIGURAS**

[0067]

35 Figuras 1A-1G: La sobreexpresión de miHDS1 produce efectos adversos en el cerebro de ratón. a) Apareamiento de miHDS1 con el ARNm de huntingtina de ratón y humano (miHDS1 es SEQ ID NO:1; Htt de ratón es SEQ ID NO:2; y Htt de humano es SEQ ID NO:3). b) Gráfico que describe los vectores lanzaderas AAV/stuffer que contienen los cassettes de expresión miHDS1 y miCtl. c) Estrategia experimental para evaluar la tolerabilidad *in vivo* a miHDS1. d) Datos de Rotarod de ratones inyectados con miHDS1 (n=13) o miCtl (n=11). Los datos se presentan como el promedio de los 2 mejores ensayos de cada ratón por día de los cuatro días consecutivos ensayados a las 7 semanas (Basal), 16 semanas y 24 semanas. La latencia de caída se muestra como media±e.s.m. (*p>0,05, test de t por desapareados a los tiempos indicados). e) Análisis de aumento de peso de ratones inyectados con miHDS1 y miCtl. Los datos se muestran como aumento de peso respecto al punto de tiempo basal a las 7 semanas. f) Análisis de *clasp*ing en ratones inyectados con miHDS1 y miCtl. Los datos se muestran como porcentaje y número de ratones que presenten *clasp*ing en los puntos indicados de tiempo. g) Desviación de hebra del vector U6/miHDS1. La desviación de hebra se evaluó por medida de la actividad de luciferasa de construcciones reporteras que contenían secuencias dianas complementarias a las hebras miHDS1 pasajero (sentido) o guía (antisentido). Los resultados son de un experimento representativo de 3 experimentos diferentes (n=4/grupo). Los datos se muestran como media±esm en relación a las células transfectadas con miCtl y demuestran que miHDS1 preferiblemente se carga en la hebra guía de miHDS1.

50 Figuras 2A-2D: caracterización de los genes no diana de miHDS1. a) Lista de genes entre el percentil 25 de genes no diana de miHDS1 predichos. Información presentada: ID del gen, secuencia de referencia, tipo de sitio de unión del miARN, posición de nucleótido 3'UTR, puntuación de contexto de escaneo de diana predicho, puntuación de ddG predicha por el algoritmo PITA. b) Gráfico que describe los sitios de unión de miHDS1:ARNm (miHDS1 es SEQ ID NO: 1) sobre los genes no diana predichos (Bcl2 es SEQ ID NO:31; Smad9 es SEQ ID NO:32; Sdf4 es SEQ ID NO:33; Map2k6 es SEQ ID NO:34). c) Análisis mediante Q-PCR cuantitativa de los niveles de ARNm de *Htt*, *Bcl2*, *Smad9*, *Sdf4* y *Map2k6* en muestras de cuerpo estriado 4 meses después de la inyección de miHDS1. Todas las muestras se normalizaron a *fi-actina* y los resultados son la media±esm en relación a los ratones inyectados con miCtl. (n=6 ratones por grupo; *p<0,05, **p<0,01, prueba de Mann Whitney) d) Análisis mediante Q-PCR cuantitativa de los niveles de ARNm de *Htt*, *Bcl2*, *Smad9*, *Sdf4* y *Map2k6* en células SthdhQ7 después de la electroporación de miHDS1. Todas las muestras se normalizaron a *fi-actina* y los resultados son la media±esm en relación a células electroporadas con plásmido que contiene el promotor U6 o el cassette de expresión miCtl. (n=8 pocillos electroporados; **p<0,01, ANOVA de una vía seguido por una prueba posterior de Bonferroni).

65 Figuras 3A-3C: Generación de variantes de semilla de miHDS1 de nucleótido único. a) Gráfico que describe la ubicación de modificaciones de nucleótido único en la región semilla de la secuencia de miHDS1 (Pri-miHDS1 es

SEQ ID NO:4; Pre-miHDS1 es SEQ ID NO:5; RISC cargado con la secuencia de miARN es SEQ ID NO:1). b) Impacto en la puntuación de SPS que depende de la posición de la no coincidencia de nucleótidos en los no diana de miHDS1. c) La tabla indica el número de genes no diana predichos (totales y específicos de cuerpo estriado) para miHDS1 y las variantes de miHDS1, así como el número de no diana compartidos.

5 Figuras 4A-4I: Eficacia de silenciamiento de variantes de semilla de miHDS1 de nucleótido único. a) Análisis cuantitativo de los niveles de ARNm de hHtt en células HEK293 transfectadas con cassettes de expresión de U6/miHDS1. El ARN total se recolectó 24 horas luego de la transfección y los niveles de hHtt se determinaron mediante Q-PCR. Todas las muestras se normalizaron a *fi-actina* y los resultados son la media±esm en relación a células transfectadas con miHDS1 (n=12 pocillos; **p<0,01, ***p<0,001, ANOVA de una vía seguido por prueba posterior de Bonferroni). b) los cassettes de expresión de miHDS1, miHDS1v5u, miHDS1v6a y miHDS1v7u se transfectaron en células HEK293 de humano, y se determinaron los niveles de proteína huntingtina endógena 48 horas después de la transfección. Se usó miCtl como un control no silenciador y fi-Catenina sirve como control de carga. c) Cuantificación de los niveles de proteína hHtt 48 horas después de la transfección de miHDS1, miHDS1v5u, miHDS1v6a y miHDS1v7u. Los datos son la media±esm en relación a células transfectadas con miCtl (n=6, tres transferencias western diferentes, * p<0,01, prueba de Mann Whitney). d) apareamiento de miHDS1v5u y miHDS1v6a con ARNm de huntingtina de ratón (miHDS1v5U es SEQ ID NO:6; Htt de ratón es SEQ ID NO:2; miHDS1v6A es SEQ ID NO:7). e) los cassettes de expresión miHDS1, miHDS1v5u, y miHDS1v6a se electroporaron en células SthdhQ7 de ratón, y se determinaron los niveles de proteína huntingtina endógena 48 horas después de la electroporación. Se usó miCtl como control no silenciador y fi-Catenina sirve como control de carga. f) Cuantificación de los niveles de proteína mHtt 48 horas después de la electroporación de miHDS1, miHDS1v5u y miHDS1v6a. Los datos son la media±esm en relación a células transfectadas con miCtl (n=6, tres transferencias western diferentes, *p<0,01, prueba de Mann Whitney). g-h-i) Análisis cuantitativo de los niveles de ARNm de mHtt, Bcl2 y Smad9 en células SthdhQ7 electroporadas con los cassettes de expresión U6/miHDS1, U6/miHDS1v5u y U6/miHDS1v6a. Se recolectó el ARN total 24 horas luego de la electroporación y se determinaron los niveles de mHtt, Bcl2 y Smad9 mediante Q-PCR. Todas las muestras se normalizaron a *fi-actina* y los resultados son la media±esm en relación a células transfectadas con promotor que contenía U6 y cassette de expresión U6/miCtl (n=12 pocillos; # p<0,01, ANOVA de una vía seguido por prueba posterior de Bonferroni).

30 Figuras 5A-5F: Generación de secuencias miHDss1-4 dirigidas contra la expresión de huntingtina de humano. a). Se generaron cuatro miARN artificiales iniciadores que contienen la secuencia semilla de miCtl permitiendo una no coincidencia de nucleótido único entre la región semilla y el ARNm de Htt de humano diana. Los sitios de unión de miHDss1 y miHDss4 están localizados en la 3'UTR, mientras que miHDss2 y 3 se unen en el límite del exón 7-8 y exón 33 del ARNm de hHtt, respectivamente. b) Par de unión miARN/ARNm entre miHDss1-4 y ARNm de huntingtina de humano. Se encontraron no coincidencias de nucleótido único en la posición de la región semilla 7,6,5 y 4 para las secuencias de miHDss1, 2, 3 y 4, respectivamente (miHDss1 es SEQ ID NO:35; miHDss2 es SEQ ID NO:36; miHDss3 es SEQ ID NO:37; miHDss4 es SEQ ID NO:38). La Figura también describe las SEQ ID NOS 40-43, respectivamente, en orden de aparición. c) Análisis cuantitativo de los niveles de ARNm de hHtt en células HEK293 transfectadas con los cassettes de expresión U6/miHDss1-4. Se recolectó el ARN total 24 horas luego de la transfección y se determinaron los niveles de hHtt mediante Q-PCR. Todas las muestras se normalizaron a *fi-actina* y los resultados son la media±esm en relación a células transfectadas con miCtl (n=8 pocillos; *p<0,001, ANOVA de una vía seguido por prueba posterior de Bonferroni). d) El cassette de expresión miHDss3 se transfectó en células HEK293 de humano, y se determinaron los niveles de proteína huntingtina endógena 48 horas después de la transfección. Se usó miCtl como control no silenciador y fi-Catenina sirve como control de carga. e) Cuantificación de los niveles de proteína hHtt 48 horas después de la transfección de miHDss3. Los datos son la media±esm en relación a células transfectadas con miCtl (n=6, dos transferencias western diferentes, *p<0,01, prueba de Mann Whitney). f) El algoritmo PITA se usó para determinar la estabilidad de unión de miHDss3 (SEQ ID NO:37) y miCtl (SEQ ID NO:39) en los sitios de unión de ARNm no intencional predichos. La región semilla de miCtl y miHDss3 están resaltadas en negrita. Los datos se muestran como puntuación de ddG (Kcal/mol) para cada gen no diana respecto a miCtl o miHDss3. La predicción de los inventores sugiere que la secuencia 3' de miHDss3 provee más estabilidad de unión en los genes no diana que miCtl.

Figuras 6A-6E: Tolerabilidad *in vivo* de secuencias de miHDS1 variante y miHDss3. a) Estrategia experimental para evaluar la tolerabilidad *in vivo* de nuevo diseño de secuencias de miARN. b) Gráfico que describe los vectores lanzaderas AAV/stuffer que contienen los cassettes de expresión miHDS1 variantes y miHDss3. c) Datos de Rotarod de ratones inyectados con formulación amortiguadora (n=7), miCtl (n=8), miHDS1 (n=9), miHDS1v5u (n=10), miHDS1v6a (n=11) o miHDss3 (n=10). Los datos se presentan como el promedio de los 2 mejores ensayos para cada ratón por día de los cuatro días consecutivos ensayados a las 7 semanas (Basal), 16 semanas y 24 semanas. La latencia de caída se muestra como media±e.s.m en relación a los ratones inyectados con miCtl. (*p<0,05, ANOVA de una vía seguido por prueba posterior de Bonferroni). d) Análisis de aumento de peso de ratones inyectados con formulación amortiguadora, miCtl, miHDS1, miHDS1v5u, miHDS1v6a o miHDss3. Los datos se muestran como aumento de peso respecto al punto de tiempo basal a las 7 semanas. e) Análisis de *clasp*ing de ratones inyectados con miHDS1 y miCtl. Los datos se muestran como porcentaje y número de ratones que presentan *clasp*ing en los puntos de tiempo indicados.

65

DESCRIPCIÓN DETALLADA DE LA INVENCION

[68] La presente invención se relaciona con un ácido nucleico que codifica para un transcripto primario de miARN artificial (pri-miARN) que consiste en, en orden de posición, una región flanqueante 5', una región no guía, una región bucle, una región guía, y una región flanqueante 3', en donde la región guía consiste en SEQ ID NO: 37 (miHDss3), SEQ ID NO:6 (miHDS1v5U) o SEQ ID NO:7 (miHDS1v6A), y la región no guía es por lo menos 80% complementaria con la región guía.

[69] La presente invención se relaciona además con un ARN codificado por el ácido nucleico de la invención.

[70] La presente invención se relaciona además con un cassette de expresión que codifica para el ácido nucleico aislado de la invención.

[71] La presente invención se relaciona además con un vector que comprende el cassette de expresión de la invención.

[72] La presente invención se relaciona además con un ácido nucleico aislado entre 80 y 4000 nucleótidos de longitud, que comprende un ácido nucleico que codifica para un transcripto primario de miARN artificial (pri-miARN) que consiste en, en orden de posición, una región flanqueante 5', una región no guía, una región bucle, una región guía, y una región flanqueante 3', en donde la región guía consiste en SEQ ID NO: 37 (miHDss3), SEQ ID NO:6 (miHDS1v5U) o SEQ ID NO:7 (miHDS1v6A), and la región no guía es por lo menos 80% complementaria con la región guía.

[73] La presente invención se relaciona además con un ácido nucleico aislado que consiste en Pri-miHDS1v5U (SEQ ID NO:8), Pri-miHDS1v6A (SEQ ID NO:9), Pre-miHDS1v5U (SEQ ID NO:10), o Pre-miHDS1v6A (SEQ ID NO:11).

[74] La presente invención se relaciona además con el ácido nucleico de la invención, el cassette de expresión de la invención, o el vector de la invención para usar para tratar la enfermedad de Huntington.

[75] La interferencia de ARN (iARN) es un proceso de regulación de genes mediado por ARNhd pequeños. La iARN se usa como una herramienta biológica habitual para estudiar la función de genes, y se encuentra en investigación como terapia para tratar diferentes enfermedades. La administración o expresión de ARNi puede ser realizada mediante la administración de ARNip exógenos (silenciamiento génico transitorio) o mediante la administración de vectores que expresan ARN tallo-bucle (silenciamiento génico persistente). La especificidad absoluta de la iARN es cuestionable. Los problemas que deben abordarse incluyen respuestas celulares al ARNhd (IFN-b, PKR, OAS1) y efectos no diana debido a la saturación de la maquinaria de iARN o mediados por complementariedad parcial con ARNm no deseados. Existe una continua necesidad por optimizar los vectores de ARNi y desarrollar potencialmente estrategias de expresión específica de tejido y regulada.

[76] El uso de la iARN como terapia depende de la dilucidación de varios factores que incluyen i) la administración y persistencia de la construcción de ARNi para el silenciamiento eficaz de la secuencia del gen diana; ii) el diseño del ARNip con el objetivo de obtener el noqueo eficaz o la supresión del gen de la secuencia diana, y iii) el sistema de expresión de ARNip óptimo (ARNhp o miARN) para la administración del ARNip terapéutico. Si bien muchos estudios han evaluado el uso de ARNi administrados como estructuras de oligonucleótidos sintetizados químicamente, para muchas condiciones clínicas y estados de enfermedad tales como la enfermedad de Huntington, se cree que para obtener un beneficio terapéutico existe una necesidad de expresión de nivel alto a largo plazo y/o persistente del ARNip terapéutico tal como lo obtenido mediante la producción endógena del ARNip expresado. Hasta la actualidad, se han comparado estrategias basadas en ARNhp y miARN artificiales con resultados conflictivos. La utilidad terapéutica del ARNi expresado está sin resolver debido a las preocupaciones de seguridad como resultado de la toxicidad fuera de diana que surge de las respuestas celulares al ARNhd (IFN-b, PKR, OAS1), saturación de la maquinaria de iARN o silenciamiento de genes no diana por medio de la complementariedad parcial con ARNm no deseados. Por lo tanto, existe una continua necesidad de optimizar los vectores de ARNi expresados que sean seguros y eficaces.

[77] Los ARNhp están comprendidos por estructuras tallo-bucle que se diseñan para que contengan una región flanqueante 5', segmentos de región de ARNip, una región bucle, una región de ARNip 3' y una región flanqueante 3'. La mayor parte de las estrategias de expresión de ARNi han utilizado ARN en horquilla pequeños (ARNhp) dirigidos por promotores basados en polIII fuertes. Muchos ARNhp han demostrado noqueo eficaz de las secuencias diana *in vitro* así como *in vivo*, sin embargo, algunos ARNhp que demostraron noqueo eficaz del gen diana también se ha encontrado que tienen toxicidad *in vivo*. Una estrategia alternativa descubierta recientemente es el uso de miARN artificiales (esqueletos de pri-miARN que transportan secuencias de ARNip) como vectores de ARNi. Los miARN artificiales se parecen de manera más natural a los sustratos de ARNi endógenos y son más susceptibles de la transcripción por Pol-II (por ejemplo, permitiendo una expresión específica de tejido del ARNi) y estrategias policistrónicas (por ejemplo, permitiendo la administración de múltiples secuencias de ARNip). Hasta la actualidad la eficacia de los sistemas de vectores basados en miARN en comparación a los ARNhp ha sido frustrada por

resultados conflictivos. Importantemente, el tema de la toxicidad fuera de diana producida por los dos sistemas no ha sido evaluado.

5 [78] Una consideración importante para el desarrollo de los ARNip expresados es el concepto de "dosificación" de la célula huésped con la construcción del ARNip expresado. "Dosificación" para un ARNip expresado en el contexto de la presente invención se refiere y puede ser dependiente del vehículo de administración (por ejemplo, viral o no viral), las cantidades o concentraciones relativas del vehículo de administración, y la potencia y especificidad del promotor utilizado para dirigir la expresión de la secuencia del ARNip.

10 [79] Los inventores han desarrollado vectores lanzaderas de miARN artificiales que incorporan las secuencias tallo-bucle contenidas en los ARNhp sin modificaciones de una secuencia microARN 30 de humano de origen natural o secuencia mi30 que sirve como transportar estas secuencias de ARN de interferencia pequeño (ARNip). Véase, por ejemplo, la Publicación PCT WO 2008/150897.

15 [80] Los inventores han desarrollado miARN artificiales, pri-miARN, pre-miARN, vectores de expresión, dúplex, y métodos para tratar la enfermedad de Huntington. Véase, por ejemplo, la Publicación PCT WO 2012/109667.

Lanzaderas de microARN para iARN

20 [81] Los miARN son ARN celulares pequeños (aproximadamente 22 nt) que son procesados a partir de transcritos tallo bucle precursores. Los miARN tallo-bucle conocidos pueden ser modificados para que contengan secuencias de ARNi específicas de genes de interés. Las moléculas de miARN pueden preferirse respecto a las moléculas de ARNhp ya que los miARN se expresan de manera endógena. Por lo tanto, es improbable que las moléculas de miARN induzcan vías de señalización de interferón en respuesta a ARNhd, son procesados de manera más eficaz que los ARNhp, y se ha mostrado que silencian 80% más eficazmente.

25 [82] Además, los roles del promotor son diferentes para las moléculas de miARN en comparación con las moléculas de ARNhp. La expresión específica de tejido, inducible, de los ARNhp involucra el truncado de los promotores de polII en el sitio de inicio de la transcripción. Por el contrario, los miARN pueden expresarse a partir de cualquier promotor de polIII ya que los sitios de inicio y fin de la transcripción pueden ser relativamente arbitrarios.

Tratamiento de la enfermedad de Huntington

30 [83] Las enfermedades de expansión de poliglutamina dominante, que incluyen ataxia espinocerebelar tipo 1 (SCA1) y la enfermedad de Huntington (HD), son trastornos neurodegenerativos progresivos, no tratables. En los modelos de HD inducible en ratón, la represión de la expresión del alelo mutante mejora los fenotipos de la enfermedad. Por lo tanto, las terapias diseñadas para inhibir la expresión del gen de la enfermedad podrían ser beneficiosas. Se describen métodos para usar iARN *in vivo* para tratar la enfermedad de Huntington. "Tratar" como se usa en la presente se refiere a aliviar por lo menos un síntoma de, curar y/o prevenir el desarrollo una enfermedad o una condición.

35 [84] En determinada forma de realización, las moléculas de ARNi se emplean para inhibir la expresión de un gen diana. Por "inhibir la expresión" se refiere a reducir, disminuir o suprimir la expresión de un gen diana. La expresión de un gen diana puede ser inhibida mediante "silenciamiento génico". El silenciamiento génico se refiere a la supresión de la expresión del gen, por ejemplo, expresión del transgén, gen heterólogo y/o gen endógeno, el cual puede estar mediado por procesos que afectan la transcripción y/o mediante procesos que afectan los mecanismos posttranscripcionales. En algunas formas de realización, el silenciamiento génico se produce cuando una molécula de ARNi inicia la inhibición o degradación del ARNm transcripto a partir de un gen de interés de un modo específico de secuencia mediante interferencia de ARN, previniendo de esta manera la traducción del producto del gen.

40 [85] La referencia a los ARNip en la presente se refiere a que incluyan ARNhp y otros ARNp que pueden o que tiene la capacidad de modular la expresión de un gen diana, por ejemplo, el gen de HD, por ejemplo mediante interferencia de ARN. Dichos ARN pequeños incluyen sin limitación, ARNhp y microARN (miARN).

45 [86] En la presente se describe una estrategia que produce un silenciamiento sustancial de genes diana mediante iARN. El uso de esta estrategia produce una disminución marcada de la expresión *in vitro* e *in vivo* de genes diana. Esta estrategia es útil para reducir la expresión de genes diana con el objetivo de modelar procesos biológicos o para proveer terapia para enfermedades humanas. Por ejemplo, esta estrategia puede ser aplicada a la enfermedad de Huntington. Como se usa en la presente el término "silenciamiento sustancial" se refiere a que el ARNm del gen diana es inhibido y/o degradado por la presencia del ARNip introducido, de manera tal que la expresión del gen diana se reduce en aproximadamente entre 10% y 100% en comparación al nivel de expresión vista cuando el ARNip no está presente. En general, cuando un gen es sustancialmente silenciado, tendrá por lo menos 40%, 50%, 60% a 70%, por ejemplo, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78% a 79%, en general por lo menos 80%, por ejemplo, entre 81% y 84%, por lo menos 85%, por ejemplo, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 50 93%, 94%, 95%, 96%, 97%, 98%, 99% o incluso 100% de reducción de la expresión en comparación a cuando no está presente el ARNip. Como se usa en la presente el término "actividad sustancialmente normal" se refiere al nivel

de expresión de un gen cuando no se ha introducido un ARNip en una célula.

[87] La enfermedad de Huntington (HD) es un gran candidato para la terapia basada en ARNip. La HD es provocada por expansiones de repeticiones CAG que codifican para poliQ en la proteína de la enfermedad. La expansión de poliQ confiere una propiedad tóxica dominante en la proteína mutante que se asocia con la acumulación aberrante de la proteína de la enfermedad en las neuronas. La HD es progresiva, con trastornos finalmente fatales que habitualmente comienzan en la edad adulta. La expansión del dominio de repetición CAG/poliQ confiere a la proteína codificada una propiedad tóxica dominante. Por lo tanto, como estrategia terapéutica, los esfuerzos por disminuir la expresión del producto génico mutante antes de la muerte de la célula podría ser muy beneficioso para los pacientes.

Moléculas de interferencia de ARN (iARN)

[88] Una "interferencia de ARN", "iARN", "ARN de interferencia pequeño" o "ARN de interferencia corto" o "ARNip" o "ARN de horquilla pequeño" o molécula de "ARNhp", o "miARN" es un ARN dúplex de nucleótidos que está dirigido a una secuencia de ácido nucleico de interés, por ejemplo, huntingtina (htt). Como se usa en la presente, el término "ARNip" es un término genérico que abarca el subconjunto de los ARNhp y miARN. Un "dúplex de ARN" se refiere a la estructura formada por el apareamiento complementario entre dos regiones de una molécula de ARN. El ARNip está "dirigido" contra un gen en donde la secuencia nucleotídica de la parte de dúplex del ARNip es complementaria a una secuencia nucleotídica del gen diana. En determinadas formas de realización, los ARNip están dirigidos a la secuencia que codifica para ataxina-1 o huntingtina. En algunas formas de realización, la longitud del dúplex de los ARNip es menor que 30 pares de bases. En algunas formas de realización, el dúplex puede tener 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 12, 11 o 10 pares de bases de longitud. En algunas formas de realización, la longitud del dúplex tiene entre 19 y 25 pares de bases de longitud. En determinada forma de realización, la longitud del dúplex tiene 19 o 21 pares de bases de longitud. La parte del ARN dúplex del ARNip puede ser parte de una estructura en horquilla. Además de la parte de dúplex, la estructura en horquilla puede contener una parte en bucle posicionada entre las dos secuencias que forman el dúplex. El bucle puede variar de longitud. En algunas formas de realización el bucle tiene 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. En determinadas formas de realización, el bucle tiene 18 nucleótidos de longitud. La estructura en horquilla también puede contener partes de sobreextensión 3' y/o 5'. En algunas formas de realización, la sobreextensión es una sobreextensión 3' y/o una 5' de 0, 1, 2, 3, 4 o 5 nucleótidos de longitud.

[89] La unidad transcripcional de un "ARNhp" está comprendida por las secuencias sentido y antisentido conectadas por un bucle de nucleótidos desapareados. Los ARNhp son exportados del núcleo mediante la Exportina-5, y una vez en el citoplasma, son procesados por Dicer para generar ARNip funcionales. Los "miARN" tallo-bucle están comprendidos por las secuencias sentido y antisentido conectadas mediante un bucle de nucleótidos desapareados habitualmente expresados como parte de transcritos primarios más grandes (pri-miARN), que son escindidos por el complejo Drosha-DGCR8 que genera intermediarios conocidos como pre-miARN, que son posteriormente exportados del núcleo por la Exportina-5, y una vez en el citoplasma, son procesados por Dicer para generar ARNip funcionales. El "miARN artificial" o un "vector lanzadera de miARN artificial", como se usan en la presente como sinónimos, se refiere a un transcrito de miARN primario que ha tenido una región del dúplex tallo bucle (por lo menos de aproximadamente entre 9 y 20 nucleótidos) que es escindida mediante procesamiento con Drosha y Dicer y reemplazada por las secuencias de ARNip para el gen diana a la vez que retiene los elementos estructurales en el tallo bucle necesarios para el procesamiento eficaz mediante Drosha. El término "artificial" surge del hecho de que las secuencias flanqueantes (-35 nucleótidos corriente arriba y -40 nucleótidos corriente abajo) surgen de los sitios de la enzima de restricción dentro del sitio de clonado múltiple del ARNip. Como se usa en la presente el término "miARN" abarca las secuencias de miARN de origen natural así como los vectores lanzadera de miARN generados de manera artificial.

[90] El ARNip puede estar codificado por una secuencia de ácido nucleico, y la secuencia de ácido nucleico también puede incluir un promotor. La secuencia de ácido nucleico también puede incluir una señal de poliadenilación. En algunas formas de realización, la señal de poliadenilación es una señal de poliadenilación mínima sintética o una secuencia de seis T.

[91] "Toxicidad fuera de diana" se refiere a cambios fenotípicos deletéreos, no deseables, o no pretendidos de una célula huésped que expresa o contiene un ARNip. La toxicidad fuera de diana puede producir la pérdida de la función deseada, ganancia de una función no deseada, o incluso la muerte a nivel celular o del organismo. La toxicidad fuera de diana puede producirse inmediatamente al expresarse el ARNip o puede producirse gradualmente en el tiempo. La toxicidad fuera de diana puede producirse como resultado directo de la expresión de ARNip o puede producirse como resultado de la inducción de la respuesta inmunológica del huésped contra la célula que expresa el ARNip. Sin desear de estar ligado a la teoría, se postula que la toxicidad fuera de diana surge de niveles altos o sobreabundancia de sustratos de iARN dentro de la célula. Estos sustratos de iARN sobreabundantes o sobreexpresados, incluyendo sin limitación sustratos pre o pri-ARNi así como ARN antisentido maduros sobreabundantes, pueden competir con la maquinaria de iARN endógena, alterando de esta manera la biogénesis y la función de los miARN naturales. La toxicidad fuera de diana también puede surgir de una probabilidad aumentada de silenciamiento de ARNm no deseados (es decir, no diana) debido a la complementariedad parcial de la

secuencia. La toxicidad fuera de diana también puede producirse por desviación de hebra no apropiada de una región no guía tal que haya una carga preferencial de la región no guía sobre la región diana o guía del ARNi. La toxicidad fuera de diana también puede surgir de la estimulación de respuestas celulares a los ARNhd que incluyen ARNhd (IFN-b, PKR, OAS1). "Toxicidad fuera de diana disminuida" se refiere a una disminución, reducción, supresión o atenuación de la toxicidad fuera de diana de manera tal que el efecto terapéutico sea más beneficioso para el huésped a que la toxicidad sea limitante o perjudicial medido según una duración o calidad de vida mejorada o una mejora en un signo o síntoma de una enfermedad o condición que está siendo tratada con el ARNip. "Toxicidad fuera de diana limitada" o "toxicidad fuera de diana baja" se usa para referirse a un cambio fenotípico no deseado no pretendido en una célula u organismo, sea detectable o no, que no impide o compensa o limita el beneficio terapéutico para el huésped tratado con el ARNip y puede considerarse un "efecto secundario" de la terapia. La toxicidad fuera de diana disminuida o limitada puede determinarse o inferirse mediante la comparación del análisis *in vitro* tal como por transferencia Northern o QPCR para los niveles de sustratos de ARNip o los efectos *in vivo* por comparación de un vector de ARNhp equivalente con el vector lanzadera de miARN de la presente divulgación.

[92] "Noqueo", "tecnología de noqueo" se refiere a una técnica de silenciamiento génico en la cual la expresión de un gen diana es reducida en comparación con la expresión del gen antes de la introducción del ARNip, que puede llevar a la inhibición de la producción del producto génico diana. El término "reducida" se usa en la presente para indicar que la expresión del gen diana disminuyó entre 1 y 100%. En otras palabras, se minimiza la cantidad de ARN disponible para la traducción a un polipéptido o proteína. Por ejemplo, la cantidad de proteína puede reducirse en un 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, o 99%. En algunas formas de realización, la expresión se reduce en aproximadamente un 90% (es decir, solo se observa aproximadamente 10% de la cantidad de proteína en una célula en comparación con una célula en la cual no se han administrado las moléculas de ARNip). El noqueo de la expresión del gen puede estar dirigido por el uso de los ARNhd o ARNip.

[93] La "interferencia de ARN (iARN)" es el proceso de silenciamiento génico postranscripcional específico de secuencia iniciado por ARNip. Durante la iARN, el ARNip induce la degradación del ARNm diana con la consecuente inhibición específica de secuencia de la expresión del gen.

[94] De acuerdo con un método de la presente divulgación, la expresión de huntingtina puede ser modificada por medio de la iARN. Por ejemplo, puede suprimirse la acumulación de huntingtina en una célula. El término "supresión" se refiere a la disminución, reducción o eliminación del número o cantidad de transcritos presentes en una célula particular. Por ejemplo, la acumulación de ARNm que codifica para huntingtina puede ser suprimida en una célula mediante interferencia de ARN (iARN), por ejemplo, el gen es silenciado por un ARN de hebra doble específico de secuencia (ARNhd), que también es llamado ARN de interferencia pequeño (ARNip). Estos ARNip pueden ser dos moléculas de ARN separadas que han hibridizado entre sí, o pueden ser una horquilla individual en donde dos partes de una molécula de ARN han hibridizado entre sí para formar un dúplex.

[95] Una proteína mutante se refiere a la proteína codificada por un gen que tiene una mutación, por ejemplo, una mutación con sentido cambiado o sin sentido en huntingtina. Una huntingtina mutante puede producir enfermedad, es decir, puede llevar a una enfermedad asociada con la presencia de huntingtina en un animal que tiene uno o dos alelos mutantes.

[96] El término "gen" se usa de manera amplia para referirse a cualquier segmento de ácido nucleico asociado con una función biológica. Por lo tanto, los genes incluyen secuencias codificantes y/o las secuencias reguladoras requeridas para su expresión. Por ejemplo, "gen" se refiere a un fragmento de ácido nucleico que expresa ARNm, ARN funcional, o proteína específica, que incluye secuencias reguladoras. "Genes" también incluyen segmentos de ADN no expresados que, por ejemplo, forman secuencias de reconocimiento de otras proteínas. Los "genes" pueden obtenerse de una variedad de fuentes, que incluyen clonado de una fuente de interés o síntesis a partir de información de secuencia conocida o predicha, y puede incluir secuencias diseñadas para que tengan parámetros deseados. Un "alelo" es uno de varias formas alternativas de un gen que ocupa un *locus* dado en un cromosoma.

[97] El término "ácido nucleico" se refiere a ácido desoxirribonucleico (ADN) o ácido ribonucleico (ARN) y polímeros de los mismos en una forma de hebra simple o doble, compuesto por monómeros (nucleótidos) que contienen un azúcar, fosfato y una base que es una purina o pirimidina. A menos que se limite de manera específica, el término abarca ácidos nucleicos que contienen análogos de nucleótidos naturales que tienen propiedades de unión similares al ácido nucleico de referencia y que son metabolizados de una manera similar a los nucleótidos de origen natural. A menos que se indique de otra manera, una secuencia de ácido nucleico particular también abarca variantes modificadas de manera conservativa de la misma (por ejemplo, sustituciones de codón degenerado) y secuencias complementarias, así como la secuencia indicada explícitamente. Específicamente, las sustituciones de codones degenerados pueden lograrse mediante la generación de secuencias en las cuales la tercera posición de uno o más codones seleccionados (o todos) está sustituida por residuos de base mezclada y/o desoxiinosina. Un "fragmento de ácido nucleico" es una parte de una molécula de ácido nucleico dado.

[98] Una "secuencia nucleotídica" es un polímero de ADN o ARN que puede ser de hebra simple o de hebra doble, que opcionalmente contiene bases nucleotídicas sintéticas, no naturales o alteradas con capacidad de

incorporarse en los polímeros de ADN o ARN.

[99] Los términos "ácido nucleico", "molécula de ácido nucleico", "fragmento de ácido nucleico", "secuencia o segmento de ácido nucleico", o "polinucleótido" se usan como sinónimos y también pueden usarse como sinónimo con gen, ADNc, ADN y ARN codificados por un gen.

[100] Se describen moléculas de ácido nucleico sustancialmente aisladas o purificadas y composiciones que contienen esas moléculas. En el contexto de la presente divulgación, una molécula de ADN "aislada" o "purificada" o molécula de ARN es una molécula de ADN o molécula de ARN que existe aparte de en su ambiente nativo y por lo tanto no es un producto de la naturaleza. Una molécula de ADN o molécula de ARN aislada puede existir en una forma purificada o puede existir en un ambiente no nativo tal como, por ejemplo, una célula huésped transgénica. Por ejemplo, una molécula de ácido nucleico "aislada" o "purificada" o parte biológicamente activa de la misma, está sustancialmente libre de otro material celular, o medio de cultivo cuando se produce mediante técnicas recombinantes, o sustancialmente libre de precursores químicos u otros químicos cuando se sintetizan de manera química. En una forma de realización, un ácido nucleico "aislado" está libre de secuencias que flanquean de manera natural el ácido nucleico (es decir, secuencias localizadas en los extremos 5' y 3' del ácido nucleico) en el ADN genómico del organismo del cual deriva el ácido nucleico. Por ejemplo, en diferentes formas de realización, la molécula de ácido nucleico aislado puede contener menos de aproximadamente 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb, o 0,1 kb de secuencias nucleotídicas que flanquean de manera natural la molécula de ácido nucleico en el ADN genómico de la célula de la cual deriva el ácido nucleico. Los fragmentos y variantes de las secuencias nucleotídicas descritas también pueden estar abarcadas por la presente divulgación. Por "fragmento" o "parte" se comprende una longitud completa o menos que la longitud completa de la secuencia nucleotídica.

[101] "De origen natural", "nativo", o "de tipo salvaje" se usa para describir un objeto que puede ser encontrado en la naturaleza de una manera diferente a la producida de manera artificial. Por ejemplo, una proteína o secuencia nucleotídica presente en un organismo (incluyendo un virus), que puede ser aislado de una fuente en la naturaleza y que no ha sido modificado de manera intencional por una persona en el laboratorio, es de origen natural.

[102] Una "variante" de una molécula es una secuencia que es sustancialmente similar a la secuencia de la molécula nativa. Para las secuencias nucleotídicas, las variantes incluyen aquellas secuencias que, debido a la degeneración del código genético, codifican para la secuencia de aminoácidos idéntica a la proteína nativa. Las variantes alélicas de origen natural tales como estas pueden identificarse con el uso de técnicas de biología molecular, como, por ejemplo, mediante la reacción en cadena de la polimerasa (PCR) y técnicas de hibridación. Las secuencias nucleotídicas variantes también incluyen secuencias nucleotídicas derivadas de manera sintética, tales como aquellas generadas, por ejemplo, mediante el uso de mutagénesis dirigida a sitio, que codifican para la proteína nativa, así como aquellas que codifican para un polipéptido que tiene sustituciones de aminoácido. En general, las variantes de secuencia nucleotídica de la invención tendrán por lo menos 40%, 50%, 60%, a 70%, por ejemplo, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, a 79%, en general por lo menos 80%, por ejemplo, entre 81% y 84%, por lo menos 85%, por ejemplo, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, a 98%, de identidad de secuencia con la secuencia nucleotídica nativa (endógena).

[103] Un "transgén" se refiere a un gen que ha sido introducido en el genoma mediante transformación. Los transgenes incluyen, por ejemplo, ADN que es heterólogo u homólogo al ADN de una célula particular a transformar. Además, los transgenes pueden incluir genes nativos insertados en un organismo no nativo, o genes quiméricos.

[104] El término "gen endógeno" se refiere a un gen nativo en su ubicación natural en el genoma de un organismo.

[105] Los términos "proteína", "péptido" y "polipéptido" se usan como sinónimos en la presente.

[106] "De tipo salvaje" se refiere al gen u organismo normal encontrado en la naturaleza.

[107] "Genoma" se refiere al material genético completo de un organismo.

[108] Un "vector" se define como que incluye, entre otros, cualquier vector viral, así como cualquier plásmido, cósmido, fago o vector binario en forma de hebra doble o simple lineal o circular que puede o no ser autotransmisible o móvil, y que puede transformar un huésped procarionota o eucariota mediante integración en el genoma celular o existe extracromosomalmente (por ejemplo, plásmido replicante autónomo con un origen de replicación).

[109] "Cassette de expresión" como se usa en la presente se refiere a una secuencia de ácido nucleico con capacidad de dirigir la expresión de una secuencia nucleotídica particular en una célula huésped apropiada, que puede incluir un promotor ligado operativamente a la secuencia nucleotídica de interés que puede estar ligada operativamente a señales de terminación. La región codificante habitualmente codifica para un ARN funcional de interés, por ejemplo un ARNip. El cassette de expresión que incluye la secuencia nucleotídica de interés puede ser quimérico. El cassette de expresión también puede ser uno que es de origen natural pero que ha sido obtenido de una forma recombinante útil para la expresión heteróloga. La expresión de la secuencia nucleotídica en el cassette

de expresión puede estar bajo el control de un promotor constitutivo o de un promotor regulable que inicia la transcripción solo cuando la célula huésped es expuesta a algún estímulo particular. En el caso de un organismo multicelular, el promotor también puede ser específico por un tejido u órgano o estadio de desarrollo particular.

5 [110] Dichos cassettes de expresión pueden incluir una región de inicio de la transcripción ligada a una secuencia nucleotídica de interés. Dicho cassette de expresión se provee con una pluralidad de sitios de restricción para que la inserción del gen de interés sea bajo la regulación de la transcripción de las regiones reguladoras. El cassette de expresión además puede contener genes marcadores de selección.

10 [111] "Secuencia codificante" se refiere a una secuencia de ADN o ARN que codifica para una secuencia de aminoácidos específica. Puede constituir una "secuencia codificante ininterrumpida", es decir, que carece de un intrón, tal como en un ADNc, o puede incluir uno o más intrones unidos mediante uniones de corte y empalme apropiadas. Un "intrón" es una secuencia de ARN que está contenida en el transcripto primario pero que es eliminada mediante escisión y vuelta a ligar del ARN en la célula para crear el ARNm maduro que puede ser
15 traducido a una proteína.

[112] El término "marco de lectura abierto" (ORF) se refiere a la secuencia entre los codones de inicio y terminación de la traducción de una secuencia codificante. Los términos "codón de inicio" y "codón de terminación" se refieren a una unidad de tres nucleótidos adyacentes (un "codón") en una secuencia codificante que especifica el inicio y terminación de cadena, respectivamente, de la síntesis de la proteína (traducción del ARNm).
20

[113] "ARN funcional" se refiere a ARN sentido, ARN antisentido, ARN ribozima, ARNip, u otro ARN que no puede ser traducido pero que aún tiene un efecto sobre al menos un proceso celular.

25 [114] El término "transcripto de ARN" o "transcripto" se refiere al producto que resulta de la transcripción catalizada por la ARN polimerasa de una secuencia de ADN. Cuando el transcripto de ARN es una copia complementaria perfecta de la secuencia de ADN, se refiere como el transcripto primario o puede ser una secuencia de ARN derivada del procesamiento postranscripcional del transcripto primario y es referido como el ARN maduro. El "ARN mensajero" (ARNm) se refiere al ARN que no tiene los intrones y que puede ser traducido a una proteína por
30 la célula.

[115] "ADNc" se refiere a un ADN de hebra simple o doble que es complementario y que deriva de ARNm.

35 [116] "Secuencias reguladoras" son secuencias nucleotídicas localizadas corriente arriba (secuencias no codificantes 5'), en, o corriente abajo (secuencias no codificantes 3') de una secuencia codificante, y que influyen en la transcripción, procesamiento o estabilidad de ARN, o traducción de la secuencia codificante asociada. Las secuencias reguladoras incluyen potenciadores, promotores, secuencias líderes de traducción, intrones, y secuencias señal de poliadenilación. Ellas incluyen secuencias naturales y sintéticas así como secuencias que pueden ser una combinación de secuencias sintéticas y naturales. Tal como se indica en la presente, el término "secuencias reguladoras adecuadas" no se limita a promotores. Sin embargo, algunas secuencias reguladoras útiles
40 adecuadas en la presente divulgación incluirán, a título enunciativo no taxativo promotores constitutivos, promotores específicos de tejido, promotores específicos de desarrollo, promotores regulables y promotores virales.

45 [117] "Secuencia no codificante 5'" se refiere a una secuencia nucleotídica ubicada 5' (corriente arriba) a la secuencia codificante. Se encuentra presente en el ARNm completamente procesado corriente arriba del codón de inicio y puede afectar al procesamiento del transcripto primario a ARNm, estabilidad del ARNm o eficacia de la traducción.

50 [118] "Secuencia no codificante 3'" se refiere a secuencias nucleotídicas ubicadas 3' (corriente abajo) a una secuencia codificante y puede incluir secuencias señal de poliadenilación y otras secuencias que codifican para señales de regulación con capacidad de afectar el procesamiento del ARNm o la expresión del gen. La señal de poliadenilación habitualmente se caracteriza por afectar el agregado de extensiones de ácido poliadenílico al extremo 3' del precursor de ARNm.

55 [119] El término "secuencia líder de traducción" se refiere a la parte de la secuencia de ADN de un gen entre el promotor y la secuencia codificante que se transcribe a ARN y que está presente en el ARNm completamente procesado corriente arriba (5') del codón de inicio de la traducción. La secuencia líder de traducción puede afectar el procesamiento del transcripto primario al ARNm, estabilidad del ARNm o eficacia de traducción.

60 [120] El término proteína "madura" se refiere a un polipéptido procesado postraduccionalmente sin su péptido señal. Proteína "precursora" se refiere al producto primario de la traducción de un ARNm. "Péptido señal" se refiere a la extensión terminal amino de un polipéptido, que es traducida junto con el polipéptido formando un péptido precursor y que se requiere para su entrada en la vía de secreción. El término "secuencia señal" se refiere a una secuencia nucleotídica que codifica para el péptido señal.
65

[121] "Promotor" se refiere a una secuencia nucleotídica, habitualmente corriente arriba (5') a su secuencia

codificante, que dirige y/o controla la expresión de la secuencia codificante al proveer el reconocimiento para la ARN polimerasa y otros factores requeridos para la transcripción apropiada. "Promotor" incluye un promotor mínimo que es una secuencia de ADN corta comprendida por una caja TATA y otras secuencias que sirven para especificar el sitio de inicio de la transcripción, al cual se agregan elementos reguladores para el control de la expresión.

5 "Promotor" también se refiere a una secuencia nucleotídica que incluye un promotor mínimo más elementos reguladores que tiene la capacidad de controlar la expresión de una secuencia codificante o ARN funcional. Este tipo de secuencia promotora consiste en elementos corriente arriba próximos y más distales, en donde los últimos elementos son referidos con frecuencia como potenciadores. En consecuencia, un "potenciador" es una secuencia de ADN que puede estimular la actividad del promotor y puede ser un elemento innato del promotor o un elemento

10 heterólogo insertado para aumentar el nivel o especificidad de tejido de un promotor. Tiene la capacidad de operar en ambas orientaciones (normal o invertido), y tiene la capacidad de funcionar cuando se mueve corriente arriba o corriente abajo del promotor. Los potenciadores y otros elementos promotores corriente arriba se unen a proteínas de unión a ADN específicas de secuencia que median sus efectos. Los promotores pueden derivar en su totalidad de un gen nativo, o pueden estar compuestos por elementos diferentes derivados de diferentes promotores encontrados en la naturaleza, o incluso pueden estar compuestos de segmentos de ADN sintéticos. Un promotor también puede contener secuencias de ADN que están involucradas en la unión de factores proteicos que controla la eficacia del inicio de la transcripción en respuesta a condiciones fisiológicas o de desarrollo. Los ejemplos de promotores que pueden usarse en la presente invención incluyen los promotores de ARN U6 de ratón, promotores H1RNA de humano sintéticos, promotores SV40, CMV, RSV, ARN polimerasa II y ARN polimerasa III.

[122] El "sitio de inicio" es la posición que rodea al primer nucleótido que es parte de la secuencia transcrita, que también se define como posición +1. Respecto a este sitio todas las otras secuencias del gen y sus regiones de control están numeradas. Las secuencias corriente abajo (es decir, otras secuencias codificantes de proteína en la dirección 3') son denominadas positivas, mientras que las secuencias corriente arriba (la mayor parte de las regiones controladoras en la dirección 5') son denominadas negativas.

[123] Los elementos promotores, particularmente un elemento TATA, que son inactivos o que tienen una actividad promotora muy reducida en ausencia de activación corriente arriba son referidos como "promotores mínimos o centrales". En presencia de un factor de transcripción adecuado, el promotor mínimo funciona para permitir la transcripción. Un "promotor mínimo o central" consiste por lo tanto en solo elementos basales requeridos para el inicio de la transcripción, por ejemplo, una caja TATA y/o un iniciador.

[124] "Expresión constitutiva" se refiere a la expresión usando un promotor constitutivo o regulado. "Expresión condicional" y "regulada" se refiere a la expresión controlada por un promotor regulado.

[125] "Ligado operativamente" se refiere a la asociación de secuencias de ácido nucleico en fragmento de ácido nucleico individual de manera que la función de una de las secuencias se afectada por otra. Por ejemplo, una secuencia de ADN reguladora se dice que está "ligada operativamente a" o "asociada con" una secuencia de ADN que codifica para un ARN o un polipéptido si las dos secuencias están ubicadas de manera que la secuencia de ADN reguladora afecta la expresión de la secuencia de ADN codificante (es decir, que la secuencia codificante o ARN funcional está bajo el control transcripcional del promotor). Las secuencias codificantes pueden estar operativamente ligadas a secuencias reguladoras en orientación sentido o antisentido.

[126] "Expresión" se refiere a la transcripción y/o traducción de un gen endógeno, heterólogo gen o segmento de ácido nucleico, o un transgén en las células. Por ejemplo, en el caso de construcciones de ARNip, la expresión puede referirse a la transcripción del ARNip solo. Además, expresión se refiere a la transcripción y acumulación estable de ARN sentido (ARNm) o funcional. La expresión también puede referirse a la producción de proteína.

[127] "Niveles alterados" se refiere al nivel de expresión en las células u organismos transgénicos que difieren de las células u organismos normales o sin transformar.

[128] "Sobreexpresión" se refiere al nivel de expresión en las células u organismos transgénicos que excede los niveles de expresión en las células u organismos normales o sin transformar.

[129] "Inhibición antisentido" se refiere a la producción de transcritos de ARN antisentido con capacidad de suprimir la expresión de proteína de un gen o transgén endógeno.

[130] "Fragmento de terminación de la transcripción" se refiere a secuencias nucleotídicas que contienen una o más señales reguladoras, tal como secuencias señales de poliadenilación, con capacidad de terminar la transcripción. Los ejemplos incluyen las regiones no reguladoras 3' de genes que codifican para nopalina sintasa y la subunidad pequeña de la ribulosa bifsosfato carboxilasa.

[131] "Fragmento de terminación de la traducción" se refiere a secuencias nucleotídicas que contienen una o más señales reguladoras, tales como uno o más codones de terminación en los tres marcos, con capacidad de terminar la traducción. La inserción de un fragmento de terminación de la traducción adyacente a o cerca del codón de inicio en el extremo 5' de la secuencia codificante resultará en ausencia de traducción o traducción incorrecta. La

escisión del fragmento de terminación de la traducción mediante recombinación específica de sitio dejará una secuencia específica de sitio en la secuencia codificante que no interfiere con la traducción apropiada usando el codón de terminación.

5 [132] Los términos "secuencia de acción cis" y "elemento de acción cis" se refiere a secuencias de ADN o ARN cuyas funciones requieren de ellas para estar en la misma molécula. Un ejemplo de una secuencia de acción cis en el replicón es el origen de replicación viral.

10 [133] Los términos "secuencia de acción trans" y "elemento de acción trans" se refieren a secuencias de ADN o ARN cuya función no requiere de ellas para estar en la misma molécula.

15 [134] "Integrado cromosómicamente" se refiere a la integración de un gen o construcción de ácido nucleico foráneo en el ADN del huésped por enlaces covalentes. Cuando los genes no están "integrados cromosómicamente" pueden ser "expresados de manera transitoria". La expresión transitoria de un gen se refiere a la expresión de un gen que no está integrado en el cromosoma del huésped pero funciona independientemente, como parte de un plásmido o cassette de expresión de replicación autónoma, por ejemplo, o como parte de otro sistema biológico tal como un virus.

20 [135] Los siguientes términos se usan para describir las relaciones de secuencias entre dos o más ácidos nucleicos o polinucleótidos: (a) "secuencia de referencia", (b) "ventana de comparación", (c) "identidad de secuencia", (d) "porcentaje de identidad de secuencia", y (e) "identidad sustancial".

25 (a) Como se usa en la presente, "secuencia de referencia" es una secuencia definida usada como una base para la comparación de secuencia. Una secuencia de referencia puede ser un subconjunto o la totalidad de una secuencia especificada; por ejemplo, como un segmento de un ADNc de longitud completa o secuencia génica, o el ADNc completo o secuencia génica.

30 (b) Como se usa en la presente, "ventana de comparación" hace referencia a un segmento contiguo y específico de una secuencia polinucleotídica, en donde la secuencia polinucleotídica en la ventana de comparación puede comprender agregados o eliminaciones (es decir, no coincidencias) en comparación con la secuencia de referencia (que no comprende agregados o eliminaciones) para el alineamiento óptimo de las dos secuencias. En general, la ventana de comparación tiene por lo menos 20 nucleótidos de longitud contiguos, y opcionalmente puede tener 30, 40, 50, 100, o más. Aquellas personas con experiencia en el arte comprenden que para evitar una alta similitud con una secuencia de referencia debido a la inclusión de no coincidencias en la secuencia polinucleotídica habitualmente se introduce una penalidad por no coincidencia y se sustrae del número de coincidencias.

35 [136] Los métodos de alineamiento de secuencias para la comparación son bien conocidos en el arte. Por lo tanto, la determinación del porcentaje de identidad entre dos secuencias cualesquiera puede realizarse usando un algoritmo matemático.

40 [137] Las implementaciones de computadora de estos algoritmos matemáticos pueden ser utilizadas para la comparación de secuencias para determinar la identidad de secuencia. Dichas implementaciones incluyen, a título enunciativo no taxativo: CLUSTAL en el programa PC/Gene (disponible de Intelligenetics, Mountain View, California); el programa ALIGN (Versión 2.0) y GAP, BESTFIT, BLAST, FASTA, y TFASTA en el paquete de programa de computación Wisconsin Genetics, Versión 8 (disponible de Genetics Computer Grupo (GCG), 575 Science Drive, Madison, Wisconsin, EE.UU.). Los alineamientos mediante el uso de estos programas pueden realizarse usando los parámetros por defecto.

50 [138] El programa para realizar los análisis por BLAST se encuentra públicamente disponible del Centro Nacional de Información de Biotecnología. Este algoritmo involucra identificar primero los pares de secuencia con alta puntuación (HSP) mediante la identificación de palabras cortas de longitud W en la secuencia de estudio, que coinciden o satisfacen alguna puntuación de umbral valorado positivo T cuando se alinea con una palabra de la misma longitud en una secuencia de bases de datos. T es referido como el umbral de puntuación de palabra vecina. Estas coincidencias de palabra vecina inicial actúa como semilla para iniciar búsquedas para encontrar HSP más largos que los contienen. Las coincidencias de palabra luego son extendidas en ambas direcciones a lo largo de cada secuencia por tanto como la puntuación de alineamiento acumulativo puede aumentarse. Las puntuaciones acumulativas se calculan usando, para secuencias nucleotídicas, los parámetros M (puntuación de recompensa para un par de residuos coincidentes; siempre > 0) y N (puntuación de penalidad para residuos no coincidentes; siempre < 0). Para las secuencias de aminoácidos, se usa una matriz de puntuación para calcular la puntuación acumulativa. La extensión de las coincidencias de palabra en cada dirección se detiene cuando la puntuación de alineamiento acumulativo cae fuera de la cantidad X de su valor alcanzado máximo, la puntuación acumulativa se va acero o por debajo de la acumulación de uno o más alineamiento de residuos con puntuación negativa, o se alcanza el extremo de una secuencia.

65 [139] Además de calcular la identidad de secuencia porcentual, el algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias. Una medida de similitud provista por el algoritmo BLAST es la probabilidad de suma más pequeña (P(N)), que provee una indicación de la probabilidad por la cual una coincidencia

entre dos secuencias nucleotídicas podría producirse por azar. Por ejemplo, una secuencia de ácido nucleico de ensayo es considerada similar a una secuencia de referencia si la probabilidad de suma más pequeña en una comparación de la secuencia de ácido nucleico de ensayo con la secuencia de ácido nucleico de referencia es menor que aproximadamente 0,1, más preferiblemente menor que aproximadamente 0,01, y aún más preferiblemente menor que aproximadamente 0,001.

[140] Para obtener alineamientos con no coincidencias con propósito de comparación, puede utilizarse Gapped BLAST (en BLAST 2.0). Como alternativa, puede utilizarse PSI-BLAST (en BLAST 2.0) para realizar una búsqueda iterada que detecte relaciones distantes entre moléculas. Cuando se utiliza BLAST, Gapped BLAST, PSI-BLAST, pueden usarse los parámetros por defecto de los respectivos programas (por ejemplo BLASTN para secuencias nucleotídicas). El programa BLASTN (para secuencias nucleotídicas) usa por defecto una longitud de palabra (W) de 11, una expectativa (E) de 10, un valor de corte de 100, M=5, N=-4, y una comparación de ambas hebras. El alineamiento también puede realizarse manualmente por inspección.

[141] Para los propósitos de la presente divulgación, la comparación de secuencias nucleotídicas para la determinación de la identidad de secuencia porcentual con las secuencias promotoras descritas en la presente preferiblemente se hace usando el programa BlastN (versión 4.7 o posterior) con sus parámetros por defecto o cualquier programa equivalente. Por "programa equivalente" se comprende cualquier programa de comparación de secuencia que, para dos secuencias cualesquiera de interés, genera un alineamiento que tiene coincidencias de nucleótido idénticas y una identidad de secuencia idéntica cuando se compara con el alineamiento correspondiente generado por el programa preferido.

(c) Como se usa en la presente, "identidad de secuencia" o "identidad" en el contexto de dos secuencias de ácido nucleico hace referencia a un porcentaje específico de nucleótidos en las dos secuencias que son los mismos cuando se alinean con la máxima correspondencia en una ventana de comparación específica, medido por algoritmos de comparación de secuencia o mediante inspección visual.

(d) Como se usa en la presente, "porcentaje de identidad de secuencia" se refiere al valor determinado por comparación de dos secuencias alineadas de manera óptima en una ventana de comparación, en donde la parte de la secuencia polinucleotídica en la ventana de comparación puede comprender agregados o eliminaciones (es decir, no coincidencias) en comparación con la secuencia de referencia (que no comprende agregados o eliminaciones) para el alineamiento óptimo de las dos secuencias. El porcentaje se calcula mediante la determinación del número de posiciones en las cuales hay bases de ácido nucleico o residuos de aminoácidos idénticos en ambas secuencias para obtener el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación, y multiplicando el resultado por 100 para obtener el porcentaje de identidad de secuencia.

(e) El término "identidad sustancial" de secuencias polinucleotídicas se refiere a que un polinucleótido comprende una secuencia que tiene por lo menos 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, o 79%, preferiblemente por lo menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, o 89%, más preferiblemente por lo menos 90%, 91%, 92%, 93%, o 94%, y aún más preferiblemente por lo menos 95%, 96%, 97%, 98%, o 99% de identidad de secuencia, en comparación con una secuencia de referencia usando uno de los programas de alineamiento descritos usando los parámetros estándares.

[142] Otra indicación de que las secuencias nucleotídicas son sustancialmente idénticas es si dos moléculas hibridizan entre sí en condiciones rigurosas. En general, las condiciones rigurosas se seleccionan para que sean aproximadamente 5°C menores que el punto de fusión térmico (Tm) para la secuencia específica a una fuerza iónica y pH definidos. Sin embargo, las condiciones rigurosas abarcan temperaturas en el rango de entre aproximadamente 1°C y aproximadamente 20°C, dependiendo del grado deseado de rigurosidad calificado de otra manera en la presente.

[143] Para la comparación de secuencia, habitualmente una secuencia actúa como secuencia de referencia contra la cual se comparan las secuencias de ensayo. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de ensayo y de referencia se ingresan en una computadora, se diseñan las coordenadas subsecuencia si es necesario, y se diseñan los parámetros del programa de algoritmo de secuencia. El algoritmo de comparación de secuencias calcula luego el porcentaje de identidad de secuencia para la o las secuencias de ensayo en relación a la secuencia de referencia, en base a los parámetros diseñados del programa.

[144] Como se indica en la presente, otra indicación de que dos secuencias de ácido nucleico son sustancialmente idénticas es que las dos moléculas hibridizan entre sí en condiciones rigurosas. La expresión "que hibridiza específicamente con" se refiere a la unión, formación de dúplex, o hibridización de una molécula solo con una secuencia nucleotídica particular en condiciones rigurosas cuando esa secuencia está presente en una mezcla compleja de ADN o ARN (por ejemplo, célula total). "Que se unen o une sustancialmente" se refiere a la hibridización complementaria entre un ácido nucleico sonda y un ácido nucleico diana y abarca no coincidencias menores que pueden acomodarse mediante reducción de la rigurosidad del medio de hibridización para obtener la detección deseada de la secuencia de ácido nucleico diana.

[145] Las "condiciones de hibridización rigurosas" y las "condiciones de lavado de hibridización rigurosas" en el contexto de los experimentos de hibridización de ácido nucleico tales como hibridaciones Southern y Northern son dependientes de secuencia, y son diferentes en parámetros ambientales diferentes. Las secuencias más largas hibridizan específicamente a temperaturas más altas. La T_m es la temperatura (en una fuerza iónica y pH definidos) a la cual el 50% de la secuencia diana hibridiza con una sonda coincidente perfectamente. La especificidad habitualmente es la función de los lavados posthibridización, siendo los factores críticos la fuerza iónica y la temperatura de la solución de lavado final. Para los híbridos de ADN-ADN, la T_m puede ser aproximada a partir de la ecuación: $T_m = 81,5^\circ\text{C} + 16,6 (\log M) + 0,41 (\%GC) - 0,61 (\% \text{ de form}) - 500/L$; en donde M es la molaridad de los cationes monovalentes, %GC es el porcentaje de nucleótidos guanosina y citosina en el ADN, % de form es el porcentaje de formamida en la solución de hibridización, y L es la longitud del híbrido en pares de bases. La T_m se reduce en aproximadamente 1°C por cada 1% de no coincidencia; por lo tanto, la T_m , condiciones de hibridización, y/o de lavado pueden ajustarse para hibridizar con secuencias de la identidad deseada. Por ejemplo, si se buscan secuencias con >90% de identidad, la T_m puede disminuirse 10°C . En general, se seleccionan condiciones rigurosas para que sean aproximadamente 5°C menores que el punto de fusión térmica (T_m) para la secuencia específica y es complementaria con una fuerza iónica y pH definidos. Sin embargo, las condiciones severamente rigurosas pueden utilizar una hibridización y/o lavado a $1, 2, 3, \text{ o } 4^\circ\text{C}$ más baja que el punto de fusión térmica (T_m); las condiciones rigurosas pueden utilizar una hibridización y/o lavado a $6, 8, 9, \text{ o } 10^\circ\text{C}$ más bajo que el punto de fusión térmica (T_m); las condiciones de baja rigurosidad pueden utilizar una hibridización y/o lavado a $11, 12, 13, 14, 15, \text{ o } 20^\circ\text{C}$ más bajo que el punto de fusión térmica (T_m). Usando la ecuación, las composiciones de hibridización y lavado, y la T_m deseada, aquellos con experiencia ordinaria comprenderán que las variaciones en la rigurosidad de las soluciones de hibridización y/o lavado se describen de manera inherente. Si el grado deseado de no coincidencias resulta en una T_m menor que 45°C (solución acuosa) o 32°C (solución de formamida), se prefiere aumentar la concentración de SSC de manera que puede usarse una temperatura más alta. En general, se seleccionan condiciones de hibridización y lavado muy rigurosas para que sean aproximadamente 5°C menores que el punto de fusión térmica (T_m) para la secuencia específica a una fuerza iónica y pH definidos.

[146] Un ejemplo de condiciones de lavado muy rigurosas es NaCl 0,15 M a 72°C durante aproximadamente 15 minutos. Un ejemplo de condiciones de lavado rigurosas es un lavado con 0,2X SSC a 65°C durante 15 minutos (véase, Sambrook and Russell 2001, para una descripción de la solución amortiguadora SSC). Con frecuencia, un lavado con alta rigurosidad es precedido por un lavado con baja rigurosidad para eliminar la señal de sonda de fondo. Para secuencias de ácido nucleico corto (por ejemplo, aproximadamente entre 10 y 50 nucleótidos), las condiciones rigurosas habitualmente involucran concentraciones salinas menores que aproximadamente 1,5 M, más preferiblemente aproximadamente entre 0,01 y 1,0 M, concentración del ion Na (u otras sales) a pH entre 7,0 y 8,3, y la temperatura habitualmente es por lo menos aproximadamente 30°C . Las condiciones rigurosas también pueden alcanzarse mediante el agregado de agentes desestabilizantes tales como formamida. En general, una relación de señal a ruido de 2X (o mayor) que la observada para una sonda no relacionada en el ensayo de hibridización particular indica la detección de una hibridización específica. Se seleccionan condiciones muy rigurosas para que sean iguales a la T_m de una molécula ácido nucleico particular.

[147] Se seleccionan condiciones muy rigurosas para que sean iguales a la T_m de una sonda particular. Un ejemplo de condiciones rigurosas para hibridización de ácidos nucleicos complementarios que tienen más de 100 residuos complementarios en un filtro en una transferencia Southern o Northern es 50% de formamida, por ejemplo, hibridización en formamida 50%, NaCl 1 M, SDS 1% a 37°C , y un lavado en 0,1X SSC a entre $60 \text{ y } 65^\circ\text{C}$. Las condiciones de baja rigurosidad ejemplificativas incluyen hibridización con una solución amortiguadora de entre 30 y 35% de formamida, NaCl 1M, SDS 1% (dodecilsulfato de sodio) a 37°C , y un lavado en 1X a 2X SSC (20X SSC = NaCl 3,0 M/citrato trisódico 0,3 M) a entre $50 \text{ y } 55^\circ\text{C}$. Las condiciones moderadamente rigurosas ejemplificativas incluyen hibridización en formamida entre 40 y 45%, NaCl 1,0 M, SDS 1% a 37°C , y un lavado en 0,5X a 1X SSC a entre $55 \text{ y } 60^\circ\text{C}$.

[148] El término "transformación" se refiere a la transferencia de un fragmento de ácido nucleico en el genoma of de una célula huésped, que resulta en una herencia estable genéticamente. Una "célula huésped" es una célula que ha sido transformada, o que tiene la capacidad de ser transformada, por una molécula de ácido nucleico exógena. Las células huéspedes que contienen los fragmentos de ácido nucleico transformados son referidas como células "transgénicas".

[149] "Transformada", "transducida", "transgénica" y "recombinante" se refiere a una célula huésped en la cual se ha introducido una molécula de ácido nucleico heteróloga. Como se usa en la presente el término "transfección" se refiere a la administración de ADN en células de eucariota (por ejemplo, mamífero). El término "transformación" se usa en la presente para referirse a la administración de ADN en células procariotas (por ejemplo, E. coli). El término "transducción" se usa en la presente para referirse a infectar células con partículas virales. La molécula de ácido nucleico puede integrarse de manera estable en el genoma conocido en general en el arte. Los métodos conocidos de PCR incluyen, a título enunciativo no taxativo, métodos que usan cebadores apareados, cebadores anidados, cebadores específicos individuales, cebadores degenerados, cebadores específicos de gen, cebadores específicos de vector, cebadores parcialmente sin coincidencia, y similares. Por ejemplo, las células "transformadas",

"transformantes", y "transgénicas" han atravesado el proceso de transformación y contienen un gen foráneo integrado en su cromosoma. El término "sin transformar" se refiere a células normales que no han pasado por el proceso de transformación.

5 [150] "Células genéticamente alteradas" indica células que han sido modificadas mediante la introducción de ácidos nucleicos recombinantes o heterólogos (por ejemplo, una o más construcciones de ADN o sus contrapartes de ARN) y que además incluyen la progenie de dichas células que retienen parte o toda la modificación genética.

10 [151] Como se usa en la presente, el término "derivado" o "dirigido a" con respecto a una molécula de nucleótidos se refiere a que la molécula tiene identidad de secuencia complementaria con una molécula particular de interés.

15 [152] Los ARNip de la presente invención pueden ser generados por cualquier método conocido en el arte, por ejemplo, mediante transcripción *in vitro*, de manera recombinante, o por medios sintéticos. En un ejemplo, los ARNip pueden ser generados *in vitro* mediante el uso de una enzima recombinante, tal como ARN polimerasa T7, y moldes de oligonucleótido de ADN.

Moléculas de ácido nucleico de la invención

20 [153] Los términos "asilado y/o purificado" se refieren al aislamiento *in vitro* de un ácido nucleico, por ejemplo, una molécula de ADN o ARN de su ambiente celular natural, y de la asociación con otros componentes de la célula, tal como ácido nucleico o polipéptido, de manera que puede ser secuenciado, replicado, y/o expresado. El ARN o ADN está "aislado" en el sentido que está libre de por lo menos un ácido nucleico contaminante con el cual está normalmente asociado en la fuente natural del ARN o ADN y preferiblemente está libre sustancialmente de cualquier otro ARN o ADN de mamífero. La expresión "libre de por lo menos un ácido nucleico contaminante de la fuente con el cual está normalmente asociado" incluye el caso en el cual el ácido nucleico es reintroducido en la fuente o célula natural pero está en una ubicación del cromosoma diferente o de otra manera está flanqueado por secuencias de ácido nucleico que no se encuentran habitualmente en la célula de origen, por ejemplo, en un vector o plásmido.

30 [154] Además de una secuencia de ADN que codifica para un ARNip, las moléculas de ácido nucleico de la invención incluyen moléculas de ARN de interferencia de hebra doble, que también son útiles para inhibir la expresión de un gen diana.

35 [155] Como se usa en la presente, el término "ácido nucleico recombinante", por ejemplo, "secuencia o segmento de ADN recombinante" se refiere a un ácido nucleico, por ejemplo, a ADN, que ha sido derivado o aislado de cualquier fuente celular apropiada, que puede ser alterada posteriormente de manera química *in vitro*, de forma que su secuencia no sea de origen natural, o corresponda a secuencias de origen natural que no están posicionadas tal como estarían posicionadas en un genoma que no ha sido transformado con ADN exógeno. Un ejemplo de ADN preseleccionado "derivado" de una fuente podría ser una secuencia de ADN que es identificada como un fragmento útil en un organismo dado, y que luego es sintetizado de manera química en forma esencialmente pura. Un ejemplo de dicho ADN "aislado" de una fuente podría ser una secuencia de ADN útil que es escindida y eliminada de una fuente por medios químicos, por ejemplo, mediante el uso de endonucleasas de restricción, de manera que pueda ser manipulada adicionalmente, por ejemplo, amplificada, para usar en la invención, mediante la metodología de la ingeniería genética. "ADN recombinante" incluye secuencias de ADN completamente sintéticas, secuencias de ADN semisintéticas, secuencias de ADN aisladas de fuentes biológicas, y secuencias de ADN derivadas de ARN, así como mezclas de las mismas.

Cassettes de expresión de la invención

50 [156] Para preparar cassettes de expresión, la secuencia o segmento de ADN recombinante puede ser circular o lineal, de hebra doble o de hebra simple. En general, la secuencia o segmento de ADN está en la forma de ADN quimérico, tal como ADN plasmídico o un vector que también pueden contener regiones codificantes flanqueadas por secuencias control que promueven la expresión del ADN recombinante presente en la célula transformada resultante.

55 [157] Un vector o cassette de expresión "quimérico", como se usa en la presente, se refiere a un vector o cassette que incluye secuencias de ácido nucleico de por lo menos dos especies diferentes, o tiene una secuencia de ácido nucleico de la misma especie que está unido o asociado de una manera que no existe en la forma "nativa" o de tipo salvaje de la especie.

60 [158] Aparte de que las secuencias de ADN recombinante que sirven como unidades de transcripción para un transcripto de ARN, o partes del mismo, una parte del ADN recombinante puede ser no transcripto, sirviendo como una función reguladora o estructural. Por ejemplo, el ADN recombinante puede tener un promotor que es activo en las células de mamífero.

65 [159] Otros elementos funcionales en las células huéspedes, tales como intrones, potenciadores, secuencias de

poliadenilación y similares, también pueden ser parte del ADN recombinante. Dichos elementos pueden o no ser necesarios para la función del ADN, pero pueden proveer una expresión mejorada del ADN afectando la transcripción, estabilidad del ARNip, o similar. Dichos elementos pueden incluirse en el ADN según se desee para obtener el rendimiento óptimo del ARNip en la célula.

[160] Las secuencias de control son secuencias de ADN necesarias para la expresión de una secuencia codificante ligada operativamente en un organismo huésped particular. Las secuencias de control que son adecuadas para las células procariotas, por ejemplo, incluyen un promotor, y opcionalmente una secuencia operadora, y un sitio de unión a ribosoma. Se sabe que las células eucariotas utilizan promotores, señales de poliadenilación, y potenciadores.

[161] Los ácidos nucleicos ligados operativamente son ácidos nucleicos ubicados en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, un promotor o potenciador está ligado operativamente a una secuencia codificante si afecta la transcripción de la secuencia; o un sitio de unión a ribosoma está ligado operativamente a una secuencia codificante si está posicionada de manera que facilite la traducción. En general, las secuencias de ADN ligadas operativamente son secuencias de ADN que están unidas de manera contigua. Sin embargo, los potenciadores no tienen que ser contiguos. La unión puede lograrse mediante el ligado en sitios de restricción convenientes. Si dichos sitios no existen, se usan adaptadores de oligonucleótidos sintéticos o conectores de acuerdo con la práctica convencional.

[162] El ADN recombinante a introducir en las células puede contener un gen marcador de selección o un gen reportero o ambos para facilitar la identificación y selección de células que expresan de la población de células que se piensan transfectadas o infectadas a través de vectores virales. En otras formas de realización, el marcador de selección puede portado en una pieza separada de ADN y usado en un procedimiento de cotransfección. Los marcadores de selección y los genes reporteros pueden estar flanqueados por secuencias reguladoras apropiadas para permitir la expresión en las células huéspedes. Los marcadores de selección útiles son conocidos en el arte e incluyen, por ejemplo, genes de resistencia a antibióticos, tales como *neo* y similares.

[163] Los genes reporteros se usan para identificar células potencialmente transfectadas y para evaluar la funcionalidad de las secuencias reguladoras. Los genes reporteros que codifican para proteínas de fácil ensayo son bien conocidos en el arte. En general, un gen reportero es un gen que no está presente en el organismo o tejido receptor o que no es expresado por el mismo y que codifica para una proteína cuya expresión se manifiesta por alguna propiedad fácilmente detectable, por ejemplo, actividad enzimática. Por ejemplo, los genes reporteros incluyen el gen de cloranfenicol acetil transferasa (*cat*) de Tn9 de *E. coli* y el gen de luciferasa de la luciérnaga *Photinus pyralis*. La expresión del gen reportero se ensaya en un tiempo adecuado después que se ha introducido el ADN en las células receptoras.

[164] Los métodos generales para la construcción de ADN recombinante que puede transfectar células diana son bien conocidos por aquellos con experiencia en el arte, y las mismas composiciones y métodos de construcción pueden utilizarse para producir el ADN útil en la presente.

[165] El ADN recombinante puede ser introducido fácilmente en las células huéspedes, por ejemplo, células de mamífero, bacterias, levaduras o de insecto mediante la transfección con un vector de expresión compuesto por ADN que codifica para el ARNip mediante cualquier procedimiento útil para la introducción en una célula particular, por ejemplo, métodos físicos o biológicos, para obtener una célula que tiene el ADN recombinante integrado de manera estable en su genoma o que existe como un elemento episomal, de manera que las moléculas o secuencias de ADN de la presente invención son expresadas por la célula huésped. Preferiblemente, el ADN es introducido en las células huéspedes mediante un vector. La célula huésped preferiblemente es de origen eucariota, por ejemplo, fuente vegetal, de mamífero, de insecto, de levadura o fúngica, pero también pueden usarse células huéspedes de origen no eucariota.

[166] Los métodos físicos para introducir un ADN preseleccionado en una célula huésped incluyen precipitación con fosfato de calcio, lipofección, bombardeo de partículas, microinyección, electroporación, y similares. Los métodos biológicos para introducir el ADN de interés en una célula huésped incluyen el uso de vectores virales de ADN y ARN. Para la terapia génica en mamíferos, como se describe más adelante en la presente, se desea usar un medio eficaz para insertar una copia de gen en el genoma del huésped. Los vectores virales, y especialmente los vectores retrovirales, se han vuelto el método más ampliamente usado para insertar genes en células de mamífero, por ejemplo, de humano. Otros vectores virales pueden derivarse de poxvirus, herpes simple virus I, adenovirus y virus adenoasociados, y similares. Véase, por ejemplo, Patente de los EE.UU. Nos. 5.350.674 y 5.585.362.

[167] Tal como se describe en la presente, una célula huésped o línea celular "transfectada" o "transducida" es una en la cual el genoma ha sido alterado o aumentado por la presencia de por lo menos una secuencia de ácido nucleico heteróloga o recombinante. Las células huéspedes de la presente invención habitualmente se producen por transfección con una secuencia de ADN en un vector de expresión plasmídico, un vector de expresión viral, o como una secuencia de ADN lineal aislada. El ADN transfectado puede volverse una secuencia de ADN recombinante integrada cromosómicamente, que está compuesta por la secuencia que codifica para el ARNip.

[168] Para confirmar la presencia de la secuencia de ADN recombinante en la célula huésped, pueden realizarse una variedad de ensayos. Dichos ensayos incluyen, por ejemplo, ensayos de "biología molecular" bien conocidos por aquellos con experiencia en el arte, tales como transferencia Southern y Northern, RT-PCR y PCR; ensayos "bioquímicos", tales como detección de la presencia o ausencia de un péptido particular, por ejemplo, mediante medios inmunológicos (ELISA y transferencias Western) o mediante ensayos descritos en la presente para identificar agentes que caen dentro del alcance de la invención.

[169] Para detectar y cuantificar el ARN producido a partir de segmentos del ADN recombinante introducido, puede emplearse RT-PCR. En esta aplicación de la PCR, primero es necesario transcribir de manera reversa el ARN a ADN, usando enzimas tales como la transcriptasa reversa, y luego mediante el uso de técnicas de PCR convencional amplificar el ADN. En la mayoría de los casos las técnicas de PCR, si bien son útiles, no demostrarán la integridad del producto de ARN. Mediante transferencia Northern puede obtenerse información adicional sobre la naturaleza del producto de ARN. Esta técnica demuestra la presencia de una especie de ARN y brinda información sobre la integridad del ARN. La presencia o ausencia de una especie de ARN también puede ser determinada usando hibridizaciones Northern dot o slot blot. Estas técnicas son modificaciones de la transferencia Northern y solo demuestran la presencia o ausencia de una especie de ARN.

[170] Si bien la transferencia Southern y la PCR pueden usarse para detectar el segmento de ADN recombinante de interés, no proveen información sobre cuál segmento de ADN preseleccionado se está expresando. La expresión puede evaluarse mediante la identificación específica de los productos peptídicos de las secuencias de ADN recombinante introducidas o la evaluación de los cambios fenotípicos generados por la expresión del segmento de ADN recombinante introducido en la célula huésped.

[171] La presente invención provee un sistema de expresión celular para la expresión de material de ácido nucleico exógeno en un receptor mamífero. El sistema de expresión, también referido como una "célula modificada genéticamente", comprende una célula y un vector de expresión para expresar el material de ácido nucleico exógeno. Las células modificadas genéticamente son adecuadas para la administración a un receptor mamífero, en donde reemplazan las células endógenas del receptor. Por lo tanto, las células modificadas genéticamente preferidas son no inmortalizadas y no tumorigénicas.

[172] De acuerdo con una forma de realización, las células son transfectadas o modificadas genéticamente de otra manera *ex vivo*. Las células son aisladas de un mamífero (preferiblemente un humano), se introduce un ácido nucleico (es decir, transducen o transfectan *in vitro*) con un vector para la expresión de un gen heterólogo (por ejemplo, recombinante) que codifica para el agente terapéutico, y luego se administran a un receptor mamífero para la administración del agente terapéutico *in situ*. El receptor mamífero puede ser un humano y las células a modificar son células autólogas, es decir, las células son aisladas del receptor mamífero.

[173] De acuerdo con otra forma de realización, las células son transfectadas o transducidas o modificadas genéticamente de otra manera *in vivo*. Las células del receptor mamífero son transducidas o transfectadas *in vivo* con un vector que contiene el material de ácido nucleico exógeno para expresa un gen heterólogo (por ejemplo, recombinante) que codifica para un agente terapéutico y el agente terapéutico es administrado *in situ*.

[174] Como se usa en la presente, "material de ácido nucleico exógeno" se refiere a un ácido nucleico o un oligonucleótido, natural o sintético, que no se encuentra de manera natural en las células; o si se encuentra naturalmente en las células, está modificado respecto a su forma original o nativa. Por lo tanto, "material de ácido nucleico exógeno" incluye, por ejemplo, un ácido nucleico de origen no natural que puede ser transcrito a un ARN antisentido, un ARNip, así como un "gen heterólogo" (es decir, un gen que codifica para una proteína que no se expresa o que se expresa en niveles biológicamente insignificantes en una célula de origen natural del mismo tipo). Para ilustrar, un gen sintético o natural que codifica para eritropoyetina (EPO) podría ser considerado "material de ácido nucleico exógeno" respecto a células de mesotelio peritoneal de humano ya que las últimas células no expresan EPO naturalmente. Otro ejemplo de "material de ácido nucleico exógeno" es la introducción de solo parte de un gen para crear un gen recombinante, tal como por combinación de un promotor regulable con una secuencia codificante endógena mediante recombinación homóloga.

[175] La condición susceptible a terapia de inhibición génica puede ser un proceso profiláctico, es decir, un proceso para prevenir una enfermedad o una condición médica no deseada. Por lo tanto, la presente invención abarca un sistema para la administración de ARNip que tiene una función profiláctica (es decir, un agente profiláctico) para el receptor mamífero.

Métodos para la introducción de cassettes de expresión de la invención en células

[176] El material de ácido nucleico inhibitorio (por ejemplo, un cassette de expresión que codifica para ARNip dirigido contra un gen de interés) puede ser introducido en la célula *ex vivo* o *in vivo* mediante métodos de transferencia genética, tales como transfección o transducción, para proveer una célula modificada genéticamente.

Una persona con experiencia en el arte conoce diferentes vectores de expresión (es decir, vehículos para facilitar la administración de ácido nucleico exógeno en una célula diana).

[177] Como se usa en la presente, "transfección de células" se refiere a la adquisición en una célula de nuevo material de ácido nucleico por incorporación del ADN agregado. Por lo tanto, la transfección se refiere a la inserción de ácido nucleico en una célula mediante el uso de métodos físicos o químicos. Diferentes técnicas de transfección son conocidas por aquellos con experiencia ordinaria en el arte que incluyen coprecipitación de ADN con fosfato de calcio, DEAE-dextrano, electroporación, transfección mediada por liposomas catiónicos, bombardeo con micropartículas facilitada con partículas de tungsteno, y coprecipitación de ADN con fosfato de estroncio.

[178] Por el contrario, "transducción de células" se refiere al proceso de transferencia de ácido nucleico a una célula usando un virus a ADN o ARN. Un virus a ARN (es decir, un retrovirus) para transferir un ácido nucleico en una célula se refiere en la presente a un retrovirus quimérico de transducción. El material de ácido nucleico exógeno contenido en el retrovirus es incorporado en el genoma de la célula transducida. Una célula que ha sido transducida con un virus a ADN quimérico (por ejemplo, un adenovirus que porta un ADNc que codifica para un agente terapéutico), no tendrá el material de ácido nucleico exógeno incorporado en su genoma pero tendrá la capacidad de expresar el material de ácido nucleico exógeno que está retenido de manera extracromosómica en la célula.

[179] El material de ácido nucleico exógeno puede incluir el ácido nucleico que codifica para el ARNip junto con un promotor para controlar la transcripción. El promotor característicamente tiene una secuencia nucleotídica específica necesaria para iniciar la transcripción. El material de ácido nucleico exógeno además puede incluir secuencias adicionales (es decir, potenciadores) requeridas para obtener la actividad de transcripción génica deseada. Con el propósito de esta descripción un "potenciador" simplemente es cualquier secuencia de ADN no traducida que funciona con la secuencia codificante (en *cis*) para cambiar el nivel de transcripción basal indicado por el promotor. El material de ácido nucleico exógeno puede ser introducido en el genoma de la célula inmediatamente corriente abajo del promotor de manera que el promotor y la secuencia codificante estén operativamente ligados de forma de permitir la transcripción de la secuencia codificante. Un vector de expresión puede incluir un elemento promotor exógeno para controlar la transcripción del gen exógeno insertado. Dichos promotores exógenos incluyen promotores constitutivos y regulables.

[180] Los promotores constitutivos de origen natural controlan la expresión de funciones celulares esenciales. Como resultado, una secuencia de ácido nucleico bajo el control de un promotor constitutivo se expresa en todas las condiciones de crecimiento de la célula. Los promotores constitutivos incluyen los promotores para los siguientes genes que codifican para determinadas funciones constitutivas o "de mantenimiento": promotor de hipoxantina fosforibosil transferasa (HPRT), dihidrofolato reductasa (DHFR), adenosina desaminasa, fosfoglicerol quinasa (PGK), piruvato quinasa, fosfoglicerol mutasa, beta-actina, y otros promotores constitutivos conocidos por aquellos con experiencia en el arte. Además, muchos promotores virales funcionan constitutivamente en células eucariotas. Estos incluyen: los promotores temprano y tardío de SV40; las repeticiones terminales largas (LTR) del virus de leucemia de Moloney y otros retrovirus; y el promotor de timidina quinasa del virus de Herpes Simple, entre muchos otros.

[181] Las secuencias de ácido nucleico que están bajo el control de promotores regulables se expresan solo o en un mayor o menor grado en presencia de un agente inductor o represor, (por ejemplo, la transcripción bajo el control del promotor de metalotioneína aumenta mucho en presencia de determinados iones metálicos). Los promotores regulables incluyen elementos de respuesta (RE) que estimulan la transcripción cuando se unen sus factores inductores. Por ejemplo, existen RE para factores séricos, hormonas esteroideas, ácido retinoico, AMP cíclico, y tetraciclina y doxiciclina. Los promotores que contienen un RE particular pueden seleccionarse con el objetivo de obtener una respuesta que puede regularse y en algunos casos, el RE mismo puede unirse a un promotor diferente, confiriendo de esta manera la capacidad de regulación de la secuencia de ácido nucleico codificada. Por lo tanto, mediante la selección del promotor apropiado (constitutivo respecto a regulable; fuerte respecto a débil), es posible controlar la existencia y el nivel de expresión de una secuencia de ácido nucleico en la célula genéticamente modificada. Si la secuencia de ácido nucleico está bajo el control de un promotor regulable, la administración del agente terapéutico *in situ* es iniciada por exposición de la célula genéticamente modificada *in situ* a condiciones que permitan la transcripción de la secuencia de ácido nucleico, por ejemplo, mediante inyección intraperitoneal de inductores específicos de promotores regulables que controlan la transcripción del agente. Por ejemplo, la expresión *in situ* de una secuencia de ácido nucleico bajo el control del promotor de metalotioneína en células genéticamente modificadas aumenta (es decir, induce) por contacto de las células genéticamente modificadas con una solución que contiene los iones metálicos apropiados *in situ*.

[182] En consecuencia, la cantidad de ARNip generado *in situ* está regulada por el control de dichos factores como la naturaleza del promotor usado para dirigir la transcripción de la secuencia del ácido nucleico, (es decir, si el promotor es constitutivo o regulable, fuerte o débil) y el número de copias de la secuencia de ácido nucleico exógeno que codifica para una secuencia de ARNip que está en la célula.

[183] Además, al por lo menos un promotor y a la por lo menos una secuencia de ácido nucleico heterólogo que codifica para el ARNip, el vector de expresión puede incluir un gen de selección, por ejemplo, un gen de resistencia a neomicina, para facilitar la selección de las células que han sido transfectadas o transducidas con el vector de

expresión.

[184] Las células también pueden ser transfectadas con dos o más vectores de expresión, por lo menos un vector que contiene la o las secuencias de ácido nucleico que codifican para el o los ARNip, el otro vector que contiene un gen de selección. La selección de un promotor adecuado, potenciador, gen de selección, y/o secuencia señal se considera dentro del alcance de una persona con experiencia ordinaria en el arte sin experimentación indebida.

[185] La siguiente descripción está dirigida a diferentes utilidades de la presente divulgación. Por ejemplo, la presente divulgación tiene utilidad como un sistema de expresión adecuado para silenciar el o los genes de interés.

[186] La presente divulgación también provee métodos para modificar genéticamente células de un receptor mamífero *in vivo*. De acuerdo con una forma de realización, el método comprende la introducción de un vector de expresión para expresar una secuencia de ARNip en células del receptor mamífero *in situ* mediante, por ejemplo, la inyección del vector en el receptor.

Vehículos de administración de los cassettes de expresión de la invención

[187] La administración de compuestos en tejidos y a través de la barrera hematoencefálica puede estar limitada por el tamaño y propiedades bioquímicas de los compuestos. En la actualidad, la administración eficaz de los compuestos en células *in vivo* puede ser realizarse solo cuando las moléculas son pequeñas (habitualmente menores a 600 Daltons). La transferencia de genes para la corrección de errores de nacimiento del metabolismo y enfermedades neurodegenerativas del sistema nervioso central (SNC), y para el tratamiento de cáncer ha sido logrado con vectores adenovirales recombinantes.

[188] La selección y optimización de un vector de expresión particular para expresar un ARNip específico en una célula puede lograrse mediante la obtención de la secuencia de ácido nucleico del ARNip, posiblemente con una o más regiones de control apropiadas (por ejemplo, secuencia promotora, de inserción); preparación de una construcción de vector que comprende el vector en el cual se inserta la secuencia de ácido nucleico que codifica para el ARNip; transfectar o transducir las células cultivadas *in vitro* con la construcción del vector; y determinar si el ARNip está presente en las células cultivadas.

[189] Los vectores para la terapia génica en células incluyen virus, tales como virus deficientes para la replicación (descritos con mayor detalle más adelante). Los vectores virales ejemplificativos derivan del virus de sarcoma Harvey, virus del sarcoma de ROUS, (MPSV), virus de leucemia murina de Moloney y virus a ADN (por ejemplo, adenovirus).

[190] Los retrovirus deficientes para la replicación tienen la capacidad de dirigir la síntesis de todas las proteínas del virión, pero no tienen la capacidad de elaborar partículas infecciosas. En consecuencia, estos vectores de expresión retrovirales genéticamente alterados tienen utilidad general para la transducción de alta eficacia de secuencias de ácido nucleico en células cultivadas, y utilidad específica para usar en el método de la presente divulgación. Dichos retrovirus además tienen utilidad para la transducción eficaz de secuencias de ácido nucleico en células *in vivo*. Los retrovirus han sido usados extensamente para transferir material de ácido nucleico a células. Los protocolos para producir retrovirus deficientes para la replicación (incluyendo los pasos de incorporación de material de ácido nucleico exógeno en un plásmido, transfección de una línea de células empaquetadoras con plásmido, producción de retrovirus recombinantes por la línea de células empaquetadoras, recolección de las partículas virales del medio de cultivo de tejido, e infección de las células diana con las partículas virales) son bien conocidos en el arte.

[191] Una ventaja de usar retrovirus para la terapia génica es que los virus insertan la secuencia de ácido nucleico que codifica para el ARNip en el genoma de la célula huésped, permitiendo de esta manera que la secuencia de ácido nucleico que codifica para el ARNip pase a la progenie de la célula cuando se divide. Las secuencias promotoras en la región LTR pueden aumentar la expresión de una secuencia codificante insertada en una variedad de tipos de células. Algunas desventajas de usar un vector de expresión a retrovirus son (1) mutagénesis por inserción, es decir, la inserción de la secuencia de ácido nucleico que codifica para el ARNip en una posición no deseada en el genoma de la célula diana que, por ejemplo, lleva al crecimiento celular no regulado y (2) la necesidad de proliferación de la célula diana con el objetivo de que la secuencia de ácido nucleico que codifica para el ARNip portada por el vector sea integrada al genoma diana.

[192] Otro candidato viral útil como un vector de expresión para la transformación de células es el adenovirus, un virus a ADN de hebra doble. El adenovirus no es eficaz en un rango amplio de tipos de células, incluyendo, por ejemplo, células musculares y endoteliales.

[193] Los adenovirus (Ad) son virus a ADN lineales de hebra doble con un genoma de 36 kb. Varias características de los adenovirus los han hecho útiles como vehículos de administración de transgenes para aplicaciones terapéuticas, tal como para facilitar la administración de genes *in vivo*. Se ha mostrado que los vectores

de adenovirus recombinantes tienen la capacidad de ser eficaces para la transferencia de genes *in situ* a células del parénquima de diferentes órganos, incluyendo los pulmones, cerebro, páncreas, vesícula biliar e hígado. Esto ha permitido el uso de estos vectores en métodos para tratar enfermedades genéticas hereditarias, tales como la fibrosis quística, en donde los vectores pueden ser administrados a un órgano blanco. Además, la capacidad del vector a adenovirus para lograr la transducción del tumor *in situ* ha permitido el desarrollo de una variedad de métodos de terapia génica contra el cáncer para enfermedades no diseminadas. En estos métodos, la contención del vector favorece la transducción específica de la célula tumoral.

[194] Al igual que los retrovirus, el genoma del adenovirus puede adaptarse para usar como un vector de expresión para la terapia génica, es decir, por eliminación de la información genética que controla la producción del virus mismo. Debido a que el adenovirus funciona de una manera extracromosómica, los adenovirus recombinantes no tienen el problema teórico de mutagénesis por inserción.

[195] Se han usado varias estrategias tradicionalmente para generar los adenovirus recombinantes. Una estrategia involucra el ligado directo de los fragmentos de endonucleasa de restricción que contienen una secuencia de ácido nucleico de interés a partes del genoma adenoviral. Como alternativa, la secuencia de ácido nucleico de interés puede insertarse en un adenovirus defectuoso mediante recombinación homóloga. Los recombinantes deseados se identifican mediante el cribado de placas individuales generadas en una capa de células de complemento.

[196] La mayor parte de los vectores de adenovirus se basan en el esqueleto del adenovirus tipo 5 (Ad5) en el cual se ha introducido un cassette de expresión que contiene la secuencia de ácido nucleico de interés en lugar de la región temprana 1 (E1) o región temprana 3 (E3). Los virus en los cuales se ha eliminado E1 son defectuosos para la replicación y se propagan en células de complemento humanas (por ejemplo, 293 o 911 células), las cuales proveen el gen E1 perdido y pIX en *trans*.

[197] En una forma de realización de la presente divulgación, una persona deseará generar ARNip en una célula de cerebro o tejido de cerebro. Un vector adecuado para esta aplicación es un vector FIV o un vector AAV. Por ejemplo, uno puede usar AAV5. Además, uno puede aplicar los vectores de poliovirus o HSV.

[198] La aplicación de ARNip en general se realiza mediante transfección de ARNip sintéticos, ARN sintetizados *in vitro*, o plásmidos que expresan ARNhp o miARN. Más recientemente, los virus se han empleado para estudios *in vitro* y para generar noqueos en ratones transgénicos de genes diana. Los adenovirus recombinantes, virus adenoasociados (AAV) y virus de inmunodeficiencia felina (FIV) pueden usarse para administrar genes *in vitro* e *in vivo*. Cada uno tiene sus propias ventajas y desventajas. Los adenovirus son virus a ADN de hebra doble con grandes genomas (36 kb) y han sido diseñados en el laboratorio de los inventores y otros para acomodar cassettes de expresión en diferentes regiones.

[199] Los virus adenoasociados tienen genomas encapsidados, similar a los Ad, pero son más chicos de tamaño y la capacidad de empaquetado (-30 nm vs. -100 nm; límite de empaquetamiento de -4,5 kb). Los AAV contienen genomas a ADN de hebra simple de hebra + o -. Se han estudiado ocho serotipos de AAV (1-8) de manera extensa, tres de los cuales han sido evaluados en el cerebro. Una consideración importante para la presente solicitud es que el AAV5 transduce neuronas del cuerpo estriado y corticales, y no se asocia con ninguna patología conocida.

[200] El virus adenoasociado (AAV) es un virus no patógeno pequeño de la familia de parvoviridae. El AAV es diferente a los otros miembros de esta familia por su dependencia de un virus colaborador para la replicación. En ausencia de un virus colaborador, el AAV puede integrarse de una manera específica de *locus* en el brazo q del cromosoma 19. El genoma de aproximadamente 5 kb del AAV consiste en un segmento de ADN de hebra simple de polaridad positiva o negativa. Los extremos del genoma son repeticiones invertidas cortas que pueden plegarse en estructuras de horquilla y servir como el origen de la replicación del ADN viral. Físicamente, el virión del parvovirus no es encapsulado y su cápside icosaédrica tiene aproximadamente 20 nm de diámetro.

[201] Además, se describen virus quiméricos en donde AAV puede combinarse con virus herpes, amplicones de virus herpes, baculovirus u otros virus para lograr un tropismo deseado asociado con otro virus. Por ejemplo, las ITR de AAV4 podrían insertarse en el virus herpes y podrían infectarse las células. Luego de la infección, las ITR de AAV4 podrían actuar sobre el rep de AAV4 provisto en el sistema o en un vehículo separado para rescatar AAV4 del genoma. Por lo tanto, el tropismo celular del virus de herpes simple puede combinarse con la integración dirigida mediada por rep de AAV4. Otros virus que podrían utilizarse para construir virus quiméricos incluyen lentivirus, retrovirus, vectores retrovirales pseudotipo, y vectores adenovirales.

[202] También se describen vectores AAV variantes. Por ejemplo, la secuencia de un AAV nativo, tal como AAV5, puede modificarse en nucleótidos individuales. La presente divulgación incluye vectores AAV nativos y mutantes. La presente divulgación además incluye todos los serotipos de AAV.

[203] El FIV es un virus con envoltura con un fuerte perfil de seguridad en humanos; los individuos mordidos o arañados por gatos infectados con FIV no seroconvierten y no se ha reportado que presenten ningún signo de

enfermedad. Como los AAV, los FIV proveen expresión del transgén duradera en neuronas de ratones y de primates no humanos, y la transducción puede ser dirigida a diferentes tipos de células por pseudotipificación, el proceso de intercambio de la envoltura nativa del virus por otra envoltura de otro virus.

5 [204] Por lo tanto, tal como resultará evidente para una persona con experiencia ordinaria en el arte, se encuentra disponible una variedad de vectores virales de expresión adecuados para la transferencia de material de ácido nucleico exógeno a las células. La selección de un vector de expresión apropiado para expresar un agente terapéutico para una condición particular sensible a la terapia de silenciamiento génico y la optimización de las condiciones para la inserción del vector de expresión seleccionado en la célula, se encuentran dentro del alcance de una persona con experiencia ordinaria en el arte sin la necesidad de experimentación indebida.

15 [205] En otra forma de realización, el vector de expresión está en la forma de un plásmido, el cual se transfiere a las células diana mediante uno de una variedad de métodos: físicos (por ejemplo, microinyección, electroporación, carga por raspado, bombardeo con micropartículas) o por incorporación celular como un complejo químico (por ejemplo, coprecipitación con calcio o estroncio, formación de complejo con lípidos, formación de complejo con ligando). Se encuentran disponibles diferentes productos comerciales para la formación de complejos con liposomas catiónicos que incluyen Lipofectin™ (Gibco-BRL, Gaithersburg, Md.) y Transfectam™ (Promega®, Madison, Wis.). Sin embargo, la eficacia de la transfección por estos métodos es muy dependiente de la naturaleza de la célula diana y en consecuencia, las condiciones para la transfección óptima de los ácidos nucleicos a las células usando los procedimientos mencionados en la presente deben optimizarse. Dicha optimización se encuentra dentro del alcance una persona con experiencia ordinaria en el arte sin la necesidad de experimentación indebida.

Virus adenoasociados (AAV)

25 [206] El virus adenoasociado (AAV) es un virus no patogénico pequeño de la familia parvoviridae. El AAV es diferente de los otros miembros de esta familia por su dependencia de un virus colaborador para la replicación. En ausencia de un virus colaborador, el AAV puede integrarse de una manera específica de *locus* en el brazo q del cromosoma 19. El genoma de aproximadamente 5 kb del AAV consiste en un segmento de ADN de hebra simple de polaridad positiva o negativa. Los extremos del genoma son repeticiones invertidas cortas que pueden plegarse en estructuras de horquilla y servir como el origen de la replicación del ADN viral. Físicamente, el virión del parvovirus no es encapsulado y su cápside icosaédrica tiene aproximadamente 20 nm de diámetro.

35 [207] Hasta la actualidad, se han identificado diferentes AAV serológicamente distintos, y más de una docena han sido aislados de humano o primates. El genoma de AAV2 tiene 4680 nucleótidos de longitud y contiene dos marcos de lectura abiertos (ORF). El ORF izquierdo codifica para las proteínas Rep no estructurales, Rep 40, Rep 52, Rep 68 y Rep 78, que están involucradas en la regulación de la y transcripción además de la producción de genomas de la progenie de hebra simple. Además, dos de las proteínas Rep han sido asociadas con la integración preferencial de genomas de AAV en una región del brazo q del cromosoma 19 de humano. También se ha mostrado que Rep68/78 posee actividad de unión a NTP así como actividades de ADN y ARN helicasa. Las proteínas Rep poseen una señal de localización nuclear así como varios sitios de fosforilación potencial. La mutación de uno de estos sitios de quinasa produce la pérdida de la actividad de replicación.

45 [208] Los extremos del genoma son repeticiones terminales invertidas cortas (ITR) que tienen el potencial de plegarse en estructuras de horquilla con forma de T que sirven como el origen de la replicación de ADN viral. Dentro de la región ITR se han descrito dos elementos que son centrales para la función de la ITR, un motivo de repetición GAGC y el sitio de resolución terminal (*trs*). Se ha mostrado que el motivo de repetición se une a Rep cuando la ITR está en conformación lineal o en horquilla. Esta unión sirve para posicionar a Rep68/78 para la escisión en el *trs* que se produce de un modo específico de sitio y de hebra. Además de su rol en la replicación, estos dos elementos parecen ser centrales para la integración viral. Existe un sitio de unión a Rep contenido en el *locus* de integración del cromosoma 19 con un *trs* adyacente. Se ha mostrado que estos elementos son funcionales y necesarios para la integración específica de *locus*.

55 [209] El virión AAV es una partícula sin envoltura, icosaédrica de aproximadamente 25 nm de diámetro, que consiste en tres proteínas relacionadas referidas como VP1, VP2 y VP3. El ORF derecho codifica para las proteínas de la cápside VP1, VP2, y VP3. Estas proteínas se encuentran en una proporción de 1:1:10 respectivamente y todas derivan del ORF con giro derecho. Las proteínas de la cápside difieren entre sí por el uso de corte y empalme alternativo y un codón de inicio inusual. El análisis de eliminación ha mostrado que la eliminación o alteración de VP1 que es traducida a partir de un mensajero con corte y empalme alternativo genera un rendimiento reducido de partículas infecciosas. Las mutaciones en la región codificante de VP3 genera una falla para producir cualquier ADN de progenie de hebra simple o partículas infecciosas. Una partícula de AAV es una partícula viral que comprende una proteína de cápside de AAV. Un polipéptido de la cápside de AAV puede codificar para el polipéptido completo de VP1, VP2 y VP3. La partícula puede ser una partícula que comprende proteínas de cápside de AAV2 y otros AAV (es decir, una proteína quimérica, tal como AAV1 y AAV2). En la presente se contemplan variaciones en la secuencia de aminoácidos de la proteína de cápside de AAV2, con la condición de que la partícula viral resultante que comprende la cápside de AAV2 permanezca antigénica o inmunológicamente diferente de AAV1, como puede ser determinado de rutina mediante los métodos estándares. Específicamente, por ejemplo, puede usarse ELISA y

transferencias Western para determinar si una partícula viral es antigénica o inmunológicamente diferente de AAV1. Además, la partícula viral de AAV2 preferiblemente retiene el tropismo de tejido diferente de AAV1.

[210] Una partícula AAV2 es una partícula viral que comprende una proteína de cápside de AAV2. Un polipéptido de la cápside de AAV2 que codifica para el polipéptido completo de VP1, VP2 y VP3 puede tener en total por lo menos aproximadamente 63% de homología (o identidad) con el polipéptido que tiene la secuencia de aminoácidos codificada por los nucleótidos descritos en NC_001401 (secuencia nucleotídica que codifica para la proteína de cápside de AAV2). La proteína de la cápside puede tener aproximadamente 70% de homología, aproximadamente 75% de homología, 80% de homología, 85% de homología, 90% de homología, 95% de homología, 98% de homología, 99% de homología, o incluso 100% de homología con la proteína codificada por la secuencia nucleotídica descrita en NC_001401. La proteína de la cápside puede tener aproximadamente 70% de identidad, aproximadamente 75% de identidad, 80% de identidad, 85% de identidad, 90% de identidad, 95% de identidad, 98% de identidad, 99% de identidad, o incluso 100% de identidad con la proteína codificada por la secuencia nucleotídica descrita en NC_001401. La partícula puede ser una partícula que comprende proteínas de cápside de otros AAV y AAV2, es decir, una proteína quimérica. En la presente están contempladas variaciones en la secuencia de aminoácidos de la proteína de cápside de AAV2, con la condición de que la partícula viral resultante que comprende la cápside de AAV2 permanezca antigénica o inmunológicamente diferente de AAV4, como puede ser determinado de rutina mediante métodos estándares. Específicamente, por ejemplo, puede usarse ELISA y transferencias Western para determinar si una partícula viral es antigénica o inmunológicamente diferente de AAV1. Además, la partícula viral de AAV2 preferiblemente retiene el tropismo de tejido diferente de AAV1, tal como lo que se ejemplifica en los ejemplos en la presente, aunque una partícula quimérica de AAV2 que comprende por lo menos una proteína de recubrimiento de AAV2 puede tener un tropismo de tejido diferente del de una partícula de AAV2 que consiste solo de proteínas de recubrimiento de AAV2.

[211] En determinadas formas de realización, la invención además provee una partícula de AAV2 que contiene, es decir, encapsida, un vector que comprende un par de repeticiones terminales invertidas de AAV2. La secuencia nucleotídica de las ITR de AAV2 es conocida en el arte. Además, la partícula puede ser una partícula que comprende proteínas de cápside de AAV1 y AAV2, es decir, una proteína quimérica. Además, la partícula puede ser una partícula que encapsida un vector que comprende un par de repeticiones terminales invertidas de AAV de otros AAV (por ejemplo, AAV1-AAV9 y AAVrh10). El vector encapsidado en la partícula además puede comprender un ácido nucleico exógeno insertado entre las repeticiones terminales invertidas.

[212] Las siguientes características de AAV lo han convertido en un vector atractivo para la transferencia de genes. Se ha mostrado *in vitro* que los vectores de AAV se integran de manera estable en el genoma celular; poseen un rango amplio de huéspedes; transducen células en división y en no división *in vitro* e *in vivo* y mantienen niveles altos de expresión de los genes transducidos. Las partículas virales son estables al calor, resistentes a solventes, detergentes, cambios en el pH, temperatura, y pueden ser concentradas en gradientes de CsCl o por otros medios. Se describen métodos para administrar partículas de AAV, vectores de AAV recombinantes, y viriones de AAV recombinantes. Por ejemplo, una partícula de AAV2 es una partícula viral que comprende una proteína de cápside de AAV2, o una partícula de AAV1 es una partícula viral que comprende una proteína de la cápside de AAV1. Un vector de AAV2 recombinante es una construcción de ácido nucleico que comprende por lo menos un ácido nucleico único de AAV2. Un virión de AAV2 recombinante es una partícula que contiene un vector de AAV2 recombinante. En el término "ITR de AAV2" se debe considerar que la secuencia nucleotídica debe retener una o ambas características descritas en la presente que distinguen a la ITR de AAV2 de la ITR de AAV1 ITR: (1) tres (en lugar de cuatro como en AAV1) repeticiones "GAGC" y en el sitio de unión a Rep de la ITR de AAV2 el cuarto nucleótido en las primeras dos repeticiones "GAGC" es una C en lugar de una T.

[213] El promotor para dirigir la expresión de la proteína o la secuencia que codifica para otro agente a administrar puede ser cualquier promotor deseado, seleccionado de acuerdo a consideraciones conocidas, tales como el nivel de expresión de un ácido nucleico ligado funcionalmente al promotor y el tipo de célula en el cual se usa el vector. Los promotores pueden ser un promotor exógeno o uno endógeno. Los promotores pueden incluir, por ejemplo, promotores fuertes conocidos tales como SV40 o el promotor inducible de metalotioneina, o un promotor de AAV, tal como un promotor p5 de AAV. Los ejemplos adicionales de promotores incluyen promotores derivados de genes de actina, genes de inmunoglobulina, citomegalovirus (CMV), adenovirus, virus de papiloma bovino, promotores adenovirales, tales como promotor tardío principal adenoviral, un promotor de choque térmico inducible, virus sincicial respiratorio, virus de sarcoma de Rous (RSV), etc. Los ejemplos adicionales incluyen promotores regulados.

[214] El vector AAV además puede comprender un ácido nucleico exógeno (heterólogo) funcionalmente ligado al promotor. Por "ácido nucleico heterólogo" se comprende que cualquier ácido nucleico heterólogo o exógeno puede ser insertado en el vector para la transferencia a una célula, tejido u organismo. El ácido nucleico puede codificar para un polipéptido o proteína o un ARN antisentido, por ejemplo. Por "funcionalmente ligado" se comprende que el promotor puede promover la expresión del ácido nucleico heterólogo, tal como es conocido en el arte, tal como la orientación apropiada del promotor en relación al ácido nucleico heterólogo. Además, el ácido nucleico heterólogo preferiblemente tiene todas las secuencias apropiadas para la expresión del ácido nucleico, como es conocido en el arte, para codificar funcionalmente, es decir, permitir que se exprese el ácido nucleico. El ácido nucleico puede

incluir, por ejemplo, secuencias de control de la expresión, tal como un potenciador, y sitios de procesamiento de información necesarios, tal como sitios de unión a ribosoma, sitios de corte y empalme de ARN, sitios de poliadenilación, y secuencias de terminación de la transcripción. El ácido nucleico puede codificar para más de un producto génico, limitado solo por el tamaño del ácido nucleico que puede ser empaquetado.

5 [215] En determinadas formas de realización de la presente divulgación, el ácido nucleico heterólogo puede codificar para proteínas beneficiosas que reemplazan proteínas perdidas o defectuosas requeridas por el sujeto en el cual se transfiere el vector, tal como Rhes o Rhes.

10 [216] Una partícula AAV1 es una partícula viral que comprende una proteína de la cápside de AAV1. En la presente están contempladas variaciones en la secuencia de aminoácidos de la proteína de cápside de AAV1, con la condición de que la partícula viral resultante que comprende la cápside de AAV1 permanezca antigénica o inmunológicamente diferente de las cápsides de otros AAV, como puede ser determinado de rutina mediante métodos estándares. Específicamente, por ejemplo, puede usarse ELISA y transferencias Western para determinar si una partícula viral es antigénica o inmunológicamente diferente de otros serotipos de AAV.

15 [217] El término "polipéptido" como se usa en la presente se refiere a un polímero de aminoácidos e incluye proteínas de longitud completa y fragmentos de las mismas. Por lo tanto, "proteína" y "polipéptido" con frecuencia se usan como sinónimos en la presente. Las sustituciones pueden seleccionarse de acuerdo a parámetros conocidos por ser neutros. Tal como será apreciado por aquellos con experiencia en el arte, la invención también incluye aquellos polipéptidos que tienen leves variaciones en las secuencias de aminoácidos u otras propiedades. Dichas variaciones pueden surgir naturalmente como variaciones alélicas (por ejemplo debido al polimorfismo genético) o pueden ser producidas por intervención del humano (por ejemplo, por mutagénesis de secuencias de ADN clonadas), tal como mutante puntual inducido, por eliminación, por inserción y por sustitución. En general se prefieren cambios menores en la secuencia de aminoácidos, tal como reemplazos conservativos de aminoácidos, eliminaciones o inserciones internas pequeñas, y agregados o eliminaciones en los extremos de las moléculas. Estas modificaciones pueden producir cambios en la secuencia de aminoácidos, proveer mutaciones silentes, modificar un sitio de restricción, o proveer otras mutaciones específicas.

20 [218] Se describe un método para administrar un ácido nucleico a una célula que comprende administrar a la célula una partícula de AAV que contiene un vector que comprende el ácido nucleico insertado entre un par de repeticiones terminales invertidas de AAV, proveyendo de esta manera el ácido nucleico a la célula. La administración a la célula puede realizarse por medio de mediante cualquier medio, que incluye simplemente poner en contacto la partícula, opcionalmente contenida en un líquido deseado tal como un medio de cultivo de tejido, o una solución salina amortiguada, con las células. Puede permitirse que la partícula permanezca en contacto con las células durante cualquier longitud de tiempo, y habitualmente la partícula se administra y se deja permanecer indefinidamente. Para dichos métodos *in vitro*, el virus puede administrarse a la célula mediante métodos de transducción viral estándares, como es conocido en el arte y se ejemplifica en la presente. Los títulos de virus a administrar pueden variar, particularmente dependiendo del tipo de célula, pero serán típicos de los usados para la transducción de AAV en general. Además, pueden utilizarse los títulos usados para transducir las células particulares en los presentes ejemplos. Las células pueden incluir cualquier célula deseada en humanos así como otros mamíferos grandes (no roedores), tales como primates, caballo, oveja, cabra, cerdo, y perro.

25 [219] Más específicamente, se describe un método para administrar un ácido nucleico a una célula en el cerebro, particularmente neuronas espinosas medianas, que comprenden el ácido nucleico insertado entre un par de repeticiones terminales invertidas de AAV, proveyendo de esta manera el ácido nucleico a la célula.

30 [220] Además, se describe un método para administrar un ácido nucleico a una célula en un sujeto que comprende la administración al sujeto de una partícula de AAV que comprende el ácido nucleico insertado entre un par de repeticiones terminales invertidas de AAV, proveyendo de esta manera el ácido nucleico a una célula en el sujeto.

35 [221] También se describe un método de administración de un ácido nucleico a una célula del cerebro, tal como una neurona en el cuerpo estriado o corteza en un sujeto que comprende administrar al sujeto una partícula AAV que comprende el ácido nucleico insertado entre un par de repeticiones terminales invertidas de AAV, administrando de esta manera el ácido nucleico a la neurona u otra célula en el sujeto.

40 [222] Determinadas formas de realización de la presente divulgación proveen una célula que comprende un vector viral como se describe en la presente.

45
60 Vectores AAV

50 [223] En una forma de realización, un vector viral de la divulgación es un vector de AAV. Un vector de "AAV" se refiere a un virus adenoasociado, y puede ser usado para referirse al virus de tipo salvaje natural mismo o derivados del mismo. El término cubre todos los subtipos, serotipos y seudotipos, y las formas de origen natural y recombinante, excepto que se requiera de otra manera. Como se usa en la presente, el término "serotipo" se refiere

a un AAV que es identificado por y distinguido de otros AAV en base a la reactividad de la proteína de la cápside con antisuero definido, por ejemplo, existen ocho serotipos conocidos de AAV de primate, AAV-1 a AAV-9 y AAVrh10. Por ejemplo, el serotipo AAV2 se usa para referirse a un AAV que contiene proteínas de la cápside codificadas por el gen cap de AAV2 y un genoma que contiene las secuencias ITR 5' y 3' del mismo serotipo de AAV2. Como se usa en la presente, por ejemplo, puede usarse rAAV1 para referirse a un AAV que tiene ambas proteínas de la cápside y las ITR 5'-3' del mismo serotipo o puede referirse a un AAV que tiene proteínas de la cápside de un serotipo y las ITR 5'-3' de un serotipo de AAV diferente, por ejemplo, la cápside del serotipo 2 de AAV y las ITR del serotipo 5 de AAV. Para cada ejemplo ilustrado en la presente la descripción del diseño y producción del vector describes el serotipo de la cápside y las secuencias ITR 5'-3'. La abreviatura "rAAV" se refiere a virus adenoasociado recombinante, también referido como un vector AAV recombinante (o "vector rAAV").

[224] Un "virus AAV" o "partícula viral de AAV" se refiere a una partícula viral compuesta por al menos una proteína de la cápside de AAV (preferiblemente por todas las proteínas de la cápside de un AAV de tipo salvaje) y un polinucleótido encapsidado. Si la partícula comprende el polinucleótido heterólogo (es decir, un polinucleótido diferente del genoma de AAV de tipo salvaje tal como un transgén a administrar a una célula de mamífero), habitualmente se lo refiere como "rAAV".

[225] En una forma de realización, los vectores AAV de expresión se construyen usando técnicas conocidas para al menos proveer como componentes operativamente ligados en la dirección de la transcripción, elementos de control que incluyen una región de inicio transcripcional, el ADN de interés y una región de terminación de la transcripción. Los elementos de control se seleccionan para que sean funcionales en una célula de mamífero. La construcción resultante que contiene los componentes operativamente ligados están flanqueados (5' y 3') con secuencias de ITR de AAV funcionales.

[226] Por "repeticiones terminales invertidas de virus adenoasociados" o "ITR de AAV" se refiere a las regiones reconocidas en el arte encontradas en cada extremo del genoma de AAV que funcionan juntas en cis como orígenes de la replicación del ADN y como señales de empaquetamiento para los virus. Las ITR de AAV, junto con la región codificante de rep de AAV, proveen la escisión eficaz y rescate de, e integración de una secuencia nucleotídica interpuesta entre dos ITR flanqueantes en un genoma de células de mamífero.

[227] Las secuencias nucleotídicas de las regiones ITR de AAV son conocidas. Como se usa en la presente, una "ITR de AAV" no requiere tener descrita la secuencia nucleotídica de tipo salvaje, pero puede estar alterada, por ejemplo, por la inserción, eliminación o sustitución de nucleótidos. Además, la ITR de AAV puede derivar de cualquiera de los diferentes serotipos de AAV, incluyendo sin limitación, AAV1, AAV2, AAV3, AAV4, AAV5, AAV7, etc. Además, las ITR 5' y 3' que flanquean una secuencia nucleotídica seleccionada en un vector de AAV no requieren ser necesariamente idénticas o derivar del mismo serotipo o aislado de AAV, con la condición de que funcionen como se pretende, es decir, para permitir la escisión y rescate de la secuencia de interés de un genoma de célula huésped o vector, y para permitir la integración de la secuencia heteróloga en el genoma de la célula receptora cuando están presente los productos génicos de Rep de AAV en la célula.

[228] En una forma de realización, las ITR de AAV pueden derivar de cualquiera de los diferentes serotipos de AAV, que incluyen sin limitación, AAV1, AAV2, AAV3, AAV4, AAV5, AAV7, etc. Además, las ITR 5' y 3' que flanquean una secuencia nucleotídica seleccionada en un vector de expresión de AAV no requieren ser necesariamente idénticas o derivadas del mismo serotipo o aislado de AAV, con la condición de que funcionen como se pretende, es decir, para permitir la escisión y rescate de la secuencia de interés de un genoma de célula huésped o vector, y para permitir la integración de la molécula de ADN en el genoma de la célula receptora cuando están presentes los productos génicos de Rep de AAV en la célula.

[229] En una forma de realización, las cápsides de AAV pueden derivar de AAV2. Las moléculas de ADN adecuadas para usar en los vectores AAV serán de menos de aproximadamente 5 kilobases (kb), menos de aproximadamente 4,5 kb, menos de aproximadamente 4kb, menos de aproximadamente 3,5 kb, menos de aproximadamente 3 kb, menos de aproximadamente 2,5 kb de tamaño y son conocidas en el arte.

[230] En una forma de realización, la secuencia nucleotídica seleccionada está ligada operativamente a elementos de control que dirigen la transcripción o expresión de la misma en el sujeto *in vivo*. Dichos elementos de control pueden comprender secuencias de control asociadas habitualmente con el gen seleccionado. Como alternativa, pueden emplearse secuencias de control heterólogas. Las secuencias de control heterólogas útiles en general incluyen aquellas derivadas de secuencias que codifican para genes de mamífero o virales. Los ejemplos incluyen, a título enunciativo no taxativo, el promotor temprano de SV40, el promotor LTR del virus del tumor mamario de ratón; el promotor tardío principal de adenovirus (Ad MLP); un promotor del virus de herpes simple (HSV), un promotor de citomegalovirus (CMV) tal como la región del promotor temprano inmediato de CMV (CMVIE), un promotor del virus de sarcoma de rous (RSV), promotores de pol II, promotores de pol III, promotores sintéticos, promotores híbridos, y similares. Además, las secuencias derivadas de genes no virales, tal como el gen de metalotioneina de murino, también será útil en la presente. Dichas secuencias promotoras se encuentran disponibles comercialmente de, por ejemplo, Stratagene (San Diego, Calif.).

[231] En una forma de realización, los promotores heterólogos y otros elementos de control, tal como los promotores específicos de SNC e inducibles, potenciadores y similares, serán de uso particular. Los ejemplos de promotores heterólogos incluyen el promotor de CMV. Los ejemplos de promotores específicos de SNC incluyen aquellos aislados de los genes de la proteína mielina básica (MBP), proteína ácida fibrilar glial (GFAP), y enolasa específica de neurona (NSE). Los ejemplos de promotores inducibles incluyen elementos de respuesta del ADN para ecdisona, tetraciclina, hipoxia y aulina.

[232] En una forma de realización, el vector de expresión de AAV que porta la molécula de ADN de interés unida por las ITR de AAV, puede ser construido por inserción directa de la o las secuencias seleccionadas en un genoma de AAV que se le han escindido los marcos de lectura abiertos principales de AAV ("ORF"). Otras partes del genoma del AAV también pueden ser eliminadas, con la condición de que permanezca una parte suficiente de las ITR para permitir las funciones de replicación y empaquetamiento. Dichas construcciones pueden ser diseñadas usando técnicas bien conocidas en el arte.

[233] Como alternativa, las ITR de AAV pueden ser escindidas del genoma viral o de un vector de AAV que contiene el mismo 5' y 3' fusionados de una construcción de ácido nucleico seleccionado que está presente en otro vector usando técnicas de ligado estándar. Por ejemplo, los ligados pueden hacerse en Tris-Cl 20 mM pH 7,5, MgCl₂ 10 mM, DTT 10 mM, BSA 33 mg/ml, NaCl entre 10 mM y 50 mM, y ATP 40 μM, entre 0,01 y 0,02 (Weiss) unidades de T4 ADN ligasa a 0°C (para ligado de "extremos cohesivos") o ATP 1 mM, entre 0,3 y 0,6 (Weiss) unidades de T4 ADN ligasa a 14°C (para ligado de "extremos romos"). Los ligados intermoleculares "de extremos cohesivos" se realizan habitualmente a entre 30 y 100 mg/ml de concentraciones de ADN total (entre 5 y 100 nM de concentración final total). Los vectores de AAV que contienen las ITR.

[234] Además, los genes quiméricos pueden ser producidos de manera sintética para que incluyan las secuencias de ITR de AAV dispuestas 5' y 3' de una o más secuencias de ácido nucleico seleccionadas. Pueden usarse codones preferidos para la expresión de la secuencia del gen quimérico en células de SNC de mamífero. La secuencia quimérica completa es ensamblada a partir de oligonucleótidos solapantes preparados por métodos estándares.

[235] Con el objetivo de producir viriones de rAAV, se introduce un vector de expresión de AAV en una célula huésped adecuada usando técnicas conocidas, tal como por transfección. En general en el arte se conocen diferentes técnicas de transfección. Véase, por ejemplo, Sambrook y col. (1989) Molecular Cloning, a laboratory manual, Cold Spring Harbor Laboratories, New York. Los métodos de transfección particularmente adecuados incluyen coprecipitación con fosfato de calcio, microinyección directa en células cultivadas, electroporación, transferencia génica mediada por liposomas, transducción mediada por lípidos, y administración de ácido nucleico usando microproyectiles de alta velocidad.

[236] En una forma de realización, las células huéspedes adecuadas para la producción de viriones de rAAV incluyen microorganismos, células de levaduras, células de insectos, y células de mamíferos, que pueden ser, o han sido, usadas como receptoras de una molécula de ADN heteróloga. El término incluye la progenie de la célula original que ha sido transfectada. Por lo tanto, una "célula huésped" como se usa en la presente en general se refiere a una célula que ha sido transfectada con una secuencia de ADN exógeno. Las células de la línea celular de humano estable, 293 (disponible fácilmente de, por ejemplo, la Colección Americana de Cultivos Tipo bajo el número de acceso ATCC CRL1573) pueden usarse en la práctica de la presente divulgación. Particularmente, la línea de células de humano 293 es una línea de células de riñón embrionario de humano que ha sido transformada con fragmentos de ADN de adenovirus tipo-5, y expresa los genes adenovirales E1a y E1b. La línea de células 293 es fácilmente transfectada, y provee una plataforma particularmente conveniente en la cual producir los viriones de rAAV.

[237] Por "región codificante de rep de AAV" se refiere a la región reconocida en el arte del genoma de AAV que codifica para las proteínas de replicación Rep 78, Rep 68, Rep 52 y Rep 40. Se ha mostrado que estos productos de expresión de Rep poseen muchas funciones, que incluyen reconocimiento, unión y corte monocatenario del origen de replicación de ADN de AAV, actividad ADN helicasa y modulación de la transcripción desde los promotores de AAV (u otro heterólogo). Los productos de expresión de Rep son colectivamente requeridos para la replicación del genoma de AAV. Los homólogos adecuados de la región codificante de rep de AAV incluyen el gen rep del herpesvirus 6 de humano (HHV-6) que también es conocido por mediar la replicación del ADN de AAV-2.

[238] Por "región codificante de cap de AAV" se refiere a la región reconocida en el arte del genoma de AAV que codifica para las proteínas de la cápside VP1, VP2 y VP3, u homólogos funcionales de las mismas. Estos productos de expresión de Cap proveen las funciones de empaquetamiento que se requieren colectivamente para el empaquetado del genoma viral.

[239] En una forma de realización, se introducen las funciones colaboradoras de AAV en la célula huésped mediante la transfección de la célula huésped con una construcción colaboradora del AAV antes de, o de manera concurrente con, la transfección del vector de expresión de AAV. Las construcciones colaboradoras del AAV se usan por lo tanto para proveer al menos una expresión transitoria de los genes rep y/o cap de AAV para complementar las

funciones perdidas de AAV que son necesarias para la infección productiva del AAV. Las construcciones colaboradoras de AAV carecen de las ITR de AAV y no pueden replicarse ni empaquetarse ellas mismas. Estas construcciones pueden estar en la forma de un plásmido, fago, transposón, cósmido, virus, o virión. Se han descrito varias construcciones colaboradoras de AAV, tales como los plásmidos usados habitualmente pAAV/Ad y pIM29+45 que codifican para los productos de expresión Rep y Cap. Se han descrito varios vectores adicionales que codifican para los productos de expresión de Rep y/o Cap.

[240] Los métodos de administración de vectores virales incluyen inyección del AAV en el sujeto. En general, los viriones de rAAV pueden introducirse en las células del SNC usando técnicas de transducción *in vivo* o *in vitro*. Si se transducen *in vitro*, la célula receptora deseada será extraída del sujeto, transducida con los viriones de rAAV y reintroducida en el sujeto. Como alternativa, pueden usarse células singénicas o xenogénicas cuando aquellas células no van a generar una respuesta inmunológica no apropiada en el sujeto.

[241] Se han descrito métodos adecuados para la administración e introducción de las células transducidas en un sujeto. Por ejemplo, las células pueden transducirse *in vitro* por combinación de viriones de AAV recombinantes con células del SNC por ejemplo, en medio apropiado, y el cribado de aquellas células que portan el ADN de interés puede realizarse usando técnicas convencionales tales como transferencias Southern y/o PCR, o mediante el uso de marcadores de selección. Las células transducidas pueden formularse luego en composiciones farmacéuticas, descrito con mayor detalle más adelante, e introducir la composición en el sujeto mediante diferentes técnicas, tal como por injerto, inyección intramuscular, intravenosa, subcutánea e intraperitoneal.

[242] En una forma de realización, las composiciones farmacéuticas comprenderán suficiente material genético para producir una cantidad terapéuticamente eficaz del ácido nucleico de interés, es decir, una cantidad suficiente para reducir o aliviar los síntomas del estado de enfermedad en cuestión o una cantidad suficiente para conferir el beneficio deseado. Las composiciones farmacéuticas también contendrán un excipiente farmacéuticamente aceptable. Dichos excipientes incluyen cualquier agente farmacéutico que no induce el mismo la producción de anticuerpos dañinos para el individuo que recibe la composición, y que pueden ser administrados sin toxicidad indebida. Los excipientes farmacéuticamente aceptables incluyen, a título enunciativo no taxativo, sorbitol, Tween80, y líquidos tales como agua, solución salina, glicerol y etanol. Las sales farmacéuticamente aceptables pueden estar incluidas en los mismos, por ejemplo, sales de ácido mineral tales como clorhidratos, bromhidratos, fosfatos, sulfatos, y similares; y las sales de ácidos orgánicos tales como acetatos, propionatos, malonatos, benzoatos, y similares. Además, en dichos vehículos pueden estar presente sustancias auxiliares, tales como agentes humectantes o emulsificantes, sustancias amortiguadoras del pH, y similares. En Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991) se encuentra disponible una descripción extensa de los excipientes farmacéuticamente aceptables.

[243] Se debe comprender que el vector viral administrado podría expresar más de un transgén. Como alternativa, también pueden administrarse vectores por separado, en donde cada uno expresa uno o más transgenes diferentes, al sujeto como se describe en la presente. Además, también se pretende que los vectores virales administrados por los métodos de la presente divulgación se combinen con otras composiciones y terapias adecuadas.

[244] Tal como resulta evidente para aquellos con experiencia en el arte en vista de las divulgaciones de esta memoria descriptiva, puede determinarse empíricamente una cantidad eficaz de vector viral que debe agregarse. La administración puede realizarse en una dosis, de manera continua o intermitente durante el curso del tratamiento. Los métodos para determinar el medio y dosificaciones de administración más eficaces son bien conocidos para aquellos con experiencia en el arte y variará con el vector viral, la composición de la terapia, las células diana, y el sujeto a tratar. Pueden realizarse administraciones individuales o múltiples en donde el nivel de la dosis y el patrón son seleccionados por el médico.

[245] En determinadas formas de realización, el rAAV se administra en una dosis de aproximadamente entre 0,3 y 2 ml de entre 1×10^5 y 1×10^{16} vg/ml. En determinadas formas de realización, el rAAV se administra en una dosis de aproximadamente entre 1 y 3 ml de entre 1×10^7 y 1×10^{14} vg/ml. En determinadas formas de realización, el rAAV se administra en una dosis de aproximadamente entre 1 y 2 ml de entre 1×10^8 y 1×10^{13} vg/ml.

[246] Las formulaciones que contienen las partículas de rAAV contendrán una cantidad eficaz de las partículas de rAAV en un vehículo, en donde la cantidad eficaz será fácilmente determinada por una persona con experiencia en el arte. Las partículas de rAAV habitualmente pueden variar en un rango entre aproximadamente 1 % y aproximadamente 95% (peso en peso) de la composición, o incluso mayor o menor si es apropiado. La cantidad a administrar depende de factores tales como la edad, peso y condición física del animal o el sujeto humano considerado para el tratamiento. Las dosificaciones eficaces pueden ser establecidas por una persona con experiencia ordinaria en el arte mediante ensayos de rutina estableciendo curvas de respuesta a dosis. El sujeto es tratado mediante la administración de las partículas de rAAV en una o más dosis. Las dosis múltiples pueden administrarse según sea requerido para mantener la actividad adecuada de la enzima.

[247] Con la invención pueden emplearse vehículos que incluyen agua, solución salina, LCR artificial, u otras

sustancias conocidas. Para preparar una formulación, la composición purificada puede aislarse, liofilizarse y estabilizarse. La composición puede luego ajustarse a una concentración apropiada, combinarse opcionalmente con un agente antiinflamatorio, y envasarse para su uso.

5 [248] Se describe un método para aumentar el nivel de una proteína diana en una célula mediante la introducción de una proteína, o molécula de ácido nucleico que codifica para una proteína descrita precedentemente en una célula en una cantidad suficiente para aumentar el nivel de la proteína diana en la célula. En determinadas formas de realización, la acumulación de la proteína diana aumenta en por lo menos 10%. La acumulación de la proteína diana aumenta en por lo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 95%, o 99%.

10 [249] Además, el vector AAV puede seleccionarse/diseñarse de acuerdo con la ruta deseada de administración, por ejemplo, y sin limitación, para la administración sistémica, puede usarse un vector AAV con capacidad de cruzar la barrera hematoencefálica (por ejemplo, AAV9, o un vector de AAV quimérico que tiene proteínas de la cápside de AAV9). También se describe un método para administrar el AAV al torrente sanguíneo ya que algunos serotipos tienen la capacidad de atravesar la barrera hematoencefálica.

15 Péptidos de direccionamiento

[250] Se han identificado péptidos que funcionan para dirigir agentes, tales como vectores virales, a células del endotelio vascular del sistema nervioso central. La presente divulgación describe un método para utilizar estos péptidos novedosos, por ejemplo, cápsides virales al tipo de célula de interés. En este caso, las células endoteliales que recubren los vasos sanguíneos del cerebro son diana de los péptidos identificados. Pueden usarse vectores que portan las proteínas de la cápside modificadas para que incluyan dichos péptidos para proveer agentes terapéuticos al sistema nervioso central (por ejemplo, el cerebro).

25 [251] Como se usa en la presente, el término "dirige" se refiere a que la proteína de la cápside de un virus, tal como un virus adenoasociado (AAV), preferiblemente se une a un tipo de tejido (por ejemplo, tejido cerebral) respecto a otro tipo de tejido (por ejemplo, tejido hepático), y/o se une a un tejido en determinado estado (por ejemplo, de tipo salvaje o enfermedad). En determinadas formas de realización, la proteína modificada genéticamente de la cápside puede "dirigir" al tejido del epitelio vascular del cerebro por unión a un nivel de entre 10% y 1000% mayor que una proteína de cápside comparable, sin modificar. Por ejemplo, un AAV que tiene una proteína genéticamente modificada de la cápside puede unirse al tejido del epitelio vascular del cerebro en un nivel de entre 50% y 100% mayor que un virus AAV sin modificar. En determinadas formas de realización, los ácidos nucleicos que codifican para las proteínas de la cápside de un virus son modificados de manera que las cápsides virales preferiblemente se unan a endotelio vascular del cerebro en un mamífero que padece de enfermedad de almacenamiento lisosomal, o, usando diferentes secuencias, al endotelio vascular del cerebro de tipo salvaje en el cerebro de la misma especie.

40 [252] Se describe una proteína modificada de la cápside de virus adenoasociado (AAV) que contiene un péptido de direccionamiento, en donde el péptido de direccionamiento tiene entre 3 y 10 aminoácidos de longitud y en donde el péptido de direccionamiento dirige un AAV al endotelio vascular de cerebro. En determinadas formas de realización, el péptido de direccionamiento tiene 3, 4, 5, 6 o 7 aminoácidos de longitud. En determinadas formas de realización, el AAV es AAV2, a pesar de que el tropismo es modificado de manera que podría ser que dichas modificaciones podrían cambiar el tropismo de cualquier AAV.

45 [253] Determinadas formas de realización de la presente divulgación proveen un vector viral que comprende una cápside modificada, en donde la cápside modificada comprende por lo menos una secuencia de aminoácidos que dirige el vector viral al endotelio vascular de cerebro.

50 [254] En determinadas formas de realización, el vector viral es un vector viral adenoasociado (AAV). En determinadas formas de realización, el AAV es AAV2.

[255] En determinadas formas de realización, el péptido de direccionamiento dirige al endotelio vascular de cerebro de tipo salvaje. En determinadas formas de realización, el péptido de direccionamiento es PXXPS (SEQ ID NO:44), SPXXP (SEQ ID NO:45), TLH (SEQ ID NO:46), o QSXY (SEQ ID NO:47), según se exprese en una orientación amino a carboxilo o en una orientación carboxilo a amino. En determinadas formas de realización, el péptido de direccionamiento es PYFPSLS (SEQ ID NO:48), YAPLTPS (SEQ ID NO:49), PLSPSAY (SEQ ID NO:50), DSPAHP (SEQ ID NO:51), GTPHPS (SEQ ID NO:52), PDAPSNH (SEQ ID NO:53), TEPHWPS (SEQ ID NO:54), SPPLPPK (SEQ ID NO:55), SPKPPPG (SEQ ID NO:56), NWSPWDP (SEQ ID NO:57), DSPAHP (SEQ ID NO:58), GWTLHMK (SEQ ID NO:59), KIPPTLH (SEQ ID NO:60), ISQTLHG (SEQ ID NO:61), QSFYILT (SEQ ID NO:62), o TTQSEYG (SEQ ID NO:63), según se exprese en una orientación amino a carboxilo o en una orientación carboxilo a amino. Se debe hacer notar que la orientación de la secuencia no es importante. Por ejemplo, el péptido puede estar orientado desde el extremo amino-terminal al carboxilo-terminal del péptido para que sea TTQSEYG (SEQ ID NO:63) o puede estar desde el extremo amino-terminal al carboxilo-terminal del péptido para que sea GYESQTT (SEQ ID NO:65).

[256] En determinadas formas de realización, el péptido de direccionamiento dirige a un endotelio vascular de cerebro enfermo. En determinadas formas de realización, el péptido de direccionamiento dirige a endotelio vascular de cerebro en un sujeto que tiene una enfermedad de almacenamiento lisosomal. En determinadas formas de realización, el péptido de direccionamiento dirige a mucopolisacárido (MPS) VII de endotelio vascular de cerebro. En determinadas formas de realización, el péptido de direccionamiento es LXSS (SEQ ID NO:66), PFXG (SEQ ID NO:67), o SIXA (SEQ ID NO:68), según se exprese en una orientación amino a carboxilo o en una orientación carboxilo a amino. En determinadas formas de realización, el péptido de direccionamiento es MLVSSPA (SEQ ID NO:69), LPSSLQK (SEQ ID NO:70), PPLLKSS (SEQ ID NO:71), PXKLDSS (SEQ ID NO:72), AWTLASS (SEQ ID NO:73), WPFYGTG (SEQ ID NO:74), GTFPFLG (SEQ ID NO:75), GQVPFMG (SEQ ID NO:76), ANFSILA (SEQ ID NO:77), GSIWAPA (SEQ ID NO:78), o SIAASFS (SEQ ID NO:79), según se exprese en una orientación amino a carboxilo o en una orientación carboxilo a amino.

[257] En determinadas formas de realización, el péptido de direccionamiento dirige a TPP1 de endotelio vascular de cerebro. En determinadas formas de realización, el péptido de direccionamiento es GMNAFRA (SEQ ID NO:64), según se exprese en una orientación amino a carboxilo o en una orientación carboxilo a amino.

[258] En determinadas formas de realización, la secuencia de aminoácidos que está dirigida al endotelio vascular del cerebro comprende por lo menos una de las SEQ ID NOs 44-47.

[259] En determinadas formas de realización, la secuencia de aminoácidos que está dirigida al endotelio vascular del cerebro comprende por lo menos una de las SEQ ID NOs 66-68.

[260] En determinadas formas de realización, la secuencia de aminoácidos que está dirigida al endotelio vascular del cerebro comprende por lo menos una de las SEQ ID NOs 48-63.

[261] En determinadas formas de realización, la secuencia de aminoácidos que está dirigida al endotelio vascular del cerebro comprende por lo menos una de las SEQ ID NOs 69-79.

[262] En determinadas formas de realización, la secuencia de aminoácidos que dirige a tejido de cerebro se selecciona de aquellas que se listan en la Tabla 1 más adelante:

Tabla 1: PM-AAV dirigido a cerebro

Nombre	Secuencia	SEQ ID NO.	Direccionamiento
THR	THRPPMWSPVWP	80	Transferrina
CRT	CRTIGPSVC	81	Transferrina
BX2	GHKVKRPKG	82	Transferrina
BX3	KDKIKMDKK	83	Transferrina
BX6	GHKAKGPRK	84	Transferrina
BX8	KWKTPKVRV	85	Transferrina
AAV-PPS	DSPAHP	51	De tipo salvaje
	PYFPSLS	48	De tipo salvaje
	YAPLTPS	49	De tipo salvaje
	PLSPSAY	50	De tipo salvaje
	GTPTHPS	52	De tipo salvaje
	PDAPSNH	53	De tipo salvaje
	TEPHWPS	54	De tipo salvaje
	SPPLPPK	55	De tipo salvaje
	SPKPPPG	56	De tipo salvaje
	NWSPWDP	57	De tipo salvaje
AAV-TLH	GWTLHNK	59	De tipo salvaje
	KIPPTLH	60	De tipo salvaje
	ISQTLHG	61	De tipo salvaje
	QSFYILT	62	De tipo salvaje
	TTQSEYG	63	De tipo salvaje
AAV-PFG	WPFYGTG	74	MPS VII
	GTFPFLG	75	MPS VII
	GQVPFMG	76	MPS VII
	PPLLKSS	71	MPS VII
	MLVSSPA	69	MPS VII
	AWTLASS	73	MPS VII
AAV-LSS	LPSSLQK	70	MPS VII
	PXKLDSS	72	MPS VII
	GSIWAPA	78	MPS VII
	ANFSILA	77	MPS VII

	SIAASFS	79	MPS VII
AAV-GMN	GMNAFRA	64	CLN2

- 5 [263] En determinadas formas de realización, la secuencia de aminoácidos que está dirigida al endotelio vascular del cerebro dirige a endotelio vascular de cerebro en un sujeto que tiene una enfermedad, por ejemplo, una enfermedad de almacenamiento lisosomal.
- [264] En determinadas formas de realización, la secuencia de aminoácidos que está dirigida al endotelio vascular del cerebro dirige a endotelio vascular de cerebro en un sujeto que no tiene una enfermedad de almacenamiento lisosomal.
- 10 [265] En determinadas formas de realización, el vector viral comprende una secuencia de ácido nucleico que codifica para un agente terapéutico. En determinadas formas de realización, el agente terapéutico es p-glucuronidasa.
- 15 [266] En determinadas formas de realización, la secuencia de aminoácidos que está dirigida al endotelio vascular del cerebro tiene como máximo diez aminoácidos de longitud.
- [267] En determinadas formas de realización, la secuencia de aminoácidos que está dirigida al endotelio vascular del cerebro tiene 3, 4, 5, 6 o 7 aminoácidos de longitud.
- 20 [268] Determinadas formas de realización de la presente divulgación proveen una secuencia de ácido nucleico que codifica para un vector viral como se describe en la presente.
- [269] Determinadas formas de realización de la presente divulgación proveen una secuencia de ácido nucleico que codifica para una cápside modificada como se describe en la presente. Determinadas formas de realización de la presente divulgación proveen una cápside modificada codificada por una secuencia de ácido nucleico descrita en la presente.
- 25 [270] Determinadas formas de realización de la presente divulgación proveen una célula que comprende un vector viral como se describe en la presente.
- 30 [271] Determinadas formas de realización de la presente divulgación proveen una célula transducida por un vector viral como se describe en la presente.
- 35 [272] En determinadas formas de realización, la célula es una célula de mamífero. En determinadas formas de realización, la célula es una célula de humano. En determinadas formas de realización, la célula es una célula no humana. En determinadas formas de realización, la célula está *in vitro*. En determinadas formas de realización, la célula está *in vivo*. En determinadas formas de realización, la célula es una célula endotelial. En determinadas formas de realización, la célula es una célula de endotelio vascular.
- 40 **Dosificaciones, formulaciones y rutas de administración de los agentes descritos**
- [273] Los agentes descritos preferiblemente se administran de manera de producir una reducción en por lo menos un síntoma asociado con una enfermedad. La cantidad administrada variará dependiendo de diferentes factores que incluyen, a título enunciativo no taxativo, la composición elegida, la enfermedad particular, el peso, la condición física, y la edad del mamífero, y si se quiere lograr prevención o tratamiento. Dichos factores pueden ser determinados por el médico que emplea modelos animales u otros sistemas de ensayo, que son bien conocidos en el arte. Como se usa en la presente, el término "ARNip terapéutico" se refiere a cualquier ARNip que tiene un efecto beneficioso en el receptor. Por lo tanto, "ARNip terapéutico" abarca al ARNip terapéutico y profiláctico.
- 45 [274] La administración de ARNip puede ser realizada mediante la administración de la molécula de ácido nucleico que codifica para el ARNip. Las formulaciones farmacéuticas, dosificaciones y rutas de administración para los ácidos nucleicos son en general conocidas.
- 50 [275] La presente divulgación prevé el tratamiento de la enfermedad de Huntington en un mamífero mediante la administración de un agente, por ejemplo, una composición de ácido nucleico, un vector de expresión, o una partícula viral de la divulgación. La administración de los agentes terapéuticos de acuerdo con la presente divulgación puede ser continua o intermitente, dependiendo, por ejemplo, de la condición fisiológica del receptor, de si el propósito de la administración es terapéutico o profiláctico, y otros factores conocidos para los médicos con experiencia. La administración de los agentes de la divulgación puede ser esencialmente continua durante un periodo preseleccionado de tiempo o puede ser en una serie de dosis espaciadas. Se contempla la administración local y sistémica.
- 55 [276] Puede administrarse una o más formas de dosificación unitaria adecuadas que tienen el o los agentes terapéuticos de la invención, las cuales, como se describe más adelante, pueden formularse opcionalmente para la
- 60

liberación sostenida (por ejemplo usando microencapsulación, véase WO 94/07529, and y la Patente de los EE.UU. No. 4.962.091), mediante una variedad de rutas que incluyen parenteral, incluyendo las rutas intravenosa e intramuscular, así como por inyección directa en el tejido enfermo. Por ejemplo, el agente terapéutico puede ser directamente inyectado en el cerebro. Como alternativa el agente terapéutico puede ser introducido por vía intratecal para condiciones del cerebro y médula espinal. En otro ejemplo, el agente terapéutico puede ser introducido por vía intramuscular para virus que trafican hacia atrás en neuronas afectadas desde el músculo, tales como AAV, lentivirus y adenovirus. Las formulaciones pueden, cuando es apropiado, ser presentadas convenientemente en formas de dosificación unitaria discretas y pueden prepararse por cualquiera de los métodos bien conocidos para un fármaco. Dichos métodos pueden incluir el paso de asociar el agente terapéutico con vehículos líquidos, matrices sólidas, vehículos semisólidos, vehículos sólidos finamente divididos o combinaciones de los mismos, y entonces, si es necesario, introducir o darle forma al producto en el sistema de administración deseado.

[277] Cuando los agentes terapéuticos descritos se preparan para la administración, preferiblemente se combinan con un vehículo, diluyente o excipiente farmacéuticamente aceptable para formar una formulación farmacéutica, o forma de dosificación unitaria. Los ingredientes activos totales en dichas formulaciones incluyen entre 0,1 y 99,9% en peso de la formulación. Un "farmacéuticamente aceptable" es un vehículo, diluyente, excipiente, y/o sal que es compatible con los otros ingredientes de la formulación, y no es deletéreo para el receptor del mismo. El ingrediente activo para la administración puede estar presente como un polvo o como gránulos, como una solución, una suspensión o una emulsión.

[278] Las formulaciones farmacéuticas que contiene los agentes terapéuticos descritos pueden prepararse mediante procedimientos conocidos en el arte usando ingredientes bien conocidos y fácilmente disponibles. Los agentes terapéuticos descritos también pueden ser formulados como soluciones apropiadas para la administración parenteral, por ejemplo a través de las rutas intramuscular, subcutánea o intravenosa.

[279] Las formulaciones farmacéuticas de los agentes terapéuticos descritos también pueden tomar la forma de una solución o dispersión acuosa o anhidra, o como alternativa la forma de una emulsión o suspensión.

[280] Por lo tanto, el agente terapéutico puede ser formulado para la administración parenteral (por ejemplo, para inyección, por ejemplo, inyección en bolo o infusión continua) y puede presentar en forma de dosis unitaria en ampollas, jeringas precargadas, recipientes de infusión de volumen pequeño o en recipientes de dosis múltiples con un conservante agregado. Los ingredientes activos pueden tomar dichas formas como suspensiones, soluciones, o emulsiones en vehículos oleosos o acuosos, y pueden contener agentes de formulación tales como agentes de suspensión, estabilizantes y/o de dispersión. Como alternativa, los ingredientes activos pueden estar en forma de polvo, obtenido por aislamiento aséptico del sólido estéril o por liofilización a partir de una solución, para la constitución con un vehículo adecuado, por ejemplo, agua estéril, libre de pirógenos, antes del uso.

[281] Se apreciará que el contenido de la unidad de ingrediente o ingredientes activos contenidos en una dosis de aerosol individual de cada forma de dosificación no requiere constituir en sí mismo una cantidad eficaz para tratar la indicación o enfermedad particular ya que la cantidad eficaz necesaria puede alcanzarse mediante la administración de una pluralidad de unidades de dosificación. Además, la cantidad eficaz puede alcanzarse usando menos que la dosis en la forma de dosificación, en forma individual, o en una serie de administraciones.

[282] Las formulaciones farmacéuticas de la presente divulgación pueden incluir, como ingredientes opcionales, vehículos farmacéuticamente aceptables, diluyentes, agentes solubilizantes o emulsificantes, y sales del tipo de las que son bien conocidas en el arte. Los ejemplos no limitantes específicos de los vehículos y/o diluyentes que son útiles en las formulaciones farmacéuticas de la presente divulgación incluyen agua y soluciones salinas amortiguadas fisiológicamente aceptables tales como soluciones salinas amortiguadas de fosfato con pH entre 7,0 y 8,0 y agua.

[283] La invención se ilustrará a continuación mediante el siguiente ejemplo no limitante.

EJEMPLO 1

La modificación de semilla de nucleótido único restaura la tolerabilidad *in vivo* de una secuencia de miARN tóxica en el cerebro de ratón

[284] La enfermedad de Huntington (HD) es una enfermedad neurodegenerativa fatal provocada por la expresión de una forma expandida de poliglutamina de huntingtina (HTT). Un trabajo reciente mostró que las estrategias de silenciamiento génico, que incluye interferencia de ARN (iARN), mejora los indicadores de la enfermedad en modelos de ratones con HD. Para avanzar en la iARN dirigida contra la HTT a la clínica los inventores diseñaron una construcción de ARNi, HDS1 con eficacia robusta de silenciamiento de diana y con silenciamiento minimizado de transcritos humano no pretendidos (McBride y col., Mol Ther. Dec 2011; 19(12): 2152-2162). En macacos Rhesus, la administración de AAV.miHDS1 en el putamen redujo la expresión de la HTT sin efectos adversos sobre el estado neurológico incluyendo las habilidades motoras finas y gruesas, sin activación inmunológica, y sin neuropatología después de 6 semanas luego de la inyección. Otros mostraron seguridad de una iARN dirigida contra la HTT

diferente en monos durante 6 meses después de la inyección.

[285] La aplicación de HDS1 a pacientes con HD requiere de ensayos de seguridad adicionales en roedores, a pesar del hecho de que fue optimizado para humanos. Para satisfacer este requerimiento de regulación, los inventores evaluaron ratones después de la inyección de AAV.miHDS1. A diferencia de los monos, se produjeron déficits neurológicos de manera aguda en el cerebro de ratones y podría atribuirse al silenciamiento no diana mediado por interacciones de miHDS1 con la región no traducida 3' de otros transcritos. Si bien los inventores resolvieron la toxicidad de miHDS1 en el cerebro de ratón y mantuvieron una eficacia de silenciamiento con miHDS1, estos estudios destacan que la optimización de las medicinas basadas en ácidos nucleicos para la seguridad del uso en humanos presenta desafíos para la evaluación de la seguridad en roedores u otras especies relacionadas de manera distante.

[286] La HD es producida por la expansión de repeticiones CAG (>36 repeticiones) en el primer exón de huntingtina. A pesar de que la huntingtina mutante (mHTT) se expresa de manera ubicua, el cerebro, y en particular el cuerpo estriado, muestra una degeneración robusta y más temprana. La incidencia de HD es entre 5 y 10 cada 100.000 individuos en Europa y en los EE.UU., mientras que el inicio en general se produce en la 3^{ra} o 4^{ta} década de edad. Hasta la fecha, el manejo de HD incluye fármacos que pueden reducir los síntomas motores y psiquiátricos.

[287] Un trabajo previo con el uso de modelos inducibles de HD mostró que los síntomas de la enfermedad mejoran una vez que la expresión de mHTT fue apagada, incluso muchas semanas luego del inicio de la enfermedad y después de la atrofia del cuerpo estriado. Esto infiere que existe una ventana de oportunidad para tratar la HD después de la aparición temprana de los síntomas. Por lo tanto, los métodos para reducir la expresión génica usando tecnologías de silenciamiento génico, que incluyen iARN, deberían ser investigadas como una alternativa terapéutica.

[288] La iARN es un proceso evolutivamente conservado de silenciamiento génico postranscripcional mediante el cual los ARN no codificantes pequeños de hebra doble (por ejemplo, miARN) producen una degradación específica de secuencia de las secuencias de ARNm diana. La vía de iARN endógena comienza con la expresión de un transcrito primario de ARN más grande (pri-miARN) que es secuencialmente escindido en el núcleo por Drosha, un componente del complejo microprocesador, para generar un miARN precursor (pre-miARN). Los pre-miARN son exportados al citoplasma y escindidos posteriormente por Dicer para liberar el miARN maduro. De las dos hebras de la secuencia de miARN, una (la hebra antisentido "guía") en general se incorpora preferiblemente al Complejo de Silenciamiento Inducido por ARN (RISC), en donde dirigirá la unión al ARNm diana e inhibirá la expresión. Los miARN habitualmente reprimen la expresión de ARNm mediante complementariedad parcial. Cuando una hebra del ARNhd que emerge de la escisión por Dicer es completamente complementario a su diana, el ARN de interferencia pequeño resultante (ARNip) dirige la escisión endonucleotídica de la diana en una base entre los nucleótidos 10 y 11 de la hebra "guía", iniciando la destrucción del ARNm. Los científicos han desarrollado diferentes sistemas basados en expresión para cooptar la vía endógena de iARN y suprimir la expresión de genes específicos. Por ejemplo, los sistemas de expresión de iARN pueden diseñarse para que expresen secuencias de ARN en horquilla pequeños con una de sus hebras complementarias al ARNm diana, y que entre en la vía en los pasos de pre-miARN (ARN en horquilla pequeño; ARNhp) o pri-miARN (miARN artificial).

[289] Para los sistemas de expresión o ARNip que son transfectados de manera aguda en las células, la hebra guía activa se diseña para que sea tan específica como sea posible con un silenciamiento fuera de secuencia mínimo. El silenciamiento fuera de secuencia surge de la interacción de la guía con otros transcritos con complementariedad completa puede evitarse usando algoritmos de búsqueda estándares. Un tipo más difícil de fuera de diana a evitar es el que se produce debido a la complementariedad parcial de la secuencia semilla de iARN, bases 2 a 7 en el extremo 5' de la hebra cargada, con otras secuencias 3'UTR del ARNm. En este caso, la represión de la expresión se produce por medio de un mecanismo similar a miARN. En estudios previos, los inventores desarrollaron un algoritmo, siSPOTR, para diseñar secuencias de iARN potentes con fuerte desviación de hebra para la carga de RISC, y potencial de silenciamiento no diana minimizado respecto a los transcritos de humano no deseados. Cuando se usa siSPOTR para diseñar iniciadores del silenciamiento de HTT, los inventores encontraron que miHDS1, expresado a partir de AAV, mostró seguridad a múltiples niveles luego de la administración en el putamen de primates no humanos.

[290] Como un prerrequisito para la aplicación en humanos, los inventores realizaron experimentos de seguimiento para evaluar la seguridad en roedores normales. Notablemente, los inventores encontraron que HDS1 indujo déficits motores robustos después de las inyecciones en el cuerpo estriado, que podrían atribuirse al silenciamiento no deseado de *Bcl2*. Además, los inventores muestran que la toxicidad fuera de diana podría resolverse mediante diferentes estrategias a la vez que se mantiene una eficacia de silenciamiento de HTT. En resumen, estos estudios destacan el desafío de optimizar las medicinas basadas en ácidos nucleicos en cuanto a la especificidad y seguridad en humanos que cuando se usan en especies relacionadas de manera distante se reflejarán perfiles fuera de diana diferentes, y quizás inductores de enfermedad.

65 RESULTADOS

miHDS1 induce déficits neurológicos en el cerebro de ratón

[291] En trabajos previos los inventores diseñaron miHDS1, una secuencia de miARN artificial dirigida contra huntingtina con eficacia alta de silenciamiento de diana y potencial fuera de diana minimizada (Figura 1A). Cuando se inyectaron los vectores de AAV que expresan miHDS1 en el putamen de primates no humanos, los niveles de HTT se redujeron significativamente y no hubieron signos de degeneración neuronal, respuestas inmunológicas o déficits motores. En resumen, estos estudios destacan el potencial de miHDS1 para la terapia contra la HD. Sin embargo, como un prerrequisito para la aplicación en humanos, se requieren ensayos adicionales en otras especies, tales como en roedores. Por lo tanto, los inventores se propusieron realizar ensayos de seguridad de AAV.miHDS1 en ratones normales, a pesar del hecho que fue diseñado para seguridad en células de humano.

[292] Como un primer paso en la elaboración de la construcción preclínica, los inventores rediseñaron el vector AAV.miHDS1 para que contenga una secuencia stuffer en lugar del cassette de expresión de eGFP, que se usó en los estudios previos de los inventores para la visualización de las regiones transducidas. La secuencia stuffer se diseñó para que esté desprovista de secuencias potenciadoras y represoras, activadores o represores de corte y empalme, y ARN antisentido u otro no codificante, y de tamaño suficiente para el empaquetado óptimo del cassette de expresión de ARNi pequeño. Los vectores de AAV2/1 expresaron finalmente miHDS1 o miCtl, un control usado previamente en muchos de los estudios *in vivo* de los investigadores (Figura 1B).

[293] Se pesaron ratones de tipo salvaje y se evaluó el rendimiento basal por rotarod a las 7 semanas de edad antes de la distribución de los animales en grupos de capacidad iguales (para evitar las diferentes pretratamiento entre los grupos). Se inyectó bilateralmente AAV.miHDS1 o AAV.miCtl en el cuerpo estriado a las 8 semanas de edad con AAVmiHDS1/Stuffer (n=13) y virus AAVmiCtl/Stuffer (n=11) (Figura 1C,D). Tan pronto como a los 2 meses después de la administración del AAV, los ratones que expresaban miHDS1 tuvieron déficits significativos en rotarod y mostraron una latencia disminuida a caerse respecto a los hermanos de camada tratados como control (Figura 1D). Y si bien todos los animales aumentaron de peso durante el curso del estudio, los ratones tratados con HDS1 aumentaron significativamente menos que los ratones tratados con miCtl (Figura 1E).

Caracterización de genes no diana de miHDS1 en el cerebro de ratón.

[294] El miHDS1 se diseñó para que tenga un silenciamiento no diana mínimo de transcriptos humanos, pero no se optimizó para seguridad en ratón y los inventores no evaluaron, previamente, la toxicidad potencial de la secuencia contra los transcriptos de ratón *in silico*. A pesar de que la construcción AAV.miHDS1.eGFP usada previamente en monos mostró una carga de hebra apropiada, los inventores ensayaron luego la fidelidad del cassette de expresión miHDS1.stuffer para la desviación de hebra, ya que cada hebra, si se cargaba, podría producir un silenciamiento no diana. Para esto los inventores diseñaron construcciones reporteras que consisten en dianas de miHDS1 clonadas corriente debajo de un reportero luciferasa. Los inventores encontraron represión de la hebra guía, y no represión de la hebra no guía (Figura 1G). Esto está de acuerdo con los análisis de expresión *in vitro* previos de los investigadores de los cassettes de expresión HDS1.eGFP, y es avalado por el hecho de que los inventores diseñaron la secuencia de miHDS1 con baja estabilidad termodinámica del extremo 5' para promover la carga de la hebra "antisentido" guía en el complejo RISC. Por lo tanto, los déficits neuronales observados por la expresión de miHDS1 posiblemente se deban a la unión de la hebra "antisentido" guía a la 3'UTR de los ARNm no deseados y el silenciamiento de la expresión por un mecanismo similar a miARN.

[295] Debido a que los estudios previos demostraron que la mayor parte de los efectos no diana son debidos a la unión mediada por semilla a otras 3'UTR del ARNm, los inventores primero identificaron posibles no dianas de miHDS1 usando una estrategia *in silico* común. Se han descrito muchos programas diferentes de predicciones de diana para identificar sitios de unión de miARN putativos, tales como los algoritmos TargetScan (TS) y PITA. El TargetScan predice dianas para los miARN específicos por búsqueda en las secuencias de 3'UTR de la presencia de sitios 8mero y 7mero complementarios con la secuencia de semilla de miARN. El algoritmo mejora la precisión de la predicción de diana al priorizar los sitios diana con apareamiento compensatorio de bases en 3', contexto de la secuencia local y conservación fuerte de secuencia que se sabe que son favorables para la regulación mediada por miARN. Debido a que en trabajos previos se ha mostrado que los efectos fuera de diana mediados por semilla son específicos de tejido, los inventores usaron TargetScan para predecir dianas en base a la complementariedad de secuencia semilla en el transcriptoma de 3'UTR de ratón. El algoritmo PITA incorpora la accesibilidad del sitio diana para predecir los sitios de unión de miARN. Para un dado sitio diana el PITA determina un valor de puntuación de ddG, la diferencia de energía libre entre la unión del miARN a la diana (dG_{duplex}) y el apareamiento de los nucleótidos del sitio diana (dG_{open}). En base a PITA, las puntuaciones de ddG por debajo de -10 son más probables de ser funcionales para las dianas de miARN endógeno, a pesar de que el umbral para una secuencia de miARN sobreexpresada podría ser mayor (entre 0 y -10). Por lo tanto, en esta estrategia los inventores usaron TargetScan para identificar todos los sitios de unión a semilla potenciales, seguido por el algoritmo PITA para determinar la puntuación de ddG, y clasificar todos los potenciales sitios de miHDS1. Utilizando la estrategia de los inventores contra el 3' UTRoma de ratón, los inventores predicen 197 transcriptos como potenciales no dianas para miHDS1, con 170 expresados en el cuerpo estriado (Figura 3c). Tal como se esperaba, la predicción de los no diana de miHDS1 en las 3'UTR del humano ortólogo y rhesus reveló que el direccionamiento fuera de diana de miHDS1 en ratón no está conservado.

[296] Los inventores identificaron Bcl2, Sdf4, Smad9, Bmi1, Mettl2, Lancl1 y Map2k6 entre los primeros del percentil 25^{to} de la lista de genes no diana (Figura 2a,b). Los inventores analizaron muestras de cuerpo estriado obtenidas de ratones tratados con miCtl o miHDS1 mediante Q-PCR para estas no dianas predichas y Htt de ratón. Como se esperaba, la expresión de Htt de ratón se redujo significativamente (hasta 70%) en ratones tratados con miHDS1 respecto a los tratados con miCtl (Figura 2c). Entre el conjunto de transcritos no diana evaluados, Bcl2, Sdf4 y Map2k6 se redujeron significativamente en muestras de tejido obtenidas de ratones tratados con AAV.miHDS1 (Figura 2c). No se predijo que ninguno de estos transcritos esté afectado por miCtl. Los inventores confirmaron estos resultados usando una línea de células de cuerpo estriado de ratón inmortalizadas que tiene un alelo de Htt normal (células SthdhQ7). Las células SthdhQ7 se electroporaron con plásmidos que expresan miHDS1, miCtl o sin transcripto (conteniendo solo el promotor U6), y 24 horas más tarde se analizaron los transcritos mediante Q-PCR. Como se observa en el cerebro de ratón, la expresión de Bcl2 se redujo en las células que expresan miHDS1, pero no en aquellas que expresan miCtl o tratadas con el plásmido U6 control (Figura 2d). Por el contrario, la expresión de Sdf4 y Map2k6 no se redujo por la sobreexpresión de miHDS1 (Figura 2d), sugiriendo que estos genes pueden no estar dirigidos a no dianas *in vivo*, y puede reflejar efectos indirectos de la supresión de Htt en el tiempo o la supresión fuera de diana en célula no neuronales; a pesar de que AAV2/1 transduce neuronas principalmente. Interesantemente, la expresión de Smad9 aumentó significativamente en células SthdhQ7, y aumentó, a pesar de no ser de manera significativa, en cuerpo estriado tratado con miHDS1 (Figura 2d). Por lo tanto, el tamizaje de los inventores reveló a Bcl2 como un potencial no diana deletéreo de HDS1 en el 3' UTRoma de ratones.

Rescate de miHDS1 para seguridad en cerebro de ratón.

[297] Cuando una secuencia de miARN se carga en el RISC que contiene una proteína catalítica argonauta (Ago2), se requiere complementariedad completa de unión entre un miARN y su secuencia diana para mediar la escisión endonucleotídica de ARNm. Sin embargo, pueden tolerarse no coincidencias producidas por mutaciones puntuales individuales en la secuencia del miARN. Por lo tanto, para modificar el perfil de no dianas de miHDS1, que está dirigido principalmente por la región semilla, los inventores introdujeron mutaciones puntuales individuales que fueron diseñadas para alterar la semilla sin afectar la eficacia de silenciamiento (Figura 3A).

[298] Como un primero paso para identificar cuales mutaciones de semilla (i) cambian de manera eficaz el perfil de no dianas, (ii) mantienen un potencial total bajo de fuera de diana y (iii) silencian la HTT de humano, los inventores repitieron el análisis de predicción fuera de diana usando todas las variantes de semilla de nucleótido único (posiciones 2-7) de miHDS1 (Figura 3C). Los mutantes de la posición 8 se descartaron, ya que el perfil de fuera de diana se superpuso extensamente con el del miHDS1. Esto era esperado, ya que el apareamiento de la posición 8 no es necesario para el silenciamiento mediado por miARN. Los mutantes de semilla en las posiciones 3 y 4 también fueron descartados, ya que estas mutaciones aumentaron significativamente el número de fuera de diana predichos. Para las variantes de semilla remanentes, el potencial fuera de diana total fue comparable a miHDS1, con menos de 10% de los fuera de diana de miHDS1 estando compartidos con las variantes de miHDS1 (Figura 3C).

[299] Por lo tanto, los inventores introdujeron mutaciones puntuales individuales en las posiciones 2, 5, 6 y 7 de la región semilla de miHDS1 para generar las variantes de miHDS1. Debido a que el objetivo de los inventores es silenciar la HTT de humano, los inventores primero cribaron todas las variantes en la línea de células HEK293 derivadas de humano y determinaron la eficacia de silenciamiento mediante Q-PCR (Figura 4A). No todas las variantes de miHDS1 redujeron la expresión de HTT de manera equivalente a la miHDS1 original. En comparación con miHDS1, las variantes de miHDS1 con no coincidencias en las posiciones 2 y 7 alteran la eficacia de silenciamiento de miHDS1. Sin embargo, no se observaron diferencias significativas para las variantes de miHDS1 que contenían no coincidencias en la posición 5 o 6. Of Para destacar, entre las diferentes variantes de miHDS1 con una no coincidencia en la posición 7 solo la variante con una sustitución C>U tuvo una eficacia equivalente respecto a miHDS1, posiblemente debido a la estabilidad termodinámica de tambaleo de G:U. Los inventores eligieron miHDS1v6A y miHDS1v5U para experimentos adicionales en base a: (1) su mayor eficacia de silenciamiento respecto a las otras variantes de miARN que contienen una no coincidencia en la misma posición semilla, y (2) el tipo de no coincidencia de nucleótido generado (U:U, miHDS1 v5U; A:G, miHDS1v6A, Figura 4D) que tiene un perfil moderado fuera de diana que difiere extensamente de HDS1 (Figuras 4B, 4C, 4E, 4F).

[300] Las secuencias de miARN cargadas en RISC son las siguientes (NOTA: 3' → 5'):

miHDS1v5U: 3'-CACGACCGAGCGUACCUGCUG-5' (SEQ ID NO:6)

miHDS1v6A: 3'-CACGACCGAGCGUACAGCUG-5' (SEQ ID NO:7)

[301] Las Pri-miHDS1 son las siguientes (5' → 3'):

Pri-miHDS1v5U (SEQ ID NO:8):

NNNAGCGAUGCUGGCUCGCAUGGUCGAUACUGUAAAGCCACAGAUGCUGUCGUCCAUGCGAGCC
AGCACCGCANNN

Pri-miHDS1v6A (SEQ ID NO:9):

NNNAGCGAUGCUGGCUCGCAUGGGUCGAUACUGUAAAGCCACAGAUGCUGUCGAAACAUGCGAGCC
AGCACCGCANNN

[302] Las Pre-miHDSI son las siguientes (5'→3'):
Pre-miHDS1v5U (SEQ ID NO:10):

5'P-

GCUGGCUCGCAUGGGUCGAUACUGUAAAGCCACAGAUGCUGUCGUCCAUGCGAGCCAGCAC-OH3'
Pre-miHDS1v6A (SEQ ID NO:11):

5'P-

5 GCUGGCUCGCAUGGGUCGAUACUGUAAAGCCACAGAUGCUGUCGAAACAUGCGAGCCAGCAC-OH3'

[303] Como se esperaba, la expresión de miHDS1v6A y miHDS1v5U redujo los niveles de la proteína Htt en la línea de células de ratón (SthdhQ7) y de humano (HEK293), son diferencias significativas para miHDS1 (Figuras 4A-4E).

10 [304] Luego, los inventores evaluaron el efecto de miHDS1v6A y miHDS1v5U sobre los transcritos de ratón fuera de diana de miHDS1 validados. La estabilidad de apareamiento de semilla (SPS), la energía libre de unión entre una semilla de miARN y su ARNm diana, influye sobre si una secuencia de miARN produce efectos de silenciamiento fuera de diana. A pesar de que las variantes de miHDS1 tuvieron un valor similar de SPS respecto a miHDS1 original para sus dianas, las no coincidencias en la región de las variantes de miHDS1 disminuyeron el valor de SPS sobre los no diana de miHDS1 (Figura 3b). En base a PITA, la introducción de una no coincidencia en la región semilla redujo el valor de puntuación de ddG en todos los genes no diana predichos de HDS1, siendo más significativo para miHDS1v6A que miHDS1v5U. Interesantemente, la secuencia semilla de las variantes de miHDS1 generó un sitio diana diferente en algunas de las no dianas de miHDS1.

20 [305] Los inventores demostraron previamente el silenciamiento de Bcl2 *in vivo* e *in vitro* por miHDS1. En base a TargetScan, miHDS1v6A y miHDS1v5U no estarán más dirigidos contra la 3'UTR de Bcl2, y PITA predice una puntuación de ddG reducida en el sitio de miHDS1, sugiriendo que el silenciamiento de Bcl2 será debilitado por miHDS1v6A o miHDS1v5U. Para evaluar esto, se electroporaron células SthdhQ7 con plásmidos que contenían los cassettes de expresión miARN o el plásmido control solo con el promotor U6, y 24 horas después se determinó la expresión de Bcl2, Htt y Smad9 mediante Q-PCR. En relación a los controles (miCtl y U6), los niveles de ARNm de Htt se redujeron significativamente en las células electroporadas con miHDS1 y variantes de miHDS1 (Figura 4g). Pero importantemente, si bien miHDS1 redujo significativamente la expresión de Bcl2 en un 40%, no se observó silenciamiento después de la electroporación de miHDS1v6A. La expresión de miHDS1v5U estuvo aún activa contra Bcl2, reteniendo los niveles de silenciamiento en 20% (Figura 4h). Interesantemente, la sobreexpresión de Smad9 asociada con la expresión de miHDS1 no se observó en células electroporadas con miHDS1v5U o miHDS1v6 (Figura 4i).

35 **Redireccionamiento de miCtl contra el ARNm de huntingtina de humano**

[306] Los resultados previos de los inventores expusieron la toxicidad de miHDS 1 debido a sus efectos fuera de diana, pero también destacaron que miCtl es tolerable cuando se expresa en el cerebro de ratón. El miCtl se diseñó con un perfil de silenciamiento bajo de fuera de diana, pero no se pretendió que estuviera dirigido contra el ARNm de huntingtina. Por lo tanto, los inventores ensayaron si los inventores podrían obtener una ventaja de la relativa seguridad de la semilla de miCtrl en el cuerpo estriado de ratón, y diseñar una iARN dirigida contra HTT alrededor de la semilla.

40 [307] Como un primer paso, los inventores cribaron el ARNm de HTT de humano, o diana clínica, para evaluar secuencias completamente complementarias con la región semilla de miCtl, pero no encontraron nada. Siguiendo la misma estrategia para el diseño de las variantes de miHDS 1, los inventores repitieron este análisis *in silico* permitiendo no coincidencias individuales entre los nucleótidos 2 a 7 de la secuencia miARN. Los inventores encontraron cuatro secuencias complementarias (miHDss1-4): MiHDss1 (no coincidencia en posición 7) y miHDss4 (no coincidencia en posición 4) dirigida contra HTT en la 3'UTR, mientras que miHDss2 (no coincidencia en posición 6) y miHDss3 (no coincidencia en posición 5) dirigidas contra HTT en la región codificante que se expande en la unión del exón7-8 o en el exón33, respectivamente (Figura 5a,b). Cuando se ensaya en células HEK 293, solo miHDss3 silenció la expresión de HTT a entre 40 y 50% de las células tratadas con control, determinado por Q-PCR (Figura 5c) y transferencia western (Figura 5d,e).

[308] Debido a que miCtl y miHDSS3 comparten la misma secuencia semilla los inventores esperan que ambos miARN tengan el mismo perfil fuera de diana. Sin embargo, como se observa en los miARN endógenos de una familia de miARN específica, la eficacia de silenciamiento podría cambiar debido a las diferencias de secuencia en la región 3' de cada miARN. Los inventores usaron el algoritmo PITA para comparar la estabilidad de unión del miARN y el potencial de silenciamiento entre los no diana de miCtl y miHDSS3. De acuerdo a la estrategia *in silico* de los inventores, se predicen 89 sitios no diana para miCtl y miHDSS3, con 67 expresados en el cuerpo estriado. Interesantemente, la región 3' de miHDSS3 aumenta la estabilidad de unión de no diana-miARN respecto a miCtl (Figura 5f).

Caracterización de las variantes de miHDS1 y la tolerabilidad de miHDSS3 en el cerebro de ratón

[309] Para determinar la tolerabilidad *in vivo* de las nuevas secuencias, se clonaron los cassettes de expresión de miARN en el vector lanzadera de AAV de los inventores (Figura 6b). Se dividieron ratones de siete semanas de edad de tipo salvaje en grupos basados en el peso equivalente y rendimiento basal en rotarod, y posteriormente se inyectaron bilateralmente en el cuerpo estriado con virus expresando miHDS1v6A, miHDS1v5U, miHDSS3, o miHDS1, miCtl, y formulación amortiguadora (FB) como controles experimentales. Se registró el peso del ratón a los dos y cuatro meses después de la inyección, y se determinaron los efectos neurológicos adversos mediante el uso de las pruebas de rotarod acelerada, clasping, y de campo abierto (Figura 6a).

[310] Consistente con los resultados previos de los inventores, los ratones que expresan miHDS1 mostraron déficits motores en el aparato de rotarod acelerado (Figura 6c). Además, no se observaron diferencias entre los ratones inyectados con solución amortiguadora FB sola o miCtl. Este resultado es importante porque sugiere que los efectos adversos no son un resultado de la cooptación de la vía endógena, sino de efectos fuera de diana de miHDS1 específicos. Interesantemente, y consistente con los estudios *in vitro* de los inventores, miHDS1v5U mostró déficits en rotarod también. Esto puede reflejar que las no coincidencias pirimidina: pirimidina (U:U, miARN: ARNm) presentan un poder de discriminación moderado y esta variante silencia aun parcialmente a Bcl2 (Figura 4H). También se predice a partir del trabajo de los inventores, toxicidad mejorada mediada por miHDS 1 de miHDS1v6A, sin diferencias significativas observadas entre miHDS1v6A y miCtl a los 2 o 4 meses después de la inyección de AAV. Además del silenciamiento de la huntingtina de humano, miHDSS3 comparte el mismo perfil fuera de diana que miCtl. Sin embargo, el algoritmo PITA sugirió que miHDSS3 es más propenso a silenciar el repertorio fuera de diana de miCtl al aumentar la estabilidad de unión del par miARN:ARNm (Figura 5f). Sin embargo, no se observaron diferencias significativas en el rotarod acelerado a 2 o 4 meses (Figura 6c).

[311] Excepto para los ratones inyectados con miHDS1v5U que perdieron peso en el tiempo (-1,7g, 8% de reducción a los 4 meses), se registró aumento de peso corporal en todos los otros grupos. Además, se redujo significativamente el aumento de peso con el tratamiento con miHDS1, como antes. A los 4 meses, los ratones inyectados con miHDS1 tuvieron 1,3 gramos (5%) de aumento de peso corporal mientras que los otros grupos tuvieron aumentos de entre 3,6 y 5,2 gramos (entre 15 y 22% de aumento a los 4 meses) (Figura 6d).

DISCUSIÓN

[312] En este trabajo los inventores proponen evaluar la seguridad en ratones normales de un iniciador de iARN diseñado para seguridad en humanos y que se mostró en un trabajo previo que es seguro en primates no humanos. La evaluación de fármacos para el uso en humanos en dos especies, en general un roedor y un mamífero más grande, es un procedimiento estándar para la aprobación de regulaciones para avanzar a los estudios de fase aguda. Si bien los inventores no encontraron una toxicidad notable en monos, que también fue informado por otros en un estudio posterior, cuando la construcción deseada se ensayó en roedores, se observó una toxicidad aguda.

[313] Los inventores encontraron que podían reducir la toxicidad de la secuencia ensayada, miHDS1, realizando mutaciones puntuales en la semilla para alterar el perfil fuera de diana. La modificación individual de la secuencia semilla cambió el perfil fuera de diana del miHDS1 original, a la vez que se mantuvo la eficacia de silenciamiento. El miHDS1v6, pero no el miHDS1v5, restauró la tolerabilidad de miHDS1 en el cerebro de ratón. El miHDS1v5U generó una no coincidencia U:U, que es una no coincidencia pirimidina:pirimidina, mientras que el miHDS1v6A generó un A:G que es una no coincidencia Purina:Pirimidina, y se encontró que era la más eficaz para discriminar.

[314] Como una alternativa a la alteración de la semilla de HDS1, los inventores indicaron previamente que su secuencia control, también diseñada para bajo potencial de fuera de diana, podría ser rediseñada para dirigirla a la HTT de humano. Ambas secuencias, cuando se ensayaron en ratones, fueron bien toleradas y no indujeron neuropatología o déficits neurológicos, como se indicó previamente para el HDS1 parental.

[315] Estos hallazgos destacan el contraste entre el desarrollo de fármacos tradicional y el nuevo campo emergente de las medicinas basadas en ácidos nucleicos. Si bien el objetivo de todo desarrollo de fármacos para el humano es la seguridad y la eficacia en la población diana, en el caso de las medicinas basadas en ácidos nucleicos que pretenden que el fármaco interactúe directamente con el genoma y/o los transcritos expresados. Por lo tanto, los fármacos que se basan en la especificidad de secuencia y que están optimizados para la seguridad en humanos

posiblemente interaccionen con los genomas de otras especies, y en particular aquellos de especies relacionadas de manera distante tales como roedores. Por otra parte, si las secuencias están optimizadas para seguridad en roedores, el riesgo de problemas en el contexto del genoma humano es mayor.

5 **SÍNTESIS**

[316] Los resultados presentes destacan que (1) el perfil de seguridad y tolerabilidad de un miARN es específico de especie, enfatizando la cuidadosa interpretación de los estudios iniciales usando modelos de ratón de enfermedad, y (2) la modificación de la secuencia semilla individual es una estrategia eficaz para resolver la toxicidad fuera de diana de una secuencia de miARN, a la vez que se mantiene la eficacia de silenciamiento.

10 **MATERIALES Y MÉTODOS**

15 **Líneas celulares y transfecciones**

[317] Se obtuvieron HEK293 de la ATCC y se cultivaron en las condiciones provistas por el fabricante. Las SthdhQ7 se obtuvieron generosamente de Marcy MacDonald. Todas las transfecciones de ADN plasmídico en HEK293 se realizaron con lipofectamina 2000 (Invitrogen) usando las instrucciones provistas por el fabricante. La transfección de ADN de las células SthdhQ7 se realizó usando un sistema de transfección Invitrogen Neon usando las condiciones de electroporación y de acuerdo a las instrucciones provistas por el fabricante.

20 **Diseño del vector y producción de AAV**

[318] Se generaron secuencias artificiales de miARN (miCtl, miHDss variantes, miHDS1 y miHDS1 variantes) mediante extensión con polimerasa de oligonucleótidos de ADN solapantes (IDT, Coralville). Los productos extendidos por la polimerasa se purificaron usando el kit de purificación de PCR Qiaquick, se digirió con XhoI-SpeI y se clonó en un sitio XhoI-XbaI en un cassette de expresión Pol-III que contenía el promotor U6 de ratón, MCS y el terminador de Pol-III (6T).

[319] Se construyeron vectores reporteros con luciferasa para iARN usando el vector psiCheck2 (Promega). Los oligonucleótidos de ADN con cola que contienen un sitio diana de iARN perfectamente complementario, individual, para las hebras sentido o antisentido de miHDS1 se hibridizaron y clonaron en sitios XhoI-NotI corriente abajo del codón de terminación de la secuencia de ADNc de luciferasa Renilla.

[320] Para los estudios *in vivo*, se transportaron los cassettes de expresión de miARN en plásmido lanzadera AAV corriente arriba de una secuencia stuffer de ADN. El cassette de expresión de miARN y la secuencia stuffer se flanquearon en cada extremo con secuencias de repeticiones terminales invertidas de 145 pb de serotipo de 2AAV.

40 **Ensayos de luciferasa in vitro**

[321] Se cotranfectaron células HEK293 con un 70% de confluencia cultivadas en una placa de 24 pocillos con plásmido que expresa miARN y plásmido reportero de luciferasa para iARN. A las 24 horas, se lavaron las células con PBS helado y se evaluaron las actividades de luciferasa Renilla y Firefly usando el sistema de ensayo de reportero luciferasa dual (Promega) de acuerdo con las instrucciones del fabricante, usando 20 ml de lisado celular. Las lecturas luminiscentes se obtuvieron con un luminómetro Monolight 3010 (Pharmingen, EE.UU.). Se calcularon las unidades de luz relativas como el cociente de las unidades de luz relativas de Renilla/Firefly y los resultados se expresaron en relación a un miARN control.

50 **Análisis mediante transferencia Western**

[322] Se transfectaron las células HEK293 con cassettes que expresan el miARN como se indicó. A las 48 horas se lavaron las células una vez con PBS helado y se lisaron con solución amortiguadora de lisis Passive (PBL, Promega). Se determinó la concentración de proteínas mediante el método de Bradford-Lowry (BioRad) y se cargaron 10 mg de proteína en un gel de NuPAGE 3-8% Tris-Acetato (Novex Life technologies). Las proteínas se transfirieron a membranas de PVDF y se incubaron con anticuerpos anti-Htt de ratón (1:5000, Millipore, CA), o anti-beta-actina de conejo (1:40000, Sigma) seguido por anticuerpos acoplados a peroxidasa de rábano picante (1:10.000, ratón; o 1:50.000, conejo; Jackson ImmunoResearch, West Grove, PA). Las membranas se desarrollaron con el reactivo ECL-Plus (Amersham Pharmacia). Se determinó la eficacia de silenciamiento mediante densitometría (n=4 experimentos independientes) de niveles de proteína en relación a la beta actina con el sistema de obtención de imágenes VersaDocTM (Biorad) y el programa de análisis Quantity OneR.

60 **Extracción de ARN y análisis mediante QPCR**

[323] Se extrajo el aislamiento de ARN total usando Trizol (Life Technologies, Grand Island, NY) de acuerdo con el protocolo del fabricante, excepto por el agregado de 1ml de Glycoblue a la fase acuosa en el paso de precipitación con isopropanol y un lavado individual con etanol 70% frío. Las muestras de ARN se cuantificaron mediante

espectrometría y posteriormente se generaron ADNc a partir de 500 ng de ARN total con hexámeros al azar (reactivos TaqMan RT, Applied Biosystems). Se diseñaron pares de cebadores SyBrGreen Q-PCR para genes no diana de ratón usando el servidor de internet RealTime PCR Custom Assay Design (IDT, Coralville). Se realizó una curva estándar de siete puntos con un ensayo de curva de fusión final realizado para validar cada par de cebadores. Solo los pares de cebadores con eficacias de amplificación de $100\pm 5\%$ y un único producto de amplificación se usaron para determinar la expresión génica relativa usando el método de ddCt.

Estudios en ratones

[324] Todos los protocolos en animales fueron aprobados por el Comité de cuidado y uso de animales de la Universidad de Iowa. Se obtuvieron ratones FBV de tipo salvaje y BACHD de Jackson Laboratories (Bar Harbor, ME, EE.UU.). Los ratones fueron genotipificados usando cebadores específicos para el transgén de huntingtina mutante de humano flanqueando la repetición CAG, y se usaron hermanos de camada transgénicos y de tipo salvaje de la misma edad para los experimentos indicados. Los ratones se albergaron en un ambiente con control de temperatura con un ciclo de 12 horas de luz/oscuridad. Se proveyó alimento y agua *ad libitum*. A los tiempos indicados los ratones se inyectaron con el virus AAV2/1-mU6-miARN/Stuffer. Para las inyecciones de AAV, los ratones se anestesiaron con una mezcla de ketamina y xilazina, y se inyectaron 5 ml del AAV bilateralmente en el cuerpo estriado a una velocidad de 0,2 ml/min (coordenadas: +0,86 mm rostral a Bregma, +/-1,8 mm lateral a medial, -2,5 mm ventral desde la superficie del cerebro). Los ratones usados para los análisis de expresión génica se anestesiaron con una mezcla de ketamina y xilazina y se perfundieron con 18 ml de solución salina fría 0,9% mezclada con 2 ml de solución RNAlater (Ambion). A los tiempos indicados se sacrificaron los ratones y se extrajo el cerebro, se bloqueó, y cortó en rodajas coronales de 1 mm de espesor. Se tomaron bocados de tejido del cuerpo estriado usando un sacabocado de tejido (1,4 mm de diámetro; Zivic Instruments, Pittsburgh, PA, EE.UU.). Todos los bocados de tejido se congelaron inmediatamente en nitrógeno líquido y se almacenaron a -80°C hasta su uso.

Análisis de comportamiento

[325] Se determinó la coordinación motora de los ratones inyectados usando el aparato Rotarod (modelo 47,600; Ugo Basile, Comerio, Italia). Se realizó una prueba basal en rotarod a las 7 semanas de edad y nuevamente a los 2 y 4 meses después de la inyección de AAV. Se evaluaron los ratones durante cuatro días consecutivos con tres ensayos por día, con un periodo de 30 min de reposo entre los ensayos y un periodo de aclimatación de 5 minutos cada día comenzando sesenta minutos antes del primer ensayo. La latencia a la caída por ratón se calculó promediando las dos mejores pruebas de cada ratón por día de los cuatro días consecutivos ensayados. Para la prueba de *clasp ing* cada ratón se suspendió por la cola durante un minuto y se clasificó como *clasp ing* si el ratón mantuvo sus patas frontales juntas cerca de su torso.

[326] El uso de los términos "un" y "uno/a" y "el/la" y referencias similares en el contexto de la descripción de la invención se deben interpretar como que cubren el singular y el plural, a menos que se indique de otra manera en la presente o que el contexto lo contradiga claramente. Los términos "que comprende", "que tiene", "que incluye", y "que contiene" se deben considerar como términos de significado amplio (es decir, que se refieren a "que incluye, pero que no se limita a") a menos que se indique de otra manera. La descripción de rangos de valores en la presente solo pretende servir como un método abreviado para referirse individualmente a cada valor por separado que cae dentro del rango, a menos que se indique de otra manera en la presente, y cada valor por separado se incorpora a la memoria descriptiva como si fuera individualmente indicado en la presente. Todos los métodos descritos en la presente pueden llevarse a cabo en cualquier orden adecuado a menos que se indique de otra manera en la presente o que el contexto claramente lo indique de otra manera. El uso de cualquiera y todos los ejemplos, o lenguaje ejemplificativo (por ejemplo, "tal como") provisto en la presente, pretende solamente ilustrar de mejor manera la invención y no posee una limitación en el alcance de la invención a menos que se invoque de otra manera. No se debe interpretar ningún lenguaje en la memoria descriptiva como indicador de cualquier elemento no reivindicado como esencial para la práctica de la invención.

[327] Las formas de realización de esta invención se describen en la presente, incluyendo el mejor modo conocido para los inventores para llevar a cabo la invención. A las personas con experiencia ordinaria en el arte le resultarán evidentes variaciones de aquellas formas de realización al leer la descripción precedente.

REIVINDICACIONES

1. Un ácido nucleico que codifica para un transcripto primario de miARN artificial (pri-miARN) que consiste en, en orden de posición, una región flanqueante 5', una región no guía, una región bucle, una región guía, y una región flanqueante 3', en donde la región guía consiste en SEQ ID NO: 37 (miHDss3), SEQ ID NO:6 (miHDS1v5U) o SEQ ID NO:7 (miHDS1v6A), y la región no guía es por lo menos 80% complementaria con la región guía.
2. El ácido nucleico de la reivindicación 1, en donde la región flanqueante 5' comprende una secuencia de unión 5' unida de manera contigua a la región no guía, en donde la secuencia de unión 5' preferiblemente consiste en entre 5 y 8 nucleótidos o codifica para GUGAGCGA (SEQ ID NO:13) o GUGAGCGC (SEQ ID NO:14).
3. El ácido nucleico de la reivindicación 2, en donde la región flanqueante 5' comprende además una secuencia protuberante 5' posicionada corriente arriba a la secuencia de unión 5', en donde la secuencia protuberante 5' preferiblemente comprende un sitio de clonado o consiste en aproximadamente entre 1 y 10 nucleótidos o codifica para UAAACUCGA (SEQ ID NO:15).
4. El ácido nucleico de la reivindicación 3, en donde la región flanqueante 5' comprende además una secuencia espaciadora 5' posicionada corriente arriba a la secuencia protuberante 5', en donde la secuencia espaciadora 5' preferiblemente consiste en entre 10 y 12 nucleótidos o codifica para UGGUACCGUU (SEQ ID NO:16).
5. El ácido nucleico de la reivindicación 4, en donde la región flanqueante 5' comprende además una secuencia 5' corriente arriba posicionada corriente arriba a la secuencia espaciadora 5', en donde la secuencia 5' corriente arriba preferiblemente tiene aproximadamente entre 30 y 2000 nucleótidos de longitud.
6. El ácido nucleico de cualquiera de las reivindicaciones 1-5, en donde la región flanqueante 3' comprende una secuencia de unión 3' unida de manera contigua a la región guía, en donde la secuencia de unión 3' preferiblemente consiste en entre 5 y 8 nucleótidos o es por lo menos aproximadamente 85% complementaria a la secuencia de unión 5' o codifica para CGCCUAC (SEQ ID NO:18).
7. El ácido nucleico de la reivindicación 6, en donde la región flanqueante 3' comprende además una secuencia protuberante 3' posicionada corriente abajo de la secuencia de unión 3', en donde la secuencia protuberante 3' preferiblemente comprende un sitio de clonado o consiste en aproximadamente entre 1 y 10 nucleótidos o codifica para UAG (SEQ ID NO:30) o en donde la secuencia protuberante 5' es complementaria a la secuencia protuberante 3' en solo un nucleótido en cada extremo de la secuencia protuberante.
8. El ácido nucleico de la reivindicación 7, en donde la región flanqueante 3' comprende además una secuencia espaciadora 3' posicionada corriente abajo de la secuencia protuberante 3', en donde la secuencia espaciadora 3' preferiblemente consiste en entre 10 y 12 nucleótidos o codifica para AGCGGCCGCCA (SEQ ID NO:19) o es por lo menos aproximadamente 70% complementaria a la secuencia espaciadora 5'.
9. El ácido nucleico de la reivindicación 8, en donde la región flanqueante 3' comprende además una secuencia 3' corriente abajo posicionada corriente abajo de la secuencia espaciadora 3', en donde la secuencia 5' corriente arriba no se aparea significativamente con la secuencia 3' corriente abajo y/o en donde la secuencia 3' corriente abajo tiene aproximadamente entre 30 y 2000 nucleótidos de longitud.
10. El ácido nucleico de cualquiera de las reivindicaciones 1-9, en donde la región bucle tiene entre 15 y 25 nucleótidos de longitud.
11. Un ARN codificado por el ácido nucleico de cualquiera de las reivindicaciones 1-10.
12. Un cassette de expresión que codifica para el ácido nucleico aislado descrito en cualquiera de las reivindicaciones 1-10, que preferiblemente además comprende un promotor unido de manera contigua al ácido nucleico, en donde el promotor preferiblemente es un promotor de polII o polIII, un promotor específico de tejido o un promotor inducible.
13. Un vector que comprende el cassette de expresión de la reivindicación 12, en donde el vector preferiblemente es un vector de virus adenoasociado (AAV), en donde el AAV preferiblemente es AAV1, AAV2, AAV5, AAV6 y/o AAV9, más preferiblemente AAV2 o AAV2/1.
14. Un ácido nucleico aislado de entre 80 y 4000 nucleótidos de longitud, que comprende un ácido nucleico que codifica para un transcripto primario de miARN artificial (pri-miARN) que consiste en, en orden de posición, una región flanqueante 5', una región no guía, una región bucle, una región guía, y una región flanqueante 3', en donde la región guía consiste en SEQ ID NO: 37 (miHDss3), SEQ ID NO:6 (miHDS1v5U) o SEQ ID NO:7 (miHDS1v6A), y la región no guía es por lo menos 80% complementaria con la región guía.
15. Un ácido nucleico aislado que consiste en Pri-miHDS1v5U (SEQ ID NO:8), Pri-miHDS1v6A (SEQ ID NO:9), Pre-

miHDS1v5U (SEQ ID NO:10), o Pre-miHDS1v6A (SEQ ID NO:11).

5 16. El ácido nucleico de cualquiera de las reivindicaciones 1-11 o 14-15, el cassette de expresión de la reivindicación 12, o el vector de la reivindicación 13 para usar en el tratamiento de la enfermedad de Huntington.

10 17. El ácido nucleico de cualquiera de las reivindicaciones 1-11 o 14-15, el cassette de expresión de la reivindicación 12, o el vector de la reivindicación 13 para usar de acuerdo con la reivindicación 16, en donde el ácido nucleico, el cassette de expresión o el vector es para administrar al cerebro del sujeto directamente o a través del torrente sanguíneo, preferiblemente el ácido nucleico, el cassette de expresión o el vector es para administrar por vía intracraneal, más preferiblemente el ácido nucleico, el cassette de expresión o el vector es para administrar en la cisterna magna, cuerpo estriado, corteza o ventrículo, espacio subaracnoideo y/o espacio intratecal del sujeto.

15 18. El ácido nucleico de cualquiera de las reivindicaciones 1-11 o 14-15, el cassette de expresión de la reivindicación 12, o el vector de la reivindicación 13 para usar de acuerdo con la reivindicación 17, en donde el sujeto es humano y/o, el ácido nucleico, el cassette de expresión o el vector es para inyectar en 1 a 5 ubicaciones en el SNC o en una ubicación individual en el cerebro.

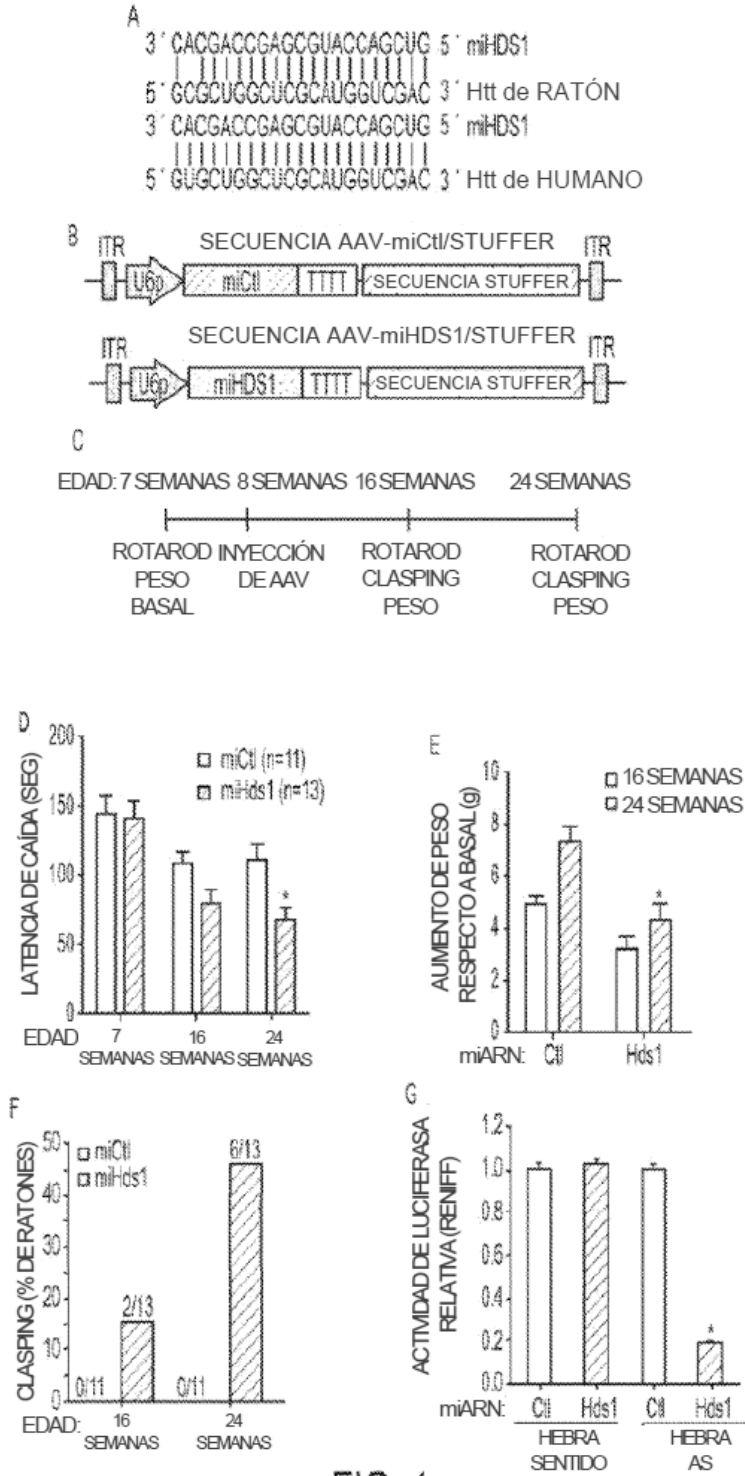


FIG. 1

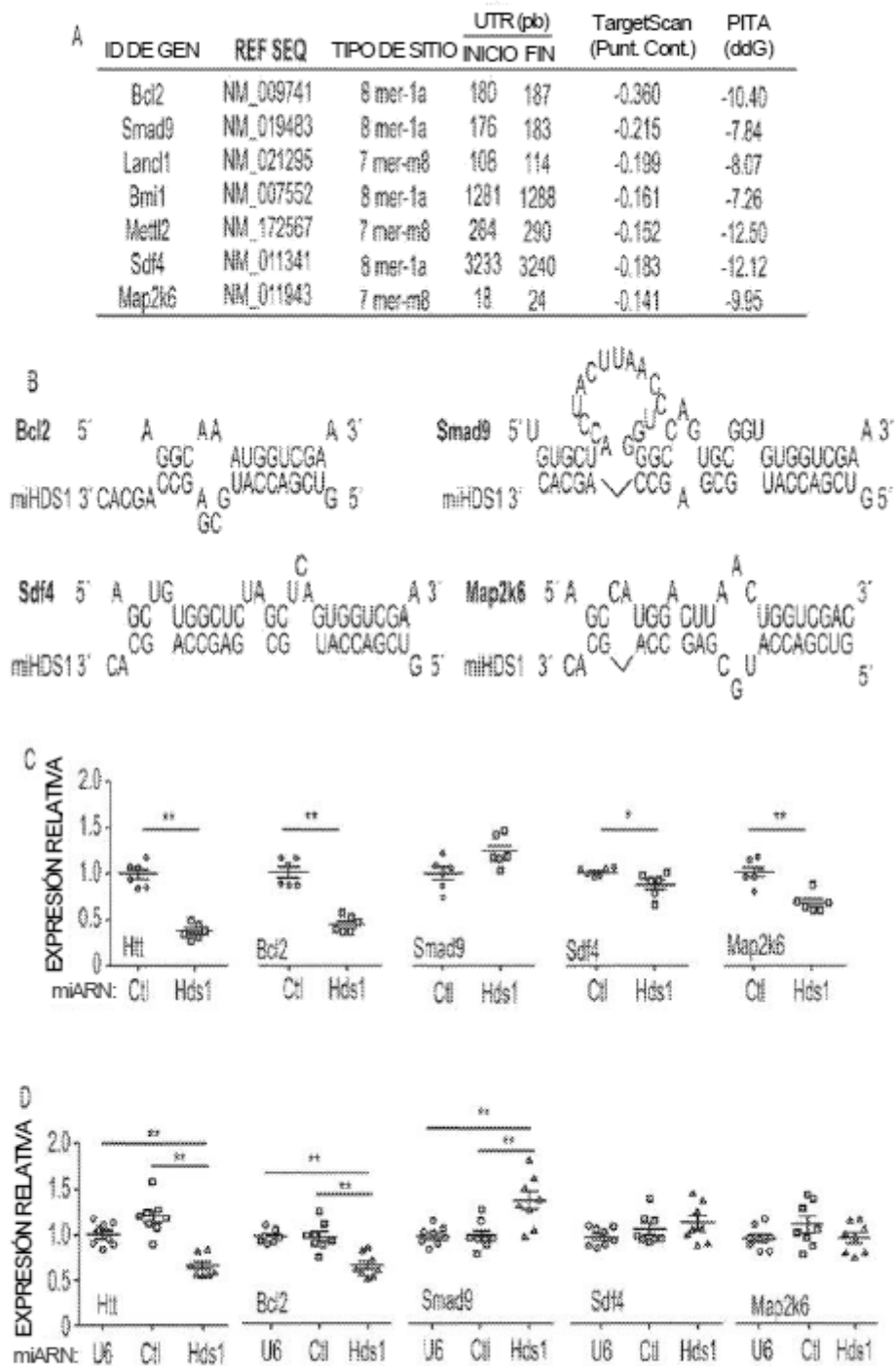
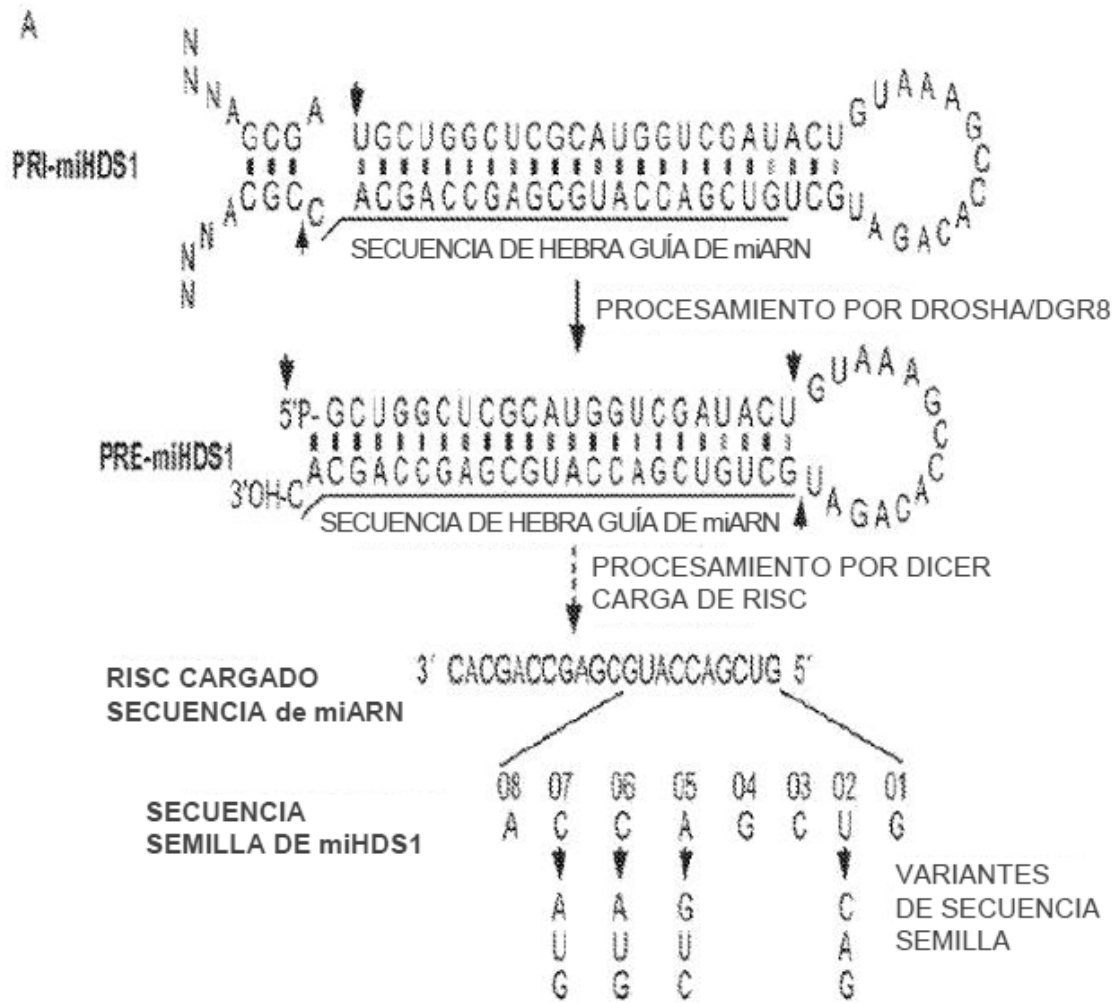


FIG. 2



B

VARIANTES DE HDS1 SPS SOBRE DIANAS DE HDS1 (ΔG : -13.13 Kcal/mol)

VARIANTES 2			VARIANTES 5			VARIANTES 6			VARIANTES 7		
HDS1v2C:	-11.55 Kcal/mol		HDS1v5G:	-6.53 Kcal/mol		HDS1v6A:	-8.1 Kcal/mol		HDS1v7A:	-9.45 Kcal/mol	
HDS1v2A:	-11.55 Kcal/mol		HDS1v5U:	-6.53 Kcal/mol		HDS1v6U:	-8.1 Kcal/mol		HDS1v7U:	-9.45 Kcal/mol	
HDS1v2G:	-11.55 Kcal/mol		HDS1v5C:	-6.53 Kcal/mol		HDS1v6G:	-8.1 Kcal/mol		HDS1v7G:	-9.45 Kcal/mol	

FIG. 3

miARN	GENES DIANA TOTALES		GENES EXPRESADOS EN CUERPO ESTRIADO	
	# DE GENES DIANA	# COMPARTIDO CON HDS1	# DE GENES DIANA	# COMPARTIDO CON HDS1
HDS1	197	197	170	170
2U>A	288	5	254	5
2U>C	350	9	318	8
2U>G	270	7	234	7
3C>A	1611	41	1381	37
3C>G	1704	31	1433	26
3C>U	1501	34	1246	33
4G>A	1803	38	1561	35
4G>C	2409	58	2088	53
4G>U	1998	41	1719	37
5A>C	329	10	279	8
5A>G	408	10	366	9
5A>U	361	13	317	13
6C>A	298	7	266	7
6C>G	346	8	310	8
6C>U	241	4	205	3
7C>A	329	12	287	11
7C>G	54	2	46	2
7C>U	228	6	188	6

FIG. 3
CONTINUACIÓN

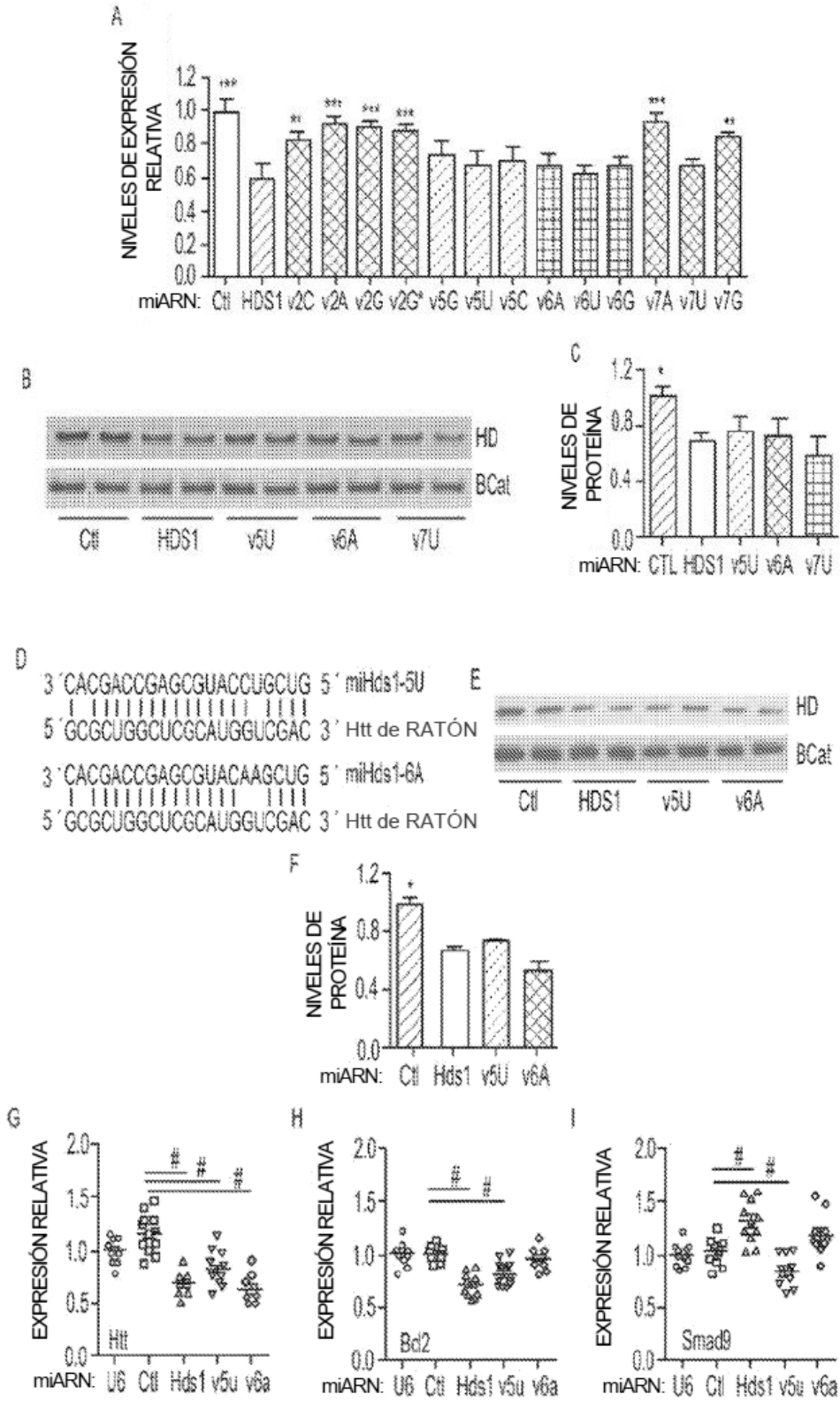


FIG. 4

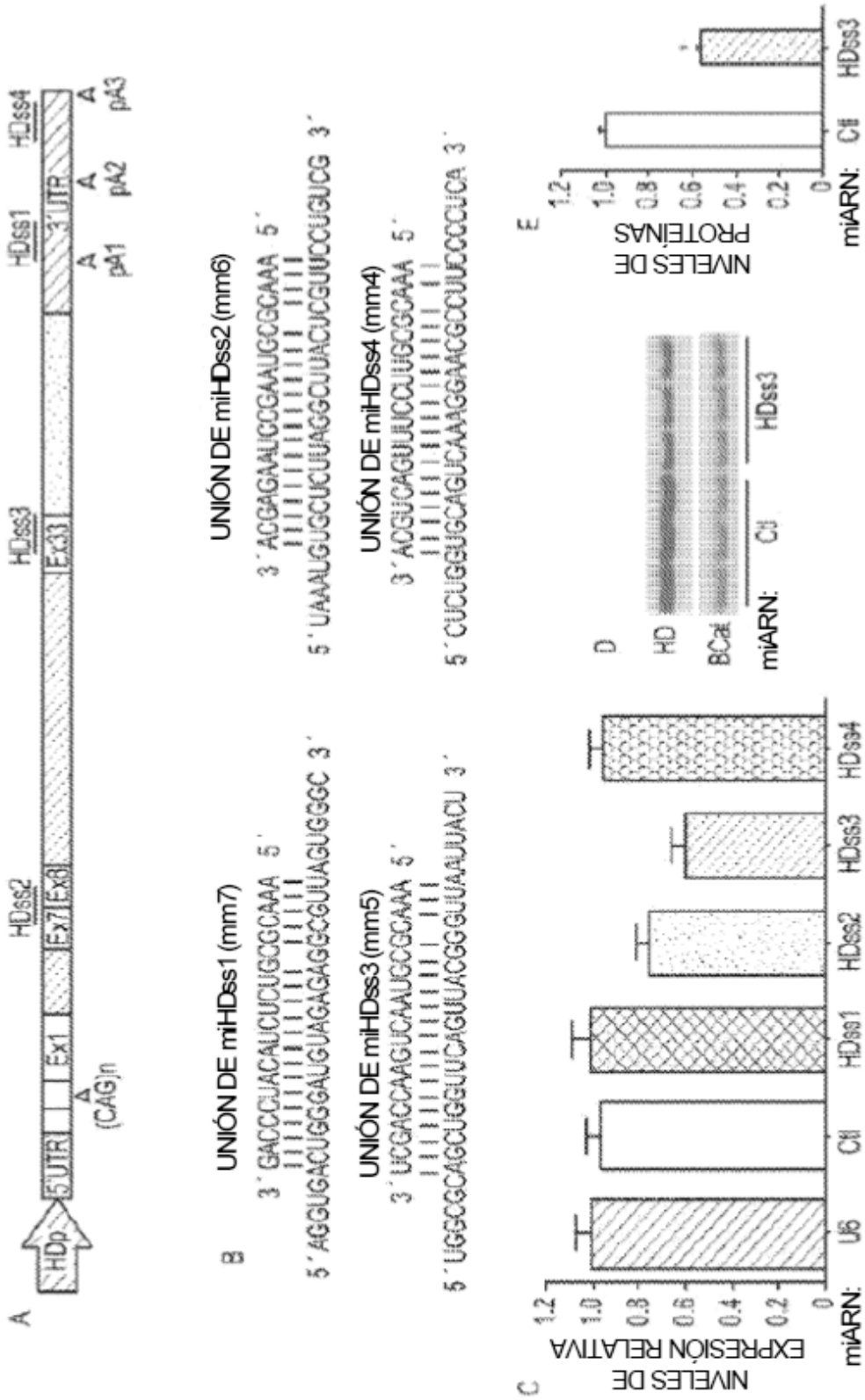


FIG. 5

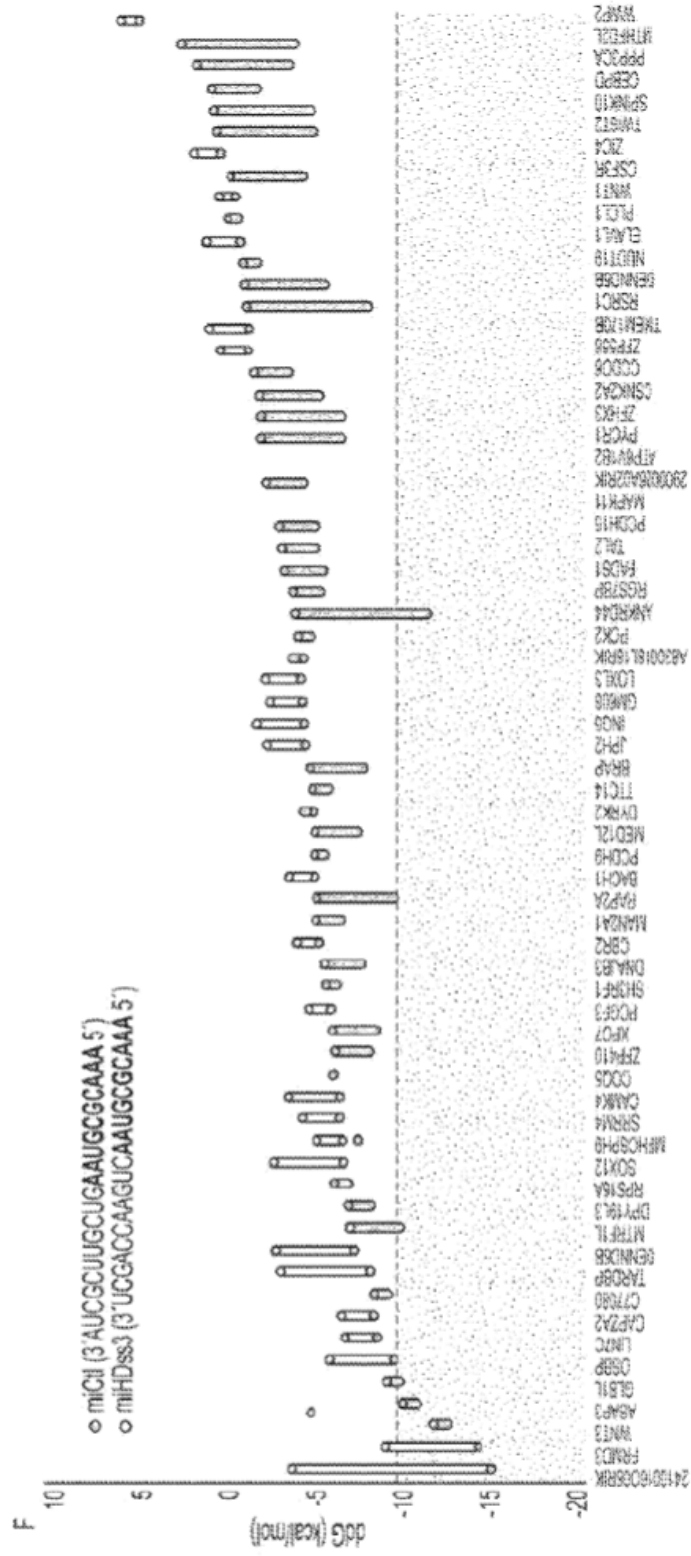


FIG. 5
CONTINUACIÓN

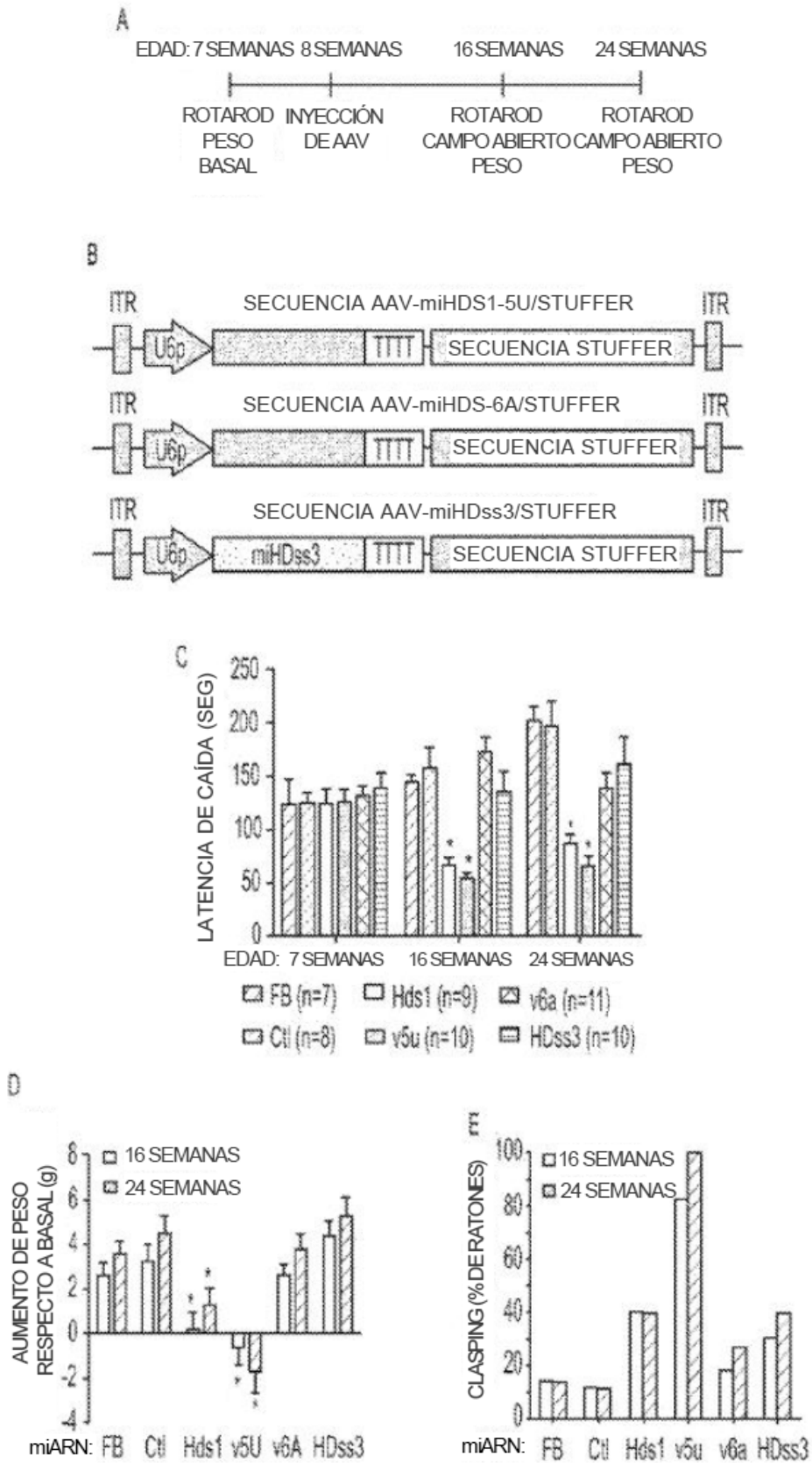


FIG. 6