

VECTORES VIRALES ADENOASOCIADOS PARA DIRIGIRSE A ESTRUCTURAS CEREBRALES PROFUNDAS

REFERENCIA A SOLICITUDES RELACIONADAS

- 5 Esta solicitud reivindica el beneficio de prioridad a la solicitud provisional de Estados Unidos No. 63/340,823, presentada el 11 de mayo de 2022, y solicitud provisional de Estados Unidos No. 63/342,324, presentada el 16 de mayo de 2022, el contenido completo de cada una de las cuales se incorpora por la presente por referencia.

REFERENCIA A LISTADO DE SECUENCIAS

- 10 Esta solicitud contiene un listado de secuencias XML, que se ha presentado electrónicamente y se incorpora por la presente por referencia en su totalidad. Este listado de secuencias XML, creado el 9 de mayo de 2023, se nombra CHOPP0054WO_ST26.xml y tiene un tamaño de 32.221 bytes.

ANTECEDENTES

- 15 1. Campo

La presente divulgación se refiere en general a los campos de medicina y virología. Más en particular, se refiere a composiciones y métodos para administración de agentes terapéuticos moleculares para estructuras cerebrales profundas.

2. Descripción de la técnica relacionada

- 20 Se han desarrollado diferentes estrategias para generar variantes de vectores de AAV, incluido el diseño racional y evolución dirigida. El enfoque de diseño racional utiliza el conocimiento de las cápsides de AAV para realizar cambios dirigidos a la cápside para alterar la eficiencia o especificidad de transducción, tal como mutaciones de tirosina en la superficie de cápside para incrementar la eficiencia de transducción. Sin embargo, ninguna variante de
- 25 AAV se dirige a estructuras cerebrales profundas de manera específica o eficiente. Las enfermedades que incluyen la enfermedad de Huntington (HD) y enfermedad de Parkinson (PD) afectan principalmente a las estructuras cerebrales profundas a las que las variantes de AAV existentes no son capaces de dirigirse de manera efectiva. Como tal, se necesitan

variantes de AAV que sean capaces de dirigirse a estructuras cerebrales profundas.

BREVE DESCRIPCIÓN

En la presente se proporcionan vectores virales que comprenden cada uno una cápside modificada, en donde la cápside modificada comprende al menos una secuencia de aminoácidos que dirige el vector viral a una estructura cerebral distinta.

En una realización, se proporcionan proteínas de cápside de virus adenoasociado modificada (AAV) que comprenden un péptido de direccionamiento que dirige un vector viral que comprende la proteína de cápside de AAV modificada a un órgano o estructura cerebral distinta, y el péptido de direccionamiento es de tres a diez aminoácidos de longitud. En algunos aspectos, las proteínas de cápside de AAV modificadas son proteínas de cápside de AAV1 modificadas, proteínas de cápside de AAV2 modificadas o proteínas de cápside de AAV9 modificadas.

En algunos aspectos, las proteínas de cápside de AAV modificadas se derivan de una proteína de cápside de AAV1 (ver SEQ ID NO: 1), y el péptido de direccionamiento se inserta después del residuo 590 de la proteína de cápside de AAV1. En algunos aspectos, el péptido de direccionamiento se flanquea por secuencias ligadoras, y las secuencias ligadoras en cada lado de los péptidos de direccionamiento tienen dos o tres aminoácidos de longitud. En algunos aspectos, las secuencias ligadoras son SSA en el lado N-terminal del péptido de direccionamiento y AS en el lado C-terminal del péptido de direccionamiento. En algunos aspectos, las proteínas de cápside de AAV1 modificadas tienen una secuencia al menos 95 % idéntica a SEQ ID NO: 4.

En algunos aspectos, las proteínas de cápside de AAV modificadas se derivan de una proteína de cápside de AAV2 (ver SEQ ID NO: 2), y el péptido de direccionamiento se inserta después del residuo 587 de la proteína de cápside de AAV2. En algunos aspectos, el péptido de direccionamiento se flanquea por secuencias ligadoras, y las secuencias ligadoras en cada lado de los péptidos de direccionamiento tienen dos o tres aminoácidos de longitud. En algunos aspectos, las secuencias ligadoras son AAA en el lado N-terminal del péptido de direccionamiento y AA en el lado C-terminal del péptido de direccionamiento. En algunos aspectos, las proteínas de cápside de AAV2 modificadas tienen una secuencia al menos 95 % idéntica a SEQ ID NO: 5.

En algunos aspectos, las proteínas de cápside de AAV modificadas se derivan de una

proteína de cápside de AAV9 (ver SEQ ID NO: 3), y el péptido de direccionamiento se inserta después del residuo 588 de la proteína de cápside de AAV9. En algunos aspectos, el péptido de direccionamiento se flanquea por secuencias ligadoras, y las secuencias ligadoras en cada lado de los péptidos de direccionamiento tienen dos o tres aminoácidos de longitud. En algunos aspectos, las secuencias ligadoras son AAA en el lado N-terminal del péptido de direccionamiento y AS en el lado C-terminal del péptido de direccionamiento. En algunos aspectos, las proteínas de cápside de AAV9 modificadas tienen una secuencia al menos 95 % idéntica a SEQ ID NO: 6.

En algunos aspectos, el péptido diana comprende una secuencia de hasta diez aminoácidos de longitud que tiene una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO: 7-13 o una secuencia de aminoácidos al menos 85 % idéntica a una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO: 8, 10, 12, 14, 16 y 18. En algunos aspectos, el péptido de direccionamiento es de siete aminoácidos de longitud. En algunos aspectos, la proteína de cápside de AAV tiene una secuencia de cualquiera de SEQ ID NO: 19-24.

En algunos aspectos, la estructura cerebral distinta es una estructura cerebral profunda, tal como el globus pallidus, putamen, cápsula interna, caudado, claustrum, sustancia negra, tálamo, hipocampo, subículo y núcleos cerebelosos profundos. En algunos aspectos, la estructura cerebral distinta es la corteza motora o corteza temporal. En algunos aspectos, la estructura cerebral distinta es neuronas de proyección V/VI de capa en la corteza motora y corteza premotora.

En una realización, en la presente se proporcionan ácidos nucleicos que comprenden una secuencia que codifica para la proteína de cápside modificada de cualquiera de las presentes realizaciones.

En una realización, se proporcionan en la presente virus adenoasociados recombinantes (rAAV) que comprenden la proteína de cápside modificada de cualquiera de las presentes realizaciones. En algunos aspectos, se proporcionan combinaciones de rAAV. En algunos aspectos, los rAAV no transducen células hepáticas.

En una realización, en la presente se proporcionan vectores virales que comprenden un ácido nucleico que codifica para la proteína de cápside modificada de cualquiera de las presentes realizaciones. En algunos aspectos, los vectores virales comprenden además una secuencia de ácido nucleico que codifica para un ácido nucleico de interés. En algunos

aspectos, el ácido nucleico de interés es un agente terapéutico. En algunos aspectos, el agente terapéutico es una enzima o una molécula de iARN.

En una realización, en la presente se proporcionan células que comprenden el vector viral de cualquiera de las presentes realizaciones. En algunos aspectos, la célula es una célula
5 de mamífero, tal como una célula humana. En algunos aspectos, la célula es *in vitro* o *in vivo*.

En una realización, en la presente se proporcionan composiciones farmacéuticas que comprenden el vector viral de las presentes realizaciones y un portador farmacéuticamente aceptable.

En una realización, se proporcionan en la presente métodos para administrar un
10 agente a una estructura cerebral distinta de un sujeto, que comprende administrar un virus AAV que codifica para el agente directamente al globus pallidus. En algunos aspectos, el virus AAV es un virus de las presentes realizaciones.

En una realización, en la presente se proporcionan métodos para administrar un agente a una estructura cerebral distinta de un sujeto, que comprende administrar un virus de
15 las presentes realizaciones al sujeto. En algunos aspectos, la estructura cerebral distinta es el globus pallidus, putamen, cápsula interna, caudado, claustrum, sustancia negra, corteza motora, ínsula, corteza temporal, tálamo, hipocampo, subículo y núcleos cerebelosos profundos. En algunos aspectos, la estructura cerebral distinta es neuronas de proyección V/VI de capa en la corteza motora y corteza premotora.

En algunos aspectos, el agente es un ARNsi, ARNsh, ARNmi, ARN no codificador, ARNinc, proteína terapéutica o sistema CRISPR. En algunos aspectos, la administración es al sistema nervioso central. En algunos aspectos, la administración es a una cisterna magna, un espacio intraventricular, un epéndima, un ventrículo cerebral, un espacio subaracnoideo y/o un espacio intratecal. En algunos aspectos, el ventrículo cerebral es el ventrículo lateral
20 rostral, y/o el ventrículo lateral caudal, y/o el ventrículo lateral derecho, y/o el ventrículo lateral izquierdo, y/o el ventrículo lateral rostral derecho, y/o el ventrículo lateral rostral izquierdo, y/o el ventrículo lateral caudal derecho, y/o el ventrículo lateral caudal izquierdo. En algunos aspectos, la administración es sistémica.

En algunos aspectos, se administra una pluralidad de partículas virales. En algunos
30 aspectos, el virus se administra a una dosis de aproximadamente 1×10^6 a aproximadamente 1×10^{18} genomas de vector por kilogramo (vg/kg). En algunos aspectos, el virus se administra

a una dosis de aproximadamente 1×10^7 - 1×10^{17} , aproximadamente 1×10^8 - 1×10^{16} , aproximadamente 1×10^9 - 1×10^{15} , aproximadamente 1×10^{10} - 1×10^{14} , aproximadamente 1×10^{10} - 1×10^{13} , aproximadamente 1×10^{10} - 1×10^{13} , aproximadamente 1×10^{10} - 1×10^{11} , aproximadamente 1×10^{11} - 1×10^{12} , aproximadamente 1×10^{12} - 1×10^{13} , o aproximadamente 1×10^{13} - 1×10^{14} vg/kg del paciente. En algunos aspectos, el sujeto es humano.

En una realización, en la presente se proporcionan métodos para tratar una enfermedad en un mamífero que comprenden administrar el virus de las presentes realizaciones al mamífero. En algunos aspectos, la enfermedad es una enfermedad neurodegenerativa. En algunos aspectos, la enfermedad neurodegenerativa es enfermedad de Huntington, esclerosis lateral amiotrófica (ALS), hemiplejía espástica hereditaria, esclerosis lateral primaria, atrofia muscular espinal, enfermedad de Kennedy, enfermedad de Alzheimer, una enfermedad de repetición de poliglutamina o enfermedad de Parkinson. En algunos aspectos, el mamífero es humano.

Otros objetos, características y ventajas de la presente invención llegarán a ser evidentes a partir de la siguiente descripción detallada de la invención. Se debe entender, sin embargo, que la descripción detallada de la invención y los ejemplos específicos, en tanto que indican realizaciones preferidas de la invención, se dan a manera de ilustración solamente, puesto que diferentes cambios y modificaciones dentro del espíritu y alcance de la invención llegarán a ser evidentes para aquellos expertos en la técnica a partir de esta descripción detallada.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Los siguientes dibujos forman parte de la presente especificación y se incluyen para demostrar adicionalmente ciertos aspectos de la presente invención. La invención se puede entender mejor con referencia a uno o más de estos dibujos en combinación con la descripción detallada de las realizaciones específicas presentadas en la presente.

Figura 1. Imágenes de hibridación *in situ* fluorescente (FISH), capturadas por microscopía confocal, para cada fluoróforo específico para localizar ARNm dentro de las células del Globus Pallidus (sitio de inyección). La señal positiva se ubica en el soma y núcleo de estas células. La mayoría de las células son positivas para las tres variantes de cápside.

Figura 2. Imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico dentro de las células del Globus Pallidus (sitio de inyección). La señal positiva se

ubica tanto en el soma como en las proyecciones de estas células. Específicamente, aunque el ARNm de mRuby3 estuvo presente en el soma de las neuronas de globo pálido por FISH, hay poca o ninguna señal positiva para la proteína en esta placa de tejido.

5 Figura 3. Imágenes de epifluorescencia, de diferentes regiones del Globus Pallidus que las imágenes de la figura 2, localizando la proteína de cada fluoróforo específico dentro de las células del Globus Pallidus caudal (sitio de inyección). La señal positiva se ubica tanto en el soma como en las proyecciones de estas células. En esta sección de tejido, está presente la señal de mRuby.

10 Figura 4. Imágenes de hibridación *in situ* fluorescente (FISH), capturadas por microscopía confocal, para cada fluoróforo específico para localizar ARNm dentro de las células del putamen. La señal positiva de ARNm se ubica en el soma y núcleos de estas células. La variante de cápside que expresa mNeonGreen (mNG) muestra la mayor transducción de células en el putamen.

15 Figura 5. Imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico dentro de las células del putamen. La señal positiva se ubica tanto en el soma como en las proyecciones de estas células. Similar a ARNm de FISH, la nueva variante de cápside que expresa mNeonGreen (mNG) muestra la señal más positiva tanto en células como en proyecciones.

20 Figura 6. Imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico dentro de la cápsula interna. Hay una separación sorprendente de la señal positiva de mTFP1 y fluorescencia NeonGreen.

 Figura 7. Imágenes de hibridación *in situ* fluorescente (FISH), capturadas por microscopía confocal, para cada fluoróforo específico para localizar ARNm dentro de las células del caudado. La señal positiva de ARNm se ubica en el soma de estas células.

25 Figura 8. Imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico dentro de las células del caudado. La señal positiva se ubica tanto en el soma como en las proyecciones de estas células.

30 Figura 9. Imágenes de hibridación *in situ* fluorescente (FISH), capturadas por microscopía confocal, para cada fluoróforo específico para localizar ARNm dentro de las células del claustrum. La señal positiva de ARNm se ubica en el soma de estas células.

Figura 10. Imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico dentro del claustrum. Una vez más, se observa una señal positiva de fluorescencia tanto en el soma celular como en las proyecciones celulares.

5 Figura 11. Imágenes de hibridación *in situ* fluorescente (FISH), capturadas por microscopía confocal, para cada fluoróforo específico para localizar ARNm dentro de las células de la sustancia negra. La señal positiva de ARNm se ubica en el soma y núcleo de estas células.

10 Figura 12. Imágenes de hibridación *in situ* fluorescente (FISH), capturadas por microscopía confocal, para cada fluoróforo específico para localizar ARNm dentro de las células de la sustancia negra, pero de diferentes áreas con respecto a las imágenes en la figura 11. La señal positiva de ARNm se ubica en el soma y núcleo de estas células.

Figura 13. Imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico dentro de la sustancia negra. Se observa una señal positiva de fluorescencia tanto en el soma celular como en las proyecciones celulares.

15 Figura 14. Imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico dentro de la sustancia negra, pero de diferentes áreas con respecto a las imágenes en la figura 13. Se observa una señal positiva de fluorescencia tanto en el soma celular como en las proyecciones celulares.

20 Figura 15. Imágenes de hibridación *in situ* fluorescente (FISH), capturadas por microscopía confocal, para cada fluoróforo específico para localizar ARNm dentro de las células de corteza motora. La señal positiva de ARNm se ubica en el soma y núcleo de estas células.

25 Figura 16. Imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico dentro de las células de la corteza motora. La señal positiva se ubica tanto en el soma como en las proyecciones de estas células. Sorprendentemente, la señal de ARNm estuvo presente tanto para mTFP1 como para mNeonGreen por FISH; sin embargo, es mTFP1 la que se expresa predominantemente como se muestra por epifluorescencia.

30 Figura 17. En la corteza cingulada, se detectaron algunas células positivas para el nuevo ARNm de cápside por FISH. Sin embargo, se detectaron células positivas para mTFP1 de amplia difusión por microscopía de epifluorescencia.

Figura 18. Imágenes de hibridación *in situ* fluorescente (FISH), capturadas por microscopía confocal, para cada fluoróforo específico para localizar ARNm dentro de las células de la ínsula de corteza temporal. La señal positiva de ARNm se ubica en el soma (citoplasma y núcleo) de estas células.

- 5 Figura 19. Imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico dentro de las células de la ínsula de la corteza temporal. La señal positiva se ubica tanto en el soma como en las proyecciones de estas células.

- 10 Figura 20. Imágenes de hibridación *in situ* fluorescente (FISH), capturadas por microscopía confocal, para cada fluoróforo específico para localizar ARNm dentro de las células del tálamo. La señal positiva de ARNm se ubica en el soma (citoplasma y núcleo) de estas células.

Figura 21. Imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico dentro de las células del tálamo. La señal positiva se ubica tanto en el soma como en las proyecciones de estas células.

- 15 Figuras 22A-D. Imágenes de epifluorescencia de las tres cápsides novedosas superiores administradas al globus pallidus del macaco rhesus. La expresión de las tres es alta en el putamen. A) Imagen que muestra la expresión de AAV1.THTD.mTFP1; B) Imagen que muestra la expresión de AAV2.SGGR.mNeonGreen; C) Imagen que muestra la expresión de AAV2.SKGQ.mRuby3; D) Imagen fusionada de A-C.

- 20 Figuras 23A-D. Imágenes de epifluorescencia de las tres cápsides novedosas superiores administradas al globus pallidus del macaco rhesus. La expresión de AAV1.THTD.mTFP1 es extremadamente alta en la corteza motora. A) Imagen que muestra la expresión de AAV1.THTD.mTFP1; B) Imagen que muestra la expresión de AAV2.SGGR.mNeonGreen; C) Imagen que muestra la expresión de AAV2.SKGQ.mRuby3;
25 D) Imagen fusionada de A-C.

- Figuras 24A-D. Imágenes de epifluorescencia de las tres cápsides novedosas superiores administradas al globus pallidus del macaco rhesus. La expresión de las tres se localiza en el hipocampo (CA1) y el subículo adyacente. A) Imagen que muestra la expresión de AAV1.THTD.mTFP1; B) Imagen que muestra la expresión de AAV2.SGGR.mNeonGreen;
30 C) Imagen que muestra la expresión de AAV2.SKGQ.mRuby3; D) Imagen fusionada de A-C.

Figura 25. Imágenes de hibridación *in situ* fluorescente (FISH), capturadas por

microscopía confocal, para cada fluoróforo específico para localizar ARNm dentro de las células del Globus Pallidus (sitio de inyección) de un segundo macaco rhesus. La señal positiva se ubica en el soma y núcleo de estas células.

5 Figura 26. Imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico dentro de las células del Globus Pallidus (sitio de inyección) de un segundo macaco rhesus. La señal positiva se ubica tanto en el soma como en las proyecciones de estas células.

10 Figura 27. Montajes enteros de cerebros y médula espinal de ratones inyectados con virus en la cápsula interna/globus pallidus medial. Los números de ratón están en el lado izquierdo, sus números de ID se codifican por colores para reflejar qué hemisferio recibió qué cápside que expresa el fluoróforo.

Figura 28. Montajes enteros de cerebros y médula espinal de ratones inyectados con virus en el tálamo. Los números de ratón están en el lado izquierdo, sus números de ID se codifican por colores para reflejar qué hemisferio recibió qué cápside que expresa el fluoróforo.

15 Figura 29. Montajes enteros de cerebros y médula espinal de ratones inyectados con virus en los núcleos cerebelosos profundos. Los números de ratón están en el lado izquierdo, sus números de ID se codifican por colores para reflejar qué hemisferio recibió qué cápside que expresa el fluoróforo.

20 Figura 30. Cuatro cortes de cerebro coronales de un ratón inyectado con AAV2.SGGR.mNeonGreen a la cápsula interna izquierda y Globus Pallidus medial y AAV2.SKGO.mRuby3 a la derecha.

Figura 31. Tres cortes de cerebro coronales de un ratón inyectado con AAV2.SKGO.mRuby3 a la cápsula interna izquierda y globus pallidus medial y AAV1.THTD.mTFP1 a la derecha.

25 Figura 32. Cuatro cortes de cerebro coronales de un ratón inyectado con AAV2.SGGR.mNeonGreen al tálamo izquierdo y AAV2.SKGO.mRuby3 a la derecha. Una sección sagital individual del cerebelo y cerebro posterior.

30 Figura 33. Cuatro cortes de cerebro coronales de un ratón inyectado con AAV2.SKGO.mRuby3 al tálamo izquierdo y AAV1.THTD.mTFP1 a la derecha. Las secciones sagitales del cerebelo y tronco encefálico muestran neuronas en los núcleos cerebelosos profundos positivas para AAV1 THTD.mTFP 1.

Figura 34. Cortes sagitales del cerebelo inyectados con AAV1.THTD.mTFP1 a los núcleos cerebelosos profundos izquierdos y AAV2.SGGR.mNeonGreen a la derecha.

Figura 35. Corte sagital del cerebro lateral inyectado con AAV1.THTD.mTFP1 a los núcleos cerebelosos profundos izquierdos y AAV2.SGGR.mNeonGreen a la derecha.

5 Figura 36. Corte sagital de todo el cerebro inyectado con AAV1.THTD.mTFP1 a los núcleos cerebelosos profundos izquierdos y AAV2.SGGR.mNeonGreen a la derecha. Hay neuronas y proyecciones en todo el cerebro positivas para AAV1.THTD.mTFP1.

10 Figura 37. Corte sagital de todo el cerebro inyectado con AAV2.SKGO.mRuby3 a los núcleos cerebelosos profundos izquierdos y AAV1.THTD.mTFP1 a la derecha. Hay neuronas y proyecciones en todo el cerebro positivas para AAV1.THTD.mTFP1.

15 Figuras 38A-38C: Representaciones de mapas de calor del enriquecimiento de UMI de variantes de cápside de AAV en regiones cerebrales muestreadas. (Figura 38A) Sec. de amplicón que abarca el área de la inserción de péptido de AAV permitió recolectar conteos de identificador molecular único (UMI) de variantes de cápside administradas en varios tejidos muestreados. El amplicón en A se generó a partir del ADN de cápside recuperado de los tejidos transducidos. El % de abundancia de cada variante de cápside se representa para cada tejido indicado. Se indican las cápsides llevadas a la validación de fluorescencia. (Figura 38B) Los mismos conteos de UMI de sec. de amplicón de ADN mostrados en A se normalizaron al calcular el enriquecimiento con respecto a la entrada. Los candidatos de 20 cápside más altamente representados son aquellos que muestran el mayor enriquecimiento con respecto a su abundancia en la mezcla de cápsides de entrada. (Figura 38C) El % de abundancia de conteos de UMI obtenidos a partir de ARN por sec. de amplicón.

25 Figura 39. Imágenes de epifluorescencia para localizar AAV1.THTD.mTFP1 en un gran número de neuronas de proyección V/VI de capa cortical en la corteza motora y corteza premotora del macaco rhesus.

Figura 40. El ADNg de AAV y expresión de proteínas estaban por debajo del umbral de detección en el hígado del macaco rhesus.

DESCRIPCIÓN DETALLADA

30 En la presente se proporcionan vectores virales que comprenden cada uno una cápside modificada, en donde la cápside modificada comprende al menos una secuencia de

aminoácidos que dirige el vector viral a una estructura cerebral profunda. En ciertas realizaciones, la estructura cerebral es el globus pallidus, putamen, cápsula interna, caudado, claustrum, sustancia negra, corteza motora, ínsula, corteza temporal, tálamo, hipocampo, subículo y núcleos cerebelosos profundos. En ciertas realizaciones, el AAV es AAV1, AAV2 o AAV9. Una secuencia de proteína de cápside de AAV1 de referencia tipo silvestre de ejemplo se proporciona en SEQ ID NO: 1. Una secuencia de proteína de cápside de AAV2 de referencia tipo silvestre de ejemplo se proporciona en SEQ ID NO: 2. Una secuencia de proteína de cápside de AAV9 de referencia tipo silvestre de ejemplo se proporciona en SEQ ID NO: 3. En ciertos aspectos, el péptido de direccionamiento se inserta en la posición 590 de cápside de AAV1, posición 587 de cápside de AAV2 o posición 588 de cápside de AAV9. Una secuencia de proteína de cápside de AAV1 modificada de ejemplo se proporciona en SEQ ID NO: 4, que muestra la inserción de péptido de direccionamiento después de la posición 590 como SSAX₇AS, donde la SSA principal y la AS posterior son secuencias ligadoras y X₇ representa el péptido de direccionamiento. Una secuencia de proteína de cápside de AAV2 modificada de ejemplo se proporciona en SEQ ID NO: 5, que muestra la inserción de péptido de direccionamiento después de la posición 587 como AAAX₇AA, donde la AAA principal y la AA posterior son secuencias ligadoras y X₇ representa el péptido de direccionamiento. Una secuencia de proteína de cápside de AAV9 modificada de ejemplo se proporciona en SEQ ID NO: 6, que muestra la inserción de péptido de direccionamiento después de la posición 588 como AAAX₇AS, donde la AAA principal y la AS posterior son secuencias ligadoras y X₇ representa el péptido de direccionamiento.

I. Vectores de virus adenoasociados (AAV)

El virus adenoasociado (AAV) es un pequeño virus no patógeno de la familia parvoviridae. Hasta la fecha, se han identificado numerosos AAV serológicamente distintos, y se han aislado más de una docena de humanos o primates. El AAV se distingue de otros miembros de esta familia por su dependencia de un virus auxiliar para replicación.

Los genomas de AAV pueden existir en un estado extracromosómico sin integrarse en los genomas celulares de hospedador; poseen una amplia gama de hospedadores; transducen células tanto divisoras como no divisoras *in vitro* e *in vivo* y mantienen altos niveles de expresión de los genes transducidos. Las partículas virales de AAV son estables al calor; resistentes a solventes, detergentes, cambios en pH y temperatura; y se pueden purificar en columna y/o concentrar en gradientes de CsCl o por otros medios. El genoma de AAV comprende un ácido desoxirribonucleico de hebra individual (ADNss), ya sea detectado

positivo o negativo. El genoma de aproximadamente 4,7 kb de AAV consiste en un segmento de ADN de hebra individual de polaridad más o menos. Los extremos del genoma son repeticiones terminales invertidas cortas (ITR) que pueden plegarse en estructuras de horquilla y servir como el origen de la replicación de ADN viral.

5 Un “genoma” de AAV se refiere a una secuencia de ácido nucleico recombinante que finalmente se empaqueta o encapsula para formar una partícula de AAV. Una partícula de AAV frecuentemente comprende un genoma de AAV empaquetado con proteínas de cápside de AAV. En los casos donde se usan plásmidos recombinantes para construir o fabricar vectores recombinantes, el genoma de vector de AAV no incluye la porción del “plásmido” que
10 no corresponde a la secuencia de genoma de vector del plásmido recombinante. Esta porción de genoma no vectorial del plásmido recombinante se refiere como la “cadena principal de plásmido”, que es importante para clonación y amplificación del plásmido, un proceso que se necesita para la propagación y producción del plásmido, pero no se empaqueta o encapsula en sí mismo en partículas virales. Por lo tanto, un “genoma” de vector de AAV se refiere a
15 ácido nucleico que se empaqueta o encapsula por proteínas de cápside de AAV.

El virión (partícula) de AAV es una partícula icosaédrica sin envoltura de aproximadamente 25 nm de diámetro que comprende una cápside de AAV. La partícula de AAV comprende una simetría icosaédrica compuesta por tres proteínas de cápside relacionadas, VP1, VP2 y VP3, que interactúan conjuntamente para formar la cápside. El
20 genoma de la mayoría de los AAV nativos frecuentemente contiene dos marcos de lectura abiertos (ORF), algunas veces referidos como ORF izquierdo y ORF derecho. El ORF derecho frecuentemente codifica para las proteínas de cápside VP1, VP2 y VP3. Estas proteínas frecuentemente se encuentran en una relación de 1:1:10 respectivamente, pero pueden estar en relaciones variadas, y todas se derivan del ORF derecho. Las proteínas de cápside VP1,
25 VP2 y VP3 difieren entre sí por el uso de empalme alternativo y un codón de inicio inusual. El análisis de delección ha mostrado que la remoción o alteración de VP1 que se traduce a partir de un mensaje empalmado de manera alternativa da por resultado un rendimiento reducido de partículas infecciosas. Las mutaciones dentro de la región codificadora de VP3 dan por resultado la falta de producción de ADN de progenie de hebra individual o partículas
30 infecciosas. En ciertas realizaciones, el genoma de una partícula de AAV codifica para uno, dos o los tres polipéptidos VP1, VP2 y VP3.

El ORF izquierdo frecuentemente codifica para las proteínas Rep no estructurales, Rep 40, Rep 52, Rep 68 y Rep 78, que se implican en la regulación de la replicación y transcripción

además de la producción de genomas de progenie de hebra individual. Dos de las proteínas Rep se han asociado con la integración preferencial de genomas de AAV en una región del brazo q del cromosoma humano 19. Se ha mostrado que Rep68/78 posee actividad de unión a NTP, así como actividades de ADN y ARN helicasa. Algunas proteínas Rep poseen una
5 señal de localización nuclear, así como varios sitios de fosforilación potenciales. En ciertas realizaciones, el genoma de un AAV (por ejemplo, un rAAV) codifica para algunas o todas las proteínas Rep. En ciertas realizaciones, el genoma de un AAV (por ejemplo, un rAAV) no codifica para las proteínas Rep. En ciertas realizaciones, una o más de las proteínas Rep se pueden administrar en trans y por lo tanto, no se incluyen en una partícula de AAV que
10 comprende un ácido nucleico que codifica para un polipéptido.

Los extremos del genoma de AAV comprenden repeticiones terminales invertidas cortas (ITR) que tienen el potencial de plegarse en estructuras de horquilla en forma de T que sirven como el origen de la replicación de ADN viral. Por consiguiente, el genoma de un AAV comprende una o más (por ejemplo, un par de) secuencias de ITR que flanquean un genoma
15 de ADN viral de hebra individual. Las secuencias de ITR frecuentemente tienen una longitud de aproximadamente 145 bases cada una. Dentro de la región ITR, se han descrito dos elementos que se cree que son fundamentales 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 una conformación lineal o en horquilla. Se cree
20 que esta unión posiciona a Rep68/78 para escisión en el trs que se presenta de una manera específica de sitio y hebra. Además de su papel en la replicación, estos dos elementos parecen ser fundamentales para la integración viral. Dentro del locus de integración de cromosoma 19 se encuentra un sitio de unión a Rep con un trs adyacente. Se ha mostrado que estos elementos son funcionales y necesarios para la integración específica de locus.

El término “recombinante”, como un modificador de vector, tal como viral recombinante, por ejemplo, vectores de lenti o parvovirus (por ejemplo, AAV), así como un modificador de secuencias tal como secuencias de ácido nucleico y polipéptidos recombinantes, significa que las composiciones se han manipulado (es decir, modificado) de una manera que en general no se presenta en la naturaleza. Un ejemplo particular de un
25 vector recombinante, tal como un vector de AAV, retroviral o lentiviral sería cuando se inserta una secuencia de ácido nucleico que normalmente no está presente en el genoma viral tipo silvestre dentro del genoma viral. Un ejemplo de una secuencia de ácido nucleico recombinante sería cuando un ácido nucleico (por ejemplo, gen) codifica para un ARN inhibidor clonado en un vector, con o sin regiones 5', 3' y/o de intrón que el gen se asocia
30

normalmente dentro del genoma viral. Aunque el término “recombinante” no siempre se usa en la presente con referencia a vectores, tal como vectores virales, así como secuencias tal como polinucleótidos, las formas “recombinantes” que incluyen secuencias de ácido nucleico, polinucleótidos, transgenes, etc. se incluyen expresamente a pesar de cualquier omisión de este tipo.

Un “vector” viral recombinante se deriva del genoma tipo silvestre de un virus al usar métodos moleculares para remover parte del genoma tipo silvestre del virus y reemplazar con un ácido nucleico no nativo, tal como una secuencia de ácido nucleico. Habitualmente, por ejemplo, para AAV, una o ambas secuencias de repetición terminal invertida (ITR) del genoma de AAV se retienen en el vector de AAV recombinante. Un vector viral “recombinante” (por ejemplo, rAAV) se distingue de un genoma viral (por ejemplo, AAV), puesto que parte del genoma viral se ha reemplazado con una secuencia no nativa con respecto al ácido nucleico genómico viral, tal como un ácido nucleico que codifica para un transactivador o ácido nucleico que codifica para un ARN inhibidor o ácido nucleico que codifica para una proteína terapéutica. Por lo tanto, la incorporación de estas secuencias de ácido nucleico no nativas define el vector viral como un vector “recombinante”, que en el caso de AAV se puede referir como un “vector de rAAV”.

En ciertas realizaciones, un AAV (por ejemplo, un rAAV) comprende dos ITR. En ciertas realizaciones, un AAV (por ejemplo, un rAAV) comprende un par de ITR. En ciertas realizaciones, un AAV (por ejemplo, un rAAV) comprende un par de ITR que flanquean (es decir, están en cada extremo 5' y 3') de una secuencia de ácido nucleico que al menos codifica para un polipéptido que tiene función o actividad.

Un vector de AAV (por ejemplo, vector de rAAV) se puede empaquetar y se refiere en la presente como una “partícula de AAV” para la posterior infección (transducción) de una célula, *ex vivo*, *in vitro* o *in vivo*. Cuando un vector de AAV recombinante se encapsula o empaqueta en una partícula de AAV, la partícula también se puede referir como una “partícula de rAAV”. En ciertas realizaciones, una partícula de AAV es una partícula de rAAV. Una partícula de rAAV frecuentemente comprende un vector de rAAV, o una porción del mismo. Una partícula de rAAV puede ser una o más partículas de rAAV (por ejemplo, una pluralidad de partículas de AAV). Las partículas de rAAV habitualmente comprenden proteínas que encapsulan o empaquetan el genoma de vector de rAAV (por ejemplo, proteínas de cápside). Se señala que la referencia a un vector de rAAV también se puede usar para hacer referencia a una partícula de rAAV.

Se puede usar cualquier partícula de AAV adecuada (por ejemplo, partícula de rAAV) para un método o uso en la presente. Una partícula de rAAV, y/o genoma comprendido en la misma, se puede derivar de cualquier serotipo o cepa adecuada de AAV. Una partícula de rAAV, y/o genoma comprendido en la misma, se puede derivar de dos o más serotipos o cepas de AAV. Por consiguiente, un rAAV puede comprender proteínas y/o ácidos nucleicos, o porciones de los mismos, de cualquier serotipo o cepa de AAV, en donde la partícula de AAV es adecuada para infección y/o transducción de una célula de mamífero. Los ejemplos no limitantes de serotipos de AAV incluyen AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV-rh74, AAV-rh10 y AAV-2i8.

En ciertas realizaciones, una pluralidad de partículas de rAAV comprende partículas de, o derivadas de, la misma cepa o serotipo (o subgrupo o variante). En ciertas realizaciones, una pluralidad de partículas de rAAV comprende una mezcla de dos o más partículas de rAAV diferentes (por ejemplo, de diferentes serotipos y/o cepas).

Como se usa en la presente, el término “serotipo” es una distinción usada para referirse a un AAV que tiene una cápside que es serológicamente distinta de otros serotipos de AAV. El carácter distintivo serológico se determina con base en la falta de reactividad cruzada entre anticuerpos para un AAV en comparación con otro AAV. Estas diferencias de reactividad cruzada usualmente se deben a diferencias en las secuencias de proteína de cápside/determinantes antigénicos (por ejemplo, debido a diferencias de secuencia de VP1, VP2 y/o VP3 de serotipos de AAV). A pesar de la posibilidad de que las variantes de AAV que incluyen variantes de cápside pueden no ser serológicamente distintas de un AAV de referencia u otro serotipo de AAV, difieren por al menos un residuo de nucleótido o aminoácido en comparación con el serotipo de AAV de referencia u otro.

En ciertas realizaciones, un vector de rAAV basado en un primer genoma de serotipo corresponde al serotipo de una o más de las proteínas de cápside que empaquetan el vector. Por ejemplo, el serotipo de uno o más ácidos nucleicos de AAV (por ejemplo, ITR) que comprende el genoma de vector de AAV corresponde al serotipo de una cápside que comprende la partícula de rAAV.

En ciertas realizaciones, un genoma de vector de rAAV se puede basar en un genoma de serotipo de AAV (por ejemplo, AAV2) distinto del serotipo de una o más de las proteínas de cápside de AAV que empaquetan el vector. Por ejemplo, un genoma de vector de rAAV puede comprender ácidos nucleicos derivados de AAV2 (por ejemplo, ITR), en tanto que al menos una o más de las tres proteínas de cápside se derivan de un serotipo diferente, por

ejemplo, un serotipo AAV1, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, Rh10, Rh74 o AAV-2i8 o variante de los mismos.

En ciertas realizaciones, una partícula de rAAV o un genoma de vector de la misma relacionado con un serotipo de referencia tiene un polinucleótido, polipéptido o subsecuencia del mismo que comprende o consiste en una secuencia al menos 60 % o más (por ejemplo, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,1 %, 99,2 %, 99,3 %, 99,4 %, 99,5 %, etc.) idéntica a un polinucleótido, polipéptido o subsecuencia de una partícula de AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, Rh10, Rh74 o AAV-2i8. En realizaciones particulares, una partícula de rAAV o un genoma de vector de la misma relacionado con un serotipo de referencia tiene una secuencia de ITR o cápside que comprende o consiste en una secuencia al menos 60 % o más (por ejemplo, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,1 %, 99,2 %, 99,3 %, 99,4 %, 99,5 %, etc.) idéntica a una secuencia de ITR o cápside de un serotipo AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, Rh10, Rh74 o AAV-2i8.

En ciertas realizaciones, un método en la presente comprende el uso, administración o suministro de una partícula de rAAV1, rAAV2, rAAV3, rAAV4, rAAV5, rAAV6, rAAV7, rAAV8, rAAV9, rAAV10, rAAV11, rAAV12, rRh10, rRh74 o rAAV-2i8.

En ciertas realizaciones, un método en la presente comprende el uso, administración o suministro de una partícula de rAAV2. En ciertas realizaciones, una partícula de rAAV2 comprende una cápside de AAV2. En ciertas realizaciones, una partícula de rAAV2 comprende una o más proteínas de cápside (por ejemplo, VP1, VP2 y/o VP3) que son al menos 60 %, 65 %, 70 %, 75 % o más idénticas, por ejemplo, 80 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,1 %, 99,2 %, 99,3 %, 99,4 %, 99,5 %, etc., hasta 100 % idénticas a una proteína de cápside correspondiente de una partícula de AAV2 nativa o tipo silvestre. En ciertas realizaciones, una partícula de rAAV2 comprende proteínas de cápside de VP1, VP2 y VP3 que son al menos 75 % o más idénticas, por ejemplo, 80 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,1 %, 99,2 %, 99,3 %, 99,4 %, 99,5 %, etc., hasta 100 % idénticas a una proteína de cápside correspondiente de una partícula de AAV2 nativa o tipo silvestre. En ciertas realizaciones, una partícula de rAAV2 es una variante de una partícula de AAV2 nativa o tipo silvestre. En algunos aspectos, una o más proteínas de cápside de una variante de AAV2 tienen 1, 2, 3, 4, 5, 5-10, 10-15, 15-20 o más sustituciones de aminoácidos en

comparación con la o las proteínas de cápside de una partícula de AAV2 nativa o tipo silvestre.

En ciertas realizaciones, una partícula de rAAV9 comprende una cápside de AAV9. En ciertas realizaciones, una partícula de rAAV9 comprende una o más proteínas de cápside (por ejemplo, VP1, VP2 y/o VP3) que son al menos 60 %, 65 %, 70 %, 75 % o más idénticas, por ejemplo, 80 %, 85 %, 85 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,1 %, 99,2 %, 99,3 %, 99,4 %, 99,5 %, etc., hasta 100 % idénticas a una proteína de cápside correspondiente de una partícula de AAV9 nativa o tipo silvestre. En ciertas realizaciones, una partícula de rAAV9 comprende proteínas de cápside de VP1, VP2 y VP3 que son al menos 75 % o más idénticas, por ejemplo, 80 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,1 %, 99,2 %, 99,3 %, 99,4 %, 99,5 %, etc., hasta 100 % idénticas a una proteína de cápside correspondiente de una partícula de AAV9 nativa o tipo silvestre. En ciertas realizaciones, una partícula de rAAV9 es una variante de una partícula de AAV9 nativa o tipo silvestre. En algunos aspectos, una o más proteínas de cápside de una variante de AAV9 tienen 1, 2, 3, 4, 5, 5-10, 10-15, 15-20 o más sustituciones de aminoácidos en comparación con la o las proteínas de cápside de una partícula de AAV9 nativa o tipo silvestre.

En ciertas realizaciones, una partícula de rAAV comprende una o dos ITR (por ejemplo, un par de ITR) que son al menos 75 % o más idénticas, por ejemplo, 80 %, 85 %, 85 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,1 %, 99,2 %, 99,3 %, 99,4 %, 99,5 %, etc., hasta 100 % idénticas a las correspondientes ITR de un AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV-rh74, AAV-rh10 o AAV-2i8 nativo o tipo silvestre, siempre que retengan una o más funciones de ITR deseadas (por ejemplo, capacidad de formar una horquilla, que permite la replicación de ADN; integración del ADN de AAV en un genoma de célula hospedadora; y/o empaquetamiento, si se desea).

En ciertas realizaciones, una partícula de rAAV2 comprende una o dos ITR (por ejemplo, un par de ITR) que son al menos 75 % o más idénticas, por ejemplo, 80 %, 85 %, 85 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,1 %, 99,2 %, 99,3 %, 99,4 %, 99,5 %, etc., hasta 100 % idénticas a las ITR correspondientes de una partícula de AAV2 nativa o tipo silvestre, siempre que retengan una o más funciones de ITR deseadas (por ejemplo, capacidad para formar una horquilla, que permite la replicación de ADN; integración del ADN de AAV en un genoma de célula hospedadora; y/o empaquetamiento, si se desea).

En ciertas realizaciones, una partícula de rAAV9 comprende una o dos ITR (por ejemplo, un par de ITR) que son al menos 75 % o más idénticas, por ejemplo, 80 %, 85 %, 85 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,1 %, 99,2 %, 99,3 %, 99,4 %, 99,5 %, etc., hasta 100 % idénticas a las ITR correspondientes de una partícula de AAV2 nativa o tipo silvestre, siempre que retengan una o más funciones de ITR deseadas (por ejemplo, capacidad para formar una horquilla, que permite la replicación de ADN; integración del ADN de AAV en un genoma de célula hospedadora; y/o empaquetamiento, si se desea).

Una partícula de rAAV puede comprender una ITR que tiene cualquier número adecuado de repeticiones de "GAGC". En ciertas realizaciones, una ITR de una partícula de AAV2 comprende 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 o más repeticiones de "GAGC". En ciertas realizaciones, una partícula de rAAV2 comprende una ITR que comprende tres repeticiones de "GAGC". En ciertas realizaciones, una partícula de rAAV2 comprende una ITR que tiene menos de cuatro repeticiones de "GAGC". En ciertas realizaciones, una partícula de rAAV2 comprende una ITR que tiene más de cuatro repeticiones de "GAGC". En ciertas realizaciones, una ITR de una partícula de rAAV2 comprende un sitio de unión a Rep en donde el cuarto nucleótido en las primeras dos repeticiones de "GAGC" es una C en lugar de una T.

Se puede incorporar una longitud de ADN adecuada de ejemplo en vectores de rAAV para empaquetamiento/encapsidación en una partícula de rAAV de aproximadamente 5 kilobases (kb) o menos. En realizaciones particulares, la longitud de ADN es menor de aproximadamente 5 kb, menor de aproximadamente 4,5 kb, menor de aproximadamente 4 kb, menor de aproximadamente 3,5 kb, menor de aproximadamente 3 kb o menor de aproximadamente 2,5 kb.

Los vectores de rAAV que incluyen una secuencia de ácido nucleico que dirige la expresión de un iARN o polipéptido se pueden generar usando técnicas recombinantes adecuadas conocidas en la técnica (por ejemplo, ver Sambrook et al., 1989). Los vectores de AAV recombinantes se empaquetan habitualmente en partículas de AAV competentes para transducción y se propagan usando un sistema de empaquetamiento viral de AAV. Una partícula de AAV competente para transducción es capaz de unirse a e ingresar en una célula de mamífero y posteriormente administrar una carga de ácido nucleico (por ejemplo, un gen heterólogo) al núcleo de la célula. Por lo tanto, una partícula de rAAV intacta que es competente para transducción se configura para transducir una célula de mamífero. Una partícula de rAAV configurada para transducir una célula de mamífero frecuentemente no es

competente para replicación y requiere maquinaria proteica adicional para autorreplicarse. Por lo tanto, una partícula de rAAV que se configura para transducir una célula de mamífero se modifica para unirse e ingresar a una célula de mamífero y administrar un ácido nucleico a la célula, en donde el ácido nucleico para administración frecuentemente se coloca entre un par de ITR de AAV en el genoma de rAAV.

Las células hospedadoras adecuadas para producir partículas de AAV competentes para transducción incluyen, pero no se limitan a, microorganismos, células de levadura, células de insecto y células de mamífero que se pueden usar, o se han usado, como receptores de vectores de rAAV heterólogos. Se pueden usar células de la línea de células humanas estables, HEK293 (fácilmente disponible a través de, por ejemplo, la American Type Culture Collection con el número de acceso ATCC CRL1573). En ciertas realizaciones, una línea de células de riñón embrionarias humanas modificadas (por ejemplo, HEK293), que se transforma con fragmentos de ADN de adenovirus tipo 5, y expresa los genes adenovirales E1a y E1b se usa para generar partículas de AAV recombinantes. La línea de células HEK293 modificada se transfecta fácilmente y proporciona una plataforma particularmente conveniente en la que producir partículas de rAAV. Los métodos para generar partículas de AAV de título alto capaces de transducir células de mamífero se conocen en la técnica. Por ejemplo, la partícula de AAV se puede hacer como se expone en Wright, 2008 y Wright, 2009.

En ciertas realizaciones, las funciones auxiliares de AAV se introducen en la célula hospedadora por transfección de la célula hospedadora con una construcción auxiliar de AAV ya sea antes de, o simultáneamente con, la transfección de un vector de expresión de AAV. Por lo tanto, algunas veces se usan construcciones auxiliares de AAV para proporcionar al menos la expresión transitoria de los genes rep y/o cap de AAV para complementar las funciones de AAV faltantes necesarias para la transducción productiva de AAV. Las construcciones auxiliares de AAV frecuentemente carecen de ITR de AAV y no pueden replicarse ni envasarse por sí mismas. Estas construcciones pueden estar en la forma de un plásmido, fago, transposón, cósmido, virus o virión. Se han descrito varias construcciones auxiliares de AAV, tal como los plásmidos comúnmente usados pAAV/Ad y pIM29+45 que codifican para los productos de expresión Rep y Cap. Se conocen varios vectores diferentes que codifican para productos de expresión de Rep y/o Cap.

Un “vector de expresión” es un vector especializado que contiene un gen o secuencia de ácido nucleico con las regiones reguladoras necesarias para la expresión en una célula hospedadora. Un vector de expresión puede contener al menos un origen de replicación para

la propagación en una célula y opcionalmente elementos adicionales, tal como una secuencia de ácido nucleico heteróloga, elemento de control de expresión (por ejemplo, un promotor, intensificador), intrón, ITR y señal de poliadenilación.

II. Agentes terapéuticos

- 5 En algunas realizaciones, los métodos de transferencia génica viral se pueden usar para introducir ácidos nucleicos en células de mamífero o tejidos diana. Estos métodos se pueden usar para administrar ácidos nucleicos que codifican para ARN inhibidores, ARN no codificadores y/o proteínas terapéuticas a células en cultivo o en un organismo hospedador.

A. ARN inhibidores

- 10 “Interferencia de ARN (iARN)” es el proceso de silenciamiento génico postraduccional específico de secuencia iniciado por ARNsi. Durante la iARN, el ARNsi induce la degradación del ARNm diana con la consiguiente inhibición específica de secuencia de la expresión génica.

- Un “ARN inhibidor”, “ARNi”, “ARN de interferencia pequeño” o “ARN interferente pequeño” o molécula de “ARNsi”, “ARN de horquilla pequeño” o molécula de “ARNsh” o
- 15 “miARN” es un dúplex de ARN de nucleótidos que se dirige a una secuencia de ácido nucleico de interés. Como se usa en la presente, el término “ARNsi” es un término genérico que abarca el subconjunto de ARNsh y miARN. Un “dúplex de ARN” se refiere a la estructura formada por el emparejamiento complementario entre dos regiones de una molécula de ARN. El ARNsi se “dirige” a un gen en el que la secuencia de nucleótidos de la porción de dúplex del ARNsi es
- 20 complementaria a una secuencia de nucleótidos del gen dirigido. En ciertas realizaciones, los ARNsi se dirigen a la secuencia que codifica para huntingtina. En algunas realizaciones, la longitud del dúplex de ARNsi es menor de 30 pares de bases. En algunas realizaciones, el dúplex puede ser de 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11 o 10 pares de bases de longitud. En algunas realizaciones, la longitud del dúplex es de 19 a 25
- 25 pares de bases de longitud. En cierta realización, la longitud del dúplex es de 19 o 21 pares de bases de longitud. La porción de dúplex de ARN del ARNsi puede ser parte de una estructura de horquilla. Además de la porción de dúplex, la estructura de horquilla puede contener una porción de bucle colocada entre las dos secuencias que forman el dúplex. El bucle puede variar en longitud. En algunas realizaciones, el bucle es de 5, 6, 7, 8, 9, 10, 11,
- 30 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 nucleótidos de longitud. En ciertas realizaciones, el bucle tiene 18 nucleótidos de longitud. La estructura de horquilla también puede contener porciones salientes 3’ y/o 5’. En algunas realizaciones, el saliente es un

saliente 3' y/o 5' de 0, 1, 2, 3, 4 o 5 nucleótidos de longitud.

Los ARNsh se componen de estructuras de tallo-bucle que se diseñan para contener una región de flanqueo 5', segmentos de región de ARNsi, una región de bucle, una región de ARNsi 3' y una región de flanqueo 3'. La mayoría de las estrategias de expresión de iARN han
5 utilizado ARN de horquilla corta (ARNsh) dirigidos por promotores fuertes basados en polIII. Muchos ARNsh han demostrado un silenciamiento génico efectivo de las secuencias diana *in vitro* así como *in vivo*, sin embargo, también se encontró que algunos ARNsh que demostraron una silenciamiento génico efectivo del gen diana tuvieron toxicidad *in vivo*.

Los miARN son ARN celulares pequeños (~22 nt) que se procesan a partir de
10 transcritos de bucle de tallo precursores. Los bucles de tallo de miARN conocidos se pueden modificar para contener secuencias de ARNi específicas para genes de interés. Las moléculas de miARN pueden ser preferibles a las moléculas de ARNsh debido a que los miARN se expresan de forma endógena. Por lo tanto, es poco probable que las moléculas de miARN induzcan vías de interferón sensibles a ARNds, se procesan de manera más eficiente que los
15 ARNsh y se ha mostrado que silencian 80 % de manera más efectiva.

Un enfoque alternativo recientemente descubierto es el uso de miARN artificiales (andamios de pri-miARN que transportan secuencias de ARNsi) como vectores de iARN. Los miARN artificiales se asemejan más naturalmente a los sustratos de ARNi endógeno y son más susceptibles a la transcripción de Pol-II (por ejemplo, permitiendo la expresión específica
20 de tejido de ARNi) y estrategias policistrónicas (por ejemplo, permitiendo la administración de múltiples secuencias de ARNsi). Ver patente de Estados Unidos No. 10,093,927, que se incorpora por referencia.

La unidad transcripcional de un "ARNsh" se compone de secuencias de sentido y antisentido conectadas por un bucle de nucleótidos no emparejados. Los ARNsh se exportan
25 desde el núcleo por Exportina-5, y una vez en el citoplasma, se procesan por Dicer para generar ARNsi funcionales. Los tallo-bucles de "miARN" se componen de secuencias sentido y antisentido conectadas por un bucle de nucleótidos no emparejados habitualmente expresados como parte de transcritos primarios más grandes (pri-miARN), que se escinden por el complejo Drosha-DGCR8 generando compuestos intermedios conocidos como pre-
30 miARN, que posteriormente se exportan desde el núcleo por Exportina-5, y una vez en el citoplasma, se procesan por Dicer para generar ARNsi funcionales. "MiARN artificial" o un "vector transportador de miARN artificial", como se usa en la presente de manera intercambiable, se refiere a un transcrito de miARN primario que ha tenido una región del

bucle-tallo dúplex (al menos aproximadamente 9-20 nucleótidos) que se escinde mediante el procesamiento de Drosha y Dicer reemplazado con las secuencias de ARNsi para el gen diana en tanto que se retienen los elementos estructurales dentro del bucle-tallo necesarios para el procesamiento de Drosha efectivo. El término “artificial” surge del hecho de que las secuencias
5 flaqueantes (~35 nucleótidos en la dirección 5' y ~40 nucleótidos en la dirección 3') surgen de sitios de enzimas de restricción dentro del sitio de clonación múltiple del ARNsi. Como se usa en la presente, el término “miARN” abarca tanto las secuencias de miARN que se presentan de manera natural como los vectores transportadores de miARN generados artificialmente.

10 El ARNsi se puede codificar 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 realizaciones, la señal de poliadenilación es una señal de poliadenilación mínima sintética o una secuencia de seis T.

En el diseño de iARN hay varios factores que se necesitan considerar, tal como la
15 naturaleza del ARNsi, la durabilidad del efecto de silenciamiento y la elección del sistema de administración. Para producir un efecto de iARN, el ARNsi que se introduce en el organismo habitualmente contendrá secuencias exónicas. Además, el proceso de iARN depende de la homología, por lo que las secuencias se deben seleccionar cuidadosamente para incrementar al máximo la especificidad génica, en tanto que se reduce al mínimo la posibilidad de
20 interferencia cruzada entre secuencias homólogas, pero no específicas de genes. Preferentemente, el ARNsi exhibe más de 80 %, 85 %, 90 %, 95 %, 98 % o incluso 100 % de identidad entre la secuencia del ARNsi y el gen que se va a inhibir. Las secuencias menos de aproximadamente 80 % idénticas al gen diana son sustancialmente menos efectivas. Por lo tanto, cuanto mayor sea la homología entre el ARNsi y el gen que se va a inhibir, menos
25 probable será que se vea afectada la expresión de genes no relacionados.

Además, el tamaño del ARNsi es una consideración importante. En algunas realizaciones, la presente invención se refiere a moléculas de ARNsi que incluyen al menos aproximadamente 19-25 nucleótidos y son capaces de modular la expresión génica. En el contexto de la presente invención, el ARNsi es preferentemente de menos de 500, 200, 100,
30 50 o 25 nucleótidos de longitud. Más preferentemente, el ARNsi es de aproximadamente 19 nucleótidos a aproximadamente 25 nucleótidos de longitud.

Una diana de ARNsi en general significa un polinucleótido que comprende una región que codifica para un polipéptido, o una región de polinucleótido que regula la replicación,

transcripción o traducción u otros procesos importantes para la expresión del polipéptido, o un polinucleótido que comprende tanto una región que codifica para un polipéptido como una región enlazada de manera operativa al mismo que regula la expresión. Se puede dirigir cualquier gen que se exprese en una célula. Preferentemente, un gen diana es uno implicado en o asociado con la progresión de actividades celulares importantes para la enfermedad o de interés particular como un objeto de investigación.

B. ARN no codificadores

Como se evidencia por los proyectos de clonación de ADNc y las matrices de mosaico genómico, más de 90 % del genoma humano experimenta transcripción pero no codifica para proteínas. Estos productos transcripcionales se refieren como ARN no codificadores de proteínas (ARNnc). Una variedad de transcritos de ARNnc, tal como ARN ribosómicos, ARN de transferencia, ARN endógeno competidor (ARNce), ARN nuclear pequeño (ARNsn) y ARN nucleolar pequeño (ARNsno), son esenciales para la función celular. De manera similar, también se conoce que un gran número de ARNnc cortos tal como microARN (miARN), ARN interferentes cortos endógenos (ARNsi), ARN que interactúan con PIWI (ARNpi) y ARN nucleolares pequeños (ARNsno) desempeñan importantes funciones reguladoras en células eucariotas. Estudios recientes han demostrado un grupo de transcritos largos de ARNnc (ARNlnc) que exhiben una expresión específica de tipo de célula y se localizan en compartimentos subcelulares específicos. También se conoce que los ARNlnc desempeñan un papel importante durante el desarrollo y diferenciación celular, lo que respalda la opinión de que han sido seleccionados durante el proceso evolutivo.

Los ARNlnc parecen tener muchas funciones diferentes. En muchos casos, parecen desempeñar un papel en la regulación de la actividad o localización de proteínas, o servir como marcos organizativos para estructuras subcelulares. En otros casos, los ARNlnc se procesan para producir múltiples ARN pequeños o pueden modular cómo se procesan otros ARN. La última edición de datos producida por el consorcio público de investigación GenCode (versión No. 27) cataloga poco menos de 16.000 ARNlnc en el genoma humano, produciendo casi 28.000 transcritos; cuando se incluyen otras bases de datos, se conocen más de 40.000 ARNlnc.

Curiosamente, los ARNlnc pueden influir en la expresión de proteínas diana específicas en loci genómicos específicos, modular la actividad de los compañeros de unión a proteínas, dirigir complejos modificadores de cromatina a sus sitios de acción y se procesan postranscripcionalmente para producir numerosos ARN pequeños con caperuza 5'. Las vías

epigenéticas también pueden regular la expresión diferencial de los ARNlnc.

Un creciente cuerpo de evidencia también sugiere que los ARNlnc expresados de manera aberrante juegan un papel importante en los procesos fisiológicos normales, así como en múltiples estados de enfermedad. Los vncRNA están mal regulados en varias enfermedades, incluida isquemia, enfermedad cardíaca, enfermedad de Alzheimer, psoriasis y ataxia espinocerebelosa tipo 8. Esta mala regulación también se ha mostrado en diferentes tipos de cánceres, tal como cáncer de mama, cáncer de colon, cáncer de próstata, carcinoma hepatocelular y leucemia. Varios ARNlnc, por ejemplo, gadd74 y ARNlnc-RoR5, modulan los reguladores de ciclo celular tal como ciclinas, cinasas dependientes de ciclina (CDK), inhibidores de CDK y p53 y por lo tanto, proporcionan una capa adicional de flexibilidad y robustez a la progresión del ciclo celular. Además, algunos ARNlnc se vinculan a procesos mitóticos tal como ARN satélite centromérico, que es esencial para la formación de cinetocoro y por lo tanto, crucial para la segregación cromosómica durante la mitosis en humanos y moscas. Otro ARNlnc nuclear, MA-lin1, regula la salida de fase M al funcionar en cis para reprimir la expresión de su gen vecino Pura, un regulador de la proliferación celular.

Los ARNlnc son un grupo que se define comúnmente como transcritos de más de 200 nucleótidos (por ejemplo, de aproximadamente 200 a aproximadamente 1200 nt, aproximadamente 2500 nt o más) que carecen de un marco de lectura abierto extendido (ORF). El término “ARN no codificador” (ARNnc) incluye ARNnc así como transcritos más cortos de, por ejemplo, menos de aproximadamente 200 nt, tal como aproximadamente 30 a 200 nt.

Por lo tanto, en algunas realizaciones, la administración de un ARNnc, tal como a una estructura cerebral específica de interés, corrige los niveles de expresión de ARN aberrantes o modula los niveles de ARNnc causante de enfermedad. Por consiguiente, en algunas realizaciones, la presente invención proporciona un rAAV, en donde el genoma viral se modifica para codificar para un ARN no codificador terapéutico (ARNnc). En algunas realizaciones, el ARNnc es un ARN no codificador largo (ARNlnc) de aproximadamente 200 nucleótidos (nt) de longitud o mayor. En algunas realizaciones, el agente terapéutico es un ARNnc de aproximadamente 25 nt o aproximadamente 30 nt a aproximadamente 200 nt de longitud. En algunas realizaciones, el ARNlnc es de aproximadamente 200 nt a aproximadamente 1.200 nt de longitud. En algunas realizaciones, el ARNlnc es de aproximadamente 200 nt a aproximadamente 1.100, aproximadamente 1.000, aproximadamente 900, aproximadamente 800, aproximadamente 700, aproximadamente

600, aproximadamente 500, aproximadamente 400 o aproximadamente 300 nt de longitud.

C. Sistemas CRISPR

La edición génica es una tecnología que permite la modificación de genes diana dentro de células vivas. Recientemente, el aprovechamiento del sistema inmunitario bacteriano de CRISPR para realizar la edición génica bajo demanda revolucionó la forma en que los científicos abordan la edición genómica. La proteína Cas9 del sistema CRISPR, que es una endonucleasa de ADN guiada por ARN, se puede modificar para dirigirse a nuevos sitios con relativa facilidad al alterar su secuencia de ARN guía. Este descubrimiento ha hecho que la edición génica específica de secuencia sea funcionalmente efectiva.

En general, “sistema CRISPR” se refiere colectivamente a transcripciones y otros elementos implicados en la expresión de o dirección de la actividad de genes asociados a CRISPR (“Cas”), incluidas secuencias que codifican para un gen Cas, una secuencia tracr (CRISPR transactivador) (por ejemplo, tracrRNA o un tracrRNA parcial activo), una secuencia tracr-mate (que abarca una “repetición directa” y una repetición directa parcial procesada por tracrRNA en el contexto de un sistema CRISPR endógeno), una secuencia guía (también referida como un “separador” en el contexto de un sistema CRISPR endógeno) y/u otras secuencias y transcriptos de un locus CRISPR.

La nucleasa CRISPR/Cas o sistema de nucleasa CRISPR/Cas puede incluir un ARN de molécula de ARN no codificador (guía), cuya secuencia se une específicamente a ADN, y una proteína Cas (por ejemplo, Cas9), con funcionalidad de nucleasa (por ejemplo, dos dominios de nucleasa). Uno o más elementos de un sistema CRISPR pueden derivar de un sistema CRISPR tipo I, tipo II o tipo III, por ejemplo, derivado de un organismo particular que comprende un sistema CRISPR endógeno, tal como *Streptococcus pyogenes*.

El sistema CRISPR puede inducir rupturas de doble hebra (DSB) en el sitio diana, seguidas de interrupciones como se analiza en la presente. En otras realizaciones, las variantes de Cas9, consideradas “nickasas”, se usan para cortar una hebra individual en el sitio diana. Se pueden usar nickasas emparejadas, por ejemplo, para mejorar la especificidad, cada una dirigida por un par de secuencias de direccionamiento a ARNg diferentes de modo que en la introducción de las nickasas de manera simultánea, se introduce un saliente 5'. En otras realizaciones, Cas9 catalíticamente inactiva se fusiona a un dominio efector heterólogo tal como un activador o represor transcripcional (por ejemplo, CRAB), para afectar la expresión génica. De manera alternativa, un sistema CRISPR con una Cas9 catalíticamente inactivada

comprende además un activador o represor transcripcional fusionado a una proteína de unión ribosómica.

En algunos aspectos, una nucleasa Cas y ARNg (incluida una fusión de ARNcr específico para la secuencia diana y tracrRNA fijo) se introducen en la célula. En general, los sitios diana en el extremo 5' del ARNg se dirigen a la nucleasa Cas al sitio diana, por ejemplo, el gen, usando emparejamiento de bases complementario. El sitio diana se puede seleccionar con base en su ubicación inmediatamente 5' de una secuencia de motivo adyacente de protoseparador (PAM), tal como habitualmente NGG o NAG. En este sentido, el ARNg se dirige a la secuencia deseada al modificar los primeros 20, 19, 18, 17, 16, 15, 14, 14, 12, 11 o 10 nucleótidos del ARN guía para que correspondan a la secuencia de ADN diana. En general, un sistema CRISPR se caracteriza por elementos que promueven la formación de un complejo CRISPR en el sitio de una secuencia diana. Habitualmente, "secuencia diana" en general se refiere a una secuencia a la cual una secuencia guía se diseña para tener complementariedad, donde la hibridación entre la secuencia diana y una secuencia guía promueve la formación de un complejo CRISPR. No se requiere necesariamente complementariedad completa, siempre que exista suficiente complementariedad para provocar hibridación y promover la formación de un complejo CRISPR.

La secuencia diana puede comprender cualquier polinucleótido, tal como polinucleótidos de ADN o ARN. La secuencia diana se puede ubicar en el núcleo o citoplasma de la célula, tal como dentro de un orgánulo de la célula. En general, una secuencia o plantilla, que se puede usar para recombinación en el locus dirigido que comprende las secuencias diana, se refiere como una "plantilla de edición" o "polinucleótido de edición" o "secuencia de edición". En algunos aspectos, un polinucleótido de plantilla exógeno se puede referir como una plantilla de edición. En algunos aspectos, la recombinación es recombinación homóloga.

Habitualmente, en el contexto de un sistema CRISPR endógeno, la formación del complejo CRISPR (que comprende la secuencia guía hibridada con la secuencia diana y vuelta compleja con una o más proteínas Cas) da por resultado la escisión de una o ambas hebras en o cerca de (por ejemplo, dentro de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50 o más pares de bases de) la secuencia diana. La secuencia tracr, que puede comprender o consistir en toda o una porción de una secuencia tracr tipo silvestre (por ejemplo, aproximadamente o más de aproximadamente 20, 26, 32, 45, 48, 54, 63, 67, 85 o más nucleótidos de una secuencia tracr tipo silvestre), también puede formar parte del complejo CRISPR, tal como por hibridación a lo largo de al menos una porción de la secuencia tracr a toda o una porción de

una secuencia tracr mate que se enlaza de manera operativa a la secuencia guía. La secuencia tracr tiene suficiente complementariedad con una secuencia tracr mate para hibridar y participar en la formación del complejo CRISPR, tal como al menos 50 %, 60 %, 70 %, 80 %, 90 %, 95 % o 99 % de complementariedad de secuencia a lo largo de la longitud de la secuencia tracr mate cuando se alinea de manera óptima.

Uno o más vectores que activan la expresión de uno o más elementos del sistema CRISPR se pueden introducir en la célula de modo que la expresión de los elementos de la formación directa de sistema CRISPR del complejo CRISPR en uno o más sitios diana. Los componentes también se pueden administrar a las células como proteínas y/o ARN. Por ejemplo, una enzima Cas, una secuencia guía enlazada a una secuencia tracr-mate y una secuencia tracr se pueden enlazar cada una de manera operativa a elementos reguladores separados en vectores separados. La enzima Cas puede ser un gen diana bajo el control de un evento de empalme alternativo regulado, como se describe en la presente, ya sea como un minigén de gen diana quimérico o como un gen diana para un transactivador de minigén quimérico. El ARNg puede estar bajo el control de un promotor constitutivo.

De manera alternativa, dos o más de los elementos expresados a partir de los mismos o diferentes elementos reguladores se pueden combinar en un vector individual, con uno o más vectores adicionales que proporcionan cualquier componente del sistema CRISPR no incluido en el primer vector. El vector puede comprender uno o más sitios de inserción, tal como una secuencia de reconocimiento de endonucleasa de restricción (también referida como un “sitio de clonación”). En algunas realizaciones, uno o más sitios de inserción se ubican en la dirección 5' y/o en la dirección 3' de uno o más elementos de secuencia de uno o más vectores. Cuando se usan múltiples secuencias guía diferentes, se puede usar una construcción de expresión individual para dirigir la actividad de CRISPR a múltiples secuencias diana correspondientes diferentes dentro de una célula.

Un vector puede comprender un elemento regulador enlazado de manera operativa a una secuencia codificadora de enzimas que codifica para la enzima CRISPR, tal como una proteína Cas. Los ejemplos no limitantes de proteínas Cas incluyen Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (también conocidas como Csn1 y Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homólogos de las mismas o versiones modificadas de las mismas. Estas enzimas se conocen; por ejemplo, la secuencia de aminoácidos de la

proteína Cas9 de *S. pyogenes* se puede encontrar en la base de datos SwissProt con el número de acceso Q99ZW2.

La enzima CRISPR puede ser Cas9 (por ejemplo, de *S. pyogenes* o *S. pneumoniae*). La enzima CRISPR puede dirigir la escisión de una o ambas hebras en la ubicación de una secuencia diana, tal como dentro de la secuencia diana y/o dentro del complemento de la secuencia diana. El vector puede codificar para una enzima CRISPR que se muta con respecto a una correspondiente enzima tipo silvestre, de modo que la enzima CRISPR mutada carece de la capacidad para escindir una o ambas hebras de un polinucleótido diana que contiene una secuencia diana. Por ejemplo, una sustitución de aspartato a alanina (D10A) en el dominio catalítico RuvC I de Cas9 de *S. pyogenes* convierte Cas9 de una nucleasa que escinde ambas hebras en una nickasa (escinde una hebra individual). En algunas realizaciones, se puede usar una nickasa Cas9 en combinación con secuencias guía, por ejemplo, dos secuencias guía, que se dirigen respectivamente a hebras sentido y antisentido de la diana de ADN. Esta combinación permite que ambas hebras se corten y se usen para inducir NHEJ o HDR.

En algunas realizaciones, una secuencia codificadora de enzimas que codifica para la enzima CRISPR se optimiza por codones para la expresión en células particulares, tal como células eucariotas. Las células eucariotas pueden ser aquellas de o derivadas de un organismo particular, tal como un mamífero, incluido pero no limitado a humano, ratón, rata, conejo, perro o primate no humano. En general, la optimización por codones se refiere a un proceso para modificar una secuencia de ácido nucleico para expresión mejorada en las células hospedadoras de interés al reemplazar al menos un codón de la secuencia nativa con codones que se usan más frecuentemente o con mayor frecuencia en los genes de esa célula hospedadora en tanto que se mantiene la secuencia de aminoácidos nativa. Diferentes especies exhiben desviación particular para ciertos codones de un aminoácido particular. La desviación de codones (diferencias en el uso de codones entre organismos) frecuentemente se correlaciona con la eficiencia de traducción de ARN mensajero (ARNm), que a su vez se cree que es dependiente, entre otras cosas, de las propiedades de los codones que se traducen y la disponibilidad de moléculas particulares de ARN de transferencia (ARNt). El predominio de ARNt seleccionados en una célula es en general un reflejo de los codones usados con mayor frecuencia en la síntesis de péptidos. Por consiguiente, los genes se pueden adaptar para expresión génica óptima en un organismo dado con base en la optimización por codones.

En general, una secuencia guía es cualquier secuencia de polinucleótidos que tiene suficiente complementariedad con una secuencia de polinucleótidos diana para hibridarse con la secuencia diana y unión específica de secuencia directa del complejo CRISPR a la secuencia diana. En algunas realizaciones, el grado de complementariedad entre una
5 secuencia guía y su correspondiente secuencia diana, cuando se alinea de manera óptima usando un algoritmo de alineación adecuado, es aproximadamente o más de aproximadamente 50 %, 60 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97,5 %, 99 % o más.

La alineación óptima se puede determinar con el uso de cualquier algoritmo adecuado para alinear secuencias, los ejemplos no limitantes de lo cual incluyen el algoritmo Smith-
10 Waterman, el algoritmo Needleman-Wunsch, algoritmos basados en la transformada Burrows-Wheeler (por ejemplo, el alineador Burrows Wheeler), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies, ELAND (Illumina, San Diego, Calif.), SOAP (disponible en soap.genomics.org.cn) y Maq (disponible en maq.sourceforge.net).

La enzima CRISPR puede ser parte de una proteína de fusión que comprende uno o
15 más dominios de proteína heterólogos. Una proteína de fusión de enzima CRISPR puede comprender cualquier secuencia de proteína adicional y opcionalmente una secuencia ligadora entre cualquiera de dos dominios. Los ejemplos de dominios de proteína que se pueden fusionar a una enzima CRISPR incluyen, sin limitación, etiquetas de epítipo, secuencias de genes indicadores y dominios de proteína que tienen una o más de las
20 siguientes actividades: actividad de metilasa, actividad de desmetilasa, actividad de activación de transcripción, actividad de represión de transcripción, actividad de factor de liberación de transcripción, actividad de modificación de histona, actividad de escisión de ARN y actividad de unión a ácido nucleico. Los ejemplos no limitantes de etiquetas de epítipo incluyen etiquetas de histidina (His), etiquetas V5, etiquetas FLAG, etiquetas de hemaglutinina de
25 influenza (HA), etiquetas Myc, etiquetas VSV-G y etiquetas de tioredoxina (Trx). Los ejemplos de genes indicadores incluyen, pero no se limitan a, glutatión-S-transferasa (GST), peroxidasa de rábano picante (HRP), cloranfenicol acetiltransferasa (CAT), beta galactosidasa, beta-glucuronidasa, luciferasa, proteína fluorescente verde (GFP), HcRed, DsRed, proteína fluorescente cian (CFP), proteína fluorescente amarilla (YFP) y proteínas autofluorescentes
30 que incluyen proteína fluorescente azul (BFP). Una enzima CRISPR se puede fusionar a una secuencia génica que codifica para una proteína o un fragmento de una proteína que se une a moléculas de ADN o se une a otras moléculas celulares, incluido pero no limitado a proteína de unión a maltosa (MBP), S-tag, fusiones de dominio de unión a ADN (DBD) Lex A, fusiones de dominio de unión a ADN GAL4A y fusiones de proteína BP16 de virus de herpes simple

(HSV). Los dominios adicionales que pueden formar parte de una proteína de fusión que comprende una enzima CRISPR se describen en US 20110059502, incorporada en la presente por referencia.

D. Proteínas terapéuticas

5 Algunas realizaciones se refieren a la expresión de proteínas y polipéptidos recombinantes. Estas proteínas pueden ser proteínas neuroprotectoras, tal como una proteína antiapoptótica, una enzima antioxidante (por ejemplo, aquellas que pertenecen a la familia de superóxido dismutasa (SOD)), un factor neurotrófico/neuroprotector o una proteína antiinflamatoria. La proteína terapéutica puede ser Bircla (NAIP), Birc2 (c-IAP1/HIAP-2), Birc3
10 (cIAP-2/HIAP-1), Birc4 (XIAP), Birc5 (survivina), Birc6 (apollon), Birc7 (livin), Birc8 (TslAP); miembros de la familia Bcl-2: Bcl-2, Bcl-XL, Bcl-w, Mcl-1, Bcl-2L10, BFL-1; inhibidores endógenos de la cinasa de extremo N de c-Jun (JNK) conocidos como proteína que interactúa con Jun (JIP), JIP-1, JIP-2, JIP-3, JIP-4; SOD1, SOD2; catalasa; peroxirredoxina-1, peroxirredoxina-2, glutatión preoxidasa 1 (Gpx1), Gpx2, Gpx3 o Gpx4; NGF, BDNF, CNTF,
15 GDNF, factor de crecimiento/diferenciación-15 (GDF-15), eritropoyetina o factor de crecimiento endotelial vascular (VEGF); interleucina-10 (IL-10); glutatión S-transferasa, anexina-1 (ANXA1), inhibidor de NF- κ B (I κ B); USH1, USH1G (también conocido como SANS), CIB2 (calcio y proteína de unión a integrina 2), ATOH-1, ACTG1, ATP2B2, CDH23 (cadherina 23), CLDN14, CLRN1, COCH, COL11A2, DFNA5, DFNB31, DFNB59, ESPN, EYA4, GJB2,
20 GJB3, GJB6, KCNQ4, LHFPL5, MT-RNR1, MT-TS1, MYO1A, MYO6, MYO7A (miosina 7a), MYO 15A, OTOF, PCDH15 (protocadherina 15), PDZD7, POU3F4, SLC26A4, STRC, TECTA, TMC1, TMC2, TMIE, TMPRSS3, TRIOBP, USH1C, VLGR1, WFS1, ACTG1, ADCY1, ATOF1, ATP6V1B1, BDNF, BDP1, BSND, DATSPER2, CABP2, CD164, CDC 14 A, CDH23, CEACAM16, CHD7, CCDC50, CIB2, CLDN14, CLIC5, CLPP, CLRN1, COCH, COL2A1,
25 COL4A3, COL4A4, COL4A5, COL9A 1, COL9A2, COL1 1 A1, COL1 1 A2, CRYM, DCDC2, DFNA5, DFNB31, DFNB59, DIAPH1, EDN3, EDNRB, ELMOD3, EMOD3, EPS8, EPS8L2, ESPN, ESRRB, EYAI, EYA4, FAM65B, FOXH, GIPC3, GJB2, GJB3, GJB6, GPR98, GRHL2, GPSM2, GRXCR1 , GRXCR2, HARS2, HGF, HOMER2, HSD17B4, ILDR1, KARS, KCNE1, KCNJ10, KCNQ1, KCNQ4, KITLG, LARS2, LHFPL5, LOXHD1, LRTOMT, MARVELD2,
30 MCM2, MET, MIR183, MIRN96, MITF, MSRB3, MT-RNR1, MT-TS 1, MYH14, MYH9, MYO15A, MYOIA, MYO3A, MYO6, MYO7A, NARS2, NDP, NF2, NT3, OSBPL2, OTOA, OTOF, OTOG, OTOGL, P2RX2, PAX3, PCDH15, PDZD7, PJKV, PNPT1, POLR1D, POLR1C, POU3F4, POU4F3, PRPS 1, PTPRQ, RDX, S1PR2, SANS, SEMA3E, SERPINB6, SLC17A8, SLC22A4, SLC26A4, SLC26A5, SIX1, SIX5, S MAC/DIABLO, SNAI2, SOX 10, STRC, SYNE4,

TBC 1 D24, TCOF1, TECTA, TIMM8A, TJP2, TNC, TMC1, TMC2, TMIE, TMEM 132E, TMPRSS3, TRPN, TRIOBP, TSPEAR, USHI C, USHI G, USH2A, USH2D, VLGR1, WFS1, WHRN o XIAP.

Cuando la presente solicitud se refiere a la función o actividad de “proteína modificada” o un “polipéptido modificado”, un experto en la técnica entendería que esto incluye, por ejemplo, una proteína o polipéptido que posee una ventaja adicional sobre la proteína o polipéptido no modificado. Se contempla específicamente que las realizaciones relacionadas con una “proteína modificada” se pueden implementar con respecto a un “polipéptido modificado” y viceversa.

Las proteínas recombinantes pueden poseer deleciones y/o sustituciones de aminoácidos; por lo tanto, una proteína con una deleción, una proteína con una sustitución y una proteína con una deleción y una sustitución son proteínas modificadas. En algunas realizaciones, estas proteínas pueden incluir además inserciones o aminoácidos adicionados, tal como, por ejemplo, con proteínas de fusión o proteínas con ligadores. Una “proteína eliminada modificada” carece de uno o más residuos de la proteína nativa, pero puede poseer la especificidad y/o actividad de la proteína nativa. Una “proteína eliminada modificada” también puede tener inmunogenicidad o antigenicidad reducida. Un ejemplo de una proteína eliminada modificada es una que tiene un residuo de aminoácido eliminado de al menos una región antigénica, es decir, una región de la proteína que se determina que es antigénica en un organismo particular, tal como el organismo al que se administra la proteína modificada.

Las variantes de sustitución o remplazo habitualmente contienen el intercambio de un aminoácido por otro en uno o más sitios dentro de la proteína y se pueden diseñar para modular una o más propiedades del polipéptido, en particular sus funciones efectoras y/o biodisponibilidad. Las sustituciones pueden o no ser conservadoras, es decir, un aminoácido se reemplaza con uno de forma y carga similares. Las sustituciones conservadoras se conocen bien en la técnica e incluyen, por ejemplo, los cambios de: alanina a serina; arginina a lisina; asparagina a glutamina o histidina; aspartato a glutamato; cisteína a serina; glutamina a asparagina; glutamato a aspartato; glicina a prolina; histidina a asparagina o glutamina; isoleucina a leucina o valina; leucina a valina o isoleucina; lisina a arginina; metionina a leucina o isoleucina; fenilalanina a tirosina, leucina o metionina; serina a treonina; treonina a serina; triptófano a tirosina; tirosina a triptófano o fenilalanina; y valina a isoleucina o leucina.

Además de una deleción o sustitución, una proteína modificada puede poseer una inserción de residuos, que habitualmente implica la adición de al menos un residuo en el

polipéptido. Esto puede incluir la inserción de un péptido o polipéptido de direccionamiento o simplemente un residuo individual. Las adiciones terminales, llamadas proteínas de fusión, se analizan más adelante.

5 El término “equivalente biológicamente funcional” se entiende bien en la técnica y se define adicionalmente con detalle en la presente. Por consiguiente, se incluyen secuencias que tienen entre aproximadamente 70 % y aproximadamente 80 %, o entre aproximadamente 81 % y aproximadamente 90 %, o incluso entre aproximadamente 91 % y aproximadamente 99 % de aminoácidos que son idénticos o funcionalmente equivalentes a los aminoácidos de un polipéptido de control, siempre que se mantenga la actividad biológica de la proteína. Una
10 proteína recombinante puede ser biológicamente funcionalmente equivalente a su contraparte nativa en ciertos aspectos.

También se entenderá que las secuencias de aminoácidos y de ácido nucleico pueden incluir residuos adicionales, tal como aminoácidos N- o C-terminales adicionales o secuencias 5' o 3', y aún así ser esencialmente como se expone en una de las secuencias divulgadas en
15 la presente, siempre que la secuencia cumpla con los criterios expuestos anteriormente, incluido el mantenimiento de la actividad de proteína biológica en lo que respecta a la expresión de proteína. La adición de secuencias terminales se aplica particularmente a secuencias de ácido nucleico que pueden, por ejemplo, incluir diferentes secuencias no codificadoras que flanquean cualquiera de las porciones 5' o 3' de la región codificadora o
20 pueden incluir diferentes secuencias internas, es decir, intrones, que se conoce que se presentan dentro de genes.

Como se usa en la presente, una proteína o péptido en general se refiere, pero no se limita a, una proteína de más de aproximadamente 200 aminoácidos, hasta una secuencia de longitud completa traducida de un gen; un polipéptido de más de aproximadamente 100
25 aminoácidos; y/o un péptido de aproximadamente 3 a aproximadamente 100 aminoácidos. Por conveniencia, los términos “proteína”, “polipéptido” y “péptido” se usan indistintamente en la presente.

Como se usa en la presente, un “residuo de aminoácido” se refiere a cualquier aminoácido que se presenta de manera natural, cualquier derivado de aminoácidos o
30 cualquier mimético de aminoácidos conocido en la técnica. En ciertas realizaciones, los residuos de la proteína o péptido son secuenciales, sin ningún aminoácido que interrumpa la secuencia de residuos de aminoácidos. En otras realizaciones, la secuencia puede comprender una o más porciones no de aminoácidos. En realizaciones particulares, la

secuencia de residuos de la proteína o péptido se puede interrumpir por una o más porciones no de aminoácidos.

Por consiguiente, el término “proteína o péptido” abarca secuencias de aminoácidos que comprenden al menos uno de los 20 aminoácidos comunes que se encuentran en las proteínas que se presentan de manera natural, o al menos un aminoácido modificado o inusual.

Ciertas realizaciones de la presente invención se refieren a proteínas de fusión. Estas moléculas pueden tener una proteína terapéutica enlazada en el extremo N-terminal o C-terminal a un dominio heterólogo. Por ejemplo, las fusiones también pueden emplear secuencias líder de otras especies para permitir la expresión recombinante de una proteína en un hospedador heterólogo. Otra fusión útil incluye la adición de una etiqueta de afinidad de proteína, tal como una etiqueta de afinidad de albúmina de suero o seis residuos de histidina, o un dominio inmunológicamente activo, tal como un epítipo de anticuerpo, preferentemente escindible, para facilitar la purificación de la proteína de fusión. Las etiquetas de afinidad no limitantes incluyen polihistidina, proteína de unión a quitina (CBP), proteína de unión a maltosa (MBP) y glutatión-S-transferasa (GST).

Los métodos para generar proteínas de fusión se conocen bien por aquellos expertos en la técnica. Estas proteínas se pueden producir, por ejemplo, por síntesis de novo de la proteína de fusión completa, o por unión de la secuencia de ADN que codifica para el dominio heterólogo, seguido de expresión de la proteína de fusión intacta.

La producción de proteínas de fusión que recuperan las actividades funcionales de las proteínas de origen se puede facilitar al conectar genes con un segmento de ADN puente que codifica para un ligador peptídico que se empalma entre los polipéptidos conectados en tándem. El ligador sería de longitud suficiente para permitir el plegamiento adecuado de la proteína de fusión resultante.

La expresión de un transgén se puede dirigir por el promotor natural del transgén (es decir, el promotor que se encuentra de forma natural con la secuencia codificadora transgénica) o la expresión de un transgén se puede dirigir por un promotor heterólogo (por ejemplo, promotor de CMV, promotor de Espin, un promotor de PCDH15, un promotor de PTPRQ y un promotor de TMHS (LHFPL5)). Por ejemplo, cualquiera de los transgenes descritos en la presente se puede usar con su promotor natural. De manera alternativa, cualquiera de los transgenes descritos en la presente se puede usar con un promotor

heterólogo. Como se usa en la presente, un promotor heterólogo se refiere a un promotor que no dirige naturalmente la expresión de esa secuencia (es decir, no se encuentra con esa secuencia en la naturaleza). Los promotores heterólogos representativos que se pueden usar para dirigir la expresión de cualquiera de los transgenes indicados en la presente incluyen, por ejemplo, un promotor de CMV, un promotor de CBA, un promotor de CASI, un promotor de P y un promotor de EF-1, un promotor de receptor nicotínico alfa9, un promotor de prestina, un promotor de Gfil y un promotor de Vglut3. Además, un promotor que dirige de forma natural la expresión de uno de los transgenes mencionados anteriormente (por ejemplo, un promotor de KCNQ4, un promotor de Myo7a, un promotor de Myo6 o un promotor de Atoh1) se puede usar como un promotor heterólogo para dirigir la expresión de un transgén. En otras realizaciones, el promotor es un promotor de Espin, un promotor de PCDH15, un promotor de PTPRQ y un promotor de TMHS (LHFPL5).

III. Métodos de tratamiento y administración

Los vectores virales en algunos aspectos se pueden administrar directamente a pacientes (*in vivo*) o se pueden usar para tratar células *in vitro* o *ex vivo*, y entonces se pueden administrar a pacientes. El término “vector” se refiere a una molécula de ácido nucleico portadora pequeña, un plásmido, virus (por ejemplo, vector de AAV, vector retroviral, vector lentiviral) u otro vehículo que se puede manipular por inserción o incorporación de un ácido nucleico. Los vectores, tal como vectores virales, se pueden usar para introducir/transferir secuencias de ácido nucleico en células, de modo que la secuencia de ácido nucleico en estas se transcriba y, si codifica para una proteína, se traduzca posteriormente por las células.

Se puede administrar o tratar cualquier célula o mamífero adecuado por un método o uso descrito en la presente. Habitualmente, se sospecha que un mamífero en necesidad de un método descrito en la presente tiene o expresa una proteína anormal o aberrante que se asocia con un estado de enfermedad. De manera alternativa, el receptor mamífero puede tener una condición que es susceptible a terapia de reemplazo génico. Como se usa en la presente, “terapia de reemplazo génico” se refiere a la administración al receptor de material genético exógeno que codifica para un agente terapéutico y la posterior expresión del material genético administrado *in situ*. Por lo tanto, la frase “condición susceptible a terapia de reemplazo génico” abarca condiciones tal como enfermedades genéticas (es decir, una condición de enfermedad que es atribuible a uno o más defectos génicos), patologías adquiridas (es decir, una condición patológica que no es atribuible a un defecto congénito), cánceres y procesos profilácticos (es decir, prevención de una enfermedad o de una condición

médica no deseada). Por consiguiente, como se usa en la presente, el término “agente terapéutico” se refiere a cualquier agente o material, que tiene un efecto beneficioso en el receptor mamífero. Por lo tanto, “agente terapéutico” abarca moléculas tanto terapéuticas como profilácticas que tienen componentes de ácido nucleico o proteína.

5 Los ejemplos no limitantes de mamíferos incluyen humanos, primates no humanos (por ejemplo, simios, gibones, chimpancés, orangutanes, monos, macacos y similares), animales domésticos (por ejemplo, perros y gatos), animales de granja (por ejemplo, caballos, vacas, cabras, ovejas, cerdos) y animales experimentales (por ejemplo, ratón, rata, conejo, conejillo de indias). En ciertas realizaciones, un mamífero es un humano. En ciertas
10 realizaciones, un mamífero es un mamífero no roedor (por ejemplo, humano, cerdo, cabra, oveja, caballo, perro o similares). En ciertas realizaciones, un mamífero no roedor es un humano. Un mamífero puede ser de cualquier edad o en cualquier etapa de desarrollo (por ejemplo, un adulto, adolescente, niño, infante o un mamífero en el útero). Un mamífero puede ser hombre o mujer. En ciertas realizaciones, un mamífero puede ser un modelo de
15 enfermedad animal, por ejemplo, modelos animales que tienen o expresan una proteína anormal o aberrante que se asocia con un estado de enfermedad o modelos animales con expresión insuficiente de una proteína, que provoca un estado de enfermedad.

 Los mamíferos (sujetos) tratados por un método o composición descrita en la presente incluyen adultos (18 años o más) y niños (menos de 18 años de edad). Los adultos incluyen
20 los ancianos. Los adultos representativos tienen 50 años o más. Los niños varían en edad de 1-2 años de edad o de 2-4, 4-6, 6-18, 8-10, 10-12, 12-15 y 15-18 años de edad. Los niños también incluyen bebés. Los bebés habitualmente varían de 1-12 meses de edad.

 En ciertas realizaciones, un método incluye administrar una pluralidad de partículas virales a un mamífero como se expone en la presente, donde la severidad, frecuencia,
25 progresión o tiempo de inicio de uno o más síntomas de un estado de enfermedad, tal como una enfermedad neurodegenerativa, disminuyó, redujo, previno, inhibió o retrasó. En ciertas realizaciones, un método incluye administrar una pluralidad de partículas virales a un mamífero para tratar un síntoma adverso de un estado de enfermedad, tal como una enfermedad neurodegenerativa. En ciertas realizaciones, un método incluye administrar una
30 pluralidad de partículas virales a un mamífero para estabilizar, retrasar o prevenir el empeoramiento, o progresión, o revertir y síntomas adversos de un estado de enfermedad, tal como una enfermedad neurodegenerativa.

 En ciertas realizaciones, un método incluye administrar una pluralidad de partículas

virales al sistema nervioso central, o una porción de las mismas como se expone en la presente, de un mamífero severidad, frecuencia, progresión o tiempo de inicio de uno o más síntomas de un estado de enfermedad, tal como una enfermedad neurodegenerativa, se disminuye, reduce, previene, inhibe o retrasa por al menos aproximadamente 5 a
5 aproximadamente 10, aproximadamente 10 a aproximadamente 25, aproximadamente 25 a aproximadamente 50, o aproximadamente 50 a aproximadamente 100 días.

En ciertas realizaciones, un síntoma o efecto adverso comprende un síntoma de etapa temprana, media o tardía; un síntoma de comportamiento, personalidad o lenguaje; síntoma de deglución, movimiento, convulsión, temblor o inquietud; ataxia; y/o un síntoma cognitivo tal
10 como memoria, capacidad de organización.

IV. Composiciones farmacéuticas

Como se usa en la presente, el término “farmacéuticamente aceptable” y “fisiológicamente aceptable” significa una composición, formulación, líquido o sólido biológicamente aceptable, o mezcla de los mismos, que es adecuado para una o más vías de
15 administración, administración *in vivo* o contacto. Una composición “farmacéuticamente aceptable” o “fisiológicamente aceptable” es un material que no es biológicamente o de otro modo indeseable, por ejemplo, el material se puede administrar a un sujeto sin provocar efectos biológicos indeseables sustanciales. Esta composición, formulaciones y composiciones “farmacéuticamente aceptables” y “fisiológicamente aceptables” pueden ser
20 estériles. Estas formulaciones y composiciones farmacéuticas se pueden usar, por ejemplo, en la administración de una partícula viral a un sujeto.

Estas formulaciones y composiciones incluyen solventes (acuosos o no acuosos), soluciones (acuosas o no acuosas), emulsiones (por ejemplo, aceite en agua o agua en aceite), suspensiones, jarabes, elixires, medios de dispersión y suspensión, revestimientos,
25 agentes isotónicos y que promueven o retrasan la absorción, compatibles con la administración farmacéutica o el contacto o administración *in vivo*. Los solventes, soluciones y suspensiones acuosas y no acuosas pueden incluir agentes de suspensión y agentes espesantes. Los compuestos activos complementarios (por ejemplo, conservadores, agentes antibacterianos, antivirales y antifúngicos) también se pueden incorporar en las formulaciones
30 y composiciones.

Las composiciones farmacéuticas habitualmente contienen un excipiente farmacéuticamente aceptable. Estos excipientes incluyen cualquier agente farmacéutico que

no induzca por sí mismo la producción de anticuerpos dañinos para el individuo que recibe la composición, y que se pueda administrar sin toxicidad indebida. Los excipientes farmacéuticamente aceptables incluyen, pero no se limitan a, sorbitol, Tween80 y líquidos tal como agua, solución salina, glicerol y etanol. Se pueden incluir sales farmacéuticamente aceptables, por ejemplo, sales de ácidos minerales tal como clorhidratos, bromhidratos, fosfatos, sulfatos y similares; y las sales de ácidos orgánicos tal como acetatos, propionatos, malonatos, benzoatos y similares. Además, sustancias auxiliares, tal como agentes tensioactivos, agentes humectantes o emulsionantes, sustancias amortiguadoras de pH y similares, pueden estar presentes en estos vehículos.

10 Las composiciones farmacéuticas se pueden formular para ser compatibles con una vía de administración o suministro particular, como se expone en la presente o se conoce por un experto en la técnica. Por lo tanto, las composiciones farmacéuticas incluyen portadores, diluyentes o excipientes adecuados para la administración o suministro por diferentes vías.

15 Las formas farmacéuticas adecuadas para inyección o infusión de partículas virales pueden incluir soluciones acuosas estériles o dispersiones que se adaptan para la preparación extemporánea de soluciones inyectables estériles o infusibles o dispersiones, opcionalmente encapsuladas en liposomas. En todos los casos, la forma final debe ser un fluido estéril y estable bajo las condiciones de fabricación, uso y almacenamiento. El portador o vehículo líquido puede ser un solvente o medio de dispersión líquido que comprende, por ejemplo, 20 agua, etanol, un poliol (por ejemplo, glicerol, propilenglicol, polietilenglicoles líquidos y similares), aceites vegetales, ésteres de glicerilo no tóxicos y mezclas adecuadas de los mismos. La fluidez apropiada se puede mantener, por ejemplo, por la formación de liposomas, por el mantenimiento del tamaño de partícula requerido en el caso de dispersiones o por el uso de agentes tensioactivos. Se pueden incluir agentes isotónicos, por ejemplo, azúcares, 25 amortiguadores o sales (por ejemplo, cloruro de sodio). La absorción prolongada de composiciones inyectables se puede provocar por el uso en las composiciones de agentes que retrasan la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Las soluciones o suspensiones de partículas virales pueden incluir opcionalmente uno o más de los siguientes componentes: un diluyente estéril tal como agua para inyección, 30 solución salina, tal como solución salina amortiguada con fosfato (PBS), CFS artificial, un agente tensioactivo, aceites fijos, un poliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido y similares), glicerina u otros solventes sintéticos; agentes antibacterianos y antifúngicos tal como parabenos, clorobutanol, fenol, ácido ascórbico y similares;

antioxidantes tal como ácido ascórbico o bisulfito de sodio; agentes quelantes tal como ácido etilendiaminotetraacético; amortiguadores tal como acetatos, citratos o fosfatos y agentes para el ajuste de tonicidad tal como cloruro de sodio o dextrosa.

5 Las formulaciones farmacéuticas, composiciones y sistemas de administración apropiados para las composiciones, métodos y usos de la invención se conocen en la técnica (ver, por ejemplo, Remington: The Science and Practice of Pharmacy (2003) 20th ed., Mack Publishing Co., Easton, PA; Remington's Pharmaceutical Sciences (1990) 18th ed., Mack Publishing Co., Easton, PA; The Merck Index (1996) 12th ed., Merck Publishing Group, Whitehouse, NJ; Pharmaceutical Principles of Solid Dosage Forms (1993), Technonic
10 Publishing Co., Inc., Lancaster, Pa.; Ansel and Stoklosa, Pharmaceutical Calculations (2001) 11th ed., Lippincott Williams & Wilkins, Baltimore, MD; y Poznansky *et al.*, Drug Delivery Systems (1980), R. L. Juliano, ed., Oxford, N.Y., pp. 253-315).

Las partículas virales y sus composiciones se pueden formular en forma unitaria de dosis para facilidad de administración y uniformidad de dosis. La forma de unidad de dosis
15 como se usa en la presente se refiere a unidades físicamente discretas adecuadas como dosificaciones unitarias para un individuo que se va a tratar; cada unidad que contiene una cantidad predeterminada de compuesto activo calculada para producir el efecto terapéutico deseado en asociación con el vehículo farmacéutico requerido. Las formas unitarias de dosis dependen del número de partículas virales que se consideran necesarias para producir el o
20 los efectos deseados. La cantidad necesaria se puede formular en una dosis individual o se puede formular en múltiples unidades de dosis. La dosis se puede ajustar a una concentración de partículas virales adecuada, opcionalmente combinada con un agente antiinflamatorio y envasada para uso.

En una realización, las composiciones farmacéuticas incluirán suficiente material
25 genético para proporcionar una cantidad terapéuticamente efectiva, es decir, una cantidad suficiente para reducir o mejorar los síntomas o un efecto adverso de un estado de enfermedad en cuestión o una cantidad suficiente para conferir el beneficio deseado.

Una "forma de dosis unitaria" como se usa en la presente se refiere a unidades físicamente discretas adecuadas como dosis unitarias para el sujeto que se va a tratar; cada
30 unidad que contiene una cantidad predeterminada opcionalmente en asociación con un portador farmacéutico (excipiente, diluyente, vehículo o agente de relleno) que, cuando se administra en una o más dosis, se calcula para producir un efecto deseado (por ejemplo, efecto profiláctico o terapéutico). Las formas de dosis unitaria pueden estar dentro de, por

ejemplo, ampollas y viales, que pueden incluir una composición líquida, o una composición en un estado liofilizado o liofilizado; un portador líquido estéril, por ejemplo, se puede adicionar antes de la administración o suministro *in vivo*. Las formas de dosis unitaria individuales se pueden incluir en kits o recipientes de dosis múltiples. Por lo tanto, por ejemplo, las partículas virales y composiciones farmacéuticas de las mismas se pueden envasar en forma de dosis unitaria individual o múltiple para facilidad de administración y uniformidad de dosis.

Las formulaciones que contienen partículas virales habitualmente contienen una cantidad efectiva, la cantidad efectiva se determina fácilmente por un experto en la técnica. Las partículas virales pueden variar habitualmente de aproximadamente 1 % a aproximadamente 95 % (p/p) de la composición, o incluso más si es adecuado. La cantidad que se va a administrar depende de factores tal como la edad, peso y condición física del mamífero o el sujeto humano considerado para el tratamiento. Las dosis efectivas se pueden establecer por un experto en la técnica a través de ensayos de rutina que establecen curvas de respuesta a la dosis.

V. Definiciones

Los términos “polinucleótido”, “ácido nucleico” y “transgén” se usan indistintamente en la presente para referirse a todas las formas de ácido nucleico, oligonucleótidos, incluidos ácido desoxirribonucleico (ADN) y ácido ribonucleico (ARN) y polímeros de los mismos. Los polinucleótidos incluyen ADN genómico, ADNc y ADN antisentido, y ARNm empalmado o no empalmado, ARNr, ARNt y ADN o ARN inhibidor (ARNi, por ejemplo, ARN de horquilla pequeña o corta (sh), microARN (miARN), ARN de interferencia pequeño o corto (si), ARN de empalme en trans o ARN antisentido). Los polinucleótidos pueden incluir polinucleótidos que se presentan de manera natural, sintéticos e intencionalmente modificados o alterados (por ejemplo, ácido nucleico variante). Los polinucleótidos pueden ser de hebra individual, de doble hebra o de triple hebra, lineales o circulares y pueden ser de cualquier longitud adecuada. Al analizar los polinucleótidos, una secuencia o estructura de un polinucleótido particular se puede describir en la presente de acuerdo con la convención de proporcionar la secuencia en la dirección 5' a 3'.

Un ácido nucleico que codifica para un polipéptido frecuentemente comprende un marco de lectura abierto que codifica para el polipéptido. A menos que se indique de otro modo, una secuencia de ácido nucleico particular también incluye sustituciones de codones degenerados.

Los ácidos nucleicos pueden incluir uno o más elementos reguladores o de control de expresión unidos de manera operativa al marco de lectura abierto, donde el uno o más elementos reguladores se configuran para dirigir la transcripción y traducción del polipéptido codificado por el marco de lectura abierto en una célula de mamífero. Los ejemplos no limitantes de elementos de control/reguladores de expresión incluyen secuencias de iniciación de transcripción (por ejemplo, promotores, intensificadores, una secuencia TATA y similares), secuencias de iniciación de traducción, secuencias de estabilidad de ARNm, secuencias de poli A, secuencias secretoras y similares. Los elementos de control/reguladores de expresión se pueden obtener del genoma de cualquier organismo adecuado.

Un “promotor” se refiere a una secuencia de nucleótidos, usualmente en la dirección 5' (5') de una secuencia codificadora, que dirige y/o controla la expresión de la secuencia codificadora al proporcionar el reconocimiento para la ARN polimerasa y otros factores requeridos para la transcripción adecuada. Un promotor de pol II incluye un promotor mínimo que es una secuencia de ADN corta compuesta por una secuencia TATA y opcionalmente otras secuencias que sirven para especificar el sitio de iniciación de transcripción, al cual se adicionan elementos reguladores para el control de expresión. Un promotor de pol III tipo 1 incluye tres elementos de secuencia de acción cis en la dirección 3' del sitio de inicio transcripcional: a) elemento de secuencia 5' (bloque A); b) un elemento de secuencia intermedio (bloque I); c) elemento de secuencia 3' (bloque C). Un promotor de pol III tipo 2 incluye dos elementos de secuencia esenciales de acción cis en la dirección 3' del sitio de inicio de transcripción: a) una secuencia A (elemento de secuencia 5'); y b) una secuencia B (elemento de secuencia 3'). Un promotor de pol III tipo 3 incluye varios elementos promotores de acción cis en la dirección 5' del sitio de inicio de transcripción, tal como una secuencia TATA tradicional, un elemento de secuencia proximal (PSE) y un elemento de secuencia distal (DSE).

Un “intensificador” es una secuencia de ADN que puede estimular la actividad de transcripción y puede ser un elemento innato del promotor o un elemento heterólogo que intensifica el nivel o especificidad tisular de expresión. Es capaz de operar en cualquier orientación (5'->3' o 3' ->5'), y puede ser capaz de funcionar incluso cuando se coloca ya sea en la dirección 5' o en la dirección 3' del promotor.

Los promotores y/o intensificadores se pueden derivar en su totalidad de un gen nativo, o se pueden componer de diferentes elementos derivados de diferentes elementos encontrados en la naturaleza, o incluso se comprenden de segmentos de ADN sintético. Un

promotor o intensificador puede comprender secuencias de ADN que se implican en la unión de factores de proteína que modulan/controlan la efectividad de la iniciación de transcripción en respuesta a estímulos, condiciones fisiológicas o de desarrollo.

Los ejemplos no limitantes de promotores incluyen promotor temprano de SV40, promotor de LTR de virus de tumor mamario de ratón; promotor tardío principal de adenovirus (Ad MLP); un promotor de virus de herpes simple (HSV), un promotor de citomegalovirus (CMV) tal como la región de promotor temprano inmediato de CMV (CMVIE), un promotor de 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 metalotioneína murina, también encontrarán uso en la presente. Los ejemplos de promotores constitutivos incluyen los promotores para los siguientes genes que codifican para ciertas funciones constitutivas o “de mantenimiento”: hipoxantina fosforribosil transferasa (HPRT), dihidrofolato reductasa (DHFR), adenosina desaminasa, fosfoglicerol cinasa (PGK), piruvato cinasa, fosfoglicerol mutasa, promotor de actina, U6 y otros promotores constitutivos conocidos por aquellos expertos en la técnica. 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 la leucemia de Moloney y otros retrovirus; y el promotor de timidina cinasa del virus de herpes simple, entre muchos otros. Por consiguiente, cualquiera de los promotores constitutivos mencionados anteriormente se puede usar para controlar la transcripción de un inserto de gen heterólogo.

Un “transgén” se usa en la presente para referirse convenientemente a una secuencia de ácido nucleico/polinucleótido que se propone o se ha introducido en una célula u organismo. Los transgenes incluyen cualquier ácido nucleico, tal como un gen que codifica para un ARN inhibidor o polipéptido o proteína, y son en general heterólogos con respecto a secuencias genómicas de AAV que se presentan de manera natural.

El término “transducir” se refiere a la introducción de una secuencia de ácido nucleico en una célula u organismo hospedador por medio de un vector (por ejemplo, una partícula viral). Por lo tanto, la introducción de un transgén en una célula por una partícula viral se puede referir como “transducción” de la célula. El transgén se puede o no integrar en el ácido nucleico genómico de una célula transducida. Si un transgén introducido se integra en el ácido nucleico (ADN genómico) de la célula u organismo receptor se puede mantener de forma estable en esa célula u organismo y transmitirse adicionalmente a o heredarse por células de

progenie u organismos de la célula u organismo receptor. Finalmente, el transgén introducido puede existir en la célula receptora u organismo hospedador de forma extracromosómica, o solo de forma transitoria. Por lo tanto, una “célula transducida” es una célula en la que se ha introducido el transgén por medio de transducción. Por lo tanto, una célula “transducida” es una célula en la que, o una progenie de la misma en la que se ha introducido un transgén. Una célula transducida se puede propagar, transcribir por transgén y expresar el ARN o proteína inhibidora codificada. Para usos y métodos de terapia génica, una célula transducida puede estar en un mamífero.

Los transgenes bajo el control de promotores inducibles se expresan solo o en mayor grado, en la presencia de un agente inductor, (por ejemplo, la transcripción bajo el control del promotor de metalotioneína se incrementa en gran medida en la presencia de ciertos iones metálicos). Los promotores inducibles incluyen elementos sensibles (RE) que estimulan la transcripción cuando sus factores inductores están unidos. Por ejemplo, existen RE para factores séricos, hormonas esteroides, ácido retinoico y AMP cíclico. Los promotores que contienen un RE particular se pueden elegir a fin de obtener una respuesta inducible y en algunos casos, el propio RE se puede unir a un promotor diferente, confiriendo de este modo inducibilidad al gen recombinante. Por lo tanto, al seleccionar un promotor adecuado (constitutivo versus inducible; fuerte versus débil), es posible controlar tanto la existencia como el nivel de expresión de un polipéptido en la célula genéticamente modificada. Si el gen que codifica para el polipéptido está bajo el control de un promotor inducible, la administración del polipéptido *in situ* se desencadena al exponer la célula genéticamente modificada *in situ* a condiciones para permitir la transcripción del polipéptido, por ejemplo, por inyección intraperitoneal de inductores específicos de los promotores inducibles que controlan la transcripción del agente. Por ejemplo, la expresión *in situ* por células modificadas genéticamente de un polipéptido codificado por un gen bajo el control del promotor de metalotioneína, se mejora al poner en contacto las células modificadas genéticamente con una solución que contiene los iones metálicos apropiados (es decir, que inducen) *in situ*.

Un ácido nucleico/transgén se “enlaza de manera operativa” cuando se coloca en una relación funcional con otra secuencia de ácido nucleico. Un ácido nucleico/transgén que codifica para iARN o un polipéptido, o un ácido nucleico que dirige la expresión de un polipéptido puede incluir un promotor inducible o un promotor específico de tejido para controlar la transcripción del polipéptido codificado. Un ácido nucleico enlazado de manera operativa a un elemento de control de expresión también se puede referir como un casete de expresión.

En ciertas realizaciones, se emplean promotores, intensificadores y similares inducibles o específicos de tipo de célula en los métodos y usos descritos en la presente. Los ejemplos no taxativos de promotores específicos de tipo de célula incluyen aquellos aislados de los genes de TMC1, TMC2, Espin, PCDH15, PTPRQ, TMHS (LHFPL5), MYO1A. Los
5 ejemplos no limitantes de promotores inducibles incluyen elementos sensibles a ADN para ecdisona, tetraciclina, hipoxia e IFN.

En ciertas realizaciones, un elemento de control de expresión comprende un intensificador de CMV. En ciertas realizaciones, un elemento de control de expresión comprende un promotor de beta actina. En ciertas realizaciones, un elemento de control de
10 expresión comprende un promotor de beta actina de pollo. En ciertas realizaciones, un elemento de control de expresión comprende un intensificador de CMV y un promotor de beta actina de pollo.

Como se usa en la presente, los términos “modificar” o “variante” y variaciones gramaticales de los mismos, significan que un ácido nucleico, polipéptido o subsecuencia de
15 los mismos se desvía de una secuencia de referencia. Por lo tanto, las secuencias modificadas y variantes pueden tener sustancialmente la misma, mayor o menor expresión, actividad o función que una secuencia de referencia, pero al menos retener actividad o función parcial de la secuencia de referencia. Un tipo particular de variante es una proteína mutante, que se refiere a una proteína codificada por un gen que tiene una mutación, por ejemplo, una
20 mutación de sentido erróneo o sin sentido.

Una variante de “ácido nucleico” o “polinucleótido” se refiere a una secuencia modificada que se ha alterado genéticamente en comparación con el tipo silvestre. La secuencia se puede modificar genéticamente sin alterar la secuencia de proteína codificada. De manera alternativa, la secuencia se puede modificar genéticamente para codificar una
25 proteína variante. Un ácido nucleico o variante de polinucleótido también puede referirse a una secuencia de combinación que se ha modificado por codones para codificar una proteína que aún retiene al menos identidad de secuencia parcial con una secuencia de referencia, tal como una secuencia de proteína tipo silvestre, y también se ha modificado por codones para codificar una proteína variante. Por ejemplo, algunos codones de esta variante de ácido
30 nucleico se cambiarán sin alterar los aminoácidos de una proteína codificada por esta, y algunos codones de la variante de ácido nucleico se cambiarán, lo que a su vez cambia los aminoácidos de una proteína codificada por la misma.

Los términos “proteína” y “polipéptido” se usan indistintamente en la presente. Los

“polipéptidos” codificados por un “ácido nucleico” o “polinucleótido” o “transgén” divulgados en la presente incluyen secuencias nativas parciales o de longitud completa, como con proteínas polimórficas que se presentan de manera natural, tipo silvestre y funcionales, subsecuencias funcionales (fragmentos) de las mismas y variantes de secuencia de las mismas, siempre que el polipéptido conserve algún grado de función o actividad. Por consiguiente, en métodos y usos de la invención, no se requiere que estos polipéptidos codificados por secuencias de ácido nucleico sean idénticos a la proteína endógena que es defectuosa, o cuya actividad, función o expresión es insuficiente, deficiente o ausente en un mamífero tratado.

Los ejemplos no limitantes de modificaciones incluyen una o más sustituciones de nucleótidos o aminoácidos (por ejemplo, aproximadamente 1 a aproximadamente 3, aproximadamente 3 a aproximadamente 5, aproximadamente 5 a aproximadamente 10, aproximadamente 10 a aproximadamente 15, aproximadamente 15 a aproximadamente 20, aproximadamente 20 a aproximadamente 25, aproximadamente 25 a aproximadamente 30, aproximadamente 30 a aproximadamente 40, aproximadamente 40 a aproximadamente 50, aproximadamente 50 a aproximadamente 100, aproximadamente 100 a aproximadamente 150, aproximadamente 150 a aproximadamente 200, aproximadamente 200 a aproximadamente 250, aproximadamente 250 a aproximadamente 500, aproximadamente 500 a aproximadamente 750, aproximadamente 750 a aproximadamente 1000 o más nucleótidos o residuos).

Un ejemplo de una modificación de aminoácidos es una sustitución o una delección de aminoácidos conservadora. En realizaciones particulares, una secuencia modificada o variante retiene al menos parte de una función o actividad de la secuencia no modificada (por ejemplo, secuencia tipo silvestre).

Otro ejemplo de una modificación de aminoácido es un péptido de direccionamiento introducido en una proteína de cápside de una partícula viral. Se han identificado péptidos que se dirigen a vectores virales recombinantes, al sistema nervioso central, tal como a distintas regiones cerebrales.

Un virus recombinante así modificado puede unirse preferentemente a un tipo de tejido (por ejemplo, tejido del CNS) sobre otro tipo de tejido (por ejemplo, tejido hepático). En ciertas realizaciones, un virus recombinante que porta una proteína de cápside modificada puede “dirigirse” al tejido del epitelio vascular cerebral al unirse a un nivel más alto que una proteína de cápside no modificada comparable. Por ejemplo, un virus recombinante que tiene una proteína de cápside modificada puede unirse a tejido de epitelio vascular cerebral a un nivel

de 50 % a 100 % mayor que un virus recombinante no modificado.

Un “fragmento de ácido nucleico” es una porción de una molécula de ácido nucleico dada. El ácido desoxirribonucleico (ADN) en la mayoría de los organismos es el material genético, en tanto que el ácido ribonucleico (ARN) se implica en la transferencia de información contenida en el ADN a las proteínas. Los fragmentos y variantes de las secuencias de nucleótidos y proteínas divulgadas o proteínas de longitud parcial codificadas por las mismas también se abarcan por la presente invención. Por “fragmento” o “porción” se entiende una longitud completa o menor que la longitud completa de la secuencia de nucleótidos que codifica para, o la secuencia de aminoácidos de, un polipéptido o proteína. En ciertas realizaciones, el fragmento o porción es biológicamente funcional (es decir, retiene 5 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 99 % o 100 % de la actividad o función del tipo silvestre).

Una “variante” de una molécula es una secuencia que es sustancialmente similar a la secuencia de la molécula nativa. Para las secuencias de nucleótidos, las variantes incluyen aquellas secuencias que, debido a la degeneración del código genético, codifican para la secuencia de aminoácidos idéntica de la proteína nativa. Las variantes alélicas que se presentan de manera natural como estas se pueden identificar con el uso de técnicas de biología molecular, como, por ejemplo, con reacción en cadena de polimerasa (PCR) y técnicas de hibridación. Las secuencias de nucleótidos variantes también incluyen secuencias de nucleótidos sintéticamente derivadas, tal como aquellas generadas, por ejemplo, al usar mutagénesis dirigida al sitio, que codifican para la proteína nativa, así como aquellas que codifican para un polipéptido que tiene sustituciones de aminoácidos. En general, las variantes de secuencia de nucleótidos de la invención tendrán al menos 40 %, 50 %, 60 %, a 70 %, por ejemplo, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, a 79 %, en general al menos 80 %, por ejemplo, 81 %-84 %, al menos 85 %, por ejemplo, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, a 98 %, de identidad de secuencia a la secuencia de nucleótidos nativa (endógena). En ciertas realizaciones, la variante es biológicamente funcional (es decir, retiene 5 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 99 % o 100 % de la actividad o función del tipo silvestre).

“Variaciones conservadoras” de una secuencia de ácido nucleico particular se refiere a aquellas secuencias de ácido nucleico que codifican para secuencias de aminoácidos idénticas o esencialmente idénticas. Debido a la degeneración del código genético, un gran

número de ácidos nucleicos funcionalmente idénticos codifican para cualquier polipéptido dado. Por ejemplo, los codones CGT, CGC, CGA, CGG, AGA y AGG codifican para el aminoácido arginina. Por lo tanto, en cada posición donde una arginina se especifica por un codón, el codón se puede alterar a cualquiera de los correspondientes codones descritos sin
5 alterar la proteína codificada. Estas variaciones de ácido nucleico son “variaciones silenciosas”, que son una especie de “variaciones modificadas de manera conservadora”. Cada secuencia de ácido nucleico descrita en la presente que codifica para un polipéptido también describe cada posible variación silenciosa, excepto donde se indique lo contrario. Un experto en la técnica reconocerá que cada codón en un ácido nucleico (excepto ATG, que
10 normalmente es el único codón para metionina) se puede modificar para producir una molécula funcionalmente idéntica por técnicas estándar. Por consiguiente, cada “variación silenciosa” de un ácido nucleico que codifica para un polipéptido está implícita en cada secuencia descrita.

El término “identidad sustancial” de secuencias de polinucleótidos significa que un
15 polinucleótido comprende una secuencia que tiene al menos 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 % o 79 %, o al menos 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 % u 89 %, o al menos 90 %, 91 %, 92 %, 93 % o 94 %, o incluso al menos 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia, en comparación con una secuencia de referencia usando uno de los programas de alineación descritos usando parámetros estándar.
20 Un experto en la técnica reconocerá que estos valores se pueden ajustar adecuadamente para determinar la identidad correspondiente de las proteínas codificadas por dos secuencias de nucleótidos teniendo en cuenta la degeneración de codones, similitud de aminoácidos, posicionamiento de marco de lectura y similares. La identidad sustancial de secuencias de aminoácidos para estos propósitos normalmente significa una identidad de secuencia de al
25 menos 70 %, al menos 80 %, 90 % o incluso al menos 95 %.

El término “identidad sustancial” en el contexto de un polipéptido indica que un polipéptido comprende una secuencia con al menos 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, o 79 %, o 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, o 89 %, o al menos 90 %, 91 %, 92 %, 93 %, o 94 %, o incluso, 95 %, 96 %, 97 %, 98 % o 99 %, de
30 identidad de secuencia a la secuencia de referencia sobre una ventana de comparación especificada. Una indicación de que dos secuencias de polipéptidos son idénticas es que un polipéptido es inmunológicamente reactivo con anticuerpos generados contra el segundo polipéptido. Por lo tanto, un polipéptido es idéntico a un segundo polipéptido, por ejemplo, donde los dos péptidos difieren solo por una sustitución conservadora.

Por “mecanosensación” se quiere decir una respuesta a un estímulo mecánico. Tacto, audición y equilibrio de ejemplos de la conversión de un estímulo mecánico en una señal neuronal. La entrada mecanosensorial se convierte en una respuesta a un estímulo mecánico a través de un proceso denominado “mecanotransducción”.

5 Por “enfermedad” se entiende cualquier condición o trastorno que daña o interfiere con la función normal de una célula, tejido u órgano. Los ejemplos de enfermedades incluyen trastornos genéticos caracterizados por una pérdida de función en una proteína que funciona en la transducción mecanosensorial que se expresa, por ejemplo, en el oído interno de un sujeto. En otra realización, la enfermedad es síndrome de Usher (por ejemplo, USHI) o pérdida
10 de audición relacionada con la edad. En una realización, una enfermedad es un trastorno auditivo asociado con un defecto genético, tal como un defecto en TMC1, TMC2, MYO7A, USCHI C, CDH23, PCDH15, SANS, CIB2, USH2A, VLGR1, WHKN, CLRN1, PDZD7, USH1C (por ejemplo, armonina-a, b o c).

Los términos “tratar” y “tratamiento” se refieren tanto al tratamiento terapéutico como
15 a medidas profilácticas o preventivas, en donde el objeto es prevenir, inhibir, reducir o disminuir un cambio o trastorno fisiológico no deseado, tal como el desarrollo, progresión o empeoramiento del trastorno. Para propósitos de esta invención, los resultados clínicos beneficiosos o deseados incluyen, pero no se limitan a, alivio de síntomas, disminución de grado de enfermedad, estabilización de un síntoma o efecto adverso (es decir, no
20 empeoramiento o progresión) de enfermedad, retraso o desaceleración de progresión de enfermedad, mejora o paliación del estado de enfermedad y remisión (ya sea parcial o total), ya sea detectable o indetectable. “Tratamiento” también puede significar prolongar la supervivencia en comparación con la supervivencia esperada si no recibe tratamiento. Aquellos en necesidad de tratamiento incluyen aquellos ya con la condición o trastorno, así
25 como aquellos predispuestos (por ejemplo, como se determina por un ensayo genético).

Como se usa en la presente, “esencialmente libre”, en términos de un componente especificado, se usa en la presente para significar que ninguno del componente especificado se ha formulado intencionalmente en una composición y/o está presente solo como un contaminante o en cantidades traza. La cantidad total del componente especificado que
30 resulta de cualquier contaminación no deseada de una composición es por lo tanto, muy por debajo de 0,05 %, preferentemente por debajo de 0,01 %. Lo más preferido es una composición en la cual no se pueda detectar ninguna cantidad del componente especificado con métodos analíticos estándar.

Como se usa en la presente especificación, “un” o “una” puede significar uno o más. Como se usa en la presente en la o las reivindicaciones, cuando se usa en conjunto con la palabra “que comprende”, las palabras “un” o “una” pueden significar uno o más de uno.

5 El uso del término “o” en las reivindicaciones se usa para significar “y/o” a menos que se indique explícitamente que se refiere solamente a alternativas o las alternativas se excluyen mutuamente, aunque la divulgación soporta una definición que se refiere solamente a alternativas y a “y/o”. Como se usa en la presente, “otro” puede significar al menos un segundo o más.

10 En todo momento de esta solicitud, el término “aproximadamente” se usa para indicar que un valor incluye la variación inherente de error para el dispositivo, la variación inherente en el método que se emplea para determinar el valor, la variación que existe entre los sujetos de estudio o un valor que está dentro de 10 % de un valor establecido.

VI. Kits

15 La invención proporciona kits con material de envase y uno o más componentes en los mismos. Un kit incluye habitualmente una etiqueta o prospecto que incluye una descripción de los componentes o instrucciones para uso *in vitro*, *in vivo* o *ex vivo*, de los componentes en el mismo. Un kit puede contener una recolección de estos componentes, por ejemplo, un ácido nucleico, vector recombinante y/o partículas virales.

20 Un kit se refiere a una estructura física que aloja uno o más componentes del kit. El material de envase puede mantener los componentes de forma estéril y se puede elaborar de material comúnmente usado para estos propósitos (por ejemplo, papel, fibra corrugada, vidrio, plástico, lámina, ampollas, viales, tubos, etc.).

25 Las etiquetas o prospectos pueden incluir información de identificación de uno o más componentes en los mismos, cantidades de dosis, farmacología clínica del o de los principios activos, incluido el mecanismo de acción, farmacocinética y farmacodinámica. Las etiquetas o prospectos pueden incluir información que identifique al fabricante, números de lote, ubicación y fecha de fabricación, fechas de vencimiento. Las etiquetas o prospectos pueden incluir información que identifique la información del fabricante, números de lote, ubicación del fabricante y fecha. Las etiquetas o prospectos pueden incluir información sobre una
30 enfermedad para la cual se puede usar un componente del kit. Las etiquetas o prospectos pueden incluir instrucciones para el médico o sujeto para usar uno o más de los componentes

del kit en un método, uso o protocolo de tratamiento o régimen terapéutico. Las instrucciones pueden incluir cantidades de dosis, frecuencia o duración, e instrucciones para practicar cualquiera de los métodos, usos, protocolos de tratamiento o regímenes profilácticos o terapéuticos descritos en la presente.

5 Las etiquetas o prospectos pueden incluir información sobre cualquier beneficio que pueda proporcionar un componente, tal como un beneficio profiláctico o terapéutico. Las etiquetas o prospectos pueden incluir información sobre posibles efectos secundarios adversos, complicaciones o reacciones, tal como advertencias al sujeto o al médico con respecto a situaciones en las que no sería apropiado usar una composición particular. Los
10 efectos secundarios adversos o complicaciones también se pueden presentar cuando el sujeto tiene, estará o está tomando actualmente uno o más medicamentos que pueden ser incompatibles con la composición, o el sujeto tiene, estará o está experimentando actualmente otro protocolo de tratamiento o régimen terapéutico que sería incompatible con la composición y por lo tanto, las instrucciones pueden incluir información con respecto a estas
15 incompatibilidades.

Las etiquetas o prospectos incluyen “material impreso”, por ejemplo, papel o cartón, o separado o fijado a un componente, un kit o material de envase (por ejemplo, una caja), o unido a una ampolla, tubo o vial que contiene un componente del kit. Las etiquetas o prospectos pueden incluir adicionalmente un medio leíble por computadora, tal como una
20 etiqueta impresa con código de barras, un disco, disco óptico tal como CD o DVD-ROM/RAM, DVD, MP3 o un medio de almacenamiento eléctrico tal como RAM y ROM o híbridos de estos tal como medio de almacenamiento magnético/óptico, memoria FLASH, híbridos y tarjetas de tipo memoria.

VII. Ejemplos

25 Los siguientes ejemplos se incluyen para demostrar las realizaciones preferidas de la invención. Se debe apreciar por aquellos expertos en la técnica que las técnicas divulgadas en los ejemplos que siguen representan técnicas descubiertas por el inventor para funcionar bien en la práctica de la invención, y por lo tanto se puede considerar que constituyen modos preferidos para su práctica. Sin embargo, los expertos en la técnica deben, a la luz de la
30 presente divulgación, apreciar que se pueden hacer muchos cambios en las realizaciones específicas que se revelan y aún así obtener un resultado similar o similar sin apartarse del espíritu y el alcance de la divulgación.

Ejemplo 1 - Administración terapéutica al Globus Pallidus

Se generaron bibliotecas de variantes de cápside de AAV modificadas con péptidos por inserción de péptidos aleatorios en AAV1, AAV2 y AAV9 por inserción de secuencias de péptidos de direccionamiento aleatorio en la posición 590 de cápside de AAV1, posición 587 de cápside de AAV2 y posición 588 de cápside de AAV9, respectivamente. Estas bibliotecas se examinaron en macacos rhesus para detectar variantes que transducen ampliamente estructuras cerebrales profundas y corteza cerebral después de la infusión en el globus pallidus (figuras 38A-38C). Se seleccionaron seis variantes (tabla 1) para validación adicional en un experimento de fluorescencia en dos macacos rhesus adicionales. Al menos cinco variantes de cápside transdujeron estructuras cerebrales profundas afectadas en HD y PD. Las transducciones de AAV1.THTD.mTFP1 fueron generalizadas en el caudado/putamen. AAV1.THTD.mTFP1 también transdujo un gran número de neuronas de proyección V/VI de capa cortical (figura 39), que son altamente vulnerables en HD. Estas cápsides de AAV se pueden usar para empaquetar una variedad de cargas útiles terapéuticas para HD y PD. La transducción hepática para estas variantes estuvo por debajo del límite de detección (figura 40) (AAV-BLD010 es AAV1.THTD.mTFP1; AAV-BLD011 es AAV2.SGGR.NES_mNG; AAV-BLD012 es AAV2.SKGQ.mRuby3).

Tabla 1. Péptidos de direccionamiento a cápside.

Serotipo	ADN	Péptido
AAV1	TCTAGTGGGGATTTCGACGAAG (SEQ ID NO: 7)	SSGDSTK (SEQ ID NO: 8)
AAV1	ACGCATACTGATCATACGCGT (SEQ ID NO: 9)	THTDDTR (SEQ ID NO: 10)
AAV2	TCTAAGGGGCAGGGGACGGGG (SEQ ID NO: 11)	SKGQGTG (SEQ ID NO: 12)
AAV2	TCGGGGGGGAGGTATGCGGAG (SEQ ID NO: 13)	SGGRYAE (SEQ ID NO: 14)
AAV9	TCTTTGGGTAGTACTAAGGTG (SEQ ID NO: 15)	SLGSTKV (SEQ ID NO: 16)
AAV9	AAGATGTTGAGTGTGTTAAT (SEQ ID NO: 17)	KMLSVVN (SEQ ID NO: 18)

Para analizar adicionalmente las variantes seleccionadas, la variante de AAV1 con el péptido de direccionamiento THTDDTR (SEQ ID NO: 20) se empaquetó con un casete de expresión de mTFP1 (AAV1.THTD.mTFP1), la variante de AAV2 con el péptido de direccionamiento SGGRYAE (SEQ ID NO: 22) se empaquetó con un casete de expresión mNeonGreen (AAV2.SGGR.NES_mNG), la variante de AAV2 con el péptido de direccionamiento SKGQGTG (SEQ ID NO: 21) se empaquetó con un casete de expresión de mRuby3 (AAV2.SKGQ.mRuby3), la variante de AAV9 con el péptido de direccionamiento KMLSVVN (SEQ ID NO: 24) se empaquetó con un casete de expresión Halo-tag (AAV9KMLS.Halo-tag) y la variante de AAV9 con el péptido de direccionamiento SLGSTKV (SEQ ID NO: 23) se empaquetó con un casete de expresión de mBFP2 (AAV9SLGS.mBFP2).

Dos macacos rhesus adultos se inyectaron con una dosis de 7,5E10 vg total de ciertas variantes de cápside en el globus pallidus. Se realizaron estudios de imagenología para determinar las propiedades de biodistribución y tráfico de cada una de las cápsides después de la administración al globus pallidus. Se usaron imágenes FISH de 10um z-stacks de tejido para localizar ARNm de cada fluoróforo específico para su respectiva cápside variante (figuras 1, 4, 7, 9, 11, 12, 15, 17, 18, 20 y 25). Se usaron imágenes de epifluorescencia para localizar la proteína de cada fluoróforo específico para su respectiva variante de cápside (figuras 2, 3, 5, 6, 8, 10, 13, 14, 16, 17, 19, 21-24 y 26).

Ejemplo 2 - Inyección en ratón de vectores de globus pallidus

Tres ratones recibieron inyecciones de diferentes virus en cada hemisferio de una estructura anatómica designada: el tálamo, la cápsula interna/globus pallidus medial (IC/MGP) o núcleos cerebelosos profundos (DCN) representados en la tabla 2. Los diferentes virus fueron: RVC1023, pmAAV2.SGGR.NeonGreen; RVC1024, pmAAV2.SKGO.mRuby; y RVC1025, pmAAV1.THTD.mTFP. Los ratones se sometieron a necropsia 22 días después de la inyección y se valoraron por microscopía epifluorescente (figuras 27-37).

Tabla 2. Paradigma de inyección en ratón de vectores de Globus Pallidus.

ID de ratón	ROA	Izquierda	Derecha	vg/hemisferio
4279	Tálamo	NeonGreen	mRuby	5,00E+09
4280	Tálamo	mRuby	TFP	5,00E+09
4281	Tálamo	TFP	NeonGreen	5,00E+09
4282	IC/MGP	NeonGreen	mRuby	4,00E+09
4285	IC/MGP	mRuby	TFP	4,00E+09

4292	IC/MGP	TFP	NeonGreen	4,00E+09
4293	IC/MGP	mRuby	TFP	4,00E+09
A16	DCN	NeonGreen	mRuby	4,00E+09
A17	DCN	mRuby	TFP	4,00E+09
A18	DCN	TFP	NeonGreen	4,00E+09

Todos los métodos divulgados y reivindicados en la presente se pueden realizar y ejecutar sin experimentación indebida a la luz de la presente divulgación. En tanto que las composiciones y métodos de esta invención se han descrito en términos de realizaciones preferidas, será evidente para aquellos expertos en la técnica que se pueden aplicar variaciones a los métodos y en los pasos o en la secuencia de pasos del método descrito en la presente sin apartarse del concepto, espíritu y alcance de la invención. Más específicamente, será evidente que ciertos agentes que están tanto química como fisiológicamente relacionados se pueden sustituir por los agentes descritos en la presente en tanto que se logran los mismos resultados o resultados similares. Todos estos sustitutos y modificaciones similares evidentes para aquellos expertos en la técnica se consideran que están dentro del espíritu, alcance y concepto de la invención como se define por las reivindicaciones anexas.