

VARIANTES DEL VECTOR VIRAL ADENOASOCIADO

REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS

[0001] Esta solicitud reclama el beneficio de prioridad de la Solicitud de Patente Provisional de EE. UU. No. 62/939,315 presentada el 22 de Noviembre de 2019 y la No. 63/084,709 presentada el 29 de Septiembre de 2020 que se incorpora en el presente documento como referencia en su totalidad.

REFERENCIA A UN LISTADO DE SECUENCIAS

[0002] La solicitud instantánea contiene una Lista de Secuencias, que se ha enviado en formato ASCII a través de EFS-Web y se incorpora al presente documento como referencia en su totalidad. Dicha copia ASCII, creada el 19 de noviembre de 2020, se denomina CHOPP0038WO_ST25.txt y tiene un tamaño de 63,8 kilobytes.

ANTECEDENTES DE LA INVENCION

1. Campo de la invención

[0003] La presente invención se refiere en general a los campos de la inmunología, microbiología y patología. Más particularmente, se trata de péptidos dirigidos que dirigen la entrega de vectores virales a distintas estructuras en el cerebro.

2. Descripción de la técnica relacionada

[0004] Se han desarrollado diferentes estrategias para generar variantes de vectores AAV (por sus siglas en inglés) que incluyen diseño racional y evolución dirigida. El enfoque de diseño racional utiliza el conocimiento de las cápsides de AAV para realizar cambios específicos en la cápside para alterar la eficiencia o especificidad de la transducción, como mutaciones de tirosina en la superficie de la cápside para aumentar la eficiencia de la transducción. El enfoque de evolución dirigida no requiere ningún conocimiento de la estructura de la cápside y se realiza mediante mutagénesis aleatoria, mezcla de cápside o inserciones aleatorias de péptidos. Estas estrategias generalmente utilizan sistemas in vitro o ratones, que son ideales para estudios basados en células o en ratones, pero no implican traslado a la clínica. De hecho, ninguna variante de AAV se dirige a distintas estructuras cerebrales de manera específica o eficiente. Como tal, se necesitan variantes de AAV que puedan dirigirse a distintas estructuras cerebrales de primates.

BREVE DESCRIPCION DE LA INVENCION

[0005] En el presente documento se proporcionan vectores virales, cada uno de los cuales comprende 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.

[0006] En una realización, se proporcionan proteínas de la cápside del virus adenoasociado modificado (AAV por sus siglas en inglés) que comprenden un péptido de dirección que se dirige a un vector viral que comprende la proteína de la cápside de AAV modificada a una estructura distinta del órgano o

cerebro y el péptido de dirección tiene una longitud de tres a diez aminoácidos. En algunos aspectos, las proteínas de la cápside de AAV modificadas son proteínas de la cápside de AAV1 modificadas, proteínas de la cápside de AAV2 modificadas o proteínas de la cápside de AAV9 modificadas.

[0007] En algunos aspectos, las proteínas de la cápside de AAV modificadas se derivan de una proteína de la cápside de AAV1 (ver SEQ ID NO: 138) y el péptido dirigido se inserta después del residuo 590 de la proteína de la cápside de AAV1. En algunos aspectos, el péptido dirigido está flanqueado por secuencias conectoras y las secuencias conectoras a cada lado de los péptidos dirigidos tienen una longitud de dos o tres aminoácidos. En algunos aspectos, las secuencias conectoras son SSA en el lado N-terminal del péptido dirigido y AS en el lado C-terminal del péptido dirigido. En algunos aspectos, las proteínas de la cápside de AAV1 modificadas tienen una secuencia idéntica en al menos un 95% a la SEQ ID NO: 141.

[0008] En algunos aspectos, las proteínas de la cápside de AAV modificadas se derivan de una proteína de la cápside de AAV2 (ver SEQ ID NO: 139) y el péptido dirigido se inserta después del residuo 587 de la proteína de la cápside de AAV2. En algunos aspectos, el péptido dirigido está flanqueado por secuencias conectoras y las secuencias conectoras a cada lado de los péptidos dirigidos tienen una longitud de dos o tres aminoácidos. En algunos aspectos, las secuencias conectoras son AAA en el lado N-terminal del péptido dirigido y AA en el lado C-terminal del péptido dirigido. En algunos aspectos, las proteínas de la cápside de AAV2 modificadas tienen una secuencia idéntica en al menos un 95% a la SEQ ID NO: 142.

[0009] En algunos aspectos, las proteínas de la cápside de AAV modificadas se derivan de una proteína de la cápside de AAV9 (ver SEQ ID NO: 140) y el péptido dirigido se inserta después del residuo 588 de la proteína de la cápside de AAV9. En algunos aspectos, el péptido dirigido está flanqueado por secuencias conectoras y las secuencias conectoras a cada lado de los péptidos dirigidos tienen una longitud de dos o tres aminoácidos. En algunos aspectos, las secuencias conectoras son AAA en el lado N-terminal del péptido dirigido y AS en el lado C-terminal del péptido dirigido. En algunos aspectos, las proteínas de la cápside de AAV9 modificadas tienen una secuencia idéntica en al menos un 95% a la SEQ ID NO: 143.

[0010] 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 consta de las SEQ ID NO: 1-137 y 144. En algunos aspectos, el péptido dirigido tiene una longitud de siete aminoácidos.

[0011] En algunos aspectos, la estructura cerebral distintiva es el tronco encefálico, el caudado, la corteza cerebelosa, la corteza cerebral, el epéndimo, el globo pálido, el hipocampo, las meninges, el nervio óptico, el putamen, la médula espinal, la sustancia negra, los núcleos subtalámicos o el tálamo. En ciertos aspectos, la proteína de la cápside de AAV modificada es una proteína de la cápside de AAV1 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 1 para apuntar a una estructura cerebral correspondiente. En ciertos aspectos, la proteína de la cápside de AAV modificada es una proteína de la cápside de AAV2 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 2 para apuntar a una estructura cerebral correspondiente. En ciertos aspectos, la proteína de la cápside de AAV modificada

es una proteína de la cápside de AAV9 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 3 para apuntar a una estructura cerebral correspondiente.

[0012] En algunos aspectos, el órgano diferenciado es el cerebro, el riñón, el corazón, el hígado, las gónadas, el bazo o el hígado. En ciertos aspectos, la proteína de la cápside de AAV modificada es una proteína de la cápside de AAV1 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 4 para dirigirse a un órgano correspondiente. En ciertos aspectos, la proteína de la cápside de AAV modificada es una proteína de la cápside de AAV2 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 5 para dirigirse a un órgano correspondiente. En ciertos aspectos, la proteína de la cápside de AAV modificada es una proteína de la cápside de AAV9 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 6 para dirigirse a un órgano correspondiente.

[0013] En una realización, en el presente documento se proporcionan ácidos nucleicos que comprenden una secuencia que codifica la proteína de la cápside modificada de cualquiera de las presentes realizaciones.

[0014] En una realización, en el presente documento se proporcionan virus adenoasociados recombinantes (rAAV, por sus siglas en inglés) que comprenden la proteína de la cápside modificada de cualquiera de las presentes realizaciones. En algunos aspectos, se proporcionan combinaciones de rAAV. Por ejemplo, la combinación de un rAAV que tiene una proteína de la cápside de AAV1 modificada y un péptido dirigido de la SEQ ID NO: 21, un rAAV que tiene una proteína de la cápside de AAV2 modificada y un péptido dirigido de la SEQ ID NO: 53, un rAAV que tiene una proteína de la cápside de AAV2 modificada y un péptido dirigido de la SEQ ID NO: 80 y un rAAV que tiene una proteína de la cápside de AAV9 modificada y un péptido dirigido de la SEQ ID NO: 113 se proporciona.

[0015] En una realización, en el presente documento se proporcionan vectores virales que comprenden un ácido nucleico que codifica la proteína de la cápside modificada de cualquiera de las presentes realizaciones. En algunos aspectos, los vectores virales comprenden además una secuencia de ácido nucleico que codifica 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 ARNi.

[0016] En una realización, en el presente documento se proporcionan células que comprenden el vector viral de cualquiera de las presentes realizaciones. En algunas realizaciones, la célula es una célula de mamífero, tal como una célula de humano. En algunos aspectos, la célula es in vitro o in vivo.

[0017] En una realización, en el presente documento se proporcionan composiciones farmacéuticas que comprenden el vector viral de las presentes realizaciones y un vehículo farmacéuticamente aceptable.

[0018] En una realización, en el presente documento se proporcionan métodos para administrar un agente a una estructura cerebral distintiva de un sujeto, que comprende administrar un virus de las presentes

realizaciones al sujeto. En algunos aspectos, la estructura cerebral distintiva es el tronco encefálico, el caudado, la corteza cerebelosa, la corteza cerebral, el epéndimo, el globo pálido, el hipocampo, las meninges, el nervio óptico, el putamen, la médula espinal, la sustancia negra, los núcleos subtalámicos o el tálamo. En ciertos aspectos, se usa un rAAV que tiene una proteína de la cápside de AAV1 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 1 para apuntar a una estructura cerebral correspondiente. En ciertos aspectos, se usa un rAAV que tiene una proteína de la cápside de AAV2 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 2 para apuntar a una estructura cerebral correspondiente. En ciertos aspectos, se usa un rAAV que tiene una proteína de la cápside de AAV9 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 3 para apuntar a una estructura cerebral correspondiente. En varios aspectos, se utilizan combinaciones de cualquiera de los rAAV. Por ejemplo, la combinación de un rAAV que tiene una proteína de la cápside de AAV1 modificada y un péptido dirigido de la SEQ ID NO: 21, un rAAV que tiene una proteína de la cápside de AAV2 modificada y un péptido dirigido de la SEQ ID NO: 53, un rAAV que tiene una proteína de la cápside de AAV2 modificada y un péptido dirigido de la SEQ ID NO: 80 y un rAAV que tiene una proteína de la cápsida de AAV9 modificada y un péptido dirigido de la SEQ ID NO: Se utiliza 113.

[0019] En una realización, en el presente documento se proporcionan métodos para administrar un agente a un órgano distinto de un sujeto, que comprende administrar un virus de las presentes realizaciones al sujeto. En algunos aspectos, el órgano es el cerebro, riñón, corazón, hígado, gónada, bazo o hígado. En ciertos aspectos, se usa un rAAV que tiene una proteína de la cápside de AAV1 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 4 para apuntar a un órgano correspondiente. En ciertos aspectos, se usa un rAAV que tiene una proteína de la cápside de AAV2 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 5 para apuntar a un órgano correspondiente. En ciertos aspectos, se usa un rAAV que tiene una proteína de la cápside de AAV9 modificada y el péptido dirigido se selecciona de los enumerados en la Tabla 6 para apuntar a un órgano correspondiente. En varios aspectos, se utilizan combinaciones de cualquiera de los rAAV.

[0020] En algunos aspectos, el agente es un siARN, shARN, miARN, ARN no codificante, lncARN, 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 epéndimo, un ventrículo cerebral, un espacio subaracnoideo y/o un espacio intratecal. En algunos aspectos, el ventrículo cerebral es el ventrículo lateral 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.

[0021] En algunos aspectos, se administra una pluralidad de partículas virales. En algunos aspectos, el virus se administra a una dosis de aproximadamente 1×10^6 a aproximadamente 1×10^{18} genomas de vectores por kilogramo (vg/kg). En algunos aspectos, el virus se administra a una dosis de aproximadamente 1×10^7 - 1×10^{17} , alrededor de 1×10^8 - 1×10^{16} , alrededor de 1×10^9 - 1×10^{15} , alrededor de 1×10^{10} -

1×10^{14} , alrededor de 1×10^{10} - 1×10^{13} , alrededor de 1×10^{10} - 1×10^{13} , alrededor de 1×10^{10} - 1×10^{11} , alrededor de 1×10^{11} - 1×10^{12} , alrededor de 1×10^{12} - 1×10^{13} , o alrededor de 1×10^{13} - 1×10^{14} vg/kg del paciente. En algunos aspectos, el sujeto es humano.

[0022] En una realización, en el presente documento se proporcionan métodos para tratar una enfermedad en un mamífero que comprende 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 la enfermedad de Huntington, ALS, hemiplejía espástica hereditaria, esclerosis lateral primaria, atrofia muscular espinal, enfermedad de Kennedy, enfermedad de Alzheimer, enfermedad de repetición de poliglutamina o enfermedad de Parkinson. En algunos aspectos, el mamífero es humano.

[0023] Como se usa aquí, "esencialmente libre", en términos de un componente específico, significa que ninguno de los componentes especificados ha sido formulado a propósito en una composición y/o está presente solo como contaminante o en cantidades traza. Por lo tanto, la cantidad total del componente especificado resultante de cualquier contaminación no intencionada de una composición está muy por debajo del 0.05%, preferiblemente por debajo del 0.01%. La más preferida es una composición en donde no se puede detectar ninguna cantidad del componente especificado con métodos analíticos estándar.

[0024] Como se usa aquí en la especificación, "un", "una" o "uno" puede significar uno, una o más. Como se usa aquí en la(s) reivindicación(es), cuando se usa junto con la palabra "que comprende", las palabras "un" "una" o "uno" pueden significar uno o más de uno.

[0025] El uso del término "o" en las reivindicaciones se utiliza para significar "y/o" a menos que se indique explícitamente que se refiere únicamente a alternativas o que las alternativas sean mutuamente excluyentes, aunque la divulgación respalda una definición que se refiere únicamente a alternativas e "y/o." Como se usa aquí, "otro" puede significar al menos un segundo o más.

[0026] A lo largo de esta solicitud, el término "sobre" se usa para indicar que un valor incluye la variación inherente del error del dispositivo, el método que se emplea para determinar el valor, la variación que existe entre los sujetos del estudio o un valor que está dentro de 10% de un valor declarado.

[0027] Otros objetos, características y ventajas de la presente invención resultarán evidentes a partir de la siguiente descripción detallada. Debe entenderse, sin embargo, que la descripción detallada y los ejemplos específicos, aunque indican realizaciones específicas de la invención, se dan únicamente a modo de ilustración ya que varios cambios y modificaciones dentro del espíritu y alcance de la invención resultarán evidentes para aquellos expertos en la técnica a partir de esta descripción detallada.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0028] Las siguientes figuras forman parte del presente capítulo descriptivo y se incluyen para demostrar aún más ciertos aspectos de la presente invención. La invención puede entenderse mejor con referencia a una o más de estas figuras en combinación con la descripción detallada de las realizaciones específicas presentadas en este documento.

[0029] **FIG. 1.** Esquema de la biblioteca de visualización de péptidos AAV.

[0030] **FIG. 2** Esquema de estrategia de pantalla *in vivo*.

[0031] **FIG. 3.** Representación gráfica de la diversidad de la biblioteca de entrada. Diversidad de la biblioteca viral de entrada medida a partir de alícuotas de los vectores virales AAV1, AAV2 y AAV9 antes de la inyección ICV de la primera ronda.

[0032] **FIG. 4.** Representación gráfica del enriquecimiento de código de barras por rondas. El número total de códigos de barras únicos recuperados después del enriquecimiento de la Ronda 1 y la Ronda 2 en el macaco Rhesus para cada tejido recolectado. Se muestran los valores de la Ronda 2 para el ADN y el ARN.

[0033] **FIG. 5.** Ilustración del enriquecimiento circular de los códigos de barras en los serotipos AAV1, AAV2 y AAV9 para la corteza cerebelosa.

[0034] **FIG. 6.** Ilustración de enriquecimiento para AAV9 1999. Representaciones de mapa de calor del enriquecimiento de código de barras de AAV9, las células están coloreadas por el porcentaje de códigos de barras detectados del tejido indicado. Los códigos de barras recuperados del ADN se muestran a la izquierda, mientras que los códigos de barras recuperados del ARN están a la derecha.

[0035] **FIGS. 7A-C.** Representaciones de mapas de calor del enriquecimiento del código de barras opool de AAV1 (FIG. 7A), AAV2 (FIG. 7B) y AAV9 (FIG. 7C).

[0036] **FIG. 8.** Validación *in vivo* del AAV9 1999 en macaco Rhesus. Se empaquetó una construcción de expresión de eGFP en el AAV9 1999 bajo control del promotor CAG. Se administraron 1.5E13 vg de AAV9 1999 a un macaco Rhesus hembra de 5 años mediante inyección ICV. Se muestran imágenes representativas del cerebelo teñido con H&E que representan el patrón de transducción de AAV9 1999.

[0037] **FIGS. 9A-D.** Validación *in vivo* del AAV9 1999 en ratón. Se empaquetó una construcción de expresión de eGFP en el AAV9 1999 bajo control del promotor CAG. Las cápsidas del AAV9 1999 y el AAV9 que contenían la construcción eGFP se administraron a crías de ratón C57BL/6 p0 mediante inyección ICV a 1E10 vg. Las imágenes representativas de la señal de fluorescencia de eGFP son de todo el cerebro (FIG. 9A), sección sagital de todo el cerebro (FIG. 9B), secciones de la corteza S1 (FIG. 9C, izquierda),

secciones de hipocampo (FIG. 9C, centro), sección sagital del cerebelo (FIG. 9C, derecha) y sección coronal de la médula espinal lumbar (FIG. 9D).

[0038] FIGS. 10A-C. Imágenes fluorescentes de la mezcla AAV *in vivo* en el ventrículo lateral del macaco Rhesus (A), cuarto ventrículo (B) y meninges (C).

5 **[0039] FIGS. 11A-C.** Imágenes fluorescentes de giros cocleares (A), células ciliadas inteARNs (B), órgano de corti (C) y modiolos distales (C) después de la administración coclear de cápsidas AAV9 1999 que contienen la construcción eGFP a ratones.

DESCRIPCIÓN DETALLADA DE LA INVENCION

[0040] En el presente documento se proporcionan vectores virales, cada uno de los cuales comprende 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 ciertas realizaciones, la estructura cerebral es el tronco encefálico, el caudado, la corteza cerebelosa, la corteza cerebral, el epéndimo, el globo pálido, el hipocampo, las meninges, el nervio óptico, el putamen, la médula espinal, la sustancia negra, los núcleos subtalámicos o el tálamo. Los péptidos dirigidos para cada estructura cerebral se proporcionan en las Tablas 1-3.

[0041] En algunas realizaciones, el vector viral es un vector viral adenoasociado (AAV por sus siglas en inglés). En ciertas realizaciones, el AAV es AAV1, AAV2 o AAV9. En la SEQ ID NO: 138 se proporciona una secuencia de proteína de la cápside de AAV1 de referencia de tipo silvestre. En la SEQ ID NO: 139 se proporciona una secuencia de proteína de la cápside de AAV2 de referencia de tipo silvestre. En la SEQ ID NO: 140 se proporciona una secuencia de proteína de cápside de AAV9 de referencia de tipo silvestre. En ciertos aspectos, el péptido dirigido se inserta en la posición 590 de la cápside de AAV1, en la posición 587 de la cápside de AAV2 o en la posición 588 de la cápside de AAV9. Un ejemplo de secuencia de proteína de la cápside de AAV1 modificada se proporciona en la SEQ ID NO: 141, que muestra la inserción del péptido objetivo después de la posición 590 como SSAX₇AS, donde la SSA inicial y la AS final son secuencias conectoras y X₇ representa el péptido objetivo. Una secuencia ejemplar de proteína de cápside de AAV2 modificada se proporciona en la SEQ ID NO: 142, que muestra la inserción del péptido objetivo después de la posición 587 como AAAX₇AA, donde la AAA inicial y la AA final son secuencias conectoras y X₇ representa el péptido objetivo. Una secuencia ejemplar de proteína de cápside de AAV9 modificada se proporciona en la SEQ ID NO: 143, que muestra la inserción del péptido objetivo después de la posición 588 como AAAX₇AS, donde la AAA inicial y la AS final son secuencias conectoras y X₇ representa el péptido objetivo.

Tabla 1. Péptidos dirigidos al AAV1 para cada estructura cerebral.

Región	Péptido	SEQ ID NO.
Tronco encefálico	RPGREQA	1
	RGVLVTT	2
	RPGRESA	3
	NESLKKK	4
	DKTRAGS	5
	TAKSKQA	6
	PVKKKDA	7
	GRETLLG	8
	PIPAGKK	9
Caudado	RPGRESA	3
	PVKKKDA	7

Región	Péptido	SEQ ID NO.
	RPGREQA	1
	NVVRAGT	10
	KATANTR	11
	RDATRSS	12
	VPTKSPK	13
	AGVARSK	14
	RSRSEVL	15
	EVKGK GK	16
	RPGREAS	144
	RDSTRQL	26
	SGVLVQR	32
	DKTRAGS	5
	TAPKSLK	25
Corteza cerebelosa	AKLNKSS	17
	NESLKKK	4
	NSVHNTA	18
	NVVRGGA	19
	NRLVAGG	20
	RPGRESA	3
	ERDRTRG	21
	PIPAGKK	9
	RPGREQA	1
Corteza cerebelosa; vermis	RPGREAS	144
	DKTRAGS	5
	NSVRPLT	23
	RDSTRQL	26
	RDATRSS	12
Cerebelo; vermis superior derecho	NSVRPLT	23
	RDATRSS	12
	DKTRAGS	5
	TAPKSLK	25
	RPGREAS	144
Cerebelo; lateral al tronco encefálico (derecho)	NSVRPLT	23
	TAPKSLK	25
	RDSTRQL	26
	RPGREAS	144
	PIPAGKK	9
	RDATRSS	12
Cerebelo; hemisferio izquierdo, adyacente a el vermis superior	RPGREAS	144
	DKTRAGS	5
	RDATRSS	12
	RDSTRQL	26
	TAPKSLK	25

Región	Péptido	SEQ ID NO.
Cerebelo; hemisferio izquierdo, punto más lateral	RPGREAS	144
	RDSTRQL	26
	DKTRAGS	5
	TAPKSLK	25
	NSVRPLT	23
	RDATRSS	12
Cerebelo; hemisferio izquierdo, adyacente al tronco cerebral	NSVRPLT	23
	RPGREAS	144
	RDATRSS	12
	DKTRAGS	5
	NSVHNTA	18
	TAPKSLK	25
Corteza cerebral	VQGSKMK	22
	RDATRSS	12
	NSVRPLT	23
	RPGRESA	3
	NKIHANP	24
	ERDRTRG	21
	DKTRAGS	5
	TAPKSLK	25
	RDSTRQL	26
	RPGREQA	1
Cóclea (oído)	RDATRSS	12
	RPGREAS	144
	TAPKSLK	25
	DKTRAGS	5
	NSVRPLT	23
	TATPRKG	49
Núcleos cerebelosos profundos	RPGREAS	144
	TAPKSLK	25
	RDATRSS	12
	RDSTRQL	26
	DKTRAGS	5
Ganglios de la raíz dorsal, lumbares	RPGREAS	144
	RDATRSS	12
	NSVRPLT	23
	DKTRAGS	5
	TAPKSLK	25
Ganglios de la raíz dorsal, cervicales	RPGREAS	144
	DKTRAGS	5
	TAPKSLK	25
	RDATRSS	12
	NSVHNTA	18

Región	Péptido	SEQ ID NO.
	NSVRPLT	23
Ganglios de la raíz dorsal, torácica	RDATRSS	12
	RPGREAS	144
	DKTRAGS	5
	TAPKSLK	25
	RDSTRQL	26
	ERDRTRG	21
Epéndimo	RGVLVTT	2
	RPGRESA	3
	PIPAGKK	9
	NESLKKK	4
	NSVKSVL	27
	PVKKKDA	7
	VQGSKMK	22
	NVTIKSK	28
	DKTRAGS	5
Epéndimo; 3er ventrículo	TAPKSLK	25
	RPGREAS	144
	RDATRSS	12
	DRLKGIV	31
	TAPKSLK	25
Epéndimo; 4to ventrículo	NSVRPLT	23
	RPGREAS	144
	KAAGRTV	46
	RDATRSS	12
	DKTRAGS	5
	RPGREAS	144
Epéndimo; Lt ventrículo	NSVRPLT	23
	RDATRSS	12
	TAPKSLK	25
	DETSRLV	30
	DKTRAGS	5
	RPGRESA	3
Globo pálido	ERDRTRG	21
	VSLKERV	29
	VQGSKMK	22
	DKTRAGS	5
	DETSRLV	30
	RDATRSS	12
	AGVARSK	14
	EVKGKGK	16
	DRLKGIV	31
	RGVLVTT	2
	DKTRAGS	5
Globo pálido externo	RGVLVTT	2

Región	Péptido	SEQ ID NO.
	SGTFVKA	33
	EVKGKGK	16
	SVASAKK	48
	VSLKERV	29
	TATPRKG	49
Globo pálido interno	RGVLVTT	2
	TAPKSLK	25
	TKTGLKL	50
	ERDRTRG	21
	SGVLVQR	32
	TATPRKG	49
	RPGREAS	144
Hipocampo	RGVLVTT	2
	RPGRESA	3
	SGVLVQR	32
	PVKKKDA	7
	NESLKKK	4
	SGTFVKA	33
	RPGREQA	1
	NSIARPV	34
Hipocampo; CA1	RGVLVTT	2
	SGVLVQR	32
	SVASAKK	48
	RDATRSS	12
	DKTRAGS	5
	TAPKSLK	25
Hipocampo; CA3	RPGREAS	144
	DETSRLV	30
	RDATRSS	12
	DKTRAGS	5
	ERDRTRG	21
	RDSTRQL	26
Hipocampo; DG	NSVRPLT	23
	RDATRSS	12
	TAPKSLK	25
	RDSTRQL	26
	RPGREAS	144
Oliva inferior	RPGREAS	144
	DKTRAGS	5
	TAPKSLK	25
	NDVRAKG	40
	RDATRSS	12
	NSVRPLT	23

Región	Péptido	SEQ ID NO.
Núcleos geniculados laterales	NSVRPLT	23
	TAPKSLK	25
	RPGREAS	144
	DKTRAGS	5
	RDATRSS	12
Meninges	RPGRESA	3
	ERDRTRG	21
	RDATRSS	12
	DKTRAGS	5
	PIPAGKK	9
	NRARAJE	35
	ARHALGG	36
	HSSRPVA	37
	PVKKKDA	7
	RPGREAS	144
	TAPKSLK	25
	NSVRPLT	23
Corteza motora	RPGREAS	144
	RGVLVTT	2
	NSVRPLT	23
	TAPKSLK	25
	DKTRAGS	5
	RDATRSS	12
Nervio óptico	AGVARSK	14
	RSRSEVL	15
	EVKGKGK	16
	DRLKGIV	31
	KTGTARL	38
	RPGRESA	3
	PVKKKDA	7
	RGVLVTT	2
	ERDRTRG	21
	RPGREAS	144
	NSVRPLT	23
	TAPKSLK	25
	DKTRAGS	5
	NDVRAKG	40
	RDATRSS	12
Corteza prefrontal	RPGREAS	144
	NSVRPLT	23
	TAPKSLK	25
	DKTRAGS	5
	RDATRSS	12

Región	Péptido	SEQ ID NO.
Puente de Varolio	RPGREAS	144
	NSVRPLT	23
	TAPKSLK	25
	RDATRSS	12
	DKTRAGS	5
Putamen	DETSRLV	30
	VKALGRP	39
	NDVRAKG	40
	NESLKKK	4
	ERDRTRG	21
	RPGREAS	144
	QGVLVVR	41
	KQYAGSQ	42
	RDATRSS	12
	VPTKSPK	13
	DKTRAGS	5
	KAAGRTV	46
	TAPKSLK	25
Formación reticular	RPGREAS	144
	RDSTRQL	26
	KGLRTPT	51
	RDATRSS	12
	NSVRPLT	23
	DKTRAGS	5
Médula espinal	PVKKKDA	7
	RGVLVTT	2
	PIPAGKK	9
	SGVLVQR	32
	NESLKKK	4
	RPGRESA	3
	TNRMALS	43
	ERDRTRG	21
	SGTFVKA	33
	NSVRPLT	23
Médula espinal, torácica	SGTFVKA	33
	RPGREAS	144
	RGVLVTT	2
	RDATRSS	12
	TAPKSLK	25
	DKTRAGS	5
	RPGREAS	144
Médula espinal, lumbar	TATPRKG	49
	TAPKSLK	25

Región	Péptido	SEQ ID NO.
	RDATRSS	12
	KAAGRTV	46
Médula espinal, cervical	SGVLVQR	32
	NSVRPLT	23
	RPGREAS	144
	RGVLVTT	2
	SGTFVKA	33
	TAPKSLK	25
	RDATRSS	12
	DKTRAGS	5
	RGVLVTT	2
Sustancia negra	GITLGRL	44
	RPGRESA	3
	PIPAGKK	9
	AGIMVRV	45
	SGVLVQR	32
	VSLKERV	29
	RDSTRQL	26
	RPGREAS	144
	RDATRSS	12
	KAAGRTV	46
Núcleos subtalámicos	RDATRSS	12
	RGVLVTT	2
	RPGRESA	3
	HVIRLPS	47
	EVKGKGK	16
	NESLKKK	4
	DETSRLV	30
	NSVRPLT	23
	SGTFVKA	33
	PIPAGKK	9
	VKALGRP	39
	NSVRPLT	23
	RPGREAS	144
Corteza temporal	TAPKSLK	25
	RDATRSS	12
	DKTRAGS	5
	GRETLLKG	8
	RPGREQA	1
Tálamo	ERDRTRG	21
	RGVLVTT	2
	SVASAKK	48
	TATPRKG	49

Región	Péptido	SEQ ID NO.
	RDATRSS	12
	NVTIKSK	28
	TKTGLKL	50
	KGLRTPT	51
Tálamo anterior	SGVLVQR	32
	NSVHNTA	18
	PIPAGKK	9
	DETSRLV	30
	AKLNKSS	17
	RPGREAS	144
Tálamo VA/VL	RGVLVTT	2
	DKTRAGS	5
	DETSRLV	30
	RDSTRQL	26
	NKIHANP	24
	TAPKSLK	25
Corteza visual	RPGREAS	144
	NSVRPLT	23
	TAPKSLK	25
	RDATRSS	12
	NSVHNTA	18
	DKTRAGS	5

Tabla 2. Péptidos dirigidos al AAV2 para cada estructura cerebral.

Región	Péptido	SEQ ID NO
Tronco encefálico	LTSRTSP	52
	DDPSARR	53
	GEQDLRR	54
	VSTALPR	55
	RDDVPLR	56
	TRVGTAG	57
	SSSKTGS	58
	SLSTGPK	59
	VQGRQGG	60
Caudado	RGASGAV	61
	NARAQGV	62
	TSNRGQV	63
	AVRGGMA	64
	RGLDKGT	65
	KGVDLKP	66
	TAVREER	67
	GNAGITK	68

Región	Péptido	SEQ ID NO
	SLSTGPK	59
	saragap	69
	GSRAGVG	105
	NARPVSA	76
	HDGGASR	103
	VTQSKGA	74
	KAQGVGG	79
	ESTGRER	73
Corteza cerebelosa	SGEFVGR	70
	SGRKLEV	71
	SARSGSV	72
	ESTGRER	73
	SSSKTGS	58
	RDDVPLR	56
	VQGRQGG	60
	VTQSKGA	74
	RGSGSAV	75
Corteza cerebelosa; vermis	HDGGASR	103
	EAQSHPR	91
	ESTGRER	73
	GRSTGMT	95
	GSRAGVG	105
Cerebelo; vermis superior derecho	HDGGASR	103
	RAVPAGG	84
	RSNAPQT	90
	TKSLSSG	92
	ESTGRER	73
	RGSTQVG	94
Cerebelo; lateral al tronco encefálico (derecha)	HDGGASR	103
	ESTGRER	73
	KAQGVGG	79
	EAQSHPR	91
	RAVPAGG	84
	RDDVPLR	56
Cerebelo; hemisferio izquierdo, adyacente al vermis superior	HDGGASR	103
	ESTGRER	73
	EAQSHPR	91
	NARAQGV	62
	LTSRTSP	52
	KAQGVGG	79
Cerebelo; hemisferio izquierdo, punto más lateral	HDGGASR	103
	ESTGRER	73
	EAQSHPR	91

Región	Péptido	SEQ ID NO
	KAQGVGG	79
	GSRAVG	105
Cerebelo; hemisferio izquierdo, adyacente al tronco encefálico	ESTGRER	73
	HDGGASR	103
	EAQSHPR	91
	GSRAVG	105
	KAQGVGG	79
	GRGGAAL	100
Corteza cerebral	DDPSARR	53
	NARPVSA	76
	TSNRQV	63
	NARAQGV	62
	TARGGGG	77
	KGVDLKP	66
	GRSASGS	78
	SSSKTGS	58
	KAQGVGG	79
	VQGRQGG	60
Cóclea (oído)	HDGGASR	103
	KAQGVGG	79
	RSNAPQT	90
	LTSRTSP	52
	EAQSHPR	91
Núcleos cerebelosos profundos	GSRAVG	105
	RAVPAGG	84
	EAQSHPR	91
	AVRGGMA	64
	VPGRTAG	81
	VTQSKGA	74
Ganglios de la raíz dorsal lumbar	HDGGASR	103
	TKSLSSG	92
	KAQGVGG	79
	GSRAVG	105
	VPGRTAG	81
Ganglios de la raíz dorsal cervical	GRSTGMT	95
	AVRGGMA	64
	HDGGASR	103
	TAAGGQR	99
	KAQGVGG	79
Ganglios de la raíz dorsal torácica	HDGGASR	103
	GSRAVG	105
	ARGSGVN	82
	RAVPAGG	84

Región	Péptido	SEQ ID NO
	KAQGVGG	79
Epéndimo	GRGPGPG	80
	DDPSARR	53
	TSNRQV	63
	RGSGSAV	75
	VQGRQGG	60
	VTQSKGA	74
	NARPVSA	76
	KGVDLKP	66
	NARAQGV	62
	TARGGGG	77
	GSRAVG	105
Epéndimo; 4to ventrículo	RDDVPLR	56
	RGSTQVG	94
	KAQGVGG	79
	GRGGAAL	100
	HDGGASR	103
Epéndimo; Lt ventrículo	ESTGRER	73
	EAQSHPR	91
	NARPVSA	76
	KAQGVGG	79
	GRSTGMT	95
Epéndimo; 3er ventrículo	LTSRTSP	52
	KAQGVGG	79
	GRGGAAL	100
	EAQSHPR	91
	ARGSGVN	82
	VQGRQGG	60
Globo pálido	VPGRTAG	81
	ARGSGVN	82
	SVRVGGQ	83
	RGSGSAV	75
	RAVPAGG	84
	VMSSGKP	85
	STPAPKS	86
	RGGAQVV	87
	EAQSHPR	91
Globo pálido externo	GSRAVG	105
	VQGRQGG	60
	ESTGRER	73
	GRGGAAL	100
	KASGAGG	88
	RDDVPLR	56
Globo pálido interno	RDDVPLR	56

Región	Péptido	SEQ ID NO
	GSRAGVG	105
	GAVGGVK	107
	DDPSARR	53
	RGSTQVG	94
	KAQGVGG	79
Hipocampo	DDPSARR	53
	KASGAGG	88
	KAQGVGG	79
	TSNRGQV	63
	VQGRQGG	60
	VSTALPR	55
	TGTAGLK	89
	NARPVSA	76
	SSSKTGS	58
Hipocampo; CA1	VAPISKS	101
	HDGGASR	103
	LTSRTSP	52
	ESTGRER	73
	KAQGVGG	79
Hipocampo; CA3	VQGRQGG	60
	GRSTGMT	95
	GRGGAAL	100
	KAQGVGG	79
	RGSTQVG	94
	TSNRGQV	63
Hipocampo; DG	HDGGASR	103
	ESTGRER	73
	RGSTQVG	94
	EAQSHPR	91
	KAQGVGG	79
	GSRAGVG	105
Oliva inferior	VPGRTAG	81
	HDGGASR	103
	VQGRQGG	60
	GSRAGVG	105
	RDDVPLR	56
	KAQGVGG	79
Núcleos geniculados laterales	HDGGASR	103
	ESTGRER	73
	GSRAGVG	105
	KAQGVGG	79
	EAQSHPR	91
	VPGRTAG	81

Región	Péptido	SEQ ID NO
Meninges	SSSKTGS	58
	NARPVSA	76
	RSNAPQT	90
	EAQSHPR	91
	TKSLSSG	92
	GRGPGPG	80
	AAGAKVM	93
	ESTGRER	73
	KGVDLKP	66
	VQGRQGG	60
	HDGGASR	103
	KAQGVGG	79
	RGSTQVG	94
	VPGRTAG	81
	GSRAGVG	105
Nervio óptico	DDPSARR	53
	RGGAQVV	87
	KASGAGG	88
	RGSGSAV	75
	VQGRQGG	60
	KAQGVGG	79
	TRVGTAG	57
	GEQDLRR	54
	RGSTQVG	94
	SSSKTGS	58
	GSRAGVG	105
	ESTGRER	73
	HDGGASR	103
	RDDVPLR	56
Puente de Varolio	ESTGRER	73
	GSRAGVG	105
	HDGGASR	103
	KAQGVGG	79
	RDDVPLR	56
	GRGGAAL	100
Corteza prefrontal	HDGGASR	103
	GSRAGVG	105
	ESTGRER	73
	KAQGVGG	79
	RGSTQVG	94
Corteza motora primaria	GSRAGVG	105
	VPGRTAG	81
	ESTGRER	73

Región	Péptido	SEQ ID NO
	RGSTQVG	94
	GRGGAAL	100
	KAQGVGG	79
Corteza visual primaria	HDGGASR	103
	ESTGRER	73
	EAQSHPR	91
	TKSLSSG	92
	GSRAVG	105
	KAQGVGG	79
Putamen	GRSTGMT	95
	RATSQST	96
	VGRSVGA	97
	VQGRQGG	60
	GEGGGGR	98
	VSTALPR	55
	RGASGAV	61
	TAAGGQR	99
	SLSTGPK	59
	GRGGAAL	100
	RAVPAGG	84
	KAQGVGG	79
	GSRAVG	105
	VPGRTAG	81
	HDGGASR	103
	RDDVPLR	56
Formación reticular	VTQSKGA	74
	DDPSARR	53
	TGTAGLK	89
	ARGSGVN	82
	RGSTQVG	94
	RDDVPLR	56
Médula espinal	SLSTGPK	59
	GRSTGMT	95
	RGASGAV	61
	TARGGGG	77
	KASGAGG	88
	VAPISKS	101
	DDPSARR	53
	TSNRGQV	63
	SSSKTGS	58
	VQGRQGG	60
Médula espinal torácica	HDGGASR	103
	ESTGRER	73

Región	Péptido	SEQ ID NO
	EAQSHPR	91
	GSRAGVG	105
	VQGRQGG	60
	KAQGVGG	79
Médula espinal lumbar	HDGGASR	103
	EAQSHPR	91
	DDPSARR	53
	KLSISGN	106
	VMSSGKP	85
	KAQGVGG	79
	TSNRGQV	63
Médula espinal cervical	ESTGRER	73
	HDGGASR	103
	EAQSHPR	91
	GSRAGVG	105
	RGSTQVG	94
	KAQGVGG	79
Sustancia negra	LTSRTSP	52
	NARPVSA	76
	SSSKTGS	58
	RGGAQVV	87
	APPVKLS	102
	DDPSARR	53
	RGSGSAV	75
	HDGGASR	103
	TRVGTAG	57
	GSRAGVG	105
	KLSISGN	106
	KAQGVGG	79
	TGTAGLK	89
	ARGSGVN	82
Núcleos subtalámicos	VQGRQGG	60
	RSGGAAV	104
	KASGAGG	88
	SSSKTGS	58
	KAQGVGG	79
	GSRAGVG	105
	RGGAQVV	87
	APPVKLS	102
	RGSGSAV	75
	TRVGTAG	57
	RGSTQVG	94
	KLSISGN	106

Región	Péptido	SEQ ID NO
	RDDVPLR	56
	GRGGAAL	100
	TSNRGQV	63
Corteza temporal	ESTGRER	73
	HDGGASR	103
	GSRAGVG	105
	KAQGVGG	79
	RDDVPLR	56
Tálamo	KASGAGG	88
	RDDVPLR	56
	LTSRTSP	52
	KLSISGN	106
	VSTALPR	55
	VMSSGKP	85
	GAVGGVK	107
	KNESGKV	108
	VTQSKGA	74
Tálamo, anterior	AGQLAGR	109
	RGSTQVG	94
	KAQGVGG	79
	TKSLSSG	92
	RAVPAGG	84
	TSNRGQV	63
Tálamo VA/VL	RDDVPLR	56
	TRVGTAG	57
	ARGSGVN	82
	GRGGAAL	100
	NARPVSA	76
	VQGRQGG	60
	GSRAGVG	105

Tabla 3. Péptidos dirigidos al AAV9 para cada estructura cerebral.

Región	Péptido	SEQ ID NO
Tronco encefálico	KGGGFHG	110
	RAKPGME	111
	GRDVTRS	112
	RGDLQWV	113
	GGDRTRG	114
	RGDLASV	115
	RDTTRNL	116
	KGGGVHG	117
Caudado	KGGGFHG	110

Región	Péptido	SEQ ID NO
	RGDLQWV	113
	RGDMYRV	118
	RGDRPVS	119
	RSDVGSL	120
	RGDLASV	115
	RDTTRNL	116
	AGVKPGR	121
	KGGGVHG	117
	RAKPGME	111
	GADRTRG	127
	GRDVTRS	112
	ARGDGWR	132
	RGGGVYG	126
	GGDRTRG	114
	GRDYTRL	133
Corteza cerebeloso	KGGGFHG	110
	RGDLQWV	113
	RAKPGME	111
	RGDWPRG	122
	RGDRPVS	119
	GGRPGSW	123
	RGDYPRS	124
	RGDLRFI	125
Corteza cerebral	KGGGFHG	110
	RAKPGME	111
	RDTTRNL	116
	RGGGVYG	126
	GGDRTRG	114
	GADRTRG	127
	RGDLQWV	113
	RGDLRFI	125
Corteza prefrontal	KGGGFHG	110
	MMGRPGR	136
	AGVKPGR	121
	RGGGVYG	126
	RGDLRFI	125
	RGDWPRG	122
	RGDLQWV	113
	RGDRPVS	119
Corteza motora	MMGRPGR	136
	KGGGFHG	110
	AGVKPGR	121
	RGGGVYG	126

Región	Péptido	SEQ ID NO
	RGDLRFI	125
	RGDWPRG	122
	RGDLQWV	113
Corteza temporal	KGGGFHG	110
	RGGGVYG	126
	MMGRPGR	136
	RGDLRFI	125
	RGDLQWV	113
	RGDWPRG	122
Corteza visual	AGVKPGR	121
	MMGRPGR	136
	RGDLRFI	125
	RGDLQWV	113
	RGDFMGL	128
	RGGGVYG	126
	KGGGFHG	110
	AWDGTRV	131
Cóclea (oído)	MMGRPGR	136
	KGGGFHG	110
	RGDLASV	115
	AGVKPGR	121
	GGRPGSW	123
	RGGGVYG	126
	GGDRTRG	114
Núcleos cerebelosos profundos	RSDVGSL	120
	KGGGFHG	110
	RGDLQWV	113
	GRDVTRS	112
	RGDLRFI	125
	RGGGVYG	126
	RAKPGME	111
Ganglios de la raíz dorsal lumbar	MMGRPGR	136
	AGVKPGR	121
	GGRPGSW	123
	GRDYTRL	133
	RGGGVYG	126
	GGDRTRG	114
Ganglios de la raíz dorsal cervical	RGDLQWV	113
	KGGGFHG	110
	RRDETRT	129
	RGGGVYG	126
	GADRTRG	127
	RGDRPVS	119

Región	Péptido	SEQ ID NO
	GGDRTRG	114
Ganglios de la raíz dorsal torácica	MMGRPGR	136
	AGVKPGR	121
	GGRPGSW	123
	GRDYTRL	133
	RGDFMGL	128
	RGGGVYG	126
	GGDRTRG	114
	RAKPGME	111
	AWDGTRV	131
	KGGGFHG	110
Epéndimo	RGDLQWV	113
	RAKPGME	111
	RGDMYRV	118
	RSDVGSL	120
	RGDFMGL	128
	RGDRPVS	119
	RGDLQWV	113
Epéndimo; 4to ventrículo	RGDLASV	115
	RSDVGSL	120
	RGDLRFI	125
	MDLTKAV	135
	RGGGVYG	126
	RAKPGME	111
	RGDRPVS	119
	GGDRTRG	114
	MMGRPGR	136
	RSDVGSL	120
Epéndimo; Lt ventrículo	AGVKPGR	121
	GGRPGSW	123
	RAKPGME	111
	RGGGVYG	126
	GRDVTRS	112
	RGDLQWV	113
	KGGGVHG	117
	GADRTRG	127
	GGDRTRG	114
Epéndimo; 3er ventrículo	MMGRPGR	136
	RGDLQWV	113
	RGGGVYG	126
	GRDVTRS	112
	KGGGFHG	110
	RRDETRT	129
Globo pálido		

Región	Péptido	SEQ ID NO
	RAKPGME	111
	GGDRTRG	114
	GRDVTRS	112
	RSDVGSL	120
	RGDRPVS	119
Globo pálido externo	RAKPGME	111
	RRGDAWS	134
	RGDMYRV	118
	RGDLQWV	113
	GRDVTRS	112
	RGGGVYG	126
Globo pálido interno	MDLTKAV	135
	RDTTRNL	116
	AWDGTRV	131
	RGGGVYG	126
	KGGGVHG	117
	RGDLQWV	113
Hipocampo	KGGGFHG	110
	AESPWER	130
	RGDLQWV	113
	RAKPGME	111
	RRDETRT	129
	RGDLRFI	125
	GGRPGSW	123
	RDTTRNL	116
Hipocampo; CA1	KGGGFHG	110
	ARGDGWR	132
	RGGGVYG	126
	AWDGTRV	131
	GADRTRG	127
	RAKPGME	111
Hipocampo; CA3	RAKPGME	111
	RRDETRT	129
	MDLTKAV	135
	AWDGTRV	131
	RDTTRNL	116
	GGDRTRG	114
	GRDVTRS	112
	RGGGVYG	126
Hipocampo; DG	KGGGFHG	110
	RGDFMGL	128
	RGGGVYG	126
	ARGDGWR	132

Región	Péptido	SEQ ID NO
	RGDLRFI	125
	RGDLQWV	113
	GRDYTRL	133
Oliva inferior	RRGDAWS	134
	GGRPGSW	123
	KGGGFHG	110
	KGGGVHG	117
	RGDWPRG	122
	RGGGVYG	126
	GRDYTRL	133
	RGDRPVS	119
Meninges	KGGGFHG	110
	RGDLQWV	113
	RAKPGME	111
	RGDWPRG	122
	GGDRTRG	114
	AWDGTRV	131
	RGDMYRV	118
	RGDRPVS	119
	RSDVGSL	120
	KGGGVHG	117
	GRDVTRS	112
	AGVKPGR	121
	RGGGVYG	126
Nervio óptico	KGGGFHG	110
	RAKPGME	111
	RRDETRT	129
	RDTTRNL	116
	GGDRTRG	114
	ARGDGWR	132
	KGGGVHG	117
	RGDLRFI	125
	RGGGVYG	126
	AGVKPGR	121
	RGDLQWV	113
	GRDYTRL	133
Puente de Varolio	KGGGFHG	110
	RGGGVYG	126
	RRDETRT	129
	AGVKPGR	121
	KGGGVHG	117
	RAKPGME	111
	RGDLQWV	113
	GRDYTRL	133

Región	Péptido	SEQ ID NO
Putamen	KGGGFHG	110
	GRDVTRS	112
	GADRTRG	127
	GRDYTRL	133
	RDTTRNL	116
	RGDLQWV	113
	RRGDAWS	134
	GGRPGSW	123
	KGGGVHG	117
	RGDMYRV	118
	RAKPGME	111
	RGGGVYG	126
	GGDRTRG	114
	RGDRPVS	119
Formación reticular	MMGRPGR	136
	GRDVTRS	112
	GADRTRG	127
	RAKPGME	111
	RSDVGSL	120
	RGGGVYG	126
	AWDGTRV	131
Médula espinal	KGGGFHG	110
	RGDLQWV	113
	RGDRPVS	119
	RGDWPRG	122
	RSDVGSL	120
	GGRPGSW	123
	RRGDAWS	134
	RGDFMGL	128
Médula espinal; torácica	KGGGFHG	110
	RGGGVYG	126
	RGDLQWV	113
	RGDLRFI	125
	RRGDAWS	134
	RGDRPVS	119
	GRDYTRL	133
Médula espinal; lumbar	RGDLASV	115
	KGGGFHG	110
	RGGGVYG	126
	MMGRPGR	136
	RGDLQWV	113
	AWDGTRV	131
	RAKPGME	111

Región	Péptido	SEQ ID NO
Médula espinal; cervical	KGGGFHG	110
	RGDLRFI	125
	RGGGVYG	126
	RRGDAWS	134
	RGDLQWV	113
	RGDWPRG	122
Sustancia negra	KGGGFHG	110
	RAKPGME	111
	GRDVTRS	112
	GGDRTRG	114
	RRDETRT	129
	RGDLQWV	113
	KGGGVHG	117
	MDLTKAV	135
	RRGDAWS	134
	GGRPGSW	123
	GADRTRG	127
	RGGGVYG	126
Núcleos subtalámicos	KGGGFHG	110
	RAKPGME	111
	MDLTKAV	135
	RGDRPVS	119
	RGDWPRG	122
	ARGDGWR	132
	RGDLQWV	113
	RSDVGSL	120
	RGDLASV	115
	RGGGVYG	126
Tálamo	KGGGFHG	110
	GRDVTRS	112
	GGDRTRG	114
	RGDLQWV	113
	GRDYTRL	133
	RGDLRFI	125
	MMGRPGR	136
	TGRPGVW	137
Tálamo anterior	MDLTKAV	135
	GGRPGSW	123
	RGDLRFI	125
	RDTTRNL	116
	RAKPGME	111
	RGGGVYG	126
	GRDYTRL	133

Región	Péptido	SEQ ID NO
	RGDWPRG	122
Tálamo VA/VL	RGDWPRG	122
	GGDRTRG	114
	RRGDAWS	134
	RGDLRFI	125
	RGDYPRS	124
	RGDRPVS	119
	RGGGVYG	126

Tabla 4. Péptidos dirigidos al AAV1 para varios órganos.

Órgano	Péptido	SEQ ID NO
Cerebro	RPGREAS	144
	TAPKSLK	25
	RDATRSS	12
	DKTRAGS	5
	NSVRPLT	23
	RDSTRQL	26
	RGVLVTT	2

Tabla 5. Péptidos dirigidos al AAV2 para varios órganos.

Órgano	Péptido	SEQ ID NO
Cerebro	HDGGASR	103
	KAQGVGG	79
	GSRAGVG	105
	ESTGRER	73
	RDDVPLR	56
	RGSTQVG	94
	EAQSHPR	91
Riñón	GSRAGVG	105
	TSNRGQV	63
	GRSTGMT	95
	GRGPGPG	80
	GAVGGVK	107
	KAQGVGG	79
	RDDVPLR	56
Corazón	HDGGASR	103
	SLSTGPK	59
	GSRAGVG	105
	RGSTQVG	94
	NARAQGV	62
	ESTGRER	73

Órgano	Péptido	SEQ ID NO
Hígado	KAQGVGG	79
	GSRAVG	105
	RDDVPLR	56
	HDGGASR	103
	NARPVSA	76
Gónada	VPGRTAG	81
	KAQGVGG	79
	RGSTQVG	94
	HDGGASR	103
	GAVGGVK	107
	RDDVPLR	56
Bazo	KAQGVGG	79
	GSRAVG	105
	GRGGAAL	100
	VPGRTAG	81
	ESTGRER	73
Pulmón	ESTGRER	73
	TKSLSSG	92
	RGSTQVG	94
	KAQGVGG	79
	NARPVSA	76

Tabla 6. Péptidos dirigidos al AAV9 para varios órganos.

Órgano	Péptido	SEQ ID NO
Cerebro	RGGGVYG	126
	RAKPGME	111
	KGGGFHG	110
	GGDRTRG	114
	MMGRPGR	136
	RGDLQWV	113
	MDLTKAV	135
	GGRPGSW	123
Riñón	MMGRPGR	136
	RGDLQWV	113
	RAKPGME	111
	RSDVGSL	120
	KGGGFHG	110
	RGGGVYG	126
Corazón	GGRPGSW	123
	RSDVGSL	120
	RGDLASV	115
	KGGGFHG	110

Órgano	Péptido	SEQ ID NO
	GRDYTRL	133
	RGGGVYG	126
	RAKPGME	111
Hígado	GGDRTRG	114
	GRDVTRS	112
	RDTTRNL	116
	RRDETRT	129
	RSDVGSL	120
	RGGGVYG	126
	RAKPGME	111
	GRDYTRL	133
Gónada	KGGGVHG	117
	KGGGFHG	110
	RSDVGSL	120
	GGRPGSW	123
	RRDETRT	129
	RGGGVYG	126
	GGDRTRG	114
	RGDRPVS	119
	RAKPGME	111

I. Vectores de virus adenoasociados (AAV)

[0042] 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 la replicación.

[0043] Los genomas de AAV pueden existir en un estado extracromosómico sin integrarse en los genomas celulares del huésped; poseer una amplia gama de huéspedes; transducir tanto células en división como células que no se dividen *in vitro* e *in vivo* y mantener altos niveles de expresión de los genes transducidos. Las partículas virales de AAV son termoestables; resistentes a solventes, detergentes, cambios de pH y temperatura; y pueden purificarse en columna y/o concentrarse en gradientes de CsCl o por otros medios. El genoma del AAV comprende un ácido desoxirribonucleico monocatenario (ssDNA, por sus siglas en inglés) ya sea de sentido positivo o negativo. El genoma de aproximadamente 4.7 kb del AAV consta de un segmento de ADN monocatenario de polaridad positiva o negativa. Los extremos del genoma son repeticiones terminales invertidas cortas (ITRs, por sus siglas en inglés) que pueden plegarse en estructuras de horquilla y servir como origen de la replicación del ADN viral.

[0044] Un "genoma" AAV se refiere a una secuencia de ácido nucleico recombinante que finalmente se empaqueta o encapsula para formar una partícula AAV. Una partícula AAV a menudo

comprende un genoma del AAV empaquetado con proteínas de la cápside del AAV. En los casos en que se utilicen plásmidos recombinantes para construir o fabricar vectores recombinantes, el genoma del vector AAV no incluye la parte del "plásmido" que no corresponde a la secuencia del genoma del vector del plásmido recombinante. Esta parte del genoma del plásmido recombinante que no es un vector se denomina "esqueleto del plásmido", que es importante para la clonación y la amplificación del plásmido, un proceso que se necesita para la propagación y producción del plásmido, pero que no está empaquetado ni encapsulado en forma viral. Por lo tanto, un "genoma" de vector AAV se refiere a un ácido nucleico que está empaquetado o encapsulado por las proteínas de la cápside del AAV.

[0045] El virión (partícula) AAV es una partícula icosaédrica sin envoltura de aproximadamente 25 nm de diámetro que comprende una cápside de AAV. La partícula AAV comprende una simetría icosaédrica compuesta por tres proteínas de la cápside relacionadas, VP1, VP2 y VP3, que interactúan entre sí para formar la cápside. El genoma de la mayoría de los AAV nativos a menudo contiene dos marcos de lectura abiertos (ORF, por sus siglas en inglés), a veces denominados ORF izquierdo y ORF derecho. El ORF derecho a menudo codifica las proteínas de la cápside VP1, VP2 y VP3. Estas proteínas a menudo se encuentran en una proporción de 1:1:10 respectivamente, pero pueden estar en proporciones variadas y todas se derivan del ORF de la derecha. Las proteínas de la cápside VP1, VP2 y VP3 se diferencian entre sí por el uso de empalme (splicing) alternativo y un codón de inicio inusual. El análisis de eliminación ha demostrado que la eliminación o alteración de VP1 que se traduce de un mensaje empalmado Alternativamente da como resultado un rendimiento reducido de partículas infecciosas. Las mutaciones dentro de la región codificante de VP3 dan como resultado que no se produzca ninguna progenie de ADN monocatenario o partículas infecciosas. En ciertas realizaciones, el genoma de una partícula de AAV codifica uno, dos o los tres polipéptidos VP1, VP2 y VP3.

[0046] El ORF izquierdo a menudo codifica las proteínas Rep no estructurales, Rep 40, Rep 52, Rep 68 y Rep 78, que están involucradas en la regulación de la replicación y la transcripción además de la producción de genomas de progenie monocatenarios. Dos de las proteínas Rep se han asociado con la integración preferencial de los genomas de AAV en una región del brazo q del cromosoma 19 humano. Se ha demostrado que Rep68/78 posee actividad de unión a NTP, así como actividades de helicasa de ADN y ARN. Algunas proteínas Rep poseen una señal de localización nuclear así como varios sitios potenciales de fosforilación. En ciertas realizaciones, el genoma de un AAV (*p.ej*, un rAAV) codifica algunas o todas las proteínas Rep. En ciertas realizaciones, el genoma de un AAV (*p.ej*, un rAAV) no codifica las proteínas Rep. En ciertas realizaciones, una o más de las proteínas Rep pueden administrarse en trans y, por lo tanto, no están incluidas en una partícula de AAV que comprende un ácido nucleico que codifica un polipéptido.

[0047] 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 origen de la replicación del ADN viral. En consecuencia, el genoma de un AAV comprende uno o más (*p.ej*, un par de) secuencias ITR que flanquean un genoma de ADN viral monocatenario. Las secuencias ITR a menudo

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 las ITR, un motivo de repetición GAGC y el sitio de resolución terminal (trs). Se ha demostrado que el motivo repetido se une a Rep cuando las ITR tiene una conformación lineal o de horquilla. Se cree que esta unión posiciona a Rep68/78 para la escisión en el trs que se produce de manera específica del sitio y de la hebra. Además de su papel en la replicación, estos dos elementos parecen ser fundamentales para la integración viral. Dentro del locus de integración del cromosoma 19 hay un sitio de unión a Rep con un trs adyacente. Se ha demostrado que estos elementos son funcionales y necesarios para la integración específica del locus.

[0048] El término "recombinante", como modificador de vector, como virus recombinante, *p.ej*, lenti- o parvo-virus (*p.ej*, AAV), así como un modificador de secuencias tales como polipéptidos y secuencias de ácidos nucleicos recombinantes, significa que las composiciones han sido manipuladas (*es decir*, diseñadas) de una manera que generalmente no ocurre en la naturaleza. Un ejemplo particular de un vector recombinante, tal como un vector AAV, retroviral o lentiviral sería donde una secuencia de ácido nucleico que normalmente no está presente en el genoma viral de tipo silvestre se inserta dentro del genoma viral. Un ejemplo de una secuencia de ácido nucleico recombinante sería donde un ácido nucleico (*p.ej*, gen) codifica un ARN inhibidor clonado en un vector, con o sin 5', 3' y/o regiones de intrones con las que el gen está normalmente asociado dentro del genoma viral. Aunque el término "recombinante" no siempre se usa aquí en referencia a vectores, como vectores virales, así como secuencias como polinucleótidos, las formas "recombinantes" que incluyen secuencias de ácidos nucleicos, polinucleótidos, transgenes, etc. se incluyen expresamente a pesar de cualquier omisión de este tipo.

[0049] Un "vector" viral recombinante se deriva del genoma de tipo silvestre de un virus mediante el uso de métodos moleculares para eliminar parte del genoma de tipo silvestre del virus y reemplazarlo con un ácido nucleico no nativo, como una secuencia de ácido nucleico. Típicamente, 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" (*p.ej*, rAAV) se distingue de un viral (*p.ej*, AAV) ya que parte del genoma viral ha sido reemplazado con una secuencia no nativa con respecto al ácido nucleico genómico viral, tal como un ácido nucleico que codifica un transactivador o un ácido nucleico que codifica un ARN inhibidor o un ácido nucleico que codifica una proteína terapéutica. Por lo tanto, la incorporación de tales secuencias de ácido nucleico no nativas define al vector viral como un vector "recombinante", que en el caso de AAV puede denominarse "vector rAAV".

[0050] En ciertas realizaciones, un AAV (*p.ej*, un rAAV) comprende dos ITR. En ciertas realizaciones, un AAV (*p.ej*, un rAAV) comprende un par de ITR. En ciertas realizaciones, un AAV (*p.ej*, 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 un polipéptido que tiene función o actividad.

[0051] Un vector AAV (*p.ej*, vector rAAV) puede empaquetarse y se denomina en el presente documento "partícula AAV" para la posterior infección (transducción) de una célula, *ex-vivo*, *in vitro* o *in*

vivo. Cuando un vector AAV recombinante se encapsula o empaqueta en una partícula AAV, la partícula también puede denominarse "partícula rAAV". En ciertas realizaciones, una partícula de AAV es una partícula de rAAV. Una partícula de rAAV a menudo comprende un vector de rAAV, o una parte del mismo. Una partícula rAAV puede ser una o más partículas rAAV (*p.ej*, una pluralidad de partículas AAV). Las partículas de rAAV generalmente comprenden proteínas que encapsulan o empaquetan el genoma del vector rAAV (*p.ej*, proteínas de la cápside). Se observa que la referencia a un vector rAAV también se puede utilizar para hacer referencia a una partícula rAAV.

[0052] Cualquier partícula AAV adecuada (*p.ej*, partícula rAAV) puede usarse para un método o uso en este documento. Una partícula de rAAV y/o el genoma comprendido en la misma, puede derivar de cualquier serotipo o cepa adecuada de AAV. Una partícula de rAAV y/o el genoma comprendido en la misma, 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, donde la partícula de AAV es adecuada para la infección y/o transducción de una célula de mamífero. Los ejemplos no limitativos de serotipos de AAV incluyen AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV-rh74, AAV-rh10 y AAV-2i8.

[0053] 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 comprenden una mezcla de dos o más partículas de rAAV diferentes (*p.ej*, de diferentes serotipos y/o cepas).

[0054] Como se usa aquí, 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 sobre la base de la falta de reactividad cruzada entre los anticuerpos contra un AAV en comparación con otro AAV. Tales diferencias de reactividad cruzada generalmente se deben a diferencias en las secuencias de proteínas de la cápside/determinantes antigénicos (*p.ej*, debido a diferencias en la secuencia de VP1, VP2 y/o VP3 de los serotipos de AAV). A pesar de la posibilidad de que las variantes de AAV, incluidas las variantes de la cápside, no sean serológicamente distintas de un AAV de referencia u otro serotipo de AAV, difieren en al menos un residuo de nucleótido o aminoácido en comparación con el serotipo de referencia u otro AAV.

[0055] En ciertas realizaciones, un vector rAAV basado en un primer genoma de serotipo corresponde al serotipo de una o más de las proteínas de la cápside que empaquetan el vector. Por ejemplo, el serotipo de uno o más ácidos nucleicos de AAV (*p.ej*, ITRs) que comprende el genoma del vector AAV corresponde al serotipo de una cápside que comprende la partícula rAAV.

[0056] En ciertas realizaciones, un genoma de vector de rAAV puede basarse en un AAV (*p.ej*, AAV2) genoma de serotipo distinto del serotipo de una o más de las proteínas de la cápside de AAV que empaquetan el vector. Por ejemplo, un genoma de vector rAAV puede comprender ácidos nucleicos

derivados de AAV2 (*p.ej.*, ITR), mientras que al menos una o más de las tres proteínas de la cápside se derivan de un serotipo diferente, *p.ej.*, un serotipo AAV1, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, Rh10, Rh74 o AAV-2i8 o una variante del mismo.

[0057] En ciertas realizaciones, una partícula de rAAV o un vector de su genoma relacionado con un serotipo de referencia tiene un polinucleótido, polipéptido o una subsecuencia del mismo que comprende o consiste en una secuencia de al menos el 60% o más (*p.ej.*, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, etc.) idéntico a un polinucleótido, polipéptido o subsecuencia de una partícula 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 cápside o secuencia ITR que comprende o consiste en una secuencia de al menos 60% o más (*p.ej.*, 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 cápside o ITR de un serotipo AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, Rh10, Rh74 o AAV-2i8.

[0058] En ciertas realizaciones, un método en este documento comprende el uso, la administración o la entrega de una partícula rAAV1, rAAV2, rAAV3, rAAV4, rAAV5, rAAV6, rAAV7, rAAV8, rAAV9, rAAV10, rAAV11, rAAV12, rRh10, rRh74 o rAAV-2i8.

[0059] En determinadas realizaciones, un método del presente documento comprende el uso, la administración o el 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 la cápside (*p.ej.*, VP1, VP2 y/o VP3) que son al menos 60%, 65%, 70%, 75% o más idénticos, *p.ej.*, 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% idénticos a una proteína de la cápside correspondiente de una partícula AAV2 nativa o de tipo silvestre. En determinadas realizaciones, una partícula de rAAV2 comprende proteínas de la cápside VP1, VP2 y VP3 que son al menos un 75% o más idénticas, *p.ej.*, 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% idénticos a una proteína de la cápside correspondiente a una partícula AAV2 nativa o de tipo silvestre. En ciertas realizaciones, una partícula de rAAV2 es una variante de una partícula de AAV2 nativa o de tipo silvestre. En algunos aspectos, una o más proteínas de la 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 las proteínas de la cápside de una variante nativa o partícula AAV2 de tipo silvestre.

[0060] 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 la cápside (*p.ej.*, VP1, VP2 y/o VP3) que son al menos 60%, 65%, 70%, 75% o más idénticos, *p.ej.*, 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% idénticos a una proteína de la cápside correspondiente a una partícula AAV9 nativa o de tipo silvestre.

En determinadas realizaciones, una partícula de rAAV9 comprende proteínas de la cápside VP1, VP2 y VP3 que son al menos un 75% o más idénticas, *p.ej.* 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% idéntica a una proteína de cápside correspondiente a una partícula AAV9 nativa o de tipo silvestre. En ciertas realizaciones, una partícula de rAAV9 es una variante de una partícula de AAV9 nativa o de tipo silvestre. En algunos aspectos, una o más proteínas de la 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 las proteínas de la cápside de una variante nativa o partícula AAV9 de tipo silvestre.

[0061] En ciertas realizaciones, una partícula de rAAV comprende uno o dos ITR (*p.ej.* un par de ITR) que son al menos 75% o más idénticos, *p.ej.* 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 un 100% idénticos a las ITR correspondientes de un AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 nativo o de tipo silvestre, AAV10, AAV11, AAV12, AAV-rh74, AAV-rh10 o AAV-2i8, siempre que conserven una o más funciones ITR deseadas (*p.ej.* capacidad de formar una horquilla, que permite la replicación del ADN; integración del ADN de AAV en el genoma de una célula huésped; y/o empaquetamiento, si se desea).

[0062] En ciertas realizaciones, una partícula de rAAV2 comprende uno o dos ITR (*p.ej.* un par de ITR) que son al menos 75% o más idénticos, *p.ej.* 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 un 100% idénticos a los ITR correspondientes de una partícula AAV2 nativa o de tipo silvestre, siempre que conserven una o más funciones ITR deseadas (*p.ej.* capacidad de formar una horquilla, que permite la replicación del ADN; integración del ADN de AAV en el genoma de una célula huésped; y/o empaquetamiento, si se desea).

[0063] En ciertas realizaciones, una partícula de rAAV9 comprende uno o dos ITR (*p.ej.* un par de ITR) que son al menos 75% o más idénticos, *p.ej.* 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 un 100% idénticos a los ITR correspondientes de una partícula AAV2 nativa o de tipo silvestre, siempre que conserven una o más funciones ITR deseadas (*p.ej.* capacidad de formar una horquilla, que permite la replicación del ADN; integración del ADN de AAV en el genoma de una célula huésped; y/o empaquetamiento, si se desea).

[0064] Una partícula de rAAV puede comprender una ITR que tenga cualquier número adecuado de repeticiones "GAGC". En determinadas realizaciones, una ITR de una partícula AAV2 comprende 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 o más repeticiones "GAGC". En ciertas realizaciones, una partícula de rAAV2 comprende una ITR que comprende tres repeticiones "GAGC". En determinadas realizaciones, una partícula de rAAV2 comprende una ITR que tiene menos de cuatro repeticiones "GAGC". En determinadas realizaciones, una partícula de rAAV2 comprende una ITR que tiene más de cuatro repeticiones "GAGC". En determinadas realizaciones, una ITR de una partícula de rAAV2 comprende un sitio de unión a Rep en donde el cuarto nucleótido de las dos primeras repeticiones "GAGC" es una C en lugar de una T.

[0065] Se puede incorporar una longitud de ADN adecuada ejemplar en vectores de rAAV para el empaquetamiento/encapsulación en una partícula de rAAV de aproximadamente 5 kilobases (kb) o menos. En particular, en las realizaciones, la longitud del ADN es inferior a aproximadamente 5 kb, inferior a aproximadamente 4.5 kb, inferior a aproximadamente 4 kb, inferior a aproximadamente 3.5 kb, inferior a aproximadamente 3 kb o inferior a aproximadamente 2.5 kb.

[0066] Los vectores de rAAV que incluyen una secuencia de ácido nucleico que dirige la expresión de un ARNi o polipéptido se pueden generar utilizando técnicas recombinantes adecuadas conocidas en la técnica (*p.ej*, ver Sambrook *et al.*, 1989). Los vectores AAV recombinantes se empaquetan típicamente en partículas AAV competentes en transducción y se propagan usando un sistema de empaquetamiento viral AAV. Una partícula de AAV competente en transducción es capaz de unirse y entrar en una célula de mamífero y, posteriormente, entregar una carga de ácido nucleico (*p.ej*, un gen heterólogo) al núcleo de la célula. Por lo tanto, una partícula de rAAV intacta que es competente en transducción se configura para transducir una célula de mamífero. Una partícula de rAAV configurada para transducir una célula de mamífero a menudo no es competente para la replicación y requiere maquinaria proteica adicional para autorreplicarse. Por lo tanto, una partícula de rAAV que está configurada para transducir una célula de mamífero está diseñada para unirse e ingresar a una célula de mamífero y administrar un ácido nucleico a la célula, donde el ácido nucleico para la administración a menudo se coloca entre un par de ITR de AAV en el genoma de rAAV.

[0067] Las células huésped adecuadas para producir partículas de AAV competentes para la 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 heterólogos de rAAV. Células de la línea celular humana estable, HEK293 (fácilmente disponible a través de, *p.ej*, la Colección Americana de Cultivos Tipo con el Número de Acceso ATCC CRL1573). En ciertas realizaciones, una línea celular de riñón embrionario humano modificada (*p.ej*, HEK293), que se transforma con fragmentos de ADN de adenovirus tipo 5 y expresa los genes adenovirales E1a y E1b, se utiliza para generar partículas de AAV recombinantes. La línea celular HEK293 modificada se transfecta fácilmente y proporciona una plataforma particularmente conveniente para producir partículas de rAAV. Los métodos para generar partículas de AAV de alto título capaces de transducir células de mamífero son conocidos en la técnica. Por ejemplo, la partícula AAV se puede fabricar como se establece en Wright, 2008 y Wright, 2009.

[0068] En determinadas realizaciones, las funciones auxiliares de AAV se introducen en la célula huésped mediante la transfección de la célula huésped con una construcción auxiliar de AAV antes de la transfección de un vector de expresión de AAV o simultáneamente con ella. Por lo tanto, las construcciones auxiliares de AAV se usan a veces para proporcionar al menos una 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 a menudo carecen de ITR de AAV y no pueden replicarse ni empaquetarse. Estas construcciones pueden estar en forma de plásmido, fago, transposón, cósmido, virus

o virión. Se han descrito una serie de construcciones auxiliares de AAV, como los plásmidos pAAV/Ad y pIM29 45 de uso común que codifican los productos de expresión de Rep y Cap. Se conocen varios otros vectores que codifican productos de expresión Rep y/o Cap.

5 **[0069]** Un "vector de expresión" es un vector especializado que contiene un gen o una secuencia de ácido nucleico con las regiones reguladoras necesarias para la expresión en una célula huésped. Un vector de expresión puede contener al menos un origen de replicación para la propagación en una célula y, opcionalmente, elementos adicionales, como una secuencia de ácido nucleico heteróloga, un elemento de control de la expresión (*p.ej.*, un promotor, potenciador), intrón, ITR(s) y señal de poliadenilación.

II. Agentes Terapéuticos

10 **[0070]** En algunas realizaciones, se pueden usar métodos de transferencia de genes virales para introducir ácidos nucleicos en células de mamífero o tejidos diana. Dichos métodos se pueden usar para administrar ácidos nucleicos que codifican ARN inhibidores, ARN no codificantes y/o proteínas terapéuticas a células en cultivo o en un organismo huésped.

A. ARN inhibidores

15 **[0071]** La "interferencia de ARN (ARNi)" es el proceso de silenciamiento génico postranscripcional específico de secuencia iniciado por ARNi. Durante la ARNi, el siARN induce la degradación del ARNm diana con la consiguiente inhibición específica de la secuencia de la expresión génica.

20 **[0072]** Una molécula de "ARN inhibidor", "ARNi", "ARN de interferencia pequeño" o "ARN de interferencia corto" o "siARN", "ARN de horquilla corta" o molécula de "ARNsh", o "miARN" es un dúplex de ARN de nucleótidos que es dirigida a una secuencia de ácido nucleico de interés. Como se usa en este documento, el término "siARN" es un término genérico que abarca el subconjunto de shARN y miARN. Un "dúplex de ARN" se refiere a la estructura formada por el apareamiento complementario entre dos regiones de una molécula de ARN. El siARN está "dirigido" a un gen en el sentido en que la secuencia de nucleótidos de la porción dúplex del siARN es complementaria a una secuencia de nucleótidos del gen diana. En ciertas realizaciones, los siARN están dirigidos a la secuencia que codifica la huntingtina. En algunas realizaciones, la longitud del dúplex de siARN es inferior a 30 pares de bases. En algunas realizaciones, el dúplex puede tener 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11 o 10 pares de bases en longitud. En algunas realizaciones, la longitud del dúplex es de 19 a 25 pares de bases. En cierta realización, la longitud del dúplex es de 19 o 21 pares de bases de longitud. La porción dúplex de ARN del siARN puede ser parte de una estructura de horquilla. Además de la porción 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 tiene una longitud de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 nucleótidos. En determinadas realizaciones, el bucle tiene una longitud de 18

nucleótidos. 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.

[0073] Los shARN se componen de estructuras de bucle de tallo que están diseñadas para contener una región flanqueante en 5', segmentos de la región de ARNi, una región en bucle, una región de ARNi en 3' y una región flanqueante en 3'. La mayoría de las estrategias de expresión de ARNi han utilizado ARN de horquilla corta (shARN) impulsados por promotores fuertes basados en polIII. Muchos shARN han demostrado una inactivación eficaz de las secuencias diana tanto in vitro como in vivo; sin embargo, también se descubrió que algunos shARN que demostraron una inactivación eficaz del gen diana tienen toxicidad in vivo.

[0074] Los miARN son ARN celulares pequeños (~22 nt) que se procesan a partir de transcritos de bucle de tallo precursor. Los bucles de tallo de miARN conocidos se pueden modificar para que contengan secuencias de ARNi específicas para los genes de interés. Las moléculas de miARN pueden ser preferibles a las moléculas de shARN porque los miARN se expresan de forma endógena. Por lo tanto, es poco probable que las moléculas de miARN induzcan vías de interferón que respondan a dsARN, se procesan de manera más eficiente que los shARN y se ha demostrado que silencian un 80% más de manera efectiva.

[0075] Un enfoque alternativo recientemente descubierto es el uso de miARN artificiales (andamios de pri-miARN que transportan secuencias de siARN) como vectores de ARNi. Los miARN artificiales se asemejan más naturalmente a los sustratos de ARNi endógeno y son más susceptibles a la transcripción de Pol-II (*p.ej.*, lo que permite la expresión específica de tejido de ARNi) y estrategias policistrónicas (*p.ej.*, lo que permite la entrega de múltiples secuencias de siARN). Véase la patente de EE.UU. No. 10,093,927, que se incorpora por referencia.

[0076] La unidad transcripcional de un "shARN" se compone de secuencias sentido y antisentido conectadas por un bucle de nucleótidos desapareados. Los shARN son exportados desde el núcleo por Exportin-5 y, una vez en el citoplasma, Dicer los procesa para generar siARN funcionales. Los bucles de tallo de los "miARN" están compuestos por secuencias con sentido y antisentido conectadas por un bucle de nucleótidos no apareados que normalmente se expresan como parte de transcritos primarios más grandes (pri-miARN), que son escindidos por el complejo Drosha-DGCR8 que genera intermediarios conocidos como pre-miARN, que posteriormente son exportados desde el núcleo por Exportin-5 y una vez en el citoplasma, son procesados por Dicer para generar siARNs funcionales. Un "miARN artificial" o un "vector lanzadera de miARN artificial", como se usa aquí de manera intercambiable, se refiere a un transcrito de miARN primario al que se le ha sustituido una región de bucle de tallo dúplex (al menos alrededor de 9-20 nucleótidos) mediante el procesamiento de Drosha y Dicer por las secuencias de ARNsi para el gen diana, al tiempo que se conservan los elementos estructurales dentro del bucle de tallo necesarios para el procesamiento eficaz de Drosha. El término "artificial" surge del hecho de que las secuencias flanqueantes (~35 nucleótidos corriente arriba y ~40 nucleótidos corriente abajo) surgen de sitios de enzimas de restricción dentro del sitio de clonación múltiple del siARN. Como se usa en el presente documento, el

término "miARN" abarca tanto las secuencias de miARN de origen natural como los vectores lanzadera de miARN generados artificialmente.

[0077] El siARN puede estar codificado por una secuencia de ácido nucleico y la secuencia de ácido nucleico también puede incluir un promotor. La secuencia de ácido nucleico también puede incluir una señal de poliadenilación. En algunas realizaciones, la señal de poliadenilación es una señal de poliadenilación mínima sintética o una secuencia de seis Ts.

[0078] Al diseñar ARNi, hay varios factores que deben tenerse en cuenta, como la naturaleza del siARN, la durabilidad del efecto silenciador y la elección del sistema de administración. Para producir un efecto de ARNi, el siARN que se introduce en el organismo normalmente contendrá secuencias exónicas. Además, el proceso de ARNi depende de la homología, por lo que las secuencias deben seleccionarse cuidadosamente para maximizar la especificidad del gen y minimizar la posibilidad de interferencia cruzada entre secuencias homólogas, pero no específicas del gen. Preferiblemente, el siARN exhibe más del 80%, 85%, 90%, 95%, 98% o incluso 100% de identidad entre la secuencia del siARN y el gen a inhibir. Las secuencias idénticas en menos de aproximadamente un 80% al gen diana son sustancialmente menos eficaces. Así, cuanto mayor sea la homología entre el siARN y el gen a inhibir, menos probable será que se vea afectada la expresión de genes no relacionados.

[0079] Además, el tamaño del siARN es una consideración importante. En algunas realizaciones, la presente invención se refiere a moléculas de siARN que incluyen al menos alrededor de 19-25 nucleótidos y son capaces de modular la expresión génica. En el contexto de la presente invención, el siARN tiene preferiblemente menos de 500, 200, 100, 50 o 25 nucleótidos de longitud. Más preferiblemente, el siARN tiene una longitud de aproximadamente 19 nucleótidos a aproximadamente 25 nucleótidos.

[0080] Un objetivo diana de siARN generalmente significa un polinucleótido que comprende una región que codifica 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 una región que codifica un polipéptido y una región unida operativamente a la misma que regula la expresión. Cualquier gen que se exprese en una célula puede ser diana. Preferiblemente, un gen diana es uno implicado o asociado con la progresión de actividades celulares importantes para la enfermedad o de particular interés como objeto de investigación.

B. ARN no codificantes

[0081] Como lo demuestran los proyectos de clonación de ADNc y las matrices de mosaico genómico, más del 90% del genoma humano sufre transcripción, pero no codifica proteínas. Estos productos transcripcionales se denominan ARN no codificantes de proteínas (ncARN). Una variedad de transcritos de ncARN, como los ARN ribosómicos, los ARN de transferencia, el ARN endógeno competitivo (ceARN), el ARN nuclear pequeño (snARN) y el ARN nucleolar pequeño (snoARN), son esenciales para la función celular. De manera similar, también se sabe que una gran cantidad de ncARNs cortos como micro-ARNs

(miARNs), ARNs de interferencia cortos endógenos (siARNs), PIWI-interact ARNs (piARNs) y ARNs nucleolares pequeños (snoARNs) juegan papeles reguladores importantes en células eucariotas. Estudios recientes han demostrado un grupo de transcritos largos de ncARN (lncARN) que exhiben una expresión específica del tipo de célula y se localizan en compartimentos subcelulares específicos. También se sabe que los lncARN desempeñan un papel importante durante el desarrollo y la diferenciación celular, lo que respalda la opinión de que han sido seleccionados durante el proceso evolutivo.

[0082] Los lncARN parecen tener muchas funciones diferentes. En muchos casos, parecen jugar un papel en la regulación de la actividad o localización de proteínas, o sirven como estructuras organizativas para estructuras subcelulares. En otros casos, los lncARN 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 n.º 27) cataloga poco menos de 16 000 lncARN en el genoma humano, produciendo casi 28 000 transcritos; cuando se incluyen otras bases de datos, se conocen más de 40, 000 lncARN.

[0083] Curiosamente, los lncARN pueden influir en la expresión de proteínas diana específicas en loci genómicos específicos, modular la actividad de proteínas de unión, dirigir los complejos modificadores de la cromatina a sus sitios de acción y se procesan después de la transcripción para producir numerosos pequeños ARN con cap-5'. Las vías epigenéticas también pueden regular la expresión diferencial de lncARN.

[0084] Un creciente cuerpo de evidencia también sugiere que los lncARN expresados de manera aberrante juegan un papel importante en los procesos fisiológicos normales, así como en múltiples estados de enfermedad. Los lncARN están mal regulados en diversas enfermedades, incluidas la isquemia, las enfermedades cardíacas, la enfermedad de Alzheimer, la psoriasis y la ataxia espinocerebelosa tipo 8. Esta mala regulación también se ha demostrado en varios tipos de cáncer, como el cáncer de mama, el cáncer de colon, el cáncer de próstata, el carcinoma hepatocelular y la leucemia. Varios lncARN, *p.ej* gadd74 y lncARN-RoR5, modulan los reguladores del ciclo celular como las ciclinas, las cinasas dependientes de ciclina (CDK), los inhibidores de CDK y p53 y, por lo tanto, brindan una capa adicional de flexibilidad y robustez para la progresión del ciclo celular. Además, algunos lncARN están vinculados a procesos mitóticos, como el ARN satélite centromérico, que es esencial para la formación de cinetocoros y, por lo tanto, crucial para la segregación cromosómica durante la mitosis en humanos y moscas. Otro lncARN nuclear, MA-linc1, regula la salida de la fase M al funcionar en cis para reprimir la expresión de su gen vecino Pura, un regulador de la proliferación celular.

[0085] Los lncARN son un grupo que comúnmente se define como transcritos de más de 200 nucleótidos (*p.ej* alrededor de 200 a alrededor de 1200 nt, alrededor de 2500 nt o más) que carecen de un marco de lectura abierto extendido (ORF). El término "ARN no codificante" (ncARN) incluye lncARN, así como transcritos más cortos de, *p.ej*, menos de alrededor de 200 nt, tal como alrededor de 30 a 200 nt.

[0086] Por lo tanto, en algunas realizaciones, la entrega de un ncARN, como a una estructura cerebral específica de interés, corrige los niveles de expresión de ARN aberrantes o modula los niveles de lncARN que causan enfermedades. En consecuencia, en algunas realizaciones, la presente invención proporciona un rAAV, en donde el genoma viral está diseñado para codificar un ARN no codificante terapéutico (ncARN). En algunas realizaciones, el ncARN es un ARN largo no codificante (lncARN) de aproximadamente 200 nucleótidos (nt) de longitud o mayor. En algunas realizaciones, el agente terapéutico es un ncARN de aproximadamente 25 nt o de aproximadamente 30 nt a aproximadamente 200 nt de longitud. En algunas realizaciones, el lncARN tiene una longitud de aproximadamente 200 nt a aproximadamente 1200 nt. En algunas realizaciones, el lncARN tiene una longitud de aproximadamente 200 nt a aproximadamente 1100, aproximadamente 1000, aproximadamente 900, aproximadamente 800, aproximadamente 700, aproximadamente 600, aproximadamente 500, aproximadamente 400 o aproximadamente 300 nt.

C. Sistemas CRISPR

[0087] La edición de genes es una tecnología que permite la modificación de genes diana dentro de las células vivas. Recientemente, el aprovechamiento del sistema inmunológico bacteriano de CRISPR para realizar la edición de genes 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, puede diseñarse para dirigirse a nuevos sitios con relativa facilidad alterando su secuencia guía de ARN. Este descubrimiento ha hecho que la edición de genes específicos de secuencia sea funcionalmente efectiva.

[0088] En general, el "sistema CRISPR" se refiere colectivamente a los transcritos y otros elementos involucrados en la expresión o la dirección de la actividad de los genes asociados a CRISPR ("Cas"), incluidas las secuencias que codifican un gen Cas, una secuencia tracr (trans-activating CRISPR) (*p.ej* tracrARN o un tracrARN parcial activo), una secuencia tracr-mate (que abarca una "repetición directa" y una repetición directa parcial procesada por tracrARN en el contexto de un sistema CRISPR endógeno), una secuencia guía (también denominada "espaciador" en el contexto de un sistema CRISPR endógeno) y/u otras secuencias y transcritos de un locus CRISPR.

[0089] El sistema de nucleasa CRISPR/Cas o CRISPR/Cas nucleasa puede incluir una molécula de ARN no codificante (guía) de ARN, que se une específicamente a la secuencia de ADN y una proteína Cas (*p.ej*, Cas9), con funcionalidad de nucleasa (*p.ej*, 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, *p.ej*, derivado de un organismo particular que comprende un sistema CRISPR endógeno, como *Streptococcus pyogenes*.

[0090] El sistema CRISPR puede inducir rupturas de doble cadena (DSB) en el sitio de destino, seguidas de interrupciones como se describe en este documento. En otras realizaciones, las variantes de Cas9, denominadas "nicasas", se usan para cortar una sola hebra en el sitio objetivo. Se pueden usar nicasas emparejados, *p.ej*, para mejorar la especificidad, cada uno dirigido por un par de secuencias de

direccionamiento de ARNg diferentes, de modo que tras la introducción de las rupturas simultáneamente, se introduce un saliente 5'. En otras realizaciones, Cas9 catalíticamente inactivo se fusiona con un dominio efector heterólogo tal como un represor transcripcional (*p.ej.* KRAB) o activador, para afectar la expresión génica. Alternativamente, un sistema CRISPR con un Cas9 inactivado catalíticamente comprende además un represor o activador transcripcional fusionado con una proteína de unión ribosómica.

[0091] En algunos aspectos, se introducen en la célula una nucleasa Cas y ARNg (incluida una fusión de crARN específico para la secuencia diana y un tracrARN fijo). En general, los sitios diana en el extremo 5' del ARNg dirigen la nucleasa Cas al sitio objetivo, *p.ej.* el gen, utilizando apareamiento de bases complementarias. El sitio diana puede seleccionarse en base a su ubicación inmediatamente 5' de una secuencia de motivo adyacente protoespaciador (PAM, por sus siglas en inglés), tal como típicamente NGG o NAG. En este sentido, el ARNg se dirige a la secuencia deseada modificando los primeros 20, 19, 18, 17, 16, 15, 14, 14, 12, 11 o 10 nucleótidos del ARN guía para que correspondan con la secuencia diana de ADN. En general, un sistema CRISPR se caracteriza por elementos que promueven la formación de un complejo CRISPR en el sitio de una secuencia diana. Típicamente, "secuencia diana" generalmente se refiere a una secuencia para la cual se diseña una secuencia guía para que tenga complementariedad, donde la hibridación entre la secuencia diana y una secuencia guía promueve la formación de un complejo CRISPR. No se requiere necesariamente la complementariedad completa, siempre que haya suficiente complementariedad para provocar la hibridación y promover la formación de un complejo CRISPR.

[0092] La secuencia diana puede comprender cualquier polinucleótido, como polinucleótidos de ADN o ARN. La secuencia diana puede estar ubicada en el núcleo o citoplasma de la célula, como dentro de un orgánulo de la célula. En general, una secuencia o plantilla que se puede usar para la recombinación en el locus diana que comprende las secuencias diana se denomina "plantilla de edición" o "polinucleótido de edición" o "secuencia de edición". En algunos aspectos, un polinucleótido plantilla exógeno puede denominarse plantilla de edición. En algunos aspectos, la recombinación es una recombinación homóloga.

[0093] Por lo general, en el contexto de un sistema CRISPR endógeno, la formación del complejo CRISPR (que comprende la secuencia guía hibridada con la secuencia diana y en complejo con una o más proteínas Cas) da como resultado la escisión de una o ambas cadenas en o cerca de (*p.ej.* dentro de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50 o más pares de bases desde) la secuencia diana. La secuencia tracr, que puede comprender o consistir en la totalidad o una parte de una secuencia tracr de tipo silvestre (*p.ej.* aproximadamente o más de aproximadamente 20, 26, 32, 45, 48, 54, 63, 67, 85 o más nucleótidos de una secuencia tracr de tipo silvestre), también pueden formar parte del complejo CRISPR, como por hibridación a lo largo de al menos una parte de la secuencia tracr con la totalidad o una parte de una secuencia pareja de tracr que está unida operativamente a la secuencia guía. La secuencia tracr tiene suficiente complementariedad con una secuencia pareja de tracr para hibridar y participar en la formación del complejo CRISPR, con al menos 50%, 60%, 70%, 80%, 90%, 95% o 99% de complementariedad de secuencia a lo largo de la longitud de la secuencia de pareja de tracr cuando se alinea de manera óptima.

[0094] Uno o más vectores que dirigen la expresión de uno o más elementos del sistema CRISPR pueden introducirse en la célula de manera que la expresión de los elementos del sistema CRISPR dirija la formación 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 unida a una secuencia pareja tracr y una secuencia tracr podrían estar cada una operativamente unida a elementos reguladores separados en vectores separados. La enzima Cas puede ser un gen diana bajo el control de un evento de corte y empalme alternativo regulado, como se describe en el presente documento ya sea como un minigen de gen diana quimérico o como un gen diana para un transactivador de minigen quimérico. El ARNg puede estar bajo el control de un promotor constitutivo.

[0095] Alternativamente, dos o más de los elementos expresados a partir de elementos reguladores iguales o diferentes pueden combinarse en un solo vector, con uno o más vectores adicionales proporcionando cualquier componente del sistema CRISPR no incluido en el primer vector. El vector puede comprender uno o más sitios de inserción, como una secuencia de reconocimiento de endonucleasa de restricción (también denominada "sitio de clonación"). En algunas realizaciones, uno o más sitios de inserción están ubicados corriente arriba y/o corriente abajo de uno o más elementos de secuencia de uno o más vectores. Cuando se usan múltiples secuencias guía diferentes, se puede usar una única construcción de expresión para dirigir la actividad CRISPR a múltiples secuencias diana correspondientes diferentes dentro de una célula.

[0096] Un vector puede comprender un elemento regulador ligado operativamente a una secuencia codificante enzimática que codifica la enzima CRISPR, como una proteína Cas. Los ejemplos no limitativos 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 los mismos o versiones modificadas de los mismos. Estas enzimas son conocidas; 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.

[0097] La enzima CRISPR puede ser Cas9 (*p.ej.* de *S. pyogenes* o *S. neumonía*). La enzima CRISPR puede dirigir la escisión de una o ambas cadenas en la ubicación de una secuencia diana, como dentro de la secuencia diana y/o dentro del complemento de la secuencia diana. El vector puede codificar una enzima CRISPR que está mutada con respecto a una enzima de tipo silvestre correspondiente, de modo que la enzima CRISPR mutada carece de la capacidad de escindir una o ambas cadenas de un polinucleótido diana que contiene una secuencia diana. Por ejemplo, una sustitución de aspartato por alanina (D10A) en el dominio catalítico RuvC I de Cas9 de *S. pyogenes* convierte a la Cas9 de una nucleasa que escinde ambas cadenas en una nicasa (escinde una sola cadena). En algunas realizaciones, se puede usar una nicasa Cas9 en combinación con la(s) secuencia(s) guía, *p.ej.* dos secuencias guía, que se dirigen respectivamente a las

hebras sentido y antisentido del ADN diana. Esta combinación permite cortar ambas hebras y usarlas para inducir NHEJ o HDR.

[0098] En algunas realizaciones, una secuencia codificante enzimática que codifica la enzima CRISPR tiene codones optimizados para la expresión en células particulares, tales como células eucariotas.

5 Las células eucariotas pueden ser las de un organismo particular, tal como un mamífero, o derivadas de él, incluidos, entre otros, seres humanos, ratones, ratas, conejos, perros o primates no humanos. En general, la optimización de codones se refiere a un proceso de modificación de una secuencia de ácido nucleico para mejorar la expresión en las células hospedero de interés reemplazando al menos un codón de la secuencia nativa con codones que se usan con mayor frecuencia o frecuencia en los genes de esa célula hospedero
10 manteniendo la secuencia de aminoácidos nativa. Varias especies exhiben un sesgo particular para ciertos codones de un aminoácido particular. El sesgo de codones (diferencias en el uso de codones entre organismos) a menudo se correlaciona con la eficiencia de la traducción del ARN mensajero (ARNm), que a su vez se cree que depende, entre otras cosas, de las propiedades de los codones que se traducen y la disponibilidad de determinados moléculas de ARN de transferencia (ARNt). El predominio de ARNt
15 seleccionados en una célula es generalmente un reflejo de los codones que se usan con mayor frecuencia en la síntesis de péptidos. En consecuencia, los genes se pueden adaptar para una expresión génica óptima en un organismo dado en función de la optimización de codones.

[0099] En general, una secuencia guía es cualquier secuencia polinucleotídica que tenga suficiente complementariedad con una secuencia polinucleotídica diana para hibridarse con la secuencia diana y la
20 unión específica de secuencia directa del complejo CRISPR a la secuencia diana. En algunas realizaciones, el grado de complementariedad entre una secuencia guía y su correspondiente secuencia diana, cuando se alinea de forma óptima utilizando un algoritmo de alineación adecuado, es aproximadamente o superior al 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99% o más.

[00100] El alineamiento óptimo se puede determinar con el uso de cualquier algoritmo adecuado
25 para alinear secuencias, entre los cuales se incluyen, entre otros, el algoritmo de Smith-Waterman, el algoritmo de Needleman-Wunsch, algoritmos basados en la transformada de Burrows-Wheeler (*p.ej* the Burrows Wheeler Aligner), Clustal W, 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).

30 **[00101]** La enzima CRISPR puede ser parte de una proteína de fusión que comprende uno o más dominios proteicos heterólogos. Una proteína de fusión de enzima CRISPR puede comprender cualquier secuencia de proteína adicional y, opcionalmente, una secuencia conectora entre dos dominios cualesquiera. Los ejemplos de dominios de proteínas que pueden fusionarse con una enzima CRISPR incluyen, entre otros, etiquetas de epítopos, secuencias de genes reporteros y dominios de proteínas que tienen una o más de las
35 siguientes actividades: actividad de metilasa, actividad de desmetilasa, actividad de activación de la transcripción, actividad de represión de la transcripción, actividad del factor de liberación de transcripción,

actividad de modificación de histonas, actividad de escisión de ARN y actividad de unión de ácidos nucleicos. Los ejemplos no limitantes de etiquetas de epítomos incluyen etiquetas de histidina (His), etiquetas V5, etiquetas FLAG, etiquetas de hemaglutinina (HA) de influenza, etiquetas Myc, etiquetas VSV-G y etiquetas de tiorredoxina (Trx). Los ejemplos de genes reporteros incluyen, entre otros, glutatión-5-transferasa (GST), peroxidasa de rábano picante (HRP), cloranfenicol acetiltransferasa (CAT), beta galactosidasa, beta-glucuronidasa, luciferasa, proteína verde fluorescente (GFP), HcRed, DsRed, proteína cian fluorescente (CFP), proteína amarilla fluorescente (YFP) y proteínas autofluorescentes, incluida la proteína azul fluorescente (BFP). Una enzima CRISPR se puede fusionar con una secuencia genética que codifica una proteína o un fragmento de una proteína que se une a moléculas de ADN o se une a otras moléculas celulares, incluidas, entre otras, la proteína de unión a maltosa (MBP), la etiqueta S, el dominio de unión al ADN Lex A (DBD), fusiones del dominio de unión al ADN GAL4A y fusiones de la proteína BP16 del virus del herpes simple (HSV). Los dominios adicionales que pueden formar parte de una proteína de fusión que comprende una enzima CRISPR se describen en el documento US 20110059502, incorporado aquí como referencia.

D. Proteínas terapéuticas

[00102] Algunas realizaciones se refieren a la expresión de proteínas y polipéptidos recombinantes. En algunos aspectos, la proteína o polipéptido puede modificarse para aumentar la estabilidad del suero. Por lo tanto, cuando la presente solicitud se refiere a la función o actividad de una "proteína modificada" o un "polipéptido modificado", un experto en la materia entenderá 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 relativas a una "proteína modificada" puedan implementarse con respecto a un "polipéptido modificado" y viceversa.

[00103] Las proteínas recombinantes pueden tener deleciones y/o sustituciones de aminoácidos; por tanto, una proteína con una delección, una proteína con una sustitución y una proteína con una delección y una sustitución son proteínas modificadas. En algunas realizaciones, estas proteínas pueden incluir además inserciones o aminoácidos añadidos, como proteínas de fusión o proteínas con conectores, por ejemplo. 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 aquella 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.

[00104] Las variantes de sustitución contienen típicamente el intercambio de un aminoácido por otro en uno o más sitios dentro de la proteína y pueden diseñarse para modular una o más propiedades del polipéptido, con o sin la pérdida de otras funciones o propiedades. Las sustituciones pueden ser conservativas, es decir, un aminoácido se reemplaza por uno de forma y carga similares. Las sustituciones

conservativas son bien conocidas 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.

[00105] Además de una delección o sustitución, una proteína modificada puede tener una inserción de residuos, lo que normalmente implica la adición de al menos un residuo en el polipéptido. Esto puede incluir la inserción de un péptido o polipéptido dirigido o simplemente un único resto. Las adiciones terminales, llamadas proteínas de fusión, se analizan a continuación.

[00106] El término "equivalente biológicamente funcional" se entiende bien en la técnica y se define con más detalle en el presente documento. En consecuencia, las secuencias que tienen entre aproximadamente el 70% y aproximadamente el 80%, o entre aproximadamente el 81% y aproximadamente el 90%, o incluso entre aproximadamente el 91% y aproximadamente el 99% de aminoácidos que son idénticos o funcionalmente equivalentes a los aminoácidos de un control polipéptido están incluidos, siempre que se mantenga la actividad biológica de la proteína. Una proteína recombinante puede ser biológicamente equivalente funcional a su contraparte nativa en ciertos aspectos.

[00107] También se entenderá que las secuencias de aminoácidos y ácidos nucleicos pueden incluir residuos adicionales, tales como aminoácidos N- o C-terminales adicionales, o secuencias 5' o 3', respectivamente y aún así ser esencialmente como se establece en una de las secuencias descritas en este documento, siempre que la secuencia cumpla con los criterios establecidos anteriormente, incluido el mantenimiento de la actividad de la proteína biológica cuando se trata de la expresión de proteínas. La adición de secuencias terminales se aplica particularmente a las secuencias de ácido nucleico que pueden, por ejemplo, incluir varias secuencias no codificantes que flanquean los extremos 5' o 3' de la región codificante o pueden incluir varias secuencias internas, es decir, intrones, que se sabe que ocurren dentro de los genes.

[00108] Como se usa aquí, una proteína o péptido generalmente 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 aminoácidos; y/o un péptido de aproximadamente 3 a aproximadamente 100 aminoácidos. Por conveniencia, los términos "proteína", "polipéptido" y "péptido" se usan indistintamente en este documento.

[00109] Tal y como se utiliza en el presente documento, un "residuo de aminoácido" se refiere a cualquier aminoácido de origen natural, cualquier derivado de aminoácido o cualquier imitación de aminoácido conocida en la técnica. En ciertas realizaciones, los residuos de la proteína o péptido son secuenciales, sin que ningún no-aminoácido interrumpa la secuencia de residuos de aminoácidos. En otras

realizaciones, la secuencia puede comprender uno o más elementos no aminoacídicos. En realizaciones particulares, la secuencia de residuos de la proteína o el péptido puede estar interrumpida por uno o más elementos no aminoacídicos.

[00110] 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 naturales, o al menos un aminoácido modificado o inusual.

[00111] Ciertas realizaciones de la presente invención se refieren a proteínas de fusión. Estas moléculas pueden tener una proteína terapéutica unida en el extremo N o C 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 huésped heterólogo. Otra fusión útil incluye la adición de una etiqueta de afinidad de proteína, como una etiqueta de afinidad de albúmina sérica o seis residuos de histidina, o un dominio inmunológicamente activo, como un epítipo de anticuerpo, preferiblemente 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).

[00112] Los métodos para generar proteínas de fusión son bien conocidos por los expertos en la técnica. Dichas proteínas pueden producirse, por ejemplo, mediante la síntesis de novo de la proteína de fusión completa, o mediante el empalme (splicing) de la secuencia de ADN que codifica el dominio heterólogo, seguido de la expresión de la proteína de fusión intacta.

[00113] La producción de proteínas de fusión que recuperan las actividades funcionales de las proteínas parentales puede facilitarse conectando genes con un segmento puente de ADN que codifica un conector peptídico que se empalma entre los polipéptidos conectados en tándem. El conector tendría la longitud suficiente para permitir el plegamiento adecuado de la proteína de fusión resultante.

III. Métodos de administración

[00114] Los vectores virales en algunos aspectos pueden administrarse directamente a los pacientes (*in vivo*) o pueden usarse para tratar células *in vitro* o *ex-vivo* y luego se administra a los pacientes. El término "vector" se refiere a una pequeña molécula portadora de ácido nucleico, un plásmido, un virus (*p.ej.* vector AAV, vector retroviral, vector lentiviral), u otro vehículo que pueda manipularse mediante la inserción o incorporación de un ácido nucleico. Los vectores, como los vectores víricos, se pueden utilizar para introducir/transferir secuencias de ácido nucleico en las células, de manera que la secuencia de ácido nucleico en ellas se transcriba y, si codifica una proteína, se traduzca posteriormente por las células.

[00115] Cualquier célula o mamífero adecuado puede administrarse o tratarse mediante un método o uso descrito en el presente documento. Típicamente, se sospecha que un mamífero que necesita un método descrito en este documento tiene o expresa una proteína anormal o aberrante que está asociada con un estado de enfermedad. Alternativamente, el receptor mamífero puede tener una condición que sea susceptible de

terapia de reemplazo de genes. Como se usa en el presente documento, "terapia de reemplazo de genes" se refiere a la administración al receptor de material genético exógeno que codifica un agente terapéutico y la posterior expresión in situ del material genético administrado. Por lo tanto, la frase "condición susceptible de terapia de reemplazo de genes" abarca condiciones tales como enfermedades genéticas (*es decir*, una condición de enfermedad atribuible a uno o más defectos genéticos), 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). En consecuencia, como se usa en el presente documento, el término "agente terapéutico" se refiere a cualquier agente o material que tenga un efecto benéfico sobre el receptor mamífero. Por lo tanto, "agente terapéutico" abarca tanto moléculas terapéuticas como profilácticas que tienen componentes de ácido nucleico o proteína.

[00116] Los ejemplos no limitantes de mamíferos incluyen humanos, primates no humanos (*p.ej*, simios, gibones, chimpancés, orangutanes, monos, macacos y similares), animales domésticos (*p.ej*, perros y gatos), animales de granja (*p.ej*, caballos, vacas, cabras, ovejas, cerdos) y animales de experimentación (*p.ej*, ratón, rata, conejo, conejillo de indias). En determinadas realizaciones, un mamífero es un ser humano. En ciertas realizaciones, un mamífero es un mamífero no roedor (*p.ej*, humano, porcino, caprino, ovino, equino, canino o similar). En determinadas realizaciones, un mamífero que no es un roedor es un ser humano. Un mamífero puede tener cualquier edad o en cualquier etapa de desarrollo (*p.ej*, un adulto, un adolescente, un niño, un bebé o un mamífero en el útero). Un mamífero puede ser macho o hembra. En ciertas realizaciones, un mamífero puede ser un modelo de enfermedad animal, por ejemplo, modelos animales que tienen o expresan una proteína anormal o aberrante que está asociada con un estado de enfermedad o modelos animales con expresión insuficiente de una proteína, que provoca un estado de enfermedad.

[00117] Los mamíferos (sujetos) tratados por un método o composición descritos en este documento incluyen adultos (18 años o más) y niños (menores de 18 años). Los adultos incluyen a los ancianos. Los adultos representativos tienen 50 años o más. Los niños tienen edades comprendidas entre 1 y 2 años, o entre 2 y 4, 4 y 6, 6 y 18, 8 y 10, 10 y 12, 12 y 15 y 15 y 18 años. Los niños también incluyen a los bebés. Los bebés suelen tener entre 1 y 12 meses de edad.

[00118] En ciertas realizaciones, un método incluye administrar una pluralidad de partículas virales a un mamífero como se establece en el presente documento, donde la gravedad, la frecuencia, la progresión o el tiempo de aparición de uno o más síntomas de un estado de enfermedad, tal como una enfermedad neurodegenerativa, disminuyen, se reducen, se previenen, se inhiben o se retrasan. 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 pluralidad de partículas virales a un mamífero para estabilizar, retrasar o prevenir el empeoramiento, o la progresión, o revertir un síntoma adverso de un estado de enfermedad, tal como una enfermedad neurodegenerativa.

[00119] En ciertas realizaciones, un método incluye la administración de una pluralidad de partículas virales al sistema nervioso central, o a una porción del mismo como se establece en el presente documento, de un mamífero y la gravedad, la frecuencia, la progresión o el tiempo de aparición de uno o más síntomas de un estado de enfermedad, como una enfermedad neurodegenerativa, disminuyen, se reducen, se previenen, se inhiben o se retrasan por lo menos de unos 5 a unos 10, de unos 10 a unos 25, de unos 25 a unos 50, o de unos 50 a unos 100 días.

[00120] En ciertas realizaciones, un síntoma o efecto adverso comprende un síntoma de etapa temprana, intermedia o tardía; un síntoma de comportamiento, personalidad o lenguaje; síntoma de deglución, movimiento, convulsión, temblor o inquietud; ataxia; y/o un síntoma cognitivo como la memoria, la capacidad de organización.

IV. Composiciones farmacéuticas

[00121] Como se usa en el presente documento, los términos "farmacéuticamente aceptable" y "fisiológicamente aceptable" significan una composición, formulación, líquido o sólido, o una mezcla de los mismos, biológicamente aceptable, que es adecuado para una o más vías de administración, el contacto o la entrega *in vivo*. Una composición "farmacéuticamente aceptable" o "fisiológicamente aceptable" es un material que no es indeseable desde el punto de vista biológico o de otro tipo, *p.ej.*, el material se puede administrar a un sujeto sin causar efectos biológicos indeseables sustanciales. Tales composiciones, formulaciones y composiciones "farmacéuticamente aceptables" y "fisiológicamente aceptables" pueden ser estériles. Dichas formulaciones y composiciones farmacéuticas pueden usarse, por ejemplo, en la administración de una partícula viral a un sujeto.

[00122] Dichas formulaciones y composiciones incluyen disolventes (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, recubrimientos, agentes isotónicos y promotores o retardadores de la absorción, compatibles con la administración farmacéutica o el contacto o la entrega *in vivo*. Los disolventes, soluciones y suspensiones acuosas y no acuosas pueden incluir agentes de suspensión y agentes espesantes. También pueden incorporarse a las formulaciones y composiciones compuestos activos suplementarios (*p. ej.*, conservadores, agentes antibacterianos, antivirales y antifúngicos).

[00123] Las composiciones farmacéuticas normalmente contienen un excipiente farmacéuticamente aceptable. Dichos excipientes incluyen cualquier agente farmacéutico que no induzca por sí mismo la producción de anticuerpos perjudiciales para el individuo que recibe la composición y que pueda administrarse sin toxicidad indebida. Los excipientes farmacéuticamente aceptables incluyen, entre otros, sorbitol, Tween80 y líquidos como agua, solución salina, glicerol y etanol. Se pueden incluir sales farmacéuticamente aceptables, por ejemplo, sales de ácidos minerales tales como clorhidratos, bromhidratos, fosfatos, sulfatos y similares; y las sales de ácidos orgánicos tales como acetatos, propionatos, malonatos,

benzoatos y similares. Además, en dichos vehículos pueden estar presentes sustancias auxiliares, como tensioactivos, agentes humectantes o emulsionantes, sustancias reguladoras del pH y similares.

[00124] Las composiciones farmacéuticas se pueden formular para que sean compatibles con una ruta particular de administración o entrega, como se establece en este documento o como es conocido por un experto en la técnica. Por lo tanto, las composiciones farmacéuticas incluyen vehículos, diluyentes o excipientes adecuados para la administración o suministro por diversas vías.

[00125] Las formas farmacéuticas adecuadas para inyección o infusión de partículas virales pueden incluir soluciones o dispersiones acuosas estériles que se adaptan para la preparación extemporánea de soluciones o dispersiones inyectables o de infusión estériles, opcionalmente encapsuladas en liposomas. En todos los casos, la forma final debe ser un fluido estéril y estable en las condiciones de fabricación, uso y almacenamiento. El portador o vehículo líquido puede ser un disolvente o medio de dispersión líquido que comprenda, por ejemplo, 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 adecuada puede mantenerse, por ejemplo, mediante la formación de liposomas, mediante el mantenimiento del tamaño de partícula requerido en el caso de las dispersiones o mediante el uso de tensioactivos. Agentes isotónicos, por ejemplo, azúcares, tampones o sales (*p.ej.*, cloruro de sodio) se pueden incluir. La absorción prolongada de las composiciones inyectables puede producirse mediante el uso en las composiciones de agentes que retrasan la absorción, por ejemplo, monoestearato de aluminio y gelatina.

[00126] Las soluciones o suspensiones de partículas virales pueden incluir opcionalmente uno o más de los siguientes componentes: un diluyente estéril, como agua para inyección, solución salina, como solución salina tamponada con fosfato (PBS), LCR artificial, tensoactivos, aceites fijos, un poliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido y similares), glicerina u otros disolventes sintéticos; agentes antibacterianos y antifúngicos tales como parabenos, clorobutanol, fenol, ácido ascórbico y similares; antioxidantes como ácido ascórbico o bisulfito de sodio; agentes quelantes tales como ácido etilendiaminotetraacético; tampones tales como acetatos, citratos o fosfatos y agentes para el ajuste de la tonicidad tales como cloruro de sodio o dextrosa.

[00127] Las formulaciones farmacéuticas, las composiciones y los sistemas de administración apropiados para las composiciones, los métodos y los usos de la invención son conocidos en la técnica (*ver, por ejemplo*, Remington: La ciencia y la práctica de la farmacia (2003) 20ª ed., Mack Publishing Co., Easton, Pensilvania; Ciencias farmacéuticas de Remington (1990) 18ª ed., Mack Publishing Co., Easton, Pensilvania; El Índice Merck (1996) 12ª ed., Merck Publishing Group, Whitehouse, Nueva Jersey; Principios Farmacéuticos de Formas de Dosificación Sólidas (1993), Technomic Publishing Co., Inc., Lancaster, Pa.; Ansel y Stoklosa, Cálculos farmacéuticos (2001) 11ª edición, Lippincott Williams & Wilkins, Baltimore, MD; y Poznansky *et al.*, Drug Delivery Systems (1980), RL Juliano, ed., Oxford, NY, págs. 253-315).

[00128] Las partículas virales y sus composiciones pueden formularse en forma de unidades de dosificación para facilitar la administración y la uniformidad de la dosis. La forma de unidad de dosificación como se usa en el presente documento se refiere a unidades físicamente discretas adecuadas como dosis unitarias para los sujetos mamíferos a tratar; conteniendo cada unidad una cantidad predeterminada de compuesto activo calculada para producir el efecto terapéutico deseado en asociación con el vehículo farmacéutico requerido. Las formas de las unidades de dosificación dependen del número de partículas virales que se cree necesarias para producir los efectos deseados. La cantidad necesaria se puede formular en una sola dosis, o se puede formular en múltiples unidades de dosificación. La dosis puede ajustarse a una concentración de partículas virales adecuada, combinarse opcionalmente con un agente antiinflamatorio y envasarse para su uso.

[00129] En una realización, las composiciones farmacéuticas incluirán material genético suficiente para proporcionar una cantidad terapéuticamente eficaz, *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.

[00130] Una "forma de dosificación unitaria" como se usa en el presente documento se refiere a unidades físicamente discretas adecuadas como dosificaciones unitarias para el sujeto a tratar; cada unidad contiene una cantidad predeterminada, opcionalmente en asociación con un vehículo farmacéutico (excipiente, diluyente, vehículo o agente de relleno) que, cuando se administra en una o más dosis, se calcula que produce el efecto deseado (*p.ej.* efecto profiláctico o terapéutico). Las formas de dosificación unitaria pueden estar dentro, por ejemplo, de ampollas y viales, que pueden incluir una composición líquida o una composición en estado liofilizado o liofilizado; se puede agregar un portador de líquido estéril, por ejemplo, antes de la administración o la entrega *in vivo*. Las formas farmacéuticas unitarias individuales pueden incluirse en kits o envases multidosis. Así, por ejemplo, las partículas víricas y las composiciones farmacéuticas de las mismas pueden envasarse en forma de dosificación unitaria única o múltiple para facilitar la administración y la uniformidad de la dosificación.

[00131] Las formulaciones que contienen partículas virales típicamente contienen una cantidad efectiva, siendo la cantidad efectiva determinada fácilmente por un experto en la técnica. Las partículas víricas pueden oscilar normalmente entre aproximadamente el 1% y aproximadamente el 95% (p/p) de la composición, o incluso más si es adecuado. La cantidad a administrar depende de factores tales como la edad, el peso y la condición física del mamífero o del sujeto humano considerado para el tratamiento. Las dosis efectivas pueden ser establecidas por un experto normal en la técnica a través de ensayos de rutina que establecen curvas de respuesta a la dosis.

V. Definiciones

[00132] Los términos "polinucleótido", "ácido nucleico" y "transgen" se usan indistintamente aquí 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, ARNr, ARNt y ADN o ARN inhibidores empalmados o sin empalmar (ARNi, *p.ej.* ARN de horquilla pequeña o corta (sh), microARN (miARN), ARN de interferencia pequeño o corto (si), ARN de empalme trans o ARN antisentido). Los polinucleótidos pueden incluir polinucleótidos naturales, sintéticos e intencionalmente modificados o alterados (*p.ej.* ácido nucleico variante). Los polinucleótidos pueden ser monocatenarios, bicatenarios o triples, lineales o circulares y pueden tener cualquier longitud adecuada. Al hablar de polinucleótidos, una secuencia o estructura de un polinucleótido particular puede describirse aquí de acuerdo con la convención de proporcionar la secuencia en la dirección 5' a 3'.

[00133] Un ácido nucleico que codifica un polipéptido a menudo comprende un marco de lectura abierto que codifica el polipéptido. A menos que se indique lo contrario, una secuencia de ácido nucleico particular también incluye sustituciones de codones degenerados.

[00134] Los ácidos nucleicos pueden incluir uno o más elementos reguladores o de control de la expresión vinculados operativamente al marco de lectura abierto, donde uno o más elementos reguladores están configurados 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/regulación de la expresión incluyen secuencias de iniciación de la transcripción (*p.ej.* promotores, potenciadores, una caja TATA y similares), secuencias de iniciación de la traducción, secuencias de estabilidad de ARNm, secuencias poli A, secuencias secretoras y similares. Los elementos de control/regulación de la expresión pueden obtenerse del genoma de cualquier organismo adecuado.

[00135] Un "promotor" se refiere a una secuencia de nucleótidos, generalmente corriente arriba (5') de una secuencia codificante, que dirige y/o controla la expresión de la secuencia codificante proporcionando el reconocimiento de la ARN polimerasa y otros factores necesarios para una transcripción adecuada. Un promotor pol II incluye un promotor mínimo que es una secuencia corta de ADN compuesta por una caja TATA y, opcionalmente, otras secuencias que sirven para especificar el sitio de inicio de la transcripción, al que se añaden elementos reguladores para controlar la expresión. Un promotor pol III de tipo 1 incluye tres elementos de secuencia que actúan en cis cadena abajo del sitio de inicio de la transcripción: a) elemento de secuencia 5' (bloque A); b) un elemento de secuencia intermedia (bloque I); c) elemento de secuencia 3' (bloque C). Un promotor pol III de tipo 2 incluye dos elementos esenciales de secuencia que actúan en cis cadena abajo del sitio de inicio de la transcripción: a) una caja A (elemento de secuencia 5'); y b) una caja B (elemento de secuencia 3'). Un promotor tipo 3 pol III incluye varios elementos promotores que actúan en cis antes del sitio de inicio de la transcripción, como una caja TATA tradicional, un elemento de secuencia proximal (PSE) y un elemento de secuencia distal (DSE).

[00136] Un "potenciador" 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 potencia el nivel o

la especificidad tisular de la expresión. Es capaz de funcionar en cualquier orientación (5'→3' o 3'→5') y puede funcionar incluso cuando se coloca corriente arriba o corriente abajo del promotor.

[00137] Los promotores y/o potenciadores pueden derivar en su totalidad de un gen nativo, o estar compuestos por diferentes elementos derivados de diferentes elementos que se encuentran en la naturaleza, o incluso estar compuestos por segmentos de ADN sintético. Un promotor o potenciador puede comprender secuencias de ADN que están implicadas en la unión de factores proteicos que modulan/controlan la eficacia del inicio de la transcripción en respuesta a estímulos, condiciones fisiológicas o de desarrollo.

[00138] Los ejemplos no limitantes de promotores incluyen el promotor temprano SV40, el promotor LTR del virus del tumor mamario de ratón; promotor tardío principal de adenovirus (Ad MLP); un promotor del virus del herpes simple (HSV), un promotor del citomegalovirus (CMV) tal como la región promotora temprana inmediata del CMV (CMVIE), un promotor del virus del sarcoma de Rous (RSV), promotores de la pol II, promotores de la pol III, promotores sintéticos, promotores híbridos y similares. Además, las secuencias derivadas de genes no virales, tales como el gen de la metalotioneína murina, también encontrarán uso aquí. Los promotores constitutivos ejemplares incluyen los promotores de los siguientes genes que codifican 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 los expertos en la técnica. Además, muchos promotores virales funcionan de forma constitutiva en células eucariotas. Estos incluyen: los promotores tempranos y tardíos de SV40; las repeticiones terminales largas (LTR) del virus de la leucemia de Moloney y otros retrovirus; y el promotor de la timidina cinasa del virus del herpes simple, entre muchos otros. Además, las secuencias derivadas de los promotores de miARN intrónicos, como, por ejemplo, el promotor miR107, miR206, miR208b, miR548f-2, miR569, miR590, miR566 y miR128, también encontrarán uso aquí (ver, *p.ej.*, Monteys *et al.*, 2010). En consecuencia, cualquiera de los promotores constitutivos mencionados anteriormente puede usarse para controlar la transcripción de un inserto de gen heterólogo.

[00139] Un "transgen" se usa aquí para referirse convenientemente a una secuencia de ácido nucleico/polinucleótido que se pretende o se ha introducido en una célula u organismo. Los transgenes incluyen cualquier ácido nucleico, tal como un gen que codifica un ARN inhibidor, un polipéptido o una proteína y generalmente son heterólogos con respecto a las secuencias genómicas de AAV de origen natural.

[00140] El término "transducir" se refiere a la introducción de una secuencia de ácido nucleico en una célula u organismo huésped por medio de un vector (*p.ej.*, una partícula viral). Por lo tanto, la introducción de un transgen en una célula por una partícula viral puede denominarse "transducción" de la célula. El transgen puede integrarse o no en el ácido nucleico genómico de una célula transducida. Si un transgen introducido se integra en el ácido nucleico (ADN genómico) de la célula u organismo receptor, puede mantenerse de forma estable en esa célula u organismo y luego transmitirse o ser heredado por células u organismos descendientes de la célula u organismo receptor. Finalmente, el transgen introducido puede

existir en la célula receptora o en el organismo huésped de forma extracromosómica, o solo de forma transitoria. Por lo tanto, una “célula transducida” es una célula en donde se ha introducido el transgen mediante transducción. Por tanto, una célula "transducida" es una célula en donde, o una progenie de la misma, se ha introducido un transgen. Una célula transducida se puede propagar, transcribir el transgen y expresar la proteína o el ARN inhibidor codificado. Para usos y métodos de terapia génica, una célula transducida puede estar en un mamífero.

[00141] Los transgenes bajo el control de promotores inducibles se expresan solo o en mayor grado, en presencia de un agente inductor, (*p.ej*, la transcripción bajo el control del promotor de la metalotioneína aumenta considerablemente en presencia de ciertos iones metálicos). Los promotores inducibles incluyen elementos sensibles (RE, por sus siglas en inglés) que estimulan la transcripción cuando se unen sus factores inductores. Por ejemplo, existen RE para factores séricos, hormonas esteroideas, ácido retinoico y AMP cíclico. Los promotores que contienen un RE particular pueden elegirse para obtener una respuesta inducible y, en algunos casos, el propio RE puede unirse a un promotor diferente, confiriendo así inducibilidad al gen recombinante. Por tanto, seleccionando un promotor adecuado (constitutivo frente a inducible; fuerte frente a débil), es posible controlar tanto la existencia como el nivel de expresión de un polipéptido en la célula modificada genéticamente. Si el gen que codifica el polipéptido está bajo el control de un promotor inducible, la administración del polipéptido *in situ* se activa al exponer la célula modificada genéticamente *in situ* a las condiciones para permitir la transcripción del polipéptido, *p.ej*, mediante inyección intraperitoneal de inductores específicos de los promotores inducibles que controlan la transcripción del agente. Por ejemplo, *in situ* la expresión por parte de células modificadas genéticamente de un polipéptido codificado por un gen bajo el control del promotor de la metalotioneína se potencia poniendo en contacto las células modificadas genéticamente con una solución que contiene (*es decir*, induciendo) iones metálicos *in situ*.

[00142] Un ácido nucleico/transgen está “unido operativamente” cuando se coloca en una relación funcional con otra secuencia de ácido nucleico. Un ácido nucleico/transgen que codifica un ARNi 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 ligado operativamente a un elemento de control de la expresión también puede denominarse casete de expresión.

[00143] En ciertas realizaciones, los promotores, potenciadores y similares inducibles o específicos del Sistema Nervioso Central (SNC) se emplean en los métodos y usos descritos en el presente documento. Los ejemplos no limitantes de promotores específicos del SNC incluyen los aislados de los genes de la proteína básica de mielina (MBP), la proteína ácida fibrilar glial (GFAP) y la enolasa específica de neuronas (NSE). Los ejemplos no limitantes de promotores inducibles incluyen elementos sensibles al ADN para ecdisona, tetraciclina, hipoxia e IFN.

[00144] En determinadas realizaciones, un elemento de control de la expresión comprende un potenciador de CMV. En ciertas realizaciones, un elemento de control de la expresión comprende un

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

[00145] Como se usa en el presente documento, los términos "modificar" o "variante" y las variaciones gramaticales de los mismos significan que un ácido nucleico, un polipéptido o una subsecuencia del mismo 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 conservan la 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, *p.ej*, una mutación contrasentido o sin sentido.

[00146] 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 puede modificarse genéticamente sin alterar la secuencia proteica codificada. Alternativamente, la secuencia puede modificarse genéticamente para codificar una proteína variante. Una variante de ácido nucleico o polinucleótido también puede hacer referencia a una secuencia de combinación que se ha modificado por codones para codificar una proteína que aún conserva al menos una identidad de secuencia parcial con una secuencia de referencia, como la secuencia de proteína de tipo silvestre y también se ha modificado por codones para codificar una proteína variante. Por ejemplo, algunos codones de dicha variante de ácido nucleico se cambiarán sin alterar los aminoácidos de una proteína codificada por ella y algunos codones de la variante de ácido nucleico se cambiarán, lo que a su vez cambia los aminoácidos de una proteína codificada por ella.

[00147] Los términos "proteína" y "polipéptido" se usan indistintamente en este documento. Los "polipéptidos" codificados por un "ácido nucleico" o "polinucleótido" o "transgen" descritos en el presente documento incluyen secuencias nativas parciales o de longitud completa, al igual que con proteínas polimórficas funcionales y de tipo silvestre de origen natural, subsecuencias funcionales (fragmentos) de las mismas y variantes de secuencia del mismo, siempre que el polipéptido conserve algún grado de función o actividad. En consecuencia, en los métodos y usos de la invención, no se requiere que tales 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.

[00148] Los ejemplos no limitantes de modificaciones incluyen una o más sustituciones de nucleótidos o aminoácidos (*p.ej*, 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).

[00149] Un ejemplo de una modificación de un aminoácido es una sustitución o eliminación conservativa de un aminoácido. En realizaciones particulares, una secuencia modificada o variante retiene al menos parte de una función o actividad de la secuencia no modificada (*p.ej.*, secuencia de tipo silvestre).

[00150] Otro ejemplo de modificación de un aminoácido es un péptido dirigido introducido en una proteína de la cápside de una partícula viral. Se han identificado péptidos que se dirigen a vectores virales recombinantes, al sistema nervioso central, como a distintas regiones del cerebro.

[00151] Un virus recombinante así modificado puede unirse preferentemente a un tipo de tejido (*p.ej.*, tejido del SNC) sobre otro tipo de tejido (*p.ej.*, tejido hepático). En ciertas realizaciones, un virus recombinante que lleva una proteína de la cápside modificada puede "dirigirse" al tejido del epitelio vascular cerebral uniéndose a un nivel superior al de una proteína de la cápside comparable, no modificada. Por ejemplo, un virus recombinante que tiene una proteína de la cápside modificada puede unirse al tejido del epitelio vascular cerebral en un nivel de 50% a 100% mayor que un virus recombinante no modificado.

[00152] Un "fragmento de ácido nucleico" es una porción de una molécula de ácido nucleico determinada. El ácido desoxirribonucleico (ADN) en la mayoría de los organismos es el material genético, mientras que el ácido ribonucleico (ARN) está involucrado en la transferencia de información contenida en el ADN a las proteínas. La presente invención también abarca fragmentos y variantes de las secuencias de nucleótidos y proteínas descritas o proteínas de longitud parcial codificadas por las mismas. Por "fragmento" o "porción" se entiende una longitud completa o menos de la longitud completa de la secuencia de nucleótidos que codifica, 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 actividad o función de tipo silvestre).

[00153] 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 la secuencia de aminoácidos idéntica de la proteína nativa. Las variantes alélicas naturales como estas pueden identificarse con el uso de técnicas de biología molecular, como, por ejemplo, con la reacción en cadena de la polimerasa (PCR) y técnicas de hibridación. Las secuencias de nucleótidos variantes también incluyen secuencias de nucleótidos derivadas sintéticamente, como las generadas, por ejemplo, usando mutagénesis sitio dirigida, que codifican la proteína nativa, así como las que codifican un polipéptido que tiene sustituciones de aminoácidos. Generalmente, las variantes de secuencia de nucleótidos de la invención tendrán al menos 40%, 50%, 60%, a 70%, *p.ej.*, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, hasta 79%, generalmente al menos 80%, *p.ej.*, 81%-84%, al menos 85%, *p.ej.*, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, hasta 98%, de identidad de

secuencia con la secuencia del nucleótido nativo (endógeno). 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 actividad o función de tipo silvestre).

[00154] "Variaciones conservativas" de una secuencia de ácido nucleico particular se refiere a aquellas secuencias de ácido nucleico que codifican secuencias de aminoácidos idénticas o esencialmente idénticas. Debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifican cualquier polipéptido dado. Por ejemplo, los codones CGT, CGC, CGA, CGG, AGA y AGG codifican todos el aminoácido arginina. Por lo tanto, en cada posición en donde un codón especifica una arginina, el codón se puede alterar a cualquiera de los codones correspondientes descritos sin alterar la proteína codificada. Tales variaciones de ácido nucleico son "variaciones silenciosas", que son una especie de "variaciones modificadas de forma conservativa". Cada secuencia de ácido nucleico descrita en el presente documento que codifica un polipéptido también describe todas las posibles variaciones silenciosas, excepto cuando se indique lo contrario. Un experto en la técnica reconocerá que cada codón en un ácido nucleico (excepto ATG, que normalmente es el único codón para la metionina) puede modificarse para producir una molécula funcionalmente idéntica mediante técnicas estándar. En consecuencia, cada "variación silenciosa" de un ácido nucleico que codifica un polipéptido está implícita en cada secuencia descrita.

[00155] El término "identidad sustancial" de secuencias de polinucleótidos significa que un polinucleótido comprende una secuencia que tiene al menos 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78% o 79%, o al menos el 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88% o 89%, o al menos el 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. Un experto en la técnica reconocerá que estos valores pueden ajustarse adecuadamente para determinar la identidad correspondiente de proteínas codificadas por dos secuencias de nucleótidos teniendo en cuenta la degeneración de codones, la similitud de aminoácidos, el posicionamiento del marco de lectura y similares. La identidad sustancial de las secuencias de aminoácidos para estos fines normalmente significa una identidad de secuencia de al menos el 70%, al menos el 80%, 90% o incluso al menos el 95%.

[00156] 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 identidad de secuencia con la secuencia de referencia en una ventana de comparación especificada. Una indicación de que dos secuencias polipeptídicas son idénticas es que un polipéptido es inmunológicamente reactivo con anticuerpos producidos 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 conservativa.

[00157] Los términos “tratar” y “tratamiento” se refieren tanto al tratamiento terapéutico como a las medidas profilácticas o preventivas, donde el objeto es prevenir, inhibir, reducir o disminuir un cambio o trastorno fisiológico no deseado, como el desarrollo, progresión o empeoramiento del trastorno. Para los fines de esta invención, los resultados clínicos benéficos o deseados incluyen, entre otros, el alivio de los síntomas, la disminución de la extensión de la enfermedad, la estabilización de un (*es decir*, sin empeorar ni progresar) síntoma o efecto adverso de la enfermedad, retraso o desaceleración de la progresión de la enfermedad, mejora o paliación del estado de la enfermedad y remisión (ya sea parcial o total) ya sea detectable o no detectable. “Tratamiento” también puede significar prolongar la supervivencia en comparación con la supervivencia esperada si no se recibe tratamiento. Aquellos que necesitan tratamiento incluyen aquellos que ya tienen la afección o el trastorno, así como aquellos predispuestos (*p.ej*, según lo determinado por un ensayo genético).

VI. Kits

[00158] La invención proporciona kits con material de empaquetamiento y uno o más componentes en el mismo. Un kit generalmente incluye una etiqueta o prospecto que incluye una descripción de los componentes o instrucciones de uso. *in vitro*, *in vivo*, o *ex-vivo*, de los componentes del mismo. Un kit puede contener una colección de tales componentes, *p.ej*, un ácido nucleico, vector recombinante y/o partículas virales.

[00159] Un kit se refiere a una estructura física que alberga uno o más componentes del kit. El material de empaquetamiento puede mantener los componentes estériles y puede estar hecho de material comúnmente utilizado para tales fines (*p.ej*, papel, fibra corrugada, vidrio, plástico, láminas, ampollas, viales, tubos, etc.).

[00160] Las etiquetas o insertos pueden incluir información de identificación de uno o más componentes, cantidades de dosis, farmacología clínica de los ingredientes activos, incluido el mecanismo de acción, farmacocinética y farmacodinámica. Las etiquetas o insertos pueden incluir información que identifique al fabricante, números de lote, ubicación y fecha de fabricación, fechas de vencimiento. Las etiquetas o insertos pueden incluir información que identifique la información del fabricante, los números de lote, la ubicación del fabricante y la fecha. Las etiquetas o insertos pueden incluir información sobre una enfermedad para la cual se puede usar un componente del kit. Las etiquetas o insertos 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 dosis, frecuencia o duración, e instrucciones para practicar cualquiera de los métodos, usos, protocolos de tratamiento o regímenes profilácticos o terapéuticos descritos en este documento.

[00161] Las etiquetas o insertos pueden incluir información sobre cualquier beneficio que pueda proporcionar un componente, como un beneficio profiláctico o terapéutico. Las etiquetas o insertos pueden incluir información sobre posibles efectos secundarios adversos, complicaciones o reacciones, como

advertencias para el sujeto o el médico sobre situaciones en las que no sería apropiado usar una composición particular. Los efectos secundarios adversos o complicaciones también podrían ocurrir cuando el sujeto tiene, tomará o está tomando actualmente uno o más medicamentos que pueden ser incompatibles con la composición, o el sujeto tiene, estará o está actualmente bajo otro protocolo de tratamiento o régimen terapéutico que sería incompatible con la composición y, por tanto, las instrucciones podrían incluir información sobre dichas incompatibilidades.

[00162] Las etiquetas o insertos incluyen "material impreso", *p.ej.*, papel o cartón, o separados o fijados a un componente, un kit o material de empaquetamiento (*p.ej.*, una caja), o unido a una ampollita, tubo o vial que contiene un componente del kit. Las etiquetas o insertos también pueden incluir un medio legible por computadora, como una etiqueta impresa con código de barras, un disco, un disco óptico como CD o DVD-ROM/RAM, DVD, MP3 o un medio de almacenamiento eléctrico como RAM y ROM o híbridos de éstos tales como medios de almacenamiento magnéticos/ópticos, memoria FLASH, híbridos y tarjetas de tipo memoria.

VII. Ejemplos

[00163] Los siguientes ejemplos se incluyen para demostrar realizaciones preferidas de la invención. Los expertos en la técnica deberían apreciar que las técnicas descritas en los ejemplos que siguen representan técnicas descubiertas por el inventor para funcionar bien en la práctica de la invención y, por lo tanto, pueden considerarse que constituyen modos preferidos para su práctica. Sin embargo, los expertos en la técnica deberían, a la luz de la presente divulgación, apreciar que se pueden realizar muchos cambios en las realizaciones específicas que se divulgan y seguir obteniendo un resultado igual o similar sin apartarse del espíritu y alcance de la invención.

Ejemplo 1: identificación de variantes de AAV que se dirigen al parénquima cerebral

[00164] Se desarrollaron bibliotecas avanzadas de AAV con código de barras utilizando cápsides de AAV1, AAV2 y AAV9 como plataformas de partida. Se generaron bibliotecas de presentación de péptidos de AAV1, AAV2 y AAV9 mediante la inserción de secuencias aleatorias en la posición 590 de la cápside de AAV1, la posición 587 de la cápside de AAV2 y la posición 588 de la cápside de AAV9, respectivamente (FIG. 1). La biblioteca tenía una diversidad de 1×10^7 clones únicos (FIG. 3).

[00165] Para probar la utilidad de las bibliotecas, se realizaron estudios piloto con AAV2 modificado con cápside de grado de laboratorio (título bajo, pureza baja). La biblioteca AAV2 se inyectó por vía intravenosa en dos ratones C57BL/6 a 8×10^{10} genomas vectoriales por animal. Después de 72 horas, se diseccionaron la corteza cerebral, el cerebelo y la médula espinal. Es de destacar que el corazón, el músculo esquelético y el diafragma se recolectaron por separado para identificar el tropismo muscular. Se aisló el ADN genómico viral y las secuencias de oligonucleótidos aleatorias recuperadas se amplificaron mediante PCR. Los productos de PCR del cerebro se agruparon para generar la biblioteca de segunda ronda,

que se inyectó en dos ratones a razón de 4×10^{10} genomas vectoriales por animal. Tras la segunda inyección, los genomas del vector se recuperaron como antes y se sometieron a la secuenciación NexGen junto con la biblioteca de partida y los tejidos de la primera ronda. Para comprobar si las secuencias que mostraban un enriquecimiento en los tejidos cerebrales podían efectivamente extender el alcance de los AAV2 al cerebro, se clonaron los resultados individuales en el plásmido de empaquetamiento de la cápside de los AAV2 y se generaron AAV2 que expresaban eGFP. Se fabricaron vectores de calidad comercial y se inyectaron en ratones 3×10^{10} genomas vectoriales de los virus modificados con la cápside de AAV2. Cuatro semanas más tarde, se observó fluorescencia de eGFP en el cerebro, incluso para estas variantes de bajo título.

[00166] Usando estas bibliotecas avanzadas de AAV con código de barras, se identificaron variantes de AAV que pueden dirigirse a distintas estructuras cerebrales de primates en primates no humanos. Las bibliotecas AAV1, AAV2 y AAV9 se administraron mediante inyección intracerebroventricular a un primate no humano (FIG. 2). Setenta y dos horas después de la infusión, las regiones cerebrales se microdisectaron para el aislamiento de ADN viral y el ADN de AAV se amplificó mediante PCR. Los productos se agruparon y se usaron para empaquetar la biblioteca de la segunda ronda, que se infundió en un NHP adicional. A continuación, las regiones del cerebro se microdisectaron 12 días después de la infusión. Después de dos rondas de selección, los genomas de los vectores se recuperaron y se sometieron a secuenciación de próxima generación. Específicamente, el ADN genómico extraído de los tejidos de la ronda 1 y ronda 2 se amplificó por PCR para generar bibliotecas de secuenciación de amplicón de Illumina en la posición del código de barras del vector. Las bibliotecas resultantes se agruparon y se ejecutaron en un solo carril de un Illumina HiSeq 4000 utilizando una química de lectura de un solo extremo de 100 pb. Para ilustrar la utilidad del enfoque, se probaron varias regiones diana como ejemplos: el epéndimo, las meninges y el cerebelo. En general, las secuencias que dirigían el AAVx al epéndimo, las meninges y el cerebelo eran diferentes y diferentes para los diversos serotipos.

[00167] Se generaron gráficos de enriquecimiento por ronda (FIG. 4) y mapas de calor (FIG. 5 y 6) para los siguientes tejidos: tronco encefálico, caudado, corteza cerebelosa (FIG. 5), corteza cerebral, epéndimo, globo pálido, hipocampo, meninges, nervio óptico, putamen, médula espinal, sustancia negra, núcleos subtalámicos y tálamo. Estos ilustran el enriquecimiento de los códigos de barras indicados al inicio (ronda 0) y después de las rondas uno y dos de pases in vivo a través del macaco rhesus. Para generarlos, los archivos de resultados fastq para cada combinación de tejido y ronda se procesaron utilizando un script de Python personalizado diseñado para extraer y cuantificar configuraciones de código de barras únicas observadas a nivel de ADN. Se utilizó un script R personalizado para calcular los porcentajes de códigos de barras presentes en cada muestra y convertir los códigos de barras de ADN en códigos de barras de aminoácidos. La Tabla 1 corresponde a las muestras tratadas con la biblioteca derivada de AAV1; La Tabla 2 representa los tejidos tratados con la biblioteca derivada de AAV2; La Tabla 3 corresponde a las muestras tratadas con la biblioteca derivada de AAV9. Los principales resultados se seleccionaron de estas tres bibliotecas y se ensamblaron y generaron en una biblioteca de validación que contenía 50 (AAV1), 58 (AAV2) y 30 (AAV9) códigos de barras derivados. Esta biblioteca de validación se entregó en un macaco

Rhesus adicional mediante inyección ICV. Los tejidos se recogieron y procesaron nuevamente para facilitar la recuperación de la abundancia de códigos de barras mediante secuenciación profunda. La abundancia de códigos de barras se evaluó en tejidos recuperados y la biblioteca viral de entrada. Los valores de enriquecimiento para cada código de barras se calcularon en relación con su abundancia en la biblioteca viral de entrada. Los valores de enriquecimiento relativo resultantes son un indicador sólido del rendimiento del vector entre los diversos tejidos evaluados, lo que facilita la identificación de variantes de vector AAV amplias y específicas (FIG. 7A-C).

[00168] Con el fin de validar la especificidad del tipo de célula identificado, AAV9-1999 (que tiene una secuencia peptídica dirigida de KGGGFHG; SEQ ID NO: 110) fue seleccionado para una validación *in vivo*. Se empaquetó una construcción de expresión de eGFP en AAV9-1999 bajo el control del promotor CAG. A un macaco Rhesus hembra de 5 años se le administró 1.5E13 vg de AAV9-1999 mediante inyección ICV en el ventrículo lateral izquierdo. El cerebro se recogió 30 días después de la inyección para el análisis histológico. Los cortes de cerebelo se tiñeron con H&E para representar el patrón de transducción de AAV9-1999 (FIG. 8). La cóclea también se recogió de este animal y sorprendentemente tenía una fuerte transducción de células ciliadas. Además, las cápsides de AAV9-1999 y AAV9 que contenían la construcción eGFP se administraron a crías de ratón C57BL/6 p0 mediante inyección ICV a 1E10 vg por hemisferio. Después de 21 días, los ratones fueron perfundidos. Se tomaron imágenes de cerebros enteros (FIG. 9A), secciones sagitales de cerebros enteros de 40 µm (FIG. 9B), secciones de la corteza S1 de 40 µm (FIG. 9C, izquierda), secciones del hipocampo de 40 µm (FIG. 9C, centro), secciones sagitales del cerebelo de 40 µm (FIG. 9C, derecha) y secciones coronales de la médula espinal lumbar de 40 µm (FIG. 9D) para la señal de fluorescencia de eGFP. El AAV9-1999 inyectado en crías de ratones neonatos BL/6 mostró una expresión ubicua mayor que las inyecciones de AAV9 de dosis equivalente.

[00169] Se inyectó a un macaco rhesus adulto una mezcla de cuatro AAV modificados: AAV9 con una secuencia peptídica de orientación RGDQLQWV (SEQ ID NO: 113) y una etiqueta mTagBFP2; AAV1 con una secuencia peptídica de orientación ERDRTRG (SEQ ID NO: 21) como etiqueta mTFP1; AAV2 con una secuencia peptídica de orientación GRGAPGG (SEQ ID NO: 80) y una etiqueta mNG; y el AAV2 con una secuencia peptídica DDPSARR (SEQ ID NO: 53) y una etiqueta mRuby3. Los virus se mezclaron directamente a volúmenes iguales para conseguir las dosis totales finales de cada uno, como se indica a continuación:

AAV9.RGDL mTagBFP2	6.13E12 vg total
AAV1.ERDR mTFP1	1.23E13 vg total
AAV2.GRGA mNG	8.8E12 vg total
AAV2.DDPS mRuby3	1.32E13 vg total

[00170] El cerebro se recogió 30 días después de la inyección para obtener imágenes de fluorescencia. Se obtuvieron imágenes de las secciones del ventrículo lateral (FIG. 10A), del cuarto

ventrículo (FIG. 10B) y de las meninges (FIG. 10C) para las señales de fluorescencia de mTagBFP2, mTFP2, mNG y mRuby3.

5 **[00171]** Se realizaron experimentos adicionales inyectando AAV9-1999 en la cóclea de macacos rhesus. Según los resultados de la transducción coclear, un animal recibió AAV9-1999 en su ventrículo lateral. Un solo animal recibió 3E11 vg de AAV9-1999 inyectado directamente en su ventana redonda con fenestraciones del canal (FIG. 11A-C).

* * *

10 **[00172]** Todos los métodos divulgados y reivindicados en el presente documento pueden realizarse y ejecutarse sin experimentación indebida a la luz de la presente divulgación. Si bien las composiciones y métodos de esta invención se han descrito en términos de realizaciones preferidas, será evidente para los 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. presente sin apartarse del concepto, espíritu y alcance de la invención. Más específicamente, será evidente que ciertos agentes que están relacionados tanto química como fisiológicamente pueden sustituir a los agentes descritos en el presente documento mientras que se lograrían resultados iguales o similares. Todos los sustitutos y modificaciones similares evidentes para los expertos en la técnica se consideran dentro del espíritu, alcance y concepto de la invención tal como se define en las reivindicaciones adjuntas.

REFERENCIAS

Las siguientes referencias, en la medida en que proporcionen un procedimiento ejemplar u otros detalles complementarios a los expuestos en el presente documento, se incorporan específicamente en el presente documento como referencia.

5

Patente de Estados Unidos 8,299,215

Patente de Estados Unidos 8,691,948

Publicación de solicitud de patente de Estados Unidos n.º 2018/0142259

10

Chen *et al.*, Molecular signatures of disease brain endothelia provide new sites for CNS-directed enzyme therapy. *Nat Med*, 2009. 15(10): p. 1215-8.

Chen *et al.*, Overcoming Limitations Inherent in Sulfamidase to Improve Mucopolysaccharidosis IIIA Gene Therapy. *Mol Ther*, 2018. 26(4): p. 1118-1126.

Deverman *et al.*, Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. *Nat Biotechnol*, 2014(2): p. 204-9.

15

Hartz *et al.*, Isolation of Cerebral Capillaries from Fresh Human Brain Tissue. *J Vis Exp*, 2018(139).

Hordeaux *et al.*, The Neurotropic Properties of AAV-PHP.B Are Limited to C57BL/6J Mice. *Mol Ther*, 2018. 26(3): p. 664-668.

Katz *et al.*, AAV gene transfer delays disease onset in a TPP1-deficient canine model of the late infantile form of Batten disease. *Sci Transl Med*, 2015. 7(313): p. 313ra180.

20

Keiser *et al.*, Broad distribution of ataxin 1 silencing in rhesus cerebella for spinocerebellar ataxia type 1 therapy. *Brain*, 2015. 138(Pt 12): p. 3555-66.

Koerber *et al.*, DNA shuffling of adeno-associated virus yields functionally diverse viral progeny. *Mol Ther*, 2008. 16(10): p. 1703-9.

25

Matsuzaki *et al.*, Intravenous administration of the adeno-associated virus-PHP.B capsid fails to upregulate transduction efficiency in the marmoset brain. *Neurosci Lett*, 2018. 665: p. 182-188.

McBride *et al.*, Preclinical safety of RNAi-mediated HTT suppression in the rhesus macaque as a potential therapy for Huntington's disease. *Mol Ther*, 2011. 19(12): p. 2152-62.

Monteys *et al.*, CRISPR/Cas9 Editing of the Mutant Huntingtin Allele In Vitro and In Vivo. *Mol Ther*, 2017. 25(1): p. 12-23.

30

Muller *et al.*, Random peptide libraries displayed on adeno-associated virus to select for targeted gene therapy vectors. *Nat Biotechnol*, 2003. 21(9): p. 1040-6.

Schaffer & Maheshri, Directed evolution of AAV mutants for enhanced gene delivery. *Conf Proc IEEE Eng Med Biol Soc*, 2004. 5: p. 3520-3.

35

Zhong *et al.*, Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc Natl Acad Sci U S A*, 2008. 105(22): p. 7827-32.