(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 739 288**

(51) Int. CI.:

C12N 15/86 (2006.01)
C07K 14/435 (2006.01)
C07K 7/06 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **12.09.2014 PCT/US2014/055490**

(87) Fecha y número de publicación internacional: **19.03.2015 WO15038958**

(96) Fecha de presentación y número de la solicitud europea: **12.09.2014 E 14843244 (6)**

(97) Fecha y número de publicación de la concesión europea: **01.05.2019 EP 3044318**

(54) Título: **Recuperación selectiva**

(30) Prioridad:

13.09.2013 US 201361877506 P
24.04.2014 US 201461983624 P
03.07.2014 US 201462020658 P
06.08.2014 US 201462034060 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
30.01.2020

(73) Titular/es:

CALIFORNIA INSTITUTE OF TECHNOLOGY (100.0%)
1200 East California Street
Pasadena, CA 91125, US

(72) Inventor/es:

DEVERMAN, BENJAMIN E.;
PATTERSON, PAUL H. y
GRADINARU, VIVIANA

(74) Agente/Representante:

IZQUIERDO BLANCO, María Alicia

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

DESCRIPCIÓN

Recuperación selectiva

CAMPO DE LA INVENCIÓN

[0001] La presente invención se refiere a métodos para la recuperación selectiva y proteínas de puntería y/o métodos.

ANTECEDENTES

[0002] Los virus recombinantes adeno-asociados (rAAV) son vectores para aplicaciones de transferencia génica *in vivo.* Varias terapias génicas basadas en rAAV han demostrado ser eficaces, sobre todo para el tratamiento de la amaurosis congénita de Leber, la hemofilia asociada con la deficiencia de factor IX y la deficiencia de lipoproteína lipasa (Simonelli et al 2010; Nathwani et al 2011; Gaudet et al. 2010). Recientemente, la primera terapia génica basada en rAAV, Glybera, fue aprobada por la Agencia Europea de Medicamentos para el tratamiento de la deficiencia de lipoproteínas lipasa. Los rAAV también han demostrado éxito en los modelos preclínicos de una gran variedad de enfermedades, como el síndrome de Rett, la ELA congénita, el Parkinson, la enfermedad de Huntington, la atrofia muscular espinal, entre otras y para la administración profiláctica de anticuerpos neutralizantes amplios contra enfermedades infecciosas como el VIH y la gripe (Garg et al 2013; Valori et al. 2010; Foust et al. 2010; Foust et al 2013; Southwell et al 2009; Balazs et al 2011; y Balazs et al 2013). Además, los rAAV también son vectores populares para el suministro *in vivo* de transgenes para estudios científicos no terapéuticos, como los optogénicos.

[0003] El documento US2013/0195801 describe vectores de AAV que se dirigen al sistema nervioso central (SNC).

RESUMEN DE LA INVENCIÓN

[0004] La invención proporciona un vector de AAV que comprende una secuencia de aminoácidos que comprende la secuencia de TLAVP- FK (SEQ ID NO: 1), en la que la secuencia de aminoácidos es parte de una proteína de la cápside del vector AAV.

[0005] La invención también proporciona un péptido diana de sistema nervioso central (SNC), comprendiendo el péptido una secuencia de aminoácidos de SEQ ID NO: 1; opcionalmente conjugado adicionalmente con una nanopartícula, una segunda molécula o una proteína de la cápside viral.

[0006] La invención también proporciona un plásmido rep-cap que codifica un gen de la cápside de un vector de AAV, comprendiendo gen de la cápside una secuencia de ácido nucleico que codifica la secuencia de aminoácidos TLAVPFK (SEQ ID NO: 1), en donde el rep-cap plásmido codifica solo aquellas secuencias dentro del gen rep requerido para la expresión génica y el empalme del producto del gen de la cápside.

[0007] La invención también proporciona una proteína de la cápside que comprende una secuencia de aminoácidos que comprende la secuencia de TLAVPFK (SEQ ID NO: 1).

RESUMEN DE LA DIVULGACIÓN

[0008] Un vector de AAV se da a conocer que comprende una secuencia de aminoácidos que comprende al menos 4 aminoácidos contiguos de la secuencia TLAVPFK (SEQ ID NO: 1) o KFPVALT (SEQ ID NO: 3).

[0009] Un sistema nervioso central se da a conocer péptido diana que comprende una secuencia de aminoácidos de SEQ ID NO: 1 (o cualquiera de las secuencias de aminoácidos en la Figura 31).

[0010] Se describe una secuencia de ácido nucleico que codifica cualquiera de los cuatro aminoácidos contiguos en TLAVPFK (SEQ ID NO: 1) o en KFPVALT (SEQ ID NO: 3) (o para cualquiera de las secuencias de la Figura 31).

[0011] Se describe un método de entrega de una secuencia de ácido nucleico a un sistema nervioso. El método puede comprender proporcionar una proteína que comprende TLAVPFK (SEQ ID NO: 1) (o cualquiera de las otras secuencias de direccionamiento descritas en el presente documento, por ejemplo, en la Figura 31), en donde la proteína es parte de una cápside de un AAV, y en donde el AAV comprende una secuencia de ácido nucleico para ser administrada a un sistema nervioso; y administrando el AAV al sujeto.

[0012] Un genoma rAAV se describe que comprende al menos una repetición terminal invertida configurada para permitir el empaquetamiento en un vector y un gen de la cápside.

[0013] Un sistema de plásmido se describe que comprende un primer plásmido que comprende un/9 rep-cap plásmido auxiliar AAV2 modificado configurado de tal manera que elimina al menos uno de VP1, VP2, VP3 o expresión y un segundo plásmido que comprende un plásmido rAAV-cap-in-cis.

[0014] Se describe un método de desarrollo de una cápside con una característica deseada. El método puede comprender proporcionar una población de genomas de rAAV (de cualquiera de los descritos en este documento), seleccionar la población mediante un conjunto específico de criterios y seleccionar el genoma de rAAV que cumpla con los criterios de selección.

[0015] Se da a conocer una proteína cápside que comprende una secuencia de aminoácidos que comprende al menos 4 aminoácidos contiguos de la secuencia TLAVPFK (SEQ ID NO: 1) o KFPVALT (SEQ ID NO: 3) (o de cualquiera de las secuencias en Figura 31).

[0016] Una biblioteca de secuencias de ácido nucleico se da a conocer. La biblioteca puede comprender un elemento seleccionable y una o más secuencias de reconocimiento de recombinasa.

[0017] Se describe un método para desarrollar una cápside con una característica deseada. El método comprende proporcionar una biblioteca de plásmidos que comprende un gen de la cápside y al menos una secuencia de reconocimiento de recombinasa, configurada de tal manera que permita un cambio dependiente de la recombinasa en una secuencia de un plásmido de la biblioteca que comprende el gen de la cápside que es detectable cambio. El método puede comprender, además, seleccionar la población según un conjunto específico de criterios y seleccionar el genoma de rAAV que cumpla con los criterios de selección.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0018]

FIG. 1A representa varias proteínas de direccionamiento (en este caso, ejemplos de proteínas de la cápside del AAV), incluyendo G2B-13, G2B-26, TH1.1-32 y TH1.1-35.
FIG. 1B representa algunos casos de recuperación selectiva de proteínas de la cápside.
FIG. 2A representa algunos casos de manipulación del genoma de AAV.
FIG. 2B representa algunos casos de manipulación del gen de la cápside.
FIG. 2C representa un diagrama de flujo para algunos casos de recuperación selectiva.
FIG. **2D** representa un esquema de los genes AAV y sus productos conocidos. El gen rep AAV hace cuatro productos de proteínas que se muestran en negro. El gen de la cápside produce tres proteínas de la cápside estructural (VP1-3) de un marco de lectura mediante una combinación de empalme alternativo y codones iniciadores alternativos. Además, el gen de la cápside también codifica una proteína adicional, la proteína activadora del ensamblaje (AAP), que se expresa desde un marco de lectura alternativo.
FIG. **2E** es un esquema para el diseño de construcciones utilizadas para algunos casos del método de selección de biblioteca de cápside basado en rAAV. Se inserta un gen de la cápside dentro de un genoma de AAV recombinante flanqueado por ITR. La expresión y el empalme de los productos del gen de la cápside del AAV están controlados por el promotor p41 del AAV5 en sentido ascendente de las secuencias rep del AAV2 que contienen el donante de empalme y las secuencias del intrón para los productos del gen de la cápside. Al eliminar la mayor parte del gen rep, el espacio (representado por las líneas de puntos) está disponible dentro del genoma cap-in-cis de rAAV para la inserción de elementos adicionales.
FIG. **2F** es un esquema que muestra los componentes AAV del plásmido auxiliar rep-AAP. Se insertaron cinco codones de parada dentro del marco de lectura de VP de cápside para garantizar que se eliminen las expresiones de VP1, VP2 y VP3 del ayudante representante AAP.
FIGS. 3A y 3B. Una estrategia para la recuperación dependiente de recombinasa de secuencias de células diana transducidas. (**FIG**. **3A**) El esquema muestra el genoma de AAV de cap-in-cis. Se puede usar un fragmento del promotor de ubiquitina C para dirigir la expresión de un indicador de mCherry seguido de una secuencia sintética de poliA. Un gen de la cápside del AAV, controlado por secuencias reguladoras de repetición, es seguido por una señal de poliA tardía de SV40 con flanco lox71 y lox66. El sitio lox66 se invierte respecto a lox71. En esta configuración, Cre media la inversión (**FIG**. **3B**) de la secuencia flanqueada por los sitios lox mutantes. Después de la inversión, se generan lox72 mutante doble y un sitio loxP, lo que reduce la eficiencia de la inversión al estado original. Usando los cebadores de PCR representados por las flechas en el esquema, las secuencias de tapa se pueden recuperar selectivamente de genomas que han sufrido una inversión dependiente de Cre.
FIGS. 4A-4D. Estrategias alternativas de lax para la recuperación dependiente de la recombinasa. (FIG. 4A) Los sitios loxP o lox71 y lox66 invertidos simples pueden reemplazarse por loxP doble (triángulo blanco) y lox2272 (triángulo negro) para una recu- peración irreversible. **(Figura 4B)** Los sitios loxP o sitios similares (lox2272 u otras variantes) insertados en la misma orientación para mediar una eliminación también permiten la amplificación selectiva dependiente de la recombinación. La recuperación específica de la recombinación se puede lograr realizando la recuperación basada en PCR con tres cebadores: un 5' de la secuencia aleatoria (flecha negra), un cebador inverso corriente abajo de la secuencia 3' loxP (flecha gris oscuro) y un segundo cebador directo que se une específicamente dentro de la secuencia eliminada (flecha gris claro). Este cebador compite con el cebador directo 5' (negro) durante la amplificación, lo que reduce la amplificación de las secuencias de cápside de secuencias no recombinadas. La recuperación de secuencias no recombinadas también puede reducirse mediante la digestión con una enzima que reconoce un sitio dentro de la secuencia eliminada por la recombinasa. Alternativamente, los productos dependientes de Cre e independientes pueden separarse por tamaño mediante electroforesis en gel. (**FIG. 4C y 4D**) Los sitios invertidos loxP, lox71 y lox66 (mostrados), o DIO, FLEX pueden ubicarse en configuraciones alternativas. (**FIG**. **4C**)

muestra los sitios lox en una orientación invertida que rodea las secuencias rep y cap, que se pueden invertir en presencia de Cre. (**FIG. 4D**) El esquema muestra la opción de flanquear el reportero con sitios lox. En este caso, el reportero está invertido y puede expresarse después de la recombinación.

FIGs. 5A-5C. Un ayudante de Rep/AAP dividido y un vector rAAV-Cap-lox producen virus de alto título. (FIG. 5A) Copias del genoma de AAV (GC) resistentes a la ADNasa producidas con el genoma cap-in AIS dividido AAV2/9 y AAV9 dividido (izquierda), el gen-AAP rep/AAV2/9 y un genoma rAAV que expresa mCherry (no cápside - medio) o un ayudante de control AAV2/9 rep/cap con el mismo AAV2: genoma mCherry (derecha). **(FIG. 5B)** GC virales resistentes a la ADNasa obtenidos de preparaciones a gran escala (placa de 7-10 150 mm) de bibliotecas con secuencias aleatorias de 7 unidades que reemplazan los aminoácidos 452-8 de la cápside del AAV9 (izquierda) o se insertan después del aminoácido 588 (derecha) (n = 4/por biblioteca ± SEM). **(FIG. 5C)** Los fragmentos de PCR que contienen las bibliotecas de variación de la secuencia de la cápside (negro, 452-8 o gris claro, 588) se generan y clonan en un vector rAAV9R-delta-X/Acap-in-cis que se ha modificado para insertar los sitios de restricción únicos XbaI (X) y AgeI (A) que flanquean la región a modificar.

FIGs. 6A-6C. Recuperación de la secuencia dependiente de Cre después de la selección en células transgénicas o células Cre+. (FIG. 6A) El esquema muestra una vista general del proceso de selección. En el ejemplo 3, los ratones GFAP-Cre+ se inyectaron con el virus AAV conteniendo AAV9-cap-in-cis, o las bibliotecas de cápside con 7mers al azar en los aminoácidos 452-8 o 588, y los productos de la PCR se recuperaron utilizando cebadores que amplifican selectivamente secuencias de genomas de cap-in-cis que se han sometido a una inversión Cre-mediada de la secuencia 3' al gen de la cápside. **(Figura 6B)** La imagen muestra un gel de agarosa teñido con bromuro de etidio de los productos de PCR recuperados después de la segunda etapa de PCR utilizando los cebadores 1331 y 1312. **(FIG. 6C)** Los productos de PCR recuperados se clonan en el rAAV9R-delta-X/EI vector A-cap-in-cis como primer paso para generar la próxima ronda de bibliotecas de virus de cápside.

FIG. 7. Las nuevas variantes de AAV generan virus con eficiencias similares a las de AAV9. El gráfico muestra los GC virales resistentes a la ADNasa generados por plato de 150 mm de células productoras 293 casi confluentes para AAV9 y los cuatro nuevos serotipos de AAV.

FIGs. 8A-8C. Las variantes de G2B13 y G2B26 median la transducción mejorada del cerebro y la médula espinal después de la administración IV en comparación con AAV9. Se empaquetó un vector pA AAV-CAG-eGFP-2A-ffLUC-WPRE-SV40 en AAV9 (izquierda) o en las nuevas variantes G2B13 (media) o G2B26 (derecha). 1e12 GC de cada virus se inyectó IV en ratones hembras C57B1/6 de 5 semanas de edad individuales y se evaluó la expresión de GFP en el cerebro de los ratones 6 días después. **(FIG. 8A)** Los paneles muestran la fluorescencia de eGFP nativa en todo el cerebro. **(FIG. 8B)** La inmunotinción para la expresión de eGFP en los cerebros seccionados de ratones inyectados con el virus indicado muestra una transducción eficiente de múltiples tipos de células, incluidas las neuronas (n) y los astrocitos (a). **(FIG. 8C)** Los paneles muestran fluorescencia de eGFP nativa en los hígados de ratones inyectados con el virus indicado.

FIGs. 9A-9I. Las variantes de G2B13 y G2B26 median la transducción mejorada de las neuronas del SNC y la glía después de la administración IV en comparación con AAV9. Se empaquetó un vector rAAV-CAG-eGFP-2A-ffLUC-WPRE-SV40-pA en G2B13 **(FIG. 9A-9C**) o G2B26 **(FIG. 9D-9I)** y se inyectó 1e12 GC del virus indicado IV en ratones hembra C57B1/6 individuales de 5 semanas de edad. **(FIG. 9A-9B, FIG. 9D-9I)** Los paneles muestran inmunotinción para eGFP en los cerebros seccionados de ratones 6 días después de haber sido inyectados IV con G2B13: CAG-GFP2A-Luc **(FIG. 9A-9B)** o G2B26: CAG -GFP2A-Luc **(FIG. 9D, FIG. 9E, FIG. 9G-9I).** Ambos vectores muestran la transducción de varios tipos de células, incluidas las neuronas (n) y los astrocitos (a). **(FIG. 9A-9B, FIG. 9D-9E, FIG. 9G-9H**) Los paneles muestran la inmunotinción eGFP (paneles superiores) y la inmunotinción NeuN (paneles inferiores) de campos de imagen emparejados. **(FIG. 9C y FIG. 9F**) Los paneles muestran la fluorescencia de eGFP nativa en la médula espinal de ratones inyectados con G2B13 **(FIG. 9C**) o G2B26 **(FIG. 9F**). **(FIG. 9I)** El panel muestra la inmunotinción de eGFP en (paneles superiores) co-localizados (flechas) con el marcador glial Sox2 (paneles inferiores) del mismo campo de imagen.

FIG. 10. Estrategia para generar mayor diversidad mediante la combinación de secuencias recuperadas en múltiples sitios. Después de una o más rondas de selección para nuevas variantes de cápside en dos sitios diferentes, los grupos de variantes seleccionadas pueden mezclarse para generar bibliotecas que combinan las secuencias aleatorias en dos o más sitios mediante PCR superpuesta. Usando la misma estrategia, los clones individuales con secuencias novedosas en 2 sitios más también se pueden combinar para generar clones con modificaciones múltiples.

FIGs. 11A-11E. Recuperación de la secuencia dependiente de Cre después de la selección in vivo. (FIG. 11A) El esquema del Brain Explorer 2 (Allen Brain Atlas) muestra una descripción general del proceso de selección. Las bibliotecas de virus de la cápside se inyectaron bilateralmente en el estriado de los ratones TH-Cre+ (los asteriscos muestran los sitios de inyección aproximados), y las secuencias de la cápside se recuperaron de la sustancia negra (resaltada con un cuadrado blanco). **(FIG. 11B)** La imagen muestra la fluorescencia mCherry nativa de cortes de 1 mm a través del cerebro anterior que contiene los sitios de inyección del cuerpo estriado. **(FIG. 11C)** mCherry + las fibras de las neuronas del estriado se pueden ver en rodajas de SNr. El SNr y el SNc localizado dorsal al SNr se recolectaron para la recuperación de la secuencia de la cápside. **(FIG. 11D)** Los paneles muestran productos de PCR teñidos con bromuro de etidio recuperados de células TH-Cre+ de ratones inyectados con virus AAV que contienen las bibliotecas con marcadores aleatorios en los aminoácidos 452-8 o 588 (ronda 1). **(FIG. 11E**) Los genomas de Cap-in-cis se recuperaron mediante una estrategia de PCR independiente de Cre de ratones Cre+ y Cre- demostrando la presencia de virus en todas las muestras.

FIGs. 12A-12H. Las variantes TH1.1-32 y -35 muestran una transducción retrógrada rápida y eficiente de las neuronas TH+ SNc, así como de las neuronas en regiones adicionales que se sabe que se proyectan hacia el

estriado. AAV-TH1.1-32: CAG-GFP o AAVTH1.1-35: CAG-GFP se inyectaron en el cuerpo estriado de ratones adultos y los ratones se sacrificaron 7 días después para el análisis de expresión de GFP. Los paneles muestran inmunotinción para eGFP **(FIG. 12A-B, FIG. 12D y FIG. 12F-H)** o TH **(FIG. 12C y FIG. 12E). (FIG. 12A)** Expresión de GFP dentro del cuerpo estriado que rodea el sitio de inyección de la variante TH1.1-35. **(FIG. 12B-E) Los** paneles muestran la inmunotinción con GFP **(FIG. 12B y FIG. 12D)** y la inmunotinción con TH **(FIG. 12C y FIG. 12E)** dentro del mismo campo de imagen dentro del SN. La co-localización de la inmunotinción con GFP y TH+ dentro de la misma célula se indica con flechas. La expresión de GFP es evidente en la SNr y en una subpoblación de neuronas TH+ en la SNc de un ratón inyectado con TH1.1-35 **(FIG 12B-C)** o TH1.1-32 **(FIG 12D-E). (FIG. 12F)** inmunotinción de GFP en la corteza frontal de un ratón inyectado con la variante TH1.1-35. La inmunotinción de GFP en el tálamo **(FIG. 12G)** y la amígdala **(FIG. 12H)** de un ratón inyectado con la variante TH1.1-32.

FIG. **13** muestra una alineación de secuencia de AAV y parvovirus relacionados que muestran diversidad en los sitios de inserción/reemplazo de 7mer. La figura se divide en tres partes: la primera hoja representa el lado izquierdo de la figura, la segunda hoja representa la parte central de la figura y la tercera hoja de la FIG. 13 que representa el lado derecho de la hoja. Es decir, los nombres de cada una de las filas en la primera hoja están destinados a pasar a las otras dos hojas (en cada fila).

FIG. **14** representa un modelo estructural de una proteína de la cápside que muestra las regiones de bucle donde se pueden agregar o sustituir secuencias de direccionamiento.

FIG. **15** representa un vector rAAV9R-delta-X/A-cap-in-cis.

FIG. **16** representa un plásmido auxiliar AAV rep/cap que se modificó al insertar un total de 5 codones de parada dentro del gen de la cápside dentro del marco de lectura VP1, 2 y 3 (1 codón de parada interrumpe VP3, 3 interrumpe VP2 y los 5 interrumpen VP1 - **FIG 2F,** SEQ ID NO: 5).

FIG. 17 representa una plantilla de ADN (pCRII-9R-X/A plásmido EK, SEQ ID NO: 6).

FIG. 18 representa AAV9R-delta-X/A-cap-in-cis, SEQ ID NO: 7, en la que se eliminó la región de codificación entre los sitios Xbal y AgeI para evitar la producción de proteína de cápside AAV9R X/A "wt" de cualquier Vector no digerido durante la producción de virus en la biblioteca.

FIG. 19 representa una secuencia de una proteína VP1 de cápside AAV-PHP.B (AAV-G2B26).

FIG. 20 representa una secuencia de ácido nucleico para una secuencia codificante del gen de la cápside AAV-PHP.B (AAV-G2B26).

FIG. 21 representa una secuencia de ácido nucleico en la que una secuencia de ácido nucleico para una proteína de direccionamiento se clona en un plásmido auxiliar AAV Rep-Cap.

FIGs. 22A-22F. Recuperación dependiente de la recombinasa de las secuencias de la cápside del AAV de las células diana transducidas. FIG. 22A es un esquema que muestra el genoma de rAAV rAAV-cap-in-cis utilizado para la generación de bibliotecas de cápside. Cre media la inversión de la secuencia flanqueada por los sitios lox mutantes y los cebadores de PCR, representados por medio de flechas en el esquema, se utilizan para amplificar selectivamente las secuencias recombinadas. **(FIG. 22B)** El esquema muestra los componentes AAV del plásmido auxiliar Rep-AAP. Los codones de parada insertados en el marco de lectura VP eliminan VP1, VP2 y VP3. **(FIG. 22C)** Copias del genoma de AAV (GC) resistentes a la ADNasa producidas con el genoma cap-in cis de AAV2/9 rep AAP2/9 dividido (izquierda) en comparación con un ayudante de rep/cap de AAV2/9 de control con un control AAV-UBC-mCherry genoma (centro) o AAV2/9 rep-AAP y control AAV-UBC-mCherry genoma (sin límite - derecha). **(FIG. 22D)** Productos de PCR representativos que muestran la amplificación dependiente de Cre (superior) e independiente de Cre (inferior) de las secuencias de bibliotecas de cápside recuperadas de ratones TH-Cre positivos o Cre negativos. **(FIG. 22E)** Las bibliotecas de variación de la secuencia de la cápside en AA452-8 o después de AA588 de AAV9 (gradiente vertical) se generan por PCR y se clonan en un vector rAAV-Δcap-in-cis que se ha modificado para insertar sitios de restricción únicos Xbal (x) y AgeI (a) que flanquean la(s) región(es) variable(s). **(FIG. 22F)** El esquema muestra una vista general de algunas instancias del proceso de selección.

FIGs. 23A-23H. AAV-PHP.R2 media la transducción retrógrada rápida y eficiente. (FIG. 23A) Las bibliotecas de cápside se inyectaron en el cuerpo estriado (izquierda) y el tejido de la sustancia negra se recogió para la recuperación de la secuencia de la cápside 10 días después. **(FIG. 23A) Las** imágenes muestran la fluorescencia nativa de mCherry expresada a partir de los genomas de la biblioteca cap-in-cis que rodean los sitios de inyección (izquierda) y los axones positivos de mCherry de las neuronas estriatales en el SNr (derecha). **(FIG. 23B-FIG. 23H)** La variante de AAV recuperado PHP.R2 se utilizó para empaquetar AAVss-CAG-GFP y $7x10^9$ VG se inyectó en el cuerpo estriado. Siete días después, se evaluó la expresión de GFP de los ratones mediante inmunotinción. Las imágenes muestran células positivas para GFP en el sitio de inyección del cuerpo estriado **(FIG. 23B)** o en las regiones del cerebro indicadas que contienen cuerpos celulares GFP + **(FIG. 23C, FIG. 23E-I). (FIG. 23D)** muestra inmunotinción para TH en el SNc. **(FIG 23E)** muestra la inmunotinción de GFP del mismo campo mostrado en la **FIG. 23D.** La co-localización de la inmunotinción de GFP y TH dentro de la misma célula se observa con flechas. Las barras de escala son 100 um en C, EH y 20 um en 23D.

FIGs. 24A-24I. AAV-PHP.B media la transducción robusta de todo el SNC después de la administración IV. Imágenes representativas de ratones transducidos con $1x10^{12}$ VG de ssAAV-CAG-GFP-2A-Luc empaquetados en AAV9 o AAV-PHP.B. La expresión de GFP se evaluó 3 semanas después mediante inmunotinción **(FIG. 24A** y **Figura 24C**) o fluorescencia de GFP nativa **(FIG. 24B, FIG. 24D-24G**). **(FIG. 24A**) Las imágenes muestran la inmunotinción con GFP en secciones de cerebro sagital de ratones que recibieron AAV9 (izquierda), una dosis equivalente de AAV-PHP.B (medio) o $1x10^{11}$ VG de AAV-PHP.B (derecha). **(FIG. 24B**) Imágenes representativas de proyección confocal máxima de 50 um corticales (izquierda) o estriatales (derecha) de fluorescencia de eGFP nativa de cerebros de ratones tratados como en **24A. (FIG. 24C**) Casi todas las células de Calbindin$^+$ Purkinje (parte inferior) son GFP$^+$ (parte superior) 3 semanas después de la inyección IV de $1x10^{12}$ VG de AAV-PHP.B **(FIG. 24D**) Imagen representativa de la

fluorescencia de GFP nativa de la columna lumbar cable. El recuadro muestra una ampliación del área de la médula espinal ventral en caja. **(FIG. 24E**) Imagen de proyección máxima confocal de fluorescencia de GFP de una retina de montaje completo. **(FIG. 24F**) Corte transversal de la retina. **(FIG. 24G**) Imágenes de CLARITY de la fluorescencia de GFP desde la corteza (izquierda), el estriado (derecha) y la médula espinal ventral de un ratón transducido con $1x10^{12}$ VG de AAV-PHP.B. Biodistribución AAV en el cerebro **(FIG. 24H)** y los órganos periféricos **(FIG. 24I)** 25 días después de la inyección de $1x10^{11}$ VG IV en ratones adultos. N = 3 para AAV-PHP.B y N = 4 para AAV9; las barras de error muestran la desviación estándar (sd); * p <0,05, *** p <0,001, ANOVA de una vía y prueba de comparación múltiple de Bonferroni. Las barras de escala son de 1 mm en la **FIG. 24A** y la **FIG. 24D,** 50 um en **24B** y 200 µm en la **FIG. 24C.** Marcas principales en la **FIG. 24G** son 100 um.

FIGs. 25A-25J. AAV-PHP.b transduce muchos tipos de células neuronales y gliales del SNC. Las imágenes representativas muestran una inmunotinción para GFP **(FIG. 25A-25D, FIG. 25G-25I**) o fluorescencia de GFP nativa **(FIG. 25E, FIG. 25F y Figura 25J**). Las imágenes son de las regiones del cerebro indicadas en el panel o cuerpo estriado (**FIG 25E**) o hipocampo **(FIG. 25H**). Para Las FIGs. **25A-25F,** la imagen de la izquierda muestra la inmunotinción del antígeno, mientras que la imagen de la derecha muestra la expresión de GFP en cada par. Para Las **FIGs**. **25G-25J,** la imagen superior muestra la inmunotinción al antígeno indicada, las imágenes del medio muestran la expresión de GFP y las imágenes pareadas inferiores muestran vistas de mayor aumento de las áreas en cajas indicadas. Los ratones recibieron $1x10^{11}$ VG **(FIG. 25A)** o $1x10^{12}$ VG **(FIG. 25B-25J)** a las 5-6 semanas de edad y se evaluaron para la expresión eGFP 3 semanas más tarde. Las barras de escala son 50 µm en la **FIG. 25A** , y 20 µm en la **FIG. 25B-25J.** En todos los paneles, las flechas indican colocalización y los asteriscos indican células que son positivas para el antígeno indicado pero negativas para GFP.

Figura 26. El etiquetado del vector informador de AAV-PHP.B sistémico y de baja dosis se puede utilizar junto con CLARITY para el fenotipado morfológico de una sola célula. (FIG. 26A) Una imagen de baldosas de la corteza muestra etiquetado escaso de células con GFP después de la administración IV de $1x10^{10}$ VG de rAAV-PHP.B: CAG-GFP. La región resaltada por el recuadro se muestra ampliada en 3 orientaciones diferentes **(FIG. 26B**). Se pueden ver dos astrocitos haciendo contacto con un vaso sanguíneo que contiene una célula endotelial con un núcleo GFP+. Los pies finales de astrocitos se pueden ver en espiral alrededor del vaso sanguíneo. Esto se puede usar para etiquetar células individuales, evaluar la morfología de las células y ver el impacto del material en las células. **(FIG. 26C)** Una imagen del cuerpo estriado muestra etiquetado escasa de células con GFP después de la administración IV de $1x10^{10}$ VG de rAAV-PHP.B: CAG-GFP. **(FIG. 26D**) Una neurona espinosa de medio individual se resalta utilizando un método de rastreo de filamentos semiautomático (Imaris, software Bitplane). **(FIG. 26E)** La misma neurona espinosa del medio resaltada en la **FIG. 26D** se muestra aislado junto con varios procesos neuronales estrechamente asociados. Esto se puede usar para marcar células escasamente y evaluar la asociación de las células marcadas.

FIGs. 27A-27C Cebadores utilizados para generar fragmentos de biblioteca de cápside y recuperación de secuencia de cápside dependiente de Cre. **(FIG. 27A**) El esquema muestra los productos de la PCR como la sección sombreada a la derecha con 7AA de secuencia aleatoria (representada por barras verticales de múltiples capas) insertadas después del aminoácido 588 (biblioteca 588i) o reemplazando AA452-8 (biblioteca 452-8r). Los cebadores utilizados para generar estas bibliotecas están indicados por nombre y media flecha. Para la generación de la segunda biblioteca, la plantilla se modificó para eliminar un sitio de restricción de EarI que se produce naturalmente dentro del fragmento del gen de la cápside (xE). De esta manera, cualquier contaminación de la secuencia de cápside AAV amplificada podría eliminarse al digerir las bibliotecas recuperadas con EarI. **(FIG. 27B**) El esquema muestra el vector rAAV-Cap-in-cis y los cebadores utilizados para cuantificar los genomas del vector (izquierda) y recuperar secuencias que han transducido células que expresan Cre (izquierda). **(FIG. 27C**) La figura proporciona las secuencias para los cebadores mostrados en la **FIG. 27A** y la **FIG. 27B.** La Tabla 0.1 también proporciona una lista de las secuencias:

TABLA 0.1

Cebador	Propósito	Secuencia
9CapF	Paso 1: hacia delante	CAGGTCTTCACGGACTCAGACTATCAG SEQ ID Nº: 16
CDF	Paso 1: inverso por Cr	CAAGTAAAACCTCTACAAATGTGGTAAAATCG SEQ ID Nº: 17
XF	Paso 2 hacia delante	ACTCATCGACCAATACTTGTACTATCTCTCTAGAAC SEQ ID NO: 18
AR	Paso 2 inverso	GGAAGTATTCCTTGGTTTTGAACCCA SEQ ID NO: 19
TF	qPCR hacia delante	GGTCGCGGTTCTTGTTTGTGGAT SEQ ID NO: 20
TR	qPCR inverso	GCACCTTGAAGCGCATGAACTACTACT ID: 21
7Xnnk	generación de biblioteca 452-8r	CATCGACCAATACTTGTACTATCTCTCTAGAACTATTNNKNNK NNKNNKNNKNNKNNKCAAACGCTAAAATTCAGTGTGGCCGGA SEQ ID NO: 22
7xMNN	generación de biblioteca 588i	GTATTCCTTGGTTTTGAACCCAACCGGTCTGCGCCTGTGCMNN MNNMNNMNNMNNMNNMNNTTGGGCACTCTGGTGGTTTGTG SEQ ID NO: 23

FIG. 27D El esquema muestra una descripción general de un proceso utilizado para introducir la aleatorización de 2 sitios después de la primera ronda de selección (bibliotecas combinatorias). Este proceso fue utilizado para desarrollar AAV-PHP.R2. Ambas bibliotecas (452-8r y 588i) se generaron por PCR, se clonaron en el vector rAAV-Cap-in-cis y se realizó la selección de la cápside en ratones TH-Cre. Las secuencias de ambas bibliotecas se recuperaron y se combinaron utilizando una estrategia de PCR superpuesta para generar una nueva biblioteca que debería contener todas las combinaciones posibles de las secuencias 7mer recuperadas de la biblioteca 452-8r con todas las secuencias 7mer recuperadas en el 588i en una. biblioteca. Esta biblioteca se sometió a una segunda ronda de selección en ratones TH-Cre y las variantes recuperadas que mostraron signos de enriquecimiento se caracterizaron individualmente.

FIG. 27E Genomas de vectores resistentes a ADNasa (VG) obtenidos de preparaciones de variantes individuales recuperadas de las selecciones de GFAP-Cre y TH-Cre. Los rendimientos se dan como copias genómicas vectoriales por placa de 150 mm de células productoras. Las barras de error muestran sd N = 3 para AAV9, PHP.A y PHP.B. N = 1 para PHP.r. Análisis unidireccional de la varianza (ANOVA).

FIG. 28A-28E AAV-PHP.A transduce de manera más eficiente y selectiva los astrocitos del SNC. **(FIG. 28A)** Imágenes representativas de GFP inmunotinción de secciones de cerebro de ratones inyectados como adultos con $3x10^{11}$ VG de un vector que expresa AAVss-CAG-eGFP empaquetado en AAV9 o PHP.A como se indica. **(FIG. 28B**) Los paneles muestran la inmunotinción con GFP (izquierda) y los núcleos celulares (derecha) en la corteza de los ratones que recibieron AAV9 o PHP.A como se indica. Biodistribución AAV en el cerebro **(FIG. 28C)** y los órganos periféricos **(FIG. 28D)** 25 días después de la inyección de $1x10^{11}$ VG IV en ratones adultos. N = 4; las barras de error muestran la desviación estándar (sd); * p <0,05, ** p <0,01, *** p <0,001, ANOVA de una vía y prueba de comparación múltiple de Bonferroni. **(FIG. 28E**) Imágenes representativas de inmunotinción con GFP de secciones de hígado de ratones inyectados como adultos con $3x10^{11}$ VG de un ssAAV-CAG-eGFP que expresa el vector empaquetado en AAV9 o PHP.A como se indica. La barra de escala es de 50 µm en **28B.**

FIG. 29. Caracterización rápida del tropismo a nivel celular con la eliminación de todo el tejido animal. Las imágenes muestran la fluorescencia de GFP nativa en los órganos de ratones indicados 3 semanas después de la inyección IV de $1x10^{12}$ AAV9 o AAV-PHP.B como se indica. Los órganos indicados se volvieron ópticamente transparentes utilizando CLARITY basado en PARS (un método de limpieza de tejidos de todo el cuerpo - Yang et al. 2014). Las imágenes confocales de la pila Z se reconstruyeron en imágenes tridimensionales utilizando el software Imaris (Bitplane).

FIG. 30 representa un conjunto de realizaciones para aminoácidos y ácidos nucleicos de AAV9.

FIG. 31 representa un conjunto de proteínas dirigidas. Cualquiera de las proteínas dirigidas descritas en la FIG**.** 31 se puede intercambiar por cualquier otra instancia de proteína de direccionamiento particular descrita en este documento. De manera similar, cualquiera de los ácidos nucleicos descritos en la FIG**.** 31 se puede intercambiar de manera similar por cualquiera de los ácidos nucleicos particulares descritos en este documento. La figura muestra las secuencias más altamente enriquecidas recuperadas de la segunda ronda de selección para las variantes de AAV que transducen astrocitos GFAP-Cre+ después de la administración intravenosa.

DESCRIPCIÓN DETALLADA

[0019] Los rAAVs han revitalizado el campo de la terapia génica y facilitan la transferencia de genes crítica para una amplia variedad de estudios científicos básicos. Varias características hacen que los rAAV sean atractivos como vehículos de transmisión de genes: (i) proporcionan una expresión transgénica a largo plazo, (ii) no están asociados con ninguna enfermedad humana conocida, (iii) provocan respuestas inmunes relativamente débiles, (iv) son capaces de la transducción de una variedad de tipos de células que se dividen y no se dividen y (v) el genoma de rAAV se puede empaquetar en una variedad de cápsides o cubierta proteica del virus, que tienen diferentes características de transducción y tropismos tisulares. A pesar de estas ventajas, el uso de AAV para muchas aplicaciones está limitado por la falta de serotipos de cápside que pueden transducir eficientemente ciertos tipos de células difíciles y por la falta de serotipos que pueden dirigirse de manera eficiente y selectiva a un tipo de célula/órgano deseado después del suministro sistémico.

[0020] Uso de la evolución dirigida para mejorar las características de la cápside del AAV. Un enfoque que se ha utilizado para desarrollar rAAV con un mejoramiento del tipo de tejido/célula es realizar una evolución dirigida en el gen de la cápside del AAV. Por lo general, esto se hace creando una biblioteca de AAV competentes para la replicación que se modifican para introducir mutaciones aleatorias en el gen de la cápside de AAV, que codifica las proteínas de la cápside que determinan el tropismo del tejido. La biblioteca de virus de la cápside del AAV se inyecta luego en un animal o se entrega a las células en cultivo. Después de un cierto tiempo, se recuperan las secuencias de la cápside que están presentes en las células/tejido de interés. Estas secuencias recuperadas se utilizan para generar un nuevo conjunto de virus y luego se repite el proceso. A través de rondas repetidas de selección/recuperación de secuencias, las secuencias que generan cápsides que funcionan mejor (es decir, aquellas que pasan repetidamente el proceso de selección) se enriquecerán. Las cápsides que exhiben una capacidad mejorada para transducir el objetivo pueden luego recuperarse y evaluarse como clones individuales o mutarse más y someterse a rondas adicionales de selección.

[0021] La evolución dirigida se ha utilizado para generar AAV que evaden anticuerpos neutralizantes (Maheshri al 2006 et) y mejores células diana de glioma (McGuire et al., 2010), el epitelio de las vías respiratorias (Excoffon et al., 2009) y fotorreceptores en la retina después de la inyección intravítrea (Dalkara et al 2013). Además, utilizando un modelo de hígado quimérico humano/ratón, Lisowski *et al.* desarrolló un rAAV que se dirigió específicamente a los

hepatocitos humanos (2013).

[0022] En el presente documento se describen métodos para el enriquecimiento y la recuperación selectiva de secuencias con rasgos deseables de bibliotecas de variantes de secuencia utilizando una estrategia de recuperación dependiente de la recombinación. Este método es ampliamente aplicable para el enriquecimiento selectivo de secuencias de bibliotecas aleatorias que median un mayor contacto entre el ácido nucleico que contiene la secuencia aleatoria y una recombinasa que reconoce una secuencia o secuencias específicas presentes en el mismo ácido nucleico que la secuencia aleatoria. La recombinasa puede expresarse en respuesta a los estímulos deseados, en un compartimento subcelular deseado o expresarse en una población diana específica de células in vitro o in vivo.

[0023] Como ejemplo, se ha descrito aquí una opción para recuperar selectivamente las secuencias de la cápside del virus adenoasociado (AAV) que codifican proteínas de la cápside que transducen de manera más eficiente y/o selectiva las específicas de la recombinasa Cre (Cre) específica de las células diana. Los sitios de reconocimiento de Cre (loxP o variantes de los sitios loxP) se pueden insertar en un genoma de rAAV adyacente o flanquean al gen de la cápside. De esta manera, cuando el gen de la cápside entra en el núcleo de una célula que expresa Cre y se convierte en dsDNA, Cre puede inducir un evento de recombinación entre los sitios lox dentro del genoma de rAAV que resulta en una inversión o eliminación (dependiendo de sus orientaciones relativas) de la secuencia flanqueada por los sitios lox. Usando una estrategia de recuperación que depende del evento de recombinación, las secuencias de la cápside que codifican las cápsides que dirigen el genoma de rAAV al núcleo de las células Cre+ pueden enriquecerse a través de una o más rondas de selección. La evolución dirigida por el gen de la cápside es solo una aplicación de ejemplo de esta tecnología. En algunos casos, el método puede adaptarse para la selección de cualquier otra secuencia codificante o no codificante con rasgos deseables dentro de un genoma de AAV o cualquier secuencia dentro de otros virus o ácidos nucleicos no virales que alteren la interacción con la recombinasa. El método puede adaptarse para la selección de cualquier otra secuencia codificante o no codificante con rasgos deseables dentro de un genoma de AAV o cualquier secuencia dentro de otros virus o ácidos nucleicos no virales que alteren la interacción con la recombinasa. El método puede adaptarse para la selección de cualquier otra secuencia codificante o no codificante con rasgos deseables dentro de un genoma de AAV o cualquier secuencia dentro de otros virus o ácidos nucleicos no virales que alteren la interacción con la recombinasa.

[0024] Las secciones siguientes proporcionan un breve conjunto de definiciones de los diversos términos y a continuación diversas realizaciones que se han producido a través de estos métodos de selección. Después de esto, se describen variantes detalladas del método de cribado, así como ejemplos de los mismos.

Definiciones:

[0025] Ha de entenderse que tanto la descripción general anterior como la siguiente descripción detallada son ejemplares y explicativas y no son restrictivas de la invención como se reivindica. En esta aplicación, el uso del singular incluye el plural, a menos que se especifique lo contrario. En esta aplicación, el uso de "o" significa "y/o" a menos que se indique lo contrario. Además, el uso del término "incluyendo", así como otras formas, tales como "incluye" e "incluido", no es limitante. Además, términos como "elemento" o "componente" abarcan tanto elementos como componentes que comprenden una unidad y elementos y componentes que comprenden más de una subunidad, a menos que se indique lo contrario. Además, el uso del término "porción" puede incluir parte de un resto o todo el resto.

[0026] Los encabezados de sección usados en este documento son sólo para fines de organización y no deben interpretarse como limitantes del objeto descrito. Tal como se utiliza de acuerdo con la presente descripción, los siguientes términos, a menos que se indique otra cosa, se entenderá que tienen los siguientes significados:

[0027] Un "plásmido" es como un ácido nucleico que puede ser utilizado para replicar secuencias de ADN recombinante dentro de un huésped organismo. La secuencia puede ser preferiblemente ADN de doble cadena.

[0028] El término "secuencia de reconocimiento de recombinasa" o "sitio de reconocimiento de recombinasa" se refiere a una secuencia de ácido nucleico que es reconocible por una recombinasa y puede servir como sustrato para un evento de recombinación catalizada por dicha recombinasa. La secuencia puede ser ADN de doble cadena.

[0029] El término "genoma del virus" se refiere a una secuencia de ácido nucleico que está flanqueada por cis que actúan secuencias de ácido nucleico que median el embalaje del ácido nucleico en una cápside viral. Para los AAV y los parvovirus, por ejemplo, se sabe que las "repeticiones terminales invertidas" (ITR) que están ubicadas en los extremos 5' y 3' del genoma viral tienen esta función y que las ITR pueden mediar en el empaquetamiento de heterólogos, por ejemplo, genomas de virus no wt, en una cápside viral.

[0030] El término "elemento" se refiere a una parte separada o distinta de algo, por ejemplo, una secuencia de ácido nucleico con una función separada dentro de una secuencia de ácido nucleico más larga.

[0031] El término "rAAV" se refiere a un "AAV recombinante". El AAV recombinante se refiere a un genoma de AAV en el que parte o todos los genes rep y cap han sido reemplazados por secuencias heterólogas.

[0032] El término "AAV" o "virus adeno-asociado" se refiere a un dependovirus dentro del género *Parvoviridae* de virus. En este documento, AAV puede referirse a un AAV derivado de un virus "natural" de origen natural, un AAV derivado de un genoma de rAAV empaquetado en una cápside derivada de proteínas de la cápside codificada por un gen de la cápside natural y/o un genoma de rAAV empaquetado en una cápside derivada de proteínas de la cápside codificadas por un gen de la cápside natural no natural, por ejemplo, AAV-PHP.B.

[0033] El término "plásmido auxiliar rep-cápside" se refiere a un plásmido que proporciona las funciones rep y cap de genes virales. Este plásmido puede ser útil para la producción de AAVs a partir de genomas de rAAV que carecen de reps funcionales y/o las secuencias de los genes de la cápside.

[0034] El término "vector" se define como un vehículo para transportar o transferir un ácido nucleico. Ejemplos de vectores incluyen plásmidos y virus.

[0035] El término "gen de la cápside" se refiere a las secuencias de ácidos nucleicos que codifican proteínas de la cápside que se forman, o contribuyen a la formación de, la cápside, o cubierta de proteína, del virus. En el caso de AAV, la proteína de la cápside puede ser VP1, VP2 o VP3. Para otros parvovirus, los nombres y números de las proteínas de la cápside pueden diferir.

[0036] El término "gen rep" se refiere a las secuencias de ácidos nucleicos que codifican las proteínas no estructurales (rep78, Rep68, Rep52 y Rep40) requeridos para la replicación y producción de virus.

[0037] Una "biblioteca" puede ser en forma de una multiplicidad de ácidos nucleicos lineales, plásmidos, partículas virales o vectores virales. Una biblioteca incluirá al menos dos ácidos nucleicos lineales.

[0038] Cuando las secuencias de ácidos nucleicos insertados son generados al azar, N = A, C, G o T; K = G o T; M = A o C.

[0039] A menos que se especifique lo contrario, el extremo izquierdo de cualquier secuencia de polinucleótidos de cadena simple discutida en el presente documento es el extremo 5'; La dirección de la izquierda de las secuencias de polinucleótidos de doble cadena se conoce como la dirección 5'.

[0040] Como se usa en este documento, "unido operativamente" significa que los componentes a los que se aplica el término están en una relación que les permite llevar a cabo sus funciones inherentes en condiciones adecuadas. Por ejemplo, una secuencia de control en un vector que está "operativamente unida" a una secuencia codificante de proteína se liga al mismo para que la expresión de la secuencia codificante de proteína se logre en condiciones compatibles con la actividad transcripcional de las secuencias de control.

[0041] El término "célula huésped" se refiere a una célula que ha sido transformada, o es capaz de ser transformada, con una secuencia de ácido nucleico y por lo tanto expresa un gen de interés. El término incluye la progenie de la célula parental, independientemente de que la progenie sea idéntica en morfología o en composición genética a la célula parental original, siempre que el gen de interés esté presente.

[0042] El término "de origen natural" como se usa aquí se refiere a materiales que se encuentran en la naturaleza o una forma de los materiales que se encuentran en la naturaleza.

[0043] El término "tratar" y "tratamiento" incluye tratamientos terapéuticos, tratamientos profilácticos, y aplicaciones en las que uno reduce el riesgo de que un sujeto desarrollará un trastorno o otro factor de riesgo. El tratamiento no requiere la curación completa de un trastorno y abarca los casos en que se reduce los síntomas o los factores de riesgo subyacentes.

[0044] El término "prevenir" no requiere la eliminación del 100% de la posibilidad de un evento. Más bien, denota que la probabilidad de que ocurra el evento se ha reducido en presencia del compuesto o método.

[0045] Se pueden usar técnicas estándar para el ADN recombinante, la síntesis de oligonucleótidos y el cultivo y transformación de tejidos (por ejemplo, electroporación, lipofección). Las reacciones enzimáticas y las técnicas de purificación se pueden realizar de acuerdo con las especificaciones del fabricante o como se realiza comúnmente en la técnica o como se describe en el presente documento. Las técnicas y procedimientos anteriores se pueden realizar generalmente de acuerdo con métodos convencionales bien conocidos en la técnica y como se describe en varias referencias generales y más específicas que se citan y discuten a lo largo de la presente memoria descriptiva. *Ver, por ejemplo,* Sambrook et al., Molecular Cloning: A Laboratory Manual (2ª ed., Prensa de laboratorio de Cold Spring Harbor, Cold Spring Harbor, NY (1989)). A menos que se proporcionen definiciones específicas, las nomenclaturas utilizadas en relación con, y los procedimientos y técnicas de laboratorio de química analítica, química orgánica sintética y química médica y farmacéutica descritas en este documento son aquellas bien conocidas y comúnmente utilizadas en la técnica. Las técnicas estándar se pueden utilizar para síntesis químicas, análisis químicos, preparación farmacéutica, formulación y administración, y tratamiento de pacientes.

[0046] Como se describe aquí, SEQ ID NO: 4 - rAAV-cap-en-cis plásmido también puede ser denominado como: AAV-cap-en-cis, rAAV- Cap-en-cis vector, rAAV-CAP-in -cis genoma, construcción rAAV-Cap-cis, rAAV9-cap-in-cis, rAAV9R-X/A-cap-in-cis, o cap-in-cis, rAAV mCherry-cap-lox71/66 genoma rAAV-CAP-en-cis-lox, AAV9 cap-in-cis genoma, rAAV-cap-en-cis rAAV genoma, o el genoma cap-in-cis rAAV.

[0047] Como se describe aquí, SEQ ID NO: 7 (por ejemplo) - rAAV-delta-cap-en-cis también puede ser denominado como: rAAV9R- delta-X/A-cap-en-cis, rAAV9R-delta -X/A-cap-in-cis vector, rAAV-Δcap-in-cis, rAAV-cap-in-cis vector aceptor, cap-in-cis aceptor constructo, rAAV9R-delta-X/A-cap-in-cis constructor aceptor, aceptor de biblioteca rAAV-cap-in-cis, o AAV9R-delta-X/A-cap-in-cis.

[0048] Como se describe aquí, SEQ ID NO: 5 (por ejemplo) - AAV Rep-AAP ayudante también puede ser denominado como el presentante AAP, rep-AAP ayudante y ayudante REP-AAP, ayudante de AAV REP-AAP, AAV2/9 rep-AAP, o AAV2/9 REP-AAP plásmido auxiliar.

[0049] Como se describe aquí, SEQ ID NO: 6 (por ejemplo) - pCRII-9R-X/A EK plásmido o pCRII-9Cap-XE son intercambiables términos capaces.

[0050] Como se describe en el presente documento, AAV-PHP.B denota lo mismo que AAV-PHP.b, que denota las mismas cosas que AAV-G2B26.

[0051] Como se describe en el presente documento, AAV-PHP.A denota las mismas cosas que AAV-PHP.a.

[0052] Como se describe aquí, AAV-PHP.R2 denota el mismo que AAV-PHP.r, que denota el mismo que AAV-TH1.1-35.

[0053] Como se describe en el presente documento, 1253 también es referido como 9CapF; 1316 también se conoce como CDF; 1331 también se conoce como XF; 1312 también se conoce como AR; 1287 también se conoce como 7xNNK; y 1286 también se conoce como 7xMNN.

[0054] El término "sistema nervioso central" o "SNC" tal como se utiliza aquí se refiere a la técnica reconocida uso del término. El SNC incluye el cerebro, los nervios ópticos, los nervios craneales y la médula espinal. El SNC también incluye el líquido cefalorraquídeo, que llena los ventrículos del cerebro y el canal central de la médula espinal.

[0055] La administración sistémica de vectores que incluye una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 de la son particularmente adecuados para la entrega de secuencias de ADN exógeno que codifican polipéptidos, proteínas, o no codificantes de ADN, ARN o oligonucleótidos a, por ejemplo, células del SNC de sujetos afectados por una enfermedad del SNC.

Secuencia de orientación:

[0056] En algunas realizaciones, se proporciona un sistema nervioso central péptido de dirección. El péptido comprende una secuencia de aminoácidos de la SEQ ID NO: 1. En algunas realizaciones, el péptido se conjuga adicionalmente con una nanopartícula, una segunda molécula, una proteína de la cápside viral, o se inserta entre los aminoácidos 588 y 589 de AAV9 (SEQ ID NO : 2, **FIG. 30)**.

[0057] En algunas realizaciones, el sistema nervioso central péptido de dirección incluye 4 o más aminoácidos de los residuos que se superponen con los residuos 585 a 595 dentro de la SEQ ID NO: 8.

[0058] En algunos casos descritos en el presente documento, el péptido dirigido al sistema nervioso central incluye 4 o más aminoácidos contiguos de la SEQ ID NO: 1 (o cualquiera de las secuencias en la Figura 31). En algunos casos, el péptido dirigido al sistema nervioso central comprende 4-7 aminoácidos de la SEQ ID NO: 1 (o cualquiera de las secuencias en la Figura 31). En algunos casos, el péptido dirigido al sistema nervioso central comprende 4-6 aminoácidos de la SEQ ID NO: 1 (o cualquiera de las secuencias en la Figura 31). En algunos casos, el péptido dirigido al sistema nervioso central incluye una o más de las siguientes opciones: TLAVPFK (SEQ ID NO: 1); LAVPFK (SEQ ID NO: 31); AVPFK (SEQ ID NO: 32); VPFK (SEQ ID NO: 33); TLAVPF (SEQ ID NO: 34); TLAVP (SEQ ID NO: 35); o TLAV (SEQ ID NO: 36). En algunos casos, el péptido dirigido puede consistir en, consistir esencialmente en,o comprenden una o más de las secuencias en la FIG. 31. En algunos casos, se pueden alterar 2 o menos aminoácidos dentro de TLAVPFK (o para cualquiera de las secuencias dentro de la Figura 31). En algunos casos, un aminoácido puede alterarse dentro de TLAVPFK (o por cualquiera de las secuencias dentro de la Figura 31). En algunos casos, la alteración es una alteración conservadora (dentro de cualquiera de los péptidos de direccionamiento descritos en este documento). En algunos casos, la alteración es una eliminación o inserción de uno o dos aminoácidos (dentro de cualquiera de los péptidos de direccionamiento descritos en este documento). En algunos casos, el aminoácido puede incluir un aminoácido no natural. En algunos casos, la secuencia peptídica dirigida al sistema nervioso central puede ser una o más de: SQTLA, QTLAV, TLAVP, LAVPK, AVPKA, VPKAQ. En algunos casos, el péptido dirigido puede ser al menos un 75% idéntico a una o más de las secuencias anteriores, por ejemplo, al menos un 80% idéntico.

[0059] En algunos casos, la secuencia peptídica dirigida al sistema nervioso central puede invertirse, como en KFPVALT (SEQ ID NO: 3) (de manera similar, cualquiera de las secuencias en la Figura 31 también puede invertirse). En tales casos, pueden emplearse 4 o más aminoácidos contiguos. Por ejemplo, la secuencia puede comprender y/o consistir en KFPV, FPVA, PVAL, VALT, etc. En algunos casos, el péptido dirigido puede ser al menos un 75% idéntico a una o más de las secuencias anteriores, por ejemplo, al menos 80%.

[0060] En algunos casos, el péptido dirigido al sistema nervioso central comprende una secuencia de aminoácidos que comprende al menos 4 aminoácidos contiguos de la secuencia TLAVPFK (SEQ ID NO: 1) o KFPVALT (SEQ ID NO: 3) o cualquiera de las secuencias en la FIG.. 31. En algunas realizaciones, la secuencia de aminoácidos produce un aumento en la transducción de células del SNC por el AAV. En algunas realizaciones, la secuencia de aminoácidos es parte de una proteína de la cápside del vector AAV. En algunas realizaciones, la secuencia TLAVPFK (SEQ ID NO: 1 se inserta entre AA588-589 de una secuencia AAV del vector (SEQ ID NO: 2). En algunas realizaciones, la secuencia TLAVPFK (SEQ ID NO: 1 se inserta entre AA586-592 de una secuencia AAV del vector (SEQ ID NO: 2) En algunas realizaciones, la secuencia TLAVPFK (SEQ ID NO: 1 comprende además al menos dos de los aminoácidos 587, 588, 589 o 590 de la SEC ID) NO: 2.En algunos casos, el péptido dirigido puede ser al menos un 75% idéntico a una o más de las secuencias anteriores.

[0061] En algunos casos, el sistema nervioso central péptido de dirección comprende, consiste, o consiste esencialmente en uno cualquiera o más de las secuencias anteriores. En algunas realizaciones, el péptido dirigido al sistema nervioso central se inserta en un péptido más largo, como se describe en este documento.

[0062] En algunas realizaciones, el péptido de direccionamiento es parte de un AAV, como se describe en el presente documento. En algunas realizaciones, el péptido dirigido es parte de un AAV9. En algunas realizaciones, el péptido de direccionamiento se puede enlazar a cualquier molécula que debería apuntarse como se desee. En algunas realizaciones, el péptido dirigido se puede unir, sin limitación, a una proteína recombinante, un anticuerpo, una célula, un diagnóstico, un agente terapéutico, una nanomolécula, etc.

[0063] En algunas realizaciones, la secuencia de dirección es una secuencia de aminoácidos. En algunas realizaciones, la secuencia de direccionamiento es una secuencia de ácido nucleico. En algunos casos, la secuencia de direccionamiento es una proteína de la cápside que comprende una secuencia de aminoácidos que comprende al menos 4 aminoácidos contiguos de la secuencia TLAVPFK (SEQ ID NO: 1) y/o KFPVALT (SEQ ID NO: 3) y/o cualquiera de las secuencias en la FIG. 31.

[0064] Algunas opciones de secuencias de direccionamiento, como se indica en los ejemplos siguientes, se dan a conocer en la figura. 1A, que incluye G2B-13, G2B-26, TH1.1-32 y TH1.1-35, a modo de ejemplo.

[0065] En algunas realizaciones, la proteína dirigida se puede insertar en cualquier sección deseada de una proteína. En algunas realizaciones, la proteína dirigida se puede insertar en una proteína de la cápside. En algunas realizaciones, la proteína dirigida se inserta en una superficie de la proteína deseada. En algunas realizaciones, la proteína dirigida se inserta en la secuencia primaria de la proteína. En algunas realizaciones, la proteína dirigida está unida a la proteína. En algunas realizaciones, la proteína dirigida está unida covalentemente a la proteína. En algunas realizaciones, la proteína dirigida se inserta en un bucle no estructurado de la proteína deseada. En algunas realizaciones, el bucle no estructurado puede identificarse a través de un modelo estructural de la proteína.

[0066] En algunas realizaciones, el bucle no estructurado puede ser uno identificado por comparación de secuencias, tal como se muestra en la FIG. 13. La FIG. 13 muestra una alineación de la secuencia amino de la cápside VP1 de AAV y parvovirus relacionados alineados con AAV9. La identidad de secuencia se muestra como un punto. Se indican los AA que difieren de AAV9. La numeración se basa en AAV9 VP1. Solo se muestra AA 418-624, aunque tal alineamiento puede ser realizado por un experto en la técnica para cualquier sección deseada de proteína. Las barras verticales sombreadas de diferente longitud representan la conservación relativa en cada AA. Las barras más largas indican una mayor conservación. Las barras sombreadas horizontales indican los sitios de los bucles no estructurados en los que se puede insertar la proteína dirigida.

[0067] En algunas realizaciones, la ubicación de la inserción de la proteína de direccionamiento en la proteína deseada se puede lograr mediante un modelo estructural. Un ejemplo de tal modelo estructural se muestra en la FIG. 14. La FIG. 14 muestra un modelo estructural que resalta los bucles de superficie aleatorizados por inserción de secuencia dirigida. El inserto (que es una explosión) muestra un diagrama de cinta del modelo de superficie AAV9 construido en PyMol a partir del archivo 3uxl.pdb del banco de datos de proteínas AAV9. La superficie de la cápside se muestra en gris y las regiones de bucle elegidas para la inserción de la secuencia se resaltan con sombreado (AA586-592) y (AA452-458). Otras regiones de inserción o reemplazo de secuencia pueden identificarse desde regiones que no están altamente conservadas. Los ejemplos adicionales incluyen las regiones de AAV9 entre AA262-269, AA464-473, AA491-495,AA546-557 y AA659-668 o las regiones homólogas de otras proteínas de la cápside de otros AAV o parvovirus.

[0068] En algunas realizaciones, la proteína de la cápside puede comprender o consistir en la secuencia mostrada en la FIG. 19, SEQ ID NO: 8. El aminoácido subrayado es una mutación de K a R que se creó para proporcionar un sitio

de restricción Xbal único, esto puede ser opcional. Para referencias, la SEQ ID NO: 1 (TLAVPFK) está en negrita. Cualquiera de los otros péptidos de direccionamiento descritos en el presente documento (por ejemplo, en la FIG. 31) también se puede insertar en lugar de la SEQ ID NO: 1. 20 representa algunas realizaciones de una secuencia de ácido nucleico para una secuencia codificante del gen de la cápside de AAV-PHP.B (AAV-G2B26)). La secuencia de codificación de ácido nucleico recuperado de SEQ ID NO: 1 (TLAVPFK) está en negrita y subrayado. Las mutaciones introducidas para insertar o eliminar sitios de restricción se resaltan con texto en cursiva con doble subrayado.

Vectores

[0069] En algunos casos, un vector viral puede incluir una o más de las secuencias de direccionamiento observadas (por ejemplo, cualquiera del sistema nervioso central de orientación péptidos indicados en el presente documento o cualquier péptido proporcionado por los métodos de cribado descritos en el presente documento). En algunas realizaciones, se puede proporcionar un vector de AAV que comprende una secuencia TLAVPFK (SEQ ID NO: 1).

[0070] En algunas realizaciones, se pueden emplear una o más secuencias de direccionamiento en un solo sistema. Por ejemplo, uno puede emplear una o más secuencias de direccionamiento y también modificar otros sitios para reducir el reconocimiento de los AAV por los anticuerpos preexistentes presentes en el huésped, como un humano. En algunas realizaciones, el vector de AAV puede incluir una cápside, que influye en el tropismo/direccionamiento, la velocidad de expresión y la posible respuesta inmune. El vector también puede incluir el rAAV, cuyo genoma transporta los aspectos transgénicos/terapéuticos (por ejemplo, secuencias) junto con secuencias reguladoras. En algunas realizaciones, el vector puede incluir la secuencia de direccionamiento dentro de/sobre un sustrato que es o transporta la molécula deseada (molécula terapéutica, molécula de diagnóstico, etc.).

[0071] En algunos casos, una cualquiera o más de las secuencias de direccionamiento descritas en este documento pueden ser incorporados en un vector.

[0072] En algunas realizaciones, la secuencia de TLAVPFK (SEQ ID NO: 1) da como resultado un aumento en la transducción de células del SNC de un virus que contiene el vector. En algunas realizaciones, el aumento es de al menos 2, 10, 20, 30, 40, 50, 60, 70, 80, 90, o al menos 100 veces más que la transducción sin la secuencia de orientación. En algunas realizaciones, hay un aumento de 40-90 veces en la transducción del SNC, en comparación con la transducción de AAV9.

[0073] En algunas realizaciones, la secuencia TLAVPFK (SEQ ID NO: 1) es parte de una proteína de la cápside del vector AAV. En algunas realizaciones, la secuencia TLAVPFK (SEQ ID NO: 1) se inserta entre AA588-589 de una secuencia AAV del vector (SEQ ID NO: 2). En algunas realizaciones, la secuencia TLAVPFK (SEQ ID NO: 1) se inserta dentro de AA452-458 de una secuencia AAV del vector (SEQ ID NO: 2). En algunas realizaciones, la secuencia TLAVPFK (SEQ ID NO: 1) se inserta dentro de AA491-495 de una secuencia AAV del vector (SEQ ID NO: 2). En algunas realizaciones, la secuencia TLAVPFK (SEQ ID NO: 1) se inserta dentro de AA546-557 de una secuencia AAV del vector (SEQ ID NO: 2). En algunos casos, cualquiera de las secuencias de direccionamiento (o combinación de las mismas) en la FIG. 31 se puede usar y/o sustituir por cualquiera de las realizaciones proporcionadas en este documento con respecto a la SEQ ID NO: 1. Así, por ejemplo,una o más de las secuencias dentro de la FIG. 31 se puede insertar entre AA588-589 de una secuencia AAV del vector (SEQ ID NO: 2). En algunos casos, la secuencia de destino puede ser una o más de: SVSKPFL (SEQ ID NO: 28); FTLTTPK (SEQ ID NO: 29); o MNATKNV (SEQ ID NO: 30). FIG. 31 representa algunas de las secuencias más altamente enriquecidas recuperadas de la segunda ronda de selección de variantes de AAV que transducen astrocitos GFAP-Cre+ después de la administración intravenosa.

[0074] En algunos casos, la secuencia de dirección que es parte del vector puede comprender cualquiera de los cuatro miembros de AA contiguos dentro de AAV9 VP1 AA585-598 de la SEQ ID NO: 8.

[0075] Mientras que la numeración no es idéntica entre los serotipos, la inserción exacta del sitio no es crítica. En algunas realizaciones, la secuencia de direccionamiento se inserta dentro de los bucles expuestos en la superficie no estructurados (ver Figura 14) y pobremente conservados (ver alineación, FIG. 13). En algunas realizaciones, la inserción de la secuencia de direccionamiento puede lograrse dentro de otras cápsides de AAV insertando la secuencia de direccionamiento dentro de los bucles no estructurados homólogos de otras secuencias de AAV.

[0076] Un genoma de rAAV se describe en este documento. El genoma puede comprender al menos una repetición terminal invertida configurada para permitir el empaquetamiento en un vector y un gen de la cápside. En algunos casos, puede incluir además una secuencia dentro de un gen rep requerido para la expresión y el empalme del gen de la cápside. En algunos casos, el genoma puede incluir además una secuencia capaz de expresar VP3.

[0077] En algunos casos, la única proteína que se expresa es VP3 (la más pequeña de las proteínas estructurales de la cápside que compone la mayor parte de la cápside ensamblada - la cápside ensamblada se compone de 60 unidades de proteínas VP, ~ 50 de los cuales son VP3). En algunos casos, la expresión de VP3 sola es adecuada para permitir que el método de detección sea adecuado.

[0078] En algunos casos, consiste en colocar el elemento seleccionable, (que en algunos casos puede ser el gen de

la cápside de AAV en el genoma de AAV) junto con uno o más sitios de reconocimiento de recombinasa (loxP o de los sitios loxP mutante se prefieren el sistema para la detección, pero otros podrían ser utilizados). En algunos casos, el genoma de AAV puede definirse por un ácido nucleico que comprende al menos una repetición terminal invertida.

[0079] En algunos casos, el rAAV genoma comprende además un casete reportero mCherry que comprende un gen de ubiquitina C, un ADNc mCherry, y una secuencia de poliA sintético mínimo.

[0080] En algunos casos, el genoma comprende además el interruptor dependiente de CRE que comprende: una secuencia de poliA y un par de sitios loxP invertidas que flanquean la secuencia poliA. En algunos casos, la secuencia poliA se encuentra corriente abajo del gen de la cápside. En algunos casos, el par de sitios loxP invertidos comprende lox71 y lox66. En algunos casos, el genoma contiene solo aquellas secuencias dentro del gen rep requerido para la expresión y el empalme del producto del gen de la cápside.

[0081] En algunos casos, AAV-PHP.B entrega genes de manera eficiente a uno o más órganos, que incluyen, entre otros, el sistema nervioso central, el hígado, los músculos, el corazón, los pulmones, el estómago, la glándula suprarrenal, el tejido adiposo y el intestino.

[0082] Una biblioteca de la cápside se describe que comprende genomas de AAV que contienen tanto la secuencia completa rep y cap que han sido modificados con el fin de no impedir la replicación del virus en condiciones en las que normalmente podría replicar (co-infección de una célula mamífera junto con un virus auxiliar, como el adenovirus). Un genoma de pseudowt puede ser uno que tenga un gen de la cápside diseñado dentro de un genoma de AAV "wt".

[0083] En algunos casos, la biblioteca de la cápside se hace dentro de un genoma de AAV "tipo pseudo-salvaje" que contiene el gen viral de replicación (rep) y gen de la cápside (cap) flanqueado por repeticiones terminales invertidas (RTI). En algunos casos, la biblioteca de la cápside no se crea dentro de un genoma de AAV "pseudo-salvaje" que contiene el gen de replicación viral (rep) y el gen de la cápside (cap) flanqueado por repeticiones terminales invertidas (ITR).

[0084] En algunos casos, el genoma de rAAV contiene el gen de la cápside y sólo aquellas secuencias dentro del gen rep se requiere para la expresión y corte y empalme de los productos génicos tapa **(**FIG. 22B**)**.

[0085] Una secuencia de reconocimiento de recombinasa de gen de la cápside se da a conocer con repeticiones terminales invertidas que flanquean estas secuencias.

[0086] En algunos casos, el sistema podría usarse para desarrollar cápsides que presentan una mayor orientación de las células/órganos específicos, para seleccionar cápsides que evaden la inmunidad, seleccionar genomas que son más en la recombinación homóloga, seleccionar los elementos del genoma que aumentan la eficiencia de la conversión del genoma de AAV monocatenario a un genoma de ADN de doble cadena dentro de una célula y/o seleccionar elementos del genoma que aumentan la conversión del genoma de AAV a una forma circularizada y persistente dentro de la célula.

Secuencias de ácidos nucleicos

[0087] Se describe una secuencia de ácido nucleico que codifica cualquiera de las secuencias de direccionamiento descritas en el presente documento. En algunos casos, la secuencia de ácido nucleico es AAGTTTCCTGTGGCGTTGACT PARA SEQ ID NO 3). ACT TTG GCG GTG CCT TTT AAG (SEQ ID NO: 49) para una secuencia que codifica la secuencia AA de SEQ ID NO: 1. En algunos casos, la secuencia de ácido nucleico se hibridará con esta secuencia en condiciones rigurosas. En algunos casos, la secuencia de ácido nucleico incluye una secuencia de ácido nucleico que codifica la SEQ ID NOS: 1 y/o 3 y la secuencia es parte de una secuencia de ácido nucleico más grande. En algunos casos, una o más de las secuencias de la FIG. 31 puede proporcionar la secuencia de ácido nucleico anotada (es decir, se describe cualquier secuencia de ácido nucleico que codifique para cualquiera de estas secuencias). En ciertas ocasiones, la secuencia de ácido nucleico es una que se hibridará con cualquiera de las secuencias dentro de la FIG. 31 (o las secuencias que codifican las secuencias de aminoácidos) en condiciones rigurosas. En algunos casos, la secuencia de ácido nucleico incluye una secuencia de ácido nucleico que codifica para cualquiera de las secuencias dentro de la FIG. 31 y la secuencia es parte de una secuencia de ácido nucleico más grande. En algunos casos, la secuencia de ácido nucleico es una o más de: AGTGTGAGTAAGCCTTTTTTG (SEQ ID NO: 24); TTTACGTTGACGACGCCTAAG (SEQ ID NO: 26); o ATGAATGCTACGAAGAATGTG (SEQ ID NO: 27).

[0088] En algunos casos, la secuencia de ácido nucleico que codifica para la SEQ ID NO: 1 se inserta entre una secuencia que codifica para los aminoácidos 588 y 589 de AAV9 (SEQ ID NO: 2).

[0089] Se describe una secuencia de ácido nucleico que codifica cualquiera de los cuatro aminoácidos contiguos en TLAVPFK (SEQ ID NO: 1) o en KFPVALT (SEQ ID NO: 3). Se describe una secuencia de ácido nucleico que codifica cualesquiera cinco aminoácidos contiguos en TLAVPFK (SEQ ID NO: 1) o en KFPVALT (SEQ ID NO: 3). Se describe una secuencia de ácido nucleico que codifica cualquiera de los seis aminoácidos contiguos en TLAVPFK (SEQ ID NO: 1) o en KFPVALT (SEQ ID NO: 3) (o cualquiera de las otras proteínas de direccionamiento descritas en este

documento, incluidas las de la Figura 31).

[0090] En algunos casos, se inserta la secuencia de ácido nucleico entre una secuencia que codifica para los aminoácidos 588 y 589 de AAV9 (SEQ ID NO: 2).

[0091] En algunas realizaciones, se proporciona un sistema de plásmido. El plásmido puede incluir un primer plásmido que comprende un plásmido coadyuvante de rep-cap AAV2/9 modificado que comprende al menos uno en el codón de parada de marco dentro de su marco de lectura VP1, VP2 y VP3. El codón de parada se coloca para interrumpir la expresión de VP sin alterar la secuencia de aminoácidos de la proteína activadora del ensamblaje. El sistema de plásmido puede incluir además un segundo plásmido que comprende un plásmido rAAV-cap-in-cis.

[0092] En algunos casos, el método no implica la expresión de proteínas VP individuales a partir de plásmidos heterólogas para generar cápsides "mosaico" hechas a partir de proteínas VP codificadas por diferentes plásmidos.

[0093] Se describe una biblioteca de secuencias de ácido nucleico. La biblioteca puede comprender un elemento seleccionable y una o más secuencias de reconocimiento de recombinasa. En algunos casos, las secuencias de ácido nucleico y una o más secuencias de reconocimiento de recombinasa se incorporan dentro de un genoma de virus. En algunos casos, el genoma viral es un genoma de AAV. En algunos casos, el elemento seleccionable codifica una cápside de AAV. En algunos casos, el elemento seleccionable es un elemento genético que aumenta la conversión a dsDNA. En algunos casos, el elemento seleccionable aumenta la eficiencia de la recombinación homóloga entre el elemento y el genoma endógeno. En algunos casos, las secuencias de reconocimiento de recombinasa comprenden uno o más sitios loxP. En algunos casos, el sitio loxP es un sitio lox71 y un sitio lox66 invertido.

[0094] En algunas realizaciones, el gen que codifica la proteína de direccionamiento y/o la cápside se puede clonar en un plásmido auxiliar de Rep-Cap AAV en lugar del gen de la cápside existente. Cuando se introducen juntos en células productoras, este plásmido se puede usar para empaquetar un genoma de rAAV en la proteína y/o cápside de direccionamiento. Las células productoras pueden ser cualquier tipo de célula que posea los genes necesarios para promover la replicación del genoma de AAV, el ensamblaje de la cápside y el empaquetamiento. Las células productoras preferidas son células 293, o derivados, células HELA o células de insecto junto con un virus auxiliar o un segundo plásmido que codifica los genes del virus auxiliar que se sabe que promueven la replicación del genoma de rAAV. En algunas realizaciones, una secuencia auxiliar de rep-cap de AAV puede modificarse para introducir un sistema de expresión inducible por tetraciclina entre la rep y el gen de la cápside para aumentar la expresión de la cápside y la producción de virus. En algunas realizaciones, un cADN transactivador de tetraciclina, secuencia de poli adenilación, el elemento sensible a tetraciclina y el promotor p41 de AAV5 y los elementos reguladores de empalme de AAV2 contenidos dentro del gen rep de AAV2 se insertan entre el gen rep y el gen que codifica la cápside o la proteína de direccionamiento. El uso de este plásmido de rep-cap inducible al hacer rAAV proporciona 1.5-2 veces más virus que el plásmido de rep-cap de AAV2/9. Algunas realizaciones de dicho ácido nucleico clonado en un plásmido se representan en FIG. 21, SEQ ID NO: 10. La secuencia del gen de la cápside está subrayada en la FIG. 21**.** Las letras mayúsculas indican los sitios donde la secuencia de la cápside difiere de AAV9.

Métodos de uso

[0095] También se describe un método de entrega de una secuencia de ácido nucleico a un sistema nervioso (u otro sistema deseado). El método puede incluir proporcionar una proteína que comprende una o más de las secuencias de direccionamiento descritas en el presente documento. La proteína puede ser parte de una cápside de un AAV. El AAV puede comprender una secuencia de ácido nucleico para administrarse a un sistema nervioso. Entonces se puede administrar el AAV al sujeto.

[0096] En algunos casos, la secuencia de ácido nucleico que se administrará a un sistema nervioso comprende una o más secuencias que serían de alguna utilidad o beneficio para el sistema nervioso y/o el local de administración o el tejido o entorno circundante. En algunos casos, puede ser un ácido nucleico que codifica un factor trófico, un factor de crecimiento u otros factores solubles que podrían liberarse de las células transducidas y afectar la supervivencia o función de esa célula y/o las células circundantes. En algunos casos, puede ser un cADN que restaura la función de la proteína a humanos o animales que albergan una mutación genética en ese gen. En algunos casos, puede ser un cADN que codifica una proteína que se puede usar para controlar o alterar la actividad o el estado de una célula. En algunos casos, puede ser un cADN que codifica una proteína o un ácido nucleico utilizado para evaluar el estado de una célula. En algunos casos, puede ser un cADN y/o un ARN asociado para realizar ingeniería genómica. En algunos casos, puede ser una secuencia para la edición del genoma mediante recombinación homóloga. En algunos casos, puede ser una secuencia de ADN que codifica un ARN terapéutico. En algunos casos, puede ser un shARN o un sistema de entrega de miARN artificial. En algunos casos, puede ser una secuencia de ADN que influye en el empalme de un gen endógeno.

[0097] En algunos casos, las moléculas de direccionamiento resultantes se pueden emplear en métodos y/o terapias relacionadas con *in vivo* aplicaciones de transferencia de genes a las poblaciones de células de larga vida. En algunos casos, estos se pueden aplicar a cualquier terapia génica basada en rAAV, que incluye, por ejemplo, atrofia muscular espinal (AMS), esclerosis lateral amiotrófica (ELA), enfermedad de Parkinson, enfermedad de Pompe, enfermedad de

Huntington, enfermedad de Alzheimer, enfermedad de Battens, trastornos del almacenamiento lisosomal, glioblastoma multiforme, síndrome de Rett, amaurosis congénita de Leber, dolor crónico, accidente cerebrovascular, lesión de la médula espinal, lesión cerebral traumática y trastornos del almacenamiento lisosomal. Además, los rAAV también pueden emplearse para *in vivo* suministro de transgenes para estudios científicos no terapéuticos, como optogenética, sobreexpresión de genes, reducción de genes con shARN o miARNs, modulación de miARNs endógenos utilizando esponjas o señuelos de miARN, entrega de recombinasa para la eliminación condicional de genes, expresión condicional (dependiente de recombinasa) o edición de genes con CRISPRs, TALENs y nucleasas de dedos de zinc.

[0098] En el presente documento se describen métodos para tratar y/o prevenir la enfermedad de Huntington utilizando los métodos y composiciones descritos en este documento. El método para tratar y/o prevenir la enfermedad de Huntington puede incluir identificar al (los) sujeto(s), proporcionar un vector para el suministro de un polinucleótido al sistema nervioso del sujeto tal como se describe en el presente documento, administrar el vector en una dosis efectiva al sujeto, tratando de ese modo y/o prevenir la enfermedad de Huntington en el sujeto. En algunos casos, los métodos para tratar a un sujeto con enfermedad de Huntington involucran composiciones en las que el vector administra la composición polinucleotídica que comprende una proteína de dedo de cinc (ZFP) diseñada para reprimir la transcripción del gen Huntingtin (HTT). En ciertas ocasiones, la ZFP reprime selectivamente la transcripción del alelo del gen HTT responsable de causar la enfermedad de Huntington en el sujeto al unirse a la región de repetición CAG del gen HTT de una manera dependiente de la longitud de repetición CAG. En algunos casos, el ZNFTR reprime selectivamente la transcripción de ambos alelos del gen HTT.

[0099] En algunos casos, el elemento terapéutico que se administra al sujeto comprende un ARN de horquilla corta (ARNhc) o microARN (miARN) que derriba expresión Huntingtin mediante la inducción de la degradación selectiva de, o inhibición de la traducción de moléculas de ARN transcritas a partir de la enfermedad que causa el alelo HTT se une a la repetición CAG. En algunos casos, el ítem terapéutico que se administrará al sujeto comprende un ARN de horquilla corta (shARN) o microARN (miARN) que anula la expresión de Huntingtin al inducir la degradación o inhibición de la traducción de las moléculas de RNA transcritas de uno o ambos alelos. del gen HTT. En algunos casos, el ítem terapéutico que se administrará al sujeto comprende un ARN de horquilla corta (shARN) o un microARN (miARN) que anula la expresión de Huntingtin induciendo la degradación selectiva de, o inhibiendo la traducción de, las moléculas de ARN transcritas de la enfermedad que causa el alelo HTT a través del reconocimiento selectivo de uno o más polimorfismos de nucleótidos presentes dentro de la enfermedad que causa el alelo. Los polimorfismos de nucleótidos pueden ser utilizados por un experto en la técnica para diferenciar entre el alelo normal y el que causa la enfermedad.

[0100] En algunos casos, el artículo terapéutico que se va a administrar al sujeto comprende un polinucleótido que codifica un ARN o proteína que altera el corte y la producción del ARN de HTT. En algunos casos, el ítem terapéutico que se administrará al sujeto comprende un polinucleótido que codifica uno o más polipéptidos y/o ARN para la edición del genoma usando una nucleasa efectora similar a un activador de la transcripción (TALEN), nucleasa de dedos de cinc o agrupada regularmente repeticiones: el sistema del gen cas9 (CRISPR/Cap9) diseñado por un experto en la técnica para inducir una ruptura de ADN o ADN de doble hebra dentro o adyacente al gen HTT para causar una alteración en la secuencia del gen HTT. En ciertas ocasiones, el elemento terapéutico que se administrará al sujeto comprende un polinucleótido que codifica un polipéptido que se une a un polipéptido del gen HTT, altera la conformación de un polipéptido del gen HTT o altera el ensamblaje de un polipéptido del gen HTT en agregados o altera la vida media de un polipéptido del gen HTT. En algunos casos, el artículo terapéutico que se administrará al sujeto comprende un polinucleótido que codifica un ARN o polipéptido que causa o evita una modificación post-transcripcional de un polipéptido del gen HTT. En algunos casos, el artículo terapéutico que se va a administrar al sujeto comprende un polinucleótido que codifica un polipéptido de una proteína chaperona conocida por los expertos en la técnica para influir en la conformación y/o estabilidad de un polipéptido del gen HTT.

[0101] En algunos casos, el elemento terapéutico que se administra al sujeto comprende elementos reguladores conocidos por un experto en la técnica para influir en la expresión del ARN y/o productos de la proteína codificada por el polinucleótido dentro de las células deseadas del sujeto.

[0102] En algunos casos, el elemento terapéutico que se administra al sujeto comprende un elemento terapéutico aplicable para cualquier enfermedad o trastorno de elección. En algunos casos, esto puede incluir composiciones para tratar y/o prevenir la enfermedad de Alzheimer usando los métodos y composiciones descritos en este documento, por ejemplo, ApoE2 o ApoE3 para la enfermedad de Alzheimer; SMN para el tratamiento de SMA; administración de frataxina para el tratamiento de la ataxia de Friedreich; y/o shARN o miARN para el tratamiento de la ELA.

[0103] En algunos casos, el elemento terapéutico para la entrega es una proteína (codifica una proteína) o una estrategia basada en ARN para reducir la agregación de sinucleína para el tratamiento de Parkinson. Por ejemplo, la entrega de un polinucleótido que codifica una variante de sinucleína que es resistente a la agregación y, por lo tanto, interrumpe la agregación de la sinucleína endógena.

[0104] En algunos casos, un transgén que codifica un factor trófico para el tratamiento de AD, PD, ALS, SMA, HD puede ser el elemento terapéutico en cuestión. En algunos casos, se puede emplear un factor trófico y puede incluir, por ejemplo, BDNF, GDNF, NGF, LIF y/o CNTF.

[0105] Las dosis de un vector viral pueden depender principalmente de factores tales como la afección que se trata, la edad, el peso y la salud del paciente, y por lo tanto pueden variar entre los pacientes. Por ejemplo, una dosis humana terapéuticamente efectiva del vector viral está generalmente en el intervalo de aproximadamente 0,1 ml a aproximadamente 100 ml de solución que contiene concentraciones de aproximadamente 1 X 10^9 a 1 X 10^{16} genomas de vector virus. Una dosis humana preferida puede ser de aproximadamente 1 X 10^{13} a 1 X 10^{16} genomas de AAV. La dosis se ajustará para equilibrar el beneficio terapéutico frente a cualquier efecto secundario y dichas dosis pueden variar dependiendo de la aplicación terapéutica para la cual se emplea el vector recombinante. Los niveles de expresión del transgén pueden monitorizarse para determinar la frecuencia de dosificación resultante del vector de la invención.

[0106] En algunos casos, el vector de polinucleótidos también incluye elementos de control reguladores conocidos para un experto en la técnica para influir en la expresión del ARN y/o productos de la proteína codificada por el polinucleótido dentro de las células deseadas del sujeto.

[0107] En algunos casos, funcionalmente, la expresión del polinucleótido es al menos en parte controlable por los elementos reguladores unidos de manera operativa, de tal manera que el elemento(s) modula la transcripción del polinucleótido, el transporte, el procesamiento y la estabilidad del ARN codificado por el polinucleótido y, como Apropiado, traducción de la transcripción. Un ejemplo específico de un elemento de control de expresión es un promotor, que generalmente se ubica en 5' de la secuencia transcrita. Otro ejemplo de un elemento de control de expresión es un potenciador, que se puede ubicar en 5' o 3' de la secuencia transcrita, o dentro de la secuencia transcrita. Otro ejemplo de un elemento regulador es una secuencia de reconocimiento para un microARN.Otro ejemplo de un elemento regulador es un intrón y el donante de empalme y las secuencias aceptadoras de empalme que regulan el empalme de dicho intrón. Otro ejemplo de un elemento regulador es una señal de terminación de la transcripción y/o una secuencia de poliadenilación.

[0108] Elementos de control de expresión y promotores incluyen aquellos activos en un tejido particular o tipo de célula, se hace referencia en este documento como un "elementos/promotores de control de expresión específicos de tejido". Los elementos de control de la expresión específicos del tejido suelen estar activos en células o tejidos específicos (por ejemplo, en el hígado, el cerebro, el sistema nervioso central, la médula espinal, el ojo, la retina o el pulmón). Los elementos de control de la expresión son típicamente activos en estas células, tejidos u órganos porque son reconocidos por proteínas activadoras de la transcripción u otros reguladores de la transcripción, que son exclusivos de una célula, tejido u tipo de órgano específico.

[0109] Elementos de control de la expresión también incluyen promotores ubicuos o promiscuos/potenciadores que son capaces de dirigir la expresión de un polinucleótido en muchos tipos celulares diferentes. Dichos elementos incluyen, pero no se limitan a, las secuencias de promotor/intensificador temprano inmediato de citomegalovirus (CMV), las secuencias de promotor/promotor del virus del sarcoma de Rous (RSV) y los otros promotores/potenciadores virales activos en una variedad de tipos de células de mamíferos; secuencias promotoras/enhancer de genes de mamíferos expresados de forma ubicua o promiscua, que incluyen, entre otros, beta actina, ubiquitina o EF1 alfa; o elementos sintéticos que no están presentes en la naturaleza.

[0110] Los elementos de control de la expresión también pueden conferir la expresión de una manera que sea regulable, es decir, una señal o estímulos aumentan o disminuyen la expresión del polinucleótido unido operativamente. Un elemento regulable que aumenta la expresión del polinucleótido unido operativamente en respuesta a una señal o estímulos también se conoce como un "elemento inducible" (es decir, es inducido por una señal). Los ejemplos particulares incluyen, pero no se limitan a, un promotor inducible por hormonas (por ejemplo, esteroides). Un elemento regulable que disminuye la expresión del polinucleótido unido operativamente en respuesta a una señal o estímulos se conoce como un "elemento reprimible" (es decir, la señal disminuye la expresión de manera que cuando la señal se elimina o está ausente, la expresión aumenta). Típicamente,la cantidad de aumento o disminución conferida por tales elementos es proporcional a la cantidad de señal o estímulos presentes; cuanto mayor sea la cantidad de señal o estímulos, mayor será el aumento o la disminución de la expresión.

[0111] Cualquier uno o más de los aspectos anteriores se puede incluir dentro de cualquiera de los vectores descritos en este documento, en combinación con cualquier proteína de direccionamiento.

Método de selección

[0112] Los actuales protocolos de evolución dirigida utilizados para mejorar las cápsides de AAV tienen varios inconvenientes. El primero es que es difícil diseñar un cribado in vivo que recupere específicamente secuencias de la célula diana de interés cuando esa célula diana es uno de los muchos tipos de células en un órgano complejo. Típicamente, después de que el virus se administra in vivo, el tejido de interés se recolecta y el ADN del virus se recupera del ADN de todo el tejido o región del tejido. Recientemente, Dalkara et al. informaron el uso de FACS para clasificar las células diana (fotorreceptores) como un medio para recuperar selectivamente las secuencias de la cápside presentes en esas células. Pero este método es laborioso y costoso, especialmente para la clasificación de grandes volúmenes de tejidos disociados. Adicionalmente, este esfuerzo de clasificación adicional no supera la otra limitación importante de la selección de las secuencias de la cápside del AAV: todas las secuencias de la cápside

presentes dentro de la célula/tejido se recuperan independientemente de si estos virus transducen funcionalmente alguna célula o no. En otras palabras, las secuencias de virus atrapados en la superficie celular o los virus que ingresaron a la célula unida a un receptor que traficaba a un compartimento intracelular no compatible con el desempaquetado y la transducción de AAV se recuperan en estos cribados junto con las secuencias que codificaron cápsides que exitosamente indujeron expresión transgénica en la población de células diana. Por lo tanto, las secuencias de virus atrapadas en la superficie celular o los virus que ingresaron a la célula unida a un receptor que traficaba a un compartimento intracelular no compatible con el desempaquetado y la transducción de AAV se recuperan en estos cribados junto con las secuencias que codificaron cápsides que indujeron con éxito la expresión transgénica en la población de células diana. Por lo tanto, las secuencias de virus atrapadas en la superficie celular o los virus que ingresaron a la célula unida a un receptor que traficaba a un compartimento intracelular no compatible con el desempaquetado y la transducción de AAV se recuperan en estos cribados junto con las secuencias que codificaron cápsides que indujeron con éxito la expresión transgénica en la población de células diana. Por lo tanto, las cápsides no funcionales también se enriquecen con los métodos de selección típicos. Finalmente, la mayoría de los métodos actuales también requieren el uso de bibliotecas hechas de AAV competentes para la replicación, lo cual es un problema potencial de bioseguridad, especialmente si el virus se introducirá en instalaciones de animales donde hay primates, ya que estos virus podrían replicarse en animales portadores de virus auxiliares. Aquí se describe una plataforma de selección de bibliotecas de cápside AAV que supera una o más de cada una de estas limitaciones.

[0113] La producción exitosa de una biblioteca variante de cápside AAV depende del empaquetado de cada gen variante de cápside por las proteínas de la cápside particulares que codifica. Por lo tanto, es útil que el gen de la cápside esté presente en cis (dentro del genoma de AAV). Sin embargo, no es esencial que los genes rep no estructurales estén presentes en cis. En este documento se describe un genoma rAAV incompetente para la replicación que expresa el gen de la cápside y en lugar de gran parte de la secuencia rep, se han agregado varios elementos recombinantes que proporcionan una manera de recuperar selectivamente solo aquellas secuencias de la cápside que han transducido funcionalmente la población de células diana de interés sin la necesidad de aislamiento de la célula diana.

Recuperación selectiva de secuencias de cápside AA V de poblaciones específicas de células Cre+

[0114] En algunos casos, el enfoque incorpora un interruptor de recombinasa Cre-dependiente que utiliza PCR (reacción en cadena de la polimerasa) para recuperar selectivamente secuencias de la cápside que han transducido células Cre+ diana. Esto se puede lograr insertando sitios de reconocimiento de Cre mutantes (lox66 y lox71 en una orientación cabeza a cabeza) en el genoma de rAAV alrededor de una secuencia adyacente al gen de la cápside **(FIG. 1B**). La recombinación de Cre da como resultado una inversión de la secuencia flanqueada por los sitios lox66 y lox71, y luego se puede usar una estrategia de recuperación de PCR que solo amplifica la secuencia del gen de la cápside de los genomas de rAAV después de la inversión mediada por Cre de la secuencia adyacente del gen de la cápside (una PCR cebador se une a la secuencia invertible y uno se une al gen de la cápside). Los sitios loxP mutantes (lox66 y lox71) se pueden elegir de manera que la inversión sea menos reversible (Alberts et al. 1995). **FIG. 1B** muestra un plásmido rAAV que se ha desarrollado.

[0115] En algunos casos, el método se aprovecha de la gran cantidad de ratones transgénicos Cre que han sido (y pueden ser) desarrollados. Estas líneas expresan Cre bajo el control de promotores específicos de células, de modo que Cre está presente solo en una subpoblación de células dentro de un órgano determinado. Cientos de líneas transgénicas Cre están disponibles en proveedores comerciales y fuentes académicas, y se pueden generar líneas personalizadas.

[0116] En algunos casos, se puede aplicar esto para desarrollar cápsides que transducen de manera más eficiente los astrocitos en el sistema nervioso central después de la administración del virus IV. Hay ratones transgénicos mGFAP-Cre que expresan Cre específicamente en astrocitos y células madre neurales (NSC) en el cerebro adulto y la médula espinal. Usando la estrategia de recuperación de secuencia dependiente de Cre, se puede administrar bibliotecas de cápside de rAAV in vivo, recolectar ADN de todo el cerebro y la médula espinal y recuperar secuencias de cápside de rAAV específicamente de astrocitos y NSC.

[0117] Otra ventaja de este enfoque es que esta estrategia dependiente de Cre sólo recupera aquellas secuencias que han *transducido* la célula diana. El AAV es un virus de ADN monocatenario y su genoma debe ingresar al núcleo y convertirse en ADN bicatenario (dsADN) para la transducción funcional. Dado que Cre solo recombina dsDNA, solo se recuperarán las secuencias de cápside que se han traficado adecuadamente al núcleo celular y se han convertido en dsDNA.

Inclusión de un casete de gen reportero para facilitar la clasificación celular

[0118] Para los casos en los que los transgénicos Cre+ no están disponibles, también se puede incorporar un casete informador impulsado por un promotor/potenciador ubicuo para facilitar la clasificación de las células que expresan transductor/transgén de una población mixta. Esta segunda opción requiere más mano de obra que la estrategia basada en Cre, ya que requiere la generación de suspensiones de células individuales y FACS o clasificación basada en perlas magnéticas/anticuerpos. Pero el método de reportero también es poderoso, ya que se puede combinar con

la clasificación de poblaciones específicas de células diana utilizando anticuerpos contra marcadores de superficie conocidos o con transgénicos GFP para limitar la recuperación a una población particular. Y al igual que la estrategia de Cre anterior, solo conducirá a la recuperación de secuencias que están presentes en células *transducidas*. El informador también facilita el seguimiento de las características de transducción de la biblioteca agrupada durante la selección (útil tanto para los métodos dependientes de Cre como para los informadores).

[0119] La tecnología descrita en el presente documento se puede utilizar junto con cualquier línea transgénica que exprese Cre en el tipo de célula diana de interés para desarrollar cápsides de AAV que transduzcan de manera más eficiente que la población de células diana. Las aplicaciones incluyen, pero no se limitan al desarrollo de cápsides que son más eficientes en la transducción de tipos de células específicos en cualquier órgano después de la administración de AAV IV, dirigidas a poblaciones específicas de neuronas, mejorando el transporte internacional, dirigiendo células tumorales, células madre hematopoyéticas, células beta productoras de insulina , epitelio pulmonar, etc. El método no se limita a ningún método de administración de virus. El vector puede administrarse por cualquier vía, que incluye, entre otras, administraciones oral, intravenosa, intraarticular, intracardíaco, intramuscular, intradérmica, tópica, intranasal, intraparitoneal, rectal, sublingual, subcutánea, epidural, intracerebral, intracerebroventricular, intratecal, intravítrea o subretiniana. El sistema también se puede usar para desarrollar virus que atraviesen mejor las barreras específicas (barrera hematoencefálica, epitelio intestinal, placenta, etc.). El método también se puede usar in vitro para desarrollar cápsides que son mejores para lograr la entrada nuclear y la síntesis de la segunda cadena (conversión a dsDNA).

[0120] Además, este sistema no está limitado a AAV9. Cualquier cápside de AAV inicial (variantes naturales o modificadas) se puede incorporar a este vector de rAAV-cap, mutagenizado por métodos estándar para crear la biblioteca de cápside y luego cribarse con esta estrategia de recuperación dependiente de Cre. Las cápsides de AAV preferidas incluyen AAV1, AAV2, AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, aislados humanos (por ejemplo, hu.31 y hu.32), aislamientos de rhesus (por ejemplo, rh.8 y rh.10), AAV o parvovirus relacionados de otras especies de primates, mamíferos y no mamíferos. Además, este método no se limita a ninguna estrategia de mutagénesis de cápside comúnmente utilizada. Se puede utilizar cualquier método para generar la diversidad de la biblioteca, que incluye, entre otros, la combinación aleatoria de dominios de la cápside, la inserción de secuencias aleatorias y la mutagénesis aleatoria por PCR propensa a errores. Finalmente, el vector ha sido diseñado para ser modular, lo que simplifica la sustitución de diversos elementos, como el casete informador o la secuencia de cápside, para personalizar aún más las opciones de detección.

[0121] Para hacer posible este sistema, se ha desarrollado una opción para la incorporación del gen de la cápside de AAV en el genoma de rAAV mientras que proporciona los otros productos génicos de AAV no estructural en trans a partir de un plásmido auxiliar. Esto no fue necesariamente sencillo ya que el gen de la cápside codifica 4 proteínas (las proteínas de cápside VP1, VP2, VP3 y la proteína activadora de ensamblaje (AAP)) mediante una combinación de empalme alternativo, codones de inicio alternativos y marcos de lectura alternativos. Mantener una regulación adecuada de la expresión de estas proteínas es relevante para la producción eficiente de virus. Se encontró que las proteínas de la cápside podrían expresarse a partir de un genoma de rAAV cuando se incluye un fragmento del extremo 3' del gen rep, que contiene promotor/potenciador crítico y señales de empalme **(FIG. 2A)**. Al retener solo el extremo 3' del gen rep junto con el gen de la cápside, queda suficiente espacio dentro del genoma de rAAV para incorporar la secuencia poliA invertible de Cre corriente abajo del gen de la cápside, así como un casete informador mCherry (proteína roja fluorescente).

[0122] Para asegurarse de que la cápside está hecha completamente del gen de la cápside codificado dentro del genoma de la biblioteca de la cápside del rAAA, un plásmido auxiliar de AAV que proporcionaría las proteínas no estructurales del AAV pero no cualquier expresión de la proteína de la cápside (típicamente el rAAV se produce suministrando ambos se desarrollaron los genes rep y cap en trans del plásmido auxiliar AAV). Utilizando un núcleo de vector de plásmido RepCap AAV2/9 como punto de partida, se insertaron 5 codones de parada dentro del gen de la cápside cerca de los sitios de inicio de la traducción para las tres proteínas de la cápside VP1, VP2 y VP3 **(FIG. 2B)**. Esto eliminó de manera efectiva la producción de rAAV a menos que las proteínas de la cápside VP1-3 se proporcionen en otro lugar (como en cis en la construcción de la biblioteca de cápside de genoma de AAV descrita anteriormente).

[0123] Un método de desarrollo de una cápside con una característica deseada se da a conocer en el presente documento. El método puede comprender proporcionar una población de genomas de rAAV descritos en el presente documento. El método puede implicar además la selección de la población por un conjunto específico de criterios. El método puede implicar además la selección del genoma rAW que cumpla con los criterios de selección.

[0124] En algunos casos, algunos de los métodos de detección descritos en este documento proporcionan al menos una de las siguientes ventajas. Primero, en algunos casos, el método hace uso de la creciente biblioteca de Cre-transgénicos para proporcionar presión selectiva para las cápsides que transducen de manera más eficiente las poblaciones de células genéticamente definidas (por ejemplo, ver cre.jax, gensat, creportal, conectividad.map cerebral (MGI) (todo ".org"). Segundo, en algunos casos, ya que Cre solo recombina el ADN de doble cadena (dsADN) y el genoma de AAV es de cadena sencilla, solo las secuencias de la cápside que median el tráfico intracelular adecuado y se recuperará la conversión del genoma empaquetado a una forma de ADNds persistente, por lo que este enfoque puede proporcionar una presión selectiva adicional para las cápsides funcionales.

[0125] Como se representa en la **FIG. 2C,** se pueden seleccionar variantes de secuencia adicionales en función de su capacidad para mediar una asociación incrementada del ácido nucleico que lleva la biblioteca de secuencias y el sitio de recombinasa con la recombinasa para aplicaciones deseadas similares o diferentes. En algunos casos, el proceso puede comenzar generando una biblioteca de variantes de secuencia de ADN (1A). En algunos casos, esto puede incluir 10 ^ 2 si no más secuencias (por ejemplo, 10 ^ 6 o más). Dentro del mismo ácido nucleico, también se puede incorporar una o más secuencias de reconocimiento de recombinasa (1B). Luego se diseña una estrategia para la recuperación/amplificación de secuencia dependiente de la recombinación de las variantes de secuencia. (1B). Esto puede implicar uno o más sitios de reconocimiento de recombinasa. Luego se puede combinar la biblioteca y los sitios de recombinación para transferir fragmentos de ADN de la biblioteca a un vector con sitios de reconocimiento de recombinasa (2). Luego se puede entregar la biblioteca (por ejemplo, in vivo y/o in vitro). La expresión de recombinasa (REC+) está restringida a una o más poblaciones de células diana o componentes (3). Luego, se puede aplicar una presión selectiva seleccionada al sistema de tal manera que se puedan recuperar/amplificar secuencias basadas en la presencia o ausencia de los eventos de recombinación mediados por recombinasa en los ácidos nucleicos que comprenden las variantes de la biblioteca (4). Este proceso se puede repetir si es necesario, transfiriendo las variantes seleccionadas recuperadas o amplificadas de nuevo al vector aceptador de la biblioteca (5B) para 1 a 5, o más rondas de selección. Luego se pueden obtener y caracterizar las secuencias variantes (5A) mediante varios métodos, como la secuenciación de Sanger o la secuenciación de la próxima generación. Finalmente, s puede caracterizar la función de cualquiera o todas las variantes individuales.

[0126] Un método para recuperar selectivamente secuencias de cápside que tienen poblaciones de células diana específicas transducidas dentro de muestras de tejido complejas. El genoma de AAV tiene dos genes-rep, que codifican 4 proteínas no estructurales relevantes para la replicación (rep78, rep68, rep52 y rep40) y cápside, que codifica tres proteínas (VP1, VP2 y VP3) que forman la cáscara, o cápside, del virus **(FIG. 2D**). Además, el gen de la cápside también codifica una proteína accesoria activadora del ensamblaje (AAP) que se requiere para el ensamblaje de la cápside. Los métodos de evolución dirigida de la cápside utilizan el AAV competente para la replicación, de modo que el gen de la cápside está presente en cis (es decir, dentro del genoma viral). Sin embargo, la producción exitosa de una biblioteca de variantes de cápside de AAV depende solo de cada gen de cápside de variante que sea empaquetado por las proteínas de cápside particulares que codifica. Por lo tanto, aunque es útil que el gen de la cápside esté presente en cis, no es esencial que los genes de replicación (rep) no estructurales estén presentes en cis. Teniendo esto en cuenta, un genoma de rAAV incompetente con la replicación que expresa el gen de la cápside y solo las regiones del gen rep necesarias para la expresión y el empalme del gen de la cápside **(FIG. 2D**) ha sido desarrollado. En lugar de las secuencias de repetición restantes, se han incorporado varios elementos recombinantes que proporcionan un medio para recuperar selectivamente solo aquellas secuencias de cápside que han transducido funcionalmente la población de células objetivo de interés sin la necesidad de aislamiento de células diana **(FIG. 3A)**. Para garantizar que las proteínas codificadas por la tapa se expresen correctamente, se incorporaron las secuencias de donante y aceptor de empalme (y todas las secuencias intermedias) presentes dentro del gen rep AAV2 aguas arriba del gen de la cápside AAV dentro del genoma recombinante **(FIG. 2E)**. Se usó un fragmento del promotor p41 de AAV5 para dirigir la traducción del gen de la cápside (SEQ ID NO: 4, **FIG**. **15,** representando el plásmido completo: la espina dorsal del plásmido, ITR de AAV, UBC-mCherry-syn-pA, promotor p41 de repetición de empalme AAV5-AAV2, cápside de AAV9, lox71-SV40 polyA-lox66-ITR). Para proporcionar la función auxiliar y la función AAP, se modificó un plásmido auxiliar AAV rep/cap insertando un total de 5 codones de parada dentro del gen de la cápside dentro del marco de lectura VP1, 2 y 3 (1 codón de parada interrumpe VP3, 3 interrumpen VP2 y todo 5 interrumpe VP1 – **FIG. 2F,** SEQ ID NO: 5, **FIG**. **16)**. Estos codones de parada se diseñaron de manera que no interrumpieran la secuencia de codificación de la proteína AAP, que se codifica dentro de un marco de lectura alternativo.

[0127] Recuperación selectiva de secuencias de cápside de AAV de poblaciones de células Cre+ específicas. Para facilitar la recuperación selectiva de solo aquellas secuencias de la cápside que codifican la proteína de la cápside que median la transducción de una población específica de células diana, se diseñó un sistema para aprovechar la gran cantidad de ratones transgénicos Cre, ratas u otros organismos transgénicos Cre que han sido (y pueden ser) desarrollados. Estas líneas expresan Cre bajo el control de promotores específicos de células, de modo que Cre está presente solo en una subpoblación de células dentro de un órgano determinado. Este enfoque incorpora un "interruptor" dependiente de la recombinasa Cre que proporciona un medio para recuperar selectivamente las secuencias de la cápside que han transvertido las células diana Cre+. La recombinación de Cre da como resultado una inversión o eliminación (según la configuración de los sitios lox utilizados - ver **FIGs. 3** y **4)** de una secuencia dentro del genoma de AAV, y se utilizó una estrategia de recuperación basada en PCR que solo amplifica la secuencia del gen de la cápside de los genomas de rAAV que se han sometido a un evento de inversión mediada por Cre. Esta estrategia también puede adaptarse para seleccionar cápsides de AAV que se dirigen a células que previamente se habían hecho para expresar Cre a través de medios no transgénicos (por ejemplo, transducción previa con un virus que expresa Cre), que podría ser útil para la selección en especies más grandes donde transgenes Cre no están disponibles.

[0128] Una ventaja de algunos de estos casos es que esta estrategia recombinasa dependiente sólo recupera aquellas secuencias que han transducido la célula diana. AAV es un virus de ADN de una sola hebra y su genoma debe convertirse a ADN de doble cadena (dsADN) para transducción funcional. Ya que Cre solo recombina dsDNA, solo se recuperarán las secuencias de cápside que se han traficado correctamente y se han convertido a dsDNA. Esto

aumenta la presión selectiva aplicada, lo que anticipamos que reducirá el número de rondas de selección que son necesarias para desarrollar virus con propiedades mejoradas.

[0129] Esta aplicación no se limita a la utilización de Cre-lox como un sistema de sitio recombinasa/diana. Otras instancias pueden incluir recombinasas/integrasas que incluyen, por ejemplo, Flp, phiC31 o Bxb1. El método también se puede adaptar para su uso con estrategias de recuperación dependientes de recombinación, no basadas en PCR. Además, se usó un genoma de AAV recombinante que carece de la mayoría de las secuencias rep para proporcionar espacio para el interruptor lox y un casete informador, un interruptor dependiente de cre podría insertarse alternativamente dentro de un genoma de AAV "competente para la replicación" de tal manera que no alteró la expresión génica del virus y el empaquetamiento.

[0130] Los esfuerzos recientes para usar rAAV como un vehículo para la terapia génica son prometedores para su aplicabilidad como tratamiento para enfermedades humanas basadas en defectos genéticos. Los vectores de rAAV proporcionan la expresión a largo plazo de genes introducidos de un genoma episomal, aunque se ha observado la integración del genoma de rAAV en los cromosomas del huésped (Kaeppel 2013). Una ventaja adicional de rAAV es su capacidad para realizar esta función en tipos de células tanto de división como de no división, incluidos hepatocitos, neuronas y miocitos esqueléticos. rAAV se ha utilizado con éxito como vehículo de terapia génica para permitir la expresión de eritropoyetina en el músculo esquelético de ratones (Kessler et al., 1996), tirosina hidroxilasa y aminoácido aromático descarboxilasa en el SNC en modelos de mono de la enfermedad de Parkinson (Kaplitt et al., 1994) y Factor IX en músculo esquelético e hígado en modelos animales de hemofilia. A nivel clínico, el vector rAAV se ha utilizado en ensayos clínicos en humanos para administrar el gen cftr a pacientes con fibrosis quística, el gen Factor IX a pacientes con hemofilia (Flotte, et al., 1998, Wagner et al, 1998).

[0131] AAV recombinante es producido in vitro mediante la introducción de construcciones de genes en las células conocidas como células productoras. Algunos sistemas para la producción de rAAV emplean tres elementos fundamentales: 1) un casete de genes que contiene el gen de interés, 2) un casete de genes que contiene los genes rep y cap de AAV y 3) una fuente de genes de virus "auxiliares".

[0132] El primer casete de genes se construye con el gen de interés flanqueado por repeticiones terminales invertidas (ITR) de AAV. Los ITR funcionan para dirigir el empaquetamiento del gen de interés en el virión AAV. El segundo casete de genes contiene repeticiones y cap, genes AAV que codifican proteínas necesarias para la replicación y el empaquetamiento de rAAV. El gen rep codifica cuatro proteínas (Rep 78, 68, 52 y 40) necesarias para la replicación del ADN. Los genes de cápside codifican tres proteínas estructurales (VP1, VP2 y VP3) que forman la cápside del virus.

[0133] El tercer elemento es relevante porque AAV-2 no se replica. Las funciones de ayuda son productos de proteínas de virus de ADN de ayuda que crean un entorno celular propicio para la replicación y el empaquetado eficientes de rAAV. El adenovirus (Ad) se ha utilizado casi exclusivamente para proporcionar funciones de ayuda para rAAV. Los productos genéticos proporcionados por Ad están codificados por los genes E1a, E1b, E2a, E4orf6 y Va.

[0134] La producción de vectores rAAV para terapia génica puede llevarse a cabo in vitro, utilizando líneas celulares productoras adecuadas, tales como 293 y HeLa. Una estrategia para entregar todos los elementos requeridos para la producción de rAAV utiliza dos plásmidos y un virus auxiliar. Este método se basa en la transfección de las células productoras con plásmidos que contienen casetes de genes que codifican los productos genéticos necesarios, así como la infección de las células con Ad para proporcionar las funciones auxiliares. Este sistema emplea plásmidos con dos casetes de genes diferentes. El primero es un plásmido proviral que codifica el ADN recombinante para ser empaquetado como rAAV. El segundo es un plásmido que codifica los genes rep y cap. Para introducir estos diversos elementos en las células, las células se infectan con Ad y se transfectan con los dos plásmidos. Alternativamente, en protocolos más recientes, el paso de la infección por Ad se puede reemplazar por transfección con un "plásmido auxiliar" de adenovirus que contiene los genes VA, E2A y E4. Como se describe en este documento, las disposiciones de rep y cap pueden estar en trans para los aspectos de selección.

[0135] Mientras Ad se ha utilizado convencionalmente como el virus auxiliar para la producción de rAAV, se sabe que otros virus de ADN, tales como virus del herpes simple tipo 1 (HSV-1) se pueden utilizar también. Se identificó el conjunto mínimo de genes de HSV-1 necesarios para la replicación y el empaquetamiento de AAV-2, e incluye los primeros genes UL5, UL8, UL52 y UL29. Estos genes codifican los componentes de la maquinaria de replicación central del VHS-1, es decir, la helicasa, la primasa, las proteínas accesorias de la primasa y la proteína de unión al ADN monocatenario. Esta propiedad auxiliar de rAAV de HSV-1 se ha utilizado en el diseño y construcción de un vector de virus de Herpes recombinante capaz de proporcionar productos de genes de virus auxiliares necesarios para la producción de rAAV.

[0136] Los siguientes ejemplos se presentan como realizaciones de ejemplo solamente, y no pretenden ser limitantes en el alcance de las reivindicaciones. Además, hay secciones adicionales de varias realizaciones proporcionadas entre algunos de los ejemplos a continuación, según corresponda, y según lo indicado por el texto y el espaciado del documento.

EJEMPLO 1

[0137] Las construcciones de rep-AAP y rAAV-cap-in-cis divididas generan un rAAV de alto título. Para ensayar si el sistema auxiliar rep-AAP y el sistema rAAV-cap-in-cis generan virus rAAV, se realizó una transfección triple de células 293T (ATCC) con el ayudante rep-AAP, genoma rAAV mCherry-cap-lox71/66 y el constructo ayudante adenoviral pHelper. Los plásmidos se transfectaron en una proporción de 2:1:4 (0,263 µg de ADN total/cm2 de área de superficie celular plaqueada), respectivamente, usando polietilenimina lineal (PEI) como reactivo de transfección con una relación N:P de 25. Con estos constructos, se pudo generar virus recombinante con una eficiencia equivalente a la observada con el ayudante AAV2/9 rep/cap estándar, un genoma rAAV2 que expresa mCherry solo y pHelper **(FIG. 5A)**. En contraste, cuando el ayudante rep-AAP y el pHelper se usaron junto con un genoma de rAAV que codifica mCherry, pero no un límite de AAV, se generó poco o ningún virus. Esto confirma que la expresión de la cápside en cis era necesaria para la producción de rAAV.

[0138] Generación de las bibliotecas de cápside: introducción de secuencias aleatorias cortas en bucles de superficie de AAV9. Se pueden usar varias estrategias para introducir la diversidad de secuencias en el gen de la cápside. Los ejemplos incluyen, pero no se limitan a (i) PCR propensa a errores, que introduce mutaciones a una velocidad controlable en una región del gen de la cápside amplificada por PCR, (ii) barajado de dominio de cápside, donde se generan bibliotecas a través de eventos de recombinación entre secuencias fragmentadas de cápside generadas a partir de un panel de diferentes serotipos de cápside y (iii) modificación de secuencia dirigida en sitios específicos utilizando cebadores con bases mixtas, que generan extensiones de secuencias aleatorias en sitios específicos dentro de la cápside. Cada uno de estos métodos tiene ventajas y desventajas. En algunos casos, se puede usar una estrategia de modificación de secuencia dirigida para reemplazar o insertar secuencias aleatorias de siete aminoácidos (21 nucleótidos) en dos bucles de superficie diferentes.

EJEMPLO 2

[0139] Para generar bibliotecas de variantes de la cápside de AAV, siete aminoácidos de secuencia aleatorizada se introdujeron en la cápside AAV9. En una biblioteca, (452-8r) AA452-8, (conteo de VP1) se reemplazó por secuencia aleatoria. En una segunda biblioteca, (588i) se insertaron siete AA de secuencia aleatoria después de AA588 en la cápside de AAV9. Usando esta estrategia de aleatorización dirigida, los sitios pueden ser aleatorizados juntos en la misma biblioteca o aleatorizados secuencialmente después de la selección en un sitio individual.

[0140] Los fragmentos de la biblioteca se generaron por PCR. AA452-458 de AAV9 se reemplazaron con 7 aminoácidos aleatorios mediante el uso de un cebador que contiene un tramo de 21 bases mezcladas a mano (7 x NNK, Cebador 1287). El cebador 1312 se usó como un cebador inverso. Para la biblioteca 588i, se insertó un tramo de 7AA después de AA588 utilizando un cebador que contenía un tramo de 21 bases mezcladas a mano (7 x MNN, cebador 1286). El cebador 1331 se usó como un cebador directo. Las reacciones de las condiciones de PCR se realizaron utilizando 200 nM de cada cebador, 0,1-0,5 ng de ADN molde (pCRII-9R-X/A EK plásmido, SEQ ID NO: 6, FIG. **17**), 200um dNTPs, 0,5ul Q5 Hot Start, ADN de alta fidelidad polimerasa (NEB), 10ul 5X buffer y 10ul GC potenciador proporcionados por el fabricante. El plásmido de plantilla contenía un fragmento del gen de la cápside AAV9 que se ha modificado para tener dos sitios de restricción únicos (XbaI y AgeI) que flanquean la región que se varió (esta región crea una superposición con el rAAV9R-X/A-cap-in plásmido aceptor cis cortado con las mismas enzimas, ver la **FIG. 5C**). Además, el fragmento de la plantilla de PCR se modificó aún más para eliminar un sitio de restricción de EarI que se produce naturalmente dentro del fragmento del gen de la cápside e insertar un sitio KpnI. La modificación para eliminar el sitio de restricción de EarI proporciona una manera de eliminar cualquier contaminación de secuencia de vector de cápside de AAV "de tipo salvaje" de las bibliotecas que puedan surgir durante la clonación al digerir las bibliotecas con la enzima EarI. El paso de la digestión con EarI puede no ser necesario si se tiene cuidado de eliminar la posibilidad de que la secuencia de la cápside del AAV wt se traslade/amplifique. La inserción del sitio XbaI causó una mutación K449R, pero las otras mutaciones introducidas en la secuencia AAV9 son silenciosas.

[0141] Para facilitar la clonación de los fragmentos de PCR que comprenden las secuencias de la biblioteca de la cápside en un genoma de AAV recombinante, se modificó el plásmido rAAV-cap-en-cis para introducir los mismos dos sitios de restricción únicos, XbaI y AgeI, dentro de la secuencia del cápside flanqueando la región que será reemplazada por las bibliotecas basadas en PCR **(FIG. 5C)**. Además, la región de codificación entre los sitios XbaI y AgeI se eliminó para prevenir la producción de proteína de la cápside AAV9R X/A "wt" de cualquier vector no digerido durante la producción de virus de la biblioteca (AAV9R-delta-X/A-cap-in-cis, SEQ ID NO: 7, **FIG. 18)**.

[0142] Para ensamblar los productos de la biblioteca de PCR en el vector aceptor, los productos de PCR pueden digerirse con enzimas de restricción XbaI y AgeI y luego ligarse en el constructo aceptor de cap-in-cis cortado con las mismas enzimas. Alternativamente, los productos de la PCR y el vector aceptor rAAV-cap-in-cis se pueden ensamblar utilizando el método de ensamblaje de Gibson (Gibson et al., 2009). Ensamblaje enzimático de moléculas de ADN hasta varios cientos de kilobases. Nature Methods, 6 (5), 343-345. doi: 10.1038/nmeth.1318). En los ejemplos presentados aquí, se utilizó el método de Ensamblaje de Gibson para ensamblar de manera consistente más de 100 ng de ADN circular resistente a la ADNasa del Plásmido Seguro a partir de una reacción de ensamblaje hecha de 400 ng de XbaI y AgeI digeridosEl vector AAV9R-delta-X/A-cap-in-cis tratado con fosfatasa alcalina y 67 ng del producto de la PCR de la biblioteca con 30 ul de la mezcla maestra Gibson Assembly Master (NEB) en un volumen total de 60

ul. Las reacciones se llevaron a cabo a 50ºC durante 120 minutos.

[0143] Después del ensamblaje de Gibson, los productos de reacción se digirieron con una ADNasa segura para plásmidos (PS) como se indica (Epicentro), que digiere moléculas de ADN linealizadas pero no circularizadas. Las reacciones de ensamblaje se incubaron con 1 ul (10 U) de PS DNasa en una reacción que contenía 2 ul ATP y 7 ul del tampón de reacción suministrado por el fabricante (Epicentro) a 37ºC durante 20 minutos, seguido de una etapa de inactivación por calor a 70ºC durante 30 minutos. Esta reacción generalmente produjo más de 100 ng de plásmido ensamblado (como se define por la cantidad medida de producto que queda después de la reacción de la ADN DNasa (medida con el kit Qubit dsDNA Broad Range de Invitrogen). 100 ng son suficientes para transfectar 10 platos de 150 mm a 10 ng/plato. fue útil para transfectar esta cantidad del plásmido de biblioteca rAAV-cap-in-cis para minimizar el número de células de empaquetamiento que fueron transfectadas con copias múltiples del plásmido rAAV-cap-in-cis, lo que podría causar la generación de cápsides en mosaico. Las cápsides de mosaico (aquellas que tienen una cubierta de la cápside compuesta por más de una variante de la proteína de la cápside) llevarían solo un genoma de la variante de la cápside. Por lo tanto, no todos los aminoácidos dentro de las cápsides serían codificados por el gen de la cápside dentro del genoma empaquetado transfectando directamente el ADN ensamblado, en lugar de transformarlo primero en células competentes y amplificarlo en bacterias, fue posible transfectar las células de empaquetamiento con una biblioteca de máxima diversidad (teóricamente> 1e10 secuencias únicas).

Transfección de 293 células para la producción de virus de la cápside.

[0144] 7-10 placas de 150 mm de células 293T casi confluentes que se habían sembrado 16-30 horas antes de la transfección se transfectaron típicamente. Además del lOng del vector de biblioteca rAAV-cap-in-cis, se co-transfectaron 5,7 ug de puc18, 11,4 ug de ayudante Rep-AAP y 22,82 ug de pHelper (por placa) usando PEI en una proporción N:P de 25 (ver Grieger et al 2006). La mezcla de transfección se realizó en solución salina tamponada con fosfato (PBS) y se incubó a temperatura ambiente durante 10 minutos y luego se añadió gota a gota en el medio. 12-18 horas después de la transfección, los medios de las células transfectadas se intercambiaron por DMEM nuevo suplementado con 5% de FBS, 1X Pen/Strep y 1X mezcla de aminoácidos no esenciales (Invitrogen). Este medio se recogió 48 horas después de la transfección y se reemplazó con medio nuevo. A las 60 horas posteriores a la transfección, se recogieron los medios y las células. El virus presente en el medio se concentró por precipitación mediante la adición de poli(etilenglicol) y cloruro de sodio al 8% y 0,5 M, respectivamente. Los sedimentos celulares se resuspendieron en Tris 10 mM, $MgCl_2$ 2 mM y los virus se liberaron de las células mediante 3 ciclos de congelación y descongelación (alternando entre un baño hecho de etanol al 100% y hielo seco y un baño de agua a 37°C. Después de la descongelación final a 37°C, los lisados se trataron con 50U de Benzonasa durante 1 hora a 37°C. El virus precipitado del medio se recogió luego por centrifugación a 4.000 xg durante 30 minutos a 4°C. El virus sedimentado del medio se resuspendió en el mismo tampón Tris-$MgCl_2$. como arriba y luego combinado con la reserva viral del lisado celular. En este momento, la desoxicolina (DOC) se añadió al 0,5% y la reserva de virus se incubó a 37ºC durante 30 minutos adicionales. La reserva de virus se ajustó luego a 500 mM de NaCl y se incubó durante 30 minutos más antes de que el lisado se aclarara girando a 4000 xg durante 15 minutos a 25ºC. Después de la hilatura, los lisados de stock vírico se purificaron en gradientes escalonados de iodixanol (Optiprep, Sigma) (15%, 25%, 40% y 60% según lo descrito por Ayuso et al., 2010). Los virus se filtraron de forma estéril y se dializaron con filtros centrífugos Amicon Ultra 100K como se indica (Invitrogen) y se concentraron en PBS. Los títulos de virus se determinaron midiendo el número de copias del genoma resistente a DNaseI (GCs) utilizando qPCR y un plásmido linealizado como control (Gray et al 2011).

[0145] La producción de virus se detuvo a las 60 horas después de la transfección para reducir la probabilidad de transducción secundaria de las células productoras por que se libera en el medio el rAAV-cap-en-cis virus.

[0146] La transducción secundaria de las células que fueron transfectadas con éxito con todos los plásmidos necesarios para la producción de virus puede conducir a la generación de virus de más de una secuencia de la cápside (es decir, mosaicos). Se puede hacer un esfuerzo para minimizar la producción del virus del mosaico para asegurar que cada gen de la cápside se empaquete solo en la variante física de la cápside que codifica.

Otros ejemplos

[0147] En base a los resultados de los ejemplos iniciales presentados a continuación, se espera que este sistema se pueda usar junto con cualquier línea transgénica que exprese una recombinasa en el tipo de células diana de interés para desarrollar cápsides de AAV que transduzcan más eficientemente esa población de células diana. Las aplicaciones incluyen, pero no se limitan al desarrollo de cápsides que son más eficientes en la transducción de tipos de células específicos en cualquier órgano después de la administración de AAV IV, dirigidas a poblaciones específicas de neuronas después de inyecciones cerebrales intraparenquimatosas, mejorando el transporte neuronal (anterógrado o retrógrado), dirigidas a células tumorales, células madre hematopoyéticas, células beta productoras de insulina, epitelio pulmonar, músculo esquelético o cardíaco. Por lo tanto, los métodos de selección aquí descritos pueden aplicarse para uno o más de estos aspectos.

[0148] Además, el enfoque puede usarse para seleccionar virus que se dirigen a células humanas en animales quiméricos humanos/de ratón (las células humanas se harían para expresar Cre antes de la administración in vivo). Este último ejemplo puede ser útil en el desarrollo exitoso de vectores eficientes para aplicaciones de terapia génica,

ya que existe evidencia de que los serotipos de AAV que funcionan mejor en modelos animales no siempre funcionan con las mismas eficiencias en humanos (Lisowski et al 2013). Por lo tanto, puede ser ventajoso seleccionar los virus que transduzcan de manera más eficiente la población de células humanas diana en el contexto *in vivo* de un modelo animal.

[0149] Además, el método se puede utilizar junto con cualquier método de administración de virus (p. ej., inyección cerebral, SC, IP, intramuscular, intranasal, icv, intratecal, oral o intracraneal/intraparenquimatosa). En algunos casos, el vector puede administrarse por cualquier vía que incluye, entre otros: administraciones oral, intravenosa, intraarticular, intracardíaco, intramuscular, intradérmica, tópica, intranasal, intraparitoneal, rectal, sublingual, subcutánea, epidural, intracerebral, intracerebroventricular, intratecal, intravítrea o subretiniana.

[0150] Aunque los ejemplos en este documento han usado AAV9 como punto de partida, cualquier cápside de AAV natural o ingerida previamente también podría usarse como punto de partida para la selección usando este enfoque. Además, este método también podría ser útil para identificar otras secuencias codificantes o no codificantes dentro de un AAV u otro genoma viral que influyó en la transducción de las células que expresan recombinasa. Los ejemplos preferidos incluyen la selección de secuencias dentro del genoma de AAV que aumentan la conversión del genoma viral a dsDNA, aumentan la persistencia de los genomas virales facilitando la recombinación o circularización, aumentan la eficiencia de integración del genoma viral en un sitio o sitios favoritos en el genoma celular o secuencias que influyen en la expresión génica en la población de células diana.

[0151] Se da a conocer en el presente documento el uso de CREATE (Evolución diana ***A***AV basada en ***C***re ***Re***combinasa), una novela plataforma para la recuperación selectiva de secuencias de la cápside que transducen poblaciones de células diana Cre+. Usando CREATE, fue posible desarrollar varias variantes nuevas de cápside AAV con propiedades útiles, incluyendo una AAV-PHP.R2, que media la transducción retrógrada eficiente en el cerebro tan pronto como siete días después de la administración, y una segunda variante, AAV-PHP.B que cruza la barrera hematoencefálica del ratón adulto (BBB) y transduce una variedad de tipos de células neuronales del SNC con una eficiencia al menos 40 veces superior a la del AAV9, el estándar actual para el suministro sistémico. Además, la eliminación del tejido de todo el animal utilizando CLARITY basado en PARS (Yang et al., 2014b) como un método más rápido para evaluar el tropismo por serotipo a nivel celular y como un método para estudiar la morfología de las células individuales en el cerebro cuando se combina con la entrega sistémica de dosis AAV-PHP.B. Usados juntos, el mapeo de transducción en tejidos intactos y el método de selección de cápside basado en Cre presentado proporcionan una nueva plataforma que debería facilitar el desarrollo de virus personalizados adicionales.

EJEMPLO 3:

SELECCIÓN IN VIVO UTILIZANDO ANIMALES TRANSGÉNICOS DE GFAP-CRE

[0152] Se desarrollaron cápsides de AAV que transducen de manera más eficiente las células dentro del SNC de ratones adultos después de la inyección IV. Para este propósito, se utilizaron ratones transgénicos mGFAP-Cre que expresan Cre específicamente dentro de astrocitos y células madre neurales (NSC) en el cerebro adulto y la médula espinal. Las bibliotecas de cápside (1.2e11 GC) se administraron por vía intravenosa (IV) a ratones a través del seno retroorbital. 7-8 días después, se recogió el ADN de todo el cerebro y la médula espinal y se recuperaron las secuencias de la cápside rAAV de las células GFAP-Cre+ (FIG. 6). El ADN del vector se recuperó de un hemisferio del cerebro y la mitad de la médula espinal utilizando 4 ml de Trizol (Invitrogen). Se siguió el protocolo del fabricante, y la fracción acuosa que contiene ARN se precipitó con isopropanol y se sometió a tres lavados en etanol al 70% hecho con agua (toda el agua utilizada para la recuperación por PCR de las secuencias de la cápside en este protocolo se trata con UV usando una caja de luz UV durante 10-15 minutos antes de su uso). El material precipitado se resuspendió luego en Tris 10 mM pH 8,0. Además del ARN, esta fracción también contiene una fracción significativa del genoma viral, así como algo de ADN mitocondrial. Para eliminar el ARN, que redujo la eficiencia de la recuperación basada en la PCR de las secuencias de la cápside, las muestras se trataron con 1 ul de RNasa (Qiagen) durante la noche. Estrategias alternativas para la recuperación selectiva. Además del ARN, esta fracción también contiene una fracción significativa del genoma viral, así como algo de ADN mitocondrial. Para eliminar el ARN, que redujo la eficiencia de la recuperación basada en la PCR de las secuencias de la cápside, las muestras se trataron con 1 ul de RNasa (Qiagen) durante la noche. Estrategias alternativas para la recuperación selectiva. Además del ARN, esta fracción también contiene una fracción significativa del genoma viral, así como algo de ADN mitocondrial. Para eliminar el ARN, que redujo la eficiencia de la recuperación basada en la PCR de las secuencias de la cápside, las muestras se trataron con 1 ul de RNasa (Qiagen) durante la noche. Estrategias alternativas para la recuperación selectiva. También se podrían usar los genomas virales lejos del ADN genómico del animal, por ejemplo, el protocolo de extracción HIRT (Hirt 1967), la purificación de gel basada en el tamaño, los métodos de captura/hibridación específicos de secuencias o la digestión selectiva del ADN genómico de ratón mediante PS DNasa después de la digestión con una enzima de restricción que no corta el genoma de rAAV-cap-in-cis.

[0153] La recuperación de la secuencia de la cápside se realizó de manera dependiente de Cre utilizando los cebadores 1253 + 1316. Las condiciones de la PCR fueron 20-28 ciclos con 95C 20 s/60C durante 20 s/72C durante 30 s usando la polimerasa de ADN de alta fidelidad Hot Start Q5. El producto de la PCR luego se diluyó 1:10 y luego se usó como plantilla para una segunda reacción de PCR que generó el fragmento X a A (usando los cebadores 1331

+ 1312) que se clonó nuevamente en el constructor aceptor rAAV9R-deltaX/A-cap-in-cis para generar la siguiente ronda de virus de biblioteca utilizando los mismos métodos descritos anteriormente. 1 ul de las reacciones de Gibson Assembly se diluyó 1:10 y se transformó en células competentes Sure2 (Agilent) según las indicaciones del fabricante. Se seleccionaron al menos 10 colonias/biblioteca 12-16 horas más tarde, se aisló el ADN mediante un kit miniprep (Qiagen) y se secuenciaron los clones. Alternativamente, el ADN de los clones puede amplificarse por PCR usando los cebadores 1253 y 1312 y secuenciarse directamente, eliminando la necesidad de realizar preparaciones de mini plásmido de ADN.

[0154] Después de la primera ronda de selección todos los clones secuenciados para ambas bibliotecas fueron únicos. Por lo tanto, se realizó una segunda ronda de selección para enriquecer aún más las secuencias más potentes. La biblioteca de rAAV-cap-in-cis ensamblada se generó después de la primera ronda de selección para generar una segunda ronda de virus que luego se inyectó en un segundo lote de ratones GFAP-Cr+ como se describió anteriormente. Después de la segunda ronda, dos secuencias, G2B13 y G2B26 mostraron evidencia de enriquecimiento (Tabla 2).

TABLA 2

Secuencias enriquecidas de selección in vivo de GFAP-Cre

Variante	Línea de ratón	Administración	Rondas de selección	sitio	Secuencia(s) de ADN 7mer	Secuencia(s) AA 7mer	% de clones totales
G2B-13	GFAP-Cre	IV	2	452-8	CAGTCGTCGCAGACGCC TAGG (SEQ ID NO: 48)	QSSQTPR (SEQ ID NO: 54)	18%
G2B-26	GFAP-Cre	IV	2	588	ACTTTGGCGGTGCCTTTT AAG (SEQ ID NO: 49)	TLAVPFK (SEQ ID NO: 1)	27%
TH1.1-32	TH-Cre	Intracraneal (estriado)	1+1	452-8 + 588	ATTCTGGGGACTGGTAC TTCG (SEQ ID NO: 50)	ILGTGTS (452-8) (SEQ ID NO: 55)	18%
					ACGCGGACTAATCCTGA GGCT (SEQ ID NO: 51)	TRTNPEA (588) (SEQ ID NO: 56)	9%
TH1.1-35	TH-Cre	Intracraneal (estriado)	1+1	452-8 + 588	ATTCTGGGGACTGGTAC TTCG (SEQ ID NO: 52)	ILGTGTS (452-8) (SEQ ID NO:57)	18%
					AATGGGGGGACTAGTAG TTCT (SEQ ID NO: 53)	NGGTSSS (588) (SEQ ID NO: 58)	36%

ES 2 739 288 T3

[0155] Para ensayar las variantes recuperadas, las secuencias se cortaron con BsiWI y AgeI y se ligaron en un ayudante AAV2/9R-X/A rep/cap (ayudante AAV2/9 rep/cap modificado con la secuencia AAV9R-X/A de la cápside del plásmido rAAV-cap-in-cis) también se cortó con BsiWI y AgeI y se transformó en células competentes DH5alfa (NEB). El ADN plasmídico se purificó utilizando un kit Endofree Plasmid Maxi (Qiagen). Los plásmidos rep/cap resultantes que llevan las nuevas secuencias variantes, o AAV2/9 rep/cap como control, se usaron para empaquetar un genoma de rAAV que contiene un casete indicador doble eGFP-2A-luciferasa dirigido por un promotor CAG ubicuo (rAAV-CAG-eGFP-2A-Luc-WPRE-SV40pA). Las nuevas cápsides empaquetaron el genoma con eficiencias comparables a las del AAV9 **(FIG. 7**). 1e12 GC de cada vector se inyectó IV en ratones hembra adultos C57B1/6. Seis días después, los ratones se perfundieron con paraformaldehído al 4% en tampón fosfato 100 mM y se examinaron los cerebros para determinar la fluorescencia de eGFP.

[0156] De manera notable, la transducción por la variante G2B26 fue lo suficientemente eficiente como para que la fluorescencia eGFP nativa en todo el cerebro intacto se pudiera ver con un objetivo 1X en un microscopio de epifluorescencia **(FIG. 8A)**. En este mismo ajuste de exposición, se observa poca o ninguna fluorescencia de eGFP en el cerebro del ratón inyectado con AAV9. En secciones preparadas a partir de un cerebro de un ratón inyectado con G2B26, se observó la transducción de neuronas y glía en todas las regiones examinadas en el cerebro y la médula espinal (**FIGs. 8** y **9)**. En ciertos núcleos talámicos, más del 90% de los cuerpos celulares NeuN + expresaron GFP **(FIG. 9G)**. La transducción de las neuronas motoras en la médula espinal ventral también fue robusta **(FIG. 9F**). Numerosos Sox2 + glia expresaron GFP **(FIG. 9I).** La variante G2B13 también transdució astrocitos y neuronas más eficientemente que AAV9, pero el efecto no fue tan dramático en comparación con la transducción por G2B26 (**FIGs. 8** y **9A-C)**. La variante G2B13 mostró una fuerte transducción de los tractos de la fibra en el tronco cerebral dorsal **(FIG. 9B)** y la médula espinal **(FIG. 9C)**, así como una transducción hepática robusta **(FIG. 8C).**). También parece que la variante G2B26 proporciona un inicio de expresión más rápido en el SNC que AAV9. La transducción por AAV9 a los seis días después de la inyección fue débil. Se observó una expresión más fuerte por célula y más células que expresaban eGFP con la misma dosis de AAV9 21 días después de la inyección.

[0157] Dado que se ha propuesto que el desempaquetado rápido es un componente importante de la eficiencia de transducción y la persistencia del genoma viral (Wang et al. 2007), la recuperación de secuencias de la cápside poco después de la inyección (en este caso, 7-8 días) puede ser un componente importante de una selección exitosa.

[0158] En el ejemplo anterior, el número de ciclos se puede determinar empíricamente con el número óptimo de ciclos dentro de un rango que produce más producto a partir de muestras tomadas de células Cre+/animales que a partir de muestras que carecen de células Cre+. Si se permite que la reacción de PCR continúe más allá de este rango óptimo mediante la realización de demasiados ciclos, los productos pueden recuperarse incluso a partir de muestras negativas de Cre. Puede ser conveniente evitar hacer demasiados ciclos.

EJEMPLO 4:

SELECCIÓN IN VIVO PARA MEJORAR LA TRANSDUCCIÓN DE RETROGRADO UTILIZANDO ANIMALES DE TH-CRE

[0159] En un segundo ensayo de la plataforma de selección Cre-dependiente, se preguntó si se podría generar serotipos novelos de AAV que llevan la transducción más eficiente de neuronas TH+ en la parte compacta de sustancia negra (SN) después de la inyección del virus en el cuerpo estriado, una estructura que recibe axones de las neuronas TH. Este esquema de selección está diseñado para desarrollar AAV que sean capaces de un rápido transporte y transducción retrógrados de células TH+ y no TH+.

[0160] Se utilizaron las mismas bibliotecas como las inicialmente generadas para la selección de Ejemplo 3 y se inyectaron 0,6ul del virus bilateralmente en el estriado de ratones macho adultos TH-Cre+ utilizando las coordenadas estereotáxicas 0,7 mm rostral, 2,0 mm lateral y 3,0 mm ventral desde Bregma. 10 días después, se recogió la región que contenía el SN y se aisló ADN de virus del tejido como se describió anteriormente. Para estas disecciones, el reportero de mCherry expresado a partir del genoma de rAAV-cap-in-cis ayudó en la identificación del SN **(FIG. 10C)** y la confirmación de que las inyecciones de las bibliotecas de virus habían apuntado a las áreas deseadas **(FIG. 10B).** El ADN del virus se obtuvo de la muestra de tejido que contiene SN utilizando Trizol (Invitrogen) como se describe anteriormente y se utilizó la misma estrategia de PCR dependiente de Cre para recuperar selectivamente las secuencias de la cápside que condujeron a la transducción de las neuronas TH+ **(FIG. 10D)**. El uso de cebadores que amplifican las secuencias de la cápside de todos los genomas independientemente del estado de recombinación (1253 + 1267) demuestra que las secuencias virales también estaban presentes en los controles de Cre **(FIG. 10D,** panel inferior). Las secuencias recuperadas a través de la estrategia dependiente de la recombinación de Cre se clonaron nuevamente en el aceptador de la biblioteca rAAV-cap-in-cis como se describe anteriormente en el Ejemplo 3.

[0161] Las colonias se recogieron para la secuenciación. Después de la primera ronda de selección, todas las secuencias ensayadas fueron únicas, por lo que se realizó una segunda ronda de selección como se describió anteriormente. Además de continuar con las bibliotecas modificadas en los dos sitios individuales, también se crearon bibliotecas combinatorias mezclando todas las secuencias recuperadas en el sitio de reemplazo 452-8 con las secuencias recuperadas en el sitio de inserción 588 utilizando la estrategia de PCR descrita en la **FIG**. **10.** Se

prepararon bibliotecas de virus de la cápside a partir de las secuencias recuperadas, se seleccionaron de nuevo en ratones TH-Cre y se recuperaron como se describió anteriormente.

[0162] Después de la segunda ronda de selección en la biblioteca combinatoria, varias secuencias en ambos sitios de aleatorización mostraron evidencia de enriquecimiento. Se seleccionaron varias secuencias de cápside novedosas para ensayarlas como variantes individuales. Las secuencias se clonaron en un ayudante de rep/cap AAV2/9R-X/A utilizando los sitios únicos BsiWI y AgeI presentes en ambos vectores. Los plásmidos rep/cap resultantes que llevan las nuevas secuencias variantes, o AAV2/9 rep/cap como control, se utilizaron luego para empaquetar un genoma de rAAV-CAG-GFP-W-pA de cadena simple (ss). Las nuevas cápsides empaquetaron el genoma con eficiencias comparables con AAV9 **(FIG. 7)**. 7e9 VGs (en 0,5ul) de cada virus fueron inyectados individualmente, y bilateralmente, en ratones C57B1/6 adultos usando las mismas coordenadas estereotáxicas descritas anteriormente. 7 días después, a los ratones se les administró una sobredosis de Euthasol y se sacrificaron mediante perfusión cardíaca con PFA al 4% como se describe anteriormente. En este momento, había pocas o ninguna neurona GFP+/TH + en la SNc de ratones que recibieron una inyección del virus de control, AAV9: CAG-GFP-W-SV40pA. En contraste, hubo numerosas neuronas GFP+/TH+ presentes en los ratones que recibieron inyecciones del mismo genoma rAAV empaquetadas en los nuevos clones TH1.1-32 (**FIG 12D)** o TH1.1-35 (**FIG 23D)**.

[0163] Una lista de los cebadores usados en los ejemplos 3 y 4 se proporciona en la Tabla 1 a continuación:

TABLA 1

Cebador	Propósito	Secuencia
1253	Amplifiación dependiente de Cre, adelante	CAGGTCTTCACGGACTCAGACTATCAG (SEQ ID NO: 16)
1254	9R-X/A delta, inverso	CAACCGGTAATAGTTCTAGAGAGATAGTACAAGTATTGGTCGATGAGTG (SEQ ID NO: 37)
1255	9R-X/A delta, adelante	CTCTCTAGAACTATTACCGGTTGGGTTCAAAACCAAGGAATACTACTTC (SEQ ID NO: 38)
1267	Recuperación de biblioteca, no recombinada, inverso	GTCCAAACTCATCAATGTATCTTATCATGTCTG (SEQ ID NO: 39)
1280	Parada VP1, inverso	GAGTCAATCTGGAAGTTAACCATCGGCA (SEQ ID NO: 40)
1281	Parada VP1, adelante	GATGGTTAACTTCCAGATTGACTCG (SEQ ID NO: 41)
1283	Parada VP2, inverso	GACTACTCTACAGGCCTCTTCTATCCAG (SEQ ID NO: 42)
1284	Parada VP2, adelante	GATAGAAGAGGCCTGTAGAGTAGTCTCC (SEQ ID NO: 43)
1285	Parada VP3, inverso	CATCGGCACCTTAGTTATTGTCTG (SEQ ID NO: 44)
1286	Parada VP3, adelante	GACAATAACTAAGGTGCCGATGGAGTGG (SEQ ID NO: 45)
1286	Aleatorización sitio 588, inverso	GTATTCCTTGGTTTTGAACCCAACCGGTCTGCGCCTGTGCMNNMNNMN NMNNMNNMNNMNNTTGGGCACTCTGGTGGTTTGTG (SEQ ID NO: 23)
1287	Aleatorización sitio 452-8, adelante	CATCGACCAATACTTGTACTATCTCTCTAGAACTATTNNKNNKNNKNNKN NKNNKNNKCAAACGCTAAAATTCAGTGTGGCCGGA (SEQ ID NO: 22)
1312	Sitio 452-8, inverso y generación de fragmento X/A, inverso	GGAAGTATTCCTTGGTTTTGAACCCA (SEQ ID NO: 19)
1316	Recuperación de biblioteca, Cre-dependiente, adelante (invertido por recombinación)	CAAGTAAAACCTCTACAAATGTGGTAAAATCG (SEQ ID NO: 17)
1331	Sitio 588, inverso y generación de fragmento X/A, adelante	ACTCATCGACCAATACTTGTACTATCTCTCTAGAAC (SEQ ID NO: 18)
1352	Generación de biblioteca combinatoria, inserción de mutación EarI a KpnI, Inverso	GTCTCTGCCGGTACCTTGTTTGCCAAAAATTAAAGATCCA (SEQ ID NO: 46)
1353	Generación de biblioteca combinatoria, inserción de mutación EarI a KpnI, Adelante	GCAAACAAGGTACCGGCAGAGACAACGTGGATGCGGACA (SEQ ID NO: 47)

[0164] Al utilizar la plataforma para la selección descrita en el presente documento, fue posible desarrollar varias cápsides que proporcionan una expresión génica generalizada y mejorada en el SNC.

[0165] En particular, una cápside (AAV-PHP.R2) era capaz de transporte rápido, retrógrada dentro de las neuronas del SNC después de la inyección intracerebral, mientras que otra cápside (AAV-PHP.B) transdució células a través de los sistemas nerviosos central con mayor eficiencia en 40-90 veces que AAV9 cuando se entrega de forma sistémica. AAV-PHP.B transduce las neuronas y la glía y, por lo tanto, es muy adecuado para la transferencia de genes a los tipos de células neuronales del SNC global, incluidas las neuronas, los astrocitos y los oligodendrocitos.

[0166] Dada la gran colección de líneas transgénicas Cre específicas de tipo celular, la presente plataforma de selección de cápside es un recurso valioso para personalizar vectores de suministro de genes para aplicaciones biomédicas.

EJEMPLO 5

[0167] Dentro del genoma recombinante rAAV cap-en-cis, se introdujeron dos elementos para facilitar la selección. El primero es un casete informador mCherry, que tiene un fragmento promotor de 398 pares de bases del gen de ubiquitina C (UBC), el cADN de 711 pb de mCherry y una región no traducida 3' de 118 pb que contiene una secuencia de poli adenilación sintética de 51 pb (poliA) (Levitt et al., 1989). El segundo elemento, y más relevante, es un "interruptor" dependiente de Cre, que tiene un par de sitios loxP modificados e invertidos (lox71 y lox66) (Araki et al., 1997) que flanquean una secuencia de poliA de SV40 corriente abajo del gen de la cápside. Este elemento flotante creó una secuencia Cre-invertible que permite la amplificación por PCR selectiva y la recuperación de solo aquellas secuencias de capuchones contenidas dentro de los genomas de AAV que han transducido células Cre+. **(FIG. 22A**).

[0168] Para proporcionar *rep,* pero no *cap,* la función de genes en trans, un plásmido auxiliar AAV2/9 REP-CAP se modificó mediante la inserción de cinco en codones de parada de marco dentro del marco de lectura para las proteínas de la cápside, VP1, VP2 y VP3. Estos codones de parada se diseñaron para interrumpir la expresión de la proteína de la cápside, pero no alterar la secuencia de aminoácidos de la proteína activadora del ensamblaje (AAP), que se expresa desde un marco de lectura alternativo dentro del gen de la cápside ***(*FIG. 22B)** (Sonntag et al., 2010). De esta manera, el plásmido auxiliar REP-AAP modificado continúa proporcionando todos los productos génicos de AAV en trans, a excepción de la expresión de la proteína de la cápside. Para ensayar si este sistema dividido rAAV-CAP-in-cis-lox y REP-AAP ayuda a generar eficientemente rAAV, se realizó una transfección triple de células HEK 293T con el gen rAAV-CAP-in-cis, el ayudante REP-AAP, y el ayudante de plásmido adenoviral, pHelper. Es importante destacar que, con estos plásmidos, fue posible generar virus recombinantes con una eficiencia equivalente, si no mayor, que la observada cuando se usó un ayudante REP-CAP AAV2/9 para empaquetar un genoma rAAV que codifica mCherry (AAV-UBC-mCherry) **(FIG. 22C)**.

[0169] En contraste, cuando se usó el ayudante AAV REP-AAP para empaquetar AAV-UBC-mCherry, que carece del gen de la cápside en cis, se generó poco o ningún virus, lo que confirma que la expresión de la proteína de la cápside del rAAV-CAP. El vector cis-lox fue requerido para la producción de rAAV. Utilizados juntos, el ayudante rAAV-CAP-in-cis-lox y AAV REP-AAP proporcionaron una plataforma novedosa, que a continuación se denomina CREATE, para la recuperación selectiva de secuencias de la cápside de poblaciones de células definidas genéticamente dentro de muestras de tejido complejas.

EJEMPLO 6

[0170] Se generaron dos bibliotecas de cápside basadas en AAV9 mediante PCR utilizando una estrategia de aleatorización de base mixta. Se creó una biblioteca insertando 7 aminoácidos de secuencia aleatoria entre AA588-9 (posición VP1) de la cápside de AAV9 y otra con 7 aminoácidos de secuencia aleatoria que reemplaza AA452-8 de AAV9. La estrategia de clonación se diseñó de tal manera que el producto de PCR recuperable contendría solo el tramo de aminoácidos que abarcan las regiones variables (secuencias entre AA450 y AA592), que abarca una porción significativa de los aminoácidos expuestos en la superficie, mientras que el resto de la secuencia de la cápside dentro de la columna vertebral el vector permanece sin modificar. Los fragmentos de la biblioteca se clonaron luego en el vector rAAV-delta-cap-in-cis y los productos ensamblados se transfectaron directamente en células de empaquetamiento para producir virus, sin pasar por el cuello de botella primario de la diversificación de la biblioteca, la transformación bacteriana. Con este enfoque, la diversidad de la biblioteca está limitada por el número de células transfectadas, en lugar del número de transformantes bacterianos que dan como resultado una diversidad estimada de $1x10^7$ -$1x10^8$ secuencias únicas. Usando este enfoque, fue posible lograr rendimientos de 5-$10x10^{11}$ VG.

[0171] Los vectores que median la transducción retrógrada eficiente de las neuronas, es decir, la captación del vector por los axones y el transporte de regreso al núcleo, son deseados para el rastreo del circuito neuronal y los enfoques interseccionales para la expresión de genes específicos del circuito, y también pueden tener usos para la entrega de genes clínicos. Si bien los virus como la rabia recombinante y el virus del herpes simple (VHS) exhiben una transducción retrógrada altamente eficiente y son útiles para estudios de trazado de circuitos a corto plazo, su toxicidad a largo plazo excluye su uso en experimentos longitudinales o experimentos en los que su impacto en la salud celular cofundada (por ejemplo, optogenética, envejecimiento, estudios de neurodegeneración). Para la expresión génica de larga duración, los AAV capaces de una eficiente transducción retrógrada serían muy valiosos, ya que permitirían que el amplio conjunto de herramientas disponible en el formato del genoma rAAV se aplique a aplicaciones que requieran

transducción retrógrada (NIH Brain Initative Working Group, 2013).

EJEMPLO 7

SELECCIÓN IN VIVO PARA VARIANTES DE AAV CON TRANSDUCCIÓN DE RETROGRADO MEJORADA EN EL ROEDOR SNC

[0172] Varios serotipos de AAV han demostrado mediar en la transducción de retrógrado de las neuronas en el SNC con diferentes eficacias (Aschauer et al, 2013;. Castillo et al, 2014A;. Castillo et al, 2014b;. Cearley y Wolfe, 2007; Hutson et al., 2012; Low et al., 2013; Salegio et al., 2013; Samaranch et al., 2012). Para desarrollar cápsides de AAV con transducción retrógrada mejorada, se estableció una selección *in vivo* de cápsides que transdujeron neuronas dopaminérgicas TH+ en la sustancia negra a través del transporte retrógrado desde sus axones dentro del estriado (Smith y Bolam, 1990). Las bibliotecas de AAV-CAP-en-cis-lox 452-8r y 588i se inyectaron por separado en el estriado de ratones adultos TH-Cre+. 10 días después, el tejido que rodea la sustancia negra (SN) **(FIG. 23A)** se diseccionó y aisló el ADN viral. Para estas disecciones, el indicador de mCherry expresado a partir de los vectores de biblioteca AAV-cap-in-cis ayudó en la identificación del SN como el SN pars reticulata (SNr) fue fácilmente identificado desde los axones mCherry$^+$ que se proyectan al SNr desde el estriado. La expresión de mCherry en el cuerpo estriado confirmó que las inyecciones de la biblioteca de virus se habían dirigido correctamente **(FIG. 23A).** Después de la primera ronda de selección, se secuenciaron 10 clones de cada biblioteca y se encontró que todas las secuencias ensayadas eran únicas, por lo que se realizó una segunda ronda de selección. Para diversificar aún más las bibliotecas después de la ronda inicial de enriquecimiento, se crearon bibliotecas combinatorias mezclando todas las secuencias recuperadas de la biblioteca 452-8r con todas las secuencias recuperadas de la biblioteca 588i mediante PCR (consulte la **Figura 27D)**. Se prepararon bibliotecas de cápside viral de la biblioteca combinatoria y se seleccionaron de nuevo en ratones TH-Cre como se describió anteriormente. Después de la segunda ronda de selección, varias secuencias en ambos sitios de asignación al azar mostraron evidencia de enriquecimiento.

[0173] La variante más altamente enriquecida, PHP.R2, se caracterizó adicionalmente mediante el ensayo de forma individual (véase la Tabla 3 para información de la secuencia y los datos de enriquecimiento).

TABLA 3

Variante	Selección	Ruta	Rondas	Sitios (s)	Secuencia de ADN 7mer	AA sec.	%
PHP.R2	TH	i.c.	1+1	452-8r	ATTCTGGGGACTGGTACTTCG	ILGTGTS (SEQ ID NO: 55)	18%
				588i	(SEQ ID NO: 50)	NGGTSSS (SEQ ID NO: 58)	36%
					AATGGGGGGACTAGTAGTTCT (SEQ ID NO: 53)		
PHP.A	GFAP	i.v.	2	588i	TATACTTTGTCGCAGGGTTGG	YTLSQGW (SEQ ID NO: 60)	40%
					(SEQ ID NO: 59)		
PHP.B	GFAP	i.v.	2	588i	ACTTTGGCGGTGCCTTTTAAG	TLAVPFK (SEQ ID NO: 1)	27%
					(SEQ ID NO: 49)		

[0174] La Tabla 3 enumera las variantes AAV-PHP, la línea transgénica Cre utilizada para realizar la selección de la biblioteca, la ruta de administración y el número de rondas de selección utilizadas para enriquecer las variantes mejoradas. 1 + 1 se refiere a una ronda de selección de las dos bibliotecas separadas y luego a una ronda adicional de selección de la biblioteca combinatoria. El sitio dentro de AAV9 que se modificó en cada variante recuperada se enumera como la(s) secuencia(s) de ADN 7mer y la(s) secuencia(s) de aminoácidos (AA seq.) que se modifican en cada variante de cápside. También se proporciona el número de ocurrencias de la secuencia enriquecida como un porcentaje del número total de clones secuenciados.

[0175] Se usó AAV-PHP.R2 para empaquetar un genoma de AAV-CAG-GFP de cadena simple (ss) y se inyectó en el cuerpo estriado de ratones adultos. En particular, después de solo 7 días, se observó una expresión robusta de GFP en el sitio de inyección del cuerpo estriado **(FIG. 23B),** así como en sitios distantes a la inyección que se sabe que envían proyecciones al cuerpo estriado, incluyendo la SNc **(FIG. 23C** y **23D),** la corteza **(FIG. 23E** y **23F),** el tálamo **(FIG. 23G)** y la amígdala **(FIG. 23H**). Estos resultados demuestran que AAV-PHP.R2 puede proporcionar una rápida transducción retrógrada de varias poblaciones neuronales distribuidas.

EJEMPLO 8

SELECCIÓN IN VIVO DE LAS VARIANTES DEL AAV CAPACES DE TRANSDUCCIÓN EN EL SNC GENERAL DESPUÉS DE LA ADMINISTRACIÓN SISTÉMICA

[0176] El presente ejemplo examina el desarrollo de cápsides de AAV que transducen las células de manera más eficiente en todo el SNC. Varios AAV, en particular AAV9, rh.10 y rh.8, transducen las neuronas del sistema nervioso central y la glia después de la administración intravenosa sistémica neonatal o en adultos (Duque et al., 2009; Foust et al., 2009; Gray et al., 2011; Samaranch et al., 2011; Yang et al., 2014a). Si bien la administración sistémica de rAAV con estos serotipos es capaz de una administración generalizada del SNC, la eficiencia de transducción se reduce significativamente en comparación con la que se puede lograr en otros órganos como el hígado, el corazón o el músculo esquelético (Pulicherla et al., 2011). El presente ejemplo demuestra el uso de la plataforma CREATE para desarrollar cápsides que transducen de manera más eficiente el SNC a nivel mundial. Esto se hizo debido a los importantes papeles que desempeñan los astrocitos en la patogénesis de la enfermedad neurodegenerativa junto con el tropismo de base de AAV9 para astrocitos.

[0177] Las bibliotecas de la cápside AA452-8r y AA588i descritas anteriormente fueron administradas a ratones mGFAP-Cre transgénicos que expresan Cre a partir del promotor GFAP de ratón, que se expresa dentro de los astrocitos y células madre neurales (NSC) en el cerebro adulto y la médula espinal (Garcia et al., 2004). $1x10^{11}$ VG de cada biblioteca de cápside se inyectaron en ratones adultos positivos GFAP-Cre positivos y ratones negativos GFAP-Cre como controles. Siete días más tarde, se aisló el ADN del virus de los cerebros y la médula espinal y las secuencias recuperadas de la cápside de los genomas virales que se habían sometido a recombinación mediada por Cre mediante PCR. Los fragmentos recuperados se clonaron nuevamente en el vector aceptor rAAV-CAP-in-cis-lox, y los clones de cada biblioteca se seleccionaron al azar para la secuenciación. Como se observó en la primera ronda de la selección TH-Cre, todas las secuencias analizadas recuperadas de ambas bibliotecas después de la primera ronda fueron únicas. Después de la segunda ronda, se identificó una secuencia única, designada como AAV-PHP.B, de la biblioteca 588i y mostró signos de enriquecimiento.

[0178] Para evaluar la variante de AAV-PHP.B individualmente, esta cápside o AAV9 se usó para empaquetar un vector indicador de Luciferasa de GFP dual y luciérnaga, ssAAV-CAG-GFP-2A-Luc. AAV-PHP.B empaquetó el genoma recombinante con una eficiencia similar a AAV9 **(FIGs. 27A-27E).** A continuación, $1x10^{12}$ VG de ssAAV-CAG-GFP-2A-Luc empaquetado en AAV-PHP.B o AAV9 se administró a ratones adultos mediante inyección IV y se evaluó la transducción por expresión de GFP tres semanas después. Sorprendentemente, esta variante transdució todo el SNC con una alta eficacia como lo indica la inmunotinción para GFP **(FIG.24A y 24C)** y el análisis de la fluorescencia de eGFP nativa en varias regiones del cerebro **(FIGs. 24B, 24D** y **24G),** la médula espinal **(FIG.. 24D** y **24G)** y retina **(FIG. 24E-24F).** La fluorescencia nativa de GFP permaneció dramáticamente aumentada sobre AAV9 incluso cuando se entregó 10 veces menos AAV-PHP.B **(FIG. 25A-**derecha y **25D).** En marcado contraste con el AAV9, que marca escasamente las neuronas y la glía, los astrocitos transducidos individuales fueron difíciles de discernir en ratones que recibieron $1x10^{12}$ VG AAV-PHP.B, pero se pudieron ver en animales que recibieron 10 veces menos virus **(FIG. 24A - 24B** y la **FIG. 25A).** AAV-PHP.B también transdució células de Purkinje cerebelosas con una eficacia sorprendentemente alta como se demuestra por la co-localización de GFP y la inmunotinción de Calbindina **(FIG. 24C).** Aprovechando los avances recientes en el aclaramiento de tejido basado en CLARITY (Yang et al., 2014b), la fluorescencia de eGFP nativa se formó en imágenes a través de varios cientos de micrones de secciones de tejido de la corteza, el estriado y la médula espinal ventral. Estas representaciones en 3D demuestran aún más la eficiencia de la transducción con la variante AAV-PHP.B y confirman la limpieza del tejido (Chung y Deisseroth, 2013; Tomer et al., 2014; Yang et al., 2014b) como un medio para evaluar la distribución tridimensional de las células transducidas dentro del cerebro **(FIG. 24G).**

[0179] Para cuantificar la transducción de SNC por esta variante en comparación con AAV9, se midieron el número de genomas virales presentes en varias regiones del cerebro 25 días después de la inyección. La transducción de cerebro y médula espinal por AAV-PHP.B fue entre 40 y 92 veces más eficiente que AAV9, dependiendo de la región examinada **(FIG. 24H),** mientras que fuera del SNC, el vector AAV-PHP.B transdució varios órganos periféricos menos eficientes que AAV9 **(FIG. 24I**). Sorprendentemente, en todas las regiones distintas del cerebelo, el número de genomas virales detectados en el SNC en ratones tratados con AAV-PHP.B fue similar al observado en el hígado y mayor que el observado en los otros órganos periféricos examinados. En marcado contraste, después de la transducción de AAV9, el número de genomas de AAV detectados en cualquiera de las regiones del SNC examinadas fue al menos 120 veces menor que el número encontrado en el hígado. Por lo tanto, si bien el tropismo de AAV-PHP.B no era específico del SNC, la transducción mejorada que muestra este vector era específica del SNC. En una selección inicial utilizando CREATE en ratones GFAP-Cre, se identificó una variante basada en AAV9, AAV-PHP.A, que exhibió una transducción de astrocitos más eficiente y un tropismo reducido para varios órganos periféricos **(FIG. 28A-28E).** Sobre la base de estos resultados, AAV-PHP.B parece aplicable para la transferencia de genes no invasiva y en todo el SNC en adultos.

[0180] AAV9 transduce preferentemente astrocitos cuando se administra sistémicamente a animales adultos, pero también transduce neuronas en varias regiones **(FIG.24A y FIG. 25A,** Foust et al. 2009, y Yang et al. 2014). Para

examinar los tipos de células transducidas por AAV-PHP.B, la colocalización de la expresión de GFP con proteínas expresadas en las poblaciones de células específicas se analizó. AAV-PHP.B transdució astrocitos $GFAP^+$ **(FIG. 25A),** oligodendrocitos $CC1^+$ **(25B),** neuronas $NeuN^+$ **(FIG. 25C** y **FIG. 25D)** pero no microglia $IBA1^+$ **(FIG. 25E).**

[0181] Varios tipos de células de importancia clínica también se dirigen con alta eficiencia, incluidas las neuronas dopaminérgicas TH^+ en la SNc **(FIG. 25F),** las neuronas motoras espinales **(FIG. 34D** y la **FIG. 24H**) y las neuronas espinosas del medio estriado **(FIG. 25D).).** Además, varias poblaciones de interneuronas también se transdujeron **(FIG. 25G-25J),** aunque en raras ocasiones se encontró que la fluorescencia de eGFP fuerte se colocalizaba con células con tinción con calretinina **(FIG. 25J).**

[0182] En suma, la administración IV adulta de AAV-PHP.B se puede utilizar para orientar, con alta eficiencia, numerosos tipos de células del SNC de interés científico y clínico.

Aclaramiento de tejidos para la caracterización de serotipos de tropismo y fenotipado de células 3D

[0183] Debido a que CLARITY permite la obtención de imágenes en 3D de células en secciones gruesas o tejido intacto (Chung et al 2013; Yang et al. 2014) y AAV-PHP.B transduce numerosos tipos de células gliales y neuronales en el cerebro, ya sea que CLARITY pueda utilizarse junto con una dosis baja de AAV-PHP.B para proporcionar un etiquetado aleatorio, similar al de Golgi, de células neurales en el SNC. Para evaluar este enfoque, $1x10^{10}$ VG de AAV-PHP.B que expresan GFP fue entregado en ratones adultos mediante inyección IV. Estos ratones se purificaron por perfusión y se obtuvo una imagen de fluorescencia de GFP nativa en los cerebros de los ratones. Las neuronas individuales, los astrocitos y las células endoteliales eran visibles y se podían visualizar a través de al menos 400 um de tejido eliminado **(FIG. 26D).**

[0184] Este enfoque puede ser útil para estudiar la morfología de células individuales en estados normales y enfermos. Este enfoque se puede usar para coexpresar un informador junto con cualquiera de los siguientes ejemplos de elementos genéticos para investigar los efectos de dicho elemento genético en la morfología o conectividad celular in vivo: un gen que codifica una proteína de interés; Cre, u otra recombinasa, para la modificación génica condicional en animales transgénicos que albergan un alelo(s) diana floxed; alelos condicionales, floxed, a animales transgénicos hechos para expresar Cre en una población de células diana definida; un gen de reducción de casetes que contiene un promotor adecuado y shARN o miARN, o una esponja o señuelo de miARN endógeno. Dada la facilidad de ajuste de la frecuencia de marcaje/modificación genética mediante la modulación de la cantidad de virus administrado, este vector también podría usarse para abordar cuestiones relacionadas con la autonomía celular mediante la generación de mosaicos genéticos.

[0185] La eliminación de todo el tejido animal utilizando CLARITY basado en PARS también puede ser útil para la evaluación del tropismo de AAV a nivel celular. Normalmente, este es un proceso que requiere mucha mano de obra y requiere el procesamiento, montaje e imagen de secciones delgadas de tejido de cada órgano de 1 a 100 micras. Se exploró el posible despeje de tejido de todo el animal para reducir esta carga. $1x10^{12}$ VG de ssAAV-CAG-GFP-2A-Luc empaquetado en AAV-PHP.B o AAV9 se administró en ratones adultos mediante inyección IV. Tres semanas después, todo el tejido de los ratones se eliminó utilizando el método CLARITY basado en PARS descrito en Yang et al. (2014) y usó imágenes confocales y reconstrucción de imágenes en 3D (software Imaris, Bitplane) para evaluar la expresión nativa de GFP como informador de la transducción vectorial. En varios órganos, incluidos el músculo esquelético, el pulmón, el páncreas y el hígado, los ratones que recibieron AAV-PHP.B mostraron una reducción en la expresión de GFP en comparación con los ratones que recibieron una dosis equivalente de AAV9 **(FIG. 29**). La expresión reducida de GFP en varios órganos periféricos observada con AAV-PHP.B en comparación con AAV9 parece coherente con el número de VG detectados para cada vector en estos mismos órganos periféricos **(FIG. 24I)**. Tenga en cuenta las fibras nerviosas positivas a GFP presentes en el músculo y el páncreas de ratones inyectados con AAV9 y AAV-PHP.B.

[0186] El etiquetado de rAAV combinado con CLARITY será un enfoque útil para estudiar las morfologías 3D de los nervios periféricos.

[0187] Los siguientes aspectos se aplican a los experimentos descritos en los Ejemplos 7 y 8 anteriores:

Ratones

[0188] Se adquirieron ratones C57B1/6 hembras de 5 semanas de edad de Jackson Labs (Maine). Los ratones GFAP-Cre que expresan Cre bajo el control del promotor GFAP de ratón (Garcia et al., 2004) y los ratones TH-Cre (Savitt et al., 2005) fueron de Jackson Labs. La selección in vivo se realizó en ratones adultos de cualquier sexo.

Plásmidos

[0189] El plásmido rAAV-cap-in-cis-lox contiene los siguientes elementos clonados en un vector que contiene ITR AAV2 (Balazs et al. 2011). Un casete de expresión mCherry (fragmento de 398 pb del gen UBC humano aguas arriba del informador mCherry seguido de una secuencia de poliA sintética de 48 pb - (Levitt et al., 1989) seguido por el

casete de cápside AAV9. La expresión del gen de cápside AAV9 se colocó bajo control de la secuencia del promotor p41 de AAV5 (1680-1974 de GenBank AF085716.1; Qiu, Nayak Pintel 2002 y Farris y Pintel 2008) y secuencias de corte y empalme tomadas del gen rep del plásmido de empaquetamiento AAV9 (U. Penn). Una secuencia poliA SV40 flanqueada por los sitios lox71 y lox66 invertidos se colocó aguas abajo de la cápside de AAV9. El plásmido rAAV-cap-in-cis-lox se modificó para introducir dos sitios de restricción únicos, XbaI y AgeI dentro de la secuencia de la cápside que flanquea la región que fue reemplazada por el fragmento de PCR. La inserción del sitio XbaI provocó una mutación K449R y las mutaciones necesarias para insertar el sitio AgeI fueron silenciosas.

[0190] En la segunda iteración de la construcción de la biblioteca (usada en la segunda selección de la cápside GFAP-Cre que produjo PHP.B), se hicieron dos modificaciones para reducir la contaminación de las bibliotecas por AAV9 o la cápside de partida en AAV9R X/A. Primero, la región de codificación entre los sitios XbaI y AgeI se eliminó en el plásmido utilizado para la clonación de la biblioteca de la cápside (aceptor de rAAV-Δcap-in-cis) para eliminar cualquier posible traspaso de plásmido no digerido. Segundo, el fragmento de PCR cubriendo la región variable de la biblioteca de cápside entre los sitios XbaI y AgeI se modificó para eliminar un sitio de restricción EarI (xE) único dentro de esta región de AAV9 e insertar un sitio KpnI único. El fragmento xE modificado se clonó TA en pCRII para generar pCRII-9Cap-xE, que sirvió como plantilla para nuestros fragmentos de PCR de biblioteca posteriores. La eliminación del sitio de EarI proporcionó una precaución secundaria que permitió la digestión de cualquier secuencia de AAV9 contaminante recuperada por PCR. No fue necesario el uso de esta etapa de digestión, ya que fue suficiente tomar precauciones de PCR estándar, incluyendo reactivos de tratamiento UV y pipetas y el uso del aceptor rAAV-Δcap-en-cis para la clonación de las bibliotecas, para evitar la contaminación de AAV9 o AAV9R X/A.

[0191] El plásmido auxiliar AAV2/9 REP-AAP se construyó mediante la introducción de 5 codones de parada en la secuencia codificante del marco de lectura VP del gen AAV9 en los AA 6, 10, 142, 148, 216 (numeración de VP1). El codón de parada en AA216 se diseñó de manera que no interrumpiera la secuencia de codificación de la proteína AAP, que se codifica dentro de un marco de lectura alternativo.

Generación de bibliotecas

[0192] Los fragmentos de la biblioteca 452-8r y 588i se generaron por PCR usando Q5 Hot Start, High Fidelity DNA Polymerase (NEB). Un esquema que muestra los sitios de unión del cebador aproximados y las secuencias del cebador se dan en las **FIGs**. **27A** y **27C,** respectivamente. Para facilitar la clonación de los fragmentos de PCR que comprenden las secuencias de la biblioteca de la cápside en un genoma de rAAV, el plásmido rAAV-cap-in-cis se modificó para introducir dos sitios de restricción únicos, XbaI y AgeI, dentro de la secuencia de la cápside que flanquea la región que fue reemplazada por el fragmento de la PCR. La inserción del sitio XbaI provocó una mutación K449R y las mutaciones necesarias para insertar el sitio AgeI fueron silenciosas. Para evitar la contaminación de las bibliotecas por la cápside AAV9R X/A "tipo salvaje", la región codificante entre los sitios XbaI y AgeI se eliminó del plásmido rAAV-cap-in-cis para crear rAAV-Cap-in-cis.

[0193] Para generar la biblioteca basada en rAAV, los productos de PCR que contienen la biblioteca y vector aceptor cap-en-cis digeridos por XbaI y AgeI fueron ensamblados utilizando Gibson Assembly. Los productos de reacción se trataron luego con PS DNasa (Epicentre) para eliminar cualquier fragmento sin ensamblar. Esta reacción generalmente produjo más de 100 ng de plásmido ensamblado (como se define por la cantidad de ADN que queda después de una etapa de digestión con PS DNasa). 100ng es suficiente para transfectar 10 platos de 150 mm a 10 ng/plato.

Producción y purificación de virus

[0194] AAV recombinantes fueron generados por transfección triple de células 293T utilizando PEI. Las partículas virales se recogieron de los medios a las 72 h después de la transfección y de las células y los medios a las 120 h. El virus presente en el medio se concentró por precipitación con poli(etilenglicol) al 8% y cloruro de sodio 500 mM (Ayuso et al., 2010) y luego el virus precipitado se añadió a los lisados preparados a partir de las células recolectadas. Los virus se purificaron en gradientes escalonados de iodixanol (Optiprep, Sigma) (15%, 25%, 40% y 60% según lo descrito por Zolotukhin et al 1999). Los virus fueron concentrados y formulados en PBS. Los títulos de virus se determinaron midiendo el número de copias del genoma del vector resistente a DNaseI (VG) utilizando qPCR y el plásmido del genoma linealizado como control (Gray et al 2011).

[0195] Para la generación de virus de la biblioteca de la cápside, se hicieron dos modificaciones en el protocolo de producción de virus para reducir la producción de cápsides de mosaico que podrían derivarse de la presencia de múltiples secuencias de la cápside en la misma célula. En primer lugar, solo se transfectaron 10 ng por placa del vector de biblioteca AAV-Cap9-en-cis por placa para asegurar que la gran mayoría de las células transfectadas solo recibieran una secuencia variante de cápside. En segundo lugar, el virus se recolectó antes (48 h y 60 horas, en lugar de 72 h y 120 h como anteriormente) para minimizar la transducción secundaria de las células productoras con el virus de biblioteca rAAV liberado en el medio.

Selección in vivo

[0196] Para las selecciones en ratones GFAP-Cre, $1x10^{11}$ vg de las bibliotecas de la cápside se inyectaron IV (ruta

retro-orbital) en ratones adultos Cre+. Siete u ocho días después de la inyección, los ratones se sometieron a eutanasia y se recogieron el cerebro y la médula espinal. El ADN del vector se recuperó de un hemisferio del cerebro y la mitad de la médula espinal utilizando 4-5 ml de Trizol (Invitrogen). Para las selecciones en ratones TH-Cre, $8x10^9$. Se inyectó vg de cada cápside intracranealmente utilizando las coordenadas estereotáxicas 0,7 mm rostral, 2,0 mm lateral y 3,0 mm ventral de bregma. 10 días después, se recogió la región que contenía la sustancia negra y se homogeneizó el tejido en 1 ml de Trizol. Para el aislamiento del ADN del virus, se siguió el protocolo de extracción de ARN del fabricante (se recogió la fracción acuosa superior que contiene el ARN). Además del ARN, se encontró que esta fracción también contiene una parte significativa del genoma viral, así como también algo de ADN mitocondrial. El ARN se eliminó tratando las muestras con 1 ul de RNasa (Qiagen) a 37°C durante la noche. La estrategia de PCR dependiente de la recombinación de Cre implicó una estrategia de amplificación de dos pasos **(FIG. 27A-27E).** La recuperación de la secuencia se realizó primero de manera dependiente de Cre utilizando los cebadores 9CAPF y CDF **(FIG. 27A-27E).** La PCR se realizó durante 20-26 ciclos de 95°C durante 20 segundos, 60°C durante 20 segundos y 72°C durante 30 segundos utilizando la polimerasa de ADN de alta fidelidad Hot Start Q5. Luego se diluyó el producto de la PCR 1:10-1:100 y luego se usa como plantilla para una segunda reacción de PCR usando el cebador XF y AR que generó un fragmento más corto que se clonó nuevamente en el constructo aceptor rAAV-delta-cap-in-cis como se describió anteriormente. Una ul de las reacciones de Gibson Assembly se diluyó 1:10 y se transformó en células competentes Sure2 (Agilent) según las indicaciones del fabricante para generar clones individuales para la secuenciación.

[0197] Los clones que mostraron evidencia de enriquecimiento se cortaron con BsiWI y AgeI y se ligaron en un ayudante de rep/cap personalizado 2/9R-X/A también se cortaron con BsiWI y AgeI y luego se transformaron en células competentes DH5alfa (NEB). Los plásmidos rep/cap resultantes que llevan las nuevas secuencias variantes, o AAV2/9 rep/cap como control, se usaron para empaquetar un genoma de rAAV que contiene un casete indicador doble eGFP-2A-luciferasa dirigido por un promotor CAG ubicuo (rAAV-CAG-eGFP-2A-Luc-WPRE-SV40pA) para las variantes inyectadas IV (AAV-PHP.A y AAV-PHP.B) o un vector similar que carece del gen Luc (rAAV-CAG-eGFP-WPRE-SV40pA) para las inyecciones intracraneales (AAV-PHP.R2).

Preparación e inmunotinción de tejidos.

[0198] Los ratones se anestesiaron con Nembutal y se perfundieron transcardialmente primero con tampón fosfato (PB) 0,1 M, pH 7,4 y luego con paraformaldehído al 4% recién preparado en PB. Los cerebros se fijaron posteriormente durante la noche y luego se seccionaron por vibratome o se crioprotegieron y se seccionaron por criostato. La inmunotinción se realizó en las secciones flotantes diluyendo los anticuerpos primarios y secundarios en PBS que contenía un 10% de suero de cabra o burro, un 0,5% de Triton X-100 o ningún detergente (tinción con GAD67). Los anticuerpos primarios utilizados fueron anti-GFP de conejo (1:1000; Invitrogen), anti-GFP de pollo (1:1000; Abcam), anti-CCl de ratón (1:200; Calbiochem), anti-GFAP de conejo (1:1000; Dako), anti-NeuN de ratón (1:500; Millipore), anti-IbaI de conejo (1:500; Biocare Medical), anti-Calbindin de ratón (1:200; Sigma), anti-Calretinin de conejo (1:1000; Chemicon), anti-GAD67 de ratón (1:1000; Millipore), Anti-Parvalbúmina de ratón (1:1000). Las incubaciones de anticuerpos primarios se realizaron durante 16-24 horas a temperatura ambiente. Las secciones se lavaron e incubaron con anticuerpos secundarios conjugados con Alexa 568 (1:1000; Invitrogen) durante 2-16 horas.

Limpieza de tejidos

[0199] Los ratones se perfundieron a través de una bomba peristáltica a través del ventrículo izquierdo con tampón fosfato (PB), seguido de una perfusión inicial con 60-80 ml de PFA al 4% en PB a un caudal de 14 ml por minuto. El caudal se redujo luego a 2-3 ml por minuto y se continuó durante 2 horas a temperatura ambiente. Luego, los ratones se colocaron en cámaras de perfusión individuales hechas a la medida y se perfundieron con 200 ml de acrilamida al 4% de reciclaje en el PB a la misma velocidad de flujo a temperatura ambiente durante la noche, seguido de una perfusión de 2 horas con PB para eliminar los polímeros/monómeros residuales La vasculatura. El proceso de polimerización se inició colocando las cámaras en un baño de agua a 42°C y entregando, por perfusión al mismo caudal, 200 ml de reciclaje, 0,25% de iniciador VA-044 desgasificado en PB. Una vez completada la polimerización, los ratones se perfundieron con una solución de eliminación de SDS al 8% en 0,1M PB, pH 7,5 durante 7 días. La solución que contenía SDS se cambió dos veces y luego se enjuagó con la perfusión de aproximadamente 2 l de PB no recirculante durante la noche. Las muestras de tejido depuradas de lípidos se incubaron en una solución de RIMS (Yang et al. 2014) hasta la obtención de imágenes (al menos una semana para la transparencia óptima de tejido cerebral de ratón no seccionado). Luego, las muestras se montaron en RIMS y se adjuntaron con un cubreobjetos en un portaobjetos utilizando un espaciador de grosor apropiado (iSpacer, SunJin Lab Co.). Las imágenes se tomaron con un microscopio de fotón único Zeiss LSM 780. Se realizaron reconstrucciones de imágenes tridimensionales utilizando el software de imágenes Imaris (Bit-plane).

Biodistribucion de vector

[0200] Los ratones fueron inyectados IV con $1x10^{11}$ VG de un vector rAAV-CAG-GFP2A-Luc-WPRE-SV40-pA empaquetado en las cápsides indicadas. 25 días después, los ratones se sometieron a eutanasia y los tejidos y las regiones cerebrales indicadas se recogieron y se congelaron a -80ºC. El ADN se aisló utilizando el kit Qiagen DNeasy Blood and Tissue. Los genomas vectoriales se detectaron usando cebadores de PCR que se unen al elemento WPRE y se normalizaron a los genomas de ratón usando cebadores específicos para el gen de glucagón de ratón. La

cuantificación absoluta se realizó comparando muestras desconocidas con diluciones en serie de estándares de concentración conocida.

EJEMPLO 9

MÉTODO DE TRATAMIENTO QUE EMPLEAN LAS PROTEÍNAS DIANA

[0201] Se identifica un sujeto que tiene un trastorno que puede ser tratado mediante la aplicación de un ácido nucleico a expresar dentro de un sujeto. A continuación, al sujeto se le administra una primera cantidad de un vector que incluye el polinucleótido a expresar. El polinucleótido codifica una proteína terapéutica. El vector incluirá una proteína de la cápside que incluye una sección de proteína dirigida que es la SEQ ID NO: 1, para permitir que la orientación adecuada de la proteína se exprese al sistema adecuado dentro del sujeto. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que se exprese una cantidad terapéuticamente eficaz de la proteína a expresar dentro del sujeto en el sistema apropiado.

EJEMPLO 10

MÉTODO DE TRATAMIENTO DE LA ENFERMEDAD DE HUNTINGTON

[0202] Se identifica un sujeto que tiene la enfermedad de Huntington. Luego, al sujeto se le administra una primera cantidad de un vector que incluye el polinucleótido a expresar. El polinucleótido codifica una proteína terapéutica. El vector incluirá una proteína de la cápside que incluye una sección de proteína dirigida que es la SEQ ID NO: 1, para permitir que la orientación adecuada de la proteína se exprese al sistema nervioso dentro del sujeto. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que se exprese una cantidad terapéuticamente eficaz de la proteína a expresar dentro del sujeto en el sistema nervioso.

EJEMPLO 11

MÉTODO DE TRATAMIENTO DE LA ENFERMEDAD DE HUNTINGTON

[0203] Se identifica un sujeto que tiene la enfermedad de Huntington. Luego, al sujeto se le administra una primera cantidad de un vector que incluye un polinucleótido que codifica un pequeño ARN no codificante (ARN de pequeña horquilla (ARN) (shARN) o microARN (miARN)) configurado para reducir la expresión de la proteína Huntingtin por su secuencia). El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1, a fin de permitir el direccionamiento adecuado de dicho polinucleótido al sistema nervioso. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva del pequeño ARN no codificante se exprese al sujeto en el sistema nervioso.

EJEMPLO 12

MÉTODO DE TRATAMIENTO

[0204] Se identifica un sujeto que tiene la enfermedad de Huntington. Luego, al sujeto se le administra sistémicamente una primera cantidad de un vector que incluye un polinucleótido que codifica para una proteína de dedo de zinc (ZFP) diseñada para reprimir la transcripción del gen de Huntingtin (HTT). El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir la orientación adecuada de la ZFP al sistema nervioso, entre otros órganos. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que se exprese al sujeto en el sistema nervioso una cantidad terapéuticamente efectiva de la ZFP.

EJEMPLO 13

MÉTODO DE TRATAMIENTO

[0205] Se identifica un sujeto que tiene la enfermedad de Huntington. Luego, al sujeto se le administra sistémicamente una primera cantidad de un vector que incluye un polinucleótido que codifica para un pequeño ARN no codificante (ARN de pequeña horquilla (shARN) o microARN (miARN)) diseñado por un experto en la técnica para reducir la expresión de Proteína de Huntingtin. El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir la orientación adecuada del polinucleótido al sistema nervioso, entre otros órganos. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva del pequeño ARN no codificante se exprese al sujeto en el sistema nervioso.

EJEMPLO 14

MÉTODO DE TRATAMIENTO

[0206] Se identifica un sujeto que tiene la enfermedad de Alzheimer. Luego, al sujeto se le administra una primera cantidad de un vector que incluye un polinucleótido que codifica un anticuerpo anti-Abeta o fragmentos de anticuerpo. El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir que la orientación adecuada del anticuerpo o fragmento de anticuerpo se exprese al sistema nervioso. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva del anticuerpo o fragmento de anticuerpo se exprese al sujeto en el sistema nervioso.

EJEMPLO 15

MÉTODO DE TRATAMIENTO

[0207] Se identifica un sujeto que tiene la enfermedad de Alzheimer. Luego, al sujeto se le administra una primera cantidad de un vector que incluye un polinucleótido que codifica una proteína apolipoproteína E (ApoE), preferiblemente el polipéptido apoE humano apoE2 o una variante modificada de apoE2. El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir que la orientación adecuada del anticuerpo o fragmento de anticuerpo se exprese al sistema nervioso. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva de la proteína ApoE se exprese al sujeto en el sistema nervioso.

EJEMPLO 16

MÉTODO DE TRATAMIENTO

[0208] Se identifica un sujeto que tiene la atrofia muscular espinal (SMA). Luego, al sujeto se le administra una primera cantidad de un vector que incluye un polinucleótido que codifica un polipéptido de neurona motora de supervivencia 1 (SMN1). El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir que la orientación adecuada de la proteína SMN se exprese al sistema nervioso. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva de la proteína SMN se exprese al sujeto en el sistema nervioso.

EJEMPLO 17

MÉTODO DE TRATAMIENTO

[0209] Se identifica un sujeto que tiene ataxia de Friedreich. Luego, al sujeto se le administra sistémicamente una primera cantidad de un vector que incluye un polinucleótido que codifica una proteína frataxina. El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, con el fin de permitir que se dirija adecuadamente la proteína frataxina al sistema nervioso y al corazón, entre otros órganos. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva de la proteína frataxina se exprese al sujeto en el sistema nervioso y el corazón.

EJEMPLO 18

MÉTODO DE TRATAMIENTO

[0210] Se identifica un sujeto que tiene la enfermedad de Pompe. Luego, al sujeto se le administra sistémicamente una primera cantidad de un vector que incluye un polinucleótido que codifica una proteína ácida alfa-glucosidasa (GAA). El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir que la orientación adecuada de la proteína GAA se exprese al sistema nervioso y al corazón, entre otros órganos. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva de la proteína GAA se exprese al sujeto en el sistema nervioso y el corazón.

EJEMPLO 19

MÉTODO DE TRATAMIENTO

[0211] Se identifica un sujeto que tiene lipofuscinosis ceroide neuronal infantil tardía (LINCL). Luego, al sujeto se le administra sistemáticamente una primera cantidad de un vector que incluye un polinucleótido CLN2 que codifica la proteína tripeptidil peptidasa 1. El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, con el fin de permitir que la diana adecuada de la proteína tripeptidil peptidasa 1 se exprese al sistema nervioso. Si es necesario,

al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva de la proteína tripeptidil peptidasa 1 se exprese al sujeto en el sistema nervioso.

EJEMPLO 20

MÉTODO DE TRATAMIENTO

[0212] Se identifica un sujeto que tiene la forma juvenil NCL de la enfermedad de Batten. El sujeto luego se administra sistémicamente una primera cantidad de un vector que incluye un polinucleótido CLN3 que codifica para la proteína battenina. El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir que se dirija al sistema nervioso la orientación adecuada de la proteína battenina. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente eficaz de la proteína battenina se exprese al sujeto en el sistema nervioso.

EJEMPLO 21

MÉTODO DE TRATAMIENTO

[0213] Se identifica un sujeto que tiene enfermedad de Canavan. Luego, al sujeto se le administra sistemáticamente una primera cantidad de un vector que incluye un polinucleótido ASPA que codifica la proteína aspartoacilasa. El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir que se dirija al sistema nervioso la orientación adecuada de la proteína aspartoacilasa. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que se exprese al sujeto en el sistema nervioso una cantidad terapéuticamente efectiva de la proteína aspartoaculasa.

EJEMPLO 22

MÉTODO DE TRATAMIENTO

[0214] Se identifica un sujeto que tiene la enfermedad de Parkinson. A continuación, al sujeto se le administra sistémicamente una primera cantidad de uno o más vectores que cada uno incluye uno o más polinucleótidos que codifican una enzima o enzimas necesarias para el aumento de la producción de dopamina a partir de células no dopaminérgicas. El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir que la orientación adecuada de dicha enzima(s) se exprese al sistema nervioso. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva de la(s) enzima(s) se exprese por el sujeto en el sistema nervioso.

EJEMPLO 23

MÉTODO DE TRATAMIENTO

[0215] Se identifica un sujeto que tiene enfermedad de Parkinson. A continuación, al sujeto se le administra sistemáticamente una primera cantidad de un vector que incluye un polinucleótido que codifica una forma modificada, resistente a la agregación, de la proteína alfa-sinucleína que reduce la agregación de la alfa-sinucleína endógena. El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir que se dirija al sistema nervioso la orientación adecuada de la proteína alfa-sinucleína resistente a la agregación. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva de la proteína se exprese al sujeto en el sistema nervioso.

EJEMPLO 24

MÉTODO DE TRATAMIENTO

[0216] Se identifica un sujeto que tiene esclerosis lateral amiotrófica o demencia frontal causada por una mutación en C9ORF72. Luego, al sujeto se le administra una primera cantidad de un vector que incluye un polinucleótido que codifica un ARN o ARN no codificantes que reducen los focos de ARN nuclear causados por la expansión del hexanucleótido (GGGGCC) en las células de los sujetos. El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, con el fin de permitir que la orientación adecuada de los ARN se exprese al sistema nervioso. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que se exprese una cantidad terapéuticamente efectiva de los ARN en el sistema nervioso.

EJEMPLO 25

MÉTODO DE TRATAMIENTO

[0217] Se identifica un sujeto que tiene esclerosis múltiple. A continuación, al sujeto se le administra sistémicamente una primera cantidad de un vector que incluye un polinucleótido que codifica un factor trófico o inmunomodulador, por ejemplo el factor inhibidor de la leucemia (LIF) o el factor eurotrófico ciliar (CNTF). El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir que la orientación adecuada de dicho factor se exprese al sistema nervioso. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva del factor se exprese al sujeto en el sistema nervioso.

EJEMPLO 26

MÉTODO DE TRATAMIENTO

[0218] Se identifica un sujeto que tiene esclerosis lateral amiotrófica causada por la mutación SOD1. Luego, al sujeto se le administra una primera cantidad de un vector que incluye un polinucleótido que codifica un pequeño ARN no codificante (ARN de pequeña horquilla (ARN) o microARN (miARN)) diseñado por un experto en la técnica para reducir la expresión de SOD1 mutante proteína. El vector incluirá una proteína de la cápside que incluye una proteína de direccionamiento de la SEQ ID NO: 1 o cualquiera de las proteínas de direccionamiento en la FIG. 31, a fin de permitir que la orientación adecuada del pequeño ARN no codificante se exprese al sistema nervioso. Si es necesario, al sujeto se le administra una segunda o tercera dosis del vector, hasta que una cantidad terapéuticamente efectiva del pequeño ARN no codificante se exprese al sujeto en el sistema nervioso.

Revelación adicional

[0219] En este documento se describe una plataforma de evolución dirigida de AAV basada en Recombinasa basada en CREATE.

[0220] En el presente documento se describe un AAV-PHP.B, que permite el suministro del gen generalizado a las neuronas del SNC y la glía a través de la vasculatura.

[0221] En el presente documento se describe un AAV-PHP.R2, que permite la transducción de Retrógado Rápido en el SNC.

Equivalentes

[0222] La especificación escrita anterior se considera que es suficiente para permitir a un experto en la técnica poner en práctica la invención. La descripción y los ejemplos anteriores detallan ciertas realizaciones preferidas de la invención y describen el mejor modo contemplado por los inventores. Sin embargo, se apreciará que, sin importar lo detallado que pueda aparecer el texto anterior, la invención puede ponerse en práctica de muchas maneras y la invención debe interpretarse de acuerdo con las reivindicaciones adjuntas y sus equivalentes.

LISTADO DE SECUENCIAS

[0223]

<110> Benjamin E. Deverman Paul H. Patterson

<120> RECUPERACIÓN SELECTIVA

<130> CALTE.103WO

<140> Desconocido
<141> Con el mismo

<150> US 61/877506
<151> 2013-09-13

<150> US 61/983624
<151> 2014-04-24

<150> US 62/020658
<151> 2014-07-03

<150> US 62/034060
<151> 2014-08-06

<160> 60

<170> PatentIn versión 3.1

<210> 1
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Proteína de cápside de virus adeno-asociado

<400> 1

Thr Leu Ala Val Pro Phe Lys
1 5

<210> 2
<211> 736
<212> PRT
<213> Secuencia de cápside AAV9

<220>
<223> Virus adeno-asociado

<400> 2

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
35 40 45

```
Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
    50                  55                  60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65                  70                  75                  80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
        115                 120                 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145                 150                 155                 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
            180                 185                 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
        195                 200                 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
    210                 215                 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225                 230                 235                 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
                245                 250                 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
            260                 265                 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
        275                 280                 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
    290                 295                 300
```

```
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305                 310                 315                 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
                325                 330                 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
            340                 345                 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
        355                 360                 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
    370                 375                 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385                 390                 395                 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
                405                 410                 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
            420                 425                 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
        435                 440                 445

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
    450                 455                 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465                 470                 475                 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
                485                 490                 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
            500                 505                 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
        515                 520                 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
    530                 535                 540
```

```
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545                 550                 555                 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
                565                 570                 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
            580                 585                 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
        595                 600                 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
    610                 615                 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625                 630                 635                 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
                645                 650                 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
            660                 665                 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
        675                 680                 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
    690                 695                 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705                 710                 715                 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
                725                 730                 735
```

<210> 3
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Proteína de cápside de virus adeno-asociado

<400> 3

```
Lys Phe Pro Val Ala Leu Thr
1               5
```

<210> 4

<211> 6672
<212> ADN
<213> Secuencia artificial

<220>
<223> Plásmido

<400> 4

```
ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc       60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg      120
gccaactcca tcactagggg ttcctactag tggcctccgc gccgggtttt ggcgcctccc      180
gcgggcgccc ccctcctcac ggcgagcgct gccacgtcag acgaagggcg cagcgagcgt      240
cctgatcctt ccgcccggac gctcaggaca gcggcccgct gctcataaga ctcggcctta      300
gaaccccagt atcagcagaa ggacatttta ggacgggact tgggtgactc tagggcactg      360
gttttctttc cagagagcgg aacaggcgag gaaaagtagt cccttctcgg cgattctgcg      420
gagggatctc cgtggggcgg tgaacgccga tgattatata aggacgcgcc gggtgtggca      480
cagctagttc cgtcgcagcc gggatttggg tcgcggttct tgtttgtgga tcgctgtgat      540
cgtcacttgg cggccgccat ggtcagcaag ggcgaggagg ataacatggc catcatcaag      600
gagttcatgc gcttcaaggt gcacatggag ggctccgtga acggccacga gttcgagatc      660
gagggcgagg gcgagggccg cccctacgag ggcacccaga ccgccaagct gaaggtgacc      720
aagggtggcc ccctgccctt cgcctgggac atcctgtccc ctcaattcat gtatggctcc      780
aaggcctacg tgaagcaccc cgccgacatc cccgactact tgaagctgtc cttccccgag      840
ggcttcaagt gggagcgcgt gatgaacttc gaggacggcg gcgtggtgac cgtgacccag      900
gactcctcct tacaagacgg cgagttcatc tacaaagtga agctgcgcgg caccaacttc      960
ccctccgacg gccccgtaat gcagaagaag accatgggct gggaggcctc ctccgagcgg     1020
atgtaccccg aggacggcgc cctgaagggc gagatcaagc agaggctgaa gctgaaggac     1080
ggcggccact acgacgctga ggtcaagacc acctacaagg ccaagaagcc cgtgcagctg     1140
cccggcgcct acaacgtcaa catcaagttg gacatcacct cccacaacga ggactacacc     1200
atcgtggaac agtacgaacg cgccgagggc cgccactcca ccggcggcat ggacgagctg     1260
tacaagtaaa ggatcccctc ccccgtgcct tccttgaccc tggaaggtgc cactcccact     1320
gtcctttcct cctcctagga ataaaatatc tttattttca ttacatctgt gtgttggttt     1380
tttgtgtaga tctgttcaaa tttgaactga ctaagcggct cccgccagat tttggcaaga     1440
ttactaagca ggaagtcaag gacttttttg cttgggcaaa ggtcaatcag gtgccggtga     1500
ctcacgagtt taaagttccc agggaattgg cgggaactaa aggggcggag aaatctctaa     1560
aacgcccact gggtgacgtc accaatacta gctataaaag tctggagaag cgggccaggc     1620
tctcatttgt tcccgagacg cctcgcagtt cagacgtgac tgttgatccc gctcctctgc     1680
gaccgctagc ttcgatcaac tacgcggaca ggtaccaaaa caaatgttct cgtcacgtgg     1740
gcatgaatct gatgctgttt ccctgcagac aatgcgagag actgaatcag aattcaaata     1800
tctgcttcac tcacggtgtc aaagactgtt tagagtgctt tcccgtgtca gaatctcaac     1860
ccgtttctgt cgtcaaaaag gcgtatcaga aactgtgcta cattcatcac atcatgggaa     1920
aggtgccaga cgcttgcact gcttgcgacc tggtcaatgt ggacttggat gactgtgttt     1980
ctgaacaata aatgacttaa accaggtatg gctgccgatg gttatcttcc agattggctc     2040
gaggacaacc ttagtgaagg aattcgcgag tggtgggctt tgaaacctgg agcccctcaa     2100
cccaaggcaa atcaacaaca tcaagacaac gctcgaggtc ttgtgcttcc gggttacaaa     2160
taccttggac ccggcaacgg actcgacaag ggggagccgg tcaacgcagc agacgcggcg     2220
gccctcgagc acgacaaggc ctacgaccag cagctcaagg ccggagacaa cccgtacctc     2280
aagtacaacc acgccgacgc cgagttccag gagcggctca aagaagatac gtcttttggg     2340
ggcaacctcg ggcgagcagt cttccaggcc aaaaagaggc ttcttgaacc tcttggtctg     2400
gttgaggaag cggctaagac ggctcctgga aagaagaggc ctgtagagca gtctcctcag     2460
gaaccggact cctccgcggg tattggcaaa tcgggtgcac agcccgctaa aaagagactc     2520
aatttcggtc agactggcga cacagagtca gtcccagacc ctcaaccaat cggagaacct     2580
cccgcagccc cctcaggtgt gggatctctt acaatggctt caggtggtgg cgcaccagtg     2640
gcagacaata acgaaggtgc cgatggagtg ggtagttcct cgggaaattg gcattgcgat     2700
tcccaatggc tgggggacag agtcatcacc accagcaccc gaacctgggc cctgcccacc     2760
tacaacaatc acctctacaa gcaaatctcc aacagcacat ctggaggatc ttcaaatgac     2820
aacgcctact tcggctacag cacccccctgg gggtattttg acttcaacag attccactgc     2880
cacttctcac cacgtgactg gcagcgactc atcaacaaca actggggatt ccggcctaag     2940
cgactcaact tcaagctctt caacattcag gtcaaagagg ttacggacaa caatggagtc     3000
aagaccatcg ccaataacct taccagcacg gtccaggtct tcacggactc agactatcag     3060
ctcccgtacg tgctcgggtc ggctcacgag ggctgcctcc cgccgttccc agcggacgtt     3120
ttcatgattc ctcagtacgg gtatctgacg cttaatgatg gaagccaggc cgtgggtcgt     3180
tcgtcctttt actgcctgga atatttcccg tcgcaaatgc taagaacggg taacaacttc     3240
cagttcagct acgagtttga gaacgtacct ttccatagca gctacgctca cagccaaagc     3300
ctggaccgac taatgaatcc actcatcgac caatacttgt actatctctc tagaactatt     3360
```

```
aacggttctg gacagaatca acaaacgcta aaattcagtg tggccggacc cagcaacatg   3420
gctgtccagg gaagaaacta catacctgga cccagctacc gacaacaacg tgtctcaacc   3480
actgtgactc aaaacaacaa cagcgaattt gcttggcctg gagcttcttc ttgggctctc   3540
aatggacgta atagcttgat gaatcctgga cctgctatgg ccagccacaa agaaggagag   3600
gaccgtttct ttcctttgtc tggatcttta atttttggca aacaaggaac tggaagagac   3660
aacgtggatg cggacaaagt catgataacc aacgaagaag aaattaaaac tactaacccg   3720
gtagcaacgg agtcctatgg acaagtggcc acaaaccacc agagtgccca agcacaggcg   3780
cagaccggtt gggttcaaaa ccaaggaata cttccgggta tggtttggca ggacagagat   3840
gtgtacctgc aaggacccat ttgggccaaa attcctcaca cggacggcaa ctttcaccct   3900
tctccgctga tgggagggtt tggaatgaag cacccgcctc ctcagatcct catcaaaaac   3960
acacctgtac ctgcggatcc tccaacggcc ttcaacaagg acaagctgaa ctctttcatc   4020
acccagtatt ctactggcca agtcagcgtg gagatcgagt gggagctgca gaaggaaaac   4080
agcaagcgct ggaacccgga gatccagtac acttccaact attacaagtc taataatgtt   4140
gaatttgctg ttaatactga aggtgtatat agtgaacccc gccccattgg caccagatac   4200
ctgactcgta atctgtaagt cgactaccgt tcgtatagca tacattatac gaagttatca   4260
tatgttcgag cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg   4320
cagtgaaaaa aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt   4380
ataagctgca ataaacaagt taacaacaac aattgcattc attttatgtt tcaggttcag   4440
ggggagatgt gggaggtttt ttaaagcaag taaaacctct acaaatgtgg taaaatcgag   4500
ctctaccgtt cgtataatgt atgctatacg aagttatgat atcaagctta ggaaccccta   4560
gtgatggagt tggccactcc ctctctgcgc gctcgctcgc tcactgaggc cgggcgacca   4620
aaggtcgccc gacgcccggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgcaga   4680
gagggagtgg ccaagctagc gggcgattaa ggaaagggct agatcattct tgaagacgaa   4740
agggcctcgt gatacgccta tttttatagg ttaatgtcat gataataatg gtttcttaga   4800
cgtcaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta ttttctaaa   4860
tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt   4920
gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg   4980
catttgcct tcctgttttt gctcacccag aaacgctggt gaaagtaaaa gatgctgaag   5040
atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggt aagatccttg   5100
agagttttcg ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg   5160
gcgcggtatt atcccgtgtt gacgccgggc aagagcaact cggtcgccgc atacactatt   5220
ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga   5280
cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaacttac   5340
ttctgacaac gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc   5400
atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc   5460
gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac   5520
tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag   5580
gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg   5640
gtgagcgtgg gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta   5700
tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg   5760
ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata   5820
tactttagat tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt   5880
ttgataatct catgaccaaa atcccttaac gtgagttttc gttccactga gcgtcagacc   5940
ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct   6000
tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa   6060
ctcttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gttcttctag   6120
tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc   6180
tgctaatcct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg   6240
actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca   6300
cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat   6360
gagaaagcgc cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg   6420
tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc   6480
ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc   6540
ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc   6600
cttttgctca catgtaataa acacacacac accaacaacc gtggttggtt gttgtgttgg   6660
tttattctcg ag                                                      6672
```

<210> 5
<211> 7330
<212> ADN
<213> Secuencia artificial

<220>
<223> Plásmido

<400> 5

```
gtcgacggta tcgggggagc tcgcagggtc tccattttga agcgggaggt ttgaacgcgc      60
agccgccatg ccggggtttt acgagattgt gattaaggtc cccagcgacc ttgacgagca     120
tctgcccggc atttctgaca gctttgtgaa ctgggtggcc gagaaggaat gggagttgcc     180
gccagattct gacatggatc tgaatctgat tgagcaggca cccctgaccg tggccgagaa     240
gctgcagcgc gactttctga cggaatggcg ccgtgtgagt aaggccccgg aggctctttt     300
ctttgtgcaa tttgagaagg gagagagcta cttccacatg cacgtgctcg tggaaaccac     360
cggggtgaaa tccatggttt tgggacgttt cctgagtcag attcgcgaaa aactgattca     420
gagaatttac cgcgggatcg agccgacttt gccaaactgg ttcgcggtca caaagaccag     480
aaatggcgcc ggaggcggga acaaggtggt ggatgagtgc tacatcccca attacttgct     540
ccccaaaacc cagcctgagc tccagtgggc gtggactaat atggaacagt atttaagcgc     600
ctgtttgaat ctcacggagc gtaaacggtt ggtggcgcag catctgacgc acgtgtcgca     660
gacgcaggag cagaacaaag agaatcagaa tcccaattct gatgcgccgg tgatcagatc     720
aaaaacttca gccaggtaca tggagctggt cgggtggctc gtggacaagg ggattacctc     780
ggagaagcag tggatccagg aggaccaggc ctcatacatc tccttcaatg cggcctccaa     840
ctcgcggtcc caaatcaagg ctgccttgga caatgcggga aagattatga gcctgactaa     900
aaccgccccc gactacctgg tgggccagca gcccgtggag gacatttcca gcaatcggat     960
ttataaaatt ttggaactaa acgggtacga tccccaatat gcggcttccg tctttctggg    1020
atgggccacg aaaaagttcg gcaagaggaa caccatctgg ctgtttgggc ctgcaactac    1080
cgggaagacc aacatcgcgg aggccatagc ccacactgtg cccttctacg ggtgcgtaaa    1140
ctggaccaat gagaactttc ccttcaacga ctgtgtcgac aagatggtga tctggtggga    1200
ggaggggaag atgaccgcca aggtcgtgga gtcggccaaa gccattctcg gaggaagcaa    1260
ggtgcgcgtg gaccagaaat gcaagtcctc ggcccagata gacccgactc ccgtgatcgt    1320
cacctccaac accaacatgt gcgccgtgat tgacgggaac tcaacgacct tcgaacacca    1380
gcagccgttg caagaccgga tgttcaaatt tgaactcacc cgccgtctgg atcatgactt    1440
tgggaaggtc accaagcagg aagtcaaaga ctttttccgg tgggcaaagg atcacgtggt    1500
tgaggtggag catgaattct acgtcaaaaa gggtggagcc aagaaaagac ccgccccag     1560
tgacgcagat ataagtgagc ccaaacgggt gcgcgagtca gttgcgcagc catcgacgtc    1620
agacgcggaa gcttcgatca actacgcgga caggtaccaa aacaaatgtt ctcgtcacgt    1680
gggcatgaat ctgatgctgt ttccctgcag acaatgcgag agactgaatc agaattcaaa    1740
tatctgcttc actcacggtg tcaaagactg tttagagtgc tttcccgtgt cagaatctca    1800
acccgtttct gtcgtcaaaa aggcgtatca gaaactgtgc tacattcatc acatcatggg    1860
aaaggtgcca gacgcttgca ctgcttgcga cctggtcaat gtggacttgg atgactgtgt    1920
ttctgaacaa taaatgactt aaaccaggta tggctgccga tggttaactt ccagattgac    1980
tcgaggacaa ccttagtgaa ggaattcgcg agtggtgggc tttgaaacct ggagcccctc    2040
aacccaaggc aaatcaacaa catcaagaca acgctcgagg tcttgtgctt ccgggttaca    2100
aataccttgg acccggcaac ggactcgaca agggggagcc ggtcaacgca gcagacgcgg    2160
cggccctcga gcacgacaag gcctacgacc agcagctcaa ggccggagac aacccgtacc    2220
tcaagtacaa ccacgccgac gccgagttcc aggagcggct caaagaagat acgtcttttg    2280
ggggcaacct cgggcgagca gtcttccagg ccaaaaagag gcttcttgaa cctcttggtc    2340
tggttgagga agcggctaag acggctcctg gatagaagag gcctgtagag tagtctcctc    2400
aggaaccgga ctcctccgcg ggtattggca aatcgggtgc acagcccgct aaaaagagac    2460
tcaatttcgg tcagactggc gacacagagt cagtcccaga ccctcaacca atcggagaac    2520
ctcccgcagc cccctcaggt gtgggatctc ttacaatggc ttcaggtggt ggcgcaccag    2580
tggcagacaa taactaaggt gccgatggag tgggtagttc ctcgggaaat tggcattgcg    2640
attcccaatg gctgggggac agagtcatca ccaccagcac ccgaacctgg gccctgccca    2700
cctacaacaa tcacctctac aagcaaatct ccaacagcac atctggagga tcttcaaatg    2760
acaacgccta cttcggctac agcacccccc gggggtattt tgacttcaac agattccact    2820
gccacttctc accacgtgac tggcagcgac tcatcaacaa caactgggga ttccggccta    2880
agcgactcaa cttcaagctc ttcaacattc aggtcaaaga ggttacggac aacaatggag    2940
tcaagaccat cgccaataac cttaccagca cggtccaggt cttcacggac tcagactatc    3000
agctcccgta cgtgctcggg tcggctcacg agggctgcct cccgccgttc ccagcggacg    3060
ttttcatgat tcctcagtac gggtatctga cgcttaatga tggaagccag gccgtgggtc    3120
gttcgtcctt ttactgcctg gaatatttcc cgtcgcaaat gctaagaacg ggtaacaact    3180
tccagttcag ctacgagttt gagaacgtac ctttccatag cagctacgct cacagccaaa    3240
gcctggaccg actaatgaat ccactcatcg accaatactt gtactatctc tcaaagacta    3300
ttaacggttc tggacagaat caacaaacgc taaaattcag tgtggccgga cccagcaaca    3360
tggctgtcca gggaagaaac tacatacctg gacccagcta ccgacaacaa cgtgtctcaa    3420
ccactgtgac tcaaaacaac aacagcgaat ttgcttggcc tggagcttct tcttgggctc    3480
```

```
tcaatggacg taatagcttg atgaatcctg gacctgctat ggccagccac aaagaaggag     3540
aggaccgttt ctttcctttg tctggatctt taatttttgg caaacaagga actggaagag     3600
acaacgtgga tgcggacaaa gtcatgataa ccaacgaaga agaaattaaa actactaacc     3660
cggtagcaac ggagtcctat ggacaagtgg ccacaaacca ccagagtgcc caagcacagg     3720
cgcagaccgg ctgggttcaa aaccaaggaa tacttccggg tatggtttgg caggacagag     3780
atgtgtacct gcaaggaccc atttgggcca aaattcctca cacggacggc aactttcacc     3840
cttctccgct gatgggaggg tttggaatga agcacccgcc tcctcagatc ctcatcaaaa     3900
acacacctgt acctgcggat cctccaacgg ccttcaacaa ggacaagctg aactctttca     3960
tcacccagta ttctactggc caagtcagcg tggagatcga gtgggagctg cagaaggaaa     4020
acagcaagcg ctggaacccg gagatccagt acacttccaa ctattacaag tctaataatg     4080
ttgaatttgc tgttaatact gaaggtgtat atagtgaacc ccgccccatt ggcaccagat     4140
acctgactcg taatctgtaa ttgcttgtta atcaataaac cgtttaattc gtttcagttg     4200
aactttggtc tctgcgaagg gcgaattcgt ttaaacctgc aggactagag gtcctgtatt     4260
agaggtcacg tgagtgtttt gcgacatttt gcgacaccat gtggtcacgc tgggtattta     4320
agcccgagtg agcacgcagg gtctccattt tgaagcggga ggtttgaacg cgcagccgcc     4380
aagccgaatt ctgcagatat ccatcacact ggcggccgct cgactagagc ggccgccacc     4440
gcggtggagc tccagctttt gttcccttta gtgagggtta attgcgcgct tggcgtaatc     4500
atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg     4560
agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac tcacattaat     4620
tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc tgcattaatg     4680
aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgct     4740
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc     4800
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg     4860
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgttttcc ataggctccg     4920
ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg     4980
actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc ctgttccgac     5040
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca     5100
tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt     5160
gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc     5220
caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag     5280
agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac     5340
tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt     5400
tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa     5460
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg     5520
gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa     5580
aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat     5640
atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc     5700
gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat     5760
acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc     5820
ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtggtcc     5880
tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag     5940
ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg tggtgtcacg     6000
ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg     6060
atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg ttgtcagaag     6120
taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt     6180
catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt cattctgaga     6240
atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata ataccgcgcc     6300
acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc     6360
aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac ccaactgatc     6420
ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc     6480
cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct tccttttca      6540
atattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat     6600
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctaaatt     6660
gtaagcgtta atattttgtt aaaattcgcg ttaaattttt gttaaatcag ctcatttttt     6720
aaccaatagg ccgaaatcgg caaaatccct tataaatcaa aagaatagac cgagataggg     6780
ttgagtgttg ttccagtttg gaacaagagt ccactattaa agaacgtgga ctccaacgtc     6840
aaagggcgaa aaaccgtcta tcagggcgat ggcccactac gtgaaccatc accctaatca     6900
agtttttgg ggtcgaggtg ccgtaaagca ctaaatcgga accctaaagg gagcccccga      6960
tttagagctt gacggggaaa gccggcgaac gtggcgagaa aggaagggaa gaaagcgaaa     7020
ggagcgggcg ctagggcgct ggcaagtgta gcggtcacgc tgcgcgtaac caccacaccc     7080
gccgcgctta atgcgccgct acagggcgcg tcccattcgc cattcaggct gcgcaactgt     7140
tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggatgt      7200
gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg     7260
```

acggccagtg agcgcgcgta atacgactca ctatagggcg aattgggtac cgggcccccc 7320
ctcgatcgag 7330

<210> 6
<211> 4468
<212> ADN
<213> Secuencia artificial

<220>
<223> Plásmido

<400> 6

```
agcgcccaat acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc       60
acgacaggtt tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc      120
tcactcatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa      180
ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac gccaagctat      240
ttaggtgaca ctatagaata ctcaagctat gcatcaagct tggtaccgag ctcggatcca      300
ctagtaacgg ccgccagtgt gctggaattc gcccttactc atcgaccaat acttgtacta      360
tctctctaga actattaacg gttctggaca gaatcaacaa acgctaaaat tcagtgtggc      420
cggacccagc aacatggctg tccagggaag aaactacata cctggaccca gctaccgaca      480
acaacgtgtc tcaaccactg tgactcaaaa caacaacagc gaatttgctt ggcctggagc      540
ttcttcttgg gctctcaatg gacgtaatag cttgatgaat cctggacctg ctatggccag      600
ccacaaagaa ggagaggacc gtttctttcc tttgtctgga tctttaattt ttggcaaaca      660
aggtaccggc agagacaacg tggatgcgga caaagtcatg ataaccaacg aagaagaaat      720
taaaactact aacccggtag caacggagtc ctatggacaa gtggccacaa accaccagag      780
tgcccaagca caggcgcaga ccggttgggt tcaaaaccaa ggaatacttc caagggcgaa      840
ttctgcagat atccatcaca ctggcggccg ctcgagcatg catctagagg gcccaattcg      900
ccctatagtg agtcgtatta caattcactg gccgtcgttt tacaacgtcg tgactgggaa      960
aaccctggcg ttacccaact taatcgcctt gcagcacatc ccccttcgcc agctggcgt     1020
aatagcgaag aggcccgcac cgatcgccct tcccaacagt tgcgcagcct gaatggcgaa     1080
tggacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg cgcagcgtga     1140
ccgctacact tgccagcgcc ctagcgcccg ctcctttcgc tttcttccct tcctttctcg     1200
ccacgttcgc cggctttccc cgtcaagctc taaatcgggg gctcccttta gggttccgat     1260
ttagtgcttt acggcacctc gaccccaaaa aacttgatta gggtgatggt tcacgtagtg     1320
ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg ttctttaata     1380
gtggactctt gttccaaact ggaacaacac tcaaccctat ctcggtctat tcttttgatt     1440
tataagggat tttgccgatt tcggcctatt ggttaaaaaa tgagctgatt taacaaaaat     1500
ttaacgcgaa ttttaacaaa attcagggcg caagggctgc taaaggaagc ggaacacgta     1560
gaaagccagt ccgcagaaac ggtgctgacc ccggatgaat gtcagctact gggctatctg     1620
gacaagggaa aacgcaagcg caaagagaaa gcaggtagct tgcagtgggc ttacatggcg     1680
atagctagac tgggcggttt tatggacagc aagcgaaccg gaattgccag ctggggcgcc     1740
ctctggtaag gttgggaagc cctgcaaagt aaactggatg gctttcttgc cgccaaggat     1800
ctgatggcgc aggggatcaa gatctgatca agagacagga tgaggatcgt ttcgcatgat     1860
tgaacaagat ggattgcacg caggttctcc ggccgcttgg gtggagaggc tattcggcta     1920
tgactgggca caacagacaa tcggctgctc tgatgccgcc gtgttccggc tgtcagcgca     1980
ggggcgcccg gttctttttg tcaagaccga cctgtccggt gccctgaatg aactgcagga     2040
cgaggcagcg cggctatcgt ggctggccac gacgggcgtt ccttgcgcag ctgtgctcga     2100
cgttgtcact gaagcgggaa gggactggct gctattgggc gaagtgccgg ggcaggatct     2160
cctgtcatcc caccttgctc ctgccgagaa agtatccatc atggctgatg caatgcggcg     2220
gctgcatacg cttgatccgg ctacctgccc attcgaccac caagcgaaac atcgcatcga     2280
gcgagcacgt actcggatgg aagccggtct tgtcgatcag gatgatctgg acgaagagca     2340
tcaggggctc gcgccagccg aactgttcgc caggctcaag gcgcgcatgc ccgacggcga     2400
ggatctcgtc gtgacccatg gcgatgcctg cttgccgaat atcatggtgg aaaatggccg     2460
cttttctgga ttcatcgact gtggccggct gggtgtggcg gaccgctatc aggacatagc     2520
gttggctacc cgtgatattg ctgaagagct tggcggcgaa tgggctgacc gcttcctcgt     2580
gctttacggt atcgccgctc ccgattcgca gcgcatcgcc ttctatcgcc ttcttgacga     2640
gttcttctga attgaaaaag gaagagtatg agtattcaac atttccgtgt cgcccttatt     2700
ccctttttg cggcattttg ccttcctgtt tttgctcacc cagaaacgct ggtgaaagta     2760
aaagatgctg aagatcagtt gggtgcacga gtgggttaca tcgaactgga tctcaacagc     2820
ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc caatgatgag cacttttaaa     2880
gttctgctat gtggcgcggt attatcccgt attgacgccg ggcaagagca actcggtcgc     2940
```

```
cgcatacact attctcagaa tgacttggtt gagtactcac cagtcacaga aaagcatctt     3000
acggatggca tgacagtaag agaattatgc agtgctgcca taaccatgag tgataacact     3060
gcggccaact tacttctgac aacgatcgga ggaccgaagg agctaaccgc ttttttgcac     3120
aacatggggg atcatgtaac tcgccttgat cgttgggaac cggagctgaa tgaagccata     3180
ccaaacgacg agcgtgacac cacgatgcct gtagcaatgg caacaacgtt gcgcaaacta     3240
ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg     3300
gataaagttg caggaccact tctgcgctcg gcccttccgg ctggctggtt tattgctgat     3360
aaatctggag ccggtgagcg tgggtctcgc ggtatcattg cagcactggg gccagatggt     3420
aagccctccc gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga     3480
aatagacaga tcgctgagat aggtgcctca ctgattaagc attggtaact gtcagaccaa     3540
gtttactcat atatacttta gattgattta aaacttcatt tttaatttaa aaggatctag     3600
gtgaagatcc tttttgataa tctcatgacc aaaatccctt aacgtgagtt ttcgttccac     3660
tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt ttttctgcgc     3720
gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cggtggtttg tttgccggat     3780
caagagctac caactctttt tccgaaggta actggcttca gcagagcgca gataccaaat     3840
actgttcttc tagtgtagcc gtagttaggc caccacttca agaactctgt agcaccgcct     3900
acatacctcg ctctgctaat cctgttacca gtggctgctg ccagtggcga taagtcgtgt     3960
cttaccgggt tggactcaag acgatagtta ccggataagg cgcagcggtc gggctgaacg     4020
gggggttcgt gcacacagcc cagcttggag cgaacgacct acaccgaact gagataccta     4080
cagcgtgagc tatgagaaag cgccacgctt cccgaaggga gaaaggcgga caggtatccg     4140
gtaagcggca gggtcggaac aggagagcgc acgagggagc ttccaggggg aaacgcctgg     4200
tatctttata gtcctgtcgg gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc     4260
tcgtcagggg ggcggagcct atggaaaaac gccagcaacg cggccttttt acggttcctg     4320
gccttttgct ggccttttgc tcacatgttc tttcctgcgt tatcccctga ttctgtggat     4380
aaccgtatta ccgcctttga gtgagctgat accgctcgcc gcagccgaac gaccgagcgc     4440
agcgagtcag tgagcgagga agcggaag                                        4468
```

<210> 7
<211> 6249
<212> ADN
<213> Secuencia artificial

<220>
<223> Plásmido

<400> 7

```
ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc       60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg      120
gccaactcca tcactagggg ttcctactag tggcctccgc gccgggtttt ggcgcctccc      180
gcgggcgccc ccctcctcac ggcgagcgct gccacgtcag acgaagggcg cagcgagcgt      240
cctgatcctt ccgcccggac gctcaggaca gcggcccgct gctcataaga ctcggcctta      300
gaaccccagt atcagcagaa ggacatttta ggacgggact tgggtgactc tagggcactg      360
gttttctttc cagagagcgg aacaggcgag gaaaagtagt cccttctcgg cgattctgcg      420
gagggatctc cgtggggcgg tgaacgccga tgattatata aggacgcgcc gggtgtggca      480
cagctagttc cgtcgcagcc gggatttggg tcgcggttct tgtttgtgga tcgctgtgat      540
cgtcacttgg cggccgccat ggtcagcaag ggcgaggagg ataacatggc catcatcaag      600
gagttcatgc gcttcaaggt gcacatggag ggctccgtga acggccacga gttcgagatc      660
gagggcgagg gcgagggccg cccctacgag ggcacccaga ccgccaagct gaaggtgacc      720
aagggtggcc ccctgccctt cgcctgggac atcctgtccc ctcaattcat gtatggctcc      780
aaggcctacg tgaagcaccc cgccgacatc cccgactact tgaagctgtc cttccccgag      840
ggcttcaagt gggagcgcgt gatgaacttc gaggacggcg gcgtggtgac cgtgacccag      900
gactcctcct tacaagacgg cgagttcatc tacaaagtga agctgcgcgg caccaacttc      960
ccctccgacg gccccgtaat gcagaagaag accatgggct gggaggcctc ctccgagcgg     1020
atgtaccccg aggacggcgc cctgaagggc gagatcaagc agaggctgaa gctgaaggac     1080
ggcggccact acgacgctga ggtcaagacc acctacaagg ccaagaagcc cgtgcagctg     1140
cccggcgcct acaacgtcaa catcaagttg gacatcacct cccacaacga ggactacacc     1200
atcgtggaac agtacgaacg cgccgagggc cgccactcca ccggcggcat ggacgagctg     1260
tacaagtaaa ggatcccctc ccccgtgcct tccttgaccc tggaaggtgc cactcccact     1320
gtcctttcct cctcctagga ataaaatatc tttattttca ttacatctgt gtgttggttt     1380
tttgtgtaga tctgttcaaa tttgaactga ctaagcggct cccgccagat tttggcaaga     1440
ttactaagca ggaagtcaag gacttttttg cttgggcaaa ggtcaatcag gtgccggtga     1500
```

ctcacgagtt taaagttccc agggaattgg cgggaactaa aggggcggag aaatctctaa 1560
aacgcccact gggtgacgtc accaatacta gctataaaag tctggagaag cgggccaggc 1620
tctcatttgt tcccgagacg cctcgcagtt cagacgtgac tgttgatccc gctcctctgc 1680
gaccgctagc ttcgatcaac tacgcggaca ggtaccaaaa caaatgttct cgtcacgtgg 1740
gcatgaatct gatgctgttt ccctgcagac aatgcgagag actgaatcag aattcaaata 1800
tctgcttcac tcacggtgtc aaagactgtt tagagtgctt tcccgtgtca gaatctcaac 1860
ccgtttctgt cgtcaaaaag gcgtatcaga aactgtgcta cattcatcac atcatgggaa 1920
aggtgccaga cgcttgcact gcttgcgacc tggtcaatgt ggacttggat gactgtgttt 1980
ctgaacaata aatgacttaa accaggtatg gctgccgatg gttatcttcc agattggctc 2040
gaggacaacc ttagtgaagg aattcgcgag tggtgggctt tgaaacctgg agcccctcaa 2100
cccaaggcaa atcaacaaca tcaagacaac gctcgaggtc ttgtgcttcc gggttacaaa 2160
taccttggac ccggcaacgg actcgacaag ggggagccgg tcaacgcagc agacgcggcg 2220
gccctcgagc acgacaaggc ctacgaccag cagctcaagg ccggagacaa cccgtacctc 2280
aagtacaacc acgccgacgc cgagttccag gagcggctca aagaagatac gtcttttggg 2340
ggcaacctcg ggcgagcagt cttccaggcc aaaaagaggc ttcttgaacc tcttggtctg 2400
gttgaggaag cggctaagac ggctcctgga aagaagaggc ctgtagagca gtctcctcag 2460
gaaccggact cctccgcggg tattggcaaa tcgggtgcac agcccgctaa aaagagactc 2520
aatttcggtc agactggcga cacagagtca gtcccagacc ctcaaccaat cggagaacct 2580
cccgcagccc cctcaggtgt gggatctctt acaatggctt caggtggtgg cgcaccagtg 2640
gcagacaata acgaaggtgc cgatggagtg ggtagttcct cgggaaattg gcattgcgat 2700
tcccaatggc tgggggacag agtcatcacc accagcaccc gaacctgggc cctgcccacc 2760
tacaacaatc acctctacaa gcaaatctcc aacagcacat ctggaggatc ttcaaatgac 2820
aacgcctact tcggctacag caccccctgg gggtattttg acttcaacag attccactgc 2880
cacttctcac cacgtgactg gcagcgactc atcaacaaca actggggatt ccggcctaag 2940
cgactcaact tcaagctctt caacattcag gtcaaagagg ttacggacaa caatggagtc 3000
aagaccatcg ccaataacct taccagcacg gtccaggtct tcacggactc agactatcag 3060
ctcccgtacg tgctcgggtc ggctcacgag ggctgcctcc cgccgttccc agcggacgtt 3120
ttcatgattc ctcagtacgg gtatctgacg cttaatgatg gaagccaggc cgtgggtcgt 3180
tcgtcctttt actgcctgga atatttcccg tcgcaaatgc taagaacggg taacaacttc 3240
cagttcagct acgagtttga gaacgtacct ttccatagca gctacgctca cagccaaagc 3300
ctggaccgac taatgaatcc actcatcgac caatacttgt actatctctc tagaactatt 3360
accggttggg ttcaaaacca aggaatactt ccgggtatgg tttggcagga cagagatgtg 3420
tacctgcaag gacccatttg ggccaaaatt cctcacacgg acggcaactt tcacccttct 3480
ccgctgatgg gagggtttgg aatgaagcac ccgcctcctc agatcctcat caaaaacaca 3540
cctgtacctg cggatcctcc aacggccttc aacaaggaca agctgaactc tttcatcacc 3600
cagtattcta ctggccaagt cagcgtggag atcgagtggg agctgcagaa ggaaaacagc 3660
aagcgctgga acccggagat ccagtacact tccaactatt acaagtctaa taatgttgaa 3720
tttgctgtta atactgaagg tgtatatagt gaaccccgcc ccattggcac cagatacctg 3780
actcgtaatc tgtaagtcga ctaccgttcg tatagcatac attatacgaa gttatcatat 3840
gttcgagcag acatgataag atacattgat gagtttggac aaaccacaac tagaatgcag 3900
tgaaaaaaat gctttatttg tgaaatttgt gatgctattg ctttatttgt aaccattata 3960
agctgcaata aacaagttaa caacaacaat tgcattcatt ttatgtttca ggttcagggg 4020
gagatgtggg aggtttttta aagcaagtaa aacctctaca aatgtggtaa aatcgagctc 4080
taccgttcgt ataatgtatg ctatacgaag ttatgatatc aagcttagga acccctagtg 4140
atggagttgg ccactccctc tctgcgcgct cgctcgctca ctgaggccgg gcgaccaaag 4200
gtcgcccgac gcccgggctt tgcccgggcg gcctcagtga gcgagcgagc gcgcagagag 4260
ggagtggcca agctagcggg cgattaagga aagggctaga tcattcttga agacgaaagg 4320
gcctcgtgat acgcctattt ttataggtta atgtcatgat aataatggtt tcttagacgt 4380
caggtggcac ttttcgggga aatgtgcgcg gaacccctat ttgtttattt ttctaaatac 4440
attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa 4500
aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt tttgcggcat 4560
tttgccttcc tgtttttgct cacccagaaa cgctggtgaa agtaaaagat gctgaagatc 4620
agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga 4680
gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagttctg ctatgtggcg 4740
cggtattatc ccgtgttgac gccgggcaag agcaactcgg tcgccgcata cactattctc 4800
agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag 4860
taagagaatt atgcagtgct gccataacca tgagtgataa cactgcggcc aacttacttc 4920
tgacaacgat cggaggaccg aaggagctaa ccgcttttt gcacaacatg ggggatcatg 4980
taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg 5040
acaccacgat gcctgtagca atggcaacaa cgttgcgcaa actattaact ggcgaactac 5100
ttactctagc ttcccggcaa caattaatag actggatgga ggcggataaa gttgcaggac 5160
cacttctgcg ctcggccctt ccggctggct ggtttattgc tgataaatct ggagccggtg 5220
agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg 5280

```
tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg     5340
agataggtgc ctcactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac     5400
tttagattga tttaaaactt cattttaat ttaaaaggat ctaggtgaag atccttttg      5460
ataatctcat gaccaaaatc ccttaacgtg agttttcgtt ccactgagcg tcagaccccg     5520
tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc     5580
aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc ggatcaagag ctaccaactc     5640
tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtt cttctagtgt     5700
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctcgctctgc     5760
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact     5820
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacgggggt tcgtgcacac     5880
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag     5940
aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc ggcagggtcg     6000
gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg     6060
tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca ggggggcgga     6120
gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt     6180
ttgctcacat gtaataaaca cacacacacc aacaaccgtg gttggttgtt gtgttggttt     6240
attctcgag                                                            6249
```

<210> 8
<211> 743
<212> PRT
<213> Secuencia artificial

<220>
<223> Proteína VP1 de cápside de virus adeno-asociado

<400> 8

```
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1               5                  10                  15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
            20                  25                  30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
        35                  40                  45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
    50                  55                  60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65                  70                  75                  80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
        115                 120                 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
```

```
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145                 150                 155                 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
            180                 185                 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
        195                 200                 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
    210                 215                 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225                 230                 235                 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
                245                 250                 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
            260                 265                 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
        275                 280                 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
    290                 295                 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305                 310                 315                 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
                325                 330                 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
            340                 345                 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
        355                 360                 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
    370                 375                 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
```

385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445

Arg Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Thr Leu Ala Val
580 585 590

Pro Phe Lys Ala Gln Ala Gln Thr Gly Trp Val Gln Asn Gln Gly Ile
595 600 605

Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro
610 615 620

Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser Pro
625 630 635 640

Leu Met Gly Gly Phe Gly Met Lys His Pro Pro Pro Gln Ile Leu Ile
645 650 655

Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Ala Phe Asn Lys Asp
660 665 670

Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val
675 680 685

Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro
690 695 700

Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys Ser Asn Asn Val Glu Phe
705 710 715 720

Ala Val Asn Thr Glu Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr
725 730 735

Arg Tyr Leu Thr Arg Asn Leu
740

<210> 9
<211> 2232
<212> ADN
<213> Secuencia artificial

<220>
<223> Gen de cápside de virus adeno-asociado

<400> 9

```
atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc       60
gagtggtggg ctttgaaacc tggagcccct caacccaagg caaatcaaca acatcaagac      120
aacgctcgag gtcttgtgct tccgggttac aaataccttg gacccggcaa cggactcgac      180
aaggggggagc cggtcaacgc agcagacgcg gcggccctcg agcacgacaa ggcctacgac     240
cagcagctca aggccggaga caacccgtac ctcaagtaca accacgccga cgccgagttc      300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag      360
gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct      420
ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggtattggc      480
aaatcgggtg cacagcccgc taaaaagaga ctcaatttcg gtcagactgg cgacacagag      540
tcagtcccag accctcaacc aatcggagaa cctcccgcag ccccctcagg tgtgggatct      600
cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga      660
gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctgggggа cagagtcatc      720
accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc      780
tccaacagca catctggagg atcttcaaat gacaacgcct acttcggcta cagcaccccc      840
tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga      900
ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt      960
caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc     1020
acggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac     1080
gagggctgcc tcccgccgtt cccagcggac gttttcatga ttcctcagta cgggtatctg     1140
acgcttaatg atggaagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc     1200
ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta     1260
cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc     1320
gaccaatact tgtactatct ctctagaact attaacggtt ctggacagaa tcaacaaacg     1380
ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct     1440
ggacccagct accgacaaca acgtgtctca accactgtga ctcaaaacaa caacagcgaa     1500

tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct     1560
ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct     1620
ttaatttttg gcaaacaagg taccggcaga gacaacgtgg atgcggacaa agtcatgata     1680
accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg     1740
gccacaaacc accagagtgc ccaaactttg gcggtgcctt ttaaggcaca ggcgcagacc     1800
ggttgggttc aaaaccaagg aatacttccg ggtatggttt ggcaggacag agatgtgtac     1860
ctgcaaggac ccatttgggc caaaattcct cacacggacg gcaactttca cccttctccg     1920
ctgatgggag ggtttggaat gaagcacccg cctcctcaga tcctcatcaa aaacacacct     1980
gtacctgcgg atcctccaac ggccttcaac aaggacaagc tgaactcttt catcacccag     2040
tattctactg gccaagtcag cgtggagatc gagtgggagc tgcagaagga aaacagcaag     2100
cgctggaacc cggagatcca gtacacttcc aactattaca agtctaataa tgttgaattt     2160
gctgttaata ctgaaggtgt atatagtgaa ccccgcccca ttggcaccag atacctgact     2220
cgtaatctgt aa                                                        2232
```

<210> 10
<211> 9378
<212> ADN
<213> Secuencia artificial

<220>
<223> Plásmido

<400> 10

```
gtcgacggta tcgggggagc tcgcagggtc tccattttga agcgggaggt ttgaacgcgc       60
agccgccatg ccggggtttt acgagattgt gattaaggtc cccagcgacc ttgacgagca      120
tctgcccggc atttctgaca gctttgtgaa ctgggtggcc gagaaggaat gggagttgcc      180
gccagattct gacatggatc tgaatctgat tgagcaggca cccctgaccg tggccgagaa      240
gctgcagcgc gactttctga cggaatggcg ccgtgtgagt aaggccccgg aggctctttt      300
ctttgtgcaa tttgagaagg gagagagcta cttccacatg cacgtgctcg tggaaaccac      360
cggggtgaaa tccatggttt tgggacgttt cctgagtcag attcgcgaaa aactgattca      420
gagaatttac cgcgggatcg agccgacttt gccaaactgg ttcgcggtca caaagaccag      480
aaatggcgcc ggaggcggga acaaggtggt ggatgagtgc tacatcccca attacttgct      540
ccccaaaacc cagcctgagc tccagtgggc gtggactaat atggaacagt atttaagcgc      600
ctgtttgaat ctcacggagc gtaaacggtt ggtggcgcag catctgacgc acgtgtcgca      660
gacgcaggag cagaacaaag agaatcagaa tcccaattct gatgcgccgg tgatcagatc      720
aaaaacttca gccaggtaca tggagctggt cgggtggctc gtggacaagg ggattacctc      780
ggagaagcag tggatccagg aggaccaggc ctcatacatc tccttcaatg cggcctccaa      840
ctcgcggtcc caaatcaagg ctgccttgga caatgcggga aagattatga gcctgactaa      900
aaccgccccc gactacctgg tgggccagca gcccgtggag gacatttcca gcaatcggat      960
ttataaaatt ttggaactaa acgggtacga tccccaatat gcggcttccg tctttctggg     1020
atgggccacg aaaaagttcg gcaagaggaa caccatctgg ctgtttgggc ctgcaactac     1080
cgggaagacc aacatcgcgg aggccatagc ccacactgtg cccttctacg ggtgcgtaaa     1140
ctggaccaat gagaactttc ccttcaacga ctgtgtcgac aagatggtga tctggtggga     1200
ggaggggaag atgaccgcca aggtcgtgga gtcggccaaa gccattctcg gaggaagcaa     1260
ggtgcgcgtg gaccagaaat gcaagtcctc ggcccagata gacccgactc ccgtgatcgt     1320
cacctccaac accaacatgt gcgccgtgat tgacgggaac tcaacgacct tcgaacacca     1380
gcagccgttg caagaccgga tgttcaaatt tgaactcacc cgccgtctgg atcatgactt     1440
tgggaaggtc accaagcagg aagtcaaaga ctttttccgg tgggcaaagg atcacgtggt     1500
tgaggtggag catgaattct acgtcaaaaa gggtggagcc aagaaaagac ccgcccccag     1560
tgacgcagat ataagtgagc ccaaacgggt gcgcgagtca gttgcgcagc catcgacgtc     1620
agacgcggaa gcttcgatca actacgcgga caggtaccaa aacaaatgtt ctcgtcacgt     1680
gggcatgaat ctgatgctgt ttccctgcag acaatgcgag agactgaatc agaattcaaa     1740
tatctgcttc actcacggtg tcaaagactg tttagagtgc tttcccgtgt cagaatctca     1800
acccgtttct gtcgtcaaaa aggcgtatca gaaactgtgc tacattcatc acatcatggg     1860
aaaggtgcca gacgcttgca ctgcttgcga cctggtcaat gtggacttgg atgactgtgt     1920
ttctgaacaa taaatgactt aaaccaggta tgagtcggct ggataaatct aaagtcataa     1980
acggcgctct ggaattactc aatgaagtcg gtatcgaagg cctgacgaca aggaaactcg     2040
ctcaaaagct gggagttgag cagcctaccc tgtactggca cgtgaagaac aagcgggccc     2100
tgctcgatgc cctggccatc gagatgctgg acaggcatca tacccacttc tgccccctgg     2160
aaggcgagtc atggcaagac tttctgcgga acaacgccaa gtcattccgc tgtgctctcc     2220
```

```
tctcacatcg cgacggggct aaagtgcatc tcggcacccg cccaacagag aaacagtacg 2280
aaaccctgga aaatcagctc gcgttcctgt gtcagcaagg cttctccctg gagaacgcac 2340
tgtacgctct gtccgccgtg ggccacttta cactgggctg cgtattggag gaacaggagc 2400
atcaagtagc aaaagaggaa agagagacac ctaccaccga ttctatgccc ccacttctga 2460
gacaagcaat tgagctgttc gaccggcagg gagccgaacc tgccttcctt ttcggcctgg 2520
aactaatcat atgtggcctg gagaaacagc taaagtgcga aagcggcggg ccggccgacg 2580
cccttgacga ttttgactta gacatgctcc cagccgatgc ccttgacgac tttgaccttg 2640
atatgctgcc tgctgacgct cttgacgatt ttgaccttga catgctcccc gggtaaatgc 2700
atgaattcga tctagagggc cctattctat agtgtcacct aaatgctaga gctcgctgat 2760
cagcctcgac tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt 2820
ccttgaccct ggaaggtgcc actcccactg tcctttccta ataaaatgag gaaattgcat 2880
cgcattgtct gagtaggtgt cattctattc tgggggggtgg ggtggggcag gacagcaagg 2940
gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggcttctg 3000
aggcggaaag aaccagctgg ggctcgaatc aagctatcaa gtgccacctg acgtctccct 3060
atcagtgata gagaagtcga cacgtctcga gctccctatc agtgatagag aaggtacgtc 3120
tagaacgtct ccctatcagt gatagagaag tcgacacgtc tcgagctccc tatcagtgat 3180
agagaaggta cgtctagaac gtctccctat cagtgataga gaagtcgaca cgtctcgagc 3240
tccctatcag tgatagagaa ggtacgtcta gaacgtctcc ctatcagtga tagagaagtc 3300
gacacgtctc gagctcccta tcagtgatag agaaggtacc ccctatataa gcagagagat 3360
ctgttcaaat ttgaactgac taagcggctc ccgccagatt ttggcaagat tactaagcag 3420
gaagtcaagg actttttgc ttgggcaaag gtcaatcagg tgccggtgac tcacgagttt 3480
aaagttccca gggaattggc gggaactaaa ggggcggaga aatctctaaa acgcccactg 3540
ggtgacgtca ccaatactag ctataaaagt ctggagaagc gggccaggct ctcatttgtt 3600
cccgagacgc ctcgcagttc agacgtgact gttgatcccg ctcctctgcg accgctagct 3660
tcgatcaact acgcggacag gtaccaaaac aaatgttctc gtcacgtggg catgaatctg 3720
atgctgtttc cctgcagaca atgcgagaga ctgaatcaga attcaaatat ctgcttcact 3780
cacggtgtca aagactgttt agagtgcttt cccgtgtcag aatctcaacc cgtttctgtc 3840
gtcaaaaagg cgtatcagaa actgtgctac attcatcaca tcatgggaaa ggtgccagac 3900
gcttgcactg cttgcgacct ggtcaatgtg gacttggatg actgtgtttc tgaacaataa 3960
atgacttaaa ccaggtatgg ctgccgatgg ttatcttcca gattggctcg aggacaacct 4020
tagtgaagga attcgcgagt ggtgggcttt gaaacctgga gcccctcaac ccaaggcaaa 4080
tcaacaacat caagacaacg ctcgaggtct tgtgcttccg ggttacaaat accttggacc 4140
cggcaacgga ctcgacaagg gggagccggt caacgcagca gacgcggcgg ccctcgagca 4200
cgacaaggcc tacgaccagc agctcaaggc cggagacaac ccgtacctca agtacaacca 4260
cgccgacgcc gagttccagg agcggctcaa agaagatacg tcttttgggg gcaacctcgg 4320
gcgagcagtc ttccaggcca aaaagaggct tcttgaacct cttggtctgg ttgaggaagc 4380
ggctaagacg gctcctggaa agaagaggcc tgtagagcag tctcctcagg aaccggactc 4440
ctccgcgggt attggcaaat cgggtgcaca gcccgctaaa aagagactca atttcggtca 4500
gactggcgac acagagtcag tcccagaccc tcaaccaatc ggagaacctc ccgcagcccc 4560
ctcaggtgtg ggatctctta caatggcttc aggtggtggc gcaccagtgg cagacaataa 4620
cgaaggtgcc gatggagtgg gtagttcctc gggaaattgg cattgcgatt cccaatggct 4680
gggggacaga gtcatcacca ccagcacccg aacctgggcc ctgcccacct acaacaatca 4740
cctctacaag caaatctcca acagcacatc tggaggatct tcaaatgaca acgcctactt 4800
cggctacagc acccccctggg ggtattttga cttcaacaga ttccactgcc acttctcacc 4860
acgtgactgg cagcgactca tcaacaacaa ctggggattc cggcctaagc gactcaactt 4920
caagctcttc aacattcagg tcaaagaggt tacggacaac aatggagtca agaccatcgc 4980
caataacctt accagcacgg tccaggtctt cacggactca gactatcagc tcccgtacgt 5040
gctcgggtcg gctcacgagg gctgcctccc gccgttccca gcggacgttt tcatgattcc 5100
tcagtacggg tatctgacgc ttaatgatgg aagccaggcc gtgggtcgtt cgtcctttta 5160
ctgcctggaa tatttcccgt cgcaaatgct aagaacgggt aacaacttcc agttcagcta 5220
cgagtttgag aacgtacctt tccatagcag ctacgctcac agccaaagcc tggaccgact 5280
aatgaatcca ctcatcgacc aatacttgta ctatctctct agaactatta acggttctgg 5340
acagaatcaa caaacgctaa aattcagtgt ggccggaccc agcaacatgg ctgtccaggg 5400
aagaaactac atacctggac ccagctaccg acaacaacgt gtctcaacca ctgtgactca 5460
aaacaacaac agcgaatttg cttggcctgg agcttcttct tgggctctca atggacgtaa 5520
tagcttgatg aatcctggac ctgctatggc cagccacaaa gaaggagagg accgtttctt 5580
tcctttgtct ggatctttaa tttttggcaa acaaggtacc ggcagagaca acgtggatgc 5640
ggacaaagtc atgataacca acgaagaaga aattaaaact actaacccgg tagcaacgga 5700
gtcctatgga caagtggcca caaaccacca gagtgcccaa actttggcgg tgccttttaa 5760
ggcacaggcg cagaccggtt gggttcaaaa ccaaggaata cttccgggta tggtttggca 5820
ggacagagat gtgtacctgc aaggacccat ttgggccaaa attcctcaca cggacggcaa 5880
ctttcaccct tctccgctga tgggagggtt tggaatgaag cacccgcctc ctcagatcct 5940
catcaaaaac acacctgtac ctgcggatcc tccaacggcc ttcaacaagg acaagctgaa 6000
```

```
ctctttcatc acccagtatt ctactggcca agtcagcgtg gagatcgagt gggagctgca  6060
gaaggaaaac agcaagcgct ggaacccgga gatccagtac acttccaact attacaagtc  6120
taataatgtt gaatttgctg ttaatactga aggtgtatat agtgaacccc gccccattgg  6180
caccagatac ctgactcgta atctgtaatt gcttgttaat caataaaccg tttaattcgt  6240
ttcagttgaa ctttggtctc tgcgaagggc gaattcgttt aaacctgcag gactagaggt  6300
cctgtattag aggtcacgtg agtgttttgc gacattttgc gacaccatgt ggtcacgctg  6360
ggtatttaag cccgagtgag cacgcagggt ctccattttg aagcgggagg tttgaacgcg  6420
cagccgccaa gccgaattct gcagatatcc atcacactgg cggccgctcg actagagcgg  6480
ccgccaccgc ggtggagctc cagcttttgt tcccttttagt gagggttaat tgcgcgcttg  6540
gcgtaatcat ggtcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac  6600
aacatacgag ccggaagcat aaagtgtaaa gcctggggtg cctaatgagt gagctaactc  6660
acattaattg cgttgcgctc actgcccgct ttccagtcgg gaaacctgtc gtgccagctg  6720
cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct  6780
tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac  6840
tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga  6900
gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gttttccat  6960
aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac  7020
ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct  7080
gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg  7140
ctttctcata gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg  7200
ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt  7260
cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg  7320
attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac  7380
ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga  7440
aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt  7500
gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt  7560
tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga  7620
ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc  7680
taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct  7740
atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata  7800
actacgatac gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca  7860
cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga  7920
agtggtcctg caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga  7980
gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg ccattgctac aggcatcgtg  8040
gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga  8100
gttacatgat cccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt  8160
gtcagaagta agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct  8220
cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca  8280
ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat  8340
accgcgccac atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga  8400
aaactctcaa ggatcttacc gctgttgaga tccagttcga tgtaacccac tcgtgcaccc  8460
aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg  8520
caaaatgccg caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc  8580
ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt  8640
gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca  8700
cctaaattgt aagcgttaat attttgttaa aattcgcgtt aaatttttgt taaatcagct  8760
catttttaa ccaataggcc gaaatcggca aaatcccctta taaatcaaaa gaatagaccg  8820
agatagggtt gagtgttgtt ccagtttgga acaagagtcc actattaaag aacgtggact  8880
ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt gaaccatcac  8940
cctaatcaag ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac cctaaaggga  9000
gcccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag gaagggaaga  9060
aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg cgcgtaacca  9120
ccacacccgc cgcgcttaat gcgccgctac agggcgcgtc ccattcgcca ttcaggctgc  9180
gcaactgttg ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaaag  9240
ggggatgtgc tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt  9300
gtaaaacgac ggccagtgag cgcgcgtaat acgactcact atagggcgaa ttgggtaccg  9360
ggcccccccct cgatcgag                                               9378
```

<210> 11
<211> 2211
<212> ADN
<213> Secuencia de cápside AAV9

<220>
<223> Gen VP1 de cápside de virus adeno-asociado

<400> 11

```
atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc       60
gagtggtggg ctttgaaacc tggagcccct caacccaagg caaatcaaca acatcaagac      120
aacgctcgag gtcttgtgct tccgggttac aaatatcttg gacccggcaa cggactcgac      180
aaggggggagc cggtcaacgc agcagacgcg gcggccctcg agcacgacaa ggcctacgac     240
cagcagctca aggccggaga caacccgtac ctcaagtaca accacgccga cgccgagttc      300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag      360
gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct      420
ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggtattggc      480
aaatcgggtg cacagcccgc taaaaagaga ctcaatttcg gtcagactgg cgacacagag      540
tcagtcccag accctcaacc aatcggagaa cctcccgcag ccccctcagg tgtgggatct      600
cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga      660
gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctgggga cagagtcatc       720
accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc      780
tccaacagca catctggagg atcttcaaat gacaacgcct acttcggcta cagcaccccc      840
tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga      900
ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt      960
caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc     1020
acggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac     1080
gagggctgcc tcccgccgtt cccagcggac gttttcatga ttcctcagta cgggtatctg     1140
acgcttaatg atggaagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc     1200
ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta     1260
cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc     1320
gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg     1380
ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct     1440
ggacccagct accgacaaca acgtgtctca accactgtga ctcaaaacaa caacagcgaa     1500
tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct     1560
ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct     1620
ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgcggacaa agtcatgata     1680
accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg     1740
gccacaaacc accagagtgc ccaagcacag gcgcagaccg gctgggttca aaaccaagga     1800
atacttccgg gtatggtttg gcaggacaga gatgtgtacc tgcaaggacc catttgggcc     1860
aaaattcctc acacggacgg caactttcac ccttctccgc tgatgggagg gtttggaatg     1920
aagcacccgc ctcctcagat cctcatcaaa aacacacctg tacctgcgga tcctccaacg     1980
gccttcaaca aggacaagct gaactctttc atcacccagt attctactgg ccaagtcagc     2040
gtggagatcg agtgggagct gcagaaggaa aacagcaagc gctggaaccc ggagatccag     2100
tacacttcca actattacaa gtctaataat gttgaatttg ctgttaatac tgaaggtgta     2160
tatagtgaac cccgccccat tggcaccaga tacctgactc gtaatctgta a               2211
```

<210> 12
<211> 736
<212> PRT
<213> Secuencia artificial

<220>
<223> Proteína de cápside de virus adeno-asociado G2B-13

<400> 12

```
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1               5                  10                  15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
            20                  25                  30
```

```
Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
        35                  40                  45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
    50                  55                  60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65                  70                  75                  80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
        115                 120                 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145                 150                 155                 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
            180                 185                 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
        195                 200                 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
    210                 215                 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225                 230                 235                 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
                245                 250                 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
            260                 265                 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
        275                 280                 285
```

```
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
    290                 295                 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305                 310                 315                 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
                325                 330                 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
            340                 345                 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
        355                 360                 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
    370                 375                 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385                 390                 395                 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
                405                 410                 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
            420                 425                 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
        435                 440                 445

Arg Thr Ile Gln Ser Ser Gln Thr Pro Arg Gln Thr Leu Lys Phe Ser
    450                 455                 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465                 470                 475                 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
                485                 490                 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
            500                 505                 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
        515                 520                 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
    530                 535                 540
```

```
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545                 550                 555                 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
                565                 570                 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
            580                 585                 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
        595                 600                 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
    610                 615                 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625                 630                 635                 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
                645                 650                 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
            660                 665                 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
        675                 680                 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
    690                 695                 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705                 710                 715                 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
                725                 730                 735
```

<210> 13
<211> 743
<212> PRT
<213> Secuencia artificial

<220>
<223> Proteína de cápside de virus adeno-asociado G2B-26

<400> 13

```
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1               5                   10                  15
```

```
Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
            20                  25                  30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
        35                  40                  45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
    50                  55                  60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65                  70                  75                  80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
        115                 120                 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145                 150                 155                 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
            180                 185                 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
        195                 200                 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
    210                 215                 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225                 230                 235                 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
                245                 250                 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
            260                 265                 270
```

```
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
        275                 280                 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
    290                 295                 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305                 310                 315                 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
                325                 330                 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
            340                 345                 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
        355                 360                 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
    370                 375                 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385                 390                 395                 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
                405                 410                 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
            420                 425                 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
        435                 440                 445

Arg Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
    450                 455                 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465                 470                 475                 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
                485                 490                 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
            500                 505                 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
```

```
          515                 520                 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
    530                 535                 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545                 550                 555                 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
                565                 570                 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Thr Leu Ala Val
            580                 585                 590

Pro Phe Lys Ala Gln Ala Gln Thr Gly Trp Val Gln Asn Gln Gly Ile
        595                 600                 605

Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro
    610                 615                 620

Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser Pro
625                 630                 635                 640

Leu Met Gly Gly Phe Gly Met Lys His Pro Pro Pro Gln Ile Leu Ile
                645                 650                 655

Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Ala Phe Asn Lys Asp
            660                 665                 670

Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val
        675                 680                 685

Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro
    690                 695                 700

Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys Ser Asn Asn Val Glu Phe
705                 710                 715                 720

Ala Val Asn Thr Glu Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr
                725                 730                 735

Arg Tyr Leu Thr Arg Asn Leu
            740
```

<210> 14
<211> 743
<212> PRT
<213> Secuencia artificial

<220>
<223> Proteína de cápside de virus adeno-asociado TH1.1-32

<400> 14

```
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1               5                  10                  15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
            20                  25                  30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
        35                  40                  45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
    50                  55                  60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65                  70                  75                  80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
        115                 120                 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145                 150                 155                 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
            180                 185                 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
        195                 200                 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
    210                 215                 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
```

```
225                 230                 235                 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
                245                 250                 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
            260                 265                 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
        275                 280                 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
    290                 295                 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305                 310                 315                 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
                325                 330                 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
            340                 345                 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
        355                 360                 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
    370                 375                 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385                 390                 395                 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
                405                 410                 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
            420                 425                 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
        435                 440                 445

Arg Thr Ile Ile Leu Gly Thr Gly Thr Ser Gln Thr Leu Lys Phe Ser
    450                 455                 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465                 470                 475                 480
```

```
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
                485                 490                 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
            500                 505                 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
        515                 520                 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
    530                 535                 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545                 550                 555                 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
                565                 570                 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Thr Arg Thr Asn
            580                 585                 590

Pro Glu Ala Ala Gln Ala Gln Thr Gly Trp Val Gln Asn Gln Gly Ile
        595                 600                 605

Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro
    610                 615                 620

Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser Pro
625                 630                 635                 640

Leu Met Gly Gly Phe Gly Met Lys His Pro Pro Pro Gln Ile Leu Ile
                645                 650                 655

Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Ala Phe Asn Lys Asp
            660                 665                 670

Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val
        675                 680                 685

Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro
    690                 695                 700

Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys Ser Asn Asn Val Glu Phe
705                 710                 715                 720

Ala Val Asn Thr Glu Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr
                725                 730                 735
```

Arg Tyr Leu Thr Arg Asn Leu
740

<210> 15
<211> 743
<212> PRT
<213> Secuencia artificial

<220>
<223> Proteína de cápside de virus adeno-asociado TH1.1-35

<400> 15

```
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1               5                  10                  15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
            20                  25                  30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
        35                  40                  45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
    50                  55                  60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65                  70                  75                  80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
        115                 120                 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145                 150                 155                 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
            180                 185                 190
```

```
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
        195                 200                 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
    210                 215                 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225                 230                 235                 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
                245                 250                 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
            260                 265                 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
        275                 280                 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
    290                 295                 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305                 310                 315                 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
                325                 330                 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
            340                 345                 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
        355                 360                 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
    370                 375                 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385                 390                 395                 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
                405                 410                 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
            420                 425                 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
        435                 440                 445
```

```
Arg Thr Ile Ile Leu Gly Thr Gly Thr Ser Gln Thr Leu Lys Phe Ser
    450                 455                 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465                 470                 475                 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
                485                 490                 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
            500                 505                 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
        515                 520                 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
    530                 535                 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545                 550                 555                 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
                565                 570                 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Asn Gly Gly Thr
            580                 585                 590

Ser Ser Ser Ala Gln Ala Gln Thr Gly Trp Val Gln Asn Gln Gly Ile
        595                 600                 605

Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro
    610                 615                 620

Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser Pro
625                 630                 635                 640

Leu Met Gly Gly Phe Gly Met Lys His Pro Pro Pro Gln Ile Leu Ile
                645                 650                 655

Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Ala Phe Asn Lys Asp
            660                 665                 670

Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val
        675                 680                 685

Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro
    690                 695                 700
```

```
Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys Ser Asn Asn Val Glu Phe
705                 710                 715                 720

Ala Val Asn Thr Glu Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr
                725                 730                 735

Arg Tyr Leu Thr Arg Asn Leu
            740
```

<210> 16
<211> 27
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador 9CapF

<400> 16
caggtcttca cggactcaga ctatcag 27

<210> 17
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador CDF

<400> 17
caagtaaaac ctctacaaat gtggtaaaat cg 32

<210> 18
<211> 36
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador XF

<400> 18
actcatcgac caatacttgt actatctctc tagaac 36

<210> 19
<211> 26
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador AR

<400> 19
ggaagtattc cttggttttg aaccca 26

<210> 20
<211> 23
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador TF

<400> 20

ggtcgcggtt cttgtttgtg gat 23

<210> 21
<211> 24
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador TR

<400> 21
gcaccttgaa gcgcatgaac tcct 24

<210> 22
<211> 85
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador 7xNNK

<220>
<221> MISC_CARACTERÍSTICA
<222> (38)..(39)
<223> n indicates a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (41)..(42)
<223> n indica a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (44)..(45)
<223> n indica a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (47)..(48)
<223> n indica a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (50)..(51)
<223> n indica a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (53)..(54)
<223> n indica a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (56)..(57)
<223> n indica a, t, c o g

<400> 22

catcgaccaa tacttgtact atctctctag aactattnnk nnknnknnkn nknnknnkca 60
aacgctaaaa ttcagtgtgg ccgga 85

<210> 23
<211> 83
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador 7xMNN

<220>
<221> MISC_CARACTERÍSTICA
<222> (42)..(43)
<223> n indica a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (45)..(46)
<223> n indica a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (48)..(49)
<223> n indica a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (51)..(52)
<223> n indica a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (54)..(55)
<223> n indica a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (57)..(58)
<223> n indica a, t, c o g

<220>
<221> MISC_CARACTERÍSTICA
<222> (60).. (61)
<223> n indica a, t, c o g

<400> 23

gtattccttg gttttgaacc caaccggtct gcgcctgtgc mnnmnnmnnm nnmnnmnnmn 60
nttgggcact ctggtggttt gtg 83

<210> 24
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante AAV

<400> 24
actttggcgg tgccttttaa g 21

<210> 25
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante AAV

<400> 25
agtgtgagta agccttttt g 21

<210> 26
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante AAV

<400> 26
tttacgttga cgacgcctaa g 21

<210> 27
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante AAV

<400> 27
atgaatgcta cgaagaatgt g 21

<210> 28
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante AAV

<400> 28

Ser Val Ser Lys Pro Phe Leu
1 5

<210> 29
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante AAV

<400> 29

Phe Thr Leu Thr Thr Pro Lys
1 5

<210> 30
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante AAV

<400> 30

Met Asn Ala Thr Lys Asn Val
1 5

<210> 31
<211> 6

<212> PRT
<213> Secuencia artificial

<220>
<223> Target peptide

<400> 31

Leu Ala Val Pro Phe Lys
1 5

<210> 32
<211> 5
<212> PRT
<213> Secuencia artificial

<220>
<223> Péptido diana

<400> 32

Ala Val Pro Phe Lys
1 5

<210> 33
<211> 4
<212> PRT
<213> Secuencia artificial

<220>
<223> Péptido diana

<400> 33

Val Pro Phe Lys
1

<210> 34
<211> 6
<212> PRT
<213> Secuencia artificial

<220>
<223> Péptido diana

<400> 34

Thr Leu Ala Val Pro Phe
1 5

<210> 35
<211> 5
<212> PRT
<213> Secuencia artificial

<220>
<223> Péptido diana

<400> 35

Thr Leu Ala Val Pro
1 5

<210> 36
<211> 4
<212> PRT
<213> Secuencia artificial

<220>
<223> Péptido diana

<400> 36
Thr Leu Ala Val
1

<210> 37
<211> 49
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 37
caaccggtaa tagttctaga gagatagtac aagtattggt cgatgagtg 49

<210> 38
<211> 46
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 38
ctctctagaa ctattaccgg ttgggttcaa aaccaaggaa tacttc 46

<210> 39
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 39
gtccaaactc atcaatgtat cttatcatgt ctg 33

<210> 40
<211> 28
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 40
gagtcaatct ggaagttaac catcggca 28

<210> 41
<211> 25
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 41
gatggttaac ttccagattg actcg 25

<210> 42
<211> 28
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 42
gactactcta caggcctctt ctatccag 28

<210> 43
<211> 28
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 43
gatagaagag gcctgtagag tagtctcc 28

<210> 44
<211> 24
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 44
catcggcacc ttagttattg tctg 24

<210> 45
<211> 28
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 45
gacaataact aaggtgccga tggagtgg 28

<210> 46
<211> 40
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 46
gtctctgccg gtaccttgtt tgccaaaaat taaagatcca 40

<210> 47
<211> 39
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador

<400> 47
gcaaacaagg taccggcaga gacaacgtgg atgcggaca 39

<210> 48
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante G2B-13

<400> 48
cagtcgtcgc agacgcctag g 21

<210> 49
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante G2B-26

<400> 49
actttggcgg tgccttttaa g 21

<210> 50
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante TH 1-32

<400> 50
attctgggga ctggtacttc g 21

<210> 51
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante TH 1-32

<400> 51
acgcggacta atcctgaggc t 21

<210> 52
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante TH 1-35

<400> 52
attctgggga ctggtacttc g 21

<210> 53
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante TH 1-35

<400> 53
aatgggggga ctagtagttc t 21

<210> 54
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante G2B-13

<400> 54

Gln Ser Ser Gln Thr Pro Arg
1 5

<210> 55
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante TH 1-32

<400> 55

Ile Leu Gly Thr Gly Thr Ser
1 5

<210> 56
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante TH 1-32

<400> 56

Thr Arg Thr Asn Pro Glu Ala
1 5

<210> 57
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante TH 1-35

<400> 57

Ile Leu Gly Thr Gly Thr Ser
1 5

<210> 58
<211> 7
<212> PRT
<213> Secuencia artificial

<220>

<223> Variante TH 1-35

<400> 58

Asn Gly Gly Thr Ser Ser Ser
1 5

<210> 59
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante PHP-A

<400> 59
tatactttgt cgcagggttg g 21

<210> 60
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante PHP-A

<400> 60

Tyr Thr Leu Ser Gln Gly Trp
1 5

REIVINDICACIONES

1. Un vector de AAV que comprende una secuencia de aminoácidos que comprende la secuencia TLAVPFK (SEQ ID NO: 1), en donde la secuencia de aminoácidos es parte de una proteína de la cápside del vector AAV.

2. El vector de AAV de la reivindicación 1, en el que la secuencia de la SEQ ID NO: 1 se inserta entre los aminoácidos 588 y 589 de una proteína de la cápside de AAV9 (SEQ ID NO: 2).

3. Un péptido dirigido al sistema nervioso central (SNC), comprendiendo el péptido una secuencia de aminoácidos de la SEQ ID NO: 1; opcionalmente conjugado adicionalmente con una nanopartícula, una segunda molécula o una proteína de la cápside viral.

4. El péptido dirigido al SNC de la reivindicación 3, en el que el péptido dirigido es parte de una proteína de la cápside del AAV; preferiblemente parte de una proteína de cápside AAV9.

5. Un plásmido rep-cap que codifica un gen de la cápside de un vector AAV, comprendiendo dicho gen de la cápside una secuencia de ácido nucleico que codifica la secuencia de aminoácidos TLAVPFK (SEQ ID NO: 1), en donde el plásmido rep-cap codifica solo aquellas secuencias dentro del gen rep requerido para la expresión génica y el empalme del producto del gen de la cápside.

6. El plásmido rep-cap de la reivindicación 5, en el que la secuencia de ácido nucleico se inserta en una secuencia que codifica los aminoácidos de un bucle de superficie no estructurado de una proteína de cápside de AAV.

7. El plásmido rep-cap de la reivindicación 6, en el que la proteína de la cápside del AAV es una proteína de la cápside de AAV1, AAV2, AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, aislado humano hu.31, aislado humano hu.32, aislado de rhesus rh.8, o rhesus aislado rh.10.

8. El plásmido rep-cap de la reivindicación 7, en el que:

(a) la secuencia de ácido nucleico se inserta dentro de una secuencia que codifica los aminoácidos 452-458, de una proteína de cápside AAV9 (SEQ ID NO: 2);
(b) la secuencia de ácido nucleico se inserta entre una secuencia que codifica para los aminoácidos 588 y 589 de una proteína de cápside AAV9 (SEQ ID NO: 2); o
(c) la secuencia de ácido nucleico se inserta entre una secuencia que codifica los aminoácidos 586-592, 262-269, 464-473, 491-495, 546-557 o 659-668 de una proteína de cápside AAV9 (SEQ ID NO: 2).

9. El vector de AAV de la reivindicación 1 para uso en el tratamiento de una afección o enfermedad del sistema nervioso central.

10. El vector AAV para uso de acuerdo con la reivindicación 9, en el que la enfermedad del sistema nervioso central se selecciona de la atrofía muscular espinal (SMA), la esclerosis lateral amiotrófica (ALS), la enfermedad de Parkinson, la enfermedad de Pompe, la enfermedad de Huntington, la enfermedad de Alzheimer, la enfermedad de Batten, trastornos de almacenamiento lisosomal, glioblastoma multiforme, síndrome de Rett, amaurosis congénita de Leber, dolor crónico, infarto, lesión de la médula espinal, lesión cerebral traumática y trastornos de almacenamiento lisosomal.

11. Una proteína de la cápside que comprende una secuencia de aminoácidos que comprende la secuencia TLAVPFK (SEQ ID NO: 1).

12. La proteína de la cápside de la reivindicación 11, en la que la proteína de la cápside es una proteína de la cápside del AAV.

13. La proteína de la cápside de la reivindicación 11 o 12, en donde la proteína de la cápside es una proteína de la cápside AAV9.

14. La proteína de la cápside de la reivindicación 13, en la que:

(a) la secuencia de SEQ ID NO: 1 se inserta dentro de los aminoácidos 452-458 de la proteína de la cápside AAV9 (SEQ ID NO: 2); o
(b) la secuencia de SEQ ID NO: 1 se inserta entre 586-592, 262-269, 464-473, 491-495, 546-557 o 659-668 de la proteína de la cápside AAV9 (SEQ ID NO: 2).

15. La proteína de la cápside de la reivindicación 13, en la que la secuencia de la SEQ ID NO: 1 se inserta entre los aminoácidos 588 y 589 de la proteína de la cápside del AAV9 (SEQ ID NO: 2).

Secuencias de aminoácidos cápside para variantes G2B-26, TH1.1-32 y TH1.1-35

G2B-13

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEF
QERLKEDTSFGGNLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTESVPDPQPIGEPPAAPSGVGS
LTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQR
LINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYF
PSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTIQSSQTPRQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSE
FAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQ
YTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL **SEQ ID NO: 12**

G2B-26

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEF
QERLKEDTSFGGNLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTESVPDPQPIGEPPAAPSGVGS
LTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQR
PSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTINGSGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSE
FAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQTLAVPFKAQAQT
GWVQNQGILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWELQKENSK
RWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL **SEQ ID NO: 13**

TH1.1-32

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEF
QERLKEDTSFGGNLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTESVPDPQPIGEPPAAPSGVGS
LTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQR
PSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTIILGTGTSQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSE
FAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQTQAQAQT
GWVQNQGILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWELQKENSK
RWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL **SEQ ID NO: 14**

TH1.1-35

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEF
QERLKEDTSFGGNLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTESVPDPQPIGEPPAAPSGVGS
LTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQR
PSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTIILGTGTSQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSE
FAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQNGGTSSSAQAQT
GWVQNQGILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWELQKENSK
RWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL **SEQ ID NO: 14**

FIG. 1A

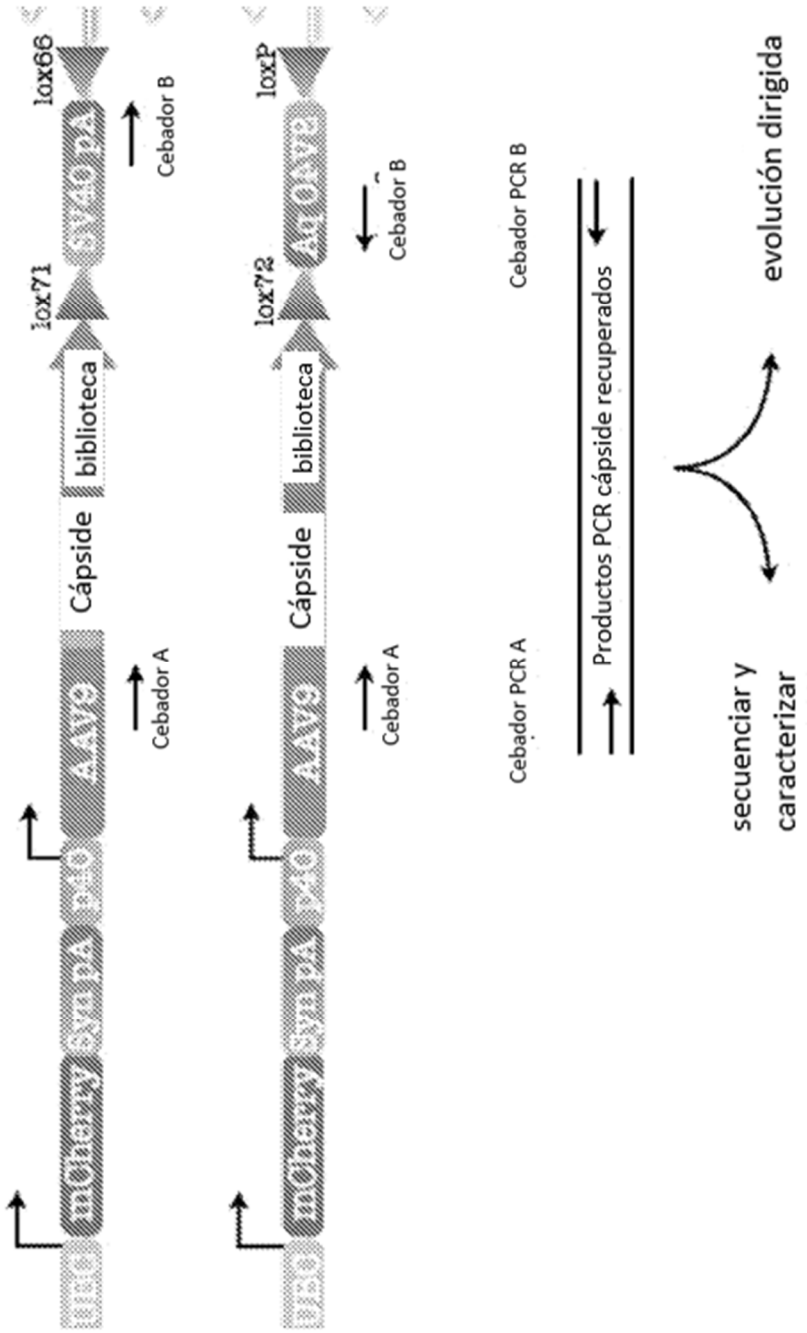

lox66
lox71
biblioteca
Cápside
AAV9
Cebador A
Cebador B
lox72
loxP
Cebador PCR A
Cebador PCR B
Productos PCR cápside recuperados
secuenciar y caracterizar
evolución dirigida

FIG. 1B

FIG. 2A

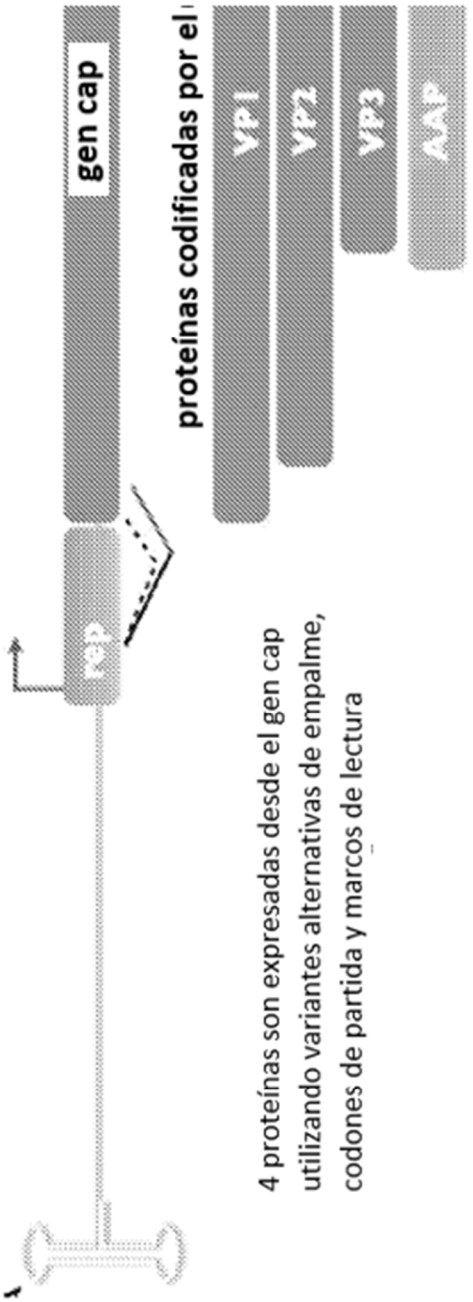
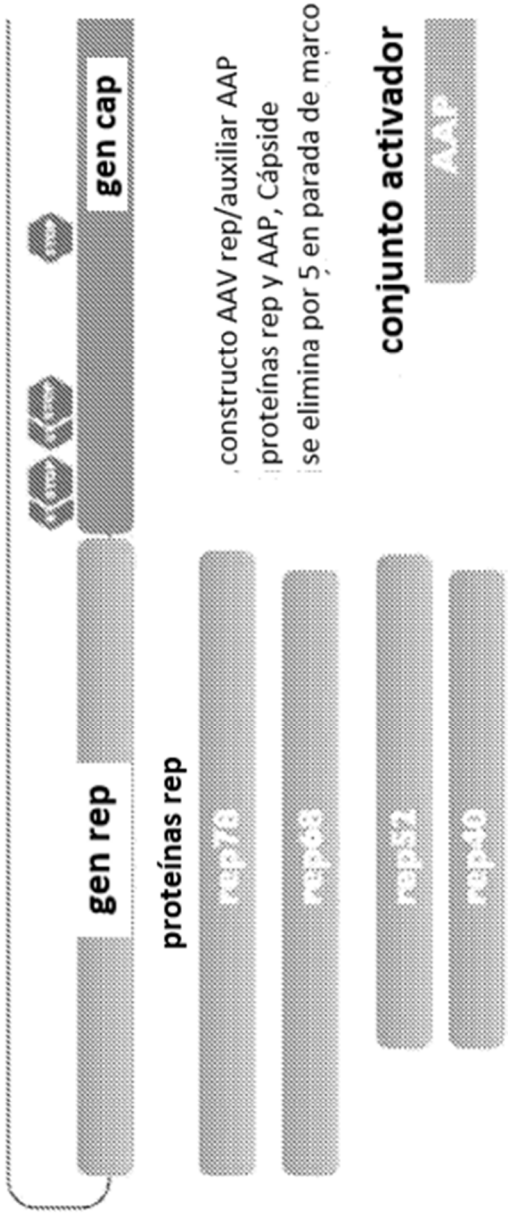

gen cap
rep
proteínas codificadas por el
VP1
VP2
VP3
AAP
4 proteínas son expresadas desde el gen cap utilizando variantes alternativas de empalme, codones de partida y marcos de lectura

FIG. 2B


gen cap
gen rep
proteínas rep
rep78
rep68
rep52
rep40
constructo AAV rep/auxiliar AAP
proteínas rep y AAP, Cápside
se elimina por 5 en parada de marco
conjunto activador
AAP

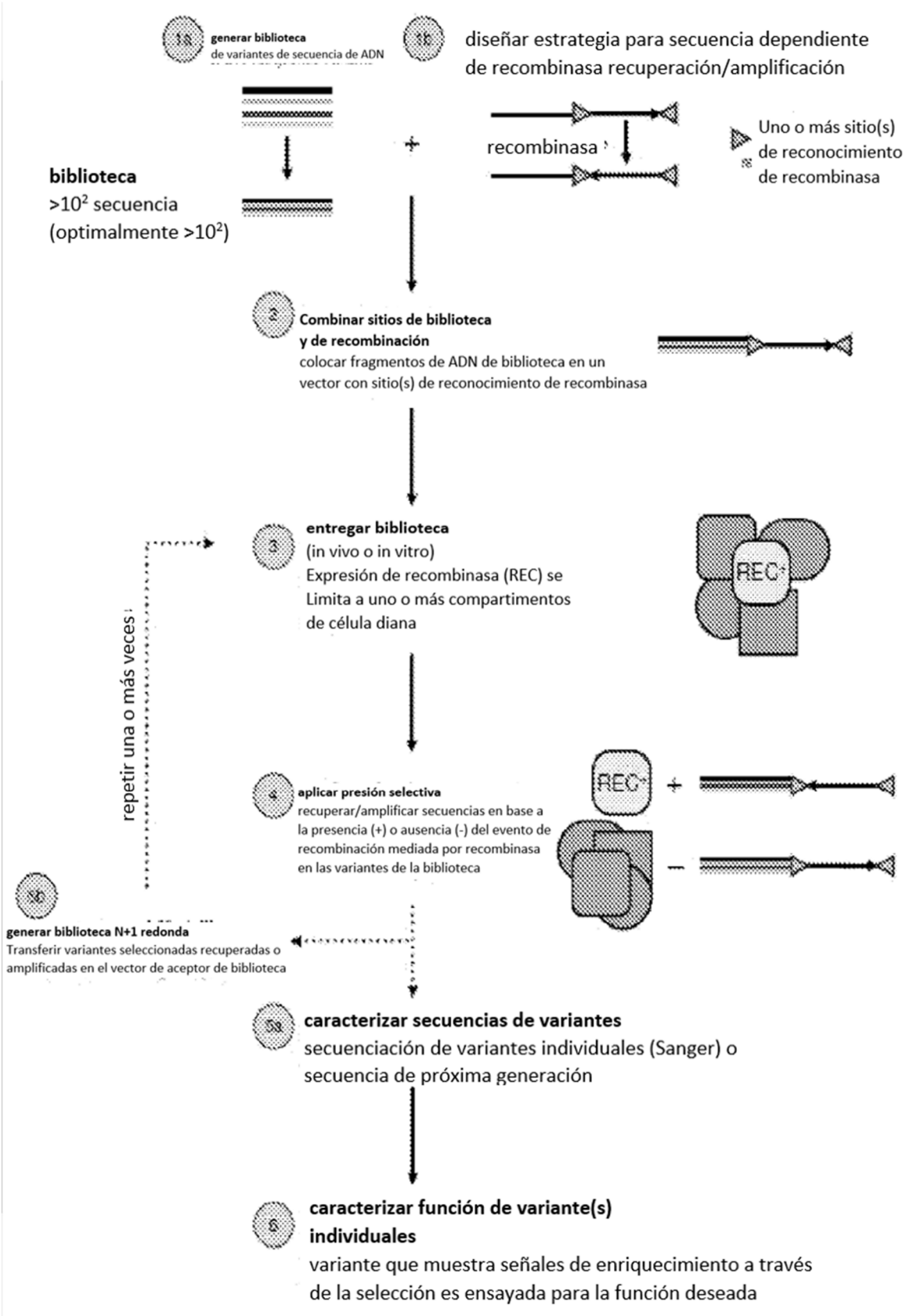

1a generar biblioteca
de variantes de secuencia de ADN
1b diseñar estrategia para secuencia dependiente
de recombinasa recuperación/amplificación
biblioteca
>10² secuencia
(optimalmente >10²)
recombinasa
Uno o más sitio(s)
de reconocimiento
de recombinasa
2 Combinar sitios de biblioteca
y de recombinación
colocar fragmentos de ADN de biblioteca en un
vector con sitio(s) de reconocimiento de recombinasa
3 entregar biblioteca
(in vivo o in vitro)
Expresión de recombinasa (REC) se
Limita a uno o más compartimentos
de célula diana
REC
repetir una o más veces
4 aplicar presión selectiva
recuperar/amplificar secuencias en base a
la presencia (+) o ausencia (-) del evento de
recombinación mediada por recombinasa
en las variantes de la biblioteca
REC
+
–
5b generar biblioteca N+1 redonda
Transferir variantes seleccionadas recuperadas o
amplificadas en el vector de aceptor de biblioteca
5a caracterizar secuencias de variantes
secuenciación de variantes individuales (Sanger) o
secuencia de próxima generación
6 caracterizar función de variante(s)
individuales
variante que muestra señales de enriquecimiento a través
de la selección es ensayada para la función deseada

FIG. 2C

FIG. 2D

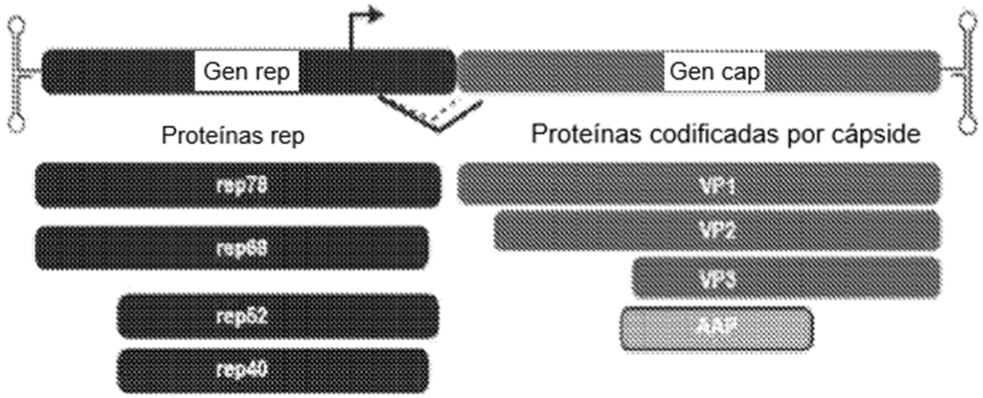


FIG.2E

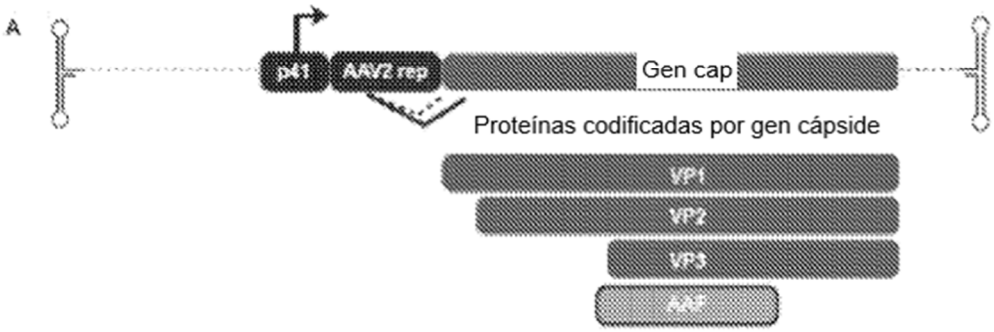


FIG. 2F

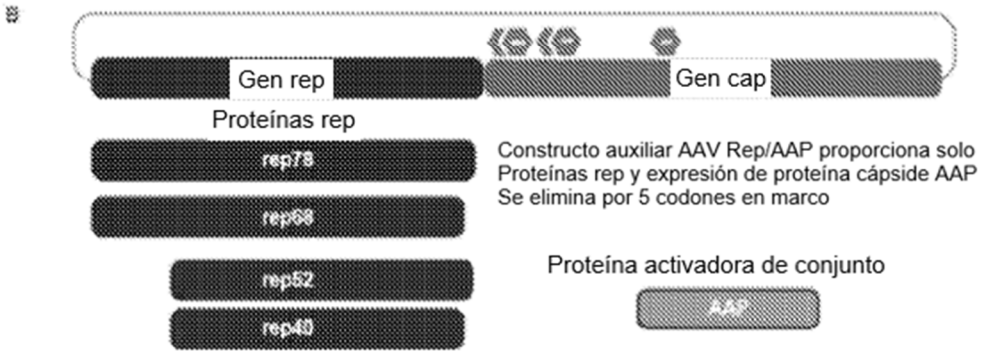


FIG. 3A

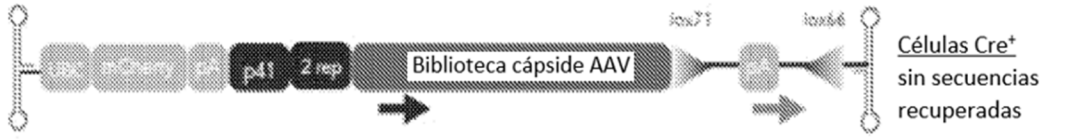

p41
2 rep
Biblioteca cápside AAV
Células Cre+
sin secuencias
recuperadas

FIG. 3B

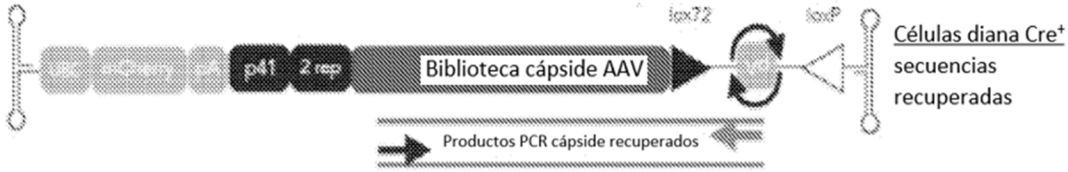

p41
2 rep
Biblioteca cápside AAV
Productos PCR cápside recuperados
Células diana Cre+
secuencias
recuperadas

FIG. 4A

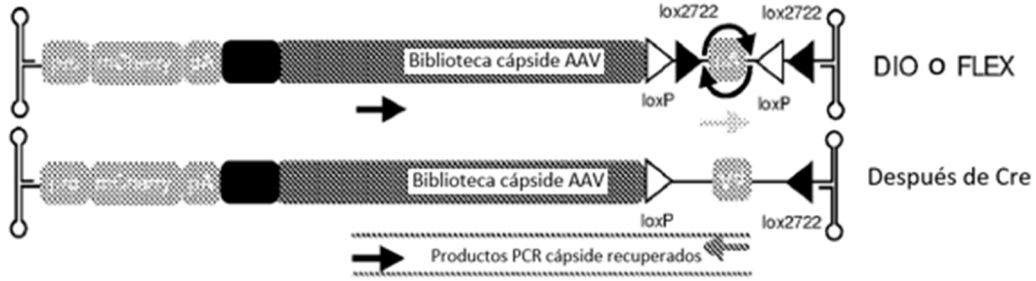


FIG. 4B

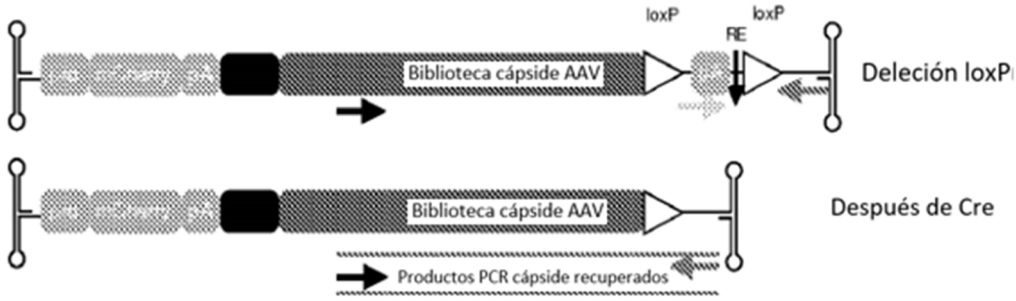


FIG. 4C

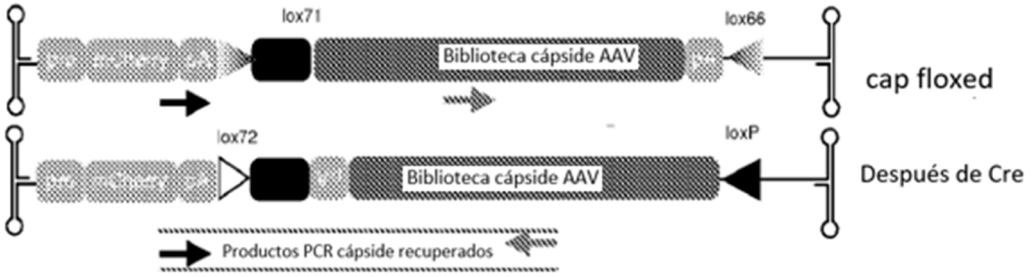


FIG. 4D

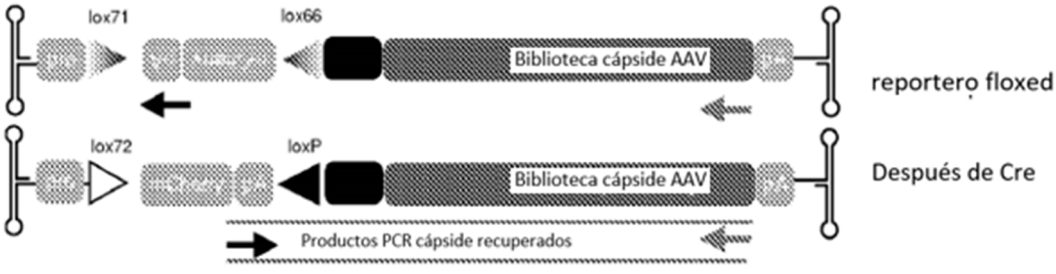


FIG. 5A

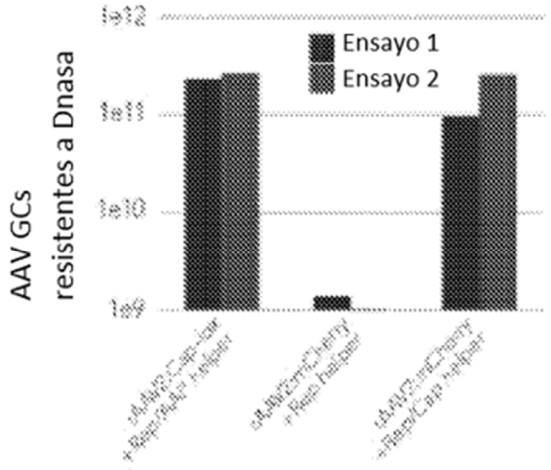


FIG. 5B

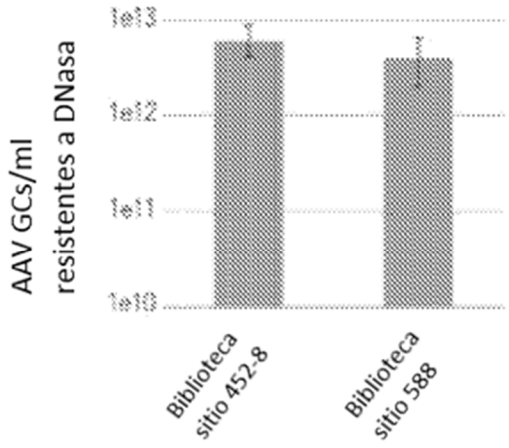
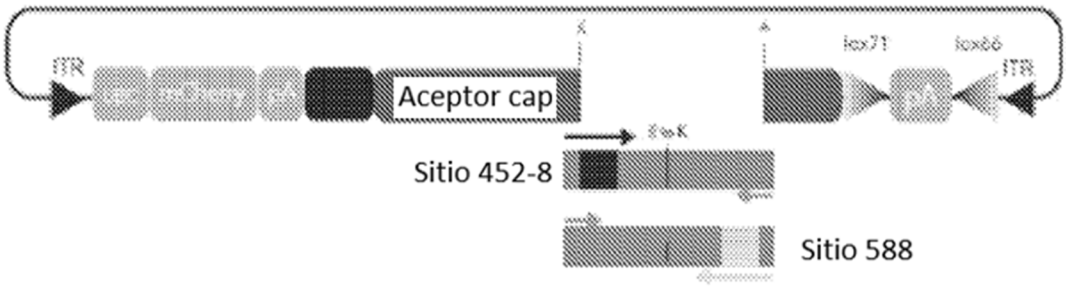


FIG. 5C



Bibliotecas basadas en PCR

FIG. 6A

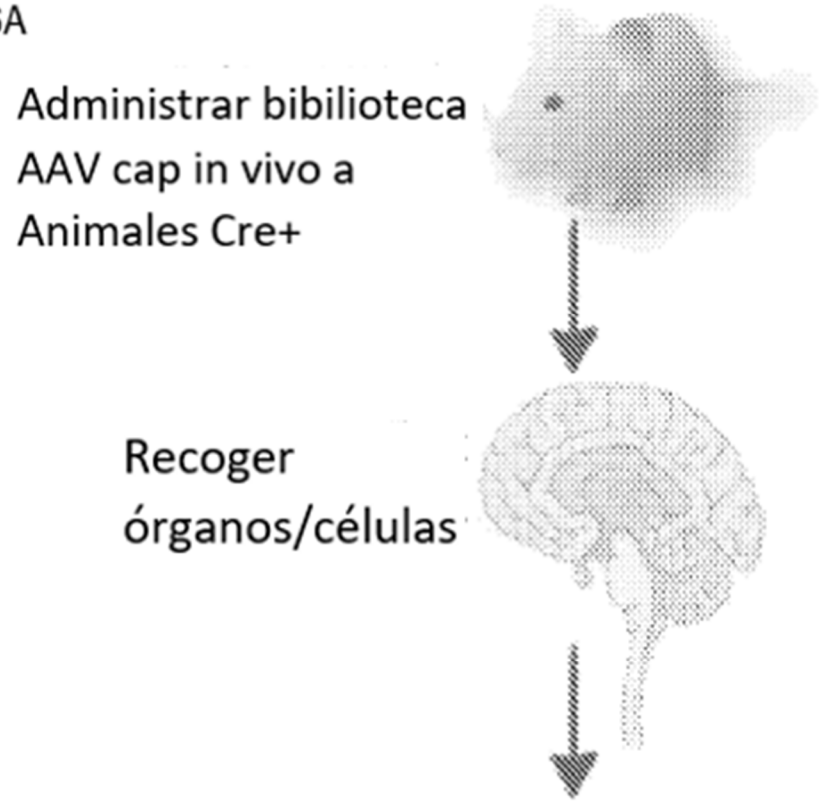


FIG. 6B

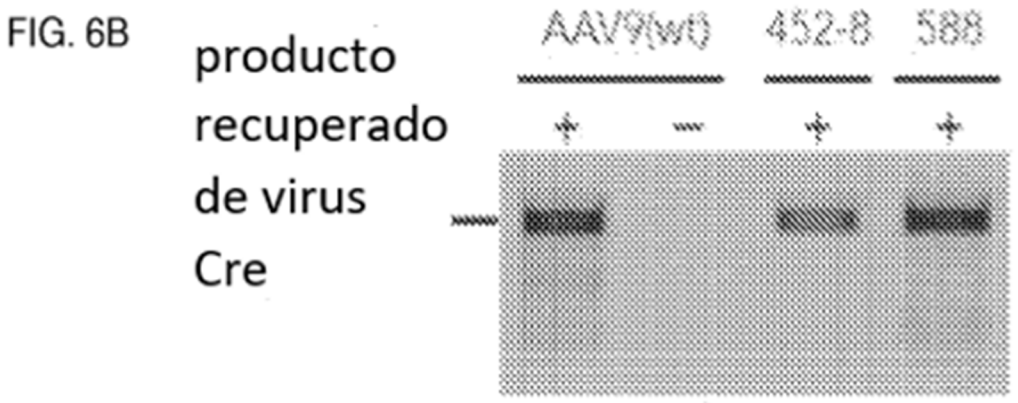
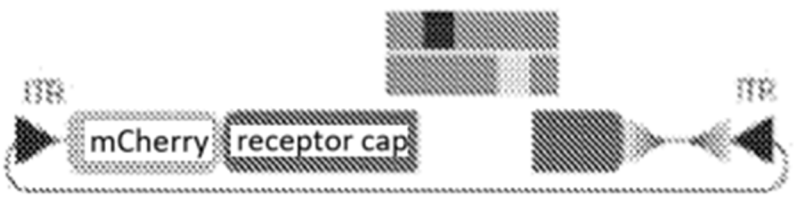


FIG. 6C



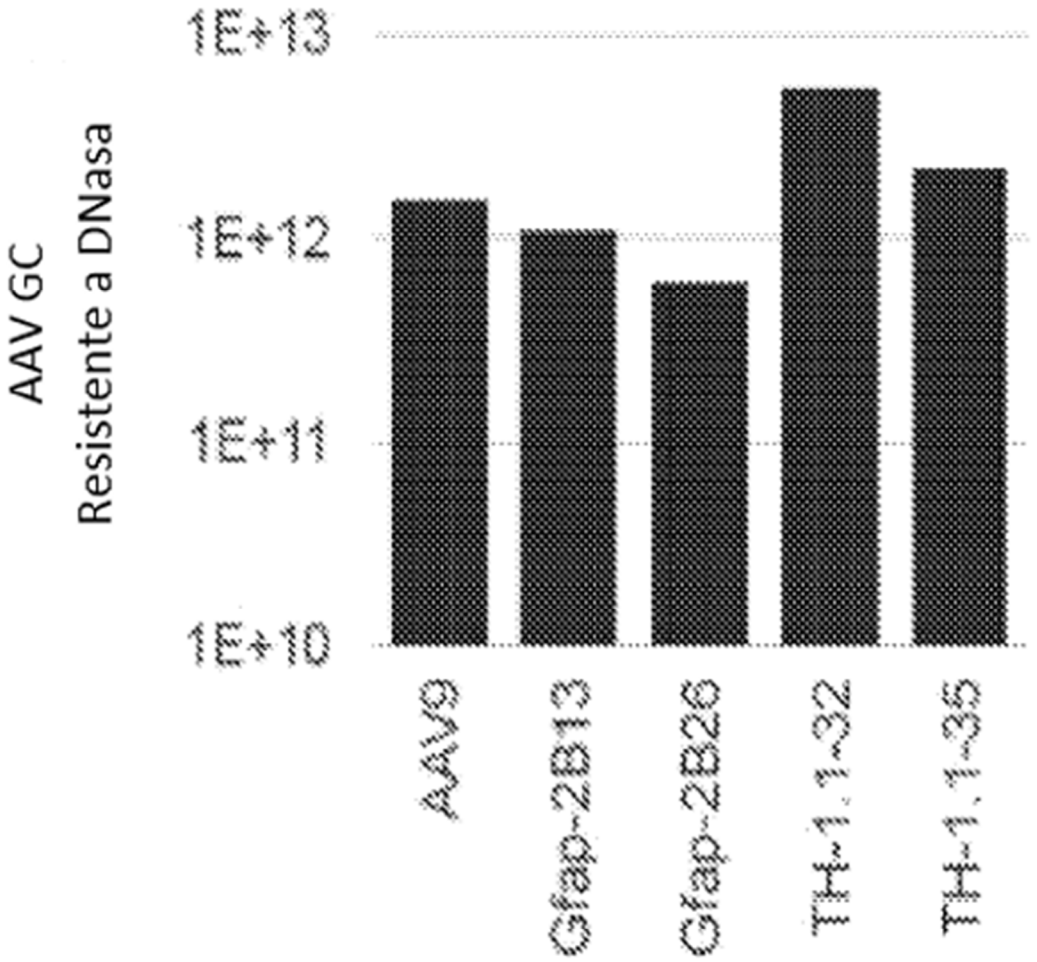

AAV GC
Resistente a DNasa
1E+10
1E+11
1E+12
1E+13
AAV9
Gfap-2B13
Gfap-2B26
TH-1.1-32
TH-1.1-35

FIG. 7

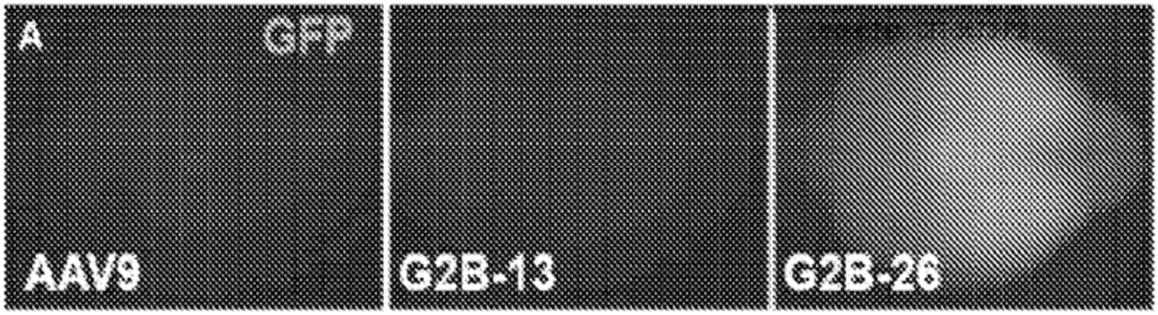


FIG. 8A

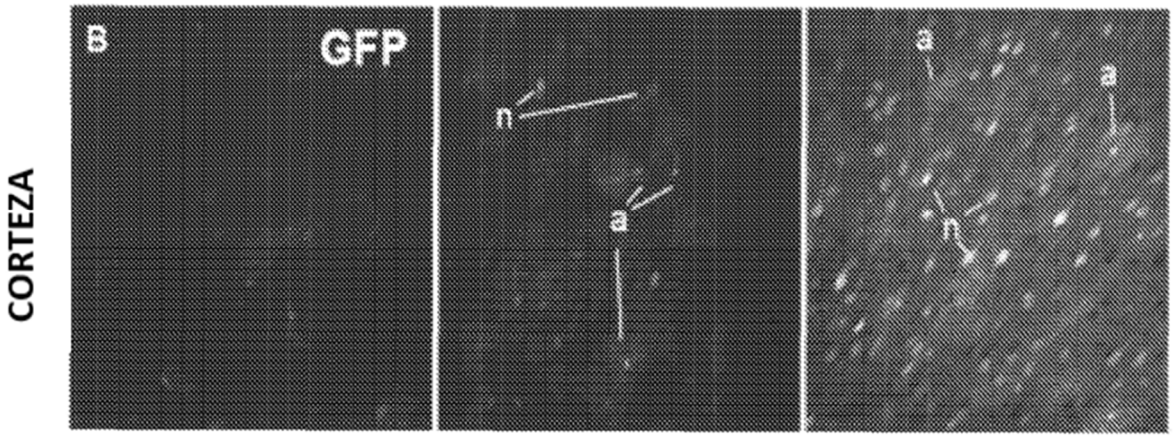


FIG. 8B

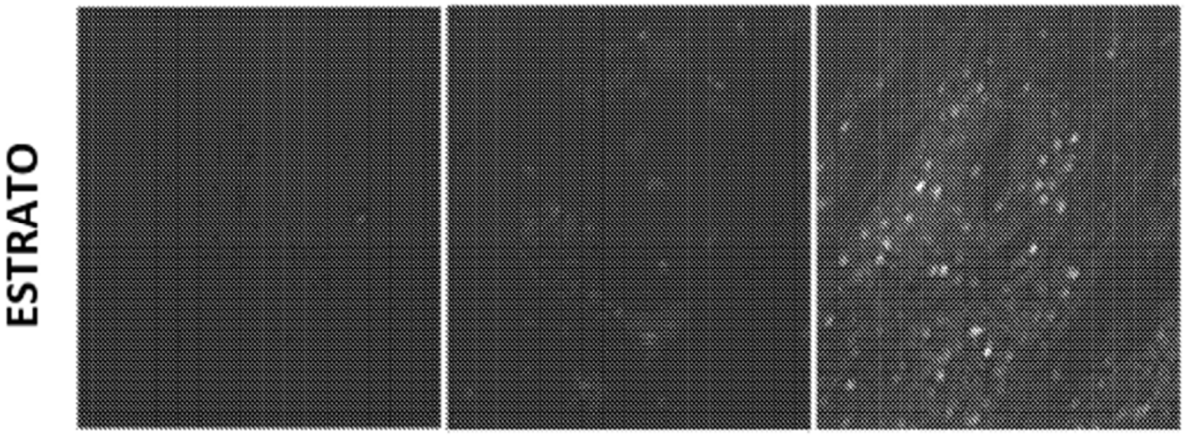


FIG. 8B
Cont.

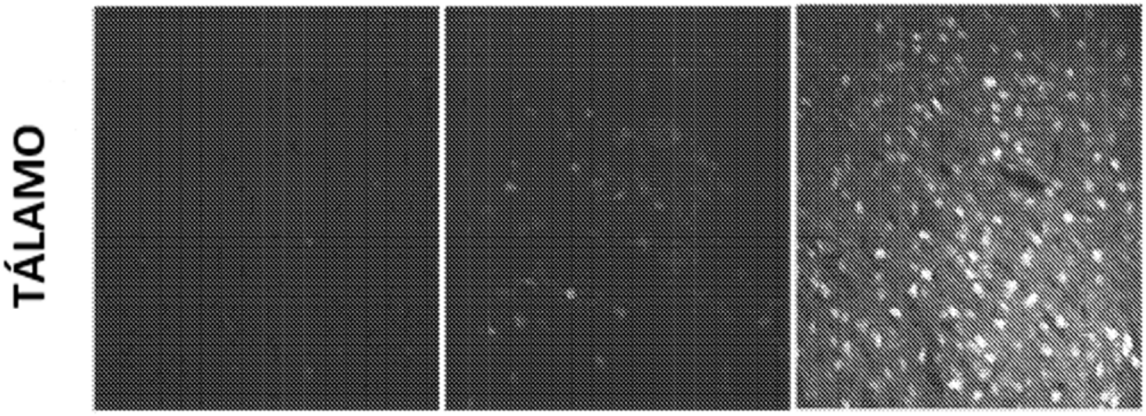


FIG. 8B
cont.

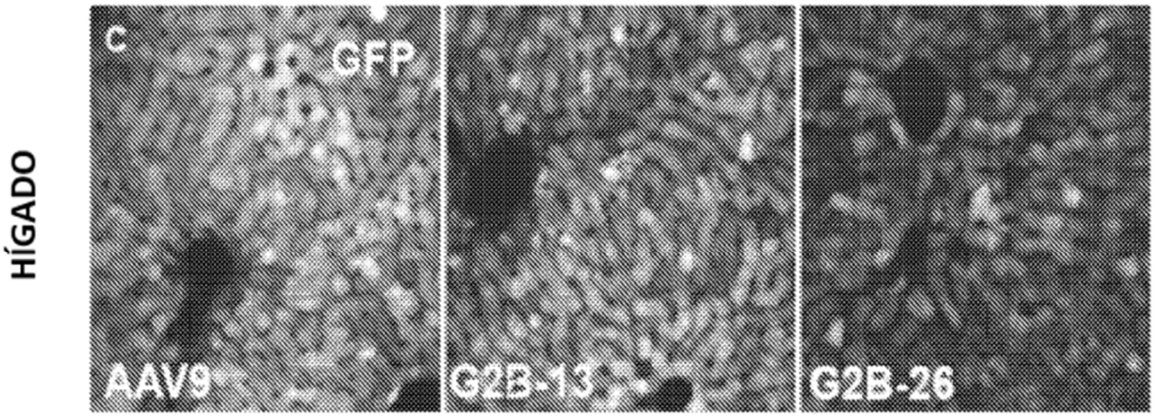
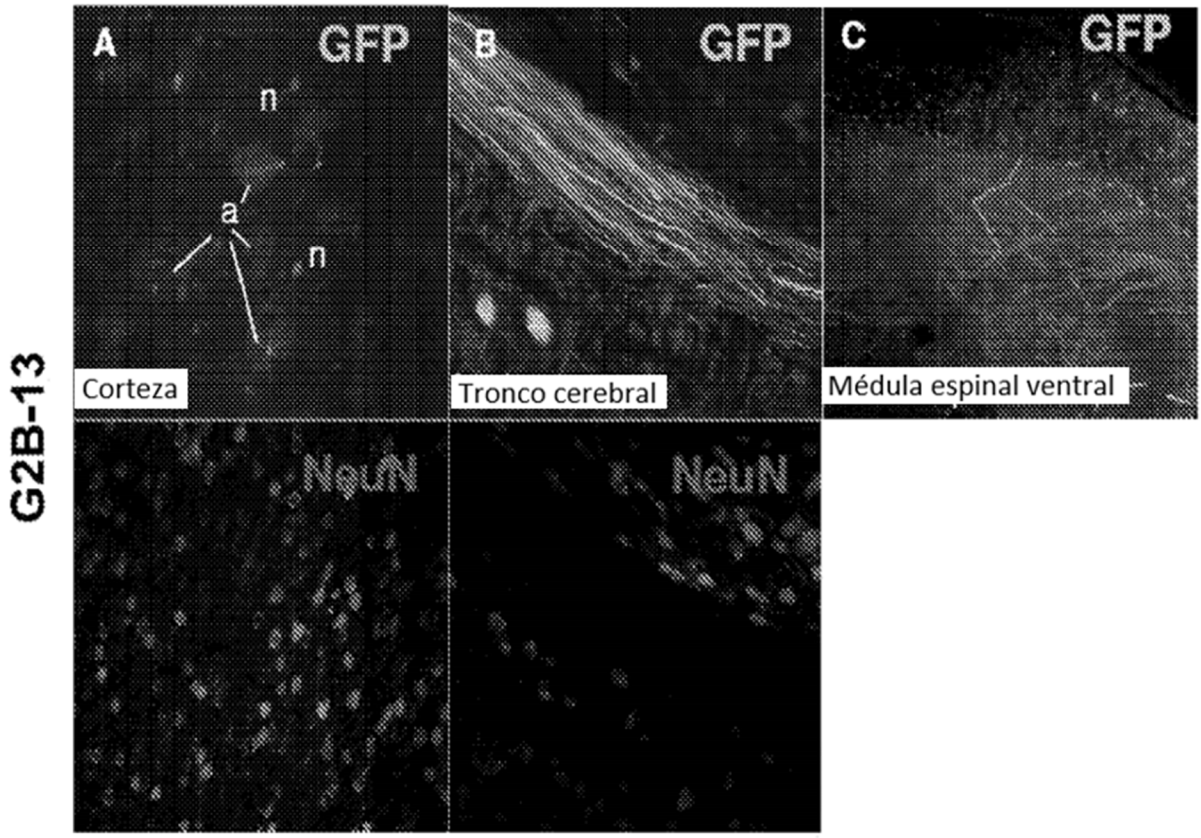
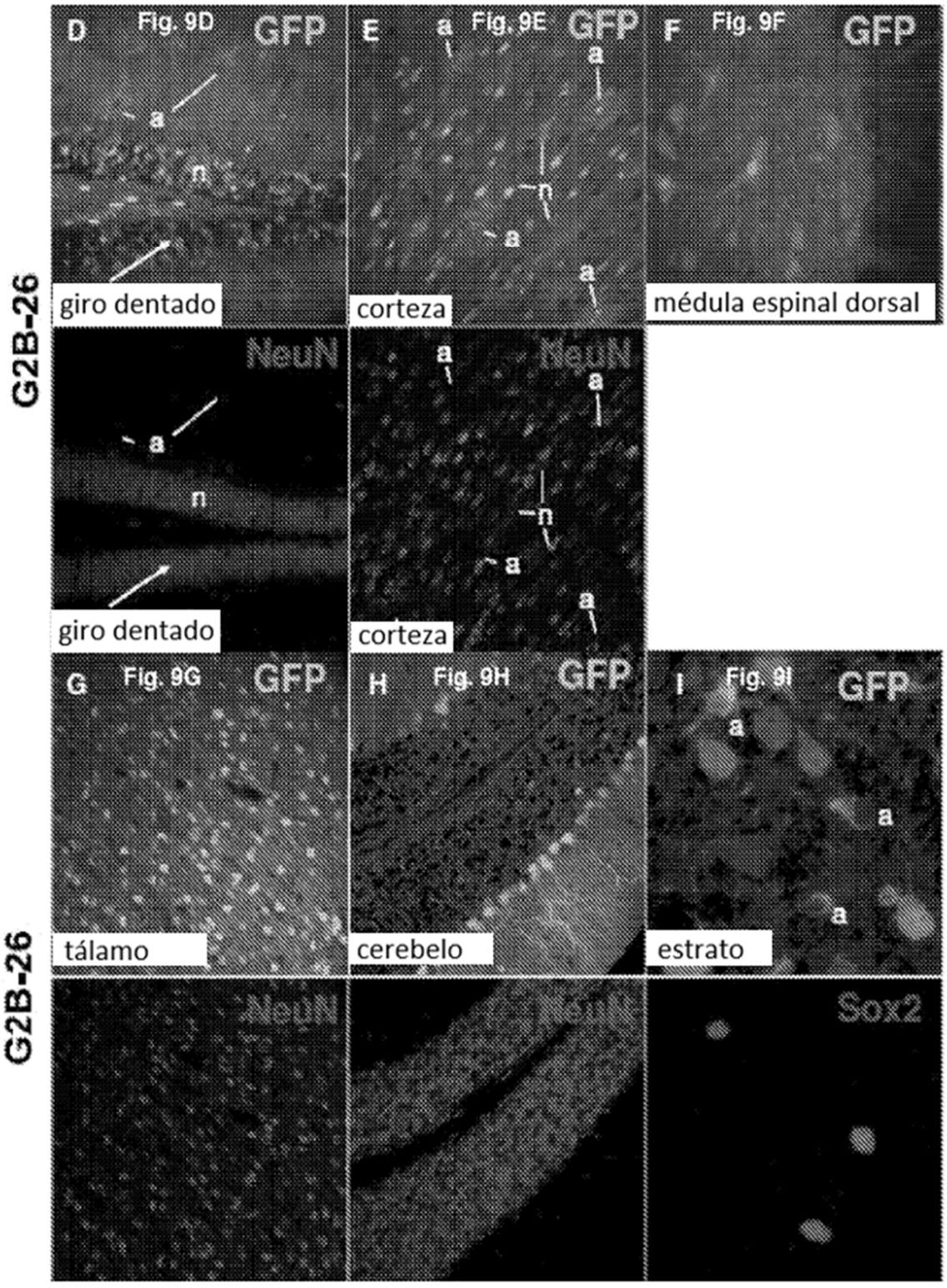


FIG. 8C

Fig. 9A Fig. 9B Fig. 9C




G2B-26
D
Fig. 9D
GFP
a
n
giro dentado
E
a
Fig. 9E
GFP
a
n
a
a
corteza
F
Fig. 9F
GFP
médula espinal dorsal
NeuN
a
n
giro dentado
a
NeuN
a
n
a
a
corteza
G2B-26
G
Fig. 9G
GFP
tálamo
H
Fig. 9H
GFP
cerebelo
I
Fig. 9I
GFP
a
a
a
estrato
NeuN
Sox2

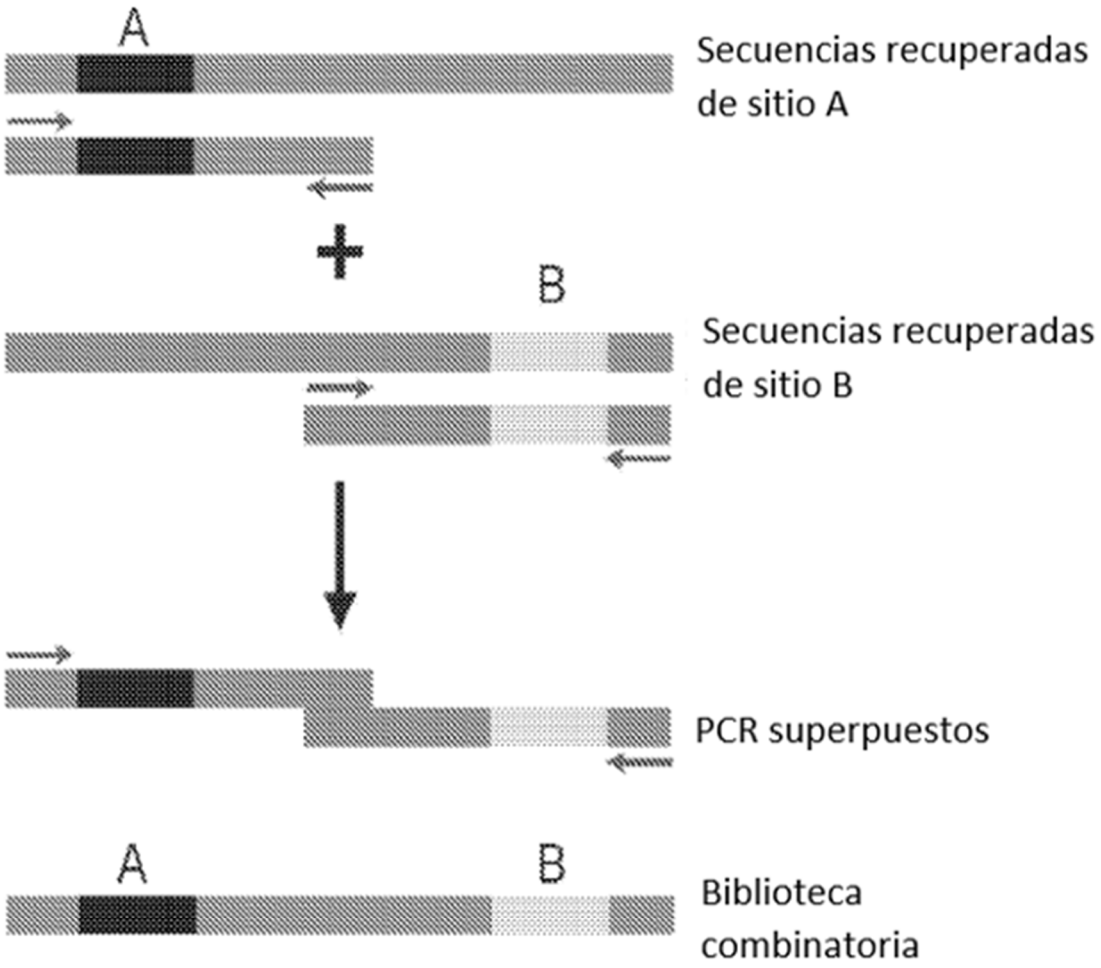

A
Secuencias recuperadas de sitio A
B
Secuencias recuperadas de sitio B
PCR superpuestos
A
B
Biblioteca combinatoria

Fig. 10

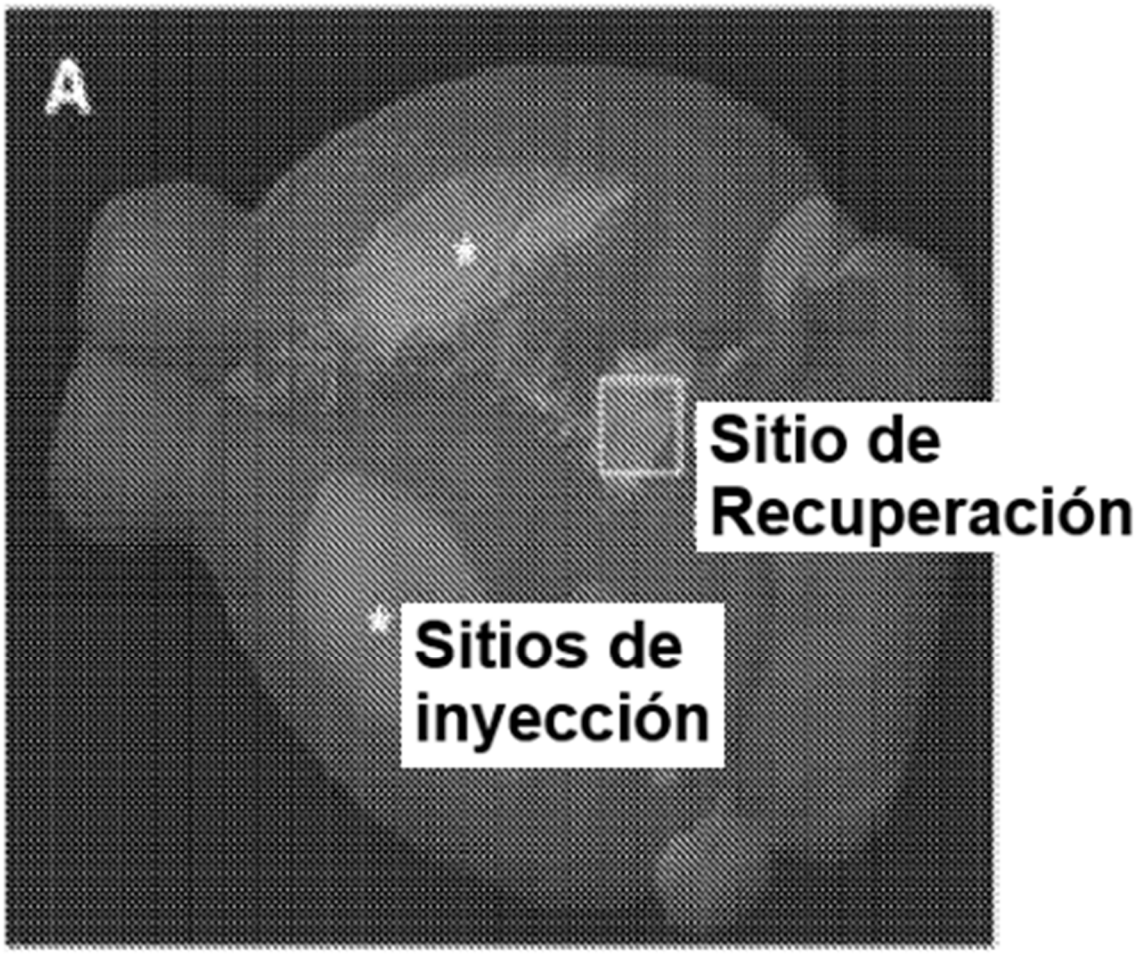


FIG. 11A

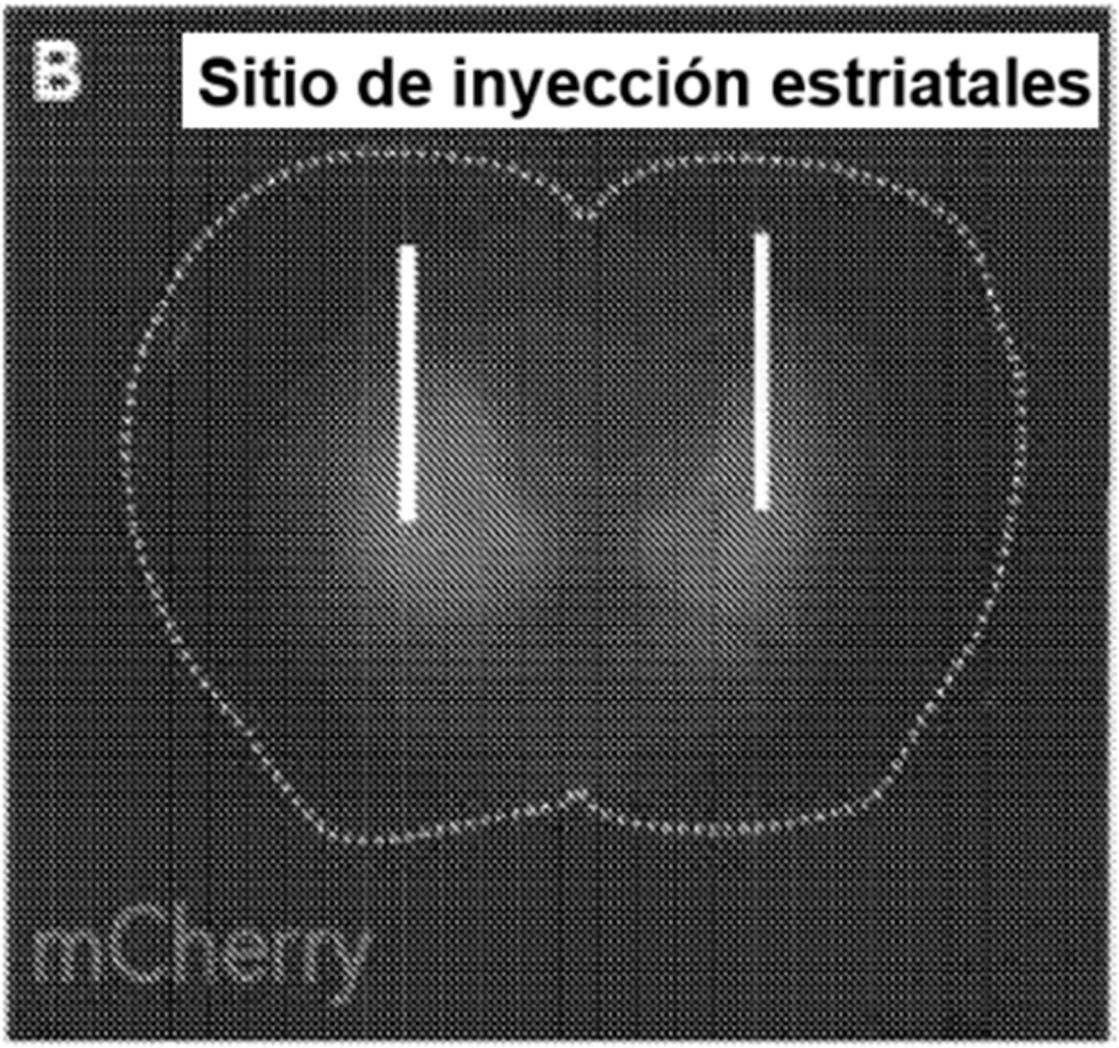


FIG. 11B

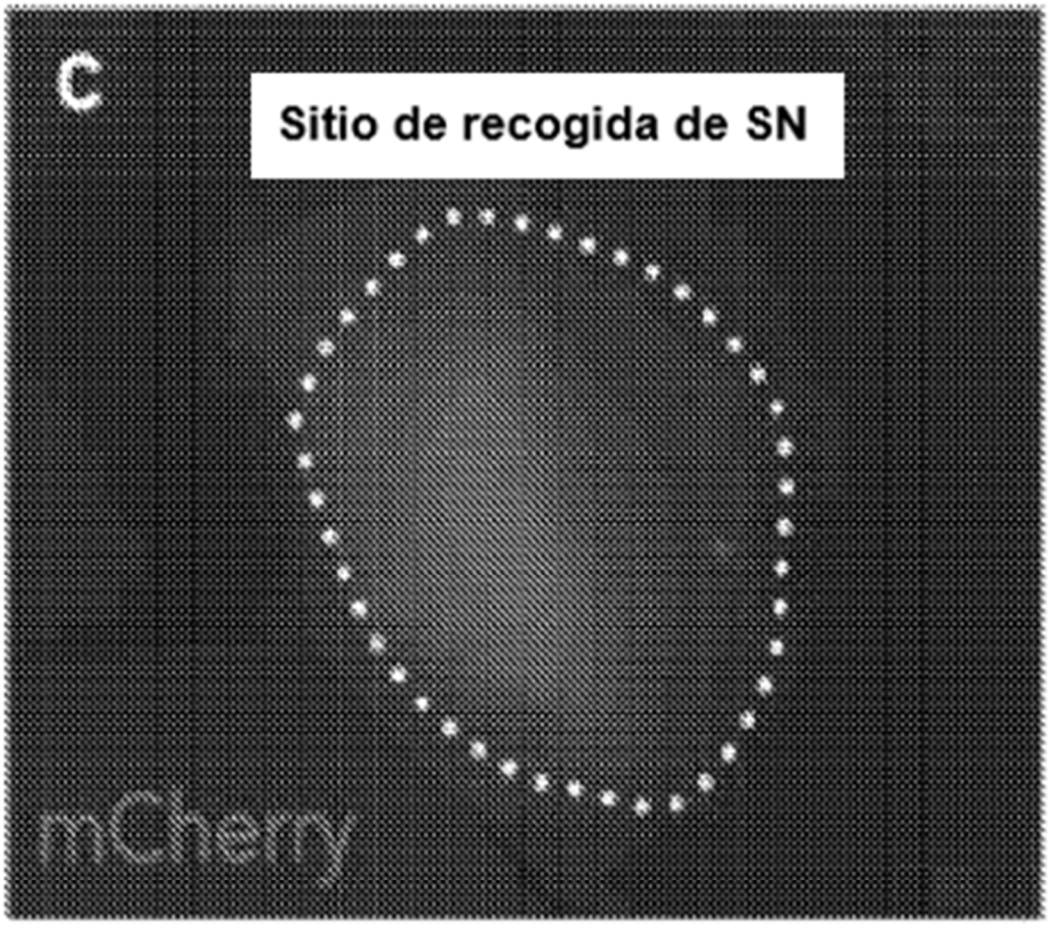


FIG. 11C

Fig. 11D

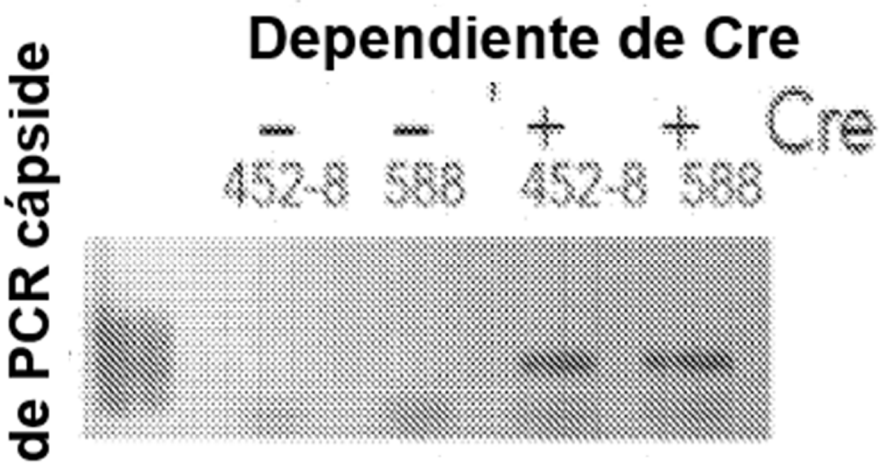


Fig. 11E

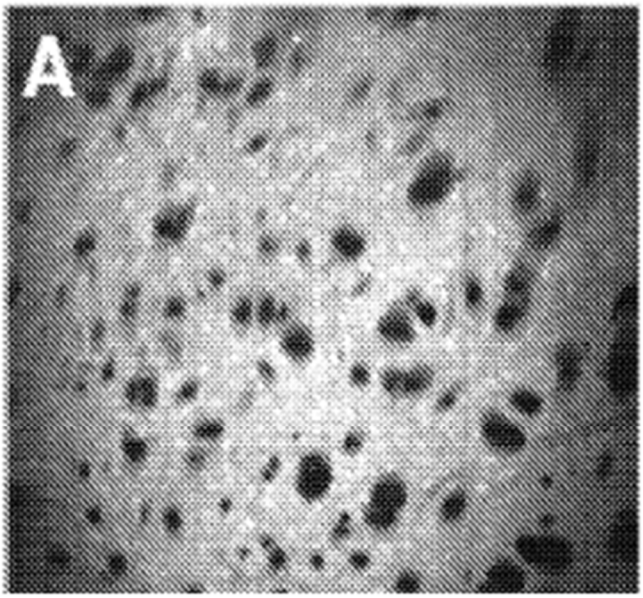


FIG. 12A

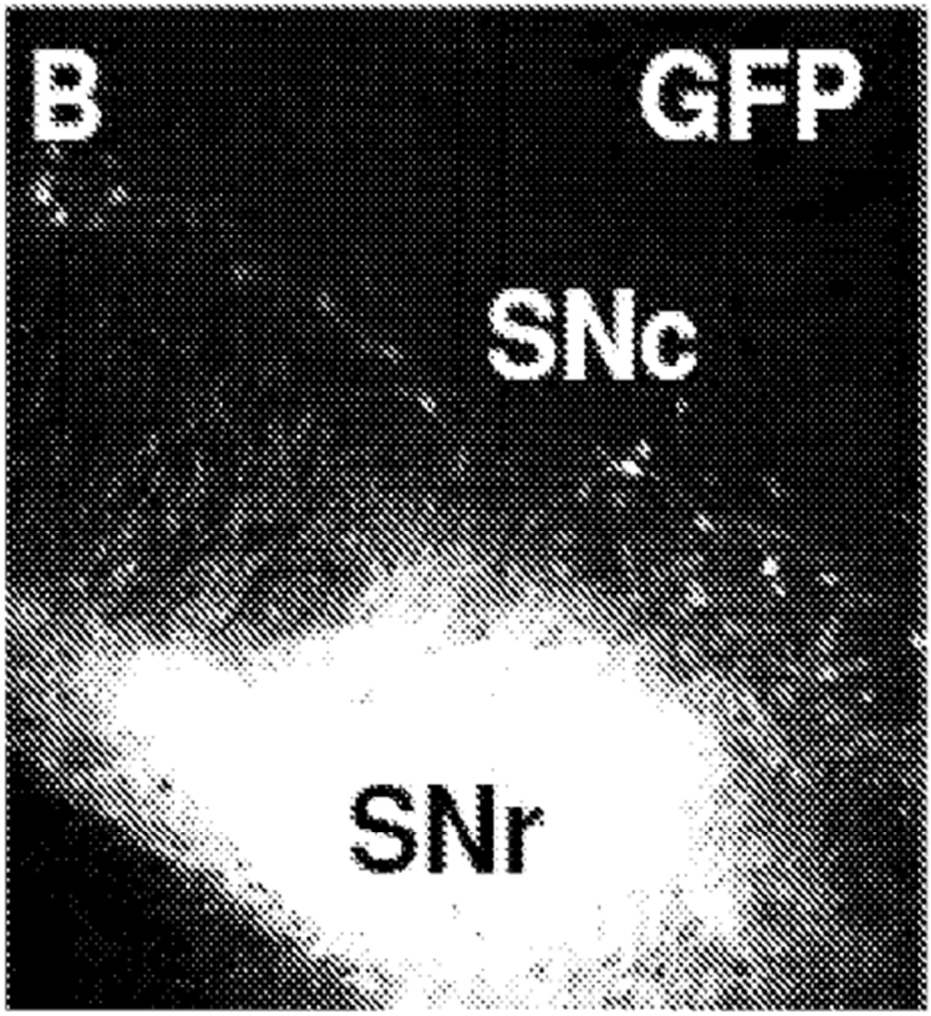


FIG. 12B

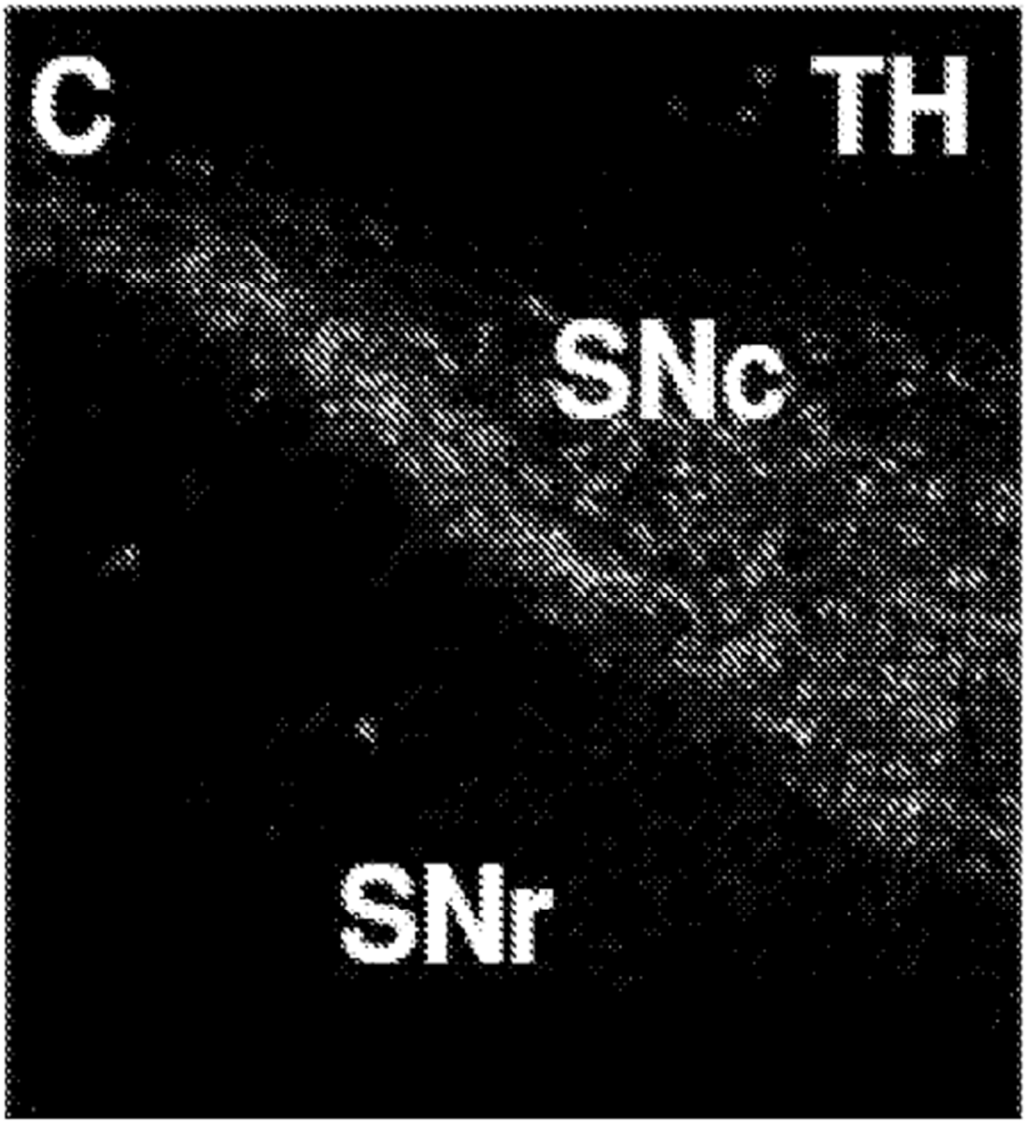


FIG. 12C

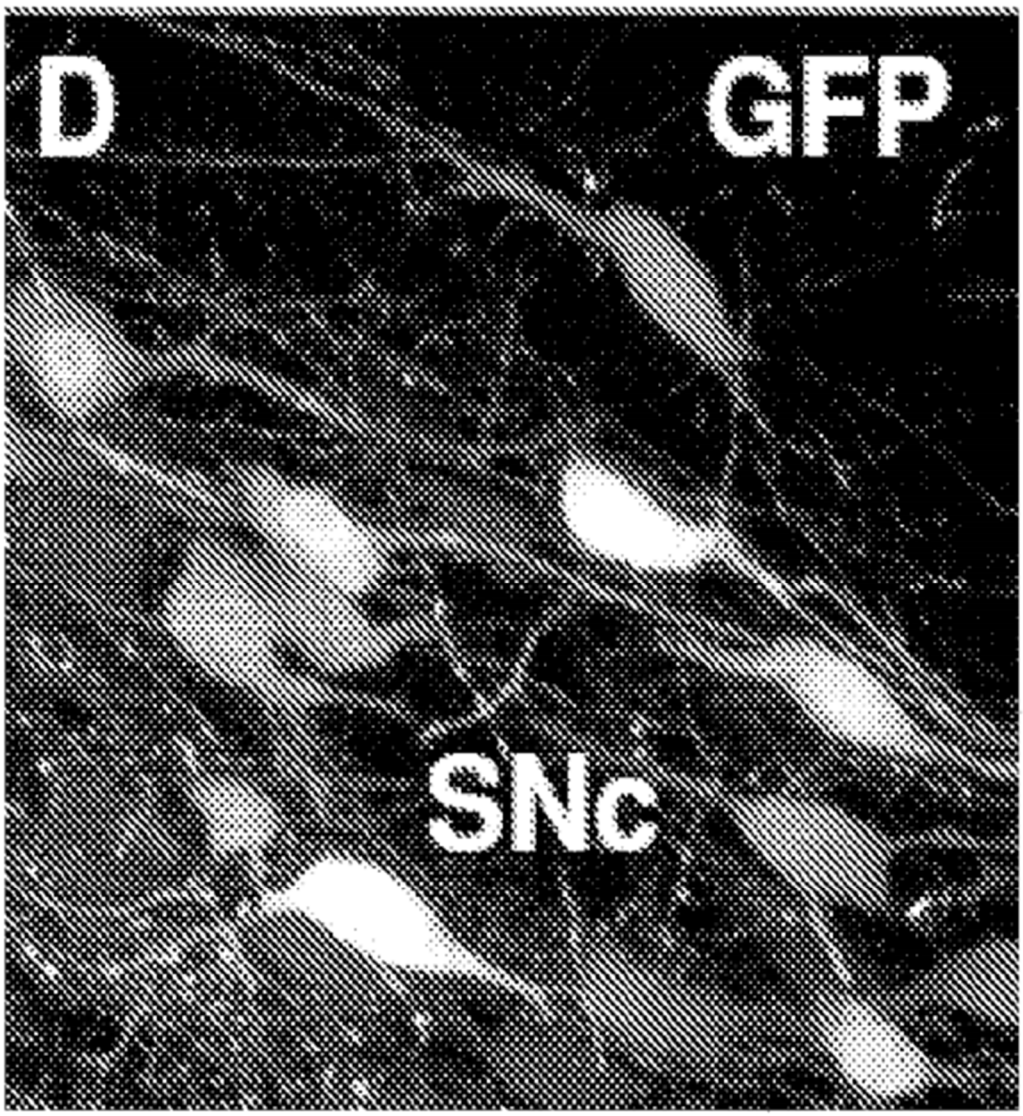


FIG. 12D

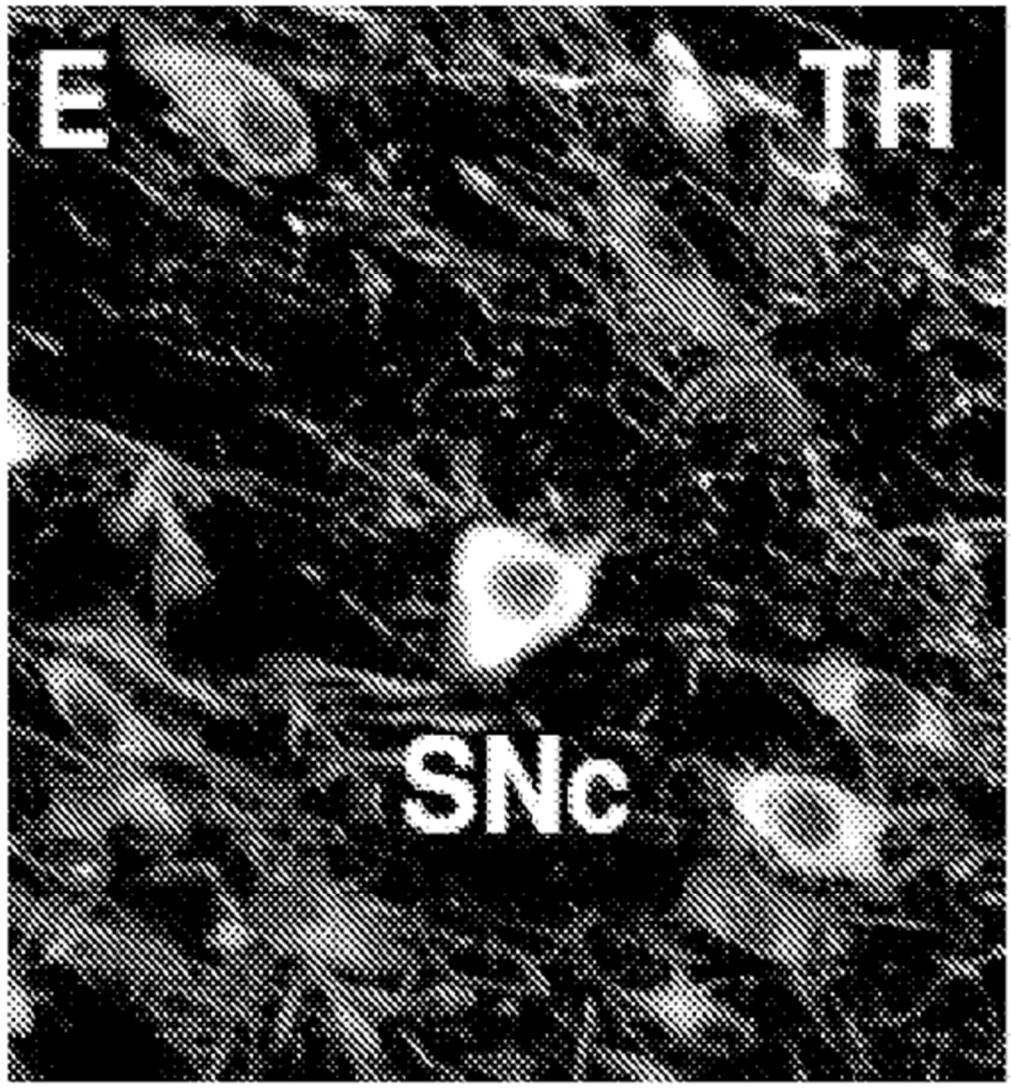


FIG. 12E

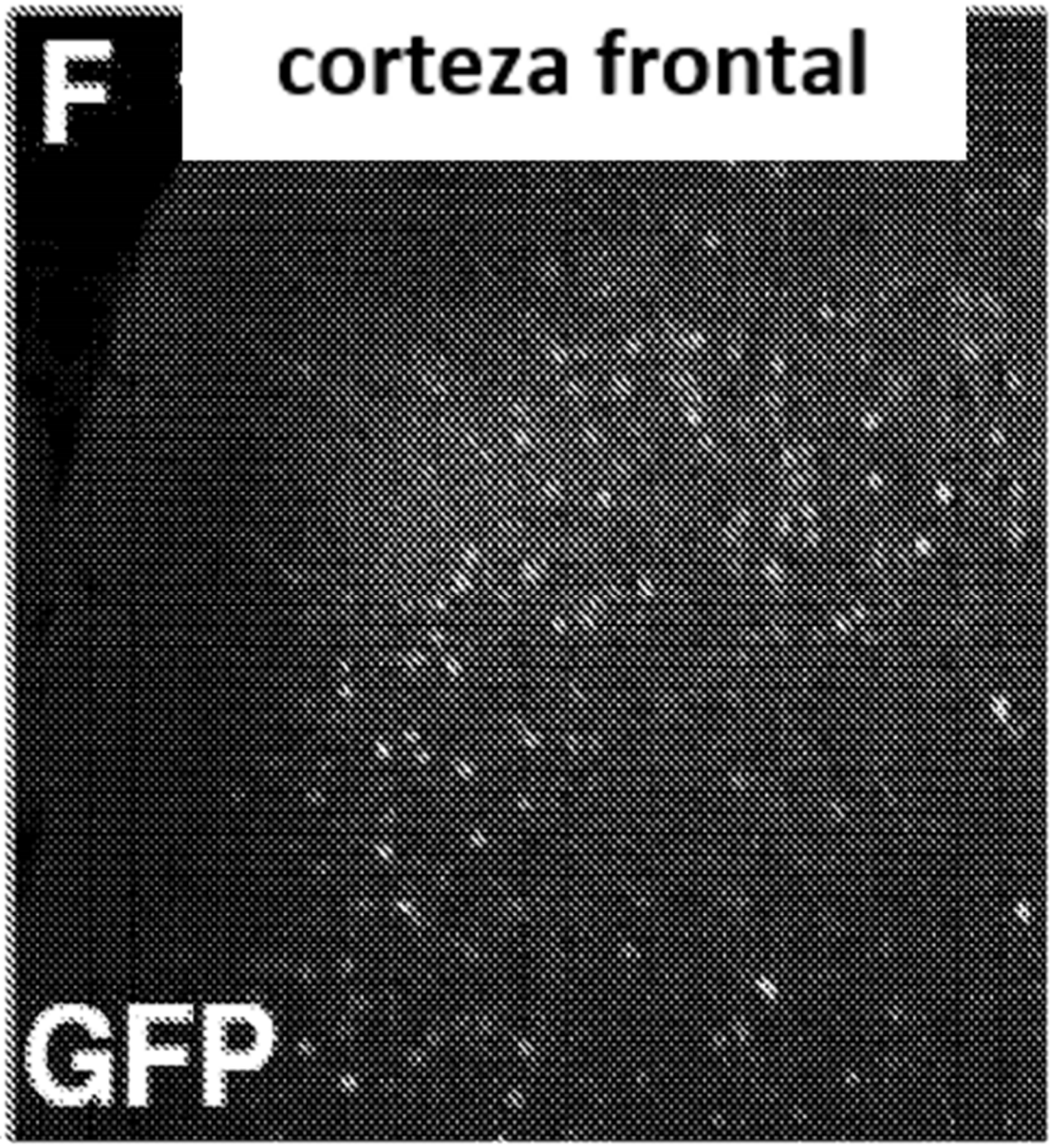


FIG. 12F

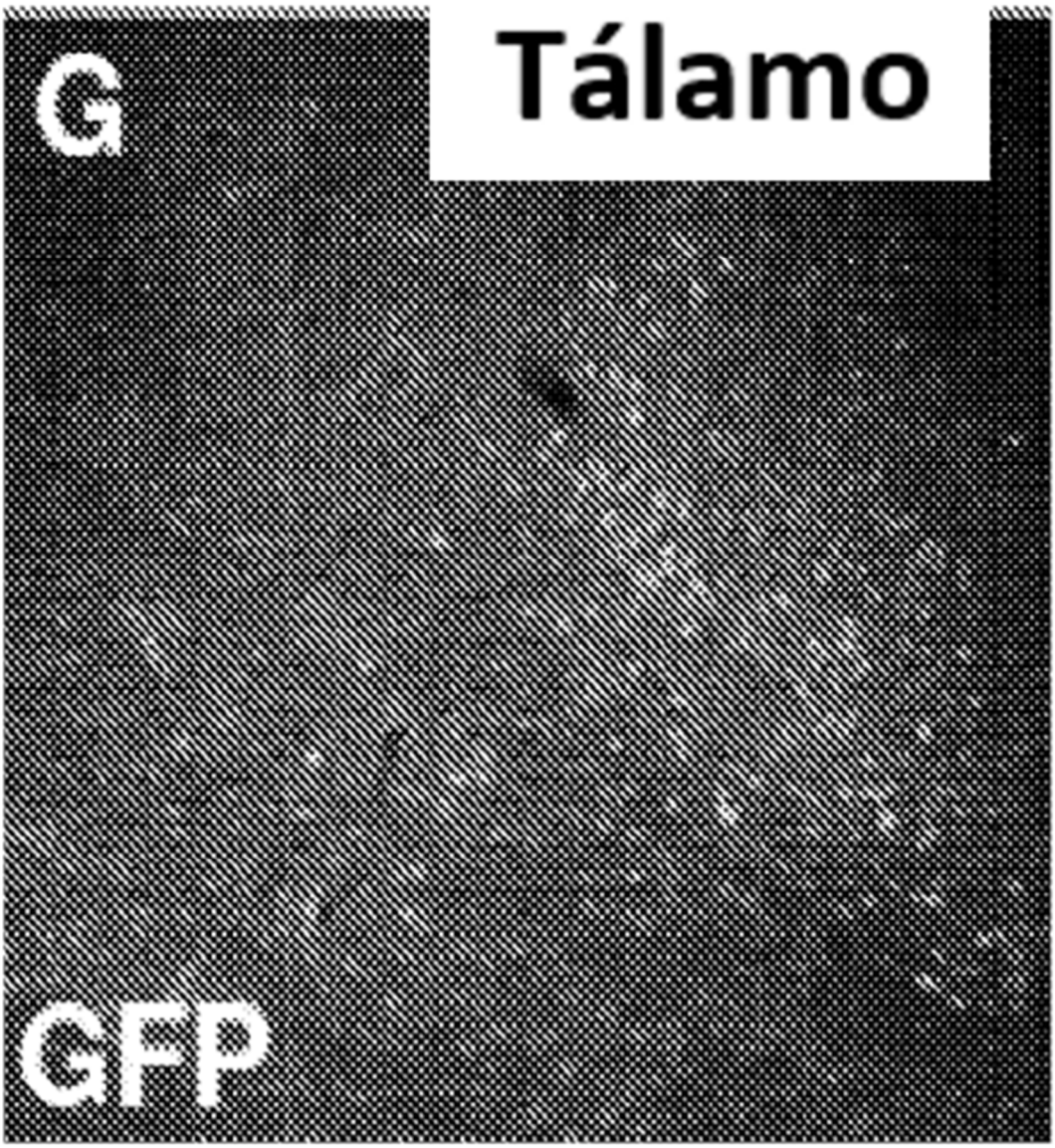


FIG. 12G

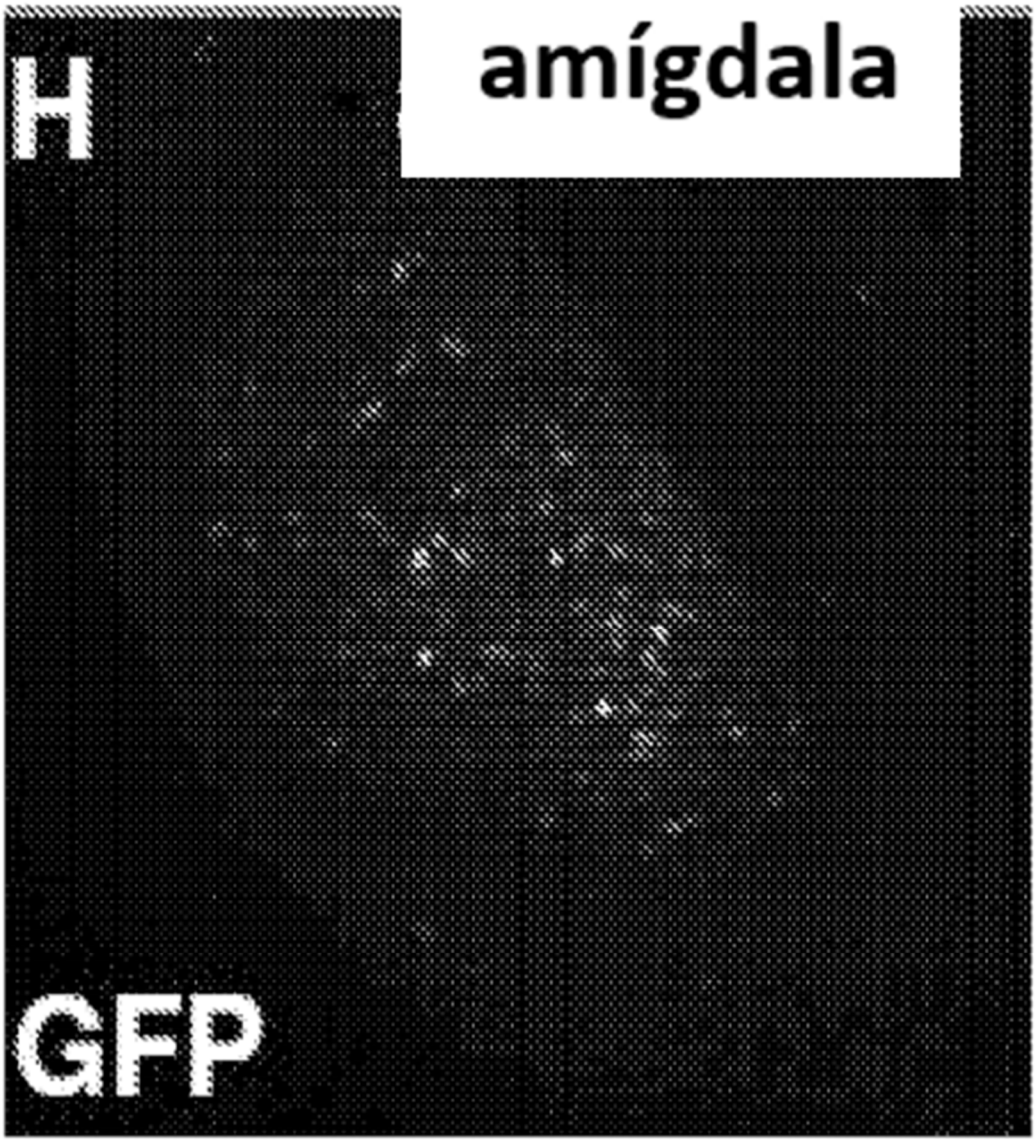


FIG. 12H

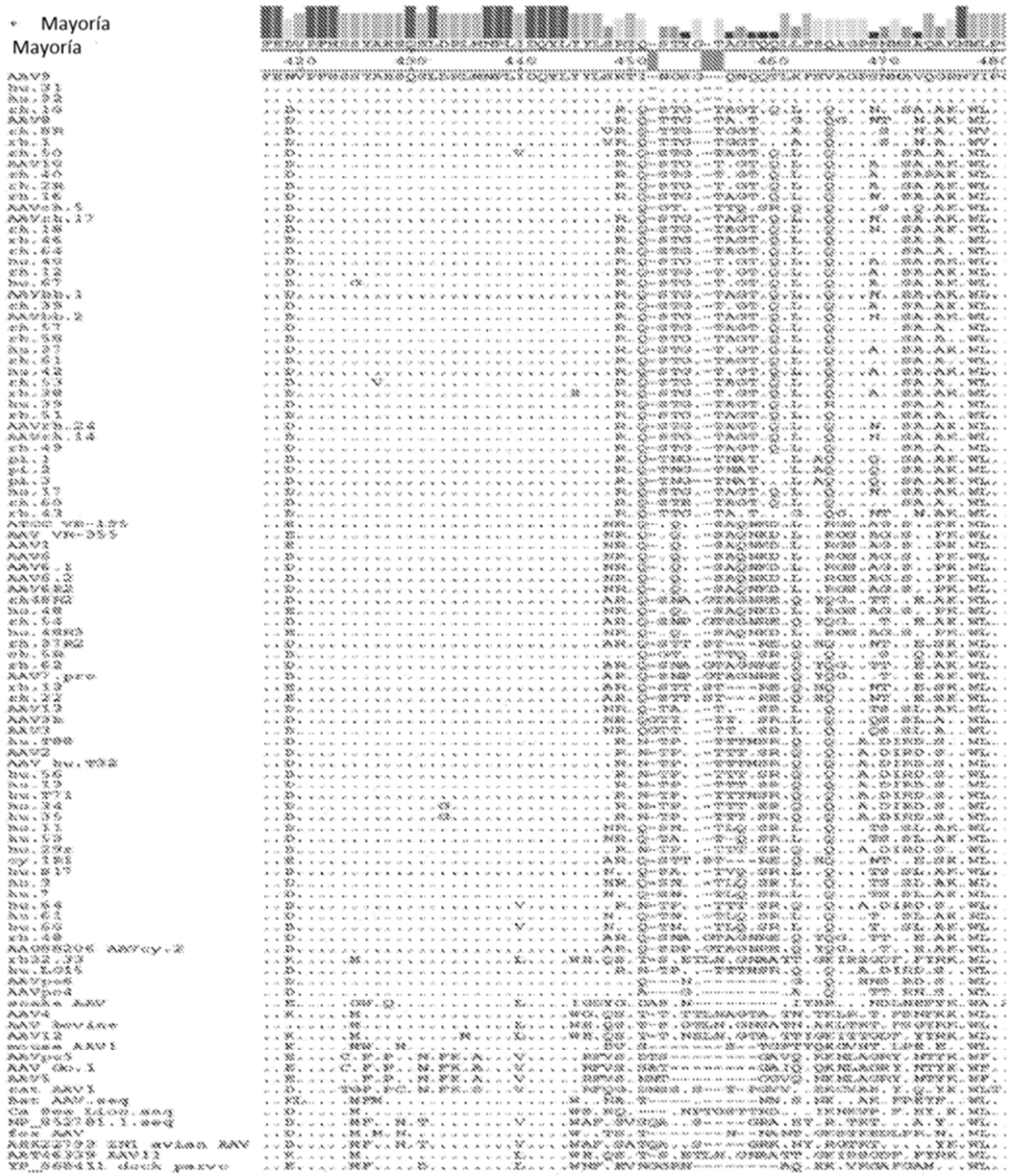


Fig. 13

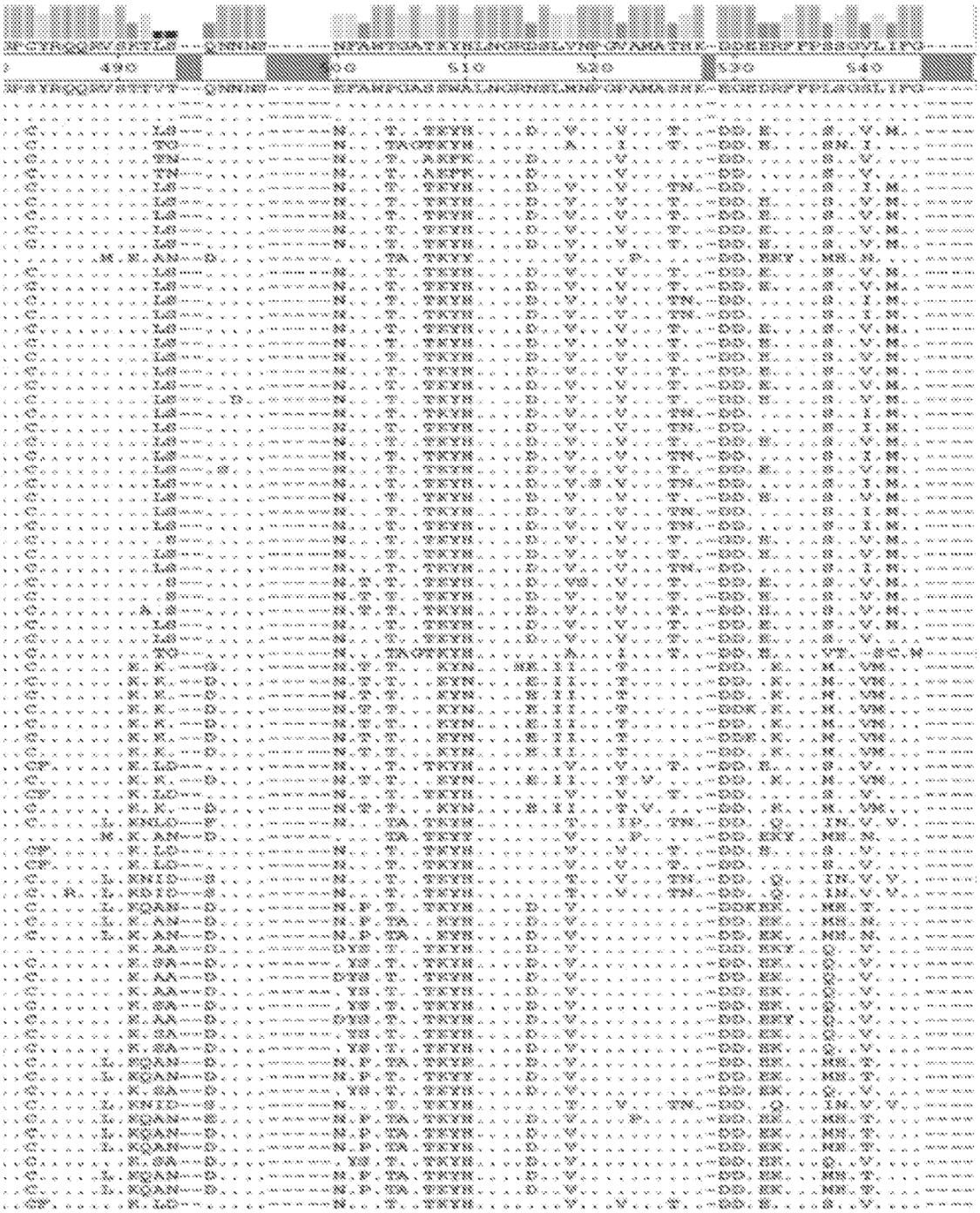


Fig. 13. Cont.

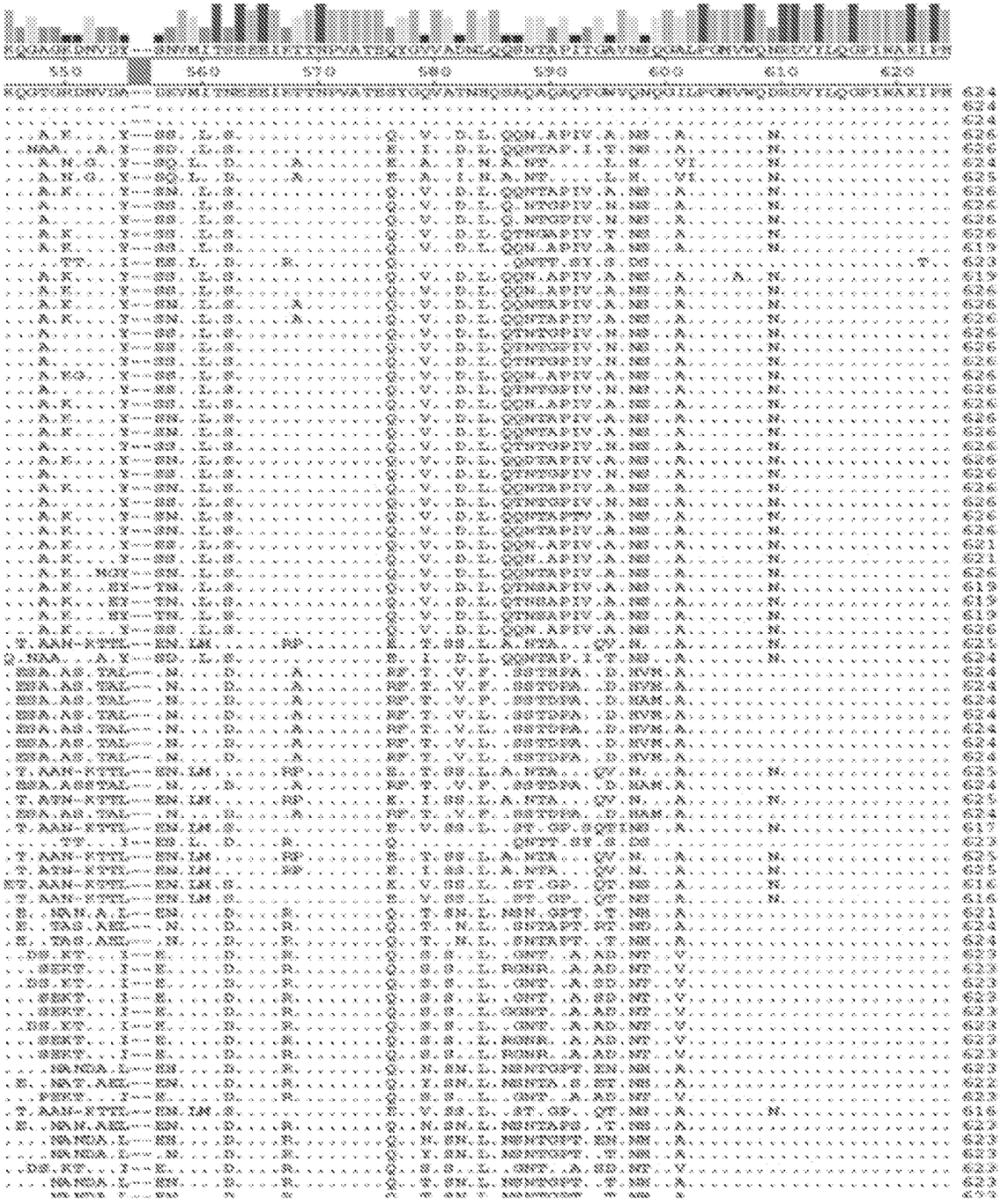


Fig. 13. Cont.

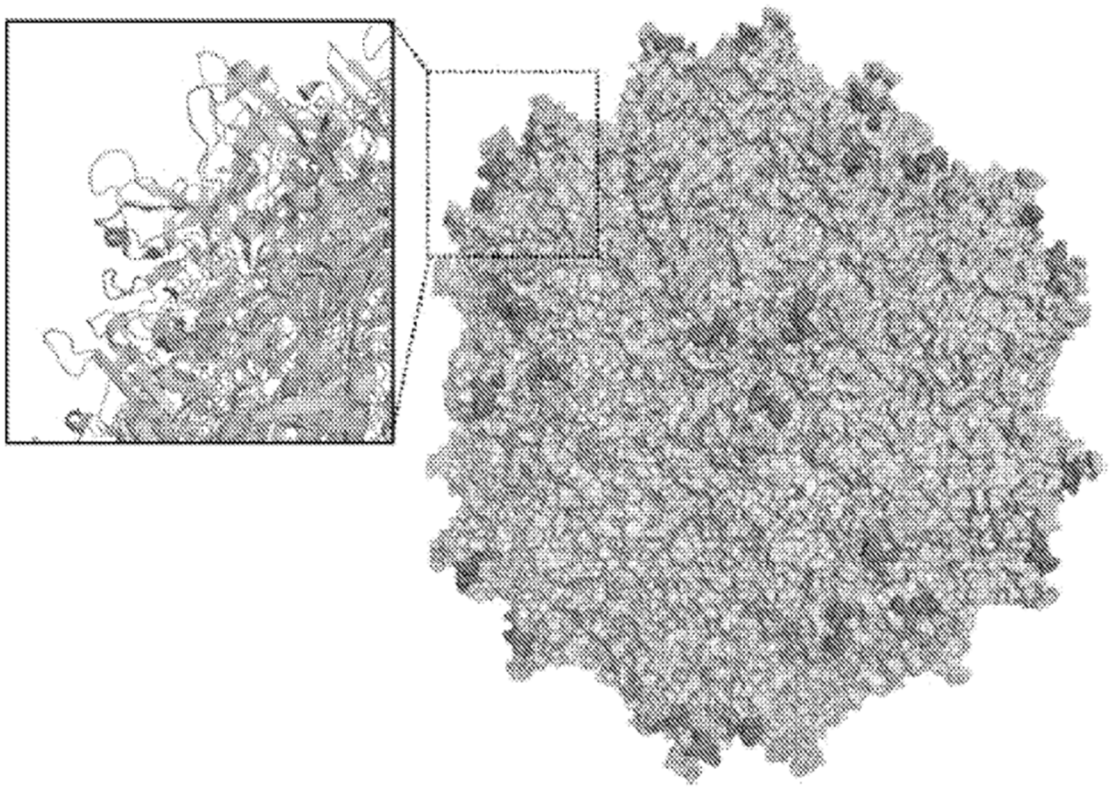


FIG. 14

FIG. 15

SEQ ID NO: 4

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGC
TTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCC
TACTAGTGGCCTCCGCGCCGGGTTTTGGCGCCTCCCGCGGGCGCCCCCCTCCTCACGGCGAGCGCTGCCACG
TCAGACGAAGGGCGCAGCGAGCGTCCTGATCCTTCCGCCCGGACGCTCAGGACAGCGGCCCGCTGCTCATAA
GACTCGGCCTTAGAACCCCAGTATCAGCAGAAGGACATTTTAGGACGGGACTTGGGTGACTCTAGGGCACTGGT
TTTCTTTCCAGAGAGCGGAACAGGCGAGGAAAAGTAGTCCCTTCTCGGCGATTCTGCGGAGGGATCTCCGTGG
GGCGGTGAACGCCGATGATTATATAAGGACGCGCCGGGTGTGGCACAGCTAGTTCCGTCGCAGCCGGGATTTG
GGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGGCGGCCGCCATGGTCAGCAAGGGCGAGGAGGA
TAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTC
GAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGG
TGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAATTCATGTATGGCTCCAAGGCCTACGTGAAGCACC
CCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGA
GGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCTTACAAGACGGCGAGTTCATCTACAAAGTGAAGCTG
CGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAG
CGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCAC
TACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACA
TCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCA
CTCCACCGGCGGCATGGACGAGCTGTACAAGTAAAGGATCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGT
GCCACTCCCACTGTCCTTTCCTCCTCCTAGGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGT
GTAGATCTGTTCAAATTTGAACTGACTAAGCGGCTCCCGCCAGATTTTGGCAAGATTACTAAGCAGGAAGTCAAG
GACTTTTTTGCTTGGGCAAAGGTCAATCAGGTGCCGGTGACTCACGAGTTTAAAGTTCCCAGGGAATTGGCGGG
AACTAAAGGGGCGGAGAAATCTCTAAAACGCCCACTGGGTGACGTCACCAATACTAGCTATAAAAGTCTGGAGA
AGCGGGCCAGGCTCTCATTTGTTCCCGAGACGCCTCGCAGTTCAGACGTGACTGTTGATCCCGCTCCTCTGCGA
CCGCTAGCTTCGATCAACTACGCGGACAGGTACCAAAACAAATGTTCTCGTCACGTGGGCATGAATCTGATGCT
GTTTCCCTGCAGACAATGCGAGAGACTGAATCAGAATTCAAATATCTGCTTCACTCACGGTGTCAAAGACTGTTT
AGAGTGCTTTCCCGTGTCAGAATCTCAACCCGTTTCTGTCGTCAAAAAGGCGTATCAGAAACTGTGCTACATTCA
TCACATCATGGGAAAGGTGCCAGACGCTTGCACTGCTTGCGACCTGGTCAATGTGGACTTGGATGACTGTGT
TTCTGAACAATAAATGACTTAAACCAGGTATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTA
GTGAAGGAATTCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCAACAACATCAAGAC
AACGCTCGAGGTCTTGTGCTTCCGGGTTACAAATACCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGG
TCAACGCAGCAGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAAGGCCGGAGACAACC
CGTACCTCAAGTACAACCACGCCGACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCAA
CCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTCTTGGTCTGGTTGAGGAAGCGGCTAAG
ACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCAAAT
CGGGTGCACAGCCCGCTAAAAAGAGACTCAATTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCCTCA
ACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGCACCA
GTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTCCCAATGGCT
GGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTGCCCACCTACAACAATCACCTCTACAAGCAA
ATCTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCCTGGGGGTATTTT
GACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACAACAACTGGGGATTCCG
GCCTAAGCGACTCAACTTCAAGCTCTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATGGAGTCAAGACCAT
CGCCAATAACCTTACCAGCACGGTCCAGGTCTTCACGGACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGG
CTCACGAGGGCTGCCTCCCGCCGTTCCCAGCGGACGTTTTCATGATTCCTCAGTACGGGTATCTGACGCTTAAT
GATGGAAGCCAGGCCGTGGGTCGTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGG
TAACAACTTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGA
CCGACTAATGAATCCACTCATCGACCAATACTTGTACTATCTCTCTAGAACTATTAACGGTTCTGGACAGAATCAA
CAAACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGGAAGAAACTACATACCTGGACCCAG
CTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCTTC
TTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGGACC
GTTTCTTTCCTTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGATGCGGACAAAG
TCATGATAACCAACGAAGAAGAAATTAAAACTACTAACCCGGTAGCAACGGAGTCCTATGGACAAGTGGCCACAA
ACCACCAGAGTGCCCAAGCACAGGCGCAGACCGGTTGGGTTCAAAACCAAGGAATACTTCCGGGTATGGTTTG
GCAGGACAGAGATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCACACGGACGGCAACTTTCACCCTT
CTCCGCTGATGGGAGGGTTTGGAATGAAGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCTGTACCTGCG
GATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTG
GAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCAACTATT
ACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAACCCCGCCCCATTGGCACCAGATA
CCTGACTCGTAATCTGTAAGTCGACTACCGTTCGTATAGCATACATTATACGAAGTTATCATATGTTCGAGCAGAC
ATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTT
GTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATG
TTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCGAGCTC
TACCGTTCGTATAATGTATGCTATACGAAGTTATGATATCAAGCTTAGGAACCCCTAGTGATGGAGTTGGCCACT
CCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTT

GCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAAGCTAGCGGGCGATTAAGGAAA
GGGCTAGATCATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATG
GTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATT
CAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTAT
TCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTG
GTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAA
GATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGT
ATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGT
ACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGA
GTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAAC
ATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTG
ACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCC
GGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGG
CTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATG
GTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATC
GCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATT
TAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGT
GAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGC
GTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACT
CTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGC
CACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGT
GGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAA
CGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCT
ATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGG
AGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGA
CTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTAATAAACACACACACACCAACAACCGTGGTTGGTTGT
TGTGTTGGTTTATTCTCGAG

SEQ ID NO:4

FIG. 15 Cont.

FIG. 16

SEQ ID NO: 5

GTCGACGGTATCGGGGGAGCTCGCAGGGTCTCCATTTTGAAGCGGGAGGTTTGAACGCGCAGCCGCCATGCC
GGGGTTTTACGAGATTGTGATTAAGGTCCCCAGCGACCTTGACGAGCATCTGCCCGGCATTTCTGACAGCTTTG
TGAACTGGGTGGCCGAGAAGGAATGGGAGTTGCCGCCAGATTCTGACATGGATCTGAATCTGATTGAGCAGGC
ACCCCTGACCGTGGCCGAGAAGCTGCAGCGCGACTTTCTGACGGAATGGCGCCGTGTGAGTAAGGCCCCGGA
GGCTCTTTTCTTTGTGCAATTTGAGAAGGGAGAGAGCTACTTCCACATGCACGTGCTCGTGGAAACCACCGGGG
TGAAATCCATGGTTTTGGGACGTTTCCTGAGTCAGATTCGCGAAAAACTGATTCAGAGAATTTACCGCGGGATCG
AGCCGACTTTGCCAAACTGGTTCGCGGTCACAAAGACCAGAAATGGCGCCGGAGGCGGGAACAAGGTGGTGGA
TGAGTGCTACATCCCCAATTACTTGCTCCCCAAAACCCAGCCTGAGCTCCAGTGGGCGTGGACTAATATGGAAC
AGTATTTAAGCGCCTGTTTGAATCTCACGGAGCGTAAACGGTTGGTGGCGCAGCATCTGACGCACGTGTCGCAG
ACGCAGGAGCAGAACAAAGAGAATCAGAATCCCAATTCTGATGCGCCGGTGATCAGATCAAAAACTTCAGCCAG
GTACATGGAGCTGGTCGGGTGGCTCGTGGACAAGGGGATTACCTCGGAGAAGCAGTGGATCCAGGAGGACCA
GGCCTCATACATCTCCTTCAATGCGGCCTCCAACTCGCGGTCCCAAATCAAGGCTGCCTTGGACAATGCGGGAA
AGATTATGAGCCTGACTAAAACCGCCCCCGACTACCTGGTGGGCCAGCAGCCCGTGGAGGACATTTCCAGCAA
TCGGATTTATAAAATTTTGGAACTAAACGGGTACGATCCCCAATATGCGGCTTCCGTCTTTCTGGGATGGGCCAC
GAAAAAGTTCGGCAAGAGGAACACCATCTGGCTGTTTGGGCCTGCAACTACCGGGAAGACCAACATCGCGGAG
GCCATAGCCCACACTGTGCCCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTTCCCTTCAACGACTGTGT
CGACAAGATGGTGATCTGGTGGGAGGAGGGGAAGATGACCGCCAAGGTCGTGGAGTCGGCCAAAGCCATTCTC
GGAGGAAGCAAGGTGCGCGTGGACCAGAAATGCAAGTCCTCGGCCCAGATAGACCCGACTCCCGTGATCGTCA
CCTCCAACACCAACATGTGCGCCGTGATTGACGGGAACTCAACGACCTTCGAACACCAGCAGCCGTTGCAAGAC
CGGATGTTCAAATTTGAACTCACCCGCCGTCTGGATCATGACTTTGGGAAGGTCACCAAGCAGGAAGTCAAAGA
CTTTTTCCGGTGGGCAAAGGATCACGTGGTTGAGGTGGAGCATGAATTCTACGTCAAAAAGGGTGGAGCCAAGA
AAAGACCCGCCCCCAGTGACGCAGATATAAGTGAGCCCAAACGGGTGCGCGAGTCAGTTGCGCAGCCATCGAC
GTCAGACGCGGAAGCTTCGATCAACTACGCGGACAGGTACCAAAACAAATGTTCTCGTCACGTGGGCATGAATC
TGATGCTGTTTCCCTGCAGACAATGCGAGAGACTGAATCAGAATTCAAATATCTGCTTCACTCACGGTGTCAAAG
ACTGTTTAGAGTGCTTTCCCGTGTCAGAATCTCAACCCGTTTCTGTCGTCAAAAAGGCGTATCAGAAACTGTGCT
ACATTCATCACATCATGGGAAAGGTGCCAGACGCTTGCACTGCTTGCGACCTGGTCAATGTGGACTTGGATGAC
TGTGTTTCTGAACAATAAATGACTTAAACCAGGTATGGCTGCCGATGGTTAACTTCCAGATTGACTCGAGGACAA
CCTTAGTGAAGGAATTCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCAACAACATC
AAGACAACGCTCGAGGTCTTGTGCTTCCGGGTTACAAATACCTTGGACCCGGCAACGGACTCGACAAGGGGGA
GCCGGTCAACGCAGCAGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAAGGCCGGAGA
CAACCCGTACCTCAAGTACAACCACGCCGACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGG
GGCAACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTCTTGGTCTGGTTGAGGAAGCGG
CTAAGACGGCTCCTGGATAGAAGAGGCCTGTAGAGTAGTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGG
CAAATCGGGTGCACAGCCCGCTAAAAAGAGACTCAATTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACC
CTCAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGC
ACCAGTGGCAGACAATAACTAAGGTGCCGATGGAGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTCCCAAT
GGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTGCCCACCTACAACAATCACCTCTACAA
GCAAATCTCCAACAGCACATCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCCTGGGGGT
ATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACAACAACTGGGGAT
TCCGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATGGAGTCAAG
ACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCACGGACTCAGACTATCAGCTCCCGTACGTGCTCGG
GTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCCAGCGGACGTTTTCATGATTCCTCAGTACGGGTATCTGACG
CTTAATGATGGAAGCCAGGCCGTGGGTCGTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGA
ACGGGTAACAACTTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCACAGCCAAAG
CCTGGACCGACTAATGAATCCACTCATCGACCAATACTTGTACTATCTCTCAAAGACTATTAACGGTTCTGGACA
GAATCAACAAACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGGAAGAAACTACATACCTG
GACCCAGCTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGA
GCTTCTTCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGG
AGAGGACCGTTTCTTTCCTTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGATG
CGGACAAAGTCATGATAACCAACGAAGAAGAAATTAAAACTACTAACCCGGTAGCAACGGAGTCCTATGGACAA
GTGGCCACAAACCACCAGAGTGCCCAAGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGAATACTTCCGG
GTATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCACACGGACGGCAAC
TTTCACCCTTCTCCGCTGATGGGAGGGTTTGGAATGAAGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCT
GTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCAGTATTCTACTGGCCAA
GTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTT
CCAACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAACCCCGCCCCATTGG
CACCAGATACCTGACTCGTAATCTGTAATTGCTTGTTAATCAATAAACCGTTTAATTCGTTTCAGTTGAACTTTGGT
CTCTGCGAAGGGCGAATTCGTTTAAACCTGCAGGACTAGAGGTCCTGTATTAGAGGTCACGTGAGTGTTTTGCG
ACATTTTGCGACACCATGTGGTCACGCTGGGTATTTAAGCCCGAGTGAGCACGCAGGGTCTCCATTTTGAAGCG
GGAGGTTTGAACGCGCAGCCGCCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGACTAGAGC
GGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGG
TCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGT
AAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGG
AAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCT
TCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGG
CGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGC

CAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATC
GACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCT
CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCG
CTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGA
ACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACT
TATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT
GAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGC
AGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA
ACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAA
ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA
CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGG
GAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAG
CAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATT
AATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGC
ATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATG
ATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAG
TGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC
TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATAC
GGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCT
CAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTA
CTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACG
GAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGAT
ACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTAAATT
GTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAAT
CGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCC
ACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAAC
CATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGA
TTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCT
AGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAG
GGCGCGTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTAC
GCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACG
TTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTC
GATCGAG

SEQ ID NO: 5

FIG. 16 Cont.

FIG. 17

SEQ ID NO: 6

AGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
CGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTAC
ACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACC
ATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCAAGCTTGGTACCGAGCTCGGAT
CCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTACTCATCGACCAATACTTGTACTATCTCTCTAGAAC
TATTAACGGTTCTGGACAGAATCAACAAACGCTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGG
GAAGAAACTACATACCTGGACCCAGCTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAACAGC
GAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTATG
GCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCCTTTGTCTGGATCTTTAATTTTTGGCAAACAAGGTACCGGC
AGAGACAACGTGGATGCGGACAAAGTCATGATAACCAACGAAGAAGAAATTAAAACTACTAACCCGGTAGCAAC
GGAGTCCTATGGACAAGTGGCCACAAACCACCAGAGTGCCCAAGCACAGGCGCAGACCGGTTGGGTTCAAAAC
CAAGGAATACTTCCAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGG
CCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCT
GGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCAC
CGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGG
CGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTT
CCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGAT
TTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGA
TAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACA
CTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGC
TGATTTAACAAAAATTTAACGCGAATTTTAACAAAATTCAGGGCGCAAGGGCTGCTAAAGGAAGCGGAACACGTA
GAAAGCCAGTCCGCAGAAACGGTGCTGACCCCGGATGAATGTCAGCTACTGGGCTATCTGGACAAGGGAAAAC
GCAAGCGCAAAGAGAAAGCAGGTAGCTTGCAGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTATGGA
CAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGAT
GGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTC
GCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTG
GGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTT
GTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACG
ACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAG
TGCCGGGGCAGGATCTCCTGTCATCCCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGG
CGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTA
CTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACT
GTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCC
GAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATC
AGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTT
TACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAATTGAAAAA
GGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGC
TCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGG
ATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTC
TGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAG
AATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGT
GCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAAC
CGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATAC
CAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA
CTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC
GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAG
CACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAA
CGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATAT
ATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACC
AAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGAT
CCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGAT
CAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTG
TAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCA
GTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGC
AGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGAT
ACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGG
CAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGG
TTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAG
CAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGAT
TCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCG
AGTCAGTGAGCGAGGAAGCGGAAG

SEQ ID NO: 6

FIG. 18

SEQ ID NO: 7

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGC
TTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCC
TACTAGTGGCCTCCGCGCCGGGTTTTGGCGCCTCCCGCGGGCGCCCCCCTCCTCACGGCGAGCGCTGCCACG
TCAGACGAAGGGCGCAGCGAGCGTCCTGATCCTTCCGCCCGGACGCTCAGGACAGCGGCCCGCTGCTCATAA
GACTCGGCCTTAGAACCCCAGTATCAGCAGAAGGACATTTTAGGACGGGACTTGGGTGACTCTAGGGCACTGGT
TTTCTTTCCAGAGAGCGGAACAGGCGAGGAAAAGTAGTCCCTTCTCGGCGATTCTGCGGAGGGATCTCCGTGG
GGCGGTGAACGCCGATGATTATATAAGGACGCGCCGGGTGTGGCACAGCTAGTTCCGTCGCAGCCGGGATTTG
GGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGGCGGCCGCCATGGTCAGCAAGGGCGAGGAGGA
TAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTC
GAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGG
TGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAATTCATGTATGGCTCCAAGGCCTACGTGAAGCACC
CCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGA
GGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCTTACAAGACGGCGAGTTCATCTACAAAGTGAAGCTG
CGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAG
CGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCAC
TACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACA
TCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCA
CTCCACCGGCGGCATGGACGAGCTGTACAAGTAAAGGATCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGT
GCCACTCCCACTGTCCTTTCCTCCTCCTAGGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGT
GTAGATCTGTTCAAATTTGAACTGACTAAGCGGCTCCCGCCAGATTTTGGCAAGATTACTAAGCAGGAAGTCAAG
GACTTTTTTGCTTGGGCAAAGGTCAATCAGGTGCCGGTGACTCACGAGTTTAAAGTTCCCAGGGAATTGGCGGG
AACTAAAGGGGCGGAGAAATCTCTAAAACGCCCACTGGGTGACGTCACCAATACTAGCTATAAAAGTCTGGAGA
AGCGGGCCAGGCTCTCATTTGTTCCCGAGACGCCTCGCAGTTCAGACGTGACTGTTGATCCCGCTCCTCTGCGA
CCGCTAGCTTCGATCAACTACGCGGACAGGTACCAAAACAAATGTTCTCGTCACGTGGGCATGAATCTGATGCT
GTTTCCCTGCAGACAATGCGAGAGACTGAATCAGAATTCAAATATCTGCTTCACTCACGGTGTCAAAGACTGTTT
AGAGTGCTTTCCCGTGTCAGAATCTCAACCCGTTTCTGTCGTCAAAAAGGCGTATCAGAAACTGTGCTACATTCA
TCACATCATGGGAAAGGTGCCAGACGCTTGCACTGCTTGCGACCTGGTCAATGTGGACTTGGATGACTGTGTTT
CTGAACAATAAATGACTTAAACCAGGTATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGT
GAAGGAATTCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCAACAACATCAAGACAA
CGCTCGAGGTCTTGTGCTTCCGGGTTACAAATACCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGGTC
AACGCAGCAGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAAGGCCGGAGACAACCCG
TACCTCAAGTACAACCACGCCGACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCAACCT
CGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTCTTGGTCTGGTTGAGGAAGCGGCTAAGACG
GCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGCAAATCGG
GTGCACAGCCCGCTAAAAAGAGACTCAATTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCCTCAACCA
ATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAGGTGGTGGCGCACCAGTGG
CAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTCCCAATGGCTGGG
GGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTGCCCACCTACAACAATCACCTCTACAAGCAAATCT
CCAACAGCACATCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCCTGGGGGTATTTTGAC
TTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACAACAACTGGGGATTCCGGCC
TAAGCGACTCAACTTCAAGCTCTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCG
CCAATAACCTTACCAGCACGGTCCAGGTCTTCACGGACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCT
CACGAGGGCTGCCTCCCGCCGTTCCCAGCGGACGTTTTCATGATTCCTCAGTACGGGTATCTGACGCTTAATGA
TGGAAGCCAGGCCGTGGGTCGTTCGTCCTTTTACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTA
ACAACTTCCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGAC
CGACTAATGAATCCACTCATCGACCAATACTTGTACTATCTCTCTAGAACTATTACCGGTTGGGTTCAAAACCAAG
GAATACTTCCGGGTATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCAC
ACGGACGGCAACTTTCACCCTTCTCCGCTGATGGGAGGGTTTGGAATGAAGCACCCGCCTCCTCAGATCCTCAT
CAAAAACACACCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCAGTA
TTCTACTGGCCAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAG
ATCCAGTACACTTCCAACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAACC
CCGCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAAGTCGACTACCGTTCGTATAGCATACATTATACGAA
GTTATCATATGTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAA
AAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACA
ACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTA
CAAATGTGGTAAAATCGAGCTCTACCGTTCGTATAATGTATGCTATACGAAGTTATGATATCAAGCTTAGGAACCC
CTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCC
GACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAAGCTAGCG
GGCGATTAAGGAAAGGGCTAGATCATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATG
TCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATT
TTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGA
AGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCA
CCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATC
TCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGC
TATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAAT
GACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCT

GCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGC
TTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAA
CGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTA
CTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCC
CTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACT
GGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGA
AATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATA
CTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAA
ATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCT
TTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAA
GAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAG
CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTG
GCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGC
GGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAG
GGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTT
CGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCA
ACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTAATAAACACACACACACCAACAAC
CGTGGTTGGTTGTTGTGTTGGTTTATTCTCGAG

SEQ ID NO: 7

FIG. 18 Cont.

Secuencia AA de AAV-PHP.B (AAV-G2B26) proteína VP1 cápside

SEQ ID NO: 8

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLDKGEPVNA
ADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQAKKRLLEPLGLVEEAA
KTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMAS
GGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDN
AYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTST
VQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNN
FQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTINGSGQNQQTLKFSVAGPSNMAVQGRNYIP
GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQG
TGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQ**TLAVPFK**AQAQTGWVQNQGILPGMVWQ
DRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQ
VSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL

SEQ ID NO: 8

FIG. 19

SEQ ID NO: 9

ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTGAAGGAATTCGCGAGTGGT
GGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCAACAACATCAAGACAACGCTCGAGGTC
TTGTGCTTCCGGGTTACAAATACCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGGTCAACG
CAGCAGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAAGGCCGGAGACAAC
CCGTACCTCAAGTACAACCACGCCGACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTT
GGGGGCAACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTCTTGGTCTGGTT
GAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACCGGA
CTCCTCCGCGGGTATTGGCAAATCGGGTGCACAGCCCGCTAAAAAGAGACTCAATTTCGGTCAGAC
TGGCGACACAGAGTCAGTCCCAGACCCTCAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGT
GGGATCTCTTACAATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGTGCCGATGG
AGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCACCAC
CAGCACCCGAACCTGGGCCCTGCCCACCTACAACAATCACCTCTACAAGCAAATCTCCAACAGCACA
TCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCCTGGGGGTATTTTGACTTCA
ACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACAACAACTGGGGATTCCG
GCCTAAGCGACTCAACTTCAAGCTCTTCAACATTCAGGTCAAAGAGGTTACGGACAACAATGGAGTC
AAGACCATCGCCAATAACCTTACCAGCACGGTCCAGGTCTTCACGGACTCAGACTATCAGCTCCCGT
ACGTGCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCCAGCGGACGTTTTCATGATTCCTC
AGTACGGGTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTCGTCCTTTTACTGCCTGGA
ATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAACTTCCAGTTCAGCTACGAGTTTGAGAACGTA
CCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATCGACCAAT
ACTTGTACTATCTCTC*TAGA*ACTATTAACGGTTCTGGACAGAATCAACAAACGCTAAAATTCAGTGTG
GCCGGACCCAGCAACATGGCTGTCCAGGGAAGAAACTACATACCTGGACCCAGCTACCGACAACAA
CGTGTCTCAACCACTGTGACTCAAAACAACAACAGCGAATTTGCTTGGCCTGGAGCTTCTTCTTGGG
CTCTCAATGGACGTAATAGCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGGA
CCGTTTCTTTCCTTTGTCTGGATCTTTAATTTTTGGCAAACAAGG*T*AC*C*GG*C*AGAGACAACGTGGATG
CGGACAAAGTCATGATAACCAACGAAGAAGAAATTAAAACTACTAACCCGGTAGCAACGGAGTCCTA
TGGACAAGTGGCCACAAACCACCAGAGTGCCCAA**ACTTTGGCGGTGCCTTTTAAG**GCACAGGCGCA
GACCGG*T*TGGGTTCAAAACCAAGGAATACTTCCGGGTATGGTTTGGCAGGACAGAGATGTGTACCT
GCAAGGACCCATTTGGGCCAAAATTCCTCACACGGACGGCAACTTTCACCCTTCTCCGCTGATGGGA
GGGTTTGGAATGAAGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCTGTACCTGCGGATCCTC
CAACGGCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGT
GGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTT
CCAACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGTGAACCCCGC
CCCATTGGCACCAGATACCTGACTCGTAATCTGTAA

SEQ ID NO: 9

FIG. 20

FIG. 21

SEQ ID NO: 10

gtcgacggtatcgggggagctcgcagggtctccattttgaagcgggaggtttgaacgcgcagccgccatgccggggttttacgagattgtgattaag
gtccccagcgaccttgacgagcatctgcccggcatttctgacagctttgtgaactgggtggccgagaaggaatgggagttgccgccagattctgaca
tggatctgaatctgattgagcaggcacccctgaccgtggccgagaagctgcagcgcgactttctgacggaatggcgccgtgtgagtaaggccccg
gaggctcttttctttgtgcaatttgagaagggagagagctacttccacatgcacgtgctcgtggaaaccaccggggtgaaatccatggttttgggacgtt
tcctgagtcagattcgcgaaaaactgattcagagaatttaccgcgggatcgagccgactttgccaaactggttcgcggtcacaaagaccagaaatg
gcgccggaggcgggaacaaggtggtggatgagtgctacatccccaattacttgctccccaaaacccagcctgagctccagtgggcgtggactaat
atggaacagtatttaagcgcctgtttgaatctcacggagcgtaaacggttggtggcgcagcatctgacgcacgtgtcgcagacgcaggagcagaa
caaagagaatcagaatcccaattctgatgcgccggtgatcagatcaaaaacttcagccaggtacatggagctggtcgggtggctcgtggacaagg
ggattacctcggagaagcagtggatccaggaggaccaggcctcatacatctccttcaatgcggcctccaactcgcggtcccaaatcaaggctgcct
tggacaatgcgggaaagattatgagcctgactaaaaccgcccccgactacctggtgggccagcagcccgtggaggacatttccagcaatcggatt
tataaaattttggaactaaacgggtacgatcccaatatgcggcttccgtctttctgggatgggccacgaaaaagttcggcaagaggaacaccatct
ggctgtttgggcctgcaactaccgggaagaccaacatcgcggaggccatagcccacactgtgcccttctacgggtgcgtaaactggaccaatgag
aactttcccttcaacgactgtgtcgacaagatggtgatctggtgggaggaggggaagatgaccgccaaggtcgtggagtcggccaaagccattctc
ggaggaagcaaggtgcgcgtggaccagaaatgcaagtcctcggcccagatagacccgactcccgtgatcgtcacctccaacaccaacatgtgc
gccgtgattgacgggaactcaacgaccttcgaacaccagcagccgttgcaagaccggatgttcaaatttgaactcacccgccgtctggatcatgact
ttgggaaggtcaccaagcaggaagtcaaagacttttccggtgggcaaaggatcacgtggttgaggtggagcatgaattctacgtcaaaaagggtg
gagccaagaaaagacccgcccccagtgacgcagatataagtgagcccaaacgggtgcgcgagtcagttgcgcagccatcgacgtcagacgc
ggaagcttcgatcaactacgcggacaggtaccaaaacaaatgttctcgtcacgtgggcatgaatctgatgctgtttccctgcagacaatgcgagag
actgaatcagaattcaaatatctgcttcactcacggtgtcaaagactgtttagagtgctttcccgtgtcagaatctcaacccgtttctgtcgtcaaaaagg
cgtatcagaaactgtgctacattcatcacatcatgggaaaggtgccagacgcttgcactgcttgcgacctggtcaatgtggacttggatgactgtgtttc
tgaacaataaatgacttaaaccaggtatgagtcggctggataaatctaaagtcataaacggcgctctggaattactcaatgaagtcggtatcgaagg
cctgacgacaaggaaactcgctcaaaagctgggagttgagcagcctaccctgtactggcacgtgaagaacaagcgggccctgctcgatgccctg
gccatcgagatgctggacaggcatcatacccacttctgcccccctggaaggcgagtcatggcaagactttctgcggaacaacgccaagtcattccgc
tgtgctctcctctcacatcgcgacggggctaaagtgcatctcggcacccgcccaacagagaaacagtacgaaaccctggaaaatcagctcgcgtt
cctgtgtcagcaaggcttctccctggagaacgcactgtacgctctgtccgccgtgggccactttacactgggctgcgtattggaggaacaggagcatc
aagtagcaaaagaggaaagagagacacctaccaccgattctatgcccccacttctgagacaagcaattgagctgttcgaccggcagggagccg
aacctgccttcctttttcggcctggaactaatcatatgtggcctggagaaacagctaaagtgcgaaagcggcgggccggccgacgcccttgacgattt
tgacttagacatgctcccagccgatgcccttgacgactttgaccttgatatgctgcctgctgacgctcttgacgattttgaccttgacatgctccccgggta
aatgcatgaattcgatctagagggccctattctatagtgtcacctaaatgctagagctcgctgatcagcctcgactgtgccttctagttgccagccatctg
ttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtccttcctaataaaatgaggaaattgcatcgcattgtctgagtaggt
gtcattctattctgggggggtggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctat
ggcttctgaggcggaaagaaccagctggggctcgaatcaagctatcaagtgccacctgacgtctccctatcagtgatagagaagtcgacacgtctc
gagctccctatcagtgatagagaaggtacgtctagaacgtctccctatcagtgatagagaagtcgacacgtctcgagctccctatcagtgatagaga
aggtacgtctagaacgtctccctatcagtgatagagaagtcgacacgtctcgagctccctatcagtgatagagaaggtacgtctagaacgtctcccta
tcagtgatagagaagtcgacacgtctcgagctccctatcagtgatagagaaggtacccccctatataagcagagagatctgttcaaatttgaactgact
aagcggctcccgccagattttggcaagattactaagcaggaagtcaaggacttttttgcttgggcaaaggtcaatcaggtgccggtgactcacgagtt
taaagttcccagggaattggcgggaactaaaggggcggagaaatctctaaaacgcccactgggtgacgtcaccaatactagctataaaagtctgg
agaagcgggccaggctctcatttgttcccgagacgcctcgcagttcagacgtgactgttgatcccgctcctctgcgaccgctagcttcgatcaactac
gcggacaggtaccaaaacaaatgttctcgtcacgtgggcatgaatctgatgctgtttccctgcagacaatgcgagagactgaatcagaattcaaata
tctgcttcactcacggtgtcaaagactgtttagagtgctttcccgtgtcagaatctcaacccgtttctgtcgtcaaaaaggcgtatcagaaactgtgctac
attcatcacatcatgggaaaggtgccagacgcttgcactgcttgcgacctggtcaatgtggacttggatgactgtgtttctgaacaataaatgacttaaa
ccaggt**atg**gctgccgatggttatcttccagattggctcgaggacaaccttagtgaaggaattcgcgagtggtgggctttgaaacctggagcccctca
acccaaggcaaatcaacaacatcaagacaacgctcgaggtcttgtgcttccgggttacaaataccttggacccggcaacggactcgacaagggg
gagccggtcaacgcagcagacgcggcggccctcgagcacgacaaggcctacgaccagcagctcaaggccggagacaacccgtacctcaag
tacaaccacgccgacgccgagttccaggagcggctcaaagaagatacgtcttttgggggcaacctcgggcgagcagtcttccaggccaaaaag
aggcttcttgaacctcttggtctggttgaggaagcggctaagacggctcctggaaagaagaggcctgtagagcagtctcctcaggaaccggactcct
ccgcgggtattggcaaatcgggtgcacagcccgctaaaaagagactcaatttcggtcagactggcgacacagagtcagtcccagaccctcaacc
aatcggagaacctcccgcagccccctcaggtgtgggatctcttacaatggcttcaggtggtggcgcaccagtggcagacaataacgaaggtgccg
atggagtgggtagttcctcgggaaattggcattgcgattcccaatggctgggggacagagtcatcaccaccagcacccgaacctgggccctgccc
acctacaacaatcacctctacaagcaaatctccaacagcacatctggaggatcttcaaatgacaacgcctacttcggctacagcacccccctgggg
gtattttgacttcaacagattccactgccacttctcaccacgtgactggcagcgactcatcaacaacaactggggattccggcctaagcgactcaactt
caagctcttcaacattcaggtcaaagaggttacggacaacaatggagtcaagaccatcgccaataaccttaccagcacggtccaggtcttcacgg
actcagactatcagctcccgtacgtgctcgggtcggctcacgagggctgcctcccgccgttcccagcggacgttttcatgattcctcagtacgggtatct
gacgcttaatgatggaagccaggccgtgggtcgttcgtcctttttactgcctggaatatttcccgtcgcaaatgctaagaacgggtaacaacttccagttc
agctacgagtttgagaacgtacctttccatagcagctacgctcacagccaaagcctggaccgactaatgaatccactcatcgaccaatacttgtacta
tctctcTaGAactattaacggttctggacagaatcaacaaacgctaaaattcagtgtggccggacccagcaacatggctgtccagggaagaaact
acatacctggacccagctaccgacaacaacgtgtctcaaccactgtgactcaaaacaacaacagcgaatttgcttggcctggagcttcttcttgggct
ctcaatggacgtaatagcttgatgaatcctggacctgctatggccagccacaaagaaggagaggaccgtttcttccttgtctggatctttaatttttggc

aaacaaggTacCggCagagacaacgtggatgcggacaaagtcatgataaccaacgaagaagaaattaaaactactaacccggtagcaacg
gagtcctatggacaagtggccacaaaccaccagagtgcccaaACTTTGGCGGTGCCTTTTAAGgcacaggcgcagaccggTtgg
gttcaaaaccaaggaatacttccgggtatggtttggcaggacagagatgtgtacctgcaaggacccatttgggccaaaattcctcacacggacggc
aactttcacccttctccgctgatgggagggtttggaatgaagcacccgcctcctcagatcctcatcaaaaacacacctgtacctgcggatcctccaac
ggccttcaacaaggacaagctgaactctttcatcacccagtattctactggccaagtcagcgtggagatcgagtgggagctgcagaaggaaaaca
gcaagcgctggaacccggagatccagtacacttccaactattacaagtctaataatgttgaatttgctgttaatactgaaggtgtatatagtgaacccc
gccccattggcaccagatacctgactcgtaatctgtaattgcttgttaatcaataaaccgtttaattcgtttcagttgaactttggtctctgcgaagggcga
attcgtttaaacctgcaggactagaggtcctgtattagaggtcacgtgagtgttttgcgacattttgcgacaccatgtggtcacgctgggtatttaagccc
gagtgagcacgcagggtctccattttgaagcgggaggtttgaacgcgcagccgccaagccgaattctgcagatatccatcacactggcggccgct
cgactagagcggccgccaccgcggtggagctccagcttttgttcccttagtgagggttaattgcgcgcttggcgtaatcatggtcatagctgtttcctgt
gtgaaattgttatccgctcacaattccacacaacatacgagccggaagcataaagtgtaaagcctggggtgcctaatgagtgagctaactcacatta
attgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtat
tgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatc
cacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttt
tccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtt
tcccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatag
ctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggt
aactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggt
gctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaag
agttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaag
aagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctag
atccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctca
gcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgat
accgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatcc
gcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggt
gtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctcc
ttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagat
gcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgc
gccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgta
acccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaaggg
aataagggcgacacggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaat
gtatttagaaaaataaacaaatagggttccgcgcacatttccccgaaaagtgccacctaaattgtaagcgttaatattttgttaaaattcgcgttaaatt
tttgttaaatcagctcatttttttaaccaataggccgaaatcggcaaaatcccttataaatcaaaagaatagaccgagatagggttgagtgttgttccagtt
tggaacaagagtccactattaaagaacgtggactccaacgtcaaagggcgaaaaaccgtctatcagggcgatggcccactacgtgaaccatcac
cctaatcaagttttttggggtcgaggtgccgtaaagcactaaatcggaaccctaaagggagcccccgatttagagcttgacggggaaagccggcg
aacgtggcgagaaaggaagggaagaaagcgaaaggagcgggcgctagggcgctggcaagtgtagcggtcacgctgcgcgtaaccaccaca
cccgccgcgcttaatgcgccgctacagggcgcgtcccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctat
tacgccagctggcgaaaggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagt
gagcgcgcgtaatacgactcactatagggcgaattgggtaccgggccccccctcgatcgag

SEQ ID NO: 10

FIG. 21 Cont.

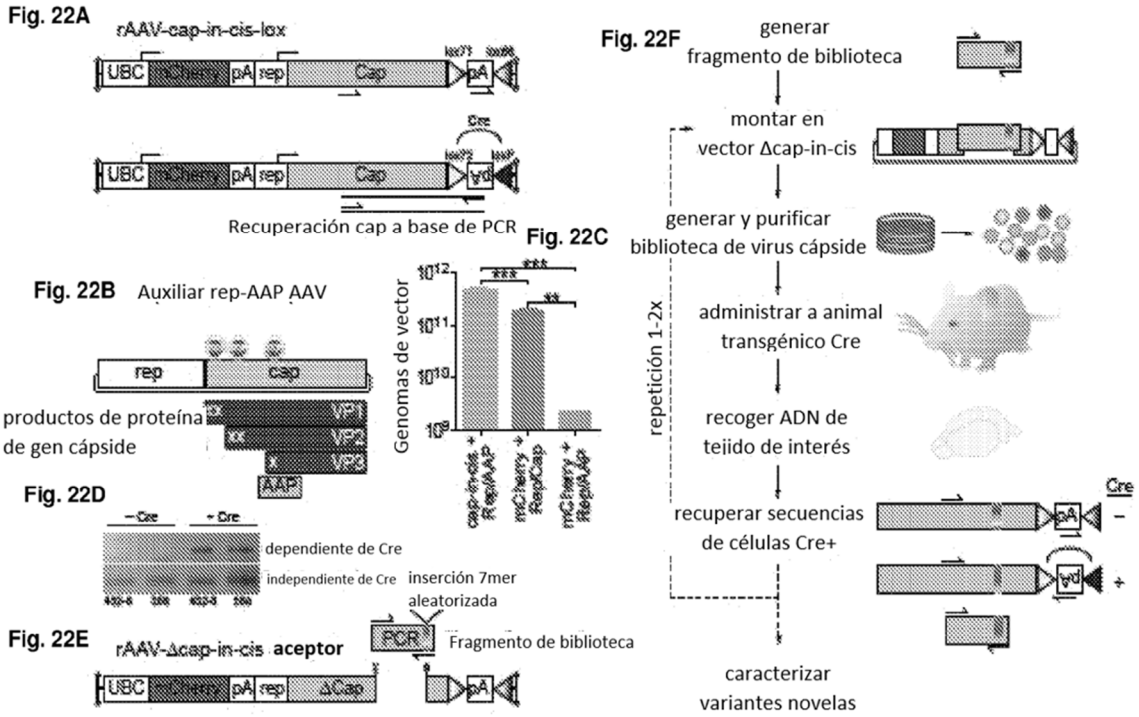


Fig. 22A
rAAV-cap-in-cis-lox
UBC
mCherry
pA
rep
Cap
Cre
Recuperación cap a base de PCR
Fig. 22B
Auxiliar rep-AAP AAV
rep
cap
productos de proteína de gen cápside
VP1
VP2
VP3
AAP
Fig. 22C
Genomas de vector
cap-in-cis + Rep/AAP
mCherry + Rep/Cap
mCherry + Rep/AAP
Fig. 22D
−Cre
+Cre
dependiente de Cre
independiente de Cre
Fig. 22E
rAAV-Δcap-in-cis aceptor
ΔCap
PCR
Fragmento de biblioteca
inserción 7mer aleatorizada
Fig. 22F
generar fragmento de biblioteca
montar en vector Δcap-in-cis
generar y purificar biblioteca de virus cápside
administrar a animal transgénico Cre
recoger ADN de tejido de interés
recuperar secuencias de células Cre+
caracterizar variantes novelas
repetición 1-2x
Cre
−
+

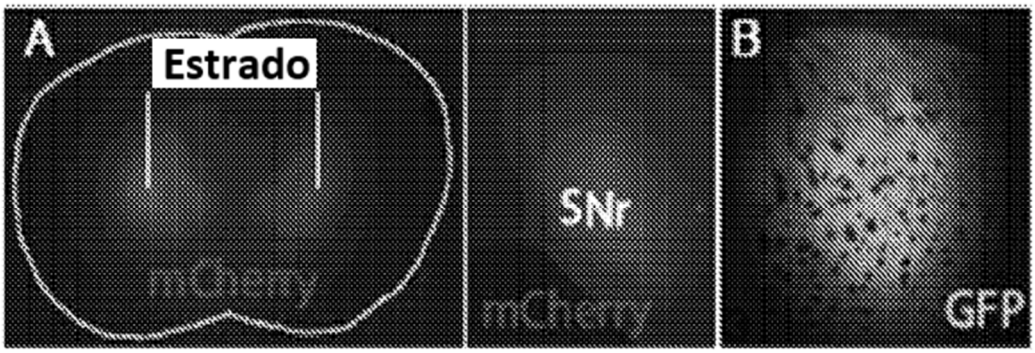


Fig. 23A

Fig. 23B

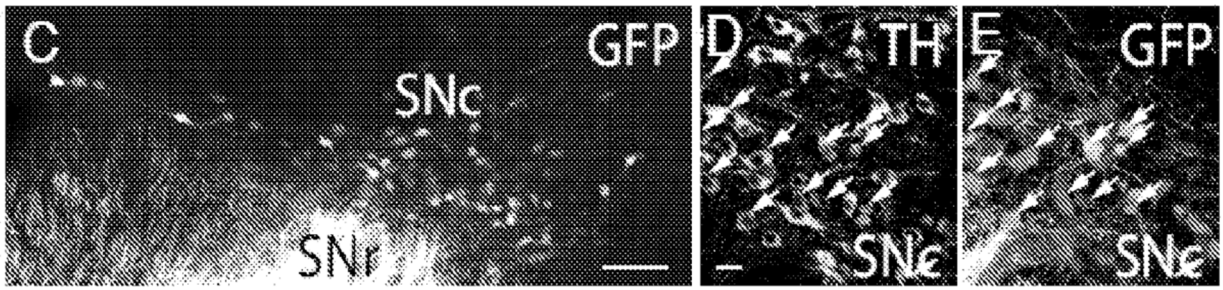


Fig. 23C Fig. 23D Fig. 23E

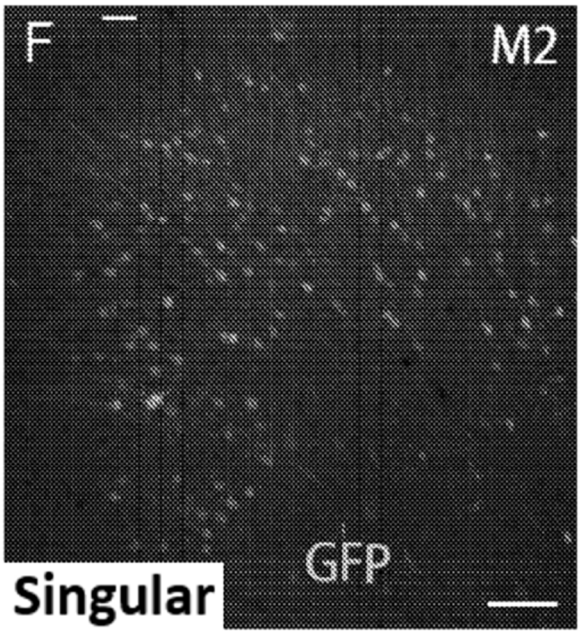


Fig. 23F

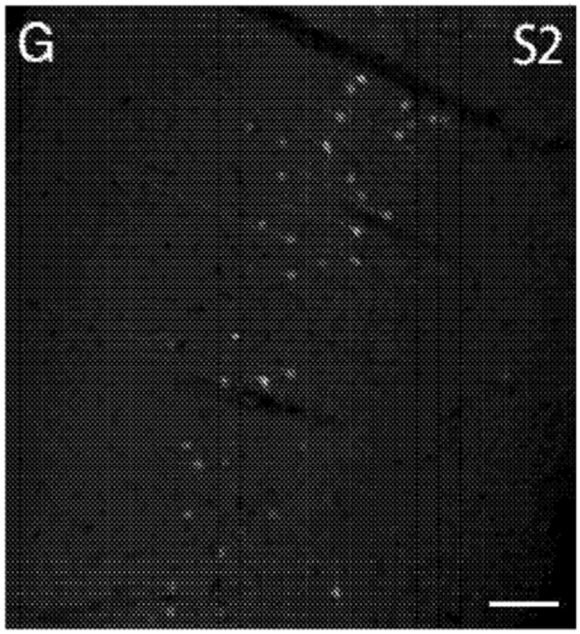


Fig. 23G

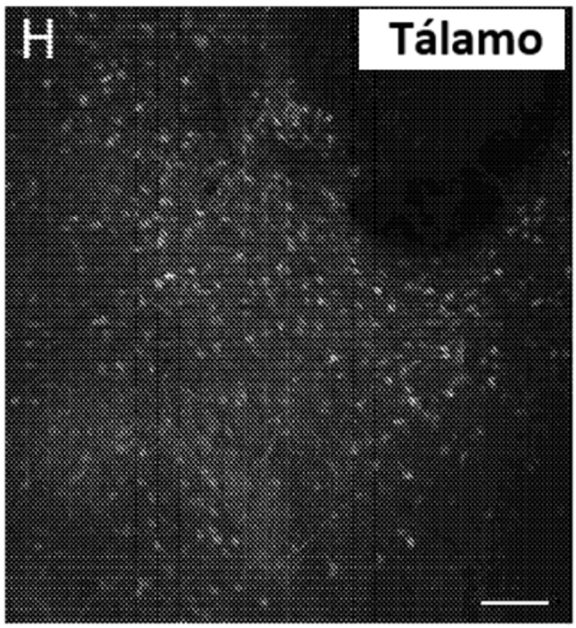


Fig. 23H

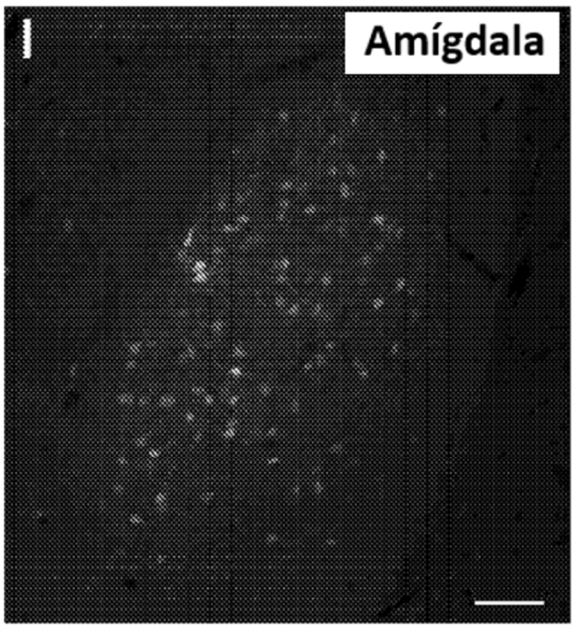


Fig. 23I

Fig. 24A

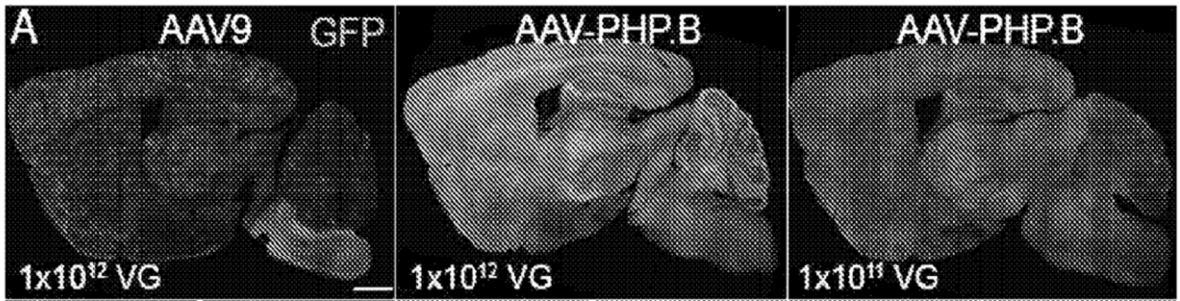
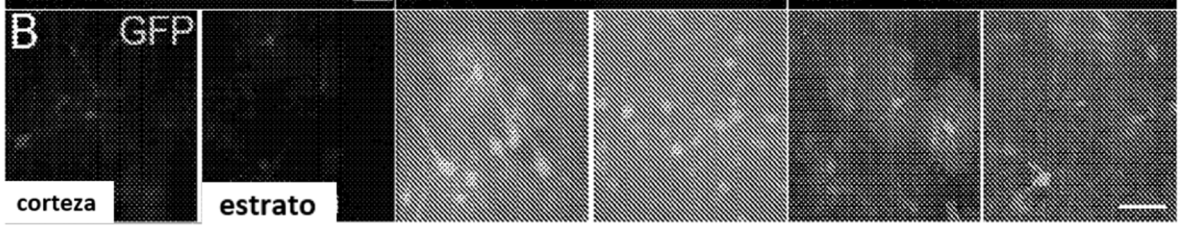


Fig. 24B



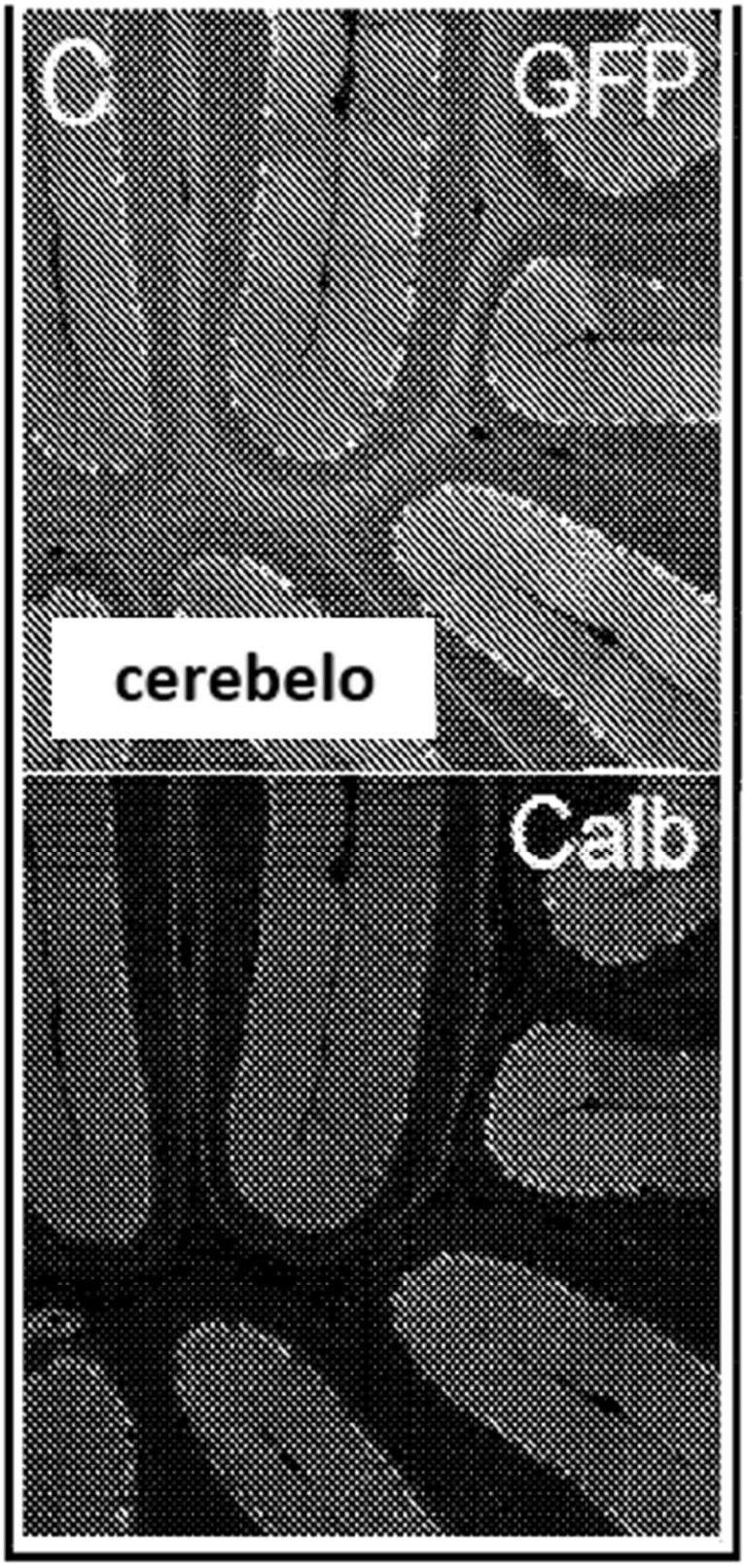


Fig. 24C

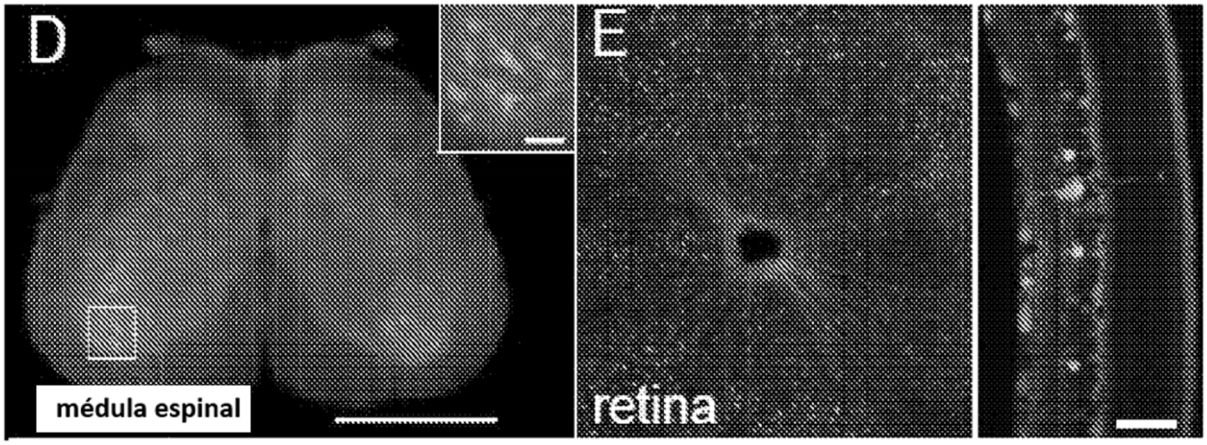


Fig. 24D Fig. 24E Fig. 24F

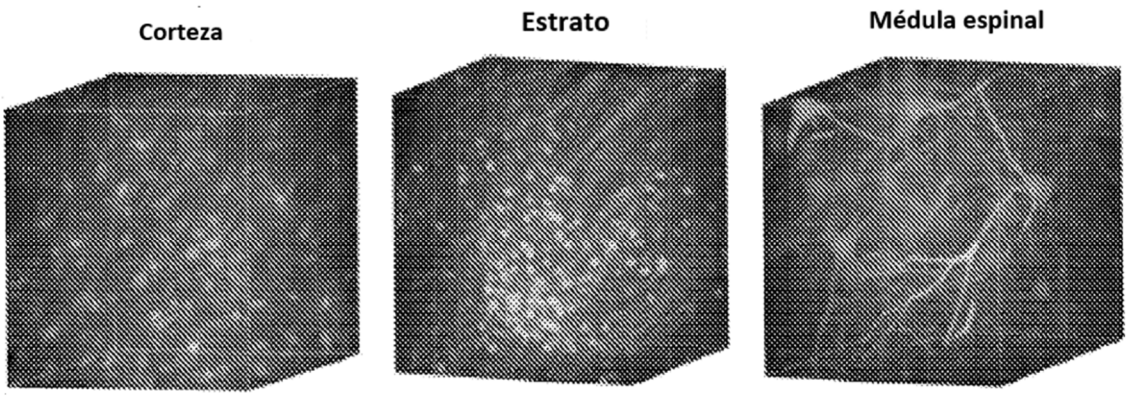


Fig. 24G

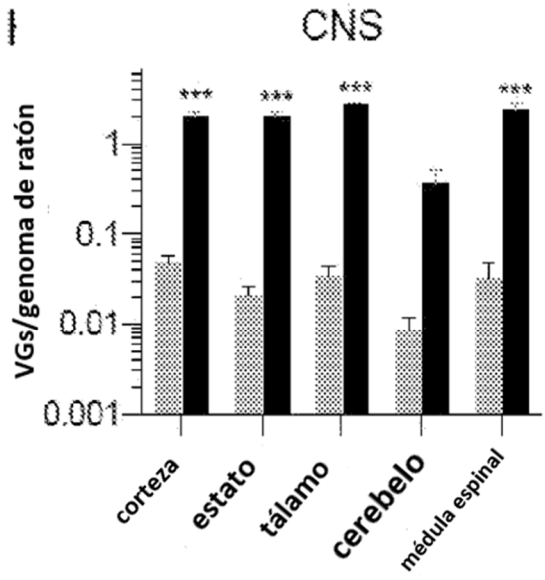


Fig. 24H

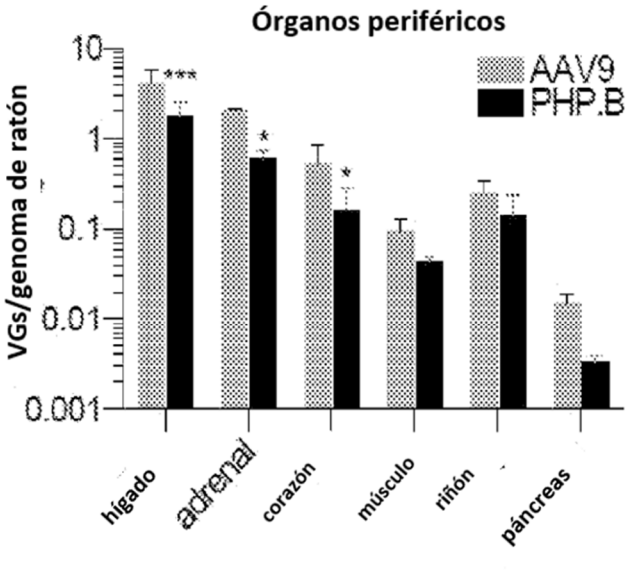


Fig. 24I

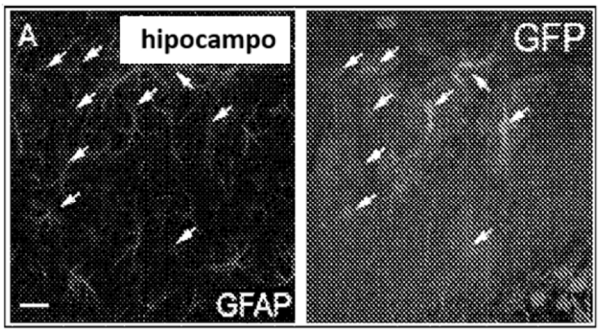


Fig. 25A

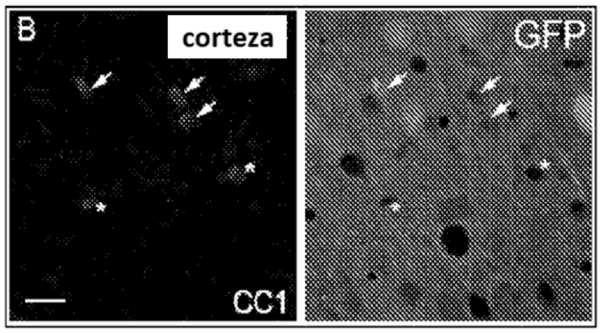


Fig. 25B

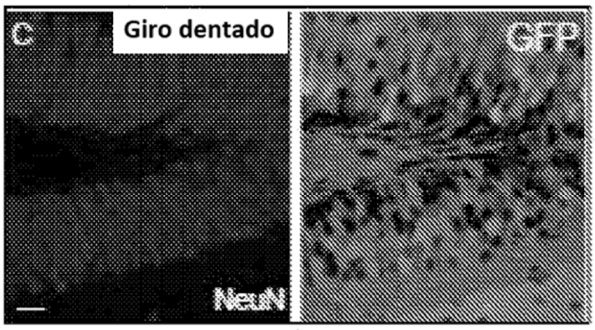


Fig. 25C

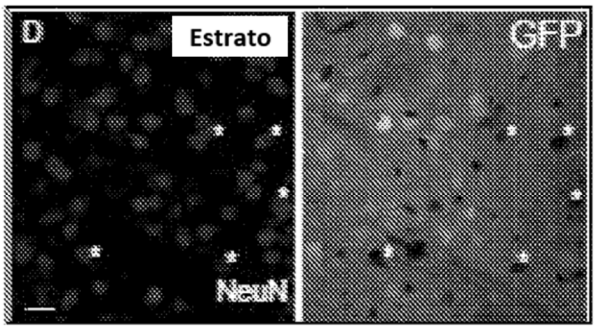


Fig. 25D

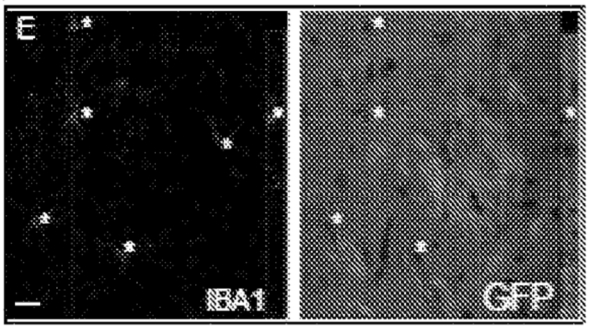


Fig. 25E

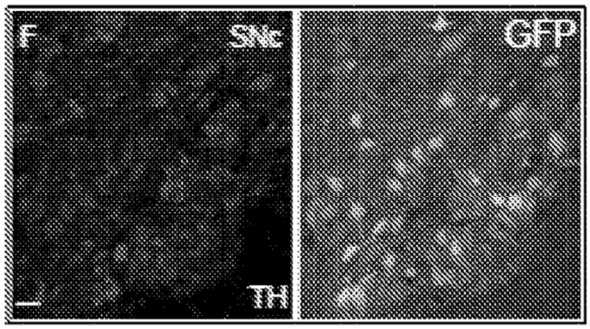
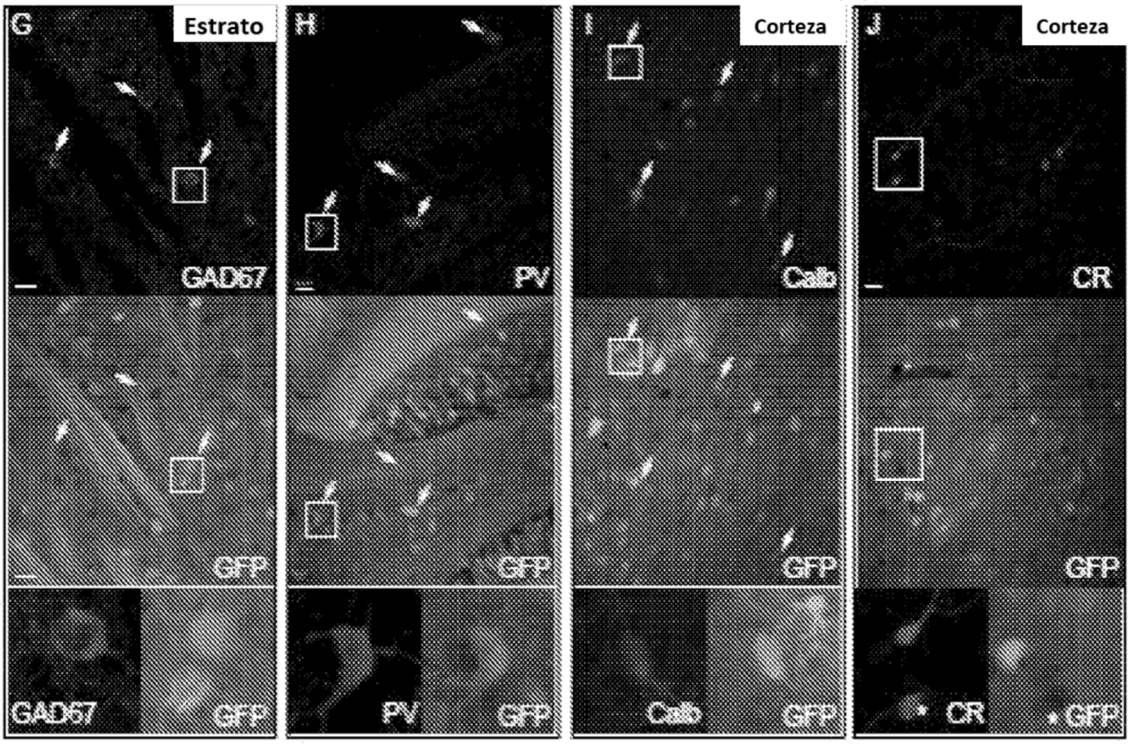


Fig. 25F

Fig. 25G Fig. 25H Fig. 25I Fig. 25J



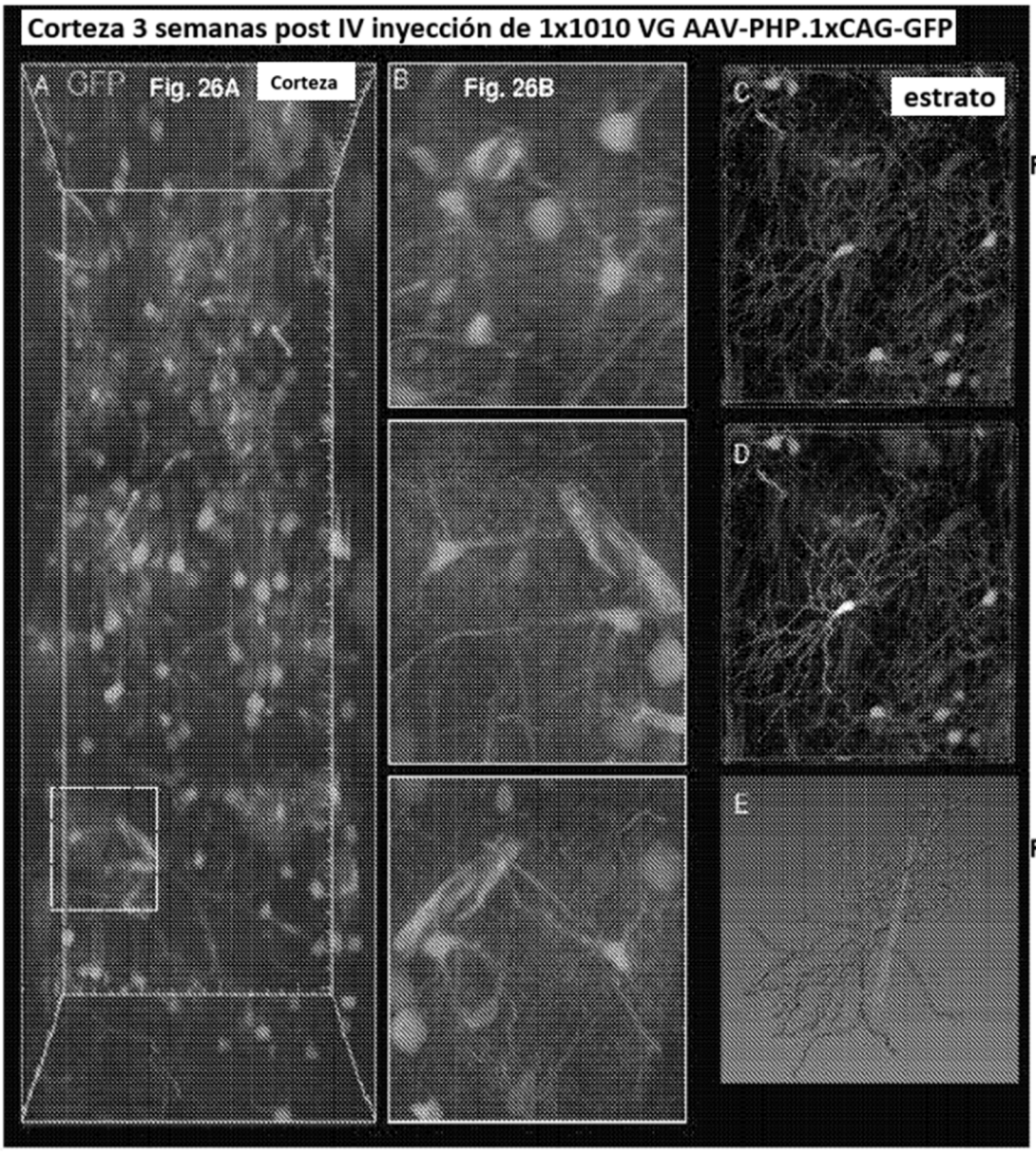

Corteza 3 semanas post IV inyección de 1x1010 VG AAV-PHP.1xCAG-GFP
A GFP
Fig. 26A
Corteza
B
Fig. 26B
C
estrato
Fig. 26C
D
Fig. 26D
E
Fig. 26E

Fig. 27A

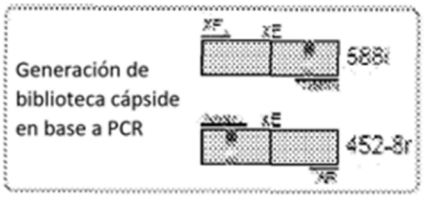
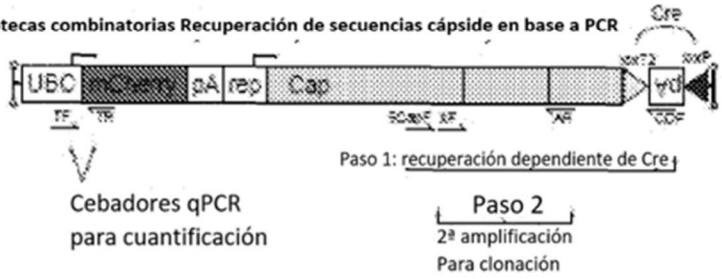


Fig. 27B



	Cebador	Propósito	Secuencia
SEQ ID NO: 16	9CapF	Paso 1: adelante	CAGGTCTTCACGGACTCAGACTATCAG
SEQ ID NO: 17	CDF	Paso 1: invertido por Cre	CAAGTAAAACCTCTACAAATGTGGTAAAATCG
SEQ ID NO: 18	XF	Paso 2: adelante	ACTCATCGACCAATACTTGTACTATCTCTCTAGAAC
SEQ ID NO: 19	AR	Paso 2: inverso	GGAAGTATTCCTTGGTTTTGAACCCA
SEQ ID NO: 20	TF	qPCR adelante	GGTCGCGGTTCTTGTTTGTGGAT
SEQ ID NO: 21	TR	qPCR inverso	GCACCTTGAAGCGCATGAACTCCT
SEQ ID NO: 22	7xNNK	452-8r generación biblioteca	CATCGACCAATACTTGTACTATCTCTCTAGAACTATTNNKNNK NNKNNKNNKNNKNNKCAAACGCTAAAATTCAGTGTGGCCGGA
SEQ ID NO: 23	7xNNK	588i generación biblioteca	GTATTCCTTGGTTTTGAACCCAACCGGTCTGCGCCTGTGCMNN MNNMNNMNNMNNMNNMNNNTTGGGCACTCTGGTGGTTTGTG

Fig. 27C

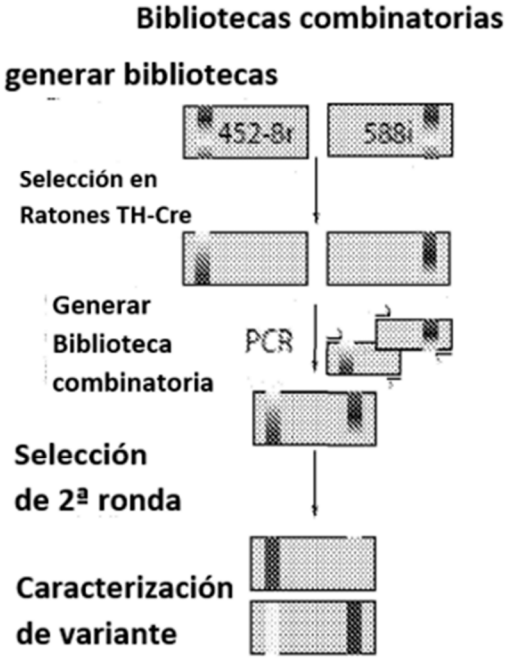


Fig. 27D

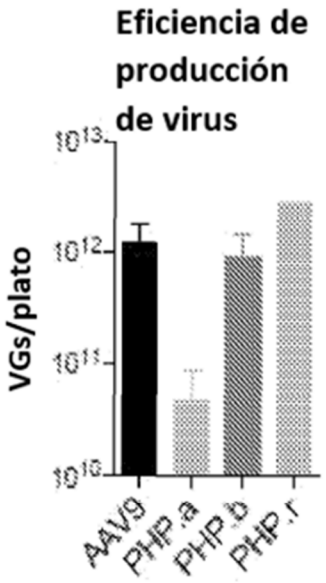


Fig. 27E

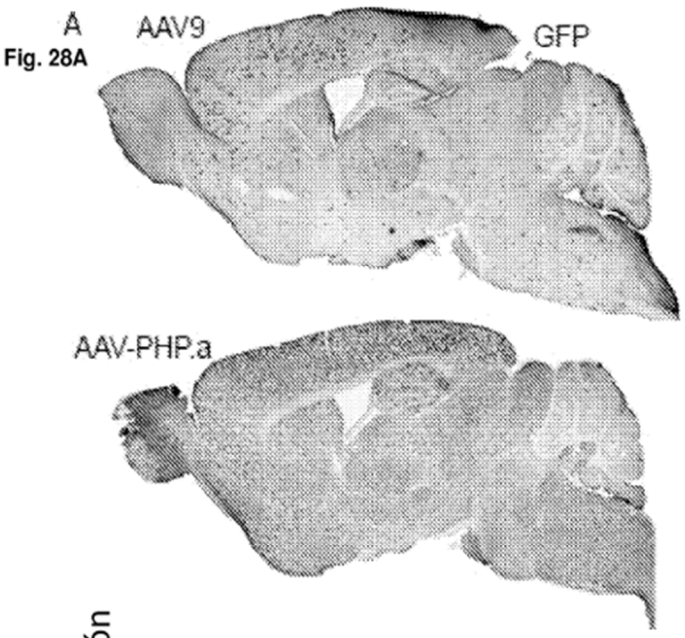


Fig. 28A

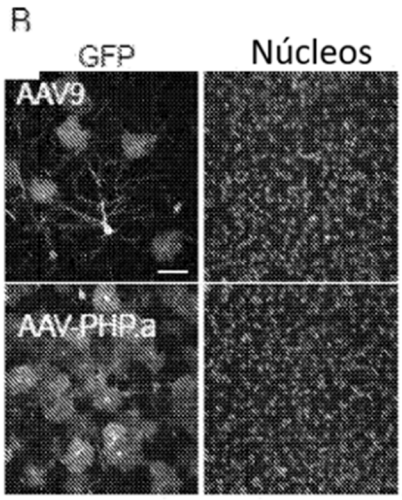


Fig. 28B

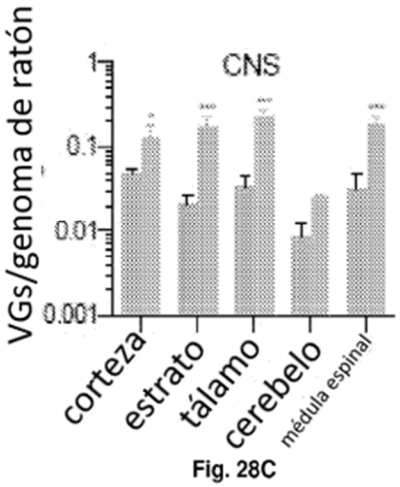


Fig. 28C

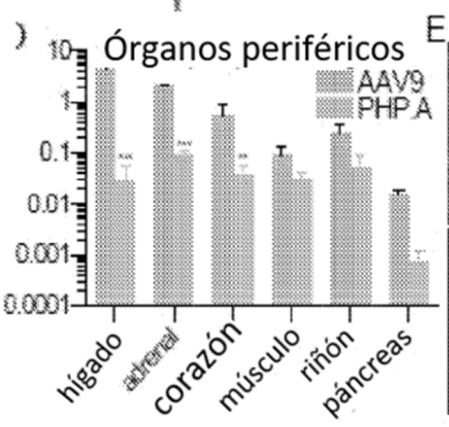


Fig. 28D

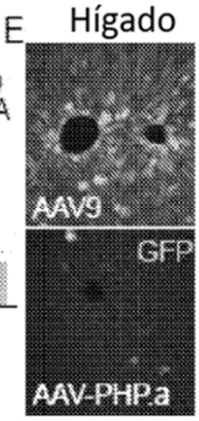


Fig. 28E

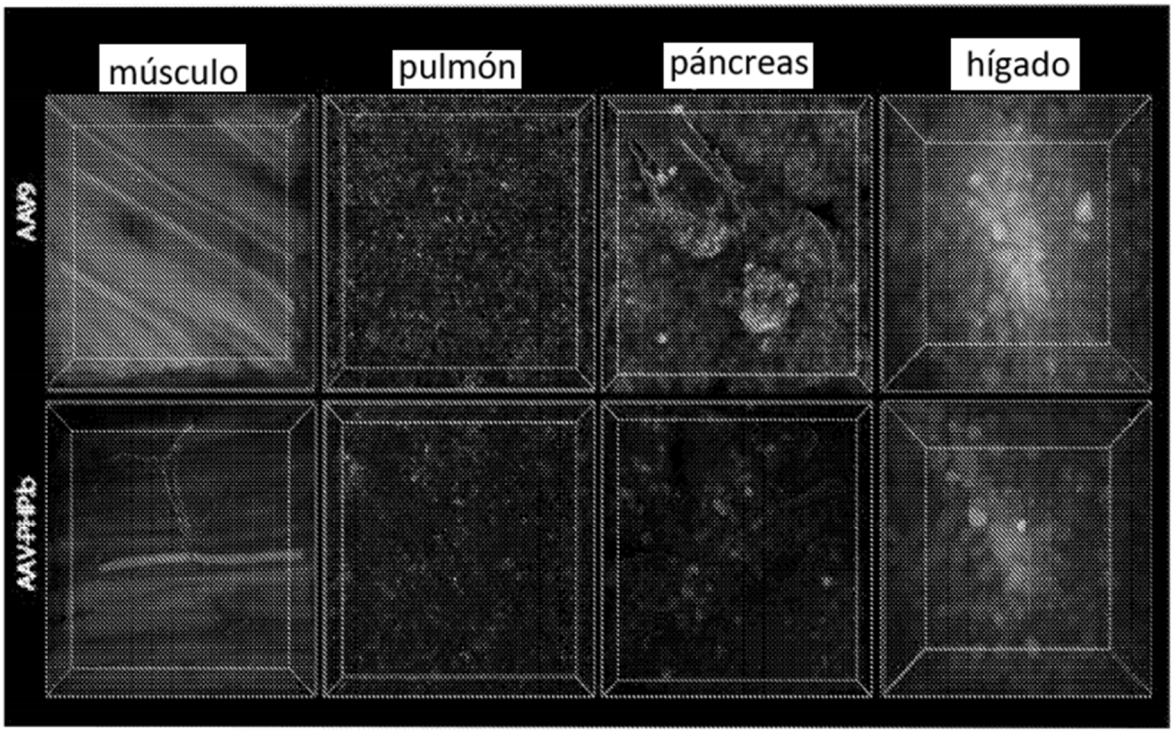


Fig. 29

Secuencia de ácido nucleico AAV9 VP1

SEQ ID NO: 2

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLDKGEPVNA
ADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQAKKRLLEPLGLVEEAA
KTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMAS
GGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDN
AYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTST
VQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNN
FQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKFSVAGPSNMAVQGRNYIP
GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQG
TGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDRDVYLQ
GPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWE
LQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL

SEQ ID NO: 2

Secuencia de aminoácido AAV9 VP1

SEQ ID NO. 11

atggctgccgatggttatcttccagattggctcgaggacaaccttagtgaaggaattcgcgagtggtgggctttgaaacctggagcccctcaacccaa
ggcaaatcaacaacatcaagacaacgctcgaggtcttgtgcttccgggttacaaataccttggacccggcaacggactcgacaaggggagccg
gtcaacgcagcagacgcggcggccctcgagcacgacaaggcctacgaccagcagctcaaggccggagacaacccgtacctcaagtacaacc
acgccgacgccgagttccaggagcggctcaaagaagatacgtcttttgggggcaacctcgggcgagcagtcttccaggccaaaaagaggcttctt
gaacctcttggtctggttgaggaagcggctaagacggctcctggaaagaagaggcctgtagagcagtctcctcaggaaccggactcctccgcggg
tattggcaaatcgggtgcacagcccgctaaaaagagactcaatttcggtcagactggcgacacagagtcagtcccagaccctcaaccaatcgga
gaacctcccgcagccccctcaggtgtgggatctcttacaatggcttcaggtggtggcgcaccagtggcagacaataacgaaggtgccgatggagt
gggtagttcctcgggaaattggcattgcgattcccaatggctgggggacagagtcatcaccaccagcacccgaacctgggccctgcccacctaca
acaatcacctctacaagcaaatctccaacagcacatctggaggatcttcaaatgacaacgcctacttcggctacagcacccccctgggggtattttga
cttcaacagattccactgccacttctcaccacgtgactggcagcgactcatcaacaacaactggggattccggcctaagcgactcaacttcaagctct
tcaacattcaggtcaaagaggttacggacaacaatggagtcaagaccatcgccaataaccttaccagcacggtccaggtcttcacggactcagac
tatcagctcccgtacgtgctcgggtcggctcacgagggctgcctcccgccgttcccagcggacgttttcatgattcctcagtacgggtatctgacgctta
atgatggaagccaggccgtgggtcgttcgtccttttactgcctggaatatttcccgtcgcaaatgctaagaacgggtaacaacttccagttcagctacg
agtttgagaacgtacctttccatagcagctacgctcacagccaaagcctggaccgactaatgaatccactcatcgaccaatacttgtactatctctcaa
agactattaacggttctggacagaatcaacaaacgctaaaattcagtgtggccggacccagcaacatggctgtccagggaagaaactacatacct
ggacccagctaccgacaacaacgtgtctcaaccactgtgactcaaaacaacaacagcgaatttgcttggcctggagcttcttcttgggctctcaatgg
acgtaatagcttgatgaatcctggacctgctatggccagccacaaagaaggagaggaccgtttctttccttgtctggatctttaatttttggcaaacaag
gaactggaagagacaacgtggatgcggacaaagtcatgataaccaacgaagaagaaattaaaactactaacccggtagcaacggagtcctat
ggacaagtggccacaaaccaccagagtgcccaagcacaggcgcagaccggctgggttcaaaaccaaggaatacttccgggtatggtttggcag
gacagagatgtgtacctgcaaggacccatttgggccaaaattcctcacacggacggcaactttcaccttctccgctgatgggagggtttggaatga
agcacccgcctcctcagatcctcatcaaaaacacacctgtacctgcggatcctccaacggccttcaacaaggacaagctgaactctttcatcaccc
agtattctactggccaagtcagcgtggagatcgagtgggagctgcagaaggaaaacagcaagcgctggaacccggagatccagtacacttcca
actattacaagtctaataatgttgaatttgctgttaatactgaaggtgtatatagtgaaccccgccccattggcaccagatacctgactcgtaatctgtaa

SEQ ID NO: 11

FIG. 30

FIG. 31

Secuencia ID	Secuencia de ácido nucleico	Secuencia AA	Porcentaje de clones totales
SEQ ID NO: 24	ACTTTGGCGGTGCCTTTTAAG	TLAVPFK (SEQ ID NO: 1)	25% (20/80)
SEQ ID NO: 25	AGTGTGAGTAAGCCTTTTTTG	SVSKPFL (SEQ ID NO: 28)	10% (8/80)
SEQ ID NO: 26	TTTACGTTGACGACGCCTAAG	FTLTTPK (SEQ ID NO: 29)	7,5% (6/80)
SEQ ID NO: 27	ATGAATGCTACGAAGAATGTG	MNATKNV (SEQ ID NO: 30)	5% (4/80)