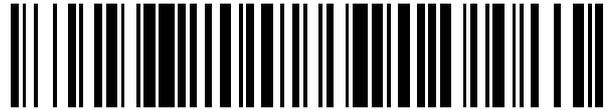


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 526 341**

51 Int. Cl.:

**C12N 7/00** (2006.01)  
**C12N 7/01** (2006.01)  
**C12N 15/85** (2006.01)  
**C12N 15/86** (2006.01)  
**C12N 15/864** (2006.01)  
**C07K 14/015** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **12.11.2002 E 02793807 (5)**

97 Fecha y número de publicación de la concesión europea: **22.10.2014 EP 1463805**

54 Título: **Secuencias de serotipo 9 de virus adeno-asociado (AAV), vectores que las contienen, y usos de las mismas**

30 Prioridad:

**17.12.2001 US 341150 P**  
**05.06.2002 US 386132 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**09.01.2015**

73 Titular/es:

**THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA (100.0%)**  
**3160 CHESTNUT STREET, SUITE 200**  
**PHILADELPHIA, PA 19104, US**

72 Inventor/es:

**GAO, GUANGPING;**  
**WILSON, JAMES M. y**  
**ALVIRA, MAURICIO**

74 Agente/Representante:

**DE PABLOS RIBA, Julio**

**ES 2 526 341 T3**

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

## DESCRIPCIÓN

Secuencias de serotipo 9 de virus adeno-asociado (AAV), vectores que las contienen, y usos de las mismas.

### Antecedentes de la invención

5 Un virus adeno-asociado (AAV), un miembro de la familia Parvovirus, es un pequeño virus icosaédrico, sin envoltura, con genomas de ADN lineales de cadena simple de 4,7 kilobases (kb) a 6 kb. El AAV se asigna al género *Dependovirus*, debido a que el virus fue descubierto como contaminante en acumulaciones de adenovirus purificados. El ciclo de vida del AAV incluye una fase latente en la que los genomas de AAV, tras la infección, son el sitio integrado específicamente en cromosomas anfitrión y una fase infecciosa en la que, a continuación de la  
10 infección ya sea por adenovirus o ya sea por virus herpes simplex, los genomas integrados son posteriormente rescatados, replicados y empaquetados en virus infecciosos. Las propiedades de no patogenicidad, amplia gama de infectividad de anfitrión, incluyendo las células sin división, y la potencial integración cromosómica específica del sitio, hacen del AAV una atractiva herramienta para transferencia de gen.

15 Estudios recientes sugieren que los vectores de AAV pueden ser el vehículo preferido para el suministro de gen. Hasta la fecha, han existido 6 serotipos diferentes de AAVs aislados a partir de humanos y de primates no humanos (NHP), y bien caracterizados. Entre ellos, el serotipo 2 humano es el primer AAV que fue desarrollado como vector de transferencia de gen; Se ha usado ampliamente para experimentos eficientes de transferencia de gen en diferentes tejidos objetivo y modelos animales. Los ensayos clínicos de la aplicación experimental de vectores a base de AAV2 a algunos modelos de enfermedad humana están en curso, e incluyen enfermedades tales como  
20 fibrosis quística y hemofilia B.

Lo que resulta deseable son construcciones a base de AAV para suministro de gen.

### Sumario de la invención

25 En un aspecto, la invención proporciona secuencias novedosas de AAV, composiciones que contienen esas secuencias, y usos de las mismas. Ventajosamente, esas composiciones son muy adecuadas para su uso en composiciones que requieren la readministración de rAAV con fines terapéuticos o profilácticos.

Estos y otros aspectos de la invención resultarán evidentes a partir de la descripción detallada que sigue de la invención.

### Breve descripción de los dibujos

Las Figuras 1A, 1B y 1C son secuencias de ácido nucleico de las regiones rep y cap de AAV9 [SEQ ID Núm. 1].

30 Las Figuras 2A, 2B y 2C son un alineamiento de las secuencias de aminoácido de la proteína vp1 de la cápside de las nuevas secuencias de AAV9 [SEQ ID Núm. 2] de la invención en comparación con las secuencias publicadas de AAV2 [SEQ ID Núm. 4], AAV1 [SEQ ID Núm. 5], y AAV3 [SEQ ID Núm. 6] y de un nuevo serotipo AAV8 [SEQ ID Núm. 7], que es el objeto de una solicitud en trámite. El alineamiento fue realizado usando el programa Clustal W, con la numeración de AAV usada como referencia. El subrayado y la negrita bajo las secuencias de AAV9 indican  
35 casetes de identidad dentro de la HVR.

Las Figuras 3A, 3B y 3C son las secuencias de aminoácido de las proteínas rep de AAV9 [SEQ ID Núm. 3].

### Descripción detallada de la invención

40 La invención proporciona las secuencias de ácido nucleico y aminoácidos de un nuevo serotipo de AAV, el AAV9. También se proporcionan fragmentos de esas secuencias de AAV. Cada uno de esos fragmentos puede ser utilizado fácilmente en una diversidad de sistemas de vector y de células anfitrión. Entre los fragmentos de AAV9 deseables están las proteínas cap, incluyendo la vp1, vp2 y vp3. Otros fragmentos de interés incluyen las regiones hipervariables, las proteínas rep, incluyendo la rep78, rep68, rep52 y rep40, y las secuencias que codifican esas proteínas. Esos fragmentos pueden ser utilizados fácilmente en una diversidad de sistemas de vector y de células anfitrión. Tales fragmentos pueden ser usados solos, en combinación con otras secuencias de AAV9 o fragmentos, o  
45 en combinación elementos procedentes de AAV viral o de AAV no viral. En una realización particularmente deseable, un vector contiene las secuencias cap y rep de AAV9 de la invención.

Las secuencias de AAV y los fragmentos de las mismas, son útiles en la producción de rAAV, y son también útiles como vectores de suministro antisentido, vectores de terapia de gen, o vectores de vacuna. La invención proporciona además moléculas de ácido nucleico, vectores de suministro de gen, y células anfitrión que contienen  
50 las secuencias de AAV9 de la invención.

Los fragmentos adecuados pueden ser determinados usando la información proporcionada en la presente memoria.

Las alineaciones se realizan usando cualquiera de una diversidad de Programas de Alineamiento de Secuencia Múltiple disponibles pública o comercialmente, tal como el "Clustal W", accesible a través de Servidores Web por Internet. Alternativamente, se usan también utilidades de vector NTI. Existe también un número de algoritmos conocidos en el estado de la técnica que pueden ser usados para medir identidad de secuencia de nucleótido, incluyendo los contenidos en los programas descritos con anterioridad. Según otro ejemplo, las secuencias de polinucleótido pueden ser comparadas usando Fasta, un programa en GCG Versión 6.1. Fasta proporciona alineamientos y porcentaje de identidad de secuencia de las regiones de mejor solapamiento entre las secuencias de consulta y búsqueda. Por ejemplo, el porcentaje de identidad de secuencia entre secuencias de ácido nucleico puede ser determinado usando Fasta con sus parámetros por defecto (un tamaño de palabra de 6 y el factor NOPAM para la matriz de puntuación) según se proporciona en GCG Versión 6.1, incorporado en la presente memoria por referencia. Se encuentran disponibles programas similares para secuencias de aminoácido, por ejemplo el programa 2Clustal X". En general, cualquiera de esos programas se utiliza en configuraciones por defecto, aunque un experto en la materia alterar esas configuraciones según se necesite. Alternativamente, un experto en la materia puede utilizar otro algoritmo o programa de ordenador que proporcione al menos un nivel de identidad o alineamiento como el proporcionado por los algoritmos y programas referenciados.

El término "homología sustancial" o "similitud sustancial", cuando se refieren a ácido nucleico, o a un fragmento del mismo, indica que cuando se alinea óptimamente con inserciones o supresiones de nucleótido apropiadas con otro ácido nucleico (o su cadena complementaria), existe identidad de secuencia de nucleótido en al menos alrededor de un 95 a un 99% de las secuencias alineadas. Con preferencia, la homología es sobre la secuencia de longitud completa, o sobre una trama de lectura abierta de la misma.

El término "homología sustancial" o "similitud sustancial", cuando se refiere a aminoácidos o a un fragmento de los mismos, indica que, cuando se alinean óptimamente con inserciones o supresiones de aminoácido apropiadas con otro aminoácido (o su cadena complementaria), existe identidad de secuencia de aminoácido en al menos alrededor de un 95 a un 99% de las secuencias alineadas. Con preferencia, la homología es sobre la secuencia de longitud completa, o sobre una proteína de la misma, por ejemplo, una proteína cap, o una proteína rep.

Mediante el término "altamente conservado" se entiende al menos una identidad del 80%, con preferencia al menos una identidad del 90%, y más preferentemente, una identidad por encima del 97%. La identidad la determina fácilmente un experto en la materia recurriendo a algoritmos y programas de ordenador conocidos por los expertos en la materia.

El término "porcentaje de identidad de secuencia" o "idéntico" en el contexto de secuencias de ácido nucleico, se refiere a los residuos en las dos secuencias que son iguales cuando se alinean respecto a una correspondencia máxima. La longitud de comparación de identidad de secuencia puede ser sobre la longitud completa del genoma, la longitud completa de una secuencia de codificación de gen, o un fragmento de al menos alrededor de 500 a 5000 nucleótidos, si se desea. De forma similar, se puede determinar fácilmente el "porcentaje de identidad de secuencia" para secuencias de aminoácido, sobre la longitud completa de una proteína.

Según se describe en la presente memoria, los vectores de la invención que contienen las proteínas de cápside de AAV de la invención son particularmente adecuados para su uso en aplicaciones en las que los anticuerpos neutralizantes disminuyen la efectividad de otros vectores basados en serotipo de AAV, así como otros vectores virales. Los vectores de AAV de la invención son particularmente ventajosos en la readministración de rAAV y en terapia de gen de repetición.

Estas y otras realizaciones y ventajas de la invención se describen con mayor detalle en lo que sigue. Según se utiliza a través de la presente descripción y de las reivindicaciones, el término "que comprende" es inclusivo de otros componentes, elementos, integradores, etapas y similares. Por el contrario, el término "que consiste" y sus variantes, son excluyentes de otros componentes, elementos, integradores, etapas y similares.

## 45 I. Secuencias de serotipo 9 de AAV

### A. Secuencias de ácido nucleico

Las secuencias de ácido nucleico de AAV9 de la invención incluyen las secuencias de ADN de la Figura 1 [SEQ ID Núm. 1], consistentes en 4382 nucleótidos. Las secuencias de ácido nucleico de AAV9 de la invención abarcan además la cadena que es complementaria con la Figura 1 [SEQ ID Núm. 1], así como las secuencias de ARN y cADN correspondientes a la Figura 1 [SEQ ID Núm. 1] y su cadena complementaria. También incluidas en las secuencias de ácido nucleico de la invención están las variantes naturales y las modificaciones de diseño de la Figura 1 [SEQ ID Núm. 1] y su cadena complementaria. Tales modificaciones incluyen, por ejemplo, etiquetas que son conocidas en el estado de la técnica, metilación y sustitución de uno o más de los nucleótidos que ocurren de forma natural con un nucleótido degenerado.

55 También de interés son las secuencias de ácido nucleico que son mayores de un 85%, con preferencia al menos un 90%, más preferiblemente al menos alrededor del 95%, y más preferiblemente al menos alrededor de un 98 a 99% idénticas u homólogas en la Figura 1 [SEQ ID Núm. 1]. También incluidos dentro de la invención están los

fragmentos de la Figura 1 [SEQ ID Núm. 1], su cadena complementaria, cADN y ARN complementarios con los mismos. Tales fragmentos incluyen las secuencias que codifican las proteínas variables (vp) de la cápside de AAV9 que son variantes de empalme alternativas: vp1 [nt 2116 a 4323 de la Figura 1, SEQ ID Núm. 1]; vp2 [nt 2527 a 4323 de la Figura 1, SEQ ID Núm. 1]; y vp3 [nt 2725 a 4323 de la Figura 1, SQ ID Núm. 1].

5 Incluso otros fragmentos incluyen los que codifican las proteínas rep, incluyendo la rep 78 [codón de iniciación en nt 228 de la Figura 1], la rep 68 [codón de iniciación en nt 228 de la Figura 1], la rep 52 [codón de iniciación en nt 900 de la Figura 1], y la rep 40 [codón de iniciación en nt 900 de la Figura 1]. Véase SEQ ID Núm. 1. Otros fragmentos de interés pueden incluir las repeticiones de terminal invertido (ITRs) de AAV, secuencias P19 de AAV, secuencias P40 de AAV, el sitio de enlace de rep, y el sitio resolutivo terminal (TRS). Otros fragmentos adecuados más resultarán  
10 fácilmente evidentes para el experto en la materia.

Adicionalmente a la inclusión de las secuencias de ácido nucleico proporcionadas en las figuras y en el Listado de Secuencias, la presente invención incluye moléculas y secuencias de ácido nucleico que están diseñadas para expresar las secuencias de aminoácido, proteínas y péptidos de los serotipos de AAV de la invención. La invención incluye secuencias de aminoácido que codifican las siguientes nuevas secuencias de aminoácido de AAV y los serotipos de AAV artificiales generados usando esas secuencias y/o fragmentos únicos de las mismas.  
15

Según se utiliza en la presente memoria, los serotipos de AAV artificiales incluyen, sin limitación, el AAV con una proteína de cápside que ocurre de forma no natural. Tal cápside artificial puede ser generada mediante cualquier técnica adecuada, usando una nueva secuencia de AAV de la invención (por ejemplo, un fragmento de una proteína de cápside de vp1) en combinación con secuencias homólogas que pueden ser obtenidas a partir de otro serotipo de AAV (conocido o nuevo), porciones no contiguas del mismo serotipo de AAV, procedente de una fuente viral de no AAV, o procedente de una fuente no viral. Un serotipo de AAV artificial pueden ser, sin limitación, una cápside de AAV quimérico, una cápside de AAV recombinante, o una cápside de AAV "humanizado".  
20

#### **B. Secuencias de aminoácido de AAV9, proteínas y péptidos**

La invención proporciona además proteínas y fragmentos de las mismas que son codificadas por ácidos nucleicos de AAV9 de la invención, y aminoácidos de AAV9 que son generados por medio de otros métodos. La invención abarca además serotipos de AAV generados usando secuencias del nuevo serotipo de AAV de la invención, que son generados usando técnicas sintéticas, recombinantes u otras técnicas conocidas por los expertos en la materia. La invención no se limita a las nuevas secuencias de aminoácido de AAV, los péptidos y las proteínas expresados a partir de las nuevas secuencias de ácido nucleico de AAV de la invención y abarca secuencias de aminoácido, péptidos y proteínas generados por medio de otros métodos conocidos en el estado de la técnica incluyendo, por ejemplo, síntesis química mediante otras técnicas sintéticas, o mediante otros métodos. Por ejemplo, cualesquiera secuencias que sean generadas fácilmente usando una diversidad de técnicas.  
25  
30

Las técnicas de producción adecuadas son bien conocidas por los expertos en la materia. Véase, por ejemplo, Sambrook et al., Clonación Molecular: Un Manual de Laboratorio, Cold Spring Harbor Press (Cold Spring Harbor, NY). Alternativamente, los péptidos pueden ser también sintetizados mediante los métodos bien conocidos de síntesis de péptidos de fase sólida (Merrifield, J. Am. Chem. Soc., 85: 2149 (1962); Steward y Young, Síntesis de Péptido de Fase Sólida (Freeman, San Francisco, 1069), pp. 27-62). Estos y otros métodos de producción adecuados están dentro del conocimiento de los expertos en la materia y no son una limitación de la presente invención.  
35

40 Las proteínas particularmente deseables incluyen las proteínas de cápside de AAV9 y las proteínas rep de AAV9.

Las proteínas particularmente deseables incluyen las proteínas de cápside de AAV, que son codificadas por las secuencias de nucleótido identificadas en lo que antecede. La cápside de AAV está compuesta por tres proteínas: las vp1, vp2 y vp3, las cuales son variantes de empalme alternativas. La secuencia de longitud completa proporcionada en la Figura 2 es la vp1. Las proteínas de cápside de AAV9 incluyen la vp1 [SEQ ID Núm. 2], la vp2 [aa 138 a 736 de SEQ ID Núm. 2], y la vp3 [aa 203 a 736 de SEQ ID Núm. 2], y fragmentos funcionales de las mismas. Otros fragmentos deseables de la proteína de cápside incluyen las regiones constantes y variables, localizadas entre regiones hipervariables (HPV). Otros fragmentos deseables de la proteína de cápside incluyen las propias HPV.  
45

Un algoritmo desarrollado para determinar áreas de divergencia de secuencia en AAV2 ha producido 12 regiones hipervariables (HVR) de las que 5 se solapan con, o son parte de, las cuatro regiones variables descritas con anterioridad. [Chiorini et al., J. Virol, 73: 1309-19 (1999); Rutledge, et al., J. Virol., 72: 309-319]. Usando este algoritmo y/o las técnicas de alineamiento descritas en la presente memoria, se determinan las HVR de los nuevos serotipos de AAV. Por ejemplo, con respecto al número de la vp1 de AAV2 [SEQ ID Núm. 4], las HVR están localizadas como sigue: HVR1, aa 146-152; HVR2, 11 182-186; HVR3, aa 262-264; HVR4, aa 381-383; HVR5, aa 450-474; HVR6, 490-495; HVR7, aa 500-504; HVR8, aa 514-522; HVR9, aa 534-555; HVR10, aa 581-594; HVR11, aa 658-667; y HVR12, aa 705-719. Usando el alineamiento proporcionado en la presente memoria llevado a cabo usando el programa Clustal X como configuración por defecto, o usando otros programas de alineamiento disponibles comercialmente o públicamente en configuraciones por defecto, un experto en la materia puede  
50  
55

determinar fácilmente fragmentos correspondientes de las nuevas cápsides de AAV de la invención.

Otros fragmentos deseables más de la proteína de cápside de AAV incluyen aminoácidos 1 a 184 de SEQ ID Núm. 2, aminoácidos 199 a 259; aminoácidos 274 a 446; aminoácidos 603 a 659; aminoácidos 670 a 706; aminoácidos 724 a 736 de SEQ ID Núm. 2; aa 185-198; aa 260-273; aa 447-477; aa 495-602; aa 660-669; y aa 707-723.

5 Adicionalmente, ejemplos de otros fragmentos adecuados de cápsides de AAV incluyen, con respecto a la numeración de AAV2 [SEQ ID Núm. 4], aa 24 a 42, aa 25 a 28; aa 81 a 85; aa 133 a 165; aa 134 a 165; aa 137 a 143; aa 154 a 156; aa 194 a 208; aa 261 a 274; aa 262 a 274; aa 171 a 173; aa 413 a 417; aa 449 a 478; aa 494 a 525; aa 534 a 571; aa 581 a 601; aa 660 a 671; aa 709 a 723. Otros fragmentos deseables más incluyen, por ejemplo, en AAV7, aminoácidos 1 a 184 de SEQ ID Núm. 2, aminoácidos 199 a 259; aminoácidos 274 a 446; aminoácidos 603 a 659; aminoácidos 670 a 706; aminoácidos 724 a 736; aa 185 a 198; aa 260 a 273; aa 447 a 477; aa 495 a 602; aa 660 a 669; y aa 707 a 723. Usando el alineamiento proporcionado en la presente memoria llevado a cabo usando el programa Clustal X en configuraciones por defecto, o usando otros programas de alineamiento disponibles comercialmente o públicamente en configuraciones por defecto, un experto en la materia puede determinar fácilmente fragmentos correspondientes de las nuevas cápsides de AAV de la invención.

15 Otras proteínas de AAV deseables más incluyen las proteínas rep incluyendo la rep68/78 y la rep40/52 [localizada dentro de SEQ ID Núm. 3]. Fragmentos adecuados de las proteínas rep pueden incluir aa 1 a 102; aa 103 a 140; aa 141 a 173; aa 174 a 226; aa 227 a 275; aa 276 a 374; aa 375 a 383; aa 384 a 446; aa 447 a 542; aa 543 a 555; aa 556 a 623, de SEQ ID Núm. 3.

20 Adecuadamente, los fragmentos son de al menos 8 aminoácidos de longitud. Sin embargo, Se pueden utilizar fácilmente fragmentos de otras longitudes deseadas. Tales fragmentos pueden ser producidos recombinantemente o mediante otros medios adecuados, por ejemplo por síntesis química.

25 La invención proporciona además otras secuencias de AAV9 que se identifican usando la información de secuencia proporcionada en la presente memoria. Por ejemplo, dadas las secuencias de AAV9 proporcionadas en la presente memoria, el AAV9 infeccioso puede ser aislado usando tecnología de genoma andante (Siebert, et al., 1995, Búsqueda de Ácido Nucleico, 23: 1087-1088; Friezner-Degen et al., 1986, J. Biol. Chem. 261: 6972-6985, BD Biosciences Clontech, Palo Alto, CA). El genoma andante es particularmente muy adecuado para identificar y aislar las secuencias adyacentes a las nuevas secuencias identificadas conforme al método de la invención. Esta técnica es también útil para aislar repeticiones de terminal invertido (ITRs) del nuevo serotipo de AAV9, en base a la nueva cápside de AAV y a las secuencias rep que se proporcionan en la presente memoria.

30 Las secuencias, proteínas y fragmentos de la invención pueden ser producidos mediante cualquier medio adecuado, incluyendo producción recombinante, síntesis química, u otro medio sintético. Tales métodos de producción caen dentro del conocimiento del experto en la materia y no son una limitación de la presente invención.

#### IV. Producción de rAAV con cápsides de AAV9

35 La invención abarca las nuevas secuencias de AAV9 de tipo natural, de las que aquellas que están libres de ADN y/o de material celular con esos virus, están asociadas en la naturaleza. En otro aspecto, la presente invención proporciona moléculas que utilizan las nuevas secuencias de AAV de la invención, incluyendo los fragmentos de las mismas, para la producción de moléculas útiles en el suministro de un gen heterólogo o de otras secuencias de ácido nucleico a una célula objetivo.

40 En otro aspecto, la presente invención proporciona moléculas que utilizan las secuencias de AAV9 de la invención, incluyendo los fragmentos de las mismas, para la producción de vectores virales útiles en el suministro de un gen heterólogo o de otras secuencias de ácido nucleico a una célula objetivo.

45 Las moléculas de la invención que contienen secuencias de AAV9 incluyen cualquier elemento (vector) genético que pueda ser suministrado a una célula anfitrión, por ejemplo ADN desnudo, un plásmido, fago, transposón, cósmido, episoma, una proteína en un vehículo de suministro no viral (por ejemplo, un portador a base de lípido), virus, etc., que transfiere las secuencias portadas por el mismo. El vector seleccionado puede ser suministrado mediante cualquier método adecuado, incluyendo transfección, electroporación, suministro de liposoma, técnicas de fusión de membrana, gránulos recubiertos de ADN de alta velocidad, infección viral y fusión de protoplasto. Los métodos usados para construir cualquier realización de la presente invención, son conocidos por los expertos en manipulación de ácido nucleico e incluyen ingeniería genética, ingeniería recombinante, y técnicas sintéticas. Véase, por ejemplo, Sambrook et al., Clonación Molecular: Un Manual de Laboratorio, Cold Spring Harbor Press, Cold Spring Harbor, NY.

50 En una realización, los vectores de la invención contienen, como mínimo, secuencias que codifican una cápside de AAV9 o un fragmento de la misma. En otra realización, los vectores de la invención contienen, como mínimo, secuencias que codifican una proteína rep de AAV9 o un fragmento de la misma. Opcionalmente, tales vectores pueden contener ambas proteínas rep y cap de AAV. En vectores en los que se proporcionan ambas proteínas rep y cap de AAV, las secuencias rep de AAV y cap de AAV pueden ser ambas el origen de AAV9. Alternativamente, la presente invención proporciona vectores en los que las secuencias rep proceden de un serotipo de AAV que difiere

del que está proporcionando las secuencias *cap*. En una realización, las secuencias *rep* y *cap* son expresadas a partir de fuentes separadas (por ejemplo, de vectores separados, o de una célula anfitrión y un vector). En otra realización, esas secuencias *rep* son fusionadas en un marco de secuencias *cap* de un serotipo de AAV diferente para formar un vector de AAV quimérico. Opcionalmente, los vectores de la invención contienen además un minigén que comprende un transgén seleccionado que está flanqueado por la ITR 5' de AAV y la ITR 3' de AAV.

De ese modo, en una realización, los vectores descritos en la presente memoria contienen secuencias de ácido nucleico que codifican una cápside de AAV intacta que puede proceder de un serotipo de AAV simple (por ejemplo, de AAV9), Dicha cápside puede comprender aminoácidos 1 a 736 de SEQ ID Núm. 2. Alternativamente, estos vectores contienen secuencias que codifican cápsides artificiales que contienen uno o más fragmentos de la cápside de AAV9 fusionados en AAV heterólogo o en proteínas de cápside que no son de AAV (o en fragmentos de las mismas). Estas proteínas de cápside artificiales se seleccionan a partir de porciones no contiguas de la cápside de AAV9 o a partir de cápsides de los otros serotipos de AAV. Por ejemplo, un rAAV puede tener una proteína de cápside que comprenda una o más de las regiones de cápside de AAV9 seleccionadas a partir de la vp2 y/o la vp3, o a partir de la vp1, o de fragmentos de las mismas seleccionados a partir de los aminoácidos 1 a 184, los aminoácidos 199 a 259; los aminoácidos 274 a 446; los aminoácidos 603 a 659; los aminoácidos 670 a 706; los aminoácidos 724 a 736 de la cápside de AAV9, SEQ ID Núm. 2. En otro ejemplo, puede resultar deseable alterar el codón de inicio de la proteína vp3 en GTG. Alternativamente, el rAAV puede contener una o más de las regiones hipervariables de proteína de cápside del serotipo 9 de AAV, las cuales están identificadas en la presente memoria, u otros fragmentos incluyendo, sin limitación, aa 185 a 198; aa 260-273; aa 447 a 477; aa 495 a 602; aa 660 a 669; y aa 707 a 723, de la cápside de AAV9. Véase SEQ ID Núm. 2. Estas modificaciones pueden estar destinadas a incrementar la expresión, producción y/o mejorar la purificación en los sistemas de expresión seleccionados, o para otros propósitos deseados (por ejemplo, para cambiar tropismo o alterar epitopes de anticuerpo neutralizante).

Los vectores descritos en la presente memoria, por ejemplo un plásmido, son útiles para una diversidad de propósitos, pero son particularmente adecuados para su uso en la producción de un rAAV que contenga una cápside que comprenda una secuencia de AAV o un fragmento de la misma. Estos vectores, incluyendo el rAAV, sus elementos, construcción y usos, se describen con detalle en la presente memoria.

En un aspecto, la invención proporciona un método de generar un virus adeno-asociado recombinante (AAV) que tiene una cápside del serotipo 9 de AAV, o una porción de la misma. Dicho método incluye cultivar una célula anfitrión que contiene una secuencia de ácido nucleico que codifica una proteína de cápside del serotipo 9 de un virus adeno-asociado (AAV), o un fragmento de la misma, según se define en la presente memoria; un gen *rep* funcional; un minigén compuesto, como mínimo, por repeticiones de terminal invertido (ITRs) de AAV y un transgén; y funciones auxiliares suficientes para permitir el empaquetamiento del minigén en la proteína de cápside de AAV9.

Los componentes que se requiere que sean cultivados en la célula anfitrión para empaquetar el minigén de AAV en una cápside de AAV, pueden ser proporcionados a la célula anfitrión en *trans*. Alternativamente, uno cualquiera o más de los componentes requeridos (por ejemplo, minigén, secuencias *rep*, secuencias *cap*, y/o funciones auxiliares) pueden ser proporcionados mediante una célula anfitrión estable que haya sido diseñada para contener uno o más de los componentes requeridos usando métodos conocidos por los expertos en la materia. Más adecuadamente, dicha célula anfitrión estable contendrá el (los) componente(s) requerido(s) bajo el control de un promotor inducible. Sin embargo, el (los) componente(s) requerido(s) puede(n) estar bajo el control de un promotor constitutivo. Ejemplos de promotores inducibles adecuados y de promotores constitutivos adecuados se proporcionan en la presente memoria, en la discusión de elementos reguladores adecuados para su uso con el transgén. En otra alternativa más, una célula anfitrión seleccionada puede contener el (los) componente(s) seleccionado(s) bajo el control de un promotor constitutivo y de otro(s) componente(s) seleccionado(s) bajo el control de uno o más promotores inducibles. Por ejemplo, se puede generar una célula anfitrión estable que sea derivada de células 293 (las cuales contienen funciones auxiliares E1 bajo el control de un promotor constitutivo), pero que contenga las proteínas *rep* y *cap* bajo el control de promotores inducibles. Otras células anfitrión adecuadas adicionales pueden ser generadas por un experto en la materia.

El minigén, las secuencias *rep*, las secuencias *cap* y las funciones auxiliares requeridos para la producción del rAAV de la invención, pueden ser suministrados a la célula anfitrión de empaquetamiento en forma de cualquier elemento genético que transfiera las secuencias portadas por el mismo. El elemento genético seleccionado puede ser suministrado mediante cualquier método adecuado, incluyendo los que se describen en la presente memoria. Los métodos usados para construir cualquier realización de la presente invención son conocidos por los expertos en manipulación de ácido nucleico e incluyen ingeniería genética, ingeniería recombinante, y técnicas sintéticas. Véase, por ejemplo, Sambrook et al., Clonación Molecular: Un Manual de Laboratorio, Cold Spring Harbor Press, Cold Spring Harbor, NY. De forma similar, los métodos de generación de viriones de rAAV son bien conocidos y la selección de un método adecuado no es una limitación de la presente invención. Véase, por ejemplo, K. Fisher et al., J. Virol., 70: 520-532 (1993) y la Patente US 5.478.745.

A menos que se especifique otra cosa, las ITRs de AAV y otros componentes de AAV seleccionados descritos en la presente memoria, pueden ser seleccionados fácilmente a partir de entre cualquier serotipo de AAV, incluyendo sin limitación el AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, y el nuevo serotipo de la invención, el AAV9. Estas

ITRs u otros componentes de AAV pueden ser aislados fácilmente usando técnicas disponibles para los expertos en la materia a partir de un serotipo de AAV. Dicho AAV puede ser aislado u obtenido a partir de fuentes académicas, comerciales o públicas (por ejemplo, la American Type Culture Collection, Manassas, VA). Alternativamente, las secuencias de AAV pueden ser obtenidas a través de medios sintéticos o de otros medios adecuados por referencia a secuencias publicadas tal como las que están disponibles en la literatura o en las bases de datos, como por ejemplo, en GenBank, PubMed o similares.

## A. El minigén

El minigén está compuesto, como mínimo, por un transgén y sus secuencias reguladoras, y las repeticiones de terminal invertido (ITRs) 5' y 3' de AAV. En una realización deseable, se utilizan las ITRs del serotipo 2 de AAV. Sin embargo, se pueden seleccionar las ITRs de otros serotipos adecuados. Este minigén es el que se empaqueta en una proteína de cápside y se suministra a una célula anfitrión seleccionada.

### 1. El transgén

El transgén es una secuencia de ácido nucleico, heteróloga respecto a las secuencias de vector que flanquean el transgén, que codifica un polipéptido, una proteína u otro producto de interés. La secuencia de codificación de ácido nucleico está operativamente enlazada a componentes reguladores de una manera que permite la transcripción del transgén, y/o la expresión en una célula anfitrión.

La composición de la secuencia de transgén dependerá del uso para el que se disponga el vector resultante. Por ejemplo, un tipo de secuencia de transgén incluye una secuencia informadora, la cual, tras la expresión, produce una señal detectable. Tales secuencias informadoras incluyen, sin limitación, secuencias de ADN que codifican  $\beta$ -lactamasa,  $\beta$ -galactosidasa (LacZ), fosfatasa alcalina, timidina quinasa, proteína fluorescente verde (GFP), cloranfenicol acetiltransferasa (CAT), luciferasa, proteína de enlace de membrana incluyendo, por ejemplo, CD2, CD4, CD8, la proteína de hemaglutinina de la influenza, y otros bien conocidos en el estado de la técnica, para los que existen anticuerpos de alta afinidad dirigidos a los mismos o pueden ser producidos con medios convencionales, y proteínas de fusión que comprenden una proteína de enlace de membrana fusionada apropiadamente con un dominio de etiqueta de antígeno procedente, entre otros, de la hemaglutinina o Myc.

Estas secuencias de codificación, cuando se asocian a elementos reguladores que activan su expresión, proporcionan señales detectables con medios convencionales, incluyendo los ensayos enzimático, radiográfico, colorimétrico, de fluorescencia u otros ensayos espectrográficos, ensayos de clasificación celular de activación fluorescente y ensayos inmunológicos, incluyendo el ensayo inmunoabsorbente enlazado a enzima (ELISA), radioinmunoensayo (RIA) e inmunohistoquímica. Por ejemplo, cuando la secuencia marcadora es el gen LacZ, la presencia del vector que porta la señal es detectada mediante ensayos respecto a actividad de beta-galactosidasa. Cuando el transgén es proteína fluorescente o luciferasa, el vector que porta la señal puede ser medido visualmente por producción de color o de luz en un luminómetro.

Sin embargo, deseablemente, el transgén es una secuencia no marcadora que codifica un producto que es útil en biología y en medicina, tal como las proteínas, los péptidos, el ARN, las enzimas o los ARNs catalíticos. Las moléculas de ARN deseables incluyen el tARN, dsARN, ARN ribosómico, ARNs catalíticos, y ARNs antisentido. Un ejemplo de una secuencia de ARN útil es una secuencia que extingue la expresión de una secuencia de ácido nucleico en el animal tratado.

El transgén puede ser usado para corregir o mejorar deficiencias genéticas, las cuales pueden incluir deficiencias en los genes normales son expresados a niveles menores que el normal, o deficiencias en las que el producto de gen funcional no está expresado. Un tipo preferido de secuencia de transgén codifica una proteína terapéutica o un polipéptido que está expresado en una célula anfitrión. La invención incluye además el uso de múltiples transgenes, por ejemplo, para corregir o mejorar un defecto genético causado por una proteína multi-subunidad. En ciertas situaciones, se puede usar un transgén diferente para codificar cada subunidad de una proteína, o para codificar diferentes péptidos o proteínas. Esto es deseable cuando el tamaño del ADN que codifica la subunidad de proteína es grande, por ejemplo para una inmunoglobulina, el factor de crecimiento derivado de plaqueta, o una proteína distrofina. Con el fin de que la célula produzca la proteína multi-subunidad, una célula se infecta con el virus recombinante que contiene cada una de las diferentes subunidades. Alternativamente, diferentes subunidades de una proteína pueden ser codificadas mediante el mismo transgén. En este caso, un solo transgén incluye el ADN que codifica cada una de las subunidades, estando el ADN para cada subunidad separado por un sitio de entrada de ribozima interno (IRES). Esto es deseable cuando el tamaño del ADN que codifica cada una de las subunidades es pequeño, por ejemplo el tamaño total del ADN que codifica las subunidades y el IRES es menor de cinco kilobases. Como alternativa a un IRES, el ADN puede estar separado por medio de secuencias que codifican un péptido 2A, el cual se auto-escinde en un evento post-traslacional. Véase, por ejemplo, M.L. Donnelly, et al., J. Gen. Virol., 78(Pt 1); 13-21 (Enero de 1997); Furter, S., et al, Gene Ther., 8(11): 864-873 (Junio de 2001); Klump H., et al., Gene Ther., 8(10): 811-817 (Mayo de 2001). Este péptido 2A es significativamente más pequeño que un IRES, lo que hace que sea muy adecuado para su uso cuando el espacio es un factor limitativo. Sin embargo, el transgén seleccionado puede codificar cualquier producto biológicamente activo u otro producto, por ejemplo, un producto deseable para su estudio.

Los transgenes adecuados pueden ser seleccionados fácilmente por un experto en la materia. La selección del transgén no se considera una limitación de la presente invención.

## 2. Elementos reguladores

5 Adicionalmente a los elementos principales identificados con anterioridad para el minigén, el vector incluye también elementos de control convencionales necesarios que están enlazados operativamente al transgén de una manera que permite su transcripción, traducción y/o expresión en una célula transfectada con el vector de plásmido o infectada con el virus producido mediante la invención. Según se utiliza en la presente memoria, las secuencias "operativamente enlazadas" incluyen tanto secuencias de control de expresión que son contiguas con el gen de interés, como secuencias de control de expresión que actúan en *trans* o a una distancia para controlar el gen de interés.

15 Las secuencias de control de expresión incluyen secuencias apropiadas de iniciación de transcripción, terminación promotoras y potenciadoras; señales de procesamiento eficiente de ARN tal como señales de empalme y poliadenilación (poliA); secuencias que estabilizan mRNA citoplásmico; secuencias que potencian la eficacia de traducción (es decir, la secuencia de consenso de Kozak); secuencias que potencian estabilidad de proteína; y cuando se desee, secuencias que potencian la secreción del producto codificado. Un gran número de secuencias de control de expresión, incluyendo los promotores que sean naturales, constitutivos, inducibles y/o específicos del tejido, son conocidos en el estado de la técnica y pueden ser utilizados.

20 Ejemplos de promotores constitutivos incluyen, sin limitación, el promotor LTR del virus de sarcoma de Rous (RSV) retroviral (opcionalmente con el potenciador de RSV), el promotor de citomegalovirus (CMV) (opcionalmente con el potenciador de CMV) [véase, por ejemplo, Boshart et al., *Célula*, 41: 521-530 (1985)], el promotor SV40, el promotor de dihidrofolato reductasa, el promotor de  $\beta$ -actina, el promotor de fosfoglicerol quinasa (PGK), y el promotor EF1 $\alpha$  [Invitrogen].

25 Los promotores inducibles permiten la regulación de expresión de gen y pueden ser regulados mediante compuestos suministrados exógenamente, factores ambientales tales como la temperatura, o mediante la presencia de un estado fisiológico específico, por ejemplo fase aguda, un estado de diferenciación particular de la célula, o en células replicantes solamente. Los promotores inducibles y los sistemas inducibles están disponibles a partir de una diversidad de fuentes comerciales incluyendo, sin limitación, Invitrogen, Clontech y Ariad. Muchos otros sistemas han sido descritos y pueden ser seleccionados fácilmente por un experto en la materia. Ejemplos de promotores inducibles regulados mediante compuestos suministrados exógenamente incluyen, por ejemplo, el promotor de metalotionina (MT) de oveja inducible por zinc, el promotor del virus de tumor mamario de ratón (MMTV) inducible por dexametasona (Dex), el sistema promotor de polimerasa T7 [WO 98/10088]; el promotor de insecto de acdysona [No et al., *Proc. Natl. Acad. Sci. USA*, 93: 3346-3351 (1996)], el sistema represible por tetraciclina [Gossen et al., *Proc. Natl. Acad. Sci. USA*, 89: 5547-5551 (1992)], el sistema inducible por tetraciclina [Gossen et al., *Science*, 268: 1766-1769 (1995), véase también Harvey et al., *Curr. Opin. Chem. Biol.*, 2: 512-518 (1998)], el sistema inducible por RU486 [Wang et al., *Nat. Biotech.*, 15: 239-243 (1997), y Wang et al., *Gene Ther.*, 4: 432-441 (1997)] y el sistema inducible por rapamicina [Magari et al., *J. Clin. Invest.*, 100: 2865-2872 (1997)]. Otros tipos de promotores inducibles que pueden ser útiles en este contexto y los que son regulados mediante un estado fisiológico específico, por ejemplo la temperatura, fase aguda, un estado de diferenciación articular de la célula o en células replicantes solamente.

40 En otra realización, se podrá utilizar el promotor natural para el transgén. Se puede preferir el promotor natural cuando se desea que la expresión del transgén imite la expresión natural. El promotor natural puede ser usado cuando la expresión del transgén deba ser regulada temporalmente o por desarrollo, o de una manera específica del tejido, o en respuesta a estímulos transcripciones específicos. En una realización adicional, otros elementos de control de expresión natural, tal como elementos potenciadores, sitios de poliadenilación o secuencias de consenso de Kozak, pueden ser también usados para mimetizar la expresión natural.

45 Otra realización del transgén incluye un transgén operativamente enlazado a un promotor específico del tejido. Por ejemplo, si se desea expresión en el músculo esquelético, se deberá usar un promotor activo en el músculo. Éstos incluyen los promotores procedentes de genes que codifican la  $\beta$ -actina esquelética, la cadena ligera de miosina 2A, la distrofina, la creatina quinasa del músculo, así como promotores musculares sintéticos con actividades más altas que los promotores que ocurren de forma natural (véase Li et al., *Nat. Biotech.*, 17: 241-245 (1999)). Ejemplos de promotores que son específicos del tejido son conocidos para el hígado (albúmina, Miyatake et al., *J. Virol.* 71: 5124-32 (1997); promotor de núcleo de virus de hepatitis B, Sandig et al., *Gene Ther.*, 3: 1002-9 (1996); alfa-fetoproteína (AFP), Arbutnot et al., *Hum. Gene Ther.*, 7: 1503-14 (1996)), osteocalcina ósea (Stein et al., *Mol. Biol. Rep.*, 24: 185-96 (1997)); sialoproteína ósea (Chen et al., *J. Bone Miner. Res.*, 11: 654-64 (1996)); linfocitos (CD2, Hansal et al., *J. Immunol.* 161: 1063-8 (1998); cadena pesada de inmunoglobulina; cadena de receptor  $\alpha$  de células T), neuronales tal como el promotor de enolasa específico de neurona (NSE) (Anderseen et al., *Célula, Mol. Neurobiol.*, 13: 503-15 (1993)), gen de cadena ligera de neurofilamento (Piccioli et al., *Proc. Natl. Acad. Sci. USA*, 88: 5611-5 (1991)), y el gen vgf específico de neurona (Piccioli et al., *Neurona*, 15: 373-84 (1995)), entre otros.

Opcionalmente, los plásmidos portadores de transgenes terapéuticamente útiles pueden incluir también genes

5 marcadores o informadores seleccionables que pueden incluir secuencias que codifiquen genitocina, higromicina o resistencia a purimicina, entre otros. Tales genes informadores o marcadores seleccionables (con preferencia, situados fuera del genoma viral que va a ser rescatado mediante el método de la invención), pueden ser usados para señalar la presencia de los plásmidos en células bacterianas, tal como resistencia a la ampicilina. Otros componentes del plásmido pueden incluir un origen de replicación. La selección de estos y de otros promotores y elementos de vector, es convencional y muchas de tales secuencias se encuentran disponibles [véase, por ejemplo, Sambrook et al., y las referencias citadas en la presente memoria].

10 La combinación del transgén, promotor/potenciador e ITRs 5' y 3', se conoce como "minigén" para facilidad de referencia en la presente memoria. Siempre según las enseñanzas de la presente invención, el diseño de un minigén de ese tipo puede hacerse recurriendo a técnicas convencionales.

### 3. Suministro del minigén a una célula anfitrión de empaquetamiento

15 El minigén puede ser portado sobre cualquier vector adecuado, por ejemplo, un plásmido, que sea suministrado a una célula anfitrión. Los plásmidos útiles en esta invención pueden estar diseñados de tal modo que sean adecuados para replicación y, opcionalmente, para integración en células procariotas, células de mamífero, o en ambas. Estos plásmidos (u otros vectores portadores de la ITR 5' de AAV-molécula heteróloga-ITR 3') contienen secuencias que permiten la replicación del minigén en eucariotas y/o procariotas y marcadores de selección para esos sistemas. Genes marcadores o informadores seleccionables pueden incluir secuencias que codifican la genitocina, higromicina o la resistencia a la purimicina, entre otros. Los plásmidos pueden contener también ciertos genes informadores o marcadores seleccionables que pueden ser usados para señalar la presencia del vector en células bacterianas, tal como la resistencia a la ampicilina. Otros componentes del plásmido pueden incluir un origen de replicación y un amplicón, tal como el sistema amplicón que emplea el antígeno nuclear del virus de Epstein Barr. Este sistema amplicón, u otros componentes de amplicón similares, permiten una replicación episómica de alta copia en las células. Con preferencia, la molécula portadora del minigén es transfectada en la célula, donde puede existir de forma transitoria. Alternativamente, el minigén (portador de la ITR 5' de AAV-molécula heteróloga-ITR 3') puede estar integrado establemente en el genoma de la célula anfitrión, ya sea cromosómicamente o ya sea como episoma. En algunas realizaciones, el minigén puede estar presente en múltiples copias, opcionalmente en concatámeros de cabeza con cabeza, cabeza con cola, o cola con cola. Las técnicas de transfección adecuadas son conocidas y pueden ser utilizadas fácilmente para suministrar el minigén a la célula anfitrión.

20 En general, cuando se suministra el vector que comprende el minigén mediante transfección, el vector es suministrado en una cantidad de aproximadamente 5 µg a aproximadamente 100 µg de ADN, y con preferencia aproximadamente 10 a aproximadamente 50 µg de ADN a aproximadamente  $1 \times 10^4$  células a aproximadamente  $1 \times 10^{13}$  células, y con preferencia aproximadamente  $10^5$  células. Sin embargo, las cantidades relativas de ADN de vector respecto a células anfitrión, pueden ser ajustadas teniendo en cuenta factores tales como el vector seleccionado, el método de suministro y las células anfitrión seleccionadas.

#### 35 Secuencias *rep* y *cap*

40 Adicionalmente al minigén, la célula anfitrión contiene las secuencias que activan expresión de la proteína de cápside de AAV9 (o una proteína de cápside que comprende un fragmento de la cápside de AAV9) en la célula anfitrión y secuencias *rep* del mismo serotipo que el serotipo de las ITRs de AAV encontradas en el minigén, o un serotipo de complementación cruzada. Las secuencias *cap* y *rep* de AAV pueden ser obtenidas independientemente a partir de una fuente de AAV según se ha descrito con anterioridad, y pueden ser introducidas en la célula anfitrión de cualquier manera conocida por un experto en la técnicas según se ha descrito con anterioridad. Adicionalmente, cuando se seudotipa un vector de AAV en una cápside de AAV9, las secuencias que codifican cada una de las proteínas *rep* esenciales pueden ser suministradas mediante AAV9, o las secuencias que codifican las proteínas *rep* pueden ser suministradas por serotipos de AAV diferentes (por ejemplo, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7). Por ejemplo, las secuencias *rep78/68* pueden proceder de AAV2, mientras que las secuencias *rep52/40* pueden proceder de AAV1.

50 En una realización, la célula anfitrión contiene establemente la proteína de cápside bajo el control de un promotor adecuado, tal como los que se han descrito con anterioridad. Más deseablemente, en esta realización, la proteína de cápside se expresa bajo el control de un promotor inducible. En otra realización, la proteína de cápside se suministra a la célula anfitrión en *trans*. Cuando se suministra a la célula anfitrión en *trans*, la proteína de cápside puede ser suministrada por medio de un plásmido que contenga las secuencias necesarias para dirigir expresión de la proteína de cápside seleccionada en la célula anfitrión. Más deseablemente, cuando se suministra a la célula anfitrión en *trans*, el plásmido que porta la proteína de cápside porta también otras secuencias requeridas para empaquetar el rAAV, por ejemplo, las secuencias *rep*.

55 En otra realización, la célula anfitrión contiene establemente las secuencias *rep* bajo el control de un promotor adecuado tal como los que se han descrito con anterioridad. Más deseablemente, en esta realización, las proteínas *rep* esenciales se expresan bajo el control de un promotor inducible. En otra realización, las proteínas *rep* son suministradas a la célula anfitrión en *trans*. Cuando se suministran a la célula anfitrión en *trans*, las proteínas *rep* pueden ser suministradas por medio de un plásmido que contenga las secuencias necesarias para dirigir expresión

de las proteínas *rep* seleccionadas en la célula anfitrión. Más deseablemente, cuando se suministra a la célula anfitrión en *trans*, el plásmido que porta la proteína de cápside porta también otras secuencias requeridas para empaquetar el rAAV, por ejemplo las secuencias *rep* y *cap*.

5 Así, en una realización, las secuencias *rep* y *cap* pueden ser transfectadas en la célula anfitrión en sola molécula de ácido nucleico y existir establemente en la célula como un episoma. En otra realización, las secuencias *rep* y *cap* están integradas establemente en el genoma de la célula. Otra realización tiene las secuencias *rep* y *cap* expresadas transitoriamente en la célula anfitrión. Por ejemplo, una molécula de ácido nucleico útil para tal transfección comprende, desde 5' hasta 3', un promotor, un espaciador opcional intercalado entre el promotor y el sitio de la secuencia de gen *rep*, una secuencia de gen *rep* de AAV, y una secuencia de gen *cap* de AAV.

10 Opcionalmente, las secuencias *rep* y *cap* pueden ser suministradas sobre un vector que contenga otras secuencias de ADN que pueden ser introducidas en las células anfitrión. Por ejemplo, el vector puede contener la construcción de rAAV que comprende el minigén. El vector puede comprender uno o más de los genes que codifican las funciones auxiliares, por ejemplo, las proteínas adenovirales E1, E2a, y E4ORF6, y el gen para VAI ARN.

15 Con preferencia, el promotor utilizado en esta construcción puede ser cualquiera de entre los promotores constitutivos, inducibles o nativos conocidos por el experto en la materia, o según se ha discutido en lo que antecede. En una realización, se emplea una secuencia de promotor P5 de AAV. La selección del AAV para proporcionar cualquiera de esas secuencias no limita la invención.

20 En otra realización preferida, el promotor para *rep* es un promotor inducible, de los que se han descrito muchos en lo que antecede en relación con los elementos reguladores de transgén. El vector que comprende el gen *rep* regulado por el promotor T7 y el gen *cap*, es transfectado o transformado en una célula que expresada ya sea constitutivamente o ya sea induciblemente la polimerasa T7. Véase el documento WO 98/10088, publicado el 12 de Marzo de 1998.

25 El espaciador es un elemento opcional en el diseño del vector. El espaciador es una secuencia de ADN intercalada entre el promotor y el sitio de inicio de ATG de gen *rep*. El espaciador puede tener cualquier diseño deseado, es decir, puede ser una secuencia aleatoria de nucleótidos, o alternativamente, puede codificar un producto de gen, tal como un gen marcador. El espaciador puede contener genes que incorporen típicamente sitios de inicio/interrupción y poliA. El espaciador puede ser una secuencia de ADN no codificadora procedente de una secuencia procariota o eucariota, una secuencia repetitiva no codificadora, una secuencia codificadora sin controles de transcripción o una secuencia codificadora con controles transcripcionales. Dos ejemplos de fuente de secuencias espaciadoras son las  
30 secuencias en escalera  $\gamma$  fago o secuencias en escalera de levadura, las cuales están disponibles comercialmente, por ejemplo a partir de Gibco o de Invitrogen, entre otros. El espaciador puede ser de cualquier tamaño suficiente para reducir la expresión de los productos de gen *rep78* y *rep68*, dejando los productos de gen *rep52*, *rep40* y *cap* expresados a niveles normales. La longitud del espaciador puede estar por lo tanto en la gama de aproximadamente 10 bp a aproximadamente 10,0 kbp, con preferencia en la gama de aproximadamente 100 bp a aproximadamente  
35 8,0 kbp. Para reducir la posibilidad de recombinación, el espaciador es con preferencia menor de 2 kbp de longitud; sin embargo, la invención no está limitada por ello.

Aunque la(s) molécula(s) que proporciona(n) *rep* y *cap* pueden existir en la célula anfitrión de forma transitoria (es decir, mediante transfección), se prefiere que una o ambas de las proteínas *rep* y *cap* y el (los) promotor(es) que controla(n) su expresión este(n) expresada(s) establemente en la célula anfitrión, por ejemplo como un episoma o  
40 por integración en el cromosoma de la célula anfitrión. Los métodos empleados para construir realizaciones de la presente invención son técnicas convencionales de ingeniería genética o de ingeniería recombinante tal como las que se han descrito en las referencias que anteceden. Mientras que esta descripción proporciona ejemplos ilustrativos de construcciones específicas, usando la información proporcionada en la presente memoria un experto en la materia puede seleccionar y diseñar otras construcciones adecuadas, usando una opción espaciadores, promotores P5, y otros elementos, incluyendo al menos una señal de inicio e interrupción transcripcional, y la adición  
45 opcional de sitios de poliadenilación.

En otra realización de la presente invención, la proteína *rep* o *cap* puede ser proporcionada de forma estable por una célula anfitrión.

### C. Las funciones auxiliares

50 La célula anfitrión de empaquetamiento requiere también funciones auxiliares con el fin de empaquetar el rAAV de la invención. Opcionalmente, esas funciones pueden ser suministradas por un herpesvirus. Más deseablemente, cada una de las funciones auxiliares necesarias es proporcionada a partir de una fuente de adenovirus de primate no humano o humano, tal como se ha descrito en la presente memoria, o que están disponibles a partir de una diversidad de fuentes, incluyendo la ATCC. En una realización normalmente preferida, la célula anfitrión está dotada  
55 y/o contiene un producto de gen E1a, un producto de gen E1b, y/o un producto de gen E4 ORF6. La célula anfitrión puede contener otros genes adenovirales como el VAI ARN, pero estos genes no se necesitan. En una realización preferida, ningún otro gen de adenovirus ni funciones de gen están presentes en la célula anfitrión.

Mediante "ADN adenoviral que expresa el producto de gen E1a", se indica cualquier secuencia de adenovirus que codifique E1a o una porción funcional de E1a. El ADN adenoviral que expresa el producto de gen E2a y el ADN adenoviral que expresa los productos de gen E4 ORF6, se definen de forma similar. También están incluidos cualesquiera alelos u otras modificaciones del gen adenoviral o de una porción funcional del mismo. Tales modificaciones pueden ser introducidas deliberadamente recurriendo a técnicas convencionales de ingeniería genética o mutagénicas para potenciar la función adenoviral de alguna manera, así como variantes alélicas de los mismos que ocurren de forma natural. Tales modificaciones y métodos para manipular ADN para conseguir esas funciones de gen de adenovirus son conocidos por los expertos en la materia.

Los productos de gen E1a, E1b, E2a y/o E4 ORF6 de adenovirus, así como cualesquiera otras funciones auxiliares deseadas, pueden ser proporcionados usando cualquier medio que permita su expresión en una célula. Cada una de las secuencias que codifican esos productos puede estar sobre un vector separado, o uno o más genes pueden estar sobre el mismo vector. El vector puede ser cualquier vector conocido en el estado de la técnica o descrito en lo que antecede, incluyendo los plásmidos, cósmidos y virus. La introducción en la célula anfitrión del vector puede lograrse mediante cualquier medio conocido en el estado de la técnica o según se ha descrito con anterioridad, incluyendo transfección, infección, electroporación, suministro de liposoma, técnicas de fusión de membrana, gránulos recubiertos de ADN de alta velocidad, infección viral y fusión de protoplasto, entre otros. Uno o más de los genes adenovirales pueden estar integrados establemente en el genoma de la célula anfitrión, expresados establemente como episomas, o expresados transitoriamente. Los productos de gen pueden estar todos expresados transitoriamente, sobre un episoma o integrados establemente, o algunos de los productos de gen pueden estar expresados de forma estable mientras otros están expresados transitoriamente. Además, se pueden seleccionar los promotores para cada uno de los genes adenovirales de forma independiente a partir de un promotor constitutivo, un promotor inducible o un promotor adenoviral nativo. Los promotores pueden estar regulados por un estado fisiológico específico del organismo o de la célula (es decir, por el estado de diferenciación o en células replicantes o quiescentes), o por factores añadidos exógenamente, por ejemplo.

#### **D. Células anfitrión y líneas de células de empaquetamiento**

La célula anfitrión en sí misma puede ser seleccionada a partir de cualquier organismo biológico, incluyendo las células procarióticas (por ejemplo, bacterianas), y las células eucarióticas, incluyendo células de insecto, células de levadura y células de mamífero. Las células anfitrión particularmente deseables se seleccionan entre cualesquiera especies de mamífero, incluyendo sin limitación células tales como A549, WEHI, 3T3, 10T1/2, BHK, MDCK, COS 1, COS 7, BSC 1, BSC 40, BMT 10, VERO WI38, HeLa, células 293 (que expresan E1 adenoviral funcional), Saos, C2C12, células L, HT1080, HepG2, y fibroblasto primario, células de hepatocito y de mioblasto derivadas de mamíferos incluyendo el ser humano, el mono, el ratón, la rata, el conejo y el hámster. La selección de las especies de mamífero que proporcionan las células no es una limitación de la presente invención, ni lo es el tipo de célula de mamífero, es decir, célula de fibroblasto, de hepatocito, tumoral, etc. Los requisitos para célula utilizada son que no porte ningún gen de adenovirus distinto de E1, E2a y/o E4 ORF6; no contenga ningún otro gen de virus que pueda dar como resultado recombinación homóloga de un virus contaminante durante la producción de rAAV; y, que sea capaz de infección o de transfección de ADN y de expresión del ADN transfectado. En una realización preferida, la célula anfitrión es una que tiene *rep* y *cap* transfectados establemente en la célula.

Una célula anfitrión útil en la presente invención es una célula anfitrión transformada establemente con las secuencias que codifican *rep* y *cap*, y que es infectada con el ADN de E1, E2a y E4 ORF6 de adenovirus, y una construcción que porta el minigén según se ha descrito con anterioridad. Líneas de células que expresan *rep* y *cap* de forma estable, tal como la B-50 (PCT/US98/19463), o las que han sido descritas en el documento de Patente U.S. núm. 5.658.785, pueden ser empleadas también de forma similar. Otra célula anfitrión deseable contiene el mínimo ADN adenoviral que sea suficiente para expresar E4 ORF6. Incluso pueden construirse otras líneas de células usando las secuencias *rep* de AAV9 y/o las secuencias *cap* de AAV9 de la invención.

La preparación de una célula anfitrión conforme a la invención incluye técnicas tales como el ensamblaje de secuencias de ADN seleccionadas. Este ensamblaje puede ser realizado utilizando técnicas convencionales. Tales técnicas incluyen clonación genómica y de ADN, las cuales son bien conocidas y han sido descritas en Sambrook et al., citado con anterioridad; el uso de secuencias solapantes de oligonucleótido de los genomas de adenovirus de AAV, combinadas con reacción de cadena de polimerasa, métodos sintéticos, y cualesquiera otros métodos adecuados que proporcionen la secuencia de nucleótido deseada.

La introducción de las moléculas (tal como plásmidos o virus) en la célula anfitrión puede ser realizada usando técnicas conocidas por el experto en la materia y que se discuten a través de la presente descripción. En una realización, se usan técnicas de transfección estándar, por ejemplo transfección de  $\text{CaPO}_4$  o electroporación, y/o infección por adenovirus híbrido/vectores de AAV en líneas de células tales como la línea de células de riñón embrionario humano HEK 293 (una línea de células de riñón humano que contiene genes E1 funcionales de adenovirus que proporciona proteínas E1 que actúan en *trans*).

Los vectores de terapia de gen basados en AAV9 que son generados por un experto en la materia, son beneficiosos para suministro de gen a células anfitrión seleccionadas y a pacientes de terapia de gen puesto que no se han encontrado anticuerpos de neutralización para el AAV9 en la población humana. Un experto en la materia puede

preparar fácilmente otros vectores virales de rAAV que contengan las proteínas de cápside de AAV9 proporcionadas en la presente memoria usando una diversidad de técnicas conocidas por los expertos en la materia. Se puede preparar también de forma similar otros vectores virales de rAAV que contengan secuencia de AAV9 y cápsides de AAV de otro serotipo.

- 5 De ese modo, un experto en la materia comprenderá fácilmente que las secuencias de AAV9 de la invención pueden ser adaptadas fácilmente para su uso en estos y otros sistemas de vector viral para suministro de gen *in vitro*, *ex vivo* o *in vivo*. De forma similar, un experto en la materia puede seleccionar fácilmente otros fragmentos del genoma de AAV9 de la invención para su uso en una diversidad de sistemas de vector de rAAV y no de AAV. Tales sistemas de vector pueden incluir, por ejemplo, lentivirus, retrovirus, poxvirus, virus de vaccinia, y sistemas adenovirales, entre otros. La selección de esos sistemas de vector no es una limitación de la presente invención.

- 10 Así, la invención proporciona además vectores generados usando las secuencias de ácido nucleico y de aminoácido del nuevo AAV de la invención. Tales vectores son útiles para una diversidad de propósitos, incluyendo para el suministro de moléculas terapéuticas y para su uso en regímenes de vacuna. Particularmente deseables para el suministro de moléculas terapéuticas son los AAV recombinantes que contienen cápsides del nuevo AAV de la invención. Estas, u otras construcciones de vector que contienen las nuevas secuencias de AAV de la invención, pueden ser usadas en regímenes de vacuna, por ejemplo, para el co-suministro de una citoquina, o para el suministro del propio inmunógeno.

#### V. Virus recombinantes y usos de los mismos

- 20 Usando las técnicas descritas en la presente memoria, un experto en la materia puede generar un rAAV que tenga una cápside de un serotipo 9 de la invención, o que tenga una cápside que contenga uno o más fragmentos de AAV9. En una realización, se puede utilizar una cápside de longitud completa procedente de un único serotipo, por ejemplo AAV9 [SEQ ID Núm. 2]. En otra realización, se puede generar una cápside de longitud completa que contenga uno o más fragmentos de AAV9 fusionados en una trama con secuencias procedentes de otro serotipo de AAV seleccionado, o procedente de porciones heterólogas de AAV9. Por ejemplo, un rAAV puede contener una o más de las nuevas secuencias de región hipervariable de AAV9. Alternativamente, las secuencias únicas de AAV9 de la invención pueden ser usadas en construcciones que contengan otras secuencias virales o no virales. Opcionalmente, un virus recombinante puede portar secuencias *rep* de AAV9 que codifiquen una o más de las proteínas *rep* de AAV9.

#### A. Suministro de virus

- 30 En otro aspecto, se pueden usar vectores conforme a la presente invención en un método para el suministro de un transgén a un anfitrión que incluye transfectar o infectar una célula anfitrión seleccionada con un vector viral recombinante generado con las secuencias de AAV9 (o con fragmentos funcionales de las mismas) de la invención. Los métodos para el suministro son bien conocidos por los expertos en la materia y no son una limitación de la presente invención.

- 35 En un modo de uso deseable, la invención proporciona un método para el suministro mediado por AAV9 de un transgén a un anfitrión. Este método incluye transfectar o infectar una célula anfitrión seleccionada con un vector viral recombinante que contiene un transgén seleccionado bajo el control de secuencias que dirigen la expresión del mismo y proteínas de cápside de AAV9.

- 40 Opcionalmente, una muestra del anfitrión puede ser previamente sometida a ensayo en cuanto a la presencia de anticuerpos en un serotipo de AAV seleccionado. Una diversidad de formatos de ensayo para la detección de anticuerpos neutralizantes son bien conocidos por los expertos en la materia. La selección de dicho ensayo no es una limitación de la presente invención. Véase, por ejemplo, Fisher et al., *Nature Med.*, 3(3): 306-312 (Marzo de 1997) y W.C. Manning et al., *Terapia de Gen Humana*, 9: 477-485 (1 de Marzo de 1998). Los resultados del ensayo pueden ser usados para determinar qué vector de AAV que contiene proteínas de cápside de un serotipo particular es el preferido para el suministro, por ejemplo por la ausencia de anticuerpos neutralizantes específicos para el serotipo de cápside.

- 45 En un aspecto de este método, el suministro de vector con proteínas de cápside de AAV9 puede preceder o seguir al suministro de un gen por medio de un vector con una proteína de cápside de AAV de serotipo diferente. Así, el suministro de gen por medio de vectores de rAAV puede ser usado para repetir el suministro de gen a una célula anfitrión seleccionada. Deseablemente, los vectores de rAAV suministrados posteriormente portan el mismo transgén que el primer vector de rAAV, pero los vectores administrados posteriormente contienen proteínas de cápside de serotipos que difieren del primer vector. Por ejemplo, si un primer vector tiene proteínas de cápside de AAV9, los vectores administrados posteriormente pueden tener proteínas de cápside seleccionadas entre los otros serotipos, incluyendo sin limitación los AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7 y AAV8.

- 55 Los vectores recombinantes descritos anteriormente pueden ser suministrados a células anfitrión de acuerdo con métodos publicados. El rAAV, preferentemente suspendido en un portador farmacéuticamente compatible, puede ser administrado a un paciente mamífero humano o no humano. Los portadores adecuados pueden ser seleccionados

fácilmente por un experto en la materia en vista de la indicación a la que está dirigido el virus transferido. Por ejemplo, un portador adecuado incluye solución salina, la cual puede ser formulada con una diversidad de soluciones tampón (por ejemplo, solución salina tamponada de fosfato). Otros ejemplos de portadores incluyen solución salina estéril, lactosa, sacarosa, fosfato de calcio, gelatina, dextrano, agar, pectina, aceite de cacahuete, y agua. La selección del portador no es una limitación de la presente invención.

Opcionalmente, las composiciones de la invención pueden contener, además del rAAV y del (de los) portador(es), otros ingredientes farmacéuticos convencionales, tal como conservantes o estabilizadores químicos. Ejemplos de conservantes adecuados incluyen clorobutanol, sorbato de potasio, ácido sórbico, dióxido de azufre, propil galato, los parabenos, etil vanilina, glicerina, fenol, y paraclorofenol. Los estabilizadores químicos adecuados incluyen gelatina y albúmina.

Los vectores virales son administrados en cantidades suficientes para transfectar las células y proporcionar niveles suficientes de transferencia de gen y expresión para proporcionar un beneficio terapéutico sin efectos adversos indebidos, o con efectos fisiológicos médicamente aceptables, los cuales pueden ser determinados por los expertos en los temas médicos. Las rutas de administración convencionales y farmacéuticamente aceptables incluyen, sin limitación, el suministro directo al hígado o pulmón, oralmente, intranasalmente, intratraquealmente, por inhalación, intravenosamente, intramuscularmente, intraocularmente, subcutáneamente, intradérmicamente, o mediante otras rutas de administración. Se pueden combinar las rutas de administración, si se desea.

Las dosis del vector viral dependerán principalmente de factores tales como la condición que va a ser tratada, la edad, el peso y el estado de salud del paciente, y pueden por tanto variar entre pacientes. Por ejemplo, una dosis humana terapéuticamente efectiva del vector viral está comprendida generalmente en la gama de desde aproximadamente 1 ml a aproximadamente 100 ml de solución conteniendo concentraciones de desde aproximadamente  $1 \times 10^9$  a  $1 \times 10^{16}$  genomas de vector del virus. Una dosis humana preferida puede ser de aproximadamente  $1 \times 10^{13}$  a  $1 \times 10^{16}$  genomas de AAV. La dosis será ajustada para equilibrar el beneficio terapéutico frente a cualesquiera efectos colaterales y tales dosis pueden variar dependiendo de la aplicación terapéutica para que se emplee el vector recombinante. Los niveles de expresión del transgén pueden ser monitorizados a efectos de determinar la frecuencia de dosificación resultante en vectores virales, con preferencia de AAV que contienen el minigén. Opcionalmente, se pueden utilizar regímenes de dosificación similares a los descritos para fines terapéuticos para inmunización usando las composiciones de la invención.

Ejemplos de productos terapéuticos y de productos inmunogénicos para su suministro mediante los vectores que contienen AAV9 según la invención, se proporcionan en lo que sigue. Estos vectores pueden ser usados para una diversidad de regímenes terapéuticos o de vacuna, según se describe en la presente memoria. Adicionalmente, esos vectores pueden ser suministrados en combinación con uno o más de otros vectores o ingredientes activos en un régimen terapéutico y/o de vacuna deseado.

## B. Transgenes terapéuticos

Los productos terapéuticos útiles codificados por el transgén incluyen hormonas y factores de crecimiento y diferenciación, incluyendo sin limitación insulina, glucagón, hormona de crecimiento (GH), hormona paratiroidea (PTH), factor de liberación de hormona de crecimiento (GRF), hormona de estimulación de folículo (FSH), hormona leutinizante (LH), gonadotropina coriónica humana (hCG), factor de crecimiento endotelial vascular (VEGF), angiopoyetinas, angiostatina, factor de estimulación de colonia de granulocito (GCSF), eritropoyetina (EPO), factor de crecimiento de tejido conectivo (CTGF), factor de crecimiento de fibroblasto básico (bFGF), factor de crecimiento de fibroblasto ácido (aFGF), factor de crecimiento epidérmico (EGF), factor de crecimiento derivado de plaqueta (PDGF), factores I y II de crecimiento de insulina (IGF-I e IGF-II), uno cualquiera de la superfamilia de factor de crecimiento  $\alpha$  de transformación, activinas, inhibinas, o cualquiera de las proteínas morfogénicas óseas (BMP) BMPs 1-15, uno cualquiera de la familia de factores de crecimiento de factor de diferenciación de herogluina/neurogluina, ARIA/neu (NDF), factor de crecimiento nérveo (NGF), factor neurotrópico derivado del cerebro (BDNF), neurotrofinas NT-3 y NT4/5, factor neurotrópico ciliar (CNTF), factor neurotrópico derivado de línea de célula glial (GDNF), neurturina, agrina, uno cualquiera de la familia de semaforinas/colapsinas, netrina-1, netrina-2, factor de crecimiento de hepatocito (HGF), efrinas, noggin, erizo sónico y tirosina hidroxilasa.

Otros productos de transgén útiles incluyen los que regulan el sistema inmune incluyendo, sin limitación, citoquinas y linfoquinas tales como trombopoyetina (TPO), interleuquinas (IL) IL-1 a IL-25 (incluyendo, por ejemplo, IL-2, IL-4, IL-6, IL-12 e IL-18), proteína quimioatrayente de monocito, factor inhibitorio de leucemia, factor de estimulación de colonia de granulocito-macrófago, ligando Fas, factores  $\alpha$  y  $\beta$  de necrosis de tumor, factor de células madre, ligando flk-2/flt3. Los productos de gen producidos por el sistema inmune son también útiles en la invención. Estos incluyen, sin limitaciones, las inmunoglobulinas IgG, IgM, IgA, IgD e IgE, inmunoglobulinas quiméricas, anticuerpos humanizados, anticuerpos de cadena simple, receptores de célula T, receptores de célula T quiméricos, receptores de célula T de cadena simple, moléculas MHC de clase I y de clase II, así como inmunoglobulinas de diseño y moléculas MHC. Los productos de gen útiles incluyen también las proteínas reguladoras de complemento tal como las proteínas reguladoras de complemento, proteína de cofactor de membrana (MCP), factor acelerador de pudrición (DAF), CR1, CF2 y CD59.

Otros productos de gen útiles adicionales incluyen uno cualquiera de los receptores para las hormonas, factores de crecimiento, citoquinas, linfoquinas, proteínas reguladoras y proteínas de sistema inmune. La invención abarca receptores para regulación del colesterol, incluyendo el receptor de lipoproteína de baja densidad (LDL), el receptor de lipoproteína de alta densidad (HDL), el receptor de lipoproteína de muy baja densidad (VLDL), y el receptor scavenger. La invención abarca también productos de gen tales como los miembros de la superfamilia del receptor de la hormona esteroide incluyendo receptores de glucocorticoide y receptores de estrógeno, receptores de Vitamina D y otros receptores nucleares. Adicionalmente, los productos de gen útiles incluyen factores de transcripción tales como *jun*, *fos*, *max*, *mad*, factor de respuesta de suero (SRF), AP-1, AP-2, *myb*, MyoD y miogenina, proteínas que contienen ETS-box, TFE3, E2F, ATF1, ATF2, ATF3, ATF4, ZF5, NFAT, CREB, HNF-4, C/EBP, SP1, proteínas de enlace de CCAAT-box, factor de regulación de interferón (IRF-1), proteína del tumor de Wilms, proteína de enlace de ETS, STAT, proteínas de enlace de GATA-box, por ejemplo GATA-3, y la familia forkhead de proteínas de hélice alada.

Otros productos de gen útiles incluyen carbamoil sintetasa I, ornitina transcarbamilasa, arginosuccinato sintetasa, arginosuccinato liasa, arginasa, fumarilacetato hidrolasa, fenilalanina hidroxilasa, alfa-1-antitripsina, glucosa-6-fosfatasa, profobilínógeno, desaminasa, factor VIII, factor IX, cistationa betasintasa, cetoácido descarboxilasa de cadena ramificada, albúmina, isovaleril-coA deshidrogenasa, propionil CoA carboxilasa, metil malonil CoA mutasa, glutaril CoA deshidrogenasa, insulina, beta-glucosidasa, piruvato carboxilato, fosforilasa hepática, fosforilasa quinasa, glicina descarboxilasa, proteína H, proteína T, una secuencia reguladora de transmembrana de fibrosis cística (CFTR), y una secuencia de cADN de distrofina. Incluso otros productos de gen útiles incluyen enzimas siempre que puedan ser útiles en terapias de sustitución de enzima, la cual es útil en una diversidad de condiciones resultantes de una actividad deficiente de la enzima. Por ejemplo, las enzimas que contienen manosa-6-fosfato pueden ser utilizadas en terapias para enfermedades de acumulación liposómica (por ejemplo, un gen adecuado incluye el que codifica la  $\beta$ -glucuronidasa (GUSB)).

Otros productos de gen útiles incluyen los polipéptidos que ocurren de forma no natural, tal como los polipéptidos quiméricos o híbridos que tienen una secuencia de aminoácido que ocurre de forma no natural, que contiene inserciones, supresiones o sustituciones de aminoácido. Por ejemplo, las inmunoglobulinas de diseño de cadena simple podrían ser útiles en determinados pacientes inmunocomprometidos. Otros tipos de secuencias de gen que ocurren de forma no natural incluyen moléculas antisentido y ácidos nucleicos catalíticos, tal como ribozimas, que podrían ser usados para reducir la sobreexpresión de un objetivo.

La reducción y/o modulación de expresión de un gen resulta particularmente deseable para el tratamiento condiciones hiperproliferativas caracterizadas por células hiperproliferantes, tal como el cáncer y la psoriasis. Los polipéptidos objetivo incluye los polipéptidos que se produce de manera exclusiva o a niveles más altos en células hiperproliferantes en comparación con las células normales. Los antígenos objetivo incluyen polipéptidos codificados por oncogenes tales como *myb*, *myc*, *fyn*, y el gen de translocación *bcr/abl*, *ras*, *src*, P53, *neu*, *trk* y EGFR. Adicionalmente a los productos de oncogén como antígenos objetivo, los polipéptidos objetivo para tratamientos anti-cáncer y regímenes protectores incluyen regiones variables de anticuerpos formados por linfomas de célula B y regiones variables de receptores de célula T o de linfomas de célula T que, en algunas realizaciones, se usan también como antígenos objetivo para enfermedad autoinmune. Otros polipéptidos asociados a tumor pueden ser usados como polipéptidos objetivo tal como polipéptidos que se encuentran a niveles más altos en células de tumor incluyendo el polipéptido reconocido por el anticuerpo monoclonal 17-1A y polipéptidos de enlace de folato.

Otros polipéptidos y proteínas terapéuticos adecuados incluyen los que pueden ser útiles para el tratamiento de individuos que adolecen de desórdenes y enfermedades autoinmunes, confiriendo una amplia respuesta inmune protectora frente a objetivos que se asocian con la autoinmunidad incluyendo los receptores de células y las células que producen anticuerpos "auto"-dirigidos. Las enfermedades autoinmunes mediadas por células T incluyen la artritis reumatoide (RA), esclerosis múltiple (MS), síndrome de Sjörger, sarcoidosis, diabetes mellitus dependiente de insulina (IDDM), tiroiditis autoinmune, artritis reactiva, espondilitis anquilosante, escleroderma, polimiositis, dermatomiositis, psoriasis, vasculitis, granulomatosos de Wegener, enfermedad de Crohn y colitis ulcerosa. Cada una de esas enfermedades se caracteriza por receptores de célula T (TCRs) que enlazan con antígenos endógenos e inician la cascada inflamatoria asociada a enfermedades autoinmunes.

### 50 C. Transgenes inmunogénicos

Alternativamente, o adicionalmente, los vectores de la invención pueden contener secuencias de AAV9 de la invención y un transgén que codifica un péptido, polipéptido o proteína que induce una respuesta inmune a un inmunógeno seleccionado. Por ejemplo, los inmunógenos pueden ser seleccionados a partir de una diversidad de familias virales. Ejemplos de familias virales deseables frente a las que pudiera ser deseable una respuesta inmune incluyen la familia picomavirus, la cual incluye los géneros rinovirus que son responsables de aproximadamente el 50% de los casos de resfriado común; los géneros enterovirus, los cuales incluyen poliovirus, coxsackievirus, ecovirus, y enterovirus humanos tal como el virus de la hepatitis A; y los géneros aptovirus, los cuales son responsables de enfermedades del pie y la boca, principalmente en animales no humanos. Dentro de la familia de virus picomavirus, los antígenos objetivo incluyen los VP1, VP2, VP3, VP4 y VPG. Otra familia viral incluye la familia calcivirus, la cual abarca el grupo de virus Norwalk, los cuales son importante agente causante de gastroenteritis

epidémica. Aún otra familia viral deseable para su uso en antígenos objetivo para inducir respuestas inmunes en humanos y animales no humanos, es la familia togavirus, la cual incluye los géneros alfavirus, la cual incluye los virus de Sindbis, el virus de Ross River, y la encefalitis Venezolana, Eastern & Western Equina, y rubivirus, incluyendo el virus Rubella. La familia flaviviridae incluye los virus del dengue, la fiebre amarilla, la encefalitis Japonesa, la encefalitis de St. Louis y de la encefalitis propagada por garrapatas. Otros antígenos objetivo pueden ser generados a partir de la hepatitis C o de la familia de coronavirus, la cual incluye un número de virus no humanos tal como el virus de la bronquitis infecciosa (aves de corral), el virus gastroentérico transmisible porcino (cerdo), el virus de la encefalomielitis hemaglutinante porcino (cerdo), el virus de la peritonitis infecciosa felina (gatos), el coronavirus entérico felino (gato), el coronavirus canino (perro), y coronavirus respiratorios humanos, los cuales pueden causar resfriado común y/o hepatitis no A, B o C. Dentro de la familia coronavirus, los antígenos objetivo incluyen la glicoproteína (no presente en todos los coronavirus) E1 (también denominada proteína M o matriz), E2 (también denominada proteína S o de Pike), E3 (también denominada HE o hemaglutina-elterosa), o N (nucleocápside). Otros antígenos adicionales pueden estar objetivos contra la familia rhabdovirus, la cual incluye los géneros vesiculovirus (por ejemplo, el Virus de la Estomatitis Vesicular), y el lisavirus general (por ejemplo, rabias). Dentro de la familia rhabdovirus, los antígenos adecuados pueden derivarse de la proteína G o de la proteína N. La familia flavoviridae, la cual incluye los virus de la fiebre hemorrágica tal como el virus de Marburg y del Ébola, puede ser una fuente de antígenos adecuada. La familia paramixovirus incluye el virus de parainfluenza de Tipo 1, el virus de parainfluenza Tipo 3, el virus de parainfluenza bovina Tipo 3, los rubaluvirus (virus de las paperas, virus de parainfluenza de Tipo 2, virus de parainfluenza de Tipo 4, virus de la enfermedad de Newcastle (pollos), rinderpest, morbilivirus, que incluyen el sarampión y el moquillo canino, y neumovirus, que incluyen el virus sincitial respiratorio. El virus de la influenza está clasificado dentro de la familia ortomixovirus y es una fuente de antígeno adecuada (por ejemplo, proteína HA, la proteína N1). La familia bunyavirus incluye los géneros bunyavirus (encefalitis de California, La Cross), fletovirus (Fiebre del Valle del Rift), hantavirus (el puremala es un virus de la fiebre de hemahagina), nairovirus (enfermedad de la oveja de Nairobi), y varios bungavirus sin asignación. La familia arenavirus proporciona una fuente de antígenos frente a LCM y al virus de la fiebre de Lassa. La familia reovirus incluye los géneros reovirus, rotavirus (que causa gastroenteritis aguda en niños), orbivirus, y cultivirus (virus de la garrapata de Colorado, Lebonbo (humanos), encefalosis equina, lengua azul). La familia retrovirus incluye la subfamilia oncovirinal que abarca enfermedades humanas y veterinarias tales como el virus de la leucemia felina, HTLV1 y HTLVII, lentivirus (que incluye el VIH, el virus de inmunodeficiencia de simios, el virus de inmunodeficiencia felina, el virus de anemia infecciosa equina, y espumavirinal). La familia papovavirus incluye la subfamilia de polioma virus (virus BKU y JCU) y la subfamilia papilomavirus (asociados con cánceres o progresión maligna de papiloma). La familia adenovirus incluye los virus (EX, AD7, ARD, O.B.) que causan enfermedad respiratoria y/o enteritis. El papovirus felino de la familia de los parvovirus (enteritis felina), el panleucovirus felino, el parvovirus canino, y el parvovirus porcino. La familia herpesvirus incluye la subfamilia alfa herpesvirus, la cual abarca los géneros simplexvirus (HSV1, HSVII), varicelovirus (pseudo-rabias, varicela zóster) y la subfamilia beta herpesvirinae, la cual incluye los géneros citomegalovirus (HCMV, muromegalovirus) y la subfamilia gamma herpesvirinae, la cual incluye los géneros linfocriptovirus, EBV (linfoma de Burkitts), rinotraqueitis infecciosa, virus de la enfermedad de Mark, y radinivirus. La familia poxvirus incluye la subfamilia cordopoxvirinae, la cual abarca los géneros ortopoxvirus (Variola major (Viruela) y Vaccinia (viruela bovina), parapoxvirus, avipoxvirus, capripoxvirus, leporipoxvirus, suipoxvirus, y la familia entomopoxvirinae. La familia hepadnavirus incluye el virus de la hepatitis B. Un virus sin clasificar que puede ser una fuente de antígeno adecuada es el virus de la hepatitis delta. Otro virus que es una fuente de antígeno es el Virus Nipán. Incluso otras fuentes virales pueden incluir el virus de la enfermedad bursal infecciosa aviar y el virus de síndrome respiratorio y reproductivo porcino. La familia alfavirus incluye el virus de la artritis y varios virus de encefalitis.

La presente invención puede abarcar también inmunógenos que sean útiles para inmunizar un humano o un animal no humano frente a otros patógenos incluyendo las bacterias, hongos, microorganismos parásitos o parásitos intracelulares que infectan a los vertebrados humanos y no humanos, o procedentes de una célula cancerígena o una célula tumoral. Ejemplos de patógenos bacterianos incluyen los cocos gram-positivos patogénicos, incluyendo los neumococos; estafilococos (y las toxinas producidas por los mismos, por ejemplo la enterotoxina B); y estreptococos. Los cocos gram-negativos patogénicos incluyen los meningococos; gonococos. Los bacilos gram-negativos entéricos patogénicos incluyen las enterobacteriaceae; pseudomonas, acinetobacteria y eikenella; melioidosis; salmonella; shingella; hemófilos; moraxella; H. ducreyi (que causa cancroide); especies de brucella (brucelosis); *Francisella tularensis* (que causa tularemia); *Yersinia pestis* (plaga) y otra yersinia (pasteurella); *Streptobacillus moniliformis* y *Spirillum*; los bacilos gram-positivos incluyen listeria monocytogenes; *Erysipelothrix rhusiopathiae*; *Corynebacterium diphtheria* (difteria); cólera; *B anthracis* (ántrax); donovanosis (granuloma inguinal), y bartonellosis. Las enfermedades causadas por bacterias anaeróbicas patogénicas incluyen el tétanos; botulismo (*Clostridium botulinum* y su toxina); *Clostridium perfringens* y su toxina épsilon; otra clostridia; tuberculosis; lepra, y otras micobacterias. Las enfermedades espiroquetales patogénicas incluyen la sífilis; trepanomatosis; sífilis pian, pinta y endémica, y leptoespirosis. Otras infecciones causadas por bacterias patógenas más altas y hongos patogénicos incluyen el muermo (*Burkholderia mallei*); actinomycosis; nocardiosis; criptococosis, blastomycosis, histoplasmosis y coccidioidomycosis; candidiasis, aspergillosis, y mucomicosis; esporotricosis; paracoccidioidomycosis, petriellidiosis, torulopsosis, micetoma y cromomicosis; y dermatofytosis. Las infecciones rickettsianas incluyen la fiebre del tifus, la fiebre de las Montañas Rocosas, la fiebre Q (*Coxiella burnetii*), y Rickettsialpox. Ejemplos de infecciones clamidiales y de micoplasma incluyen: *Mycoplasma pneumoniae*;

lymphogranuloma venereum; psitacosis; e infecciones clamidiales perinatales. Las eucariotas patogénicas abarcan protozoos patogénicos y helmintos y las infecciones producidas por los mismos incluyen: amebiasis; malaria; leishmaniasis; tripanosomiasis; toxoplasmosis; *pneumocystis carinii*; *trichans*; *toxoplasma gondii*; babesiosis; giardiasis; triquinosis; filariasis; esquistosomiasis; nematodos; trematodos o fasciolas; e infecciones por cestodos (tenia).

Muchos de esos organismos y/o de las toxinas producidas por los mismos han sido identificados por los Centros de Control de Enfermedades [(CDC), Departamento de Salud y Servicios Humanos, USA], como agentes que tienen potencial para su uso en ataques biológicos. Por ejemplo, algunos de esos agentes biológicos incluyen: *Bacillus anthracis* (ántrax), *clostridium botulinum* y su toxina (botulismo), *yersinia pestis* (plaga), variola mayor (viruela), *francisella tularensis* (tularemia), y fiebres hemorrágicas virales [flavivirus (por ejemplo, Ébola, Marburg), y arenavirus [por ejemplo, Lassa, Machupol], todos los cuales están actualmente clasificados como agentes de Categoría A; *coxiella burnetti* (fiebre Q); especies de brucella (brucelosis), *burkholderia mallei* (muermo), *burkholderia pseudomallei* (meloidosis), *ricinus communis* y su toxina (toxina ricina), *clostridium perfringens* y su toxina (toxina épsilon), especies de *staphylococcus* y sus toxinas (enterotoxina B), *clamydia psittaci* (psitacosis), ataques contra la seguridad del agua (por ejemplo, *Vibrio cholera*, *cryptosporidium parvum*), fiebre del tifus (*rickettsia powazeli*), y encefalitis viral (alfavirus, por ejemplo, encefalitis equina Venezolana; encefalitis equina Eastern; encefalitis equina western), todos los cuales están actualmente clasificados como agentes de Categoría B; y el virus Nipán y hantavirus, los cuales están actualmente clasificados como agentes de categoría C. Adicionalmente, otros organismos que están clasificados de ese modo o de forma diferentes, pueden ser identificados y/o usados para tales propósitos en el futuro. Se comprenderá fácilmente que los vectores virales y otras construcciones que se describen en la presente memoria, son útiles para suministrar antígenos a partir de esos organismos, virus, sus toxinas u otros subproductos, que evitarán y/o tratarán la infección u otras reacciones adversas con esos agentes biológicos.

La administración de los vectores de la invención para suministrar inmunógenos frente a la región variable de las células T provoca una respuesta inmune que incluye CTLs para eliminar esas células T. En artritis reumatoide (RA), se han caracterizado varias regiones variables específicas de receptores de células T (TCRs) que están involucradas en la enfermedad. Esas TCRs incluyen las V-3, V-14, V-17 y V-17. De ese modo, el suministro de una secuencia de ácido nucleico que codifica al menos uno de esos polipéptidos provocará una respuesta inmune que objetivará células T en RA. En esclerosis múltiple (MS), han sido caracterizadas varias regiones variables específicas de TCRs que están involucradas en la enfermedad. Esas TCRs incluyen las V-7 y V-10. De ese modo, el suministro de una secuencia de ácido nucleico que codifica al menos uno de esos polipéptidos provocará una respuesta inmune que objetivará células T involucradas en MS. En escleroderma, varias regiones variables específicas de TCRs que están involucradas en la enfermedad, han sido caracterizadas. Esas TCRs incluyen las V-6, V-8, V-14 y V-16, V-3C, V-7, V-14, V-15, V-16, V-28 y V-12. Así, el suministro de una molécula de ácido nucleico que codifica al menos uno de esos polipéptidos provocará una respuesta inmune que objetivará células T involucradas en escleroderma.

De ese modo, un vector viral recombinante derivado de rAAV9 según la invención proporciona un vehículo eficiente de transferencia de gen que puede suministrar un transgén seleccionado a una célula anfitrión seleccionada *in vivo* o *ex vivo* incluso donde el organismo tenga anticuerpos neutralizantes para uno o más serotipos de AAV. En una realización, el rAAV y las células se mezclan *ex vivo*; las células infectadas se cultivan usando metodologías convencionales; y, las células transfectadas son re-infundidas en el paciente.

Los vectores de la invención son particularmente adecuados para suministro de gen con fines terapéuticos y de inmunización, incluyendo la inducción de inmunidad protectora. Además, los vectores de la invención pueden ser también usados para la producción de un producto de gen deseado *in vitro*. Para la producción *in vitro*, un producto deseado (por ejemplo, una proteína) puede ser obtenido a partir de un cultivo deseado a continuación de una transfección de células anfitrión con un rAAV que contiene la molécula que codifica el producto deseado, y cultivando el cultivo celular bajo condiciones que permitan la expresión. El producto expresado pueden ser purificado y aislado a continuación, según se desee. Las técnicas adecuadas para transfección, cultivo celular, purificación y aislamiento, son bien conocidas por los expertos en la materia.

Los ejemplos que siguen ilustran varios aspectos y realizaciones de la invención.

## Ejemplos

### Ejemplo 1: Producción de genoma viral de AAV9 recombinante equipado con ITRs de AAV2

Las construcciones de empaquetamiento quiméricas son generadas mediante fusión de secuencias rep de AAV2 con secuencias cap de los nuevos serotipos de AAV. Estas construcciones de empaquetamiento quiméricas se usan, en principio, para seudotipar genomas de AAV recombinante que portan ITRs de AAV2 mediante triple transfección en células 293 usando plásmido auxiliar de Ad5. Esos vectores seudotipados se usan para evaluar en comportamiento en estudios serológicos basados en transducción y para evaluar eficiencia de transferencia de gen de nuevos serotipos de AAV en diferentes modelos animales incluyendo NHP y roedores, antes de que virus intactos e infecciosos de esos nuevos serotipos sean aislados.

**A. pAAV2GFP**

Es el plásmido de AAV2 que contiene las ITRs de AAV2 y la proteína fluorescente verde expresada bajo el control de un promotor constitutivo. Este plásmido contiene los siguientes elementos: las ITRs de AAV2, un promotor de CMV, las secuencias de codificación de GFP.

**5 B. Clonación de plásmido trans**

Para construir el plásmido trans quimérico para la producción de vectores de AAV9 seudotipados recombinantes, el plásmido p5E18 (Xiac et al., J. Virol. 73: 3994-4003) fue digerido parcialmente con Xho I para linealizar el plásmido en el sitio Xho I en la posición de 3169 bp solamente. Los extremos de corte del Xho I fueron rellenados y enlazados de nuevo. Este plásmido p5E18 modificado fue limitado con Xba I y Xhb I en una digestión completa para eliminar la secuencia de gen cap de AAV2 y reemplazarla con un fragmento Spe I/Xho I de 2267 bp que contiene el gen cap de AAV9 que fue aislado a partir del plásmido pCRAAV9- 6-5+15-4.

El plásmido resultante contiene las secuencias rep de AAV2 para Rep78/68 bajo el control del promotor de AAV2P5, y las secuencias rep de AAV2 para Rep52/40 bajo el control del promotor P19 de AAV2. Las secuencias de cápside de AAV9 están bajo el control del promotor P40 de AAV2, el cual está situado dentro de las secuencias rep. El plásmido contiene además un espaciador 5' de la rep ORF.

Alternativamente, se puede construir un plásmido similar que utilice las secuencias rep de AAV9 y las secuencias de promotor nativo de AAV9. Este plásmido se usa después para la producción de rAAV9, según se ha descrito en la presente memoria.

**C. Producción de rAAV seudotipado**

Las partículas de rAAV (vector de AAV2 en cápside de AAV9) son generadas usando un método libre de adenovirus. De forma resumida, el plásmido cis (plásmido pAAV2.1 lacZ que contiene ITRs de AAV2), y el plásmido trans pCR!!V9 6-5+15-4 (que contiene la rep de AAV2 y la cap de AAV9) y un plásmido auxiliar, respectivamente, son co-transfectados simultáneamente en células 293 en una relación de 1:1:2 mediante precipitación de fosfato de calcio.

Para la construcción del plásmido auxiliar pAD, se adquirió plásmido pBG10 en Microbix (Canadá). Un fragmento RsrII que contiene L2 y L3, fue suprimido del pBHG10, dando como resultado el primer plásmido auxiliar, el pAd F13. El plásmido Ad F1 fue construido mediante clonación del fragmento Asp700/SalI con una supresión Pmel/Sgf1, y aislamiento del pBHG10 en Bluescript. Se suprimió MLP, L2, L2 y L3 en el pAd F1. Supresiones adicionales de un fragmento de NruI de 2,3 kb y, posteriormente, un fragmento de RsrI/NruI de 0,5 kb generaron los plásmidos auxiliares pAd F5 y pAd F6, respectivamente. El plásmido auxiliar, denominado pF6, proporciona las funciones auxiliares esenciales de E2a y E4 ORF6 no proporcionadas por la célula auxiliar de expresión de E1, pero está exento de proteínas de cápside adenovirales y de regiones funcionales E1).

Típicamente, 50 µg de ADN (cis:trans:auxiliar) son transfectadas sobre un plato de cultivo de tejido de 150 mm. Las células 293 son cultivadas durante 72 horas post-transfección, sonicadas y tratadas con desoxicolato de sodio al 0,5% (37 °C durante 10 min.). Los lisatos de las células son sometidos a continuación a dos rondas de un gradiente de CsCl. Las fracciones de pico que contienen el vector rAAV, son recogidas, agrupadas y dializadas frente a PBS.

**Ejemplo 2 – Modelo de ratón de hipercolesterolemia familiar**

El experimento que sigue demostrará que el rAAV de AAV2/9 construido según se ha descrito en la presente memoria, suministra el receptor de LDL y expresa el receptor de LDL en una cantidad suficiente para reducir los niveles de colesterol del plasma y de triglicéridos en modelos animales de hipercolesterolemia familiar.

Las construcciones de rAAV2/9 usadas en el experimento que sigue contienen el cADN para el receptor VLDL de murina (LDLR) inducido a partir de un promotor de β-actina y potenciado a partir del potenciador de citomegalovirus (AAV-CB-VLDLR). Los ratones deficientes en receptor de LDL (Laboratorios Jackson) en una dieta alta en colesterol que presentan los síntomas de hipercolesterolemia familiar (ratones de FH), fueron infundidos intravenosamente con  $4 \times 10^{12}$  AAV-CB-LDLR o con un control salino y se monitorizó el nivel de colesterol del plasma.

Los perfiles de colesterol y triglicéridos del plasma de ratones FH infundidos con virus de VLDLR, fueron determinados como sigue. Los ratones tratados según se ha descrito anteriormente se mantuvieron en ayunas durante seis (6) horas, y se recogieron muestras de sangre mediante punción de plexus venoso retro-orbital con tubos capilares heparinizados. El plasma se separó por centrifugación. La concentración total de colesterol y triglicéridos se determinó usando kits adquiridos en Wako-Chemicals.

Los ratones del grupo de control muestran niveles de control gradualmente crecientes hasta alcanzar un estado estable en aproximadamente 4 semanas post-inyección. En ratones que recibieron un virus de LDR, los niveles de colesterol del plasma se redujeron sostenidamente a partir de 3 semanas post-administración del virus.

Se obtuvo un perfil de colesterol del plasma a partir de ratones FH infundidos con virus de LDLR mediante análisis

FPLC. Se agrupó el plasma de 4 ratones en cada grupo. Las muestras de plasma se aislaron del control y los animales fueron infundidos con AAV-VLDRL en preinyección (PI) o 2 meses (2M) de post-infusión del virus y se analizaron mediante fraccionamiento FPLC seguido de ensayo de colesterol.

### Ejemplo 3 – Transducción *in vivo* con vectores de serotipo de AAV9

5 Se evaluó el comportamiento del vector basado en el nuevo serotipo de AAV9 en modelos murina de transferencia de gen dirigida al músculo y al hígado, y se comparó con vectores basados en los serotipos AAV1, AAV2 y AAV5 conocidos. Los vectores que expresan proteínas encubiertas son utilizados para cuantificar eficacias de transducción relativa entre diferentes serotipos mediante análisis ELISA de los sueros. Se evaluó la distribución celular de transducción dentro del órgano objetivo usando vectores de expresión de lacZ e histoquímica X-gal.

10 Para este experimento, genomas de AAV recombinante, los AAV2CBhA1AT, AAV2AlbA1AT, AAV2CMVrhCG, AAV2TBGrhCG, AAV2TBGcFIX, AAV2CMVLacZ y AAV2TBGLacZ, se empaquetaron con proteínas de cápside de AAV9, o de AAV1, AAV2 o AAV5. En todas las construcciones, las casetas de minigén están flanqueadas con ITRs de AAV2. Se usaron genes de cADNs de anti-tripsina humana (A1AT) [Xiado, W., et al., (1999) J. Virol. 73, 3994-4003], subunidad-β de hormona coriogonadotrópica (CG) de mono Rhesus [Zoltick, P. W. & Wilson, J. M. (2000), Mol. Ther. 2, 657-9], factor canino IX [Wang, L., et al., (1997), Proc. Natl. Acad. Sic. USA 94, 11563-6] y β-galactosidasa bacteriana (es decir, LacZ) como genes informadores. Para transferencia de gen dirigida al hígado, se usó ya sea el promotor de gen de albúmina de ratón (Alb) [Xiao, W. (1999), citado anteriormente] o ya sea el promotor de globulina de enlace de hormona tiroidea humana (TBG) [Wang (1997), citado anteriormente] para activar expresión específica en el hígado de genes informadores. En experimentos de transferencia de gen dirigida al músculo, se empleó ya sea promotor prematuro de citomegalovirus (CMV) o ya sea promotor de β-actina de pollo con potenciador CMV (CB) para dirigir la expresión de informadores.

Para transferencia de gen dirigida al músculo, los vectores se inyectan en el tibial anterior derecho de ratones desnudos NCR o C57BL/6 de 4-6 semanas de edad (Taconic, Germantown, NY). En estudios de transferencia de gen dirigida al hígado, los vectores son infundidos intraportalmente en ratones NCR desnudos o C57BL/6 de 7-9 semanas de edad [Taconic, Germantown, NY]. Se recogieron muestras de suero intraorbitalmente en diferentes momentos tras la administración del vector. Los tejidos del músculo y del hígado fueron cultivados en diferentes momentos en cuanto a crioseccionamiento y manchado histoquímico Xgal a partir de animales que recibieron los vectores lacZ. Para el experimento de readministración, ratones C56BL/6 recibieron inicialmente vectores AAV2/1, 2/2, 2/5 y 2/9 vectores CBA1AT intramuscularmente y siguió expresión de gen A1AT durante 7 semanas. Los estudios previos indicaron que los ratones C57BL/6 inmunocompetentes provocaron respuestas humorales limitadas a la proteína A1AT humana cuando se expresan a partir de vectores de AAV [Xiao, “., et al., (1999), J. Virol. 73, 3994-4003]. Los animales fueron tratados a continuación con AAV2/9 TBGcFIX intraportalmente y estudiados en cuanto a expresión de gen cFIX.

35 Se realizaron ensayos basados en ELISA para cuantificar los niveles en el suero de proteínas hA1AT, rhCG y cFIX según se ha descrito anteriormente [Gao, G.P. et al., (1996), J. Virol. 70, 8934-43; Zoltlick. P.W. & Wilson, J.M. (2000), Mol. Ther. 2, 647-9; Wang, L., et al., Proc. Natl. Acad. Sci. USA 94, 11563-6]. Los experimentos estuvieron completos cuando los animales fueron sacrificados para el cultivo de tejidos del músculo y del hígado en cuanto a extracción de ADN y análisis cuantitativo de copias de genoma de vectores presentes en tejidos objetivo mediante TaqMan usando el mismo conjunto de cebadores y sondas que en la titulación de preparación del vector [Zhang, Y., et al., (2001), Mol. Ther. 3, 697-707].

Mientras que la invención ha sido descrita con referencia a realizaciones particularmente preferidas, se apreciará que se pueden realizar modificaciones. Tales modificaciones está previsto que caigan dentro del alcance de las reivindicaciones.

45

50

**Listado de secuencias**

- <110> Los Fideicomisarios de la Universidad de Pennsylvania Gao, Guangping Wilson, James M. Alvira, Mauricio
- <120> Secuencias de serotipo 9 de virus adeno-asociado (AAV), vectores que las contienen, y uso de las mismas
- 5 <130> UPN-02734PCT
  - <150> US 60/341.150
  - <151> 17-12-2001
  - <150> US 60/386.132
  - <151> 05-06-2002
- 10 <160> 7
  - <170> Patentin versión 3.1
  - <210> 1
  - <211> 4382
  - <212> ADN
- 15 <213> serotipo 9 de virus adeno-asociado
  - <400> 1

20

25

30

35

ES 2 526 341 T3

cagagagggg gtggccaact ccatcactag gggtaatgca gaagcgctc ccacgctgcc 60  
 gcgtcagggc tgacgtagat taogtcatag gggagtgggc ctgtattagc tgtcacgtga 120  
 gtgcttttgc gacattttgc gacaccacat ggcatttga ggtatatatg gccgagtgag 180  
 cgagcaggat ctccattttg accgcgaaat ttgaacgagc agcagccatg ccgggcttct 240  
 acgagattgt gatcaagggt ccgagcgacc tggacgagca cctgccgggc atttctgact 300  
 cttttgtgaa ctgggtggcc gagaaggaat gggagctgcc ccgggattct gacatggatc 360  
 ggaatctgat cgagcaggca ccctgacog tggccgagaa gctgtagcgc gacttctctg 420  
 tccaatggcg ccgctgtagt aaggccccg aggcctctt ctttgttcag ttcgagaagg 480  
 gcgagagcta ctttcacctg cacgttctgg tcgagaccac gggggtcaag tccatggtgc 540  
 taggcgctt cctgagtcag attcgggaga agctgggtca gaccatctac ccggggatcg 600  
 agccgacct gcccaactgg ttgcgggtga ccaagacgcg taatggcgcc ggcgggggga 660  
 acaagggtgt ggacgagtgc tacatcccc aactacctct gcccaagact cagcccagc 720  
 tgcagtgggc gtggactaac atggaggagt atataagcgc gtgcttgaac ctggccgagc 780  
 gcaaacggct cgtggcgcag cacctgacct acgtcagcca gacgcaggag cagaacaagg 840  
 agaatctgaa cccaattct gacgcgccc tgatcaggtc aaaaacctcc gcgcgctaca 900

ES 2 526 341 T3

tggagctggt cgggtggctg gtggaccggg gcatcacctc cgagaagcag tggatccagg 960  
 aggaccaggc ctcgtacatc tccttcaacg ccgcctocaa ctogcggtoe cagatcaagg 1020  
 ccgcgctgga caatgcgggc aagatcatgg cgctgaccaa atccgcgccc gactacctgg 1080  
 taggcccctc acttcoogtg gacattacgc agaaccgeat ctaccgcac cctgcagctca 1140  
 acggctacga cctgcctac gccggctocg tctttctcgg ctgggcacaa aagaagttcg 1200  
 ggaaacgcaa caccatctgg ctgtttgggc cggccaccac gggaaagacc aacatcgag 1260  
 aagccattgc ccacgcgctg cccttctacg gctgcgtcaa ctggaccaat gagaactttc 1320  
 ccttcaacga ttgcgtcgac aagatggtga tctggtggga ggagggcaag atgacggcca 1380  
 aggtcgtgga gtccgccaag gccattctcg gcggcagcaa ggtgcgctg gacccaaagt 1440  
 gcaagtcgtc cgcaccagac gacccactc ccgtgatcgt cacctccac aaccaatgt 1500  
 gcgcgctgat tgacgggaac agcaccacct tcgagcacca gcagcctctc caggaccgga 1560  
 tgtttaagtt cgaactcacc cgcctctggt agcacgactt tggcaaggtg acaaagcagg 1620  
 aagtcaaaga gttcttccgc tgggccagtg atcaogtgac cgaggtggcg catgagtttt 1680  
 acgtcagaaa gggcggagcc agcaaaagac ccgccccga tgacgcggat aaaagcgagc 1740  
 ccaagcgggc ctgcccctca gtccggtac catcgacgtc agacgcggaa ggagctccgg 1800  
 tggactttgc cgacaggtac caaaacaaat gttctcgtca cgcgggcatg cttcagatgc 1860  
 tgcttcctg caaaacgtgc gagagaatga atcagaattt caacatttgc ttcacacacg 1920  
 gggtcagaga ctgctcagag tgtttcccg gogtgcaga atctcaaccg gtcgtcagaa 1980  
 agaggacgta tcggaaactc tgtgcgattc atcatctgct ggggcgggct cccgagattg 2040  
 cttgctcggc ctgcgatctg gtcaacgtgg acctggatga ctgtgtttct gagcaataaa 2100  
 tgacttaaac caggtatggc tgccgatggt tatcttccag attggctcga ggacaacctc 2160  
 totgagggca ttcgcgagtg gtgggacctg aaacctggag ccccgaaacc caaagccaac 2220  
 cagcaaaagc aggacgacgg ccggggtctg gtgcttctg gctacaagta cctcggacct 2280  
 ttoaacggac tcgacaaggg ggagcccgtc aacgcggcgg acgcagcggc cctcagacac 2340  
 gacaaggcct acgaccagca gctcaaagcg ggtgacaatc cgtacctgcg gtataaccac 2400  
 gccgacgccg agtttcagga gogtctgcaa gaagatacgt cttttggggg caacctcggg 2460  
 cgagcagtct tccaggcca gaagcgggtt ctgaaacctc tcggtctggt tgaggaaggc 2520  
 gctaagacgg ctcttgaaa gaagagaccg gtagagcagt caccccaga accagactca 2580

ES 2 526 341 T3

tcctcgggca	tcggcaaatac	aggccagcag	cccgctaaaa	agagactcaa	ttttggtcag	2640
actggcgact	cagagtcagt	ccccgaccca	caacctctcg	gagaacctcc	agaagcccc	2700
tcaggtctgg	gacctaatc	aatggcttca	ggcggtgccg	ctccaatggc	agacaataac	2760
gaaggcgccg	acggagtggg	taattcctcg	ggaaattggc	attgctgattc	cacatggctg	2820
ggggacagag	tcatcaccac	cagcacccca	acctgggcat	tgcccacctc	caacaaccac	2880
ctctacaagc	aaatctccaa	tggaacatcg	ggaggaagca	ccaacgacaa	cacctacttt	2940
ggctacagca	ccccctgggg	gtattttgac	ttcaacagat	tccactgcc	cttctoacca	3000
cgtgactggc	agcgactcat	caacaacaac	tggggattcc	ggccaaagag	actcaacttc	3060
aagctgttca	acatccaggt	caaggaggtt	acgacgaacg	aaggcaccaa	gaccatcgcc	3120
aataacctta	ccagcaacct	ccaggtcttt	acggactcgg	agtaccagct	acctgactgc	3180
ctaggctctg	cccaccaagg	atgcctgcc	ccgtttcctg	cagacgtctt	catggttctt	3240
cagtacggct	acctgacgct	caacaatgga	agtcaagcgt	taggacgttc	ttctttctac	3300
tgtctggaat	acttcccttc	tcagatgctg	agaaccggca	acaactttca	gttcagctac	3360
actttcgagg	acgtgccttt	ccacagcagc	tacgcacaca	gccagagtct	agatcgactg	3420
atgaaccccc	tcatcgacca	gtacctatac	tacctggcca	gaacacagac	aactggaact	3480
gggggaactc	aaactttggc	attcagccaa	gcaggcccta	gctcaatggc	caatcaggct	3540
agaaactggg	taccggggcc	ttgctaccgt	cagcagcggg	tctccacaac	caccaaccaa	3600
aataacaaca	gcaactttgc	gtggacggga	gctgctaaat	tcaagctgaa	cgggagagac	3660
tcgctaatag	atcctggcgt	ggctatggca	tcgcacaaag	acgacgagga	ccgcttcttt	3720
ccatcaagtg	gcgttctcat	at ttggcaag	caaggagccg	ggaacgatgg	agtgcactac	3780
agccaggtgc	tgattacaga	tgaggaagaa	attaaagcca	ccaacctgt	agccacagag	3840
gaatacggag	cagtggccat	caacaaccag	gcccgtacaa	cgcaggcgca	aactggactt	3900
gtgcataacc	agggagttat	tcctggatg	gtctggcaga	accgggacgt	gtacctgcag	3960
ggccctattt	gggctaaaaat	acctcacaca	gatggcaact	ttcaccctgc	tcctctgatg	4020
ggtggatttg	gactgaaaca	cccacctcca	cagattctaa	ttaaaaatac	accagtgcgg	4080
gcagatcctc	ctcttaacct	caatcaagcc	aagctgaact	ctttcatcac	gcagtacagc	4140
acgggacaag	tcagcgtgga	aatcgagtgg	gagctgcaga	aagaaaacag	caagcgctgg	4200
aatccagaga	tccagtatac	ttcaaaactac	tacaaatcta	caaatgtgga	ctttgctgtc	4260
aataccgaag	gtgtttactc	tgagcctcgc	cccattggta	ctcgttacct	caccogtaat	4320
ttgtaattgc	ctgttaatac	ataaacgggt	taattcgttt	cagttgaact	ttggtctctg	4380
cg						4382

ES 2 526 341 T3

<210> 2  
 <211> 736  
 <212> PRT  
 <213> proteína de cápside de serotipo 9 de virus adeno-asociado

5 <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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
          20          25          30

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

Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
          85          90          95

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

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

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

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

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

ES 2 526 341 T3

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro  
 180 185 190

Glu Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly  
 195 200 205

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

Ser Gly Asn Trp His Cys Asp Ser Thr 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 Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn  
 260 265 270

Thr 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 Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn  
 325 330 335

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

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

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

Gly Ser Gln Ala Leu 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 Thr

ES 2 526 341 T3

				405					410					415			
Phe	Glu	Asp	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	Val		
		435					440					445					
Arg	Thr	Gln	Thr	Thr	Gly	Thr	Gly	Gly	Thr	Gln	Thr	Leu	Ala	Phe	Ser		
	450					455						460					
Gln	Ala	Gly	Pro	Ser	Ser	Met	Ala	Asn	Gln	Ala	Arg	Asn	Trp	Val	Pro		
465					470					475					480		
Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Thr	Asn	Gln	Asn		
				485					490					495			
Asn	Asn	Ser	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Ala	Lys	Phe	Lys	Leu	Asn		
			500					505					510				
Gly	Arg	Asp	Ser	Leu	Met	Asn	Pro	Gly	Val	Ala	Met	Ala	Ser	His	Lys		
		515					520					525					
Asp	Asp	Glu	Asp	Arg	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Ile	Phe	Gly		
	530					535						540					
Lys	Gln	Gly	Ala	Gly	Asn	Asp	Gly	Val	Asp	Tyr	Ser	Gln	Val	Leu	Ile		
545					550					555					560		
Thr	Asp	Glu	Glu	Glu	Ile	Lys	Ala	Thr	Asn	Pro	Val	Ala	Thr	Glu	Glu		
				565					570					575			
Tyr	Gly	Ala	Val	Ala	Ile	Asn	Asn	Gln	Ala	Ala	Asn	Thr	Gln	Ala	Gln		
			580					585					590				
Thr	Gly	Leu	Val	His	Asn	Gln	Gly	Val	Ile	Pro	Gly	Met	Val	Trp	Gln		
		595					600					605					
Asn	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	Leu		
625					630					635					640		

ES 2 526 341 T3

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

Asp Pro Pro Leu Thr Phe Asn Gln Ala 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 Thr Asn Val Asp 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> 623
- <212> PRT
- <213> proteína rep de serotipo 9 de virus adeno-asociado
- <400> 3

5

Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15

Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30

Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Arg Asn Leu Ile  
35 40 45

Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60

Val Gln Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80

Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Leu His Val Leu Val Glu  
85 90 95

ES 2 526 341 T3

Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110

Arg Glu Lys Leu Val Gln Thr Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125

Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
130 135 140

Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160

Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Glu Tyr Ile  
165 170 175

Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190

Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Leu Asn  
195 200 205

Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220

Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys  
225 230 235 240

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270

Ile Met Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ser  
275 280 285

Leu Pro Val Asp Ile Thr Gln Asn Arg Ile Tyr Arg Ile Leu Gln Leu  
290 295 300

Asn Gly Tyr Asp Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala  
305 310 315 320

Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala

ES 2 526 341 T3

				325					330					335			
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Ala	Val	Pro		
			340					345					350				
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp		
		355					360					365					
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala		
	370					375					380						
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg		
	385				390					395					400		
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val		
				405					410						415		
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser		
			420					425					430				
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe		
		435					440					445					
Glu	Leu	Thr	Arg	Arg	Leu	Glu	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln		
	450					455					460						
Glu	Val	Lys	Glu	Phe	Phe	Arg	Trp	Ala	Ser	Asp	His	Val	Thr	Glu	Val		
	465				470					475					480		
Ala	His	Glu	Phe	Tyr	Val	Arg	Lys	Gly	Gly	Ala	Ser	Lys	Arg	Pro	Ala		
				485					490					495			
Pro	Asp	Asp	Ala	Asp	Lys	Ser	Glu	Pro	Lys	Arg	Ala	Cys	Pro	Ser	Val		
			500					505					510				
Ala	Asp	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Gly	Ala	Pro	Val	Asp	Phe	Ala		
		515					520					525					
Asp	Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Ala	Gly	Met	Leu	Gln	Met		
	530					535					540						
Leu	Leu	Pro	Cys	Lys	Thr	Cys	Glu	Arg	Met	Asn	Gln	Asn	Phe	Asn	Ile		
	545				550					555					560		

ES 2 526 341 T3

Cys Phe Thr His Gly Val Arg Asp Cys Ser Glu Cys Phe Pro Gly Val  
565 570 575

Ser Glu Ser Gln Pro Val Val Arg Lys Arg Thr Tyr Arg Lys Leu Cys  
580 585 590

Ala Ile His His Leu Leu Gly Arg Ala Pro Glu Ile Ala Cys Ser Ala  
595 600 605

Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Val Ser Glu Gln  
610 615 620

<210> 4

<211> 735

<212> PRT

5 <213> serotipo 2 de virus adeno-asociado

<400> 4

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

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

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro  
35 40 45

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

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

Arg Gln Leu Asp Ser 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 Val Leu Glu Pro  
115 120 125

ES 2 526 341 T3

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

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly  
 145 150 155 160

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

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro  
 180 185 190

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

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

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met 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 Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr  
 260 265 270

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

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

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

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr  
 340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp

ES 2 526 341 T3

355		360		365											
Val	Phe	Met	Val	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asn	Gly	Ser
370						375					380				
Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe	Pro	Ser
385					390					395					400
Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Thr	Phe	Ser	Tyr	Thr	Phe	Glu
				405					410					415	
Asp	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	Leu	Asp	Arg
			420					425						430	
Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ser	Arg	Thr
		435					440					445			
Asn	Thr	Pro	Ser	Gly	Thr	Thr	Thr	Gln	Ser	Arg	Leu	Gln	Phe	Ser	Gln
	450						455				460				
Ala	Gly	Ala	Ser	Asp	Ile	Arg	Asp	Gln	Ser	Arg	Asn	Trp	Leu	Pro	Gly
465					470					475					480
Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Lys	Thr	Ser	Ala	Asp	Asn	Asn
				485					490					495	
Asn	Ser	Glu	Tyr	Ser	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly
			500					505					510		
Arg	Asp	Ser	Leu	Val	Asn	Pro	Gly	Pro	Ala	Met	Ala	Ser	His	Lys	Asp
		515					520						525		
Asp	Glu	Glu	Lys	Phe	Phe	Pro	Gln	Ser	Gly	Val	Leu	Ile	Phe	Gly	Lys
	530					535					540				
Gln	Gly	Ser	Glu	Lys	Thr	Asn	Val	Asp	Ile	Glu	Lys	Val	Met	Ile	Thr
545					550					555					560
Asp	Glu	Glu	Glu	Ile	Arg	Thr	Thr	Asn	Pro	Val	Ala	Thr	Glu	Gln	Tyr
				565					570					575	
Gly	Ser	Val	Ser	Thr	Asn	Leu	Gln	Arg	Gly	Asn	Arg	Gln	Ala	Ala	Thr
			580					585					590		

ES 2 526 341 T3

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

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

Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys  
625 630 635 640 645

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

Pro Ser Thr Thr Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln  
660 665 670

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

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

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

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

- <210> 5
- <211> 735
- 5 <212> PRT
- <213> serotipo 1 de virus adeno-asociado
- <400> 5

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 Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

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

ES 2 526 341 T3

Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala  
85 90 95

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

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

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

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

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

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro  
180 185 190

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

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

Ser Gly Asn Trp His Cys Asp Ser Thr 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 Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His  
260 265 270

Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe



ES 2 526 341 T3

Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys  
 515 520 525

Asp Asp Glu Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe Gly  
 530 535 540

Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met Ile  
 545 550 555 560

Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Arg  
 565 570 575

Phe Gly Thr Val Ala Val Asn Phe Gln Ser Ser Ser Thr Asp Pro Ala  
 580 585 590

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

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

Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala 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 Val Gln Tyr Thr Ser Asn  
 690 695 700

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

Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Pro Leu  
 725 730 735

- <210> 6
- <211> 736
- 5 <212> PRT
- <213> serotipo 3 de virus adeno-asociado
- <400> 6

ES 2 526 341 T3

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 Val Pro Gln Pro  
 20 25 30

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

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

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

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

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

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro  
 180 185 190

Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly

ES 2 526 341 T3

	195							200							205
Ala	Pro	Met	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Asn	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	Ser	Gln	Ser	Gly	Ala	Ser	Asn	Asp	Asn	His	Tyr
			260					265					270		
Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg	Phe	His
		275					280					285			
Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	Asn	Trp
	290					295					300				
Gly	Phe	Arg	Pro	Lys	Lys	Leu	Ser	Phe	Lys	Leu	Phe	Asn	Ile	Gln	Val
305					310					315					320
Arg	Gly	Val	Thr	Gln	Asn	Asp	Gly	Thr	Thr	Thr	Ile	Ala	Asn	Asn	Leu
				325					330						335
Thr	Ser	Thr	Val	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln	Leu	Pro	Tyr
			340					345					350		
Val	Leu	Gly	Ser	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	Pro	Ala	Asp
		355					360					365			
Val	Phe	Met	Val	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asn	Gly	Ser
	370					375					380				
Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe	Pro	Ser
385					390					395					400
Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Gln	Phe	Ser	Tyr	Thr	Phe	Glu
				405					410					415	
Asp	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	Leu	Asp	Arg
			420					425					430		

ES 2 526 341 T3

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

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

Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro  
 465 470 475 480

Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn  
 485 490 495

Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn  
 500 505 510

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

Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly  
 530 535 540

Lys Glu Gly Thr Thr Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile  
 545 550 555 560

Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln  
 565 570 575

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

Thr Gly Thr Val Asn His Gln Gly Ala 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 His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu  
 625 630 635 640

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

ES 2 526 341 T3

Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala 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 Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn 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> 7
- <211> 738
- 5 <212> PRT
- <213> serotipo 8 de virus adeno-asociado
- <400> 7

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 Lys Pro  
20 25 30

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

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

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

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro

ES 2 526 341 T3

	115						120							125			
Leu	Gly	Leu	Val	Glu	Glu	Gly	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg		
	130					135					140						
Pro	Val	Glu	Pro	Ser	Pro	Gln	Arg	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile		
145					150					155					160		
Gly	Lys	Lys	Gly	Gln	Gln	Pro	Ala	Arg	Lys	Arg	Leu	Asn	Phe	Gly	Gln		
				165					170					175			
Thr	Gly	Asp	Ser	Glu	Ser	Val	Pro	Asp	Pro	Gln	Pro	Leu	Gly	Glu	Pro		
			180					185					190				
Pro	Ala	Ala	Pro	Ser	Gly	Val	Gly	Pro	Asn	Thr	Met	Ala	Ala	Gly	Gly		
		195					200					205					
Gly	Ala	Pro	Met	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser		
	210					215					220						
Ser	Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val		
225					230					235					240		
Ile	Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His		
				245					250					255			
Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ala	Thr	Asn	Asp		
			260					265					270				
Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn		
		275					280					285					
Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn		
	290					295					300						
Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Ser	Phe	Lys	Leu	Phe	Asn		
305					310					315					320		
Ile	Gln	Val	Lys	Glu	Val	Thr	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala		
				325					330					335			
Asn	Asn	Leu	Thr	Ser	Thr	Ile	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln		
			340					345					350				

ES 2 526 341 T3

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

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

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

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

Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser  
 420 425 430

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

Ser Arg Thr Gln Thr Thr Gly Gly Thr Ala Asn Thr Gln Thr Leu Gly  
 450 455 460

Phe Ser Gln Gly Gly Pro Asn Thr Met Ala Asn Gln Ala Lys Asn Trp  
 465 470 475 480

Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Gly  
 485 490 495

Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Ala Gly Thr Lys Tyr His  
 500 505 510

Leu Asn Gly Arg Asn Ser Leu Ala Asn Pro Gly Ile Ala Met Ala Thr  
 515 520 525

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

Phe Gly Lys Gln Asn Ala Ala Arg Asp Asn Ala Asp Tyr Ser Asp Val  
 545 550 555 560

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

ES 2 526 341 T3

Glu Glu Tyr Gly Ile Val Ala Asp Asn Leu Gln Gln Gln Asn Thr Ala  
 580 585 590

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

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

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

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

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

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

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

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

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

Asn Leu

**REIVINDICACIONES**

- 5 1.- Un virus adeno-asociado (AAV) recombinante que comprende una cápside de AAV9 que tiene la secuencia de aminoácido de SEQ ID Núm. 2 o una secuencia de aminoácido al menos un 95% idéntica con la misma, donde dicho AAV recombinante comprende además un minigén que tiene repeticiones de terminal invertido de AAV y un gen heterólogo enlazado operativamente a secuencias reguladoras que dirigen su expresión en una célula anfitrión.
- 2.- El AAV recombinante según la reivindicación 1, en el que la secuencia de aminoácido es al menos un 99% idéntica con los aminoácidos 1 a 736 de SEQ ID Núm. 2.
- 10 3.- El AAV recombinante según la reivindicación 1, en el que la secuencia de aminoácido es la de aminoácidos 1 a 736 de SEQ ID Núm. 2.
- 4.- Una proteína aislada que comprende una proteína de cápside de AAV9 seleccionada en el grupo consistente en:  
proteína de cápside vp1, aminoácidos (aa) 1 a 736;  
proteína de cápside vp2, aa 138 a 736, y  
proteína de cápside vp3, aa 203 a 736;
- 15 en el que los números de aminoácidos son los de la cápside de AAV9, SEQ ID Núm. 2.
- 5.- Una molécula que comprende una secuencia de ácido nucleico que codifica una proteína según la reivindicación 4, en la que dicha secuencia de ácido nucleico se elige en el grupo consistente en:  
vp1, nt 2116 a 4326;  
vp2, nt 2527 a 4326, y  
20 vp3, nt 2722 a 4326,
- en la que los números de nucleótidos son de AAV9, SEQ ID Núm. 1.
- 6.- La molécula según la reivindicación 5, en la que dicha molécula comprende una secuencia de AAV que codifica una proteína de cápside de AAV y una proteína rep funcional de AAV.
- 25 7.- La molécula según la reivindicación 5 ó 6, en la que dicha secuencia de ácido nucleico comprende una secuencia que codifica una proteína rep seleccionada en el grupo consistente en rep 78, rep 68, rep 52 y rep 40.
- 8.- La molécula según la reivindicación 6, en la que la secuencia de AAV que codifica la proteína rep se elige en el grupo consistente en AAV1, AAV2, AAV3, AAV4, AAV5 y AAV6.
- 9.- La molécula según cualquiera de las reivindicaciones 5 ó 6, en la que dicha molécula es un plásmido.
- 30 10.- Un método de generación de un virus adeno-asociado (AAV) recombinante que comprende una cápside de serotipo de AAV, que comprende las etapas de cultivar una célula anfitrión que contiene: (a) una molécula que codifica una proteína de cápside de AAV; (b) un gen rep funcional; (c) un minigén que comprende repeticiones de terminal invertido (ITRs) de AAV y un transgén, y (d) funciones auxiliares suficientes para permitir el empaquetamiento del minigén en la proteína de cápside de AAV, en el que dicha célula anfitrión comprende una molécula según cualquiera de las reivindicaciones 5 a 9.
- 35 11.- Una célula anfitrión en cultivo transfectada con un virus adeno-asociado según cualquiera de las reivindicaciones 1 a 3, o con una molécula según cualquiera de las reivindicaciones 5 a 9.
- 12.- Una composición que comprende un AAV según cualquiera de las reivindicaciones 1 a 3 y un portador fisiológicamente compatible.
- 40 13.- Un virus adeno-asociado según cualquiera de las reivindicaciones 1 a 3 o una molécula según cualquiera de las reivindicaciones 5 a 9, para su uso en el suministro de un transgén a una célula.

Fig. 1A

```

cagagagggg gtggccaact ccatcactag gggtaatcgc gaagcgcctc ccacgctgcc      60
gcgtcagcgc tgacgtagat tacgtcatag gggagtggtc ctgtattagc tgtcacgtga      120
gtgcttttgc gacatthttgc gacaccacat ggccatthtga ggtatataatg gccgagtga      180
cgagcaggat ctccatthttg accgcgaaat ttgaacgagc agcagccatg cccggcttct      240
          Rep68/78 start
acgagattgt gatcaaggtg ccgagcagacc tggacgagca cctgccgggc atthctgact      300
ctthttgtgaa ctgggtggcc gagaaggaat gggagctgcc cccgattct gacatggatc      360
ggaatctgat cgagcaggca cccctgaccg tggcagagaa gctgcagcgc gacttctctg      420
tccaatggcg ccgctgagat aaggccccgg aggcctctct cthttgtcag ttcgagaagg      480
gcgagagcta cthttcacctg cacgttctgg tcgagaccac ggggtcaag tccatgggtc      540
taggcgctt cctgagtcag attcgggaga agctgggtcca gaccatctac ccggggatcg      600
agccgaccct gcccaactgg ttcgcggtga ccaagacgcg taatggcgcc ggcgggggga      660
acaaggtggg ggacgagtgac tacatcccca actacctct gcccaagact cagcccagac      720
tgagtgggc gtggactaac atggaggagt atataagcgc gtgcttgaac ctggccgagc      780
gcaaacggct cgtggcgcag cacctgacct acgtcagcca gacgcaggag cagaacaagg      840
agaatctgaa ccccaattct gacgcgcccg tgatcaggtc aaaaacctoc gcgcgctaca      900
Rep40/52 start
tgagctggg cgggtggctg gtggaccggg goatcacctc cgagaagcag tggatccagg      960
aggaccaggc ctcgtaacac tcttcaacg ccgctccaa ctcgcggtcc cagatcaagg      1020
ccgctctgga caatgcggc aagatcatgg cgtgaccaa atccgcgcc gactacctg      1080
taggccttct acttccggtg gacattacgc agaaccgat ctaccgcat ctgcagctca      1140
acggctacga cctgcctac gccggctccg tctttctcgg ctgggcacaa aagaagttc      1200
ggaaacgcaa caccatctgg ctgtttggc cggccaccac gggaaagacc aacatcgcag      1260
aagccattgc ccacgcctg ccttctacg gctgcgtcaa ctggaccaat gagaactttc      1320
ccttcaacga ttgcgtgac aagatgggtg tctggtggga ggagggcaag atgacggcca      1380
aggtcgtgga gtccgccaag gccattctcg gcggcagcaa ggtgcgcgtg gacccaaagt      1440
gcaagtctgc ccccagatc gacccactc ccgtgatcgt cacctccaac accaactatg      1500

```

Fig. 1B

gcgcogtgat tgacgggaac agcaaccaact tcgagcacca gcagcctctc caggaccgga 1560  
 tgtttaagtt cgaactcacc cgcogtctgg agcacgactt tggcaaggtg acaaagcagg 1620  
 aagtcaaaga gttcttccgc tgggocagtg atcacgtgac cgagggtggcg catgagtttt 1680  
 acgtcagaaa gggcgggagc agcaaaaagc cggccccga tgacgggat aaaagcgagc 1740  
 ccaagcgggc ctgcccctca gtogcgggato oatogacgtc agacgggaa ggagctccgg 1800  
 tggactttgc cgacaggtac caaaacaaat gttctcgtca cgcgggcatg cttcagatgc 1860  
 tgcttccctg caaaacgtgc gagagaatga atcagaattt caacatttgc ttcacacacg 1920  
 gggtcagaga ctgctcagag tgtttcccg gogtgcaga atctcaaccg gtctcagaa 1980  
 agaggacgta toggaaacte tgtgcgattc atcatctgct ggggcgggct cccgagattg 2040  
 ctgctcggc ctgcgatctg gtcaacgtgg aootggatga ctgtgtttct gagcaat<sup>Rep 78</sup>aaa<sup>stop</sup> 2100  
 tgacttaaac <sup>vp1 start</sup> caggtatggc tgccgatggt tatcttccag attggetcga ggacaacctc 2160  
 tctgagggca ttcgcgagtg gtgggacctg aaacctggag ccccgaaacc caaagccaac 2220  
 cagcaaaagc aggacgacgg ccggggtctg gtgcttctg gctacaagta cctcggacc 2280  
 ttcaacggac togacaaggg ggagcccgtc aacgcggcgg acgcagcggc cctcagcac 2340  
 gacaaggcct acgaccagca gctcaaaagc ggtgacaato cgtacctgog gtataaccac 2400  
 gcogaogcog agtttcagga gcgtctgcaa gaagatacgt cttttggggg caacctcggg 2460  
 cgagcagttt tccaggccaa gaagcggggt ctcgaaacctc toggctcggg tgaggaaggo 2520  
 gctaag<sup>vp2 start</sup>acgg ctctcggaaa gaagagaccg gtagagcagt caccocaaaga acccagactca 2580  
 tctcgggca toggcaate agccagcag ccogctaaaa agagactca ttttggtcag 2640  
 actggcgaet cagagtcagt ccccgaccca caacctctcg gagaacctcc agaagcccc 2700  
 tcaggtctgg gacctaatc <sup>vp3 start</sup>aatggcttca ggcggtggcg ctccaatggc agacaataac 2760  
 gaaggcgcog accgagtggg taattootog ggaaattggc attgcgatte cacatggctg 2820  
 gggacagag tcatcaccac cagcaccoga aootgggcat tgcccaccta caacaaccac 2880  
 ctctacaagc aaatctccaa tgyaacatog ggaggaagca ccaacgaaa cacctacttt 2940

Fig. 1C

```

ggctacagca ccccoctgggg gtatthttgac ttcaacagat tccactgcca cttctcacca 3000
cgtgactggc agcgactcat caacaacaac tggggattcc ggccaaagag actcaacttc 3060
aagctgttca acatccaggt caaggaggtt acgacgaacg aaggcaccoc gaccatcgcc 3120
aataacotta ccagcaccgt ccaggtcttt acggactcgg agtaccagct accgtacgtc 3180
ctaggctctg cccaccaagg atgootgcca cggtttctct gacacgtott catggttctt 3240
cagtacggct acctgacgct caacaatgga agtcaagcgt taggaegtte ttottttctac 3300
tgtctggaat acttcccttc tcagatgctg agaaccggca acaactttca gttcagctac 3360
actttogagg acgtgccttt ccacagcagc tacgcacaca gdcagagtct agatcgactg 3420
atgaaccccc tcctcgacca gtacctatac tacctggtca gaacacagac aactggaact 3480
gggggaactc aaactttggc attcagccca gcaggcccta gctcaatggc caatcagget 3540
agaaactggg taccctgggc ttgctaccgt cagcagcgcg totcoacaac caccacccca 3600
aataacaaca gcaactttgc gtggacggga gctgctaaat tcaagctgaa cgggagagac 3660
tcgctaataa atcctggcgt ggotatggca togcacaaag acgacgagga ccgcttcttt 3720
ccatcaagtg gogttctcat atttggcaag caaggagcog ggaacgatgg agtcgactac 3780
agccaggtgc tgattacaga tgaggaagaa attaaagcca ccaaccctgt agccacagag 3840
gaatacggag cagtggccat caacaaccag gccgctaaca cgcaggcgcga aactggactt 3900
gtgcataacc agggagttat tcctggtatg gtctggcaga accgggacgt gtacctgcag 3960
ggccctatth gggctaaaat acctcaocaa gatggcaact ttoaccogtc tcctctgatg 4020
ggtggatttg gactgaaaca cccacctcca cagattctaa ttaaaaatac accagtgccg 4080
gcagatcctc ctcttacctt caatcaagcc aagctgaact ctttcatcac gcagtacagc 4140
acgggacaag tcagcgtgga aatcgagtgg gagctgcaga aagaaaacag caagcgtctg 4200
aatocagaga tccagtatac ttcaaaactac taaaaatcta caaatgtgga ctttgotgtc 4260
aatacogaag gtgtttactc tgagcctcgc ccatttgga ctcgttacct caccogtaat 4320
  vpl-3 stop          polyA
ttgtaattgc ctgttaatca ataaacoggt taattcgtht cagttgaact ttggtctctg 4380
cg                                                                 4382

```

Fig. 2A

	1				50
AAV_2	MAADGYLPDW	LEDTLSEGIR	QWWK <b>L</b> KPGPP	PPKPAERHKD	DSRGLVLPGY
AAV_8	MAADGYLPDW	LEDNLSEGIR	EW <b>W</b> AL <b>K</b> PGAP	KPKAN <b>Q</b> QKQD	DGRGLVLPGY
AAV_1	MAADGYLPDW	LEDNLSEGIR	EW <b>W</b> DL <b>K</b> PGAP	KPKAN <b>Q</b> QKQD	DGRGLVLPGY
AAV_3	MAADGYLPDW	LEDNLSEGIR	EW <b>W</b> AL <b>K</b> PGVP	Q <b>P</b> KAN <b>Q</b> QH <b>Q</b> D	NRRGLVLPGY
AAV_9	MAADGYLPDW	LEDNLSEGIR	EW <b>W</b> DL <b>K</b> PGAP	K <b>P</b> KAN <b>Q</b> Q <b>K</b> OD	DGRGLVLPGY
	51				100
AAV_2	KYLGPFNGLD	KGEPVNEADA	AALEHDKAYD	RQLDSGDNPY	LKYNHADAEF
AAV_8	KYLGPFNGLD	KGEPVNAADA	AALEHDKAYD	QQLQAGDNPY	LRYNHADAEF
AAV_1	KYLGPFNGLD	KGEPVNAADA	AALEHDKAYD	QQLKAGDNPY	LRYNHADAEF
AAV_3	KYLGPFNGLD	KGEPVNEADA	AALEHDKAYD	QQLKAGDNPY	LKYNHADAEF
AAV_9	KYLGPFNGLD	KGEPVNAADA	AALEHDKAYD	QQLKAGDNPY	LRYNHADAEF
	101				150
AAV_2	QERLKEDTSF	GGNLGRAVFQ	AKKRVLEPLG	LVEEPVKTAP	GKKRPVEHSP
AAV_8	QERLQEDTSF	GGNLGRAVFQ	AKKRVLEPLG	LVEEGAKTAP	GKKRPVEPSP
AAV_1	QERLQEDTSF	GGNLGRAVFQ	AKKRVLEPLG	LVEEGAKTAP	GKKRPVEQSP
AAV_3	QERLQEDTSF	GGNLGRAVFQ	AKKRILEPLG	LVEEAAKTAP	GKKGAVDQSP
AAV_9	QERLQEDTSF	GGNLGRAVFQ	AKKRVLEPLG	LVEEGAKTAP	GKKRPVEQSP
	151				200
AAV_2	.VEPDSSSGT	GKAGQQPARK	RLNFGQTGDA	DSVPDPQPLG	QPPAAPSGLG
AAV_8	QRSPDSSTGI	GKKGQQPARK	RLNFGQTGDS	ESVPDPQPLG	EPPAAPSGVG
AAV_1	.QEPDSSSGI	GKTGQQPAK	RLNFGQTGDS	ESVPDPQPLG	EPPATPAAVG
AAV_3	.QEPDSSSGV	GKSGKQPARK	RLNFGQTGDS	ESVPDPQPLG	EPPAAPTSLG
AAV_9	QE. <u>PDSSSGI</u>	<u>GKSGQQPAK</u>	RLNFGQTGDS	ESVPDPQPLG	EPPEA <u>PSGLG</u>
	201				250
AAV_2	TNTMATGSGA	PMADNNEGAD	GVGNSGNWH	CDSTWMGDRV	ITTSTRTWAL
AAV_8	PNTMAAGGGA	PMADNNEGAD	GVGSSGNWH	CDSTWLGDRV	ITTSTRTWAL
AAV_1	PTTMAAGGGA	PMADNNEGAD	GVGNASGNWH	CDSTWLGDRV	ITTSTRTWAL
AAV_3	SNTMAAGGGA	PMADNNEGAD	GVGNSGNWH	CDSQWLGDRV	ITTSTRTWAL
AAV_9	PNTMAAGGGA	PMADNNEGAD	GVGNSGNWH	CDSTWLGDRV	ITTSTRTWAL
	251				300
AAV_2	PTYNNHLYKQ	ISSQS--GAS	NDNHYFGYST	PWGYDFNRF	HCHFSPRDWQ
AAV_8	PTYNNHLYKQ	ISNGTSGGAT	NDNTYFGYST	PWGYDFNRF	HCHFSPRDWQ
AAV_1	PTYNNHLYKQ	ISSAST.GAS	NDNHYFGYST	PWGYDFNRF	HCHFSPRDWQ
AAV_3	PTYNNHLYKQ	ISSQS.GAS	NDNHYFGYST	PWGYDFNRF	HCHFSPRDWQ
AAV_9	PTYNNHLYKQ	ISNGTSGGST	<u>NDNTYFGYST</u>	PWGYDFNRF	HCHFSPRDWQ

Fig. 2B

	301				350
AAV_2	RLINNNWGFR	PKRLNFKLFN	IQVKEVTQND	GTTTIANNLT	STVQVFTDSE
AAV_8	RLINNNWGFR	PKRLSFKLFN	IQVKEVTQNE	GTKTIANNLT	STIQVFTDSE
AAV_1	RLINNNWGFR	PKRLNFKLFN	IQVKEVTTND	GVTTIANNLT	STVQVFSNSE
AAV_3	RLINNNWGFR	PKKLSFKLFN	IQVRGVTQND	GTTTIANNLT	STVQVFTDSE
AAV_9	RLINNNWGFR	PKRLNFKLFN	IQVKEVTTNE	GTKTIANNLT	STVQVFTDSE

	351				400
AAV_2	YQLPYVLGSA	HQGCLPPFPA	DVFMVPQYGY	LTLNNGSQAV	GRSSFYCLEY
AAV_8	YQLPYVLGSA	HQGCLPPFPA	DVFMIPQYGY	LTLNNGSQAV	GRSSFYCLEY
AAV_1	YQLPYVLGSA	HQGCLPPFPA	DVFMIPQYGY	LTLNNGSQAV	GRSSFYCLEY
AAV_3	YQLPYVLGSA	HQGCLPPFPA	DVFMVPQYGY	LTLNNGSQAV	GRSSFYCLEY
AAV_9	YQLPYVLGSA	HQGCLPPFPA	DVFMVPQYGY	LTLNNGSQAL	GRSSFYCLEY

	401				450
AAV_2	FPSQMLRTGN	NFTFSYTFED	VPFHSSYAHS	QSLDRLMNPL	IDQYLYYLSR
AAV_8	FPSQMLRTGN	NFQFTYTFED	VPFHSSYAHS	QSLDRLMNPL	IDQYLYYLSR
AAV_1	FPSQMLRTGN	NFTFSYTFEE	VPFHSSYAHS	QSLDRLMNPL	IDQYLYYLSR
AAV_3	FPSQMLRTGN	NFQFSYTFED	VPFHSSYAHS	QSLDRLMNPL	IDQYLYYLSR
AAV_9	FPSQMLRTGN	NFQFSYTFED	VPFHSSYAHS	QSLDRLMNPL	IDQYLYYLSR

	451				500
AAV_2	TNTPSG.TTT	QSRLQFSQAG	ASDIRDQS	RNWLPGPCYRQQ	RVSKTSADNN
AAV_8	TQTTGG.TAN	TQTLGFSQGG	PNTMANQA	KNWLPGPCYRQQ	RVSTTTGQNN
AAV_1	TQ.NQSGSAQ	NKDLLFSRGS	PAGMSVQP	KNWLPGPCYRQQ	RVSKTKTDNN
AAV_3	TQGTSGTTN	QSRLLEFSQAG	PQSMSLQA	RNWLPGPCYRQQ	RLSKTANDNN
AAV_9	<u>TQTTG</u> .TGG	<u>TQTLA</u> FSQAG	<u>PSSMANQA</u>	<u>RNWV</u> PGPCYRQQ	<u>RVSTTTN</u> QNN

	501				550
AAV_2	NSEYSWTGAT	KYHLNGRDSL	VNPGPAMASH	KDDEEKFFPQ	SGVLI <u>FGKQ</u> G
AAV_8	NSNFAWTAGT	KYHLNGRNSL	ANPGIAMATH	KDDEERFFPS	NGILIFGKQN
AAV_1	NSNFTWTGAS	KYNLNGRESI	INPGTAMASH	KDDEDKFFPM	SGVMI <u>FGKES</u>
AAV_3	NSNFPWTAAS	KYHLNGRDSL	VNPGPAMASH	KDDEEKFFPM	HGNLI <u>FGKES</u>
AAV_9	<u>NSNFAWT</u> GAA	<u>KFKLNGR</u> DSL	<u>MNPG</u> VAMASH	<u>KDDEDR</u> FFPS	<u>SGVLI</u> <u>FGKQ</u> G

	551				600
AAV_2	SEKTNVDIEK	VMITDEEEIR	T TNPVATEQY	GSVSTNLORG	NRQAATADVN
AAV_8	AARDNADYSD	VMLTSEEEIK	T TNPVATEEY	GIVADNLQQQ	NTAPQIGTVN
AAV_1	AGASNTALDN	VMITDEEEIK	A TNPVATERF	GTVAVNFQSS	STDPATGDVH
AAV_3	TTASNAELDN	VMITDEEEIR	T TNPVATEQY	GTVANNLQSS	NTAPTGTGVN
AAV_9	AGNDGVDYSQ	<u>VLI</u> <u>TDEEEIK</u>	A TNPVATEEY	<u>GAVAIN</u> NQAA	NTQAQTGLVH

Fig. 2C

	601				650
AAV_2	TQGVLP	QDRDVYLQGP	IWAKIPHTDG	HFHPSPLMGG	FGLKHPPPQI
AAV_8	SQGALP	QNRDVYLQGP	IWAKIPHTDG	NFHPSPLMGG	FGLKHPPPQI
AAV_1	AMGALP	QDRDVYLQGP	IWAKIPHTDG	HFHPSPLMGG	FGLKNPPPQI
AAV_3	HQGALP	QDRDVYLQGP	IWAKIPHTDG	HFHPSPLMGG	FGLKHPPPQI
AAV_9	NOGVIP	QNRDVYLQGP	IWAKIPHTDG	NFHPSPLMGG	FGLKHPPPQI

	651				700
AAV_2	LIKNTVPVA	NPSTTFSAAKF	ASFITQYSTG	QVSVEIEWEL	QKENSKRWNP
AAV_8	LIKNTVPVA	DPPTTFNQSKL	NSFITQYSTG	QVSVEIEWEL	QKENSKRWNP
AAV_1	LIKNTVPVA	NPPAEFSATKF	ASFITQYSTG	QVSVEIEWEL	QKENSKRWNP
AAV_3	MIKNTVPVA	NPPTTFSPAKF	ASFITQYSTG	QVSVEIEWEL	QKENSKRWNP
AAV_9	LIKNTVPVA	<u>DP</u> <u>PL</u> <u>T</u> <u>E</u> <u>N</u> <u>O</u> <u>A</u> <u>K</u> <u>L</u>	NSFITQYSTG	QVSVEIEWEL	QKENSKRWNP

	701				739
AAV_2	EIQYTSN	SVNVDFTVDT	NGVYSEPRPI	GTRYLTRNL	
AAV_8	EIQYTSN	STSVDFAVNT	EGVYSEPRPI	GTRYLTRNL	
AAV_1	EVQYTSN	SANVDFTVDN	NGLYTEPRPI	GTRYLTRPL	
AAV_3	EIQYTSN	SVNVDFTVDT	NGVYSEPRPI	GTRYLTRNL	
AAV_9	EIQYTSN	<u>Y</u> <u>K</u> <u> </u> <u>S</u> <u>T</u> <u>N</u> <u>V</u> <u>D</u> <u>F</u> <u>A</u> <u>V</u> <u>N</u> <u>T</u>	<u>E</u> <u>G</u> <u>V</u> <u>Y</u> <u>S</u> <u>E</u> <u>P</u> <u>R</u> <u>P</u> <u>I</u>	GTRYLTRNL	

Fig. 3A

Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15

Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30

Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Arg Asn Leu Ile  
 35 40 45

Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60

Val Gln Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80

Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Leu His Val Leu Val Glu  
 85 90 95

Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110

Arg Glu Lys Leu Val Gln Thr Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125

Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140

Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160

Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Glu Tyr Ile  
 165 170 175

Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190

Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Leu Asn  
 195 200 205

Fig. 3B

Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220

Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys  
 225 230 235 240

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270

Ile Met Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ser  
 275 280 285

Leu Pro Val Asp Ile Thr Gln Asn Arg Ile Tyr Arg Ile Leu Gln Leu  
 290 295 300

Asn Gly Tyr Asp Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320

Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro  
 340 345 350

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365

Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415

