

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 644 880**

51 Int. Cl.:

C12N 7/04 (2006.01)

C07K 14/015 (2006.01)

C12N 5/10 (2006.01)

C12N 15/85 (2006.01)

C12N 15/864 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.02.2009 PCT/NL2009/050076**

87 Fecha y número de publicación internacional: **27.08.2009 WO09104964**

96 Fecha de presentación y número de la solicitud europea: **19.02.2009 E 09713345 (8)**

97 Fecha y número de publicación de la concesión europea: **02.08.2017 EP 2250256**

54 Título: **Optimización de expresión de proteínas Rep y Cap parvovirales en las células de insecto**

30 Prioridad:

19.02.2008 EP 08151634

19.02.2008 US 29673 P

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

30.11.2017

73 Titular/es:

UNIQUE IP B.V. (100.0%)

Paasheuvelweg 25a

1105 BP Amsterdam, NL

72 Inventor/es:

BAKKER, ANDREW CHRISTIAN;

HERMENS, WILHELMUS THEODORUS

JOHANNES MARIACHRISTIAAN y

NOORDMAN, YVET

74 Agente/Representante:

TOMAS GIL, Tesifonte Enrique

ES 2 644 880 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

Optimización de expresión de proteínas Rep y Cap parvovirales en las células de insecto

5 Campo de la invención

[0001] Esta invención se refiere a la producción de vectores de parvovirus, especialmente a la producción de virus adeno-asociados recombinantes (rAAV) en las células de insecto, vectores de expresión baculovirales que comprenden un constructo de la invención y una célula que comprende tal vector de expresión baculoviral.

10

Antecedentes de la invención

[0002] El sistema de expresión de baculovirus es bien conocido por su uso como clonación eucariótica y vector de expresión (rey, L.A., y F D. Possee, 1992, "The baculovirus expression system", Chapman and Hall, United Kingdom; O'Reilly, D. R., et al., 1992. Baculovirus Expression Vectors: A Laboratory Manual. Nueva York: W. H. Freeman.).

15

Las ventajas del sistema de expresión de baculovirus son entre otras que las proteínas expresadas son casi siempre solubles, correctamente plegadas y biológicamente activas.

Otras ventajas incluyen altos niveles de expresión de proteína, producción más rápida, idoneidad para la expresión de proteínas grandes e idoneidad para producción a gran escala.

20

Sin embargo, en la producción a gran escala o continua de proteínas heterólogas utilizando el sistema de expresión de baculovirus en los biorreactores de célula de insecto, la inestabilidad de niveles de producción, también conocida como el efecto de pasaje, es un gran obstáculo.

Este efecto es al menos en parte debido a la recombinación entre secuencias homólogas repetidas en el ADN baculoviral.

25

[0003] El sistema de expresión de baculovirus también se ha usado exitosamente para la producción de vectores (AAV) de virus adeno-asociados recombinantes (Urabe et al., 2002, Hum. Gene Ther. 13: 1935-1943; US 6,723,551 y US 20040197895).

AAV se puede considerar como uno de los vectores virales más prometedores para terapia genética humana.

30

AAV tiene la capacidad de infectar eficazmente la división al igual que células humanas que no se dividen, el genoma vírico AAV se integra en un sitio cromosómico único en el genoma de la célula huésped, y lo más importante, aunque AAV esté presente en muchos seres humanos nunca se ha asociado a ninguna enfermedad.

En vista de estas ventajas, el virus adeno-asociado recombinante (rAAV) está siendo evaluado en las pruebas clínicas de terapia genética para hemofilia B, melanoma maligno, fibrosis quística, hiperlipoproteinemia tipo I y otras enfermedades.

35

[0004] Para superar problemas con sistemas de producciones mamíferas para AAV Urabe et al. (2002; supra) desarrolló un sistema de producción AAV en las células de insecto.

Para la producción de AAV en las células de insecto algunas modificaciones fueron necesarias para conseguir la estequiometría correcta de las tres proteínas cápsidas AAV (VP1, VP2 y VP3), que dependen de una combinación de uso alternativo de dos sitios de aceptor de empalme y la utilización subóptima de un ACG codón iniciador para VP2 que no está reproducido con precisión por células de insecto.

40

Para imitar la estequiometría correcta de las proteínas cápsidas en las células de insecto Urabe et al. (2002; supra) se usa un constructo que se transcribe en un mensajero policistrónico único que es capaz de expresar todas las tres proteínas de vasopresina sin requerir empalme y donde el codón de inicio aguas arriba se sustituye por el codón de inicio subóptimo ACG.

45

La WO2007/046703 divulga otra mejora de la infectividad de la producción basada en vectores rAAV producidos por baculovirus por optimización de la estequiometría de proteínas cápsidas AAV como se ha producido en las células de insecto.

50

[0005] Para la expresión de las proteínas AAV Rep en el sistema de expresión de célula de insecto AAV como se ha desarrollado inicialmente por Urabe et al. (2002; supra), un constructo de baculovirus recombinante se usa que alberga dos unidades de expresión Rep independientes (una para Rep78 y una para Rep52), cada una bajo el control de un promotor de célula de insecto separado, los promotores \otimes IE1 y PolH, respectivamente.

55

Sin embargo, Kohlbrenner et al. (2005, Mol. Ther. 12: 1217-25; WO 2005/072364) declaró que el constructo de baculovirus para la expresión de las dos proteínas Rep, como se ha usado por Urabe et al., padece de una inestabilidad inherente.

Por la separación de la orientación palíndroma de los dos genes Rep en el vector original Urabe y diseñar dos vectores de baculovirus separados para la expresión Rep52 y Rep78, Kohlbrenner et al. (2005; supra) aumentó la estabilidad de pasaje del vector.

60

Sin embargo, a pesar de la expresión que consiste en Rep78 y Rep52 de los dos constructos de baculovirus-Rep independientes en las células de insecto sobre al menos 5 pasajes, el rendimiento de vector rAAV es 5 a 10-plegues inferior en comparación con el constructo de baculovirus-Rep original diseñado por Urabe et al. (2002, supra).

[0006] En la aplicación WO2007/148971, los presentes inventores han mejorado significativamente la estabilidad de producción de vector rAAV en las células de insecto usando una secuencia codificante única para las proteínas Rep78 y Rep52 donde un codón de inicio subóptimo se usa para la proteína Rep78 que es omitida parcialmente por el escaneado de ribosomas para permitir la iniciación de traducción para también producirse en el codón de inicio aguas abajo de la proteína Rep52.

[0007] La solicitud de patente internacional WO 2007/084773 divulga un método de producción rAAV en las células de insecto, donde la producción de partículas víricas infecciosas se aumenta por la complementación VP1 con respecto a VP2 y VP3.

La suplementación se puede efectuar por la introducción en la célula de insecto de una expresión de secuencias de nucleótidos que comprenden el vector de cápsida VP1, VP2 y VP3 y adicionalmente introducción en la célula de insecto de una expresión de secuencias de nucleótidos VP1, que puede estar bien en el mismo vector cápsida o en un vector diferente.

[0008] Sin embargo, todavía hay una necesidad de otras mejoras en la producción (comercial) a gran escala de vectores parvovirales en las células de insecto.

Así es un objeto de la presente invención proveer de medios y métodos que permiten una producción de rendimiento elevado y estable (a gran escala) de vectores parvovirales y para la producción que resulta en una proporción de partícula mejorada completa:vacía (es decir, una proporción superior de partículas rellenas).

Resumen de la invención

[0009] La invención se refiere a un método para la producción de un virión parvoviral recombinante.

El uso de tal método puede permitir la producción de tales viriones en valoraciones de producción aumentadas.

El uso del método puede alternativamente, o adicionalmente, permitir la producción con una proporción superior de partículas rellenas, es decir, una proporción completa: vacía o proporción total:completa más favorable.

[0010] La invención proporciona un método para la producción de un virión parvoviral recombinante, cuyo método comprende los pasos de:

(a) proporcionar una célula de insecto que comprende uno o más constructos de ácidos nucleicos que comprende:

(i) una secuencia de nucleótidos que comprende un transgen que se flanquea por al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral;

(ii) un primer casete de expresión que comprende una secuencia de nucleótidos que codifica una proteína Rep parvoviral que está operativamente enlazada a un primer promotor que es capaz de dirigir la expresión de la proteína Rep en la célula de insecto;

(iii) un segundo casete de expresión que comprende una secuencia de nucleótidos que codifica una proteína cápsida parvoviral que está operativamente enlazada a un segundo promotor que es capaz de dirigir la expresión de la proteína cápsida en la célula de insecto;

(b) cultivo de la célula definido en (a) bajo una condición propicia para la expresión de la proteína Rep y la proteína cápsida; y,

(c) opcionalmente recuperación del virión parvoviral recombinante;

donde los primeros y segundos casetes de expresión están presentes en un constructo de ácidos nucleicos único y donde los primeros y segundos casetes de expresión están presentes en la cantidad equimolar en la célula de insecto, donde el primer promotor se selecciona del grupo que consiste en un promotor PolH, promotor p10 o promotor básico y donde el segundo promotor es un promotor deltaE1 o un promotor E1.

[0011] La invención también proporciona:

- un constructo de ácidos nucleicos que comprende un primer y un segundo casete de expresión tal como se ha definido anteriormente, donde:

(d) el primer promotor es un promotor PolH y el segundo promotor es un promotor deltaE1 o un E1;

(e) el primer promotor es un promotor p10 y el segundo promotor es un promotor deltaE1 o un E1; o

y donde el primer casete de expresión opcionalmente comprende un elemento intensificador;

- una célula de insecto tal y como se define por arriba; y

- un equipo que comprende

(a) un constructo de ácidos nucleicos que comprende el primer y segundo casete de expresión tal como se ha definido anteriormente; y

(b) un constructo de ácidos nucleicos que comprende una secuencia de nucleótidos que codifica un sitio de clonación múltiple para un transgen que se flanquea por al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral, donde el transgen está operativamente enlazado a un promotor capaz de dirigir la expresión del transgen en una célula huésped.

Breve descripción de las figuras

[0012]

La Figura 1 muestra una representación esquemática de la construcción de pVD118(nueva).
 La Figura 2 muestra una representación esquemática de la construcción de pVD118(nueva)+HR1.
 La Figura 3 muestra una representación esquemática de la construcción de pVD165.
 La Figura 4 muestra una representación esquemática de la construcción de pVD165+HR1.
 La Figura 5 muestra una representación esquemática de la construcción de pVD165 + 4xEcRE Cap.
 La Figura 6 muestra una representación esquemática de la construcción de pVD165 + 4*EcRE Rep78.
 La Figura 7 muestra una representación esquemática de la construcción de pVD165 + deltaE1 Cap.
 La Figura 8 muestra una representación esquemática de la construcción de deltaE1 Cap + pPolh Rep (pVD190).
 La Figura 9 muestra una representación esquemática de la construcción de p10 de Cap + pPolh Rep.
 La Figura 10 muestra una representación esquemática de la construcción de pVD194.
 La Figura 11A muestra análisis de electrotransferencia de proteínas Cap y Rep expresadas de Bac.VD194, tres días después de la infección con las proporciones de baculovirus 1:1 y 5:1 (C=Bac.VD88:Bac.VD84:Bac.VD43 a 5:1:1 respectivamente).
 La Figura 11B muestra las valoraciones víricas para las mismas producciones analizadas en la figura 11A.

Descripción de la invención

Definiciones

[0013] Como se utiliza en este caso, el término "operativamente enlazado" se refiere a una conexión de elementos de polinucleótido (o polipéptido) en una relación funcional.

Un ácido nucleico está "operativamente enlazado" cuando se coloca en una relación funcional con otra secuencia de ácidos nucleicos.

Por ejemplo, una secuencia reguladora de transcripción está operativamente enlazada a una secuencia codificante si esta afecta a la transcripción de la secuencia codificante.

Operativamente enlazado significa que las secuencias de ADN que están enlazadas son contiguas típicamente y, donde es necesario para unir dos regiones de codificación de proteína, contiguas y en el marco de lectura.

[0014] "Secuencia de control de expresión" se refiere a una secuencia de ácidos nucleicos que regula la expresión de una secuencia de nucleótidos a la que está operativamente enlazada.

[0015] Una secuencia de control de expresión está "operativamente enlazada" a una secuencia de nucleótidos cuando la secuencia de control de expresión controla y regula la transcripción y/o la traducción de la secuencia de nucleótidos.

Así, una secuencia de control de expresión puede incluir promotores, potenciadores, sitios de introducción de ribosoma interno (IRES), terminadores de transcripción, un codón de inicio delante de un gen de codificación de proteína, señal de empalme para intrones y codones de parada.

El término "secuencia de control de expresión" se destina a incluir, como mínimo, una secuencia cuya presencia está diseñada para influir en la expresión y también puede incluir componentes ventajas adicionales.

Por ejemplo, las secuencias líder y secuencias de pareja de fusión son secuencias de control de expresión.

El término puede también incluir el diseño de la secuencia de ácidos nucleicos de manera que, codones de iniciación indeseables potenciales dentro y fuera del bastidor se retiran de la secuencia.

Esto puede también incluir el diseño de la secuencia de ácidos nucleicos de manera que se quiten los sitios de empalme potencial indeseable.

Incluye secuencias o secuencias de poliadenilación (pA) que dirigen la adición de una cola polyA, es decir, una cadena de residuos de adenina en el 3'-terminal de un ARNm, secuencias referidas como secuencias polyA.

También se puede diseñar para mejorar la estabilidad de ARNm.

Las secuencias de control de expresión que afectan a la transcripción y estabilidad de traducción, por ejemplo, promotores, al igual que las secuencias que efectúan la traducción, por ejemplo, las secuencias Kozak, se conocen en las células de insecto.

Las secuencias de control de expresión pueden ser de tal naturaleza en cuanto a modular la secuencia de nucleótidos a la que están operativamente enlazadas de manera que niveles de expresión inferiores o niveles de expresión más altos se consiguen.

[0016] Como se utiliza en este caso, el término "promotor" o "secuencia de regulación de transcripción" se refiere a un fragmento de ácido nucleico que funciona para controlar la transcripción de una o más secuencias codificantes y está localizada arriba con respecto a la dirección de transcripción del sitio de iniciación de transcripción de la secuencia codificante y se identifica estructuralmente por la presencia de un sitio de unión para RNA-polimerasa que depende del ADN, sitios de iniciación de transcripción y cualquier otra secuencia de ADN, que incluya, pero no de forma limitada sitios de unión de factor de transcripción, represor y sitios de unión a proteínas activadoras, y cualquier otra secuencia de nucleótidos conocida por un experto en la técnica para actuar directamente o regular indirectamente la cantidad de transcripción del promotor.

Un promotor "constitutivo" es un promotor que es activo en la mayoría de tejidos bajo las mayores condiciones fisiológicas y de desarrollo.

Un promotor "inducible" es un promotor que es fisiológicamente o de desarrollo regulado, por ejemplo por la aplicación de un inductor químico.

Un promotor de "tejido específico" solo es activo en tipos específicos de tejidos o células.

5 [0017] "Identidad de secuencia" y "similaridad de secuencia" se puede determinar por alineamiento de dos péptidos o dos secuencias de nucleótidos que usan algoritmos de alineamiento globales o locales, dependiendo de la longitud de las dos secuencias.

Las secuencias de longitudes similares son preferiblemente alineadas utilizando unos algoritmos de alineamiento global (por ejemplo Needleman Wunsch) que alinean las secuencias óptimamente sobre la longitud entera, mientras
10 las secuencias de longitudes diferentes sustancialmente son preferiblemente alineadas utilizando un algoritmo de alineamiento local (por ejemplo Smith Waterman).

Las secuencias pueden luego denominarse "sustancialmente idénticas" o "esencialmente similares" cuando estas (cuando se han alineado óptimamente mediante por ejemplo los programas GAP o BESTFIT usando parámetros por defecto) comparten al menos un determinado porcentaje mínimo de identidad de secuencia (como se ha definido
15 abajo).

GAP usa el algoritmo de alineamiento global de Needleman y Wunsch para alinear dos secuencias sobre su longitud entera (longitud total), maximizando el número de coincidencias y minimizando el número de espacios.

Un alineamiento global se usa adecuadamente para determinar la identidad de secuencia cuando las dos secuencias tienen longitudes similares.

20 Generalmente, se usan los parámetros por defecto GAP, con una penalización de creación del espacio=50 (nucleótidos)/8 (proteínas) y penalización de extensión del espacio=3 (nucleótidos) / 2 (proteínas).

Para los nucleótidos de la matriz de puntuación por defecto usada es nwsgapdna y para las proteínas de la matriz de puntuación por defecto es Blossum62 (Henikoff & Henikoff, 1992, PNAS 89,915-919).

Los alineamientos de secuencia y puntuaciones para identidad de secuencia de porcentaje se pueden determinar utilizando los programas informáticos, tal como el paquete GCG Wisconsin, versión 10.3, disponible de Accelrys Inc., 9685 Scranton Road, San Diego, CA 92121-3752 USA o usando el software de fuente abierta, tal como el programa "needle" (utilizando el algoritmo global Needleman Wunsch) o "water" (utilizando el algoritmo local Smith Waterman) en EmbossWIN versión 2.10.0, usando los mismos parámetros que para GAP de arriba o utilizando los ajustes por defecto (ambos para 'needle' y 'water' y ambos para alineamientos de proteína y de ADN, la penalización de apertura de espacio predeterminada es 10.0 y la penalización por extensión de espacio predeterminada es 0.5; los matrices de puntuación por defecto son Blossum62 para proteínas y DNAFull para ADN).

30 Cuando las secuencias tienen unas longitudes diferentes sustancialmente en general, los alineamientos locales, tal como aquellos que utilizan el algoritmo de Smith Waterman, son preferidos.

Alternativamente, la similitud de porcentaje o identidad se puede determinar por la búsqueda contra las bases de datos públicas, usando algoritmos tal como FASTA, BLAST, etc.

[0018] Las secuencias de nucleótidos que codifican proteínas Cap y/o Rep parvovirales de la invención también se pueden definir por su capacidad para hibridar con las secuencias de nucleótidos de la SEC ID del nº 20, 22,24 y 1 - 4, respectivamente, bajo condiciones moderadas o preferiblemente bajo hibridación rigurosa.

40 [0019] Las condiciones de hibridación rigurosa se definen aquí como condiciones que permiten una secuencia de ácidos nucleicos de al menos aproximadamente 25, preferiblemente aproximadamente 50 nucleótidos, 75 o 100 y de la forma más preferible de aproximadamente 200 o más nucleótidos, para hibridar a una temperatura de aproximadamente 65°C en una solución que comprende aproximadamente 1 M sal, preferiblemente 6 x SSC o cualquier otra solución con una fuerza iónica comparable y lavada a 65°C en una solución que comprende aproximadamente 0.1 M sal, o menos, preferiblemente 0.2 x SSC o cualquier otra solución con una fuerza iónica comparable.

Preferiblemente, la hibridación se realiza durante toda la noche, es decir, al menos durante 10 horas y preferiblemente el lavado se realiza durante al menos una hora con al menos dos cambios de la solución de lavado.

50 Estas condiciones normalmente permitirán la hibridación específica de las secuencias con aproximadamente 90% o más identidad de secuencia.

[0020] Las condiciones moderadas se definen aquí como condiciones que permiten unas secuencias de ácidos nucleicos de al menos 50 nucleótidos, preferiblemente de aproximadamente 200 o más nucleótidos, para hibridar a una temperatura de aproximadamente 45°C en una solución que comprende aproximadamente 1 M sal, preferiblemente 6 x SSC o cualquier otra solución con una fuerza iónica comparable, y lavado a temperatura ambiente en una solución que comprende aproximadamente 1 M sal, preferiblemente 6 x SSC o cualquier otra solución con una fuerza iónica comparable.

60 Preferiblemente, la hibridación se realiza durante toda la noche, es decir, al menos durante 10 horas, y preferiblemente el lavado se realiza durante al menos una hora con al menos dos cambios de la solución de lavado.

Estas condiciones normalmente permitirán la hibridación específica de secuencias con hasta 50% de identidad de secuencia.

El experto en la técnica será capaz de modificar estas condiciones de hibridación para específicamente identificar secuencias variables en identidad entre 50% y 90%.

65

[0021] Un "vector" es una molécula de ácido nucleico (típicamente ADN o ARN) que sirve para transferir una secuencia de ácidos nucleicos pasajeros (es decir, ADN o ARN) en una célula huésped.

Tres tipos comunes de vectores incluyen plásmidos, fagos y virus.

Preferiblemente, el vector es un virus.

5 Los vectores que contienen ambos un promotor y un sitio de clonación en que un polinucleótido se puede enlazar operativamente se conocen en la técnica. Tales vectores son capaces de la transcripción ARN in vitro o in vivo, y están disponibles comercialmente de fuentes tal como Stratagene (La Jolla, Calif.) y Promega Biotech (Madison, Wis.).

10 Para optimizar la expresión y/o transcripción in vitro, puede ser necesario eliminar, añadir o alterar 5' y/o 3' partes no traducidas de los clones para eliminar codones de iniciación de traducción alternativos inapropiados extra potenciales u otras secuencias que pueden interferir con o reducir la expresión, bien en el nivel de transcripción o traducción.

Alternativamente, los sitios de unión de ribosoma de consenso se pueden insertar inmediatamente 5' del codón de inicio para mejorar la expresión.

15 [0022] Un "vector viral" se refiere a un vector que comprende alguno o todos los siguientes: genes víricos que codifican un producto génico, secuencias de control y secuencias de embalaje víricas.

20 [0023] Un "vector parvoviral" se define como un parvovirus recombinantemente producido o partícula parvoviral que comprende un polinucleótido para entregar en una célula huésped, bien in vivo, ex vivo o in vitro.

Ejemplos de vectores parvovirales incluyen por ejemplo, vectores de virus adeno-asociados.

Aquí, una construcción de vector parvoviral se refiere al polinucleótido que comprende el genoma vírico o parte del mismo, y un transgen.

25 [0024] La capacidad de adaptación de una secuencia de nucleótidos que codifica una enzima al uso de codón de una célula huésped se puede expresar como índice de adaptación de codón (CAI).

El índice de adaptación de codón se define aquí como una medición de la adaptación relativa del uso de codón de un gen hacia el uso de codón de genes expresados altamente en una célula huésped particular u organismo.

30 La capacidad de adaptación relativa (w) de cada codón es la proporción del uso de cada codón, a aquel del codón más abundante para el mismo aminoácido.

El índice CAI se define como la media geométrica de estos valores de capacidad de adaptación relativos.

Los codones no sinónimos y codones de terminación (dependiendo del código genético) se excluyen.

35 Los valores CAI varían de 0 a 1, con valores más altos que indican una proporción más alta de los codones más abundantes (ver Sharp y Li, 1987, Nucleic Acids Research 15: 1281-1295; también ver: Jansen et al, 2003, Nucleic Acids Res. 31(8):2242-51).

40 [0025] El término "indicador" (o gen reportero o proteína) se usa principalmente para referirse a secuencias de nucleótidos que codifican proteínas de marcador visibles, tales como (GFP), eGFP, otras proteínas fluorescentes, luciferasa, fosfatasa alcalina segregada (SEAP), GUS y similar, al igual que marcadores nptII y similares

Descripción detallada de la invención

45 [0026] La presente invención se refiere al uso de parvovirus, en particular dependovirus tal como humano infeccioso o AAV símico y sus componentes (por ejemplo, un genoma de parvovirus) para uso como vectores para introducción y/o expresión de ácidos nucleicos en células mamíferas, preferiblemente células humanas.

En particular, la invención se refiere a mejoras en la productividad de tales vectores parvovirales cuando se ha producido en las células de insecto.

50 [0027] La productividad en este contexto abarca mejoras en las valoraciones de producción y mejoras en la calidad del producto resultante, por ejemplo un producto que ha mejorado una proporción total:completa (una medida del número de partículas que comprende ácido nucleico).

Es decir, el producto final puede tener una proporción aumentada de partículas llenas, donde llenas implica que la partícula comprende ácido nucleico.

55 [0028] Virus de la familia Parvoviridae son virus de ADN pequeño.

La familia Parvoviridae se puede dividir entre dos subfamilias: la Parvovirinae, que infecta los vertebrados y la Densovirinae, que infecta invertebrados, incluyendo insectos.

Los miembros de la subfamilia Parvovirinae se refieren aquí como los parvovirus e incluyen el género Dependovirus.

60 Como se puede deducir del nombre de su género, los miembros del Dependovirus son únicos en el hecho de que estos normalmente requieren coinfección con un virus auxiliar tal como adenovirus o virus herpes para infección productiva en el cultivo celular.

El género Dependovirus incluye AAV, que infecta normalmente seres humanos (por ejemplo, serotipos 1, 2,3A, 3B, 4, 5 y 6) o primates (por ejemplo, serotipos 1 y 4), y virus relativos que infectan otros animales de sangre caliente (por ejemplo, bovino, canino, equino y virus ovinos adeno-asociados).

Otra información en parvovirus y otros miembros del Parvoviridae está descrita en Kenneth I. Bernas, "Parvoviridae: The Viruses and Their Replication," Chapter 69 in Fields Virology (3d Ed. 1996).

Por conveniencia, la presente invención es posteriormente ejemplificada y descrita aquí por referencia para AAV.

5 Sin embargo, se entiende que la invención no está limitada a AAV pero puede igualmente aplicarse a otros parvovirus.

[0029] La organización genómica de todos los serotipos AAV conocidos es muy similar.

El genoma de AAV es una molécula de ADN monocatenaria, lineal que es menor de aproximadamente 5,000 nucleótidos (nt) en longitud.

10 Las secuencias repetidoras terminales invertidas (ITR) flanquean las secuencias de nucleótidos codificantes únicas para las proteínas de replicación no estructurales (rep) y proteínas estructurales (VP).

Las proteínas de vasopresina (VP1; -2 y -3) forman la cápsida.

Los terminales 145 nt son autocomplementarios y son programados de modo que un dúplex intramolecular estable energéticamente que forma una horquilla en forma de T se puede formar.

15 Estas estructuras de horquilla funcionan como un origen para replicación de ADN vírico, que sirven como cebadores para el complejo de polimerasa de DNA celular.

La siguiente infección wtAAV en células mamíferas de los genes Rep (es decir Rep78 y Rep52) se expresa del P5 promotor y el P19 promotor, respectivamente y ambas proteínas Rep tienen una función en la replicación del genoma vírico.

20 Un evento de empalme en los Rep ORF produce la expresión de cuatro Proteína Reps reales (es decir, Rep78; Rep68; Rep52 y Rep40).

[0030] Sin embargo, se ha demostrado que el ARNm no empalmado, que codifica proteínas Rep78 y Rep52, en células mamíferas son suficientes para la producción de vector AAV.

25 También en las células de insecto, las proteínas Rep78 y Rep52 bastan para la producción de vector AAV.

[0031] Un "recombinante parvoviral o vector AAV" (o "vector rAAV") aquí se refiere a un vector que comprende una o más secuencias polinucleótidas de interés, genes de interés o "transgenes" que es/son flanqueados por al menos una secuencia de repetición de terminal parvoviral o AAV invertido (ITR). Preferiblemente, el transgen(s) es/son

30 flanqueados por ITR, uno a cada lado del transgen(s).

Tales vectores rAAV se pueden replicar y empaquetar en partículas víricas infecciosas en caso de existir en una célula huésped de insecto que exprese productos genéticos AAV rep y cap (es decir, proteínas AAV Rep y Cap).

35 Cuando un vector rAAV se incorpora en un constructo de ácidos nucleicos mayor (por ejemplo en un cromosoma o en otro vector tal como un plásmido o baculovirus usado para la clonación o transfección), luego el vector rAAV es típicamente referido como un "pro-vector" que puede ser "salvado" por replicación y encapsidación en presencia de las funciones de embalaje AAV y funciones auxiliares.

[0032] La invención se refiere a un método para la producción de un virión (rAAV) parvoviral recombinante, que comprende un vector (rAAV) parvoviral recombinante, en una célula de insecto.

40

[0033] En un primer aspecto, la invención se refiere a un método para la producción de un virión parvoviral recombinante que incluye las etapas de: (a) proporcionar una célula de insecto que comprende uno o más constructos de ácidos nucleicos que comprenden: (i) una secuencia de nucleótidos que comprende un transgén que se flanquea por al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral; (ii) un primer casete de expresión que comprende una secuencia de nucleótidos que codifica una proteína Rep parvoviral que está operativamente enlazada a un primer promotor que es capaz de la expresión de transmisión de la proteína Rep en la célula de insecto; (iii) un segundo casete de expresión que comprende una secuencia de nucleótidos que codifica una proteína cápsida parvoviral que está operativamente enlazada a un segundo promotor que es capaz de la expresión de transmisión de la proteína cápsida en la célula de insecto; (b) cultivo de la célula definida en (a) bajo una condición propicia para la expresión de la proteína Rep y cápsida; y, (c) opcionalmente, recuperación del virión parvoviral recombinante.

45

Preferiblemente, en la célula de insecto, los primeros y segundos casetes de expresión están presentes en un constructo de ácidos nucleicos único.

55 [0034] Preferiblemente, los primeros y segundos casetes de expresión, en caso de existir en la cantidad equimolar en una célula de insecto producen una proporción de niveles de ARNm que codifica la proteína Rep contra el ARNm que codifica la proteína cápsida de al menos aproximadamente 0.5, al menos aproximadamente 0.75, al menos aproximadamente 1.0, al menos aproximadamente 1.5, al menos aproximadamente 2.0, al menos aproximadamente 5.0 o al menos aproximadamente 10.0.

60 Los niveles de ARNm que codifican Rep y cápsida son preferiblemente determinados por transcripción inversa cuantitativa PCR, transferencia northern u otros medios para cuantificación de ARN conocidos en la técnica.

[0035] Preferiblemente, la proporción de los niveles de ARNm que codifican Rep y cápsida se producen en la célula de insecto en el momento de 24, 30, 36, 40,44 o 46 horas después de la transfección a 72, 66, 60, 54, 52 o 50 horas

65 después de la transfección.

Más preferiblemente la proporción de los niveles de ARNm que codifican Rep y cápsida son al menos producidos en la célula de insecto 2 o 1 hora alrededor de 48 horas después de la transfección.

El marco de lectura puede ser capaz de la codificación de más de una proteína, por ejemplo dos, tres o cuatro proteínas.

5 [0036] Alternativamente preferidos los primeros y segundos casetes de expresión, en caso de existir en la cantidad equimolar en una célula de insecto producen una proporción de niveles de Proteína Rep contra proteína cápsida de al menos aproximadamente 0.5, al menos aproximadamente 0.75, al menos aproximadamente 1.0, al menos aproximadamente 1.5, al menos aproximadamente 2.0, al menos aproximadamente 5.0 o al menos aproximadamente 10.0.

Los niveles de proteína Rep y cápsida son preferiblemente determinados usando anticuerpos de referencia contra las proteínas Rep y cápsida en un inmunoensayo cuantitativo (por ejemplo ELISA o transferencia Western cuantitativa).

15 Preferiblemente, la proporción de los niveles de proteína Rep y cápsida se producen en la célula de insecto en el momento de 24, 30, 36, 40,44 o 46 horas después de la transfección a 72, 66, 60, 54, 52, o 50 horas después de la transfección.

Más preferiblemente la proporción de los niveles de proteína Rep y cápsida se producen en la célula de insecto 2 o 1 horas alrededor de 48 horas después de la transfección.

20 Los anticuerpos de referencia adecuados para cuantificación de niveles de proteína AAV Rep y cápsida son p. ej Anti-AAV-rep ratón monoclonal, clon 303.9 o anti-AAV VP1/VP2/VP3, ratón monoclonal, clon B1 obtenible de PROGEN Biotechnik GmbH.

[0037] En el método de la invención, los primeros y segundos casetes de expresión están presentes en un constructo de ácidos nucleicos único.

25 Una de las ventajas de ambos casetes de expresión que están presentes en un constructo de ácidos nucleicos único es que las células de insecto transfectado expresarán ambas Proteína Rep y proteína cápsida.

Solo la expresión de ambas Proteína Rep y proteína cápsida en una célula de insecto dará lugar a la formación de viriones parvovirales.

30 Otra ventaja es que la proporción requerida mínima de la expresión de la Proteína Rep contra la proteína cápsida se puede controlar cuando los primeros y segundos casetes de expresión son en una construcción única.

[0038] Es una forma de realización de la invención que otro constructo de ácidos nucleicos que comprende una secuencia de nucleótidos que codifica un Proteína Rep también puede estar presente en la célula.

35 [0039] En un método de la invención, la secuencia de nucleótidos del primer casete de expresión (ii) puede preferiblemente comprender solo una codificación de secuencias de nucleótidos que comprenden el marco de lectura abierto al menos una de las proteínas Rep78 y Rep 68.

Es decir, la proteína Rep o proteínas serán codificadas por un marco de lectura abierto único, no por dos o más marcos de lectura abiertos.

40 Para poner otra vía, el primer casete de expresión típicamente no tendrá dos o más marcos de lectura abiertos separados que codifican las mismas o diferentes Proteína Reps.

Para la evitar la duda, sin embargo, un marco de lectura abierto único puede ser capaz de la codificación de más de una proteína, por ejemplo dos, tres o cuatro proteínas.

45 [0040] Todo lo anterior se puede aplicar igualmente a las VP1, VP2 y VP3 proteínas, es decir dos o todas aquellas proteínas pueden convenientemente todas ser codificadas por un marco de lectura abierto único.

[0041] Típicamente, en un método de la invención, al menos unas secuencias de nucleótidos que comprenden el marco de lectura abierto que codifica las VP1, VP2 y VP3 proteínas cápsidas o al menos un marco de lectura abierto que incluye una codificación de secuencias de nucleótidos que comprende el marco de lectura abierto en al menos una de las proteínas Rep78 y Rep68 no comprende un intrón artificial (o una secuencia derivada a partir de un intrón artificial).

Es decir, al menos marcos de lectura abiertos usados para codificar Rep o proteínas de vasopresina no comprenderán un intrón artificial.

55 Por intrón artificial se entiende un intrón que no se produciría naturalmente en un virus adeno-asociado Rep o secuencia Cap, por ejemplo un intrón que ha sido diseñado para permitir un empalme funcional dentro de una célula de insecto.

Un intrón artificial en este contexto por lo tanto abarca intrones de célula de insecto tipo salvaje.

60 Un casete de expresión de la invención puede comprender una secuencia de intrón truncada nativa (por nativa se entiende una secuencia de origen natural en un virus adeno-asociado) tales secuencias no se destinan a incluirse en el significado de intrón artificial tal y como se define aquí.

[0042] En la invención, una posibilidad es que ninguna secuencia de nucleótidos que comprenda el marco de lectura abierto que codifica las proteínas cápsidas VP1, VP2 y VP3 y/o ningún marco de lectura abierto que comprenda la codificación de secuencias de nucleótidos que comprenden al menos una de las proteínas Rep78 y Rep68 comprende un intrón artificial.

[0043] En una forma de realización preferida de la invención, la secuencia de nucleótidos que comprende el transgen está también en el mismo constructo de ácidos nucleicos como primeros y segundos casetes de expresión. Una ventaja de la misma es, que todas las células transfectadas comprenden cada una de las tres secuencias de nucleótidos, que se necesitan para la producción de virión parvoviral.

[0044] Además, los presentes inventores descubrieron que la mayor proporción de proteína Rep:proteína cápsida, mayor será la proporción de virión completo vs. virión vacío.

El término "virión completo" se refiere a una partícula de virión que comprende un vector parvoviral.

El término "virión vacío" se refiere a una partícula de virión que no comprende un vector parvoviral.

En una forma de realización preferida de la invención, la proporción de virión completo vs. virión vacío es al menos 1:100 más preferiblemente al menos 1:10 e incluso más preferiblemente al menos 1:1.

Aún más preferiblemente, ningunos viriones vacíos se pueden detectar y de la forma más preferible no están presentes viriones vacíos.

El experto en la técnica sabrá determinar la proporción de virión completo vs virión vacío, por ejemplo por la división del número de copias de gen por (cápsida total - número de copias de genoma), ya que por virión habrá solo una copia presente en el genoma.

Inversamente, la mayor proporción proteína Rep:proteína cápsida, la menor proporción de viriones totales vs viriones vacíos (es decir, viriones llenos+vacíos).

El experto en la materia sabrá cómo determinar tal proporción.

Por ejemplo, la proporción de viriones vacíos vs. cápsidas totales se puede determinar por la división de la cantidad de copias de genoma (es decir número de copias de genoma) por la cantidad de partículas parvovirales total (es decir número de partículas parvovirales), donde la cantidad de copias de genoma por ml se mide por PCR cuantitativo y la cantidad de partículas parvovirales totales por ml se mide con un enzimoimmunoanálisis, por ejemplo de Progen.

En una forma de realización preferida de la invención, la proporción de viriones totales vs. viriones vacíos es inferior a 100:1, más preferiblemente menos de 10:1 e incluso más preferiblemente menos de 1:1.. Aún más preferiblemente, ningún virión vacío puede ser detectado y de la forma más preferible ningún virión vacío está presente.

[0045] En una forma de realización preferida, la proporción de expresión del Rep contra la proteína cápsida se regula por uno o más de los siguientes: (a) el primer promotor es fuerte igualmente o más fuerte que el segundo promotor, como se ha determinado por la expresión de gen reportero (por ejemplo luciferasa o SEAP), o transferencia Northern; (b) la presencia de más y/o elementos intensificadores más fuertes en el primer casete de expresión en comparación con el segundo casete de expresión; (c) la secuencia de nucleótidos codificante para la proteína Rep parvoviral tiene un índice de adaptación de codón más alto en comparación con la secuencia de nucleótidos codificante para la proteína cápsida; (d) optimización de temperatura de la proteína Rep parvoviral; y/o (e) proteínas Rep variantes con una o más alteraciones en la secuencia de aminoácidos en comparación con una proteína Rep tipo salvaje correspondiente y donde uno o más de la alteración de aminoácido produce el aumento en la actividad de la función Rep como se ha evaluado por la detección de la producción AAV aumentada en las células de insecto. Los métodos para la generación, selección y/o selección de variantes de proteínas Rep con actividad aumentada de la función Rep como se ha evaluado por la detección de la producción AAV aumentada en las células de insecto se pueden obtener por adaptación a células de insecto de los métodos descritos en la US20030134351 para obtener proteínas Rep variantes con función aumentada respecto a la producción AAV en células mamíferas.

La variante de Proteína Reps con una o más alteraciones en la secuencia de aminoácidos en comparación con una proteína tipo salvaje correspondiente Rep se entienden aquí para incluir Proteína Reps con una o más sustituciones de aminoácidos, inserciones y/o deleciones en la secuencia de aminoácidos variante en comparación con la secuencia de aminoácidos de una proteína Rep tipo salvaje correspondiente.

[0046] El primer promotor es igualmente fuerte o más fuerte que el segundo promotor significa que en caso de más secuencias de nucleótidos que codifican una Proteína Rep que las secuencias de nucleótidos que codifican una proteína cápsida, se puede utilizar igualmente un promotor fuerte, ya que la expresión de Proteína Rep luego aumentará en comparación con la expresión de proteína cápsida, mientras que en caso de cantidades similares de secuencias de nucleótidos que codifican Rep y codifican proteína cápsida, un promotor más fuerte se puede utilizar para la expresión de Proteína Rep que para la expresión de proteína cápsida.

La fuerza del promotor se puede determinar por la expresión que se obtiene bajo condiciones que se usan en el método de la invención.

En una forma de realización preferida, el primer promotor o el segundo promotor se selecciona del grupo que consiste en un Promotor PolH, promotor p10, promotor de proteína básica, un promotor inducible o un promotor deltaE1 o un Promotor E1, o cualquier otro promotor de gen de Baculovirus tardío o muy tardío.

Más preferiblemente, el primer promotor se selecciona del grupo que consiste en un Promotor PolH, promotor p10 o promotor de proteína básica y donde el segundo promotor es un promotor deltaE1 o un Promotor E1, o cualquier otro promotor de gen de baculovirus temprano o tardío.

Preferiblemente, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor p10 y el segundo promotor es un promotor PolH o un promotor 4xHsp27 EcRE+minimal Hsp70.

En otra forma de realización, el primer promotor en el constructo de ácidos nucleicos de la invención es un 4xHsp27 EcRE+minimal Hsp70 promotor y el segundo promotor es un Promotor PolH.

En otra forma de realización, el primer promotor en el constructo de ácidos nucleicos de la invención es un Promotor PolH y el segundo promotor es un p10, un deltaE1 o un Promotor E1.

5 En otra forma de realización, el primer promotor en el constructo de ácidos nucleicos de la invención es un Promotor PolH y el segundo promotor es un deltaE1 o un Promotor E1.

En otra forma de realización, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor p10 y el segundo promotor es un deltaE1 o un Promotor E1.

10 En otra forma de realización, el primer promotor en el constructo de ácidos nucleicos de la invención es un Promotor PolH y el segundo promotor es un Promotor PolH.

De la forma más preferible, el primer promotor en el constructo de ácidos nucleicos de la invención es un Promotor PolH y el segundo promotor es un promotor deltaE1.

15 [0047] Un "elemento intensificador" o "intensificador" se entiende que define una secuencia que intensifica la actividad de un promotor (es decir aumenta el índice de transcripción de una secuencia aguas abajo del promotor) que, a diferencia de un promotor, no posee actividad promotora, y que puede normalmente funcionar sin tener en cuenta su ubicación con respecto al promotor (es decir aguas arriba o abajo del promotor).

Los elementos intensificadores son bien conocidos en la técnica. Ejemplos no limitativos de elementos intensificadores (o partes de los mismos) que podrían ser usados en la presente invención incluyen potenciadores de baculovirus y elementos intensificadores encontrados en células de insecto.

20 Se prefiere que el elemento intensificador aumente en una célula la expresión de ARNm de un gen, al que el promotor está operativamente enlazado, por al menos 25%, más preferiblemente al menos 50%, aún más preferiblemente al menos 100% y de la forma más preferible al menos 200% en comparación con la expresión de ARNm del gen en ausencia del elemento intensificador. La expresión de ARNm se puede determinar por ejemplo por RT-PCR cuantitativo.

[0048] Aquí se prefiere usar un elemento intensificador para mejorar la expresión de la proteína Rep parvoviral. Así, en otra forma de realización preferida, el primer casete de expresión comprende al menos un elemento intensificador de baculovirus y/o al menos un elemento sensible ecdisona.

30 Preferiblemente el elemento intensificador es seleccionado del grupo que consiste en hr1, hr2, hr3, hr4 y hr5.

[0049] La optimización de codón de la proteína Rep parvoviral se discute con más detalle de aquí en adelante.

35 [0050] La optimización de temperatura de la proteína Rep parvoviral se refiere al uso de la condición óptima con respecto a la temperatura a la que la célula de insecto crecerá y Rep está en funcionamiento.

Una proteína Rep puede por ejemplo ser activa óptimamente a 37°C, mientras que una célula de insecto puede crecer óptimamente a 28°C.

Una temperatura a la que la proteína Rep es activa y la célula de insecto crece puede ser 30°C.

40 En una forma de realización preferida, la temperatura optimizada es más del 27, 28, 29, 30, 31, 32, 33,34 o 35°C y/o menos de 37, 36, 35, 34, 33, 32, 31,30 o 29°C.

[0051] Como se entenderá por la persona experta en la técnica, la proporción virión completo:virión vacío también se puede mejorar por la expresión Cap atenuada, por ejemplo mediante un promotor más débil, en comparación con la expresión Rep moderada a alta.

45 [0052] Preferiblemente un constructo de ácidos nucleicos de la invención, es un vector compatible con célula de insecto.

Por "vector compatible con la célula de insecto" o "vector" se entiende una molécula de ácido nucleico capaz de la transformación productiva o transfección de un insecto o célula de insecto.

50 Los vectores biológicos ejemplares incluyen plásmidos, moléculas lineales de ácido nucleico y virus recombinantes.

Cualquier vector se puede emplear en tanto que sea compatible con la célula de insecto.

El vector se puede integrar en el genoma de células de insecto pero la presencia del vector en la célula de insecto no necesita ser permanente y los vectores episómicos transitorios también se incluyen.

55 Los vectores se pueden introducir por cualquier medio conocido, por ejemplo por tratamiento químico de las células, electroporación o infección.

En una forma de realización preferida, el vector es un baculovirus, un vector vírico o un plásmido.

En una forma de realización más preferida, el vector es un baculovirus, es decir el constructo de ácidos nucleicos es un vector de expresión de baculovirus.

60 Los vectores de expresión de baculovirus y métodos para su uso son descritos por ejemplo en Summers y Smith. 1986. A Manual of Methods for Baculovirus Vectors and Insect Culture Procedures, Texas Agricultural Experimental Station Bull. No. 7555, College Station, Tex.; Luckow. 1991. En Prokop et al., Cloning and Expression of Heterologous Genes in Insect Cells with Baculovirus Vectors' Recombinant DNA Technology and Applications, 97-152; King, L. A. and R. D. Possee, 1992, The baculovirus expression system, Chapman and Hall, United Kingdom; O'Reilly, D. R., L. K. Miller, V. A. Luckow, 1992, Baculovirus Expression Vectors: A Laboratory Manual, New York; W.

H. Freeman and Richardson, C. D., 1995, Baculovirus Expression Protocols, Methods in Molecular Biology, volume 39; US 4,745,051; US2003148506; and WO 03/074714.

[0053] El número de constructos de ácidos nucleicos empleado en la célula de insecto para la producción del vector (rAAV) parvoviral recombinante no es limitativo en la invención.

Por ejemplo, uno, dos, tres o más constructos separados se pueden emplear para producir rAAV en las células de insecto conforme a los métodos de la presente invención.

Si se emplean dos constructos, una construcción puede comprender la secuencia de nucleótidos que comprende el transgen que se flanquea por al menos una secuencia ITR parvoviral y la otra construcción puede luego comprender un primero y un segundo casete de expresión.

Si tres constructos son empleados, un constructo puede comprender la secuencia de nucleótidos que comprende el transgen que se flanquea por al menos una secuencia ITR parvoviral, otro constructo puede comprender los primeros y segundos casetes de expresión y todavía otro constructo puede comprender una secuencia de nucleótidos adicional que codifica una proteína Rep, opcionalmente bien el codón optimizado, AT optimizado o GC optimizado, para minimizar o prevenir la recombinación, como se describe de ahora en adelante.

[0054] Una secuencia de nucleótidos que codifica proteínas Rep parvovirales, se entiende aquí como una secuencia de nucleótidos que codifica las proteínas Rep no estructurales que se requieren y son suficientes para la producción de vector parvoviral en las células de insecto tal como las proteínas Rep78 o Rep68, y/o el Rep52 o Rep40.

La secuencia de nucleótidos de parvovirus es preferiblemente a partir de un dependovirus, más preferiblemente de un virus humano o simio adeno-asociado (AAV) y de la forma más preferible a partir de un AAV que infecta normalmente averses humanos (por ejemplo, serotipos 1, 2,3A, 3B, 4, 5, y 6) o primates (por ejemplo, serotipos 1 y 4). Un ejemplo de una secuencia de nucleótidos que codifica proteínas parvovirus Rep se da en SEQ ID N.º 5, que representa una parte del genoma de secuencia AAV serotipo 2 que codifica las proteínas Rep.

La Rep78 secuencia codificante comprende nucleótidos 11 - 1876 y la Rep52 secuencia codificante comprende nucleótidos 683 - 1876, representada también separadamente en la SEC ID No.5 y 7.

Se entiende que los pesos moleculares exactos de las proteínas Rep78 y Rep52, al igual que las posiciones exactas de los codones de iniciación de traducción pueden diferir entre parvovirus diferentes.

Sin embargo, la persona experta sabrá identificar la posición correspondiente en la secuencia de nucleótidos de otros parvovirus que AAV-2.

[0055] En una forma de realización preferida, el primer casete de expresión comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos de una proteína Rep parvoviral52 o 40 y/o una secuencia de nucleótidos que codifica una secuencia de aminoácidos de una proteína Rep parvoviral78 o 68.

Preferiblemente, las proteínas Rep parvovirales son proteínas Rep de virus adeno-asociado (AAV).

[0056] En una forma de realización preferida, la invención se refiere a una célula de insecto que comprende no más de un tipo de secuencia de nucleótidos que comprende un marco de lectura abierto único que codifica una proteína Rep parvoviral.

Preferiblemente el marco de lectura abierto único codifica una o más de las proteínas Rep parvovirales, más preferiblemente el marco de lectura abierto codifica todas las proteínas Rep parvovirales, de la forma más preferible el marco de lectura abierto codifica la proteína en toda su longitud Rep 78 donde preferiblemente al menos ambas proteínas Rep 52 y Rep 78 se pueden expresar en la célula de insecto.

Se entiende aquí que la célula de insecto puede comprender más de una copia del tipo único de secuencia de nucleótidos, por ejemplo en un vector episómico multicopia, pero que estos son copias múltiples de esencialmente una y la misma molécula de ácido nucleico, o al menos moléculas de ácido nucleico que codifican una y la misma Rep secuencia de aminoácidos, por ejemplo moléculas de ácido nucleico que solo difieren entre sí debido a la degeneración del código genético.

La presencia de solo un tipo único de molécula de ácido nucleico que codifica las proteínas Rep parvovirales evita la recombinación entre las secuencias homólogas como pueden estar presentes en diferentes tipos de vectores que comprenden Rep secuencias, que pueden dar lugar a constructos de expresión Rep defectuosos que afecten a (la estabilidad de) niveles de producción parvovirales en las células de insecto.

[0057] En una forma de realización alternativa de la invención, el primer casete de expresión comprende más de una secuencia de nucleótidos que codifica una proteína Rep parvoviral.

Se prefiere que la secuencia de nucleótidos de (ii) comprenda un marco de lectura abierto que comprende secuencias de nucleótidos que codifican al menos una de las proteínas Rep78 y Rep68.

Preferiblemente, las secuencias de nucleótidos son del mismo serotipo.

Más preferiblemente, las secuencias de nucleótidos difieren entre sí en que estas pueden ser bien codón optimizado, AT optimizado o GC optimizado, para minimizar o prevenir la recombinación.

Preferiblemente, el primer casete de expresión comprende dos secuencias de nucleótidos que codifican una proteína Rep parvoviral, es decir, una primera secuencia de nucleótidos y una segunda secuencia de nucleótidos.

Preferiblemente, la diferencia en el primer y la segunda secuencia de nucleótidos codificante para las secuencias de aminoácidos comunes de una proteína Rep parvoviral se maximiza (es decir, la identidad de nucleótido se minimiza) por uno o más de: a) cambio de la preferencia codónica de la primera secuencia de nucleótidos codificante para la secuencia de aminoácidos común Rep parvoviral; b) cambio de la preferencia codónica de la segunda secuencia de

nucleótidos codificante para la secuencia de aminoácidos común Rep parvoviral; c) cambio del contenido GC de la primera secuencia de nucleótidos que codifica la secuencia de aminoácidos común; y d) cambio del contenido GC de la segunda secuencia de nucleótidos que codifica la secuencia de aminoácidos común.

La optimización de codón se puede realizar basándose en el uso de un codón de la célula de insecto usada en el método de la invención, preferiblemente *Spodoptera frugiperda*, como puede ser descubierta en una base de datos de uso de codón (ver por ejemplo <http://www.kazusa.or.jp/codon/>).

Los programas informáticos adecuados para la optimización de codón están disponibles para la persona experta (ver por ejemplo Jayaraj et al., 2005, Nucl. Acido Res. 33(9):3011-3016; y en internet).

Alternativamente, las optimizaciones pueden hacerse a mano, usando la misma base de datos de uso de codón.

[0058] Una secuencia de nucleótidos del primer casete de expresión que codifica una proteína Rep52 parvoviral se puede definir como una secuencia de nucleótidos:

- a) que codifica un polipéptido que incluye una secuencia de aminoácidos que tiene al menos 50, 60, 70, 80, 88, 89, 90, 95, 97, 98, o 99% de identidad de secuencia con la secuencia de aminoácidos de la SEC ID No: 6
- b) que tiene al menos 50, 60, 70, 80, 81, 82, 85, 90, 95, 97, 98, o 99% de identidad de secuencia con la secuencia de nucleótidos de cualquier SEC ID del nº 1 - 5;
- c) la cadena complementaria de la cual hibrida a una secuencia de molécula de ácido nucleico de (a) o (b);
- d) secuencias de nucleótidos la secuencia de las cuales difiere de la secuencia de una molécula de ácido nucleico de (c) debido a la degeneración del código genético.

[0059] Una secuencia de nucleótidos del primer casete de expresión que codifica una proteína Rep78 parvoviral se puede definir como una secuencia de nucleótidos:

- a) que codifica un polipéptido que incluye una secuencia de aminoácidos que tiene al menos 50, 60, 70, 80, 88, 89, 90, 95, 97, 98, o 99% de identidad de secuencia con la secuencia de aminoácidos de la SEC ID No: 8
- b) que tiene al menos 50, 60, 70, 80, 81, 82, 85, 90, 95, 97, 98, o 99% de identidad de secuencia con la secuencia de nucleótidos de posiciones 11 - 1876 de la SEC ID No: 7
- c) la cadena complementaria de la cual hibrida a una secuencia de molécula de ácido nucleico de (a) o (b);
- d) secuencias de nucleótidos la secuencia de las cuales difiere de la secuencia de una molécula de ácido nucleico de (c) debido a la degeneración del código genético.

Preferiblemente, la secuencia de nucleótidos codifica proteínas parvovirus Rep que se requieren y son suficientes para la producción de vector parvoviral en las células de insecto.

[0060] La eliminación de sitios de iniciación de traducción falsa posible en las secuencias codificantes de proteína Rep, diferentes de los sitios de iniciación de traducción Rep78 y Rep52, de otros parvovirus la entenderá bien un experto en la técnica, como será la eliminación de sitios de empalme putativo que se puede reconocer en las células de insecto.

[0061] En una forma de realización preferida, el codón iniciador para traducción de la proteína Rep78 parvoviral es un codón iniciador subóptimo.

El codón iniciador subóptimo es preferiblemente un codón iniciador que proporciona salto de exón parcial. Salto de exón parcial se entiende aquí que significa que al menos una parte de los ribosomas no inician la traducción en el codón iniciador subóptimo de la proteína Rep78 pero en un codón iniciador más aguas abajo, por lo cual preferiblemente el codón iniciador más aguas abajo es el codón iniciador de la Rep52 proteína.

El codón iniciador subóptimo efectúa preferiblemente un salto de exón parcial en la expresión de la secuencia de nucleótidos en una célula de insecto.

Preferiblemente, el codón iniciador subóptimo efectúa el salto de exón parcial en una célula de insecto para producir en la célula de insecto una proporción molar de Rep78 a Rep52 en el rango de 1:10 a 10:1, 1:5 a 5:1, o 1:3 a 3:1, preferiblemente en alrededor de 20 - 40 horas después de la infección, más preferiblemente en alrededor de 30 - 40 horas después de la infección, utilizando una expresión de baculovirus.

La ración molar de la Rep78 y Rep52 se puede determinar mediante método de transferencia Western, preferiblemente utilizando un anticuerpo monoclonal que reconoce un epítipo común tanto de Rep78 como de Rep52 o usando por ejemplo un ratón anti-rep anticuerpo (303.9, Progen, Germany; dilution 1:50).

[0062] El término "codón de iniciación subóptimo" aquí no solo se refiere al codón de inicio de trinucleótido en sí mismo sino también a su contexto.

Así, un codón iniciador subóptimo puede consistir en un codón ATG "óptimo" en un contexto subóptimo, por ejemplo un contexto no Kozak.

Sin embargo, más preferidos son los codones de iniciación subóptima donde el codón de inicio de trinucleótido mismo es subóptimo, es decir no es ATG.

Subóptimo se entiende aquí que significa que el codón es menos eficaz en la iniciación de traducción en un contexto idéntico de otro modo en comparación con el codón ATG normal.

Preferiblemente, la eficiencia de codón subóptimo es inferior a 90, 80, 60,40 o 20% de la eficiencia del codón ATG normal en un contexto idéntico de otro modo.

Los métodos para la comparación de la eficiencia relativa de la iniciación de traducción son conocidos de por sí para la persona experta.

Los codones de iniciación subóptima preferidos se pueden seleccionar de ACG, TTG, CTG y GTG.
Más preferido es ACG.

Una secuencia de nucleótidos que codifica roteína Rep parvovirus se entiende aquí como una secuencia de nucleótidos que codifica las proteínas no estructurales Rep que se requieren y son suficientes para la producción de vector parvoviral en las células de insecto tales como las proteínas Rep78 y Rep52.

[0063] También es posible reemplazar otros ATG en la secuencia con un codón diferente que codifica metionina de modo que la posibilidad menor de la iniciación de traducción de tal ATG.

[0064] Varias modificaciones de la secuencia de nucleótidos de codificación tal como se ha definido anteriormente, incluyendo por ejemplo las secuencias parvovirales de tipo salvaje, para la expresión apropiada en las células de insecto se consigue por aplicación de técnicas de ingeniería genéticas bien conocidas tales como las descritas por ejemplo en Sambrook y Russell (2001) "Molecular Cloning: A Laboratory Manual (3rd edition), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, New York. Varias otras modificaciones de las regiones de codificación se conocen por el experto en la materia, lo que podría aumentar el rendimiento de las proteínas de codificación. Estas modificaciones se incluyen en el campo de la presente invención.

[0065] Una secuencia de nucleótidos que codifica una proteína (Cap) cápsida parvoviral se entiende aquí como que comprende las secuencias de nucleótidos que codifican una o más de las tres proteínas cápsidas parvovirales, VP1; -2 y -3.

La secuencia de nucleótidos de parvovirus es preferiblemente a partir de un dependovirus, más preferiblemente de un virus humano o símico adeno-asociado (AAV) y de la forma más preferible de un AAV que infecta normalmente seres humanos (por ejemplo, serotipos 1, 2,3A, 3B, 4, 5 y 6) o primates (por ejemplo, serotipos 1 y 4).

Ejemplos de una secuencia de nucleótidos que codifica proteínas cápsidas de parvovirus se da en SEQ ID N.º 20,22 y 24.

[0066] En una forma de realización preferida la secuencia de nucleótidos de (iii) comprende un marco de lectura abierto que comprende secuencias de nucleótidos que comprenden las proteínas cápsidas VP1, VP2 y VP3.

Las secuencias codificantes de proteína cápsida pueden estar presentes en varias formas.

Por ejemplo las secuencias codificantes separadas para cada una de las proteínas cápsidas VP1; -2 y -3 se pueden usar, por lo cual cada una de las secuencias codificantes está operativamente enlazada a las secuencias de control de expresión para la expresión en una célula de insecto.

Más preferiblemente, sin embargo, el segundo casete de expresión comprende una secuencia de nucleótidos que comprende un marco de lectura abierto único que codifica todos los tres (AAV) parvoviral VP1, VP2 y VP3 proteínas cápsidas, donde el codón iniciador para la traducción de la VP1 proteína cápsida es un codón iniciador subóptimo que no es ATG como por ejemplo descrito por Urabe et al. (2002; supra) y en WO2007/046703.

Un codón iniciador subóptimo para la VP1 proteína cápsida puede ser tal como se ha definido anteriormente para la proteína Rep78.

Los codones de iniciación subóptima más preferidos para la VP1 proteína cápsida se pueden seleccionar de ACG, TTG, CTG y GTG, de los cuales CTG y GTG son más preferidos.

La secuencia de nucleótidos comprendida en el segundo casete de expresión para la expresión de las proteínas cápsidas puede comprender además una o más modificaciones como se describe en WO2007/046703.

Varias otras modificaciones de regiones de codificación de vasopresina las conoce el experto en la materia que podría bien aumentar la producción de vasopresina y virión o tener otros efectos deseados, tal como tropismo alterado o reducir la antigenicidad del virión.

Estas modificaciones están dentro del campo de la presente invención.

[0067] En una forma de realización preferida, la expresión de VP1 se aumenta en comparación con la expresión de VP2 y VP3.

La VP1 expresión se puede aumentar por suplementación de VP1, por introducción en la célula de insecto de un vector de insecto único que comprende las secuencias de nucleótidos para el VP1 como ha sido descrito en la WO 2007/084773.

[0068] En el contexto de la invención "al menos una secuencia de nucleótido de repetición terminal invertida parvoviral" se entiende que significa una secuencia palíndroma, que comprende secuencias dispuestas simétricamente, complementarias en su mayoría también referidas como regiones "A", "B" y "C".

Las funciones ITR como un origen de replicación, un sitio con un papel "cis" en la replicación, es decir, siendo un sitio de reconocimiento para proteínas de replicación de actuación trans tal como por ejemplo Rep 78 (o Rep68) que reconocen el palíndromo y secuencias específicas internas al palíndromo.

Una excepción a la simetría de la secuencia ITR es la "D" región del ITR.

Es único (sin un complemento dentro de un ITR).

Mellar el ADN monocatenario ocurre en la juntura entre las regiones A y D.

Es la región donde se inicia la nueva síntesis de ADN.

La D región yace normalmente a un lado del palíndromo y proporciona direccionalidad al paso de replicación de ácido nucleico.

Un parvovirus que replica en una célula mamífera típicamente tiene dos secuencias ITR.

Sin embargo, es posible para el ingeniero un ITR, de modo que los sitios de unión estén en ambos hilos de las regiones A y D se sitúan simétricamente, en cada lado del palíndromo.

5 En un molde de ADN circular bicatenario (por ejemplo, un plásmido), la replicación de ácido nucleico Rep78- o Rep68-asistida luego procede en ambas direcciones y un ITR único basta para la replicación parvoviral de un vector circular. Así, una secuencia de nucleótidos ITR se puede usar en el contexto de la presente invención.

Preferiblemente, sin embargo, se usan dos o incluso otro número de ITR regulares.

De la forma más preferible, se usan dos secuencias ITR.

Un ITR parvoviral preferido es un ITR AAV.

10 Por razones de seguridad puede ser deseable una construcción de un vector (rAAV) parvoviral recombinante que es incapaz de propagar adicionalmente después de la introducción inicial en una célula en presencia de un segundo AAV.

Tal mecanismo de seguridad para limitar la propagación de vector indeseable en un receptor se puede proporcionar usando rAAV con un ITR quimérico como se describe en US2003148506.

15 [0069] El término "flanqueado" respecto a una secuencia que se flanquea por otro elemento(s) aquí indica la presencia de uno o más de los elementos flanqueantes aguas arriba y/o abajo, es decir, 5' y/o 3', relativamente a la secuencia.

El término "flanqueado" no se destina a indicar que las secuencias son contiguas necesariamente.

20 Por ejemplo, puede haber secuencias de intervención entre el ácido nucleico que codifica el transgen y un elemento flanqueante.

Una secuencia que se "flanquea" por otros dos elementos (por ejemplo ITR), indican que un elemento está localizado 5' a la secuencia y el otro está localizado 3' a la secuencia; sin embargo, puede haber secuencias de intervención entre ellos.

25 En una forma de realización preferida una secuencia de nucleótidos de (i) se flanquea en cada lado por secuencias de nucleótidos terminales invertidas de repetición parvoviral.

[0070] En las formas de realización de la invención, la secuencia de nucleótidos que comprende el transgen (que codifica un producto génico de interés) que se flanquea por al menos una secuencia ITR parvoviral preferiblemente se vuelve incorporada en el genoma de un vector (rAAV) parvoviral recombinante producido en la célula de insecto.

Preferiblemente, el transgen codifica un producto génico de interés para la expresión en una célula mamífera.

Preferiblemente, la secuencia de nucleótidos que comprende el transgen se flanquea por dos secuencias de nucleótidos ITR (AAV) parvoviral y donde el transgen está localizado entremedias de dos secuencias de nucleótidos ITR (AAV) parvoviral.

35 Preferiblemente, la secuencia de nucleótidos que codifica un producto génico de interés (para la expresión en la célula mamífera) será incorporada en el vector (rAAV) parvoviral recombinante producido en la célula de insecto si está situada entre dos ITR regulares o se localiza en cada lado de un diseñado ITR con dos D regiones.

[0071] Las secuencias AAV que se pueden utilizar en la presente invención para la producción de un virión AAV recombinante en las células de insecto se pueden derivar del genoma de cualquier serotipo AAV.

Generalmente, los serotipos AAV tienen secuencias genómicas de homología significativa en los niveles de aminoácido y de ácido nucleico, proporcionan un conjunto idéntico de funciones genéticas, producen viriones que son esencialmente físicamente y funcionalmente equivalentes, y se replican y ensamblan por mecanismos prácticamente idénticos.

45 Para la secuencia genómica de los varios serotipos AAV y una visión de conjunto de las similitudes genómicas ver por ejemplo el número de registro GenBank U89790, número de registro GenBank J01901, número de registro GenBank AF043303, número de registro GenBank AF085716 Chlorini et al. (1997, J. Vir. 71: 6823-33); Srivastava et al. (1983, J. Vir.45:555-64); Chlorini et al. (1999, J. Vir. 73:1309-1319); Rutledge et al. (1998, J.Vir. 72:309-319); y Wu et al. (2000, J. Vir. 74: 8635-47). Serotipos AAV 1, 2, 3, 4 y 5 son una fuente preferida de secuencias de nucleótidos AAV para usar en el contexto de la presente invención.

50 Preferiblemente las secuencias ITR AAV para usar en el contexto de la presente invención son derivadas de AAV1, AAV2 y/o AAV4.

Asimismo, las Rep (Rep78/68 y Rep52/40) secuencias codificantes son preferiblemente derivadas de AAV1, AAV2 y/o AAV4.

55 Las secuencias codificantes para las VP1, VP2 y VP3 proteínas cápsidas para usar en el contexto de la presente invención pueden sin embargo ser tomadas de cualquiera de los 42 serotipos conocidos, más preferiblemente de AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 o AAV9 o partículas tipo AAV recién desarrolladas obtenidas por ejemplo técnicas de redistribución de cápsida y bibliotecas de cápsida AAV o de ITR recién diseñados, desarrollados o evolucionados.

60 [0072] AAV Rep y secuencias ITR están particularmente conservadas entre más serotipos.

Las Rep78 proteínas de varios serotipos AAV son por ejemplo más del 89% idénticas y la identidad de secuencia de nucleótidos totales en el nivel de genoma entre AAV2; AAV3A, AAV3B y AAV6 es alrededor de 82% (Bantel-Schaal et al., 1999, J. Virol., 73(2):939-947). Además, las secuencias Rep e ITR de muchos serotipos AAV se conocen por complementar eficazmente de forma cruzada (es decir, sustituto funcionalmente) las secuencias correspondientes de otros serotipos en la producción de partículas AAV en células mamíferas.

65

La US2003148506 informa que AAV Rep y secuencias ITR eficaz también complementan eficientemente de forma cruzada otro AAV Rep y secuencias ITR en las células de insecto.

[0073] Las proteínas de vasopresina AAV se conocen por determinar la tropicidad celular del virión AAV.

5 Las secuencias de codificación de proteína de vasopresina están menos conservadas significativamente que las Proteína Rep y genes entre serotipos AAV diferentes.

La capacidad de las secuencias Rep y ITR para complementar de forma cruzada secuencias correspondientes de otros serotipos permite la producción de partículas rAAV pseudotipadas que comprenden las proteínas cápsidas de un serotipo (por ejemplo; AAV3) y las secuencias Rep y/o ITR de otro serotipo AAV (por ejemplo, AAV2).

10 Tales partículas rAAV pseudotipadas son una parte de la presente invención.

[0074] Las "AAV" secuencias modificadas también se pueden usar en el contexto de la presente invención, por ejemplo para la producción de vectores rAAV en las células de insecto.

15 Tales secuencias modificadas por ejemplo incluyen secuencias con al menos aproximadamente 70%, al menos aproximadamente 75%, al menos aproximadamente 80%, al menos aproximadamente 85%, al menos aproximadamente 90%, al menos aproximadamente 95% o más nucleótidos y/o identidad de secuencia de aminoácidos (por ejemplo, una secuencia con aproximadamente 75-99% de identidad de secuencia de nucleótidos) a un AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 o AAV9 ITR, Rep o vasopresina se puede usar en lugar de

20 secuencias de tipo salvaje AAV, ITR, Rep o de vasopresina.

[0075] Aunque similar a los otros serotipos AAV en muchos aspectos, AAV5 difiere de otros serotipos AAV humanos y símicos más del otros serotipos humanos y símicos conocidos.

En vista de ello, la producción de rAAV5 puede diferir de la producción de otros serotipos en las células de insecto.

25 Donde métodos de la invención se emplean para producir rAAV5, se prefiere que uno o más constructos que comprendan, colectivamente en el caso de un constructo, una secuencia de nucleótidos que incluye un AAV5 ITR, una secuencia de nucleótidos que comprenda una secuencia AAV5 Rep codificante (es decir una secuencia de nucleótidos comprende un AAV5 Rep78).

30 Tales secuencias ITR y Rep se pueden modificar como se desea para obtener una producción eficaz de rAAV5 o vectores pseudotipados rAAV5 en las células de insecto.

Por ejemplo, el codón de inicio de las secuencias Rep puede ser modificado, sitios de empalme de vasopresina se pueden modificar o eliminar y/o el codón de inicio VP1 y nucleótidos cercanos se pueden modificar para mejorar la producción de vectores rAAV5 en la célula de insecto.

35 [0076] La secuencia de nucleótidos comprende el transgen tal y como se define aquí arriba puede así comprender una secuencia de nucleótidos que codifica al menos un "producto génico de interés" para la expresión en una célula mamífera, situada de manera que será incorporada en un vector (rAAV) parvoviral recombinante replicado en la célula de insecto.

40 En el contexto de la invención se entiende que una célula de mamífero preferida particularmente donde el "producto génico de interés" debe ser expresado, es una célula humana.

Cualquier secuencia de nucleótidos se puede incorporar para una expresión tardía en una célula mamífera transfectada con el vector (rAAV) parvoviral recombinante producido conforme a la presente invención.

45 La secuencia de nucleótidos puede por ejemplo codificar una proteína, esta puede expresar un agente de ARNi, es decir una molécula de ARN que es capaz de la interferencia de ARN tal como por ejemplo un ARNhc (hairpinARN corto) o un ARNsi (interferencia corta ARN). "ARNsi" significa un ARN de interferencia pequeño que es un ARN bicatenario de longitud corta que no es tóxico en células mamíferas (Elbashir et al., 2001, Nature 411: 494-98; Caplen et al., 2001, Proc. Natl. Acad.Sci.EE.UU 98: 9742-47). En una forma de realización preferida, la secuencia de nucleótidos que comprende el transgen puede comprender dos secuencias de codificación de nucleótidos, cada codificación un producto génico de interés para la expresión en una célula mamífera.

50 Cada una de las dos secuencias de nucleótidos que codifican un producto de interés está localizada de manera que será incorporada en un vector (rAAV) parvoviral recombinante replicado en la célula de insecto.

[0077] El producto de interés para la expresión en una célula mamífera puede ser un producto génico terapéutico.

55 Un producto génico terapéutico puede ser un polipéptido o una (ARNsi) molécula de ARN, u otro producto génico que, cuando se ha expresado en una célula objetivo, proporciona un efecto terapéutico deseado tal como por ejemplo ablación de una actividad no deseada, por ejemplo la ablación de una célula infectada, o la complementación de un defecto genético, por ejemplo causando una deficiencia en una actividad enzimática.

60 Los ejemplos de productos genéticos de polipéptido terapéutico incluyen cFTR Factor IX, lipoproteína-lipasa, (LPL preferiblemente LPL S447X; ver WO 01/00220), apolipoproteína A1, glucuronosiltransferasa de difosfato de uridina (UGT), proteína de interacción con el regulador GTPase de la retinitis pigmentosa (RP-GRIP) y citocinas o interleuquinas como por ejemplo IL-10, porfobilinogeno desaminasa (PBGD), un factor neurotrófico tal como factor neurotrófico derivado de línea celular glial (GDNF) y alanina:glioxilato aminotransferasa (AGT).

65 [0078] Alternativamente o además como otro producto génico, la secuencia de nucleótidos que comprende el transgen tal y como se define aquí arriba puede comprender además una secuencia de nucleótidos que codifica un polipéptido que sirve como una proteína marcadora para valorar la transformación y expresión celular.

Proteínas de marcador adecuado para este propósito son por ejemplo la proteína GFP fluorescente y la timidina quinasa HSV de genes marcadores seleccionables (para la selección en el medio HAT), higromicina bacteriana B fosfotransferasa (para la selección en la higromicina B), Tn5 fosfotransferasa de aminoglicósido (para la selección en G418) y dihidrofolato reductasa (DHFR) (para la selección en el metotrexato), CD20, el gen de factor de crecimiento nervioso de baja afinidad.

Las fuentes para obtener estos genes marcadores y métodos para su uso se proporcionan en Sambrook y Russel, supra. Además, la secuencia de nucleótidos que comprende el transgen tal y como se define aquí arriba puede comprender otra secuencia de nucleótidos que codifica un polipéptido que puede servir como un mecanismo de fallo de seguridad que permite polimerizar un sustrato de células transducidas con el vector (rAAV) parvoviral recombinante de la invención, si se cree necesario.

Tal secuencia de nucleótidos, frecuentemente referida como un gen suicida, codifica una proteína que es capaz de la conversión de un profármaco en una sustancia tóxica que es capaz de matar las células transgénicas donde la proteína se expresa.

Los ejemplos adecuados de tales genes suicidas incluyen por ejemplo el gen E.coli de citosina deaminasa o uno de los genes de timidina quinasa de virus herpes simplex, citomegalovirus y virus Varicella-Zoster, en cuyo caso ganciclovir se puede utilizar como profármaco para matar las células transgénicas en el sujeto (ver por ejemplo Clair et al., 1987, Antimicrob. Agents Chemoter. 31: 844-849).

[0079] En otra forma de realización uno de los productos genéticos de interés pueden ser una proteína AAV.

En particular, una proteína Rep, tal como Rep78 o Rep68, o un fragmento funcional del mismo.

Una secuencia de nucleótidos que codifica un Rep78 y/o un Rep68, si está presente en el genoma de un vector (rAAV) parvoviral recombinante de la invención y expresado en una célula mamífera transducida con el vector, permite la integración del vector (rAAV) parvoviral recombinante en el genoma de la célula de mamífero transducida. La expresión de Rep78 y/o Rep68 en una célula de mamífero transducida rAAV o infectada puede proporcionar una ventaja para ciertos usos del vector (rAAV) parvoviral recombinante, lo que permite la expresión a largo plazo o permanente de cualquier otro producto génico de interés introducido en la célula por el vector.

[0080] En los vectores (rAAV) parvoviral recombinantes de la invención la al menos una secuencia(s) de nucleótido que codifica un producto génico de interés para la expresión en una célula mamífera, preferiblemente está/están operativamente enlazadas a al menos una secuencia de control de expresión compatible con la célula mamífera, por ejemplo, un promotor.

Muchos de estos promotores se conocen en la técnica (ver Sambrook y Russel, 2001, supra).

Los promotores constitutivos que son en términos generales expresados en muchos tipos de célula, tal como el promotor CMV pueden ser utilizados.

Sin embargo, más preferidos serán los promotores que son, inducibles de tejido específico, específico de tipo celular o específico de ciclo celular.

Por ejemplo, para la expresión específica de hígado un promotor se puede seleccionar de un promotor (1-anti-tripsina (AAT), un promotor de globulina de unión de hormona de tiroides, un promotor de albúmina, un promotor LPS (globulina de unión de tiroxina), un promotor híbrido HCR-ApoCII, un promotor híbrido HCR-hAAT, un AAT promotor combinado con el elemento intensificador de gen de albúmina de ratón (Ealb) y un promotor de apolipoproteína E.

Otros ejemplos incluyen el promotor E2F para tumor selectivo y, en particular, expresión selectiva de tumor de célula neurológica (Parr et al., 1997, Nat.Med. 3:1145-9) o el IL-2 promotor para usar en células mononucleares de sangre (Hagenbaugh et al., 1997, J Exp Med; 185: 2101-10).

[0081] AAV es capaz de infectar un número de células mamíferas.

Ver, por ejemplo, Tratschin et al. (1985, Mol.Cell Biol. 5:3251-3260) y Grimm et al. (1999, Hum.Gene Ther. 10:2445-2450). Sin embargo, la transducción AAV de fibroblastos sinoviales humanos es significativamente más eficaz que en células murinas similares, Jennings et al., Arthritis Res, 3:1 (2001) y la tropicidad celular de AAV difiere de entre los serotipos. Ver, por ejemplo, Davidson et al. (2000, Proc.Natl.Acad.Sci.EE.UU, 97:3428-3432), que discute las diferencias entre AAV2, AAV4 y AAV5 respecto al tropismo de célula CNS mamífera y eficiencia de transducción.

En una forma de realización preferida, una célula huésped de la invención es cualquier célula mamífera que se puede infectar por un virión parvoviral, por ejemplo, pero no limitado a una célula muscular, unas células hepáticas, una célula nerviosa, una célula glial y una célula epitelial.

En una forma de realización preferida una célula huésped de la invención es una célula humana.

[0082] Preferiblemente, en la construcción, la secuencia de nucleótidos que codifica una proteína Rep parvoviral y/o una proteína cápsida parvoviral está operativamente enlazada a las secuencias de control de expresión para la expresión en una célula de insecto.

Estas secuencias de control de expresión al menos incluirán un promotor que es activo en las células de insecto.

Las técnicas conocidas por un experto en la técnica para los genes extranjeros de expresión en las células huésped de insecto pueden utilizarse para la práctica de la invención.

La metodología para ingeniería molecular y expresión de polipéptidos en las células de insecto se describe, por ejemplo, en Summers and Smith. 1986.

A Manual of Methods for Baculovirus Vectors and Insect Culture Procedures, Texas Agricultural Experimental Station Bull.No. 7555, College Station, Tex.; Luckow. 1991. En Prokop et al., Cloning and Expression of Heterologous Genes in Insect Cells with Baculovirus Vectors' Recombinant DNA Technology and Applications, 97-152; King, L.A. y R. D. Possee, 1992, The baculovirus expression system, Chapman and Hall, United Kingdom; O'Reilly, D. R., L.K. Miller, V.A. Luckow, 1992, Baculovirus Expression Vectors: A Laboratory Manual, New York; W.H. Freeman and Richardson, C.D., 1995, Baculovirus Expression Protocols, Methods in Molecular Biology, volume 39; US 4,745,051; US2003148506; y WO 03/074714.

Los promotores adecuados para la transcripción de las secuencias de nucleótidos comprendidas en la primera y la segunda construcción de la invención incluyen por ejemplo el poliedro (PolH), p10; p35, IE-1 o promotores \otimes IE-1 y otros promotores descritos en las referencias anteriores.

[0083] El constructo de ácidos nucleicos que comprende al menos los primeros y/o segundos casetes de expresión, puede comprender además una secuencia de control de expresión que comprende una novena secuencia de nucleótidos de SEC.

ID N°: 9 o una secuencia de nucleótidos sustancialmente homóloga para SEC.

ID N°: 9, aguas arriba del codón iniciador de la secuencia de nucleótidos que codifica la proteína Rep parvoviral y/o codón iniciador de la secuencia de nucleótidos que codifica la proteína cápsida parvoviral VP1.

Una secuencia con identidad sustancial a la secuencia de nucleótidos de SEC.

ID N°: 9 y que ayudará a aumentar la expresión de la proteína Rep parvoviral es por ejemplo una secuencia que tiene al menos de 60%, 70%, 80% o 90% de identidad a la novena secuencia de nucleótidos de SEQ ID N.º: 9.

[0084] La célula de insecto puede ser cualquier célula que se adecua a la producción de proteínas heterólogas.

Preferiblemente la célula de insecto permite para replicación de vectores baculovirales y se pueden mantener en el cultivo.

Más preferiblemente la célula de insecto también permite la replicación de vectores parvovirales recombinantes, incluyendo vectores rAAV.

Por ejemplo, la línea celular usada puede ser de Spodoptera frugiperda, Líneas celulares de Drosophila o líneas celulares de mosquito, por ejemplo, líneas celulares derivadas Aedes albopictus.

Las células de insecto preferidas o líneas celulares son células de las especies de insecto que son susceptibles de infección de baculovirus, incluyendo por ejemplo S2 (CRL-1963 ATCC), Se301; SeIZD2109, SeUCR1, Sf9; Sf900+, Sf21, BTI-TN-5B1-4, MG-1; Tn368, HzAm1; Ha2302, Hz2E5, High Five (Invitrogen, CA, EE.UU) y expresSF+ \otimes (US 6,103,526; Protein Sciences Corp., CT, USA).

Una célula de insecto preferida según la invención es una célula de insecto para la producción de vectores parvovirales recombinantes.

[0085] El uno o más constructos de ácidos nucleicos del método de la invención pueden estar integrados de forma estable en el genoma de la célula de insecto.

Un técnico en la materia sabe cómo introducir de forma estable una secuencia de nucleótidos en el genoma de insecto y como identificar una célula con tal secuencia de nucleótidos en el genoma.

La incorporación en el genoma se puede asistir por, por ejemplo, el uso de un vector que comprende las secuencias de nucleótidos altamente homólogas para regiones del genoma de insecto.

El uso de secuencias específicas, tales como transposones, es otra vía para introducir una secuencia de nucleótidos en un genoma.

[0086] Las condiciones de crecimiento para células de insecto en el cultivo y producción de productos heterólogos en las células de insecto en el cultivo se conocen bien en la técnica y se describen por ejemplo en las referencias citadas anteriores en ingeniería molecular de células de insectos (ver también WO2007/046703).

[0087] En una forma de realización preferida del método de la invención, el virión parvoviral recombinante se recupera.

La recuperación preferiblemente comprende la etapa de purificación de afinidad de (viriones que comprenden el) vector (rAAV) parvoviral recombinante que utiliza un anticuerpo anti-AAV, preferiblemente un anticuerpo inmovilizado.

El anticuerpo anti-AAV es preferiblemente un anticuerpo monoclonal.

Un anticuerpo especialmente adecuado es un anticuerpo de camélido monocatenario o un fragmento del mismo como por ejemplo obtenible de camellos o llamas (ver por ejemplo Muyldermans, 2001, Biotechnol. 74: 277-302).

El anticuerpo para purificación de afinidad de rAAV preferiblemente es un anticuerpo que específicamente enlaza un epítipo en una proteína cápsida AAV, por lo cual preferiblemente el epítipo es un epítipo que está presente en la proteína cápsida de más de un serotipo AAV.

Por ejemplo el anticuerpo se puede elevar o seleccionar basándose en enlace específico para AAV2 cápsida pero al mismo tiempo también éste puede también específicamente enlazar con cápsidas AAV1, AAV3 y AAV5.

[0088] En un segundo aspecto, la invención se refiere a un constructo de ácidos nucleicos que comprende uno o más casetes de expresión de la invención como se ha definido aquí arriba.

- En una forma de realización preferida, el constructo de ácidos nucleicos de la invención comprende un primero y un segundo casete de expresión de la invención y opcionalmente una secuencia de ácidos nucleicos de (i). Preferiblemente, el primer promotor en el constructo de ácido nucleico de la invención es un promotor p10 y el segundo promotor es un Promotor PolH o un promotor 4xHsp27 EcRE+mínimo Hsp70.
- 5 Más preferiblemente, el primer promotor está operativamente enlazado con un intensificador, preferiblemente un intensificador HR1.
- En otra forma de realización, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor 4xHsp27 EcRE+mínimo Hsp70 y el segundo promotor es un promotor PolH.
- 10 En otra forma de realización, el primer promotor en el constructo de ácido nucleico de la invención es un promotor PolH y el segundo promotor es un promotor p10, un deltaE1 o un E1.
- En otra forma de realización, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor PolH y el segundo promotor es un promotor deltaE1 o E1.
- 15 En otra forma de realización, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor p10 y el segundo promotor es un promotor deltaE1 o un E1.
- En otra forma de realización, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor PolH y el segundo promotor es un promotor PolH.
- En otras formas de realización, cualquier otra combinación del primer promotor y el segundo promotor son parte de la invención.
- 20 El primer promotor se puede seleccionar del grupo que consiste en un Promotor PolH, promotor p10, promotor de proteína básica, un promotor inducible o un promotor deltaE1 o un Promotor E1, o cualquier otro promotor de gen de baculovirus tardío o muy tardío.
- El segundo promotor se puede seleccionar del grupo que consiste en un Promotor PolH, promotor p10, promotor de proteína básica, un promotor inducible o un promotor deltaE1 o un Promotor E1, o cualquier otro promotor de gen de baculovirus tardío o muy tardío.
- 25 Más preferiblemente, el primer promotor se selecciona del grupo que consiste en un Promotor PolH, promotor p10 o promotor de proteína básica y donde el segundo promotor es un promotor deltaE1 o un Promotor E1, o cualquier otro promotor de gen de baculovirus temprano o tardío.
- [0089] En una forma de realización preferida el primer casete de expresión comprimido en el constructo de ácidos nucleicos de la invención, comprende un elemento intensificador tal como se ha definido anteriormente.
- 30 [0090] En un tercer aspecto, la invención se refiere a una célula de insecto tal como se ha definido anteriormente.
- [0091] En un cuarto aspecto, la invención se refiere a un equipo que comprende (a) un constructo de ácidos nucleicos que comprende el primer y segundo casete de expresión tal como se ha definido anteriormente; y (b) un constructo de ácidos nucleicos que comprende una secuencia de nucleótidos que codifica un sitio de clonación múltiple para un transgen que se flanquea por al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral, cuyo transgen está operativamente enlazado a un promotor capaz de dirigir la expresión del transgen en una célula huésped.
- 35 [0092] En una forma de realización preferida, el constructo de ácidos nucleicos (b) comprende una secuencia de nucleótidos que codifica un transgen que se flanquea por al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral, cuyo transgen está operativamente enlazado a un promotor capaz de dirigir la expresión del transgen en una célula huésped.
- 40 [0093] El equipo puede comprender además células de insecto y una secuencia de ácidos nucleicos que codifica funciones auxiliares de baculovirus para la expresión en la célula de insecto.
- [0094] En otro aspecto la invención se refiere a un lote de viriones parvovirales producido en los métodos descritos anteriormente de la invención.
- 50 Un "lote de viriones parvoviral" aquí se define como todos los viriones parvovirales que se producen en el mismo ciclo de producción, opcionalmente por el contenedor de células de insecto.
- En una forma de realización preferida, el lote de viriones parvovirales de la invención comprende una proporción virión completo:virión total como se ha descrito anteriormente y/o una proporción virión completo: vacío como se ha descrito anteriormente.
- 55 [0095] En este documento y en sus reivindicaciones, la palabra "comprender" y sus conjugaciones se usan en su sentido no limitativo para significar que los aspectos seguidos a la palabra se incluyen, pero los aspectos no específicamente mencionados no se excluyen.
- 60 Además, la referencia a un elemento por el artículo indefinido "un" o "una" no excluye la posibilidad de que más de uno de los elementos estén presentes, a menos que el contexto requiera claramente que haya uno y solo uno de los elementos.
- El artículo indefinido "uno" o "una" así normalmente significa "al menos uno".
- 65 [0096] Los ejemplos siguientes ilustran la invención:

Ejemplos

Ejemplo 1

- 5 1.1 Materiales y métodos
- 1.1.1 Generación de baculovirus recombinante
- 10 [0097] Los constructos siguientes fueron generados:
- 1.1.1.1 Construcción de pVD118(new)
- [0098] PVD118 (new) es un vector de control que incluye un casete de expresión que comprende la secuencia de nucleótidos de Rep78 bajo el control del promotor p10 y un casete de expresión que comprende la secuencia de nucleótidos de genes Cap bajo el control del poliedro promotor (PolH o pPH). PVD118 (new) fue construido como se muestra en la figura 1.
- 15 Antes del plásmido final pVD118(new) se podría hacer el predecesor pFastBac doble Rep78/ACG y pVD118 (lot#1) fue construido.
- 20 En resumen, el plásmido REP-ACG/PSC (solicitud de patente WO2007148971 aquí también referido como pVD88) pVD88 fue digerido con salientes SpeI* XbaI y 5' salientes se hicieron romos.
- El 2057bp fragmento fue aislado a partir de un gel de agarosa, purificado y ligado en el con SmaI linealizado pFastBac doble (Invitrogen), dando como resultado pFastBac doble Rep78/ACG.
- 25 Luego, este plásmido fue digerido con BstZ17I y SnaBI y el fragmento 2537bp p10 Rep78/ACG fue aislado y ligado en el BstZ17I linealizado pVD84, generando pVD118 (lot#1).
- Sin embargo, la orientación del casete de expresión Rep78/ACG fue incorrecta y por lo tanto eliminada mediante la realización de una digestión con NheI* B1pI.
- Después de producir los salientes 5' romos, el fragmento de vector 12431bp fue aislado y purificado de gel.
- 30 Posteriormente, el 2057bp fragmento purificado de REP-ACG/PSC fue ligado en el vector y la mezcla de transformación fue transformada para células competentes químicamente TOP10 (Invitrogen) y colocada en placas que contienen ampicilina.
- El análisis de restricción con NaeI*SacI fue realizado en el ADN aislado de cultivos de minipreparación. Los clones correctos proporcionan fragmentos con un tamaño de 2204bp y 12292bp.
- 35 1.1.1.2 Construcción de pVD118 (new) + HR1
- [0099] PVD118 (new) + HR1 es lo mismo que pVD118, con la adición de un hr1 intensificador que está situado entremedias del p10 y promotor de poliedro y fue construido como se muestra en la figura 2.
- 40 Brevemente, un PCR realizado en ADN vírico AcMNPV (Protein Sciences Corporation, Meriden, USA) con cebadores HR1-F w 5' -gtatcgtatgacact atcgatgttgac-3 (SEQ ID N.º: 10) y H R 1-R v 5' -gtatcgtatgacattgtctccaatactag-3 (SEC ID No:11) resultó en un producto de 904bp que se clona al vector de pCRII-blunt-TOPO (Invitrogen).
- Después de la digestión con BstZ17I el fragmento 898bp fue aislado de gel, purificado y ligado en el vector pVD118(new) que se abrió con BstZ17I y defosforilado.
- 45 Las digestiones de control de clones correctos con SpeI*EcoNI resultaron en fragmentos de 1269bp y 14125bp.
- 1.1.1.3 Construcción de pVD84 (+p10 Rep) (=pVD165)
- [0100] PVD165 es un vector de control que incluye un casete de expresión que comprende la secuencia de nucleótidos de Rep78 donde el codón de inicio ha sido mutado en ACG bajo el control del promotor p10 y un casete de expresión que comprende la secuencia de nucleótidos de genes Cap bajo el control del promotor PolH.
- 50 Los casetes de expresión están en dirección opuesta en el plásmido. PVD165 fue construido como se muestra en la figura 3.
- Brevemente, un PCR realizado en pVD118 (new) con cebadores poliA Fw 5'-agatct gtatggctatggcagggc-3 (SEC ID No:12) y p10 Rv 5'-agatct cccggg acggaccttaattcaaccaac-3 (SEQ ID N.º: 13) resultó en un producto de 2566bp que fue clonado al vector (Invitrogen) pCRII-blunt-TOPO.
- 55 Después de la digestión con Bg1II el fragmento 2541bp fue aislado de gel, purificado y ligado en el pVD84 vector que se abrió con Bg1II y defosforilado.
- Las digestiones de control de clones correctos con SpeI*XbaI resultaron en fragmentos de 1269bp y 11810bp.
- 60 1.1.1.4 Construcción de pVD165 + HR1
- [0101] PVD165 + HR1 es lo mismo que pVD165, con la adición de un hr1 intensificador aguas abajo del promotor p10 y fue construido como se muestra en la figura 4.

Brevemente, un PCR realizado en ADN vírico AcMNPV con cebadores HR1-Fw 5'-gtatacgtatgacact atcgtatgtgac-3 (SEQ ID N.º: 10) y HR1-Rv 5'-gtatacgtatgacattatgtctccaactactag-3' (SEQ ID N.º: 11) resultó en un producto de 904bp que fue clonado al (Invitrogen) vector pCRII-blunt-TOPO.

Después de la digestión con BstZ171 el 898bp fragmento fue aislado del gel, purificado y ligado en el pVD165 vector que se abrió con SmaI y defosforilado.

Las digestiones de control de clones correctos con ClaI resultaron en fragmentos de 875bp, 1184bp, 4160bp y 9180bp.

1.1.1.5 Construcción de pVD165 con promotor inducible aguas arriba Cap (pVD165 + 4xEcRE CAP)

[0102] PVD165 + 4xEcRE tapa es similar al pVD165, pero comprende un promotor inducible en vez del promotor poliedro (pPH). PVD165 + 4xEcRE CAP fue construida como se muestra en la figura 5.

Brevemente, el promotor inducible, que comprende 4 EcRE consecutivos, un promotor mínimo hsp70 (Poels et al. *Insect Biochem Mol Biol* 2004,34(5):451-458) y parte de la secuencia codificante de CAP, fue sintetizada y ligada en el pCRII-blunt-TOPO (Invitrogen).

Después de la digestión con BstZ171 y EcoNI el 375bp fragmento fue aislado de gel, purificado y ligado en el pVD165 vector que se abrió con BstZ171 y EcoNI.

Las digestiones de control de clones correctos con PmeI*EcoRV resultaron en fragmentos de 2554bp y 12116bp.

1.1.1.6 Construcción de pVD165 con promotor inducible aguas arriba de Rep (pVD165 + 4*EcRE Rep78-52)

[0103] PVD165 + 4*EcRE Rep78 es similar al pVD165, pero comprende un promotor inducible en vez del promotor p10. PVD165 + 4*EcRE Rep78 fue construido como se muestra en la figura 6.

Brevemente, el promotor inducible que consiste fuera de 4 EcRE consecutivos y un promotor mínimo hsp70 fue sintetizado y ligado en el pCRII-blunt-TOPO (Invitrogen).

Después de la digestión con BstZ171 y SpeI el 290bp fragmento fue aislado de gel, purificado y ligado en el vector pVD165 que se abrió con SmaI y SpeI.

Las digestiones de control de clones correctos con SpeI*BglII resultaron en los fragmentos de 298bp, 2376bp y 11954bp.

1.1.1.7 Construcción de pVD190 construcción (deltaIE1 Cap + pPolh Rep)

[0104] DeltaIE1 Cap + pPolh Rep es similar al pVD165, pero comprende un promotor delta IE1 en vez del pPH promotor y un pPromotor PolH en vez de del promotor p10.

DeltaIE1 Cap + pPolh Rep fue construido como se muestra en la figura 7 y 8.

Aquí, un vector precursor fue construido de la siguiente manera.

Un PCR realizado en pFBDSLRL (Urabe et al. *Human gene therapy*. 2002,13(16):1935-1943) con cebadores BstZ171-deltaIE1 Fw: 5'-gggcc gtatacgcgataacgcggttggtggcg-3 (SEC ID No:14) y AfIII-delta IE1 Rv: 5'-cgacttaagacggcgaattctgc agatggc-3 (SEQ ID N.º: 15) resultó en un producto de 181bp que fue clonado al vector de pCRII-blunt-TOPO (Invitrogen).

Después de la digestión con BstZ171*AfIII el 165bp fragmento fue aislado de gel, purificado y ligado en el pVD165 + 4xEcRE vector Cap que se abrió con BstZ171 y AfIII.

El plásmido resultante es pVD165 + CAP deltaIE1 y la digestión de control con BstZ171*AfIII debería suponer los fragmentos 165bp y 14374bp.

Posteriormente, el promotor de polihedrina fue clonado delante del Rep casete de expresión en pVD165 + deltaIE1 Cap. Aquí, un PCR con cebadores SmaI-pPolh Fw: 5'-tctcccgggagatcatggaga taattaaatgataac-3 (SEC ID No:16) y SpeI-pPolh Rv: 5'-gttactagtgcgctcgtcgac-3 (SEC ID No:17) fue realizado en pVD88.

Esto generó un producto PCR de 198bp que fue clonado al vector de pCRII-blunt-TOPO (Invitrogen).

Después de la digestión con SpeI*SmaI el 188bp fragmento fue aislado de gel, purificado y ligado en el vector pVD165 + deltaIE1 Cap que se abrió con SpeI y SmaI.

Finalmente, esto resultó en el constructo pVD165 + deltaIE1 CAP.

1.1.1.8 Construcción del constructo p10-Cap-pPolh-Rep

[0105] El constructo p10 Cap + pPolh Rep es similar al constructo delta-IE1-Cap-pPolh-Rep, pero comprende un promotor p10 en vez del promotor delta-IE1 delante del casete de expresión Cap y fue construido como se muestra en la figura 9.

Un PCR realizado en pFastBac doble con cebadores BstZ171-p10 Fw: 5'-agtatacggaccttaattcaac-3' (SEC ID No:18) y AfIII-p10 Rv: 5'-cgacttaagacggcggcgttcgaatc-3' (SEQ ID N.º: 19) resultó en un producto de 171bp que fue clonado al vector pCRII-blunt-TOPO (Invitrogen).

Después de la digestión con BstZ171*AfIII el 160bp fragmento fue aislado de gel, purificado y ligado en el constructo delta-IE1-Cap-pPolh-Rep que fue abierto con BstZ171 y AfIII.

1.1.1.9 Construcción de pVD194 (Rep78/CTG(delta ATG))

[0106] Un plásmido rep con un CTG codón de inicio sin algunos sitios ATG internos fue primero construido por la sintetización del gen necesario según la secuencia expuesta en SEQ ID N.º: 31.

El gen rep fue luego eliminado del plásmido pVD88 usando enzimas de restricción RsrII y XbaI.

El gen sintetizado fue luego ligado en el plásmido pVD88 usando sitios de restricción RsrII y XbaI para proporcionar pVD195. PVD195 fue luego digerido con BgIII., dando como resultado fragmentos 9297bp y 2231bp.

Los salientes 5' fueron luego llenados de Klenow y luego digeridos con SexAI.

Esto produce fragmentos 9297bp, 1372 y 859bp.

El 859bp fragmento luego se aisló.

El vector necesario fue generado digiriendo pVD165 con SpeI para dar un 14501bp fragmento lineal.

Los 5' salientes fueron llenados de Klenow y luego digeridos con SexAI para producir fragmentos 13600bp y 901bp.

El fragmento 13600bp fue aislado.

El 859bp BgIII(Klenow)*SexAI fragmento fue luego ligado en pVD165 (SpeI(Klenow)*SexAI) para producir pVD194 (ver Fig. 10).

Una digestión de control con SexAI*XmaI debería suponer fragmentos de 13435bp y 1029bp

1.1.1.10 Construcción de pVD84

[0107] Para convertir el sitio de inicio del vector de expresión de baculovirus pFBAAV1VPm11 (Urabe et al., 2002, Hum. Gene Ther. 13: 1935-1943) de ACG a GTG un PCR fue realizado utilizando los cebadores siguientes:

La secuencia de cebador directo contiene un sitio de BamHI (cebador AMT #169 SEC ID NO:29)

5'- TTAGGATCCTGTTAAGGTGGCTGCCGACGG -3

La secuencia de cebador inverso contiene un sitio Stul (cebador AMT #158 SEC ID NO:30)

5'- GTCGTAGCCTTGTCGTGCTCGAGGGCCGC -3

sitio de inicio

Sitio de inicio	Cebador directo	Cebador inverso	AMT plasmid# (Bac-toBac)	AMTplasmid# (PSC)
GTG	169	158	pVD63	pVD84

[0108] El PCR que usa cebadores #169 y #158 se realizó y el producto PCR (250bp) fue purificado utilizando el equipo de purificación PCR (Qiagen Lot#11879372) y cortado con enzimas de restricción BamHI y Stul.

El vector de expresión de baculovirus fue también digerido con BamHI y Stul.

La defosforilación siguiente del vector con SAP (Promega Lot#17501504), tanto el inserto (Cap con sitio de inicio GTG) como el vector fueron purificados en el gel.

Después del ligamiento de vector e inserto, fueron transformados en células competentes químicamente DH5 (Invitrogen Lot#1241753) y estriados sobre placas LB que contienen ampicilina.

El ADN de pVD63 (Cap/ sitio de inicio GTG) fue purificado y examinado para identificar usando análisis de secuencias por BaseClear y restricción.

[0109] Para clonar el gen Cap con el GTG sitio de inicio en el vector de expresión de baculovirus pPSC10 (Protein Sciences) fue cortado de pVD63 con SmaI y AvrII y ligado en el vector de expresión defosforilado que se abrió con EcoRV y XbaI.

Posteriormente, la mezcla de ligamiento fue transformada en células competentes químicamente DH10® (Invitrogen lot# 1268527) y estriadas sobre placas LB de ampicilina.

Después del aislamiento de minipreparación de ADN utilizando el equipo de minipreparación Qiaprep Spin (Qiagen lot #12180218) se seleccionó una minipreparación (clon #13) por análisis de restricción con SphI.

Con este clon ADN (pVD84) fue purificado con el equipo midiprep SNAP (solución de Invitrogen Lot#1256921 y columna Lot#1259367).

La identidad de pVD84 fue controlada por análisis de secuencias alrededor del codón de inicio realizado por BaseClear y por análisis de restricción con BamHI y SphI (SEC ID NO:28).

1.1.1.11 Producción de baculovirus recombinante

[0110] Recombinante Bac.VD118(new), Bac.VD118(new)+HR1; Bac.VD165, Bac.VD165+HR1, Bac.VD165+4xEcRE tapa, Bac.VD165+4*EcRE Rep78; Bac.VD190 y Bac.VD194 (p0) fueron generados con el Protein sciences system (Protein Sciences Corporation, Meriden, USA).

El Baculovirus recombinante fue amplificado por la dilución de estos 1:100 en 2x10⁶ células SF+ por ml.

Tres días después de la infección las células fueron centrifugadas y se recuperó el sobrenadante que contenía el virus. La amplificación de pasajes siguientes fue realizada de la misma manera.

1.1.2 Producción rAAV

[0111] Los lotes rAAV fueron producidos según Urabe et al., 2002 (supra), pero con la excepción de que dos baculovirus recombinantes fueron usados en vez de tres.

Un baculovirus albergó un constructo de expresión bajo el control del promotor CMV y se flanquea por ITR AAV.

El otro baculovirus es el Rep-Cap baculovirus que albergó dos casetes de expresión, uno para el gen de replicación AAV y uno para la cápsida AAV.

La expresión de la replicación o gen cápsida bajo el control del promotor inducible fue regulada por adición de 0.001-1uM Ponasterone A al medio de cultivo.

5 Los experimentos de producción diferentes rAAV1 fueron realizados con existencias de baculovirus Bac.VD43 p5 con el transgen LPL bajo control del promotor CMV y pasajes diferentes (p3, p4 o p5) del Rep/Cap baculovirus. En cada experimento la rAAV1 producción estándar (Bac.VD88:Bac.VD84:Bac.VD43 con proporción 5:1:1) se llevó como un control.

10 1.1.3 Determinación completa/total de partícula AAV

[0112] Para determinar la proporción de cápsidas totales frente a completas, la cantidad de copias de genoma (gc) se divide por la cantidad de partículas AAV totales.

15 La cantidad de GC/ml se midió por ensayo Q-PCR y la cantidad de partículas AAV totales se determinó con un enzimoimmunoanálisis de Progen (ver debajo).

Para la mezcla de reacción Q-PCR SYBR Green PCR Master Mix (Applied Biosystems, #4309155) se usó según las instrucciones del productor (25 µl volumen total; programa PCR: 10 min 95°C, 40 ciclos de 15 seg 95°C y 1 min 60°C) usando uno de los siguientes conjuntos de cebador:

Pr59	AATGGGCGGTAGGCGTGTA	CMV	SEQ ID N.º: 26
pr60	AGGCGATCTGACGGTTCCTAA	CMV	SEQ ID N.º: 27

20 [0113] La proporción se compara con la proporción obtenida bajo las condiciones de producción estándar, utilizando una 5:1:1 proporción en volumen de Bac.Rep, Bac.Cap y Bac.ITR.

25 1.1.4 Análisis de electrotransferencia

[0114] Tres días después, las células de producción rAAV fueron lisadas añadiendo 0.1V 10 x TRIS tampón de lisis (1.5M NaCl, 0.5M TRIS, 0.01M MgCl, 1% Tritón X-100, pH8.5, filtro esterilizado) e incubadas en hielo durante 30 minutos.

ADN y ARN desnudos fue degradado por incubación con benzonasa a 37°C durante 15 minutos.

30 El lisado celular fue centrifugado (1,900 x g; 15 min; 4°C).

El tampón de muestra NuPage LDS (4x; Invitrogen) fue adicionado a una muestra del sobrenadante y fue cargado sobre un gel 4-12% Bis-Tris (120V).

Las proteínas fueron transferidas sobre una membrana PVDF (BioRad) durante 30 minutos, 10V (transferencia semiseca).

35 Inmunoquímica Western fue realizada por el bloqueo de la membrana con el tampón de bloqueo Superblock-PBS (PIERCE) e incubación posterior con ratón anti-Rep (303.9, Progen, Germany; dilution 1:50) y conejo anti-ratón HRP (DAKO, dilución 1:500).

Las proteínas fueron visualizadas por coloración quimioluminiscente con lumi-luz más sustrato de transferencia Western (Roche).

40 1.1.5 Total rAAV1 partícula ELISA

[0115] La cantidad total de rAAV1 partículas (tp) hecha en cada producción se determinó con el AAV1 equipo de titulación de ELISA (Progen, Heidelberg, Alemania) y se realizó según el protocolo suministrado por el fabricante, pero con la excepción de que todas las muestras y controles fueron pre-diluidos en el volumen de lisado crudo (CLB).

Brevemente, el CLB fue hecho por la cosecha de expressSF+ células después de tres días, añadiendo 10X tampón de lisis y la incubación durante 1h a 28°C.

45 Después del tratamiento con benzonasa durante 1h a 37°C, el lisado fue centrifugado a 1900g y el sobrenadante fue almacenado a 4°C.

Para determinar las rAAV1 partículas totales en el lisado crudo de cada producción las muestras fueron pre-diluidas 50-plegues en CLB, esta es la dilución de inicio.

Luego las diluciones extra de 250,1250 y 6250-plegues se hicieron en el CLB.

La línea estándar fue también diluida en el CLB.

55 Las muestras de producción con Bac.VD190 fueron solo unos 50 y 100 pliegues diluidos, debido a los niveles de expresión bajos de las proteínas cápsida.

1.1.6 Análisis de partícula total usando HPLC

60 [0116] Las materias primas AAV-1 desconocidas están inyectadas en el sistema de HPLC, dando como resultado un valor máximo en el cromatograma.

Este valor máximo se puede integrar por Chemstation-software y representa la cantidad de partículas totales inyectadas.

Una materia prima concentrada AAV-1 fue valorada para su cantidad de partículas totales por mililitro por microscopía electrónica y ajustada en el método.

Usando este estándar como un calibrador, la cantidad de partículas totales del desconocido pueden ser calculadas.

Además, cuando se inyecta una cantidad particular de copias genómicas (que aproximadamente representa la cantidad de partículas completas) en el rango de lo estándar, la proporción de partículas completas y vacías puede ser estimada.

1.2 Resultados

Proporción de partícula AAV completa: vacía mejora tras el uso de dos sistemas baculovirus, en particular con una expresión de proteína alta Rep y expresión de proteína Cap moderada

[0117] Tres constructos fueron investigados en detalle, pVD165; pVD190 y pVD194.

[0118] Para determinar la proporción total/completa para partículas rAAV1 producidas con Bac.VD165 la concentración de partícula total en los lisatos crudos fue determinada en dos experimentos independientes. Los resultados de estos ELISA y las proporciones totales/completas correspondientes se muestran en la tabla 1.

Tabla 1: la proporción total/completa de rAAV1 producida con la materia prima de baculovirus Bac.VD165.

La concentración de partícula total se determinó con la partícula total AAV1 ELISA.

La concentración de partícula completa se determinó por Q-PCR. Las proporciones totales/completas pueden solo ser en comparación con el control producido en el mismo experimento.

Bac.VD165: Bac.VD43		Concentración de partícula Total (TP/ml)	Concentración de partícula completa (GC/ml)	Proporción total/completa (TP/GC)
#1	Control	7.07E+12	2.05E+10	345
	1:1	4.91E+11	2.24E+09	219
	5:1	8.3E+11	5.0E+09	166
#2	Control	3.19E+13	1.99E+10	1603
	1:1	6.32E+11	6.33E+08	998
	5:1	1.30E+12	1.10E+09	1182

[0119] La proporción total/completa para las 1:1 producciones es en ambos experimentos 1.6 veces mejorada en comparación con el control.

Para la producción 5:1 la proporción total/completa es 2.1 y 1.4 veces mejor en comparación con el control.

En conclusión, la proporción total/completa se mejora para las rAAV1 partículas producidas con Bac.VD165.

[0120] En Bac.VD190, el casete de expresión Cap está bajo control del \otimes IPromotor E1 débil.

Esto puede suponer una expresión de cápsida inferior que en las producciones realizadas con el otro Rep/Cap baculovirus o en la situación de control, porque en todas aquellas condiciones la expresión CAP se induce por el promotor de polihedrina fuerte.

Para probar esta hipótesis de dos experimentos diferentes realizados con Bac.VD190 la concentración de partícula total fue determinada.

Los resultados se muestran en la tabla 2.

Tabla 2: proporción total/completa de rAAV1 producida con la materia prima de baculovirus Bac.VD190.

La concentración de partícula total fue determinada con la partícula ELISA total AAV1.

La concentración de partícula completa fue determinada por Q-PCR.

Las proporciones totales/completas pueden solo ser en comparación con el control producido en el mismo experimento

Bac.VD190: Bac.VD43		Concentración de partícula Total (TP/ml)	Concentración de partícula completa (GC/ml)	Proporción total/completa (TP/GC)
#1	Control	3.19E+13	1.99E+10	1603
	1:1	8.6E+10	5.97E+09	14
	5:1	3.62E+11	1.8E+09	201
#2	Control	6.11E+12	4.03E+10	152
	1:1	3.75E+11	1.74E+09	216
	5:1	3.44E+11	3.12E+09	110

Los resultados de producción # 1 (tabla 2) muestran que las proporciones totales/completas obtenidas por las producciones 1:1 y 5:1 fueron 114 y 8 veces mejoradas en comparación con el control, respectivamente.

En la producción # 2, la proporción total/completa es 1.4 veces inferior o 1.4 veces más alta para la 1:1 o 5:1 producción.

En conclusión, la proporción total/completa parece también estar mejorada para las partículas rAAV1 producidas

con Bac.VD190 (como fue observada para Bac.VD165).

[0121] Para Bac.VD194, el análisis de electrotransferencia se efectuó para determinar la expresión Rep y Cap. Por consiguiente, tres días después las células fueron infectadas con 2 proporciones de baculovirus diferentes (es decir 1:1 y 9:1, con 5:1:1 para los tres componentes de control) lisatos celulares fueron cosechados y sometidos al análisis de electrotransferencia.

Como se muestra en la figura 11A, la expresión CAP es baja en la producción con la proporción 1:1 de baculovirus, pero la producción con la proporción 9:1 es comparable a los tres componentes de control.

Sin embargo, para las producciones de proporción 1:1 y 9:1, la expresión Rep se reduce espectacularmente en comparación con los tres componentes de control.

La Figura 11B, sin embargo, muestra que la producción de virus fue superior en las producciones con Bac.VD194.

El hecho de que la producción sea más alta, pero la expresión CAP sea inferior indica un total más favorable: proporción completa.

Nuevamente, la conclusión es que los dos sistemas de componentes con cantidades equimolares Rep y Cap, en particular con expresión alta Rep y Cap moderada, lleva a una proporción inferior total:parcial (es decir, un porcentaje superior de partículas rellenas).

Ejemplo 2

Materiales y métodos

[0122] Tres matraces de agitación (todos cultivados a 28°C antes de la infección) fueron inoculados con inóculo P5 4 mL Bac.VD43, 4 mL Bac.VD84 y 20 mL Bac.VD88.

Los agitadores fueron incubados durante el 72 horas de producción vírica a tres temperaturas diferentes (26, 28 y 30°C).

[0123] El tratamiento de lisis de todos los matraces de agitación se efectuó dentro de 1 incubador de agitación utilizando 0.9% (v/v) Tritón X-100.

Después de la adición de tampón de lisis, los agitadores fueron incubados durante 30 minutos a un valor de ajuste de 28°C, la benzonasa fue añadida (16 µL por 400 mL) y los agitadores fueron incubados durante 1 hora a un valor de ajuste de 37°C.

Después del tratamiento de benzonasa, todos los matraces de agitación fueron colocados para aclaramiento vírico durante toda la noche a 29°C, seguidos de almacenamiento a 4°C durante máximo 3 días.

[0124] Fuera de cada incubador de agitación, 200 ml de volumen lisado crudo fue procesado en una columna de afinidad de 1 mL utilizando un caudal de 1 mL/min.

Después del lavado, el producto fue eluido utilizando PBS, pH 3.5.

Resultados

[0125] Las concentración de células en proceso de los matraces de agitación usados para el estudio de temperatura se muestran en la tabla 1.

Un aumento de temperatura correlaciona con una reducción de concentración de células total y viabilidad en el momento de la cosecha.

La viabilidad obtenida a 28°C fue comparable a los resultados obtenidos con el sistema de biorreactor de onda.

Tabla 1: producciones de agitación de recuentos de célula en proceso

Matraz de coctelera (Lote N°)	Antes de la infección		Antes de la adición del tampón de lisis	
	Célula Total conc. (#/ml)	Viabilidad (%)	Célula Total conc. (#/ml)	Viabilidad (%)
Temperatura de infección de matraz de agitación 26°C (Lote N° 0078)	1.89x10 ⁶	99.5	2.74x10 ⁶	85.7
Temperatura de infección de matraz de agitación 28°C (Lote N° 0078)	1.89x10 ⁶	99.5	2.51 X10 ⁶	74.9
Temperatura de infección de matraz de agitación 30°C (Lote N° 0078)	1.89x10 ⁶	99.5	2.42 X10 ⁶	61.0

ES 2 644 880 T3

[0126] Los resultados de la prueba de las masas lisadas crudas y eluatos de los matraces de agitación se muestran en la tabla 2.

Tabla 2: proporción total:completa de partículas rAAV producidas a temperaturas diferentes.

5

Matraz de agitación (Lote N°)	Partículas Totales (TP/ml) método de HPLC	Q-PCR CLB (GC/ml)	Eluato Q-PCR (GC/ml)	Proporción total:completa (TP/GC)
Matraz de agitación 26°C (Lote N° 0078)	6.14×10^{12}	1.28×10^{10}	5.52×10^{10}	111
Matraz de agitación 28°C (Lote N° 0078)	6.58×10^{12}	2.00×10^{10}	1.43×10^{11}	46
Matraz de agitación 30°C (Lote N° 0078)	7.24×10^{12}	3.94×10^{10}	1.73×10^{11}	42

Discusión y conclusión:

10 [0127] Aumentar la temperatura durante la producción rAAV en células SF⁺, aumenta ligeramente la cantidad total de partículas rAAV que se producen, pero aumenta la cantidad de partículas completas que se producen más, conduciendo a una proporción disminuida total:completa.

LISTADO DE SECUENCIAS

[0128]

<110> AMSTERDAM MOLECULAR THERAPEUTICS B.V.

<120> Optimización de expresión de proteínas Rep y Cap parvovirales en las células de insecto

5 <130> P6019170

<150> EP 08151634.6

<151> 2008-02-19

<150> US 61/029,673

<151> 2008-02-19

10 <160> 31

<170> PatentIn version 3.3

<210> 1

<211> 1194

<212> ADN

15 <213> virus adeno-asociado 2

<220>

<221> misc_feature

<223> rep52 tipo salvaje

<400> 1

```

atggagctgg tcgggtggct cgtggacaag gggattacct cggagaagca gtggatccag    60
gaggaccagg cctcatatc ctccttcaat gcggcctcca actcgcggtc ccaaatacaag    120
gctgccttgg acaatgcggg aaagattatg agcctgacta aaaccgcccc cgactacctg    180
gtgggccagc agcccgtgga ggacatttcc agcaatcgga tttataaaat tttggaacta    240
aacgggtacg atccccaata tgcggcttcc gtctttctgg gatgggccac gaaaaagttc    300
ggcaagagga acaccatctg gctgtttggg cctgcaacta ccggaagac caacatcggg    360
gaggccatag cccacactgt gcccttctac gggcgcgtaa actggacca tgagaacttt    420
cccttcaacg actgtgtcga caagatggtg atctgggtgg aggaggggaa gatgaccgcc    480
aaggtcgtgg agtcggccaa agccattctc ggaggaagca aggtgcgcgt ggaccagaaa    540
tgcaagtcct cggcccagat agaccgact cccgtgatcg tcacctcaa caccaacatg    600
tgcgcctgta ttgacgggaa ctcaacgacc ttcgaacacc agcagccggt gcaagaccgg    660
atgttcaaat ttgaactcac ccgccgtctg gatcatgact ttgggaaggt caccaagcag    720
gaagtcaaag actttttccg gtgggcaaag gatcacgtgg ttgaggtgga gcatgaattc    780
tacgtcaaaa aggggtggagc caagaaaaga cccgccccca gtgacgcaga tataagtgag    840
cccaaaccgg tgcgcgagtc agttgcgag ccatcgacgt cagacgcgga agcttcgatc    900
aactacgcag accgctacca aaacaaatgt tctcgtcacg tgggcatgaa tctgatgctg    960
tttcctgca gacaatgcga gagaatgaat cagaattcaa atatctgctt cactcacgga   1020
cagaaagact gtttagagtg ctttcccgtg tcagaatctc aaccggtttc tgtcgtcaaa   1080
aaggcgtatc agaaactgtg ctacattcat catatcatgg gaaaggtgcc agacgcttgc   1140
actgcctgcg atctggtcaa tgtggatttg gatgactgca tctttgaaca ataa         1194

```

20 <210> 2

<211> 1194

ES 2 644 880 T3

<212> ADN

<213> adeno-associated virus 2

<220>

<221> misc_feature

5 <223> rep52 sf9 (célula de insecto) optimizado

<400> 2

```

atggagctgg tgggttggt ggtggacaag ggtatcacct ccgagaagca gtggatccag    60
gaggaccagg cttcctacat ctccctcaac gctgcttcca actcccgttc ccagatcaag    120
gctgctctgg acaacgctgg taagatcatg tccttgacca agaccgctcc tgactacctg    180
gtgggtcagc agcctgtgga ggacatctcc tccaaccgta tctacaagat cctggagctg    240
aacggttacg accctcagta cgctgcttcc gtttctctgg gttgggctac caagaagttc    300
ggtaagcgta acaccatctg gctgttcggt cctgctacca ccgtaagac caacatcgct    360
gaggctatcg ctcacaccgt gcctttctac ggtgctgta actggaccaa cgagaacttc    420
ccttcaacg actgctgga caagatggtg atctggtggg aggagggtaa gatgaccgct    480
aagggtggtg agtccgctaa ggctatcctg ggtggttcca aggtgctgtg ggaccagaag    540
tgcaagtctc ccgctcagat cgaccctacc cctgtgatcg tgacctcaa caccaacatg    600
tgcgctgtga tcgacggtaa ctccaccacc ttcgagcacc agcagcctct gcaggaccgt    660
atgttcaagt tcgagctgac ccgctgctg gaccacgact tcggttaagg gaccaagcag    720
gagggtgaag atttcttccg ttgggctaag gaccacgtgg tggagggtga gcacgagttc    780
tacgtgaaga agggtggtgc taagaagcgt cctgctcctt ccgacgctga catctccgag    840
cctaagcgtg tgcgtgagtc cgtggctcag ccttccacct ccgacgctga ggcttccatc    900
aactacgctg accgttacca gaacaagtgc tccgctcacg tgggtatgaa cctgatgctg    960
ttcccttgcc gtcagtgcga gcgatgaac cagaactcca acatctgctt caccacggt    1020
cagaaggact gcctggagtg cttccctgtg tccgagctcc agcctgtgtc cgtggtgaag    1080
aaggcttacc agaagctgtg ctacatccac cacatcatgg gtaagggtgcc tgacgcttgc    1140
accgcttgcg acctggtgaa cgtggacctg gacgactgca tcttcgagca gtaa        1194

```

<210> 3

<211> 1194

10 <212> ADN

<213> virus adeno-asociado 2

<220>

<221> misc_feature

<223> rep52 AT optimizado

15 <400> 3

```

atggaattag taggatggtt agtagataaa ggaataacat cagaaaaaca atggatacaa    60
gaagatcaag catcatatat atcatttaat gcagcatcaa attcaagatc acaataaaaa    120

```

ES 2 644 880 T3

gcagcattag ataatgcagg aaaaataatg tcattaacaa aaacagcacc agattattta 180
 gtaggacaac aaccagtaga agatatatca tcaaatagaa tatataaaat attagaatta 240
 aatggatag atccacaata tgcagcatca gtatttttag gatgggcaac aaaaaattt 300
 ggaaaaagaa atacaatatg gttatttggg ccagcaacaa caggaaaaac aaatatagca 360
 gaagcaatag cacatacagt accattttat ggatgtgtaa attggacaaa tgaaaatttt 420
 ccatttaatg attgtgtaga taaaatggta atatgggtggg aagaaggaaa aatgacagca 480
 aaagtagtag aatcagcaaa agcaatatta ggaggatcaa aagtaagagt agatcaaaaa 540
 tgtaaatcat cagcacaat agatccaaca ccagtaatag taacatcaaa tacaatatg 600
 tgtgcagtaa tagatggaaa ttcaacaaca tttgaacatc aacaaccatt acaagataga 660
 atgtttaa atgtgaattaac aagaagatta gatcatgatt ttggaaaagt aacaaaacaa 720
 gaagtaaaag attttttttag atgggcaaaa gatcatgtag tagaagtaga acatgaattt 780
 tatgtaaaa aaggaggagc aaaaaaaga ccagcaccat cagatgcaga tatatcagaa 840
 ccaaaaagag taagagaatc agtagcaca ccatcaacat cagatgcaga agcatcaata 900
 aattatgcag atagatatca aaataaatgt tcaagacatg taggaatgaa tttaatgtta 960
 tttccatgta gacaatgtga aagaatgaat caaaattcaa atatatgttt tacacatgga 1020
 caaaaagatt gtttagaatg ttttccagta tcagaatcac aaccagtatc agtagtaaaa 1080
 aaagcatatc aaaaattatg ttatatacat catataatgg gaaaagtacc agatgcatgt 1140
 acagcatgtg atttagtaaa tgtagattta gatgattgta tatttgaaca ataa 1194

<210> 4

<211> 1194

<212> ADN

5 <213> virus adeno-asociado 2

<220>

<221> misc_feature

<222> (1)..(1194)

<223> rep52 GC optimizado

10 <400> 4

atggagctgg tggggtggct ggtggacaag gggatcacga gcgagaagca gtggatccag 60
 gaggaccagg cgagctacat cagcttcaac gcggcgagca acagccggag ccagatcaag 120
 gcggcgctgg acaacgcggg gaagatcatg agcctgacga agacggcgcc ggactacctg 180
 gtggggcagc agccggtgga ggacatcagc agcaaccgga tctacaagat cctggagctg 240
 aacgggtacg acccgcagta cgcggcgagc gtgttcctgg ggtgggagac gaagaagttc 300
 gggaaagcga acacgatctg gctgttcggg ccggcgacga cggggaagac gaacatcgcg 360
 gaggcgatcg cgcacacggt gccgttctac ggggtcgtga actggacgaa cgagaacttc 420
 ccgttcaacg actgcgtgga caagatggtg atctgggtggg aggaggggaa gatgacggcg 480
 aagtggtgag agagcgcgaa ggcgatcctg ggggggagca aggtgcgggt ggaccagaag 540
 tgcaagagca gcgcgcagat cgacccgacg ccggtgatcg tgacgagcaa cacgaacatg 600

ES 2 644 880 T3

tgcgcggtga tgcacgggaa cagcacgacg ttcgagcacc agcagccgct gcaggaccgg 660
 atgttcaagt tcgagctgac gcggcggtcg gaccacgact tcgggaaggt gacgaagcag 720
 gaggtgaagg acttcttccg gtgggcgaag gaccacgtgg tggaggtgga gcacgagttc 780
 tacgtgaaga aggggggggc gaagaagcgg ccggcgccga gcgacgcgga catcagcgag 840
 ccgaagcggg tgcgggagag cgtggcgacg ccgagcacga gcgacgcgga ggcgagcatc 900
 aactacgcgg accggtacca gaacaagtgc agccggcacg tggggatgaa cctgatgctg 960
 ttcccgtgcc ggcagtgcga gcggatgaac cagaacagca acatctgctt cacgcacggg 1020
 cagaaggact gcctggagtg cttcccggtg agcgagagcc agccggtgag cgtggtgaag 1080
 aaggcgtacc agaagctgtg ctacatccac cacatcatgg ggaaggtgcc ggacgcgtgc 1140
 acggcgtgcg acctggtgaa cgtggacctg gacgactgca tcttcgagca gtaa 1194

<210> 5

<211> 1194

<212> ADN

5 <213> virus adeno-asociado 2

<220>

<221> CDS

<222> (1)..(1194)

<223> Rep52

10 <400> 5

atg gag ctg gtc ggg tgg ctc gtg gac aag ggg att acc tcg gag aag 48
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
 1 5 10 15
 cag tgg atc cag gag gac cag gcc tca tac atc tcc ttc aat gcg gcc 96
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
 20 25 30
 tcc aac tcg cgg tcc caa atc aag gct gcc ttg gac aat gcg gga aag 144
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
 35 40 45
 att atg agc ctg act aaa acc gcc ccc gac tac ctg gtg ggc cag cag 192
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
 50 55 60
 ccc gtg gag gac att tcc agc aat cgg att tat aaa att ttg gaa cta 240
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
 65 70 75 80
 aac ggg tac gat ccc caa tat gcg gct tcc gtc ttt ctg gga tgg gcc 288
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
 85 90 95
 acg aaa aag ttc ggc aag agg aac acc atc tgg ctg ttt ggg cct gca 336
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
 100 105 110
 act acc ggg aag acc aac atc gcg gag gcc ata gcc cac act gtg ccc 384
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
 115 120 125
 ttc tac ggg tgc gta aac tgg acc aat gag aac ttt ccc ttc aac gac 432
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
 130 135 140

ES 2 644 880 T3

tgt gtc gac aag atg gtg atc tgg tgg gag gag ggg aag atg acc gcc 480
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
 145 150 155 160
 aag gtc gtg gag tcg gcc aaa gcc att ctc gga gga agc aag gtg cgc 528
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
 165 170 175
 gtg gac cag aaa tgc aag tcc tcg gcc cag ata gac ccg act ccc gtg 576
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
 180 185 190
 atc gtc acc tcc aac acc aac atg tgc gcc gtg att gac ggg aac tca 624
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
 195 200 205
 acg acc ttc gaa cac cag cag ccg ttg caa gac cgg atg ttc aaa ttt 672
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
 210 215 220
 gaa ctc acc cgc cgt ctg gat cat gac ttt ggg aag gtc acc aag cag 720
 Glu Leu Thr Arg Arg Arg Asp His Asp Phe Gly Lys Val Thr Lys Gln
 225 230 235
 gaa gtc aaa gac ttt ttc cgg tgg gca aag gat cac gtg gtt gag gtg 768
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Val Glu Val
 245 250 255
 gag cat gaa ttc tac gtc aaa aag ggt gga gcc aag aaa aga ccc gcc 816
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
 260 265 270
 ccc agt gac gca gat ata agt gag ccc aaa cgg gtg cgc gag tca gtt 864
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
 275 280 285
 gcg cag cca tcg acg tca gac gcg gaa gct tcg atc aac tac gca gac 912
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
 290 295 300
 cgc tac caa aac aaa tgt tct cgt cac gtg ggc atg aat ctg atg ctg 960
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
 305 310 315 320
 ttt ccc tgc aga caa tgc gag aga atg aat cag aat tca aat atc tgc 1008
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
 325 330 335
 ttc act cac gga cag aaa gac tgt tta gag tgc ttt ccc gtg tca gaa 1056
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu
 340 345 350
 tct caa ccc gtt tct gtc gtc aaa aag gcg tat cag aaa ctg tgc tac 1104
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
 355 360 365
 att cat cat atc atg gga aag gtg cca gac gct tgc act gcc tgc gat 1152
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
 370 375 380
 ctg gtc aat gtg gat ttg gat gac tgc atc ttt gaa caa taa 1194
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
 385 390 395

<210> 6

<211> 397

<212> PRT

5 <213> virus adeno-asociado 2

<400> 6

ES 2 644 880 T3

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
 1 5 10 15

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
 20 25 30

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
 35 40 45

Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
 50 55 60

Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
 65 70 75 80

Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
 85 90 95

Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
 100 105 110

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
 115 120 125

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
 130 135 140

Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
 145 150 155 160

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
 165 170 175

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
 180 185 190

Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
 195 200 205

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
 210 215 220

Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
 225 230 235 240

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
 245 250 255

Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
 260 265 270

ES 2 644 880 T3

Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
 275 280 285

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
 290 295 300

Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
 305 310 315 320

Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
 325 330 335

Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu
 340 345 350

Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
 355 360 365

Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
 370 375 380

Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
 385 390 395

<210> 7

<211> 1876

<212> ADN

5 <213> virus adeno-asociado 2

<220>

<221> CDS

<222> (11)..(1876)

<223> Rep78

10 <400> 7

cgcagccgcc	atg ccg ggg ttt tac gag att gtg att aag gtc ccc agc	49
	Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser	
	1 5 10	
gac ctt gac gag cat ctg ccc ggc att tct gac agc ttt gtg aac tgg	97	
Asp Leu Asp Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp		
	15 20 25	
gtg gcc gag aag gaa tgg gag ttg ccg cca gat tct gac atg gat ctg	145	
Val Ala Glu Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu		
	30 35 40 45	
aat ctg att gag cag gca ccc ctg acc gtg gcc gag aag ctg cag cgc	193	
Asn Leu Ile Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg		
	50 55 60	
gac ttt ctg acg gaa tgg cgc cgt gtg agt aag gcc ccg gag gcc ctt	241	
Asp Phe Leu Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu		
	65 70 75	
ttc ttt gtg caa ttt gag aag gga gag agc tac ttc cac atg cac gtg	289	
Phe Phe Val Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val		
	80 85 90	

ES 2 644 880 T3

ctc Leu	gtg Val 95	gaa Glu	acc Thr	acc Thr	ggg Gly	gtg Val 100	aaa Lys	tcc Ser	atg Met	gtt Val	ttg Leu 105	gga Gly	cgt Arg	ttc Phe	ctg Leu	337
agt Ser 110	cag Gln	att Ile	cgc Arg	gaa Glu	aaa Lys 115	ctg Leu	att Ile	cag Gln	aga Arg	att Ile 120	tac Tyr	cgc Arg	ggg Gly	atc Ile	gag Glu 125	385
ccg Pro	act Thr	ttg Leu	cca Pro	aac Asn 130	tgg Trp	ttc Phe	gcg Ala	gtc Val	aca Thr 135	aag Lys	acc Thr	aga Arg	aat Asn	ggc Gly 140	gcc Ala	433
gga Gly	ggc Gly	ggg Gly	aac Asn 145	aag Lys	gtg Val	gtg Val	gat Asp	gag Glu 150	tgc Cys	tac Tyr	atc Ile	ccc Pro	aat Asn 155	tac Tyr	ttg Leu	481
ctc Leu	ccc Pro	aaa Lys 160	acc Thr	cag Gln	cct Pro	gag Glu	ctc Leu 165	cag Gln	tgg Trp	gcg Ala	tgg Trp	act Thr 170	aat Asn	atg Met	gaa Glu	529
cag Gln	tat Tyr 175	tta Leu	agc Ser	gcc Ala	tgt Cys	ttg Leu 180	aat Asn	ctc Leu	acg Thr	gag Glu	cgt Arg 185	aaa Lys	cgg Arg	ttg Leu	gtg Val	577
gcg Ala 190	cag Gln	cat His	ctg Leu	acg Thr	cac His 195	gtg Val	tcg Ser	cag Gln	acg Thr	cag Gln 200	gag Glu	cag Gln	aac Asn	aaa Lys	gag Glu 205	625
aat Asn	cag Gln	aat Asn	ccc Pro	aat Asn 210	tct Ser	gat Asp	gcg Ala	ccg Pro	gtg Val 215	atc Ile	aga Arg	tca Ser	aaa Lys	act Thr 220	tca Ser	673
gcc Ala	agg Arg	tac Tyr	atg Met 225	gag Glu	ctg Leu	gtc Val	ggg Gly	tgg Trp 230	ctc Leu	gtg Val	gac Asp	aag Lys	ggg Gly 235	att Ile	acc Thr	721
tcg Ser	gag Glu	aag Lys 240	cag Gln	tgg Trp	atc Ile	cag Gln	gag Glu 245	gac Asp	cag Gln	gcc Ala	tca Ser	tac Tyr 250	atc Ile	tcc Ser	ttc Phe	769
aat Asn	gcg Ala 255	gcc Ala	tcc Ser	aac Asn	tcg Ser	cgg Arg 260	tcc Ser	caa Gln	atc Ile	aag Lys	gct Ala 265	gcc Ala	ttg Leu	gac Asp	aat Asn	817
gcg Ala 270	gga Gly	aag Lys	att Ile	atg Met 275	agc Ser	ctg Leu	act Thr	aaa Lys	acc Thr	gcc Ala 280	ccc Pro	gac Asp	tac Tyr	ctg Leu	gtg Val 285	865
ggc Gly	cag Gln	cag Gln	ccc Pro	gtg Val 290	gag Glu	gac Asp	att Ile	tcc Ser	agc Ser 295	aat Asn	cgg Arg	att Ile	tat Tyr	aaa Lys 300	att Ile	913
ttg Leu	gaa Glu	cta Leu	aac Asn 305	ggg Gly	tac Tyr	gat Asp	ccc Pro	caa Gln 310	tat Tyr	gcg Ala	gct Ala	tcc Ser	gtc Val 315	ttt Phe	ctg Leu	961
gga Gly	tgg Trp	gcc Ala 320	acg Thr	aaa Lys	aag Lys	ttc Phe	ggc Gly 325	aag Lys	agg Arg	aac Asn	acc Thr	atc Ile 330	tgg Trp	ctg Leu	ttt Phe	1009
ggg Gly	cct Pro 335	gca Ala	act Thr	acc Thr	ggg Gly	aag Lys 340	acc Thr	aac Asn	atc Ile	gcg Ala	gag Glu 345	gcc Ala	ata Ile	gcc Ala	cac His	1057
act Thr 350	gtg Val	ccc Pro	ttc Phe	tac Tyr	ggg Gly 355	tgc Cys	gta Val	aac Asn	tgg Trp	acc Thr 360	aat Asn	gag Glu	aac Asn	ttt Phe	ccc Pro 365	1105

ES 2 644 880 T3

ttc aac gac tgt gtc gac aag atg gtg atc tgg tgg gag gag ggg aag 1153
 Phe Asn Asp Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys
 370 375 380
 atg acc gcc aag gtc gtg gag tcg gcc aaa gcc att ctc gga gga agc 1201
 Met Thr Ala Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser
 385 390 395
 aag gtg cgc gtg gac cag aaa tgc aag tcc tcg gcc cag ata gac ccg 1249
 Lys Val Arg Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro
 400 405 410
 act ccc gtg atc gtc acc tcc aac acc aac atg tgc gcc gtg att gac 1297
 Thr Pro Val Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp
 415 420 425
 ggg aac tca acg acc ttc gaa cac cag cag ccg ttg caa gac cgg atg 1345
 Gly Asn Ser Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met
 430 435 440 445
 ttc aaa ttt gaa ctc acc cgc cgt ctg gat cat gac ttt ggg aag gtc 1393
 Phe Lys Phe Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val
 450 455 460
 acc aag cag gaa gtc aaa gac ttt ttc cgg tgg gca aag gat cac gtg 1441
 Thr Lys Gln Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val
 465 470 475
 gtt gag gtg gag cat gaa ttc tac gtc aaa aag ggt gga gcc aag aaa 1489
 Val Glu Val Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys
 480 485 490
 aga ccc gcc ccc agt gac gca gat ata agt gag ccc aaa cgg gtg cgc 1537
 Arg Pro Ala Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg
 495 500 505
 gag tca gtt gcg cag cca tcg acg tca gac gcg gaa gct tcg atc aac 1585
 Glu Ser Val Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn
 510 515 520 525
 tac gca gac agg tac caa aac aaa tgt tct cgt cac gtg ggc atg aat 1633
 Tyr Ala Asp Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn
 530 535 540
 ctg atg ctg ttt ccc tgc aga caa tgc gag aga atg aat cag aat tca 1681
 Leu Met Leu Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser
 545 550 555
 aat atc tgc ttc act cac gga cag aaa gac tgt tta gag tgc ttt ccc 1729
 Asn Ile Cys Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro
 560 565 570
 gtg tca gaa tct caa ccc gtt tct gtc gtc aaa aag gcg tat cag aaa 1777
 Val Ser Glu Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys
 575 580 585
 ctg tgc tac att cat cat atc atg gga aag gtg cca gac gct tgc act 1825
 Leu Cys Tyr Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr
 590 595 600 605
 gcc tgc gat ctg gtc aat gtg gat ttg gat gac tgc atc ttt gaa caa 1873
 Ala Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
 610 615 620
 taa 1876

<210> 8

<211> 621

<212> PRT

5 <213> virus adeno-asociado 2

<400> 8

ES 2 644 880 T3

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 Leu 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

Thr Glu 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 Met 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 Ile Gln Arg 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 Gln Tyr Leu
165 170 175

Ser Ala Cys Leu Asn Leu Thr 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 Gln 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 Lys 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

ES 2 644 880 T3

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

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

Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
 290 295 300

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

Thr 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 Thr 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 Asp His Asp Phe Gly Lys Val Thr Lys Gln
 450 455 460

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
 465 470 475 480

Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
 485 490 495

Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
 500 505 510

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
 515 520 525

Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu

ES 2 644 880 T3

530

535

540

Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
545 550 555 560

Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu
565 570 575

Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
580 585 590

Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
595 600 605

Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
610 615 620

<210> 9

<211> 9

<212> ADN

5 <213> virus adeno-asociado 2

<220>

<221> misc_feature

<223> 7 nt iniciación

<400> 9

10 cctgttaag 9

<210> 10

<211> 28

<212> ADN

<213> Artificial

15 <220>

<223> primera secuencia

<220>

<221> misc_feature

<223> HR1-Fw

20 <400> 10

gtatacgtat gacactatcg atgtgac 28

<210> 11

<211> 30

<212> ADN

25 <213> Artificial

<220>

<223> primera secuencia

<220>

<221> misc_feature

<223> HR1-Rv
 <400> 11
 gtatacgatc gattattgct ccaatactag 30
 <210> 12
 5 <211> 25
 <212> ADN
 <213> Artificial
 <220>
 <223> primera secuencia
 10 <220>
 <221> misc_feature
 <223> polyA Fw
 <400> 12
 agatctgtag tggctatggc agggc 25
 15 <210> 13
 <211> 35
 <212> DNA
 <213> Artificial
 <220>
 20 <223> primera secuencia
 <220>
 <221> misc_feature
 <223> p10 Rv
 <400> 13
 25 agatctcccg ggacggacct ttaattcaac ccaac 35
 <210> 14
 <211> 33
 <212> ADN
 <213> Artificial
 30 <220>
 <223> primera secuencia
 <220>
 <221> misc_feature
 <223> BstZ17I-deltaIE1 Fw
 35 <400> 14
 ggtccgtata cgacgataac gccgttggtg gcg 33
 <210> 15
 <211> 30
 <212> DNA
 40 <213> Artificial

<220>
 <223> primera secuencia
 <220>
 <221> misc_feature
 5 <223> AfIII-delta IE1 Rv
 <400> 15
 cgacttaaga cggcgaattc tgcagatggc 30
 <210> 16
 <211> 37
 10 <212> ADN
 <213> Artificial
 <220>
 <223> primera secuencia
 <220>
 15 <221> misc_feature
 <223> Smal-ppolh Fw
 <400> 16
 tctcccggga gatcatggag ataattaaaa tgataac 37
 <210> 17
 20 <211> 21
 <212> ADN
 <213> Artificial
 <220>
 <223> primera secuencia
 25 <220>
 <221> misc_feature
 <223> Spel-ppolh Rv
 <400> 17
 gttactagtg agctcgtcga c 21
 30 <210> 18
 <211> 23
 <212> ADN
 <213> Artificial
 <220>
 35 <223> primera secuencia
 <220>
 <221> misc_feature
 <223> BstZ17I-p10 Fw
 <400> 18
 40 agtatacggga cctttaattc aac 23

<210> 19
 <211> 28
 <212> ADN
 <213> Artificial
 5 <220>
 <223> primera secuencia
 <220>
 <221> misc_feature
 <223> AfIII-p10 Rv
 10 <400> 19
 cgacttaaga gcgggcccgt ttcgaatc 28
 <210> 20
 <211> 2211
 <212> ADN
 15 <213> virus 1 adeno-asociado; VP1, VP2, VP3; codón de inicio VP1 alterado (GTG)
 <220>
 <221> CDS
 <222> (1)..(2208)
 <223> VP1
 20 <220>
 <221> misc_feature
 <222> (412)..(2208)
 <223> VP2
 <220>
 25 <221> misc_feature
 <222> (607)..(2208)
 <223> VP3
 <400> 20

ES 2 644 880 T3

gtg Val 1	gct Ala	gcc Ala	gac Asp	ggt Gly 5	tat Tyr	cta Leu	ccc Pro	gat Asp	tgg Trp 10	ctc Leu	gag Glu	gac Asp	aac Asn	ctc Leu 15	tct Ser	48
gag Glu	ggc Gly	att Ile	cgc Arg 20	gag Glu	tgg Trp	tgg Trp	gac Asp	ttg Leu 25	aaa Lys	cct Pro	gga Gly	gcc Ala	ccg Pro 30	aag Lys	ccc Pro	96
aaa Lys	gcc Ala	aac Asn 35	cag Gln	caa Gln	aag Lys	cag Gln	gac Asp 40	gac Asp	ggc Gly	cgg Arg	ggt Gly	ctg Leu 45	gtg Val	ctt Leu	cct Pro	144
ggc Gly	tac Tyr 50	aag Lys	tac Tyr	ctc Leu	gga Gly	ccc Pro 55	ttc Phe	aac Asn	gga Gly	ctc Leu	gac Asp 60	aag Lys	ggg Gly	gag Glu	ccc Pro	192
gtc Val 65	aac Asn	gcg Ala	gcg Ala	gac Asp	gca Ala 70	gcg Ala	gcc Ala	ctc Leu	gag Glu	cac His 75	gac Asp	aag Lys	gcc Ala	tac Tyr	gac Asp 80	240
cag Gln	cag Gln	ctc Leu	aaa Lys	gcg Ala 85	ggt Gly	gac Asp	aat Asn	ccg Pro	tac Tyr 90	ctg Leu	cgg Arg	tat Tyr	aac Asn	cac His 95	gcc Ala	288
gac Asp	gcc Ala	gag Glu	ttt Phe 100	cag Gln	gag Glu	cg Arg	ctg Leu	caa Gln 105	gaa Glu	gat Asp	acg Thr	tct Ser	ttt Phe 110	ggg Gly	ggc Gly	336
aac Asn	ctc Leu	ggg Gly 115	cga Arg	gca Ala	gtc Val	ttc Phe	cag Gln 120	gcc Ala	aag Lys	aag Lys	cgg Arg	gtt Val 125	ctc Leu	gaa Glu	cct Pro	384
ctc Leu	ggt Gly 130	ctg Leu	gtt Val	gag Glu	gaa Glu	ggc Gly 135	gct Ala	aag Lys	acg Thr	gct Ala	cct Pro 140	gga Gly	aag Lys	aaa Lys	cg Arg	432
ccg Pro 145	gta Val	gag Glu	cag Gln	tcg Ser	cca Pro 150	caa Gln	gag Glu	cca Pro	gac Asp	tcc Ser 155	tcc Ser	tcg Ser	ggc Gly	atc Ile	ggc Gly 160	480
aag Lys	aca Thr	ggc Gly	cag Gln	cag Gln	ccc Pro	gct Ala	aaa Lys	aag Lys	aga Arg 170	ctc Leu	aat Asn	ttt Phe	ggt Gly	cag Gln 175	act Thr	528
ggc Gly	gac Asp	tca Ser	gag Glu 180	tca Ser	gtc Val	ccc Pro	gat Asp	cca Pro 185	caa Gln	cct Pro	ctc Leu	gga Gly	gaa Glu 190	cct Pro	cca Pro	576

ES 2 644 880 T3

gca Ala	acc Thr	ccc Pro 195	gct Ala	gct Ala	gtg Val	gga Gly	cct Pro 200	act Thr	aca Thr	atg Met	gct Ala	tca Ser 205	ggc Gly	ggg Gly	ggc Gly	624
gca Ala	cca Pro 210	atg Met	gca Ala	gac Asp	aat Asn	aac Asn 215	gaa Glu	ggc Gly	gcc Ala	gac Asp	gga Gly 220	gtg Val	ggt Gly	aat Asn	gcc Ala	672
tca Ser 225	gga Gly	aat Asn	tgg Trp	cat His	tgc Cys 230	gat Asp	tcc Ser	aca Thr	tgg Trp	ctg Leu 235	ggc Gly	gac Asp	aga Arg	gtc Val	atc Ile 240	720
acc Thr	acc Thr	agc Ser	acc Thr	cgc Arg 245	acc Thr	tgg Trp	gcc Ala	ttg Leu	ccc Pro 250	acc Thr	tac Tyr	aat Asn	aac Asn	cac His 255	ctc Leu	768
tac Tyr	aag Lys	caa Gln	atc Ile 260	tcc Ser	agt Ser	gct Ala	tca Ser	acg Thr 265	ggg Gly	gcc Ala	agc Ser	aac Asn	gac Asp 270	aac Asn	cac His	816
tac Tyr	ttc Phe	ggc Gly 275	tac Tyr	agc Ser	acc Thr	ccc Pro	tgg Trp 280	ggg Gly	tat Tyr	ttt Phe	gat Asp	ttc Phe 285	aac Asn	aga Arg	ttc Phe	864
cac His	tgc Cys 290	cac His	ttt Phe	tca Ser	cca Pro	cgt Arg 295	gac Asp	tgg Trp	cag Gln	cga Arg	ctc Leu 300	atc Ile	aac Asn	aac Asn	aat Asn	912
tgg Trp 305	gga Gly	ttc Phe	cgg Arg	ccc Pro	aag Lys 310	aga Arg	ctc Leu	aac Asn	ttc Phe	aaa Lys 315	ctc Leu	ttc Phe	aac Asn	atc Ile	caa Gln 320	960
gtc Val	aag Lys	gag Glu	gtc Val	acg Thr 325	acg Thr	aat Asn	gat Asp	ggc Gly	gtc Val 330	aca Thr	acc Thr	atc Ile	gct Ala	aat Asn 335	aac Asn	1008
ctt Leu	acc Thr	agc Ser	acg Thr 340	gtt Val	caa Gln	gtc Val	ttc Phe	tcg Ser 345	gac Asp	tcg Ser	gag Glu	tac Tyr	cag Gln 350	ctt Leu	ccg Pro	1056
tac Tyr	gtc Val	ctc Leu 355	ggc Gly	tct Ser	gcg Ala	cac His	cag Gln 360	ggc Gly	tgc Cys	ctc Leu	cct Pro	ccg Pro 365	ttc Phe	ccg Pro	gcg Ala	1104
gac Asp	gtg Val 370	ttc Phe	atg Met	att Ile	ccg Pro	caa Gln 375	tac Tyr	ggc Gly	tac Tyr	ctg Leu	acg Thr 380	ctc Leu	aac Asn	aat Asn	ggc Gly	1152
agc Ser 385	caa Gln	gcc Ala	gtg Val	gga Gly	cgt Arg 390	tca Ser	tcc Ser	ttt Phe	tac Tyr	tgc Cys 395	ctg Leu	gaa Glu	tat Tyr	ttc Phe	cct Pro 400	1200
tct Ser	cag Gln	atg Met	ctg Leu	aga Arg 405	acg Thr	ggc Gly	aac Asn	aac Asn	ttt Phe 410	acc Thr	ttc Phe	agc Ser	tac Tyr	acc Thr 415	ttt Phe	1248
gag Glu	gaa Glu	gtg Val	cct Pro 420	ttc Phe	cac His	agc Ser	agc Ser	tac Tyr 425	gcg Ala	cac His	agc Ser	cag Gln 430	agc Ser	ctg Leu	gac Asp	1296
cgg Arg	ctg Leu	atg Met 435	aat Asn	cct Pro	ctc Leu	atc Ile	gac Asp 440	caa Gln	tac Tyr	ctg Leu	tat Tyr	tac Tyr 445	ctg Leu	aac Asn	aga Arg	1344
act Thr	caa Gln 450	aat Asn	cag Gln	tcc Ser	gga Gly	agt Ser 455	gcc Ala	caa Gln	aac Asn	aag Lys	gac Asp 460	ttg Leu	ctg Leu	ttt Phe	agc Ser	1392

ES 2 644 880 T3

cgt Arg 465	ggg Gly	tct Ser	cca Pro	gct Ala 470	ggc Gly	atg Met	tct Ser	gtt Val	cag Gln	ccc Pro 475	aaa Lys	aac Asn	tgg Trp	cta Leu	cct Pro 480	1440
gga Gly	ccc Pro	tgt Cys	tat Tyr	cgg Arg 485	cag Gln	cag Gln	cgc Arg	gtt Val	tct Ala 490	aaa Lys	aca Thr	aaa Lys	aca Thr	gac Asp 495	aac Asn	1488
aac Asn	aac Asn	agc Ser	aat Asn 500	ttt Phe	acc Thr	tgg Trp	act Thr	ggt Gly 505	gct Ala	tca Ser	aaa Lys	tat Tyr	aac Asn 510	ctc Leu	aat Asn	1536
ggg Gly	cgt Arg	gaa Glu 515	tcc Ser	atc Ile	atc Ile	aac Asn	cct Pro 520	ggc Gly	act Thr	gct Ala	atg Met	gcc Ala 525	tca Ser	cac His	aaa Lys	1584
gac Asp 530	gac Asp	gaa Glu	gac Asp	aag Lys	ttc Phe	ttt Phe 535	ccc Pro	atg Met	agc Ser	ggt Gly	gtc Val 540	atg Met	att Ile	ttt Phe	gga Gly	1632
aaa Lys 545	gag Glu	agc Ser	gcc Ala	gga Gly	gct Ala 550	tca Ser	aac Asn	act Thr	gca Ala	ttg Leu 555	gac Asp	aat Asn	gtc Val	atg Met	att Ile 560	1680
aca Thr	gac Asp	gaa Glu	gag Glu	gaa Glu 565	att Ile	aaa Lys	gcc Ala	act Thr	aac Asn 570	cct Pro	gtg Val	gcc Ala	acc Thr	gaa Glu 575	aga Arg	1728
ttt Phe	ggg Gly	acc Thr	gtg Val 580	gca Ala	gtc Val	aat Asn	ttc Phe	cag Gln 585	agc Ser	agc Ser	agc Ser	aca Thr	gac Asp 590	cct Pro	gcg Ala	1776
acc Thr	gga Gly	gat Asp 595	gtg Val	cat His	gct Ala	atg Met	gga Gly 600	gca Ala	tta Leu	cct Pro	ggc Gly	atg Met 605	gtg Val	tgg Trp	caa Gln	1824
gat Asp 610	aga Arg	gac Asp	gtg Val	tac Tyr	ctg Leu	cag Gln 615	ggt Gly	ccc Pro	att Ile	tgg Trp	gcc Ala 620	aaa Lys	att Ile	cct Pro	cac His	1872
aca Thr 625	gat Asp	gga Gly	cac His	ttt Phe	cac His 630	ccg Pro	tct Ser	cct Pro	ctt Leu	atg Met 635	ggc Gly	ggc Gly	ttt Phe	gga Gly	ctc Leu 640	1920
aag Lys	aac Asn	ccg Pro	cct Pro	cct Pro 645	cag Gln	atc Ile	ctc Leu	atc Ile	aaa Lys 650	aac Asn	acg Thr	cct Pro	gtt Val	cct Pro 655	gcg Ala	1968
aat Asn	cct Pro	ccg Pro	gcg Ala 660	gag Glu	ttt Phe	tca Ser	gct Ala	aca Thr 665	aag Lys	ttt Phe	gct Ala	tca Ser	ttc Phe 670	atc Ile	acc Thr	2016
caa Gln	tac Tyr	tcc Ser	aca Thr	gga Gly	caa Gln	gtg Val	agt Ser 680	gtg Val	gaa Glu	att Ile	gaa Glu	tgg Trp 685	gag Glu	ctg Leu	cag Gln	2064
aaa Lys 690	gaa Glu	aac Asn	agc Ser	aag Lys	cgc Arg	tgg Trp 695	aat Asn	ccc Pro	gaa Glu	gtg Val 700	cag Gln	tac Tyr	aca Thr	tcc Ser	aat Asn	2112
tat Tyr 705	gca Ala	aaa Lys	tct Ser	gcc Ala	aac Asn 710	gtt Val	gat Asp	ttt Phe	act Thr	gtg Val 715	gac Asp	aac Asn	aat Asn	gga Gly	ctt Leu 720	2160
tat Tyr	act Thr	gag Glu	cct Pro	cgc Arg 725	ccc Pro	att Ile	ggc Gly	acc Thr	cgc Arg 730	tac Tyr	ctt Leu	acc Thr	cgc Arg	ccc Pro 735	ctg Leu	2208
taa																2211
<210> 21																
<211> 736																
<212> PRT																
5	<213> virus 1 adeno-asociado; VP1, VP2, VP3; codón de inicio VP1 alterado (GTG)															
<400> 21																

ES 2 644 880 T3

Val 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 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

ES 2 644 880 T3

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
 275 280 285
 His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn
 290 300
 Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln
 305 310 315 320
 Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn
 325 330 335
 Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro
 340 345
 Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
 355 360 365
 Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
 370 380
 Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
 385 390 395 400
 Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe
 405 410 415
 Glu Glu Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp
 420 425 430
 Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg
 435 440 445
 Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu Leu Phe Ser
 450 455 460
 Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys Asn Trp Leu Pro
 465 470 475 480
 Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Lys Thr Asp Asn
 485 490 495
 Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu Asn
 500 505 510
 Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys

ES 2 644 880 T3

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> 22

<211> 1800

<212> ADN

5 <213> virus 1 adeno-asociado; VP2, VP3

<220>

<221> CDS

<222> (1)..(1797)

<223> VP2

10 <400> 22

acg gct cct gga aag aaa cgt ccg gta gag cag tcg cca caa gag cca
Thr Ala Pro Gly Lys Lys Arg Pro Val Glu Gln Ser Pro Gln Glu Pro

48

ES 2 644 880 T3

1	5	10	15													
gac Asp	tcc Ser	tcc Ser	tcg Ser 20	ggc Gly	atc Ile	ggc Gly	aag Lys	aca Thr 25	ggc Gly	cag Gln	cag Gln	ccc Pro	gct Ala 30	aaa Lys	aag Lys	96
aga Arg	ctc Leu	aat Asn 35	ttt Phe	ggt Gly	cag Gln	act Thr	ggc Gly 40	gac Asp	tca Ser	gag Glu	tca Ser	gtc Val 45	ccc Pro	gat Asp	cca Pro	144
caa Gln	cct Pro 50	ctc Leu	gga Gly	gaa Glu	cct Pro	cca Pro 55	gca Ala	acc Thr	ccc Pro	gct Ala	gct Ala 60	gtg Val	gga Gly	cct Pro	act Thr	192
aca Thr 65	atg Met	gct Ala	tca Ser	ggc Gly	ggt Gly 70	ggc Gly	gca Ala	cca Pro	atg Met	gca Ala 75	gac Asp	aat Asn	aac Asn	gaa Glu	ggc Gly 80	240
gcc Ala	gac Asp	gga Gly	gtg Val	ggt Gly 85	aat Asn	gcc Ala	tca Ser	gga Gly	aat Asn 90	tgg Trp	cat His	tgc Cys	gat Asp	tcc Ser 95	aca Thr	288
tgg Trp	ctg Leu	ggc Gly	gac Asp 100	aga Arg	gtc Val	atc Ile	acc Thr 105	acc Thr 105	agc Ser	acc Thr	cgc Arg	acc Thr	tgg Trp 110	gcc Ala	ttg Leu	336
ccc Pro	acc Thr	tac Tyr 115	aat Asn	aac Asn	cac His	ctc Leu	tac Tyr 120	aag Lys	caa Gln	atc Ile	tcc Ser	agt Ser 125	gct Ala	tca Ser	acg Thr	384
ggg Gly	gcc Ala 130	agc Ser	aac Asn	gac Asp	aac Asn	cac His 135	tac Tyr	ttc Phe	ggc Gly	tac Tyr	agc Ser 140	acc Thr	ccc Pro	tgg Trp	ggg Gly	432
tat Tyr 145	ttt Phe	gat Asp	ttc Phe	aac Asn	aga Arg 150	ttc Phe	cac His	tgc Cys	cac His	ttt Phe 155	tca Ser	cca Pro	cgt Arg	gac Asp	tgg Trp 160	480
cag Gln	cga Arg	ctc Leu	atc Ile	aac Asn 165	aac Asn	aat Asn	tgg Trp	gga Gly	ttc Phe 170	cgg Arg	ccc Pro	aag Lys	aga Arg	ctc Leu 175	aac Asn	528
ttc Phe	aaa Lys	ctc Leu	ttc Phe 180	aac Asn	atc Ile	caa Gln	gtc Val	aag Lys 185	gag Glu	gtc Val	acg Thr	acg Thr	aat Asn 190	gat Asp	ggc Gly	576
gtc Val	aca Thr 195	acc Thr	atc Ile	gct Ala	aat Asn	aac Asn	ctt Leu 200	acc Thr	agc Ser	acg Thr	gtt Val	caa Gln 205	gtc Val	ttc Phe	tcg Ser	624
gac Asp	tcg Ser 210	gag Glu	tac Tyr	cag Gln	ctt Leu 215	ccg Pro	tac Tyr	gtc Val	ctc Leu	ggc Gly	tct Ser 220	gcg Ala	cac His	cag Gln	ggc Gly	672
tgc Cys 225	ctc Leu	cct Pro	ccg Pro	ttc Phe 230	ccg Pro	gcg Ala	gac Asp	gtg Val	ttc Phe	atg Met 235	att Ile	ccg Pro	caa Gln	tac Tyr	ggc Gly 240	720
tac Tyr	ctg Leu	acg Thr	ctc Leu	aac Asn 245	aat Asn	ggc Gly	agc Ser	caa Gln	gcc Ala 250	gtg Val	gga Gly	cgt Arg	tca Ser	tcc Ser 255	ttt Phe	768
tac Tyr	tgc Cys	ctg Leu	gaa Glu 260	tat Tyr	ttc Phe	cct Pro	tct Ser	cag Gln 265	atg Met	ctg Leu	aga Arg	acg Thr	ggc Gly 270	aac Asn	aac Asn	816
ttt Phe	acc Thr	ttc Phe 275	agc Ser	tac Tyr	acc Thr	ttt Phe	gag Glu 280	gaa Glu	gtg Val	cct Pro	ttc Phe	cac His 285	agc Ser	agc Ser	tac Tyr	864

ES 2 644 880 T3

gcg cac agc cag agc ctg gac cgg ctg atg aat cct ctc atc gac caa 912
Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln
290 295 300

tac ctg tat tac ctg aac aga act caa aat cag tcc gga agt gcc caa 960
Tyr Leu Tyr Tyr Leu Asn Arg Thr Gln Asn Ser Gly Ser Ala Gln
305 310 315 320

aac aag gac ttg ctg ttt agc cgt ggg tct cca gct ggc atg tct gtt 1008
Asn Lys Asp Leu Leu Phe Ser Arg Gly Ser Pro Ala Gly Met Ser Val
325 330 335

cag ccc aaa aac tgg cta cct gga ccc tgt tat cgg cag cag cgc gtt 1056
Gln Pro Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val
340 345 350

tct aaa aca aaa aca gac aac aac aac agc aat ttt acc tgg act ggt 1104
Ser Lys Thr Lys Thr Asp Asn Asn Asn Ser Asn Phe Thr Trp Thr Gly
355 360 365

gct tca aaa tat aac ctc aat ggg cgt gaa tcc atc atc aac cct ggc 1152
Ala Ser Lys Tyr Asn Leu Asn Gly Arg Glu Ser Ile Ile Asn Pro Gly
370 375 380

act gct atg gcc tca cac aaa gac gac gaa gac aag ttc ttt ccc atg 1200
Thr Ala Met Ala Ser His Lys Asp Asp Glu Asp Lys Phe Phe Pro Met
385 390 395 400

agc ggt gtc atg att ttt gga aaa gag agc gcc gga gct tca aac act 1248
Ser Gly Val Met Ile Phe Gly Lys Glu Ser Ala Gly Ala Ser Asn Thr
405 410 415

gca ttg gac aat gtc atg att aca gac gaa gag gaa att aaa gcc act 1296
Ala Leu Asp Asn Val Met Ile Thr Asp Glu Glu Glu Ile Lys Ala Thr
420 425 430

aac cct gtg gcc acc gaa aga ttt ggg acc gtg gca gtc aat ttc cag 1344
Asn Pro Val Ala Thr Glu Arg Phe Gly Thr Val Ala Val Asn Phe Gln
435 440 445

agc agc agc aca gac cct gcg acc gga gat gtg cat gct atg gga gca 1392
Ser Ser Ser Thr Asp Pro Ala Thr Gly Asp Val His Ala Met Gly Ala
450 455 460

tta cct ggc atg gtg tgg caa gat aga gac gtg tac ctg cag ggt ccc 1440
Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro
465 470 475 480

att tgg gcc aaa att cct cac aca gat gga cac ttt cac ccg tct cct 1488
Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro Ser Pro
485 490 495

ctt atg ggc ggc ttt gga ctc aag aac ccg cct cct cag atc ctc atc 1536
Leu Met Gly Gly Phe Gly Leu Lys Asn Pro Pro Pro Gln Ile Leu Ile
500 505 510

aaa aac acg cct gtt cct gcg aat cct ccg gcg gag ttt tca gct aca 1584
Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Ala Glu Phe Ser Ala Thr
515 520 525

aag ttt gct tca ttc atc acc caa tac tcc aca gga caa gtg agt gtg 1632
Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val
530 535 540 545

gaa att gaa tgg gag ctg cag aaa gaa aac agc aag cgc tgg aat ccc 1680
Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro
545 550 555 560

gaa gtg cag tac aca tcc aat tat gca aaa tct gcc aac gtt gat ttt 1728
Glu Val Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Ala Asn Val Asp Phe
565 570 575

act gtg gac aac aat gga ctt tat act gag cct cgc ccc att ggc acc 1776
Thr Val Asp Asn Asn Gly Leu Tyr Thr Thr Pro Arg Pro Ile Gly Thr
580 585 590

cgt tac ctt acc cgt ccc ctg taa 1800
Arg Tyr Leu Thr Arg Pro Leu
595

<210> 23

<211> 599

ES 2 644 880 T3

<212> PRT

<213> virus 1 adeno-asociado; VP2, VP3

<400> 23

Thr Ala Pro Gly Lys Lys Arg Pro Val Glu Gln Ser Pro Gln Glu Pro
 1 5 10 15

Asp Ser Ser Ser Gly Ile Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys
 20 25 30

Arg Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu Ser Val Pro Asp Pro
 35 40 45

Gln Pro Leu Gly Glu Pro Pro Ala Thr Pro Ala Ala Val Gly Pro Thr
 50 55 60

Thr Met Ala Ser Gly Gly Gly Ala Pro Met Ala Asp Asn Asn Glu Gly
 65 70 75 80

Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp His Cys Asp Ser Thr
 85 90 95

Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu
 100 105 110

Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile Ser Ser Ala Ser Thr
 115 120 125

Gly Ala Ser Asn Asp Asn His Tyr Phe Gly Tyr Ser Thr Pro Trp Gly
 130 135 140

Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp
 145 150 155 160

Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn
 165 170 175

Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr Thr Asn Asp Gly
 180 185 190

Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Val Gln Val Phe Ser

ES 2 644 880 T3

195	200	205
Asp Ser Glu Tyr Gln Leu 210	Pro Tyr Val Leu Gly 215	Ser Ala His Gln Gly 220
Cys Leu Pro Pro Phe 225	Pro Ala Asp Val Phe 230	Met Ile Pro Gln Tyr Gly 235
Tyr Leu Thr Leu 245	Asn Asn Gly Ser Gln 245	Ala Val Gly Arg Ser Ser Phe 250
Tyr Cys Leu Glu Tyr Phe 260	Pro Ser Gln Met Leu 265	Arg Thr Gly Asn Asn 270
Phe Thr Phe Ser Tyr Thr 275	Phe Glu Glu Val Pro 280	Phe His Ser Ser Tyr 285
Ala His Ser Gln Ser Leu 290	Asp Arg Leu Met Asn 295	Pro Leu Ile Asp Gln 300
Tyr Leu Tyr Tyr Leu 305	Asn Arg Thr Gln Asn 310	Gln Ser Gly Ser Ala Gln 315
Asn Lys Asp Leu Leu Phe 325	Ser Arg Gly Ser Pro 330	Ala Gly Met Ser Val 335
Gln Pro Lys Asn Trp Leu 340	Pro Gly Pro Cys Tyr 345	Arg Gln Gln Arg Val 350
Ser Lys Thr Lys Thr Asp 355	Asn Asn Asn Ser Asn 360	Phe Thr Trp Thr Gly 365
Ala Ser Lys Tyr Asn Leu 370	Asn Gly Arg Glu Ser 375	Ile Ile Asn Pro Gly 380
Thr Ala Met Ala Ser His 385	Lys Asp Asp Glu Asp 390	Lys Phe Phe Pro Met 395
Ser Gly Val Met Ile Phe 405	Gly Lys Glu Ser Ala 410	Gly Ala Ser Asn Thr 415
Ala Leu Asp Asn Val Met 420	Ile Thr Asp Glu Glu 425	Glu Ile Lys Ala Thr 430
Asn Pro Val Ala Thr Glu 435	Arg Phe Gly Thr Val 440	Ala Val Asn Phe Gln 445
Ser Ser Ser Thr Asp Pro 450	Ala Thr Gly Asp Val 455	His Ala Met Gly Ala 460
Leu Pro Gly Met Val Trp 465	Gln Asp Arg Asp Val 470	Tyr Leu Gln Gly Pro 475
		480

ES 2 644 880 T3

Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro Ser Pro
485 490 495

Leu Met Gly Gly Phe Gly Leu Lys Asn Pro Pro Pro Gln Ile Leu Ile
500 505 510 515

Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Ala Glu Phe Ser Ala Thr
515 520 525

Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val
530 535 540

Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro
545 550 555 560

Glu Val Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Ala Asn Val Asp Phe
565 570 575

Thr Val Asp Asn Asn Gly Leu Tyr Thr Glu Pro Arg Pro Ile Gly Thr
580 585 590

Arg Tyr Leu Thr Arg Pro Leu
595

<210> 24

<211> 1605

<212> ADN

5 <213> virus 1 adeno-asociado; VP3

<220>

<221> CDS

<222> (1)..(1602)

<223> VP3

10 <400> 24

atg gct tca ggc ggt ggc gca cca atg gca gac aat aac gaa ggc gcc 48
Met Ala Ser Gly Gly Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala
1 5 10 15

gac gga gtg ggt aat gcc tca gga aat tgg cat tgc gat tcc aca tgg 96
Asp Gly Val Gly Asn Ala Ser Gly Asn Trp His Cys Asp Ser Thr Trp
20 25 30

ctg ggc gac aga gtc atc acc acc agc acc cgc acc tgg gcc ttg ccc 144
Leu Gly Asp Arg Val Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro
35 40 45

acc tac aat aac cac ctc tac aag caa atc tcc agt gct tca acg ggg 192
Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly
50 55 60

gcc agc aac gac aac cac tac ttc ggc tac agc acc ccc tgg ggg tat 240
Ala Ser Asn Asp Asn His Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr
65 70 75 80

ttt gat ttc aac aga ttc cac tgc cac ttt tca cca cgt gac tgg cag 288
Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln
85 90 95

ES 2 644 880 T3

cga	ctc	atc	aac	aac	aat	tgg	gga	ttc	cgg	ccc	aag	aga	ctc	aac	ttc	336
Arg	Leu	Ile	Asn	Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Asn	Phe	
			100					105					110			
aaa	ctc	ttc	aac	atc	caa	gtc	aag	gag	gtc	acg	acg	aat	gat	ggc	gtc	384
Lys	Leu	Phe	Asn	Ile	Gln	Val	Lys	Glu	Val	Thr	Thr	Asn	Asp	Gly	Val	
		115					120					125				
aca	acc	atc	gct	aat	aac	ctt	acc	agc	acg	ggt	caa	gtc	ttc	tcg	gac	432
Thr	Thr	Ile	Ala	Asn	Asn	Leu	Thr	Ser	Thr	Val	Gln	Val	Phe	Ser	Asp	
		130				135					140					
tcg	gag	tac	cag	ctt	ccg	tac	gtc	ctc	ggc	tct	gcg	cac	cag	ggc	tgc	480
Ser	Glu	Tyr	Gln	Leu	Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Gln	Gly	Cys	
		145			150					155					160	
ctc	cct	ccg	ttc	ccg	gcg	gac	gtg	ttc	atg	att	ccg	caa	tac	ggc	tac	528
Leu	Pro	Pro	Phe	Pro	Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	
				165					170					175		
ctg	acg	ctc	aac	aat	ggc	agc	caa	gcc	gtg	gga	cgt	tca	tcc	ttt	tac	576
Leu	Thr	Leu	Asn	Asn	Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	
			180					185					190			
tgc	ctg	gaa	tat	ttc	cct	tct	cag	atg	ctg	aga	acg	ggc	aac	aac	ttt	624
Cys	Leu	Glu	Tyr	Phe	Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	
		195					200					205				
acc	ttc	agc	tac	acc	ttt	gag	gaa	gtg	cct	ttc	cac	agc	agc	tac	gcg	672
Thr	Phe	Ser	Tyr	Thr	Phe	Glu	Glu	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	
		210				215					220					
cac	agc	cag	agc	ctg	gac	cgg	ctg	atg	aat	cct	ctc	atc	gac	caa	tac	720
His	Ser	Gln	Ser	Leu	Asp	Arg	Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	
		225			230					235					240	
ctg	tat	tac	ctg	aac	aga	act	caa	aat	cag	tcc	gga	agt	gcc	caa	aac	768
Leu	Tyr	Tyr	Leu	Asn	Arg	Thr	Gln	Asn	Gln	Ser	Gly	Ser	Ala	Gln	Asn	
				245					250					255		
aag	gac	ttg	ctg	ttt	agc	cgt	ggg	tct	cca	gct	ggc	atg	tct	ggt	cag	816
Lys	Asp	Leu	Leu	Phe	Ser	Arg	Gly	Ser	Pro	Ala	Gly	Met	Ser	Val	Gln	
			260				265						270			
ccc	aaa	aac	tgg	cta	cct	gga	ccc	tgt	tat	cgg	cag	cag	cgc	ggt	tct	864
Pro	Lys	Asn	Trp	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	
		275				280						285				
aaa	aca	aaa	aca	gac	aac	aac	aac	agc	aat	ttt	acc	tgg	act	ggt	gct	912
Lys	Thr	Lys	Thr	Asp	Asn	Asn	Asn	Ser	Asn	Phe	Thr	Trp	Thr	Gly	Ala	
		290			295						300					
tca	aaa	tat	aac	ctc	aat	ggg	cgt	gaa	tcc	atc	atc	aac	cct	ggc	act	960
Ser	Lys	Tyr	Asn	Leu	Asn	Gly	Arg	Glu	Ser	Ile	Ile	Asn	Pro	Gly	Thr	
					310					315					320	
gct	atg	gcc	tca	cac	aaa	gac	gac	gaa	gac	aag	ttc	ttt	ccc	atg	agc	1008
Ala	Met	Ala	Ser	His	Lys	Asp	Asp	Glu	Asp	Lys	Phe	Phe	Pro	Met	Ser	
				325					330					335		
ggt	gtc	atg	att	ttt	gga	aaa	gag	agc	gcc	gga	gct	tca	aac	act	gca	1056
Gly	Val	Met	Ile	Phe	Gly	Lys	Glu	Ser	Ala	Gly	Ala	Ser	Asn	Thr	Ala	
			340					345					350			
ttg	gac	aat	gtc	atg	att	aca	gac	gaa	gag	gaa	att	aaa	gcc	act	aac	1104
Leu	Asp	Asn	Val	Met	Ile	Thr	Asp	Glu	Glu	Glu	Ile	Lys	Ala	Thr	Asn	
		355					360					365				

ES 2 644 880 T3

cct gtg gcc acc gaa aga ttt ggg acc gtg gca gtc aat ttc cag agc 1152
 Pro Val Ala Thr Glu Arg Phe Gly Thr Val Ala Val Asn Phe Gln Ser
 370 375 380

agc agc aca gac cct gcg acc gga gat gtg cat gct atg gga gca tta 1200
 Ser Ser Thr Asp Pro Ala Thr Gly Asp Val His Ala Met Gly Ala Leu
 385 390 395 400

cct ggc atg gtg tgg caa gat aga gac gtg tac ctg cag ggt ccc att 1248
 Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile
 405 410 415

tgg gcc aaa att cct cac aca gat gga cac ttt cac ccg tct cct ctt 1296
 Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro Ser Pro Leu
 420 425 430

atg ggc ggc ttt gga ctc aag aac ccg cct cct cag atc ctc atc aaa 1344
 Met Gly Gly Phe Gly Leu Lys Asn Pro Pro Pro Gln Ile Leu Ile Lys
 435 440 445

aac acg cct gtt cct gcg aat cct ccg gcg gag ttt tca gct aca aag 1392
 Asn Thr Pro Val Pro Ala Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys
 450 455 460

ttt gct tca ttc atc acc caa tac tcc aca gga caa gtg agt gtg gaa 1440
 Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu
 465 470 475 480

att gaa tgg gag ctg cag aaa gaa aac agc aag cgc tgg aat ccc gaa 1488
 Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu
 485 490 495

gtg cag tac aca tcc aat tat gca aaa tct gcc aac gtt gat ttt act 1536
 Val Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr
 500 505 510

gtg gac aac aat gga ctt tat act gag cct cgc ccc att ggc acc cgt 1584
 Val Asp Asn Asn Gly Leu Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg
 515 520 525

tac ctt acc cgt ccc ctg taa 1605
 Tyr Leu Thr Arg Pro Leu
 530

<210> 25

<211> 534

<212> PRT

5 <213> virus 1 adeno-asociado; VP3

<400> 25

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

Asp Gly Val Gly Asn Ala Ser Gly Asn Trp His Cys Asp Ser Thr Trp
 20 25 30

Leu Gly Asp Arg Val Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro
 35 40 45

Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly
 50 55 60

Ala Ser Asn Asp Asn His Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr

ES 2 644 880 T3

Leu Asp Asn Val Met Ile Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn
 355 360 365

Pro Val Ala Thr Glu Arg Phe Gly Thr Val Ala Val Asn Phe Gln Ser
 370 375 380

Ser Ser Thr Asp Pro Ala Thr Gly Asp Val His Ala Met Gly Ala Leu
 385 390 395 400

Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile
 405 410 415

Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro Ser Pro Leu
 420 425 430

Met Gly Gly Phe Gly Leu Lys Asn Pro Pro Pro Gln Ile Leu Ile Lys
 435 440 445

Asn Thr Pro Val Pro Ala Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys
 450 455 460

Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu
 465 470 475 480

Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu
 485 490 495

Val Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr
 500 505 510

Val Asp Asn Asn Gly Leu Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg
 515 520 525

Tyr Leu Thr Arg Pro Leu
 530

<210> 26

<211> 20

<212> ADN

5 <213> Artificial

<220>

<223> primera secuencia

<220>

<221> misc_feature

10 <223> CMV gen

<400> 26

aatgggcggt aggcgtgta 19

<210> 27

<211> 21

15 <212> ADN

<213> Artificial
 <220>
 <223> primera secuencia
 <220>
 5 <221> misc_feature
 <223> CMV gen
 <400> 27
 agcgatctg acggtcact aa 22
 <210> 28
 10 <211> 11954
 <212> ADN
 <213> Artificial
 <220>
 <223> pVD84 vector
 15 <400> 28
 gggatgatcaa gtcttcgctg agtgattgta aataaaatgt aatttacagt atagtatattt 60
 aattaatata caaatgattt gataataatt cttatttaac tataatata tgtggtgggt 120
 tgaattaaag gtccgtatac tccggaatat taatagatca tggagataat taaaatgata 180
 accatctcgc aaataaataa gtattttact gttttcgtaa cagttttgta ataaaaaac 240
 ctataaatat tccggattat tcataccgct ccaccatcgg gcgcggatcc tgtaaggtg 300
 gctgccgacg gttatctacc cgattggctc gaggacaacc tctctgaggg cattcgcgag 360
 tgggtgggact tgaaacctgg agccccgaag cccaaagcca accagcaaaa gcaggacgac 420
 ggccggggctc tgggtgcttc tggctacaag tacctcggac cttcaacgg actcgacaag 480
 ggggagcccg tcaacgcggc ggacgcagcg gccctcgagc acgacaaggc ctacgaccag 540
 cagctcaaag cgggtgacaa tccgtacctg cggtataacc acgccgacgc cgagtctcag 600
 gagcgtctgc aagaagatac gtcttttggg ggcaacctcg ggcgagcagt cttccaggcc 660
 aagaagcggg ttctcgaacc tctcggctcg gttgaggaag gcgctaagac ggctcctgga 720
 aagaacgctc cggtagagca gtcgccacaa gagccagact cctcctcggg catcggcaag 780
 acaggccagc agcccgctaa aaagagactc aattttggtc agactggcga ctcagagtca 840
 gtccccgatc cacaacctct cgggagaacct ccagcaacct ccgctgctgt gggacctact 900
 acaatggctt caggcgggtg cgcaccaatg gcagacaata acgaaggcgc cgacggagtg 960
 ggtaatgcct caggaaattg gcattgcgat tccacatggc tgggcgacag agtcatcacc 1020
 accagcacc gcacctgggc cttgcccacc tacaataacc acctctaca gcaaatctcc 1080
 agtgcttcaa cgggggccag caacgacaac cactacttcg gctacagcac cccctggggg 1140
 tattttgatt tcaacagatt cactgccac ttttcaccac gtgactggca gcgactcatc 1200
 aacaacaatt ggggattccg gcccaagaga ctcaacttca aactcttcaa catccaagtc 1260
 aaggaggctc cgacgaatga tggcgtcaca accatcgcta ataaccttac cagcacgggt 1320
 caagtcttct cggactcggg gtaccagctt ccgtacgtcc tcggctctgc gcaccagggc 1380

ES 2 644 880 T3

tgctccctc	cgttcccggc	ggacgtgttc	atgattccgc	aatacggcta	cctgacgctc	1440
aacaatggca	gccaaagccgt	gggacgttca	tccttttact	gcctggaata	tttcccttct	1500
cagatgctga	gaacgggcaa	caactttacc	ttcagctaca	cctttgagga	agtgcctttc	1560
cacagcagct	acgcgcacag	ccagagcctg	gaccggctga	tgaatcctct	catcgaccaa	1620
tacctgtatt	acctgaacag	aactcaaaat	cagtccggaa	gtgcccaaaa	caaggacttg	1680
ctgtttagcc	gtgggtctcc	agctggcatg	tctgttcagc	ccaaaaactg	gctacctgga	1740
ccctgttatc	ggcagcagcg	cgtttctaaa	acaaaaacag	acaacaacaa	cagcaatttt	1800
acctggactg	gtgcttcaaa	atataacctc	aatgggctgt	aatccatcat	caaccctggc	1860
actgctatgg	cctcacacaa	agacgacgaa	gacaagttct	ttcccatgag	cggtgtcatg	1920
atTTTTgaa	aagagagcgc	cggagcttca	aacactgcat	tggacaatgt	catgattaca	1980
gacgaagagg	aaattaaagc	cactaacctt	gtggccaccg	aaagatttgg	gaccgtggca	2040
gtcaatttcc	agagcagcag	cacagaccct	gcgaccggag	atgtgcatgc	tatgggagca	2100
ttacctggca	tgggtgtggca	agatagagac	gtgtacctgc	agggtcccat	ttgggcaaaa	2160
attcctcaca	cagatggaca	ctttcacccg	tctcctctta	tgggcggtct	tggactcaag	2220
aaccgcctc	ctcagatcct	catcaaaaac	acgcctgttc	ctgcgaatcc	tccggcggag	2280
ttttcagcta	caaagtttgc	ttcattcatc	acccaatact	ccacaggaca	agtgagtgtg	2340
gaaattgaat	gggagctgca	gaaagaaaac	agcaagcgc	ggaatcccga	agtgcagtac	2400
acatccaatt	atgcaaaatc	tgccaacggt	gattttactg	tggacaacaa	tggactttat	2460
actgagcctc	gccccattgg	cacccttac	cttaccctgc	ccctgtaagc	aagggcgaat	2520
tctgcagata	tccatcacac	tggggccgc	tttcaaatct	agagcctgca	gtctcgacaa	2580
gcttgtcgag	aagtactaga	ggatcataat	cagccatacc	acattttag	aggttttact	2640
tgctttaaaa	aacctcccac	acctccccct	gaacctgaaa	cataaaatga	atgcaattgt	2700
tgttgttaac	ttgtttattg	cagcttataa	tggttacaaa	taaagcaata	gcatcacaaa	2760
tttcacaaat	aaagcatttt	tttactgca	ttctagttgt	ggtttgtcca	aactcatcaa	2820
tgtatcttat	catgtctgga	tctgatcact	gcttgagcct	agaggcctcg	cgagatctta	2880
attaattaag	taccgactct	gctgaagagg	aggaaattct	ccttgaagtt	tccctggtgt	2940
tcaaagtaaa	ggagtttgca	ccagacgcac	ctctgttcac	tgggtccggcg	tattaaacaa	3000
cgatacattg	ttattagtac	atTTattaag	cgtagattc	tgtgcgttgt	tgatttacag	3060
acaattgttg	tacgtatttt	aataattcat	taaatttata	atctttaggg	tggtatgtta	3120
gagcgaanaa	caaatgattt	tcagcgtctt	tatatctgaa	tttaaatatt	aaatcctcaa	3180
tagatttgta	aaataggttt	cgattagttt	caaacaaggg	ttgtttttcc	gaaccgatgg	3240
ctggactatc	taatggattt	tcgctcaacg	ccacaaaact	tgccaaatct	tgtagcagca	3300
atctagcttt	gtcgaatttc	gtttgtgttt	tgtttttaa	taaagggttcg	acgtcgttca	3360
aaatattatg	cgcttttgta	tttctttcat	cactgtcgtt	agtgtaaat	tgactcgacg	3420

ES 2 644 880 T3

taaacacggt aaataaagct tggacatatt taacatcggg cgtgttagct ttattagcc 3480
 gattatcgtc gtcgtcccaa ccctcgtcgt tagaagttgc ttccgaagac gattttgcca 3540
 tagccacacg acgcctatta attgtgtcgg ctaacacgtc cgcgatcaaa tttgtagtgtg 3600
 agcttttttg aattatttct gattgcgggc gtttttgggc gggtttcaat ctaactgtgc 3660
 ccgattttta ttcagacaac acgttagaaa gcgatggtgc aggcggtggt aacatttcag 3720
 acggcaaate tactaatggc ggcgggtggt gagctgatga taaatctacc atcggtgagg 3780
 gcgcaggcgg ggctggcggc ggaggcggag gcggaggtgg tggcgggtgat gcagacggcg 3840
 gtttaggctc aatgtctct ttaggcaaca cagtcggcac ctcaactatt gtactggttt 3900
 cgggcgcgct ttttggtttg accggtctga gacgagtgcg atttttttcg tttctaatag 3960
 cttcaacaa ttgttgtctg tcgtctaaag gtgcagcggg ttgaggttcc gtcggcattg 4020
 gtggagcggg cggcaattca gacatcgatg gtggtggtgg tggaggaggc gctggaatgt 4080
 taggcacggg agaaggtggt ggcggcgggt ccgccggtat aatttgttct ggttttagttt 4140
 gttcgcgcac gattgtgggc accggcgcag gcgccgctgg ctgcacaacg gaaggtcgtc 4200
 tgcttcgagg cagcgttgg ggtggtggca attcaatatt ataattggaa tacaatcgt 4260
 aaaaatctgc tataagcatt gtaatttcgc tatcgtttac cgtgccgata tttacaacc 4320
 gctcaatgta agcaattgta ttgtaaagag attgtctcaa gctcggatcc cgcacgccga 4380
 taacaagcct tttcattttt actacagcat tgtagtggcg agacacttcg ctgtcgtcga 4440
 cgtacatgta tgctttggtg tcaaaaacgt cgttggcaag ctttaaata tttaaaagaa 4500
 catctctgtt cagcaccact gtgttctcgt aaatgttgtt tttgataatt tgcgcttccg 4560
 cagtatcgac acgttcaaaa aattgatgcg catcaatttt gttgttccta ttattgaata 4620
 aataagattg tacagattca tatctacgat tcgtcatggc caccacaaat gctacgctgc 4680
 aaacgctggt acaattttac gaaaactgca aaaacgtcaa aactcgggtat aaaataatca 4740
 acgggcgctt tggcaaaata tctattttat cgcacaagcc cactagcaaa ttgtatttgc 4800
 agaaaacaat ttcggcgcac aattttaacg ctgacgaaat aaaagttcac cagttaatga 4860
 gcgaccaccc aaattttata aaaatctatt ttaatcacgg ttccatcaac aaccaagtga 4920
 tcgtgatgga ctacattgac tgtcccgatt tatttgaaac actacaaatt aaaggcgagc 4980
 tttcgtacca acttgttagc aatattatta gacagctgtg tgaagcgtc aacgatttgc 5040
 acaagcacia tttcatacac aacgacataa aactcgaaaa tgtcttata ttcgaagcac 5100
 ttgatcgcgt gtatgtttgc gattacggat tgtgcaaca cgaaaactca cttagcgtgc 5160
 acgacggcac gttggagtat tttagtccgg aaaaaattcg acacacaact atgcacgttt 5220
 cgtttgactg gtacgcggcg tgtaacata caagttgcta accggcggcc gacaccatt 5280
 tgaaaaaagc gaagacgaaa tgttggactt gaatagcatg aagcgtcgtc agcaatacaa 5340
 tgacattggc gttttaaaac acgttcgtaa cgtaaacgct cgtgactttg tgtactgcct 5400
 aacaagatac aacatagatt gtagactcac aaattacaaa caaattataa aacatgagtt 5460
 tttgtcgtaa aaatgccact tgttttacga gtagaattcg taatcatggt catagctgtt 5520

ES 2 644 880 T3

tcctgtgtga aattgttata cgctcacaat tccacacaac atacgagccg gaagcataaa 5580
 gtgtaaagcc tggggtgccct aatgagtgag ctaactcaca ttaattgctg tgcgctcact 5640
 gcccgctttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc 5700
 ggggagagcc ggtttgcgta ttgggcgctc ttcgcttcc tcgctcactg actcgtgcy 5760
 ctcggtcgtt cggctgcggc gagcggatc agctcactca aaggcggtaa tacggttatc 5820
 cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 5880
 gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg ctccgcccc ctgacgagca 5940
 tcacaaaaat cgacgctcaa gtcagaggtg gcgaaaccg acaggactat aaagatacca 6000
 ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg 6060
 atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag 6120
 gtatctcagt tcgggttagg tcggtcgcct caagctgggc tgtgtgcacg aacccccgt 6180
 tcagcccgcg cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaaagaca 6240
 cgacttatcg ccaactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 6300
 cgggtctaca gagttcttga agtggtggcc taactacggc taaactagaa ggacagtatt 6360
 tggatctcgc gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc 6420
 cggcaaaaa accaccgctg gtagcgggtg ttttttgtt tgcaagcagc agattacgcg 6480
 cagaaaaaaa ggatctcaag aagatcctt gatcttttct acggggctc acgctcagtg 6540
 gaacgaaaac tcacgttaag ggattttgtt catgagatta tcaaaaagga tcttcaccta 6600
 gatcctttta aattaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg 6660
 gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg 6720
 ttcattcata gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc 6780
 atctggcccc agtgctgcaa tgataccgcg agaccacgc tcaccggctc cagatttatc 6840
 agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc 6900
 ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag 6960
 tttgcgcaac gttggtgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat 7020
 ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgty 7080
 caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt 7140
 gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag 7200
 atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg 7260
 accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt 7320
 aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct 7380
 gttgagatcc agttcagatg aaccactcg tgcaccaac tgatcttcag catcttttac 7440
 tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaa 7500
 aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat 7560

ES 2 644 880 T3

ttatcagggg tattgtctca tgagcggata catatgtgaa tgtattaga aaaataaaca 7620
 aatagggggt ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat 7680
 tatcatgaca ttaacctata aaaataggcg taccacgagg ccctttcgtc tcgcgcggtt 7740
 cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca cagcttgtct 7800
 gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg 7860
 tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc accatattgcg 7920
 gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc attcgcatt 7980
 caggctgcgc aactgttggg aagggcgatc ggtgccccgc tcttcgctat tacgccagct 8040
 ggcgaaaggg ggatgtgctg caaggcgatt aagtgggta acgccagggt tttcccagtc 8100
 acgacgttgt aaaacgaccg agttgtttgc gtacgtgact agcgaagaag atgtgtggac 8160
 cgcagaacag atagtaaac aaaaccctag tattggagca ataatcgatt taaccaacac 8220
 gtctaaatat tatgatggtg tgcatttttt gcggcggggc ctgttataca aaaaaattca 8280
 agtacctggc cagactttgc cgctgaaag catagttcaa gaatttattg acacggtaaa 8340
 agaatttaca gaaaagtgtc ccggcatggt ggtgggcgtg cactgcacac acggatttaa 8400
 tcgcaccggt tacatggtgt gcagatattt aatgcacacc ctgggtattg cgccgcagga 8460
 agccatagat agattcgaaa aagccagagg tcacaaaatt gaaagacaaa attacgttca 8520
 agatttatta atttaattaa tattatttgc attctttaa aaatacttta tcctattttc 8580
 aaattgttgc gcttcttcca gcgaaccaa actatgcttc gcttctccg tttagcttgt 8640
 agccgatcag tggcgttgtt ccaatcgacg gtaggattag gccgatatt ctccaccaca 8700
 atgttgcaa cgttgatggt acgtttatgc ttttggtttt ccacgtacgt cttttggccg 8760
 gtaatagccg taaacgtagt gccgtcgcgc gtcacgaca acaccggatg tttgcgcttg 8820
 tccgcggggg attgaaccgc gcgatccgac aaatccacca ctttggcaac taaatcgggtg 8880
 acctgcgctt ctttttctg cattatttct tcttctttt gcattggtttc ctggaagccg 8940
 gtgtacatgc ggttttagatc agtcatgacg cgcgtgacct gcaaactttt ggctcgcgac 9000
 tgcttgcctt tgatggcaac gatgcgttca ataaactctt gttttttaa aagttcctcg 9060
 gttttttgcg ccaccaccgc ttgcagcgcg tttgtgtgct cggatgaatgt cgcaatcagc 9120
 ttagtcacca actgtttgct ctctctctc cgttgtttga tcgcgggatc gtacttgccg 9180
 gtgcagagca cttgaggaat tacttcttct aaaagccatt cttgtaattc tatggcgtaa 9240
 ggcaatttgg acttcataat cagctgaatc acgccggatt tagtaatgag cactgtatgc 9300
 ggctgcaaat acagcgggtc gcccttttc acgacgctgt tagaggtagg gcccccattt 9360
 tggatggtct gctcaaaaa cgatttgtat ttattgtcta catgaacacg tatagcttta 9420
 tcacaaactg tatattttaa actgttagcg acgtccttgg ccacgaaccg gacctgttg 9480
 tcgcgctcta gcacgtaccg caggttgaac gtatcttctc caaattttaa ttctcaatt 9540
 ttaacgcgag ccattttgat acacgtgtgt cgattttgca acaactattg ttttttaacg 9600
 caaactaac ttattgtggt aagcaataat taaatatggg ggaacatgcg ccgctacaac 9660

ES 2 644 880 T3

actcgtcgtt atgaacgcag acggcgccgg tctcggcgca agcggctaaa acgtgttgcg 9720
cgttcaacgc ggcaaacatc gcaaaagcca atagtacagt tttgatttgc atattaacgg 9780
cgatttttta aattatctta ttaataaat agttatgacg cctacaactc cccgcccgcg 9840
ttgactcgtc gcacctcgag cagttcgttg acgccttcct ccgtgtggcc gaacacgtcg 9900
agcgggtggg cgatgaccag cggcgtgccc cacgcgacgc acaagtatct gtacaccgaa 9960
tgatcgtcgg gcgaaggcac gtcggcctcc aagtggcaat attggcaaat tcgaaaatat 10020
atacagttgg gttgtttgcg catatctatc gtggcgttgg gcatgtacgt ccgaacgttg 10080
atgtgcatgc aagccgaaat taaatcattg cgattagtgc gattaaaacg ttgtacatcc 10140
tcgcttttaa tcatgccgtc gattaaatcg cgcaatcgag tcaagtgatc aaagtgtgga 10200
ataatgtttt ctttgtattc ccgagtcaag cgacgcgctt attttaaca actagccatc 10260
ttgtaagtta gtttcattta atgcaacttt atccaataat atattatgta tcgcacgtca 10320
agaattaaca atgccccgtt tgtcgcattc caacacgact atgatagaga tcaataaag 10380
cgcaattaa atagcttgcg acgcaacgtg cacgatctgt gcacgcggtc cggcacgagc 10440
tttgattgta ataagttttt acgaagcgat gacatgacc ccgtagtgac aacgatcacg 10500
ccaaaagaa ctgccgacta caaaattacc gagtatgtcg gtgacgttaa aactattaag 10560
ccatccaatc gaccgttagt cgaatcagga ccgctgggtc gagaagccgc gaagtatggc 10620
gaatgcatcg tataacgtgt ggagtccgct cattagagcg tcatgttttag acaagaaagc 10680
tacatattta attgatcccg atgattttat tgataaattg accctaactc catacacggt 10740
attctacaat ggccgggttt tgggtcaaat ttccggactg cgattgtaca tgctgttaac 10800
ggctccgccc actattaatg aaattaataa ttccaatttt aaaaaacgca gcaagagaaa 10860
catttgatg aaagaatgcg tagaaggaaa gaaaatgtc gtcgacatgc tgaacaaca 10920
gattaatatg cctccgtgta taaaaaaaaat attgaacgat ttgaaagaaa acaatgtacc 10980
gcgcgccggt atgtacagga agaggtttat actaaactgt tacattgcaa acgtggtttc 11040
gtgtgccaa gttgaaaacc gatgtttaat caaggctctg acgcatttct acaaccacga 11100
ctccaagtgt gtgggtgaag tcatgcatct ttaatacaa tccaagatg tgtataaacc 11160
accaaactgc caaaaaatga aaactgtcga caagctctgt ccgtttgctg gcaactgcaa 11220
gggtctcaat cctatttga attattgaat aataaaaca ttataaatgc taaatttgtt 11280
ttttattaac gatacaaac aaacgcaaca agaacatttg tagtattatc tataattgaa 11340
aacgcgtagt tataatcgct gaggtaatat ttaaatcat tttcaaatga ttcacagtta 11400
atgtgcgaca atataatttt attttcatat aaactagacg cttgtcgtc ttcttcttcg 11460
tattccttct ctttttcatt tttctcctca taaaaattaa catagttatt atcgtatcca 11520
tatatgtatc tatcgtatag agtaaatfff ttgttgcatt aaatataat gtctttttta 11580
atggggtgta tagtaccgct gcgcatagtt tttctgtaat ttacaacagt gctattttct 11640
ggtagttctt cggagtgtgt tgctttaatt attaaattta tataatcaat gaatttggga 11700
tcgtcggttt tgtacaatat gttgccggca tagtacgacg cttcttctag ttcaattaca 11760
ccatttttta gcagcaccgg attaacataa ctttcaaaa tgttgtagca accgttaaac 11820
aaaaacagtt cacctccctt ttctatacta ttgtctgcga gcagttgttt gttgttaaaa 11880
ataacagcca ttgtaatgag acgcacaac taatatcaca aactggaaat gtctatcaat 11940
atatagttgc tgat 11954

<210> 29

<211> 30

5 <212> ADN

ES 2 644 880 T3

<213> Artificial
 <220>
 <223> primera secuencia (directa, comprende sitio BamHI)
 <400> 29
 5 ttaggatcct gtaaggtgg ctgccgacgg 30
 <210> 30
 <211> 30
 <212> ADN
 <213> Artificial
 10 <220>
 <223> primera secuencia (inversa, que comprende sitio Stul)
 <400> 30
 gtcgtaggcc ttgtcgtgct cgagggccgc 30
 <210> 31
 15 <211> 1966
 <212> ADN
 <213> virus 2 adeno-asociado
 <220>
 <221> misc_feature
 20 <223> rep con CTG codón de inicio ATG no interno
 <400> 31
 cgggccgaag cgcgcggaat tcaaaggcct acgtcgacga ggggagatct gccgccctgg 60
 cgggggtttta cgagattgtg attaaggtcc ccagcgacct tgacgagcat ctgcccggca 120
 tttctgacag ctttgtgaac tgggtggccg agaaggagtg ggagttgccg ccagattctg 180
 acttgatct gaactctgatt gagcaggcac ccctgaccgt ggccgagaag ctgcagcgcg 240
 actttctgac ggagtggcgc cgtgtgagta aggccccgga ggcccttttc tttgtgcaat 300
 ttgagaaggg agagagctac ttccacttac acgtgctcgt ggaaccacc ggggtgaaat 360
 ccttagtttt gggacgtttc ctgagtcaga ttcgcgaaaa actgattcag agaatttacc 420
 gcgggatcga gccgactttg ccaaactggt tcgcggtcac aaagaccaga aacggcgccg 480
 gaggcgggaa caaggtggtg gacgagtgt acatcccaa ttacttgctc cccaaaacct 540
 agcctgagct ccagtggcg tggactaatt tagaacagta ttaagcgcc tgtttgaatc 600
 tcacggagcg taaacggtt gtggcgcagc atctgacgca cgtgtcgag acgcaggagc 660
 agaacaaaga gaatcagaat cccaattctg acgcgccggt gatcagatca aaaacttcag 720

ES 2 644 880 T3

ccaggtagat ggagctggc gggaggctcg tggacaaggg gattacctcg gagaagcagt	780
ggatccagga ggaccaggcc tcatacatct cttcaatgc ggcctccaac tcgaggctcc	840
aaatcaaggc tgccttgac aatgaggaa agattatgag cctgactaaa accgccccg	900
actacctgt gggcagcag cccgtggagg acatttccag caatcggatt tataaaattt	960
tggaactaaa cgggtacgat cccaatatg cggcttccgt ctttctgga tggccacga	1020
aaaagtccg caagaggaac accatctggc tgttgggccc tgcaactacc ggaagacca	1080
acatcgagg ggcacatagc cacactgtgc cttctacgg gtgcgtaaac tggaccaatg	1140
agaactttc cttcaacgac tgtgtcgaca agatggtgat ctggtgggag gagggaaga	1200
tgaccgcaa ggtcgtggag tcgcccagg ccattctcgg aggaagcaag gtgcgctgg	1260
accagaaatg caagtcctcg gccagatag acccgactcc cgtgatcgtc acctcaaca	1320
ccaacatgtg cggcgtgatt gacgggaact caacgacctt cgaacaccag cagccgttg	1380
aagaccggat gttcaaattt gaactcacc gccgtctgga tcatgacttt ggaagggtca	1440
ccaagcagga agtcaaagac ttttccggg gggcaaagga tcacgtggtt gaggtggagc	1500
atgaattcta cgtcaaaaag ggtggagcca agaaaagacc cggccccagt gacgagata	1560
taagtgagcc caaacgggtg cgcgagtcag ttgcgagcc atcgacgtca gacgaggaag	1620
cttcgatcaa ctacgagac aggtaccaa acaaatgttc tcgtcacgtg ggcagatc	1680
tgatgctgt tccctgcaga caatgcgaga gaatgaatca gaattcaat atctgcttca	1740
ctcacggaca gaaagactgt ttagagtgtc ttcccgtgtc agaattctca cccgtttctg	1800
tcgtcaaaa ggcgtatcag aaactgtgct acattcatca tatcatggga aagggtccag	1860
acgcttgac tgcctcgat ctggtcaatg tggatttga tgactgcatc tttgaacaat	1920
aaactcagg acaatcaagc ttgcatgcct gcaggctgac tctaga	1966

REIVINDICACIONES

1. Método para la producción de un virión parvoviral recombinante, donde el método comprende los pasos de:
- 5 (a) proporcionar una célula de insecto que comprende uno o más constructos de ácidos nucleicos que comprenden:
- (i) una secuencia de nucleótidos que comprende un transgen que se flanquea por al menos una secuencia de nucleótidos terminales invertidas de repetición parvoviral;
- 10 (ii) un primer casete de expresión que comprende una secuencia de nucleótidos que codifica una proteína Rep parvoviral que está operativamente enlazada a un primer promotor que es capaz de dirigir la expresión de la proteína Rep en la célula de insecto;
- (iii) un segundo casete de expresión que comprende una secuencia de nucleótidos que codifica una proteína cápsida parvoviral que está operativamente enlazada a un segundo promotor que es capaz de dirigir la expresión de la proteína cápsida en la célula de insecto;
- 15 (b) cultivo de la célula definida en (a) bajo una condición propicia para la expresión de la proteína Rep y cápsida; y,
- (c) opcionalmente recuperación del virión parvoviral recombinante;
- donde los primeros y segundos casetes de expresión están presentes en un constructo de ácidos nucleicos único, donde los primeros y segundos casetes de expresión están presentes en la cantidad equimolar en la célula de insecto, donde el primer promotor es seleccionado del grupo que consiste en un promotor PolH, promotor p10 o promotor básico y donde el segundo promotor es un promotor deltaE1 o un promotor E1.
- 20 2. Un método, según la reivindicación 1, donde la proporción de expresión de la proteína Rep contra la proteína cápsida se regula por uno o más de lo siguiente:
- (a) la presencia de más elementos intensificadores y/o más fuertes en el primer casete de expresión en comparación con el segundo casete de expresión;
- 25 (b) la secuencia de nucleótidos que codifica la proteína Rep parvoviral tiene un índice de adaptación de codón más alto en comparación con la secuencia de nucleótidos que codifica la proteína cápsida; y/o
- (c) optimización de la temperatura de la proteína Rep parvoviral.
- 30 3. Método según las reivindicaciones 1 o 2, donde la secuencia de nucleótidos de (iii) comprende un marco de lectura abierto que comprende las secuencias de nucleótidos que codifican las proteínas cápsidas VP1, VP2 y VP3.
4. Método según cualquiera de las reivindicaciones precedentes, donde la secuencia de nucleótidos de (ii) comprende un marco de lectura abierto que comprende unas secuencias de nucleótidos que codifican al menos una de las proteínas Rep78 y Rep68.
- 35 5. Método según la reivindicación 4, donde la secuencia de nucleótidos de (ii) comprende un marco de lectura abierto que comprende secuencias de nucleótidos que codifican al menos una de las proteínas Rep78 y Rep 68.
- 40 6. Método según cualquiera de las reivindicaciones precedentes, donde al menos un marco de lectura abierto que comprende unas secuencias de nucleótidos que codifican las proteínas cápsidas VP1, VP2 y VP3 o al menos un marco de lectura abierto que comprende unas secuencias de nucleótidos que codifican al menos una de las proteínas Rep78 y Rep68 no comprende un intrón artificial.
- 45 7. Método, según la reivindicación 6, donde ningún marco de lectura abierto que comprende las secuencias de nucleótidos que codifican las proteínas cápsidas VP1, VP2 y VP3 y/o ningún un marco de lectura abierto que comprende las secuencias de nucleótidos que codifican al menos una de las proteínas Rep78 y Rep68 comprende un intrón artificial.
- 50 8. Método según cualquiera de las reivindicaciones precedentes, donde el primer casete de expresión comprende al menos un elemento intensificador de baculovirus y/o al menos un elemento que reacciona a la ecdisona.
9. Método según la reivindicación 8, donde el elemento intensificador se selecciona del grupo que consiste en hr1, hr2, hr3, hr4 y hr5.
- 55 10. Método según cualquiera de las reivindicaciones precedentes, donde la secuencia ITR parvoviral, proteína Rep parvoviral y/o proteína Cap parvoviral provienen de un virus adeno-asociado (AAV).
- 60 11. Constructo de ácidos nucleicos que comprende un primer y un segundo casete de expresión tal y como se define en cualquiera de las reivindicaciones 1 - 10, donde:
- (a) el primer promotor es un promotor PolH y el segundo promotor es un promotor deltaE1 o E1;
- (b) el primer promotor es un promotor p10 y el segundo promotor es un promotor deltaE1 o un E1;
- y donde el primer casete de expresión opcionalmente comprende un elemento intensificador.

12. Célula de insecto tal y como se define en cualquiera de las reivindicaciones 1 - 10, donde la célula de insecto comprende uno o más constructos de ácidos nucleicos que comprenden:

(i) una secuencia de nucleótidos que comprende un transgen que se flanquea por al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral;

5 (ii) un primer casete de expresión que comprende una secuencia de nucleótidos que codifica una proteína Rep parvoviral que está operativamente enlazada a un primer promotor que es capaz de dirigir la expresión de la proteína Rep en la célula de insecto;

10 (iii) un segundo casete de expresión que comprende una secuencia de nucleótidos que codifica una proteína cápsida parvoviral que está operativamente enlazada a un segundo promotor que es capaz de dirigir la expresión de la proteína cápsida en la célula de insecto;

donde los primeros y segundos casetes de expresión están presentes en un constructo de ácidos nucleicos único, donde los primeros y segundos casetes de expresión están presentes en la cantidad equimolar en la célula de insecto, donde el primer promotor está seleccionado del grupo que consiste en un promotor PolH, promotor p10 o promotor básico y donde el segundo promotor es un promotor deltaE1 o un promotor E1.

15 13. Un equipo que comprende

(a) un constructo de ácidos nucleicos que comprende el primer y segundo casete de expresión tal y como se define en cualquiera de las reivindicaciones 1-10; y

20 (b) un constructo de ácidos nucleicos que comprende una secuencia de nucleótidos que codifica un sitio de clonación múltiple para un transgen, que se flanquea por al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral, donde el transgen está operativamente enlazado a un promotor capaz de dirigir la expresión del transgen en una célula huésped.

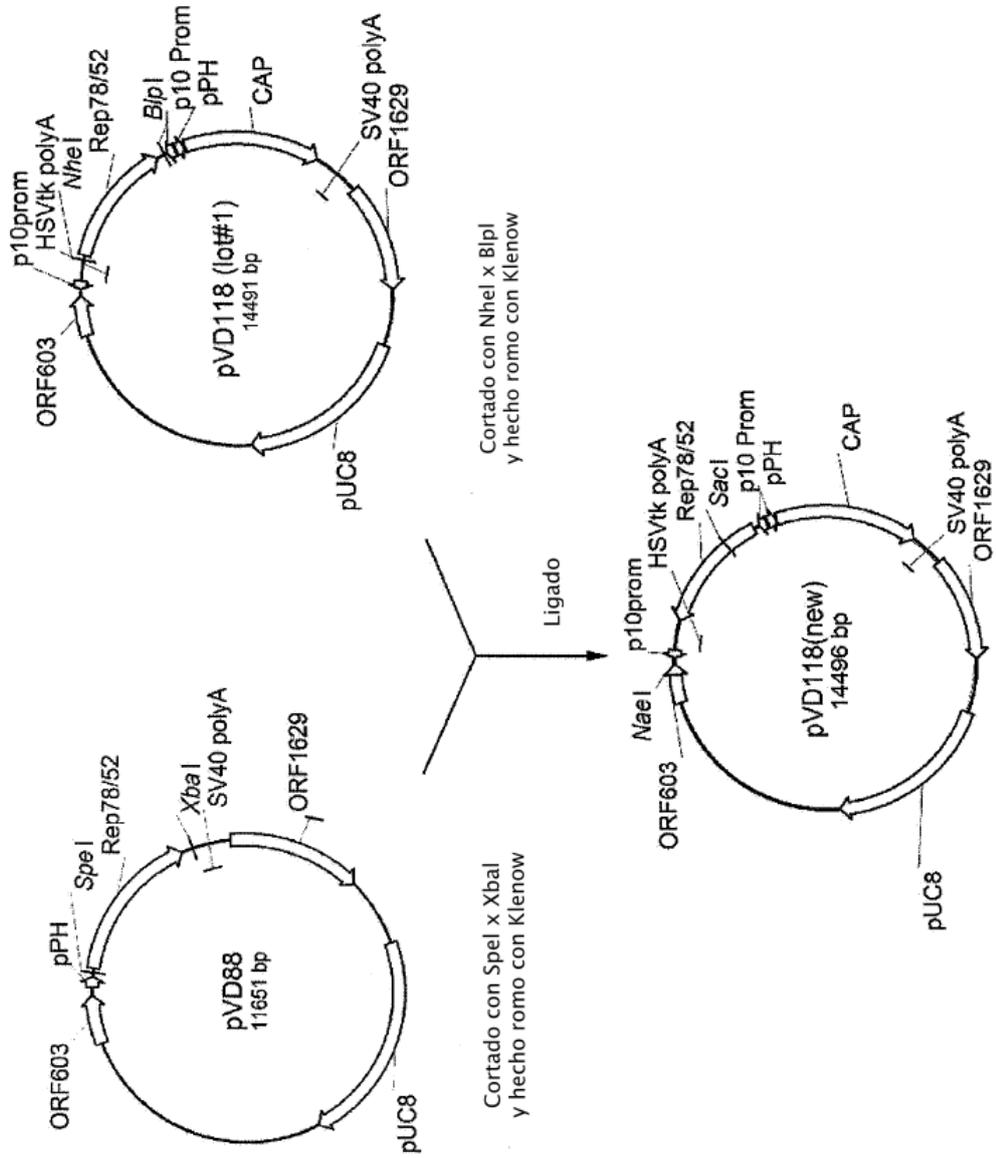
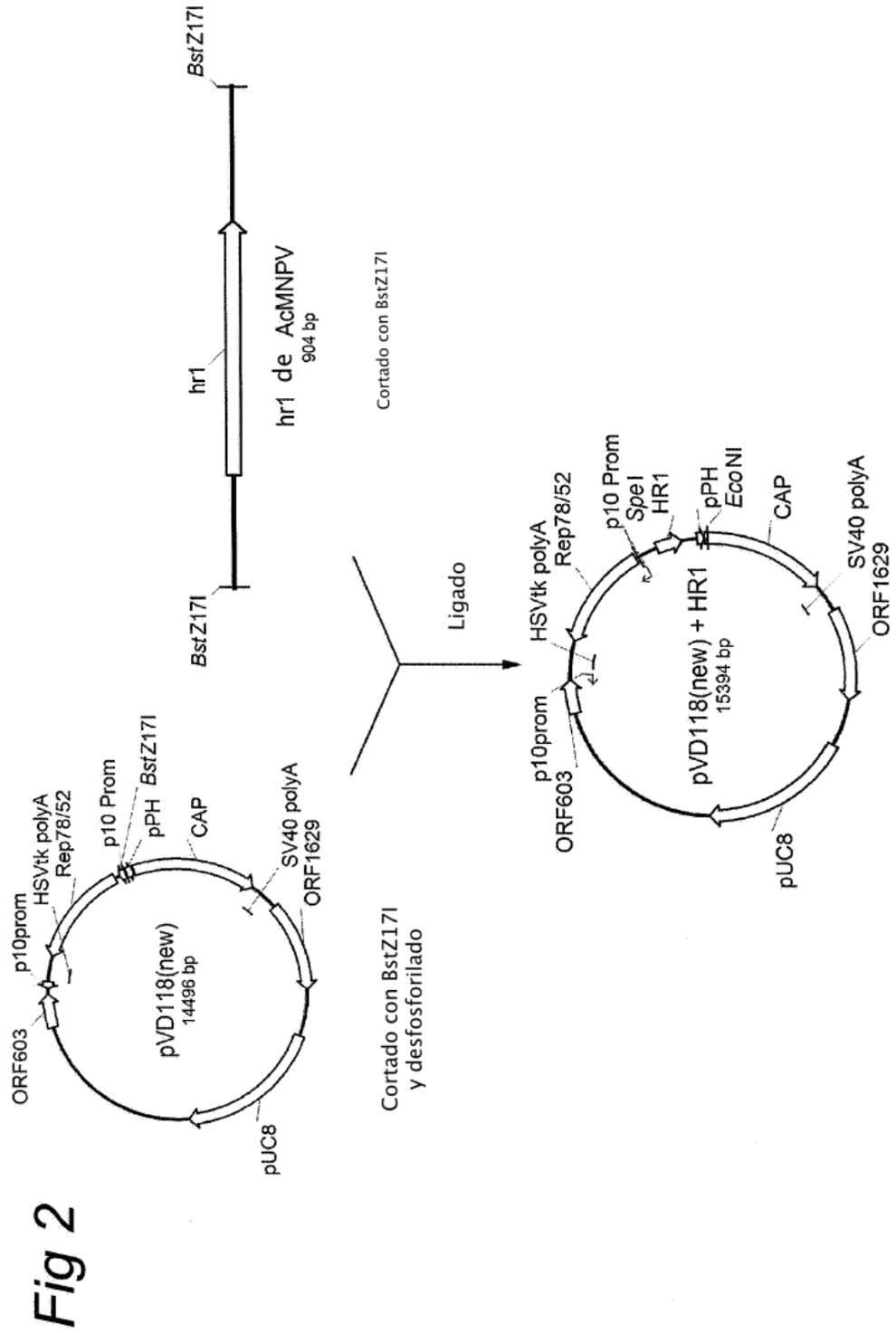


Fig 1



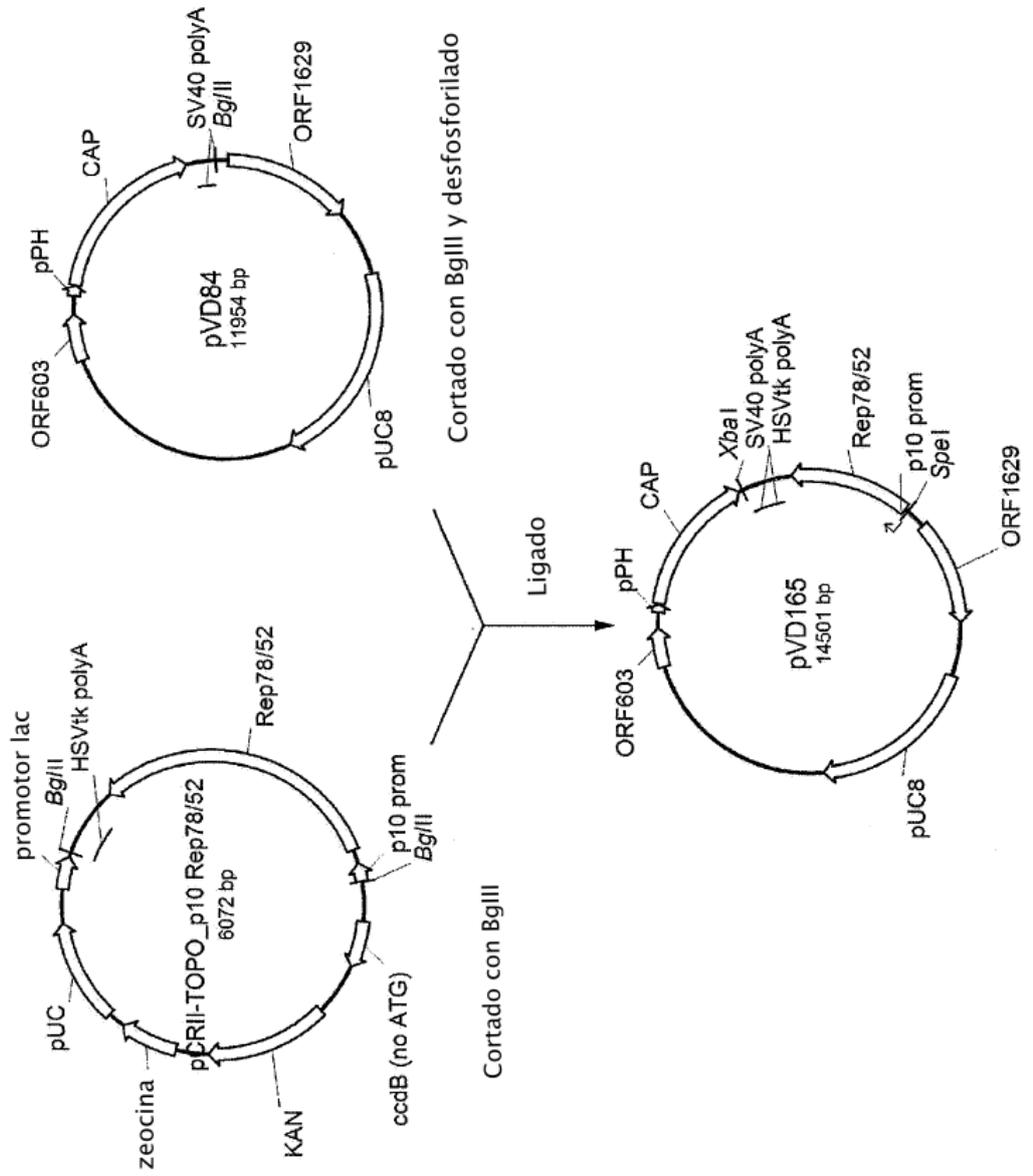


Fig 3

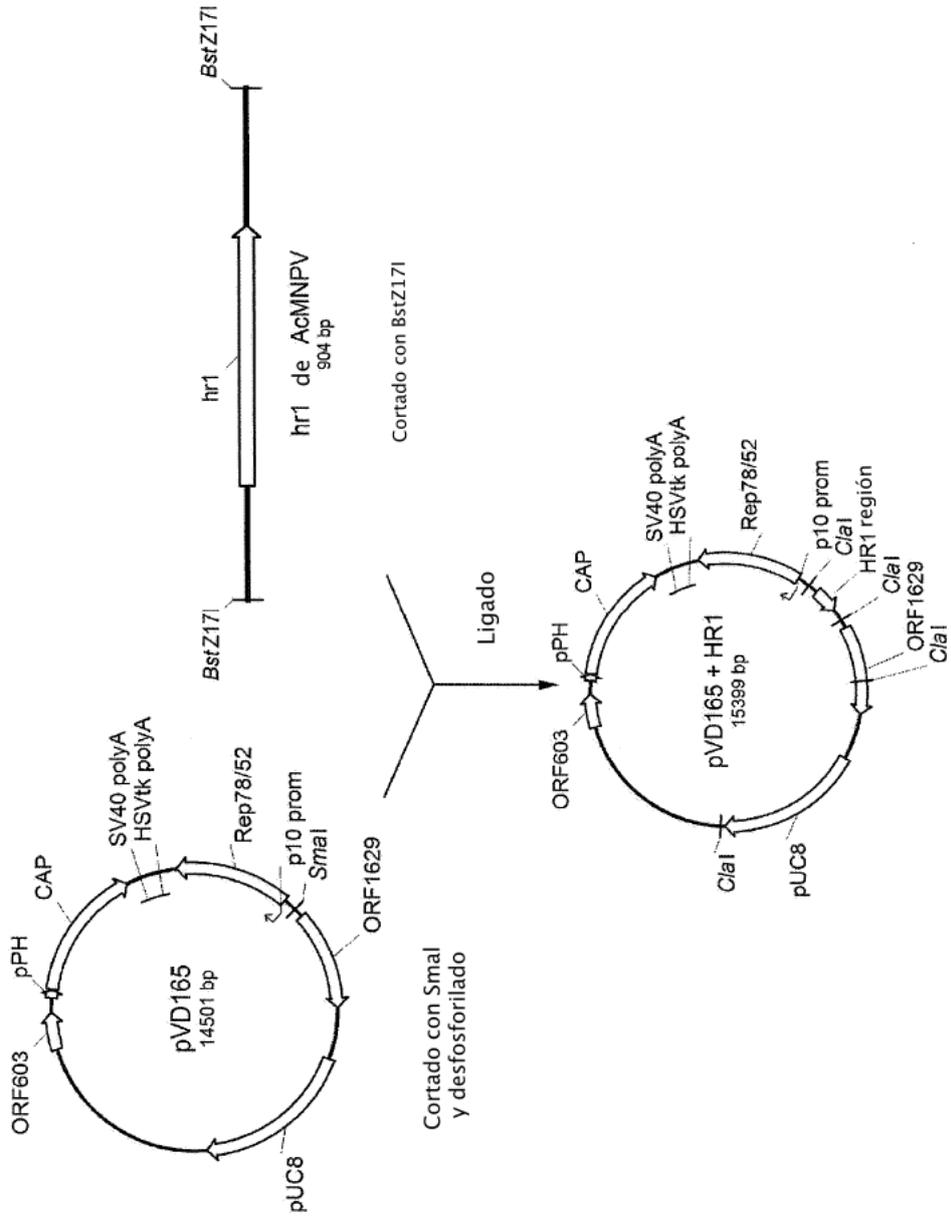
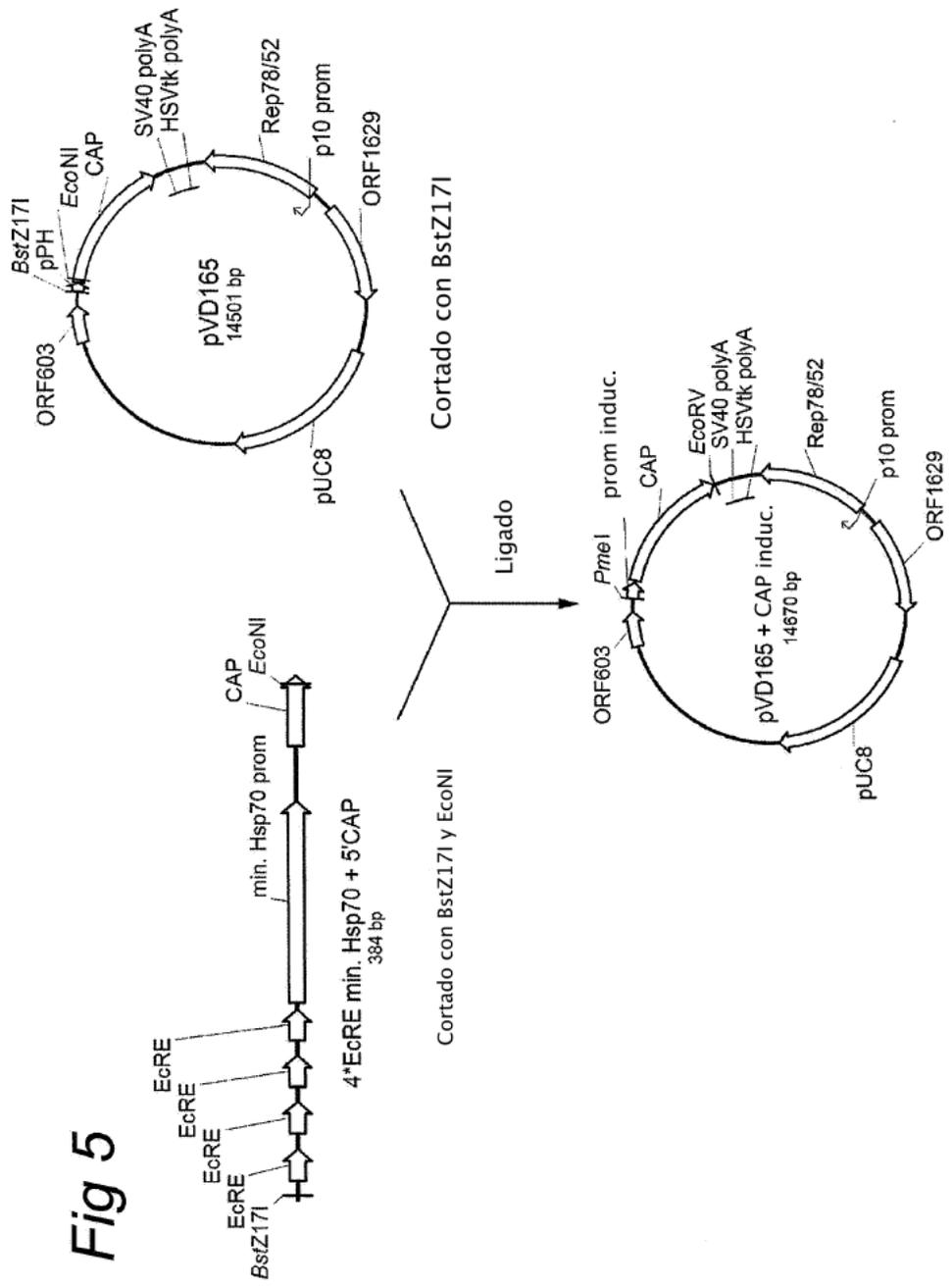
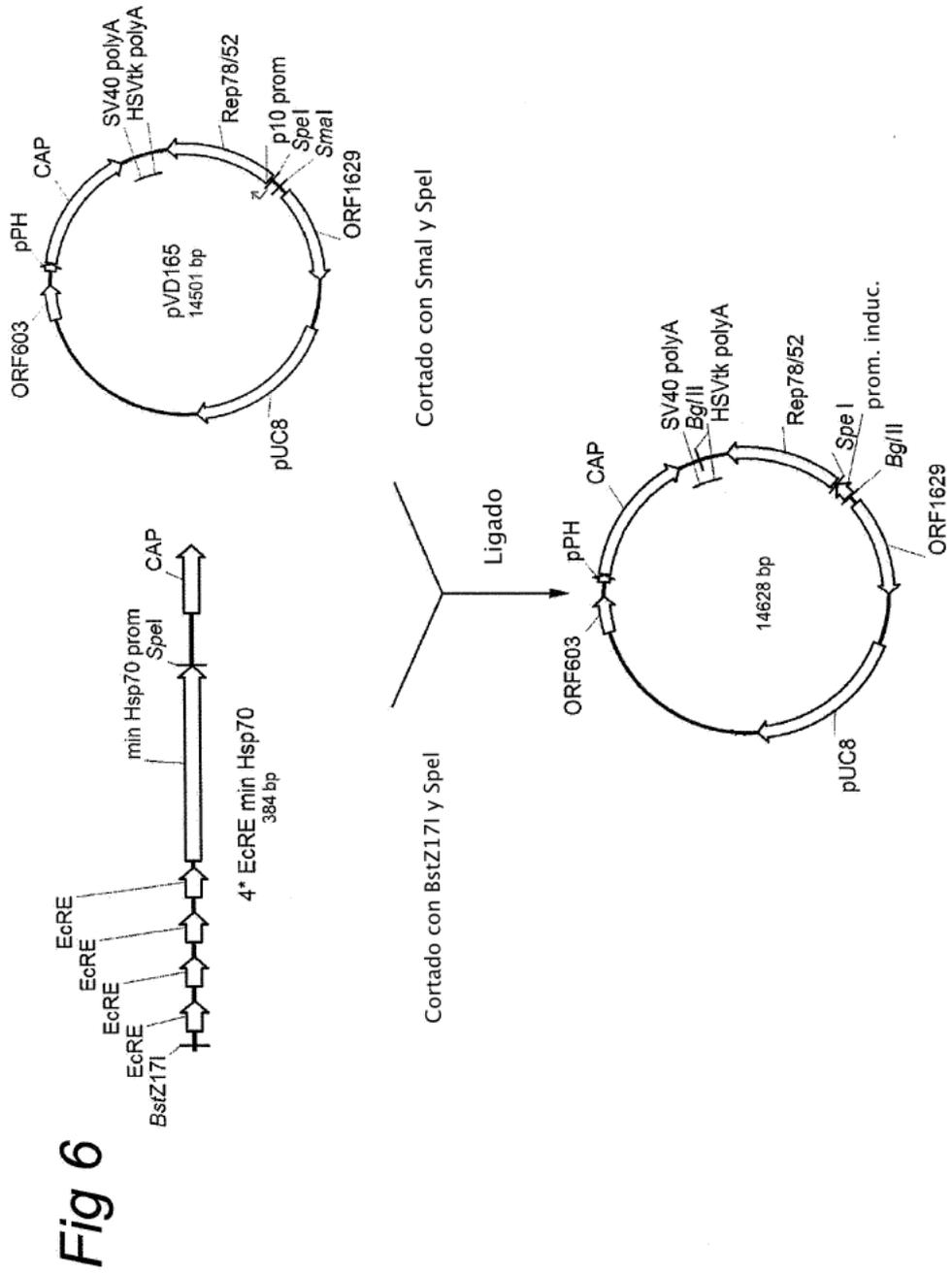
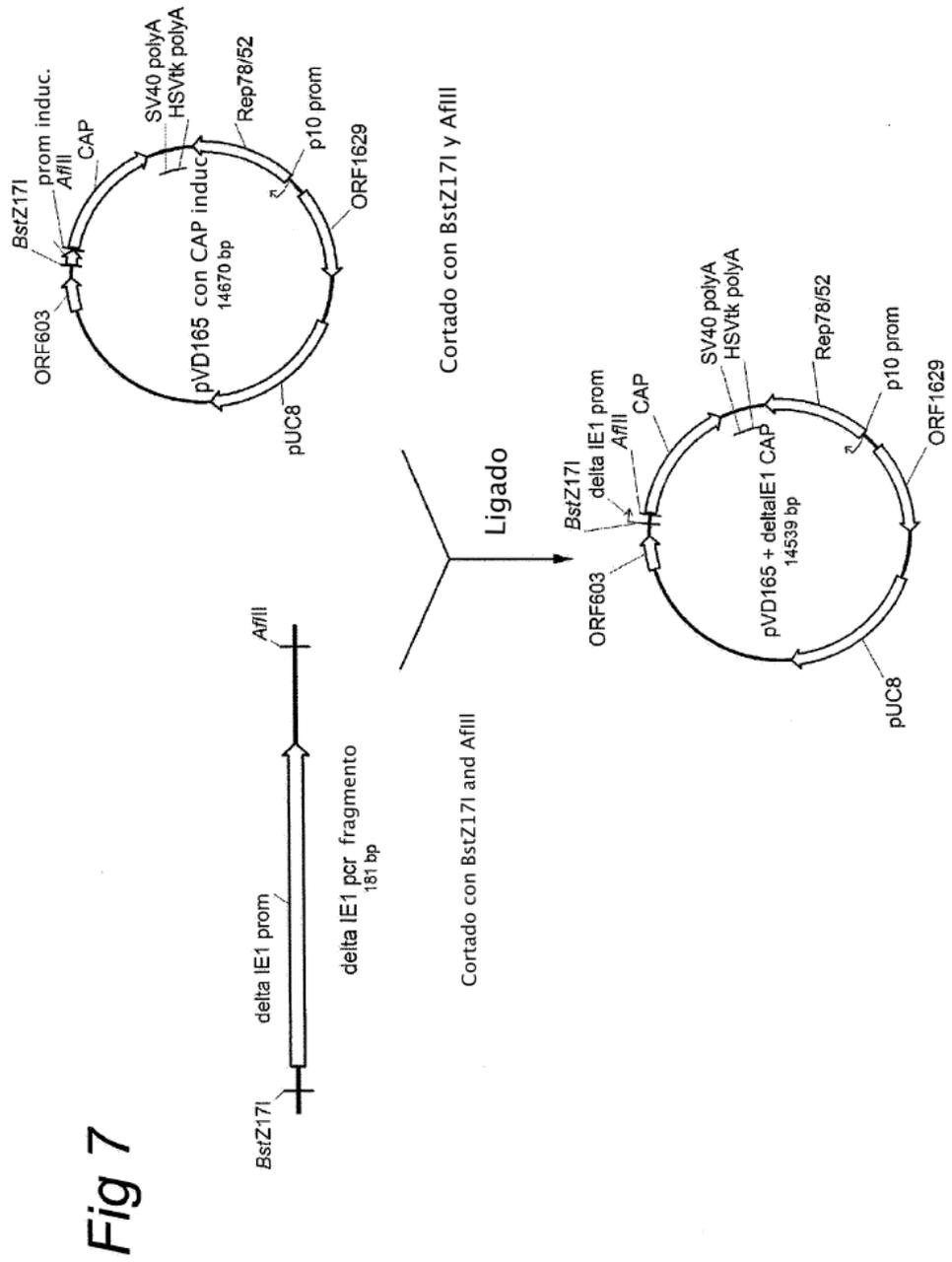
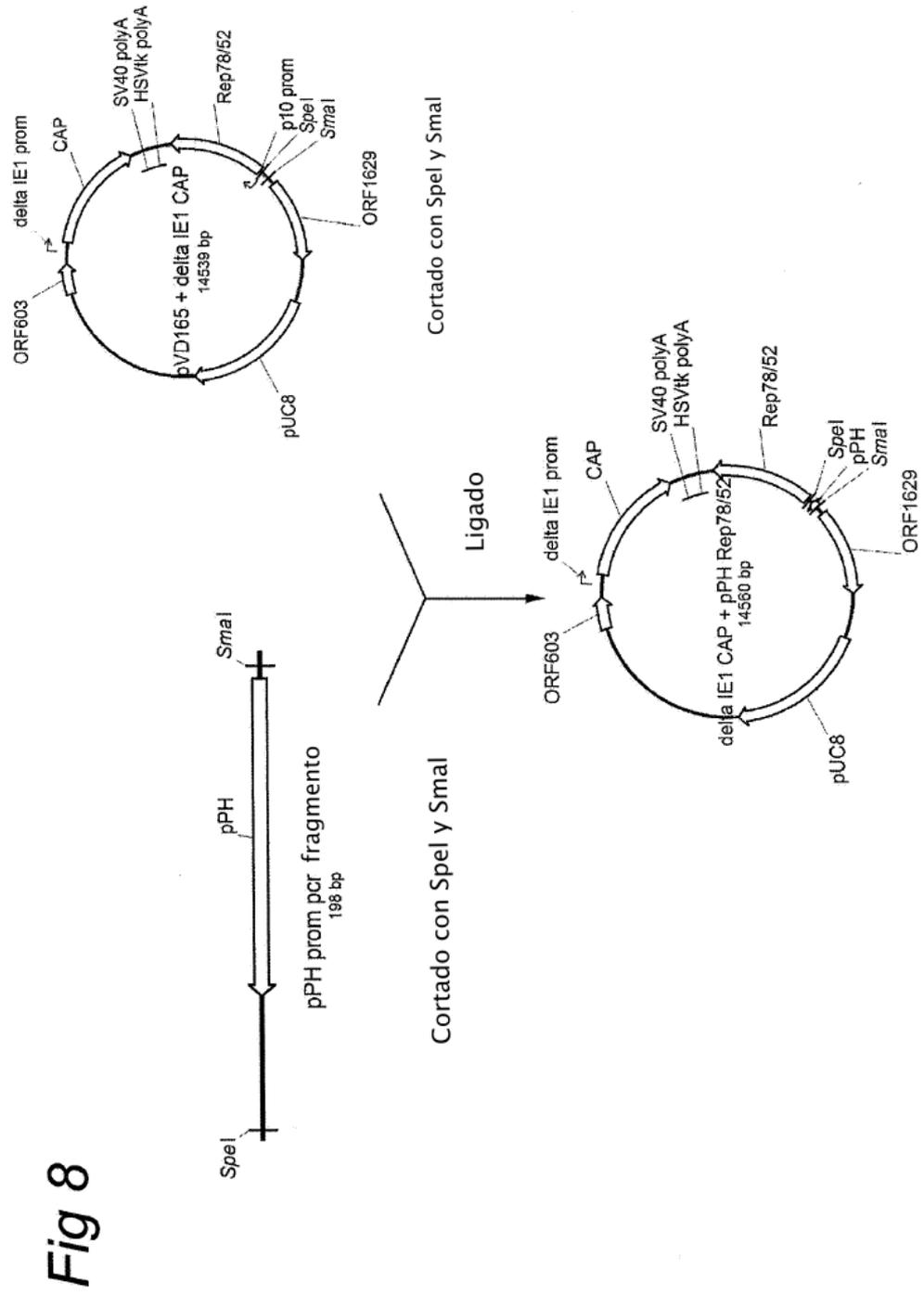


Fig 4









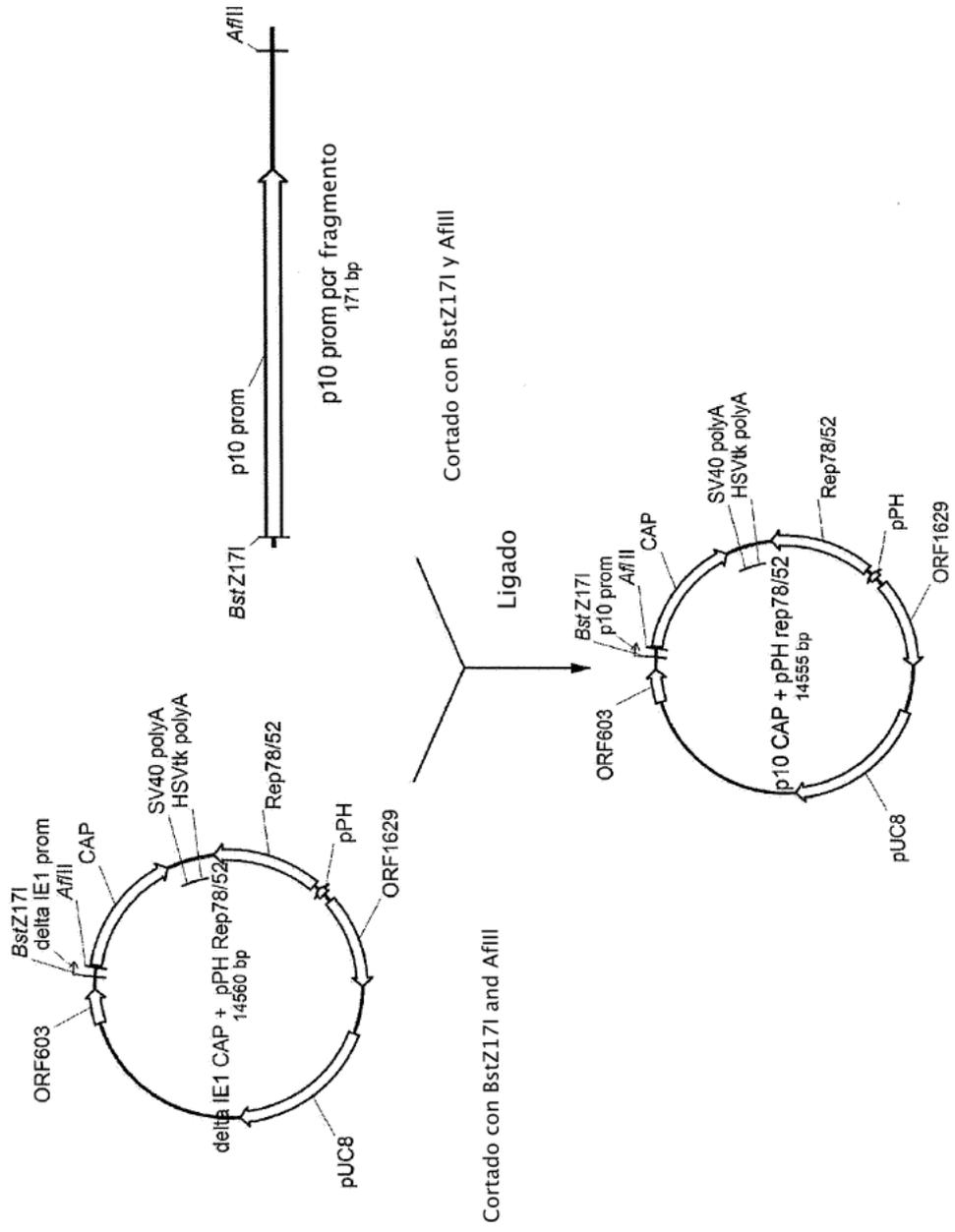


Fig 9

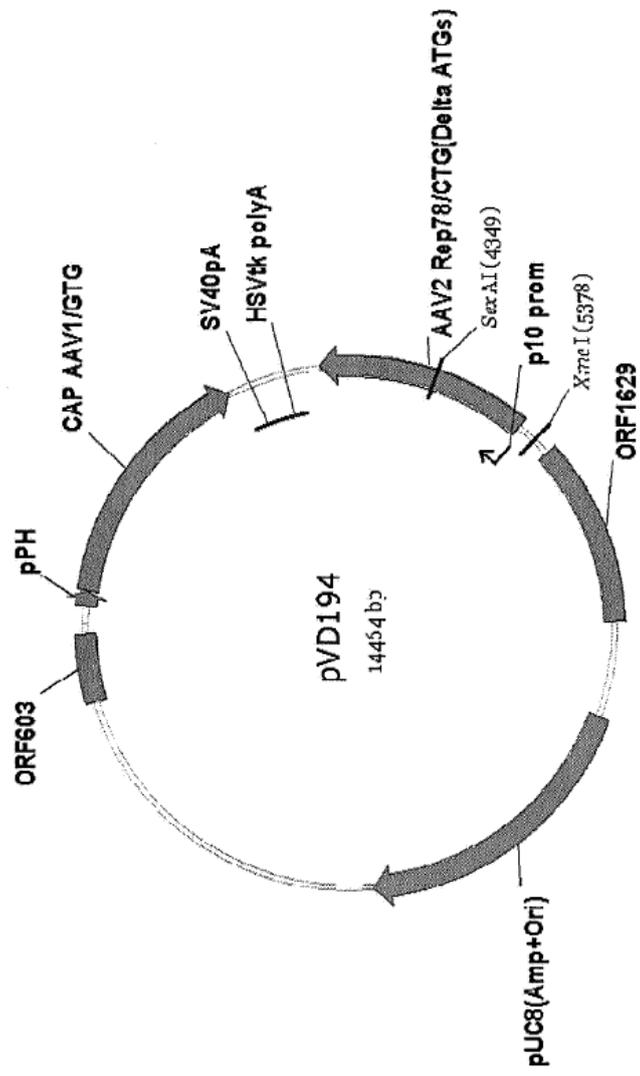


Fig. 10

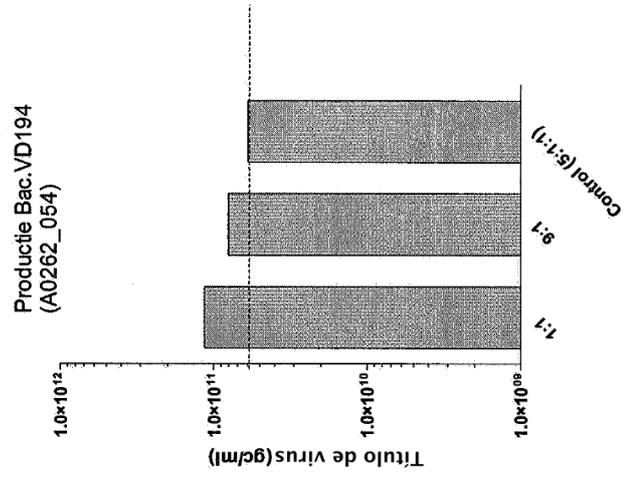


Fig. 11B

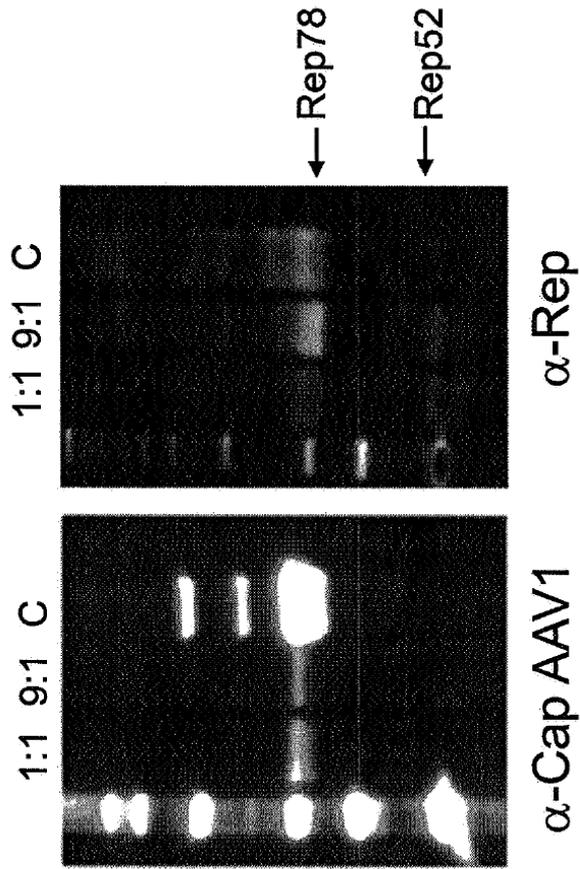


Fig. 11A