(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 929 031**

(51) Int. CI.:

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

(96) Fecha de presentación y número de la solicitud europea: **19.02.2009 E 17181630 (9)**

(97) Fecha y número de publicación de la concesión europea: **03.08.2022 EP 3257937**

(54) Título: **Optimización de la expresión de las proteínas rep y cap parvovirales en 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:
24.11.2022

(73) Titular/es:

UNIQURE IP B.V. (100.0%)
Paasheuvelweg 25
1105 BP Amsterdam, NL

(72) Inventor/es:

BAKKER, ANDREW CHRISTIAN;
HERMENS, WILHELMUS THEODORUS JOHANNES MARIA y
NOORDMAN, YVET

(74) Agente/Representante:

TOMAS GIL, Tesifonte Enrique

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 la expresión de las proteínas rep y cap parvovirales en células de insecto

Campo de la invención

[0001] Esta invención se refiere a la producción de vectores de parvovirus, especialmente a la producción de virus adenoasociados recombinantes (rAAV) en células de insectos, vectores de expresión baculoviral que comprenden un constructo de la invención y una célula que comprende dicho vector de expresión baculoviral.

Antecedentes de la invención

[0002] El sistema de expresión de baculovirus es bien conocido por su uso como vector de clonación y expresión de eucariotas. King, L. A. y R. D. Possee, 1992, "The baculovirus expression system", Chapman and Hall, Reino Unido; O'Reilly, D. R., et al., 1992. Baculovirus Expression Vectors: A Laboratory Manual. Nueva York: WH Freeman.). Las ventajas del sistema de expresión de baculovirus son, entre otras, que las proteínas expresadas son casi siempre solubles, están correctamente plegadas y son biológicamente activas. Otras ventajas incluyen altos niveles de expresión de proteínas, producción más rápida, idoneidad para la expresión de proteínas grandes e idoneidad para la producción a gran escala. Sin embargo, en la producción a gran escala o continua de proteínas heterólogas utilizando el sistema de expresión de baculovirus en biorreactores de células de insectos, la inestabilidad de los niveles de producción, también conocida como efecto de paso, es un obstáculo importante. Este efecto se debe, al menos en parte, a la recombinación entre secuencias homólogas repetidas en el ADN baculoviral.

[0003] El sistema de expresión de baculovirus también se ha utilizado con éxito para la producción de vectores recombinantes de virus adenoasociados (AAV) (Urabe et al., 2002, Hum. Gene Ther. 13: 1935-1943; EE. UU. 6.723.551 y EE. UU. 20040197895). El AAV puede considerarse como uno de los vectores virales más prometedores para la terapia génica humana. El AAV tiene la capacidad de infectar eficientemente células humanas, tanto las que se dividen como las que no. El genoma viral del AAV se integra en un solo sitio cromosómico en el genoma de la célula huésped y, lo que es más importante, aunque el AAV está presente en muchos humanos, nunca se ha asociado con ninguna enfermedad. En vista de estas ventajas, el virus adenoasociado recombinante (rAAV) se está evaluando en ensayos clínicos de terapia génica para hemofilia B, melanoma maligno, fibrosis quística, hiperlipoproteinemia tipo I y otras enfermedades.

[0004] Para superar los problemas con los sistemas de producción de mamíferos para AAV, Urabe et al. (2002, *supra)* desarrollaron un sistema de producción de AAV en células de insecto. Para la producción de AAV en células de insecto, fueron necesarias algunas modificaciones para lograr la estequiometría correcta de las tres proteínas de la cápside de los AAV (VP1, VP2 y VP3), que se basa en una combinación del uso alternativo de dos sitios aceptores de empalme y la utilización subóptima de un codón de iniciación ACG para VP2 que no se reproduce con precisión en las células de insecto. Para imitar la estequiometría correcta de las proteínas de la cápside en células de insectos, Urabe et al. (2002, *supra)* usan un constructo que se transcribe en un solo mensajero policistrónico que puede expresar las tres proteínas VP sin necesidad de corte y empalme y en el que el codón iniciador más alejado aguas arriba se reemplaza por el codón iniciador subóptimo ACG. WO2007/046703 describe una mejora adicional de la infectividad de la producción basada en vectores de rAAV producidos por baculovirus mediante la optimización de la estequiometría de las proteínas de la cápside de AAV tal como se producen en células de insecto.

[0005] Para la expresión de las proteínas Rep de AAV en el sistema de expresión de células de insectos AAV desarrollado inicialmente por Urabe et al. (2002, supra), se usa un constructo de baculovirus recombinante que alberga dos unidades de expresión de Rep independientes (una para Rep78 y otra para Rep52), cada una de ellas bajo el control de un promotor de células de insecto separado, los promotores ΔIE1 y PolH, respectivamente. Sin embargo, Kohlbrenner et al. (2005, Mol. Ther. 12: 1217-25; WO 2005/072364) informaron de que el constructo de baculovirus para la expresión de la proteína de dos Rep, tal como lo utilizan Urabe et al., adolece de una inestabilidad inherente. Al dividir la orientación palindrómica de los dos genes Rep en el vector original de Urabe y diseñar dos vectores de baculovirus separados para expresar Rep52 y Rep78, Kohlbrenner et al. (2005, supra) aumentaron la estabilidad de paso del vector. Sin embargo, a pesar de la expresión consistente de Rep78 y Rep52 a partir de los dos constructos independientes de baculovirus Rep en células de insecto durante al menos 5 pases, el rendimiento del vector rAAV es de 5 a 10 veces menor en comparación con el constructo original de baculovirus Rep diseñado por Urabe et al. (2002, supra).

[0006] En la solicitud WO2007/148971 los presentes inventores han mejorado significativamente la estabilidad de la producción del vector rAAV en células de insecto mediante el uso de una sola secuencia codificante para las proteínas Rep78 y Rep52 en la que se usa un codón iniciador subóptimo para la proteína Rep78 omitida parcialmente por los ribosomas que se están explorando para permitir que el inicio de traducción también ocurra aguas abajo en el codón de iniciación de la proteína Rep52.

[0007] La solicitud de patente internacional WO 2007/084773 describe un método de producción de rAAV en células de insecto, en el que la producción de partículas víricas infecciosas aumenta complementando VP1 con respecto a VP2 y VP3. La suplementación se puede efectuar introduciendo en la célula de insecto un vector de cápside que comprende secuencias de nucleótidos que expresan VP1, VP2 y VP3 y, además, introduciendo en la célula de insecto secuencias de nucleótidos que expresan VP1, que pueden estar en el mismo vector de cápside o en un vector diferente.

[0008] Sin embargo, todavía existe la necesidad de mejoras adicionales en la producción a gran escala (comercial) de vectores parvovirales en células de insectos. Por lo tanto, es un objetivo de la presente invención proporcionar medios y métodos que permitan una producción estable y de alto rendimiento (a gran escala) de vectores parvovirales y una producción que resulte en una proporción mejorada de partículas llenas:vacías (es decir, una mayor proporción de partículas llenas).

Resumen de la invención

[0009] La invención se refiere a un método para la producción de un virión parvoviral recombinante. El uso de dicho método puede permitir la producción de dichos viriones con títulos de producción incrementados. El uso del método puede, alternativa o adicionalmente, permitir una producción que tenga una mayor proporción de partículas llenas, es decir, una relación llenas:vacías o una relación totales:llenas más favorable.

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

(a) proporcionar una célula de insecto que comprende uno o varios constructos de ácidos nucleicos que comprenden:
(i) una secuencia de nucleótidos que comprende un transgén que está flanqueado 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 Rep78 parvoviral a partir de la cual se expresan las proteínas Rep52 y Rep78 en la célula de insecto, cuya secuencia de nucleótidos está operativamente enlazada a un primer promotor que es capaz de impulsar la expresión de la proteína Rep en la célula de insecto, y en el que la secuencia de nucleótidos que codifica la proteína parvoviral Rep78 no comprende un intrón artificial;
(iii) un segundo casete de expresión que comprende una secuencia de nucleótidos que codifica una proteína de la cápside parvoviral que está operativamente enlazada a un segundo promotor que es capaz de dirigir la expresión de la proteína de la cápside en la célula de insecto; y
(b) cultivar la célula definida en (a) en condiciones propicias para la expresión de la Rep y la proteína de la cápside;

donde los casetes de expresión primero y segundo están presentes en un solo constructo de ácidos nucleicos, donde los casetes de expresión primero y segundo están presentes en cantidad equimolar en la célula de insecto, y donde el primer promotor es tan fuerte como el segundo promotor.

[0011] La invención también proporciona:

una célula de insecto tal como se define en las reivindicaciones adjuntas; y
un kit que comprende

(a) un constructo de ácidos nucleicos que comprende el primer y segundo casete de expresión tal como se define en las reivindicaciones adjuntas; y
(b) un constructo de ácidos nucleicos que comprende una secuencia de nucleótidos que codifica un sitio de clonación múltiple para un transgén que está flanqueado por al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral, cuyo transgén está operativamente enlazado a un promotor capaz de impulsar la expresión del transgén en una célula huésped.

Breve descripción de las figuras.

[0012]

La figura 1 muestra una representación esquemática de la construcción de pVD118 (nuevo).
La figura 2 muestra una representación esquemática de la construcción de pVD118 (nuevo) + 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 + deltaIE1 CAP.
La figura 8 muestra una representación esquemática de la construcción de deltaIE1 Cap + pPolh Rep (pVD190).
La figura 9 muestra una representación esquemática de la construcción de p10 Cap + pPolh Rep.
La figura 10 muestra una representación esquemática de la construcción de pVD194.
La figura 11A muestra el análisis de transferencia Western de las proteínas Cap y Rep expresadas a partir de Bac.VD194, tres días después de la infección con el baculovirus en proporciones 1:1 y 5:1 (C=Bac.VD88:Bac.VD84:Bac.VD43 a 5:1:1 respectivamente). La figura 11B muestra los títulos virales para las mismas producciones analizadas en la figura 11A.

Descripción de la invención

Definiciones

[0013] Tal como se usa en el presente documento, el término "operativamente enlazado" se refiere a un enlace de elementos polinucleotídicos (o polipeptídicos) en una relación funcional. Un ácido nucleico está "operativamente enlazado" cuando se coloca en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, una secuencia reguladora de la transcripción está operativamente enlazada a una secuencia codificante si afecta a la transcripción de la secuencia codificante. Operativamente enlazado significa que las secuencias de ADN que se enlazan son típicamente contiguas y, cuando es necesario unir dos regiones codificantes de proteínas, son contiguas y están en el marco de lectura.

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

[0015] Una secuencia de control de la expresión está operativamente enlazada a una secuencia de nucleótidos cuando la secuencia de control de la expresión controla y regula la transcripción y/o la traducción de la secuencia de nucleótidos. Por lo tanto, una secuencia de control de la expresión puede incluir promotores, potenciadores, sitios internos de entrada al ribosoma (IRES), terminadores de la transcripción, un codón de inicio frente a un gen que codifica una proteína, una señal de empalme para intrones y codones de terminación. El término "secuencia de control de la expresión" pretende incluir, como mínimo, una secuencia cuya presencia está diseñada para influir en la expresión y también puede incluir componentes ventajosos adicionales. Por ejemplo, las secuencias líder y las secuencias asociadas de fusión son secuencias de control de la expresión. El término también puede incluir el diseño de la secuencia de ácidos nucleicos de manera que los codones de iniciación potenciales no deseables dentro y fuera del marco se eliminen de la secuencia. También puede incluir el diseño de la secuencia de ácidos nucleicos de manera que se eliminen los sitios de empalme potenciales no deseables. Incluye secuencias, entre ellas de poliadenilación (pA) que dirigen la adición de una cola poliA, es decir, una cadena de residuos de adenina en el extremo 3' de un ARNm, secuencias denominadas secuencias poliA. También se puede diseñar para mejorar la estabilidad del ARNm. Las secuencias de control de la expresión que afectan a la estabilidad de la transcripción y la traducción, por ejemplo, los promotores, así como las secuencias que afectan a la traducción, por ejemplo, las secuencias de Kozak, son conocidas en la células de insectos. Las secuencias de control de la expresión pueden ser de tal naturaleza que modulen la secuencia de nucleótidos a la que están operativamente enlazadas de modo que se consigan niveles de expresión más bajos o más altos.

[0016] Según se usa en el presente documento, los términos "promotor" o "secuencia reguladora de la transcripción" se refieren a un fragmento de ácidos nucleicos que funciona para controlar la transcripción de una o más secuencias codificantes, y está ubicado corriente arriba con respecto a la dirección de transcripción del sitio de inicio de la transcripción de la secuencia codificante, y se identifica estructuralmente por la presencia de un sitio de unión para la ARN polimerasa dependiente de ADN, sitios de iniciación de la transcripción y cualquier otra secuencia de ADN, incluyendo, entre otros, los sitios de unión al factor de transcripción, los sitios de unión a las proteínas represora y activadora y cualquier otra secuencia de nucleótidos conocida por un experto en la técnica que actúe directa o indirectamente para regular la cantidad de transcripción del promotor. Un promotor "constitutivo" es un promotor que está activo en la mayoría de los tejidos en la mayoría de las condiciones fisiológicas y de desarrollo. Un promotor "inducible" es un promotor que está regulado fisiológica o evolutivamente, p. ej., mediante la aplicación de un inductor químico. Un promotor "específico de tejido" solo está activo en tipos específicos de tejidos o células.

[0017] La "identidad de secuencia" y la "similitud de secuencia" se pueden determinar mediante la alineación de dos secuencias de péptidos o de dos nucleótidos usando algoritmos de alineación globales o locales, dependiendo de la longitud de las dos secuencias. Las secuencias de longitudes similares se alinean preferiblemente usando algoritmos de alineación globales (por ejemplo, Needleman Wunsch) que alinea las secuencias de manera óptima en toda la longitud, mientras que las secuencias de longitudes sustancialmente diferentes se alinean preferiblemente usando un algoritmo de alineación local (por ejemplo, Smith Waterman). Las secuencias pueden entonces denominarse "sustancialmente idénticas" o "esencialmente similares" cuando comparten al menos un cierto porcentaje mínimo de identidad de secuencia (como se define a continuación) (cuando se alinean de manera óptima, por ejemplo, mediante los programas GAP o BESTFIT utilizando

parámetros predeterminados). GAP utiliza el algoritmo de alineación global de Needleman y Wunsch para alinear dos secuencias en toda su longitud (longitud completa), maximizando el número de coincidencias y minimizando el número de gaps. Una alineación global se usa adecuadamente para determinar la identidad de secuencia cuando las dos secuencias tienen longitudes similares. Generalmente, se utilizan los parámetros predeterminados de GAP, con una penalización por creación de gap = 50 (nucleótidos)/8 (proteínas) y una penalización por extensión de gap = 3 (nucleótidos)/2 (proteínas). Para los nucleótidos, la matriz de puntuación predeterminada utilizada es nwsgapdna y para las proteínas, la matriz de puntuación predeterminada es Blosum62 (Henikoff y Henikoff, 1992, PNAS 89, 915-919). Las alineaciones de secuencias y las puntuaciones para el porcentaje de identidad de secuencias se pueden determinar utilizando programas informáticos, como el paquete Wisconsin de GCG, versión 10.3, disponible en Accelrys Inc., 9685 Scranton Road, San Diego, CA 92121-3752 EE. UU., o utilizando software de código abierto, como el programa "needle" (usando el algoritmo global de Needleman Wunsch) o "water" (usando el algoritmo local de Smith Waterman) en EmbossWIN versión 2.10.0, usando los mismos parámetros que para GAP anteriormente, o usando la configuración predeterminada (ambos para "needle" y para "water" y tanto para proteínas como para alineaciones de ADN, la penalización por abertura de gap predeterminada es 10.0 y la penalización por extensión de gap predeterminada es 0.5; las matrices de puntuación predeterminadas son Blossum62 para proteínas y DNAFull para ADN). Cuando las secuencias tienen longitudes totales sustancialmente diferentes, se prefieren las alineaciones locales, como las que utiliza el algoritmo de Smith Waterman. Alternativamente, el porcentaje de similitud o identidad puede determinarse mediante la búsqueda en bases de datos públicas, utilizando algoritmos como FASTA, BLAST, etc.

[0018] Las secuencias de nucleótidos que codifican las proteínas Cap y/o Rep parvovirales de la invención también pueden definirse por su capacidad para hibridarse con las secuencias de nucleótidos de la SEC ID N.º bajo estrictas condiciones de hibridación.

[0019] Las condiciones de hibridación estrictas se definen en el presente documento como condiciones que permiten que una secuencia de ácidos nucleicos de al menos aproximadamente 25, preferiblemente aproximadamente 50, 75 o 100 y más preferiblemente aproximadamente 200 o más nucleótidos, se hibride a una temperatura de aproximadamente 65 °C en una solución que comprende aproximadamente 1 M de sal, preferiblemente 6 X SSC o cualquier otra solución que tenga una fuerza iónica comparable, y se lave a 65 °C en una solución que comprende aproximadamente 0.1 M de sal o menos, preferiblemente 0.2 X SSC o cualquier otra solución que tenga una fuerza iónica comparable. Preferiblemente, la hibridación se realiza durante 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 que tengan aproximadamente un 90 % o más de identidad de secuencia.

[0020] Las condiciones moderadas se definen en el presente documento como condiciones que permiten que una secuencia de ácidos nucleicos de al menos aproximadamente 50 nucleótidos, preferiblemente de aproximadamente 200 o más nucleótidos, se hibride a una temperatura de aproximadamente 45 °C en una solución que comprende aproximadamente 1 M de sal, preferiblemente 6 X SSC o cualquier otra solución que tenga una fuerza iónica comparable, y se lave a temperatura ambiente en una solución que comprende aproximadamente 1 M de sal, preferiblemente 6 X SSC o cualquier otra solución que tenga una fuerza iónica comparable. Preferiblemente, la hibridación se realiza durante 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 que tengan hasta un 50 % o más de identidad de secuencia. La persona experta en la materia podrá modificar estas condiciones de hibridación para identificar específicamente secuencias cuya identidad varía entre el 50 % y el 90 %.

[0021] Un "vector" es una molécula de ácido nucleico (típicamente ADN o ARN) que sirve para transferir una secuencia de ácido nucleico pasajero (es decir, ADN o ARN) a una célula huésped. Tres tipos comunes de vectores son los plásmidos, los fagos y los virus. Preferiblemente, el vector es un virus. Los vectores que contienen tanto un promotor como un sitio de clonación en el que se puede enlazar operativamente un polinucleótido son bien conocidos en la técnica. Dichos vectores son capaces de transcribir ARN *in vitro* o *in vivo*, y están comercialmente disponibles en fuentes tales como Stratagene (La Jolla, Calif.) y Promega Biotech (Madison, Wis.). Para optimizar la expresión y/o la transcripción in vitro, puede ser necesario eliminar, añadir o alterar las porciones no traducidas de 5' y/o 3' de los clones para eliminar codones alternativos potenciales de iniciación de la traducción adicionales e inapropiados u otras secuencias que puedan interferir con la expresión o reducirla, ya sea a nivel de transcripción o de traducción. Alternativamente, los sitios de unión a ribosomas de consenso se pueden insertar inmediatamente en 5' del codón de inicio para mejorar la expresión.

[0022] Un "vector viral" se refiere a un vector que comprende algunos o todos los siguientes: genes virales que codifican un producto génico, secuencias de control y secuencias de empaquetamiento viral.

[0023] Un "vector parvoviral" se define como un parvovirus producido de forma recombinante o una partícula parvoviral que comprende un polinucleótido que se administrará en una célula huésped, ya sea *in vivo*, *ex vivo* o *in vitro.* Los ejemplos de vectores parvovirales incluyen, por ejemplo, vectores de virus adenoasociados. En el

presente documento, un constructo de vector parvoviral se refiere al polinucleótido que comprende el genoma viral o parte del mismo, y un transgén.

[0024] La adaptabilidad de una secuencia de nucleótidos que codifica una enzima al uso de codones de una célula huésped puede expresarse como índice de adaptación de codones (CAI). El índice de adaptación de codones se define en el presente documento como una medida de la adaptabilidad relativa del uso de codones de un gen al uso de codones de genes altamente expresados en una célula huésped u organismo particulares. La adaptabilidad relativa (w) de cada codón es la relación entre el uso de cada codón y el del codón más abundante para el mismo aminoácido. El índice CAI se define como la media geométrica de estos valores de adaptabilidad relativa. Se excluyen los codones no sinónimos y los codones de terminación (que dependen del código genético). Los valores de CAI varían de 0 a 1; y los valores más altos indican una mayor proporción de los codones más abundantes (ver Sharp y Li, 1987, Nucleic Acids Research 15: 1281-1295; ver también: Jansen et al., 2003, Nucleic Acids Res. 31(8):2242-51).

[0025] El término "reportero" (o gen reportero o proteína reportera) se utiliza principalmente para referirse a la secuencia de nucleótidos que codifica proteínas marcadoras visibles, como la proteína verde fluorescente (GFP), eGFP, otras proteínas fluorescentes, la luciferasa, la fosfatasa alcalina secretada (SEAP), el GUS y similares, así como marcadores nptII y similares

Descripción detallada de la invención

[0026] La presente invención se refiere al uso de parvovirus, en particular dependovirus tales como AAV humano o de simio infeccioso, y los componentes de los mismos (p. ej., un genoma de parvovirus) para usar como vectores para la introducción y/o expresión de ácidos nucleicos en células de mamífero, preferiblemente células humanas. En particular, la invención se refiere a mejoras en la productividad de dichos vectores parvovirales cuando se producen en células de insectos.

[0027] La productividad en este contexto abarca mejoras en los títulos de producción y mejoras en la calidad del producto resultante, por ejemplo, un producto que ha mejorado la relación totales:llenas (una medida del número de partículas que comprenden ácidos nucleicos). Es decir, el producto final puede tener una mayor proporción de partículas llenas, donde lleno implica que la partícula comprende ácidos nucleicos.

[0028] Los virus de la familia Parvoviridae son pequeños virus de ADN. La familia Parvoviridae se puede dividir en dos subfamilias: Parvovirinae, que infecta a los vertebrados, y Densovirinae, que infecta a los invertebrados, incluidos los insectos. Los miembros de la subfamilia Parvovirinae se denominan aquí parvovirus e incluyen el género Dependovirus. Como puede deducirse del nombre de su género, los miembros del Dependovirus son únicos en el sentido de que normalmente requieren la coinfección con un virus auxiliar como el adenovirus o el virus del herpes para la infección productiva en cultivo celular. El género Dependovirus incluye AAV, que normalmente infecta a humanos (p. ej., serotipos 1, 2, 3A, 3B, 4, 5 y 6) o primates (p. ej., serotipos 1 y 4), y virus relacionados que infectan a otros animales de sangre caliente (por ejemplo, virus adenoasociados bovinos, caninos, equinos y ovinos). Se describe más información sobre parvovirus y otros miembros de Parvoviridae en Kenneth I. Berns, "Parvoviridae: The Viruses and Their Replication", capítulo 69 en Fields Virology (3.ª ed. 1996). Por conveniencia, la presente invención se ejemplifica y describe adicionalmente en el presente documento con referencia a AAV. Sin embargo, se entiende que la invención no se limita a AAV sino que puede aplicarse igualmente a otros parvovirus.

[0029] La organización genómica de todos los serotipos de AAV conocidos es muy similar. El genoma de AAV es una molécula de ADN monocatenario lineal que tiene menos de aproximadamente 5000 nucleótidos (nt) de longitud. Las repeticiones terminales invertidas (ITR) flanquean las secuencias codificantes de nucleótidos únicas para las proteínas de replicación no estructural (Rep) y las proteínas estructurales (VP). Las proteínas VP (VP1, -2 y -3) forman la cápside. Los 145 nt terminales son autocomplementarios y están organizados de manera que se puede formar un dúplex intramolecular energéticamente estable que forma una horquilla en forma de T. Estas estructuras en horquilla funcionan como un origen para la replicación del ADN viral, sirviendo como cebadores para el complejo celular de ADN polimerasa. Después de la infección con wtAAV en células de mamífero, los genes Rep (es decir, Rep78 y Rep52) se expresan a partir del promotor P5 y el promotor P19, respectivamente, y ambas proteínas Rep tienen una función en la replicación del genoma viral. Un evento de empalme en el ORF de Rep da como resultado la expresión de cuatro proteínas Rep (es decir, Rep78, Rep68, Rep52 y Rep40).

[0030] Sin embargo, se ha demostrado que el ARNm no empalmado, que codifica las proteínas Rep78 y Rep52, en células de mamífero es suficiente para la producción del vector AAV. También en células de insecto, las proteínas Rep78 y Rep52 son suficientes para la producción del vector AAV.

[0031] Un "vector parvoviral o AAV recombinante" (o "vector rAAV") en el presente documento se refiere a un vector que comprende una o más secuencias de polinucleótidos de interés, genes de interés o "transgenes" que están flanqueados por al menos una secuencia de repetición terminal invertida (ITR) parvoviral o de AAV. Preferiblemente, el o los transgenes están flanqueados por ITR, uno a cada lado del o de los transgenes. Dichos

vectores de rAAV se pueden replicar y empaquetar en partículas virales infecciosas cuando están presentes en una célula huésped de insecto que expresa productos de los genes rep y cap de AAV (es decir, proteínas AAV Rep y Cap). Cuando un vector rAAV se incorpora a un constructo de ácidos nucleicos más grande (p. ej., en un cromosoma o en otro vector, como un plásmido o baculovirus utilizado para la clonación o la transfección), el vector rAAV generalmente se denomina "pro-vector" que puede ser "rescatado" por replicación y encapsidación en presencia de funciones de empaquetamiento AAV y funciones auxiliares necesarias.

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

[0033] En un primer aspecto, la invención está relacionada con un método para la producción de un virión parvoviral recombinante que comprende los pasos de: (a) proporcionar una célula de insecto que comprende uno o varios constructos de ácidos nucleicos que comprenden: (i) una secuencia de nucleótidos que comprende un transgén que está flanqueado 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 Rep78 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, y en el que la secuencia de nucleótidos que codifica la proteína Rep78 parvoviral no comprende un intrón artificial; (iii) un segundo casete de expresión que comprende una secuencia de nucleótidos que codifica una proteína de la cápside parvoviral que está operativamente enlazada a un segundo promotor que es capaz de dirigir la expresión de la proteína de la cápside en la célula de insecto; y (b) cultivar la célula definida en (a) en condiciones propicias para la expresión de Rep y la proteína de la cápside; donde los casetes de expresión primero y segundo están presentes en un solo constructo de ácidos nucleicos, donde los casetes de expresión primero y segundo están presentes en cantidad equimolar en la célula de insecto, y donde el primer promotor es tan fuerte como el segundo promotor.

[0034] Preferiblemente, los casetes de expresión primero y segundo presentes en cantidad equimolar en una célula de insecto producen una proporción de niveles de ARNm que codifica proteínas Rep frente a ARNm que codifica proteínas de la cápside 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 ARNm que codifican Rep y cápsides se determinan preferiblemente mediante PCR de transcripción inversa cuantitativa, transferencia Northern u otros medios para la cuantificación de ARN conocidos en la técnica.

[0035] Preferiblemente, la proporción de los niveles de ARNm que codifican Rep y cápsides se produce en la célula de insecto en un periodo de tiempo 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 ARNm que codifican Rep y cápsides se producen al menos en la célula de insecto 2 o 1 hora alrededor de 48 horas después de la transfección.

[0036] En otra alternativa preferida, los casetes de expresión primero y segundo presentes en cantidad equimolar en una célula de insecto producen una proporción de niveles de proteínas Rep frente a proteínas de la cápside 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 de la cápside se determinan preferiblemente utilizando anticuerpos de referencia contra las proteínas Rep y de la cápside en un inmunoensayo cuantitativo (p. ej., ELISA o transferencia Western cuantitativa). Preferiblemente, la proporción de los niveles de proteína Rep y de la cápside se produce en la célula de insecto en un periodo de tiempo 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 de la cápside se produce en la célula de insecto 2 o 1 hora alrededor de 48 horas después de la transfección. Los anticuerpos de referencia adecuados para la cuantificación de Rep de AAV y los niveles de proteína de la cápside son, por ejemplo, anti-AAV-rep, monoclonal de ratón, clon 303.9, o anti-AAV VP1/VP2/VP3, monoclonal de ratón, clon B1 que se pueden obtener de PROGEN Biotechnik GmbH.

[0037] En el método de la invención, los casetes de expresión primero y segundo están presentes en un solo constructo de ácidos nucleicos. Una de las ventajas de que ambos casetes de expresión estén presentes en un único constructo de ácidos nucleicos es que las células de insecto transfectadas expresarán tanto la proteína Rep como la proteína de la cápside. Solo la expresión tanto de la proteína Rep como de la proteína de la cápside en una célula de insecto dará como resultado la formación de viriones parvovirales. Otra ventaja es que la proporción mínima requerida de expresión de la proteína Rep frente a la proteína de la cápside se puede controlar cuando los casetes de expresión primero y segundo están en un solo constructo.

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

[0039] En un método de la invención, la secuencia de nucleótidos del primer casete de expresión (ii) puede comprender preferiblemente solo un marco de lectura abierto que comprende secuencias de nucleótidos que

codifican al menos la proteína Rep78. Es decir, la proteína o proteínas Rep estarán codificadas por un solo marco de lectura abierto, no por dos o más marcos de lectura abiertos. Dicho de otra manera, el primer casete de expresión típicamente no tendrá dos o más marcos de lectura abiertos separados que codifiquen las mismas o diferentes proteínas Rep. Sin embargo, para evitar dudas, un único marco de lectura abierto puede ser capaz de codificar más de una proteína, por ejemplo, dos, tres o cuatro proteínas.

[0040] Todo lo anterior puede aplicarse por igual a las proteínas VP1, VP2 y VP3, es decir, dos o todas esas proteínas pueden estar codificadas convenientemente por un único marco de lectura abierto.

[0041] Normalmente, en un método de la invención, al menos un marco de lectura abierto que comprende secuencias de nucleótidos que codifican las proteínas de la cápside VP1, VP2 y VP3 o al menos un marco de lectura abierto que comprende un marco de lectura abierto que comprende secuencias de nucleótidos que codifican al menos la proteína Rep78 no comprende un intrón artificial (o una secuencia derivada de un intrón artificial). Es decir, al menos los marcos de lectura abiertos usados para codificar proteínas Rep o VP no comprenderán un intrón artificial. Por intrón artificial se entiende un intrón que no se encontraría de forma natural en una secuencia Rep o Cap de un virus adenoasociado, por ejemplo, un intrón que se ha diseñado para permitir el empalme funcional dentro de una célula de insecto. Por lo tanto, un intrón artificial en este contexto abarca intrones de células de insectos de tipo salvaje. Un casete de expresión de la invención puede comprender una secuencia de intrones truncada nativa (por nativa se entiende una secuencia que ocurre de manera natural en un virus adenoasociado) - tales secuencias no pretenden caer dentro del significado de intrón artificial tal como se define en el presente documento.

[0042] En la invención, una posibilidad es que ningún marco de lectura abierto que comprenda secuencias de nucleótidos que codifiquen las proteínas de la cápside VP1, VP2 y VP3 y/o ningún marco de lectura abierto que comprenda secuencias de nucleótidos que codifiquen al menos una de las proteínas Rep78 y Rep68 comprenda un intrón artificial.

[0043] En una realización preferida de la invención, la secuencia de nucleótidos que comprende el transgén también está en el mismo constructo de ácidos nucleicos que el primer y el segundo casete de expresión. Una ventaja de esto es que todas las células transfectadas comprenden cada una de las tres secuencias de nucleótidos que se necesitan para la producción de viriones parvovirales.

[0044] Además, los presentes inventores descubrieron que cuanto mayor es la proporción de proteína Rep:proteína de la cápside, mayor es la proporción de virión lleno frente a virión vacío. El término "virión lleno" 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 realización preferida de la invención, la proporción de virión lleno frente a virión vacío es de al menos 1:100, más preferiblemente de al menos 1:10 e incluso más preferiblemente de al menos 1:1. Aún más preferiblemente, no se pueden detectar viriones vacíos y lo más preferiblemente no hay viriones vacíos presentes. El experto en la materia sabrá determinar la relación virión lleno frente a virión vacío, por ejemplo dividiendo el número de copias del gen entre (cápside total - número de copias del genoma), ya que por virión habrá una sola copia del genoma presente. A la inversa, cuanto mayor sea la proporción de proteína Rep:proteína de la cápside, menor será la proporción de viriones vacíos frente a viriones totales (es decir, viriones llenos + vacíos). El experto en la materia sabrá cómo determinar dichas proporciones. Por ejemplo, la proporción de viriones vacíos frente a cápsides totales se puede determinar dividiendo la cantidad de copias del genoma (es decir, el número de copias del genoma) por la cantidad de partículas parvovirales totales (es decir, el número de partículas parvovirales), donde la cantidad de copias del genoma por ml se mide mediante PCR cuantitativa y la cantidad de partículas parvovirales totales por ml se mide con un inmunoensayo enzimático, p. ej., de Progen. En una realización preferida de la invención, la proporción de viriones vacíos frente a viriones totales es inferior a 100:1, más preferentemente inferior a 10:1, e incluso más preferentemente inferior a 1:1. Aún más preferiblemente, no se pueden detectar viriones vacíos y lo más preferiblemente no hay viriones vacíos presentes.

[0045] En una realización preferida, la proporción de expresión de Rep frente a la proteína de la cápside se regula adicionalmente mediante uno o más de los siguientes: (a) la presencia de elementos potenciadores en mayor cantidad y/o más fuertes en el primer casete de expresión en comparación con el segundo casete de expresión; (b) la secuencia de nucleótidos que codifica la proteína Rep parvoviral tiene un índice de adaptación de codones más alto en comparación con la secuencia de nucleótidos que codifica la proteína de la cápside; (c) la optimización de la temperatura de la proteína Rep parvoviral; y/o (d) las proteínas Rep variantes con una o varias alteraciones en la secuencia de aminoácidos en comparación con una proteína Rep de tipo salvaje correspondiente y donde la alteración de uno o más aminoácidos da como resultado aumentos en la actividad de la función Rep según lo evaluado mediante la detección de una mayor producción de AAV en células de insecto. Los métodos para la generación, selección y/o cribado de proteínas Rep variantes con una actividad incrementada de la función Rep según lo evaluado mediante la detección de una producción incrementada de AAV en células de insecto pueden obtenerse mediante la adaptación a células de insecto de los métodos descritos en EE. UU. 20030134351 para la obtención de proteínas Rep variantes con función aumentada con respecto a la producción de AAV en células de mamífero. En el presente documento, se entiende que las

proteínas Rep variantes con una o varias alteraciones en la secuencia de aminoácidos en comparación con una proteína Rep de tipo salvaje correspondiente incluyen proteínas Rep con una o varias sustituciones, inserciones y/o deleciones de aminoácidos en la secuencia de aminoácidos variante en comparación con la secuencia de aminoácidos de una proteína Rep de tipo salvaje correspondiente.

[0046] Que el primer promotor sea igual o más fuerte que el segundo promotor, según lo determinado mediante la expresión del gen informador (por ejemplo, luciferasa o SEAP), o mediante transferencia Northern, significa que en el caso de que haya más secuencias de nucleótidos que codifiquen una proteína Rep que secuencias de nucleótidos que codifiquen una proteína de la cápside, se puede usar un promotor igualmente fuerte, ya que la expresión de la proteína Rep aumentará en comparación con la expresión de la proteína de la cápside, mientras que en el caso de cantidades similares de secuencias de nucleótidos que codifiquen Rep y codifiquen la proteína de la cápside, se puede usar un promotor más fuerte para la expresión de la proteína Rep que para la expresión de la proteína de la cápside. La fuerza del promotor puede determinarse mediante la expresión que se obtiene en las condiciones que se usan en el método de la invención. En una realización preferida, el primer promotor o el segundo promotor se seleccionan del grupo que consiste en un promotor PolH, un promotor p10, un promotor de proteína básica, un promotor inducible o un promotor deltaE1 o un promotor E1, o cualquier otro promotor génico de baculovirus tardío o muy tardío. Más preferiblemente, el primer promotor se selecciona del grupo que consiste en un promotor PolH, un promotor p10 o un promotor de proteína básica y donde el segundo promotor es un promotor deltaE1 o un promotor E1, o cualquier otro promotor génico 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+Hsp70 mínimo. En otra realización, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor 4xHsp27 EcRE+Hsp70 mínimo y el segundo promotor es un promotor PolH. En otra realización más, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor PolH y el segundo promotor es un promotor p10, deltaE1 o E1. En otra realización más, 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. En otra realización más, 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 E1. En otra realización más, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor PolH y el segundo promotor es un promotor PolH. Lo más preferiblemente, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor PolH y el segundo promotor es un promotor deltaE1.

[0047] Un "elemento potenciador" o un "potenciador" tiene por objeto definir una secuencia que potencia la actividad de un promotor (es decir, aumenta la tasa de transcripción de una secuencia cadena abajo del promotor) que, a diferencia de un promotor, no posee actividad promotora, y que normalmente puede funcionar independientemente de su ubicación con respecto al promotor (es decir, aguas arriba o aguas abajo del promotor). Los elementos potenciadores son bien conocidos en la técnica. Los ejemplos no limitantes de elementos potenciadores (o partes de los mismos) que podrían usarse en la presente invención incluyen potenciadores de baculovirus y elementos potenciadores que se encuentran en células de insectos. Se prefiere que el elemento potenciador aumente en una célula la expresión de ARNm de un gen, al que está operativamente enlazado el promotor, en al menos un 25 %, más preferiblemente al menos un 50 %, incluso más preferiblemente al menos un 100 % y lo más preferiblemente al menos un 200 % en comparación con la expresión del ARNm del gen en ausencia del elemento potenciador. La expresión de ARNm se puede determinar, por ejemplo, mediante RT-PCR cuantitativa.

[0048] En el presente documento se prefiere usar un elemento potenciador para potenciar la expresión de la proteína Rep parvoviral. Por lo tanto, en una realización preferida adicional, el primer casete de expresión comprende al menos un elemento potenciador de baculovirus y/o al menos un elemento sensible a la ecdisona. Preferiblemente, el elemento potenciador se selecciona del grupo que consta de hr1, hr2, hr3, hr4 y hr5.

[0049] La optimización de codones de la proteína Rep parvoviral se analiza con más detalle a continuación.

[0050] La optimización de la temperatura de la proteína Rep parvoviral se refiere al uso de la condición óptima con respecto tanto a la temperatura a la que crecerá la célula de insecto como a la que Rep está funcionando. Una proteína Rep puede, por ejemplo, ser óptimamente activa a 37 °C, mientras que una célula de insecto puede crecer de manera óptima a 28 °C. Una temperatura a la que tanto la proteína Rep es activa como a la que la célula de insecto crece puede ser de 30 °C. En una realización preferida, la temperatura optimizada es superior a 27, 28, 29, 30, 31, 32, 33, 34 o 35 °C y/o inferior a 37, 36, 35, 34, 33, 32, 31, 30 o 29 °C.

[0051] Como comprenderá el experto en la técnica, la proporción virión lleno:virión vacío también puede mejorarse mediante la expresión de Cap atenuada, por ejemplo mediante un promotor más débil, en comparación con la expresión de Rep de moderada a alta.

[0052] Preferiblemente, un constructo de ácidos nucleicos de la invención es un vector compatible con células de insecto. Por "vector compatible con células de insectos" o "vector" se entiende una molécula de ácido nucleico capaz de transformar o transfectar productivamente un insecto o una célula de insecto. Los ejemplos de vectores

biológicos incluyen plásmidos, moléculas de ácidos nucleicos lineales y virus recombinantes. Puede emplearse cualquier vector siempre que sea compatible con células de insecto. El vector puede integrarse en el genoma de las células de insecto pero la presencia del vector en la célula de insecto no necesita ser permanente y también se incluyen vectores episomales transitorios. Los vectores pueden introducirse por cualquier medio conocido, por ejemplo, por tratamiento químico de las células, electroporación o infección. En una realización preferida, el vector es un baculovirus, un vector viral o un plásmido. En una realización más preferida, el vector es un baculovirus, es decir, el constructo de ácidos nucleicos es un vector de expresión de baculovirus. Los vectores de expresión de baculovirus y los métodos para su uso se describen, por ejemplo, en Summers y Smith. 1986. A Manual of Methods for Baculovirus Vectors and Insect Culture Procedures, Texas Agricultural Experimental Station Bull. N.º 7555, College Station, Texas; 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, Reino Unido; O'Reilly, D. R., L. K. Miller, V. A. Luckow, 1992, Baculovirus Expression Vectors: A Laboratory Manual, Nueva York; W. H. Freeman y Richardson, C. D., 1995, Baculovirus Expression Protocols, Methods in Molecular Biology, volumen 39; EE. UU. 4.745.051; EE. UU. 2003148506; y WO 03/074714.

[0053] El número de constructos de ácidos nucleicos empleados en la célula de insecto para la producción del vector parvoviral recombinante (rAAV) no es limitante en la invención. Por ejemplo, se pueden emplear uno, dos, tres o más constructos separados para producir rAAV en células de insecto de acuerdo con los métodos de la presente invención. Si se emplean dos constructos, un constructo puede comprender la secuencia de nucleótidos que comprende el transgén que está flanqueado por al menos una secuencia ITR parvoviral y el otro constructo puede comprender un primer y un segundo casete de expresión. Si se emplean tres constructos, un constructo puede comprender la secuencia de nucleótidos que comprende el transgén que está flanqueado por al menos una secuencia ITR parvoviral, otro constructo puede comprender el primer y el segundo casete de expresión y otro constructo puede comprender una secuencia de nucleótidos adicional que codifica un proteína Rep, opcionalmente optimizada ya sea por codón, por AT o por GC, para minimizar o prevenir la recombinación, tal como se describe a continuación.

[0054] Una secuencia de nucleótidos que codifica proteínas Rep parvovirales, se entiende en el presente documento como una secuencia de nucleótidos que codifica las proteínas Rep no estructurales que son necesarias y suficientes para la producción de vectores parvovirales en células de insectos tales como las proteínas Rep78 o Rep68 y/o Rep52 o Rep40. La secuencia de nucleótidos del parvovirus es preferiblemente de un dependovirus, más preferiblemente de un virus adenoasociado (AAV) humano o de simio y lo más preferiblemente de un AAV que normalmente infecta a humanos (p. ej., serotipos 1, 2, 3A, 3B, 4, 5, y 6) o primates (p. ej., serotipos 1 y 4). Un ejemplo de una secuencia de nucleótidos que codifica proteínas Rep de parvovirus se proporciona en la SEC ID N.º 5, que representa una parte del genoma de la secuencia del serotipo 2 de AAV que codifica las proteínas Rep. La secuencia de codificación de Rep78 comprende los nucleótidos 11 - 1876 y la secuencia de codificación de Rep52 comprende los nucleótidos 683 - 1876, también representados por separado en las SEC ID N.º 5 y 7. Se entiende que los pesos moleculares exactos de las proteínas Rep78 y Rep52, así como las posiciones exactas de los codones de iniciación de la traducción pueden diferir entre diferentes parvovirus. Sin embargo, el experto en la materia sabrá cómo identificar la posición correspondiente en la secuencia de nucleótidos de otros parvovirus distintos del AAV-2.

[0055] En una realización preferida, el primer casete de expresión comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos de una proteína Rep78 parvoviral. Preferiblemente, la proteína Rep parvoviral es una proteína Rep de virus adenoasociados (AAV).

[0056] En una 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 único marco de lectura abierto que codifica una proteína Rep78 parvoviral. Preferiblemente, el único marco de lectura abierto codifica la proteína Rep 78 de longitud completa a partir de la cual preferiblemente al menos las proteínas Rep 52 y Rep 78 pueden expresarse en la célula de insecto. Se entiende en el presente documento que la célula de insecto puede comprender más de una copia del único tipo de secuencia de nucleótidos, p. ej., en un vector episomal multicopia, pero que estas son copias múltiples de esencialmente una única molécula de ácido nucleico, o al menos moléculas de ácido nucleico que codifican una única secuencia de aminoácidos de Rep, p. ej., moléculas de ácidos nucleicos que solo difieren entre sí debido a la degeneración del código genético. La presencia de un solo tipo de molécula de ácido nucleico que codifica las proteínas Rep parvovirales evita la recombinación entre secuencias homólogas que pueden estar presentes en diferentes tipos de vectores que comprenden secuencias Rep, lo que puede dar lugar a constructos de expresión Rep defectuosos que afectan a (la estabilidad de) los niveles de producción parvovirales en células de insecto.

[0057] En una 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 comprenda secuencias de nucleótidos que codifiquen al menos 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 pueden estar optimizadas por codón,

optimizadas por AT u optimizadas por GC, para minimizar o evitar 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 la primera y la segunda secuencia de nucleótidos que codifican las secuencias de aminoácidos comunes de una proteína Rep parvoviral se maximiza (es decir, se minimiza la identidad de nucleótidos) mediante una o más de las siguientes opciones: a) cambiando el sesgo del uso de codones de la primera secuencia de nucleótidos que codifica la secuencia de aminoácidos comunes de Rep parvoviral; b) cambiando el sesgo del uso de codones de la segunda secuencia de nucleótidos que codifica la secuencia de aminoácidos común de Rep parvoviral; c) cambiando el contenido de GC de la primera secuencia de nucleótidos que codifica la secuencia de aminoácidos común; y d) cambiando el contenido de GC de la segunda secuencia de nucleótidos que codifica la secuencia de aminoácidos común. La optimización de codones se puede realizar sobre la base del uso de codones de la célula de insecto usada en el método de la invención, preferiblemente *Spodoptera frugiperda*, tal como puede verse en una base de datos de uso de codones (véase, por ejemplo, http://www.kazusa.or.jp/codon/). Los programas informáticos adecuados para la optimización de codones están disponibles para los expertos (véase, p. ej., Jayaraj et al., 2005, Nucl. Acids Res. 33(9):3011-3016; y en internet). Alternativamente, las optimizaciones se pueden hacer a mano, utilizando la misma base de datos de uso de codones.

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

a) que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene al menos un 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 N.º 6
b) que tiene al menos un 50, 60, 70, 80, 81, 82, 85, 90, 95, 97, 98 o 99 % de identidad de secuencia con la secuencia de nucleótidos de cualquiera de las SEC ID N.º 1 - 5;
c) cuya cadena complementaria se hibrida con una secuencia de moléculas de ácidos nucleicos de (a) o (b);
d) secuencias de nucleótidos cuya secuencia difiere de la secuencia de una molécula de ácidos nucleicos 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 puede definirse como una secuencia de nucleótidos:

a) que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene al menos un 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 N.º 8
b) que tiene al menos un 50, 60, 70, 80, 81, 82, 85, 90, 95, 97, 98 o 99 % de identidad de secuencia con la secuencia de nucleótidos de las posiciones 11 - 1876 de la SEC ID N.º 7
c) cuya cadena complementaria se hibrida con una secuencia de moléculas de ácidos nucleicos de (a) o (b);
d) secuencias de nucleótidos cuya secuencia difiere de la secuencia de una molécula de ácidos nucleicos de (c) debido a la degeneración del código genético.
1.

Preferiblemente, la secuencia de nucleótidos codifica proteínas Rep de parvovirus que son necesarias y suficientes para la producción de vectores parvovirales en células de insectos.

[0060] La eliminación de posibles sitios de iniciación de la traducción falsos en las secuencias codificantes de la proteína Rep, distintos de los sitios de iniciación de la traducción Rep78 y Rep52 de otros parvovirus será bien comprendida por un experto en la técnica, al igual que la eliminación de los sitios de empalme putativos que puede reconocerse en células de insecto.

[0061] En una realización preferida, el codón de iniciación para la traducción de la proteína Rep78 parvoviral es un codón de iniciación subóptimo. El codón de iniciación subóptimo preferiblemente es un codón de iniciación que efectúa la omisión parcial de exón. La omisión parcial de exón se entiende en el presente documento como que al menos una parte de los ribosomas no inician la traducción en el codón de iniciación subóptimo de la proteína Rep78 sino en un codón de iniciación aguas abajo, siendo preferiblemente el codón de iniciación aguas abajo el codón de iniciación de la proteína Rep52. El codón de iniciación subóptimo realiza preferentemente una omisión parcial de exón tras la expresión de la secuencia de nucleótidos en una célula de insecto. Preferiblemente, el codón de iniciación subóptimo efectúa la omisión parcial de exón en una célula de insecto para producir en la célula de insecto una relación molar de Rep78 con respecto a Rep52 en el rango de 1:10 a 10:1, 1:5 a 5:1, o 1:3 a 3:1, preferiblemente aproximadamente 20 - 40 horas después de la infección, más preferiblemente aproximadamente 30 - 40 horas después de la infección, usando una expresión de baculovirus. La relación molar de Rep78 y Rep52 puede determinarse mediante transferencia Western, preferiblemente usando un anticuerpo monoclonal que reconozca un epítopo común de Rep78 y de Rep52, o usando p. ej., un anticuerpo anti-Rep de ratón (303.9, Progen, Alemania; dilución 1:50).

[0062] El término "codón de iniciación subóptimo" en el presente documento no solo se refiere al propio codón de

iniciación de trinucleótido sino también a su contexto. Por lo tanto, un codón de iniciación subóptimo puede consistir en un codón ATG "óptimo" en un contexto subóptimo, p. ej., un contexto no Kozak. Sin embargo, son más preferidos los codones de iniciación subóptimos en los que el propio codón de iniciación de trinucleótido es subóptimo, es decir, no es ATG. Se entiende en el presente documento que subóptimo significa que el codón es menos eficaz en la iniciación de la traducción en un contexto por lo demás idéntico en comparación con el codón ATG normal. Preferiblemente, la eficacia del codón subóptimo es inferior al 90, 80, 60, 40 o 20 % de la eficacia del codón ATG normal en un contexto por lo demás idéntico. Los métodos para comparar la eficacia relativa de la iniciación de la traducción son conocidos per se por el experto en la materia. Los codones de iniciación subóptimos preferidos pueden seleccionarse entre ACG, TTG, CTG y GTG. El más preferido es ACG. Una secuencia de nucleótidos que codifica proteínas Rep de parvovirus, se entiende en el presente documento como una secuencia de nucleótidos que codifica proteínas Rep no estructurales que son necesarias y suficientes para la producción de vectores parvovirales en células de insectos tales como las proteínas Rep78 y Rep52.

[0063] También es posible reemplazar otros ATG en la secuencia con un codón diferente que codifique metionina de modo que haya menos posibilidad de iniciar la traducción desde dichos ATG.

[0064] Se logran diversas modificaciones de la secuencia codificante de nucleótidos tal como se ha definido anteriormente, que incluyen, p. ej., las secuencias de parvovirus de tipo salvaje, para la adecuada expresión en células de insecto mediante la aplicación de técnicas de ingeniería genética bien conocidas, tales como las descritas, p. ej., en Sambrook y Russell (2001) "Molecular Cloning: A Laboratory Manual" (3.ª edición), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Nueva York. El experto en la materia conoce varias modificaciones adicionales de las regiones codificantes que podrían aumentar el rendimiento de las proteínas codificantes. Estas modificaciones están dentro del alcance de la presente invención.

[0065] Una secuencia de nucleótidos que codifica una proteína de la cápside parvoviral (Cap) se entiende en el presente documento que comprende secuencias de nucleótidos que codifican una o varias de las tres proteínas de la cápside parvoviral, VP1, -2 y -3. La secuencia de nucleótidos del parvovirus es preferiblemente de un dependovirus, más preferiblemente de un virus adenoasociado (AAV) humano o de simio y lo más preferiblemente de un AAV que normalmente infecta a humanos (p. ej., serotipos 1, 2, 3A, 3B, 4, 5, y 6) o primates (p. ej., serotipos 1 y 4). Se proporcionan ejemplos de una secuencia de nucleótidos que codifica proteínas de la cápside de parvovirus en las SEC ID N.º 20, 22 y 24.

[0066] En una realización preferida, la secuencia de nucleótidos de (iii) comprende un marco de lectura abierto que comprende secuencias de nucleótidos que codifican las proteínas de la cápside VP1, VP2 y VP3. Las secuencias codificantes de la proteína de la cápside pueden estar presentes en varias formas. P. ej., pueden usarse secuencias de codificación separadas para cada una de las proteínas de la cápside VP1, -2 y -3, por lo que cada secuencia codificante está operativamente enlazada a secuencias de control de la 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 único marco de lectura abierto que codifica las tres proteínas de la cápside parvoviral (AAV) VP1, VP2 y VP3, donde el codón de iniciación para la traducción de la proteína de la cápside VP1 es un codón de iniciación subóptimo que no es ATG tal como se describe, p. ej., en Urabe et al. (2002, *supra)* y en WO2007/046703. Un codón de iniciación subóptimo para la proteína de la cápside VP1 puede ser como se ha definido anteriormente para la proteína Rep78. Los codones de iniciación subóptimos de mayor preferencia para la proteína de la cápside de VP1 se pueden seleccionar entre ACG, TTG, CTG y GTG, de los cuales CTG y GTG son los preferidos. La secuencia de nucleótidos comprendida en el segundo casete de expresión para la expresión de las proteínas de la cápside puede comprender además una o varias modificaciones tal como se describe en WO2007/046703. El experto en la materia conoce varias modificaciones adicionales de las regiones codificantes de VP que podrían aumentar el rendimiento de VP y el virión o tener otros efectos deseados, como alterar el tropismo o reducir la antigenicidad del virión. Estas modificaciones están dentro del alcance de la presente invención.

[0067] En una realización preferida, la expresión de VP1 aumenta en comparación con la expresión de VP2 y VP3. La expresión de VP1 puede incrementarse mediante la suplementación de VP1, introduciendo en la célula de insecto un solo vector que comprende secuencias de nucleótidos para VP1 tal como se ha descrito en WO 2007/084773.

[0068] En el contexto de la invención, "al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral" se entiende como una secuencia palindrómica, que comprende secuencias dispuestas simétricamente, en su mayoría complementarias, también denominadas regiones "A", "B" y "C". La ITR funciona como un origen de replicación, un sitio que tiene un papel "cis" en la replicación, es decir, es un sitio de reconocimiento para las proteínas de replicación que actúan en trans, como p. ej., 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 región "D" de la ITR. Es única (no tiene un complemento dentro de una ITR). El corte del ADN monocatenario se produce en la unión entre las regiones A y D. Es la región donde se inicia la síntesis de ADN nuevo. La región D normalmente se encuentra a un lado del palíndromo y proporciona direccionalidad al paso de replicación del ácido nucleico. Un parvovirus que se replica en una célula de mamífero normalmente

tiene dos secuencias ITR. Sin embargo, es posible diseñar una ITR de modo que los sitios de unión en ambas cadenas de las regiones A y las regiones D estén ubicadas simétricamente, una a cada lado del palíndromo. En una plantilla de ADN circular bicatenario (p. ej., un plásmido), la replicación del ácido nucleico asistida por Rep78 o Rep68 procede en ambas direcciones y una única ITR es suficiente para la replicación parvoviral de un vector circular. Por lo tanto, se puede usar una secuencia de nucleótidos de ITR en el contexto de la presente invención. Preferiblemente, sin embargo, se utilizan dos u otro número par de ITR regulares. Más preferiblemente, se usan dos secuencias ITR. Una ITR parvoviral preferida es una ITR de AAV. Por razones de seguridad, puede ser deseable construir un vector parvoviral recombinante (rAAV) que no pueda seguir propagándose después de la introducción inicial en una célula en presencia de un segundo AAV. Dicho mecanismo de seguridad para limitar la propagación de vectores no deseados en un receptor puede proporcionarse mediante el uso de rAAV con una ITR quimérica como se describe en US2003148506.

[0069] El término "flanqueado" con respecto a una secuencia que está flanqueada por otro(s) elemento(s) en el presente documento indica la presencia de uno o varios de los elementos flanqueantes aguas arriba y/o aguas abajo, es decir, 5' y/o 3', en relación con la secuencia. El término "flanqueado" no pretende indicar que las secuencias son necesariamente contiguas. Por ejemplo, puede haber secuencias intermedias entre el ácido nucleico que codifica el transgén y un elemento flanqueante. Una secuencia que está "flanqueada" por otros dos elementos (por ejemplo, secuencias ITR), indica que un elemento está ubicado en 5' respecto de la secuencia y el otro está ubicado en 3' respecto de la secuencia; sin embargo, puede haber secuencias intermedias entre ellos. En una realización preferida, una secuencia de nucleótidos de (i) está flanqueada a cada lado por secuencias de nucleótidos de repetición terminal invertida parvoviral.

[0070] En las realizaciones de la invención, la secuencia de nucleótidos que comprende el transgén (que codifica un producto génico de interés) que está flanqueada por al menos una secuencia ITR parvoviral, preferiblemente se incorpora al genoma de un vector parvoviral recombinante (rAAV) producido en la célula de insecto. . Preferiblemente, el transgén codifica un producto génico de interés para la expresión en una célula de mamífero. Preferiblemente, la secuencia de nucleótidos que comprende el transgén está flanqueada por dos secuencias de nucleótidos ITR parvovirales (AAV) y en donde el transgén está ubicado entre las dos secuencias de nucleótidos ITR parvovirales (AAV). Preferiblemente, la secuencia de nucleótidos que codifica un producto génico de interés (para la expresión en la célula de mamífero) se incorporará al vector parvoviral recombinante (rAAV) producido en la célula de insecto si está ubicado entre dos ITR regulares, o si está ubicado a ambos lados de una ITR diseñada con dos regiones D.

[0071] Las secuencias de AAV que pueden usarse en la presente invención para la producción de un virión de AAV recombinante en células de insectos pueden derivarse del genoma de cualquier serotipo de AAV. Generalmente, los serotipos de AAV tienen secuencias genómicas de homología significativa a nivel de aminoácidos y ácidos nucleicos, proporcionan un conjunto idéntico de funciones genéticas, producen viriones que son esencialmente física y funcionalmente equivalentes, y se replican y ensamblan mediante mecanismos prácticamente idénticos. Para conocer la secuencia genómica de los diversos serotipos de AAV y una descripción general de las similitudes genómicas, véase, p. ej., número de acceso de GenBank U89790; número de acceso de GenBank J01901; número de acceso de GenBank AF043303; Número de acceso de 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). Los serotipos 1, 2, 3, 4 y 5 de AAV son la fuente preferida de secuencias de nucleótidos de AAV para el uso en el contexto de la presente invención. Preferiblemente, las secuencias ITR de AAV para su uso en el contexto de la presente invención se derivan de AAV1, AAV2 y/o AAV4. Asimismo, las secuencias codificantes de Rep (Rep78/68 y Rep52/40) se derivan preferentemente de AAV1, AAV2 y/o AAV4. Sin embargo, las secuencias que codifican las proteínas de la cápside VP1, VP2 y VP3 para su uso en el contexto de la presente invención pueden tomarse de cualquiera de los 42 serotipos conocidos, más preferiblemente de AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 o AAV9 o partículas similares a AAV recientemente desarrolladas obtenidas, p. ej., mediante técnicas de barajado de cápside y bibliotecas de cápside de AAV, o de ITR recientemente diseñadas, desarrolladas o evolucionadas.

[0072] Las secuencias Rep e ITR de AAV se conservan particularmente entre la mayoría de los serotipos. Las proteínas Rep78 de varios serotipos de AAV son, p. ej., son idénticas en más de un 89 % y la identidad total de la secuencia de nucleótidos a nivel del genoma entre AAV2, AAV3A, AAV3B y AAV6 es de alrededor del 82 % (Bantel-Schaal et al., 1999, J. Virol., 73(2):939-947). Además, se sabe que las secuencias Rep y las ITR de muchos serotipos de AAV complementan con eficiencia de forma cruzada (es decir, sustituyen funcionalmente) las secuencias correspondientes de otros serotipos en la producción de partículas de AAV en células de mamífero. EE. UU. 2003148506 informa de que las secuencias AAV Rep e ITR también complementan con eficiencia de manera cruzada otras secuencias AAV Rep e ITR en células de insectos.

[0073] Se sabe que las proteínas VP de AAV determinan la tropicidad celular del virión de AAV. Las secuencias que codifican la proteína VP están significativamente menos conservadas que las proteínas y los genes Rep entre los diferentes serotipos de AAV. La capacidad de las secuencias Rep e ITR para complementar de forma cruzada las secuencias correspondientes de otros serotipos permite la producción de partículas de rAAV

pseudotipadas que comprenden las proteínas de la cápside de un serotipo (por ejemplo, AAV3) y las secuencias Rep y/o ITR de otro serotipo de AAV (por ejemplo, AAV2). Dichas partículas de rAAV pseudotipadas son parte de la presente invención.

[0074] Las secuencias "AAV" modificadas también se pueden usar en el contexto de la presente invención, p. ej., para la producción de vectores rAAV en células de insectos. Dichas secuencias modificadas incluyen, p. ej., secuencias que tienen al menos aproximadamente un 70 %, al menos aproximadamente un 75 %, al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos alrededor del 90 %, al menos aproximadamente un 95 % o más de identidad de secuencia de nucleótidos y/o aminoácidos (por ejemplo, una secuencia que tiene aproximadamente un 75-99 % de identidad de secuencia de nucleótidos) con una ITR de AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 o AAV9, Rep o VP se pueden usar en lugar de secuencias de tipo salvaje ITR de AAV, Rep o VP.

[0075] Aunque es similar a otros serotipos de AAV en muchos aspectos, AAV5 difiere de otros serotipos de AAV humanos y simios más que otros serotipos humanos y simios conocidos. En vista de ello, la producción de rAAV5 puede diferir de la producción de otros serotipos en células de insecto. Cuando se emplean métodos de la invención para producir rAAV5, se prefiere que uno o varios constructos que comprendan, colectivamente en el caso de más de un constructo, una secuencia de nucleótidos que comprenda una ITR de AAV5, una secuencia de nucleótidos comprende una secuencia codificante de Rep de AAV5 (es decir, una secuencia de nucleótidos comprende una Rep78 de AAV5). Dichas secuencias ITR y Rep pueden modificarse según se desee para obtener una producción eficaz de vectores rAAV5 o rAAV5 seudotipados en células de insecto. Por ejemplo, el codón de inicio de las secuencias de Rep se puede modificar, los sitios de empalme de VP se pueden modificar o eliminar y/o el codón de inicio de VP1 y los nucleótidos cercanos se pueden modificar para mejorar la producción de vectores rAAV5 en la célula de insecto.

[0076] La secuencia de nucleótidos que comprende el transgén como se ha definido anteriormente en el presente documento puede comprender, por lo tanto, una secuencia de nucleótidos que codifica al menos un "producto génico de interés" para la expresión en una célula de mamífero, ubicada de manera que se incorporará en un vector parvoviral recombinante (rAAV) replicado en la célula de insecto. En el contexto de la invención se entiende que una célula de mamífero particularmente preferida en la que se va a expresar el "producto génico de interés" es una célula humana. Puede incorporarse cualquier secuencia de nucleótidos para expresión posterior en una célula de mamífero transfectada con el vector parvoviral recombinante (rAAV) producido de acuerdo con la presente invención. La secuencia de nucleótidos puede, p. ej., codificar una proteína, que puede expresar un agente de ARNi, es decir, una molécula de ARN que es capaz de interferir con el ARN, como p. ej., un ARNhc (ARN de horquilla corta) o un ARNpi (ARN pequeño de interferencia). "ARNpi" significa ARN pequeño de interferencia que es un ARN de doble cadena de longitud corta que no es tóxico en las células de mamíferos (Elbashir et al., 2001, Nature 411: 494-98; Caplen et al., 2001, Proc. Natl. Acad. Sci. EE. UU. 98: 9742-47). En una realización preferida, la secuencia de nucleótidos que comprende el transgén puede comprender dos secuencias de nucleótidos codificantes, cada una de las cuales codifica un producto génico de interés para la expresión en una célula de mamífero. Cada una de las dos secuencias de nucleótidos que codifican un producto de interés se ubica de manera que se incorporará en un vector parvoviral recombinante (rAAV) replicado en la célula de insecto.

[0077] El producto de interés para la expresión en una célula de mamífero puede ser un producto génico terapéutico. Un producto génico terapéutico puede ser un polipéptido, una molécula de ARN (ARNpi) u otro producto génico que, cuando se expresa en una célula diana, proporciona un efecto terapéutico deseado, como p. ej., ablación de una actividad no deseada, p. ej., la ablación de una célula infectada, o la complementación de un defecto genético, p. ej., causando una deficiencia en una actividad enzimática. Los ejemplos de productos génicos de polipéptidos terapéuticos incluyen CFTR, Factor IX, lipoproteína lipasa (LPL, preferiblemente LPL S447X; véase WO 01/00220), apolipoproteína A1, uridina difosfato glucuronosiltransferasa (UGT), proteína que interactúa con el regulador de la GTPasa de la retinosis pigmentaria (RP-GRIP) y citoquinas o interleucinas como p. IL-10, porfobilinógeno desaminasa (PBGD), un factor neurotrófico como el factor neurotrófico derivado de la línea de células gliales (GDNF) y alanina:glioxilato aminotransferasa (AGT).

[0078] Alternativamente, o adicionalmente como otro producto génico, la secuencia de nucleótidos que comprende el transgén como se ha definido en el presente documento anteriormente puede comprender además una secuencia de nucleótidos que codifica un polipéptido que sirve como proteína marcadora para evaluar la transformación y la expresión celulares. Proteínas marcadoras adecuadas para este fin son, p. ej., la proteína fluorescente GFP y los genes marcadores seleccionables HSV timidina quinasa (para selección en medio HAT), la higromicina B fosfotransferasa bacteriana (para selección en higromicina B), la aminoglucósido fosfotransferasa Tn5 (para selección en G418) y la dihidrofolato reductasa (DHFR) (para selección en metotrexato), CD20, el gen del factor de crecimiento nervioso de baja afinidad. Las fuentes para obtener estos genes marcadores y los métodos para su uso se proporcionan en Sambrook y Russel, *supra*. Además, la secuencia de nucleótidos que comprende el transgén como se ha definido anteriormente en el presente documento puede comprender una secuencia de nucleótidos adicional que codifica un polipéptido que puede servir como mecanismo a prueba de fallos que permite curar a un sujeto a partir de células transducidas con el

vector parvoviral recombinante (rAAV) de la invención, si se considera necesario. Dicha secuencia de nucleótidos, a menudo denominada gen suicida, codifica una proteína que es capaz de convertir un profármaco en una sustancia tóxica que es capaz de destruir las células transgénicas en las que se expresa la proteína. Los ejemplos adecuados de dichos genes suicidas incluyen p. ej., el gen de la citosina desaminasa de *E. coli* o uno de los genes de la timidina quinasa del virus del herpes simple, el citomegalovirus y el virus de la varicela-zoster, en cuyo caso se puede usar ganciclovir como profármaco para matar las células transgénicas en el sujeto (ver p. Clair et al., 1987, Antimicrob. Agents Chemother. 31: 844-849).

[0079] En otra realización, uno de los productos génicos de interés puede ser una proteína AAV. En particular, una proteína Rep, por ejemplo, Rep78 o Rep68, o un fragmento funcional de las mismas. Una secuencia de nucleótidos que codifica un Rep78 y/o un Rep68, si está presente en el genoma de un vector parvoviral recombinante (rAAV) de la invención y se expresa en una célula de mamífero transducida con el vector, permite la integración del vector parvoviral recombinante (rAAV) en el genoma de la célula de mamífero transducida. La expresión de Rep78 y/o Rep68 en una célula de mamífero transducida o infectada con rAAV puede proporcionar una ventaja para ciertos usos del vector parvoviral recombinante (rAAV), al permitir 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 de parvovirus recombinantes (rAAV) de la invención, la secuencia o secuencias de nucleótidos que codifican un producto génico de interés para la expresión en una célula de mamífero, preferiblemente está operativamente enlazada a al menos una secuencia de control de expresión compatible con células de mamífero, por ejemplo, un promotor. Muchos de estos promotores son conocidos en la técnica (ver Sambrook y Russel, 2001, *supra).* Pueden usarse promotores constitutivos que se expresan ampliamente en muchos tipos de células, como el promotor CMV. Sin embargo, se preferirán más los promotores que sean inducibles, específicos de tejido, específicos de tipo celular o específicos de ciclo celular. Por ejemplo, para la expresión específica del hígado, un promotor puede seleccionarse entre un promotor de α1-antitripsina (AAT), un promotor de globulina fijadora de hormona tiroidea, un promotor de albúmina, un promotor de LPS (globulina fijadora de tiroxina), un promotor híbrido ApoCII, un promotor híbrido HCR-hAAT, un promotor AAT combinado con el elemento potenciador del gen de la albúmina de ratón (Ealb) y un promotor de la apolipoproteína E. Otros ejemplos incluyen el promotor E2F para la expresión selectiva de tumores y, en particular, la expresión selectiva de tumores de células neurológicas (Parr et al., 1997, Nat. Med. 3:1145-9) o el promotor de IL-2 para uso en células sanguíneas mononucleares (Hagenbaugh et al., 1997, J Exp Med; 185: 2101-10).

[0081] El AAV es capaz de infectar varias células de mamíferos. Véase, 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 de AAV de fibroblastos sinoviales humanos es significativamente más eficiente que en células murinas similares, Jennings et al., Arthritis Res, 3:1 (2001), y la tropicidad celular de AAV difiere entre los serotipos. Véase, por ejemplo, Davidson et al. (2000, Proc. Natl. Acad. Sci. EE. UU., 97: 3428-3432), que discuten las diferencias entre AAV2, AAV4 y AAV5 con respecto al tropismo de células del SNC de mamíferos y la eficiencia de transducción. En una realización preferida, una célula huésped de la invención es cualquier célula de mamífero que puede ser infectada por un virión parvoviral, por ejemplo, entre otras, una célula muscular, una célula hepática, una célula nerviosa, una célula glial y una célula epitelial. En una realización preferida, una célula huésped de la invención es una célula humana.

[0082] Preferiblemente, en el constructo, la secuencia de nucleótidos que codifica una proteína Rep parvoviral y/o una proteína de la cápside parvoviral está operativamente enlazada a las secuencias de control de la expresión para la expresión en una célula de insecto. Estas secuencias de control de la expresión incluirán al menos un promotor que sea activo en células de insectos. Las técnicas conocidas por los expertos en la técnica para expresar genes foráneos en células huésped de insectos se pueden utilizar para poner en práctica la invención. La metodología para la ingeniería molecular y la expresión de polipéptidos en células de insectos se describe, por ejemplo, en Summers y Smith. 1986. A Manual of Methods for Baculovirus Vectors and Insect Culture Procedures, Texas Agricultural Experimental Station Bull. N.º 7555, College Station, Texas; 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, Reino Unido; O'Reilly, D. R., L. K. Miller, V. A. Luckow, 1992, Baculovirus Expression Vectors: A Laboratory Manual, Nueva York; W. H. Freeman y Richardson, C. D., 1995, Baculovirus Expression Protocols, Methods in Molecular Biology, volumen 39; EE. UU. 4.745.051; EE. UU. 2003148506; y WO 03/074714. Los promotores adecuados para la transcripción de las secuencias de nucleótidos comprendidas en los constructos primero y segundo de la invención incluyen, p. ej., los promotores poliedro (PolH), p10, p35, IE-1 o ΔIE-1 y otros promotores descritos en las referencias anteriores.

[0083] El constructo de ácidos nucleicos que comprende al menos el primer y/o segundo casete de expresión, puede comprender además una secuencia de control de expresión que comprende una secuencia de nueve nucleótidos de la SEC. ID N.º: 9 o una secuencia de nucleótidos sustancialmente homóloga a la SEC. ID N.º: 9, corriente arriba del codón de iniciación de la secuencia de nucleótidos que codifica la proteína Rep parvoviral y/o del codón de iniciación de la secuencia de nucleótidos que codifica la proteína de la cápside VP1 parvoviral. Una secuencia con identidad sustancial con la secuencia de nucleótidos de la SEC. ID N.º: 9 y que ayudará a

aumentar la expresión de la proteína Rep parvoviral es, p. ej., una secuencia que tiene al menos un 60 %, 70 %, 80 % o 90 % de identidad con la secuencia de nueve nucleótidos de la SEC ID N.º: 9.

[0084] La célula de insecto puede ser cualquier célula que sea adecuada para la producción de proteínas heterólogas. Preferiblemente, la célula de insecto permite la replicación de vectores baculovirales y puede mantenerse en cultivo. Más preferiblemente, la célula de insecto también permite la replicación de vectores parvovirales recombinantes, incluidos los vectores rAAV. Por ejemplo, la línea celular utilizada puede ser de las líneas celulares de *Spodoptera frugiperda, Drosophila*, o líneas celulares de mosquito, por ejemplo, líneas celulares derivadas de *Aedes albopictus*. Las células o líneas celulares de insectos preferidas son células de las especies de insectos que son susceptibles a la infección por baculovirus, incluidas, p. ej., 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 *expres*SF+® (EE. UU. 6.103.526; Protein Sciences Corp., CT, EE. UU.). Una célula de insecto preferida según la invención es una célula de insecto para la producción de vectores parvovirales recombinantes.

[0085] El constructo o constructos de ácidos nucleicos del método de la invención pueden integrarse de forma estable en el genoma de la célula de insecto. Un experto normal en la técnica sabe cómo introducir de forma estable una secuencia de nucleótidos en el genoma de un insecto y cómo identificar una célula que tiene dicha secuencia de nucleótidos en el genoma. La incorporación en el genoma puede ayudarse, por ejemplo, del uso de un vector que comprenda secuencias de nucleótidos altamente homólogas a las regiones del genoma del insecto. El uso de secuencias específicas, como los transposones, es otra forma de introducir una secuencia de nucleótidos en un genoma.

[0086] Las condiciones de crecimiento para células de insecto en cultivo y la producción de productos heterólogos en células de insecto en cultivo son bien conocidas en la técnica y se describen, p. ej., en las referencias citadas anteriormente sobre ingeniería molecular de células de insectos (ver también WO2007/046703).

[0087] En una realización preferida del método de la invención, se recupera el virión parvoviral recombinante. La recuperación comprende preferentemente la etapa de purificación por afinidad de (viriones que comprenden) el vector parvoviral recombinante (rAAV) usando un anticuerpo anti-AAV, preferentemente un anticuerpo inmovilizado. El anticuerpo anti-AAV es preferentemente un anticuerpo monoclonal. Un anticuerpo particularmente adecuado es un anticuerpo de camélido de cadena sencilla o un fragmento del mismo como, p. ej., el que se puede obtener de camellos o llamas (ver p. ej., Muyldermans, 2001, Biotechnol. 74: 277-302). El anticuerpo para la purificación por afinidad de rAAV preferiblemente es un anticuerpo que se une específicamente a un epítopo en una proteína de la cápside de AAV, siendo preferiblemente el epítopo un epítopo que está presente en la proteína de la cápside de más de un serotipo de AAV. Por ejemplo, el anticuerpo puede generarse o seleccionarse sobre la base de la unión específica a la cápside de AAV2 pero, al mismo tiempo, también puede unirse específicamente a las cápsides de 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 tal como se ha definido anteriormente en el presente documento. En una realización preferida, el constructo de ácidos nucleicos de la invención comprende un primer 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 ácidos nucleicos de la invención es un promotor p10 y el segundo promotor es un promotor PolH o un promotor 4xHsp27 EcRE+Hsp70 mínimo. Más preferiblemente, el primer promotor está operativamente enlazado con un potenciador, preferiblemente un potenciador HR1. En otra realización, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor 4xHsp27 EcRE+Hsp70 mínimo y el segundo promotor es un promotor PolH. En otra realización más, el primer promotor en el constructo de ácidos nucleicos de la invención es un promotor PolH y el segundo promotor es un promotor p10, deltaE1 o E1. En otra realización más, 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. En otra realización más, 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 E1. En otra realización más, 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 realizaciones adicionales, cualquier otra combinación del primer promotor y el segundo promotor son parte de la invención. El primer promotor puede seleccionarse del grupo que consiste en un promotor PolH, un promotor p10, un promotor de proteína básica, un promotor inducible o un promotor deltaE1 o un promotor E1, o cualquier otro promotor génico de baculovirus tardío o muy tardío. El segundo promotor puede seleccionarse del grupo que consiste en un promotor PolH, un promotor p10, un promotor de proteína básica, un promotor inducible o un promotor deltaE1 o un promotor E1, o cualquier otro promotor génico de baculovirus tardío o muy tardío. Más preferiblemente, el primer promotor se selecciona del grupo que consiste en un promotor PolH, un promotor p10 o un promotor de proteína básica y donde el segundo promotor es un promotor deltaE1 o un promotor E1, o cualquier otro promotor génico de baculovirus temprano o tardío.

[0089] En una realización preferida, el primer casete de expresión comprendido en el constructo de ácidos nucleicos de la invención comprende un elemento potenciador tal como se ha definido anteriormente.

[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 kit 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 transgén que está flanqueado por al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral, cuyo transgén está operativamente enlazado a un promotor capaz de impulsar la expresión del transgén en una célula huésped.

[0092] En una realización preferida, el constructo de ácidos nucleicos (b) comprende una secuencia de nucleótidos que codifica un transgén que está flanqueado por al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral, cuyo transgén está operativamente enlazado a un promotor capaz de impulsar la expresión del transgén en una célula huésped.

[0093] El kit 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 producidos en los métodos de la invención descritos anteriormente. Un "lote de viriones parvovirales" se define en el presente documento como todos los viriones parvovirales que se producen en la misma ronda de producción, opcionalmente por contenedor de células de insecto. En una realización preferida, el lote de viriones parvovirales de la invención comprende una relación virión lleno:virión total como se ha descrito anteriormente y/o una relación virión lleno:vacío como se ha descrito anteriormente.

[0095] En este documento y en sus reivindicaciones, el verbo «comprender» y sus conjugaciones se utilizan en su sentido no limitante para indicar que se incluyen los elementos que siguen a la palabra, pero no se excluyen los elementos que no se mencionan específicamente. Además, la referencia a un elemento mediante los artículos indefinidos "un" o "uno" no excluye la posibilidad de que esté presente más de uno de los elementos, a menos que el contexto claramente requiera que haya uno y solo uno de los elementos. De este modo, los artículos indefinidos «un» o «una» significan generalmente «al menos uno».

[0096] Los siguientes ejemplos ilustran la invención:

Ejemplos

Ejemplo 1

1.1 Materiales y métodos

1.1.1 Generación de baculovirus recombinantes

[0097] Se generaron los siguientes constructos:

1.1.1.1 Construcción de pVD118 (nuevo)

[0098] El pVD118 (nuevo) es un vector de control que comprende 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 los genes Cap bajo el control del promotor poliedro (PolH o pPH). El pVD118 (nuevo) se construyó como se muestra en la figura 1. Antes de que pudiera fabricarse el plásmido final pVD118 (nuevo), se construyeron los predecesores pFastBac Dual Rep78/ACG y pVD118 (lote n.º 1). En resumen, el plásmido REP-ACG/PSC (solicitud de patente WO2007148971; también denominado en el presente documento pVD88) pVD88 fue digerido con SpeI* XbaI y los extremos salientes 5' se hicieron romos. El fragmento de 2057 bp se aisló de un gel de agarosa, se purificó y se ligó en pFastBac Dual linealizado con SmaI (Invitrogen), dando como resultado pFastBac Dual Rep78/ACG. Posteriormente, este plásmido se digirió con BstZ17I y SnaBI y el fragmento p10_Rep78/ACG de 2537 bp se aisló y se ligó en el pVD84 linealizado con BstZ17I, generando pVD118 (lote n.º 1). Sin embargo, la orientación del casete de expresión Rep78/ACG era incorrecta y, por lo tanto, se eliminó realizando una digestión con NheI.* BlpI. Después de hacer romos los salientes 5', el fragmento de vector de 12 431 bp se aisló y purificó del gel. Posteriormente, el fragmento purificado de 2057 bp de REP-ACG/PSC se ligó en el vector y la mezcla de transformación se transformó en células TOP10 químicamente competentes (Invitrogen) y se sembró en placas que contenían ampicilina. Se realizó un análisis de restricciones

con NaeI*SacI en ADN aislado de cultivos miniprep. Los clones correctos dan fragmentos con un tamaño de 2204 bp y 12 292 bp.

1.1.1.2 Construcción de pVD118 (nuevo) + HR1

[0099] El pVD118 (nuevo) + HR1 es el mismo que el pVD118, con la adición de un potenciador hr1 que se encuentra entre los promotores p10 y poliedro y se construyó como se muestra en la figura 2. En resumen, una PCR realizada en el ADN viral de AcMNPV (Protein Sciences Corporation, Meriden, EE. UU.) con los cebadores HR1-Fw 5'-gtatacgtatgacact atcgatgttgac-3' (SEC ID N.º: 10) y HR1-Rv 5'-gtatacgramoatcgattattgctccaatactag-3' (SEC ID N.º: 11) dio como resultado un producto de 904 bp que se clonó en el vector pCRII-blunt-TOPO (Invitrogen). Después de la digestión con BstZ17I, el fragmento de 898 bp se aisló del gel, se purificó y se ligó en el vector pVD118 (nuevo) que se abrió con BstZ17I y se desfosforiló. El control de las digestiones de los clones correctos con SpeI*EcoNI dio como resultado fragmentos de 1269 bp y 14 125 bp.

1.1.1.3 Construcción de pVD84 (+p10 Rep) (=pVD165)

[0100] El pVD165 es un vector de control que comprende un casete de expresión que comprende la secuencia de nucleótidos de Rep78 donde el codón de inicio se ha mutado en ACG bajo el control del promotor p10 y un casete de expresión que comprende la secuencia de nucleótidos de los genes Cap bajo el control del promotor PolH. Los casetes de expresión están en dirección opuesta en el plásmido. El pVD165 se construyó como se muestra en la figura 3. En resumen, se realizó una PCR en pVD118 (nuevo) con los cebadores poliA Fw 5'-agatct gtagtggctatggcagggc-3' (SEC ID N.º: 12) y p10 Rv 5'-agatct ccgggg acggacctttaattcaacccaac-3' (SEC ID N.º: 13) dio como resultado un producto de 2566 bp que se clonó en el vector pCRII-blunt-TOPO (Invitrogen). Después de la digestión con BglII, el fragmento de 2541 bp se aisló del gel, se purificó y se ligó en el vector pVD84 que se abrió con BglII y se desfosforiló. El control de las digestiones de los clones correctos con SpeI*XbaI dio como resultado fragmentos de 1269 bp y 11 810 bp.

1.1.1.4 Construcción de pVD165 + HR1

[0101] El pVD165 + HR1 es el mismo que el pVD165, con la adición de un potenciador hr1 detrás del promotor p10 y se construyó como se muestra en la figura 4. En resumen, una PCR realizada en el ADN viral de AcMNPV con los cebadores HR1-Fw 5'-gtatacgtatgacact atcgatgttgac-3' (SEC ID N.º: 10) y HR1-Rv 5'-gtatacgramoatcgattattgctccaatactag-3' (SEC ID N.º: 11) dio como resultado un producto de 904 bp que se clonó en el vector pCRII-blunt-TOPO (Invitrogen). Después de la digestión con BstZ17I, el fragmento de 898 bp se aisló del gel, se purificó y se ligó en el vector pVD165 que se abrió con SmaI y se desfosforiló. Las digestiones de control de los clones correctos con ClaI dieron como resultado fragmentos de 875 bp, 1184 bp, 4160 bp y 9180 bp.

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

[0102] El pVD165 + 4xEcRE CAP es similar al pVD165, pero comprende un promotor inducible en lugar del promotor poliedro (pPH). El pVD165 + 4xEcRE CAP se construyó como se muestra en la figura 5. En resumen, el promotor inducible, que comprende 4 EcRE consecutivos, un promotor hsp70 mínimo (Poels et al. Insect Biochem Mol Biol 2004, 34(5):451-458) y parte de la secuencia codificante de CAP, se sintetizó y se ligó en pCRII-blunt-TOPO (Invitrogen). Después de la digestión con BstZ17I y EcoNI, el fragmento de 375 bp se aisló del gel, se purificó y se ligó en el vector pVD165 que se abrió con BstZ17I y EcoNI. El control de las digestiones de los clones correctos con PmeI*EcoRV dio como resultado fragmentos de 2554 bp y 12 116 bp.

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

[0103] El pVD165 + 4*EcRE Rep78 es similar al pVD165, pero comprende un promotor inducible en lugar del promotor p10. El pVD165 + 4*EcRE Rep78 se construyó como se muestra en la figura 6. En resumen, el promotor inducible que consiste en 4 EcRE consecutivos y un promotor hsp70 mínimo se sintetizó y se ligó en pCRII-blunt-TOPO (Invitrogen). Después de la digestión con BstZ17I y SpeI, el fragmento de 290 bp se aisló del gel, se purificó y se ligó en el vector pVD165 que se abrió con SmaI y SpeI. El control de las digestiones de los clones correctos con SpeI*BglII dio como resultado fragmentos de 298 bp, 2376 bp y 11 954 bp.

1.1.1.7 Construcción del constructo pVD190 (deltaIE1 Cap + pPolic Rep)

[0104] El DeltaIE1 Cap + pPolh Rep es similar al pVD165, pero comprende un promotor delta IE1 en lugar del promotor pPH y un promotor pPolH en lugar del promotor p10. El DeltaIE1 Cap + pPolh Rep se construyó como se muestra en las figuras 7 y 8. Para ello, se construyó un vector precursor de la siguiente manera. Una PCR realizada en pFBDSLR (Urabe et al. Human gene therapy 2002;13(16):1935-1943) con los cebadores BstZ17I-deltaIE1 Fw: 5'-ggtcc gtatacgacgataacgccgttggtggcg-3' (SEC ID N.º: 14) y AflII-delta IE1 Rv: 5'-cgacttaagacggcgaattctgc agatggc-3' (SEC ID N.º: 15) dio como resultado un producto de 181 bp que se clonó en

el vector pCRII-blunt-TOPO (Invitrogen). Después de la digestión con BstZ17I*AflII, el fragmento de 165 bp se aisló del gel, se purificó y se ligó en el vector pVD165 + 4xEcRE CAP que se abrió con BstZ17I y AflII. El plásmido resultante es pVD165 + deltaIE1 CAP y el control de digestión con BstZ17I*AflII debería dar como resultado fragmentos de 165 bp y 14 374 bp. Posteriormente, el promotor de poliedrina se clonó frente al casete de expresión de Rep en pVD165 + deltaIE1 CAP. A este fin, se realizó una PCR con los cebadores SmaI-pPolh Fw: 5'-tctcccgggagatcatggaga taattaaaatgataac-3' (SEC ID N.º: 16) y SpeI-pPolh Rv: 5'- gttactuargagctcgtcgac-3' (SEC ID N.º: 17) en el pVD88. Esto generó un producto de PCR de 198 bp que se clonó en el vector pCRII-blunt-TOPO (Invitrogen). Después de la digestión con SpeI*SmaII, el fragmento de 188 bp se aisló del gel, se purificó y se ligó en el vector pVD165 + deltaIE1_Cap que se abrió con SpeI y SmaI. Finalmente, esto dio como resultado la construcción 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 lugar del promotor delta-IE1 frente al casete de expresión Cap y se construyó como se muestra en la figura 9. Una PCR realizada en pFastBac Dual con los cebadores BstZ17I-p10 Fw: 5'-agtatacggacctttaattcaac-3' (SEC ID N.º: 18) y AflII-p10 Rv: 5'-cgacttaagagcgggccgctttcgaatc-3' (SEC ID N.º: 19) dio como resultado un producto de 171 bp que se clonó en el vector pCRII-blunt-TOPO (Invitrogen). Después de la digestión con BstZ17I*AflII, el fragmento de 160 bp se aisló del gel, se purificó y se ligó en el constructo delta-IE1-Cap-pPolh-Rep que se abrió con BstZ17I y AflII.

1.1.1.9 Construcción de pVD194 (*rep78*/*CTG (delta ATG))*

[0106] Primero se construyó un plásmido rep con un codón de inicio CTG sin ningún sitio ATG interno sintetizando el gen necesario de acuerdo con la secuencia establecida en la SEC ID N.º: 31. Luego se eliminó el gen rep del plásmido pVD88 usando las enzimas de restricción RsrII y XbaI. A continuación, el gen sintetizado se ligó en el plásmido pVD88 usando los sitios de restricción RsrII y XbaI para dar pVD195. Luego se digirió pVD195 con BglII, dando como resultado fragmentos de 9297 bp y 2231 bp. Los salientes 5' se llenaron luego con Klenow y se digirieron con SexAI. Esto da como resultado fragmentos de 9297 bp, 1372 y 859 bp. A continuación, se aisló el fragmento de 859 bp. El vector necesario se generó digiriendo pVD165 con SpeI para dar un fragmento lineal de 14 501 bp. Los salientes 5' se llenaron con Klenow y luego se digirieron con SexAI para producir fragmentos de 13 600 bp y 901 bp. Se aisló el fragmento de 13 600 bp. El fragmento BglII(Klenow)*SexAI de 859 bp se ligó luego en pVD165 (SpeI(Klenow)*SexAI) para producir pVD194 (ver Fig. 10). Una digestión de control con SexAI*XmaI debería dar como resultado fragmentos de 13 435 bp y 1029 bp.

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 se realizó una PCR utilizando los siguientes cebadores:
La secuencia del cebador directo contiene un sitio BamHI (cebador AMT n.º 169; SEC ID N.º:29)
5'- TTAGGATCCTGTTAAG**GTG**GCTGCCGACGG -3'
La secuencia del cebador inverso contiene un sitio StuI (cebador AMT n.º 158; SEC ID N.º:30)
5'- GTCGTAGGCCTTGTCGTGCTCGAGGGCCGC -3'

sitio de inicio	Cebador directo	Cebador inverso	Plásmido AMT # (Bac-toBac)	*Plásmido AMT # (PSC)*
GTG	*169*	*158*	*pVD63*	*pVD84*

[0108] Se realizó la PCR con los cebadores n.º 169 y n.º 158 y el producto de la PCR (250 bp) se purificó con el kit de purificación de PCR (Qiagen, lote n.º 11879372) y se cortó con las enzimas de restricción BamHI y StuI. El vector de expresión de baculovirus también se digirió con BamHI y StuI. Después de la desfosforilación del vector con SAP (Promega, lote n.º 17501504), tanto el inserto (Cap con sitio de inicio GTG) como el vector se purificaron en gel. Después de ligar el vector y el inserto, se transformaron en células DH5α químicamente competentes (Invitrogen, lote n.º 1241753) y se sembraron en placas LB que contenían ampicilina. Se purificó el ADN de pVD63 (sitio de inicio Cap/GTG) y se examinó su identidad usando el análisis de secuencia mediante BaseClear y restricción.

[0109] Para clonar el gen Cap con el sitio de inicio GTG en el vector de expresión de baculovirus pPSC10 (Protein Sciences), se separó del pVD63 con SmaI y AvrII y se ligó en el vector de expresión desfosforilado que se abrió con EcoRV y XbaI. Posteriormente, la mezcla de ligadura se transformó en células DH10β químicamente competentes (Invitrogen, lote n.º 1268527) y se sembró en placas de LB-ampicilina. Después del aislamiento del ADN en miniprep utilizando el kit de miniprep QIAprep Spin (Qiagen, lote n.° 12180218), se seleccionó un miniprep (clon n.° 13) mediante análisis de restricción con SphI. Con este clon, el ADN (pVD84) se purificó con el kit SNAP midiprep (Invitrogen solution, lote n.° 1256921 y columna, lote n.° 1259367). La identidad

de pVD84 se comprobó a través del análisis de secuencias alrededor del codón de inicio realizado mediante BaseClear y mediante análisis de restricción con BamHI y SphI (SEC ID N.º:28).

1.1.1.11 Producción de baculovirus recombinantes

[0110] Los Bac.VD118 (nuevo), Bac.VD118 (nuevo)+HR1, Bac.VD165, Bac.VD165+HR1, Bac.VD165+4xEcRE CAP, Bac.VD165+4+EcRE Rep78, Bac.VD190 y Bac.VD194 (p0) recombinantes se generaron con el sistema de Protein Sciences (Protein Sciences Corporation, Meriden, EE. UU.). Los baculovirus recombinantes se amplificaron diluyéndolos 1:100 en 2×10^6 células SF^+ por ml. Tres días después de la infección, las células se centrifugaron y se recuperó el sobrenadante que contenía el virus. La amplificación de los siguientes pasajes se realizó de la misma manera.

1.1.2 Producción de rAAV

[0111] Los lotes de rAAV se produjeron de acuerdo con Urabe et al., 2002 (*supra*), pero con la excepción de que se utilizaron dos baculovirus recombinantes en lugar de tres. Un baculovirus albergaba un constructo de expresión bajo el control del promotor de CMV y está flanqueado por ITR de AAV. El otro baculovirus es el baculovirus Rep-Cap que albergaba dos casetes de expresión, uno para el gen de replicación de AAV y otro para la cápside de AAV. La expresión del gen de replicación o cápside bajo el control del promotor inducible se reguló mediante la adición de 0.001-1 uM de Ponasterona A al medio de cultivo. Los diferentes experimentos de producción de rAAV1 se realizaron con stocks de baculovirus Bac.VD43 p5 que contenían el transgén LPL bajo el control del promotor CMV y diferentes pasajes (p3, p4 o p5) de los baculovirus Rep/Cap. En cada experimento se tomó como control la producción estándar de rAAV1 (Bac.VD88:Bac.VD84:Bac.VD43 con una proporción de 5:1:1).

1.1.3 Determinación de particulas AAV llenas/totales

[0112] Para determinar la proporción de cápsides llenas frente a totales, la cantidad de copias del genoma (gc) se divide por la cantidad de partículas AAV totales. La cantidad de gc/ml se midió mediante un ensayo Q-PCR y la cantidad de partículas de AAV totales se determinó con un inmunoensayo enzimático de Progen (ver más adelante). Para la reacción de Q-PCR, se utilizó SYBR Green PCR Master Mix (Applied Biosystems, n.° 4309155) de acuerdo con las instrucciones del fabricante (25 µl de volumen total; programa de PCR: 10 min a 95 °C, 40 ciclos de 15 s a 95 °C y 1 min a 60 °C) utilizando uno de los siguientes conjuntos de cebadores:

pr59	AATGGGCGGTAGGCGTGTA	CMV	SEC ID N.º: 26
pr60	AGGCGATCTGACGGTTCACTAA	CMV	SEC ID N.º: 27

[0113] La relación se compara con la relación obtenida en condiciones de producción estándar, usando una relación de volumen de 5:1:1 de Bac.Rep, Bac.Cap y Bac.ITR.

1.1.4 Análisis de transferencia Western

[0114] Tres días después de la producción de rAAV, las células se lisaron añadiendo tampón de lisis 0.1 V 10 × TRIS (NaCl 1.5 M, TRIS 0.5 M, MgCl 0.01 M, TRITON X-100 al 1 %, pH 8.5, filtro esterilizado) y se incubaron en hielo durante 30 minutos. El ADN y el ARN libres se degradaron mediante incubación con benzonasa a 37 °C durante 15 minutos. El lisado celular se centrifugó (1900 × g; 15 min; 4 °C). Se añadió tampón de muestra NuPAGE LDS (4x, Invitrogen) a una muestra del sobrenadante y se cargó en un gel Bis-Tris al 4-12 % (120 V). Las proteínas se transfirieron a una membrana de PVDF (BioRad) durante 30 minutos, 10 V (transferencia semiseca). La inmunoquímica Western se realizó mediante el bloqueo de la membrana con tampón de bloqueo Superblock-PBS (PIERCE) y la posterior incubación con ratón anti-Rep (303.9, Progen, Alemania; dilución 1:50) y conejo anti-ratón-HRP (DAKO, dilución 1: 500). Las proteínas Rep se visualizaron mediante tinción quimioluminiscente con lumi-light más sustrato de transferencia Western (Roche).

1.1.5 ELISA de partículas de rAAV1 totales

[0115] La cantidad total de partículas de rAAV1 (tp) realizadas en cada producción se determinó con el kit AAV1 Titration ELISA (Progen, Heidelberg, Alemania) y se realizó de acuerdo con el protocolo proporcionado por el fabricante, pero con la excepción de que todas las muestras y controles fueron prediluidos en lisado crudo a granel (CLB). En resumen, el CLB se hizo cosechando células *express*SF+ después de tres días, añadiendo tampón de lisis 10X e incubando durante 1 hora a 28 °C. Después del tratamiento con benzonasa durante 1 hora a 37 °C, el lisado se centrifugó a 1900 g y el sobrenadante se almacenó a 4 °C. Para determinar las partículas de rAAV1 totales en el lisado crudo de cada producción, las muestras se prediluyeron 50 veces en CLB, esta es la dilución inicial. Posteriormente se realizaron diluciones adicionales de 250, 1250 y 6250 veces en el CLB. La

línea estándar también se diluyó en el CLB. Las muestras de la producción con Bac.VD190 solo se diluyeron 50 y 100 veces, debido a los bajos niveles de expresión de las proteínas de la cápside.

1.1.6 Análisis de partículas totales mediante HPLC

[0116] Se inyectan stocks de AAV-1 desconocidos en el sistema de HPLC, lo que da como resultado un pico en el cromatograma. Este pico puede ser integrado por el software Chemstation y representa la cantidad de partículas totales inyectadas. Una reserva concentrada de AAV-1 se tituló por su cantidad de partículas totales por mililitro mediante microscopía electrónica y se le aplicó el método. Usando este estándar como calibrador, se puede calcular la cantidad de partículas totales de los stocks desconocidos. Además, al inyectar una cantidad particular de copias genómicas (que representa aproximadamente la cantidad de partículas llenas) dentro del rango del estándar, se puede estimar la relación de partículas llenas y vacías.

1.2 Resultados

La proporción de partículas de AAV llenas:vacías mejora con el uso de dos sistemas de baculovirus, en particular con una alta expresión de proteína Rep y una expresión moderada de proteína Cap

[0117] Se investigaron en detalle tres constructos, pVD165, pVD190 y pVD194.

[0118] Para determinar la proporción totales/llenas de las partículas rAAV1 producidas con Bac.VD165, se determinó la concentración de partículas totales en los lisados crudos en dos experimentos independientes. Los resultados de estos ELISA y las proporciones totales/llenas correspondientes se muestran en la tabla 1.

Tabla 1: Proporción totales/llenas de rAAV1 producida con el stock de baculovirus Bac.VD165. La concentración de partículas totales se determinó con el ELISA de partículas de AAV1 totales. La concentración de partículas llenas se determinó mediante Q-PCR. Las proporciones totales/llenas solo se pueden comparar con el control producido en el mismo experimento.

Bac.VD165: Bac.VD43		Concentración de partículas totales (tp/ml)	Concentración de partículas llenas (gc/ml)	Proporción totales/llenas (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 relación totales/llenas para las producciones 1:1 es en ambos experimentos 1.6 veces mejor en comparación con el control. Para la producción 5:1 la relación totales/llenas es 2.1 y 1.4 veces mejor en comparación con el control. En conclusión, la relación totales/llenas mejora para las partículas rAAV1 producidas con Bac.VD165.

[0120] En Bac.VD190, el casete de expresión Cap está bajo el control del promotor débil ΔIE1. Esto puede resultar en una expresión de la cápside más baja que en las producciones realizadas con otros baculovirus Rep/Cap o en la situación del control, porque en todas esas condiciones la expresión de Cap es inducida por el promotor de poliedrina fuerte. Para probar esta hipótesis a partir de dos experimentos diferentes realizados con Bac.VD190 se determinó la concentración de partículas totales. Los resultados se muestran en la tabla 2.

Tabla 2: Proporción totales/llenas de rAAV1 producida con el stock de baculovirus Bac.VD190. La concentración de partículas totales se determinó con el ELISA de partículas de AAV1 totales. La concentración de partículas llenas se determinó mediante Q-PCR. Las proporciones totales/llenas solo se pueden comparar con el control producido en el mismo experimento

Bac.VD190: Bac.VD43		Concentración de partículas totales (tp/ml)	Concentración de partículas llenas (gc/ml)	Proporción totales/llenas (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

Bac.VD190: Bac.VD43	Concentración de partículas totales (tp/ml)	Concentración de partículas llenas (gc/ml)	Proporción totales/llenas (tp/gc)
1:1	3.75E+11	1.74E+09	216
5:1	3.44E+11	3.12E+09	110

[0121] Los resultados de la producción # 1 (tabla 2) muestran que las proporciones totales/llenas obtenidas por las producciones 1:1 y 5:1 mejoraron 114 y 8 veces en comparación con el control, respectivamente. En la producción #2, la proporción totales/llenas es 1.4 veces menor o 1.4 veces mayor para la producción 1:1 o 5:1. En conclusión, la relación totales/llenas parece mejorar también para las partículas rAAV1 producidas con Bac.VD190 (como se observó para Bac.VD165).

[0122] Para Bac.VD194, se llevó a cabo un análisis de transferencia Western para determinar la expresión de Rep y Cap. En consecuencia, tres días después de que las células se infectaran con dos proporciones diferentes de baculovirus (es decir, 1:1 y 9:1, con 5:1:1 para el control de tres componentes), se recolectaron los lisados celulares y se sometieron a análisis de transferencia Western. Como se muestra en la figura 11A, la expresión de Cap es baja en la producción con una proporción de baculovirus de 1:1, pero la producción con una proporción de 9:1 es comparable al control de tres componentes. Sin embargo, para las producciones de proporción 1:1 y 9:1, la expresión de rep se reduce drásticamente en comparación con el control de tres componentes. La figura 11B, sin embargo, muestra que la producción de virus fue mayor en las producciones con Bac.VD194. El hecho de que la producción sea mayor, pero la expresión de Cap sea menor indica una proporción totales:llenas más favorable. Una vez más, la conclusión es que el sistema de dos componentes con cantidades equimolares de Rep y Cap, en particular con una expresión de Rep alta y Cap moderada, conduce a una relación totales:llenas más baja (es decir, un mayor porcentaje de partículas llenas).

Ejemplo 2

Materiales y métodos

[0123] Tres matraces agitadores (todos cultivados a 28 °C antes de la infección) se inocularon con el inóculo P5; 4 ml de Bac.VD43, 4 ml de Bac.VD84 y 20 ml de Bac.VD88. Los agitadores se incubaron durante las 72 horas de producción viral a tres temperaturas diferentes (26, 28 y 30 °C).

[0124] El tratamiento de lisis de todos los matraces con agitación se llevó a cabo dentro de una incubadora con agitación utilizando Triton X-100 al 0.9 % (v/v). Después de la adición del tampón de lisis, los agitadores se incubaron durante 30 minutos a una temperatura de 28 °C, se añadió benzonasa (16 µl por 400 ml) y los agitadores se incubaron durante 1 hora a una temperatura de 37 °C. Después del tratamiento con benzonasa, todos los matraces agitadores se colocaron durante la noche a 29 °C para la eliminación viral, y después se almacenaron a 4 °C durante un máximo de 3 días.

[0125] De cada incubador con agitador, se procesaron 200 ml de lisado crudo a granel en una columna de afinidad de 1 ml usando una velocidad de flujo de 1 ml/min. Después del lavado, el producto se eluyó con PBS, pH 3.5.

Resultados

[0126] El recuento de células en el proceso de los matraces agitadores utilizados para el estudio de temperatura se muestra en la tabla 1. Un aumento de la temperatura se correlaciona con una disminución de la concentración celular total y la viabilidad en el momento de la recolección. La viabilidad obtenida a 28°C fue comparable a los resultados obtenidos con el sistema de biorreactor Wave.

Tabla 1: Recuento de células en el proceso de producción de los agitadores

Matraz agitador (lote n.º)	Antes de la infección		*Antes de la adición del tampón de lisis*	
	Conc. celular total (#/ml)	Viabilidad (%)	Conc. celular total (#/ml)	*Viabilidad (%)*
Temperatura de infección del matraz agitador 26°C (Lote n.º 0078)	1.89×10^6	99.5	2.74×10^6	*85.7*
Temperatura de infección del matraz agitador 28°C (Lote n.º 0078)	1.89×10^6	99.5	2.51×10^6	74.9
Temperatura de infección del matraz	*1.89×10^6*	*99.5*	*2.42×10^6*	*61.0*

Matraz agitador (lote n.º)	Antes de la infección		*Antes de la adición del tampón de lisis*	
	Conc. celular total (#/ml)	Viabilidad (%)	Conc. celular total (#/ml)	*Viabilidad (%)*
agitador 30°C (Lote n.º 0078)				

[0127] Los resultados de las pruebas de los lisados crudos a granel y los eluatos de los matraces de agitación se muestran en la tabla 2.

Tabla 2: Proporción totales:llenas de partículas de rAAV producidas a diferentes temperaturas.

Matraz agitador (lote n.º)	Partículas totales (tp/ml) método HPLC	Q-PCR CLB (gc/ml)	Eluato Q-PCR (gc/ml)	Proporción totales/llenas (tp/gc)
Matraz agitador 26 °C (lote n.º 0078)	6.14×10^{12}	1.28×10^{10}	5.52×10^{10}	*111*
Matraz agitador 28 °C (lote n.º 0078)	6.58×10^{12}	2.00×10^{10}	1.43×10^{11}	46
Matraz agitador 30 °C (lote n.º 0078)	*7.24×10^{12}*	*3.94×10^{10}*	*1.73×10^{11}*	42

Comentarios y conclusión:

[0128] El aumento de la temperatura durante la producción de rAAV en células SF^+, aumenta ligeramente la cantidad total de partículas rAAV que se producen, pero aumenta más la cantidad de partículas llenas que se producen, lo que conduce a una disminución de la proporción totales:llenas.

LISTA DE SECUENCIAS

[0129]

<110> uniQure IP B.V.

<120> Optimización de la expresión de las proteínas rep y cap parvovirales en células de insecto

<130> P6019170EP1

<140> PCT/NL2009/050076
<141> 2009-02-19

<150> EP 08151634.6
<151> 2008-02-19

<150> US 61/029,673
<151> 2008-02-19

<160> 31

<170> versión de PatentIn 3.3

<210> 1
<211> 1194
<212> ADN
<213> virus adenoasociado 2

<220>
<221> misc_feature
<222> (1)..(1194)
<223> rep52 de tipo salvaje

<400> 1
atggagctgg tcgggtggct cgtggacaag gggattacct cggagaagca gtggatccag 60

gaggaccagg cctcatacat ctccttcaat gcggcctcca actcgcggtc ccaaatcaag 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 ccgggaagac caacatcgcg 360

gaggccatag cccacactgt gcccttctac gggtgcgtaa actggaccaa tgagaacttt 420

cccttcaacg actgtgtcga caagatggtg atctggtggg aggaggggaa gatgaccgcc 480

aaggtcgtgg agtcggccaa agccattctc ggaggaagca aggtgcgcgt ggaccagaaa 540

tgcaagtcct cggcccagat agacccgact cccgtgatcg tcacctccaa caccaacatg 600

tgcgccgtga ttgacgggaa ctcaacgacc ttcgaacacc agcagccgtt gcaagaccgg 660

atgttcaaat ttgaactcac ccgccgtctg gatcatgact ttgggaaggt caccaagcag 720

gaagtcaaag actttttccg gtgggcaaag gatcacgtgg ttgaggtgga gcatgaattc 780

tacgtcaaaa agggtggagc caagaaaaga cccgccccca gtgacgcaga tataagtgag 840

cccaaacggg tgcgcgagtc agttgcgcag ccatcgacgt cagacgcgga agcttcgatc 900

aactacgcag accgctacca aaacaaatgt tctcgtcacg tgggcatgaa tctgatgctg 960

tttccctgca gacaatgcga gagaatgaat cagaattcaa atatctgctt cactcacgga 1020

cagaaagact gtttagagtg ctttcccgtg tcagaatctc aacccgtttc tgtcgtcaaa 1080

aaggcgtatc agaaactgtg ctacattcat catatcatgg gaaaggtgcc agacgcttgc 1140

actgcctgcg atctggtcaa tgtggatttg gatgactgca tctttgaaca ataa 1194

<210> 2
<211> 1194
<212> ADN
<213> virus adenoasociado 2

<220>
<221> misc_feature
<222> (1)..(1194)
<223> rep52 sf9 (célula de insecto) optimizada

<400> 2
atggagctgg tgggttggct ggtggacaag ggtatcacct ccgagaagca gtggatccag 60

gaggaccagg cttcctacat ctccttcaac gctgcttcca actcccgttc ccagatcaag 120

gctgctctgg acaacgctgg taagatcatg tccctgacca agaccgctcc tgactacctg 180

gtgggtcagc agcctgtgga ggacatctcc tccaaccgta tctacaagat cctggagctg 240

aacggttacg accctcagta cgctgcttcc gtgttcctgg gttgggctac caagaagttc 300

ggtaagcgta acaccatctg gctgttcggt cctgctacca ccggtaagac caacatcgct 360

gaggctatcg ctcacaccgt gcctttctac ggttgcgtga actggaccaa cgagaacttc 420

cctttcaacg actgcgtgga caagatggtg atctggtggg aggagggtaa gatgaccgct 480

aaggtggtgg agtccgctaa ggctatcctg ggtggttcca aggtgcgtgt ggaccagaag 540

tgcaagtcct ccgctcagat cgaccctacc cctgtgatcg tgacctccaa caccaacatg 600

tgcgctgtga tcgacggtaa ctccaccacc ttcgagcacc agcagcctct gcaggaccgt 660

atgttcaagt tcgagctgac ccgtcgtctg gaccacgact tcggtaaggt gaccaagcag 720

gaggtgaagg acttcttccg ttgggctaag gaccacgtgg tggaggtgga gcacgagttc 780

tacgtgaaga agggtggtgc taagaagcgt cctgctcctt ccgacgctga catctccgag 840

cctaagcgtg tgcgtgagtc cgtggctcag ccttccacct ccgacgctga ggcttccatc 900

aactacgctg accgttacca gaacaagtgc tcccgtcacg tgggtatgaa cctgatgctg 960

ttcccttgcc gtcagtgcga gcgtatgaac cagaactcca acatctgctt cacccacggt 1020

cagaaggact gcctggagtg cttccctgtg tccgagtccc agcctgtgtc cgtggtgaag 1080

aaggcttacc agaagctgtg ctacatccac cacatcatgg gtaaggtgcc tgacgcttgc 1140

accgcttgcg acctggtgaa cgtggacctg gacgactgca tcttcgagca gtaa 1194

<210> 3
<211> 1194
<212> ADN
<213> virus adenoasociado 2

<220>
<221> misc_feature
<222> (1)..(1194)
<223> rep52 AT optimizada

<400> 3
atggaattag taggatggtt agtagataaa ggaataacat cagaaaaaca atggatacaa 60

gaagatcaag catcatatat atcatttaat gcagcatcaa attcaagatc acaaataaaa 120

gcagcattag ataatgcagg aaaaataatg tcattaacaa aaacagcacc agattattta 180

gtaggacaac aaccagtaga agatatatca tcaaatagaa tatataaaat attagaatta 240

aatggatatg atccacaata tgcagcatca gtatttttag gatgggcaac aaaaaaattt 300

ggaaaaagaa atacaatatg gttatttgga ccagcaacaa caggaaaaac aaatatagca 360

gaagcaatag cacatacagt accattttat ggatgtgtaa attggacaaa tgaaaatttt 420

ccatttaatg attgtgtaga taaaatggta atatggtggg aagaaggaaa aatgacagca 480

aaagtagtag aatcagcaaa agcaatatta ggaggatcaa aagtaagagt agatcaaaaa 540

tgtaaatcat cagcacaaat agatccaaca ccagtaatag taacatcaaa tacaaatatg 600

tgtgcagtaa tagatggaaa ttcaacaaca tttgaacatc aacaaccatt acaagataga 660

atgtttaaat ttgaattaac aagaagatta gatcatgatt ttggaaaagt aacaaaacaa 720

gaagtaaaag atttttttag atgggcaaaa gatcatgtag tagaagtaga acatgaattt 780

tatgtaaaaa aaggaggagc aaaaaaaaga ccagcaccat cagatgcaga tatatcagaa 840

ccaaaaagag taagagaatc agtagcacaa 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
<213> virus adenoasociado 2

<220>
<221> misc_feature
<222> (1)..(1194)
<223> rep52 GC optimizada

<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 ggtgggcgac gaagaagttc 300

gggaagcgga acacgatctg gctgttcggg ccggcgacga cggggaagac gaacatcgcg 360

gaggcgatcg cgcacacggt gccgttctac gggtgcgtga actggacgaa cgagaacttc 420

ccgttcaacg actgcgtgga caagatggtg atctggtggg aggaggggaa gatgacggcg 480

aaggtggtgg agagcgcgaa ggcgatcctg gggggagca aggtgcgggt ggaccagaag 540

tgcaagagca gcgcgcagat cgacccgacg ccggtgatcg tgacgagcaa cacgaacatg 600

tgcgcggtga tcgacgggaa cagcacgacg ttcgagcacc agcagccgct gcaggaccgg 660

atgttcaagt tcgagctgac gcggcggctg gaccacgact tcgggaaggt gacgaagcag 720

gaggtgaagg acttcttccg gtgggcgaag gaccacgtgg tggaggtgga gcacgagttc 780

tacgtgaaga agggggggc gaagaagcgg ccggcgccga gcgacgcgga catcagcgag 840

ccgaagcggg tgcgggagag cgtggcgcag 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
<213> virus adenoasociado 2

<220>
<221> CDS
<222> (1)..(1194)
<223> Rep52

<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

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

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 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
<213> virus adenoasociado 2

<400> 6

```
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

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
<213> virus adenoasociado 2

<220>
<221> CDS
<222> (11)..(1876)
<223> Rep78

<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

ctc gtg gaa acc acc ggg gtg aaa tcc atg gtt ttg gga cgt ttc ctg 337
Leu Val Glu Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu
95 100 105

agt cag att cgc gaa aaa ctg att cag aga att tac cgc ggg atc gag 385
Ser Gln Ile Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu
110 115 120 125

ccg act ttg cca aac tgg ttc gcg gtc aca aag acc aga aat ggc gcc 433
Pro Thr Leu Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala
130 135 140

```
gga ggc ggg aac aag gtg gtg gat gag tgc tac atc ccc aat tac ttg      481
Gly Gly Gly Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu
            145                 150                 155

ctc ccc aaa acc cag cct gag ctc cag tgg gcg tgg act aat atg gaa      529
Leu Pro Lys Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu
        160                 165                 170

cag tat tta agc gcc tgt ttg aat ctc acg gag cgt aaa cgg ttg gtg      577
Gln Tyr Leu Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val
    175                 180                 185

gcg cag cat ctg acg cac gtg tcg cag acg cag gag cag aac aaa gag      625
Ala Gln His Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu
190                 195                 200                 205

aat cag aat ccc aat tct gat gcg ccg gtg atc aga tca aaa act tca      673
Asn Gln Asn Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser
                210                 215                 220

gcc agg tac atg gag ctg gtc ggg tgg ctc gtg gac aag ggg att acc      721
Ala Arg Tyr Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr
            225                 230                 235

tcg gag aag cag tgg atc cag gag gac cag gcc tca tac atc tcc ttc      769
Ser Glu Lys Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe
        240                 245                 250

aat gcg gcc tcc aac tcg cgg tcc caa atc aag gct gcc ttg gac aat      817
Asn Ala Ala Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn
    255                 260                 265

gcg gga aag att atg agc ctg act aaa acc gcc ccc gac tac ctg gtg      865
Ala Gly Lys Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val
270                 275                 280                 285

ggc cag cag ccc gtg gag gac att tcc agc aat cgg att tat aaa att      913
Gly Gln Gln Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile
                290                 295                 300

ttg gaa cta aac ggg tac gat ccc caa tat gcg gct tcc gtc ttt ctg      961
Leu Glu Leu Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu
            305                 310                 315

gga tgg gcc acg aaa aag ttc ggc aag agg aac acc atc tgg ctg ttt     1009
Gly Trp Ala Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe
        320                 325                 330

ggg cct gca act acc ggg aag acc aac atc gcg gag gcc ata gcc cac     1057
Gly Pro Ala Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His
    335                 340                 345

act gtg ccc ttc tac ggg tgc gta aac tgg acc aat gag aac ttt ccc     1105
Thr Val Pro Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro
350                 355                 360                 365

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
<213> virus adenoasociado 2

<400> 8

```
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

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
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
<213> virus adenoasociado 2

<220>
<221> misc_feature
<222> (1)..(9)
<223> iniciación 7 nt

<400> 9
cctgttaag 9

<210> 10
<211> 28
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador

<220>
<221> misc_feature
<222> (1)..(28)
<223> HR1-FW

<400> 10
gtatacgtat gacactatcg atgttgac 28

<210> 11
<211> 30
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador

<220>
<221> misc_feature
<222> (1)..(30)
<223> HR1-RV

<400> 11
gtatacgatc gattattgct ccaatactag 30

<210> 12
<211> 25
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador

<220>
<221> misc_feature
<222> (1)..(25)
<223> poliA FW

<400> 12
agatctgtag tggctatggc agggc 25

<210> 13
<211> 35
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador

<220>
<221> misc_feature
<222> (1)..(35)
<223> p10 RV

<400> 13
agatctcccg ggacggacct ttaattcaac ccaac 35

<210> 14
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador

<220>
<221> misc_feature
<222> (1)..(33)
<223> BstZ17I-deltaIE1 FW

<400> 14
ggtccgtata cgacgataac gccgttggtg gcg 33

<210> 15
<211> 30
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador

<220>
<221> misc_feature

<222> (1)..(30)
<223> AflII-delta IE1 RV

<400> 15
cgacttaaga cggcgaattc tgcagatggc 30

<210> 16
<211> 37
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador

<220>
<221> misc_feature
<222> (1)..(37)
<223> SmaI-pPolh FW

<400> 16
tctcccggga gatcatggag ataattaaaa tgataac 37

<210> 17
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador

<220>
<221> misc_feature
<222> (1)..(21)
<223> SpeI-pPolh RV

<400> 17
gttactagtg agctcgtcga c 21

<210> 18
<211> 23
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador

<220>
<221> misc_feature
<222> (1)..(23)
<223> BstZ17I-p10 FW

<400> 18
agtatacgga cctttaattc aac 23

<210> 19
<211> 28
<212> ADN
<213> Secuencia artificial

<220>

<223> Secuencia del cebador

<220>
<221> misc_feature
<222> (1)..(28)
<223> AflII-p10 RV

<400> 19
cgacttaaga gcgggccgct ttcgaatc 28

<210> 20
<211> 2211
<212> ADN
<213> Secuencia artificial

<220>
<223> virus adenoasociado 1; VP1, VP2, VP3; codón de iniciación VP1 alterado (GTG)

<220>
<221> CDS
<222> (1)..(2208)
<223> VP1

<220>
<221> misc_feature
<222> (412)..(2208)
<223> VP2

<220>
<221> misc_feature
<222> (607)..(2208)
<223> VP3

<400> 20

```
gtg gct gcc gac ggt tat cta ccc gat tgg ctc gag gac aac ctc tct       48
Val Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1               5                   10                  15

gag ggc att cgc gag tgg tgg gac ttg aaa cct gga gcc ccg aag ccc       96
Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30

aaa gcc aac cag caa aag cag gac gac ggc cgg ggt ctg gtg ctt cct      144
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45

ggc tac aag tac ctc gga ccc ttc aac gga ctc gac aag ggg gag ccc      192
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
    50                  55                  60

gtc aac gcg gcg gac gca gcg gcc ctc gag cac gac aag gcc tac gac      240
Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65                  70                  75                  80

cag cag ctc aaa gcg ggt gac aat ccg tac ctg cgg tat aac cac gcc      288
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
                85                  90                  95

gac gcc gag ttt cag gag cgt ctg caa gaa gat acg tct ttt ggg ggc      336
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110

aac ctc ggg cga gca gtc ttc cag gcc aag aag cgg gtt ctc gaa cct      384
```

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

ctc ggt ctg gtt gag gaa ggc gct aag acg gct cct gga aag aaa cgt      432
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140

ccg gta gag cag tcg cca caa gag cca gac tcc tcc tcg ggc atc ggc      480
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145                 150                 155                 160

aag aca ggc cag cag ccc gct aaa aag aga ctc aat ttt ggt cag act      528
Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175

ggc gac tca gag tca gtc ccc gat cca caa cct ctc gga gaa cct cca      576
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
            180                 185                 190

gca acc ccc gct gct gtg gga cct act aca atg gct tca ggc ggt ggc      624
Ala Thr Pro Ala Ala Val Gly Pro Thr Thr Met Ala Ser Gly Gly Gly
        195                 200                 205

gca cca atg gca gac aat aac gaa ggc gcc gac gga gtg ggt aat gcc      672
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala
    210                 215                 220

tca gga aat tgg cat tgc gat tcc aca tgg ctg ggc gac aga gtc atc      720
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile
225                 230                 235                 240

acc acc agc acc cgc acc tgg gcc ttg ccc acc tac aat aac cac ctc      768
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
                245                 250                 255

tac aag caa atc tcc agt gct tca acg ggg gcc agc aac gac aac cac      816
Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His
            260                 265                 270

tac ttc ggc tac agc acc ccc tgg ggg tat ttt gat ttc aac aga ttc      864
Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe
        275                 280                 285

cac tgc cac ttt tca cca cgt gac tgg cag cga ctc atc aac aac aat      912
His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn
    290                 295                 300

tgg gga ttc cgg ccc aag aga ctc aac ttc aaa ctc ttc aac atc caa      960
Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln
305                 310                 315                 320

gtc aag gag gtc acg acg aat gat ggc gtc aca acc atc gct aat aac     1008
Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn
                325                 330                 335

ctt acc agc acg gtt caa gtc ttc tcg gac tcg gag tac cag ctt ccg     1056
Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro
            340                 345                 350

tac gtc ctc ggc tct gcg cac cag ggc tgc ctc cct ccg ttc ccg gcg     1104
Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
        355                 360                 365

gac gtg ttc atg att ccg caa tac ggc tac ctg acg ctc aac aat ggc     1152
Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
    370                 375                 380
```

```
agc caa gcc gtg gga cgt tca tcc ttt tac tgc ctg gaa tat ttc cct     1200
Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
385                 390                 395                 400

tct cag atg ctg aga acg ggc aac aac ttt acc ttc agc tac acc ttt     1248
Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe
                405                 410                 415

gag gaa gtg cct ttc cac agc agc tac gcg cac agc cag agc ctg gac     1296
Glu Glu Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp
            420                 425                 430

cgg ctg atg aat cct ctc atc gac caa tac ctg tat tac ctg aac aga     1344
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg
        435                 440                 445

act caa aat cag tcc gga agt gcc caa aac aag gac ttg ctg ttt agc     1392
Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu Leu Phe Ser
    450                 455                 460

cgt ggg tct cca gct ggc atg tct gtt cag ccc aaa aac tgg cta cct     1440
Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys Asn Trp Leu Pro
465                 470                 475                 480

gga ccc tgt tat cgg cag cag cgc gtt tct aaa aca aaa aca gac aac     1488
Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Lys Thr Asp Asn
                485                 490                 495

aac aac agc aat ttt acc tgg act ggt gct tca aaa tat aac ctc aat     1536
Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu Asn
            500                 505                 510

ggg cgt gaa tcc atc atc aac cct ggc act gct atg gcc tca cac aaa     1584
Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys
        515                 520                 525

gac gac gaa gac aag ttc ttt ccc atg agc ggt gtc atg att ttt gga     1632
Asp Asp Glu Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe Gly
    530                 535                 540

aaa gag agc gcc gga gct tca aac act gca ttg gac aat gtc atg att     1680
Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met Ile
545                 550                 555                 560

aca gac gaa gag gaa att aaa gcc act aac cct gtg gcc acc gaa aga     1728
Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Arg
                565                 570                 575

ttt ggg acc gtg gca gtc aat ttc cag agc agc agc aca gac cct gcg     1776
Phe Gly Thr Val Ala Val Asn Phe Gln Ser Ser Ser Thr Asp Pro Ala
            580                 585                 590

acc gga gat gtg cat gct atg gga gca tta cct ggc atg gtg tgg caa     1824
Thr Gly Asp Val His Ala Met Gly Ala Leu Pro Gly Met Val Trp Gln
        595                 600                 605

gat aga gac gtg tac ctg cag ggt ccc att tgg gcc aaa att cct cac     1872
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
    610                 615                 620

aca gat gga cac ttt cac ccg tct cct ctt atg ggc ggc ttt gga ctc     1920
Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
625                 630                 635                 640

aag aac ccg cct cct cag atc ctc atc aaa aac acg cct gtt cct gcg     1968
Lys Asn Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
                645                 650                 655
```

```
aat cct ccg gcg gag ttt tca gct aca aag ttt gct tca ttc atc acc     2016
Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala Ser Phe Ile Thr
            660                 665                 670

caa tac tcc aca gga caa gtg agt gtg gaa att gaa tgg gag ctg cag     2064
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
        675                 680                 685

aaa gaa aac agc aag cgc tgg aat ccc gaa gtg cag tac aca tcc aat     2112
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Val Gln Tyr Thr Ser Asn
    690                 695                 700

tat gca aaa tct gcc aac gtt gat ttt act gtg gac aac aat gga ctt     2160
Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr Val Asp Asn Asn Gly Leu
705                 710                 715                 720

tat act gag cct cgc ccc att ggc acc cgt tac ctt acc cgt ccc ctg     2208
Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Pro Leu
                725                 730                 735

taa                                                                 2211
```

<210> 21
<211> 736
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción sintética

<400> 21

```
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

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                 295                 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                 350

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                 375                 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
        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
<213> Secuencia artificial

<220>
<223> virus adenoasociado 1; VP2, VP3

<220>
<221> CDS
<222> (1)..(1797)
<223> VP2

<400> 22

```
acg gct cct gga aag aaa cgt ccg gta gag cag tcg cca caa gag cca       48
Thr Ala Pro Gly Lys Lys Arg Pro Val Glu Gln Ser Pro Gln Glu Pro
1               5                   10                  15

gac tcc tcc tcg ggc atc ggc aag aca ggc cag cag ccc gct aaa aag       96
Asp Ser Ser Ser Gly Ile Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys
            20                  25                  30

aga ctc aat ttt ggt cag act ggc gac tca gag tca gtc ccc gat cca      144
Arg Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu Ser Val Pro Asp Pro
        35                  40                  45

caa cct ctc gga gaa cct cca gca acc ccc gct gct gtg gga cct act      192
Gln Pro Leu Gly Glu Pro Pro Ala Thr Pro Ala Ala Val Gly Pro Thr
    50                  55                  60

aca atg gct tca ggc ggt ggc gca cca atg gca gac aat aac gaa ggc      240
Thr Met Ala Ser Gly Gly Gly Ala Pro Met Ala Asp Asn Asn Glu Gly
65                  70                  75                  80

gcc gac gga gtg ggt aat gcc tca gga aat tgg cat tgc gat tcc aca      288
Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp His Cys Asp Ser Thr
                85                  90                  95

tgg ctg ggc gac aga gtc atc acc acc agc acc cgc acc tgg gcc ttg      336
Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu
            100                 105                 110

ccc acc tac aat aac cac ctc tac aag caa atc tcc agt gct tca acg      384
Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile Ser Ser Ala Ser Thr
        115                 120                 125

ggg gcc agc aac gac aac cac tac ttc ggc tac agc acc ccc tgg ggg      432
Gly Ala Ser Asn Asp Asn His Tyr Phe Gly Tyr Ser Thr Pro Trp Gly
    130                 135                 140
```

```
tat ttt gat ttc aac aga ttc cac tgc cac ttt tca cca cgt gac tgg      480
Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp
145                 150                 155                 160

cag cga ctc atc aac aac aat tgg gga ttc cgg ccc aag aga ctc aac      528
Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn
                165                 170                 175

ttc aaa ctc ttc aac atc caa gtc aag gag gtc acg acg aat gat ggc      576
Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr Thr Asn Asp Gly
            180                 185                 190

gtc aca acc atc gct aat aac ctt acc agc acg gtt caa gtc ttc tcg      624
Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Val Gln Val Phe Ser
        195                 200                 205

gac tcg gag tac cag ctt ccg tac gtc ctc ggc tct gcg cac cag ggc      672
Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly
    210                 215                 220

tgc ctc cct ccg ttc ccg gcg gac gtg ttc atg att ccg caa tac ggc      720
Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly
225                 230                 235                 240

tac ctg acg ctc aac aat ggc agc caa gcc gtg gga cgt tca tcc ttt      768
Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe
                245                 250                 255

tac tgc ctg gaa tat ttc cct tct cag atg ctg aga acg ggc aac aac      816
Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn
            260                 265                 270

ttt acc ttc agc tac acc ttt gag gaa gtg cct ttc cac agc agc tac      864
Phe Thr Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe His Ser Ser Tyr
        275                 280                 285

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 Gln 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

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 Glu 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
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción sintética

<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
        195                 200                 205

Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly
    210                 215                 220

Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly
225                 230                 235                 240

Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe
                245                 250                 255

Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn
            260                 265                 270

Phe Thr Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe His Ser Ser Tyr
        275                 280                 285

Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln
```

```
     290                 295                 300

Tyr Leu Tyr Tyr Leu Asn Arg Thr Gln Asn Gln Ser Gly Ser Ala Gln
305                 310                 315                 320

Asn Lys Asp Leu Leu Phe Ser Arg Gly Ser Pro Ala Gly Met Ser Val
                325                 330                 335

Gln Pro Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val
            340                 345                 350

Ser Lys Thr Lys Thr Asp Asn Asn Asn Ser Asn Phe Thr Trp Thr Gly
        355                 360                 365

Ala Ser Lys Tyr Asn Leu Asn Gly Arg Glu Ser Ile Ile Asn Pro Gly
    370                 375                 380

Thr Ala Met Ala Ser His Lys Asp Asp Glu Asp Lys Phe Phe Pro Met
385                 390                 395                 400

Ser Gly Val Met Ile Phe Gly Lys Glu Ser Ala Gly Ala Ser Asn Thr
                405                 410                 415

Ala Leu Asp Asn Val Met Ile Thr Asp Glu Glu Glu Ile Lys Ala Thr
            420                 425                 430

Asn Pro Val Ala Thr Glu Arg Phe Gly Thr Val Ala Val Asn Phe Gln
        435                 440                 445

Ser Ser Ser Thr Asp Pro Ala Thr Gly Asp Val His Ala Met Gly Ala
    450                 455                 460

Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro
465                 470                 475                 480

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

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
<213> Secuencia artificial

<220>
<223> virus adenoasociado 1; VP3

<220>
<221> CDS
<222> (1)..(1602)
<223> VP3

<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

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 gtt 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 gtt 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 gtt 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
305                 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

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
<213> Secuencia artificial

<220>
<223> Construcción sintética

<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
65                  70                  75                  80

Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln
                85                  90                  95

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

```
Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr Thr Asn Asp Gly Val
        115                 120                 125

Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Val Gln Val Phe Ser Asp
    130                 135                 140

Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys
145                 150                 155                 160

Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr
                165                 170                 175

Leu Thr Leu Asn Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr
            180                 185                 190

Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe
        195                 200                 205

Thr Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe His Ser Ser Tyr Ala
    210                 215                 220

His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr
225                 230                 235                 240

Leu Tyr Tyr Leu Asn Arg Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn
                245                 250                 255

Lys Asp Leu Leu Phe Ser Arg Gly Ser Pro Ala Gly Met Ser Val Gln
            260                 265                 270

Pro Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser
        275                 280                 285

Lys Thr Lys Thr Asp Asn Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala
    290                 295                 300

Ser Lys Tyr Asn Leu Asn Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr
305                 310                 315                 320

Ala Met Ala Ser His Lys Asp Asp Glu Asp Lys Phe Phe Pro Met Ser
                325                 330                 335

Gly Val Met Ile Phe Gly Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala
            340                 345                 350

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> 19
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador

<220>
<221> misc_feature
<222> (1)..(19)
<223> Gen CMV

<400> 26
aatgggcggt aggcgtgta 19

<210> 27
<211> 22
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador

<220>
<221> misc_feature
<222> (1)..(22)
<223> Gen CMV

<400> 27
aggcgatctg acggttcact aa 22

<210> 28
<211> 11954
<212> ADN
<213> Secuencia artificial

<220>
<223> vector pVD84

<400> 28
gggtgatcaa gtcttcgtcg agtgattgta aataaaatgt aatttacagt atagtatttt 60

aattaatata caaatgattt gataataatt cttatttaac tataatatat tgtgttgggt 120

tgaattaaag gtccgtatac tccggaatat taatagatca tggagataat taaaatgata 180

accatctcgc aaataaataa gtattttact gttttcgtaa cagttttgta ataaaaaaac 240

ctataaatat tccggattat tcataccgtc ccaccatcgg gcgcggatcc tgttaaggtg 300

gctgccgacg gttatctacc cgattggctc gaggacaacc tctctgaggg cattcgcgag 360

tggtgggact tgaaacctgg agccccgaag cccaaagcca accagcaaaa gcaggacgac 420

ggccggggtc tggtgcttcc tggctacaag tacctcggac ccttcaacgg actcgacaag 480

ggggagcccg tcaacgcggc ggacgcagcg gccctcgagc acgacaaggc ctacgaccag 540

cagctcaaag cgggtgacaa tccgtacctg cggtataacc acgccgacgc cgagtttcag 600

gagcgtctgc aagaagatac gtcttttggg ggcaacctcg ggcgagcagt cttccaggcc 660

aagaagcggg ttctcgaacc tctcggtctg gttgaggaag gcgctaagac ggctcctgga 720

aagaaacgtc cggtagagca gtcgccacaa gagccagact cctcctcggg catcggcaag 780

acaggccagc agcccgctaa aaagagactc aattttggtc agactggcga ctcagagtca 840

gtccccgatc cacaacctct cggagaacct ccagcaaccc ccgctgctgt gggacctact 900

acaatggctt caggcggtgg cgcaccaatg gcagacaata acgaaggcgc cgacggagtg 960

ggtaatgcct caggaaattg gcattgcgat tccacatggc tgggcgacag agtcatcacc 1020

accagcaccc gcacctgggc cttgcccacc tacaataacc acctctacaa gcaaatctcc 1080

agtgcttcaa cgggggccag caacgacaac cactacttcg gctacagcac cccctggggg 1140

tattttgatt tcaacagatt ccactgccac ttttcaccac gtgactggca gcgactcatc 1200

aacaacaatt ggggattccg gcccaagaga ctcaacttca aactcttcaa catccaagtc 1260

aaggaggtca cgacgaatga tggcgtcaca accatcgcta ataaccttac cagcacggtt 1320

caagtcttct cggactcgga gtaccagctt ccgtacgtcc tcggctctgc gcaccagggc 1380

tgcctccctc cgttcccggc ggacgtgttc atgattccgc aatacggcta cctgacgctc 1440

aacaatggca gccaagccgt 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 aatgggcgtg aatccatcat caaccctggc 1860

actgctatgg cctcacacaa agacgacgaa gacaagttct ttcccatgag cggtgtcatg 1920

atttttggaa aagagagcgc cggagcttca aacactgcat tggacaatgt catgattaca 1980

gacgaagagg aaattaaagc cactaaccct gtggccaccg aaagatttgg gaccgtggca 2040

gtcaatttcc agagcagcag cacagaccct gcgaccggag atgtgcatgc tatgggagca 2100

ttacctggca tggtgtggca agatagagac gtgtacctgc agggtcccat ttgggccaaa 2160

attcctcaca cagatggaca ctttcacccg tctcctctta tgggcggctt tggactcaag 2220

aacccgcctc ctcagatcct catcaaaaac acgcctgttc ctgcgaatcc tccggcggag 2280

ttttcagcta caaagtttgc ttcattcatc acccaatact ccacaggaca agtgagtgtg 2340

gaaattgaat gggagctgca gaaagaaaac agcaagcgct ggaatcccga agtgcagtac 2400

acatccaatt atgcaaaatc tgccaacgtt gattttactg tggacaacaa tggactttat 2460

actgagcctc gccccattgg cacccgttac cttacccgtc ccctgtaagc aagggcgaat 2520

tctgcagata tccatcacac tggcggccgc tttcgaatct agagcctgca gtctcgacaa 2580

gcttgtcgag aagtactaga ggatcataat cagccatacc acatttgtag aggttttact 2640

tgctttaaaa aacctcccac acctcccccc gaacctgaaa cataaaatga atgcaattgt 2700

tgttgttaac ttgtttattg cagcttataa tggttacaaa taaagcaata gcatcacaaa 2760

tttcacaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca aactcatcaa 2820

tgtatcttat catgtctgga tctgatcact gcttgagcct agaggcctcg cgagatctta 2880

attaattaag taccgactct gctgaagagg aggaaattct ccttgaagtt tccctggtgt 2940

tcaaagtaaa ggagtttgca ccagacgcac ctctgttcac tggtccggcg tattaaaaca 3000

cgatacattg ttattagtac atttattaag cgctagattc tgtgcgttgt tgatttacag 3060

acaattgttg tacgtatttt aataattcat taaatttata atctttaggg tggtatgtta 3120

gagcgaaaat caaatgattt tcagcgtctt tatatctgaa tttaaatatt aaatcctcaa 3180

tagatttgta aaataggttt cgattagttt caaacaaggg ttgtttttcc gaaccgatgg 3240

ctggactatc taatggattt tcgctcaacg ccacaaaact tgccaaatct tgtagcagca 3300

atctagcttt gtcgatattc gtttgtgttt tgttttgtaa taaaggttcg acgtcgttca 3360

aaatattatg cgcttttgta tttctttcat cactgtcgtt agtgtacaat tgactcgacg 3420

taaacacgtt aaataaagct tggacatatt taacatcggg cgtgttagct ttattaggcc 3480

gattatcgtc gtcgtcccaa ccctcgtcgt tagaagttgc ttccgaagac gattttgcca 3540

tagccacacg acgcctatta attgtgtcgg ctaacacgtc cgcgatcaaa tttgtagttg 3600

agctttttgg aattatttct gattgcgggc gtttttgggc gggtttcaat ctaactgtgc 3660

ccgattttaa ttcagacaac acgttagaaa gcgatggtgc aggcggtggt aacatttcag 3720

acggcaaatc tactaatggc ggcggtggtg gagctgatga taaatctacc atcggtggag 3780

gcgcaggcgg ggctggcggc ggaggcggag gcggaggtgg tggcggtgat gcagacggcg 3840

gtttaggctc aaatgtctct ttaggcaaca cagtcggcac ctcaactatt gtactggttt 3900

cgggcgccgt ttttggtttg accggtctga gacgagtgcg atttttttcg tttctaatag 3960

cttccaacaa ttgttgtctg tcgtctaaag gtgcagcggg ttgaggttcc gtcggcattg 4020

gtggagcggg cggcaattca gacatcgatg gtggtggtgg tggtggaggc gctggaatgt 4080

taggcacggg agaaggtggt ggcggcggtg ccgccggtat aatttgttct ggtttagttt 4140

gttcgcgcac gattgtgggc accggcgcag gcgccgctgg ctgcacaacg gaaggtcgtc 4200

tgcttcgagg cagcgcttgg ggtggtggca attcaatatt ataattggaa tacaaatcgt 4260

aaaaatctgc tataagcatt gtaatttcgc tatcgtttac cgtgccgata tttaacaacc 4320

gctcaatgta agcaattgta ttgtaaagag attgtctcaa gctcggatcc cgcacgccga 4380

taacaagcct tttcattttt actacagcat tgtagtggcg agacacttcg ctgtcgtcga 4440

cgtacatgta tgctttgttg tcaaaaacgt cgttggcaag ctttaaaata tttaaaagaa 4500

catctctgtt cagcaccact gtgttgtcgt aaatgttgtt tttgataatt tgcgcttccg 4560

cagtatcgac acgttcaaaa aattgatgcg catcaatttt gttgttccta ttattgaata 4620

aataagattg tacagattca tatctacgat tcgtcatggc caccacaaat gctacgctgc 4680

aaacgctggt acaattttac gaaaactgca aaaacgtcaa aactcggtat 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 tgaagcgctc aacgatttgc 5040

acaagcacaa tttcatacac aacgacataa aactcgaaaa tgtcttatat ttcgaagcac 5100

ttgatcgcgt gtatgtttgc gattacggat tgtgcaaaca cgaaaactca cttagcgtgc 5160

acgacggcac gttggagtat tttagtccgg aaaaaattcg acacacaact atgcacgttt 5220

cgtttgactg gtacgcggcg tgttaacata caagttgcta accggcggcc gacacccatt 5280

tgaaaaaagc gaagacgaaa tgttggactt gaatagcatg aagcgtcgtc agcaatacaa 5340

tgacattggc gttttaaaac acgttcgtaa cgttaacgct cgtgactttg tgtactgcct 5400

aacaagatac aacatagatt gtagactcac aaattacaaa caaattataa aacatgagtt 5460

tttgtcgtaa aaatgccact tgttttacga gtagaattcg taatcatggt catagctgtt 5520

tcctgtgtga aattgttatc cgctcacaat tccacacaac atacgagccg gaagcataaa 5580

gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca ttaattgcgt tgcgctcact 5640

gcccgctttc agtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc 5700

ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg 5760

ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc 5820

cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 5880

gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca 5940

tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca 6000

ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg 6060

atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag 6120

gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccccgt 6180

tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca 6240

cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 6300

cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt 6360

tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc 6420

cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt tgcaagcagc agattacgcg 6480

cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg 6540

gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta 6600

gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg 6660

gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg 6720

ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc 6780

atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc 6840

agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc 6900

ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag 6960

tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat 7020

ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg 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 agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac 7440

tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat 7500

aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat 7560

ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca 7620

aatagggttt ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat 7680

tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc tcgcgcgttt 7740

cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct 7800

gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg 7860

tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc accatatgcg 7920

gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc attcgccatt 7980

caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat tacgccagct 8040

ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt tttcccagtc 8100

acgacgttgt aaaacgaccg agttgtttgc gtacgtgact agcgaagaag atgtgtggac 8160

cgcagaacag atagtaaaac aaaaccctag tattggagca ataatcgatt taaccaacac 8220

gtctaaatat tatgatggtg tgcatttttt gcgggcgggc ctgttataca aaaaaattca 8280

agtacctggc cagactttgc cgcctgaaag catagttcaa gaatttattg acacggtaaa 8340

agaatttaca gaaaagtgtc ccggcatgtt ggtgggcgtg cactgcacac acggtattaa 8400

tcgcaccggt tacatggtgt gcagatattt aatgcacacc ctgggtattg cgccgcagga 8460

agccatagat agattcgaaa aagccagagg tcacaaaatt gaaagacaaa attacgttca 8520

agatttatta atttaattaa tattatttgc attctttaac aaatacttta tcctattttc 8580

aaattgttgc gcttcttcca gcgaaccaaa actatgcttc gcttgctccg tttagcttgt 8640

agccgatcag tggcgttgtt ccaatcgacg gtaggattag gccggatatt ctccaccaca 8700

atgttggcaa cgttgatgtt acgtttatgc ttttggtttt ccacgtacgt cttttggccg 8760

gtaatagccg taaacgtagt gccgtcgcgc gtcacgcaca acaccggatg tttgcgcttg 8820

tccgcggggt attgaaccgc gcgatccgac aaatccacca ctttggcaac taaatcggtg 8880

acctgcgcgt ctttttttctg cattatttcg tctttctttt gcatggtttc ctggaagccg 8940

gtgtacatgc ggtttagatc agtcatgacg cgcgtgacct gcaaatcttt ggcctcgatc 9000

tgcttgtcct tgatggcaac gatgcgttca ataaactctt gtttttttaac aagttcctcg 9060

gtttttttgcg ccaccaccgc ttgcagcgcg tttgtgtgct cggtgaatgt cgcaatcagc 9120

ttagtcacca actgtttgct ctcctcctcc cgttgtttga tcgcgggatc gtacttgccg 9180

gtgcagagca cttgaggaat tacttcttct aaaagccatt cttgtaattc tatggcgtaa 9240

ggcaatttgg acttcataat cagctgaatc acgccggatt tagtaatgag cactgtatgc 9300

ggctgcaaat acagcgggtc gccccttttc acgacgctgt tagaggtagg gcccccattt 9360

tggatggtct gctcaaataa cgatttgtat ttattgtcta catgaacacg tatagcttta 9420

tcacaaactg tatattttaa actgttagcg acgtccttgg ccacgaaccg gacctgttgg 9480

tcgcgctcta gcacgtaccg caggttgaac gtatcttctc caaatttaaa ttctccaatt 9540

ttaacgcgag ccattttgat acacgtgtgt cgattttgca acaactattg ttttttaacg 9600

caaactaaac ttattgtggt aagcaataat taaatatggg ggaacatgcg ccgctacaac 9660

actcgtcgtt atgaacgcag acggcgccgg tctcggcgca agcggctaaa acgtgttgcg 9720

cgttcaacgc ggcaaacatc gcaaaagcca atagtacagt tttgatttgc atattaacgg 9780

cgatttttta aattatctta tttaataaat agttatgacg cctacaactc cccgcccgcg 9840

ttgactcgct gcacctcgag cagttcgttg acgccttcct ccgtgtggcc gaacacgtcg 9900

agcgggtggt cgatgaccag cggcgtgccg cacgcgacgc acaagtatct gtacaccgaa 9960

tgatcgtcgg gcgaaggcac gtcggcctcc aagtggcaat attggcaaat tcgaaaatat 10020

atacagttgg gttgtttgcg catatctatc gtggcgttgg gcatgtacgt ccgaacgttg 10080

atttgcatgc aagccgaaat taaatcattg cgattagtgc gattaaaacg ttgtacatcc 10140

tcgcttttaa tcatgccgtc gattaaatcg cgcaatcgag tcaagtgatc aaagtgtgga 10200

ataatgtttt ctttgtattc ccgagtcaag cgcagcgcgt attttaacaa actagccatc 10260

ttgtaagtta gtttcattta atgcaacttt atccaataat atattatgta tcgcacgtca 10320

agaattaaca atgcgcccgt tgtcgcatct caacacgact atgatagaga tcaaataaag 10380

cgcgaattaa atagcttgcg acgcaacgtg cacgatctgt gcacgcgttc cggcacgagc 10440

tttgattgta ataagttttt acgaagcgat gacatgaccc ccgtagtgac aacgatcacg 10500

cccaaaagaa ctgccgacta caaaattacc gagtatgtcg gtgacgttaa aactattaag 10560

ccatccaatc gaccgttagt cgaatcagga ccgctggtgc gagaagccgc gaagtatggc 10620

gaatgcatcg tataacgtgt ggagtccgct cattagagcg tcatgtttag acaagaaagc 10680

tacatattta attgatcccg atgattttat tgataaattg accctaactc catacacggt 10740

attctacaat ggcggggttt tggtcaaaat ttccggactg cgattgtaca tgctgttaac 10800

ggctccgccc actattaatg aaattaaaaa ttccaatttt aaaaaacgca gcaagagaaa 10860

catttgtatg aaagaatgcg tagaaggaaa gaaaaatgtc gtcgacatgc tgaacaacaa 10920

gattaatatg cctccgtgta taaaaaaaat attgaacgat ttgaaagaaa acaatgtacc 10980

gcgcggcggt atgtacagga agaggtttat actaaactgt tacattgcaa acgtggtttc 11040

gtgtgccaag tgtgaaaacc gatgtttaat caaggctctg acgcatttct acaaccacga 11100

ctccaagtgt gtgggtgaag tcatgcatct tttaatcaaa tcccaagatg tgtataaacc 11160

accaaactgc caaaaaatga aaactgtcga caagctctgt ccgtttgctg gcaactgcaa 11220

gggtctcaat cctatttgta attattgaat aataaaacaa ttataaatgc taaatttgtt 11280

ttttattaac gatacaaacc aaacgcaaca agaacatttg tagtattatc tataattgaa 11340

aacgcgtagt tataatcgct gaggtaatat ttaaaatcat tttcaaatga ttcacagtta 11400

atttgcgaca atataatttt attttcacat aaactagacg ccttgtcgtc ttcttcttcg 11460

tattccttct ctttttcatt tttctcctca taaaaattaa catagttatt atcgtatcca 11520

tatatgtatc tatcgtatag agtaaatttt ttgttgtcat aaatatatat gtcttttta 11580

atggggtgta tagtaccgct gcgcatagtt tttctgtaat ttacaacagt gctattttct 11640

ggtagttctt cggagtgtgt tgctttaatt attaaattta tataatcaat gaatttggga 11700

tcgtcggttt tgtacaatat gttgccggca tagtacgcag cttcttctag ttcaattaca 11760

ccatttttta gcagcaccgg attaacataa ctttccaaaa tgttgtacga accgttaaac 11820

aaaaacagtt cacctccctt ttctatacta ttgtctgcga gcagttgttt gttgttaaaa 11880

ataacagcca ttgtaatgag acgcacaaac taatatcaca aactggaaat gtctatcaat 11940

atatagttgc tgat 11954

<210> 29
<211> 30
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador (directo, comprende el sitio BamHI)

<400> 29
ttaggatcct gttaaggtgg ctgccgacgg 30

<210> 30
<211> 30
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia del cebador (inverso, comprende el sitio StuI)

<400> 30
gtcgtaggcc ttgtcgtgct cgagggccgc 30

<210> 31
<211> 1966
<212> ADN
<213> virus adenoasociado 2

<220>
<221> misc_feature
<222> (1)..(1966)
<223> rep con codón de iniciación CTG y sin ATG interno

<400> 31
cggtccgaag cgcgcggaat tcaaaggcct acgtcgacga ggggagatct gccgccctgg 60

cggggtttta cgagattgtg attaaggtcc ccagcgacct tgacgagcat ctgcccggca 120

tttctgacag ctttgtgaac tgggtggccg agaaggagtg ggagttgccg ccagattctg 180

acttggatct gaatctgatt gagcaggcac ccctgaccgt ggccgagaag ctgcagcgcg 240

actttctgac ggagtggcgc cgtgtgagta aggccccgga ggccctttc tttgtgcaat 300

ttgagaaggg agagagctac ttccacttac acgtgctcgt ggaaaccacc ggggtgaaat 360

ccttagtttt gggacgtttc ctgagtcaga ttcgcgaaaa actgattcag agaatttacc 420

gcgggatcga gccgactttg ccaaactggt tcgcggtcac aaagaccaga aacggcgccg 480

gaggcgggaa caaggtggtg gacgagtgct acatccccaa ttacttgctc cccaaaaccc 540

agcctgagct ccagtgggcg tggactaatt tagaacagta tttaagcgcc tgtttgaatc 600

tcacggagcg taaacggttg gtggcgcagc atctgacgca cgtgtcgcag acgcaggagc 660

agaacaaaga gaatcagaat cccaattctg acgcgccggt gatcagatca aaaacttcag 720

ccaggtacat ggagctggtc gggtggctcg tggacaaggg gattacctcg gagaagcagt 780

ggatccagga ggaccaggcc tcatacatct ccttcaatgc ggcctccaac tcgcggtccc 840

aaatcaaggc tgccttggac aatgcgggaa agattatgag cctgactaaa accgcccccg 900

actacctggt gggccagcag cccgtggagg acatttccag caatcggatt tataaaattt 960

tggaactaaa cgggtacgat ccccaatatg cggcttccgt ctttctggga tgggccacga 1020

aaaagttcgg caagaggaac accatctggc tgtttgggcc tgcaactacc gggaagacca 1080

acatcgcgga ggccatagcc cacactgtgc ccttctacgg gtgcgtaaac tggaccaatg 1140

agaactttcc cttcaacgac tgtgtcgaca agatggtgat ctggtgggag gaggggaaga 1200

tgaccgccaa ggtcgtggag tcggccaaag ccattctcgg aggaagcaag gtgcgcgtgg 1260

accagaaatg caagtcctcg gcccagatag acccgactcc cgtgatcgtc acctccaaca 1320

ccaacatgtg cgccgtgatt gacgggaact caacgacctt cgaacaccag cagccgttgc 1380

aagaccggat gttcaaattt gaactcaccc gccgtctgga tcatgacttt gggaaggtca 1440

ccaagcagga agtcaaagac tttttccggt gggcaaagga tcacgtggtt gaggtggagc 1500

atgaattcta cgtcaaaaag ggtggagcca agaaaagacc cgcccccagt gacgcagata 1560

taagtgagcc caaacgggtg cgcgagtcag ttgcgcagcc atcgacgtca gacgcggaag 1620

cttcgatcaa ctacgcagac aggtaccaaa acaaatgttc tcgtcacgtg ggcatgaatc 1680

tgatgctgtt tccctgcaga caatgcgaga gaatgaatca gaattcaaat atctgcttca 1740

ctcacggaca gaaagactgt ttagagtgct ttcccgtgtc agaatctcaa cccgtttctg 1800

tcgtcaaaaa ggcgtatcag aaactgtgct acattcatca tatcatggga aaggtgccag 1860

acgcttgcac tgcctgcgat ctggtcaatg tggatttgga tgactgcatc tttgaacaat 1920

aaactcgagg acaatcaagc ttgcatgcct gcaggtcgac tctaga 1966

REIVINDICACIONES

1. Método para la producción de un virión parvoviral recombinante, el cual comprende los pasos de:

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

(i) una secuencia de nucleótidos que comprende un transgén que está flanqueado 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 comprende un único marco de lectura abierto que codifica una proteína Rep78 parvoviral de longitud completa a partir de la cual se expresan las proteínas Rep52 y Rep78 en la célula de insecto, cuya secuencia de nucleótidos está operativamente enlazada a un primer promotor que es capaz de impulsar la expresión de las proteínas Rep en la célula de insecto, y en el que la secuencia de nucleótidos que codifica la proteína parvoviral Rep78 no comprende un intrón artificial;
(iii) un segundo casete de expresión que comprende una secuencia de nucleótidos que codifica una proteína de la cápside parvoviral que está operativamente enlazada a un segundo promotor que es capaz de dirigir la expresión de la proteína de la cápside en la célula de insecto; y

(b) cultivar la célula definida en (a) en condiciones propicias para la expresión de la Rep y la proteína de la cápside;

donde los casetes de expresión primero y segundo están presentes en un solo constructo de ácidos nucleicos, donde los casetes de expresión primero y segundo están presentes en cantidad equimolar en la célula de insecto, y donde el primer promotor es tan fuerte como el segundo promotor.

2. Método según la reivindicación 1, en el que el método comprende además la etapa de recuperación del virión parvoviral recombinante.

3. Método según las reivindicaciones 1 o 2, en el que la secuencia de nucleótidos de (iii) que codifica una proteína de la cápside parvoviral comprende un marco de lectura abierto que comprende secuencias de nucleótidos que codifican las proteínas de la cápside VP1, VP2 y VP3.

4. Método según la reivindicación 3, en el que el marco de lectura abierto que comprende secuencias de nucleótidos que codifican las proteínas de la cápside VP1, VP2 y VP3 no comprende un intrón artificial.

5. Método según cualquiera de las reivindicaciones 1-4, en el que el primer promotor o el segundo promotor se selecciona del grupo que consiste en un promotor PolH, un promotor p10, un promotor básico, un promotor inducible, un promotor E1 y un promotor deltaE1.

6. Método según cualquiera de las reivindicaciones 1-5, en el que el primer casete de expresión comprende al menos un elemento potenciador de baculovirus y/o al menos un elemento sensible a la ecdisona.

7. Método según la reivindicación 6, en el que el elemento potenciador se selecciona del grupo que consta de hr1, hr2, hr3, hr4 y hr5.

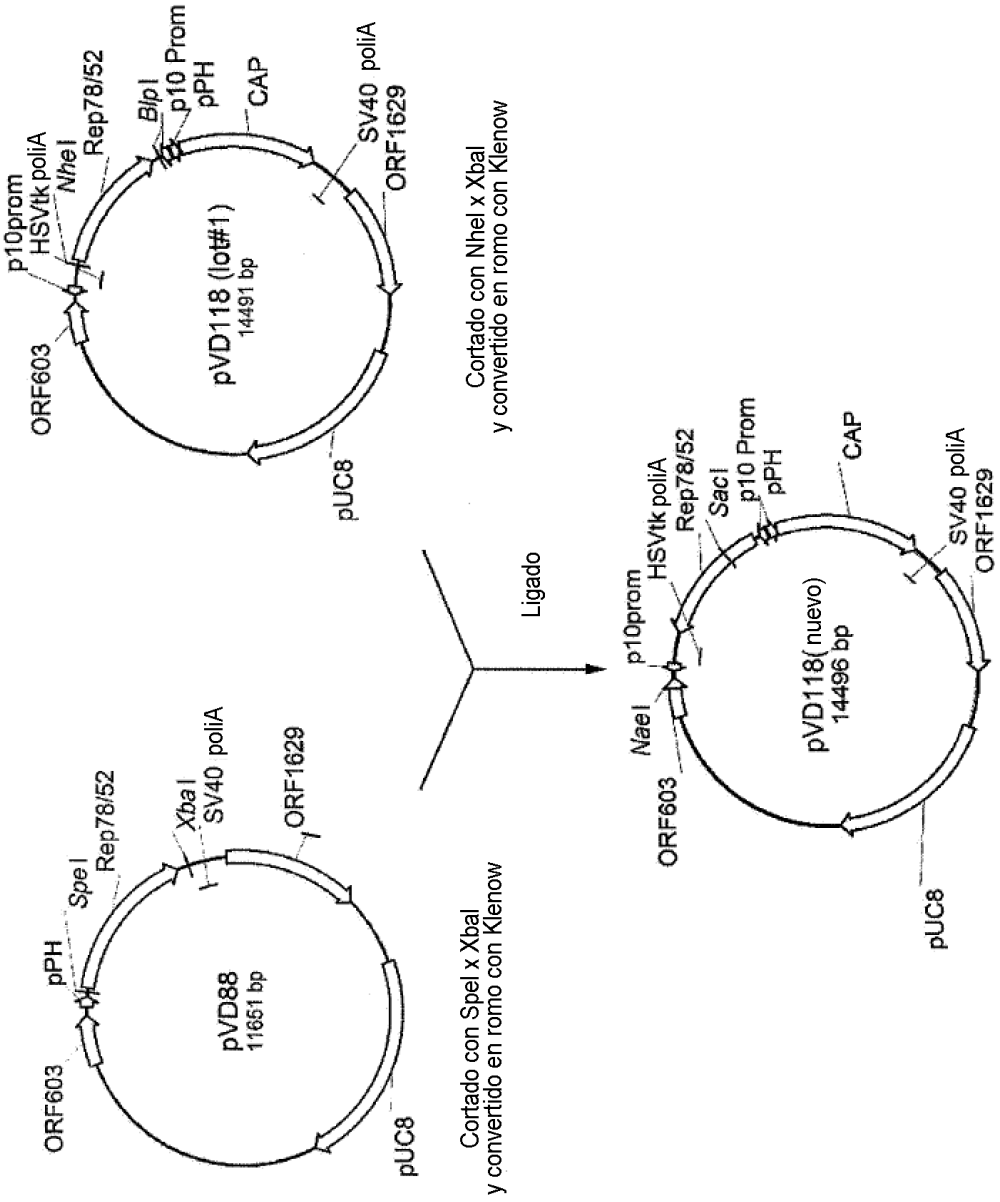
8. Método según cualquiera de las reivindicaciones 1-7, en el que la secuencia de nucleótidos de repetición terminal invertida parvoviral, la proteína Rep78 parvoviral y/o la proteína de la cápside parvoviral son de un virus adenoasociado (AAV).

9. Célula de insecto tal como se define en cualquiera de las reivindicaciones 1-8.

10. Kit que comprende

(a) un constructo de ácidos nucleicos que comprende el primer y segundo casete de expresión tal como se define en cualquiera de las reivindicaciones 1-9; y
(b) un constructo de ácidos nucleicos que comprende una secuencia de nucleótidos que codifica un sitio de clonación múltiple para un transgén que está flanqueado por al menos una secuencia de nucleótidos de repetición terminal invertida parvoviral, cuyo transgén está operativamente enlazado a un promotor capaz de impulsar la expresión del transgén en una célula huésped.

Fig 1


pVD118 (lot#1)
14491 bp
p10prom
HSVtk polyA
NheI
Rep78/52
BlpI
p10 Prom
pPH
CAP
SV40 polyA
ORF1629
ORF603
pUC8
Cortado con NheI x XbaI
y convertido en romo con Klenow
Ligado
pVD88
11651 bp
pPH
SpeI
Rep78/52
XbaI
SV40 polyA
ORF1629
ORF603
pUC8
Cortado con SpeI x XbaI
y convertido en romo con Klenow
pVD118(nuevo)
14496 bp
p10prom
HSVtk polyA
Rep78/52
SacI
p10 Prom
pPH
CAP
SV40 polyA
ORF1629
NaeI
ORF603
pUC8

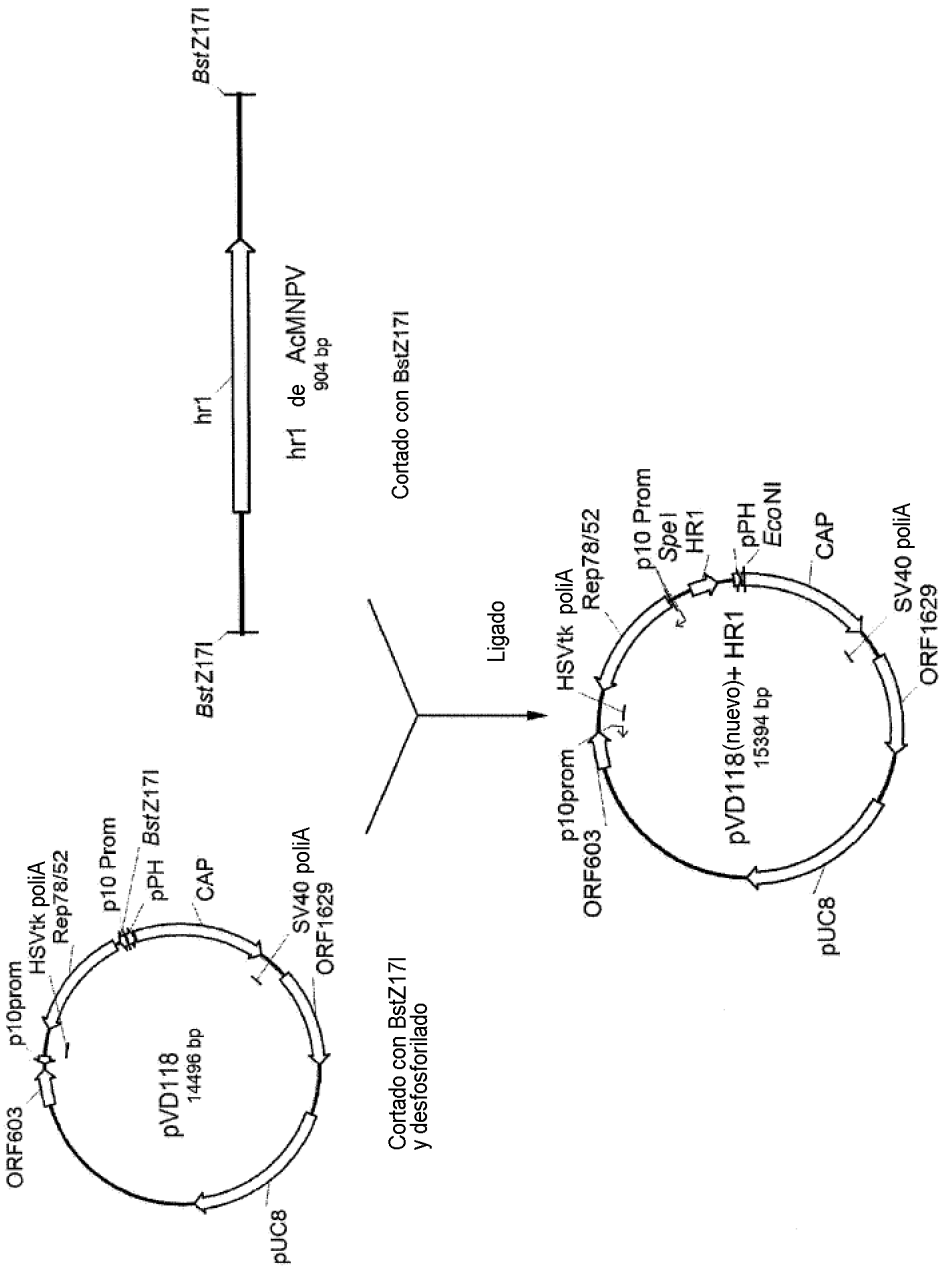

BstZ17I
hr1
hr1 de AcMNPV
904 bp
BstZ17I
Cortado con BstZ17I
Ligado
Cortado con BstZ17I
y desfosforilado
ORF603
p10prom
HSVtk polìA
Rep78/52
p10 Prom
pPH
BstZ17I
CAP
SV40 polìA
ORF1629
pUC8
pVD118
14496 bp
p10prom
ORF603
HSVtk polìA
Rep78/52
p10 Prom
SpeI
HR1
pPH
EcoNI
CAP
SV40 polìA
ORF1629
pUC8
pVD118(nuevo)+ HR1
15394 bp

Fig 2

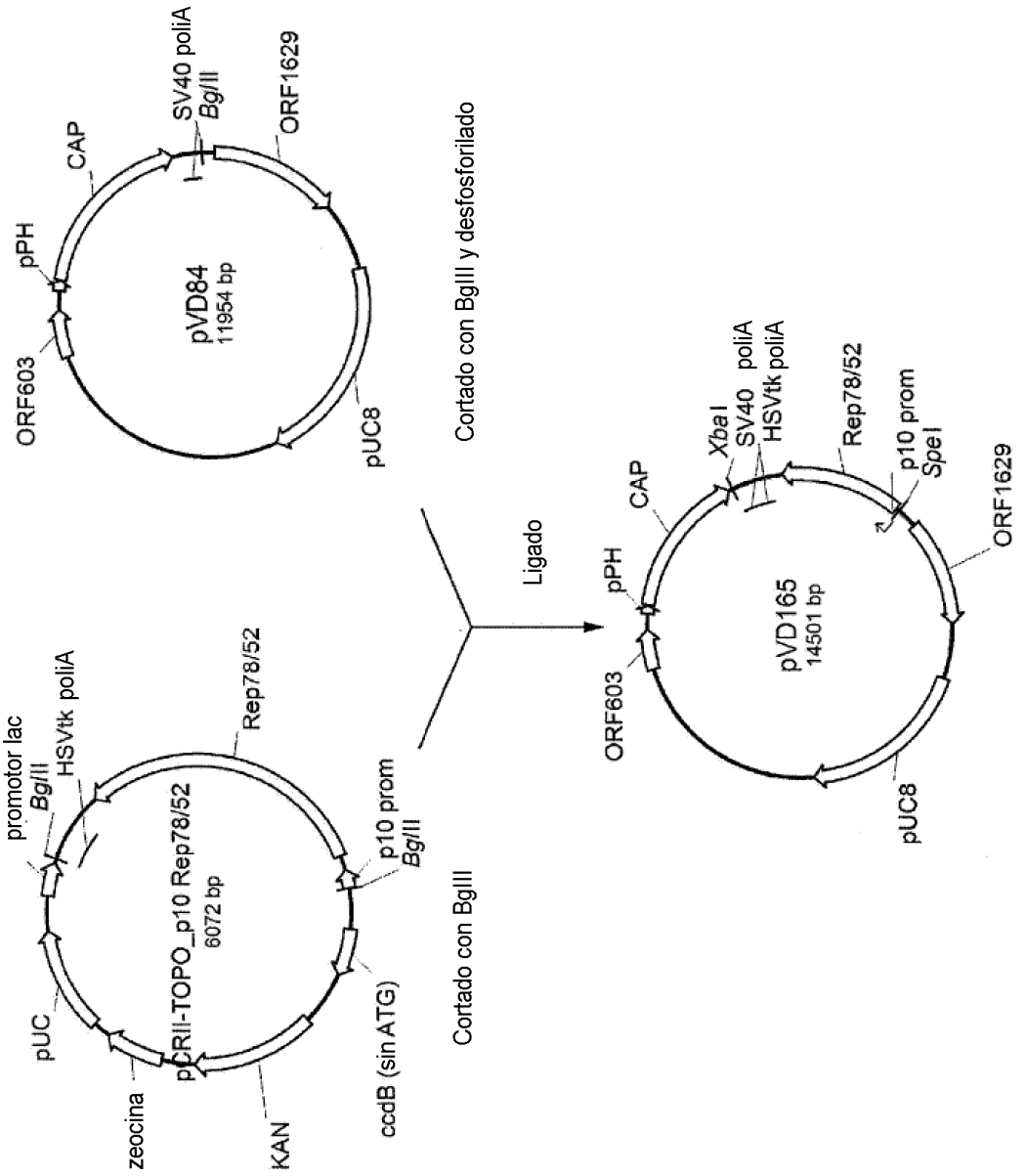

pVD84
11954 bp
ORF603
pPH
CAP
SV40 poliA
BglII
ORF1629
pUC8
Cortado con BglII y desfosforilado
Ligado
pCRII-TOPO_p10 Rep78/52
6072 bp
promotor lac
BglII
HSVtk poliA
Rep78/52
p10 prom
BglII
pUC
zeocina
KAN
ccdB (sin ATG)
Cortado con BglII
pVD165
14501 bp
ORF603
pPH
CAP
XbaI
SV40 poliA
HSVtk poliA
Rep78/52
p10 prom
SpeI
ORF1629
pUC8

Fig 3

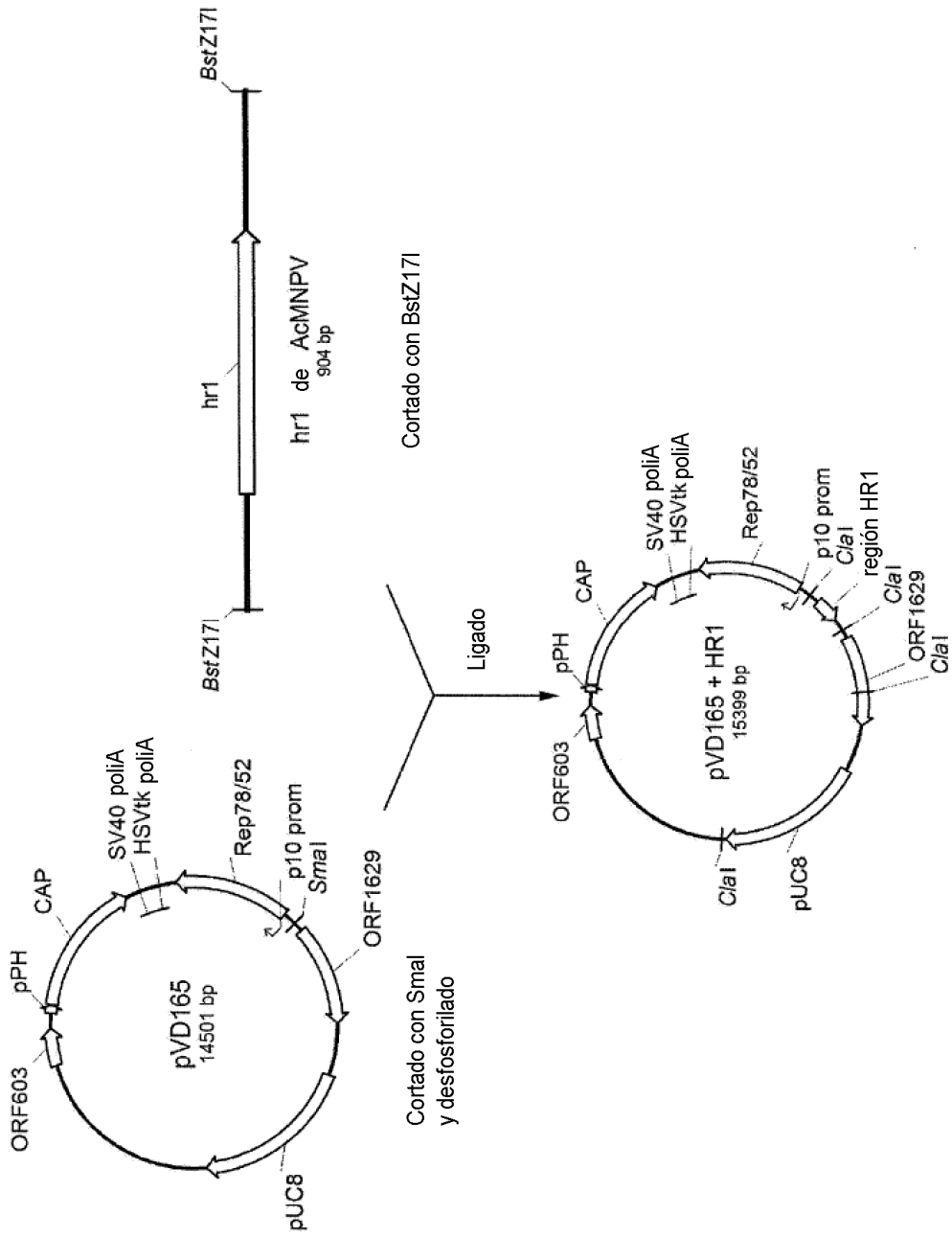

BstZ17I
hr1
hr1 de AcMNPV
904 bp
BstZ17I
Cortado con BstZ17I
Ligado
pVD165
14501 bp
pPH
CAP
SV40 poliA
HSVtk poliA
Rep78/52
p10 prom
SmaI
ORF1629
ORF603
pUC8
Cortado con SmaI
y desfosforilado
pVD165 + HR1
15399 bp
ClaI
región HR1

Fig 4

Fig 5

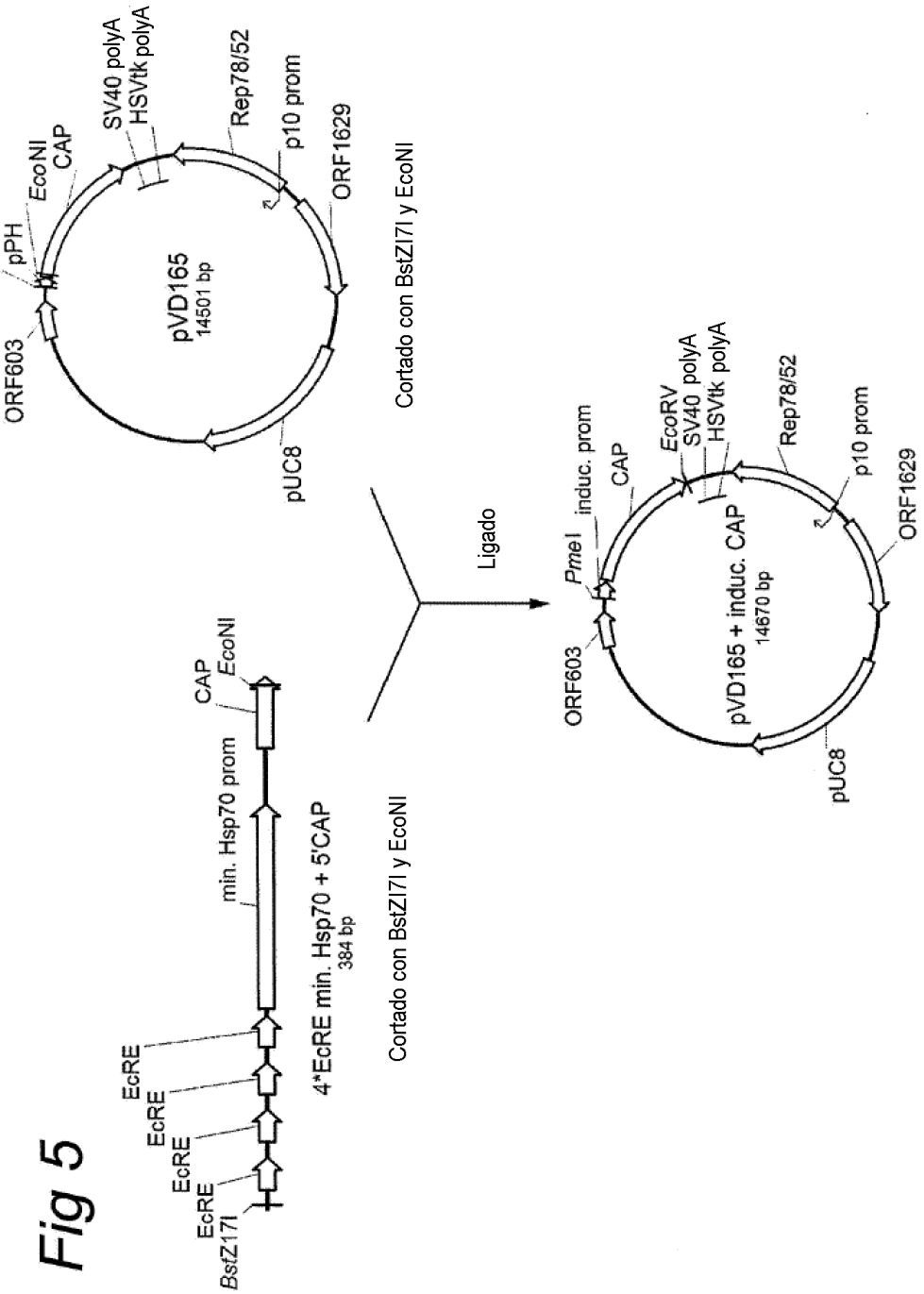
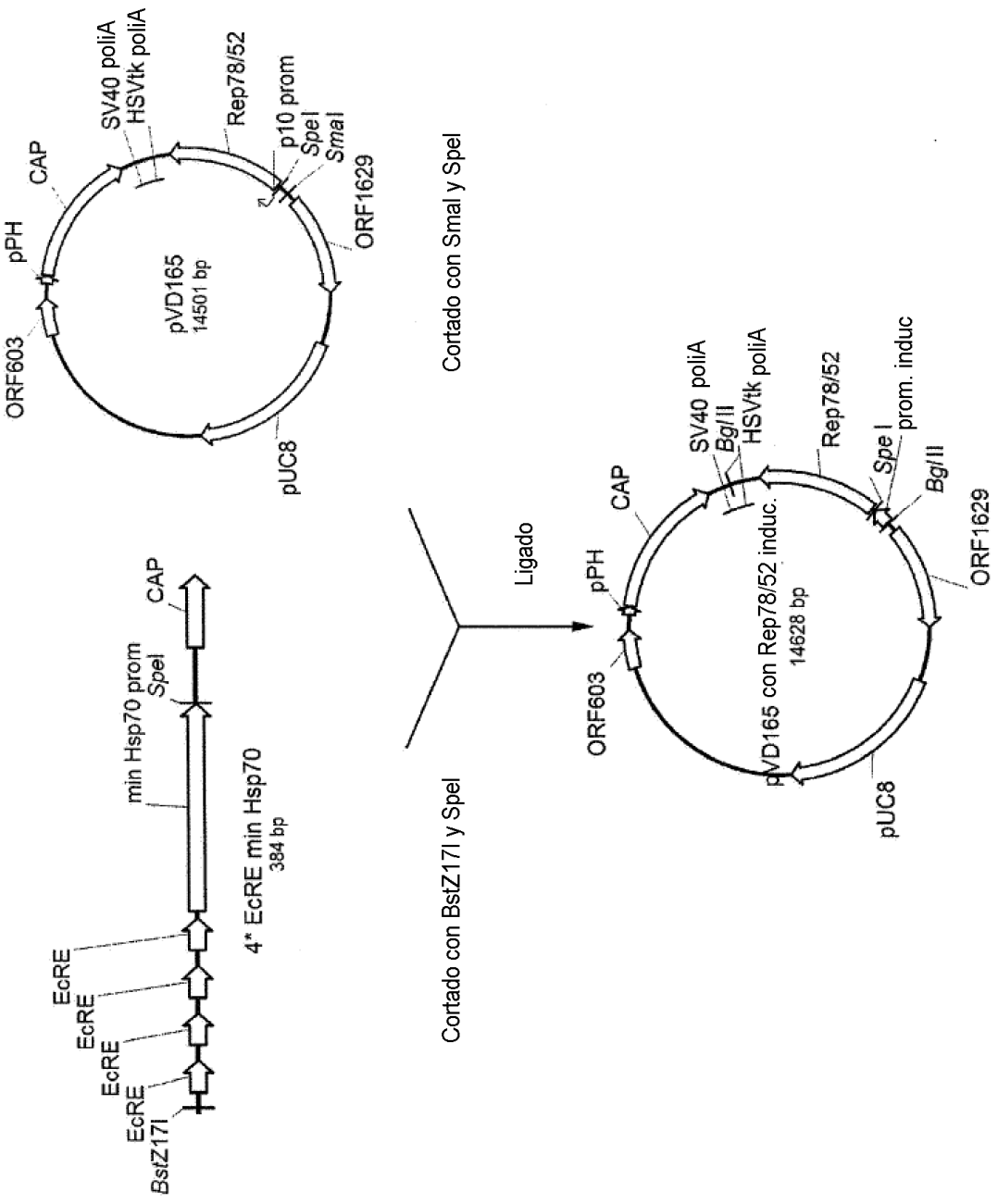

pVD165
14501 bp
ORF603
BstZ17I
pPH
EcoNI
CAP
SV40 polyA
HSVtk polyA
Rep78/52
p10 prom
ORF1629
pUC8
Cortado con BstZI7I y EcoNI
4*EcRE min. Hsp70 + 5'CAP
384 bp
EcRE
EcRE
EcRE
EcRE
BstZ17I
min. Hsp70 prom
CAP
EcoNI
Cortado con BstZI7I y EcoNI
Ligado
pVD165 + induc. CAP
14670 bp
ORF603
PmeI
induc. prom
CAP
EcoRV
SV40 polyA
HSVtk polyA
Rep78/52
p10 prom
ORF1629
pUC8

Fig 6


pVD165
14501 bp
ORF603
pPH
CAP
SV40 poliA
HSVtk poliA
Rep78/52
p10 prom
SpeI
SmaI
ORF1629
pUC8
Cortado con SmaI y SpeI
Cortado con BstZ17I y SpeI
Ligado
4* EcRE min Hsp70
384 bp
EcRE
EcRE
EcRE
EcRE
BstZ17I
min Hsp70 prom
SpeI
CAP
pVD165 con Rep78/52 induc.
14628 bp
SV40 poliA
BglII
HSVtk poliA
Rep78/52
Spe I
prom. induc
BglII
ORF1629
pUC8

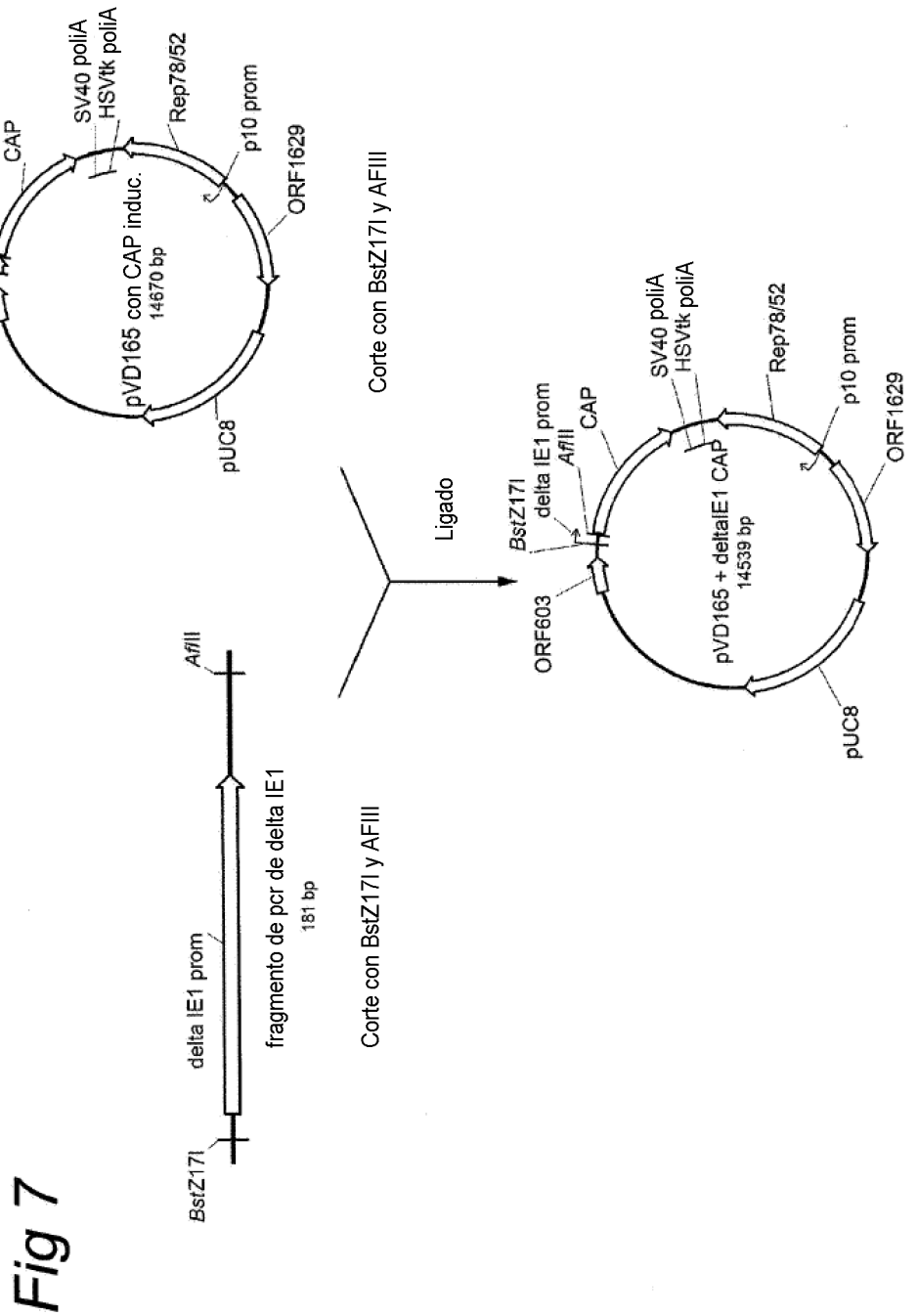

BstZ17I
induc. prom
AflII
CAP
SV40 poliA
HSVtk poliA
Rep78/52
p10 prom
ORF1629
ORF603
pVD165 con CAP induc.
14670 bp
pUC8
Corte con BstZ17I y AFIII
Ligado
AflII
delta IE1 prom
fragmento de pcr de delta IE1
181 bp
BstZ17I
Corte con BstZ17I y AFIII
BstZ17I
delta IE1 prom
AflII
CAP
SV40 poliA
HSVtk poliA
Rep78/52
p10 prom
ORF1629
ORF603
pVD165 + deltaIE1 CAP
14539 bp
pUC8

Fig 7

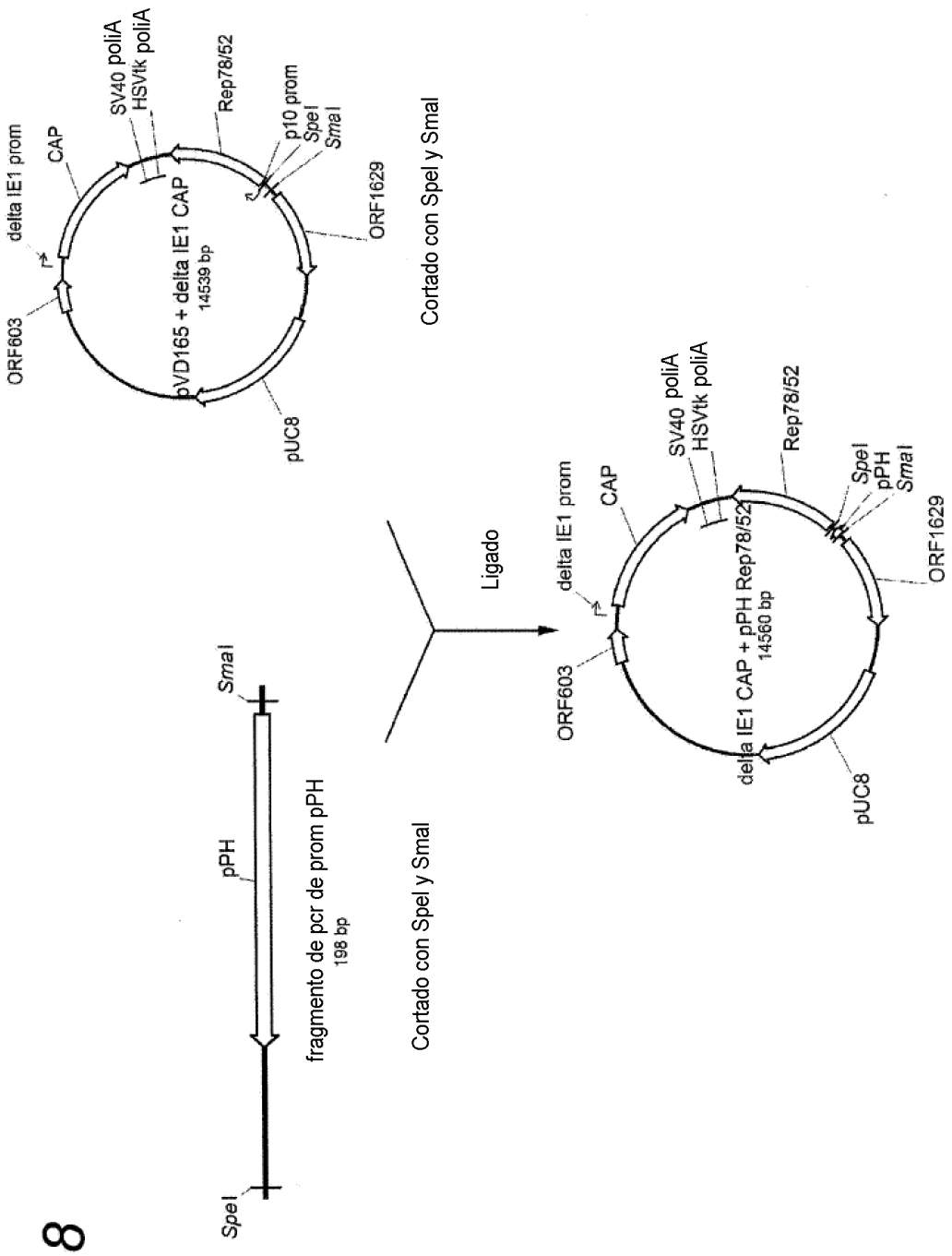

pVD165 + delta IE1 CAP
14539 bp
delta IE1 prom
CAP
SV40 poliA
HSVtk poliA
Rep78/52
p10 prom
SpeI
SmaI
ORF1629
ORF603
pUC8
Cortado con SpeI y SmaI
Ligado
Cortado con SpeI y SmaI
SmaI
pPH
SpeI
fragmento de pcr de prom pPH
198 bp
delta IE1 CAP + pPH Rep78/52
14560 bp

Fig 8

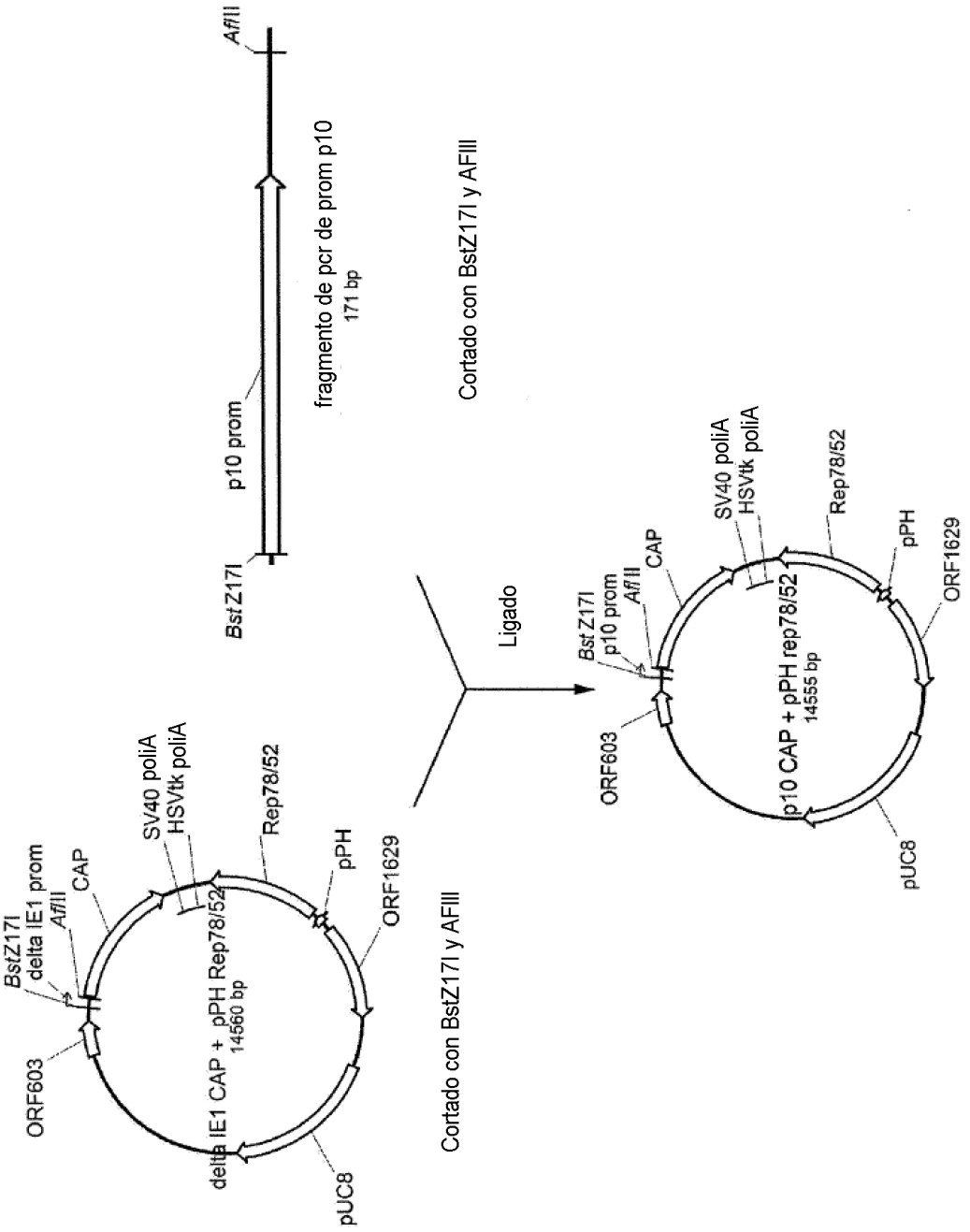

AflII
p10 prom
BstZ17I
fragmento de pcr de prom p10
171 bp
Cortado con BstZ17I y AFIII
Ligado
Cortado con BstZ17I y AFIII
BstZ17I
delta IE1 prom
AflII
CAP
SV40 poliA
HSVtk poliA
Rep78/52
pPH
ORF1629
ORF603
delta IE1 CAP + pPH Rep78/52
14560 bp
pUC8
BstZ17I
p10 prom
AflII
CAP
SV40 poliA
HSVtk poliA
Rep78/52
pPH
ORF1629
ORF603
p10 CAP + pPH rep78/52
14555 bp
pUC8

Fig 9

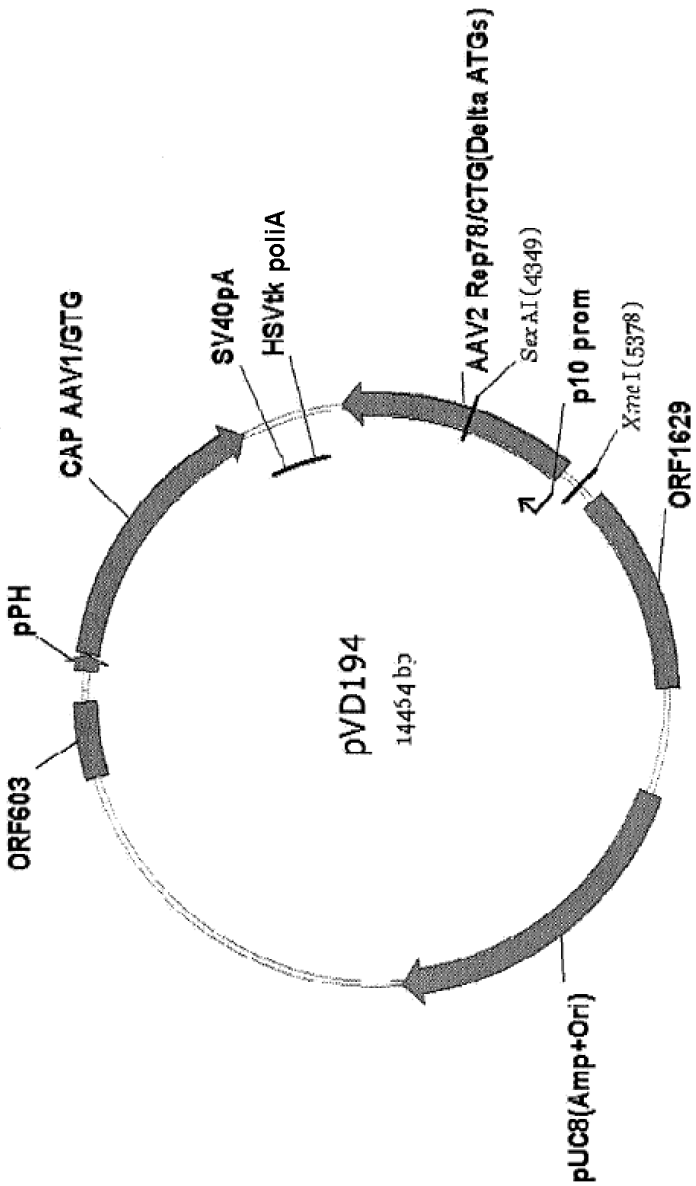

pVD194
14454 bp
AAV2 Rep78/CTG(Delta ATGs)
SexAI (4349)
p10 prom
XmnI (5378)
ORF1629
pUC8(Amp+Ori)
ORF603
pPH
CAP AAV1/GTG
SV40pA
HSVtk poliA

Fig. 10

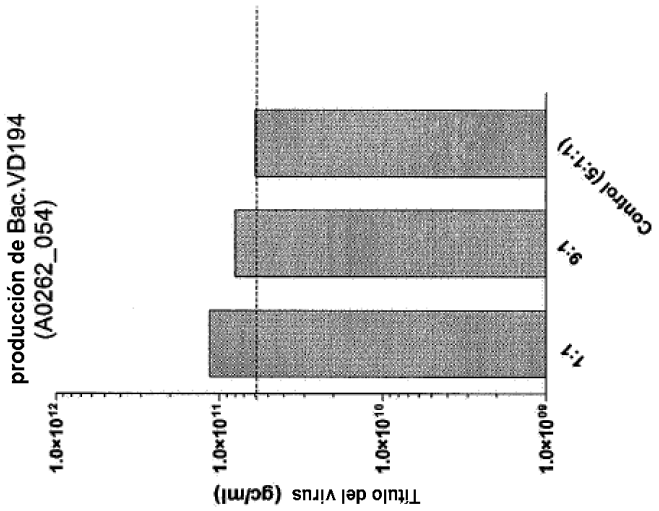


Fig. 11B

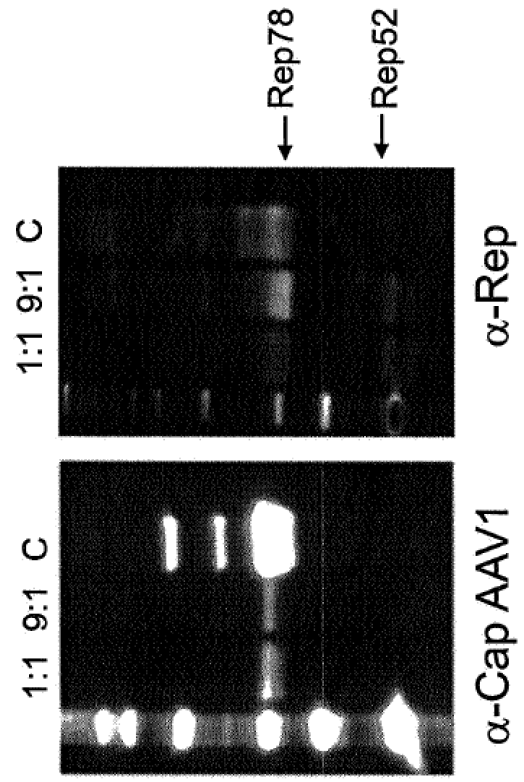


Fig. 11A