

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 602 585**

51 Int. Cl.:

**C07K 14/005** (2006.01)

**C12N 7/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.07.2008 PCT/NL2008/050512**

87 Fecha y número de publicación internacional: **29.01.2009 WO09014445**

96 Fecha de presentación y número de la solicitud europea: **25.07.2008 E 08779058 (0)**

97 Fecha y número de publicación de la concesión europea: **17.08.2016 EP 2173888**

54 Título: **Vectores de baculovirus que comprenden secuencias de codificación repetidas con sesgos de codones diferenciales**

30 Prioridad:

**26.07.2007 US 952081 P**  
**26.07.2007 EP 07113257**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**21.02.2017**

73 Titular/es:

**UNIQUE IP B.V. (100.0%)**  
**MEIBERGDREEF 61**  
**1105 BA Amsterdam Zuidoost, NL**

72 Inventor/es:

**BAKKER, ANDREW, CHRISTIAN y**  
**HERMENS, WILHELMUS, THEODORUS,**  
**JOHANNES**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

**Observaciones:**

**Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes**

**ES 2 602 585 T3**

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

## DESCRIPCIÓN

Vectores de baculovirus que comprenden secuencias de codificación repetidas con sesgos de codones diferenciales

**Campo de la invención**

5 La presente invención se refiere a los campos de la medicina, la biología molecular y la terapia génica. La invención se refiere a la producción de proteínas en células de insectos mediante la cual se utilizan secuencias de codificación repetidas en vectores de baculovirus. En particular, la invención se refiere a la producción de vectores parvovíricos que pueden utilizarse en la terapia génica y a mejoras en la expresión de las proteínas rep virales que incrementen la productividad de los vectores parvovíricos.

**Antecedentes de la invención**

10 El sistema de expresión de baculovirus es bien conocido por su uso como vector de clonación y expresión eucariótico (King, L. A., y R. D. Possee, 1992, "The baculovirus expression system", Chapman y Hall, United Kingdom; O'Reilly, D. R., y col., 1992. Baculovirus Expression Vectors: A Laboratory Manual. New York: W. H. Freeman.). Entre otras ventajas del sistema de expresión de baculovirus se encuentran que las proteínas expresadas son casi siempre solubles, están correctamente plegadas y son biológicamente activas. Otras ventajas  
15 incluyen el conocimiento de las preferencias de codones de las células de insectos (véase, por ejemplo, el documento US 2003/0228696), los altos niveles de expresión de proteínas, una producción más rápida, la idoneidad para la expresión de proteínas de gran tamaño y la 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 conocido como el 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 de baculovirus.

20 El sistema de expresión den baculovirus se ha usado con éxito también para la producción de vectores de virus adenoasociados (VAA) recombinantes (Urabe y col., 2002, Hum. Gene Ther. 13: 1935-1943; documento US 6.723.551 y documento US 20040197895). Los VAA pueden considerarse no de los vectores virales más prometedores para la terapia génica humana. Los VAA tienen la capacidad de infectar eficazmente dividiendo así como no dividiendo las células humanas, el genoma del virus VAA se integra en un único sitio cromosómico en el genoma de la célula huésped, y lo más importante, aunque el VAA está presente en muchos seres humanos, nunca se ha asociado con ninguna enfermedad. En vista de estas ventajas, se está evaluando el virus adenoasociado recombinante (VAAr) en ensayos clínicos de terapia génica para la hemofilia B, el melanoma maligno, la fibrosis  
25 quística, la hiperlipoproteinemia de tipo I y otras enfermedades.

30 Para superar los problemas con los sistemas de producción de mamífero para VAA Urabe y col., (2002, mencionado anteriormente) desarrollaron un sistema de producción de VAA en células de insecto. Para la producción de VAA en células de insecto fue necesario realizar algunas modificaciones con el fin de lograr la estequiometría correcta de las tres proteínas de la cápside de VAA (VP1, VP2 y VP3), que se basa en una combinación del uso alterno de dos sitios aceptores de corte y empalme y la utilización subóptima de un codón de iniciación ACG para VP2 que no se reproduce con precisión por células de insectos. Para imitar la estequiometría correcta de las proteínas de la cápside en células de insectos, Urabe y col., ((2002, citado anteriormente) utilizan una construcción que se transcribe en un único mensajero policistrónico que es capaz de expresar las tres proteínas VP sin necesidad de corte y empalme y en el que el codón iniciador en dirección 5' se sustituye por el codón iniciador subóptimo ACG. El documento  
35 WO2007/046703 da a conocer una mejora adicional de la infectividad de los vectores VAAr producidos en baculovirus basados en la producción mediante optimización de la estequiometría de las proteínas de la cápside de VAA tal como se producen en células de insectos. El documento WO2007/084773 da a conocer la suplementación de VP1 en relación con VP2 y VP3 para aumentar la producción de partículas virales infecciosas, por ejemplo mediante la introducción en la célula de insecto de un vector Cap que comprende secuencias de nucleótidos de expresión de VP, VP2 y VP3 y un vector de VP1 que comprende secuencias de nucleótidos que expresan VP1.

40 Para la expresión de las proteínas Rep de VAA en el sistema de expresión de células de insecto de VAA como han desarrollado inicialmente Urabe y col., (2002, citado anteriormente), se usa una construcción de baculovirus recombinante que alberga dos unidades independientes de expresión de Rep (una para Rep78 y una para Rep52), cada una bajo el control de un promotor de células de insecto distinto, los promotores  $\Delta$ IE1 y PolH, respectivamente.

50 No obstante, Kohlbrenner y col., (2005, Mol. Ther. 12: 1217-25; documento WO 2005/072364) notificaron que la construcción de baculovirus para la expresión de las dos proteínas Rep, tal como utilizan Urabe y col., sufre de una inestabilidad inherente. Al dividir la orientación palindrómica de los dos genes de Rep en el vector original de Urabe y diseñando dos vectores de baculovirus independientes para expresar Rep52 y Rep78, Kohlbrenner y col., (2005, citado anteriormente) se aumentó la estabilidad al pase del vector. Sin embargo, a pesar de la expresión coherente de Rep78 y Rep52 a partir de las dos construcciones independientes de Rep en baculovirus en células de insecto en  
55 al menos 5 pases, el rendimiento del vector VAAr es de 5 a 10 veces menor en comparación con la construcción de baculovirus-Rep original diseñado por Urabe y col., (2002, mencionado anteriormente).

En el documento WO2007/148971, los presentes inventores han mejorado significativamente la estabilidad de la producción del vector de VAAr en células de insecto mediante el uso de una sola secuencia de codificación para las proteínas Rep78 y Rep52 en el que se utiliza un codón iniciador subóptimo para la proteína Rep78 que está parcialmente omitido por los ribosomas de exploración para permitir la iniciación de la traducción también en dirección 3' en el codón de iniciación de la proteína Rep52.

Sin embargo, existe todavía la necesidad de mejoras adicionales en la producción a gran escala (comercial) de proteínas heterólogas, incluyendo los vectores de VAAr, en células de insecto. Por lo tanto es un objeto de la presente invención proporcionar medios y procedimientos que permitan un rendimiento estable y alto (a gran escala) de la producción de proteínas heterólogas y vectores parvovíricos.

## 10 **Descripción de la invención**

### Definiciones

Tal como se utiliza en el presente documento, la expresión "unido operativamente" se refiere a un enlace de elementos de polinucleótido (o polipéptido) en una relación funcional. Un ácido nucleico está "unido operativamente" cuando se coloca en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, una secuencia reguladora de la transcripción está operativamente unida a una secuencia de codificación si afecta a la transcripción de la secuencia de codificación. Unido operativamente significa que las secuencias de ADN que están unidas son típicamente contiguas y, cuando es necesario, la unión de dos regiones de codificación de proteínas, contiguas y en el marco de lectura.

"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 unida operativamente. Una secuencia de control de la expresión está "unido operativamente" 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 de entrada al ribosoma interno (IRES), terminadores de la transcripción, un codón de iniciación en el frente de un gen que codifica la proteína, señal de corte y empalme para intrones, y codones de terminación. La expresión "secuencia de control de la expresión" se pretende que incluya, como mínimo, una secuencia cuya presencia está diseñada para influir sobre la expresión, y también puede incluir componentes ventajosos adicionales. Por ejemplo, secuencias líder y secuencias de pareja de fusión son secuencias de control de la expresión. La expresión también puede incluir el diseño de la secuencia de ácido nucleico de tal manera que se eliminan de la secuencia posibles codones de iniciación dentro y fuera del marco. También puede incluir el diseño de la secuencia de ácido nucleico de tal manera que se eliminan los sitios de corte y empalme potenciales no deseados. Incluye secuencias o secuencias de poliadenilación (pA) que dirigen la adición de una cola de poliA, es decir, una cadena de residuos de adenina en el extremo 3' de un ARNm, secuencias denominadas secuencias poliA. También pueden diseñarse para mejorar la estabilidad del ARNm. Las secuencias de control de la expresión que afectan a la estabilidad de la transcripción y traducción de estabilidad, por ejemplo, promotores, así como las secuencias que efectúan la traducción, por ejemplo, secuencias de Kozak, se conocen en células de insecto. Las secuencias de control de la expresión pueden ser de tal naturaleza que modulan la secuencia de nucleótidos a la que están unidas operativamente de tal manera que se consiguen niveles de expresión más bajos o niveles de expresión más altos.

Tal como se utiliza en el presente documento, el término "promotor" o "secuencia reguladora de la transcripción" se refiere a un fragmento de ácido nucleico que funciona controlando la transcripción de una o más secuencias de codificación, y está situado en dirección 5' con respecto a la dirección de la transcripción del sitio de iniciación de la transcripción de la secuencia de codificación y está estructuralmente identificado 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, pero sin limitaciones, los sitios de unión para el factor de transcripción, sitios de unión para la proteína represora y activadora, y cualquier otra secuencia de nucleótidos conocida para un experto en la técnica para actuar directa o indirectamente para regular la cantidad de transcripción a partir del promotor. Un promotor "constitutivo" es un promotor que es activo en la mayoría de los tejidos en las condiciones ambientales y fisiológicas del desarrollo. Un promotor "inducible" es un promotor que está regulado fisiológicamente en el desarrollo, por ejemplo, por la aplicación de un inductor químico. Un promotor "específico de tejido" es activo únicamente en tipos específicos de células o tejidos.

Las expresiones "sustancialmente idéntico/a", "identidad sustancial" o "esencialmente similar/es" o "similitud esencial" significa que dos péptidos o dos secuencias de nucleótidos, cuando están alineadas óptimamente, tales como mediante los programas GAP o BESTFIT usando parámetros por defecto, comparten al menos un cierto porcentaje de identidad de secuencia como se define en otro lugar del presente documento. GAP usa el algoritmo de alineación global de Needleman y Wunsch para alinear dos secuencias sobre la totalidad de su longitud, de modo que se maximiza el número de coincidencias y minimiza el número de huecos. Generalmente se usan los parámetros por defecto de GAP con una penalización por creación de hueco= 50 (nucleótidos)/8 (proteínas) y penalización por extensión de hueco= 3 (nucleótidos)/2 (proteínas). Para los nucleótidos, la matriz de puntuación por defecto usada es nwsgapdna y para las proteínas, la matriz de puntuación por defecto es Blosum62 (Henikoff y Henikoff, 1992, PNAS 89, 915-919). Está claro que cuando se dice que las secuencias de ARN son esencialmente

similares o tienen un cierto grado de identidad de secuencia con secuencias de ADN, la timina (T) en la secuencia de ADN se considera igual a uracilo (U) en la secuencia de ARN. Las alineaciones y puntuaciones de las secuencias para el porcentaje de la identidad de secuencia se puede determinar, por ejemplo, usando programas de ordenador, tal como el paquete GCG Wisconsin Versión 10.3, disponible en Accelrys Inc., 9685 Scranton Road, San Diego, CA 92121-3752 USA o el software de fuente abierta EMBOSSE para Windows (versión actual 2.7, 1-07). Como alternativa se puede determinar el porcentaje de similitud o identidad mediante la búsqueda en bases de datos tales como FASTA, BLAST, etc.

Las secuencias de nucleótidos que codifican proteínas Rep de parvovirus de la invención también pueden definirse por su capacidad para hibridarse con las secuencias de nucleótidos de SEQ ID NO. de 1 a 4, respectivamente, en condiciones de hibridación moderadas o, preferentemente, rigurosas. Las condiciones de hibridación rigurosas se definen en el presente documento como condiciones que permiten que una secuencia de ácido nucleico de al menos aproximadamente 25, preferentemente aproximadamente 50 nucleótidos, 75 o 100 y, lo más preferentemente, de aproximadamente 200 o más nucleótidos, hibride a una temperatura de aproximadamente 65 °C en una solución que comprende aproximadamente 1 M de sal, preferentemente 6 x SSC o cualquier otra solución con una fuerza iónica comparable, y lavado a 65 °C en una solución que comprende aproximadamente 0,1 M de sal, o menos, preferentemente 0,2 x SSC o cualquier otra solución con una comparable fuerza iónica. Preferentemente, la hibridación se realiza durante la noche, es decir al menos durante 10 horas y, preferentemente, el lavado se lleva a cabo durante al menos una hora con al menos dos cambios de la solución de lavado. Estas condiciones generalmente permitirán la hibridación específica de secuencias que tienen aproximadamente 90 % o más de identidad de secuencia.

Las condiciones moderadas se definen en el presente documento como condiciones que permiten que secuencias de ácido nucleico de al menos 50 nucleótidos, preferentemente aproximadamente 200 o más nucleótidos, hibriden a una temperatura de aproximadamente 45 °C en una solución que comprende aproximadamente 1 M de sal, preferentemente 6 x SSC o cualquier otra solución con una fuerza iónica comparable, y lavado a temperatura ambiente en una solución que comprende aproximadamente 1 M de sal, preferentemente 6 x SSC o cualquier otra solución con una fuerza iónica comparable. Preferentemente, la hibridación se realiza durante la noche, es decir al menos durante 10 horas y, preferentemente, el lavado se lleva a cabo durante al menos una hora con al menos dos cambios de la solución de lavado. Estas condiciones generalmente permitirán la hibridación específica de secuencias que tienen hasta un 50 % de identidad de secuencia. El experto en la técnica será capaz de modificar estas condiciones de hibridación con el fin de identificar específicamente las secuencias con una identidad variable entre 50 % y 90 %.

### **Descripción detallada de la invención**

En algunos aspectos, la presente divulgación se refiere a la utilización de parvovirus animales, en particular dependovirus, tales como VAA infecciosos de simio o humanos, y los componentes de los mismos (por ejemplo, un genoma de parvovirus de animal) para su uso como vectores para la introducción y/o expresión de ácidos nucleicos en células de mamífero. En particular, la invención se refiere a mejoras en la productividad de tales vectores de parvovirus cuando se producen en células de insecto.

Los virus de la familia Parvoviridae son pequeños virus animales de ADN. La familia Parvoviridae puede dividirse en dos subfamilias: la Parvovirinae, que infectan a los vertebrados, y la Densovirinae, que infecta a insectos. Los miembros de la subfamilia Parvovirinae se denominan en el presente documento los parvovirus e incluyen el género Dependovirus. Como se deduce a partir del nombre de su género, los miembros de Dependovirus son únicos en cuanto a que, por lo general, requieren coinfección con un virus auxiliar, tal como adenovirus o virus herpes, para la infección productiva en cultivo celular. El género Dependovirus incluye VAA, que normalmente infecta a seres humanos (por ejemplo, los serotipos 1, 2, 3A, 3B, 4, 5, y 6) o primates (por ejemplo, los serotipos 1 y 4), y virus relacionados que infectan a otros animales de sangre caliente (por ejemplo, virus adenoasociados bovinos, caninos, equinos y ovinos). Información adicional sobre los parvovirus y otros miembros de la familia Parvoviridae se describe en Kenneth I. Berns, "Parvoviridae: The Viruses and Their Replication," Capítulo 69 en Fields Virology (3ª Ed. 1996). Por conveniencia, la presente invención se ilustra adicionalmente con ejemplos y se describe en el presente documento por referencia a VAA. Sin embargo, se entiende que la invención no se limita a VAA sino que puede aplicarse igualmente a otros parvovirus.

La organización genómica de todos los serotipos de VAA conocidos es muy similar. El genoma de los VAA es una molécula de ADN monocatenario lineal que tiene menos de aproximadamente 5.000 nucleótidos (nt) de longitud. Repeticiones terminales invertidas (ITR) flanquean las secuencias de nucleótidos de codificación únicas para las proteínas de replicación no estructurales (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 formando una horquilla en forma de T. Estas estructuras de horquilla funcionan como un origen para la replicación del ADN viral, sirviendo como cebadores para el complejo de la ADN polimerasa celular. Después de la infección de la VAA silvestre en células de mamífero, los genes de 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 acontecimiento de corte y empalme en el ORF de Rep da lugar a la expresión de, en realidad, cuatro proteínas Rep (es decir, Rep78, Rep68, Rep52 y

Rep40). Sin embargo, se ha demostrado que el ARNm sin cortar, que codifica las proteínas Rep78 y Rep52, en células de mamífero, son suficientes para la producción del vector VAA. También en las células de insecto, las proteínas Rep78 y Rep52 son suficientes para la producción del vector VAA.

Un "vector VAA o de parvovirus recombinante" (o "vector de VAAr") 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 secuencias de repetición de repetición terminales invertidas (ITR) de parvovirus o de VAA. Tales vectores de VAAr pueden replicarse y empaquetarse en partículas víricas infecciosas cuando están presentes en una célula huésped de insecto que se expresan productos génicos rep y cap de VAA (es decir, proteínas Rep y Cap de VAA). Cuando un vector de VAAr se incorpora en una construcción de ácido nucleico más grande (por ejemplo, en un cromosoma o en otro vector tal como un plásmido o baculovirus utilizado para la clonación o transfección), el vector de VAAr se denomina típicamente "provector" que puede "rescatarse" mediante replicación y encapsidación en presencia de funciones de empaquetamiento de VAA y funciones auxiliares necesarias.

En un primer aspecto, la invención se refiere a una célula de insecto que comprende una primera secuencia de nucleótidos que codifica una primera secuencia de aminoácidos y una segunda secuencia de nucleótidos que codifica una segunda secuencia de aminoácidos, en las que: (a) la primera y la segunda secuencia de nucleótidos están cada una unida de manera operativa a secuencias de control de la expresión para la expresión en una célula de insecto; (b) la primera y la segunda secuencias de aminoácidos comprenden una secuencia de aminoácidos común de la cual por lo menos 100 aminoácidos tienen una identidad de aminoácidos de al menos 90 % entre las secuencias primera y segunda de aminoácidos; (c) las secuencias de nucleótidos que codifican la secuencia de aminoácidos común en las primera y segunda secuencias de aminoácidos tienen menos del 90 % de identidades; (d) el uso de codones de la primera secuencia de nucleótidos que codifica la secuencia de aminoácidos común está adaptado para el uso de codones de la célula huésped de insectos; (e) la primera secuencia de nucleótidos codifica la secuencia de aminoácidos de una proteína Rep52 o 40 de parvovirus y la segunda secuencia de nucleótidos codifica la secuencia de aminoácidos de una proteína Rep78 o 68 de parvovirus; y (e) las secuencias de aminoácidos comunes comprenden las secuencias de aminoácidos del segundo aminoácido al aminoácido más C-terminal de la proteína Rep52 o 40 de parvovirus. La célula de insecto comprende al menos una primera secuencia de nucleótidos que codifica una primera secuencia de aminoácidos y una segunda secuencia de nucleótidos que codifica una segunda secuencia de aminoácidos. Preferentemente, las primera y segunda secuencias de aminoácidos comprenden cada una secuencia de aminoácidos común de la cual al menos 100, 200, 300, 350 o 398 aminoácidos tienen al menos 90, 95, 98, 99 o 100 % de identidad de aminoácidos entre la primera y la segunda secuencias de aminoácidos. Por el contrario, las secuencias de nucleótidos que codifican las secuencias de aminoácidos comunes en las primera y segunda secuencias de aminoácidos (como está presente en las primera y segunda secuencias de nucleótidos, respectivamente) tienen menos de 90, 89, 88,4, 85, 80, 75, 70, 65, 60, o 55 % de identidad.

Por lo general, las primera y segunda secuencias de aminoácidos serán heterólogas para la célula de insecto. Preferentemente, al menos una de las secuencias de aminoácidos comunes en las primera y segunda secuencias de aminoácidos es una secuencia de aminoácidos de origen natural. Más preferentemente, al menos una de las primera y segunda secuencias de aminoácidos son secuencias de aminoácidos de origen natural. Lo más preferentemente, las primera y segunda secuencias de aminoácidos son secuencias de aminoácidos de origen natural.

En una realización preferida, la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la primera secuencia de nucleótidos tiene un sesgo de uso de codones mejorado para la célula de insecto, en comparación con la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la segunda secuencia de nucleótidos, en la que la diferencia en el índice de adaptación de codones entre la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en el primero y segundo nucleótido es de al menos 0,2. Se entiende en el presente documento que cada vez que se haga referencia al sesgo del uso de codones para una célula de insecto, esto incluye el sesgo del uso de codones para una célula de insecto infectada con baculovirus, incluyendo, en particular, el sesgo de uso de codones para una célula infectada con el virus de la polihedrosis nuclear múltiple *Autographa californica* (AcMNPV). El uso de codones de la primera secuencia de nucleótidos que codifica la secuencia de aminoácidos común preferentemente está adaptado u optimizado al uso de los codones de la célula huésped de insecto. La capacidad de adaptación de una secuencia de nucleótidos que codifica la secuencia de aminoácidos común al uso de los codones de la célula huésped puede expresarse como el índice de adaptación de codones (IAC). Preferentemente, el uso de codones está adaptado a la célula de insecto en la que la primera y segunda secuencia de nucleótidos están presentes. Por lo general, esta será una célula del género *Spodoptera*. Más preferentemente una célula de *Spodoptera frugiperda*. Por tanto, el uso de codones está, preferentemente, adaptado para *Spodoptera frugiperda* o para una célula infectada por el virus de la polihedrosis nuclear múltiple *Autographa californica* (AcMNPV). En el presente documento, un índice de adaptación de codones se define como una medida de la capacidad de adaptación relativa del uso de codones de un gen hacia el uso de codones de genes altamente expresados. La capacidad de adaptación relativa (w) de cada codón es la relación entre el uso de cada codón con el del codón más abundante para el mismo aminoácido. El índice de IAC se define como la media geométrica de estos valores de capacidad de adaptación relativa. Se excluyen los codones no sinónimos y los codones de terminación (en función del código genético). Los valores de IAC varían de 0 a 1, indicando los valores más altos una mayor proporción de los codones más abundantes (véase Sharp y Li, 1987,

Nucleic Acids Research 15: 1281-1295; véase también: Kim y col., Gene. 1997, 199:293-301; zur Megede y col., Journal of Virology, 2000, 74: 2628-2635). En una célula de insecto preferente, la diferencia en el IAC entre la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la primera y segunda secuencias de nucleótidos es de al menos 0,3, 0,4, 0,5, 0,6, 0,7, 0,8. Preferentemente, además del IAC de la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la primera secuencia de nucleótidos es al menos 0,5, 0,6, 0,7, 0,8, 0,9 o 1,0.

Una secuencia de nucleótidos preferida que codifica la secuencia de aminoácidos común en la primera secuencia de nucleótidos es una secuencia de codificación en la que al menos 50, 75, 90, 95, 98 o 99 %, y preferentemente todos los codones no comunes o codones menos comunes están sustituidos por un codón común que codifica el mismo aminoácido como se indica en la Tabla 1 o en la Tabla 2. En el presente documento se tiende que un codón común es el codón más frecuentemente utilizado que codifica cada residuo de aminoácido particular en genes de *Spodoptera frugiperda* altamente expresados, como se muestra en la Tabla 1 o en células infectadas por MNPV con genes de *Autographa californica* altamente expresados, como se muestra en la Tabla 2. Todos los codones distintos de los codones comunes y de los codones menos comunes son "codones no comunes". Los codones no comunes incluyen los "segundos codones más frecuentes", entendidos como codones que tienen la frecuencia más alta en la Tabla 1 o la Tabla 2. Preferentemente, la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la primera secuencia de nucleótidos tiene un tramo continuo de al menos 25, 50, 100, 200 o 300 codones, todos los cuales son codones comunes. La secuencia de codificación puede adaptarse además para mejorar la expresión en la célula huésped de insecto por procedimientos descritos en el documento WO 2004/059556 y mediante la modificación del contenido de CpG de la secuencia de codificación como se describe en el documento WO 2006/015789. Se entiende que tales adaptaciones adicionales pueden hacer que no todos los codones en la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la primera secuencia de nucleótidos son codones comunes.

En una realización preferida de la célula de insecto, todos los codones en la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la primera secuencia de nucleótidos son codones comunes de conformidad con (una de ellas) las Tablas 1 o 2. Más preferentemente, en una célula de insecto de este tipo, todos los codones de la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la segunda secuencia de nucleótidos son segundos codones más frecuentes de conformidad con (cualquiera de ellas) las Tablas 1 o 2, de modo que se entiende que si en la primera secuencia de nucleótidos, los codones comunes están de acuerdo con la Tabla 1, los segundos codones más frecuentes en la segunda secuencia de nucleótidos también están de acuerdo con la Tabla 1, o que si en la primera secuencia de nucleótidos, los codones comunes están de acuerdo con la Tabla 2, los segundos codones más frecuentes en la segunda secuencia de nucleótidos también están de acuerdo con la Tabla 2.

La optimización de codones puede realizarse sobre la base del uso de codones del organismo *Spodoptera frugiperda* como se puede encontrar en una base de datos de uso de codones (véase, por ejemplo, <http://www.kazusa.or.jp/codon/>). Los programas de ordenador adecuados para la optimización de codones están disponibles para el experto en la materia (véase, por ejemplo Jayaraj y col., 2005, Nucl. Acids Res. 33(9):3011-3016; y en internet). Como alternativa, las optimizaciones se pueden hacer a mano, usando la misma base de datos de uso de codones.

En una realización de la célula de insecto de la invención, al menos 50, 60, 80, 90 o 100 % de los codones en la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la segunda secuencia de nucleótidos están alterados en comparación con el codón correspondiente de la primera secuencia de nucleótidos para maximizar el contenido de AT o GC de la segunda secuencia de nucleótidos.

Por lo tanto, en una realización preferida de la invención, la diferencia en la secuencia de nucleótidos entre la primera y segunda secuencia de nucleótidos que codifica las secuencias de aminoácidos comunes se maximiza (es decir, la identidad de nucleótidos se reduce al mínimo) mediante uno o más de: a) cambiar el sesgo de codones de la primera secuencia de nucleótidos que codifica la secuencia de aminoácidos común; b) cambiar el sesgo de codones de la segunda secuencia de nucleótidos que codifica la secuencia de aminoácidos común; c) cambiar el contenido de GC de la primera secuencia de nucleótidos que codifica la secuencia de aminoácidos común; y d) cambiar el contenido de GC de la segunda secuencia de nucleótidos que codifica la secuencia de aminoácidos común.

Una realización preferida de la invención de la célula de insecto se refiere a la producción de proteínas de parvovirus en las células de insecto de la invención. En particular, las proteínas de parvovirus se producen en las células de insecto en el contexto de la preparación de vectores de parvovirus recombinantes, más preferentemente vectores de parvovirus de animales recombinantes y, lo más preferentemente, vectores VAA recombinantes. Por lo tanto, en esta realización preferida de las células de insecto de la invención, la primera secuencia de nucleótidos codifica una secuencia de aminoácidos de una proteína Rep52 o 40 de parvovirus y la segunda secuencia de nucleótidos codifica una secuencia de aminoácidos de una proteína Rep78 o 68 de parvovirus. Se entiende, sin embargo, que las realizaciones en las que la primera secuencia de nucleótidos codifica una secuencia de aminoácidos de una proteína Rep78 o 68 de parvovirus y la segunda secuencia de nucleótidos codifica una secuencia de aminoácidos de una proteína Rep52 o proteína 40 de parvovirus están incluidas expresamente en la invención. Por conveniencia en las

realizaciones se utilizará la secuencia de nucleótidos que codifica una proteína Rep52 o 40 de parvovirus como primera secuencia de nucleótidos y la secuencia de nucleótidos que codifica una proteína Rep78 o 68 de parvovirus como segunda secuencia de nucleótidos, pero la inversa de estas realizaciones se incluye expresamente en la invención. La secuencia de aminoácidos común codificada por la primera y segunda secuencias de nucleótidos comprenden o consisten en las secuencias de aminoácidos de al menos el segundo aminoácido al aminoácido más C-terminal de una proteína Rep52 o 40 de parvovirus. Preferentemente, las secuencias de aminoácidos comunes comprenden o consisten en el primer aminoácido al aminoácido más C-terminal de la proteína Rep52 o 40 de parvovirus. Las identidades de aminoácidos entre las secuencias de aminoácidos comunes de parvovirus son como se han definido anteriormente para las secuencias de aminoácidos comunes. Preferentemente, en la célula de insecto, las proteínas Rep de parvovirus son proteínas Rep de virus adenoasociados (VAA). Más preferentemente, las proteínas Rep de parvovirus codificadas en la primera y segunda secuencias de nucleótidos son del mismo serotipo.

Una secuencia de nucleótidos que codifica proteínas Rep de parvovirus, 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 de parvovirus en las células de insectos, tales como las proteínas Rep78 o Rep68, y Rep52 o Rep40. La secuencia de nucleótidos de parvovirus animales es, preferentemente, de un dependovirus, más preferentemente de virus adenoasociado (VAA) de ser humano o de simio y, lo más preferentemente, de un VAA que normalmente infecta a los seres humanos (por ejemplo, los serotipos 1, 2, 3A, 3B, 4, 5 y 6) o primates (por ejemplo, los serotipos 1 y 4). Un ejemplo de una secuencia de nucleótidos que codifica las proteínas Rep de parvovirus animales se da en la SEQ ID No.7, que representa una parte del genoma de la secuencia de serotipo 2 de VAA que codifica las proteínas Rep. La secuencia de codificación de Rep78 comprende los nucleótidos 11 a 1876 y la secuencia de codificación de Rep52 comprende los nucleótidos 683 - 1876, también representados por separado en la SEQ ID No. 1 y 5. 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 sabrá cómo identificar la posición correspondiente en la secuencia de nucleótidos de otros parvovirus distintos de VAA-2.

Una (primera) secuencia de nucleótidos que codifica una proteína Rep52 de parvovirus puede, por tanto, definirse como una secuencia de nucleótidos:

- a) que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene al menos 50, 60, 70, 80, 88, 89, 90, 95, 97, 98, o 99 % de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO. 6;
- b) que tiene al menos 50, 60, 70, 80, 81, 82, 85, 90, 95, 97, 98, o 99 % de identidad de secuencia con la secuencia de nucleótidos de una cualquiera de las SEQ ID NO de 1 - 5 y 10;
- c) la cadena complementaria de la que se hibrida a una secuencia de molécula de ácido nucleico de (a) o (b);
- d) secuencias de nucleótidos cuya secuencia difiere de la secuencia de una molécula de ácido nucleico de (c) debido a la degeneración del código genético.

Una (segunda) secuencia de nucleótidos que codifica una proteína Rep78 de parvovirus puede, por tanto, definirse como una secuencia de nucleótidos:

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

Preferentemente, la secuencia de nucleótidos codifica las proteínas Rep de parvovirus humanos que son necesarias y suficientes para la producción de vectores de parvovirus en células de insecto.

Las diversas modificaciones de la primera y segunda secuencia de nucleótidos de codificación como se ha definido anteriormente, incluyendo, por ejemplo, las secuencias de parvovirus de tipo silvestre, para la expresión apropiada en células de insectos se consigue mediante la aplicación de técnicas de ingeniería genética bien conocidas, tales como se describe en, por ejemplo, Sambrook y Russell (2001) "Molecular Cloning: A Laboratory Manual, (3ª edición), Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory Press, New York. El experto en la técnica conoce varias modificaciones adicionales de las regiones de codificación que podrían aumentar el rendimiento de las proteínas que codifican. Estas modificaciones están dentro del ámbito de la presente invención.

En las células de insecto de la invención. La primera y segunda secuencias de nucleótidos son, preferentemente, parte de una construcción de ácido nucleico. La célula de insecto puede comprender dos construcciones de ácido nucleico distintas, una para cada una de las primera y segunda secuencias de nucleótidos, o de la célula de insecto puede comprender un solo tipo de construcción de ácido nucleico que comprende tanto la primera como la segunda secuencias de nucleótidos.

En un aspecto adicional, la invención se refiere a una construcción de ácido nucleico que comprende una primera y/o una segunda secuencia de nucleótidos que codifica para una primera y una segunda secuencia de aminoácidos, respectivamente, que comprenden una secuencia de aminoácidos común como se ha definido anteriormente. Preferentemente, las primera y/o segunda secuencias de nucleótidos en la construcción codifican proteínas Rep de parvovirus como se han definido anteriormente. Preferentemente, en la construcción, la secuencia de nucleótidos que codifica la primera y la segunda secuencias de aminoácidos están unidas operativamente a secuencias de control de la expresión para la expresión en una célula de insecto. Estas secuencias de control de la expresión incluirán, como mínimo, un promotor que es activo en las células de insecto. Las técnicas conocidas para un experto en la materia para expresar genes extraños en células huésped de insecto se pueden utilizar para practicar la invención. La metodología para la ingeniería molecular y la expresión de polipéptidos en células de insecto 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. No. 7555, College Station, Tex.; Luckow. 1991. En Prokop y col., 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, New York; W. H. Freeman y Richardson, C. D., 1995, Baculovirus Expression Protocols, Methods in Molecular Biology, volumen 39; documento US 4,745,051; documento US2003148506; y documento WO 03/074714. Los promotores adecuados para la transcripción de la primera y la segunda secuencias de nucleótidos de la invención incluyen, por ejemplo, los promotores de poliedro (pool), p10, p35, IE-1 o ΔIE-1 y otros promotores descritos en las referencias anteriores. Puesto que se sabe que en células de mamífero una expresión menos abundante de Rep78 en comparación con Rep52 favorece altos rendimientos de vectores (Li y col., 1997, J Virol. 71: 5236-43; Grimm y col., 1998, Hum Gene Ther. 9, 2745-2760), preferentemente se usa un promotor más débil para dirigir la expresión de la proteína Rep78 o 68 que el promotor utilizado para la expresión de la proteína Rep52 o 40. Por ejemplo, se puede usar el promotor de poliedro más fuerte para la expresión de la proteína Rep52 o 40, el promotor ΔIE1, un promotor mucho más débil que el promotor Polh, puede elegirse para dirigir la expresión de la proteína Rep78 o 68. Preferentemente, la elección de promotores para la proteína Rep52 o 40 y la proteína Rep78 o 68, respectivamente, es tal que en una célula de insecto con el fin de producir en la célula de insecto una relación molar de Rep78/68 a Rep52/40 en el intervalo de 1:10 a 10:1, 1:5 k.o. 5:1, o 1:3 a 3:1, preferentemente a aproximadamente 20 - 40 horas después de la infección, más preferentemente a aproximadamente 30 - 40 horas después de la infección, utilizando una expresión en baculovirus. La relación molar de Rep78 y Rep52 se puede determinar por medio de transferencia de tipo Western, utilizando preferentemente un anticuerpo monoclonal que reconoce un epítipo común tanto de Rep78/68 como de Rep52/40, o el uso de, por ejemplo, un anticuerpo anti-Rep de ratón (303.9, Progen, Alemania; dilución 1:50).

Preferentemente, la construcción de ácido nucleico para la expresión de la primera y la segunda secuencias de nucleótidos de la invención en células de insecto es un vector compatible con células de insectos. Un "vector compatible con células de insecto" o "vector" se entiende que es una molécula de ácido nucleico capaz de transformación o transfección productiva de un insecto o célula de insecto. Los vectores biológicos de ejemplo incluyen plásmidos, moléculas de ácido nucleico lineales y virus recombinantes. Se puede usar cualquier vector siempre que sea compatible con las células de insecto. El vector puede integrarse en el genoma de las células de insecto, pero el vector también puede ser episomal. La presencia del vector en la célula de insecto no tiene que permanente y también se incluyen vectores episomales transitorios. Los vectores pueden introducirse por cualquier medio conocido, por ejemplo mediante 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, la construcción es un vector de baculovirus. Los vectores de baculovirus y los procedimientos para su uso se describen en las referencias citadas anteriormente en la ingeniería molecular de células de insecto.

Las construcciones de ácido nucleico de la invención pueden comprender además una secuencia de control de la expresión que comprende una secuencia de nueve nucleótidos de la SEQ. ID NO: 9 o una secuencia de nucleótidos sustancialmente homóloga a la SEQ. ID NO: 9, en dirección 5' de los codones de iniciación de la secuencia de nucleótidos que codifica la primera y/o la segunda secuencias de aminoácidos. Una secuencia con una identidad sustancial con la secuencia de nucleótidos de SEQ. ID NO: 9 y que ayudará a aumentar la expresión de la primera y/o las secuencias de aminoácidos es, por ejemplo, una secuencia que tiene al menos 60 %, 70 %, 80 % o 90 % de identidad con la secuencia de nueve nucleótidos de de SEQ ID NO: 9.

La célula de insecto puede ser cualquier célula que sea adecuada para la producción de proteínas heterólogas. Preferentemente, la célula de insecto permite la replicación de los vectores de baculovirus y se puede mantener en cultivo. Más preferentemente, la célula de insecto también permite la replicación de los vectores de parvovirus recombinantes, incluyendo los vectores de VAAr. Por ejemplo, la línea celular utilizada puede ser las líneas celulares de *Spodoptera frugiperda*, *Drosophila* o las líneas celulares de mosquitos, por ejemplo, las líneas celulares de *Aedes albopictus*. Las células de insecto o líneas de células preferentes son células de la especie de insecto que son susceptibles a la infección por baculovirus, incluyendo, por ejemplo, Se301, SeIZD2109, SeUCR1, Sf9, Sf900+, Sf21, BTI-TN-5B1-4, MG-1, Tn368, HzAm1 Ha2302, Hz2E5, High Five (Invitrogen, CA, USA) y expresSF+® (documento US 6.103.526; Protein Sciences Corp., CT, Estados Unidos).



Una célula de insecto preferida de acuerdo con la invención es una célula de insecto para la producción de vectores de parvovirus recombinantes. Esta célula de insecto comprende adicionalmente, además de la "primera" y la "segunda" secuencias de nucleótidos o la primera y la segunda secuencias de nucleótidos de las construcciones de ácido nucleico:

- 5 a) una tercera secuencia de nucleótidos que comprende al menos una secuencia de nucleótidos de repetición terminal invertida (secuencia ITR) de parvovirus; y
- b) una cuarta secuencia de nucleótidos que comprende las secuencias de codificación de la proteína Cap de parvovirus unida operativamente a secuencias de control de la expresión para la expresión en una célula de insecto.

10 En el contexto de la invención se entiende que "al menos una secuencia de nucleótidos ITR de parvovirus" significa una secuencia palindrómica, que comprende secuencias principalmente complementarias y dispuestas simétricamente, también denominada 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 proteínas de replicación que actúan en trans, tales como, por ejemplo, Rep 78 (o Rep68) que reconocen el palíndromo y secuencias específicas internas del palíndromo. Una excepción a la simetría de la secuencia ITR es la región "D" de la ITR. Es única (no tiene una complementaria dentro de una ITR). La realización de muescas en el ADN monocatenario se produce en la unión entre las regiones A y D. Es la región en la que se inicia una nueva síntesis de ADN. La región D normalmente se encuentra a un lado del palíndromo y proporciona direccionalidad de la etapa de replicación del ácido nucleico. Un parvovirus que se replica en una célula de mamífero tiene típicamente dos secuencias ITR. Sin embargo, es posible diseñar una ITR de modo que los sitios de unión estén en ambas hebras de las regiones A y las regiones D están situadas simétricamente, una a cada lado del palíndromo. En un molde de ADN circular de doble cadena (por ejemplo, un plásmido), la replicación de ácidos nucleicos asistida por Rep78 o Rep68 procede en ambas direcciones y una sola ITR es suficiente para la replicación del parvovirus de un vector circular. Por lo tanto, una secuencia de nucleótidos ITR se puede utilizar en el contexto de la presente invención. Preferentemente, sin embargo, se utilizan dos u otro número par de ITR regulares. Lo más preferentemente, se utilizan dos secuencias ITR. Una ITR de parvovirus preferente es una ITR de VAA. Por razones de seguridad, puede ser deseable construir un vector de parvovirus recombinante (VAAr) que no es capaz de propagarse adicionalmente después de la introducción inicial en una célula en presencia de un segundo VAA. Tal mecanismo de seguridad para limitar la propagación de vectores indeseable en un receptor puede proporcionarse mediante el uso de VAAr con una ITR quimérica como se describe en el documento US2003148506.

El número de construcciones de ácido nucleico empleadas en la célula de insecto para la producción del vector de parvovirus recombinante (VAAr) no es limitante en la invención. Por ejemplo, uno, dos, tres, cuatro, cinco, o más construcciones distintas se pueden emplear para producir VAAr en células de insecto de acuerdo con los procedimientos de la presente invención. Si se emplean cinco construcciones, una construcción codifica VP 1 de VAA, otra construcción codifica VP2 de VAA, otra construcción más codifica VP3 de VAA, todavía otra construcción más codifica la proteína Rep como se ha definido anteriormente y una construcción final comprende al menos una ITR de VAA. Si se utilizan menos de cinco construcciones, las construcciones pueden comprender varias combinaciones de la al menos una ITR de VAA y las secuencias de codificación de VP1, VP2, VP3, y la proteína Rep.

40 Preferentemente, se utilizan dos, tres o cuatro construcciones. Si se utilizan dos construcciones, preferentemente la célula de insecto comprende: (A) una primera construcción de ácido nucleico para la expresión de las proteínas Rep como se ha definido anteriormente, construcción que comprende además la cuarta secuencia de nucleótidos como se define en (b) anteriormente (que comprende las secuencias de codificación de la proteína Cap de parvovirus unidas operativamente a al menos una secuencia de control de la expresión para la expresión en una célula de insecto; véase también más adelante); y (B) una tercera construcción de ácido nucleico que comprende la tercera secuencia de nucleótidos como se ha definido en (a) anterior (que comprende al menos una secuencia de nucleótidos de ITR de parvovirus/VAA). Si se utilizan tres construcciones, se usa, preferentemente, la misma configuración que se usa para dos construcciones, excepto porque se usan construcciones separadas para la expresión de las proteínas de la cápside y para la expresión de las proteínas Rep. Si se utilizan cuatro construcciones, se usa, preferentemente, la misma configuración que se usa para tres construcciones, excepto porque se usan construcciones separadas para la expresión de las proteínas Rep78/68 y para la expresión de las proteínas Rep 52/40. Las secuencias de cada construcción pueden estar en cualquier orden una respecto a la otra. Por ejemplo, si una construcción comprende ITR y un ORF que comprende secuencias de nucleótidos que codifican proteínas de la cápside VP, el ORF de VP puede estar ubicado en la construcción de manera que, tras la replicación del ADN entre las secuencias ITR, el ORF de VP está replicado o no replicado. Para otro ejemplo, las secuencias de codificación y/o el ORF de Rep que comprenden las secuencias de nucleótidos que codifican proteínas de la cápside VP pueden estar en cualquier orden en una construcción. Se entiende que también las construcciones de ácido nucleico tercera, cuarta y más son, preferentemente, vectores compatibles con una célula de insecto, preferentemente vectores de baculovirus como se ha descrito anteriormente. Como alternativa, en la célula de insecto de la invención, una o más de la primera secuencia de nucleótidos, la tercera secuencia de nucleótido, la cuarta secuencia de nucleótidos, y la quinta secuencia de nucleótidos y otras secuencias de nucleótidos opcionales pueden estar integradas de forma estable en el genoma de la célula de insecto. Un experto en la técnica sabe cómo introducir de manera estable una secuencia de nucleótidos en el genoma de insecto y cómo identificar una célula

que tiene dicha secuencia de nucleótidos en el genoma. La incorporación en el genoma puede añadirse mediante, por ejemplo, el uso de un vector que comprende las secuencias de nucleótidos altamente homólogas a regiones del genoma de insecto. El uso de secuencias específicas, tales como transposones, es otra manera de introducir una secuencia de nucleótidos en un genoma.

5 En la invención, en el presente documento se entiende la cuarta secuencia de nucleótidos que comprende las secuencias de codificación de la proteína de la cápside (Cap) de parvovirus comprende secuencias que codifican cada una de las tres proteínas de la cápside de parvovirus, VP1, -2 y -3. La cuarta secuencia de nucleótidos que comprende las secuencias de codificación de la proteína de la cápside puede estar presente en diversas formas. Por ejemplo, se pueden usar secuencias de codificación distintas para cada una de las proteínas de la cápside VP1, -2 y  
10 -3, de modo que cada secuencia de codificación está unida operativamente a las secuencias de control de la expresión para la expresión en una célula de insecto. Más preferentemente, sin embargo, la cuarta secuencia de nucleótidos comprende un único marco de lectura abierto que codifica las tres proteínas de la cápside de parvovirus animal (VAA) VP1, VP2 y VP3, en la que el codón de iniciación para la traducción de la proteína de la cápside VP1 es un codón de iniciación subóptimo que no es ATG como se describe, por ejemplo, en Urabe y col., (2002, mencionado anteriormente) y en el documento WO2007/046703. El codón de iniciación subóptimo para la proteína de la cápside VP1 se puede seleccionar de ACG, TTG, CTG y GTG, de los cuales son los más preferidos CTG y GTG. La cuarta secuencia de nucleótidos para la expresión de las proteínas de la cápside puede comprender además una o más modificaciones como se describe en el documento WO2007/046703.

El experto en la materia conoce varias modificaciones adicionales de las regiones de codificación de VP, que podrían aumentar el rendimiento de VP y el virión o tener otros efectos deseados, tales como tropismo alterado o reducir la antigenicidad del virión. Estas modificaciones están dentro del ámbito de la presente invención. Preferentemente, la secuencia de nucleótidos de la invención que codifica las proteínas de la cápside de parvovirus está unida operativamente a secuencias de control de la expresión para la expresión en una célula de insecto, que, como mínimo, incluirá un promotor que es activo en las células de insecto. Tales secuencias de control y otras técnicas y materiales (por ejemplo, vectores) para la expresión de proteínas de la cápside de parvovirus en las células huésped de insectos ya se han descrito anteriormente para las proteínas Rep.

En una realización preferida de la invención, la tercera secuencia de nucleótidos presente en las células de insecto de la invención, es decir, la secuencia que comprende al menos una ITR de parvovirus (VAA), comprende además al menos una secuencia de nucleótidos que codifica un producto génico de interés, de modo que, preferentemente, la  
30 al menos una secuencia de nucleótidos que codifica un producto génico de interés se incorpora en el genoma de un vector de parvovirus recombinante (VAAR) producido en la célula de insecto. Preferentemente, al menos una secuencia de nucleótidos que codifica un producto génico de interés es una secuencia para la expresión en una célula de mamífero. Preferentemente, la tercera secuencia de nucleótidos comprende dos secuencias de nucleótidos ITR de parvovirus (VAA) y en la que la al menos una secuencia de nucleótidos que codifica un producto génico de interés se encuentra entre las dos secuencias de nucleótidos ITR de parvovirus (VAA). Preferentemente, la secuencia de nucleótidos que codifica un producto génico de interés (para la expresión en la célula de mamífero) se incorporará en el vector de parvovirus recombinante (VAAR) producido en la célula de insecto si se encuentra entre dos ITR regulares o se encuentra en cualquiera de los lados de una ITR diseñada con dos regiones D.

La tercera secuencia de nucleótidos definida anteriormente en el presente documento puede, por lo tanto, comprender una secuencia nucleotídica que codifica al menos un "producto génico de interés" para la expresión en una célula de mamífero, que se encuentra de tal manera que se incorporará en un vector de parvovirus recombinante (VAAR) replicado en la célula de insecto. Cualquier secuencia de nucleótidos se puede incorporar para su expresión posterior en una célula de mamífero transfectada con el vector de parvovirus recombinante (VAAR) producido de acuerdo con la presente invención. La secuencia de nucleótidos puede, por ejemplo, codificar una proteína que puede expresar un agente de ARNi, es decir, una molécula de ARN que es capaz de ARN de interferencia, tal como, por ejemplo, un ARNch (ARN corto en horquilla) o un ARNic (ARN corto de interferencia). "ARNpi" significa un ARN pequeño de interferencia que es un ARN bicatenario de longitud corta que no son tóxicos en las células de mamífero (Elbashir y col., 2001, Nature 411: 494-98; Caplen y col., 2001, Proc. Natl. Acad. Sci. USA 98: 9742-47). En una realización preferida, la tercera secuencia de nucleótidos puede comprender dos  
40 secuencias de nucleótidos y cada una 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 encuentra localizada de tal manera que se incorporará en un vector de parvovirus recombinante (VAAR) replicado en la célula de insecto.

El producto de interés para la expresión en una célula de mamífero puede ser un producto génico terapéutico. Un  
55 producto génico terapéutico puede ser un polipéptido, o una molécula de ARN (ARNpi), u otro producto génico que, cuando se expresa en una célula diana, proporciona un efecto terapéutico deseado, tal como, por ejemplo, ablación de una actividad no deseada, por ejemplo, la ablación de una célula infectada, o la complementación de un defecto genético, por ejemplo, provocando una deficiencia en una actividad enzimática. Ejemplos de productos génicos de polipéptidos terapéuticos incluyen CFTR, factor IX, lipoproteína lipasa (LPL, preferentemente LPL S447X; véase el documento WO 01/00220), apolipoproteína A1, uridina difosfato glucuronosiltransferasa (UGT), proteína de interacción reguladora de la GTPasa en la retinitis pigmentosa (RP-GRIP) y citocinas o interleucinas, como por ejemplo, IL-10, porfobilinógeno desaminasa (PBGD) y alanina: glioxilato aminotransferasa.

Como alternativa, o además, como tercer producto génico, la tercera secuencia de nucleótidos definida anteriormente en el presente documento puede comprender una secuencia de nucleótidos que codifica un polipéptido que sirve como proteínas marcadoras para evaluar la transformación y la expresión celular. Las proteínas marcadoras adecuadas para este fin son, por ejemplo, la proteína GFP fluorescente y los genes marcadores seleccionables HSV timidina cinasa (para la selección en medio HAT), higromicina B fosfotransferasa bacteriana (para la selección en higromicina B), Tn5 aminoglucósido fosfotransferasa (para la selección en G418) y la dihidrofolato reductasa (DHFR) (para la selección en metotrexato), CD20, el gen del factor de crecimiento nervioso de baja afinidad. Las fuentes para la obtención de estos genes marcadores y procedimientos para su uso se proporcionan en Sambrook y Russel (2001) "Molecular Cloning: A Laboratory Manual, (3ª edición), Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory Press, New York. Además, la tercera secuencia de nucleótidos definida anteriormente en el presente documento puede comprender una secuencia de nucleótidos que codifica un polipéptido que puede servir como un mecanismo de seguridad que permite curar a un sujeto de las células transducidas con el vector de parvovirus recombinante (VAAr) de la invención, si se considera necesario. Tal secuencia de nucleótidos, a menudo referido como un gen suicida, codifica una proteína que es capaz de convertir un profármaco en una sustancia tóxica que es capaz de matar las células transgénicas en las que se expresa la proteína. Los ejemplos adecuados de tales genes suicidas incluyen, por ejemplo el gen de la citosina desaminasa de *E. coli* o uno de los genes de la timidina cinasa del virus del herpes simple, citomegalovirus y el virus de varicela zoster, en cuyo caso se puede usar ganciclovir como profármaco para matar las células transgénicas en el sujeto (véase, por ejemplo, Clair y col., 1987, Antimicrob. Agents Chemother. 31: 844-849).

En otra realización, uno de los productos génicos de interés puede ser una proteína de VAA. En particular, una proteína Rep, tal como Rep78 o Rep68, o un fragmento funcional de la misma. Una secuencia de nucleótidos que codifica una Rep78 y/o Rep68, si está presente en el genoma de un vector de parvovirus recombinante (VAAr) de la invención y se expresa en una célula de mamífero transducida con el vector, permite la integración del vector de parvovirus recombinante (VAAr) 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 VAAr puede proporcionar una ventaja para ciertos usos del vector de parvovirus recombinante (VAAr), permitiendo la expresión a largo plazo o permanente de cualquier otro producto génico de interés introducido en la célula por el vector.

En los vectores de parvovirus recombinante (VAAr) de la invención, la al menos una o más secuencia de nucleótidos que codifica un producto génico de interés para la expresión en una célula de mamífero, preferentemente está/están unido operativamente a al menos una secuencia de control de la expresión compatibles con células de mamífero, por ejemplo un promotor. Muchos de estos promotores son conocidos en la técnica (véase, Sambrook y Russel, 2001, citado anteriormente). Pueden usarse los promotores constitutivos que se expresan ampliamente en muchos tipos de células, tales como el promotor de CMV. Sin embargo, son más preferentes los promotores que son inducibles, específicos de tejido, específicos de tipo de célula o específicos del ciclo celular. Por ejemplo, para la expresión específica en el hígado, se puede seleccionar un promotor a partir de un promotor de  $\alpha$ 1-antitripsina (AAT), un promotor de globulina de unión a hormonas tiroideas, un promotor de albúmina, un promotor de LPS (globina de unión a la tiroxina), un promotor híbrido de HCR- promotor ApoCII, un promotor híbrido HCR-hAAT, un promotor de AAT en combinación con el elemento potenciador del gen de ka albúmina de ratón (Ealb) y un promotor de la apolipoproteína E. Otros ejemplos incluyen el promotor E2F Lara la expresión selectiva de tumor, y, en particular, expresión selectiva de células tumorales neurológicas (Parr y col., 1997, Nat. Medicina. 3: 1145-9) o el promotor de IL-2 para su uso en células sanguíneas mononucleares (Hagenbaugh y col., 1997, J Exp Med.; 185: 2101-10).

El VAA es capaz de infectar una serie de células de mamífero. Véase, por ejemplo, Tratschin y col., (1985, Mol. Cell Biol. 5:3251-3260) y Grimm y col., (1999, Hum. Gene Ther. 10:2445-2450). Sin embargo, la transducción de VAA de fibroblastos sinoviales humanos es significativamente más eficiente que en las células murinas similares, Jennings y col., Arthritis Res., 3: 1 (2001), y la tropicidad celular del VAA se diferencia entre los serotipos. Véase, por ejemplo, Davidson y col., (2000, Proc. Natl. Acad. Sci. USA, 97:3428-3432), que analizan las diferencias entre VAA2, VAA4 y VAA5 con respecto al tropismo celular del SNC de mamífero y la eficiencia de la transducción.

Las secuencias de VAA que se pueden utilizar en la presente invención para la producción de vectores VAA recombinantes en células de insectos se pueden derivar del genoma de cualquier serotipo de VAA. En general, los serotipos de VAA tienen secuencias genómicas de homología significativa a nivel de aminoácidos y de ácido nucleico proporcionan un conjunto idéntico de funciones genéticas, producen viriones que son esencialmente físicamente y funcionalmente equivalentes y se replican y ensamblan mediante mecanismos prácticamente idénticos. Para la secuencia genómica de los diversos serotipos de VAA y una visión general de las similitudes genómicas véase, por ejemplo número de acceso en GenBank U89790; número de acceso en GenBank J01901; número de acceso en GenBank AF043303; número de acceso en GenBank AF085716; Chlorini y col., (1997, J. Vir. 71: 6823-33); Srivastava y col., (1983, J. Vir. 45:555-64); Chlorini y col., (1999, J. Vir. 73:1309-1319); Rutledge y col., (1998, J. Vir. 72:309-319); y Wu y col., (2000, J. Vir. 74: 8635-47). Los serotipos 1, 2, 3, 4 y 5 de VAA son una fuente preferente de secuencias de nucleótidos de VAA para su uso en el contexto de la presente invención. Preferentemente, las secuencias ITR de VAA para su uso en el contexto de la presente invención derivan de VAA1, VAA2, y/o VAA4. Del mismo modo, las secuencias de codificación de Rep (Rep78/68 y Rep52/40) derivan, preferentemente, de VAA1, VAA2, y/o VAA4. Las secuencias de codificación de las proteínas de la cápside VP1, VP2, y VP3 para su uso en el contexto de la presente invención pueden, no obstante, tomarse de cualquiera de los

42 serotipos conocidos, más preferentemente de VAA1, VAA2, VAA3, VAA4, VAA5, VAA6, VAA7, VAA8 o VAA9 o partículas de tipo VAA recién desarrolladas obtenidas mediante, por ejemplo, técnicas de barajado de la cápside y bibliotecas de la cápside de VAA.

5 Las secuencias de ITR y Rep de VAA y están particularmente conservadas entre la mayoría de los serotipos. Las proteínas Rep78 de diversos serotipos de VAA son, por ejemplo, más del 89 % idéntica y la secuencia de nucleótidos total a nivel del genoma entre VAA2, VAA3A, VAA3B, y VAA6 es de aproximadamente el 82 % (Bantel-Schaal y col., 1999, J. Virol., 73(2):939-947). Por otra parte, se sabe que las secuencias de Rep e ITR de muchos serotipos de VAA se complementan de forma cruzada con eficiencia (es decir, sustituyen funcionalmente sustituto) las secuencias correspondientes de otros serotipos en la producción de partículas de VAA en las células de mamífero. El documento US2003148506 notifica que las secuencias de Rep e ITR del VAA también complementan de forma cruzada con eficiencia otras secuencias de VAA de Rep e ITR en células de insecto.

10 Se sabe que las proteínas VP de VAA determinan la tropicidad celular del virión del VAA. Las secuencias que codifican las proteínas VP están significativamente menos conservadas que las proteínas y los genes Rep entre diferentes serotipos de VAA. La capacidad de las secuencias Rep e ITR complementan de forma cruzada a las correspondientes secuencias de otros serotipos permite la producción de partículas de VAAr seudotipadas que comprenden las proteínas de la cápside de un serotipo (por ejemplo, VAA3) y las secuencias de Rep y/o ITR de otro serotipo de VAA (por ejemplo, VAA2). Tales partículas de VAAr seudotipadas son una parte de la presente invención.

15 Las secuencias de "VAA" modificadas también se pueden utilizar en el contexto de la presente invención, por ejemplo, para la producción de vectores de VAAr en células de insecto. Tales secuencias modificadas incluyen, por ejemplo, secuencias que tienen al menos aproximadamente 70 %, al menos aproximadamente 75 %, al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 90 %, al menos aproximadamente 95 %, o más de identidad de secuencia de aminoácidos de y/o nucleótidos (por ejemplo, una secuencia que tiene de aproximadamente 75 a 99 % de identidad de la secuencia de nucleótidos) con una VP, Rep o ITR de VAA1, VAA2, VAA3, VAA4, VAA5, VAA6, VAA7, VAA8 o VAA9 se pueden utilizar en lugar de secuencias salvajes de VP, Rep o ITR del VAA.

20 Aunque es similar a otros serotipos de VAA en muchos aspectos, el VAA5 difiere de otros serotipos de VAA humanos y de simio más de otros serotipos conocidos humanos y de simio. En vista de ello, la producción de VAA5 puede diferir de la producción de otros serotipos en células de insecto. Cuando se emplean procedimientos de la invención para producir VAA5, se prefiere que una o más construcciones que comprende, de manera colectiva en el caso de más de una construcción, una secuencia de nucleótidos que comprende una ITR de VAA5, una secuencia de nucleótidos comprende una secuencia de codificación de Rep de VAA5 (es decir, una secuencia de nucleótidos comprende una Rep78 de VAA5). Tales secuencias de ITR y Rep pueden modificarse como se desee para obtener una producción eficiente de VAA5r o vectores de VAA5r seudotipados en células de insecto. Por ejemplo, el codón de iniciación de las secuencias de Rep puede modificarse, los sitios de corte y empalme de VP pueden modificarse o eliminarse y/o el codón de iniciación de VP1 y los nucleótidos cercanos pueden modificarse para mejorar la producción de vectores VAA5r en la célula de insecto.

25 En otro aspecto, la invención se refiere a un procedimiento para producir un virión de parvovirus recombinante (VAAr) (que comprende un vector de parvovirus recombinante (VAAr) como se ha definido anteriormente) en una célula de insecto, comprendiendo el procedimiento las etapas de: a) cultivar una célula de insecto como se ha definido en el presente documento anteriormente en condiciones tales que se produce el vector de parvovirus recombinante (VAAr); y (b) recuperación del vector de parvovirus recombinante (VAAr). En el presente documento se entiende que el vector de parvovirus recombinante (VAAr) producido en el procedimiento es, preferentemente, un virión infeccioso de VAA o de parvovirus que comprende los ácidos nucleicos del vector de parvovirus recombinante (VAAr). Las condiciones de crecimiento para células de insecto en cultivo y producción de productos heterólogos en células de insecto en cultivo son bien conocidas en la técnica y se describen, por ejemplo, en las referencias citadas anteriormente sobre la ingeniería molecular de las células de insecto (véase también el documento WO2007/046703).

30 Preferentemente, el procedimiento comprende además la etapa de purificación por afinidad del vector de parvovirus recombinante (VAAr) (los viriones que lo comprenden) utilizando un anticuerpo anti-VAA, preferentemente un anticuerpo inmovilizado. El anticuerpo anti-VAA es, preferentemente, un anticuerpo monoclonal. Un anticuerpo particularmente adecuado es un anticuerpo de camélido monocatenario o un fragmento del mismo como, por ejemplo, se puede obtener a partir de camellos o llamas (véase, por ejemplo Muyldermans, 2001, Biotechnol. 74: 277-302). El anticuerpo para la purificación por afinidad del VAAr es, preferentemente, un anticuerpo que se une específicamente a un epítipo en una proteína de la cápside de VAA, de modo que, preferentemente, el epítipo es un epítipo que está presente en la proteína de la cápside de más de un serotipo de VAA. Por ejemplo, el anticuerpo puede producirse o seleccionarse sobre la base de la unión específica a la cápside de VAA2, pero, al mismo tiempo, también puede unirse específicamente a las cápsides de VAA1, VAA3 y VAA5.

35 En aún otro aspecto, la invención se refiere una construcción de ácido nucleico que comprende una primera y una segunda secuencias de nucleótidos como se ha definido anteriormente en el presente documento.

Se divulga un procedimiento para la producción de un virión de parvovirus recombinante (VAAr) (que comprende un vector de parvovirus recombinante (VAAr) como se ha definido anteriormente) en una célula de insecto. Preferentemente, el procedimiento comprende las etapas de: (a) cultivar una célula de insecto tal como se ha definido en el presente documento anteriormente en condiciones tales que se produce el vector de parvovirus recombinante (VAAr), en el que la célula de insecto comprende al menos una construcción de ácido nucleico para la expresión de las proteínas Rep78/68 y Rep52/40 de parvovirus (tales como, por ejemplo, una construcción de ácido nucleico que comprende la primera y la segunda secuencias de nucleótidos como se ha definido en el presente documento anteriormente) y comprende además una tercera y cuarta secuencias de nucleótidos como se ha definido en el presente documento anteriormente, y en el que la construcción o construcciones de ácido nucleico para la expresión de las proteínas Rep78/68 and Rep52/40 de parvovirus produce un nivel de expresión de Rep52/40 en la célula de insecto que es más alto que el nivel de expresión de Rep78/68 en una base molar; y, (b) recuperación del vector de parvovirus recombinante (VAAr). Preferentemente, en el procedimiento, la relación molar entre las proteínas Rep52/40 y Rep78/68 en la célula de insecto es mayor que 10:1, preferentemente al menos 11:1, 15:1, 20:1, 30:1, 40:1, 50:1 o 60:1. Una relación molar de las proteínas Rep52/40 y Rep78/68 en la célula de insecto mayor que 10:1 da como resultado ventajosamente una mejor relación entre viriones completos (es decir, que comprenden un genoma de VAAr) y viriones vacíos (véase, por ejemplo, la Figura 8). Sin embargo, una relación molar entre las proteínas Rep52/40 y Rep78/68 demasiado alta puede dar lugar a un título más bajo del VAAr producido tal como se determina por el número de copias de genes. En una realización, por tanto, la relación molar entre las proteínas Rep52/40 y Rep78/68 es menor que 100:1, 80:1, 70:1, 60:1, 50:1, 40:1, 30:1, o 20:1. La relación molar entre las proteínas Rep78/68 y Rep52/40 se puede determinar por medio de transferencia Western como se describe en el documento WO2007/148971, preferentemente usando un anticuerpo monoclonal que reconoce un epítipo común tanto de Rep78/68 como de Rep52/40, o utilizando el anticuerpo descrito en el documento WO2007/148971. Preferentemente, la relación molar mínima entre las proteínas Rep52/40 y Rep78/68, como se ha indicado anteriormente se consigue aproximadamente a las 20 - 40 horas después de la infección, más preferentemente aproximadamente 30 - 40 horas después de la infección, utilizando un baculovirus o un sistema de expresión similar.

Existen varios medios para aumentar el nivel de expresión relativa de las proteínas Rep52/40 en comparación con el de la proteína Rep78/68. En caso de que se use una unidad de transcripción única para la expresión de las proteínas Rep78/68 y Rep52/40, la secuencia de codificación de las proteínas Rep78/68 y Rep52/40 puede adaptarse del siguiente modo para obtener una relación molar de las proteínas Rep52/40 y Rep78/68 en la célula de insecto mayor que 10:1:

- a) el codón de iniciación de la traducción de la proteína Rep78/68 se puede cambiar a un codón de iniciación subóptimo y/o contexto subóptimo del mismo como se describe en el documento WO2007/148971;
- b) eliminación de uno o más o todas (9) las secuencias ATG que se producen entre los inicios de la traducción de los genes de Rep78/68 y Rep 52/40, respectivamente, preferentemente mediante cambios en la isocodificación en la secuencia de nucleótidos. Este se consigue, por ejemplo, en la secuencia de codificación pVD189 Rep descrita en el Ejemplo 3 y en la SEQ ID NO: 11;
- c) optimización del contexto del codón de iniciación de la traducción de la proteína Rep52/40 de acuerdo con el contexto iniciador óptimo de 5'-N N N N N A U G A a/c/g N-3' para la iniciación eficiente de la traducción en células de lepidópteros (como se describe en Chang y col., 1999, *Virology* 259: 369-383);
- d) mediante la incorporación de una secuencia de control de la expresión que comprende una secuencia de nueve nucleótidos de la SEQ. ID NO: 9 o una secuencia de nucleótidos sustancialmente homóloga a la SEQ. ID NO: 9, en dirección 5' de los codones de iniciación de la proteína Rep52/40;
- e) mejorar el sesgo del uso de codones de la parte de la secuencia de codificación que codifica la proteína Rep52/40 para la expresión en células de insecto (como se ha descrito anteriormente); y,
- f) cambiar el uso de codones de la parte de la secuencia de codificación entre los inicios de la traducción de las proteínas Rep78/68 y Rep 52/40 de modo que esté menos adaptado a la expresión en células de insecto (como se ha descrito anteriormente). La combinación de a) a f) se incluye en la invención y en una a) preferida se combina con al menos uno de b) a f).

Como alternativa, y/o además de, se puede usar una segunda unidad de transcripción genética para la expresión de la proteína Rep52/40. La expresión de la proteína Rep52/40 de esta segunda unidad de transcripción genética puede incrementarse por una o más de

- a) usar un promotor más fuerte para la unidad de transcripción de Rep52/40 en comparación con el promotor de la unidad de Rep78/68 (véase más adelante);
- b) aumentar el número de copias de la unidad de transcripción Rep52/40 en comparación con la de la unidad Rep78/68;
- c) mejorar el sesgo del uso de codones de la secuencia de codificación que codifica la proteína Rep52/40 para la expresión en células de insecto (como se ha descrito anteriormente, por ejemplo las SEQ ID NO: 2 o 10);
- d) optimización del contexto del codón de iniciación de la traducción de la proteína Rep52/40 de acuerdo con el contexto iniciador óptimo de 5'-N N N N N A U G A a/c/g N-3' para la iniciación eficiente de la traducción en células de lepidópteros (como se describe en Chang y col., citado anteriormente); y
- e) mediante la incorporación de una secuencia de control de la expresión que comprende una secuencia de nueve nucleótidos de la SEQ. ID NO: 9 o una secuencia de nucleótidos sustancialmente homóloga a la SEQ. ID

NO: 9, en dirección 5' de los codones de iniciación de la proteína Rep52/40.

Un ejemplo de una construcción en la que dos unidades de transcripción separadas se utilizan para la expresión de las proteínas Rep78/68 y Rep 52/40 es la construcción pVD183 como se ha descrito en los Ejemplos 2 y 3 en el presente documento. Las construcciones de ácido nucleico para su uso en el procedimiento para producir un virión de parvovirus recombinante (y que producen un nivel de expresión de Rep52/40 en la célula de insecto que es más alto que el nivel de expresión de Rep78/68 sobre una base molar) son un aspecto adicional de la presente invención.

En el presente documento se entiende que el vector de parvovirus recombinante (VAAr) producido en el procedimiento es, preferentemente, un virión infeccioso de VAA o de parvovirus que comprende los ácidos nucleicos del vector de parvovirus recombinante (VAAr). Las condiciones de crecimiento para células de insecto en cultivo y producción de productos heterólogos en células de insecto en cultivo son bien conocidas en la técnica y se describen, por ejemplo, en las referencias citadas anteriormente sobre la ingeniería molecular de las células de insecto (véase también el documento WO2007/046703). Preferentemente, el procedimiento comprende además la etapa de purificación por afinidad del vector de parvovirus recombinante (VAAr) (los viriones que lo comprenden) utilizando un anticuerpo anti-VAA como se ha descrito anteriormente.

Un primer promotor igualmente fuerte o más fuerte que un segundo promotor para uso en la invención puede definirse del siguiente modo. La fuerza del promotor se puede determinar mediante la expresión que se obtiene en las condiciones que se utilizan en el procedimiento de la invención. En una realización preferida, el primer promotor o el segundo promotor se seleccionan del grupo que consiste en un promotor Polh, el promotor p10, el promotor de la 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 preferentemente, el primer promotor se selecciona del grupo que consiste en un promotor Polh, el promotor p10 o el promotor de la proteína básica y en el que el segundo promotor es un promotor deltaE1 o un promotor E1, o cualquier otro promotor génico de baculovirus temprano o tardío. Preferentemente, el primer promotor en la construcción de ácido nucleico de la invención es un promotor p10 y el segundo promotor es un promotor Polh o un promotor Hsp70 mínimo 4xHsp27 EcRE +. En otra realización, la primera promotor en la construcción de ácido nucleico de la invención es un promotor 4xHsp27 EcRE+mínimo Hsp70 y el segundo promotor es un promotor Polh. En aún otra realización, el primer promotor en la construcción de ácido nucleico de la invención es un promotor Polh y el segundo promotor es un promotor p10, deltaE1 o E1. e) mediante la incorporación de una secuencia de control de la expresión que comprende una secuencia de nueve nucleótidos de la SEQ. En otra realización más, el primer promotor en la construcción de ácido nucleico 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 la construcción de ácido nucleico de la invención es un promotor PolH y el segundo promotor es un promotor PolH. Lo más preferentemente, el primer promotor en la construcción de ácido nucleico de la invención es un promotor PolH y el segundo promotor es un promotor deltaE1.

Se entiende que un "elemento potenciador" o "potenciador" define una secuencia que potencia la actividad de un promotor (es decir, aumenta la tasa de transcripción de una secuencia en dirección 3' del promotor), que, a diferencia de un promotor, no posee actividad promotora y que, por lo general, puede funcionar independientemente de su ubicación con respecto al promotor (es decir, en dirección 5' o 3' del promotor). Los elementos potenciadores son bien conocidos en la técnica. Ejemplos no limitantes de elementos potenciadores (o partes de los mismos) que podrían utilizarse en la presente invención incluyen potenciadores de baculovirus y elementos potenciadores que se encuentran en células de insecto. Se prefiere que el elemento potenciador aumente en una célula la expresión de ARNm de un gen, al que el promotor que está unido operativamente, en al menos un 25 %, más preferentemente al menos un 50 %, incluso más preferentemente al menos un 100 %, y, lo más preferentemente, al menos un 200 % en comparación con la expresión de ARNm del gen en ausencia del elemento potenciador. La expresión del ARNm se puede determinar, por ejemplo, mediante RT-PCR cuantitativa.

En el presente documento se prefiere el uso de un elemento potenciador para aumentar la expresión de la proteína Rep de parvovirus. 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 de respuesta a ecdisona. Preferentemente, el elemento potenciador se selecciona del grupo que consiste en hr1, hr2, hr3, hr4 y hr5.

En el presente documento y sus reivindicaciones, el verbo "comprender" y sus conjugaciones se usan en su sentido no limitante para indicar que los artículos que siguen a la palabra están incluidos pero que los artículos no mencionados específicamente no están excluidos. Además, la referencia a un elemento con un artículo indefinido "un" o "uno/una" no excluye la posibilidad de que haya presente más de uno de los elementos, a menos que el contexto requiera claramente que haya uno y solo uno de los elementos. El artículo indefinido "un" o "uno/una" normalmente significa "al menos uno/una".

Los ejemplos siguientes se ofrecen únicamente con fines ilustrativos y no están destinados a limitar el ámbito de la presente invención de ningún modo.

### **Descripción de las figuras**

Figura 1 Mapa físico de pVD183

Figura 2 Relación de las copias genómicas del gen ORF 1629 y el gen Rep en las muestras de baculovirus tomadas en diferentes pases de la construcción de baculovirus Bac.FBDSLRL (Urabe y col., 2002, Hum Gene Ther. 13(16):1935-43). Las copias genómicas se midieron por QPCR.

Figura 3 Relación de las copias genómicas del gen ORF 1629 y el gen Rep en las muestras de baculovirus tomadas en diferentes pases de la construcción de baculovirus pVD183 de la invención. Las copias genómicas se midieron por QPCR.

Figura 4. Producción de VAAr con BacVD183. La caída de la producción se debe a una reducción en la cantidad de baculovirus presentes.

Figura 5 ORF QPCR en los pases de Bac.VD183.

Figura 6. Transferencia de tipo Western para la expresión de varios pases de Bac.VD183. "88" indica la construcción Bac.VD88, que se denomina REP-ACG/PSC en el documento WO2007/148971, que se utiliza en el presente documento como control. La cantidad de expresión de Rep se relaciona con la concentración de Bac.VD183.

Figura 7. Q-PCR sobre volumen celular bruto (CLB) de producciones de VAAr1 utilizando tres construcciones diferentes para las proteínas Rep: VD88, VD183 y VD189. 5: 1: 1 se refiere a la relación de los diferentes baculovirus utilizados en la producción, 5 se refiere a Bac.VD88, Bac.VD183, o Bac.VD189, el primer 1 se refiere a Bac.VD84 (que contiene el gen de la cápside de VAA1) y el segundo 1 se refiere al baculovirus que contiene la construcción de ITR, Bac.VD43.

Figura 8. Los CLB de la producción de tres VAAr1 diferentes se purificaron en una columna de Llama específica para la cápside de VAA y en los lotes purificados se midieron las copias genómicas y las partículas de VAAr totales. La división de las partículas de VAAr totales por el número de Q-PCR da lugar a la relación total mencionada anteriormente. 5:1:1 se refiere a la relación de los diferentes baculovirus utilizados en la producción, 5 se refiere a Bac.VD88, Bac.VD183, o Bac.VD189, el primer 1 se refiere a Bac.VD84 (que contiene el gen de la cápside de VAA1) y el segundo 1 se refiere al baculovirus que contiene la construcción de ITR, Bac.VD43.

Figura 9. Transferencia de tipo Western de Rep. Se tomaron muestras a varios pases del baculovirus Bac.VD88 o Bac.VD189 y se realizó una transferencia de tipo Western. La cantidad de Rep52 respecto a la cantidad de Rep78 es consistentemente superior para Bac.VD189.

Figura 10. Transferencia de tipo Western de Rep. Se tomaron muestras a varios pases del baculovirus Bac.VD183 y se realizó amplificación y transferencia de tipo Western de Rep. La cantidad de Rep52 con respecto a la cantidad de Rep78 es mucho más alta para Bac.VD183 que para Bac.VD189 y Bac.VD88.

## Ejemplos

### 1. Ejemplo 1

#### 1.1. Materiales y procedimientos

##### 1.1.1 Construcción del plásmido de baculovirus

pFBDSLRL (Urabe y col., 2002, mencionado anteriormente) es un vector de expresión pFastBacDual (Invitrogen) que comprende 2 casetes de expresión separados para las proteínas Rep78 y Rep52 de VAA2, de modo que la expresión de las proteínas Rep52 está dirigida por el promotor Polh y la expresión de la proteína Rep78 a partir del promotor ΔIE. Esta construcción se ha subclonado en pPSC10, un plásmido que es compatible con el GenExpress BaculoKIT (Protein Sciences Corporation).

La secuencia de codificación de Rep52 de tipo salvaje en el casete de expresión de Rep 52 se reemplaza con el la secuencia de codificación de Rep52 optimizada por codones de la SEQ ID NO. 2 para producir pPSC10Rep-52CD.

La secuencia de codificación de Rep52 de tipo salvaje en el casete de expresión de Rep 78 de pPSC10Rep-52CD se reemplaza con la secuencia de codificación de Rep52 optimizada por AT de la SEQ ID NO. 3 para producir pPSC10Rep-52CD/78AT.

La secuencia de codificación de Rep52 de tipo salvaje en el casete de expresión de Rep 78 de pPSC10Rep-52CD se reemplaza con la secuencia de codificación de Rep52 optimizada por GC de la SEQ ID NO. 4 para producir pPSC10Rep-52CD/78GC.

##### 1.1.2 Producción de baculovirus recombinante

Los baculovirus recombinantes derivados del virus de de la polihedrosis nuclear múltiple de *Autographa californica* (AcMNPV) se producen utilizando el GenExpress BaculoKIT (Protein Sciences Corporation). La transfección se lleva a cabo de la siguiente manera: En un tubo de 14 ml de fondo redondo se mezclan 200 μl de medio GRACE con 6 μl de cellfectine (Invitrogen) y en un tubo eppendorf, 200 μl de medio GRACE se mezclan con 50 μl de ADN viral (Protein Sciences) y 2 μg del plásmido de transferencia (REP). El contenido del tubo de eppendorf se añade al tubo y se mezcla cuidadosamente. Después de un período de incubación de 30 minutos a temperatura ambiente se añaden 1.300 μl de GRACE a la mezcla de transfección. Las células de insecto en un matraz T25 se lavan con medio GRACE y se añade la mezcla de transfección gota a gota a la capa de células. Después de una incubación de 6 horas a 28 °C se añade suero SF900II suplementado con 10 % de SFB cuidadosamente y el matraz T25 se introduce en una estufa de 28 °C durante 5 días, después de lo cual se cosecha el baculovirus recombinante.

## 1.2 Resultados

El rendimiento de pPSC10Rep-52CD, pPSC10Rep-52CD/78AT and pPSC10Rep-52CD/78GC pPSC10Rep recién diseñados se compara con las construcciones pFBDSLRL de Rep originales de de Urabe y col., (2002, mencionado anteriormente). Las cuatro construcciones se pasan en serie hasta el paso 5. Se realizan experimentos de producción de VAA1 recombinante usando las construcciones de Rep de los pases 2, 3, 4 y 5 en combinación con un baculovirus que contiene un casete de expresión de mamífero de un gen indicador entre de las ITR de VAA (VAA-LPL) y un baculovirus que contiene un casete de expresión de células de insecto para el VAA1-Cap (VAA-cap) de los pases respectivos 2, 3, 4 y 5. Los baculovirus recombinantes de VAA-LPL y VAA-Cap, tal como se utilizan en el presente documento, se describen en el documento WO2007/046703. Se determinan los rendimientos de producción de VAA1-LPL mediante QPCR. El baculovirus original diseñado por Urabe y col., 2002 (REP/Bac-to-Bac original) se traduce en una disminución rápida de la producción de VAA en 5 pases. Sin embargo, el baculovirus con las unidades de expresión de REP de pPSC10Rep-52CD, pPSC10Rep-52CD/78AT y pPSC10Rep-52CD/s 78GC da lugar a la producción de VAA estable durante al menos 5 pases. Por lo tanto, los rendimientos de producción reproducibles de VAA-LPL en varios pases (por ejemplo, de 2 a 5), solo se obtienen usando baculovirus que contienen las construcciones pPSC10Rep-52CD, pPSC10Rep-52CD/78AT y pPSC10Rep-52CD/78GC.

## 2. Ejemplo 2

Previamente se ha descrito que los vectores de expresión de baculovirus que contiene de 2 casetes de expresión separados para las proteínas Rep78 and Rep52 de VAA son genéticamente inestables en los baculovirus (véase, por ejemplo, el documento WO2007/148971 y Kohlbrenner y col., 2005, Mol Ther- 12(6):1217-25). Ahora, los inventores han establecido la aplicación de la optimización del uso de codones (con respecto al uso de codones del virus de la polihedrosis nuclear múltiple por *Autographa californica* (AcMNPV) de solo la secuencia de codificación de Rep52 y no la secuencia de codificación de Rep78 a fin de introducir cambios suficiente entre las partes anteriormente idénticas de las secuencias de codificación de Rep52 y Rep78 para reducir los acontecimientos de recombinación. A continuación, demostraron que realmente este era el caso.

### 2.1 Clonación

Un plásmido que contiene los casetes de expresión doble de Rep originales en el plásmido pPSC10, pVD42 de Protein Sciences Corporation se modificó. pVD42 contiene el gen de rep78 dirigido por el promotor deltaE1 y el gen de rep52 dirigido por el promotor PolH, como en la construcción pFBDSLRL original (Urabe y col., 2002, Hum Gene Ther. 13(16):1935-43). La secuencia de codificación de rep52 en pVD42 se sustituyó por una secuencia de codificación de rep52 sintética cuyo uso de codones se adaptó al uso de codones del virus de la polihedrosis nuclear múltiple de *Autographa californica* (AcMNPV) (véase la Tabla 2; y <http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=46015>). Esta secuencia de codificación de rep52 de VAA2 optimizada por codones se representa en la SEQ ID NO: 10. En la figura 1 se muestra un mapa físico del plásmido pVD183 resultante, que comprende el la secuencia de codificación de rep52 de VAA2 optimizada por codones dirigida desde el promotor Polh y la secuencia de codificación de rep78 del VAA2 dirigida desde el promotor deltaE1.

## 2.2 Resultados

Los inventores han realizado un clon de baculovirus recombinante del plásmido pVD183 y se pasó el baculovirus 10 veces para analizar su estabilidad genética. Los inventores analizaron la estabilidad genética de la construcción mediante QPCR en el genoma del baculovirus y el gen de Rep52 mediante transferencia de tipo Western y mediante la eficiencia de producción del VAAr del baculovirus. Al mismo tiempo, el baculovirus Bac.VD42 original se hizo pasar al pase 7 para la comparación. Los datos anteriores sobre la estabilidad del Bac.VD42(o Bac.FBDSLRL) también se mencionan en el documento WO2007/148971 (denominado REPBac-to-Bac original).

### 2.2.1 QPCR

Estabilidad medida mediante QPCR en los genomas de baculovirus. El número de copias de un gen que es esencial para la replicación de baculovirus y que se utiliza para la producción de BacVD183 a partir de pVD183 mediante recombinación en ORF1629 y ORF603 entre el pVD183 y el esqueleto del baculovirus de Protein Sciences. El ORF 1629 (ORF) se ha medido mediante QPCR y el número de copias de los genes Rep también se ha medido mediante QPCR. La relación entre estos 2 genes debe permanecer igual durante los pases posteriores del baculovirus. La figura 2 muestra para la comparación que Bac.FBDSLRL es más bien inestable. Figura 3 muestra que Bac.VD183 es significativamente más estable. Los inventores observaron que la eficiencia de los 2 conjuntos de cebadores utilizados en la QPCR no es necesariamente igual, por lo tanto, se puede obtener una relación diferente de 1. Sin embargo, un indicador más importante de la estabilidad es que la relación debe permanecer relativamente constante durante múltiples pases. El pase 3 del Bac.FBDSLRL ya es subóptimo, dado que la relación es de aproximadamente 0,25 y solo empeora. Bac.VD183 también comienza alrededor de 0,3 pero fluctúa alrededor de esa relación, lo que indica que hay una situación estable. Las deleciones en el genoma de baculovirus da lugar a un baculovirus que crece más rápido que los baculovirus que tiene un genoma de longitud completa, por lo tanto, cuando se produce una deleción, esos clones crecerán más que los demás. Las variaciones en el procedimiento de QPCR pueden dar lugar a las fluctuaciones observadas en la Figura 3.



2.2.2 Producción de VAAR

La Figura 4 muestra la producción de VAAR con la construcción Bac.VD183 estable. La caída en la producción en los pases más altos se debe a una reducción en la cantidad de baculovirus utilizados en la producción de VAAR (véase la Figura 5). La figura 5 muestra la QPCR en el ORF de Bac.VD183, que está directamente relacionado con la cantidad de baculovirus presentes en la muestra. La cantidad de baculovirus utilizados en la producción de VAAR se correlaciona con la cantidad de VAAR producido.

2.2.3 Transferencia de tipo Western de Rep

La Figura 6 muestra la expresión de la proteína Rep durante los pases de Bac.VD183 analizada mediante transferencia de tipo Western.

**3. Ejemplo 3**

El efecto del nivel de expresión de Rep52 se determinó en dos parámetros de producción de VAAR. En particular, el efecto del nivel de expresión relativo de Rep52 en comparación con el nivel de expresión de Rep78 en 1) el nivel de producción de VAAR como se expresa en las copias del genoma por ml de volumen celular bruto (g/m de CLB); y 2) la relación entre los viriones de VAAR totales y los viriones de VAAR completos (los viriones de VAAR completos son viriones que comprenden una copia del genoma de VAAR). Estos parámetros se compararon para tres diferentes construcciones de Rep de VAAR diferentes que dan lugar cada una a diferentes niveles de expresión de Rep52 y en proporción diferente entre los niveles de Rep52 y Rep78. Las tres construcciones fueron pVD88 (denominados REP-ACG/PSC en el documento WO2007/148971), pVD183 (descrito en el Ejemplo 2 en el presente documento anteriormente) y pVD189 (véase más adelante).

3.1 Construcción de pVD189

La construcción pVD88 se volvió a diseñar mediante la eliminación de 9 secuencias de ATG entre el inicio de la traducción de los genes de Rep78 y Rep 52 y cambiando el codón de iniciación de la traducción ACG de Rep78 a CTG. Véase la secuencia siguiente. Baseclear (Leiden, Países Bajos) sintetizó el nuevo gen y lo clonó en pVD88 reemplazando el gen Rep existente para obtener pVD189. La secuencia de nucleótidos de la secuencia de codificación de Rep en pVD189 se representa en la SEQ ID NO: 11.

3.2 Producción de VAAR

Los baculovirus se realizaron con las construcciones de VD88, VD183, y VD189 y estos se utilizaron para la producción de VAAR1. La comparación de las construcciones VD88, VD183 y VD189 en la producción de VAAR dio lugar a una producción de VAAR (copias del genoma) mejor, medida mediante Q-PCR en el volumen celular bruto (CLB). La Figura 7 muestra que la construcción de Rep estándar VD88, que da como resultado una cantidad menor de Rep52 (Figura 9) da como resultado aproximadamente  $4 \times 10^{10}$  de GC/ml medidos en el CLB. VD189 que conduce a una cantidad de Rep52 ligeramente mayor (Figura 9) dio como resultado una producción de VAAR medida en el CLB de aproximadamente  $6 \times 10^{10}$  de GC/ml. VD183 que conduce a una cantidad de Rep52 claramente mayor (Figura 10) dio como resultado una producción de VAAR medida en el CLB de aproximadamente  $6 \times 10^{10}$  de GC/ml.

Un parámetro de calidad muy importante es la relación total: completos del lote de VAAR. La Figura 8 muestra que la mejor relación de totales (viriones): completos (viriones) se obtiene con la construcción VD183 que muestra la cantidad de Rep52 más alta respecto a la cantidad de Rep78 en comparación con las construcciones Bac.VD189 y Bac.VD88 en la Figura 9.

3.2 Construcciones adicionales

Se construyen y se analizan siguientes construcciones, y parte de la invención:

Construcciones	Promotor(es)	Codones de iniciación y secuencias de codificación
1) VD88	PolH	ACG-78-----ATG-52-----*
2) VD189	PolH	CTG-78-atg ATG-52 eliminado-----*
3) VD183	delta E1 PolH	ATG-78-----* + ATG--52---SEQ ID NO: 10-----*
4) VD196	PolH	CTG-78-----ATG-52-----*
5) VD197	PolH	ACG -78-atg ATG-52 eliminado-----*

ES 2 602 585 T3

(continuación)

Construcciones	Promotor(es)	Codones de iniciación y secuencias de codificación
6) VD197/52	P10 PolH	ACG -78-atg ATG-52 eliminado-----*+ ATG--52---SEQ ID NO: 10-----*
7) VD189/52	P10 PolH	CTG-78-atg ATG-52 eliminado-----*+ ATG--52---SEQ ID NO: 10-----*
8) VD183/10	p10 PolH	ATG--78-----* + ATG--52---SEQ ID NO: 10-----*
9) VD197/52cd	PolH	ACG -78-atg ATG-52 eliminado-SEQ ID NO: 10*

1,2, 4, 5, 8 y 9 tienen 1 unidad de transcripción para la expresión de las proteínas Rep 78 y 52.  
3, 6 y 7 tienen 2 unidades de transcripción para la expresión las proteínas Rep 78 y 52.

5 Una estimación aproximada de las cantidades y las relaciones de las proteínas rep 78 y rep 52 para las diferentes construcciones durante la producción de VAAr (rep78:rep52):

	78	52
1)	1	:1
2)	1.5	:2
3)	1	: 20
4)	5	: 0.25
5)	1	:5
6)	0.5	: 30
7)	0.75	: 30
8)	5	: 20
9)	1	: 10

**Tabla 1.** Frecuencias de los codones de *Spodoptera frugiperda* basadas en 127 secuencias de codificación (33.098 codones)

campos: [triple] [frecuencia: por mil] ([número])							
TTT	9,7 (320)	TCT	10,5 (347)	TAT	10,1 (334)	TGT	6,9 (227)
TTC	26,9 (889)	TCC	13,0 (430)	TAC	24,4 (807)	TGC	12,4 (409)
TTA	7,0 (233)	TCA	9,9 (329)	TAA	2,5 (83)	TGA	0,6 (21)
TTG	16,2 (536)	TCG	7,2 (237)	TAG	0,7 (23)	TGG	12,7 (420)
CTT	9,9 (327)	CCT	14,3 (472)	CAT	8,7 (289)	CGT	15,9 (525)
CTC	17,0 (564)	CCC	13,7 (453)	CAC	16,2 (535)	CGC	15,1 (500)
CTA	6,8 (226)	CCA	13,4 (445)	CAA	16,2 (535)	CGA	5,3 (175)
CTG	24,5 (810)	CCG	7,7 (255)	CAG	21,8 (723)	CGG	3,6 (118)
ATT	15,5 (512)	ACT	13,6 (451)	AAT	12,8 (424)	AGT	8,1 (267)
ATC	28,9 (958)	ACC	17,2 (569)	AAC	27,8 (921)	AGC	10,7 (354)
ATA	7,6 (253)	ACA	11,9 (393)	AAA	26,7 (883)	AGA	11,8 (392)
ATG	27,3 (902)	ACG	8,8 (290)	AAG	53,1 (1757)	AGG	13,5 (446)
GTT	14,7 (488)	GCT	26,3 (872)	GAT	21,8 (723)	GGT	22,0 (728)
GTC	20,4 (676)	GCC	21,1 (697)	GAC	32,3 (1070)	GGC	19,9 (659)
GTA	12,3 (406)	GCA	12,4 (411)	GAA	27,2 (901)	GGA	18,2 (603)
GTG	24,8 (822)	GCG	12,2 (404)	GAG	34,1 (1128)	GGG	4,3 (141)

## ES 2 602 585 T3

GC de codificación 50,58% 1ª letra GC 53,42% 2ª letra GC 39,40% 3ª letra GC 58,93%

**Tabla 2.** Tabla de uso de codones de la polihedrosis nuclear múltiple de *Autographa californica* (AcMNPV) en base a 277 secuencias de codificación (77.487 codones)

campos: [tripleto] [frecuencia: por mil] ([número])							
UUU	37,6 (2916)	UCU	10,3 (799)	UAU	22,2 (1721)	UGU	11,2 (865)
UUC	11,3 (879)	UCC	7,2 (556)	UAC	26,1 (2019)	UGC	12,5 (967)
UUA	20,6 (1594)	UCA	7,2 (557)	UAA	2,7 (209)	UGA	0,5 (38)
UUG	34,3 (2659)	UCG	14,2 (1100)	UAG	0,4 (29)	UGG	7,5 (579)
CUU	8,2 (637)	CCU	8,2 (636)	CAU	10,2 (789)	CGU	8,1 (630)
CUC	7,2 (555)	CCC	11,3 (879)	CAC	12,8 (991)	CGC	13,2 (1024)
CUA	8,2 (632)	CCA	8,0 (621)	CAA	26,6 (2063)	CGA	7,4 (576)
CUG	13,0 (1007)	CCG	12,7 (985)	CAG	11,5 (892)	CGG	3,9 (304)
AUU	31,2 (2416)	ACU	12,4 (962)	AAU	34,5 (2671)	AGU	10,3 (800)
AUC	14,3 (1111)	ACC	13,5 (1043)	AAC	44,3 (3433)	AGC	16,1 (1251)
AUA	19,7 (1527)	ACA	12,4 (961)	AAA	52,4 (4057)	AGA	9,7 (748)
AUG	26,7 (2071)	ACG	18,5 (1434)	AAG	18,3 (1418)	AGG	4,0 (309)
GUU	16,5 (1277)	GCU	11,0 (850)	GAU	25,4 (1968)	GGU	7,8 (603)
GUC	11,7 (904)	GCC	15,4 (1196)	GAC	33,8 (2619)	GGC	16,1 (1251)
GUA	12,6 (973)	GCA	10,0 (771)	GAA	37,2 (2885)	GGA	7,0 (541)
GUG	25,7 (1990)	GCG	16,3 (1261)	GAG	16,2 (1253)	GGG	2,9 (225)

5 GC de codificación 41,86% 1ª letra GC 43,60% 2ª letra GC 32,68% 3ª letra GC 49,29%

### LISTADO DE SECUENCIAS

- <110> AMT B.V.
- 10 <120> Vectores de baculovirus que comprenden secuencias de codificación repetidas
- <130> P6016639PCT
- 15 <150> EP 07113257.5
- <151> 26-07-2007
- <150> US 60/952,081
- <151> 26-07-2007
- 20 <160> 11
- <170> PatentIn versión 3.3
- 25 <210> 1
- <211> 1194
- <212> ADN
- <213> virus adenoasociado 2
- 30 <220>
- <221> misc\_feature
- <223> tipo silvestre
- <400> 1

# ES 2 602 585 T3

```

atggagctgg tgggtggct cgtggacaag gggattacct cggagaagca gtggatccag      60
gaggaccagc cctcata cat ctccttcaat ggggcctcca actcgcggtc ccaaatcaag     120
gctgccttgg acaatgcggg aaagattatg agcctgacta aaaccgcccc cgactacctg     180
gtgggccagc agcccgtgga ggacatttcc agcaatcgga tttataaaat tttggaacta     240
aacgggtacg atccccaaata tgcggcttcc gtctttctgg ga tgggccac gaaaaagttc     300
ggcaagagga acaccatctg gctgtttggg cctgcaacta cggggaagac caacatcgcg     360
gaggccatag cccacactgt gcccttctac gggtgcgtaa actggacca a tgagaacttt     420
cccttcaacg actgtgtcga caagatgggt atctgggtggg aggaggggaa gatgaccgcc     480
aaggtcgtgg agtcggccaa agccattctc ggaggaagca aggtgcgcgt ggaccagaaa     540
tgcaagtcct cggcccagat agaccgcact cccgtgatcg tcacctcaa caccaacatg     600
tgcgcctgta ttgacgggaa ctcaacgacc ttogaacacc agcagccgtt gcaagaccgg     660
atgttcaa at ttgaactcac ccgc cgtctg gatcatgact ttgggaaggt caccaagcag     720
gaagtcaaag actttttcog gtgggcaaag gatcacgtgg ttgaggtgga gcatgaattc     780
tacgtcaaaa aggggtggagc caagaaaaga cccgccccca gtgacgcaga tataagtgag     840

cccaaacggg tgcgcgagtc agttgcgcag ccategcagc cagacgcgga agcttcgatc     900
aactacgcag accgctacca aaacaaatgt tctcgtcacg tgggcatgaa tctgatgctg     960
tttccttcca gacaatgcga gagaatgaat cagaattcaa atatctgctt cactcacgga    1020
cagaaagact gtttagagtg ctttcccggt tcagaatctc aaccgtttc tgcgtcaaaa    1080
aaggcgtatc agaaactgtg ctacattcat catatcatgg gaaaggtgcc agacgcttgc    1140
actgcctgcg atctggtcaa tgtggatttg gatgactgca tctttgaaca ataa          1194

```

- <210> 2
- <211> 1194
- 5 <212> ADN
- <213> virus adenoasociado 2
  
- <220>
- <221> misc\_feature
- 10 <223> sf9 optimizado
  
- <400> 2

ES 2 602 585 T3

atggagctgg tgggttggct ggtggacaag ggtateacct cegagaagca gtggatccag 60  
gaggaccagc ctccctacat ctccctcaac getgcttcca actcccgttc ccagatcaag 120  
getgctctgg acaacgetgg taagatcatg tccctgacca agaccgetcc tgactacctg 180  
gtgggtcagc agcctgtgga ggacatctcc tccaaccgta tctacaagat cctggagctg 240  
aacggttaac accctcagta cgetgettcc gtgttccctgg gttgggctac caagaagttc 300  
gytaagcgtc acaccatctg getgttccgt cctgctacca ccggttaagac caacatcget 360  
gaggctatcg ctacaccgtg gcccttct ac ggttgcgtga actggacca cagagaacttc 420  
cctttcaacg actgcgtgga caagatggtg atctggtggg aggagggtaa gatgaccgct 480  
aagggtggtg agtccgctaa ggctatcctg ggtggttcca aggtgcgtgt ggaccagaag 540  
tgcaagtcc cgcctcagat cgaacctacc cctgtgatcg tgacctcaa cac caacatg 600  
tgcgctgtga tegacggtaa ctccaccacc ttcgagcacc agcagcctct gcaggaccgt 660  
atgttcaagt tegagctgac ccgtcgtctg gaccacgact tcggtaaggt gaccaagcag 720  
gagggtgaagg acttcttccg ttgggctaag gaccacgtgg tggaggtgga gcacgagttc 780  
tacgtgaaga aggggtggtc taagaagcgt cctgctcctt ccgaacgtga catctccgag 840  
cctaagcgtg tgcgtgagtc cgtggctcag ccttccacct ccgaacgtga ggcttccatc 900  
aaactacgtg accgttacca gaacaagtgc tcccgtcacg tgggtatgaa cctgatgctg 960  
ttcccttgc gtcagtgcga gcgatgaac cagaa ctcca acatctgctt caccacggt 1020  
cagaaggaact gcctggagtg ctccctgtg tccagctccc agcctgtgtc cgtggtgaag 1080  
aaggcttacc agaagctgtg ctacatccac cacatcatgg gtaaggtgcc tgacgcttgc 1140  
accgcttgcg acctggtgaa cgtggacctg gacgactgca tcttcgagca gtaa 1194

<210> 3  
<211> 1194  
5 <212> ADN  
<213> virus adenoasociado 2

<220>  
10 <221> misc\_feature  
<223> AT optimizado

<400> 3

ES 2 602 585 T3

```

atggaattag taggatgggt agtagataaa ggaataecat cagaaaaaca atggatacaa      60
gaagatcaag catcatatat atcatttaat gc agcatcaa attcaagatc acaataaaaa      120
gcagcattag ataatgcagg aaaaaaatg tcattaacaa aaacagcacc agattattta      180
gtaggacaac aaccagtaga agatatatca tcaaatagaa tatataaaat attegaatta      240
aatggatag atocacaata tgcagcatca gtatTTTTtag gatgggcaac aaaaaaat tt      300
ggaaaaagaa atacaatatg gttatTTTgga ccagcaacaa caggaaaaac aeatatagca      360
gaagcaatag cacatacagt accattttat ggatgtgtaa attggacaaa tgaaaatttt      420
ccatttaatg attgtgtaga taaaatggta atatgggtggg aagaaggaaa aatgacagca      480
aaagtagtag aatc agcaaa agcaatatta ggaggatcaa aagtaagagt agatcaaaaa      540
tgtaaactat cagcacaat agatccaaca ccagtaatag taacatcaaa tacaatatg      600
tgtgcagtaa tagatggaaa ttcaacaaca tttgaacatc acaaccatt acaagataga      660
atgtttaaat ttgaattaac aagaagatta gatcatgatt ttggaaaagt acaaaaaaca      720
gaagtaaaag atTTTTTtag atgggcaaaa gatcatgtag tagaagtaga acatgaattt      780
tatgtaaaaa aaggaggagc aaaaaaaga ccagcaccat cagatgcaga tatatcagaa      840
ccaaaaagag taagagaatc agtagcacia ccatcaacat cagatgcaga agcatcaata      90 0
aattatgcag atagatatca aaataaatgt tcaagacatg taggaatgaa tttaatgtta      960
tttccatgta gacaatgtga aagaatgaat caaaattcaa atatatgttt tacacatgga      1020
caaaaagatt gtttagaatg ttttccagta tcagaatcac aaccagtatc agtagtaaaa      1080
aaagcatatc aaaaattatg t tatatacat catataatgg gaaaagtacc agatgcatgt      1140

acagcatgtg atttagtaaa tgtagattta gatgattgta ttttgaaca ataa      1194

```

- <210> 4
- <211> 1194
- <212> ADN
- <213> virus adenoasociado 2
  
- <220>
- <221> misc\_feature
- <222> (1)..(1194)
- <223> GC optimizado
  
- <400> 4

ES 2 602 585 T3

atggagctgg tggggtggct ggtggacaag gggatcacga gcgagaagca gtggatccag 60  
 gaggaccagg cgagctacat cagcttcaac gcggcgagca acagccggag ccagatcaag 120  
 gcggcgctgg acaacgcggg gaagatcatg agcctgacga agacggcgcc ggactacctg 180  
 gtggggcagc agccggtgga ggacatcagc agcaaccgga tctacaagat cctggagctg 240  
 aacgggtacg acccgcagta cgcggcgagc gtgttcctgg ggtgggagac gaagaagttc 300  
 gggaaagcga acacgatctg gctgttcggg ccggcgacga cggggaagac gaacatcgcg 360  
 gaggcgatcg cgcacacggg gccggt ctac ggtgctgga actggacgaa cgagaacttc 420  
 ccgttcaacg actgctgga caagatggtg atctggtggg aggaggggaa gatgacggcg 480  
 aagggtggtg agagcgcgaa ggcgatcctg ggggggagca aggtgcgggt ggaccagaag 540  
 tgcaagagca gcgcgcagat cgaccgcagc ccggtgatcg tgacgagcaa c acgaacatg 600  
 tgcgcggtga tcgacgggaa cagcacgacg ttcgagcacc agcagccgct gcaggaccgg 660  
 atgttcaagt tcgagctgac gcggcggtg gaccacgact tcgggaaggt gacgaagcag 720  
 gagggtgaag acttcttccg gtgggcgaag gaccacgtgg tggaggtgga gcacgagttc 780  
 tacgtgaaga aggggggggc gaagaagcgg ccggcgccga gcgacgcgga catcagcgag 840  
 ccgaagcggg tcggggagag cgtggcgag ccgagcacga gcgacgcgga ggcgagcatc 900  
 aactacgcgg accggtacca gaacaagtgc agccggcacg tggggatgaa cctgatgctg 960  
 ttcccgtgcc ggcagtgcga gcggatgaac cag aacagca acatctgctt cagcacggg 1020  
 cagaaggact gcctggagtg cttcccgtg agcgagagcc agccggtgag cgtggtgaag 1080  
 aaggcgtacc agaagctgtg ctacatccac cacatcatgg ggaaggtgcc ggacgcgtgc 1140  
 acggcgtgcg acctggtgaa cgtggacctg gacgactgca tcttcgagca gtaa 1194

- 5 <210> 5
- <211> 1194
- <212> ADN
- <213> virus adenoasociado 2
  
- 10 <220>
- <221> CDS
- <222> (1)..(1194)
- <223> Rep52
  
- <400> 5

ES 2 602 585 T3

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 t gg 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	



ES 2 602 585 T3

180	185	1 90	
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 g gg 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 t at 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

ES 2 602 585 T3

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

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

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

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

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

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

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

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

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

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

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

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

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

ES 2 602 585 T3

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

ES 2 602 585 T3

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 tet 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 tet 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 ccg 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 tet 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

ES 2 602 585 T3

	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

ES 2 602 585 T3

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 got 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

10

ES 2 602 585 T3

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

ES 2 602 585 T3

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 L ys  
 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 L ys 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 A sn 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



ES 2 602 585 T3

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 L eu 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 G lu 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
- 5 <212> ADN
- <213> virus adenoasociado 2
- <400> 9
- cctgttaag 9
- 10 <210> 10
- <211> 1194
- <212> ADN
- <213> virus adenoasociado 2
- 15 <220>
- <221> misc\_feature
- <223> Secuencia codificante de Rep5 2 AcMNPV optimizado
- 20 <400> 10

ES 2 602 585 T3

atggaattgg tcggttggtt ggtggacaag ggtattacct cggagaagca atggatacaa 60  
 gaagatcaag cctcatacat ctcgtttaat ggggcatcca actcgcgtag ccaaatcaaa 120  
 gctgccttgg acaatgcggg caagattatg agcctgacta aaaccgccc cgactacctg 180  
 gtggggccagc aaccctgga agacatttcc agcaatcgca tctataagat tttggagtta 240  
 aacggctacg atcctcaata tgggcttcc gtatTTTTTg gctgggcgac gaaaaagttt 300  
 ggcaaaagaa acaccatttg gttgTTTgga cctgcaacta cgggaaaaac aaacatagcg 360  
 gagggcatag cccacactgt ac cTTTTat ggtgcgTTa actggacca tgagaacttt 420  
 ccattcaacg actgtgtcga caagatggtt atttggtggg aggaaggcaa aatgaccgct 480  
 aaagtogtgg agtgggcaa agcaatttta ggaggcagca aagtgcgct agaccagaaa 540  
 tgcaaaagct ctgogcagat agaccogaca cgggtgatcg ttacaagc aa cacgaacatg 600  
 tggcgcgtga ttgaogtaa cagtacgaca ttogaacacc aacaaccgtt gcaagaccga 660  
 atgttcaaat ttgaattgac ggcgcgactg gatcatgatt ttggcaaggT acaaaaacaa 720  
 gaagtcaaaag acttctttcg ttgggcaaaag gatcacgttg ttgaagtga acatgaattt 780  
 taogtcaaaa aaggtggtgc taagaaaaga cccgccccga gtgatgcaga tataagtga 840  
 cccaaacgag tgagagaatc ggttgcgcag ccaagcacgt cagatgcgga agcttcgata 900  
 aactacgcag accgctacca aaacaaatgt tctcgtcacg taggcatgaa cttaatgttg 960  
 tttccctgca gacaatgtga gagaatgaat cagaatagta atatctgttt cactcacggc 1020  
 cagaaagact gtttagaatg ctttccggtg tcagaatctc aaccogtttc tgtcgtaaaa 1080  
 aaggcgtatc aaaaattatg ctatattcat catatcatgg gaaaagtgc agacgcttgt 1140  
 actgcctgog atctggttaa tgtggatttg gatgactgta tctttgaaca ataa 1194

5 <210> 11  
 <211> 1866  
 <212> ADN  
 <213> Artificial

10 <220>  
 <223> Secuencia codificante de Rep pVD189

<400> 11

ctggcgggggt ttacgagat tgtgattaag gtcccagcg accttgacga gcatctgcc 60  
 ggcatttctg acagcttTgt gaactgggtg gccgagaagg agtggg agtt gccgccagat 120

ES 2 602 585 T3

totgacttgg atctgaatct gattgagcag gcacccctga ccgtggccga gaagctgcag 180  
 cgcgactttc tgacggagtg ggcgcgtgtg agtaaggccc cggaggccct tttctttgtg 240  
 caatttgaga agggagagag ctacttccac ttacacgtgc tcgtggaaac caccggggtg 300  
 aaatccttag ttttgggacg tttcctgagt cagattccgg aaaaactgat tcagagaatt 360  
 taccgcggga tcgagccgac tttgccaaac tggttccggg tcacaaagac cagaaacggc 420  
 gccggaggcg ggaacaaggt ggtggacgag tgctacatcc ccaattactt gctccccaaa 480  
 acccagcctg agctccagtg ggogtggg ct aatttagaac agtatttaag cgcctgtttg 540  
 aatctcacgg agcgtaaacg gttgggtggc cagcatctga cgcacgtgtc gcagacgcag 600  
 gagcagaaca aagagaatca gaatcccaat tctgacggc ccgtgatcag atcaaaaact 660  
 tcagccaggt acatggagct ggtggggtgg ctggtggaca aggggattac ctg ggagaag 720  
 cagtggatcc aggaggacca ggctcatac atctcctca atgcccctc caactcggcg 780  
 tcccaaatca aggctgcctt ggacaatgcg gaaagatta tgagcctgac taaaaccgcc 840  
 cccgactacc tgggtggcca gcagcccgtg gaggacattt ccagcaatcg gatttataaa 900  
 attttggaac taaacgggta cgatcccaa tatgcccgtt ccgtctttct gggatgggcc 960  
 acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggctgcaac taccgggaag 1020  
 accaacatcg cggaggccat agcccacact gtgcccctt acgggtgcgt aaactggacc 1080  
 aatgagaact ttccttcaa cgaactgtgtc gacaa gatgg tgatctggtg ggaggagggg 1140  
 aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc tcggaggaag caaggtgcgc 1200  
 gtggaccaga aatgcaagtc ctggcccag atagaccga ctcccgtgat cgtcacctcc 1260  
 aacaccaaca tgtgcgccgt gattgacggg aactcaacga ctttgcgaca ccagcagccg 1320  
 ttgcaagacc ggatgttcaa atttgaactc accgcgcgtc tggatcatga ctttgggaag 1380  
 gtcaccaagc aggaagtcaa agactttttc cgggtgggcaa aggatcacgt ggttgagggtg 1440  
 gagcatgaat tctacgtcaa aaaggttggg gccaaagaaa gaccgcgcc cagtgcgca 1500  
 gatataagtg agcccaa acg ggtgcgcgag tcagttgcgc agccatcgac gtcagacggc 1560  
 gaagcttcga tcaactacgc agacaggtag caaaacaaat gttctcgtca cgtgggcatg 1620  
 aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaatc aaatatctgc 1680  
 ttoactcacg gacagaaaga ctgtttagag tgctttcccg tg tcagaatc tcaaccggtt 1740  
 tctgtcgtca aaaagggcgt tcagaaactg tgctacattc atcatatcat gggaaagggtg 1800  
 ccagacgctt gcactgcctg cgatctggtc aatgtggatt tggatgactg catctttgaa 1860  
 caataa 1866

## REIVINDICACIONES

1. Una célula de insecto que comprende una primera secuencia de nucleótidos que codifica una primera secuencia de aminoácidos y una segunda secuencia de nucleótidos que codifica una segunda secuencia de aminoácidos, en la que:
- 5 (a) la primera y la segunda secuencias de nucleótidos están cada una unida operativamente a secuencias de control de la expresión para la expresión en una célula de insecto;
- (b) la primera y la segunda secuencias de aminoácidos comprenden una secuencia de aminoácidos común de la cual al menos 100 aminoácidos tienen una identidad de aminoácidos de al menos un 90 % entre la primera y la
- 10 (c) las secuencias de nucleótidos que codifican la secuencia de aminoácidos común en la primera y la segunda secuencias de aminoácidos son menos del 90% idénticas;
- (d) el uso de codones de la primera secuencia de nucleótidos que codifica la secuencia de aminoácidos común está adaptado al uso de codones de la célula huésped de insecto;
- 15 (e) la primera secuencia de nucleótidos codifica la secuencia de aminoácidos de una proteína Rep52 o 40 de parvovirus y la segunda secuencia de nucleótidos codifica la secuencia de aminoácidos de una proteína Rep78 o 68 de parvovirus; y
- (f) las secuencias de aminoácidos comunes comprenden las secuencias de aminoácidos del segundo aminoácido más C-terminal de la proteína Rep52 o 40 de parvovirus.
2. Una célula de insecto de acuerdo con la reivindicación 1, en la que las secuencias de nucleótidos que codifican la
- 20 secuencia de aminoácidos común en la primera y la segunda secuencias de aminoácidos son menos del 80 % idénticas.
3. Una célula de insecto de acuerdo con la reivindicación 1 o la reivindicación 2, en la que las secuencias de aminoácidos comunes en la primera y la segunda secuencias de aminoácidos comparten una identidad de aminoácidos de al menos un 99 %, preferentemente una identidad de aminoácidos del 100 %.
- 25 4. Una célula de insecto de acuerdo con una cualquiera de las reivindicaciones anteriores, en la que la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la primera secuencia de nucleótidos tiene un sesgo de uso de codones mejorado para la célula de insecto, en comparación con la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la segunda secuencia de nucleótidos, en la que la diferencia en el índice de adaptación de codones entre la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la
- 30 primera y la segunda secuencias de nucleótidos es al menos 0,2.
5. Una célula de insecto de acuerdo con una cualquiera de las reivindicaciones precedentes, en la que la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la secuencia de nucleótidos con el sesgo de uso de codones mejorado comprende un tramo continuo de al menos 25 codones todos los cuales son codones comunes de conformidad con la Tabla 1 o la Tabla 2, en el que, preferentemente, todos los codones en la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la secuencia de nucleótidos con el sesgo de uso mejorado son codones comunes de acuerdo con la Tabla 1 o la Tabla 2, y en la que, más preferentemente, todos los codones en la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la otra secuencia de nucleótidos son segundos codones más frecuentes de acuerdo con la Tabla 1 o la Tabla 2.
- 35 6. Una célula de insecto de acuerdo con una cualquiera de las reivindicaciones 1 - 5, en la que al menos el 50 % de los codones en la secuencia de nucleótidos que codifica la secuencia de aminoácidos común en la segunda secuencia de nucleótidos está alterada en comparación con el codón correspondiente en la primera secuencia de nucleótidos para maximizar el contenido de AT o GC de la segunda secuencia de nucleótidos.
7. Una célula de insecto de acuerdo con una cualquiera de las reivindicaciones precedentes, en la que la primera y la segunda secuencias de nucleótidos son parte de una única construcción de ácido nucleico.
- 45 8. Una célula de insecto de acuerdo con una cualquiera de las reivindicaciones precedentes, en la que las proteínas Rep de parvovirus son proteínas Rep del virus adenoasociado (VAA), en la que, preferentemente, las proteínas Rep de parvovirus codificadas en la primera y la segunda secuencias de nucleótidos son del mismo serotipo.
9. Una célula de insecto de acuerdo con una cualquiera de las reivindicaciones 1 - 8, en la que la primera secuencia de nucleótidos codifica una proteína Rep52 de parvovirus y se selecciona del grupo que consiste en:
- 50 a) una secuencia de nucleótidos que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos un 70 % con la secuencia de aminoácidos de la SEQ ID NO. 6;
- b) una secuencia de nucleótidos que tiene al menos un 70 % de identidad de secuencia con la secuencia de nucleótidos de una cualquiera de las SEQ ID NO de 1 - 5 y 10;
- c) una secuencia de nucleótidos cuya cadena complementaria hibrida a una secuencia de molécula de ácido nucleico de (a) o (b); y
- 55 d) una secuencia de nucleótidos cuya secuencia difiere de la secuencia de una molécula de ácido nucleico de (c) debido a la degeneración del código genético,

y en la que la segunda secuencia de nucleótidos codifica una proteína Rep78 de parvovirus y se selecciona del grupo que consiste en:

- 5 a) una secuencia de nucleótidos que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos un 70 % con la secuencia de aminoácidos de la SEQ ID NO. 8;
- b) una secuencia de nucleótidos que tiene al menos un 70 % de identidad de secuencia con la secuencia de nucleótidos de las posiciones 11-1.876 de la SEQ ID NO. 7;
- c) una secuencia de nucleótidos cuya cadena complementaria hibrida a una secuencia de molécula de ácido nucleico de (a) o (b);
- 10 d) una secuencia de nucleótidos cuya secuencia difiere de la secuencia de una molécula de ácido nucleico de (c) debido a la degeneración del código genético.

10. Una célula de insecto de acuerdo con una cualquiera de las reivindicaciones 1-9, en la que la célula de insecto comprende además:

- 15 a) una tercera secuencia de nucleótidos que comprende al menos una secuencia de nucleótidos de repetición terminal invertida (ITR) de parvovirus; y
- b) una cuarta secuencia de nucleótidos que comprende las secuencias de codificación de la proteína de la cápside de parvovirus unida operativamente a secuencias de control de la expresión para la expresión en una célula de insecto,

20 en la que, preferentemente, una o más de la primera, segunda, tercera y cuarta secuencias de nucleótidos son parte de una construcción de ácido nucleico que es un vector compatible con células de insectos, preferentemente un vector de baculovirus.

25 11. Una célula de insecto de acuerdo con la reivindicación 10, en la que la tercera secuencia de nucleótidos comprende además al menos una secuencia de nucleótidos que codifica un producto génico de interés y en el que la al menos una secuencia de nucleótidos que codifica un producto génico de interés se incorpora en el genoma de un vector de parvovirus producido en la célula de insecto, en la que, preferentemente, la tercera secuencia de nucleótidos comprende dos secuencias de nucleótidos ITR de parvovirus y en la que la al menos una secuencia de nucleótidos que codifica un producto génico de interés se localiza entre las dos secuencias de ITR de nucleótidos de parvovirus.

30 12. Una célula de insecto de acuerdo con la reivindicación 10 o la reivindicación 11, en la que al menos una de la primera, segunda secuencia, tercera y cuarta secuencias de nucleótidos se integran de forma estable en el genoma de la célula de insecto.

13. Una célula de insecto de acuerdo con una cualquiera de las reivindicaciones 10-12, en la que el parvovirus es VAA.

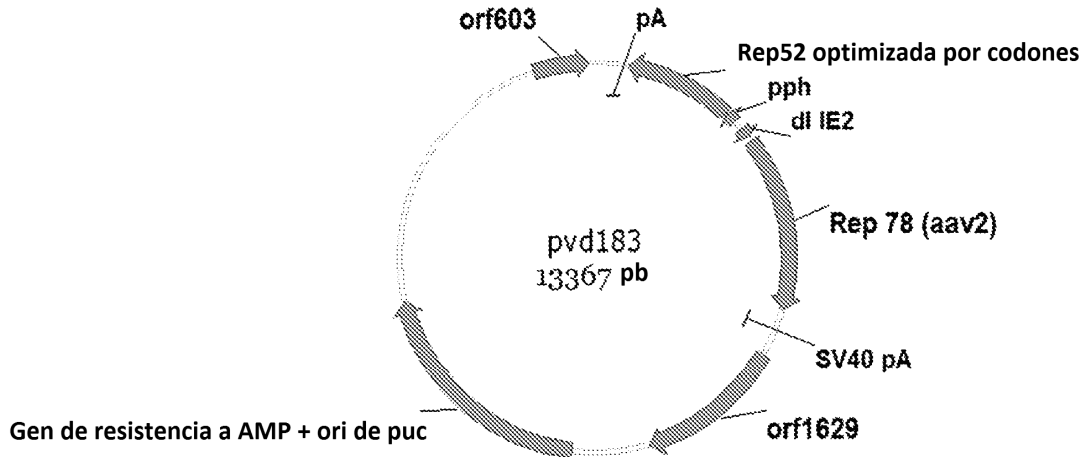
35 14. Un procedimiento para producir un virión de parvovirus recombinante en una célula de insecto que comprende las etapas de:

- a) cultivar una célula de insecto como se define en una cualquiera de las reivindicaciones 10-13 en condiciones tales que se produce el virión de parvovirus recombinante; y
- b) la recuperación del virión de parvovirus recombinante,

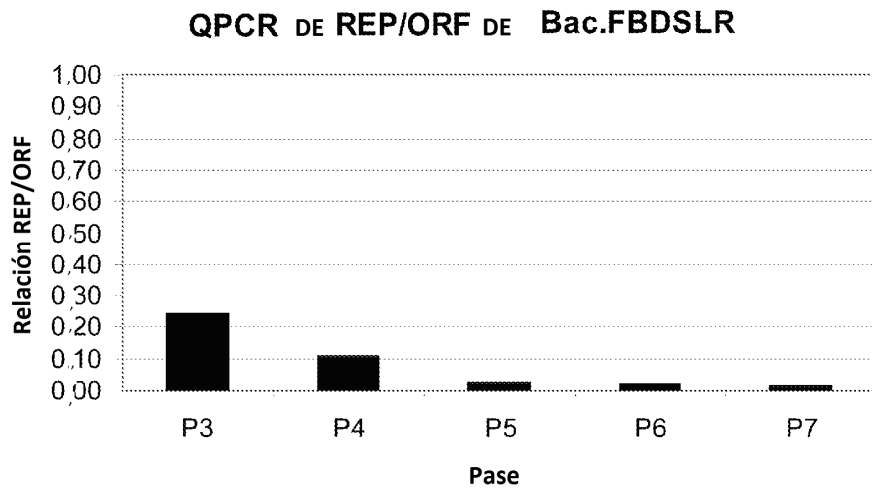
40 comprendiendo, preferentemente, el procedimiento además la etapa de purificación por afinidad del virión utilizando un anticuerpo anti-parvovirus inmovilizado, más preferentemente un anticuerpo de camélido monocatenario o un fragmento del mismo y en el que, preferentemente, el virión de parvovirus recombinante es un virión de VAA recombinante.

15. Una construcción de ácido nucleico que comprende una primera y una segunda secuencia de nucleótidos como se define en una cualquiera de las reivindicaciones 1-9.

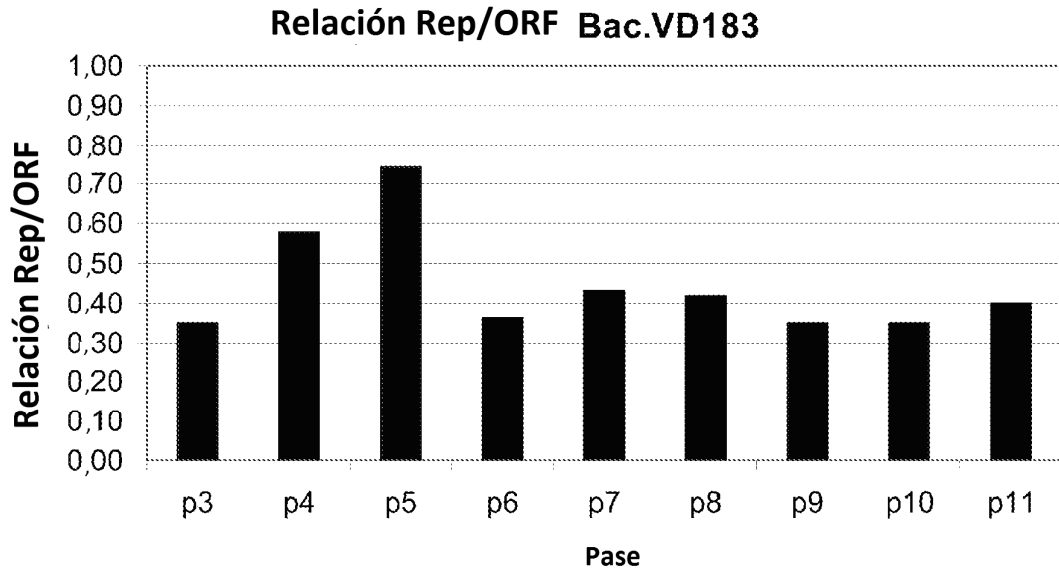
*Fig 1*



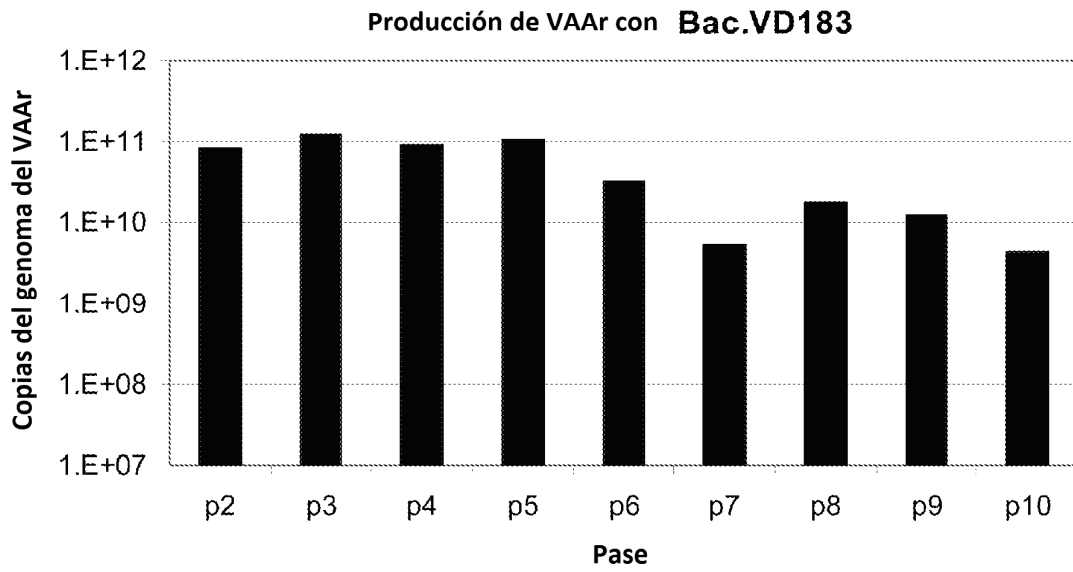
*Fig 2*



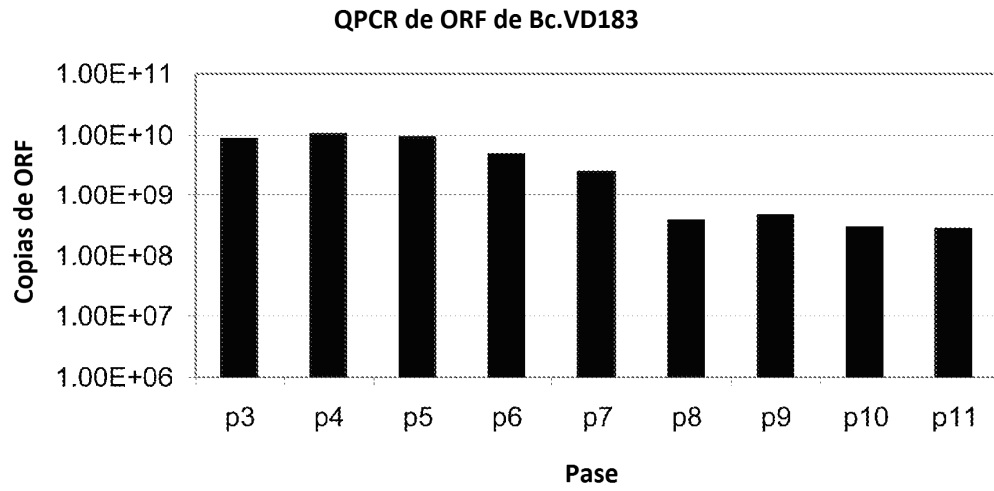
*Fig 3*



*Fig 4*



*Fig 5*



*Fig 6*

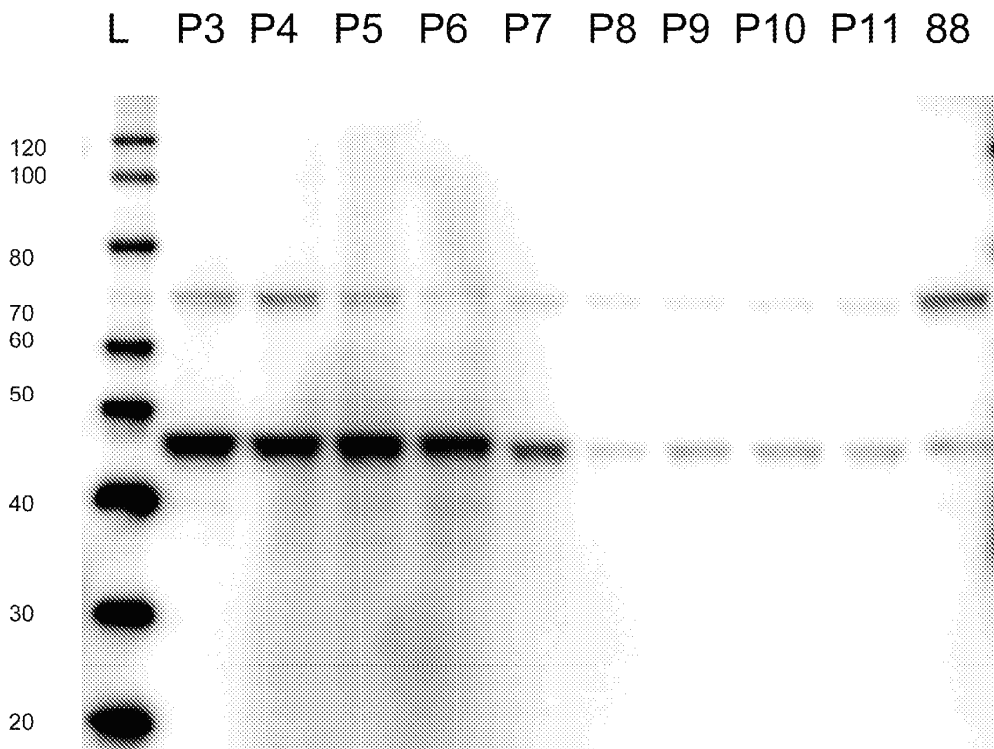




Fig 7

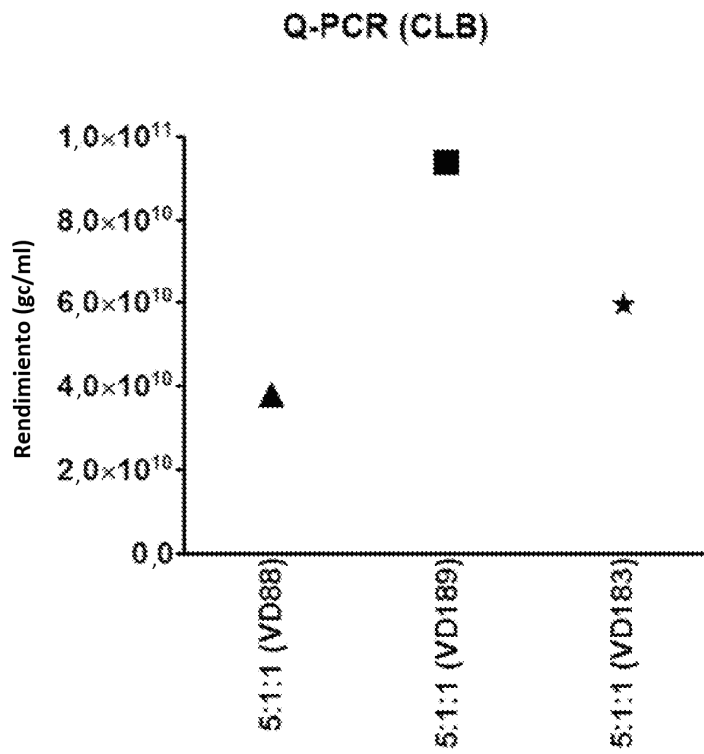
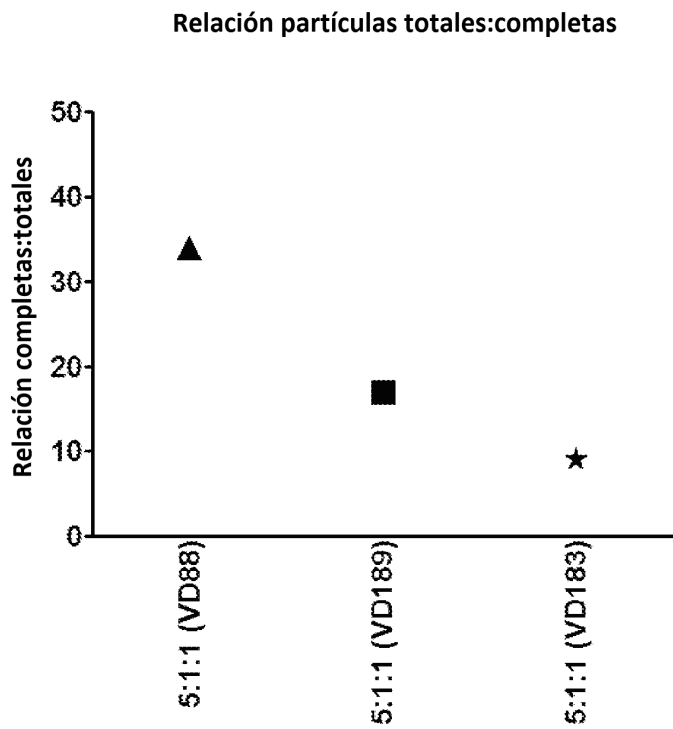
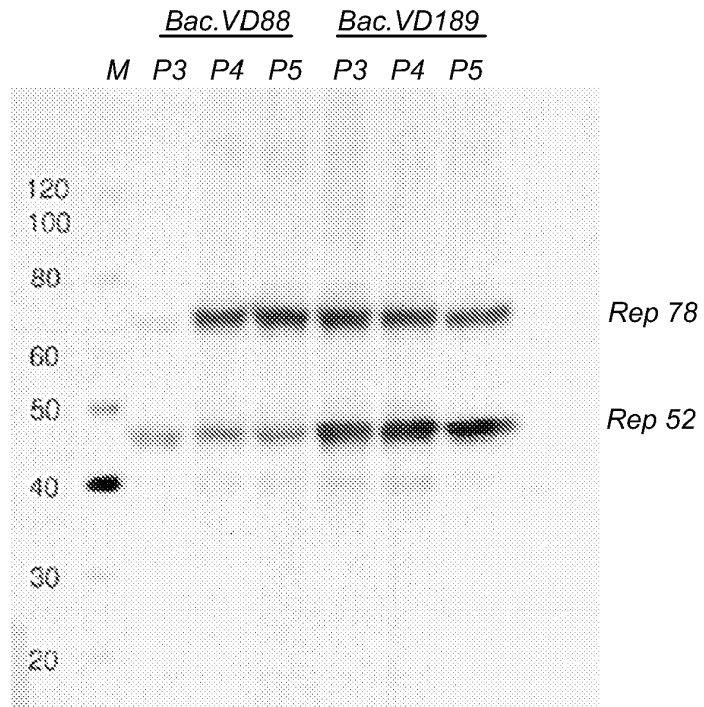


Fig 8



*Fig 9*



*Fig 10*

