(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 534 440**

(51) Int. CI.:

C12N 15/867 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **17.05.2010 E 10721147 (6)**

(97) Fecha y número de publicación de la concesión europea: **07.01.2015 EP 2430167**

(54) Título: **Sistema de vector de ASLV**

(30) Prioridad:

15.05.2009 DE 102009021592

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
22.04.2015

(73) Titular/es:

MEDIZINISCHE HOCHSCHULE HANNOVER (100.0%)
Carl-Neuberg-Strasse 1
30625 Hannover, DE

(72) Inventor/es:

SCHAMBACH, AXEL;
BAUM, CHRISTOPHER y
SÜRTH, JULIA, DEBORA

(74) Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

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

Sistema de vector de ASLV

A continuación se describe un vector autoinactivante (SIN) viral a base del Virus de la Leucosis/Sarcoma Aviar (ASLV), partículas virales que contienen este vector, un sistema de separación-empaquetamiento que comprende o que está compuesto del vector SIN, un primer plásmido auxiliar que sirve para la expresión de la proteína de fusión gag-pol viral y un segundo plásmido auxiliar que sirve para la expresión de la proteína de envoltura (env) retroviral, así como un procedimiento para la producción de la partícula viral mediante el sistema de separación-empaquetamiento, particularmente mediante expresión en células humanas cultivadas. Además, las partículas virales se usan como medicamento, en particular para la transferencia de un transgén a células, así como para la producción del medicamento. El primer y el segundo plásmido auxiliar, por ejemplo contenidos en una línea celular de empaquetamiento o transfectados de forma transitoria, sirven para la generación de partículas virales que no se replican (incompetentes en cuanto a RCR), que contienen ARN con una LTR SIN de acuerdo con la invención en el extremo 3', pudiendo presentar el ARN una sección terapéuticamente eficaz que se denomina, por ejemplo, transgén. Esta LTR SIN 3' incluye una extensa deleción de la región U3 que se copia durante la transcripción inversa en la LTR 5'. Para la expresión de un transgén que, por ejemplo, para el uso para la terapia génica puede ser un gen humano o heterólogo, el vector SIN viral contiene, preferentemente, un casete de expresión que está dispuesto en 5' con respecto a la LTR SIN 3'.

El vector permite también en células de empaquetamiento humanas la generación de partículas virales con un elevado título y se caracteriza por que durante la transducción de células diana, en particular células humanas, mediante puesta en contacto de las células de forma extracorpórea, es decir, *in vitro*, o mediante administración de una partícula preparada para esto a un paciente se generan células transducidas *in vivo* que expresan el transgén a lo largo de un largo periodo de tiempo de forma estable o con reducido debilitamiento, siendo la frecuencia de la transformación de células diana por activación de oncogén reducida. El casete de expresión presenta en 5' delante de la secuencia de ácido nucleico que codifica el transgén un promotor y, preferentemente, en 3' con respecto a la secuencia que codifica el transgén, un elemento regulador post-transcripcional (PRE), en particular un WPRE (PRE del Virus de la Hepatitis de la Marmota Canadiense (WHV)).

Estado de la técnica

Hughes (Folia Biologika 50, 107-119 (2004)) describe una serie de vectores virales a base del Virus de la Leucosis Aviar (ALV) que presentan, para la mejor producción de retrovirus con competencia de replicación (RCR), una LTR funcional. (RCAS: Replication-Competent ASLV LTR with a Splice Acceptor (LTR de ASLV con competencia de replicación con aceptor de empalme)). En este sistema, todos los componentes para la producción del virus se encuentran en un plásmido (como en el virus de longitud completa) y se incluye un posible transgén en la región no traducida (UTR) 3'.

Hu *et al*. (Molecular Therapy, 1617-1623 (2008)) muestran que el ASLV no tiende a insertarse preferentemente dentro o cerca de proto-oncogenes y que las LTR del ASLV no tienden a activar genes adyacentes y, por tanto, llegan a la conclusión de que el ASLV forma una base adecuada para vectores virales. Se señala que una deleción de la región potenciadora (enhancer) en la LTR o la evitación de la transcripción de parte a parte a través de una señal de terminación se podría conseguir mediante una poliadenilación mejorada.

El documento EP 1 757 703 A2 describe vectores retrovirales autoinactivantes, en particular vectores lentivirales y gammarretrovirales, que se caracterizan por que un promotor fuerte sustituye el elemento U3 en la LTR 5' para controlar la transcripción del ARNm de longitud completa para el empaquetamiento.

Schambach *et al*. en Molecular Therapy 391-400 (2006) describen vectores autoinactivantes gammarretrovirales y lentivirales que para la terapia génica contienen un casete de expresión de O^6-metilguanin-ADN-metiltransferasa.

Hildinger *et al*. en J. Virol. 4083-4089 (1999) describen que la presencia de un donador de empalme en combinación con la señal de empaquetamiento ψ es una condición esencial para la generación de altos títulos de partículas virales en células de empaquetamiento.

Hu *et al*. (Human Gene Therapy 691-700 (2007)) y Mitchell *et al*. (Plos Biology, e234, 2004) muestran que el ASLV no tiende a insertarse preferentemente dentro o cerca de proto-oncogenes, en particular sin clara preferencia por zonas próximas al promotor. Por consiguiente, se tiene que estimar menor el riesgo de mutagénesis de inserción en comparación con los sistemas de vector usados clínicamente del virus de la leucemia murina (MLV) y el virus de la inmunodeficiencia humana (VIH). Se señala que una deleción de la región potenciadora (enhancer) en la LTR o la evitación de la transcripción de parte a parte a través de una señal de terminación sería deseable. A causa de la arquitectura del vector (1 plásmido, replicante), por norma general, este sistema no es adecuado para aplicaciones clínicas.

Hu *et al*. muestran que los vectores virales derivados de ASLV con LTR intactas en células de mamífero no se replican. No está clara la causa exacta de esto.

A causa de la arquitectura retroviral de 2 LTR, los retrovirus tienen señales de poliA intrínsecamente débiles. Para vectores lentivirales es sabido que mediante la transcripción de parte a parte del extremo 3' del vector viral insertado puede conducir a la activación génica indeseada, así por ejemplo Zaiss *et al*. (Journal of Virology 7209-7219 (2002)) y Schambach, Baum *et al*. (Molecular Therapy 2007, 15(6): 1167-73). Esto es importante también debido a que la región U3, además de elementos promotores/potenciadores, contiene también reforzadores de la poliadenilación y, de este modo, las posibles deleciones para cada familia de virus se tienen que validar de nuevo con respecto a actividad potenciadora residual y eficiencia de la poliadenilación.

Para vectores gammarretrovirales, Kraunus *et al*. (Gene Therapy 568-1578 (2004)) proponen delecionar toda la zona del elemento potenciador y del promotor de la región U3 3' y dejar solo los primeros 22 pb cadena arriba del potenciador y los últimos 14 pb cadena abajo. En el ASLV o en el Virus del Sarcoma de Rous (RSV), la región LTR se complica por una región R muy corta, de tal manera que la señal de poliA (AATAAA) llega a encontrarse en la región U3. Esto es muy poco habitual para retrovirus (localizados por norma en la región R) y dificulta la construcción de un vector SIN.

Aubert *et al*. (The J. of Cell Biology 113, 497-506 (1991) y Flamant *et al*. (Journal of General Virology 74, 39-46 (1993)) describen deleciones de la región U3 de ALV de -149 a -15, en relación con la región R, o de -149 a -15 y -98 a -59, también en relación con la región R. En todas estas regiones U3 están contenidos todavía elementos potenciadores remanentes (Cullen *et al*., MCB, 1985, 438-47), lo que es problemático para posibles aplicaciones clínicas. De hecho, estos elementos potenciadores pueden regular erróneamente mediante mutagénesis de inserción posiblemente genes adyacentes (por ejemplo, oncogenes).

Objetivo de la invención

Por tanto, la invención se plantea como objetivo facilitar un sistema de separación-empaquetamiento que se pueda emplear clínicamente para la producción en células de mamífero (por ejemplo, células humanas, por ejemplo, HEK293T) con un vector viral y plásmidos auxiliares con el que se pueda producir el vector viral empaquetado como partícula viral, en el que esté minimizado el riesgo de la recombinación hasta dar un retrovirus con competencia de replicación (RCR) y también esté minimizado el riesgo de que la inserción del vector viral conduzca a la activación de genes adyacentes de la célula diana (mutagénesis de inserción). Preferentemente, el vector viral o el sistema con plásmidos auxiliares además debe permitir la generación de partículas virales con un elevado título, habiendo de poder contener el vector viral un casete de expresión por el cual un transgén y secuencias reguladoras de genes (por ejemplo, ARNhc, miARN etc.) se transcriben y/o traduzcan de forma eficaz. Preferentemente, estos vectores se pseudotipifican con proteínas de envoltura que se pueden emplear clínicamente (por ejemplo, RD114/TR, env anfotrópico de MLV, glucoproteína del virus de la estomatitis vesicular (VSVg)). Todos esos criterios se pretenden en combinación, ya que son esenciales para posibles aplicaciones clínicas.

Descripción general de la invención

La invención resuelve este objetivo con las características de las reivindicaciones y en particular con un vector que presenta las propiedades pretendidas. Por un plásmido se puede transcribir de acuerdo con las reivindicaciones un ARN viral que facilita un vector viral a base de ASLV, cuya región U3 3' genera o es una LTR 3' autoinactivante (LTR SIN) así como la partícula viral que contiene una secuencia de ácido nucleico que codifica el vector viral, en particular ARN. Además, la invención facilita un sistema con un primer plásmido auxiliar que en un casete de expresión codifica la secuencia codificante para la proteína de fusión gag-pro-pol de ASLV y con un segundo plásmido auxiliar que codifica en un casete de expresión una proteína de envoltura (env), por ejemplo, env de ASLV, RD114/TR, anfotrópico de MLV, VSVg o ecotrópico de MLV.

Con presencia del primer y del segundo plásmido auxiliar en una célula eucariota, denominada en lo sucesivo también célula de empaquetamiento, se generan partículas virales con un elevado título, cuyos ARN virales debido a las deleciones en la región U3 5' que se produce durante la transcripción inversa hasta ADN viral desde la región U3 3' en una célula diana por sí mismos son autoinactivantes (SIN), sin embargo, permiten un elevado índice de transcripción y una traducción eficaz del transgén que está dispuesto en un casete de expresión en 5' delante de la LTR SIN 3'. En esta configuración está significativamente reducida la activación de secuencias genómicas adyacentes debido a la inserción del vector viral. La reducción de la activación de secuencias adyacentes se ha mostrado para vectores γ-retrovirales por de Zychlinski *et al*. (Mol. Ther. 16: 718-725 (2008)).

Las mutaciones que son conocidas para vectores SIN virales a base de otros retrovirus, por ejemplo, lentivirus, que combinan las propiedades del elevado título en células de empaquetamiento por un lado, ninguna actividad promotora en la LTR así como una eficaz transcripción de un transgén de un casete de expresión interno del vector sin transcripción de parte a parte en secuencias genómicas en el lugar de la inserción, no se pueden transferir a la construcción de vectores de LTR SIN a base de ASLV. De hecho, la disposición de elementos funcionales en los virus individuales y su interacción con factores de hospedador celulares no son idénticos entre sí y los efectos del

cambio de secuencias sobre el título, la evitación de la activación de secuencias genómicas adyacentes en el lugar de la inserción y sobre una transcripción del transgén de un casete de expresión interno no se pueden predecir.

El vector viral contiene, en lugar de la LTR de tipo natural, una LTR 3' cuya región U3 está delecionada a lo largo de una gran sección, de tal manera que está eliminada la actividad transcripcional de las regiones promotoras y potenciadoras de la región U3. La eliminación de las regiones promotoras y potenciadoras de la región U3 está determinada, por ejemplo, por deleción de acuerdo con la reivindicación 1. En particular, de acuerdo con la invención, la región U3 de los 229 nucleótidos de la región U3 de tipo natural (SEC ID Nº 1) de ASLV presenta solo los nucleótidos 1-27 y los nucleótidos 187-229, preferentemente solo los nucleótidos 1-27 y 207-229. Esta región U3 3' modificada del vector viral, en la que de la región U3 de tipo natural están delecionados los nucleótidos 28-186 y están delecionados, preferentemente, además los nucleótidos 187-202, después de la transcripción inversa y la integración en el genoma de una célula diana en esencia no presenta ninguna actividad promotora y/o potenciadora. El ARN viral empaquetable es dirigido por un promotor fuerte dispuesto en 5' que se pierde durante la transcripción inversa o no está contenido en el ARN viral generado por la transcripción. Esto permite la generación de un transcrito que, junto con un primer y segundo plásmido auxiliar, es adecuado para la generación de partículas virales con un elevado título que contienen el ARN viral.

La deleción de las regiones promotoras y potenciadoras opcionalmente puede estar sustituida por una inserción con una secuencia de nucleótidos que no presenta una actividad promotora o potenciadora.

La LTR SIN 3' del vector puede estar compuesta de las secciones que se han mencionado anteriormente de la región U3 de tipo natural delecionada, una región R de tipo natural y una región U5 de tipo natural. La región U3 de acuerdo con la invención con deleciones puede contener opcionalmente entre las secciones remanentes de la región U3 de tipo natural secuencias de nucleótidos adicionales sin actividad promotora que están seleccionadas, por ejemplo, del grupo que comprende miARN, ARNhc, potenciadores de la poliadenilación (USE), secuencias de reconocimiento de recombinasa (por ejemplo, LoxP, Rox, FRT) y elementos de aislador (por ejemplo, cHS4).

Para la inserción de secuencias de nucleótidos sin actividad promotora en la región U3 de acuerdo con la invención, la misma presenta preferentemente un sitio de corte único para una restrictasa, de forma particularmente preferente debido a que los nucleótidos 187-190 en la numeración de la región U3 de tipo natural mediante mutación dan la secuencia de bases CGTA, por lo que se genera un sitio de corte secundario para SnaBI con la secuencia de reconocimiento TACGTA.

Como alternativa a la deleción de los nucleótidos 187-202 se pueden mutar también los nucleótidos 200-206 (TATTTAA), por ejemplo, en la secuencia TGTCTAA. Se supone que los nucleótidos 200-206 forman la caja TATA de la región U3 de tipo natural, de tal manera que una mutación que cambia esta caja TATA reduce o elimina esta parte de la actividad promotora de la región U3.

La región U3 delecionada de acuerdo con la invención presenta también un sitio de ataque de integrasa (integrase attachment), en particular en los nucleótidos 1-25, así como una señal de poliA, por ejemplo, en los nucleótidos 223-228, respectivamente en la numeración de los nucleótidos de la región U3 de tipo natural. Se supone que la conservación del sitio de ataque de integrasa y de la señal de poliA es importante para la infección viral. Sorprendentemente, el vector permite por un lado una producción de partículas virales con elevado título en una célula de empaquetamiento o línea celular de empaquetamiento, por otro lado, después de transcripción inversa e integración en el genoma en una célula diana, una transcripción eficaz de un transgén de un casete de expresión por un promotor fuerte interno, sin que se produzca una transcripción de parte a parte a través de la LTR SIN 3' en regiones limitantes del sitio de integración o la activación de secuencias genómicas adyacentes.

Las partículas virales generadas en células de empaquetamiento, que contenían ARN viral con una LTR SIN 3' de acuerdo con la invención o una región U3 3' de acuerdo con la invención, sin concentración presentaban títulos de aproximadamente $1x10^6$ partículas infecciosas/ml. A este respecto, el título se vio influido por el gen que codifica la proteína de fusión gag-pro-pol o la proteína de envoltura del primer o del segundo plásmido auxiliar o la línea celular de empaquetamiento y no de forma significativa por el tipo o el grado de la deleción en la región U3.

Sorprendentemente, la expresión de un transgén por un casete de expresión en vectores virales de acuerdo con la invención con deleción SIN era aproximadamente un factor 3 mayor que en un vector comparativo en el que la LTR 3' presentaba una región U3 con secuencia de tipo natural. Además, de acuerdo con la invención se prefiere que las secuencias codificantes del primer y/o segundo plásmido auxiliar estén optimizadas en cuanto a codones, de tal manera que los codones virales estén sustituidos por codones de la célula diana, en particular estén sustituidos por codones humanos, por ejemplo, en procedimientos para la producción de partículas virales y vectores virales de acuerdo con la invención que se usan, por ejemplo, como medicamento o en la producción de composiciones farmacéuticas. Un plásmido auxiliar preferente de acuerdo con la invención contiene una secuencia que codifica la proteína de fusión gag-pro-pol, en la que los codones virales están adaptados al uso de codones humano.

En el uso como medicamento, la partícula viral que contiene el vector viral puede estar producida para la administración, por ejemplo, a un paciente/ser humano o un animal no humano. Como alternativa, la partícula viral o

el vector viral se puede poner en contacto con células que proceden de un paciente de este tipo, es decir, para la transducción *in vitro* de células que a continuación, opcionalmente con la etapa del cultivo y/o de la selección, se preparan o facilitan como células transducidas para la transferencia a un paciente, siendo preferentemente el paciente el mismo del cual proceden las células.

En la secuencia codificante optimizada en cuanto a codones para la proteína de fusión gag-pro-pol, de acuerdo con la invención, el desplazamiento original de la fase de lectura en -1 y la transición de pro a pol del ASLV de tipo natural se han conservado.

Se ha mostrado que la adaptación del uso de codones de ASLV al uso de codones de la célula diana, en particular al uso de codones humano, tiene, para las zonas codificantes de los plásmidos auxiliares, como consecuencia un aumento significativo del título de partículas virales, conduciendo por ejemplo únicamente la adaptación del uso de codones de la proteína de fusión gag-pro-pol al uso de codones humano a un título un factor 50 mayor de partículas virales en células de empaquetamiento humanas.

La adaptación del uso de codones al uso de codones humano se realiza preferentemente al agotarse preferentemente las reservas de ARNt no limitados, óptimas para *Homo sapiens*; de este modo, por ejemplo, se adaptó el codón de glicina GGT preferentemente a GGC. Además se eliminaron motivos de inestabilidad de ARN negativos así como sitios de poliadenilación crípticos así como estructuras secundarias de ARN, de tal manera que se generó un ARNm más estable con una mejor capacidad de traducción en células humanas. Además, en la secuencia codificante se inactivó el sitio donador de empalme natural, el sitio aceptor de empalme y otras señales de empalme crípticas. Por consiguiente, un plásmido auxiliar preferente presenta un sitio donador de empalme inactivo en el gen gag-pol, por ejemplo, de forma correspondiente a los nucleótidos 1501 a 1502 en la SEC ID Nº 14. Preferentemente, gag se codifica por los nucleótidos 1475 a 3436 de la SEC ID Nº 14 y pol, por los nucleótidos 3738 a 6291 de la SEC ID Nº 14, ya que en estas secciones se han retirado sustancialmente los motivos de inestabilidad de ARN crípticos y sitios de poliadenilación así como estructuras secundarias de ARN, de tal manera que se puede generar un elevado título de partículas virales que presentan una elevada eficacia de transducción.

Finalmente se prefiere que el vector viral, con preferencia adicionalmente el primer plásmido auxiliar y, con mayor preferencia, adicionalmente el segundo plásmido auxiliar no presenten ninguna homología de secuencia con la célula diana y además preferentemente no tengan en particular ningún solapamiento de secuencia, por ejemplo, ninguna sección de secuencia homóloga entre sí de más de 5 nucleótidos, preferentemente ninguna sección de secuencia homóloga entre sí de más de 10 a 20 nucleótidos. De este modo se minimiza la probabilidad de una recombinación en células de empaquetamiento y/o las células diana y, por tanto, también la probabilidad de la generación de RCR.

La invención pone a disposición con ello un ARN viral que presenta o está compuesto de, de 5' a 3', una región R 5' y una región U5 5', un sitio de unión de cebador (PBS), una señal de empaquetamiento (ψ) que, a diferencia del estado de la técnica en el caso de vectores gammarretrovirales o lentivirales trabaja sin un donador de empalme retroviral, un casete de expresión para un o con un transgén y una región U3 3' acortada de acuerdo con la invención (denominada Δ o d), una región R 3' y una sección de poliA unida (cola de poliA), así como un vector para la transcripción del ARN viral que, adicionalmente a una secuencia de ácido nucleico que codifica el ARN viral, en 5' delante de su región R 5' presenta un promotor eucariota o viral y en el que la secuencia de ácido nucleico que codifica el ARN viral está presente en forma de ADN que en la LTR SIN 3' presenta una sección de poliA.

Además, la invención facilita el uso del ARN viral y una partícula viral que contiene el ARN viral como medicamento y para la producción de una composición farmacéutica para el uso para la transducción de células diana, por ejemplo, de células somáticas de un ser humano, así como composiciones farmacéuticas que comprenden ARN viral o partículas virales que contienen el ARN viral, y el uso de tales composiciones farmacéuticas como medicamento para la transferencia del transgén a células, en particular para la terapia génica.

Además, la invención se refiere a células que presentan una secuencia de ácido nucleico integrada o no integrada que se han generado mediante puesta en contacto con partículas virales de acuerdo con la invención y que contienen una región U3 de acuerdo con la invención, por ejemplo, una sección de ADN que se ha producido mediante transcripción inversa a partir de un ARN viral de acuerdo con la invención. Esto podrían ser células animales extracorpóreas o intracorpóreas, a excepción de células germinales humanas.

El casete de expresión presenta un promotor y al menos una secuencia transcribible y/o traducible de un transgén, que puede ser un gen homólogo o heterólogo a la célula diana, y preferentemente cadena abajo de la secuencia transcribible del transgén, un PRE u otras secciones o módulos de ácido nucleico reguladores post-transcripcionales. La secuencia transcribible puede presentar también sitios de entrada de ribosomas internos (IRES) junto con secuencias codificantes de proteína, dispuestas cadena abajo con respecto a los mismos, así como otras secuencias de ácido nucleico funcionales o secuencias de ácido nucleico que codifican enzimas (por ejemplo, promotores bidireccionales, secuencias codificantes para sitios de corte de proteasa 2A o genes de proteasa 2A). Además, opcionalmente en la región U3 3' delecionada, por ejemplo, en el sitio de corte por restricción contenido allí se puede incluir una secuencia de aislador que debilita una interacción del transgén con secuencias genómicas

adyacentes, o puede estar introducida una secuencia de reconocimiento de recombinasa que comprende, por ejemplo, una secuencia de reconocimiento del grupo que comprende rox, loxP, FRT. Después de la transcripción inversa realizada y la copia en la LTR 5', mediante adición o expresión de recombinasa en células diana se puede activar la secuencia de reconocimiento de recombinasa y, por ejemplo, se puede invertir o retirar una secuencia de ácido nucleico dispuesta entre secuencias de reconocimiento de recombinasa en el ADN integrado (por ejemplo, una parte del casete de expresión). Otras secuencias que se pueden incluir en la región U3 3' comprenden, por ejemplo, secuencias diana de micro-ARN, ARN horquillado corto o secciones de secuencia que facilitan la comprobación de la inserción a través de PCR.

La región U3 delecionada contenida en la LTR SIN 3' de acuerdo con la invención permite, a pesar de la deleción de los elementos potenciadores y promotores de la región U3 de tipo natural, la generación de partículas virales con títulos que se consiguen frente a los títulos para vectores con LTR de tipo natural en células de empaquetamiento, de tal manera que se evitan las reducciones espectaculares del título viral observadas con frecuencia en caso de deleciones de regiones U3.

La transcripción de transgenes que están dispuestos en el casete de expresión se controla mediante el promotor del casete de expresión, que se denomina también promotor interno. Actualmente se supone que las regiones U3 delecionadas de acuerdo con la invención presentan una señal de poliA que evita una transcripción de parte a parte más allá de la secuencia del vector y que se aprovecha la señal de poliA correcta. Esta suposición se respalda por la expresión 3 veces mejor de un transgén del casete de expresión de un vector de acuerdo con la invención en comparación con la expresión del transgén de un vector con LTR 3' de tipo natural.

De forma particularmente preferente, la secuencia de ácido nucleico del vector viral presenta una deleción completa de las secuencias que codifican proteínas estructurales virales, incluyendo el sitio donador de empalme (SD), por ejemplo, una deleción completa de las secuencias gag, env, pol y src. Se ha mostrado que con vectores de ASLV con deleción de las secuencias que codifican las proteínas estructurales virales se deleciona también el sitio donador de empalme. Sorprendentemente, se ha mostrado que un vector de este tipo a base de ASLV permite el empaquetamiento del ARN viral en células de empaquetamiento mediante el sistema de separación-empaquetamiento. Esto es inesperado en el sentido de que en vectores retrovirales conocidos, partes de las secuencias virales codificantes son necesarias para el empaquetamiento, ya que en la región de las secuencias que codifican las proteínas estructurales virales están contenidas señales para la exportación del ARN, la interacción con la nucleocápside viral y la transcripción inversa, tal como, por ejemplo, en VIH. Por tanto, el vector puede contener la señal de empaquetamiento (ψ), que está comprendida en la SEC ID Nº 5 por los nucleótidos 1042-1292, sin secciones de secuencias que codifican proteínas estructurales virales.

La forma de realización preferente de acuerdo con la invención en la que la secuencia de ácido nucleico del vector viral no presenta ningún donador de empalme viral es una diferencia significativa con respecto a vectores lentivirales o γ-retrovirales conocidos que presentan el donador de empalme viral natural cadena abajo del sitio de unión de cebador. Estos vectores conocidos necesitan el donador de empalme para la generación de elevados títulos de partículas virales. Sorprendentemente, se ha mostrado que el vector también sin su donador de empalme natural u otro da títulos elevados en células de empaquetamiento. Ya que el donador de empalme natural de ASLV se encuentra en el gen gag, su deleción completa conduce también a la eliminación del donador de empalme. Además, preferentemente de acuerdo con la invención tampoco el aceptor de empalme viral está contenido en el vector. El aceptor de empalme de ASLV se encuentra en el gen pol, de tal manera que su deleción completa también conduce a la eliminación del aceptor de empalme.

En el vector, la deleción de donador de empalme y aceptor de empalme conduce a la optimización de la producción de partículas virales en células de empaquetamiento y otorga al vector adicionalmente una mejor seguridad biológica, ya que se dificulta la recombinación hasta retrovirus con competencia de replicación. También en el caso de inserciones intragénicas del vector es menos probable una posible interferencia con señales de empalme celulares.

El vector en la zona 5' presenta una región R, una región U5 y la señal de empaquetamiento en una secuencia de ácido nucleico que no codifica ninguna proteína estructural viral y a la que le sigue por ejemplo directamente, en particular después de la región líder, en 3' el casete de expresión del transgén. Se ha mostrado que el vector viral en su extremo 5' puede presentar una región R, región U5 y una señal de empaquetamiento que está comprendida por los nucleótidos Nº 1042 a 1292 de la SEC ID Nº: 5 o está compuesta de los mismos, estando dispuesto preferentemente en 3' limitando con esta secuencia el casete de expresión con el transgén.

Para los fines de la invención se indican las disposiciones de elementos genéticos de 5' a 3', siempre que no sed describa de otro modo; las secuencias de nucleótidos están indicadas de 5' a 3'.

Descripción detallada de la invención

La invención se describe ahora de forma más detallada mediante ejemplos con referencia a las figuras, en las que

- La Figura 1 muestra la estructura esquemática del tipo natural de ASLV,
- La Figura 2 muestra la estructura esquemática de un vector SIN y un plásmido que contiene el vector,
- La Figura 3 muestra la estructura esquemática de un primer plásmido auxiliar,
- La Figura 4 muestra la estructura esquemática de un segundo plásmido auxiliar,
- La Figura 5 muestra una comparación de la región U3 de tipo natural (U3) con la región U3 delecionada (dU3) de acuerdo con la invención, de la región U3 delecionada de acuerdo con la invención sin caja TATA (dU3 noTATA) y con la región U3 delecionada de acuerdo con la invención con caja TATA mutada (dU3 TATAmut),
- La Figura 6 muestra un mapa de plásmido de un vector de acuerdo con la invención en el que el casete de expresión EGFP codifica como ejemplo un transgén,
- La Figura 7 muestra un mapa de plásmido con el vector de acuerdo con la invención en el que está mutada la caja TATA de la región U3,
- La Figura 8 muestra un mapa de plásmido con vector de acuerdo con la invención, en cuya región U3 está delecionada adicionalmente la caja TATA,
- La Figura 9 muestra un mapa de plásmido de un primer plásmido auxiliar para la traducción de la proteína de fusión gag-pol,
- La Figura 10 muestra un mapa de plásmido de un primer plásmido auxiliar en el que la secuencia que codifica la proteína de fusión gag-pol está optimizada en cuanto a codones para una célula de empaquetamiento y/o diana humana,
- La Figura 11 muestra gráficamente la magnitud de la expresión de un transgén (EGFP) del vector de acuerdo con la invención en comparación con la expresión de un vector viral con LTR 3' de tipo natural,
- La Figura 12 muestra gráficamente la magnitud de la expresión de un transgén (EGFP) de vectores SIN de acuerdo con la invención,
- La Figura 13 muestra una representación de la estructura integrada de un vector de acuerdo con la invención en comparación con la estructura de un vector SIN lentiviral o γ-retroviral,
- La Figura 14 muestra esquemáticamente construcciones de ácido nucleico de ADN para la transcripción de ARN viral con casetes de expresión que contiene funcionalmente en 5' con respecto a un gen indicador la secuencia de ácido nucleico que se debe ensayar en cuanto a actividad promotora y potenciadora y
- La Figura 15 muestra la magnitud de la expresión del gen indicador en células que están transfectadas con ARN viral de la Figura 14 y que mide la actividad promotora/potenciadora.

En la descripción de la U3 SIN de acuerdo con la invención se efectúa la numeración de los nucleótidos en relación con la secuencia de nucleótidos de la U3 de tipo natural de ASLV y correspondientemente, las deleciones o secciones de la región U3 SIN están indicadas correspondientemente a la numeración de los nucleótidos de la U3 de tipo natural.

La Figura 1 muestra la estructura esquemática del ASLV de tipo natural en el que las LTR contienen una región U3, una región R y una región U5 limitando entre sí. Las LTR son respectivamente idénticas, ya que en la replicación viral la región U3 de la LTR 3' se copia en el extremo 5' y las regiones R y U5 de la LTR 5' se copian en la LTR 3'. El inicio de la transcripción +1 define el extremo 5' de la región R 5'.

La Figura 2 muestra el vector a nivel de plásmido, en cuyo extremo 5' está dispuesto un promotor delante de la región R para posibilitar la transcripción a partir de la región R 5'. Este promotor 5' de la LTR 5', a causa del inicio de la transcripción en +1 como primer nucleótido de la región R 5' no está contenido en el ARN viral; el ARN viral que se genera está compuesto de las zonas entre el nucleótido +1 de la región R 5' y el extremo 3' de la región R 3' más cola de poliA. Correspondientemente, las partículas virales contienen un ARN viral que se produce a partir de la sección del vector que comprende la región R 5' (incluyendo el nucleótido +1) hasta el extremo 3' de la región R 3' con cola de poliA adjunta. El promotor mostrado en la Figura 2 cadena arriba de la región R 5' del vector viral, que preferentemente está compuesto de un potenciador de SV40 y una región U3 del tipo natural de RSV (Virus del Sarcoma de Rous), no se convierte en parte del ARN viral, sino que controla la transcripción eficaz del ARN viral de una secuencia de ADN correspondiente en la Figura 2. En el presente caso está mostrado el vector en el nivel de plásmido, como una secuencia de ADN que codifica el ARN viral que puede ser parte de un plásmido del cual se transcribe el ARN viral.

Como se indica en la Figura 2, el vector viral entre las LTR y correspondientemente después de la región líder contiene un casete de expresión que presenta al menos un promotor interno, por ejemplo, el promotor de SFFV y EFS (factor de elongación 1α), con una secuencia de nucleótidos que codifica un transgén (gen), preferentemente con un PRE dispuesto en 3'.

El PRE, que preferentemente es WPRE, del casete de expresión presenta preferentemente una deleción X-ORF y/o mutaciones de ATG:

El vector y el ARN viral de acuerdo con la invención presentan al menos en la LTR 3' una región U3 con deleciones que inactivan cualquier actividad promotora de la LTR. En 5' delante de la LTR SIN 3' y/o dentro de la región U3 3', el vector viral puede contener elementos genéticos homólogos y/o heterólogos, por ejemplo, secuencias que comprenden miARN, ARNhc, ARN potenciadores de la poliadenilación, secuencias de reconocimiento de recombinasa, USE y/o aisladores.

Entre los elementos de la LTR 5' y de la LTR 3', preferentemente entre las regiones de la LTR 5' y el casete de expresión, el vector presenta preferentemente un sitio de unión de cebador y una señal de empaquetamiento (ψ).

A diferencia de los retrovirus con competencia de replicación, el vector viral no presenta ninguna secuencia que codifique gag-pol y/o proteínas de envoltura.

Las proteínas virales que se requieren para la generación de una partícula viral del ARN viral son codificadas por plásmidos auxiliares independientes, preferentemente un primer plásmido auxiliar que presenta una secuencia que codifica la proteína de fusión gag-pro-pol en un casete de expresión entre un promotor y una secuencia de poliadenilación, tal como se muestra en la Figura 3, y un segundo plásmido auxiliar que contiene una secuencia que codifica la proteína de envoltura (env) en un casete de expresión, tal como se muestra por ejemplo en la Figura 4, entre un promotor de CMV (citomegalovirus) y una secuencia de poliadenilación. Para la pseuditipificación, la secuencia que codifica la proteína de envoltura puede estar compuesta de una secuencia que codifica la proteína de envoltura de MLV ecotrópico/anfotrópico, VSVg, RD114, GALV o quimeras de los mismos, por ejemplo, RD114/TR.

El primer plásmido auxiliar presenta en su secuencia codificante preferentemente el desplazamiento de fase de lectura de -1 entre el extremo 3' de la secuencia que codifica gag-pro y el extremo 5' de la secuencia que codifica pol, lo que se muestra en la Figura 3 mediante las cajas desplazadas entre sí para gag-pro o pol.

De forma particularmente preferente, las secuencias codificantes de los plásmidos auxiliares están adaptadas al uso de codones de la célula diana, es decir, optimizadas en cuanto a codones, ya que para la invención se ha mostrado que el título debido a esta optimización de codones para la proteína de fusión gag-pro-pol, preferentemente también para la proteína de envoltura env, obtiene un aumento de título en células auxiliares de aproximadamente el factor 50. Por tanto, la invención posibilita un procedimiento de producción en títulos o cantidades de partículas virales, tal como se requieren para aplicaciones clínicas, mediante producción de partículas virales que contienen el ARN viral de acuerdo con la invención, con un primer y un segundo plásmido auxiliar, por ejemplo, mediante cultivo de una célula de empaquetamiento.

La Figura 5 muestra una visión general de las secuencias preferentes de la región U3 3' de acuerdo con la invención del vector viral en comparación con la secuencia de la U3 de tipo natural (SEC ID Nº: 1) de ASLV. Las regiones U3 de acuerdo con la invención presentan preferentemente una deleción de los nucleótidos 28-186 de la U3 de tipo natural. La forma de realización SEC ID Nº: 2, denominada delta (d) U3, preferentemente está compuesta de los nucleótidos 1-27 y los nucleótidos 187-229 de la U3 de tipo natural, estando mutados también los nucleótidos 187 y 189 en C o T, de tal manera que resulta allí un sitio de corte por restricción para SnaBI (TACGTA). Esta región U3 delecionada, denominada también U3 SIN, presenta todavía la caja TATA original que se identificó como los nucleótidos 200-206 de la U3 de tipo natural.

En una forma de realización preferente está mutada también la caja TATA de la U3 SIN, de tal manera que se elimina la función de una caja TATA. En la SEC ID Nº: 3 (dU3 TATAmut) está mostrada una realización preferente de una caja TATA mutada, en la que están mutados los nucleótidos 201-203, de tal manera que se destruye la función de la caja TATA.

En una forma de realización más preferente, la deleción en relación con la región U3 de tipo natural comprende los nucleótidos 28-205 (dU3 no TATA; SEC ID Nº: 4), estando mutados los nucleótidos 203-204 para evitar la generación de una caja TATA mediante secciones que limitan entre sí y/o para producir un sitio de corte en los nucleótidos adyacentes entre sí 26-27 y 203-206, por ejemplo, mediante mutación de los nucleótidos 203-205 en CGT para producir un sitio de corte singular para la enzima de restricción SnaBI (TACGTA).

Los vectores virales comprenden, por tanto, también variantes en las que en la zona de la deleción de los nucleótidos 28-222, preferentemente en la zona de los nucleótidos 28-190 o 28-206 en la SEC ID Nº: 1 pueden estar insertados elementos genéticos heterólogos u homólogos, por ejemplo, heterólogos u homólogos en relación con la célula diana, preferentemente mediante un sitio de corte por restricción en las secciones que se han mencionado anteriormente.

Ejemplo 1: Producción de partículas virales con vector SIN de ASLV

Como ejemplos de vectores de acuerdo con la invención que codifican un ARN viral de acuerdo con la invención se construyeron plásmidos que estaban bajo el control de un promotor que estaba compuesto de un elemento potenciador de SV40 (SV40 enhancer) y el promotor de RSV (RSV) y directamente limitando con el promotor presentaban la región R, la región U5, un casete de expresión para un transgén y una LTR SIN 3' de una región U3 SIN de acuerdo con la invención, una región R y una región U5. El casete de expresión comprendía el promotor U3 del virus formador de focos esplénicos (SFFV), una secuencia codificante de EGFP como ejemplo de un transgén y un WPRE (PRE).

La región U3 SIN se correspondía con la SEC ID Nº: 2 para el vector mostrado en la Figura 6, como alternativa con la SEC ID Nº: 3 para el vector mostrado en la Figura 7 o con la SEC ID Nº: 4 para el vector mostrado en la Figura 8.

La secuencia del vector de la Figura 6 está contenida en la SEC ID Nº: 5. Correspondientemente, un vector o un plásmido que codifica un ARN de acuerdo con la invención contiene una LTR SIN 3' con una secuencia que se corresponde con los nucleótidos 3282-3452 de la SEC ID Nº: 5. El ARN viral transcrito por ello contiene la región U3 3' de acuerdo con la invención contenida en su interior.

La secuencia del vector de la Figura 7 está contenida en la SEC ID Nº: 9. Correspondientemente, un vector por el cual se transcribe un ARN viral de acuerdo con la invención contiene una LTR SIN 3' con una secuencia correspondiente a los nucleótidos 3282-3452 de la SEC ID Nº: 9. El ARN viral contiene en la LTR SIN 3' la región U3 3' de acuerdo con la invención y la región R, sin embargo no la región U5.

La secuencia del vector de la Figura 8 está contenida en la SEC ID Nº: 7. Correspondientemente, un vector por el cual se transcribe un ARN viral de acuerdo con la invención contiene una LTR SIN 3' con una secuencia correspondiente a los nucleótidos 3282-3436 de la SEC ID Nº: 7.

Además, un vector viral o un ARN viral contiene una región R de tipo natural y una región U5 de tipo natural como LTR 5' y, limitando directamente con esto en 5', en la forma en la que la secuencia de nucleótidos para el ARN viral está contenida en un plásmido, un promotor eucariota o viral, preferentemente la secuencia potenciadora de SV40 junto con el promotor de RSV, correspondiente a los nucleótidos 452-921 de la SEC ID Nº: 9.

La secuencia que codifica eGFP de la SEC ID Nº: 5 está indicada como secuencia de aminoácidos SEC ID Nº: 6, la eGFP de la SEC ID Nº: 7 como SEC ID Nº: 8 y la eGFP de la SEC ID Nº: 9 como SEC ID Nº: 10.

Para la generación de partículas virales se introdujeron los plásmidos según las Figuras 6 a 8 en células HEK293T que mediante transfección transitoria o integración estable contenían también un primer plásmido auxiliar para la expresión de la proteína de fusión gag-pol de ASLV. En la Figura 9 está mostrado un mapa esquemático de plásmido, la secuencia se adjunta como SEC ID Nº: 11 y contiene como secuencia codificante la proteína de fusión gag-pol en los nucleótidos 1475-6291. La secuencia de aminoácidos codificada por la SEC ID Nº: 11 de gag se corresponde con la SEC ID Nº: 12, la secuencia de aminoácidos codificada por la SEC ID Nº: 11 de pol se corresponde con la SEC ID Nº: 13.

De forma particularmente preferente, el primer plásmido auxiliar estaba provisto de una secuencia codificante para la proteína de fusión gag-pol, cuyos codones estaban adaptados al uso de codones de la célula de empaquetamiento, en particular de células de mamífero, tales como, por ejemplo, células humanas. Un plásmido de este tipo está mostrado esquemáticamente en la Figura 10; la secuencia de ácido nucleico está contenida como SEC ID Nº: 14. Correspondientemente, un primer plásmido auxiliar preferente o una célula de empaquetamiento humana en forma integrada contiene la secuencia codificante para la proteína gag optimizada en cuanto a codones correspondiente a los nucleótidos 1475-3436 de la SEC ID Nº: 14, cuya secuencia de aminoácidos se corresponde con la SEC ID Nº: 15, y/o la secuencia codificante para la proteína pol optimizada en cuanto a codones correspondiente a los nucleótidos 3738-6291 de la SEC ID Nº: 14, cuya secuencia de aminoácidos se corresponde con la SEC ID Nº: 16, preferentemente la secuencia codificante para la proteína de fusión gag-pol optimizada en cuanto a codones correspondiente a los nucleótidos 1475-6291 de la SEC ID Nº 14.

En condiciones de cultivo habituales de la célula de empaquetamiento, para las partículas virales de acuerdo con la invención con una de las regiones U3 SIN mostradas de los vectores de la SEC ID Nº: 7, 9 u 11 se obtuvo respectivamente el mismo título que para una construcción comparativa (no mostrada) en la que la región U3 3' se correspondía con el tipo natural de ASLV, en concreto aproximadamente 1 x 10^6/ml.

Por tanto, para el procedimiento de acuerdo con la invención para la producción de partículas virales se prefiere que las células de empaquetamiento sean células humanas cultivadas que presenten una secuencia de ácido nucleico que codifique los genes estructurales virales, por ejemplo, gag-pol, con uso de codones humano, no presentando estos genes estructurales preferentemente ningún donador de empalme activo y/o sitio aceptor de empalme para evitar un empalme dentro de las secuencias codificantes y alcanzar, por tanto, una elevada expresión en las células de empaquetamiento.

Para el control de la transfección de células diana con ARN viral se infectaron fibroblastos HT1080 humanos. A continuación del cultivo, después de la transfección se determinó la expresión del transgén (EGFP) mediante FACS. En la comparación con partículas virales, cuya LTR 3' era de ASLV de tipo natural, se pudo mostrar para los vectores virales descritos con LTR SIN 3' una expresión un factor 3 veces mayor del transgén.

La Figura 11 muestra la expresión del transgén ilustrativo EGFP de un casete de expresión correspondiente a la Figura 7 en células diana que estaban transducidas con vector viral. Se obtuvieron partículas virales del sobrenadante de células de empaquetamiento (pseudotipificadas con VSVg), que contenían un plásmido que codificaba ARN viral de acuerdo con la invención de acuerdo con la SEC ID Nº: 5 (ASLV dU3, barras derechas claras) o un ARN viral con U3 3' de tipo natural (ASLV U3, barras izquierdas oscuras; U3 de tipo natural véase la SEC ID Nº: 1). Para la transducción se ajustaron partículas virales en células HT1080E, se cultivaron durante 6 días y a continuación se analizaron mediante citometría de flujo. En el experimento 1, la proteína de fusión gag-pol de

tipo natural (nucleótidos Nº 1475 a 6291 en la SEC ID Nº: 11), en el experimento 2 estaba optimizada en cuanto a codones (nucleótidos Nº 1475 a 6291 en la SEC ID Nº 14). El resultado, que la partículas viral de acuerdo con la invención con una región U3 3' de acuerdo con la invención frente a la partícula viral con región U3 de tipo natural causaba una expresión aproximadamente 3 veces mayor del transgén en células diana. Esto se cumple tanto para el uso de gag/pol de tipo natural (experimento 1) como del plásmido auxiliar gag-pol optimizado en cuanto a codones (experimento 2). En ambos casos se produce un aumento de aproximadamente 3 veces de la expresión del transgén en células diana para partículas virales de acuerdo con la invención (dU3) en comparación con partículas comparativas (U3) de tipo natural.

	Intensidad de fluorescencia media	
Experimento	ASLV U3	ASLV dU3
1 (gag-pol de tipo natural)	503,8 +/- 15,1	1748,2 +/- 112,6
2 (gag-pol optimizado en cuanto a codones)	791,3 +/- 48,7	22275,6 +/- 120,6

La Figura 12 muestra los valores de la expresión del transgén ilustrativo EGFP de un casete de expresión en células diana que estaban transducidas con vector viral que contenían ARN viral transcrito de un plásmido según la SEC ID Nº: 5 (SIN, de acuerdo con la invención), la SEC. ID Nº: 9 (TATAmut) o SEC ID Nº: 7 (no TATA). Como se ha descrito anteriormente, se generaron partículas virales en células de empaquetamiento que traducían proteína de fusión gag-pol optimizada en cuanto a codones y proteína de envoltura de VSVg. Las células HT1080 se transdujeron con partículas virales, se cultivaron durante 6 días y después se analizaron mediante FACS. Los valores mostrados hacen evidente que la región U3 3' de acuerdo con la invención en todos los ARN virales o partículas virales de acuerdo con la invención da lugar a una elevada eficacia de la expresión del transgén en células diana. Los resultados están representados también en la siguiente tabla:

	Intensidad de fluorescencia media
ASLV dU3	3616,9 +/- 15,8
ASLV dU3 TATAmut	3020,4 +/- 227,8
ASLV dU3 no TATA	3463,2 +/- 252,8

Una realización preferente del vector, por ejemplo, de su ARN viral, contenido opcionalmente en partículas virales, no presenta secuencias virales codificantes, por ejemplo, ninguna secuencia que comprenda completa o parcialmente uno de los genes estructurales gag, env, src o pol. Por tanto, se prefiere que el ARN viral del vector o la secuencia transcrita de ácido nucleico de un plásmido que codifica el ARN viral del vector, la región R 5', la región U5 5' y la zona que comprende la señal de empaquetamiento y que se extiende hasta limitando con el casete de expresión del transgén esté compuesto de estas zonas y por ejemplo tampoco contenga ningún sitio donador de empalme. Una secuencia preferente para la región R 5', la región U5 5' y la zona que comprende la señal de empaquetamiento y que se extiende hasta limitando con el casete de expresión del transgén comprende los nucleótidos Nº 922 a 1292 de la SEC ID Nº: 5 o está compuesta de esta secuencia de nucleótidos. Esta secuencia de nucleótidos puede estar acortada opcionalmente en 3', por ejemplo, en 140 nucleótidos, preferentemente en 10 a 56 nucleótidos.

En esta forma en la que la región del ARN viral en su extremo 5' y, por tanto, en 5' directamente limitando con el casete de expresión del transgén está compuesta de secuencias no traducidas, en particular de una región R, una región U5 y la señal de empaquetamiento, por tanto, no están contenidas secciones de gag, env, src o pol.

Esta forma de realización está mostrada esquemáticamente en la Figura 13 que aclara que esta sección 5' del vector, que se denomina también líder, es claramente más corta que la región 5' de vectores SIN lentivirales o γ-retrovirales. Mientras que en vectores SIN lentivirales y muchas versiones γ-retrovirales la señal de empaquetamiento ψ se solapa con genes estructurales virales, sorprendentemente, en el vector alfarretroviral puede estar contenida la señal de empaquetamiento sin secciones de secuencias virales codificantes. Por tanto, el vector de acuerdo con la invención (vector SIN alfarretroviral) se denomina líder sin gag. Además, en el vector se puede haber eliminado el sitio donador de empalme viral o haberse convertido en no funcional mediante mutación. Esto es posible a causa del solapamiento con funciones de empaquetamiento, por ejemplo, en vectores lenti- o γ-retrovirales solo con reducción del título. Además, la retirada completa de secciones de secuencias virales codificantes permite el acortamiento de la sección de la secuencia de vector que está dispuesta en 5' delante del casete de expresión del transgén, por ejemplo, de 371 nucleótidos (nucleótidos, pb). Por tanto, en comparación con vectores lentivirales y γ-retrovirales se crea particularmente mucho espacio para recibir secuencias de transgenes. El casete de expresión está compuesto en estos ejemplos respectivamente del promotor interno fuerte de SFFV, la secuencia codificante para la proteína fluorescente verde (EGFP) como gen indicador y el WPRE.

Ejemplo 2: Expresión a largo plazo del transgén en células diana eucariotas

Como ejemplo del uso del vector viral como medicamento, por ejemplo, para la terapia génica, se transdujeron células madre sanguíneas con partículas virales de acuerdo con la invención. Se transdujeron células de médula ósea aisladas de ratones (cepa C57BL6) con partículas virales de acuerdo con la invención y para la comparación con partículas lentivirales. Como transgén estaba contenido respectivamente EGFP en el mismo casete de expresión bajo el control del promotor de SFFV. En condiciones de igual eficacia se transfectaron las células por el vector lentiviral hasta el 39 % y, por tanto, un factor de 1,6 más que por el vector descrito (24 %). Esto se puede deber a que las partículas lentivirales pueden transducir también células en reposo inactivas en cuando a división, mientras que las partículas α-retrovirales de acuerdo con la invención pueden transducir células preferentemente durante la división.

Las células madre sanguíneas transducidas se trasplantaron en respectivamente 8 ratones irradiados de forma letal de la misma cepa. Se identificaron células transducidas mediante FACS de sangre periférica. Después de 31 semanas, la parte de leucocitos positivos a EGFP transducidos para vectores lentivirales había aumentado con el 32,9 +/- 8,9 % aproximadamente al doble del valor para los vectores de acuerdo con la invención (16,1 +/- 12 %).

De ambos grupos se extrajeron de los animales con los mayores títulos de células positivas a EGFP médula ósea y se trasplantó respectivamente en 5 animales irradiados letalmente de un segundo grupo. 6 semanas después de este trasplante, la parte de leucocitos positivos a EGFP para las células madre transducidas con lentivirus con 40,6 +/- 5,9 se encontraba todavía aproximadamente el 30 % por encima de la parte en células madre transducidas de acuerdo con la invención (32,6 +/- 8,1 %). Esto muestra que la velocidad de silenciamiento epigenético de la expresión del transgén para los vectores no es sustancialmente mayor que en vectores lentivirales.

Ejemplo 3: Comprobación del menor riesgo por mutagénesis de inserción de los vectores

En un ensayo *in vitro* desarrollado por los inventores (Mol. Ther. 1919-1928 (2009)) se comprobó el riesgo de la activación de oncogén en la transducción o transfección con el vector viral. Una activación de oncogén, por ejemplo, mediante inserción de secuencias virales en la proximidad genómica, a este respecto conduce a la transformación de células hematopoyéticas primarias. Dicho brevemente, en este ensayo se examina con qué frecuencia y con qué grado de manifestación se produce la activación por inserción de proto-oncogenes celulares que favorecen la transformación tales como *Evil*. En este ensayo se ponen en contacto células hematopoyéticas primarias del ratón no tratado, por ejemplo, cepa C57BL6, con partículas virales. Después de esta transducción se determina la frecuencia de la transformación mediante cultivo en condiciones que causan la diferenciación mieloica, realizándose la nueva siembra de tales diluciones celulares en las que las células no transformadas a causa de la supresión desencadenada por la dilución de la capacidad de división remanente no proliferan, sino que proliferan solo las células transformadas. En células transformadas está regulada al alza por ejemplo la expresión de proto-oncogenes por mutagénesis de inserción. Este ensayo cuantifica por tanto la frecuencia de transformación en relación con células transducidas. La vitalidad o estabilidad de los transformantes se puede determinar como el resultado del cultivo de células transformadas después de la nueva siembra o resiembra en placas.

A este respecto se ha mostrado que los vectores γ-retrovirales y los vectores lentivirales conducen a una frecuencia de transformación de aproximadamente 1×10^{-5} o 5×10^{-6} (Mol. Ther. 1919-1928 (2009)). En enfoques paralelos se comprobó para los vectores descritos una frecuencia de transformación de aproximadamente $3,6 \times 10^{-6}$ que, por tanto, se encontraba todavía por debajo del valor comprobado para vectores lentivirales. Además de la frecuencia de transformación, en este ensayo se analiza la vitalidad o viabilidad de las células transformadas. El valor para esta viabilidad se encontraba para 3 de 4 células aisladas hasta ahora que se habían transformado por el vector aproximadamente 10 veces menor que el valor promedio para células transformadas por lentivirus. Estos resultados muestran mediante la reducida frecuencia de transformación y la reducida vitalidad de células transformadas por el vector α-retroviral de acuerdo con la invención que el mismo ofrece una mejor seguridad frente a la activación de oncogén que los vectores anteriores γ-retrovirales o lentivirales. Se puede reducir adicionalmente el riesgo de la transformación de inserción mediante la selección de secuencias promotoras internas apropiadas.

Ejemplo 4: Determinación de deleciones de la región U3 3', que eliminan su actividad promotora y potenciadora

La deleción o la sustitución de nucleótidos de la región U3 para la reducción o eliminación de la actividad promotora y potenciadora se determina preferentemente con un ensayo de actividad en el que se mide la expresión de un gen indicador que está dispuesto funcionalmente en 3' con respecto a la región U3 provista de deleciones. Cuando la expresión de un gen indicador que está dispuesto en 3' funcionalmente con respecto a la región U3 delecionada o sustituida se determina en el nivel de la actividad de fondo del gen indicador expresado sin promotor, se elimina de acuerdo con la invención la actividad promotora. Preferentemente, en 3' del gen indicador está dispuesta una secuencia de poliadenilación. Como actividad de fondo del gen indicador expresado se define por ejemplo la actividad que se mide sin cualquier región U3 o sin cualquier promotor en 5' del gen indicador.

La Figura 14 muestra esquemáticamente las construcciones de ácido nucleico por recortes con las que se han llevado a cabo el ensayo de actividad para la actividad promotora y potenciadora delecionada o sustituida de

acuerdo con la invención de la región U3. La construcción representada arriba en la Figura 13 muestra como gen indicador el gen de luciferasa de *Renilla* (Rluc) que en 3' está flanqueado por una secuencia de señal de poliadenilación de SV40 (pA de SV40) y en 5' por la secuencia de ácido nucleico que se tiene que ensayar en cuanto a secuencia promotora y potenciadora, aquí la LTR de tipo natural (U3-R-U5) de ASLV. Bajo esto está mostrada una construcción por lo demás idéntica con deleción (ΔLTR), cuya región U3 está parcialmente delecionada (ΔU3). En esta construcción están delecionados los nucleótidos 28-186. Aún por debajo está mostrada una construcción por lo demás idéntica en la que la LTR está sustituida por completo por una secuencia aleatoria (spacer, espaciador derivado del gen de β-lactamasa procariota). En 5' delante de la sección de secuencia viral está dispuesta una señal de poliadenilación sintética (pA sin).

En la Figura 15 están mostradas las actividades luciferasa de células 293T que están transfectadas con las construcciones de la Figura 13 y cultivadas en condiciones idénticas. La medición de la actividad luciferasa se realizó en homogenados celulares, la representación muestra valores de luminiscencia relativos que se han normalizado a proteína total y eficacia de la transfección. Es evidente que la deleción de acuerdo con la invención hasta ΔU3 (ASLV ΔU3) causa una espectacular reducción (de un factor de aproximadamente 400) de la actividad promotora frente a la LTR de tipo natural (ASLV U3). Los datos después de análisis estadístico unos con respecto a otros son muy significativos (**, P<0,01). La sustitución de la LTR por la secuencia aleatoria (spacer) conduce a una reducción similar de la actividad promotora. Por tanto, de acuerdo con la invención se puede usar una secuencia aleatoria que no presenta una función conocida promotora y potenciadora como secuencia con actividad de fondo. Se supone que la actividad de fondo todavía medida de la luciferasa es causada por otros constituyentes del plásmido en el que está contenido el casete de expresión con el gen indicador. Preferentemente, el gen indicador está adaptado al uso de codones de la célula transfectada.

LISTADO DE SECUENCIAS

<110> Universidad de Medicina de Hannover
Axel, Schambach, Dr.
Christopher, Baum, Prof. Dr.
Julia, Sürth

<120> Sistema de vector de ASLV

<130> M1013PCT

<160> 16

<170> PatentIn versión 3.5

<210> 1
<211> 229
<212> ADN
<213> Virus de sarcoma aviar

<220>
<221> misc_feature
<222> (1)..(229)
<223> Sec ID Nº 1: Región U3 3' de tipo natural de ASLV

<400> 1

```
aatgtagtct tatgcaatac tcttgtagtc ttgcaacatg gtaacgatga gttagcaaca      60
tgccttacaa ggagagaaaa agcaccgtgc atgccgattg gtggaagtaa ggtggtacga     120
tcgtgcctta ttaggaaggc aacagacggg tctgacatgg attggacgaa ccactgaatt     180
ccgcattgca gagatattgt atttaagtgc ctagctcgat acaataaac                229
```

<210> 2
<211> 70
<212> ADN
<213> Secuencia artificial

<220>
<223> /note = "Sec ID Nº 2: Región U3 con deleción"

<400> 2

aatgtagtct tatgcaatac tcttgtacgt agagatattg tatttaagtg cctagctcga 60

tacaataaac 70

<210> 3
<211> 70
<212> ADN
<213> Secuencia artificial

<220>
<223> /note = "Sec ID Nº 3: Región U3 con deleción"

<400> 3

aatgtagtct tatgcaatac tcttgtacgt agagatattg tgtctaagtg cctagctcga 60

tacaataaac 70

<210> 4
<211> 54
<212> ADN
<213> Secuencia artificial

<220>
<223> /note = "Sec ID Nº 4: Región U3 con deleción"

<400> 4
aatgtagtct tatgcaatac tcttgtacgt agtgcctagc tcgatacaat aaac 54

<210> 5
<211> 5704
<212> ADN
<213> Secuencia artificial

<220>
<223> /note = "Sec ID Nº 5: vector viral con región U3 delecionada"

<220>
<221> Potenciador
<222> (452)..(684)
<223> Potenciador de SV40

<220>
<221> Promotor
<222> (693)..(921)
<223> Promotor de RSV

<220>
<221> misc_feature
<222> (922)..(942)
<223> LTR 5': región R

<220>
<221> misc_feature
<222> (943)..(1022)
<223> LTR 5': región U5

<220>
<221> PBS
<222> (1023)..(1041)
<223> Sitio de unión de cebador

<220>
<221> promotor
<222> (1293)..(1684)
<223> Promotor interno del casete de expresión: SFFV

<220>
<221> CDS
<222> (1764)..(2483)
<223> EGFP: secuencia codificante para eGFP

<220>
<221> misc_feature
<222> (2522)..(3124)
<223> PRE del casete de expresión

<220>
<221> misc_feature
<222> (3282)..(3351)
<223> U3 3': Región U3 con deleción

<220>
<221> misc_feature
<222> (3352)..(3372)
<223> Región R 3'

<220>
<221> misc_feature
<222> (3373)..(3452)
<223> Región U5 3'

<400> 5

```
ctgcctcgcg cgtttcggtg atgacggtga aaacctctga cacatgcagc tcccggagac       60
ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg gcgcgtcagc      120
gggtgttggc gggtgtcggg gcgcagccat gacccagtca cgtagcgata gttactatgc      180
ggcatcagag cagattgtac tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg      240
cgtaaggaga aaataccgca tcaggcgcca ttcgccattc aggctgcgca actgttggga      300
agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg gatgtgctgc      360
aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc      420
cagtgaatta gtactctagc ttaagacgcg tggcctgaaa taacctctga aagaggaact      480
tggttaggta ccttctgagg cggaaagaac cagctgtgga atgtgtgtca gttagggtgt      540
ggaaagtccc caggctcccc agcaggcaga agtatgcaaa gcatgcatct caattagtca      600
gcaaccaggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat      660
ctcaattagt cagcaaccat agtcccttaa gaatgtagtc ttatgcaata ctcttgtagt      720
cttgcaacat ggtaacgatg agttagcaac atgccttaca aggagagaaa aagcaccgtg      780
catgccgatt ggtggaagta aggtggtacg atcgtgcctt attaggaagg caacagacgg      840
gtctgacatg gattggacga accactgaat tgccgcattg cagagatatt gtatttaagt      900
gcctagctcg atacaataaa cgccatttga ccattcacca cattggtgtg cacctgggtt      960
gatggccgga ccgttgattc cctgacgact acgagcacct gcatgaagca gaaggcttca     1020
tttggtgacc ccgacgtgat agttagggaa tagtggtcgg ccacagacgg cgtggcgatc     1080
ctgtctccat ccgtctcgtc tatcgggagg cgagttcgat gaccctggtg gaggggctg      1140
cggcttaggg aggcagaagc tgagtaccgt cggagggagc tccagggccc ggagcgactg     1200
acccctgccg agaactcaga gggtcgtcgg aagacggaga gtgagcccga cgaccacccc     1260
aggcacgtct ttggtcggcc tgcggatcaa gcagcggccg ctagctgcag taacgccatt     1320
ttgcaaggca tggaaaaata ccaaaccaag aatagagaag ttcagatcaa gggcgggtac     1380
atgaaaatag ctaacgttgg gccaaacagg atatctgcgg tgagcagttt cggccccggc     1440
ccggggccaa gaacagatgg tcaccgcagt ttcggccccg gcccgaggcc aagaacagat     1500
ggtccccaga tatggcccaa ccctcagcag tttcttaaga cccatcagat gtttccaggc     1560
tcccccaagg acctgaaatg accctgcgcc ttatttgaat taaccaatca gcctgcttct     1620
cgcttctgtt cgcgcgcttc tgcttcccga gctctataaa agagctcaca acccctcact     1680
cggcgcgcca gtcctccgac agactgagtc ggccggtcga atcaagctta tcgataccgt     1740
cgacggatcc accggtcgcc acc atg gtg agc aag ggc gag gag ctg ttc acc     1793
                         Met Val Ser Lys Gly Glu Glu Leu Phe Thr
                         1               5                  10

ggg gtg gtg ccc atc ctg gtc gag ctg gac ggc gac gta aac ggc cac      1841
Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His
                15                  20                  25
```

```
aag ttc agc gtg tcc ggc gag ggc gag ggc gat gcc acc tac ggc aag        1889
Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys
            30                  35                  40

ctg acc ctg aag ttc atc tgc acc acc ggc aag ctg ccc gtg ccc tgg        1937
Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp
        45                  50                  55

ccc acc ctc gtg acc acc ctg acc tac ggc gtg cag tgc ttc agc cgc        1985
Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg
    60                  65                  70

tac ccc gac cac atg aag cag cac gac ttc ttc aag tcc gcc atg ccc        2033
Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro
75                  80                  85                  90

gaa ggc tac gtc cag gag cgc acc atc ttc ttc aag gac gac ggc aac        2081
Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn
                95                  100                 105

tac aag acc cgc gcc gag gtg aag ttc gag ggc gac acc ctg gtg aac        2129
Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn
            110                 115                 120

cgc atc gag ctg aag ggc atc gac ttc aag gag gac ggc aac atc ctg        2177
Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu
        125                 130                 135

ggg cac aag ctg gag tac aac tac aac agc cac aac gtc tat atc atg        2225
Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met
    140                 145                 150

gcc gac aag cag aag aac ggc atc aag gtg aac ttc aag atc cgc cac        2273
Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His
155                 160                 165                 170

aac atc gag gac ggc agc gtg cag ctc gcc gac cac tac cag cag aac        2321
Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn
                175                 180                 185

acc ccc atc ggc gac ggc ccc gtg ctg ctg ccc gac aac cac tac ctg        2369
Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu
            190                 195                 200

agc acc cag tcc gcc ctg agc aaa gac ccc aac gag aag cgc gat cac        2417
Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His
        205                 210                 215

atg gtc ctg ctg gag ttc gtg acc gcc gcc ggg atc act ctc ggc atg        2465
Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met
    220                 225                 230

gac gag ctg tac aag taa agcggccgcg tcgacggatc ccccgggctg               2513
Asp Glu Leu Tyr Lys
235

caggaattcg agcatcttac cgccatttat acccatattt gttctgtttt tcttgatttg      2573

ggtatacatt taaatgttaa taaaacaaaa tggtggggca atcatttaca tttttaggga      2633

tatgtaatta ctagttcagg tgtattgcca caagacaaac atgttaagaa actttcccgt      2693

tatttacgct ctgttcctgt taatcaacct ctggattaca aaatttgtga aagattgact      2753

gatattctta actatgttgc tccttttacg ctgtgtggat atgctgcttt aatgcctctg      2813

tatcatgcta ttgcttcccg tacggctttc gttttctcct ccttgtataa atcctggttg      2873

ctgtctcttt atgaggagtt gtggcccgtt gtccgtcaac gtggcgtggt gtgctctgtg      2933
```

tttgctgacg caacccccac tggctggggc attgccacca cctgtcaact cctttctggg 2993
actttcgctt tccccctccc gatcgccacg gcagaactca tcgccgcctg ccttgcccgc 3053
tgctggacag gggctaggtt gctgggcact gataattccg tggtgttgtc ggggaagctg 3113
acgtcctttc gaattcgaaa gcttttaaat atcgatgcga tgtacgggcc agatatacgc 3173
gtatctgagg ggactagggt gtgtttaggc gaaaagcggg gcttcggttg tacgcggtta 3233
ggagtcccct taggatatag tagtttcgct tttgcatagg gaggggggaaa tgtagtctta 3293
tgcaatactc ttgtacgtag agatattgta tttaagtgcc tagctcgata caataaacgc 3353
catttgacca ttcaccacat tggtgtgcac ctgggttgat ggccggaccg ttgattccct 3413
gacgactacg agcacctgca tgaagcagaa ggcttcattc tcgagagctt tggcgtaatc 3473
atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg 3533
agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac tcacattaat 3593
tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc tgcattaatg 3653
aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgct 3713
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 3773
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 3833
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 3893
ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 3953
actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc ctgttccgac 4013
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 4073
atgctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt 4133
gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 4193
caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 4253
agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 4313
tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 4373
tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa 4433
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 4493
gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa 4553
aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat 4613
atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 4673
gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat 4733
acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc 4793
ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtggtcc 4853
tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag 4913
ttcgccagtt aatagtttgc gcaacgttgt tgccattgct gctggcatcg tggtgtcacg 4973

```
ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg     5033
atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg ttgtcagaag     5093
taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt     5153
catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt cattctgaga     5213
atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata ataccgcgcc     5273
acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc     5333
aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac ccaactgatc     5393
ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc     5453
cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct tccttttca      5513
atattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat     5573
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt     5633
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt     5693
tcgtcttcaa g                                                         5704
```

<210> 6
<211> 239
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción sintética

<400> 6

```
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                  10                  15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
            100                 105                 110

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

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130                 135                 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                 150                 155                 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
        195                 200                 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 7
<211> 5688
<212> ADN
<213> Secuencia artificial

<220>
<223> /note = "Sec ID Nº 7: vector viral con región U3 delecionada"

<220>
<221> Potenciador
<222> (452)..(684)
<223> Potenciador de SV40

<220>
<221> promotor
<222> (693)..(921)
<223> Promotor de RSV

<220>
<221> misc_feature
<222> (922)..(942)
<223> Región R 5'

<220>
<221> misc_feature
<222> (943)..(1022)
<223> Región U5 5'

<220>
<221> PBS
<222> (1023)..(1041)
<223> Sitio de unión de cebador

<220>
<221> promotor
<222> (1293)..(1684)
<223> Promotor interno del casete de expresión: SFFV

<220>
<221> CDS
<222> (1764)..(2483)
<223> EGFP: secuencia codificante para eGFP

<220>
<221> misc_feature
<222> (2522)..(3124)
<223> PRE del casete de expresión

<220>
<221> misc_feature
<222> (3282)..(3335)
<223> Región U3 3' con deleción

<220>
<221> misc_feature
<222> (3336)..(3356)
<223> Región R 3'

<220>
<221> misc_feature
<222> (3357)..(3436)
<223> Región U5 3'

<400> 7

ctgcctcgcg cgtttcggtg atgacggtga aaacctctga cacatgcagc tcccggagac 60
ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg gcgcgtcagc 120
gggtgttggc gggtgtcggg gcgcagccat gacccagtca cgtagcgata gttactatgc 180
ggcatcagag cagattgtac tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg 240
cgtaaggaga aaataccgca tcaggcgcca ttcgccattc aggctgcgca actgttggga 300
agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg gatgtgctgc 360
aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc 420
cagtgaatta gtactctagc ttaagacgcg tggcctgaaa taacctctga aagaggaact 480
tggttaggta ccttctgagg cggaaagaac cagctgtgga atgtgtgtca gttagggtgt 540
ggaaagtccc caggctcccc agcaggcaga agtatgcaaa gcatgcatct caattagtca 600
gcaaccaggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat 660
ctcaattagt cagcaaccat agtcccttaa gaatgtagtc ttatgcaata ctcttgtagt 720
cttgcaacat ggtaacgatg agttagcaac atgccttaca aggagagaaa aagcaccgtg 780
catgccgatt ggtggaagta aggtggtacg atcgtgcctt attaggaagg caacagacgg 840
gtctgacatg gattggacga accactgaat tgccgcattg cagagatatt gtatttaagt 900
gcctagctcg atacaataaa cgccatttga ccattcacca cattggtgtg cacctgggtt 960
gatggccgga ccgttgattc cctgacgact acgagcacct gcatgaagca gaaggcttca 1020
tttggtgacc ccgacgtgat agttagggaa tagtggtcgg ccacagacgg cgtggcgatc 1080
ctgtctccat ccgtctcgtc tatcgggagg cgagttcgat gaccctggtg gaggggctg 1140
cggcttaggg aggcagaagc tgagtaccgt cggagggagc tccagggccc ggagcgactg 1200
acccctgccg agaactcaga gggtcgtcgg aagacggaga gtgagcccga cgaccacccc 1260
aggcacgtct ttggtcggcc tgcggatcaa gcagcggccg ctagctgcag taacgccatt 1320
ttgcaaggca tggaaaaata ccaaaccaag aatagagaag ttcagatcaa gggcgggtac 1380

```
atgaaaatag ctaacgttgg gccaaacagg atatctgcgg tgagcagttt cggccccggc     1440

ccggggccaa gaacagatgg tcaccgcagt ttcggccccg gcccgaggcc aagaacagat     1500

ggtccccaga tatggcccaa ccctcagcag tttcttaaga cccatcagat gtttccaggc     1560

tcccccaagg acctgaaatg accctgcgcc ttatttgaat taaccaatca gcctgcttct     1620

cgcttctgtt cgcgcgcttc tgcttcccga gctctataaa agagctcaca acccctcact     1680

cggcgcgcca gtcctccgac agactgagtc ggccggtcga atcaagctta tcgataccgt     1740

cgacggatcc accggtcgcc acc atg gtg agc aag ggc gag gag ctg ttc acc     1793
                          Met Val Ser Lys Gly Glu Glu Leu Phe Thr
                          1               5                  10

ggg gtg gtg ccc atc ctg gtc gag ctg gac ggc gac gta aac ggc cac       1841
Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His
                15                  20                  25

aag ttc agc gtg tcc ggc gag ggc gag ggc gat gcc acc tac ggc aag       1889
Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys
            30                  35                  40

ctg acc ctg aag ttc atc tgc acc acc ggc aag ctg ccc gtg ccc tgg       1937
Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp
        45                  50                  55

ccc acc ctc gtg acc acc ctg acc tac ggc gtg cag tgc ttc agc cgc       1985
Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg
    60                  65                  70

tac ccc gac cac atg aag cag cac gac ttc ttc aag tcc gcc atg ccc       2033
Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro
75                  80                  85                  90

gaa ggc tac gtc cag gag cgc acc atc ttc ttc aag gac gac ggc aac       2081
Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn
                95                  100                 105

tac aag acc cgc gcc gag gtg aag ttc gag ggc gac acc ctg gtg aac       2129
Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn
            110                 115                 120

cgc atc gag ctg aag ggc atc gac ttc aag gag gac ggc aac atc ctg       2177
Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu
        125                 130                 135

ggg cac aag ctg gag tac aac tac aac agc cac aac gtc tat atc atg       2225
Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met
    140                 145                 150

gcc gac aag cag aag aac ggc atc aag gtg aac ttc aag atc cgc cac       2273
Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His
155                 160                 165                 170

aac atc gag gac ggc agc gtg cag ctc gcc gac cac tac cag cag aac       2321
Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn
                175                 180                 185

acc ccc atc ggc gac ggc ccc gtg ctg ctg ccc gac aac cac tac ctg       2369
Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu
            190                 195                 200

agc acc cag tcc gcc ctg agc aaa gac ccc aac gag aag cgc gat cac       2417
Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His
        205                 210                 215
```

atg gtc ctg ctg gag ttc gtg acc gcc gcc ggg atc act ctc ggc atg 2465
Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met
220 225 230

gac gag ctg tac aag taa agcggccgcg tcgacggatc ccccgggctg 2513
Asp Glu Leu Tyr Lys
235

caggaattcg agcatcttac cgccatttat acccatattt gttctgtttt tcttgatttg 2573

ggtatacatt taaatgttaa taaaacaaaa tggtggggca atcatttaca tttttaggga 2633

tatgtaatta ctagttcagg tgtattgcca caagacaaac atgttaagaa actttcccgt 2693

tatttacgct ctgttcctgt taatcaacct ctggattaca aaatttgtga aagattgact 2753

gatattctta actatgttgc tccttttacg ctgtgtggat atgctgcttt aatgcctctg 2813

tatcatgcta ttgcttcccg tacggctttc gttttctcct ccttgtataa atcctggttg 2873

ctgtctcttt atgaggagtt gtggcccgtt gtccgtcaac gtggcgtggt gtgctctgtg 2933

tttgctgacg caacccccac tggctggggc attgccacca cctgtcaact cctttctggg 2993

actttcgctt tccccctccc gatcgccacg gcagaactca tcgccgcctg ccttgcccgc 3053

tgctggacag gggctaggtt gctgggcact gataattccg tggtgttgtc ggggaagctg 3113

acgtcctttc gaattcgaaa gcttttaaat atcgatgcga tgtacgggcc agatatacgc 3173

gtatctgagg ggactagggt gtgtttaggc gaaaagcggg gcttcggttg tacgcggtta 3233

ggagtcccct taggatatag tagtttcgct tttgcatagg gaggggggaaa tgtagtctta 3293

tgcaatactc ttgtacgtag tgcctagctc gatacaataa acgccatttg accattcacc 3353

acattggtgt gcacctgggt tgatggccgg accgttgatt ccctgacgac tacgagcacc 3413

tgcatgaagc agaaggcttc attctcgaga gctttggcgt aatcatggtc atagctgttt 3473

cctgtgtgaa attgttatcc gctcacaatt ccacacaaca tacgagccgg aagcataaag 3533

tgtaaagcct ggggtgccta atgagtgagc taactcacat taattgcgtt gcgctcactg 3593

cccgctttcc agtcgggaaa cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg 3653

gggagaggcg gtttgcgtat tgggcgctct tccgcttcct cgctcactga ctcgctgcgc 3713

tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat acggttatcc 3773

acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca aaaggccagg 3833

aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgcccccc tgacgagcat 3893

cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata aagataccag 3953

gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga 4013

tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcaatgctc acgctgtagg 4073

tatctcagtt cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga acccccgtt 4133

cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc ggtaagacac 4193

gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc 4253

ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag gacagtattt 4313

ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc 4373

ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt gcaagcagca gattacgcgc 4433
agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg 4493
aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat cttcacctag 4553
atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaacttgg 4613
tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttcgt 4673
tcatccatag ttgcctgact ccccgtcgtg tagataacta cgatacggga gggcttacca 4733
tctggcccca gtgctgcaat gataccgcga gacccacgct caccggctcc agatttatca 4793
gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcctgcaac tttatccgcc 4853
tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt 4913
ttgcgcaacg ttgttgccat tgctgctggc atcgtggtgt cacgctcgtc gtttggtatg 4973
gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc 5033
aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt ggccgcagtg 5093
ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga 5153
tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg tatgcggcga 5213
ccgagttgct cttgcccggc gtcaatacgg gataataccg cgccacatag cagaacttta 5273
aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg 5333
ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc atcttttact 5393
ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata 5453
agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta ttgaagcatt 5513
tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacaa 5573
atagggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt 5633
atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtct tcaag 5688

<210> 8
<211> 239
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción sintética

<400> 8

```
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                  10                  15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr

    50                  55                  60

Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
            100                 105                 110

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

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130                 135                 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                 150                 155                 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
        195                 200                 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 9
<211> 5704
<212> ADN
<213> Secuencia artificial

<220>
<223> /note = "Sec ID Nº 9: vector viral con deleción y caja TATA mutada en región U3 3' "

<220>
<221> Potenciador
<222> (452)..(684)
<223> Potenciador de SV40

<220>
<221> promotor
<222> (693)..(921)
<223> Promotor de RSV

<220>
<221> misc_feature
<222> (922)..(942)
<223> Región R 5'

<220>
<221> misc_feature
<222> (943)..(1022)
<223> Región U5 5'

<220>
<221> PBS
<222> (1023)..(1041)

<220>
<221> promotor
<222> (1293)..(1684)
<223> Promotor de SFFV

<220>
<221> CDS
<222> (1764)..(2483)
<223> EGFP: secuencia codificante para eGFP

<220>
<221> misc_signal
<222> (2522)..(3124)
<223> PRE

<220>
<221> misc_feature
<222> (3282)..(3351)
<223> U3 3' delecionada

<220>
<221> misc_feature
<222> (3322)..(3328)
<223> caja TATA mutada en U3 3' delecionada

<220>
<221> misc_feature
<222> (3352)..(3372)
<223> Región R 3'

<220>
<221> misc_feature
<222> (3373)..(3452)
<223> Región U5 3'

<400> 9

ctgcctcgcg cgtttcggtg atgacggtga aaacctctga cacatgcagc tcccggagac 60
ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg gcgcgtcagc 120
gggtgttggc gggtgtcggg gcgcagccat gacccagtca cgtagcgata gttactatgc 180
ggcatcagag cagattgtac tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg 240
cgtaaggaga aaataccgca tcaggcgcca ttcgccattc aggctgcgca actgttggga 300
agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg gatgtgctgc 360
aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc 420
cagtgaatta gtactctagc ttaagacgcg tggcctgaaa taacctctga aagaggaact 480
tggttaggta ccttctgagg cggaaagaac cagctgtgga atgtgtgtca gttagggtgt 540
ggaaagtccc caggctcccc agcaggcaga agtatgcaaa gcatgcatct caattagtca 600
gcaaccaggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat 660

```
ctcaattagt cagcaaccat agtcccttaa gaatgtagtc ttatgcaata ctcttgtagt      720

cttgcaacat ggtaacgatg agttagcaac atgccttaca aggagagaaa aagcaccgtg      780

catgccgatt ggtggaagta aggtggtacg atcgtgcctt attaggaagg caacagacgg      840

gtctgacatg gattggacga accactgaat tgccgcattg cagagatatt gtatttaagt      900

gcctagctcg atacaataaa cgccatttga ccattcacca cattggtgtg cacctgggtt      960

gatggccgga ccgttgattc cctgacgact acgagcacct gcatgaagca gaaggcttca     1020

tttggtgacc ccgacgtgat agttagggaa tagtggtcgg ccacagacgg cgtggcgatc     1080

ctgtctccat ccgtctcgtc tatcgggagg cgagttcgat gaccctggtg gaggggctg      1140

cggcttaggg aggcagaagc tgagtaccgt cggagggagc tccagggccc ggagcgactg     1200

acccctgccg agaactcaga gggtcgtcgg aagacggaga gtgagcccga cgaccacccc     1260

aggcacgtct ttggtcggcc tgcggatcaa gcagcggccg ctagctgcag taacgccatt     1320

ttgcaaggca tggaaaaata ccaaaccaag aatagagaag ttcagatcaa gggcgggtac     1380

atgaaaatag ctaacgttgg gccaaacagg atatctgcgg tgagcagttt cggccccggc     1440

ccggggccaa gaacagatgg tcaccgcagt ttcggccccg gcccgaggcc aagaacagat     1500

ggtccccaga tatggcccaa ccctcagcag tttcttaaga cccatcagat gtttccaggc     1560

tcccccaagg acctgaaatg accctgcgcc ttatttgaat taaccaatca gcctgcttct     1620

cgcttctgtt cgcgcgcttc tgcttcccga gctctataaa agagctcaca acccctcact     1680

cggcgcgcca gtcctccgac agactgagtc ggccggtcga atcaagctta tcgataccgt     1740

cgacggatcc accggtcgcc acc atg gtg agc aag ggc gag gag ctg ttc acc     1793
                          Met Val Ser Lys Gly Glu Glu Leu Phe Thr
                          1               5                  10

ggg gtg gtg ccc atc ctg gtc gag ctg gac ggc gac gta aac ggc cac      1841
Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His
            15                  20                  25

aag ttc agc gtg tcc ggc gag ggc gag ggc gat gcc acc tac ggc aag      1889
Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys
        30                  35                  40

ctg acc ctg aag ttc atc tgc acc acc ggc aag ctg ccc gtg ccc tgg      1937
Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp
    45                  50                  55

ccc acc ctc gtg acc acc ctg acc tac ggc gtg cag tgc ttc agc cgc      1985
Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg
60                  65                  70

tac ccc gac cac atg aag cag cac gac ttc ttc aag tcc gcc atg ccc      2033
Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro
75                  80                  85                  90

gaa ggc tac gtc cag gag cgc acc atc ttc ttc aag gac gac ggc aac      2081
Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn
            95                  100                 105

tac aag acc cgc gcc gag gtg aag ttc gag ggc gac acc ctg gtg aac      2129
Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn
        110                 115                 120
```

```
cgc atc gag ctg aag ggc atc gac ttc aag gag gac ggc aac atc ctg       2177
Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu
        125                 130                 135

ggg cac aag ctg gag tac aac tac aac agc cac aac gtc tat atc atg       2225
Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met
    140                 145                 150

gcc gac aag cag aag aac ggc atc aag gtg aac ttc aag atc cgc cac       2273
Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His
155                 160                 165                 170

aac atc gag gac ggc agc gtg cag ctc gcc gac cac tac cag cag aac       2321
Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn
                175                 180                 185

acc ccc atc ggc gac ggc ccc gtg ctg ctg ccc gac aac cac tac ctg       2369
Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu
            190                 195                 200

agc acc cag tcc gcc ctg agc aaa gac ccc aac gag aag cgc gat cac       2417
Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His
        205                 210                 215

atg gtc ctg ctg gag ttc gtg acc gcc gcc ggg atc act ctc ggc atg       2465
Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met
    220                 225                 230

gac gag ctg tac aag taa agcggccgcg tcgacggatc ccccgggctg              2513
Asp Glu Leu Tyr Lys
235
```

caggaattcg agcatcttac cgccatttat acccatattt gttctgtttt tcttgatttg 2573

ggtatacatt taaatgttaa taaaacaaaa tggtggggca atcatttaca tttttaggga 2633

tatgtaatta ctagttcagg tgtattgcca caagacaaac atgttaagaa actttcccgt 2693

tatttacgct ctgttcctgt taatcaacct ctggattaca aaatttgtga aagattgact 2753

gatattctta actatgttgc tccttttacg ctgtgtggat atgctgcttt aatgcctctg 2813

tatcatgcta ttgcttcccg tacggctttc gttttctcct ccttgtataa atcctggttg 2873

ctgtctcttt atgaggagtt gtggcccgtt gtccgtcaac gtggcgtggt gtgctctgtg 2933

tttgctgacg caacccccac tggctggggc attgccacca cctgtcaact cctttctggg 2993

actttcgctt tccccctccc gatcgccacg gcagaactca tcgccgcctg ccttgcccgc 3053

tgctggacag gggctaggtt gctgggcact gataattccg tggtgttgtc ggggaagctg 3113

acgtcctttc gaattcgaaa gcttttaaat atcgatgcga tgtacgggcc agatatacgc 3173

gtatctgagg ggactagggt gtgtttaggc gaaaagcggg gcttcggttg tacgcggtta 3233

ggagtcccct taggatatag tagtttcgct tttgcatagg gaggggggaaa tgtagtctta 3293

tgcaatactc ttgtacgtag agatattgtg tctaagtgcc tagctcgata caataaacgc 3353

catttgacca ttcaccacat tggtgtgcac ctgggttgat ggccggaccg ttgattccct 3413

gacgactacg agcacctgca tgaagcagaa ggcttcattc tcgagagctt tggcgtaatc 3473

atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg 3533

agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac tcacattaat 3593

tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc tgcattaatg 3653

aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgct 3713
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 3773
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 3833
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 3893
ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 3953
actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc ctgttccgac 4013
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 4073
atgctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt 4133
gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 4193
caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 4253
agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 4313
tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 4373
tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa 4433
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 4493
gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa 4553
aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat 4613
atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 4673
gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat 4733
acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc 4793
ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtggtcc 4853
tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag 4913
ttcgccagtt aatagtttgc gcaacgttgt tgccattgct gctggcatcg tggtgtcacg 4973
ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg 5033
atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg ttgtcagaag 5093
taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt 5153
catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt cattctgaga 5213
atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata ataccgcgcc 5273
acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc 5333
aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac ccaactgatc 5393
ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc 5453
cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct tccttttttca 5513
atattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 5573
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt 5633
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 5693

tcgtcttcaa g 5704

<210> 10
<211> 239
<212> PRT

<213> Secuencia artificial

<220>
<223> Construcción sintética

<400> 10

```
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                  10                  15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
            100                 105                 110

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

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130                 135                 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                 150                 155                 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
        195                 200                 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220

  Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
  225                 230                 235
```

<210> 11
<211> 10768
<212> ADN
<213> Secuencia artificial

<220>
<223> /note = "Sec ID Nº 11: primer plásmido auxiliar gag-pol (tipo natural)"

<220>
<221> Promotor
<222> (232)..(819)
<223> Promotor de CMV

<220>
<221> Intrón
<222> (907)..(1463)
<223> Intrón de beta globina

<220>
<221> CDS
<222> (1475)..(3586)
<223> gag: secuencia codificante para Gag, tipo natural

<220>
<221> CDS
<222> (3604)..(6291)
<223> pol: secuencia codificante para Pol, tipo natural

<220>
<221> Señal de poli A
<222> (6368)..(6588)
<223> Elemento de poli-adenilación de bGH

<400> 11

```
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg     120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc     180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt     240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata     300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc     360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc     420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt     480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt     540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca     600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg     660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc     720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg     780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca     840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc     900
```

```
gagctcggat cctgagaact tcagggtgag tctatgggac ccttgatgtt ttctttcccc      960

ttcttttcta tggttaagtt catgtcatag gaaggggaga agtaacaggg tacacatatt     1020

gaccaaatca gggtaatttt gcatttgtaa ttttaaaaaa tgctttcttc ttttaatata     1080

ctttttgtt tatcttattt ctaatacttt ccctaatctc tttctttcag ggcaataatg      1140

atacaatgta tcatgcctct ttgcaccatt ctaaagaata acagtgataa tttctgggtt     1200

aaggcaatag caatatttct gcatataaat atttctgcat ataaattgta actgatgtaa     1260

gaggtttcat attgctaata gcagctacaa tccagctacc attctgcttt tattttatgg     1320

ttgggataag gctggattat tctgagtcca agctaggccc ttttgctaat catgttcata     1380

cctcttatct tcctcccaca gctcctgggc aacgtgctgg tctgtgtgct ggcccatcac     1440

tttggcaaag cacgtgagat ctgaattcgc cacc atg gaa gcc gtc att aag gtg      1495
                                        Met Glu Ala Val Ile Lys Val
                                        1               5

att tcg tcc gcg tgt aaa acc tat tgc ggg aaa atc tct cct tct aag       1543
Ile Ser Ser Ala Cys Lys Thr Tyr Cys Gly Lys Ile Ser Pro Ser Lys
        10                  15                  20

aag gaa ata ggg gcc atg ttg tcc ctg tta caa aag gaa ggg ttg ctt       1591
Lys Glu Ile Gly Ala Met Leu Ser Leu Leu Gln Lys Glu Gly Leu Leu
    25                  30                  35

atg tct ccc tca gat tta tat tct ccg ggg tcc tgg gat ccc atc act       1639
Met Ser Pro Ser Asp Leu Tyr Ser Pro Gly Ser Trp Asp Pro Ile Thr
40                  45                  50                  55

gcg gcg ctc tcc cag cgg gca atg gta ctt gga aaa tcg gga gag tta       1687
Ala Ala Leu Ser Gln Arg Ala Met Val Leu Gly Lys Ser Gly Glu Leu
                60                  65                  70

aaa acc tgg gga ttg gtt ttg ggg gca ttg aag gcg gct cga gag gaa       1735
Lys Thr Trp Gly Leu Val Leu Gly Ala Leu Lys Ala Ala Arg Glu Glu
            75                  80                  85

cag gtt aca tct gag caa gca aag ttt tgg ttg gga tta ggg gga ggg       1783
Gln Val Thr Ser Glu Gln Ala Lys Phe Trp Leu Gly Leu Gly Gly Gly
        90                  95                  100

agg gtc tct ccc cca ggt ccg gag tgc atc gag aaa cca gct acg gag       1831
Arg Val Ser Pro Pro Gly Pro Glu Cys Ile Glu Lys Pro Ala Thr Glu
    105                 110                 115

cgg cga atc gac aaa ggg gag gag gtg gga gaa aca act gtg cag cga       1879
Arg Arg Ile Asp Lys Gly Glu Glu Val Gly Glu Thr Thr Val Gln Arg
120                 125                 130                 135

gat gcg aag atg gcg cca gag gaa gcg gcc aca cct aaa acc gtt ggc       1927
Asp Ala Lys Met Ala Pro Glu Glu Ala Ala Thr Pro Lys Thr Val Gly
                140                 145                 150

aca tcc tgc tat cat tgc gga aca gct gtt ggc tgc aat tgc gcc acc       1975
Thr Ser Cys Tyr His Cys Gly Thr Ala Val Gly Cys Asn Cys Ala Thr
            155                 160                 165

gcc aca gcc tcg gcc cct cct ccc cct tat gtg ggg agt ggt ttg tat       2023
Ala Thr Ala Ser Ala Pro Pro Pro Pro Tyr Val Gly Ser Gly Leu Tyr
        170                 175                 180

cct tcc ctg gcg ggg gtg gga gag cag cag ggc cag gga gat aac acg       2071
Pro Ser Leu Ala Gly Val Gly Glu Gln Gln Gly Gln Gly Asp Asn Thr
    185                 190                 195
```

```
tct cgg ggg gcg gag cag cca agg gag gag cca ggg cac gcg ggt cag        2119
Ser Arg Gly Ala Glu Gln Pro Arg Glu Glu Pro Gly His Ala Gly Gln
200                 205                 210                 215

gcc cct ggg ccg gcc ctg act gac tgg gca agg gta agg gag gag ctt        2167
Ala Pro Gly Pro Ala Leu Thr Asp Trp Ala Arg Val Arg Glu Glu Leu
                220                 225                 230

gcg agt act ggt ccg ccc gtg gtg gcc atg cct gta gtg att aag aca        2215
Ala Ser Thr Gly Pro Pro Val Val Ala Met Pro Val Val Ile Lys Thr
            235                 240                 245

gag gga ccc gcc tgg acc cct ctg gag cca aaa ttg atc aca aga ctg        2263
Glu Gly Pro Ala Trp Thr Pro Leu Glu Pro Lys Leu Ile Thr Arg Leu
        250                 255                 260

gct gat acg gtc agg acc aag ggc tta cga tcc ccg atc act atg gca        2311
Ala Asp Thr Val Arg Thr Lys Gly Leu Arg Ser Pro Ile Thr Met Ala
    265                 270                 275

gaa gtg gaa gcg ctc atg tcc tcc ccg ttg ctg ccg cat gac gtc acg        2359
Glu Val Glu Ala Leu Met Ser Ser Pro Leu Leu Pro His Asp Val Thr
280                 285                 290                 295

aat cta atg aga gtg att tta gga cct gcc cca tat gcc tta tgg atg        2407
Asn Leu Met Arg Val Ile Leu Gly Pro Ala Pro Tyr Ala Leu Trp Met
                300                 305                 310

gac gct tgg gga gtc caa ctc cag acg gtt ata gcg gca gcc act cgc        2455
Asp Ala Trp Gly Val Gln Leu Gln Thr Val Ile Ala Ala Ala Thr Arg
            315                 320                 325

gac ccc cga cac cca gcg aac ggt caa ggg cgg ggg gaa cgg act aac        2503
Asp Pro Arg His Pro Ala Asn Gly Gln Gly Arg Gly Glu Arg Thr Asn
        330                 335                 340

ttg gat cga tta aag ggc tta gct gat ggg atg gtg ggc aac cca cag        2551
Leu Asp Arg Leu Lys Gly Leu Ala Asp Gly Met Val Gly Asn Pro Gln
    345                 350                 355

ggt cag gcc gca tta tta aga ccg ggg gaa ttg gtt gct att acg gcg        2599
Gly Gln Ala Ala Leu Leu Arg Pro Gly Glu Leu Val Ala Ile Thr Ala
360                 365                 370                 375

tcg gct ctc cag gcg ttt aga gaa gtt gcc cgg ctg gcg gaa cct gca        2647
Ser Ala Leu Gln Ala Phe Arg Glu Val Ala Arg Leu Ala Glu Pro Ala
                380                 385                 390

ggt cca tgg gcg gac atc acg cag gga cca tct gag tcc ttt gtt gat        2695
Gly Pro Trp Ala Asp Ile Thr Gln Gly Pro Ser Glu Ser Phe Val Asp
            395                 400                 405

ttt gcc aat cgg ctt ata aag gcg gtt gag ggg tca gat ctc ccg cct        2743
Phe Ala Asn Arg Leu Ile Lys Ala Val Glu Gly Ser Asp Leu Pro Pro
        410                 415                 420

tcc gcg cgg gct ccg gtg atc att gac tgc ttt agg cag aag tca cag        2791
Ser Ala Arg Ala Pro Val Ile Ile Asp Cys Phe Arg Gln Lys Ser Gln
    425                 430                 435

cca gat att cag cag ctt ata cgg gca gca ccc tcc acg ctg acc acc        2839
Pro Asp Ile Gln Gln Leu Ile Arg Ala Ala Pro Ser Thr Leu Thr Thr
440                 445                 450                 455

cca gga gag ata atc aaa tat gtg cta gac agg cag aag att gcc cct        2887
Pro Gly Glu Ile Ile Lys Tyr Val Leu Asp Arg Gln Lys Ile Ala Pro
                460                 465                 470
```

```
ctt acg gat caa ggc ata gcc gcg gcc atg tcg tct gct atc cag ccc      2935
Leu Thr Asp Gln Gly Ile Ala Ala Ala Met Ser Ser Ala Ile Gln Pro
            475                 480                 485

tta gtt atg gca gta gtc aat aga gag agg gat gga caa act ggg tcg      2983
Leu Val Met Ala Val Val Asn Arg Glu Arg Asp Gly Gln Thr Gly Ser
        490                 495                 500

ggt ggt cgt gcc cga ggg ctc tgc tac act tgt gga tcc ccg gga cat      3031
Gly Gly Arg Ala Arg Gly Leu Cys Tyr Thr Cys Gly Ser Pro Gly His
    505                 510                 515

tat cag gca cag tgc ccg aaa aaa cga aag tca gga aac agc cgt gag      3079
Tyr Gln Ala Gln Cys Pro Lys Lys Arg Lys Ser Gly Asn Ser Arg Glu
520                 525                 530                 535

cga tgt cag ctg tgt gac ggg atg gga cac aac gct aaa cag tgt agg      3127
Arg Cys Gln Leu Cys Asp Gly Met Gly His Asn Ala Lys Gln Cys Arg
                540                 545                 550

aag cgg gat ggc aac cag ggc caa cgc cca gga aga ggt ctc tct tcg      3175
Lys Arg Asp Gly Asn Gln Gly Gln Arg Pro Gly Arg Gly Leu Ser Ser
            555                 560                 565

ggg ccg tgg ccc ggc cct gag cag cct gcc gtc tcg tta gcg atg aca      3223
Gly Pro Trp Pro Gly Pro Glu Gln Pro Ala Val Ser Leu Ala Met Thr
        570                 575                 580

atg gaa cat aaa gat cgc ccc ttg gtt agg gtc att ctg act aac act      3271
Met Glu His Lys Asp Arg Pro Leu Val Arg Val Ile Leu Thr Asn Thr
    585                 590                 595

ggg agt cat cca gtc aaa caa cgt tcg gtg tat atc acc gcg ctg ttg      3319
Gly Ser His Pro Val Lys Gln Arg Ser Val Tyr Ile Thr Ala Leu Leu
600                 605                 610                 615

gac tcc gga gcg gac atc act att att tcg gag gag gat tgg cct act      3367
Asp Ser Gly Ala Asp Ile Thr Ile Ile Ser Glu Glu Asp Trp Pro Thr
                620                 625                 630

gat tgg ccg gtg gtg gac acc gcg aac cca cag atc cat ggc ata gga      3415
Asp Trp Pro Val Val Asp Thr Ala Asn Pro Gln Ile His Gly Ile Gly
            635                 640                 645

ggg gga att ccc atg cga aaa tcc cgg gat atg ata gag gtg ggg gtt      3463
Gly Gly Ile Pro Met Arg Lys Ser Arg Asp Met Ile Glu Val Gly Val
        650                 655                 660

att aac cga gac ggg tcg ttg gag cga ccc ctg ctc ctc ttc ccc gca      3511
Ile Asn Arg Asp Gly Ser Leu Glu Arg Pro Leu Leu Leu Phe Pro Ala
    665                 670                 675

gtc gct atg gtt aga ggg agt atc cta gga aga gat tgt ctg cag ggc      3559
Val Ala Met Val Arg Gly Ser Ile Leu Gly Arg Asp Cys Leu Gln Gly
680                 685                 690                 695

cta ggg ctc cgc ttg aca aat tta tag ggagggccac tgttctt act gtt      3609
Leu Gly Leu Arg Leu Thr Asn Leu                         Thr Val
                700                                         705

gcg cta cat ctg gct att ccg ctc aaa tgg aag tca gac cgc acg cct      3657
Ala Leu His Leu Ala Ile Pro Leu Lys Trp Lys Ser Asp Arg Thr Pro
                710                 715                 720

gtg tgg att gac cag tgg ccc ctt cct gaa ggt aaa ctt gta gcg cta      3705
Val Trp Ile Asp Gln Trp Pro Leu Pro Glu Gly Lys Leu Val Ala Leu
            725                 730                 735

acg caa tta gtg gaa aaa gaa tta cag tta gga cat ata gag ccc tca      3753
```

```
Thr Gln Leu Val Glu Lys Glu Leu Gln Leu Gly His Ile Glu Pro Ser
        740                 745                 750

ctt agt tgt tgg aac aca cct gtt ttt gtg atc cgg aag gct tcc ggg      3801
Leu Ser Cys Trp Asn Thr Pro Val Phe Val Ile Arg Lys Ala Ser Gly
    755                 760                 765

tct tat cgc tta ttg cat gat ttg cgc gct gtt aac gcc aag ctt gtc      3849
Ser Tyr Arg Leu Leu His Asp Leu Arg Ala Val Asn Ala Lys Leu Val
770                 775                 780                 785

cct ttt ggg gcc gtc caa cag ggg gcg cca gtt ctc tcc gcg ctc ccg      3897
Pro Phe Gly Ala Val Gln Gln Gly Ala Pro Val Leu Ser Ala Leu Pro
                790                 795                 800

cgt ggc tgg ccc ctg atg gtc cta gac ctc aag gat tgc ttc ttt tct      3945
Arg Gly Trp Pro Leu Met Val Leu Asp Leu Lys Asp Cys Phe Phe Ser
            805                 810                 815

atc cct ctt gcg gaa caa gat cgc gaa gct ttt gca ttt acg ctc ccc      3993
Ile Pro Leu Ala Glu Gln Asp Arg Glu Ala Phe Ala Phe Thr Leu Pro
        820                 825                 830

tct gtg aat aac cag gcc ccc gct cga aga ttc caa tgg aag gtc ttg      4041
Ser Val Asn Asn Gln Ala Pro Ala Arg Arg Phe Gln Trp Lys Val Leu
    835                 840                 845

ccc caa ggg atg acc tgt tct ccc act atc tgt cag ttg gta gtg ggt      4089
Pro Gln Gly Met Thr Cys Ser Pro Thr Ile Cys Gln Leu Val Val Gly
850                 855                 860                 865

cag gtg ctc gag ccc ttg cga ctc aag cac cca gct ctg cgc atg ttg      4137
Gln Val Leu Glu Pro Leu Arg Leu Lys His Pro Ala Leu Arg Met Leu
                870                 875                 880

cat tat atg gac gat ctt ttg cta gcc gcc tca agt cat gat ggg ttg      4185
His Tyr Met Asp Asp Leu Leu Leu Ala Ala Ser Ser His Asp Gly Leu
            885                 890                 895

gaa gcg gca ggg aag gag gtt atc ggt aca ttg gaa aga gcc ggg ttc      4233
Glu Ala Ala Gly Lys Glu Val Ile Gly Thr Leu Glu Arg Ala Gly Phe
        900                 905                 910

act att tcg ccg gat aag atc cag agg gag ccc gga gta caa tat ctt      4281
Thr Ile Ser Pro Asp Lys Ile Gln Arg Glu Pro Gly Val Gln Tyr Leu
    915                 920                 925

ggg tac aag tta ggc agt acg tat gta gca ccc gta ggc ttg gta gca      4329
Gly Tyr Lys Leu Gly Ser Thr Tyr Val Ala Pro Val Gly Leu Val Ala
930                 935                 940                 945

gaa ccc agg ata gcc acc ttg tgg gat gtt caa aag ctg gtg ggg tca      4377
Glu Pro Arg Ile Ala Thr Leu Trp Asp Val Gln Lys Leu Val Gly Ser
                950                 955                 960

ctt cag tgg ctt cgc cca gcg tta ggg atc ccg cca cga ctg atg ggt      4425
Leu Gln Trp Leu Arg Pro Ala Leu Gly Ile Pro Pro Arg Leu Met Gly
            965                 970                 975

ccc ttt tat gag cag tta cga ggg tca gat cct aac gag gcg agg gaa      4473
Pro Phe Tyr Glu Gln Leu Arg Gly Ser Asp Pro Asn Glu Ala Arg Glu
        980                 985                 990

tgg aat cta gac atg aaa atg  gcc tgg aga gag atc  gta cag ctt agc    4521
Trp Asn Leu Asp Met Lys Met  Ala Trp Arg Glu Ile  Val Gln Leu Ser
    995                 1000                1005

act  act gct gcc ttg gaa  cga tgg gac cct gcc  cag cct ctg gaa       4566
Thr  Thr Ala Ala Leu Glu  Arg Trp Asp Pro Ala  Gln Pro Leu Glu
```

```
1010                1015                1020

gga gcg gtc gct aga tgt gaa cag ggg gca ata ggg gtc ctg gga 4611
Gly Ala Val Ala Arg Cys Glu Gln Gly Ala Ile Gly Val Leu Gly
1025                1030                1035

cag gga ctg tcc aca cac cca agg cca tgt ttg tgg tta ttc tcc 4656
Gln Gly Leu Ser Thr His Pro Arg Pro Cys Leu Trp Leu Phe Ser
1040                1045                1050

acc caa ccc acc aag gcg ttt act gct tgg tta gaa gtg ctc acc 4701
Thr Gln Pro Thr Lys Ala Phe Thr Ala Trp Leu Glu Val Leu Thr
1055                1060                1065

ctt ttg att act aag cta cgc gct tcg gca gtg cga acc ttt ggc 4746
Leu Leu Ile Thr Lys Leu Arg Ala Ser Ala Val Arg Thr Phe Gly
1070                1075                1080

aag gag gtt gat atc ctc ctg ttg cct gca tgc ttc cgg gag gac 4791
Lys Glu Val Asp Ile Leu Leu Leu Pro Ala Cys Phe Arg Glu Asp
1085                1090                1095

ctt ccg ctc ccg gag ggg atc ctg tta gca ctt agg ggg ttt gca 4836
Leu Pro Leu Pro Glu Gly Ile Leu Leu Ala Leu Arg Gly Phe Ala
1100                1105                1110

gga aaa atc agg agt agt gac acg cca tct att ttt gac att gcg 4881
Gly Lys Ile Arg Ser Ser Asp Thr Pro Ser Ile Phe Asp Ile Ala
1115                1120                1125

cgt cca ctg cat gtt tct ctg aaa gtg agg gtt acc gac cac cct 4926
Arg Pro Leu His Val Ser Leu Lys Val Arg Val Thr Asp His Pro
1130                1135                1140

gtg ccg gga ccc act gtc ttt acc gac gcc tcc tca agc acc cat 4971
Val Pro Gly Pro Thr Val Phe Thr Asp Ala Ser Ser Ser Thr His
1145                1150                1155

aaa ggg gtg gta gtc tgg agg gag ggc cca agg tgg gag ata aaa 5016
Lys Gly Val Val Val Trp Arg Glu Gly Pro Arg Trp Glu Ile Lys
1160                1165                1170

gaa ata gtt gat ttg ggg gca agt gta caa caa ctg gag gca cgc 5061
Glu Ile Val Asp Leu Gly Ala Ser Val Gln Gln Leu Glu Ala Arg
1175                1180                1185

gct gtg gcc atg gca ctt ctg ctg tgg ccg aca acg ccc act aat 5106
Ala Val Ala Met Ala Leu Leu Leu Trp Pro Thr Thr Pro Thr Asn
1190                1195                1200

gta gtg act gac tct gcg ttt gtt gcg aaa atg tta ctc aag atg 5151
Val Val Thr Asp Ser Ala Phe Val Ala Lys Met Leu Leu Lys Met
1205                1210                1215

gga cag gag gga gtc ccg tct aca gcg gca gct ttt att tta gag 5196
Gly Gln Glu Gly Val Pro Ser Thr Ala Ala Ala Phe Ile Leu Glu
1220                1225                1230

gat gcg tta agc caa agg tca gcc atg gcc gcc gtt ctc cac gtg 5241
Asp Ala Leu Ser Gln Arg Ser Ala Met Ala Ala Val Leu His Val
1235                1240                1245

cgg agt cat tct gaa gtg cca ggg ttt ttc aca gaa gga aat gac 5286
Arg Ser His Ser Glu Val Pro Gly Phe Phe Thr Glu Gly Asn Asp
1250                1255                1260

gtg gca gat agc caa gcc acc ttt caa gcg tat ccc ttg aga gag 5331
Val Ala Asp Ser Gln Ala Thr Phe Gln Ala Tyr Pro Leu Arg Glu
1265                1270                1275
```

```
gct  aaa gat ctt cat acc  gct ctc cat att gga  ccc cgc gcg cta        5376
Ala  Lys Asp Leu His Thr  Ala Leu His Ile Gly  Pro Arg Ala Leu
1280                 1285                 1290

tcc  aaa gcg tgt aat ata  tct atg cag cag gct  agg gag gtt gtt        5421
Ser  Lys Ala Cys Asn Ile  Ser Met Gln Gln Ala  Arg Glu Val Val
1295                 1300                 1305

cag  acc tgc ccg cat tgt  aat tca gcc cct gcg  ttg gag gcc ggg        5466
Gln  Thr Cys Pro His Cys  Asn Ser Ala Pro Ala  Leu Glu Ala Gly
1310                 1315                 1320

gta  aac cct agg ggt ttg  gga ccc cta cag ata  tgg cag aca gac        5511
Val  Asn Pro Arg Gly Leu  Gly Pro Leu Gln Ile  Trp Gln Thr Asp
1325                 1330                 1335

ttt  acg ctt gag cct aga  atg gct ccc cgt tcc  tgg ctc gct gtt        5556
Phe  Thr Leu Glu Pro Arg  Met Ala Pro Arg Ser  Trp Leu Ala Val
1340                 1345                 1350

act  gtg gac acc gcc tca  tca gcg ata gtc gta  act cag cat ggc        5601
Thr  Val Asp Thr Ala Ser  Ser Ala Ile Val Val  Thr Gln His Gly
1355                 1360                 1365

cgt  gtt aca tcg gtt gct  gca caa cat cat tgg  gcc acg gct atc        5646
Arg  Val Thr Ser Val Ala  Ala Gln His His Trp  Ala Thr Ala Ile
1370                 1375                 1380

gcc  gtt ttg gga aga cca  aag gcc ata aaa aca  gat aac ggg tcc        5691
Ala  Val Leu Gly Arg Pro  Lys Ala Ile Lys Thr  Asp Asn Gly Ser
1385                 1390                 1395

tgc  ttc acg tct aga tcc  acg cgg gag tgg ctc  gcg aga tgg ggg        5736
Cys  Phe Thr Ser Arg Ser  Thr Arg Glu Trp Leu  Ala Arg Trp Gly
1400                 1405                 1410

ata  gca cac acc acc ggg  att ccg gga aat tcc  cag ggt caa gct        5781
Ile  Ala His Thr Thr Gly  Ile Pro Gly Asn Ser  Gln Gly Gln Ala
1415                 1420                 1425

atg  gta gag cgg gcc aac  cgg ctc ctg aaa gat  aag atc cgt gtg        5826
Met  Val Glu Arg Ala Asn  Arg Leu Leu Lys Asp  Lys Ile Arg Val
1430                 1435                 1440

ctt  gcg gag ggg gac ggc  ttt atg aaa aga atc  ccc gcc agc aaa        5871
Leu  Ala Glu Gly Asp Gly  Phe Met Lys Arg Ile  Pro Ala Ser Lys
1445                 1450                 1455

cag  ggg gaa cta cta gcc  aag gca atg tat gcc  ctc aat cac ttt        5916
Gln  Gly Glu Leu Leu Ala  Lys Ala Met Tyr Ala  Leu Asn His Phe
1460                 1465                 1470

gag  cgt ggt gaa aac aca  aaa aca ccg gta caa  aaa cac tgg aga        5961
Glu  Arg Gly Glu Asn Thr  Lys Thr Pro Val Gln  Lys His Trp Arg
1475                 1480                 1485

cct  acc gtt ctt aca gaa  gga ccc ccg gtt aaa  ata cga ata gag        6006
Pro  Thr Val Leu Thr Glu  Gly Pro Pro Val Lys  Ile Arg Ile Glu
1490                 1495                 1500

aca  ggg gag tgg gaa aaa  gga tgg aac gtg ctg  gtc tgg gga cga        6051
Thr  Gly Glu Trp Glu Lys  Gly Trp Asn Val Leu  Val Trp Gly Arg
1505                 1510                 1515

ggt  tat gcc gct gtg aaa  aac agg gac act gat  aag gtt att tgg        6096
Gly  Tyr Ala Ala Val Lys  Asn Arg Asp Thr Asp  Lys Val Ile Trp
1520                 1525                 1530
```

```
gta  ccc tct cgg aaa gtt  aaa ccg gat gtc acc  caa aag gat gag        6141
Val  Pro Ser Arg Lys Val  Lys Pro Asp Val Thr  Gln Lys Asp Glu
1535                 1540                 1545

gtg  act aag aaa gat gag  gcg agc cct ctt ttt  gca ggc att tct        6186
Val  Thr Lys Lys Asp Glu  Ala Ser Pro Leu Phe  Ala Gly Ile Ser
1550                 1555                 1560

gac  tgg ata ccc tgg gaa  gac gag caa gaa gga  ctc caa gga gaa        6231
Asp  Trp Ile Pro Trp Glu  Asp Glu Gln Glu Gly  Leu Gln Gly Glu
1565                 1570                 1575

acc  gct agc aac aag caa  gaa aga ccc gga gaa  gac acc ctt gct        6276
Thr  Ala Ser Asn Lys Gln  Glu Arg Pro Gly Glu  Asp Thr Leu Ala
1580                 1585                 1590

gcc  aac gag agt taa gaattctgca gatatccagc acagtggcgg ccgctcgagt      6331
Ala  Asn Glu Ser
1595

ctagagggcc cgtttaaacc cgctgatcag cctcgactgt gccttctagt tgccagccat      6391

ctgttgtttg cccctcccc gtgccttcct tgaccctgga aggtgccact cccactgtcc      6451

tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat tctattctgg      6511

ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc aggcatgctg      6571

gggatgcggt gggctctatg gcttctgagg cggaaagaac cagctggggc tctaggggtg      6631

atccccacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg      6691

tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc      6751

tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctccct ttagggttcc      6811

gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat ggttcacgta      6871

gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta      6931

atagtggact cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg      6991

atttataagg gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa      7051

aatttaacgc gaattaattc tgtggaatgt gtgtcagtta gggtgtggaa agtccccagg      7111

ctccccagca ggcagaagta tgcaaagcat gcatctcaat tagtcagcaa ccaggtgtgg      7171

aaagtcccca ggctccccag caggcagaag tatgcaaagc atgcatctca attagtcagc      7231

aaccatagtc ccgcccctaa ctccgcccat cccgccccta actccgccca gttccgccca      7291

ttctccgccc catggctgac taatttttt tatttatgca gaggccgagg ccgcctctgc      7351

ctctgagcta ttccagaagt agtgaggagg ctttttttgga ggcctaggct tttgcaaaaa      7411

gctcccggga gcttgtatat ccattttcgg atctgatcaa gagacaggat gaggatcgtt      7471

tcgcatgatt gaacaagatg gattgcacgc aggttctccg gccgcttggg tggagaggct      7531

attcggctat gactgggcac aacagacaat cggctgctct gatgccgccg tgttccggct      7591

gtcagcgcag gggcgcccgg ttctttttgt caagaccgac ctgtccggtg ccctgaatga      7651

actgcaggac gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc      7711

tgtgctcgac gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg      7771

gcaggatctc ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc      7831
```

aatgcggcgg ctgcatacgc ttgatccggc tacctgccca ttcgaccacc aagcgaaaca 7891
tcgcatcgag cgagcacgta ctcggatgga agccggtctt gtcgatcagg atgatctgga 7951
cgaagagcat caggggctcg cgccagccga actgttcgcc aggctcaagg cgcgcatgcc 8011
cgacggcgag gatctcgtcg tgacccatgg cgatgcctgc ttgccgaata tcatggtgga 8071
aaatggccgc ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca 8131
ggacatagcg ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg 8191
cttcctcgtg ctttacggta tcgccgctcc cgattcgcag cgcatcgcct tctatcgcct 8251
tcttgacgag ttcttctgag cgggactctg gggttcgaaa tgaccgacca agcgacgccc 8311
aacctgccat cacgagattt cgattccacc gccgccttct atgaaaggtt gggcttcgga 8371
atcgttttcc gggacgccgg ctggatgatc ctccagcgcg gggatctcat gctggagttc 8431
ttcgcccacc ccaacttgtt tattgcagct tataatggtt acaaataaag caatagcatc 8491
acaaatttca caaataaagc attttttca ctgcattcta gttgtggttt gtccaaactc 8551
atcaatgtat cttatcatgt ctgtataccg tcgacctcta gctagagctt ggcgtaatca 8611
tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga 8671
gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact cacattaatt 8731
gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga 8791
atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc 8851
actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg 8911
gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc 8971
cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc 9031
cccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga 9091
ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc 9151
ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat 9211
agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg 9271
cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc 9331
aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga 9391
gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact 9451
agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt 9511
ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtttttttgt ttgcaagcag 9571
cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct 9631
gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg 9691
atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat 9751
gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc 9811
tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac tacgatacgg 9871

```
gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg ctcaccggct     9931
ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggtcctgca     9991
actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg    10051
ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt gtcacgctcg    10111
tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt tacatgatcc    10171
cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt cagaagtaag    10231
ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg    10291
ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt ctgagaatag    10351
tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac cgcgccacat    10411
agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa actctcaagg    10471
atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa ctgatcttca    10531
gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgccgca    10591
aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct ttttcaatat    10651
tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag    10711
aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtc       10768
```

<210> 12
<211> 703
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción sintética

<400> 12

```
Met Glu Ala Val Ile Lys Val Ile Ser Ser Ala Cys Lys Thr Tyr Cys
1               5                  10                  15

Gly Lys Ile Ser Pro Ser Lys Lys Glu Ile Gly Ala Met Leu Ser Leu
            20                  25                  30

Leu Gln Lys Glu Gly Leu Leu Met Ser Pro Ser Asp Leu Tyr Ser Pro
        35                  40                  45

Gly Ser Trp Asp Pro Ile Thr Ala Ala Leu Ser Gln Arg Ala Met Val
    50                  55                  60

Leu Gly Lys Ser Gly Glu Leu Lys Thr Trp Gly Leu Val Leu Gly Ala
65                  70                  75                  80

Leu Lys Ala Ala Arg Glu Glu Gln Val Thr Ser Glu Gln Ala Lys Phe
                85                  90                  95

Trp Leu Gly Leu Gly Gly Gly Arg Val Ser Pro Pro Gly Pro Glu Cys
            100                 105                 110
```

```
Ile Glu Lys Pro Ala Thr Glu Arg Arg Ile Asp Lys Gly Glu Glu Val
        115                 120                 125

Gly Glu Thr Thr Val Gln Arg Asp Ala Lys Met Ala Pro Glu Glu Ala
    130                 135                 140

Ala Thr Pro Lys Thr Val Gly Thr Ser Cys Tyr His Cys Gly Thr Ala
145                 150                 155                 160

Val Gly Cys Asn Cys Ala Thr Ala Thr Ala Ser Ala Pro Pro Pro Pro
                165                 170                 175

Tyr Val Gly Ser Gly Leu Tyr Pro Ser Leu Ala Gly Val Gly Glu Gln
            180                 185                 190

Gln Gly Gln Gly Asp Asn Thr Ser Arg Gly Ala Glu Gln Pro Arg Glu
        195                 200                 205

Glu Pro Gly His Ala Gly Gln Ala Pro Gly Pro Ala Leu Thr Asp Trp
    210                 215                 220

Ala Arg Val Arg Glu Glu Leu Ala Ser Thr Gly Pro Pro Val Val Ala
225                 230                 235                 240

Met Pro Val Val Ile Lys Thr Glu Gly Pro Ala Trp Thr Pro Leu Glu
                245                 250                 255

Pro Lys Leu Ile Thr Arg Leu Ala Asp Thr Val Arg Thr Lys Gly Leu
            260                 265                 270

Arg Ser Pro Ile Thr Met Ala Glu Val Glu Ala Leu Met Ser Ser Pro
        275                 280                 285

Leu Leu Pro His Asp Val Thr Asn Leu Met Arg Val Ile Leu Gly Pro
    290                 295                 300

Ala Pro Tyr Ala Leu Trp Met Asp Ala Trp Gly Val Gln Leu Gln Thr
305                 310                 315                 320

Val Ile Ala Ala Ala Thr Arg Asp Pro Arg His Pro Ala Asn Gly Gln
                325                 330                 335

Gly Arg Gly Glu Arg Thr Asn Leu Asp Arg Leu Lys Gly Leu Ala Asp
            340                 345                 350

Gly Met Val Gly Asn Pro Gln Gly Gln Ala Ala Leu Leu Arg Pro Gly
        355                 360                 365

Glu Leu Val Ala Ile Thr Ala Ser Ala Leu Gln Ala Phe Arg Glu Val
    370                 375                 380

Ala Arg Leu Ala Glu Pro Ala Gly Pro Trp Ala Asp Ile Thr Gln Gly
```

385 390 395 400

Pro Ser Glu Ser Phe Val Asp Phe Ala Asn Arg Leu Ile Lys Ala Val
405 410 415

Glu Gly Ser Asp Leu Pro Pro Ser Ala Arg Ala Pro Val Ile Ile Asp
420 425 430

Cys Phe Arg Gln Lys Ser Gln Pro Asp Ile Gln Gln Leu Ile Arg Ala
435 440 445

Ala Pro Ser Thr Leu Thr Thr Pro Gly Glu Ile Ile Lys Tyr Val Leu
450 455 460

Asp Arg Gln Lys Ile Ala Pro Leu Thr Asp Gln Gly Ile Ala Ala Ala
465 470 475 480

Met Ser Ser Ala Ile Gln Pro Leu Val Met Ala Val Val Asn Arg Glu
485 490 495

Arg Asp Gly Gln Thr Gly Ser Gly Gly Arg Ala Arg Gly Leu Cys Tyr
500 505 510

Thr Cys Gly Ser Pro Gly His Tyr Gln Ala Gln Cys Pro Lys Lys Arg
515 520 525

Lys Ser Gly Asn Ser Arg Glu Arg Cys Gln Leu Cys Asp Gly Met Gly
530 535 540

His Asn Ala Lys Gln Cys Arg Lys Arg Asp Gly Asn Gln Gly Gln Arg
545 550 555 560

Pro Gly Arg Gly Leu Ser Ser Gly Pro Trp Pro Gly Pro Glu Gln Pro
565 570 575

Ala Val Ser Leu Ala Met Thr Met Glu His Lys Asp Arg Pro Leu Val
580 585 590

Arg Val Ile Leu Thr Asn Thr Gly Ser His Pro Val Lys Gln Arg Ser
595 600 605

Val Tyr Ile Thr Ala Leu Leu Asp Ser Gly Ala Asp Ile Thr Ile Ile
610 615 620

Ser Glu Glu Asp Trp Pro Thr Asp Trp Pro Val Val Asp Thr Ala Asn
625 630 635 640

Pro Gln Ile His Gly Ile Gly Gly Gly Ile Pro Met Arg Lys Ser Arg
645 650 655

Asp Met Ile Glu Val Gly Val Ile Asn Arg Asp Gly Ser Leu Glu Arg
660 665 670

Pro Leu Leu Leu Phe Pro Ala Val Ala Met Val Arg Gly Ser Ile Leu
675 680 685

Gly Arg Asp Cys Leu Gln Gly Leu Gly Leu Arg Leu Thr Asn Leu
690 695 700

<210> 13
<211> 895
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción sintética

<400> 13

```
Thr Val Ala Leu His Leu Ala Ile Pro Leu Lys Trp Lys Ser Asp Arg
1               5                  10                  15

Thr Pro Val Trp Ile Asp Gln Trp Pro Leu Pro Glu Gly Lys Leu Val
            20                  25                  30

Ala Leu Thr Gln Leu Val Glu Lys Glu Leu Gln Leu Gly His Ile Glu
        35                  40                  45

Pro Ser Leu Ser Cys Trp Asn Thr Pro Val Phe Val Ile Arg Lys Ala
    50                  55                  60

Ser Gly Ser Tyr Arg Leu Leu His Asp Leu Arg Ala Val Asn Ala Lys
65                  70                  75                  80

Leu Val Pro Phe Gly Ala Val Gln Gln Gly Ala Pro Val Leu Ser Ala
                85                  90                  95

Leu Pro Arg Gly Trp Pro Leu Met Val Leu Asp Leu Lys Asp Cys Phe
            100                 105                 110

Phe Ser Ile Pro Leu Ala Glu Gln Asp Arg Glu Ala Phe Ala Phe Thr
        115                 120                 125

Leu Pro Ser Val Asn Asn Gln Ala Pro Ala Arg Arg Phe Gln Trp Lys
    130                 135                 140

Val Leu Pro Gln Gly Met Thr Cys Ser Pro Thr Ile Cys Gln Leu Val
145                 150                 155                 160

Val Gly Gln Val Leu Glu Pro Leu Arg Leu Lys His Pro Ala Leu Arg
                165                 170                 175

Met Leu His Tyr Met Asp Asp Leu Leu Leu Ala Ala Ser Ser His Asp
            180                 185                 190

Gly Leu Glu Ala Ala Gly Lys Glu Val Ile Gly Thr Leu Glu Arg Ala
```

195 200 205

Gly Phe Thr Ile Ser Pro Asp Lys Ile Gln Arg Glu Pro Gly Val Gln
210 215 220

Tyr Leu Gly Tyr Lys Leu Gly Ser Thr Tyr Val Ala Pro Val Gly Leu
225 230 235 240

Val Ala Glu Pro Arg Ile Ala Thr Leu Trp Asp Val Gln Lys Leu Val
245 250 255

Gly Ser Leu Gln Trp Leu Arg Pro Ala Leu Gly Ile Pro Pro Arg Leu
260 265 270

Met Gly Pro Phe Tyr Glu Gln Leu Arg Gly Ser Asp Pro Asn Glu Ala
275 280 285

Arg Glu Trp Asn Leu Asp Met Lys Met Ala Trp Arg Glu Ile Val Gln
290 295 300

Leu Ser Thr Thr Ala Ala Leu Glu Arg Trp Asp Pro Ala Gln Pro Leu
305 310 315 320

Glu Gly Ala Val Ala Arg Cys Glu Gln Gly Ala Ile Gly Val Leu Gly
325 330 335

Gln Gly Leu Ser Thr His Pro Arg Pro Cys Leu Trp Leu Phe Ser Thr
340 345 350

Gln Pro Thr Lys Ala Phe Thr Ala Trp Leu Glu Val Leu Thr Leu Leu
355 360 365

Ile Thr Lys Leu Arg Ala Ser Ala Val Arg Thr Phe Gly Lys Glu Val
370 375 380

Asp Ile Leu Leu Leu Pro Ala Cys Phe Arg Glu Asp Leu Pro Leu Pro
385 390 395 400

Glu Gly Ile Leu Leu Ala Leu Arg Gly Phe Ala Gly Lys Ile Arg Ser
405 410 415

Ser Asp Thr Pro Ser Ile Phe Asp Ile Ala Arg Pro Leu His Val Ser
420 425 430

Leu Lys Val Arg Val Thr Asp His Pro Val Pro Gly Pro Thr Val Phe
435 440 445

Thr Asp Ala Ser Ser Ser Thr His Lys Gly Val Val Val Trp Arg Glu
450 455 460

Gly Pro Arg Trp Glu Ile Lys Glu Ile Val Asp Leu Gly Ala Ser Val
465 470 475 480

Gln Gln Leu Glu Ala Arg Ala Val Ala Met Ala Leu Leu Leu Trp Pro
485 490 495

Thr Thr Pro Thr Asn Val Val Thr Asp Ser Ala Phe Val Ala Lys Met
500 505 510

Leu Leu Lys Met Gly Gln Glu Gly Val Pro Ser Thr Ala Ala Ala Phe
515 520 525

Ile Leu Glu Asp Ala Leu Ser Gln Arg Ser Ala Met Ala Ala Val Leu
530 535 540

His Val Arg Ser His Ser Glu Val Pro Gly Phe Phe Thr Glu Gly Asn
545 550 555 560

Asp Val Ala Asp Ser Gln Ala Thr Phe Gln Ala Tyr Pro Leu Arg Glu
565 570 575

Ala Lys Asp Leu His Thr Ala Leu His Ile Gly Pro Arg Ala Leu Ser
580 585 590

Lys Ala Cys Asn Ile Ser Met Gln Gln Ala Arg Glu Val Val Gln Thr
595 600 605

Cys Pro His Cys Asn Ser Ala Pro Ala Leu Glu Ala Gly Val Asn Pro
610 615 620

Arg Gly Leu Gly Pro Leu Gln Ile Trp Gln Thr Asp Phe Thr Leu Glu
625 630 635 640

Pro Arg Met Ala Pro Arg Ser Trp Leu Ala Val Thr Val Asp Thr Ala
645 650 655

Ser Ser Ala Ile Val Val Thr Gln His Gly Arg Val Thr Ser Val Ala
660 665 670

Ala Gln His His Trp Ala Thr Ala Ile Ala Val Leu Gly Arg Pro Lys
675 680 685

Ala Ile Lys Thr Asp Asn Gly Ser Cys Phe Thr Ser Arg Ser Thr Arg
690 695 700

Glu Trp Leu Ala Arg Trp Gly Ile Ala His Thr Thr Gly Ile Pro Gly
705 710 715 720

Asn Ser Gln Gly Gln Ala Met Val Glu Arg Ala Asn Arg Leu Leu Lys
725 730 735

Asp Lys Ile Arg Val Leu Ala Glu Gly Asp Gly Phe Met Lys Arg Ile
740 745 750

```
Pro Ala Ser Lys Gln Gly Glu Leu Leu Ala Lys Ala Met Tyr Ala Leu
        755                 760                 765

Asn His Phe Glu Arg Gly Glu Asn Thr Lys Thr Pro Val Gln Lys His
    770                 775                 780

Trp Arg Pro Thr Val Leu Thr Glu Gly Pro Pro Val Lys Ile Arg Ile
785                 790                 795                 800

Glu Thr Gly Glu Trp Glu Lys Gly Trp Asn Val Leu Val Trp Gly Arg
                805                 810                 815

Gly Tyr Ala Ala Val Lys Asn Arg Asp Thr Asp Lys Val Ile Trp Val
            820                 825                 830

Pro Ser Arg Lys Val Lys Pro Asp Val Thr Gln Lys Asp Glu Val Thr
        835                 840                 845

Lys Lys Asp Glu Ala Ser Pro Leu Phe Ala Gly Ile Ser Asp Trp Ile
    850                 855                 860

Pro Trp Glu Asp Glu Gln Glu Gly Leu Gln Gly Glu Thr Ala Ser Asn
865                 870                 875                 880

Lys Gln Glu Arg Pro Gly Glu Asp Thr Leu Ala Ala Asn Glu Ser
                885                 890                 895
```

<210> 14
<211> 10768
<212> ADN
<213> Secuencia artificial

<220>
<223> /note = "Sec ID Nº 14: primer plásmido auxiliar para gag-pol optimizado en cuanto a codones"

<220>
<221> Promotor
<222> (232)..(819)
<223> Promotor de CMV

<220>
<221> Intrón
<222> (907)..(1463)
<223> Intrón de beta globina

<220>
<221> CDS
<222> (1475)..(3586)
<223> gal: secuencia codificante para Gal, optimizada en cuanto a codones

<220>
<221> misc_feature
<222> (3437)..(3737)
<223> Zona de transición entre zonas que codifican gag y pol (tipo natural)

<220>
<221> misc_feature
<222> (3574)..(3723)

<223> Señal para cambio de fase de lectura (frameshift signal)

<220>
<221> CDS
<222> (3604)..(6291)
<223> pol: secuencia codificante para Pol, optimizada en cuanto a codones

<220>
<221> Señal de poli A
<222> (6368)..(6588)
<223> Señal de poliadenilación de bGH

<400> 14

```
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg       60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg      120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc      180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt      240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata      300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc      360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc      420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt      480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt      540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca      600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg      660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc      720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg      780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca      840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc      900
gagctcggat cctgagaact tcagggtgag tctatgggac ccttgatgtt ttctttcccc      960
ttcttttcta tggttaagtt catgtcatag gaaggggaga agtaacaggg tacacatatt     1020
gaccaaatca gggtaatttt gcatttgtaa ttttaaaaaa tgctttcttc ttttaatata     1080
ctttttttgtt tatcttattt ctaatacttt ccctaatctc tttctttcag ggcaataatg     1140
atacaatgta tcatgcctct ttgcaccatt ctaaagaata acagtgataa tttctgggtt     1200
aaggcaatag caatatttct gcatataaat atttctgcat ataaattgta actgatgtaa     1260
gaggtttcat attgctaata gcagctacaa tccagctacc attctgcttt tattttatgg     1320
ttgggataag gctggattat tctgagtcca agctaggccc ttttgctaat catgttcata     1380
cctcttatct tcctcccaca gctcctgggc aacgtgctgg tctgtgtgct ggcccatcac     1440
tttggcaaag cacgtgagat ctgaattcgc cacc atg gaa gcc gtg atc aaa gtg     1495
                                       Met Glu Ala Val Ile Lys Val
                                       1               5

atc agc agc gcc tgc aag acc tac tgc ggc aag atc agc ccc agc aag       1543
Ile Ser Ser Ala Cys Lys Thr Tyr Cys Gly Lys Ile Ser Pro Ser Lys
```

```
        10                  15                  20

aaa gaa atc ggc gct atg ctg tct ctc ctg caa aag gaa ggc ctg ctg      1591
Lys Glu Ile Gly Ala Met Leu Ser Leu Leu Gln Lys Glu Gly Leu Leu
    25                  30                  35

atg agc ccc agc gac ctg tac agc ccc ggc agc tgg gat cct atc aca      1639
Met Ser Pro Ser Asp Leu Tyr Ser Pro Gly Ser Trp Asp Pro Ile Thr
40                  45                  50                  55

gcc gcc ctg agc cag agg gcc atg gtg ctg ggc aag agc ggc gag ctg      1687
Ala Ala Leu Ser Gln Arg Ala Met Val Leu Gly Lys Ser Gly Glu Leu
                60                  65                  70

aaa acc tgg ggg ctg gtg ctg gga gcc ctg aag gcc gcc aga gaa gaa      1735
Lys Thr Trp Gly Leu Val Leu Gly Ala Leu Lys Ala Ala Arg Glu Glu
            75                  80                  85

cag gtg acc agc gag cag gcc aag ttt tgg ctg ggc ctg ggc gga gga      1783
Gln Val Thr Ser Glu Gln Ala Lys Phe Trp Leu Gly Leu Gly Gly Gly
        90                  95                  100

aga gtg tct ccc cct ggc ccc gag tgt atc gag aag ccc gcc acc gag      1831
Arg Val Ser Pro Pro Gly Pro Glu Cys Ile Glu Lys Pro Ala Thr Glu
    105                 110                 115

cgg aga atc gac aag ggc gag gaa gtg ggc gag aca acc gtg cag cgg      1879
Arg Arg Ile Asp Lys Gly Glu Glu Val Gly Glu Thr Thr Val Gln Arg
120                 125                 130                 135

gac gcc aag atg gcc cct gag gaa gcc gcc acc cct aag acc gtg ggc      1927
Asp Ala Lys Met Ala Pro Glu Glu Ala Ala Thr Pro Lys Thr Val Gly
                140                 145                 150

acc agc tgc tac cac tgt ggc acc gcc gtg ggc tgt aat tgt gcc acc      1975
Thr Ser Cys Tyr His Cys Gly Thr Ala Val Gly Cys Asn Cys Ala Thr
            155                 160                 165

gcc acc gcc tct gcc cct cct cct cct tac gtg ggc agc ggc ctg tat      2023
Ala Thr Ala Ser Ala Pro Pro Pro Pro Tyr Val Gly Ser Gly Leu Tyr
        170                 175                 180

cct tct ctg gcc ggc gtg ggc gag cag cag gga cag ggc gac aat acc      2071
Pro Ser Leu Ala Gly Val Gly Glu Gln Gln Gly Gln Gly Asp Asn Thr
    185                 190                 195

agc aga ggc gcc gag cag cct aga gag gaa cct gga cac gct ggc cag      2119
Ser Arg Gly Ala Glu Gln Pro Arg Glu Glu Pro Gly His Ala Gly Gln
200                 205                 210                 215

gcc cca gga cct gcc ctg aca gat tgg gcc aga gtg cgg gag gaa ctg      2167
Ala Pro Gly Pro Ala Leu Thr Asp Trp Ala Arg Val Arg Glu Glu Leu
                220                 225                 230

gcc tct acc gga ccc cct gtg gtg gcc atg ccc gtg gtg atc aag aca      2215
Ala Ser Thr Gly Pro Pro Val Val Ala Met Pro Val Val Ile Lys Thr
            235                 240                 245

gag ggc cct gcc tgg acc cct ctg gaa ccc aag ctg atc acc cgg ctg      2263
Glu Gly Pro Ala Trp Thr Pro Leu Glu Pro Lys Leu Ile Thr Arg Leu
        250                 255                 260

gcc gat aca gtg cgg acc aag ggc ctg aga agc ccc atc acc atg gcc      2311
Ala Asp Thr Val Arg Thr Lys Gly Leu Arg Ser Pro Ile Thr Met Ala
    265                 270                 275

gag gtg gag gcc ctg atg agc agc ccc ctg ctg ccc cac gac gtg acc      2359
Glu Val Glu Ala Leu Met Ser Ser Pro Leu Leu Pro His Asp Val Thr
280                 285                 290                 295
```

```
aac ctg atg aga gtg atc ctg ggc cct gct ccc tac gcc ctg tgg atg        2407
Asn Leu Met Arg Val Ile Leu Gly Pro Ala Pro Tyr Ala Leu Trp Met
                300                 305                 310

gat gcc tgg ggc gtg cag ctg cag aca gtg atc gcc gct gcc acc aga        2455
Asp Ala Trp Gly Val Gln Leu Gln Thr Val Ile Ala Ala Ala Thr Arg
            315                 320                 325

gat ccc aga cac ccc gcc aat ggc cag ggc aga ggc gag cgg acc aac        2503
Asp Pro Arg His Pro Ala Asn Gly Gln Gly Arg Gly Glu Arg Thr Asn
        330                 335                 340

ctg gac aga ctg aag ggc ctg gcc gac ggc atg gtg ggc aat cct cag        2551
Leu Asp Arg Leu Lys Gly Leu Ala Asp Gly Met Val Gly Asn Pro Gln
    345                 350                 355

gga cag gcc gct ctg ctg agg cct ggc gaa ctg gtg gcc atc aca gcc        2599
Gly Gln Ala Ala Leu Leu Arg Pro Gly Glu Leu Val Ala Ile Thr Ala
360                 365                 370                 375

agc gcc ctg cag gcc ttc aga gag gtg gcc aga ctg gcc gag cct gcc        2647
Ser Ala Leu Gln Ala Phe Arg Glu Val Ala Arg Leu Ala Glu Pro Ala
                380                 385                 390

ggc cct tgg gcc gat atc acc cag ggc ccc agc gag agc ttc gtg gac        2695
Gly Pro Trp Ala Asp Ile Thr Gln Gly Pro Ser Glu Ser Phe Val Asp
            395                 400                 405

ttc gcc aac cgg ctg atc aag gcc gtg gag ggc agc gat ctg cct cct        2743
Phe Ala Asn Arg Leu Ile Lys Ala Val Glu Gly Ser Asp Leu Pro Pro
        410                 415                 420

agc gcc aga gcc ccc gtg atc atc gac tgc ttc cgg cag aag tcc cag        2791
Ser Ala Arg Ala Pro Val Ile Ile Asp Cys Phe Arg Gln Lys Ser Gln
    425                 430                 435

ccc gac atc cag cag ctg atc aga gcc gcc cct agc acc ctg acc acc        2839
Pro Asp Ile Gln Gln Leu Ile Arg Ala Ala Pro Ser Thr Leu Thr Thr
440                 445                 450                 455

cct ggc gag atc atc aaa tac gtg ctg gac cgg cag aag atc gcc cct        2887
Pro Gly Glu Ile Ile Lys Tyr Val Leu Asp Arg Gln Lys Ile Ala Pro
                460                 465                 470

ctg acc gat cag ggc att gcc gcc gct atg agc agc gcc atc cag cct        2935
Leu Thr Asp Gln Gly Ile Ala Ala Ala Met Ser Ser Ala Ile Gln Pro
            475                 480                 485

ctg gtg atg gcc gtg gtg aac aga gag agg gac ggc cag aca gga tct        2983
Leu Val Met Ala Val Val Asn Arg Glu Arg Asp Gly Gln Thr Gly Ser
        490                 495                 500

ggc ggc aga gcc aga ggc ctg tgc tac acc tgt ggc agc cct ggc cac        3031
Gly Gly Arg Ala Arg Gly Leu Cys Tyr Thr Cys Gly Ser Pro Gly His
    505                 510                 515

tac cag gcc cag tgc ccc aag aag cgg aag tcc ggc aac agc cgg gag        3079
Tyr Gln Ala Gln Cys Pro Lys Lys Arg Lys Ser Gly Asn Ser Arg Glu
520                 525                 530                 535

aga tgc cag ctg tgt gac ggc atg ggc cac aac gcc aag cag tgc cgg        3127
Arg Cys Gln Leu Cys Asp Gly Met Gly His Asn Ala Lys Gln Cys Arg
                540                 545                 550

aag cgg gat ggc aat cag ggc cag agg ccc ggc aga gga ctg tct agc        3175
Lys Arg Asp Gly Asn Gln Gly Gln Arg Pro Gly Arg Gly Leu Ser Ser
            555                 560                 565
```

```
ggc cct tgg cct gga cct gag cag cct gcc gtg agc ctg gcc atg acc        3223
Gly Pro Trp Pro Gly Pro Glu Gln Pro Ala Val Ser Leu Ala Met Thr
        570                 575                 580

atg gaa cac aag gac cgg ccc ctg gtg cgc gtg atc ctg acc aac acc        3271
Met Glu His Lys Asp Arg Pro Leu Val Arg Val Ile Leu Thr Asn Thr
    585                 590                 595

ggc agc cac ccc gtg aag cag cgg agc gtg tac atc acc gcc ctg ctg        3319
Gly Ser His Pro Val Lys Gln Arg Ser Val Tyr Ile Thr Ala Leu Leu
600                 605                 610                 615

gac agc gga gcc gac atc acc atc atc agc gaa gag gac tgg ccc acc        3367
Asp Ser Gly Ala Asp Ile Thr Ile Ile Ser Glu Glu Asp Trp Pro Thr
                620                 625                 630

gac tgg cct gtg gtg gac acc gcc aac ccc cag atc cac ggc atc ggc        3415
Asp Trp Pro Val Val Asp Thr Ala Asn Pro Gln Ile His Gly Ile Gly
            635                 640                 645

gga gga atc ccc atg cgg aag tcc cgg gat atg ata gag gtg ggg gtt        3463
Gly Gly Ile Pro Met Arg Lys Ser Arg Asp Met Ile Glu Val Gly Val
        650                 655                 660

att aac cga gac ggg tcg ttg gag cga ccc ctg ctc ctc ttc ccc gca        3511
Ile Asn Arg Asp Gly Ser Leu Glu Arg Pro Leu Leu Leu Phe Pro Ala
    665                 670                 675

gtc gct atg gtt aga ggg agt atc cta gga aga gat tgt ctg cag ggc        3559
Val Ala Met Val Arg Gly Ser Ile Leu Gly Arg Asp Cys Leu Gln Gly
680                 685                 690                 695

cta ggg ctc cgc ttg aca aat tta tag ggagggccac tgttctt act gtt        3609
Leu Gly Leu Arg Leu Thr Asn Leu                         Thr Val
                700                                         705

gcg cta cat ctg gct att ccg ctc aaa tgg aag tca gac cgc acg cct        3657
Ala Leu His Leu Ala Ile Pro Leu Lys Trp Lys Ser Asp Arg Thr Pro
                710                 715                 720

gtg tgg att gac cag tgg ccc ctt cct gaa ggt aaa ctt gta gcg cta        3705
Val Trp Ile Asp Gln Trp Pro Leu Pro Glu Gly Lys Leu Val Ala Leu
            725                 730                 735

acg caa tta gtg gaa aaa gaa tta cag tta ggc cac atc gag ccc agc        3753
Thr Gln Leu Val Glu Lys Glu Leu Gln Leu Gly His Ile Glu Pro Ser
        740                 745                 750

ctg agc tgc tgg aac acc ccc gtg ttc gtg atc aga aag gcc agc ggc        3801
Leu Ser Cys Trp Asn Thr Pro Val Phe Val Ile Arg Lys Ala Ser Gly
    755                 760                 765

agc tac aga ctg ctg cac gac ctg cgg gcc gtg aat gcc aag ctg gtg        3849
Ser Tyr Arg Leu Leu His Asp Leu Arg Ala Val Asn Ala Lys Leu Val
770                 775                 780                 785

ccc ttt ggc gcc gtg cag cag ggc gct cct gtg ctg tct gcc ctg cca        3897
Pro Phe Gly Ala Val Gln Gln Gly Ala Pro Val Leu Ser Ala Leu Pro
                790                 795                 800

aga gga tgg ccc ctg atg gtg ctg gac ctg aag gac tgc ttc ttc agc        3945
Arg Gly Trp Pro Leu Met Val Leu Asp Leu Lys Asp Cys Phe Phe Ser
            805                 810                 815

atc cct ctg gcc gag cag gac aga gag gcc ttc gcc ttc acc ctg cct        3993
Ile Pro Leu Ala Glu Gln Asp Arg Glu Ala Phe Ala Phe Thr Leu Pro
        820                 825                 830

agc gtg aac aac cag gcc cca gcc aga cgg ttc cag tgg aag gtg ctg        4041
```

```
Ser Val Asn Asn Gln Ala Pro Ala Arg Arg Phe Gln Trp Lys Val Leu
    835                 840                 845

ccc cag ggc atg acc tgc agc ccc acc atc tgt cag ctg gtc gtg gga      4089
Pro Gln Gly Met Thr Cys Ser Pro Thr Ile Cys Gln Leu Val Val Gly
850                 855                 860                 865

cag gtg ctg gaa ccc ctg aga ctg aag cac ccc gcc ctg cgg atg ctg      4137
Gln Val Leu Glu Pro Leu Arg Leu Lys His Pro Ala Leu Arg Met Leu
                870                 875                 880

cac tac atg gac gac ctg ctc ctc gcc gct tct agc cac gac ggc ctg      4185
His Tyr Met Asp Asp Leu Leu Leu Ala Ala Ser Ser His Asp Gly Leu
            885                 890                 895

gaa gcc gcc gga aaa gaa gtg atc ggc acc ctg gaa aga gcc ggc ttc      4233
Glu Ala Ala Gly Lys Glu Val Ile Gly Thr Leu Glu Arg Ala Gly Phe
        900                 905                 910

acc atc agc ccc gac aag atc cag agg gaa ccc ggc gtg cag tac ctg      4281
Thr Ile Ser Pro Asp Lys Ile Gln Arg Glu Pro Gly Val Gln Tyr Leu
    915                 920                 925

ggc tac aag ctg ggc agc acc tat gtg gcc cct gtg ggc ctg gtg gcc      4329
Gly Tyr Lys Leu Gly Ser Thr Tyr Val Ala Pro Val Gly Leu Val Ala
930                 935                 940                 945

gag cca aga atc gcc acc ctg tgg gac gtg cag aaa ctg gtg ggc agc      4377
Glu Pro Arg Ile Ala Thr Leu Trp Asp Val Gln Lys Leu Val Gly Ser
                950                 955                 960

ctg cag tgg ctg agg cca gcc ctg ggc atc cct ccc aga ctg atg ggc      4425
Leu Gln Trp Leu Arg Pro Ala Leu Gly Ile Pro Pro Arg Leu Met Gly
            965                 970                 975

ccc ttc tac gag cag ctg cgg ggc tcc gac cct aat gag gcc cgg gag      4473
Pro Phe Tyr Glu Gln Leu Arg Gly Ser Asp Pro Asn Glu Ala Arg Glu
        980                 985                 990

tgg aac ctg gat atg aag atg  gcc tgg cgg gag atc  gtg cag ctg tcc    4521
Trp Asn Leu Asp Met Lys Met  Ala Trp Arg Glu Ile  Val Gln Leu Ser
    995                 1000                1005

acc  act gcc gcc ctg gaa  cgc tgg gat cca gca  cag cca ctc gaa       4566
Thr  Thr Ala Ala Leu Glu  Arg Trp Asp Pro Ala  Gln Pro Leu Glu
1010                1015                1020

ggt  gcc gtg gcc aga tgt  gaa cag ggc gcc atc  ggc gtg ctg ggc       4611
Gly  Ala Val Ala Arg Cys  Glu Gln Gly Ala Ile  Gly Val Leu Gly
1025                1030                1035

cag  gga ctg agc acc cac  ccc aga ccc tgc ctg  tgg ctg ttc agc       4656
Gln  Gly Leu Ser Thr His  Pro Arg Pro Cys Leu  Trp Leu Phe Ser
1040                1045                1050

acc  cag ccc acc aag gcc  ttc acc gcc tgg ctg  gaa gtg ctg acc       4701
Thr  Gln Pro Thr Lys Ala  Phe Thr Ala Trp Leu  Glu Val Leu Thr
1055                1060                1065

ctg  ctg atc acc aag ctg  cgg gcc agc gcc gtg  aga acc ttc ggc       4746
Leu  Leu Ile Thr Lys Leu  Arg Ala Ser Ala Val  Arg Thr Phe Gly
1070                1075                1080

aaa  gag gtg gac atc ctg  ctg ctg cca gcc tgc  ttt aga gag gac       4791
Lys  Glu Val Asp Ile Leu  Leu Leu Pro Ala Cys  Phe Arg Glu Asp
1085                1090                1095

ctg  ccc ctg cct gag gga  atc ctg ctg gcc ctg  aga ggc ttt gcc       4836
Leu  Pro Leu Pro Glu Gly  Ile Leu Leu Ala Leu  Arg Gly Phe Ala
```

```
1100                1105                1110

ggc aag atc cgg tcc agc gac acc ccc agc atc ttc gat atc gcc  4881
Gly Lys Ile Arg Ser Ser Asp Thr Pro Ser Ile Phe Asp Ile Ala
1115                1120                1125

aga ccc ctg cac gtg tcc ctg aaa gtg cgc gtg acc gat cac cct  4926
Arg Pro Leu His Val Ser Leu Lys Val Arg Val Thr Asp His Pro
1130                1135                1140

gtg cct ggc ccc acc gtg ttt acc gac gcc agc agc agc acc cac  4971
Val Pro Gly Pro Thr Val Phe Thr Asp Ala Ser Ser Ser Thr His
1145                1150                1155

aag ggc gtg gtg gtg tgg aga gaa ggc ccc aga tgg gag atc aaa  5016
Lys Gly Val Val Val Trp Arg Glu Gly Pro Arg Trp Glu Ile Lys
1160                1165                1170

gaa atc gtg gac ctg ggc gcc tct gtg cag cag ctg gaa gcc aga  5061
Glu Ile Val Asp Leu Gly Ala Ser Val Gln Gln Leu Glu Ala Arg
1175                1180                1185

gcc gtg gcc atg gct ctg ctg ctg tgg ccc acc acc ccc acc aac  5106
Ala Val Ala Met Ala Leu Leu Leu Trp Pro Thr Thr Pro Thr Asn
1190                1195                1200

gtg gtg acc gac agc gcc ttc gtg gcc aag atg ctg ctg aag atg  5151
Val Val Thr Asp Ser Ala Phe Val Ala Lys Met Leu Leu Lys Met
1205                1210                1215

ggc cag gaa ggc gtc cct agc acc gcc gct gcc ttc atc ctg gaa  5196
Gly Gln Glu Gly Val Pro Ser Thr Ala Ala Ala Phe Ile Leu Glu
1220                1225                1230

gat gcc ctg tcc cag cgg agc gct atg gct gct gtg ctg cac gtg  5241
Asp Ala Leu Ser Gln Arg Ser Ala Met Ala Ala Val Leu His Val
1235                1240                1245

cgg agc cac agc gag gtg ccc ggc ttc ttc acc gag ggc aac gac  5286
Arg Ser His Ser Glu Val Pro Gly Phe Phe Thr Glu Gly Asn Asp
1250                1255                1260

gtg gcc gac agc cag gcc acc ttc cag gcc tac ccc ctg aga gag  5331
Val Ala Asp Ser Gln Ala Thr Phe Gln Ala Tyr Pro Leu Arg Glu
1265                1270                1275

gcc aag gac ctg cac aca gcc ctg cac atc ggc ccc aga gcc ctg  5376
Ala Lys Asp Leu His Thr Ala Leu His Ile Gly Pro Arg Ala Leu
1280                1285                1290

agc aag gcc tgc aac atc agc atg cag cag gcc aga gaa gtg gtg  5421
Ser Lys Ala Cys Asn Ile Ser Met Gln Gln Ala Arg Glu Val Val
1295                1300                1305

cag acc tgt ccc cac tgt aat agc gcc cct gct ctg gaa gcc ggc  5466
Gln Thr Cys Pro His Cys Asn Ser Ala Pro Ala Leu Glu Ala Gly
1310                1315                1320

gtg aac cct aga ggc ctc ggc cct ctg cag atc tgg cag acc gac  5511
Val Asn Pro Arg Gly Leu Gly Pro Leu Gln Ile Trp Gln Thr Asp
1325                1330                1335

ttc acc ctg gaa ccc cgg atg gcc cct aga agc tgg ctg gcc gtg  5556
Phe Thr Leu Glu Pro Arg Met Ala Pro Arg Ser Trp Leu Ala Val
1340                1345                1350

acc gtg gat acc gcc agc tcc gcc atc gtg gtg acc cag cac ggc  5601
Thr Val Asp Thr Ala Ser Ser Ala Ile Val Val Thr Gln His Gly
1355                1360                1365
```

```
aga  gtg aca tct gtg gcc  gcc cag cac cac tgg  gcc aca gcc att        5646
Arg  Val Thr Ser Val Ala  Ala Gln His His Trp  Ala Thr Ala Ile
1370                 1375                 1380

gcc  gtg ctg ggc aga ccc  aag gcc atc aag acc  gac aac ggc agc        5691
Ala  Val Leu Gly Arg Pro  Lys Ala Ile Lys Thr  Asp Asn Gly Ser
1385                 1390                 1395

tgc  ttc acc agc cgg tcc  acc aga gaa tgg ctg  gcc aga tgg ggc        5736
Cys  Phe Thr Ser Arg Ser  Thr Arg Glu Trp Leu  Ala Arg Trp Gly
1400                 1405                 1410

att  gct cac acc acc ggc  atc cca ggc aat agt  cag ggc cag gcc        5781
Ile  Ala His Thr Thr Gly  Ile Pro Gly Asn Ser  Gln Gly Gln Ala
1415                 1420                 1425

atg  gtg gag aga gcc aac  agg ctg ctg aag gac  aag atc cgg gtg        5826
Met  Val Glu Arg Ala Asn  Arg Leu Leu Lys Asp  Lys Ile Arg Val
1430                 1435                 1440

ctg  gcc gag ggc gac ggc  ttc atg aag cgg atc  ccc gcc agc aaa        5871
Leu  Ala Glu Gly Asp Gly  Phe Met Lys Arg Ile  Pro Ala Ser Lys
1445                 1450                 1455

cag  ggc gag ctg ctg gcc  aag gcc atg tac gcc  ctg aac cac ttc        5916
Gln  Gly Glu Leu Leu Ala  Lys Ala Met Tyr Ala  Leu Asn His Phe
1460                 1465                 1470

gag  cgg ggc gag aac acc  aag acc ccc gtg cag  aag cac tgg cgg        5961
Glu  Arg Gly Glu Asn Thr  Lys Thr Pro Val Gln  Lys His Trp Arg
1475                 1480                 1485

cct  acc gtg ctg aca gag  gga ccc ccc gtg aag  atc cgg atc gag        6006
Pro  Thr Val Leu Thr Glu  Gly Pro Pro Val Lys  Ile Arg Ile Glu
1490                 1495                 1500

aca  ggc gag tgg gag aag  ggg tgg aac gtg ctg  gtg tgg ggc aga        6051
Thr  Gly Glu Trp Glu Lys  Gly Trp Asn Val Leu  Val Trp Gly Arg
1505                 1510                 1515

ggc  tac gcc gct gtc aag  aac cgg gac acc gac  aaa gtg atc tgg        6096
Gly  Tyr Ala Ala Val Lys  Asn Arg Asp Thr Asp  Lys Val Ile Trp
1520                 1525                 1530

gtg  ccc agc cgg aaa gtg  aag ccc gac gtg acc  cag aaa gac gaa        6141
Val  Pro Ser Arg Lys Val  Lys Pro Asp Val Thr  Gln Lys Asp Glu
1535                 1540                 1545

gtg  acc aag aag gac gag  gcc tcc ccc ctc ttt  gcc ggc atc agc        6186
Val  Thr Lys Lys Asp Glu  Ala Ser Pro Leu Phe  Ala Gly Ile Ser
1550                 1555                 1560

gac  tgg atc cct tgg gag  gac gag cag gaa ggc  ctg cag ggc gag        6231
Asp  Trp Ile Pro Trp Glu  Asp Glu Gln Glu Gly  Leu Gln Gly Glu
1565                 1570                 1575

aca  gcc agc aac aag cag  gaa aga ccc ggc gag  gac acc ctg gcc        6276
Thr  Ala Ser Asn Lys Gln  Glu Arg Pro Gly Glu  Asp Thr Leu Ala
1580                 1585                 1590

gcc  aat gag agc tga gaattctgca gatatccagc acagtggcgg ccgctcgagt      6331
Ala  Asn Glu Ser
1595

ctagagggcc cgtttaaacc cgctgatcag cctcgactgt gccttctagt tgccagccat      6391

ctgttgtttg cccctccccc gtgccttcct tgaccctgga aggtgccact cccactgtcc      6451
```

tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat tctattctgg 6511
ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc aggcatgctg 6571
gggatgcggt gggctctatg gcttctgagg cggaaagaac cagctggggc tctaggggtg 6631
atccccacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg 6691
tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc 6751
tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctccct ttagggttcc 6811
gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat ggttcacgta 6871
gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta 6931
atagtggact cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg 6991
atttataagg gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa 7051
aatttaacgc gaattaattc tgtggaatgt gtgtcagtta gggtgtggaa agtccccagg 7111
ctccccagca ggcagaagta tgcaaagcat gcatctcaat tagtcagcaa ccaggtgtgg 7171
aaagtcccca ggctccccag caggcagaag tatgcaaagc atgcatctca attagtcagc 7231
aaccatagtc ccgcccctaa ctccgcccat cccgccccta actccgccca gttccgccca 7291
ttctccgccc catggctgac taatttttt tatttatgca gaggccgagg ccgcctctgc 7351
ctctgagcta ttccagaagt agtgaggagg ctttttttgga ggcctaggct tttgcaaaaa 7411
gctcccggga gcttgtatat ccattttcgg atctgatcaa gagacaggat gaggatcgtt 7471
tcgcatgatt gaacaagatg gattgcacgc aggttctccg gccgcttggg tggagaggct 7531
attcggctat gactgggcac aacagacaat cggctgctct gatgccgccg tgttccggct 7591
gtcagcgcag gggcgcccgg ttctttttgt caagaccgac ctgtccggtg ccctgaatga 7651
actgcaggac gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc 7711
tgtgctcgac gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg 7771
gcaggatctc ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc 7831
aatgcggcgg ctgcatacgc ttgatccggc tacctgccca ttcgaccacc aagcgaaaca 7891
tcgcatcgag cgagcacgta ctcggatgga agccggtctt gtcgatcagg atgatctgga 7951
cgaagagcat caggggctcg cgccagccga actgttcgcc aggctcaagg cgcgcatgcc 8011
cgacggcgag gatctcgtcg tgacccatgg cgatgcctgc ttgccgaata tcatggtgga 8071
aaatggccgc ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca 8131
ggacatagcg ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg 8191
cttcctcgtg ctttacggta tcgccgctcc cgattcgcag cgcatcgcct tctatcgcct 8251
tcttgacgag ttcttctgag cgggactctg gggttcgaaa tgaccgacca agcgacgccc 8311
aacctgccat cacgagattt cgattccacc gccgccttct atgaaaggtt gggcttcgga 8371
atcgttttcc gggacgccgg ctggatgatc ctccagcgcg gggatctcat gctggagttc 8431
ttcgcccacc ccaacttgtt tattgcagct tataatggtt acaaataaag caatagcatc 8491
acaaatttca caaataaagc attttttttca ctgcattcta gttgtggttt gtccaaactc 8551

atcaatgtat cttatcatgt ctgtataccg tcgacctcta gctagagctt ggcgtaatca 8611
tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga 8671
gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact cacattaatt 8731
gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga 8791
atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc 8851
actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg 8911
gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc 8971
cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc 9031
cccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga 9091
ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc 9151
ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat 9211
agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg 9271
cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc 9331
aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga 9391
gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact 9451
agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt 9511
ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtttttttgt ttgcaagcag 9571
cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct 9631
gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg 9691
atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat 9751
gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc 9811
tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac tacgatacgg 9871
gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg ctcaccggct 9931
ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggtcctgca 9991
actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg 10051
ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt gtcacgctcg 10111
tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt tacatgatcc 10171
cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt cagaagtaag 10231
ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg 10291
ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt ctgagaatag 10351
tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac cgcgccacat 10411
agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa actctcaagg 10471
atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa ctgatcttca 10531
gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgccgca 10591

aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct ttttcaatat 10651

tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 10711

aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtc 10768

<210> 15
<211> 703
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción sintética

<400> 15

Met Glu Ala Val Ile Lys Val Ile Ser Ser Ala Cys Lys Thr Tyr Cys
1 5 10 15

Gly Lys Ile Ser Pro Ser Lys Lys Glu Ile Gly Ala Met Leu Ser Leu
20 25 30

Leu Gln Lys Glu Gly Leu Leu Met Ser Pro Ser Asp Leu Tyr Ser Pro
35 40 45

Gly Ser Trp Asp Pro Ile Thr Ala Ala Leu Ser Gln Arg Ala Met Val
50 55 60

Leu Gly Lys Ser Gly Glu Leu Lys Thr Trp Gly Leu Val Leu Gly Ala
65 70 75 80

Leu Lys Ala Ala Arg Glu Glu Gln Val Thr Ser Glu Gln Ala Lys Phe
85 90 95

Trp Leu Gly Leu Gly Gly Gly Arg Val Ser Pro Pro Gly Pro Glu Cys
100 105 110

Ile Glu Lys Pro Ala Thr Glu Arg Arg Ile Asp Lys Gly Glu Glu Val
115 120 125

Gly Glu Thr Thr Val Gln Arg Asp Ala Lys Met Ala Pro Glu Glu Ala
130 135 140

Ala Thr Pro Lys Thr Val Gly Thr Ser Cys Tyr His Cys Gly Thr Ala
145 150 155 160

Val Gly Cys Asn Cys Ala Thr Ala Thr Ala Ser Ala Pro Pro Pro Pro
165 170 175

Tyr Val Gly Ser Gly Leu Tyr Pro Ser Leu Ala Gly Val Gly Glu Gln
180 185 190

Gln Gly Gln Gly Asp Asn Thr Ser Arg Gly Ala Glu Gln Pro Arg Glu
195 200 205

Glu Pro Gly His Ala Gly Gln Ala Pro Gly Pro Ala Leu Thr Asp Trp
210 215 220

Ala Arg Val Arg Glu Glu Leu Ala Ser Thr Gly Pro Pro Val Val Ala
225 230 235 240

Met Pro Val Val Ile Lys Thr Glu Gly Pro Ala Trp Thr Pro Leu Glu
245 250 255

Pro Lys Leu Ile Thr Arg Leu Ala Asp Thr Val Arg Thr Lys Gly Leu
260 265 270

Arg Ser Pro Ile Thr Met Ala Glu Val Glu Ala Leu Met Ser Ser Pro
275 280 285

Leu Leu Pro His Asp Val Thr Asn Leu Met Arg Val Ile Leu Gly Pro
290 295 300

Ala Pro Tyr Ala Leu Trp Met Asp Ala Trp Gly Val Gln Leu Gln Thr
305 310 315 320

Val Ile Ala Ala Ala Thr Arg Asp Pro Arg His Pro Ala Asn Gly Gln
325 330 335

Gly Arg Gly Glu Arg Thr Asn Leu Asp Arg Leu Lys Gly Leu Ala Asp
340 345 350

Gly Met Val Gly Asn Pro Gln Gly Gln Ala Ala Leu Leu Arg Pro Gly
355 360 365

Glu Leu Val Ala Ile Thr Ala Ser Ala Leu Gln Ala Phe Arg Glu Val
370 375 380

Ala Arg Leu Ala Glu Pro Ala Gly Pro Trp Ala Asp Ile Thr Gln Gly
385 390 395 400

Pro Ser Glu Ser Phe Val Asp Phe Ala Asn Arg Leu Ile Lys Ala Val
405 410 415

Glu Gly Ser Asp Leu Pro Pro Ser Ala Arg Ala Pro Val Ile Ile Asp
420 425 430

Cys Phe Arg Gln Lys Ser Gln Pro Asp Ile Gln Gln Leu Ile Arg Ala
435 440 445

Ala Pro Ser Thr Leu Thr Thr Pro Gly Glu Ile Ile Lys Tyr Val Leu
450 455 460

Asp Arg Gln Lys Ile Ala Pro Leu Thr Asp Gln Gly Ile Ala Ala Ala
465 470 475 480

Met Ser Ser Ala Ile Gln Pro Leu Val Met Ala Val Val Asn Arg Glu

485 490 495

Arg Asp Gly Gln Thr Gly Ser Gly Gly Arg Ala Arg Gly Leu Cys Tyr
500 505 510

Thr Cys Gly Ser Pro Gly His Tyr Gln Ala Gln Cys Pro Lys Lys Arg
515 520 525

Lys Ser Gly Asn Ser Arg Glu Arg Cys Gln Leu Cys Asp Gly Met Gly
530 535 540

His Asn Ala Lys Gln Cys Arg Lys Arg Asp Gly Asn Gln Gly Gln Arg
545 550 555 560

Pro Gly Arg Gly Leu Ser Ser Gly Pro Trp Pro Gly Pro Glu Gln Pro
565 570 575

Ala Val Ser Leu Ala Met Thr Met Glu His Lys Asp Arg Pro Leu Val
580 585 590

Arg Val Ile Leu Thr Asn Thr Gly Ser His Pro Val Lys Gln Arg Ser
595 600 605

Val Tyr Ile Thr Ala Leu Leu Asp Ser Gly Ala Asp Ile Thr Ile Ile
610 615 620

Ser Glu Glu Asp Trp Pro Thr Asp Trp Pro Val Val Asp Thr Ala Asn
625 630 635 640

Pro Gln Ile His Gly Ile Gly Gly Gly Ile Pro Met Arg Lys Ser Arg
645 650 655

Asp Met Ile Glu Val Gly Val Ile Asn Arg Asp Gly Ser Leu Glu Arg
660 665 670

Pro Leu Leu Leu Phe Pro Ala Val Ala Met Val Arg Gly Ser Ile Leu
675 680 685

Gly Arg Asp Cys Leu Gln Gly Leu Gly Leu Arg Leu Thr Asn Leu
690 695 700

<210> 16
<211> 895
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción sintética

<400> 16

Thr Val Ala Leu His Leu Ala Ile Pro Leu Lys Trp Lys Ser Asp Arg
1 5 10 15

```
Thr Pro Val Trp Ile Asp Gln Trp Pro Leu Pro Glu Gly Lys Leu Val
            20                  25                  30

Ala Leu Thr Gln Leu Val Glu Lys Glu Leu Gln Leu Gly His Ile Glu
        35                  40                  45

Pro Ser Leu Ser Cys Trp Asn Thr Pro Val Phe Val Ile Arg Lys Ala
    50                  55                  60

Ser Gly Ser Tyr Arg Leu Leu His Asp Leu Arg Ala Val Asn Ala Lys
65                  70                  75                  80

Leu Val Pro Phe Gly Ala Val Gln Gln Gly Ala Pro Val Leu Ser Ala
                85                  90                  95

Leu Pro Arg Gly Trp Pro Leu Met Val Leu Asp Leu Lys Asp Cys Phe
            100                 105                 110

Phe Ser Ile Pro Leu Ala Glu Gln Asp Arg Glu Ala Phe Ala Phe Thr
        115                 120                 125

Leu Pro Ser Val Asn Asn Gln Ala Pro Ala Arg Arg Phe Gln Trp Lys
    130                 135                 140

Val Leu Pro Gln Gly Met Thr Cys Ser Pro Thr Ile Cys Gln Leu Val
145                 150                 155                 160

Val Gly Gln Val Leu Glu Pro Leu Arg Leu Lys His Pro Ala Leu Arg
                165                 170                 175

Met Leu His Tyr Met Asp Asp Leu Leu Leu Ala Ala Ser Ser His Asp
            180                 185                 190

Gly Leu Glu Ala Ala Gly Lys Glu Val Ile Gly Thr Leu Glu Arg Ala
        195                 200                 205

Gly Phe Thr Ile Ser Pro Asp Lys Ile Gln Arg Glu Pro Gly Val Gln
    210                 215                 220

Tyr Leu Gly Tyr Lys Leu Gly Ser Thr Tyr Val Ala Pro Val Gly Leu
225                 230                 235                 240

Val Ala Glu Pro Arg Ile Ala Thr Leu Trp Asp Val Gln Lys Leu Val
                245                 250                 255

Gly Ser Leu Gln Trp Leu Arg Pro Ala Leu Gly Ile Pro Pro Arg Leu
            260                 265                 270

Met Gly Pro Phe Tyr Glu Gln Leu Arg Gly Ser Asp Pro Asn Glu Ala
        275                 280                 285

Arg Glu Trp Asn Leu Asp Met Lys Met Ala Trp Arg Glu Ile Val Gln
```

290 295 300

Leu Ser Thr Thr Ala Ala Leu Glu Arg Trp Asp Pro Ala Gln Pro Leu
305 310 315 320

Glu Gly Ala Val Ala Arg Cys Glu Gln Gly Ala Ile Gly Val Leu Gly
325 330 335

Gln Gly Leu Ser Thr His Pro Arg Pro Cys Leu Trp Leu Phe Ser Thr
340 345 350

Gln Pro Thr Lys Ala Phe Thr Ala Trp Leu Glu Val Leu Thr Leu Leu
355 360 365

Ile Thr Lys Leu Arg Ala Ser Ala Val Arg Thr Phe Gly Lys Glu Val
370 375 380

Asp Ile Leu Leu Leu Pro Ala Cys Phe Arg Glu Asp Leu Pro Leu Pro
385 390 395 400

Glu Gly Ile Leu Leu Ala Leu Arg Gly Phe Ala Gly Lys Ile Arg Ser
405 410 415

Ser Asp Thr Pro Ser Ile Phe Asp Ile Ala Arg Pro Leu His Val Ser
420 425 430

Leu Lys Val Arg Val Thr Asp His Pro Val Pro Gly Pro Thr Val Phe
435 440 445

Thr Asp Ala Ser Ser Ser Thr His Lys Gly Val Val Val Trp Arg Glu
450 455 460

Gly Pro Arg Trp Glu Ile Lys Glu Ile Val Asp Leu Gly Ala Ser Val
465 470 475 480

Gln Gln Leu Glu Ala Arg Ala Val Ala Met Ala Leu Leu Leu Trp Pro
485 490 495

Thr Thr Pro Thr Asn Val Val Thr Asp Ser Ala Phe Val Ala Lys Met
500 505 510

Leu Leu Lys Met Gly Gln Glu Gly Val Pro Ser Thr Ala Ala Ala Phe
515 520 525

Ile Leu Glu Asp Ala Leu Ser Gln Arg Ser Ala Met Ala Ala Val Leu
530 535 540

His Val Arg Ser His Ser Glu Val Pro Gly Phe Phe Thr Glu Gly Asn
545 550 555 560

Asp Val Ala Asp Ser Gln Ala Thr Phe Gln Ala Tyr Pro Leu Arg Glu
565 570 575

Ala Lys Asp Leu His Thr Ala Leu His Ile Gly Pro Arg Ala Leu Ser
580 585 590

Lys Ala Cys Asn Ile Ser Met Gln Gln Ala Arg Glu Val Val Gln Thr
595 600 605

Cys Pro His Cys Asn Ser Ala Pro Ala Leu Glu Ala Gly Val Asn Pro
610 615 620

Arg Gly Leu Gly Pro Leu Gln Ile Trp Gln Thr Asp Phe Thr Leu Glu
625 630 635 640

Pro Arg Met Ala Pro Arg Ser Trp Leu Ala Val Thr Val Asp Thr Ala
645 650 655

Ser Ser Ala Ile Val Val Thr Gln His Gly Arg Val Thr Ser Val Ala
660 665 670

Ala Gln His His Trp Ala Thr Ala Ile Ala Val Leu Gly Arg Pro Lys
675 680 685

Ala Ile Lys Thr Asp Asn Gly Ser Cys Phe Thr Ser Arg Ser Thr Arg
690 695 700

Glu Trp Leu Ala Arg Trp Gly Ile Ala His Thr Thr Gly Ile Pro Gly
705 710 715 720

Asn Ser Gln Gly Gln Ala Met Val Glu Arg Ala Asn Arg Leu Leu Lys
725 730 735

Asp Lys Ile Arg Val Leu Ala Glu Gly Asp Gly Phe Met Lys Arg Ile
740 745 750

Pro Ala Ser Lys Gln Gly Glu Leu Leu Ala Lys Ala Met Tyr Ala Leu
755 760 765

Asn His Phe Glu Arg Gly Glu Asn Thr Lys Thr Pro Val Gln Lys His
770 775 780

Trp Arg Pro Thr Val Leu Thr Glu Gly Pro Pro Val Lys Ile Arg Ile
785 790 795 800

Glu Thr Gly Glu Trp Glu Lys Gly Trp Asn Val Leu Val Trp Gly Arg
805 810 815

Gly Tyr Ala Ala Val Lys Asn Arg Asp Thr Asp Lys Val Ile Trp Val
820 825 830

Pro Ser Arg Lys Val Lys Pro Asp Val Thr Gln Lys Asp Glu Val Thr
835 840 845

```
Lys Lys Asp Glu Ala Ser Pro Leu Phe Ala Gly Ile Ser Asp Trp Ile
    850                 855                 860

Pro Trp Glu Asp Glu Gln Glu Gly Leu Gln Gly Glu Thr Ala Ser Asn
865                 870                 875                 880

Lys Gln Glu Arg Pro Gly Glu Asp Thr Leu Ala Ala Asn Glu Ser
                885                 890                 895
```

REIVINDICACIONES

1. ARN retroviral con una zona 5' que presenta una región R y una región U5 del Virus de la Leucosis/Sarcoma Aviar (ASLV) y dispuesto en 3' con respecto a esto un sitio de unión de cebador (PBS), una señal de empaquetamiento (ψ), un casete de expresión, una región U3 3' y una región R, **caracterizado por que** la región U3 3' presenta una secuencia de nucleótidos correspondiente a la SEC ID Nº: 1, de la cual están delecionados al menos los nucleótidos Nº 28 hasta el Nº 186 inclusive, por lo que se elimina la actividad promotora y potenciadora de la región U3 3' y por que no están contenidas secuencias de ácido nucleico que codifiquen al menos en parte proteínas estructurales virales o un sitio donador de empalme (SD).

2. ARN retroviral de acuerdo con la reivindicación 1, **caracterizado por que** en 3' para la deleción está dispuesta una secuencia de poliadenilación.

3. ARN retroviral de acuerdo con una de las reivindicaciones anteriores, **caracterizado por que** la región U3 3' presenta una secuencia de nucleótidos correspondiente a la SEC ID Nº: 1 y adicionalmente los nucleótidos Nº 200 a Nº 206 están mutados en una subsecuencia que no es una caja TATA.

4. ARN retroviral de acuerdo con la reivindicación 3, **caracterizado por que** los nucleótidos Nº 200 a Nº 206 están mutados en una secuencia TGTCTAA.

5. ARN retroviral de acuerdo con una de las reivindicaciones anteriores, **caracterizado por que** en la región U3 3' adicionalmente el nucleótido Nº 187 está mutado en C y el nucleótido Nº 189, en T.

6. ARN retroviral de acuerdo con una de las reivindicaciones anteriores, **caracterizado por que** la región U3 3' presenta una secuencia de nucleótidos correspondiente a la SEC ID Nº: 1, de la cual están delecionados adicionalmente los nucleótidos Nº 187 a Nº 202.

7. ARN retroviral de acuerdo con una de las reivindicaciones anteriores, **caracterizado por que** la región U3 3' presenta una secuencia correspondiente a las SEC ID Nº: 2, SEC ID Nº: 3 o SEC ID Nº: 4 o está compuesta de las mismas.

8. ARN retroviral de acuerdo con una de las reivindicaciones anteriores, **caracterizado por que** la región U3 3' inclusive la región R 3' presenta una secuencia correspondiente a los nucleótidos Nº 3282 a Nº 3372 de la SEC ID Nº: 5, una secuencia correspondiente a los nucleótidos Nº 3282 a Nº 3356 de la SEC ID Nº: 7 o una secuencia correspondiente a los nucleótidos Nº 3282 a Nº 3372 de la SEC ID Nº: 9 o está compuesta de las mismas.

9. ARN retroviral de acuerdo con una de las reivindicaciones anteriores, **caracterizado por que** la deleción de la región U3 3' está sustituida por una inserción con una secuencia de nucleótidos que no presenta ninguna secuencia promotora o potenciadora.

10. ARN retroviral de acuerdo con una de las reivindicaciones anteriores, **caracterizado por que** la región U3 3' presenta una secuencia de ácido nucleico insertada que codifica miARN, ARNhc, un potenciador de la poliadenilación, aislador y/o secuencia de reconocimiento para una ADN-recombinasa.

11. ARN retroviral de acuerdo con una de las reivindicaciones anteriores, **caracterizado por que** el casete de expresión presenta de 5' a 3' un promotor, una secuencia de ácido nucleico que codifica un transgén y un elemento regulador post-transcripcional (PRE).

12. ARN retroviral de acuerdo con una de las reivindicaciones anteriores, **caracterizado por que** la secuencia de ácido nucleico que está dispuesta en 5' limitando con el promotor codifica de 5' a 3' un elemento R, limitando directamente con esto un elemento U5 y una señal de empaquetamiento.

13. ARN retroviral de acuerdo con una de las reivindicaciones 1 a 11, **caracterizado por que** la secuencia de ácido nucleico que está dispuesta en 5' con respecto al casete de expresión presenta la secuencia de los nucleótidos Nº 922 a 1292 de la SEC ID Nº: 5 o está compuesta de la misma.

14. ARN retroviral de acuerdo con una de las reivindicaciones anteriores, **caracterizado por que** la región U3 3' presenta un sitio de ataque de integrasa en los nucleótidos 1-25 así como una señal de poliA.

15. ADN con un elemento promotor que está dispuesto funcionalmente en 5' con respecto a una secuencia de nucleótidos que codifica un ARN retroviral de acuerdo con una de las reivindicaciones anteriores.

16. ADN de acuerdo con la reivindicación 15, **caracterizado por que** la secuencia de nucleótidos que codifica ARN retroviral presenta una región R 5' y una región U5 5' con una secuencia correspondiente a los nucleótidos Nº 922 a 1022 de la SEC ID Nº: 5 y en 3' con respecto a esto una LTR 3' con una secuencia correspondiente a los nucleótidos Nº 3282 a 3452 de la SEC ID Nº: 5 o con una secuencia correspondiente a los nucleótidos Nº 3282 a 3436 de la

SEC ID Nº: 7 o con una secuencia correspondiente a los nucleótidos Nº 3282 a 3452 de la SEC ID Nº: 9.

17. Partícula viral con una proteína de envoltura, **caracterizada por que** la partícula presenta un ARN retroviral de acuerdo con una de las reivindicaciones 1 a 14.

18. Composición farmacéutica para su uso en la terapia génica, **caracterizada por** una partícula viral de acuerdo con la reivindicación 17.

19. Sistema de separación-empaquetamiento para la generación de partículas retrovirales que comprende un ADN que codifica un ARN retroviral, un primer plásmido auxiliar que codifica una proteína de fusión gag-pro-pol y un segundo plásmido auxiliar que codifica una proteína de envoltura viral, **caracterizado por que** el ARN retroviral presenta una secuencia de nucleótidos de acuerdo con una de las reivindicaciones 1 a 14.

20. Sistema de acuerdo con la reivindicación 19, **caracterizado por que** el ADN que codifica el ARN retroviral, el primer plásmido auxiliar y/o el segundo plásmido auxiliar están introducidos por transfección en una célula eucariota o están integrados en el genoma de una célula eucariota.

21. Sistema de acuerdo con una de las reivindicaciones 19 a 20, **caracterizado por que** el primer plásmido auxiliar codifica la proteína de fusión gag-pro-pol y/o el segundo plásmido auxiliar la proteína de envoltura viral con codones que están adaptados al uso de codones de la célula eucariota que los contiene.

22. Sistema de acuerdo con una de las reivindicaciones 19 a 21, **caracterizado por que** el primer plásmido auxiliar codifica la proteína de fusión gag-pro-pol con una secuencia de nucleótidos correspondiente a los nucleótidos Nº 1475 a Nº 6291 de la SEC ID Nº: 14.

23. Procedimiento para la producción de partículas virales mediante expresión de una proteína de fusión gag-pro-pol y de una proteína de envoltura viral en una célula de empaquetamiento, **caracterizado por que** el ARN retroviral en la célula de empaquetamiento se transcribe de un ADN de acuerdo con una de las reivindicaciones 15 o 16.

24. Partícula viral de acuerdo con la reivindicación 17 para el uso en la terapia génica.

25. Partícula viral de acuerdo con la reivindicación 24 para el uso en la terapia génica para la modificación genética de células adultas somáticas, células madre hematopoyéticas o células madre embrionarias no humanas.

26. Transducción de células diana mediante puesta en contacto de las células *in vitro* con un ARN retroviral de acuerdo con una de las reivindicaciones 1 a 14.

Fig. 1

Fig. 2

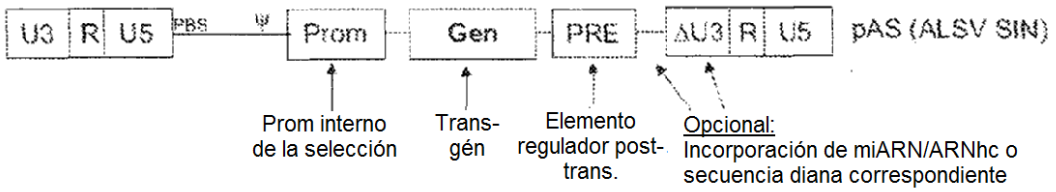


Fig. 3

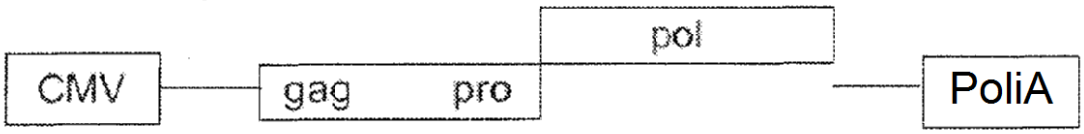


Fig. 4

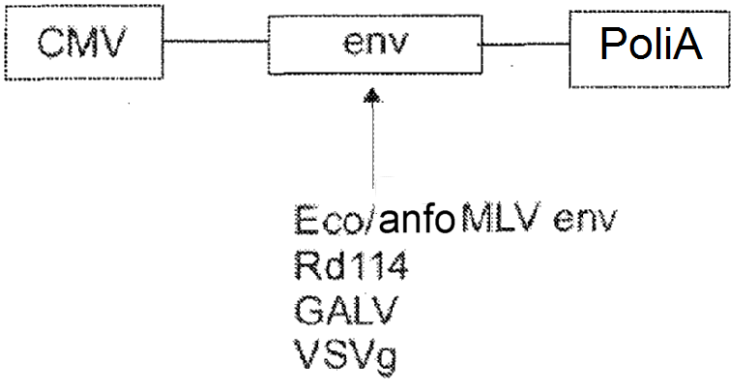


Fig. 5

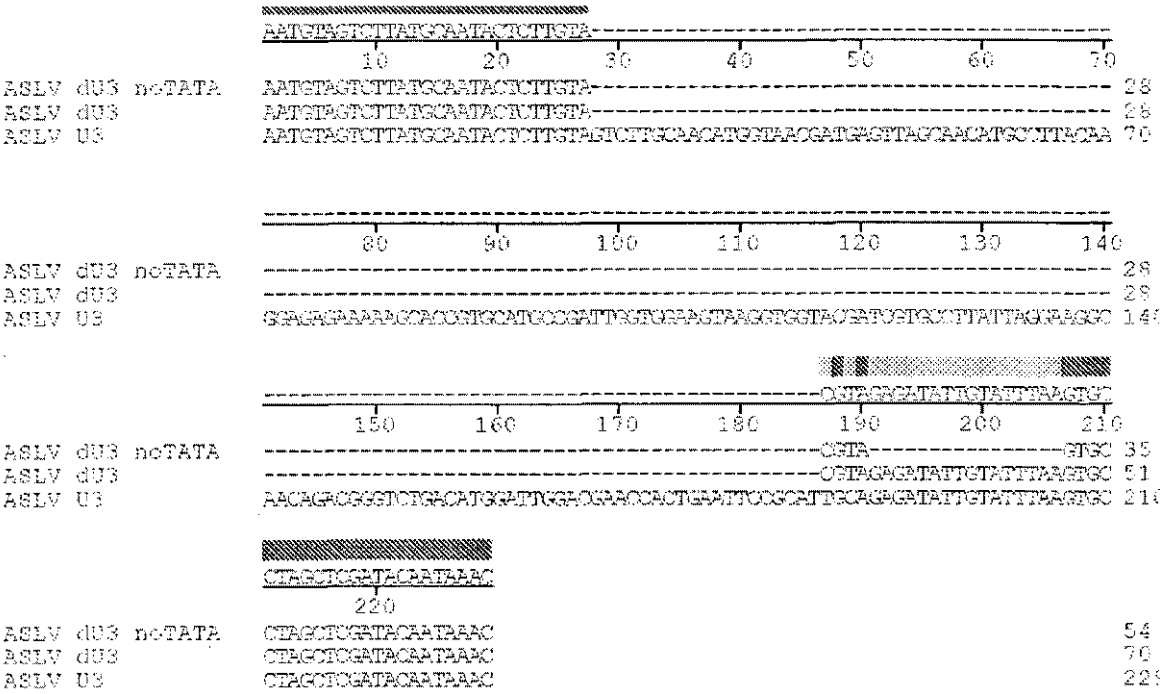


Fig. 6

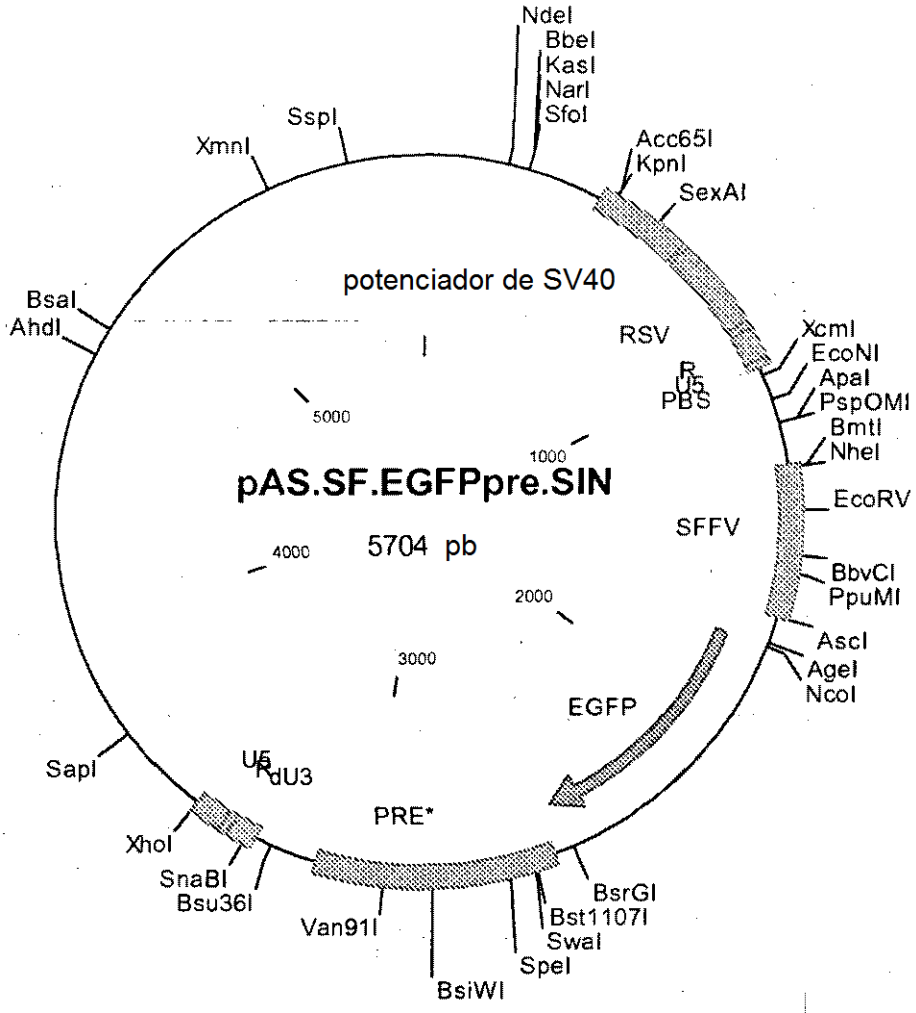

NdeI
BbeI
KasI
NarI
SfoI
Acc65I
KpnI
SexAI
XcmI
EcoNI
ApaI
PspOMI
BmtI
NheI
EcoRV
BbvCI
PpuMI
AscI
AgeI
NcoI
potenciador de SV40
RSV
R
U5
PBS
SFFV
EGFP
PRE*
U5
RdU3
SapI
BsaI
AhdI
XmnI
SspI
XhoI
SnaBI
Bsu36I
Van91I
BsiWI
SpeI
BsrGI
Bst1107I
SwaI
1000
2000
3000
4000
5000
pAS.SF.EGFPpre.SIN
5704 pb

Fig. 7

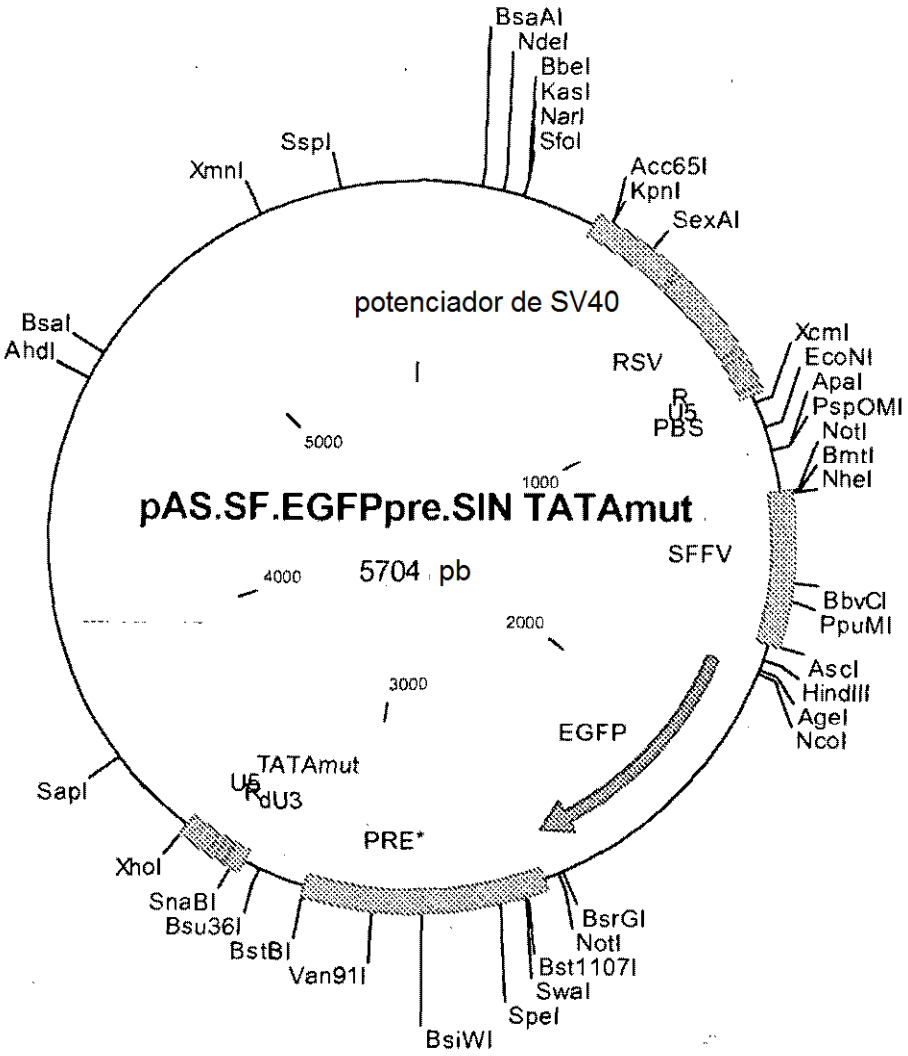

BsaAI
NdeI
BbeI
KasI
NarI
SfoI
Acc65I
KpnI
SexAI
XcmI
EcoNI
ApaI
PspOMI
NotI
BmtI
NheI
BbvCI
PpuMI
AscI
HindIII
AgeI
NcoI
SapI
XhoI
SnaBI
Bsu36I
BstBI
Van91I
BsiWI
SpeI
Bst1107I
SwaI
BsrGI
NotI
XmnI
SspI
BsaI
AhdI
potenciador de SV40
RSV
R
U5
PBS
SFFV
EGFP
TATAmut
U5
R
dU3
PRE*
1000
2000
3000
4000
5000
pAS.SF.EGFPpre.SIN TATAmut
5704 pb

Fig. 8

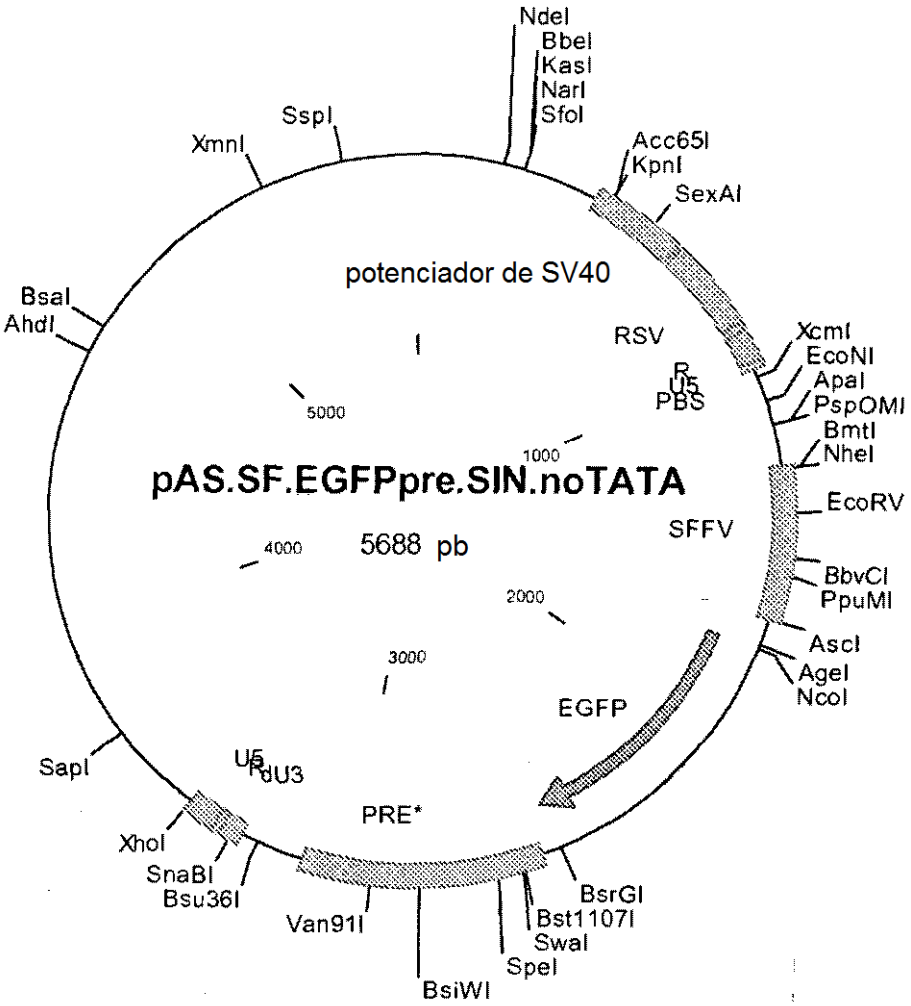

NdeI
BbeI
KasI
NarI
SfoI
Acc65I
KpnI
SexAI
potenciador de SV40
RSV
R
U5
PBS
XcmI
EcoNI
ApaI
PspOMI
BmtI
NheI
EcoRV
SFFV
BbvCI
PpuMI
AscI
AgeI
NcoI
EGFP
PRE*
BsrGI
Bst1107I
SwaI
SpeI
BsiWI
Van91I
SnaBI
Bsu36I
XhoI
SapI
U5
RdU3
BsaI
AhdI
XmnI
SspI
1000
2000
3000
4000
5000
pAS.SF.EGFPpre.SIN.noTATA
5688 pb

Fig. 9

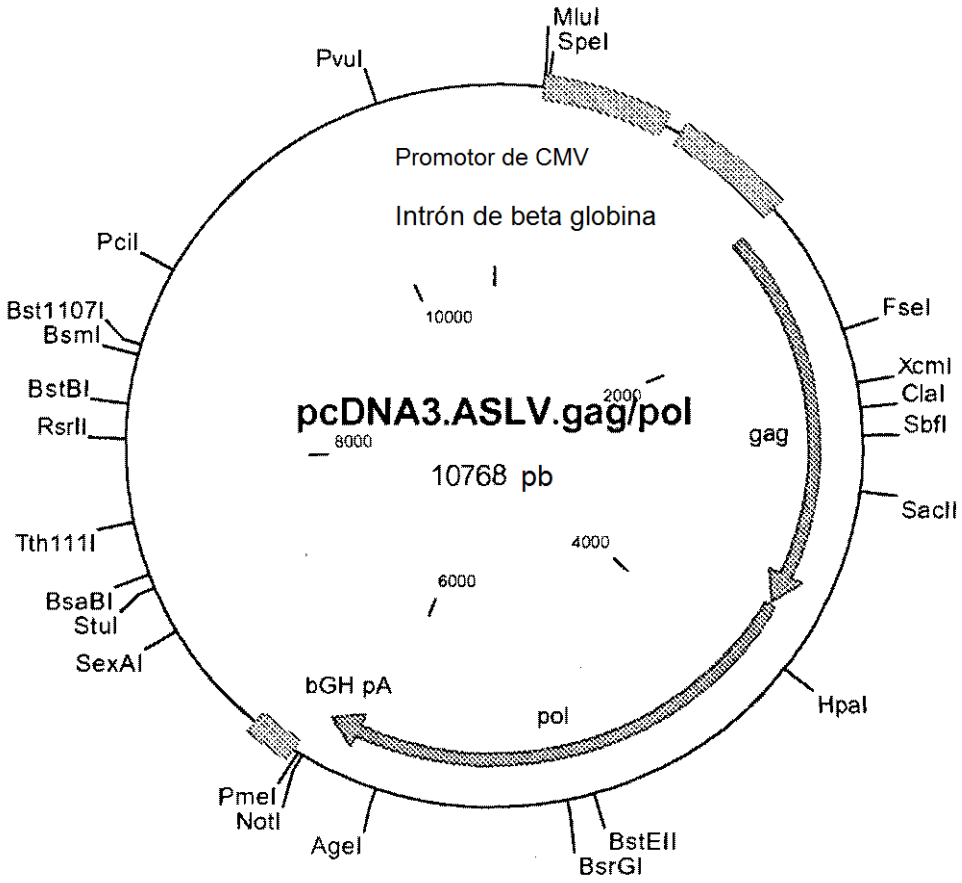

MluI
SpeI
PvuI
Promotor de CMV
Intrón de beta globina
PciI
10000
Bst1107I
BsmI
FseI
XcmI
ClaI
SbfI
BstBI
RsrII
2000
pcDNA3.ASLV.gag/pol
8000
gag
10768 pb
SacII
Tth111I
4000
6000
BsaBI
StuI
SexAI
bGH pA
pol
HpaI
PmeI
NotI
AgeI
BstEII
BsrGI

Fig. 10

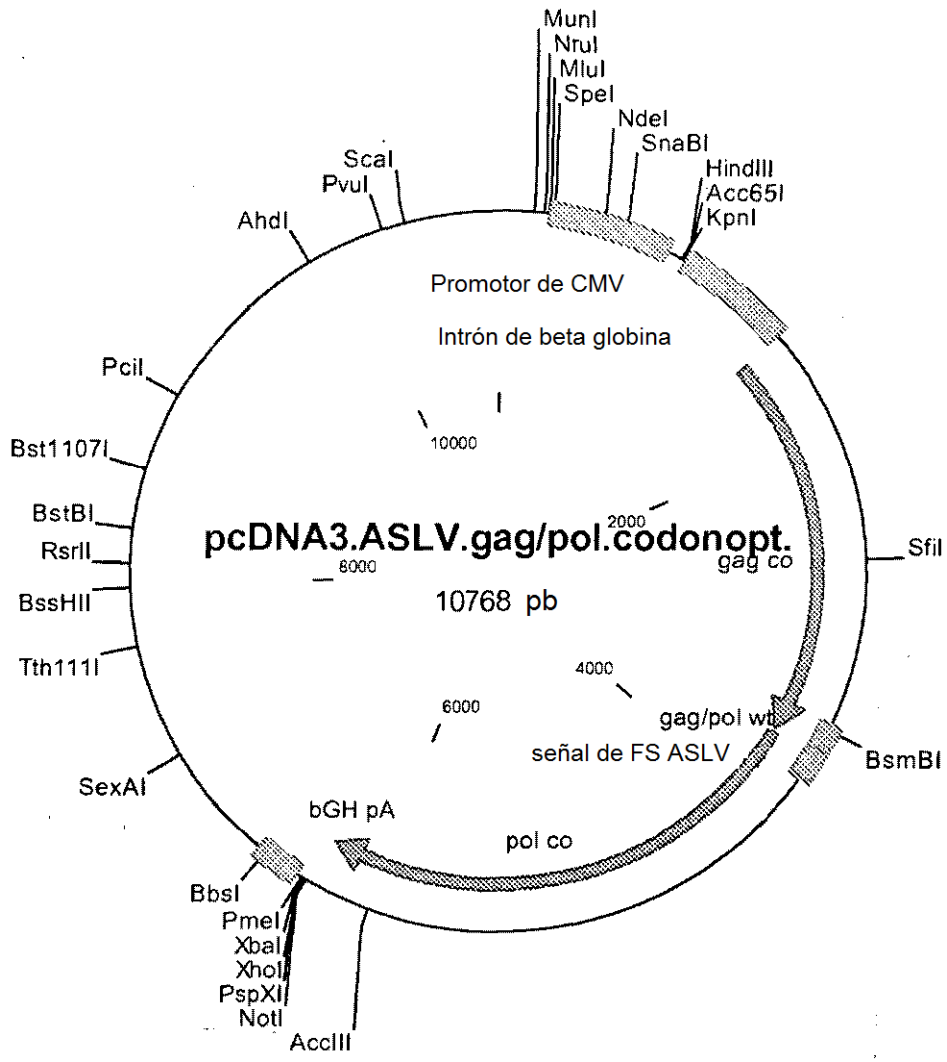

MunI
NruI
MluI
SpeI
NdeI
SnaBI
HindIII
Acc65I
KpnI
ScaI
PvuI
AhdI
Promotor de CMV
Intrón de beta globina
PciI
Bst1107I
BstBI
RsrII
BssHII
Tth111I
SexAI
BbsI
PmeI
XbaI
XhoI
PspXI
NotI
AccIII
bGH pA
pol co
señal de FS ASLV
gag/pol wt
BsmBI
SfiI
gag co
pcDNA3.ASLV.gag/pol.codonopt.
10768 pb
10000
2000
8000
4000
6000

Fig. 11

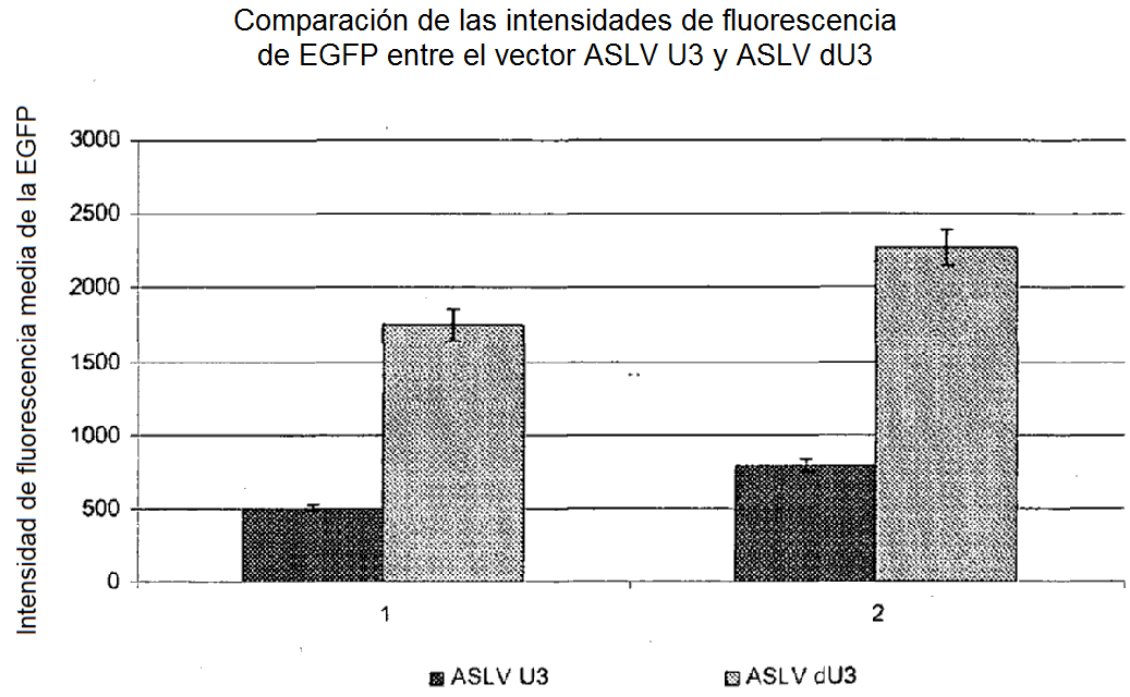

Comparación de las intensidades de fluorescencia de EGFP entre el vector ASLV U3 y ASLV dU3
Intensidad de fluorescencia media de la EGFP
3000
2500
2000
1500
1000
500
0
1
2
ASLV U3
ASLV dU3

Fig. 12

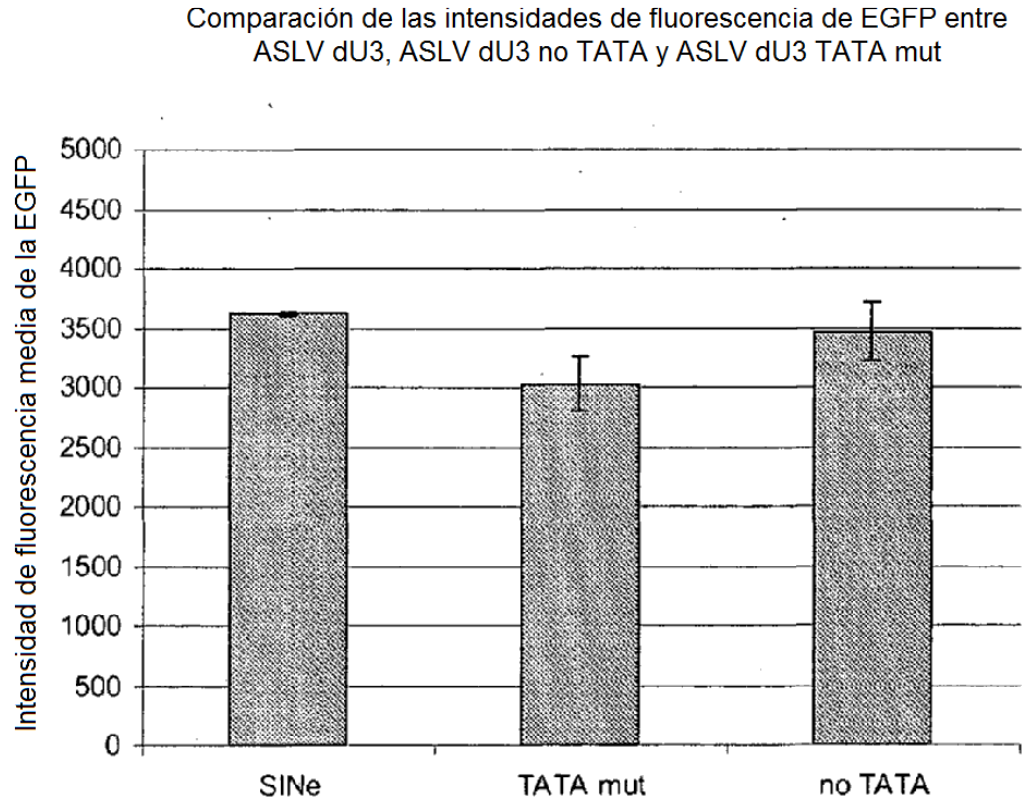

Comparación de las intensidades de fluorescencia de EGFP entre ASLV dU3, ASLV dU3 no TATA y ASLV dU3 TATA mut
Intensidad de fluorescencia media de la EGFP
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
SINe
TATA mut
no TATA

Fig. 13

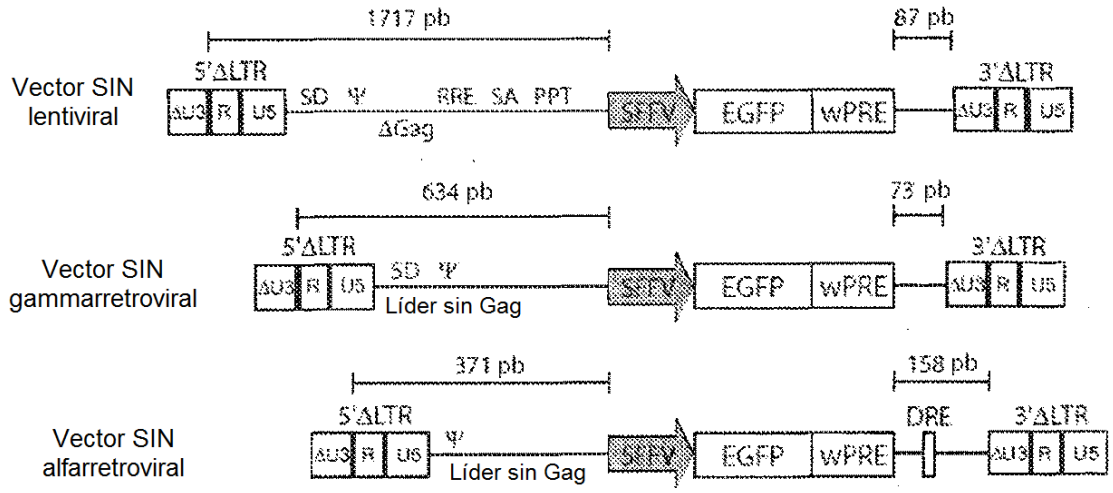
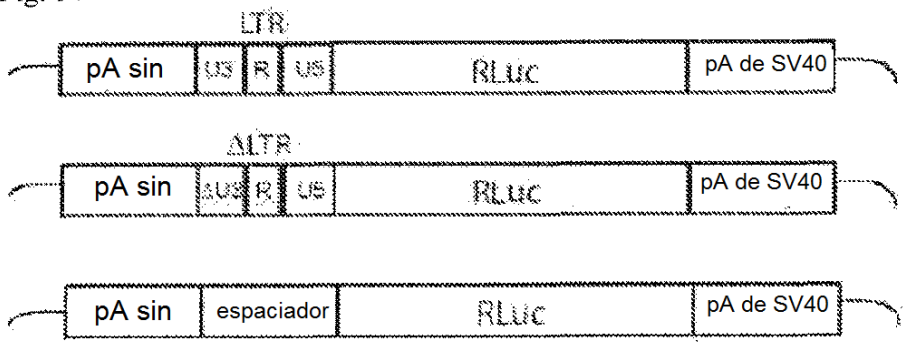

Vector SIN lentiviral
1717 pb
87 pb
5'ΔLTR
3'ΔLTR
ΔU3
R
U5
SD
Ψ
RRE
SA
PPT
ΔGag
SFFV
EGFP
WPRE
Vector SIN gammarretroviral
634 pb
73 pb
Líder sin Gag
Vector SIN alfarretroviral
371 pb
158 pb
DRE

Fig. 14


LTR
pA sin
U3
R
U5
RLuc
pA de SV40
ΔLTR
espaciador

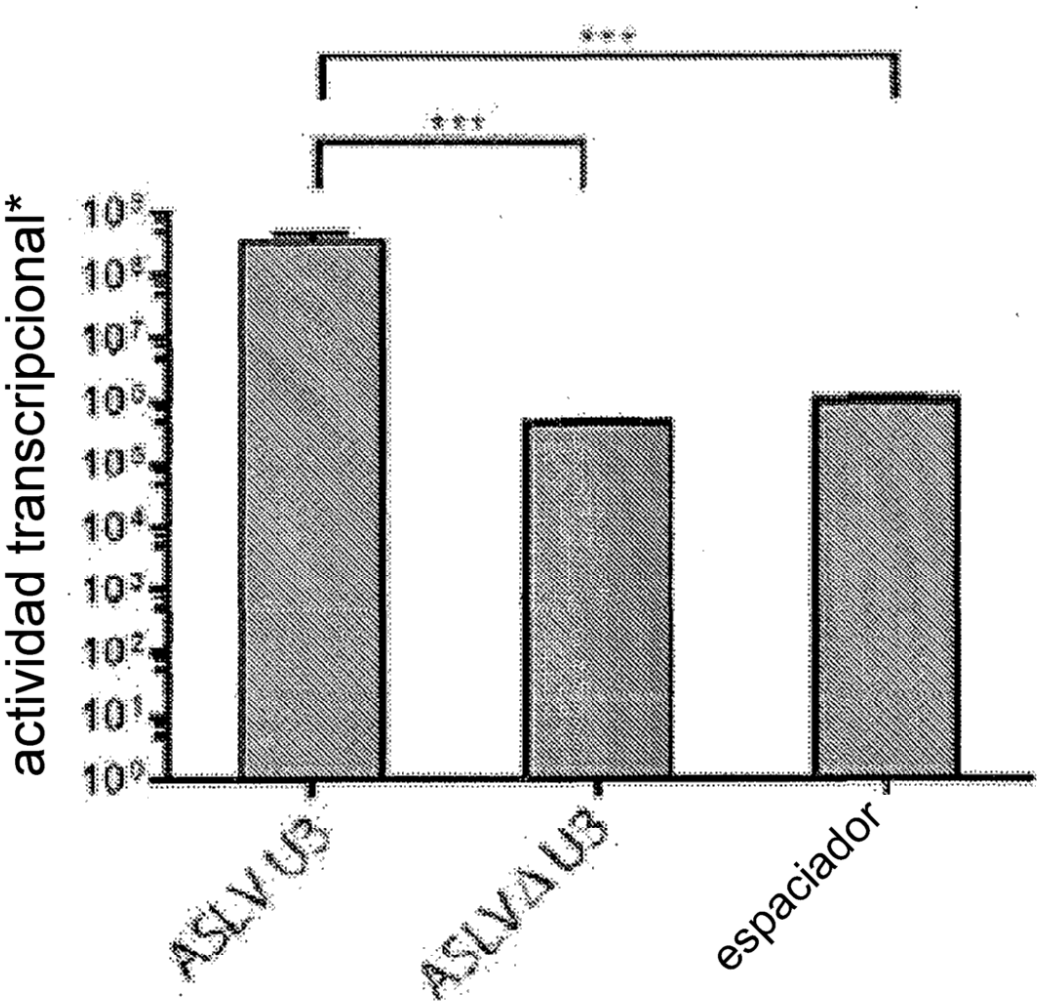



Fig. 15