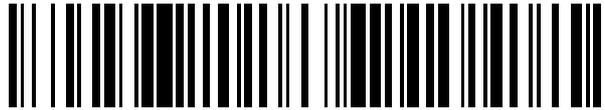


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 540 753**

51 Int. Cl.:

C12N 15/85 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **24.09.2012 E 12185728 (8)**

97 Fecha y número de publicación de la concesión europea: **08.04.2015 EP 2711426**

54 Título: **Vectores de expresión que comprenden secuencias quiméricas de promotor y amplificador de citomegalovirus**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
13.07.2015

73 Titular/es:

**LONZA BIOLOGICS PLC. (100.0%)
228-230 Bath Road
SloughSL1 4DX, GB**

72 Inventor/es:

**PAYNE, TOM;
YOUNG, ROBERT y
FEARY, MARC**

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 540 753 T3

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

DESCRIPCIÓN

Vectores de expresión que comprenden secuencias quiméricas de promotor y amplificador de citomegalovirus

5 **Campo de la invención**

La presente invención se refiere a sistemas de expresión de mamífero, y en particular a construcciones de expresión que comprenden secuencias reguladoras quiméricas para la expresión heteróloga de una secuencia de ácido nucleico de interés en células de mamífero. Las secuencias reguladoras del promotor quiméricas están compuestas por una secuencia de promotor del promotor IE1 de citomegalovirus murino y una secuencia de amplificador de la región IE1 del citomegalovirus humano y/o de simio.

Antecedentes

15 Los (poli)péptidos y proteínas recombinantes para aplicaciones en investigación básica, diagnóstico y terapia, tales como moléculas de anticuerpos, vacunas, hormonas, y factores de crecimiento, se producen utilizando una amplia variedad de organismos modificados genéticamente que incluyen tanto células procariontas como eucariotas. Sin embargo, la inmensa mayoría de péptidos o proteínas recombinantes incluyen modificaciones pos-traduccionales que no se pueden imitar o reproducir cuando se utilizan células huésped procariontas. Por esta razón, los sistemas de expresión genética de mamíferos han terminado representando la elección preferida.

Los sistemas de expresión en mamíferos que se basan en células ováricas de hámster (CHO) se utilizan ampliamente en la producción de proteínas recombinantes. Además de las líneas celulares linfoides, las células CHO representan uno de los pocos tipos celulares que permiten un cultivo discontinuo en suspensión de alta densidad simple y eficaz de células animales. Además, el uso de células CHO da como resultado rendimientos de producto muy altos, mientras que las células linfoides son más difíciles de cultivar a escala industrial. En vista de los costes considerables de la producción recombinante de polipéptidos y proteínas, es de la máxima importancia maximizar el rendimiento de proteína recombinante por proceso del biorreactor. Los parámetros del proceso que tienen un impacto considerable en el rendimiento de la producción incluyen *inter alia* las condiciones de cultivo celular, el número de copias de los ácidos nucleicos (genes) que se van a expresar, la eficacia con la que se transcriben estos genes y se traducen los ARNm correspondientes, la estabilidad del ARNm, y similares.

En consecuencia, las mejoras en la potencia o actividad de la transcripción de los elementos genéticos reguladores que controlan la expresión genética constituyen un factor particularmente crítico con el fin de aumentar el rendimiento de la proteína recombinante que se produce. Incluso pequeños aumentos crecientes de la actividad de transcripción a nivel de una única célula se traducirá finalmente en mejoras considerables del rendimiento de producto en cultivos discontinuos de alta densidad a escala industrial.

La inmensa mayoría de los sistemas de expresión genética de mamífero emplean vectores de expresión que codifican secuencias de ácido nucleico heterólogas que se van a expresar bajo el control de secuencias reguladores que derivan de un virus. Dos de los elementos reguladores víricos que se utilizan más frecuentemente en estos casetes de expresión son los genes inmediatos tempranos 1 y 2 (IE1 e IE2) de citomegalovirus humano (hCMV). Sin embargo, una desventaja que se asocia con el uso de los elementos reguladores IE1 e IE2 de hCMV es su pronunciada especificidad de especie.

La patente de EE. UU. 5.866.359 desvela que la expresión genética de tal promotor hCMV puede mejorarse con la co-expresión de proteína E1A adenovírica bajo el control de un promotor débil. La E1A es un factor de transcripción multifuncional que puede actuar en la regulación del ciclo celular y tiene dominios funcionales independientes tanto de activación como de represión de la transcripción. El ajuste preciso de la expresión de E1A es crucial para conseguir el equilibrio ideal entre la transactivación genética y cualquier impacto negativo sobre la progresión del ciclo celular. Sin embargo, la sobre-expresión de la expresión de E1A podría reducir la capacidad de la célula para sintetizar la proteína recombinante de interés.

La patente de EE. UU. 5.591.639 describe vectores que comprenden, corriente arriba (5') de la secuencia de ácido nucleico heteróloga que se va a expresar, el amplificador, promotor, y región completa 5' sin traducir del gen inmediato temprano principal del citomegalovirus humano (hCMV-MIE) que incluye el intrón A (es decir, el primer intrón natural). Sin embargo, si estaban presentes los primeros 400 pb (extremo 5') de esta secuencia (de longitud total 2100 pb), se observaban tasas bajas de expresión genética tanto en células COS7 como en CHO (Chapman, B.S. et al. (1991) Nucl. Acids Res. 19, 3979-3986).

La actividad de transcripción de los elementos reguladores de los genes inmediatos tempranos del citomegalovirus murino (mCMV) es mayor que los equivalentes del hCMV sin que manifieste la pronunciada preferencia por especies que se observa en las secuencias humanas (Addison, C.L. et al. (1987) J. Gen. Virol. 78. 1653-1661). La publicación de patente internacional WO 2004/081167 desvela vectores de expresión que comprenden el promotor IE2 de mCMV que se puede utilizar en combinación con un elemento amplificador IE2 de CMV. Se han descrito vectores de expresión para la expresión simultánea de proteína dual de alto nivel en células de vertebrados e insectos que

emplean elementos de promotor y amplificador de IE1 de mCMV en combinación con una construcción de promotor de baculovirus por Keil y sus colaboradores (Keil, G.M. et al. (2009) J. Viro/. Meth. 160, 132-137). Finalmente, se desvelan construcciones quiméricas de promotor que comprenden secuencias de promotor y amplificador, cada una de ellas de genes IE de citomegalovirus humano, murino o de simio, en la solicitud de patente europea 0 214 161.

5 Sin embargo, los intentos para aumentar la actividad de los elementos reguladores del promotor IE de mCMV, análogamente a los equivalentes del hCMV, mediante la inserción del primer intrón natural del gen inmediato temprano principal murino corriente abajo (3') del promotor IE de mCMV no funcionaron (cf. *inter alia* en patente EP 1 525 320 B1). Sin embargo, la generación de vectores de expresión que comprendían un casete quimérico compuesto por los elementos reguladores IE de mCMV y el primer intrón natural del gen inmediato temprano principal humano daba como resultado rendimientos en producto comparables con el uso de las secuencias completamente humanas (cf., por ejemplo, WO 2006/111387 A2). Se obtenían también tasas de expresión genética similares con vectores de expresión que comprendían las secuencias reguladoras IE2 de mCMV (cf. *inter alia* la patente EP 1 601 776 B1).

15 Por lo tanto, sigue existiendo la necesidad de sistemas de expresión de mamífero mejorados que tengan altos rendimientos de los polipéptidos o proteínas recombinantes que se producen. En particular, existe la necesidad de sistemas de expresión genética de mamífero que superen las limitaciones mencionadas anteriormente, es decir, sistemas de expresión que se basen en secuencias de promotor del mCMV (debido a su menor especificidad de especie) pero que alcancen mayores tasas de expresión (y por tanto de rendimientos de producto) que con el sistema disponible.

20 En consecuencia, un objetivo de la presente invención es proporcionar tales sistemas de expresión genética, construcciones de expresión adecuadas primariamente y las correspondientes células huésped de mamífero.

25 **Sumario de la invención**

En un aspecto, la presente invención se refiere a un vector de expresión para la expresión heteróloga de una secuencia de ácido nucleico de interés en células de mamífero, comprendiendo el vector una primera secuencia quimérica reguladora de promotor que está unida operativamente a una primera secuencia de ácido nucleico que se va a expresar, en donde la secuencia reguladora de promotor comprende:

- 35 (i) una secuencia promotora del promotor IE1 de citomegalovirus murino y que está unida operativamente al sitio de inicio de transcripción de la secuencia de ácido nucleico que se va a expresar; y
- (ii) una secuencia amplificadora de la región IE1 de citomegalovirus humano y/o de simio, localizándose la secuencia amplificadora 5' y unida operativamente a la secuencia de promotor IE1 de citomegalovirus murino; y en donde la secuencia quimérica reguladora de promotor comprende una secuencia de nucleótidos que se selecciona de entre el grupo que consiste en la SEC ID N° 7, SEC ID N° 9, y SEC ID N° 10.

40 En realizaciones particulares, el vector de expresión comprende además una segunda secuencia quimérica reguladora de promotor que está unida operativamente a una segunda secuencia de ácido nucleico que se va a expresar, en donde la segunda secuencia quimérica reguladora de promotor es idéntica a la primera secuencia quimérica reguladora de promotor.

45 En realizaciones particulares alternativas, el vector de expresión comprende además una segunda secuencia quimérica reguladora de promotor que está unida operativamente a una segunda secuencia de ácido nucleico que se va a expresar, en donde la segunda secuencia quimérica reguladora de promotor es diferente de la primera secuencia quimérica reguladora de promotor.

50 Preferentemente, las primera y segunda secuencias de ácido nucleico que se van a expresar codifican distintos polipéptidos. En realizaciones específicas, los distintos polipéptidos representan subunidades de una proteína dimérica o multimérica. Es particularmente preferible que la proteína dimérica o multimérica sea una molécula de anticuerpo.

55 En otro aspecto, la presente invención se refiere a una célula de mamífero transfectada con un vector de expresión como se ha definido anteriormente en el presente documento. Preferentemente, la célula huésped es una célula CHO.

60 En otro aspecto más, la presente invención se refiere a un método para la expresión heteróloga de una secuencia de ácido nucleico de interés en una célula huésped de mamífero, que comprende:

- 65 (i) transfectar la célula huésped de mamífero con un vector de expresión como se ha definido anteriormente en el presente documento; y
- (ii) cultivar la célula huésped de mamífero transfectada bajo condiciones que permitan la expresión de la secuencia de ácido nucleico de interés.

En realizaciones preferidas, la transfección es una transfección estable.

En un aspecto más, la presente invención se refiere al uso de un vector de expresión como se ha definido anteriormente en el presente documento para la expresión heteróloga *in vitro* de una secuencia de ácido nucleico de interés en una célula huésped de mamífero.

Otras realizaciones de la presente invención serán evidentes a partir de la descripción que se detalla de aquí en adelante.

10 Descripción de los dibujos

La **Figura 1** ilustra el vector de expresión pRY42 (SEC ID N° 1) que se utiliza como "vector parental" para generar los vectores de expresión de mamífero como se define en el presente documento. El pRY42 engloba dos casetes reguladores para dirigir la expresión genética heteróloga (entre "mCMV" y "pA", respectivamente): Sitios de clonación múltiples localizados 3' (es decir, "corriente abajo") de las regiones "Ex2" para la inserción de las secuencias de ácido nucleico heterólogas que se van a expresar. Los casetes reguladores están flanqueados por los sitios diana de reconocimiento de la flipasa de tipo mutante ("F5-mFRT") y tipo silvestre ("wFRT"). Se ha añadido un codón de inicio metionina en fase de lectura al extremo 5' del sitio wFRT ("ATG+F"). Se localiza un promotor temprano SV40 ("SV") 5' ("corriente arriba") de ATG+F. La transcripción de secuencias de ácido nucleico heterólogas está dirigida por el promotor del gen IE1 del citomegalovirus murino ("mCMV") que está seguido por la 5'UTR, en donde el exón 1 ("Ex1") es un híbrido de secuencias derivadas de CMV murino ("m") y humano ("h"), y donde el exón 2 ("Ex2") y la secuencia intrón A ("IntA") se derivan de la secuencia de hMCMV. El gen marcador de selección por β -lactamasa se señala como "bla". Se utiliza el pRY42 como un vector diana para la clonación de las distintas secuencias reguladoras de promotor quiméricas como se define en el presente documento.

La **Figura 2** ilustra el vector de expresión pRY57 (SEC ID N° 2) que codifica las cadenas ligera ("LC") y pesada ("HC") del anticuerpo monoclonal quimérico ratón-humano (mAb) cB72.3 (Whittle, N. et al. (1987) Protein Eng. 1, 499-505), que se localizan cada uno entre los sitios de clonación 3' de las regiones "Ex2" y los sitios de poliadenilación ("pA"), respectivamente. Por lo demás, el pRY57 es idéntico al pRY42. Las secuencias de ácido nucleico LC y HC del pRY57 se retiraron y se clonaron en variantes de pRY42 en el que la secuencia promotora original de mCMV se sustituyó con diferentes secuencias reguladoras de promotor quiméricas como se define en el presente documento.

La **Figura 3** representa esquemáticamente la secuencia promotora original de mCMV (SEC ID N° 3) que se engloba en el pRY42 (parte superior) así como las cinco diferentes secuencias reguladoras de promotor quiméricas (construcciones "1 a 5"), a saber las SEC ID N° 6 a SEC ID N° 10, respectivamente, en donde la SEC ID N° 6 y la SEC ID N° 8 no son parte de la presente invención. Las construcciones quiméricas comprenden secuencias promotoras IE1 de CMV murino (mCMV) localizadas 3' de los elementos amplificadores de la región IE1 de CMV de simio (sCMV) y/o CMV humano (hCMV). Todas las secuencias promotoras de mCMV empleadas específicamente en el presente documento incluyen en sus extremos 3' un nucleótido guanosina ("G") adicional, que representa el sitio de inicio de la transcripción.

La **Figura 4** muestra una comparación de las concentraciones del mAb cB72.3 producido en líneas CHO estables en las que la expresión genética de las secuencias del mAb estaba bajo el control de las construcciones quiméricas 1 a 5, como se ha ilustrado en la Figura 3. Se llevó a cabo la determinación por HPLC de proteína A tras 15 días de cultivo en matraces de cultivo de 50 ml con agitado utilizando un protocolo de sobrecrecimiento con alimentación discontinua (FOG). Para cada construcción quimérica, n = 4, representa los análisis de alimentación discontinua por duplicado para transfecciones duplicadas, con la excepción de la construcción 4, donde n = 6 (análisis FOG duplicados para transfecciones por triplicado) y pRY57 (mCMV original), donde n = 8, por la combinación de los puntos de datos de los experimentos 1 y 2.

Descripción detallada de la invención

La presente invención se basa en el hallazgo inesperado de que los vectores de expresión de mamífero comprenden secuencias reguladoras de promotor quiméricas (es decir, híbridas) que están compuestas por una secuencia promotora IE1 de mCMV en unión operativa con el sitio de inicio de la transcripción de la secuencia de ácido nucleico que se va a expresar y una secuencia amplificadora IE1 de hCMV y/o IE1 de sCMV que está localizada 5' de la secuencia promotora de mCMV, en donde la secuencia quimérica reguladora de promotor comprende una secuencia de nucleótidos que se selecciona de entre el grupo que consiste en las SEC ID N° 7, SEC ID N° 9, y SEC ID N° 10, que daba como resultado unas tasas de expresión genética mejores en comparación con los sistemas de expresión existentes que se basaban solo en las secuencias promotoras de mCMV, y por tanto también rendimientos mucho más altos (hasta un aumento de casi 3 veces) de las proteínas recombinantes que se producían.

En consecuencia, los vectores de expresión de mamífero como se definen en el presente documento representan

herramientas moleculares superiores para la producción de proteínas recombinantes, particularmente a escala industrial.

5 Cuando se utiliza la expresión “que comprende” en la presente descripción y las reivindicaciones, no se excluyen otros elementos o etapas. Para los fines de la presente invención, la expresión “que consiste en” se considera que es una realización preferida de la expresión “que comprende”. Si de aquí en adelante un grupo que se define como que comprende al menos un cierto número de realizaciones, se entenderá también que desvela un grupo que consiste preferentemente solo en esas realizaciones.

10 Cuando se utiliza un artículo definido o indefinido en referencia a un nombre en singular, por ejemplo, “un”, “una” o “el”, incluye un plural de ese nombre a menos de que se establezca específicamente otra cosa.

15 En el que caso en el que se indiquen valores numéricos en el contexto de la presente invención, el experto entenderá que el efecto técnico de la característica en cuestión se asegura con un intervalo de precisión, que normalmente engloba una desviación del valor numérico dado del $\pm 10\%$, y preferentemente del $\pm 5\%$.

20 Además, los términos primero, segundo, tercero, (a), (b), (c), y similares, en la descripción y las reivindicaciones, se utilizan para distinguir entre elementos similares y no necesariamente para describir un orden secuencial o cronológico. Se tiene que entender que los términos que se utilizan de esta manera son intercambiables bajo circunstancia adecuadas y que las realizaciones de la invención que se describen en el presente documento son capaces de ejecutarse en otras secuencias que las que se describen o ilustran en el presente documento.

25 Otras definiciones de términos se darán a continuación en el contexto en que se utilizan tales términos. Los siguientes términos o definiciones se proporcionan solo para ayudar a entender la invención. Estas definiciones no se deben considerar como que tienen un ámbito menor que el entendimiento por parte de un experto en la técnica.

30 En un aspecto, la presente invención se refiere a un vector de expresión para la expresión heteróloga de una secuencia de ácido nucleico de interés en células de mamífero, comprendiendo el vector una primera secuencia química reguladora de promotor que está unida operativamente a una primera secuencia de ácido nucleico que se va a expresar, en donde la secuencia química reguladora de promotor comprende:

- (i) una secuencia promotora del promotor IE1 de citomegalovirus murino y que está unida operativamente al sitio de inicio de la transcripción de la secuencia de ácido nucleico que se va a expresar; y
- 35 (ii) una secuencia amplificadora de la región IE1 del citomegalovirus humano y/o de simio, estando localizada la secuencia amplificadora 5' y unida operativamente a la secuencia promotora IE1 de citomegalovirus murino; y

en donde la secuencia química reguladora de promotor comprende una secuencia de nucleótidos que se selecciona de entre el grupo que consiste en la SEC ID N° 7, SEC ID N° 9 y SEC ID N° 10.

40 La expresión “vector de expresión”, como se utiliza en el presente documento, denota un vehículo de ácido nucleico (plásmido) que se propaga autónomamente en una célula huésped (es decir, independiente del ácido nucleico cromosómico) que se caracteriza por la presencia de al menos un “casete de expresión”. La expresión “casete de expresión”, como se utiliza en el presente documento, se refiere a una construcción genética que es capaz de permitir la expresión genética de una secuencia de ácido nucleico de interés (es decir, una secuencia de ácido nucleico “heteróloga”). Esto requiere que tal casete de expresión comprenda elementos reguladores de secuencia que contengan la información con respecto a la regulación de la transcripción y/o la traducción, y que tales secuencias reguladoras están “unidas operativamente” a la secuencia de ácido nucleico de interés. Una unión operativa es una unión en la que los elementos de la secuencia reguladora y la secuencia de ácido nucleico que se va a expresar están conectados de una manera que se haga posible la expresión genética.

50 La naturaleza precisa de las regiones reguladoras de un “casete de expresión” que son necesarias para controlar y dirigir la expresión genética puede variar entre las especies, pero en general estas regiones comprenden secuencias reguladoras de promotor (es decir, una región de secuencia localizada 5' (“corriente arriba”) de la secuencia de ácido nucleico de interés) y secuencias reguladoras no traducidas 3' (es decir, una región de secuencia localizada 3' (“corriente abajo”) de la secuencia de ácido nucleico de interés).

60 El término “promotor”, (también designado “centro promotor”) como se utiliza en el presente documento, denota elementos de secuencia que *per se* dirigen el inicio de la transcripción (por ejemplo, sitios de unión para los factores de transcripción y para la ARN polimerasa dependiente de ADN, TATA box, secuencias CAAT, y elementos de taponamiento 5'). Siempre que se mantenga esta funcionalidad de promoción de inicio de la transcripción (por ejemplo, al menos un 70 %, al menos un 80 %, al menos un 90 % o al menos un 95 % de la actividad del tipo silvestre, es decir, la actividad de una secuencia de longitud completa), cualquier variante truncada, mutada o modificada de otra manera de una secuencia promotora de tipo silvestre (de origen natural), también están dentro de la definición anterior. Como se utiliza en el presente documento, la expresión “centro promotor” se refiere a una secuencia de mínima longitud que mantiene la actividad promotora.

Los vectores de expresión de la presente invención comprenden (como parte de un casete de expresión) una primera secuencia (quimérica) reguladora de promotor (es decir al menos una de tales secuencias), la cual a su vez, engloba una secuencia de (centro) promotor que se deriva del promotor IE1 del citomegalovirus murino (mCMV). Esta secuencia promotora de mCMV está unida operativamente al sitio de inicio de la transcripción de una primera secuencia de ácido nucleico que se va a expresar.

El promotor IE1 de mCMV se conoce bien en la técnica y se puede derivar fácilmente del genoma del mCMV depositado en la base de datos de Genomas víricos NCBI bajo el nº de registro U68299.1 (<http://www.ncbi.nlm.nih.gov/genomes/GenomesHome>; Bao, Y. et al. (2004) J. Virol. 78, 7291-7298).

La secuencia promotora IE1 de mCMV que está comprendida en los vectores de expresión como se define en el presente documento tiene una secuencia de ácido nucleico de 103 pb de longitud (también denominada "centro promotor") como se muestra en la SEC ID N° 4:

```

1 acaccgcccc ggttttcccc tggaaattcc atattggcac gcattctatt
51 ggctgagctg cgttctacgt gggataaaga ggcgcgacca gcgtcggtag
101 cg

```

La SEC ID N° 4 incluye en el extremo 3' un nucleótido guanosina ("G") adicional, que representa el sitio de inicio de la transcripción.

Además, las secuencias reguladoras de promotor de un casete de expresión comprenden habitualmente una secuencia "amplificadora". El término "amplificador", como se utiliza en el presente documento, denota elementos de secuencia que aumentan, potencian o mejoran la transcripción de una secuencia de ácido nucleico independientemente de su localización y orientación con respecto a la secuencia de ácido nucleico que se va a expresar. Un amplificador puede aumentar la transcripción a partir de un promotor único o simultáneamente a partir de más de un promotor. Siempre que se mantenga esta funcionalidad de mejorar la transcripción o se mantenga sustancialmente (por ejemplo, al menos un 70 %, al menos un 80 %, al menos un 90 % o al menos un 95 % de la actividad del tipo silvestre, es decir, la actividad de una secuencia de longitud completa), cualquier variante truncada, mutada o modificada de otra manera de una secuencia amplificadora de tipo silvestre (de origen natural) está incluida también en la definición anterior.

Los vectores de expresión de la presente invención comprenden (como parte de un casete de expresión) en su primera secuencia (quimérica) reguladora de promotor respectiva, una secuencia amplificadora que se deriva de la región IE1 del citomegalovirus humano (hCMV) y/o citomegalovirus de simio (sCMV) (Meier, J.L. y Stinski, M.F. (1996) Intervirology 39, 331-342; Kim, G.Y. et al. (2011) Biotechnol. Lett. 33, 1319-1326). En otras palabras, la secuencia reguladora de promotor puede comprender una secuencia amplificadora que se derive solamente del hCMV (en la SEC ID N° 7) o una secuencia amplificadora que solo se deriva de sCMV (en la SEC ID N° 9). En la secuencia reguladora de promotor, las secuencias amplificadoras se localizan 5' (es decir, "corriente arriba") de las secuencias del centro promotor del mCMV. Las secuencias amplificadoras se disponen en la misma orientación que las secuencias promotoras.

Estas secuencias amplificadoras IE1 de hCMV y/o sCMV se conocen bien en la técnica y se pueden derivar fácilmente de los genomas de hCMV y sCMV depositados en las bases de datos de Genomas víricos NCBI bajo los nºs de registro X17403.1 y U38308.1, respectivamente.

La secuencia del elemento amplificador puede comprender la secuencia de nucleótidos de 452 pb de longitud, como se muestra en la SEC ID N° 5, que es del amplificador IE1 del sCMV.

ES 2 540 753 T3

```
1 gaccatagcc aattcaatat ggcgtatatg gactcatgcc aattcaatat
51 ggtggatctg gacctgtgcc aattcaatat ggcgtatatg gactcgtgcc
101 aattcaatat ggtggatctg gaccccagcc aattcaatat ggcggacttg
151 gcacatgcc aattcaatat ggcggacctg gactgtgcc aactggggag
201 ggttctactt ggcacggtgc caagtttgag gaggggtctt ggccctgtgc
251 caagtcgcc atattgaatt ggcattgtgc caataatggc ggccatattg
301 gctatatgcc aggatcaata tataggcaat atccaatatg gccctatgcc
351 aatattgcta ttggccaggt tcaatactat gtattggccc tatgccatat
401 agtattccat atattgggtt tctatttgac gtagatagcc cctcccaatg
451 gg
```

5 La secuencia de ácido nucleico SEC ID N° 5 puede ser una parte integral de un elemento de secuencia más larga de la región IE1 de sCMV (cf., por ejemplo, la SEC ID N° 10). De manera alternativa, la secuencia del elemento que tiene la SEC ID N° 5 puede estar presente en combinación con una secuencia de elemento más de la región IE1 de hCMV (cf., por ejemplo, SEC ID N° 9).

10 La secuencia química reguladora de promotor como se define en el presente documento comprende una secuencia de nucleótidos que se selecciona de entre el grupo que consiste en la SEC ID N° 7, SEC ID N° 9, y SEC ID N° 10.

15 La secuencia química reguladora de promotor de acuerdo con la SEC ID N° 6 (denominada también "construcción 1" en el presente documento, la cual no forma parte de la presente invención) tiene una longitud total de 1074 pb y está compuesta por 582 pb de la secuencia amplificadora IE1 de hCMV (que se muestra en cursiva) y 492 pb de la secuencia promotora IE1 de mCMV (que se muestra también como SEC ID N° 3).

```
1 ctgcagtgaa taataaaatg tgtgtttgtc cgaaatagc gttttgagat
51 ttctgtcgcc gactaaattc atgtcgcgcg atagtgggtg ttatcgccga
101 tagagatggc gatattggaa aaatcgatat ttgaaaatat ggcattattga
151 aaatgtcgcc gatgtgagtt tctgtgtaac tgatatcgcc atttttccaa
```

ES 2 540 753 T3

201 aagtgatttt tgggcatacg cgatatctgg cgatagcgct tatatcgttt
 251 acgggggatg gcgatagacg actttggtga cttgggcat tctgtgtgtc
 301 gcaaataatcg cagtttcgat ataggtgaca gacgatatga ggctatatcg
 351 ccgatagagg cgacatcaag ctggcacatg gccaatgcat atcgatctat
 401 acattgaatc aatattggcc attagccata ttattcattg gttatatagc
 451 ataaatcaat attggtatt ggccattgca tacgttgtat ccatatcata
 501 atatgtacat ttatattggc tcatgtccaa cattaccgcc atggtgacat
 551 tgattattga ctagttatta atagtaatca attactgagt ctaggggac
 601 tttccaatgg gttttgccca gtacataagg tcaatagggg tgaatcaaca
 651 ggaaagtccc attggagcca agtacactga gtcaataggg actttccatt
 701 gggttttgcc cagtacaaaa ggtcaatagg gggtgagtca atgggttttt
 751 cccattattg gcacgtacat aaggtcaata ggggtgagtc attgggtttt
 801 tccagccaat ttaattaa. ccgatgtac tttcccacca ttgacgtcaa
 851 tgggctattg aaactaatgc aacgtgacct ttaaaccgta ctttccata
 901 gctgattaat gggaaagtac cgttctcgag ccaatacacg tcaatgggaa
 951 gtgaaagggc agccaaaacg taacaccgcc ccggttttcc cctggaaatt
 1001 ccatattggc acgcattcta ttggctgagc tgcggtctac gtgggtataa
 1051 gaggcgogac cagcgtcggc accg

5 La secuencia quimérica reguladora de promotor de acuerdo con la SEC ID N° 7 (también denominada en el presente documento "construcción 2") tiene una longitud total de 1128 pb y está compuesta por 1026 pb de la secuencia amplificadora IE1 de hCMV (que se muestra en cursiva) y 102 pb de la secuencia del "centro" promotor IE1 de mCMV (que se muestra también como la SEC ID N° 4).

1 ctgcagtgaa taataaaatg tgtgtttgtc cgaataacgc gttttgagat
 51 ttctgtcgcc gactaaatc atgtcgcgcg atagtgggtg ttatcgccga
 101 tagagatggc gatattggaa aaatcgatat ttgaaaatat ggcatattga
 151 aaatgtcgcc gatgtgagtt tctgtgtaac tgatatcgcc atttttccaa
 201 aagtgatttt tgggcatacg cgatatctgg cgatagcgct tatatcgttt
 251 acgggggatg gcgatagacg actttggtga cttgggcat tctgtgtgtc
 301 gcaaataatcg cagtttcgat ataggtgaca gacgatatga ggctatatcg
 351 ccgatagagg cgacatcaag ctggcacatg gccaatgcat atcgatctat
 401 acattgaatc aatattggcc attagccata ttattcattg gttatatagc

ES 2 540 753 T3

451 ataaatcaat attggctatt ggccattgca tacgttgtat ccatatcata
 501 atatgtacat ttatatggc tcatgtccaa cattaccgcc atgttgacat
 551 tgattattga ctagttatta atagtaatca attacggggc cattagttca
 601 tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc
 651 ctggctgacc gcccaacgac ccccgcccat tgacgtcaat aatgacgtat
 701 gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga
 751 gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc
 801 caagtacgcc ccctattgac gtcaatgacg gtaaattggcc cgcctggcat
 851 tatgcccagt acatgacctt atgggacttt cctacttggc agtacatcta
 901 cgtattagtc atcgctatta ccatgggtgat gcggtttttg cagtacatca
 951 atgggcgtgg atagcggttt gactcacggg gatttccaag tctccacccc
 1001 attgacgtca atgggagttt gttttgacac cgccccggtt tccccctgga
 1051 aattccatat tggcagcat tctattggct gagctgcgtt ctacgtgggt
 1101 ataagaggcg cgaccagcgt cggtaccg

5 La secuencia quimérica reguladora de promotor de acuerdo con la SEC ID N° 8 (denominada también en el presente documento "construcción 3" la cual no es parte de la presente invención) tiene una longitud total de 509 pb y está compuesta por 407 pb de la secuencia amplificadora IE1 de hCMV (que se muestra en cursiva) y 102 pb de la secuencia del "centro" promotor IE1 de mCMV (que se muestra también como la SEC ID N° 4).

1 *cgcgttacat aacttacggt aatggcccg cctggctgac cgcccaacga*
 51 *ccccgccc ttgacgtcaa taatgacgta tgttcccata gtaacgcaa*
 101 *tagggacttt ccattgacgt caatgggtgg agtatctacg gtaaactgcc*
 151 *cacttggcag tacatcaagt gtatcatatg ccaagtacgc ccctattga*
 201 *cgtcaatgac ggtaaatggc ccgcctggca ttatgccag tacatgacct*
 251 *tatgggactt tctacttgg cagtacatct acgtattagt catcgctatt*
 301 *accatggtga tgcggttttg gcagtacatc aatgggcgtg gatagcggtt*
 351 *tgactcacgg ggatttcaa gtctccacc cattgacgtc aatgggagtt*
 401 *tgttttgaca ccgccccggt tttccccctgg aaattccata ttggcagca*
 451 *ttctattggc tgagctgcgt tctacgtggg tataagaggc gcgaccagcg*
 501 *tcggtaccg*

10 La secuencia quimérica reguladora de promotor de acuerdo con la SEC ID N° 9 (también denominada en el presente documento "construcción 4") tiene una longitud total de 961 pb y está compuesta por 452 pb de una secuencia amplificadora IE1 de mCMV (que se muestra en negrita; SEC ID N° 5), 407 pb de la secuencia amplificadora IE1 de hCMV (que se muestra en cursiva) y 102 pb de la secuencia del "centro" promotor IE1 de mCMV (que se muestra también como la SEC ID N° 4).

15

ES 2 540 753 T3

1 gaccatagcc aattcaatat ggcgtatatg gactcatgcc aattcaatat
 51 ggtggatctg gacctgtgcc aattcaatat ggcgtatatg gactcgtgcc
 101 aattcaatat ggtggatctg gaccccagcc aattcaatat ggcggacttg
 151 gcaccatgcc aattcaatat ggcggacctg gcaactgtgcc aactggggag
 201 gggctctactt ggcacgggtgc caagtttgag gaggggtctt ggccctgtgc
 251 caagtcgcgc atattgaatt ggcactgtgc caataatggc ggccatattg
 301 gctatatgcc aggatcaata tataggcaat atccaatatg gccctatgcc
 351 aatatggcta ttggccaggt tcaatactat gtattggccc tatgccatat
 401 agtattccat atatgggttt tcctattgac gtagatagcc cctcccaatg
 451 **ggcgcgttac** ataacttacg gtaaatggcc cgcctggctg accgcccac
 501 gacccccgc cattgacgtc aataatgacg tatgttccca tagtaacgcc
 551 aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg
 601 cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt
 651 gacgtcaatg acggtaaagt gcccgcctgg cattatgccc agtacatgac
 701 cttatgggac tttcctactt ggcagtacat ctacgtatta gtcatcgcta
 751 ttaccatggt gatgcggttt tggcagtaca tcaatgggcg tggatagcgg
 801 tttgactcac ggggatttcc aagtctccac ccattgacg tcaatgggag
 851 tttgttttga caccgccccg gttttcccct ggaaattcca tattggcacg
 901 cattctattg gctgagctgc gttctacgtg ggtataagag gcgcgaccag
 951 cgtcggtacc g

5 La secuencia química reguladora de promotor de acuerdo con la SEC ID N° 10 (denominada también en el presente documento "construcción 5") tiene una longitud total de 909 pb y está compuesta por 807 pb de la secuencia amplificadora IE1 de sCMV (que se muestra en **negrita**; SEC ID N° 5 que está subrayada) y 102 pb de la secuencia del "centro" promotor IE1 de mCMV (que se muestra como la SEC ID N° 4).

1 gaccatagcc aattcaatat ggcgtatatg gactcatgcc aattcaatat
 51 ggtggatctg gacctgtgcc aattcaatat ggcgtatatg gactcgtgcc
 101 aattcaatat ggtggatctg gaccccagcc aattcaatat ggcggacttg
 151 gcaccatgcc aattcaatat ggcggacctg gcactgtgcc aactggggag
 201 gggtctactt ggcacgggtgc caagtttgag gaggggtctt ggcctgtgc
 251 caagtccgcc atattgaatt ggcattggtgc caataatggc ggccatattg
 301 gctatatgcc aggatcaata tataggcaat atccaatatg gcctatgcc
 351 aatatggcta ttggccaggt tcaatactat gtattggccc tatgccatat
 401 agtattccat atatggggtt tcctattgac gtagatagcc cctcccaatg
 451 ggcgggtcca tataccatat atggggcttc ctaataccgc ccatagccac
 501 tccccattg acgtcaatgg tctctatata tggctcttcc tattgacgtc
 551 atatggggcgg tcctattgac gtatatggcg cctcccccat tgacgtcaat
 601 tacggtaaat ggcccgcctg gctcaatgcc cattgacgtc aataggacca
 651 cccaccattg acgtcaatgg gatggctcat tgcccattca tatccgttct
 701 cacgccccct attgacgtca atgacggtaa atggcccact tggcagtaca
 751 tcaatatcta ttaatagtaa cttggcaagt acattactat tggaagtacg
 801 ccagggtaga ccgccccggt ttccccctgg aaattccata ttggcacgca
 851 ttctattggc tgagctgctg tctacgtggg tataagaggc gcgaccagcg
 901 tcggtaccg

5 Las secuencias reguladoras 3' de un "casete de expresión" como se definen en el presente documento engloban normalmente elementos reguladores que están implicados en la terminación de la transcripción, poliadenilación, o similares. Sin embargo, si estas secuencias de terminación no son funcionales satisfactoriamente en una célula huésped particular, pueden sustituirse entonces con señales funcionales de esa célula.

10 Otros elementos reguladores que están comprendidos en tal "casete de expresión" incluyen *inter alia* sitios internos de entrada en el ribosoma (IRES; que permiten la expresión de secuencias de ácido nucleico "policistrónicas") así como secuencias de señal traducidas para dirigir el polipéptido nativo a un compartimento específico de una célula huésped. Las secuencias de señal ejemplares adecuadas para las células CHO se desvelan, por ejemplo, en el documento WO 2008/148519 A2. El experto en la técnica también es bien consciente de todos estos elementos reguladores y de cómo seleccionar tales elementos adecuados para la expresión de una secuencia de ácido nucleico en un entorno celular en particular.

15 El vector de expresión de la presente invención puede ser, por ejemplo, un plásmido, cósmido, fagémido, cromosoma artificial, u otro vehículo que se utiliza habitualmente en modificación genética.

20 Tales vectores de expresión incluyen normalmente, junto con uno o más "casetes de expresión" que engloban las secuencias reguladoras descritas anteriormente, uno o más sitios de clonación múltiples, con el fin de facilitar la inserción y/o retirada de secuencias de ácido nucleico. Los sitios de clonación múltiples se pueden localizar corriente arriba o corriente abajo de los casetes de expresión descritos anteriormente, permitiendo de esta manera el remplazo del casete completo. Los sitios de clonación múltiples se pueden localizar también en el casete de expresión corriente abajo de las secuencias reguladoras de promotor y corriente arriba de las secuencias reguladoras 3', permitiendo de esta manera la inserción o remplazo de una secuencia de ácido nucleico que se va a expresar. Para las transfecciones estables (cf. posteriormente), los vectores de expresión pueden comprender además secuencias de reconocimiento para integrasas o recombinasas específicas del sitio con el fin de facilitar la recombinación y la integración estable en el genoma de la célula huésped.

30 Además, los vectores de expresión como se definen en el presente documento comprenden normalmente al menos un origen de replicación así como secuencias de control derivadas de una especie compatible con la célula huésped que se emplee con el fin de asegurar el mantenimiento autónomo de replicación/episómico del vector de expresión (en particular, para su uso en transfecciones transitorias; cf. posteriormente). Los orígenes de replicación ejemplares

en mamíferos incluyen el origen de replicación SV40 o el EBV. Los vectores de expresión diseñados específicamente (por ejemplo, los vectores lanzadera) comprenden más de un origen de replicación que permiten el lanzamiento entre huéspedes diferentes, tal como entre células bacterianas y animales. Los orígenes de replicación adecuados para células procariotas incluyen, por ejemplo, los orígenes de replicación ColE1 y M13.

5 Además, un vector de expresión como se define en el presente documento puede comprender uno o más marcadores de selección que confieren un fenotipo seleccionable a las células transfectadas. Los marcadores de selección adecuados incluyen *inter alia* el gen de fosfatasa B higromicina, el gen de timidina quinasa, el gen de ornitina descarboxilasa, gen de dihidrofolato reductasa, y el gen de glutamina sintasa. Preferentemente se emplean el gen de glutamina sintasa (GS) (Cockett, D.K. et al. (1990) *Bio/Technology* 8, 662-667; Bebbington, C.R. et al. (1992) *Bio/Technology* 10:169-175). Los numerosos métodos que se pueden utilizar para diseñar y/o modificar los vectores de expresión recombinantes están bien establecidos en la técnica (véase, por ejemplo, Sambrook, J., y Russel, D.W. (2001), *Molecular cloning: A laboratory manual* (3rd Ed.) Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press; Ausubel, F.M. et al. (2001) *Current Protocols in Molecular Biology*, Wiley & Sons, Hoboken, NJ, USA). También hay disponible comercialmente una gran cantidad de vectores de expresión de mamífero adecuados y bien conocidos por el experto que también es capaz de determinar qué vectores son adecuados para expresar una molécula de ácido nucleico de interés en un entorno determinado. Ejemplos de tales vectores, incluyen *inter alia*, pcDNA3, pFRT, pTARGET, pSV2-dhfr así como derivados de los vectores pRY42 (SEC ID N° 1) y pRY57 (SEC ID N° 2) descritos anteriormente en el presente documento.

20 Las secuencias de ácido nucleico que se van a expresar empleando los vectores de expresión de la invención pueden ser monocistrónicos (es decir, que codifican un único polipéptido o proteína incluyendo proteínas de fusión) o policistrónicos (es decir, que codifican dos o más polipéptidos o proteínas individuales).

25 En realizaciones particulares, el vector de expresión comprende una única (es decir, primera) secuencia quimérica reguladora de promotor que está unida operativamente a una (primera) secuencia de ácido nucleico que se va a expresar.

30 En otras realizaciones particulares, el vector de expresión comprende además una segunda secuencia quimérica reguladora de promotor que está unida operativamente a una segunda secuencia de ácido nucleico que se va a expresar, en donde la segunda secuencia quimérica reguladora de promotor es idéntica a la primera secuencia quimérica reguladora de promotor.

35 En realizaciones particulares alternativas, el vector de expresión comprende además una segunda secuencia quimérica reguladora de promotor que está unida operativamente a una segunda secuencia de ácido nucleico que se va a expresar, en donde la segunda secuencia quimérica reguladora de promotor es diferente de la primera secuencia quimérica reguladora de promotor.

40 En realizaciones específicas, el vector de expresión comprende una tercera secuencia quimérica reguladora de promotor que está unida operativamente a una tercera secuencia de ácido nucleico que se va a expresar, en donde la tercera secuencia quimérica reguladora de promotor puede ser idéntica a la primera y/o segunda secuencia quimérica reguladora de promotor o puede ser diferente tanto de la primera como de la segunda secuencia quimérica reguladora de promotor. En otras realizaciones específicas, el vector de expresión comprende más de tres secuencias reguladoras de promotor quiméricas.

45 Las secuencias de ácido nucleico que se van a expresar empleando los vectores de expresión de la presente invención pueden codificar cualquiera de los polipéptidos o proteínas de interés, en particular polipéptidos o proteínas que tienen una aplicabilidad diagnóstica o terapéutica, tales como *inter alia* factores de crecimiento, citoquinas (interferones, interleucinas), hormonas, tirosina quinasas, receptores (GPCR), integrinas, factores de transcripción, factores de coagulación sanguínea, anticuerpos de cadena sencilla, fragmentos de anticuerpo o moléculas tipo anticuerpo (anticálidas), y similares.

50 En el caso de los vectores de expresión como se definen en el presente documento que comprenden dos secuencias reguladoras de promotor (como parte de los casetes de expresión), la primera y segunda secuencia de ácido nucleico que se van a expresar codifican diferentes polipéptidos (o proteínas). En realizaciones específicas, los diferentes polipéptidos representan subunidades de una proteína dimérica o multimérica, tales como *inter alia* moléculas de receptor homoméricas o heteroméricas, hormonas peptídicas, ARN/ADN polimerasas, hemoglobinas, vacunas, y similares.

60 En realizaciones particularmente preferibles, la proteína dimérica o multimérica es una molécula de anticuerpo "clásica" que comprende la cadena ligera como la primera subunidad y la cadena pesada como la segunda subunidad. La molécula de anticuerpo puede ser de origen natural o un anticuerpo modificado genéticamente, sea un anticuerpo de longitud completa o una variante truncada del mismo (tal como fragmentos Fab, fragmentos Fab', fragmentos F(ab')₂). Los anticuerpos inmunoglobulinas IgG son particularmente preferidos. Dependiendo de la aplicación específica, las moléculas de anticuerpo pueden ser quiméricas (por ejemplo murinas/humanas), humanizadas, o completamente humanas.

En otras realizaciones que emplean vectores de expresión que comprenden dos secuencias reguladoras de promotor quiméricas, la primera y segunda secuencia de ácido nucleico que se van a expresar codifican una "proteína diana" que se va a analizar y una correspondiente proteína indicadora (tal como una proteína fluorescente verde, luciferasa, fosfatasa alcalina, β -galactosidasa, y peroxidasa de rábano rústico) para controlar, por ejemplo, la localización celular o la actividad funcional de una proteína diana. Cuando se utilizan dos (o más) secuencias reguladoras de promotor que muestran diferentes tasas de expresión genética puede ser posible producir ciertas proporciones molares de las correspondientes proteínas de interés.

En otro aspecto, la presente invención se refiere a una célula huésped de mamífero que se transfecta con un vector de expresión como se ha definido anteriormente en el presente documento.

Las células huésped adecuadas incluyen cualquier tipo de células de mamífero, siendo las células de origen humano o no humano. Las células de mamífero de origen no humano incluyen *inter alia* células derivadas de ratón, rata, hámster, conejo, gato, perro, cerdo, vaca, caballo o mono.

Las células huésped de mamífero incluyen líneas celulares inmortalizadas tales como Hela, fibroblastos MRC5, 983M de melanoma, HEK293, H9, MCF7, y células Jurkat humanas; células renales caninas MDCK; fibroblastos de pulmón de rata cultivadas RF aisladas de ratas Sprague-Dawley; NIH3T3 murinas, C127, P815 de mastocitoma, MT1A2 de adenocarcinoma mamario, y células L; células COS1 y COS7 de simio; células QC1-3 de codorniz; y células o líneas celulares de ovario de hámster chino (CHO).

En realizaciones preferidas, las células huésped que se emplean son células CHO o líneas celulares de CHO.

Las líneas celulares de CHO adecuadas incluyen *inter alia* CHO KI (Tjio, J.T. y Puck, T.T. (1958) J. Exp. Med. 108, 945-955), CHO pro3-, CHO DG44, CHO P12, DUK-B11 dhfr-negativa (Urlaub, G. y Chasin L.A. (1980) Proc. Natl. Acad. Sci. USA 77, 4216-4220), y particularmente CHOK1SV (Lonza Ltd. Basilea, Suiza). CHOK1SV es una suspensión derivada de CHOK1 adaptada libre de proteínas que utiliza el sistema de expresión del gen glutamina sintasa (GS): se obtuvieron los transfectantes positivos bajo una selección doble en medios de metionina sulfoximina y libres de glutamina.

Todas estas células huésped o líneas celulares se pueden obtener de depósitos tales como la Colección Americana de Cultivos Tipo (Manassas, VA, EE. UU.) o la Deutsche Sammlung von Mikroorganismen und Zellkulturen (Braunschweig, Alemania) así como en varios suministradores comerciales. También en la presente invención hay células primarias de mamífero, es decir, células que se obtienen directamente de un organismo (en cualquier nivel de desarrollo, que incluyen *inter alia* blastocitos, embriones, estados larvarios, y adultos). Ejemplos de células primarias adecuadas comprenden cardiomiocitos, hepatocitos primarios, fibroblastos, células neuronales, así como células madre. También en la presente invención hay líneas celulares estables derivadas de células primarias.

En algunas realizaciones, la célula huésped de la presente invención constituye parte de un organismo multicelular. En otras palabras, la invención también se refiere a organismos mamíferos transgénicos no humanos que comprenden al menos una célula huésped como se ha definido en el presente documento.

En la presente invención, el vector de expresión que se introduce se puede propagar y mantener en la célula huésped como una unidad genética independiente (es decir, episómicamente) (denominada también en el presente documento "transfección transitoria") o los fragmentos de vector pueden llegar a integrarse establemente en el genoma de la célula huésped por medio de recombinación genética (denominada también en el presente documento "transfección estable"). Tal recombinación se puede producir o bien en posiciones aleatorias del genoma por recombinación no homóloga o bien en posiciones específicas del genoma por recombinación homóloga o por medio de integrasas específicas del sitio. Preferentemente, los fragmentos del vector (incluyendo las secuencias heterólogas de ácido nucleico que se van a expresar) se integran en el genoma de la célula huésped como una copia sencilla.

Para introducir los vectores de expresión como se ha definido en el presente documento en una célula huésped de mamífero se puede emplear cualquier técnica de transfección que sea apropiada para el tipo particular de célula que se emplee. Se han establecido bien numerosos métodos de transfección en la técnica que incluyen *inter alia* electroporación, co-precipitación en fosfato cálcico, transfección química (por ejemplo ciclodextrina, DEAE-dextrano, polietilenimina), lipofección, magnetofección, y "pistola genética" (véase, por ejemplo, Sambrook, J., y Russel, D.W. (2001), *supra*; Ausubel, F.M. et al. (2001), *supra*).

En otro aspecto más, la presente invención se refiere a un método para la expresión heteróloga de una secuencia de ácido nucleico de interés en una célula huésped de mamífero, que comprende:

- (i) transfectar la célula huésped de mamífero con un vector de expresión como se ha definido anteriormente en el presente documento; y
- (ii) cultivar la célula huésped de mamífero transfectada bajo condiciones que permitan la expresión de la secuencia de ácido nucleico de interés.

En otras palabras, la presente invención se dirige también a un proceso para la producción recombinante (es decir, heteróloga) de polipéptidos o proteínas de interés en células huésped de mamífero que se han transfectado con un vector de expresión como se ha definido en el presente documento que comprende las correspondientes secuencias de ácido nucleico que codifican dichos polipéptidos o proteínas. La transfección se puede llevar a cabo con un solo vector de expresión, o como una co-transfección con dos o más vectores de expresión diferentes.

Como ya se ha resaltado anteriormente, hay muchos métodos disponibles para la transfección transitoria o estable con el fin de establecer que las células o líneas celulares expresen continuamente las secuencias heterólogas de ácido nucleico que codifican los polipéptidos o proteínas de interés.

En algunas realizaciones, el método comprende además la etapa de recolección (y opcionalmente de purificación) de los polipéptidos o proteínas recombinantes que se producen. Dependiendo de la naturaleza de dichos polipéptidos o proteínas puede segregarse en el sobrenadante del cultivo celular, integrarse en la membrana de la célula huésped, o permanecer en un compartimento intracelular.

Normalmente, si se emplea una célula huésped de mamífero unicelular el experto en la técnica puede devolverla a varias condiciones de cultivo celular que permitan la expresión de la secuencia de ácido nucleico de interés. Convenientemente, los polipéptidos o proteínas que se producen se recolectan (y opcionalmente se purifican) a partir del medio de cultivo, lisados o extractos de células cultivadas o de membranas aisladas (biológicas) por técnicas establecidas, tales como *inter alia* precipitación fraccionada con sales o disolventes orgánicos, cromatografía de intercambio iónico, cromatografía en gel, cromatografía de exclusión por tamaño, HPLC, cromatografía de afinidad (véase, por ejemplo, Sambrook, J., y Russel, D.W. (2001), *supra*). En el caso de que la célula huésped sea parte de un organismo multicelular, una fracción de estas células puede servir como fuente para aislar el péptido de la invención.

Las condiciones y medios de cultivo adecuados para las células huésped descritas anteriormente se conocen bien en la técnica (cf., por ejemplo, Fresney, R. (2000) *Culture of Animal cells. A manual* (4th Ed.) Wiley-Liss, New York). Dependiendo de los requisitos específicos de crecimiento de la célula huésped que se emplee, se puede llevar a cabo el cultivo de las células de mamífero, por ejemplo en medio RPMI 1640, medio F12 de Ham o DMEM (Medio Eagle Modificado de Dulbecco). De manera alternativa, se puede utilizar un medio de cultivo con una concentración de suero reducida, tal como OptiMEM. Se puede suplementar el medio opcionalmente con un 10 % (v/v) de FCS (suero fetal bovino), varios factores de crecimiento, aminoácidos, antibióticos y otros aditivos. Los medios de cultivo celular especialmente adaptados a células CHO se describen por ejemplo, en los documentos EP 0 481 791 B1 y EP 1 525 320 B1. Las células huésped de mamífero transfectadas se pueden incubar a 37 °C en una atmósfera con un 5 % de CO₂ y saturada de humedad. Los medios de cultivo respectivos, kits, y reactivos están disponibles comercialmente en varios suministradores.

En un aspecto más, la presente invención se refiere al uso de un vector de expresión como se define anteriormente en el presente documento para la expresión heteróloga *in vitro* de una secuencia de ácido nucleico de interés en una célula huésped de mamífero. Preferentemente, la secuencia de ácido nucleico de interés puede codificar un polipéptido o proteína que se pretende utilizar en aplicaciones diagnósticas o terapéuticas.

En realizaciones específicas, el vector de expresión se utiliza para la expresión concomitante de dos o más secuencias de ácido nucleico de interés que se insertan en el vector de expresión bajo el control de secuencias reguladoras de promotor quiméricas separadas. Por ejemplo, un vector de expresión se puede utilizar para la expresión de un gen de interés junto con un gen indicador para controlar la dirección a las células y/o la funcionalidad del gen de interés.

Preferentemente de manera particular, los vectores de expresión se utilizan para la expresión concomitante de dos o más secuencias de ácido nucleico de interés que codifican subunidades de una proteína dimerica o multimérica, por ejemplo, las cadenas ligera o pesada de una molécula de anticuerpo o subunidades de una vacuna. Empleando secuencias reguladoras de promotor quiméricas que dan como resultado diferentes tasas de expresión genética se puede utilizar un vector de expresión como se define en el presente documento para la expresión de dos o más secuencias de ácido nucleico de interés en una relación (molar) particular.

En una realización específica, los vectores de expresión como se definen en el presente documento se usan como medicamentos (o como partes de un medicamento o kit de partes) para terapia genética.

La invención se describe además por las figuras y los siguientes ejemplos, que solamente tienen el propósito de ilustrar realizaciones específicas de esta invención, y no son para que se consideren como que limitan el alcance de la invención de alguna manera.

Ejemplos

Base

Se generaron cinco secuencias reguladoras quiméricas distintas, como se especifica en las reivindicaciones, basándose en las secuencias de los genomas de citomegalovirus murino, humano y de simio (mCMV, hCMV, y sCMV) (véase la Tabla 1, Figura 3). Estas construcciones se analizaron en cuanto a su eficacia para controlar la expresión genética heteróloga de las cadenas ligera y pesada (LC y HC) de un anticuerpo monoclonal en células ováricas de hámster chino (CHO).

Ejemplo 1: Construcción del vector

Se utilizó síntesis genética para generar las cinco construcciones de promotor/amplificador quiméricas diferentes (es decir, las construcciones “1-5”, en donde las construcciones 1 y 3 no forman parte de la presente invención). Las construcciones se proporcionaban listas para clonar, por medio de integración específica del sitio (SSI), en el vector de dirección “vacío” pRY42 (cf. **Figura 1**, SEC ID N° 1), con el fin de reemplazar los promotores de citomegalovirus murino (mCMV) original (SEC ID N° 3) que están contenidos en él. El vector parental pRY42 comprende dos casetes de expresión, cada uno bajo el control de una secuencia reguladora de promotor (derivada originalmente del mCMV).

Se sintetizaron las construcciones quiméricas (cf. **Tabla 1**, **Figura 3**. SEC ID N° 6 a SEC ID N° 10) para incluir secuencias adicionales de ADN en los extremos 5' y 3' que flanquean las secuencias promotoras/amplificadoras, permitiendo de esta manera la incorporación de sitios de restricción de endonucleasas (es decir, también presentes en el pRY42) con el fin de facilitar el intercambio de fragmentos de ácido nucleico.

Tabla 1 que ilustra las distintas secuencias de elementos que están comprendidos en las construcciones promotoras/amplificadoras quiméricas 1-5 que se emplean en el presente documento, en donde las construcciones 1 y 3 no forman parte de la presente invención. Lo que se muestra son las respectivas longitudes y localizaciones genéticas de las secuencias reguladoras individuales, como se indica en la base de datos de Genomas víricos de la NCBI (<http://www.ncbi.nlm.nih.gov/genomes/GenomesHome>; Bao, Y. et al (2004) *J. Virol.* 78, 7291-7298). En particular, todas las secuencias promotoras de mCMV que se emplean en el presente documento incluyen un nucleótido guanosina (“G”) adicional en su extremo 3', que representa el sitio de inicio de la transcripción.

| Construcción | longitud de sCMV | Nº de registro del genoma vírico NCBI | Coordinados | longitud de hCMV | Nº de registro del genoma vírico NCBI | Coordinados |
|--------------|------------------|---------------------------------------|-------------|------------------|---------------------------------------|---------------|
| 1 | - | - | - | 582 | X17403.1 | 174292-174873 |
| 2 | - | - | - | 1026 | X17403.1 | 173848-174873 |
| 3 | - | - | - | 407 | X17403.1 | 173848-174254 |
| 4 | 452 | U38308.1 | 3873-4324 | 407 | X17403.1 | 173848-174254 |
| 5 | 807 | U38308.1 | 3873-4679 | - | - | - |

| Construcción | longitud de mCMV | Nº de registro del genoma vírico NCBI | Coordinados |
|--------------|------------------|---------------------------------------|---------------|
| 1 | 492 | U68299.1 | 182895-183386 |
| 2 | 102 | U68299.1 | 182895-182996 |
| 3 | 102 | U68299.1 | 182895-182996 |
| 4 | 102 | U68299.1 | 182895-182996 |
| 5 | 102 | U68299.1 | 182895-182996 |

Los nuevos vectores que comprenden las construcciones quiméricas 1-5 se construyeron cada uno mediante una reacción de ligadura de cuatro maneras de acuerdo con el esquema que se ilustra en la **Tabla 2**. La secuencia quimérica reguladora de promotor para la expresión de la secuencia de ácido nucleico de la cadena ligera del anticuerpo se insertó 3' (es decir, “corriente abajo”) del sitio FRT mutado (diana de reconocimiento de flipasa) (F5-mFRT) del pRY42, seguido por un fragmento intercatenario y la secuencia quimérica reguladora de promotor para la expresión de la secuencia de ácido nucleico de la cadena pesada.

Las reacciones de ligadura se llevaron a cabo utilizando el kit de ligadura de ADN rápido (Roche Diagnostics GmbH,

Mannheim, Alemania) de acuerdo con las instrucciones del fabricante. Los vectores plásmidos resultantes se transformaron en células de *E. coli* DH5 α competentes (Invitrogen/Life Technologies GmbH; Darmstadt, Alemania). Las muestras se colocaron en placas en agar Luria-Bertani (LB) suplementado con 50 μ g/ml de ampicilina y se incubaron durante una noche a 37 $^{\circ}$ C. Las colonias únicas se utilizaron para inocular matraces de 500 ml con agitado que contenían 200 ml de medio LB líquido más 50 μ g/ml de ampicilina. Los matraces se incubaron una noche en una incubadora con agitado a 37 $^{\circ}$ C, 200 rpm. Los cultivos resultantes se utilizaron para la preparación del plásmido ADN utilizando el kit Maxi Nucleobond (Macherey-Nagel GmbH, Düren, Alemania) de acuerdo con las instrucciones del fabricante. Los plásmidos producidos se verificaron por digestión de restricción diagnóstica.

- 10 **Tabla 2** que muestra el esquema de clonación para generar las construcciones de secuencia quimérica reguladora de promotor 1-5, en donde las construcciones 1 y 3 no forman parte de la presente invención.

| Construcción | Elemento regulador LC | Fragmento intercatenario (derivado de pRY42) | Elemento regulador HC | Armazón del vector (derivado de pRY42) |
|--------------|-----------------------|--|-----------------------|--|
| 1 | AatII-PacI | PacI-EcoRI | EcoRI-XhoI | XhoI-AatII |
| 2 | PvuI-SacII | SacII-EcoRI | EcoRI-Bsu36I | Bsu36I-PvuI |
| 3 | SspI-SacII | SacII-EcoRI | EcoRI-Bsu36I | Bsu36I-SspI |
| 4 | SspI-SacII | SacII-EcoRI | EcoRI-Bsu36I | Bsu36I-SspI |
| 5 | SspI-SacII | SacII-EcoRI | EcoRI-Bsu36I | Bsu36I-SspI |

- 15 En una etapa posterior, la LC y HC del anticuerpo monoclonal IgG4 cB72.3 (Whittle, N. et al. (1987) Protein Eng. 1, 499-505) se clonaron en cada uno de los cinco vectores que comprendían las construcciones quiméricas 1-5. Las secuencias de ácido nucleico que codifican las cadenas de anticuerpo se derivaron del vector pRY57 (**Figura 2**, SEC ID N $^{\circ}$ 2) como fragmentos de restricción EcoRI/HindIII (LC) y BamHI/NruI (HC) y se clonaron secuencialmente en los vectores para cada construcción promotora, utilizando las mismas endonucleasas de restricción. Los vectores plásmidos resultantes se verificaron por digestión de restricción diagnóstica y secuenciación del ADN de las regiones promotoras, respectivamente.

20 **Ejemplo 2: Transfección y construcción de la línea celular**

- 25 Se utilizó una línea celular derivada de la suspensión adaptada de la línea celular ovárica de hámster chino CHOK1SV para todos los experimentos (Lonza Ltd., Basilea, Suiza). En esta línea celular, las construcciones de ácido nucleico se insertaron en un locus genómico específico de la célula huésped en un único número de copias por medio de un sistema de integración específica del sitio (SSI).

- 30 La selección positiva para la integración se consigue por restauración funcional de un casete de resistencia a la higromicina B en el sitio SSI. La integración de la construcción de expresión en el genoma huésped también retira una copia del gen de timidina quinasa, lo que convierte el profármaco ganciclovir en un nucleótido análogo fosforilado tóxico. Por lo tanto, la adición de higromicina B y ganciclovir durante la construcción de la línea celular proporciona presiones de selección positivas y negativas, respectivamente, por la integración en el locus genómico diana.

- 35 La línea celular huésped se revivió a partir de viales crioconservados y se sub-cultivó. Todos los cultivos celulares: se llevaron a cabo en medio CD-CHO (Invitrogen/Life Technologies GmbH; Darmstadt, Alemania). Las células se sembraron 48 horas antes de la transfección en 30 ml de medio CD-CHO, a una concentración final de $0,3 \times 10^6$ células/ml.

- 40 El día de la transfección, se aglomeraron $1,2 \times 10^7$ células y se resuspendieron en 1 ml de CD-CHO antes de la co-transfección con 45 μ g del plásmido pOG44 (Invitrogen/Life Technologies GmbH; Darmstadt, Alemania) y 5 μ g de cada uno de los vectores dirigidos LC/HC (que comprendían las construcciones quiméricas 1-5, respectivamente) que se llevó a cabo utilizando electroporación en una cubeta BioRad GenePulser Xcell de 0,4 cm (pulso único de 300 V/900 μ F, a tiempo constante de 12-16 mseg). El plásmido pOG44 engloba un casete de expresión para la recombinasa (flipasa) de levaduras que se necesita para facilitar la recombinación en los sitios FRT presentes en el locus diana. Todos los experimentos de transfección se llevaron a cabo al menos por duplicado.

- 50 Cada muestra de electroporación se transfirió a 20 ml de medio CD-CHO en un matraz T75 (BD Biosciences, Heidelberg, Alemania) y se incubó en modo estático a 36,5 $^{\circ}$ C, en una incubadora humidificada (5 % (v/v) de CO $_2$ en el aire). Las células se aglomeraron 48 h post-transfección (150 x g, 5 min) y se resuspendieron en 20 ml de medio CD-CHO que contenía 200 μ g/ml de higromicina B (selección positiva). El cultivo se mantuvo en modo estático y tras 72 h se cambió el medio con CD-CHO recién preparado que contenía 200 μ g/ml de higromicina B.

Posteriormente; cada 72 h se determinó la concentración celular viable y se cambió el medio con 20 ml de CD-CHO

recién preparado que contenía 200 µg/ml de higromicina B y ,3 µM de ganciclovir (selección negativa). Una vez que el cultivo en los matraces T75 alcanzaba una concentración celular total de 9×10^6 células/ml, se ajustó el volumen final a 30 ml de CD-CHO que contenía 200 µg/ml de higromicina B y 3 µM de ganciclovir. Cada cultivo diluido se transfirió entonces a un matraz E125 con agitado.

5 **Ejemplo 3: Cultivo de sobrecrecimiento con alimentación discontinua (FOG) en suspensión para determinar la concentración de anticuerpo monoclonal que se produce**

10 El análisis del sobrecrecimiento con alimentación discontinua (FOG) en matraces con agitado se llevó a cabo como se describe en la publicación internacional de patente WO 2008/148519 A2. Todos los experimentos FOG para una determinada suspensión de células transfectadas se llevaron a cabo al menos por duplicado.

15 En resumen, se sembraron las células transfectadas a una concentración de 2×10^5 células/ml en matraces de 250 ml con agitado, que contenía cada uno 50 ml de medio de cultivo CM42/SPE (Lonza Ltd., Basilea, Suiza) y se incubaron a 37 °C en una incubadora humidificada orbital (5 % (v/v) de CO₂ en el aire) con agitado a 140 rpm. Las células se alimentaron, comenzando el día 3 del cultivo, con una alimentación que consistía en una mezcla de aminoácidos y elementos traza. Se determinaron las viabilidades diarias y las concentraciones celulares viables utilizando un Analizador de Viabilidad Celular Automático Cedex (Roche Diagnostics GmbH, Mannheim, Alemania). Se determinó la concentración de anticuerpo en el medio por HPLC-proteína A el día 15 del cultivo (recolectado de los cultivos de "sobrecrecimiento").

20 **Ejemplo 4: Determinación de la concentración del anticuerpo monoclonal que se produce por medio de HPLC-Proteína A**

25 Las concentraciones del anticuerpo monoclonal (mAb) IgG4 cB72.3 que se producía por las respectivas líneas celulares que albergaban los casetes de expresión LC/HC bajo el control de las diferentes construcciones quiméricas 1-5 y que se segregan en el medio de cultivo celular se determinaron por cromatografía de altas prestaciones (HPLC)-proteína A. Los sobrenadantes libres de células (que se pasaron a través de una unidad de filtro de 0,22 µm) se cargaron en una columna de inmunodetección de Proteína A POROS (applied Biosystems Inc., Foster City, Ca, USA), conectado a un HPLC Agilent 1100. Se lavó la columna y se eluyó el mAb unido bajando el pH del disolvente.

30 La concentración del mAb se determinó por comparación con una curva de referencia generada con diluciones seriadas de cB72.3 IgG4 MabSelect SuRe-purificado (GE Healthcare GmbH, Freiburg, Alemania) (Intervalo de la curva de referencia: 1025 ng/µl a 200 ng/ml).

35 **Ejemplo 5: Resultados**

40 Los resultados de los experimentos anteriores se resumen en la **Tabla 3 y la Figura 4**, respectivamente: para cada una de las construcciones quiméricas que se utilizan en el presente documento, $n = 4$, representa los análisis FOG por duplicado para las transfecciones duplicadas, con la excepción de la construcción 4, en donde $n = 6$ (análisis FOG por duplicado para las transfecciones por triplicado) y pRY57 (mCMV original), en donde, $n = 8$, es por combinación de los puntos de datos de los experimentos 1 y 2. Los cálculos de los parámetros del cultivo celular se llevaron a cabo como había descrito anteriormente Porter et al (2010) *Biotechnol. Progr.* 26, 1446-1454).

45 **Tabla 3** que ilustra las tasas de crecimiento y las cantidades de mAb producidas por las diferentes líneas celulares transfectadas que se emplean en el presente documento (que comprenden las secuencias/construcciones reguladoras de promotor quiméricas 1-5, en donde las construcciones 1 y 3 no forman parte de la presente invención).

50

| Grupo de transfectantes / FOG | | Máx. (10^6 células/ml) | m (1/h) | IVC (10^6 células/ml) | ρ_P (pg/cél.h) | [mAb] (mg/l) |
|-------------------------------|----------------|---------------------------|---------------------|--------------------------|---------------------|---------------------|
| Experimento 1 | mCMV | $9,36 \pm 1,02$ | $0,0183 \pm 0,0018$ | $1523,35 \pm 164,78$ | $0,29 \pm 0,02$ | $439,74 \pm 23,51$ |
| | Construcción 1 | $8,5 \pm 0,75$ | $0,0206 \pm 0,0013$ | $1445,11 \pm 70,55$ | $0,46 \pm 0,07$ | $638,87 \pm 110,18$ |
| | Construcción 2 | $7,97 \pm 0,41$ | $0,0198 \pm 0,0008$ | $1409,29 \pm 87,92$ | $0,71 \pm 0,06$ | $883,10 \pm 109,78$ |
| | Construcción | $11,17 \pm 1,02$ | $0,0191 \pm 0,0009$ | $1873,58 \pm 250,14$ | $0,40 \pm 0,03$ | $707,98 \pm 80,28$ |

| | | | | | | |
|---------------|----------------|--------------|-----------------|------------------|-------------|------------------|
| Experimento 2 | mCMV | 9,29 ± 0,82 | 0,0143 ± 0,0013 | 1790,34 ± 104,97 | 0,29 ± 0,05 | 539,31 ± 76,24 |
| | Construcción 4 | 10,39 ± 0,90 | 0,0165 ± 0,0023 | 1848,05 ± 306,29 | 0,73 ± 0,22 | 1408,44 ± 337,47 |
| | Construcción 5 | 11,58 ± 0,46 | 0,0139 ± 0,0028 | 2261,28 ± 50,84 | 0,46 ± 0,03 | 1244,15 ± 74,90 |

Leyenda: Máx. - Concentración celular máxima (10⁶ células/ml); μ - tasa de crecimiento específica (1/h); IVC - tiempo integral de la concentración celular viable (106 células h/ml); p_p - tasa de producción específica del mAb (pg/cél·h); y [mAb] - concentración del mAb producido en la recolección (mg/l)

A partir de la **Figura 4**, es evidente que el día de la recolección del cultivo (es decir, el día 15), el uso de una cualquiera de las construcciones quiméricas n^{os} 2-5 (no siendo la construcción 3 parte de la presente invención) resulta en la producción de concentraciones de anticuerpos más altas que con el uso de la secuencia promotora original de mCMV (es decir, el vector pRY57). El uso de la construcción quimérica n^o 1 también daba como resultado una concentración de anticuerpo más alta que con el promotor mCMV (un factor de 1,31), incluso aunque el resultado no alcanzaba una significación estadística.

Los mejores resultados se obtuvieron con la construcción quimérica n^o 4 que daba como resultado una expresión genética de aproximadamente 2,88 veces más alta en comparación con el promotor de mCMV, seguida (en orden descendente) por la construcción quimérica n^o 5 (factor de aproximadamente 2,54), la construcción quimérica n^o 2 (factor de aproximadamente 1,80), y la construcción quimérica n^o 3 (factor de aproximadamente 1,45; que no forma parte de la presente invención).

Los datos anteriores muestran que el uso de secuencias reguladoras de promotor quiméricas que comprenden elementos amplificadores E11 de sCMV y/o hCMV en combinación con elementos promotores IE1 de mCMV como se definen en el presente documento representan herramientas genéticas superiores para obtener sistemas de expresión genética heterólogos altamente eficaces para células de mamífero.

20 LISTADO DE SECUENCIAS

<110> Lonza Biologics PLC

<120> Vectores de expresión que comprenden secuencias quiméricas de promotor y amplificador de citomegalovirus

<130> 827-1

<140> EP12185728.8

<141> 24-09-2012

<160> 12

<170> BiSSAP Versión 1.1

<210> 1

<211> 5908

<212> ADN

<213> Secuencia artificial

<220>

<223> vector plasmídico pRY42

<220>

<221 > polyA_signal

<222> (7)...(245)

<223> señal de polyA de SV40

<220>

<221 > promotor

<222> (266)...(756)

<223> fragmento de promotor mCMV-MIE (IE1)

ES 2 540 753 T3

- 5 <220>
<221 > 5'UTR
<222> (758)...(791)
<223> parte del exón 1 de mCMV-MIE (IE1) 5'UTR
- 10 <220>
<221 > 5'UTR
<222> (803)...(884)
<223> parte del exón 1 de hCMV-MIE (IE1) 5'UTR
- 15 <220>
<221 > intrón
<222> (885)...(1711)
<223> intrón A de hCMV-MIE (IE1)
- 20 <220>
<221 > 5'UTR
<222> (1712)...(1729)
<223> exón 2 de hCMV-MIE (IE1) 5'UTR
- 25 <220>
<221 > polyA_signal
<222> (1767)...(2005)
<223> señal de polyA de SV40
- 30 <220>
<221 > promotor
<222> (2019)..(2345)
<223> promotor temprano de SV40, derivado de pFRT/lacZeo (Invitrogen)
- 35 <220>
<221 > fuente
<222> (2350)...(2354)
<223> secuencia que contiene el sitio de ATG encontrado cadena arriba del FRT de tipo silvestre en pFRT/lacZeo (Invitrogen)
- 40 <220>
<221 > misc_recomb
<222> (2355)...(2402)
<223> sitio de diana de reconocimiento de flipasa de tipo silvestre (FRT), derivado de pcDNA5-Frt (Invitrogen)
- 45 <220>
<221 > fuente
<222> (3183)...(4043)
<223> complemento inverso CDS del gen de beta-lactamasa
- 50 <220>
<221 > misc_recomb
<222> (4319)...(4366)
<223> sitio de diana de reconocimiento de flipasa (FRT) mutante
- 55 <220>
<221 > promotor
<222> (4410)...(4900)
<223> fragmento de promotor de mCMV-MIE (IE1)
- 60 <220>
<221 > 5'UTR
<222> (4902)...(4935)
<223> parte de exón 1 de mCMV-MIE (IE1) 5'UTR
- <220>
<221 > 5'UTR

ES 2 540 753 T3

<222> (4947)...(5028)
 <223> parte de exón 1 de hCMV-MIE (IE1) 5'UTR

5 <220>
 <221 > intrón
 <222> (5029)...(5855)
 <223> intrón A de hCMV-MIE (IE1)

10 <220>
 <221 > 5'UTR
 <222> (5856)...(5873)
 <223> exón 2 de hCMV-MIE (IE1)5'UTR

<400> 1

| | | | | | | | |
|----|------------|-------------|-------------|-------------|-------------|-------------|------|
| | gaattcattg | atcataatca | gccataccac | atttgtagag | gttttacttg | ctttaaaaaa | 60 |
| | cctcccacac | ctccccctga | acctgaaaca | taaaatgaat | gcaattgttg | ttgttaactt | 120 |
| | gtttattgca | gcttataatg | gttacaaata | aagcaatagc | atcacaaatt | tcacaaataa | 180 |
| | agcatttttt | tcactgcatt | ctagttgtgg | tttgtccaaa | ctcatcaatg | tatcttatca | 240 |
| | tgtctggcgg | ccgctgaggc | gcgcctactg | agtcattagg | gactttccaa | tgggttttgc | 300 |
| | ccagtacata | aggtaaatag | gggtgaatca | acaggaaagt | cccattggag | ccaagtacac | 360 |
| | tgagtcaata | gggactttcc | attgggtttt | gcccagtaca | aaagggtcaat | aggggggtgag | 420 |
| | tcaatggggt | tttcccatta | ttggcacgta | cataagggtca | ataggggtga | gtcattgggt | 480 |
| | ttttccagcc | aatttaatta | aaacgccatg | tactttccca | ccattgacgt | caatgggcta | 540 |
| | ttgaaactaa | tgcaacgtga | cctttaaacg | gtactttccc | atagctgatt | aatgggaaag | 600 |
| | taccgttctc | gagccaatac | acgtcaatgg | gaagtgaaag | ggcagccaaa | acgtaacacc | 660 |
| | gccccggttt | ttcccctggaa | attccatatt | ggcacgcatt | ctattggctg | agctgcgttc | 720 |
| | tacgtgggta | taagaggcgc | gaccagcgtc | ggtaccgctg | cagtcttcgg | tctgaccacc | 780 |
| | gtagaacgca | gcctcaggac | ctccatagaa | gacaccggga | ccgatccagc | ctccgcggcc | 840 |
| | gggaacggtg | cattggaacg | cggattcccc | gtgccaagag | tgacgtaagt | accgcctata | 900 |
| | gagtctatag | gcccaccccc | ttggcttctt | atgcatgcta | tactgttttt | ggcttggggg | 960 |
| | ctatacacc | ccgcttccctc | atggttatagg | tgatggtata | gcttagccta | taggtgtggg | 1020 |
| | ttattgacca | ttattgacca | ctcccctatt | ggtgacgata | ctttccatta | ctaattccata | 1080 |
| | acatggctct | ttgccacaac | tctctttatt | ggctatatgc | caatacactg | tccttcagag | 1140 |
| | actgacacgg | actctgtatt | tttacaggat | ggggtctcat | ttattattta | caaattcaca | 1200 |
| | tatacaaac | caccgtcccc | agtgccccga | gtttttatta | aacataacgt | gggatctcca | 1260 |
| | cgcaaatctc | gggtacgtgt | tccggacatg | ggctcttctc | cggtagcggc | ggagcttcta | 1320 |
| | catccgagcc | ctgctcccat | gcctccagcg | actcatggtc | gctcggcagc | tccttgctcc | 1380 |
| | taacagtgga | ggccagactt | aggcacagca | cgatgccac | caccaccagt | gtgccgcaca | 1440 |
| 15 | aggccgtggc | ggtaggggat | gtgtctgaaa | atgagctcgg | ggagcgggct | tgcaccgctg | 1500 |

ES 2 540 753 T3

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| acgcatttgg | aagacttaag | gcagcggcag | aagaagatgc | aggcagctga | gttggtgtgt | 1560 |
| tctgataaga | gtcagaggta | actcccgttg | cggtgctggt | aacgggtggag | ggcagtgtag | 1620 |
| tctgagcagt | actcgttgct | gccgcgcgcg | ccaccagaca | taatagctga | cagactaaca | 1680 |
| gactgttctt | ttccatgggt | cttttctgca | gtcaccgtcc | ttgacacggg | atccggcgcg | 1740 |
| cccctagggg | taccgtcgac | tcgcgaattg | atcataatca | gccataccac | atttgtagag | 1800 |
| gttttacttg | ctttaaaaaa | cctcccacac | ctccccctga | acctgaaaca | taaaatgaat | 1860 |
| gcaattgttg | ttgttaactt | gtttattgca | gcttataatg | gttacaataa | aagcaatagc | 1920 |
| atcacaaatt | tcacaaataa | agcatttttt | tcactgcatt | ctagttgtgg | tttgtccaaa | 1980 |
| ctcatcaatg | tatcttatca | tgtctggatc | agcttgagca | gctgtggaat | gtgtgtcagt | 2040 |
| taggggtgtg | aaagtcecca | ggctccccag | caggcagaag | tatgcaaagc | atgcatctca | 2100 |
| attagtcagc | aaccagggtg | ggaaagtccc | caggctcccc | agcaggcaga | agtatgcaaa | 2160 |
| gcatgcatct | caattagtca | gcaacctatg | tcccgcacct | aactccgccc | atcccgcacc | 2220 |
| taactccgcc | cagttccgcc | cattctccgc | cccatggctg | actaattttt | tttatattatg | 2280 |
| cagaggccga | ggccgcctcg | gcctctgagc | tattccagaa | gtagtgagg | ggcttttttg | 2340 |
| gaggctacca | tgagagaagt | actattccga | agttcctatt | ctctagaaag | tataggaact | 2400 |
| tctcggggcg | cggtgctggc | gtttttccat | aggctccgcc | cccctgacga | gcatccaaaa | 2460 |
| aatcgacgct | caagtccagag | gtggcgaaac | ccgacaggac | tataaagata | ccaggcgttt | 2520 |
| ccccctggaa | gctcccctct | gcgctctcct | gttccgaccc | tgccgcttac | cggataacctg | 2580 |
| tccgcctttc | tcccttcggg | aagcgtggcg | ctttctcata | gctcacgctg | taggtatctc | 2640 |
| agttcggtgt | aggctggtcg | ctccaagctg | ggctgtgtgc | acgaaccccc | cgttcagccc | 2700 |
| gaccgctgcg | ccttatcccg | taactatcgt | cttgagtcca | acccggtaag | acacgactta | 2760 |
| tccgactggg | cagcagccac | tggtaacagg | attagcagag | cgaggtatgt | aggcgggtct | 2820 |
| acagagtctt | tgaagtgggt | gcctaactac | ggctacacta | gaagaacagt | atttggatct | 2880 |
| tgctctctgc | tgaagccagt | taccttcgga | aaaagagtgt | gtagctcttg | atccggcaaa | 2940 |
| caaaccaccg | ctggtagcgg | tggttttttt | gtttgcaagc | agcagattac | gcbgagaaaa | 3000 |
| aaaggatctc | aagaagatcc | ttgatctttt | tctacggggg | ctgacgctca | gtggaacgaa | 3060 |
| aactcacggt | aagggatttt | ggctcatgaga | ttatcaaaaa | ggatcttcac | ctagatcctt | 3120 |
| ttaaattaaa | aatgaagttt | taaatacaat | taaagtatat | atgagtaaac | ttggtctgac | 3180 |
| agttaccaat | tgaggcacct | tgaggcacct | atctcagcga | tctgtctatt | tcgttcatcc | 3240 |
| atagttgcct | gactccccgt | cgtgtagata | actacgatac | gggagggcct | accatctggc | 3300 |
| cccagtgctg | caatgatacc | gcgagaccca | cgctcacccg | ctccagattt | atcagcaata | 3360 |
| aaccagccag | ccggaagggc | cgagcgcaga | agtggctcct | caactttatc | cgctccatc | 3420 |
| cagctctatta | attgttgccg | ggaagctaga | gtaagtatgt | cgccagttaa | tagtttgccg | 3480 |
| aacgttgttg | ccattgctac | aggcatcgtg | gtgtcacgct | cgctgtttgg | tatggcttca | 3540 |
| ttcagctccg | gttcccacag | atcaaggcga | gttacatgat | ccccatggtt | gtgcaaaaaa | 3600 |
| gcggttagct | ccttcggtec | tccgatcgtt | gtcagaagta | agttggccgc | agtgttatca | 3660 |
| ctcatggtta | tggcagcact | gcataattct | cttactgtca | tgccatccgt | aagatgcctt | 3720 |
| tctgtgactg | gtgagtactc | aaccaagtca | ttctgagaat | agtgtatgcg | gcbgaccgag | 3780 |
| tgctcttgcc | cggcgtcaat | acgggataat | accgcgccac | atagcagaac | tttaaaagtg | 3840 |
| ctcatcattg | gaaaacgttc | ttcggggcga | aaactctcaa | ggatcttacc | gctgttgaga | 3900 |
| tccagttcga | ctaateccac | tcgtgcaccc | aactgatctt | cagcatcttt | tactttcacc | 3960 |
| agcgtttctg | ggtgagcaaa | aacaggaagg | caaaatgccg | caaaaaagg | aataagggcg | 4020 |
| acacggaaat | gttgaatact | catactcttc | ctttttcaat | attattgaag | catttatcag | 4080 |
| ggttattgtc | tcatgagcgg | atacatattt | gaatgtattt | agaaaaataa | acaaataggg | 4140 |
| gttccgcgca | catttcccgc | aaaagtcca | ctgacgtctt | aagaaacat | tattatcag | 4200 |
| acattaacct | ataaaaaatg | gcgtatcacg | aggccctgat | ggctctttgc | ggcacccatc | 4260 |
| gttcgtaatg | ttccgtggca | ccgaggacaa | ccctcaagag | aaaatgtaat | cacactggga | 4320 |
| agttcctatt | ccgaagttcc | tattcttcaa | aaggtatagg | aacttctctg | agtgaataat | 4380 |
| aaaatgagtg | tttgtccgaa | atagcgcctt | actgagtcac | tagggacttt | ccaatgggtt | 4440 |
| ttgcccagta | cataaggctca | ataggggtga | atcaacagga | aagtcccat | ggagccaagt | 4500 |
| acactgagtc | aatagggact | ttccattggg | ttttgcccag | tacaaaagg | caataggggg | 4560 |
| tgagtcaatg | ggtttttccc | attattggca | cgtaacataag | gtcaataggg | gtgagtcatt | 4620 |
| gggtttttcc | agccaattta | attaaaacgc | catgtacttt | cccaccattg | acgtcaatgg | 4680 |
| gctattgaaa | ctaattgcaac | gtgaccttta | aacggtactt | tcccatagct | gattaatggg | 4740 |
| aaagtaccgt | tctcgagcca | atacacgtca | atgggaagtg | aaagggcagc | caaaacgtaa | 4800 |
| caccgccccg | gttttcccct | ggaaattcca | tattggcacg | cattctattg | gctgagctgc | 4860 |
| gttctacgtg | ggtataagag | gcgagaccag | cgctcgttac | gtcgcagctc | tcggtctgac | 4920 |
| caccgtagaa | gctagcctca | ggacctccat | agaagacacc | gggaccgatc | cagcctccgc | 4980 |
| ggccgggaac | ggtgcattgg | aacgcggatt | ccccgtgcca | agagtgcagt | aagtaccgcc | 5040 |
| tatagagtct | ataggcccac | ccccttggct | tcttatgcat | gctatactgt | ttttggcttg | 5100 |
| gggtctatac | acccccgctt | cctcatgta | taggtgatgg | tatagcttag | cctataggtg | 5160 |
| tgggttatgg | accattatgg | accactccc | tattgggtac | gatactttcc | attactaatc | 5220 |
| cataaacatg | ctctttgcca | caactctctt | tattggctat | atgccaatac | actgtccttc | 5280 |
| agagactgac | acggactctg | tatttttaca | ggatggggtc | tcatttatta | tttacaat | 5340 |
| cacatataca | acaccaccgt | ccccagtgcc | cgcagttttt | attaaacata | acgtgggatc | 5400 |
| tccacgcgaa | tctcgggtac | gtgttccgga | catgggctct | tcctcggtag | cggcggagct | 5460 |
| tctacatccg | agccctgctc | ccatgcctcc | agcgactcat | ggtcgtcgg | cagctccttg | 5520 |
| ctcctaacag | tgagggccag | acttaggcac | agcacgatgc | ccaccaccac | cagtgtgccg | 5580 |
| cacaaggccg | tggcggtagg | gtatgtgtct | gaaaatgagc | tcggggagcg | ggcttgacc | 5640 |

ES 2 540 753 T3

```

gctgacgcat ttggaagact taaggcagcg gcagaagaag atgcaggcag ctgagttggt      5700
gtgttctgat aagagtcaga ggtaactccc gttgCGgtgc tgTTaacggt ggagggcagt      5760
gtagtctgag cagtactcgt tgctgCCcgg cgcGCCacca gacataatag ctgacagact      5820
aacagactgt tcctttccat gggTcTtttc tgcagtcacc gTccttgaca cgaagcttac      5880
cggtagatct gctagcacat gtaggcct                                         5908

```

5 <210> 2
 <211> 7948
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> vector plasmídico de pRY57

15 <220>
 <221> CDS
 <222> (67)...(711)
 <223> gen optimizado cB72.3 de cadena ligera kappa

20 <220>
 <221> polyA_signal
 <222> (720)...(958)
 <223> Señal de polyA de SV40

25 <220>
 <221> promotor
 <222> (979)...(1470)
 <223> fragmento de promotor de mCMV-MIE (IE1)

30 <220>
 <221> 5'UTR
 <222> (1471)...(1504)
 <223> parte de exón 1 de mCMV-MIE (IE1) 5'UTR

35 <220>
 <221> 5'UTR
 <222> (1516)...(1601)
 <223> parte de exón 1 de hCMV-MIE (IE1) 5'UTR

40 <220>
 <221> intrón
 <222> (1598)...(2424)
 <223> intrón A de hCMV-MIE (IE1)

45 <220>
 <221> 5'UTR
 <222> (2425)...(2441)
 <223> exón 2 de hCMV-MIE (IE1) 5'UTR

50 <220>
 <221> CDS
 <222> (2508)...(3812)
 <223> gen optimizado cB72.3 de cadena pesada gamma-4

55 <220>
 <221> polyA_signal
 <222> (3843)...(4081)
 <223> Señal de polyA de SV40

<220>
 <221> promotor
 <222> (4095)...(4421)
 <223> Promotor temprano de SV40, derivado de pFRT/lacZeo (Invitrogen)

ES 2 540 753 T3

- <220>
<221 > fuente
<222> (4426)...(4430)
5 <223> secuencia que contiene el sitio de ATG encontrado cadena arriba del FRT de tipo silvestre en pFRT/lacZeo (Invitrogen)
- <220>
<221 > misc_recomb
10 <222> (4431)...(4478)
<223> sitio de diana de reconocimiento de flipasa de tipo silvestre (FRT), derivado de pcDNA5-Frt (Invitrogen)
- <220>
<221 > fuente
15 <222> (5259)...(6119)
<223> complemento inverso CDS del gen de beta-lactamasa
- <220>
<221 > misc_recomb
20 <222> (6395)...(6442)
<223> sitio de diana de reconocimiento de flipasa (FRT) mutante
- <220>
<221 > promotor
25 <222> (6486)...(6977)
<223> fragmento de promotor de mCMV-MIE (IE1)
- <220>
<221 > 5'UTR
30 <222> (6978)...(7011)
<223> parte de exón 1 de mCMV-MIE (IE1) 5'UTR
- <220>
<221 > 5'UTR
35 <222> (7023)...(7108)
<223> parte de exón 1 de hCMV-MIE (IE1) 5'UTR
- <220>
<221 > intrón
40 <222> (7105)...(7931)
<223> intrón A de hCMV-MIE (IE1)
- <220>
<221 > 5'UTR
45 <222> (7932)...(7948)
<223> exón 2 de hCMV-MIE (IE1)5'UTR
- <400> 2

ES 2 540 753 T3

| | | | | | | | | |
|------------|------------|-------------|-------------|------------|------------|---------|---------|-----|
| aagcttgccg | ccaccatgat | gctggcctatc | gtgctgggtgc | tgctgttcgc | cacctctgcc | 60 | | |
| ctggcc | gac atc | cag atg | acc cag | tcc ccc | gcc tcc | ctg tct | gtg tcc | 108 |
| | Asp Ile | Gln Met | Thr Gln | Ser Pro | Ala Ser | Leu Ser | Val Ser | |
| | 1 | | 5 | | 10 | | | |
| gtg ggc | gag aca | gtg acc | atc acc | tgt cgg | gcc tcc | gag aac | atc tac | 156 |
| Val Gly | Asp Thr | Val Thr | Ile Thr | Cys Arg | Ala Ser | Glu Asn | Ile Tyr | |
| 15 | | 20 | | | 25 | | 30 | |
| tcc aac | ctg gcc | tgg tat | cag cag | aag cag | ggc aag | tcc cct | cag ctg | 204 |
| Ser Asn | Leu Ala | Trp Tyr | Gln Gln | Lys Gln | Gly Lys | Ser Pro | Gln Leu | |
| | | 35 | | | 40 | | 45 | |
| ctg gtg | tac gcc | gcc acc | aac ctg | gct gac | ggc gtg | ccc tcc | agg ttc | 252 |
| Leu Val | Tyr Ala | Ala Thr | Asn Leu | Ala Asp | Gly Val | Pro Ser | Arg Phe | |
| | 50 | | | 55 | | 60 | | |
| tcc ggc | tct ggc | tcc ggc | acc cag | tac tcc | ctg aag | atc aac | tcc ctg | 300 |
| Ser Gly | Ser Gly | Ser Gly | Thr Gln | Tyr Ser | Leu Lys | Ile Asn | Ser Leu | |
| | 65 | | 70 | | | 75 | | |
| cag tcc | gag gac | ttc ggc | tcc tac | tac tgc | cag cac | ttc tgg | ggc acc | 348 |
| Gln Ser | Glu Asp | Phe Gly | Ser Tyr | Tyr Cys | Gln His | Phe Trp | Gly Thr | |
| | 80 | | 85 | | 90 | | | |
| cct tac | acc ttc | ggc gga | ggc acc | cgg ctg | gaa atc | aag cgg | acc gtg | 396 |
| Pro Tyr | Thr Phe | Gly Gly | Gly Thr | Arg Leu | Glu Ile | Lys Arg | Thr Val | |
| | | 100 | | | 105 | | 110 | |
| gcc gct | cct tcc | gtg ttc | atc ttc | cca cct | tcc gac | gag cag | ctg aag | 444 |
| Ala Ala | Pro Ser | Val Phe | Ile Phe | Pro Pro | Ser Asp | Glu Gln | Leu Lys | |
| | | 115 | | 120 | | 125 | | |

ES 2 540 753 T3

| | |
|---|------|
| tcc ggc acc gcc tct gtg gtg tgc ctg ctg aac aac ttc tac cct cgg | 492 |
| Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg | |
| gag gcc aag gtg cag tgg aag gtg gac aac gcc ctg cag agc ggc aac | 540 |
| Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn | |
| tcc cag gaa tcc gtc acc gag cag gac tcc aag gac tct acc tac tcc | 588 |
| Ser Gln Asp Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser | |
| ctg tcc tcc acc ctg acc ctg tcc aag gcc gac tac gag aag cac aag | 636 |
| Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys | |
| gtg tac gcc tgc gaa gtg acc cac cag ggc ctg tcc agc cct gtg acc | 684 |
| Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr | |
| aag tcc ttc aac cgg ggc gag tgc tga tag aattcattga tcataatcag | 734 |
| Lys Ser Phe Asn Arg Gly Cys | |
| ccataccaca tttgtagagg ttttacttgc tttaaaaaac ctcccacacc tccccctgaa | 794 |
| cctgaaacat aaaatgaatg caattgttgt tgtaacttgc tttattgcag cttataatgg | 854 |
| ttacaaataa agcaatagca tcacaaattt cacaaataaa gcattttttt cactgcattc | 914 |
| tagttgtgtt ttgtccaaac tcataaatgt atcttatcat gtctggcggc cgctgaggcg | 974 |
| cgctactgga gtcattaggg actttccaat gggttttgcc cagtacataa ggtcaatagg | 1034 |
| ggtgaatcaa caggaaagtc ccattggagc caagtacact gagtcaatag ggactttcca | 1094 |
| ttgggttttg cccagtacaa aaggtcaata ggggggtgagt caatgggttt tccccattat | 1154 |
| tggcacgtac ataaggtcaa taggggtgag tcattggggtt ttccagcca atttaattaa | 1214 |
| aacgccatgt actttcccac cattgacgtc aatgggctat tgaaactaat gcaacgtgac | 1274 |
| ctttaaaccg tactttccca tagctgatta atgggaaagt accgttctcg agccaataca | 1334 |
| cgtaaatggg aagtgaaagg gcagcaaaa cgtaacaccg ccccgtttt cccctggaaa | 1394 |
| ttccatattg gcacgcattc tattggctga gctgcgttct acgtgggtat aagaggcgcg | 1454 |
| accagctcg gtaccgtcgc agtcttcggt ctgaccaccg tagaacgcag cctcaggacc | 1514 |
| tccatagaag acaccgggac cgatccagcc tccgcggccg ggaacggtgc attggaacgc | 1574 |
| ggattccccg tgccaagagt gacgtaagta ccgcctatag agtctatagg cccaccccct | 1634 |
| tggcttctta tgcattgctat actgtttttg gcttgggggtc tatacacccc cgcttctca | 1694 |
| tgttataggt gatggtatag cttagcctat aggtgtgggt tattgacct tattgaccac | 1754 |
| tcccctattg gtgacgatac tttccattac taatccataa catggctctt tgccacaact | 1814 |
| ctctttattg gctatatgcc aatacactgt ccttcagaga ctgacacgga ctctgtattt | 1874 |
| ttacaggatg gggctctcatt tattatttac aaattccacat atacaacacc accgtcccca | 1934 |
| gtgccgcag tttttattaa acataacgtg ggatctccac ggatctctac atccgagccc | 1994 |
| ccggacatgg gctctttctc ggtagcggcg gagcttctac atccgagccc tgctcccatg | 2054 |
| ccaccagcga ctcattggtcg ctcggcagct ccttgctcct aacagtggag gccagactta | 2114 |
| ggcacagcac gatgccacc accaccagtg tgccgcacaa ggccgtggcg gtaggggatg | 2174 |
| tgctgaaaaa tgagctcggg gagcgggctt gcaccgctga cgcatttga agacttaagg | 2234 |
| cagcggcaga agaagatgca ggcagctgag ttgttgtgtt ctgataagag tcagaggtaa | 2294 |
| ctccggttgc ggtgctgtta acggtggagg gcagtgtag ctgagcagta ctctgttctg | 2354 |
| ccgcgcgcgc caccagacat aatagctgac agactaacag actgttccct tccatgggtc | 2414 |
| ttttctgag tcaccgtcct tgacacggga tccgcccca ccatgatgcg gcctatcgtc | 2474 |
| ctgggtgctg tgttcgccac aagcgcctctg gct cag gtg cag ctg cag cag agc | 2528 |
| Gln Val Gln Leu Gln Gln Ser | |
| gac gcc gag ctg gtg aag cct ggc gct agc gtg aag atc agc tgc aag | 2576 |
| Asp Ala Glu Leu Val Lys Pro Gly Ala Ser Val Lys Ile Ser Cys Lys | |
| gcc agc ggc tac acc ttc acc gat cac gcc atc cac tgg gct aag cag | 2624 |
| Ala Ser Gly Tyr Thr Phe Thr Asp His Ala Ile His Trp Ala Lys Gln | |
| aag ccc gag cag ggc ctg gag tgg atc ggc tac atc agc ccc ggc aac | 2672 |
| Lys pro Glu Gln Gly Leu Glu Trp Ile Gly Tyr Ile Ser Pro Gly Asn | |
| gac gac atc aag tac aac gag aag ttc aag ggc aag gcc acc ctg acc | 2720 |
| Asp Asp Ile Lys Tyr Asn Glu Lys Phe Lys Gly Lys Ala Thr Leu Thr | |
| gcc gac aag agc agc acc gcc tac atg cag ctg aac agc ctg acc | 2768 |
| Ala Asp Lys Ser Ser Thr Ala Tyr Met Gln Leu Asn Ser Leu Thr | |
| agc gag gac agc gcc gtg tac ttc tgc aag cgg agc tac tac ggc cac | 2816 |
| Ser Glu Asp Ser Ala Val Tyr Phe Cys Lys Arg Ser Tyr Tyr Gly His | |
| tgg ggc cag ggc acc acc ctg aca gtg agc agc gct agc acc aag ggc | 2864 |

ES 2 540 753 T3

| | | | | | | | | | | | | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|-----|-----|-----|-----|-----|----------|-----|-----|-----|--|------|
| Trp | Gly | Gln | Gly | Thr | Thr | Leu | Thr | Val | Ser | Ser | Ala | Ser | Thr | Lys | Gly | | |
| cca | agc | gtg | ttc | cca | ctg | gcc | ccc | tgc | agc | aga | agc | acc | agc | gag | agc | | 2912 |
| Pro | Ser | Val | Phe | Pro | Leu | Ala | Pro | Cys | Ser | Arg | Ser | Thr | Ser | Glu | Ser | | |
| 120 | | | | | 125 | | | | 130 | | | | | 135 | | | |
| aca | gcc | gcc | ctg | ggc | tgc | ctg | gtg | aag | gac | tac | ttc | ccc | gag | ccc | gtg | | 2960 |
| Thr | Ala | Ala | Leu | Gly | Cys | Leu | Val | Lys | Glu | Tyr | Phe | Pro | Glu | Pro | Val | | |
| | | | | 140 | | | | | 145 | | | | | 150 | | | |
| acc | gtg | tcc | tgg | aac | agc | gga | gcc | ctg | aca | agc | gga | gtg | cac | acc | ttc | | 3008 |
| Thr | Val | Ser | Trp | Asn | Ser | Gly | Ala | Leu | Thr | Ser | Gly | Val | His | Thr | Phe | | |
| | | | 155 | | | | | 160 | | | | | 165 | | | | |
| ccc | gcc | gtg | ctg | cag | agc | agc | ggc | ctg | tac | tcc | ctg | agc | agc | gtg | gtg | | 3056 |
| Pro | Ala | Val | Leu | Gln | Ser | Ser | Gly | Leu | Tyr | Ser | Leu | Ser | Ser | Val | Val | | |
| | | | 170 | | | | 175 | | | | | 180 | | | | | |
| acc | gtg | cca | agc | agc | agc | ctg | ggc | acc | aag | acc | tac | acc | tgc | aac | gtg | | 3104 |
| Thr | Val | Pro | Ser | Ser | Ser | Leu | Gly | Thr | Lys | Thr | Tyr | Thr | cys | Asn | Val | | |
| | | | | | | 190 | | | | | 195 | | | | | | |
| gac | cac | aag | ccc | agc | aac | acc | aaa | gtg | gac | aag | cgc | gtg | gag | agc | aag | | 3152 |
| Asp | His | Lys | Pro | Ser | Asn | Thr | Lys | Val | Asp | Lys | Arg | Val | Glu | Ser | Lys | | |
| 200 | | | | | 205 | | | | | 210 | | | | | 215 | | |
| tac | ggc | cct | ccc | tgc | ccc | agc | tgt | ccc | gcc | cca | gag | ttc | ctg | ggc | gga | | 3200 |
| Tyr | Gly | Pro | Pro | Cys | Pro | Ser | Cys | Pro | Ala | Pro | Glu | Glu | Leu | Gly | Gly | | |
| | | | | 220 | | | | | 225 | | | | | 230 | | | |
| ccc | tca | gtg | ttt | ctg | ttc | cca | ccc | aag | ccc | aag | gat | acc | ctg | atg | atc | | 3248 |
| Pro | Ser | Val | Phe | Leu | Phe | Pro | Pro | Lys | Pro | Lys | Asp | Thr | Leu | Met | Ile | | |
| | | | 235 | | | | | 240 | | | | | 245 | | | | |
| agc | cgg | acc | cct | gaa | gtg | acc | tgc | gtg | gtg | gat | gtg | gtg | agc | cag | gag | | 3296 |
| Ser | Arg | Thr | Pro | Glu | Val | Thr | Cys | Val | Val | Val | Asp | Val | Ser | Gln | Glu | | |
| | | | 250 | | | | 255 | | | | | 260 | | | | | |
| gac | ccc | gaa | gtc | cag | ttc | aat | tgg | tac | gtg | gac | ggc | gtg | gaa | gtg | cac | | 3344 |
| Asp | Pro | Glu | Val | Gln | Phe | Asn | Trp | Tyr | Val | Asp | Gly | Val | Glu | Val | His | | |
| | | | 265 | | | 270 | | | | | 275 | | | | | | |
| aac | gcc | aag | acc | aag | ccc | aga | gag | gag | cag | ttc | aac | agc | acc | tac | cgc | | 3392 |
| Asn | Ala | Lys | Thr | Lys | Pro | Arg | Glu | Glu | Gln | Phe | Asn | Ser | Thr | Tyr | Arg | | |
| 280 | | | | | 285 | | | | | 290 | | | | | 295 | | |
| gtg | gtg | tct | gtg | ctg | acc | gtg | ctg | cac | cag | gat | tgg | ctg | aac | ggc | aaa | | 3440 |
| Val | Val | Ser | Val | Leu | Thr | Val | Leu | His | Gln | Asp | Trp | Leu | Asn | Gly | Lys | | |
| | | | | 300 | | | | | 305 | | | | | 310 | | | |
| gag | tac | aag | tgc | aag | gtc | tcc | aac | aag | ggc | ctg | cct | agc | agc | atc | gag | | 3488 |
| Glu | Tyr | Lys | Cys | Lys | Val | Ser | Asn | Lys | Gly | Leu | Pro | Ser | Ser | Ile | Glu | | |
| | | | 315 | | | | | 320 | | | | | | 325 | | | |
| aaa | acc | atc | agc | aag | gcc | aag | ggc | cag | cca | cgc | gag | ccc | cag | gtg | tac | | 3536 |
| Lys | Thr | Ile | Ser | Lys | Ala | Lys | Gly | Gln | Pro | Arg | Glu | Pro | Gln | Val | Tyr | | |
| | | | 330 | | | | 335 | | | | | 340 | | | | | |
| acc | ctg | ccc | ccc | agc | caa | gag | gag | atg | acc | aag | aac | cag | gtg | tcc | ctg | | 3584 |
| Thr | Leu | Pro | Pro | Ser | Gln | Glu | Glu | Met | Thr | Lys | Asn | Gln | Val | Ser | Leu | | |
| | | | | | 350 | | | | 355 | | | | | | | | |
| acc | tgt | ctg | gtg | aag | ggc | ttc | tac | ccc | agc | gac | atc | gcc | gtg | gag | tgg | | 3632 |
| Thr | Cys | Leu | Val | Lys | Gly | Phe | Tyr | Pro | Ser | Asp | Ile | Ala | Val | Glu | Trp | | |
| 360 | | | | | 365 | | | | | 370 | | | | | 375 | | |
| gag | agc | aac | ggc | cag | ccc | gag | aac | aac | tac | aag | acc | acc | ccc | cct | gtg | | 3680 |
| Glu | Ser | Asn | Gly | Gln | Pro | Glu | Asn | Asn | Tyr | Lys | Thr | Thr | pro | Pro | Val | | |
| | | | | 380 | | | | | 385 | | | | | 390 | | | |
| ctg | gac | agc | gat | ggc | agc | ttc | ttc | ctg | tac | tca | cgg | ctg | acc | gtg | gat | | 3728 |
| Leu | Asp | Ser | Asp | Gly | Ser | Phe | Phe | Leu | Tyr | Ser | Arg | Leu | Thr | Val | Asp | | |
| | | | 395 | | | | | 400 | | | | | 405 | | | | |
| aag | agc | aga | tgg | caa | gag | ggc | aat | gtc | ttt | agc | tgc | agc | gtg | atg | cac | | 3776 |
| Lys | Ser | Arg | Trp | Gln | Glu | Gly | Asn | Val | Phe | Ser | Cys | Ser | Val | Met | His | | |
| | | | 410 | | | 415 | | | | | | 420 | | | | | |
| gag | gcc | ctg | cac | aat | cct | aca | ccc | aga | aga | gcc | tga | gcctgtcc | | | | | 3820 |
| Glu | Ala | Leu | His | Asn | Pro | Thr | Pro | Arg | Arg | Ala | | | | | | | |
| | | | 425 | | | 430 | | | | | | | | | | | |
| cctgggcaag | tgatagtcgc | gaattgatca | taatcagcca | taccacattt | gtagaggttt | | | | | | | | | | | | 3880 |
| tacttgcttt | aaaaaacctc | ccacacctcc | ccctgaacct | gaaacataaa | atgaatgcaa | | | | | | | | | | | | 3940 |
| ttgtgttgt | taacttgttt | attgcagctt | ataatggtta | caaataaagc | aatagcatca | | | | | | | | | | | | 4000 |
| caaatttcac | aaataaagca | ttttttcac | tgcattctag | ttgtggtttg | tccaaactca | | | | | | | | | | | | 4060 |
| tcaatgtatc | ttatcatgtc | tggatcagct | tgagcagctg | tggaatgtgt | gtcagttagg | | | | | | | | | | | | 4120 |
| gtgtggaaag | tccccaggct | ccccagcagg | cagaagtatg | caaagcatgc | atctcaatta | | | | | | | | | | | | 4180 |
| gtcagcaacc | aggtgtggaa | agtccccagg | ctccccagca | ggcagaagta | tgcaaagcat | | | | | | | | | | | | 4240 |

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|-------------|------|
| gcattctcaat | tagtcagcaa | ccatagtcgcc | gcccctaact | ccgcccattc | cgcccctaac | 4300 |
| tccgcccagt | tccgcccatt | ctccgcccc | tggctgacta | atftttttta | tttatgcaga | 4360 |
| ggccgaggcc | gcctcggcct | ctgagctatt | ccagaagtag | tgaggaggct | tttttgagg | 4420 |
| ctaccatgga | gaagtacta | ttccgaagtt | cctattctct | agaaagtata | ggaacttctc | 4480 |
| gggcccgttt | gctggcgttt | ttccataggc | tccgcccc | tgacgagcat | cacaaaaatc | 4540 |
| gacgctcaag | tcagagggtg | cgaaacccga | caggactata | aagataccag | gcgtttcccc | 4600 |
| ctggaagctc | cctcgtgctc | tctcctgttc | cgaccctgcc | gcttaccgga | tacctgtccg | 4660 |
| cctttctccc | ttcggggaagc | gtggcgcttt | ctcatagctc | acgctgtagg | tatctcagtt | 4720 |
| cggtgtaggt | cgttcgtctc | aagctgggct | gtgtgcacga | acccccggt | cagcccagcc | 4780 |
| gctgcgcctt | atccggtaac | tatcgtcttg | agtccaacct | ggtaagacac | gacttatcgc | 4840 |
| cactggcagc | agccactggt | aacaggatta | gcagagcgag | gtatgtaggc | ggtgctacag | 4900 |
| agttcttgaa | gtgggtggcct | aactacggct | acactagaag | aacagtatft | ggtatctgcg | 4960 |
| ctctgctgaa | gccagttacc | ttcggaaaa | gagttggtag | ctcttgatcc | ggcaaacaaa | 5020 |
| ccaccgtgtt | tagcgggtgt | ttttttggtt | gcaagcagca | gattacgcgc | agaaaaaaag | 5080 |
| gatctcaaga | agatcctttg | atcttttcta | cggggctgta | cgctcagttg | aacgaaaact | 5140 |
| cacgttaagg | gatttttggtc | atgagattat | caaaaaggat | cttcacctag | atccttttaa | 5200 |
| atataaaaatg | aagttttaaa | tcaatctaaa | gtatatatga | gtaaacttgg | tctgacagtt | 5260 |
| accaatgctt | aatcagtgag | gcacctatct | cagcgtctg | tctatttctg | tcatccatag | 5320 |
| ttgcctgact | ccccgtctg | tagataacta | cgatacggga | gggcttacca | tctggcccca | 5380 |
| gtgctgcaat | gataccgcga | gacccacgct | caccggctcc | agatttatca | gcaataaac | 5440 |
| agccagccgg | aagggcggag | cgcagaagtg | gtcctgcaac | tttatccgcc | tccatccagt | 5500 |
| ctattaatgt | ttgccgggaa | gctagagtaa | gtagttcgcc | agtttaatag | ttgcgcaacg | 5560 |
| ttgttgccat | tgtctacaggc | atcgtgggtg | cacgctcgct | gtttggtag | gcttcattca | 5620 |
| gctccggttc | ccaacgatca | aggcgagtta | catgatcccc | catgtttgtc | aaaaaagcgg | 5680 |
| ttagctcctt | cggtcctcgg | atcgtttgca | gaagtaagtt | ggccgcagtg | ttatcactca | 5740 |
| tggttatggc | agcactgcat | aattctctta | ctgtcatgcc | atccgtaaga | tgcttttctg | 5800 |
| tgactggtga | gtactcaacc | aagtcattct | gagaatagtg | tatgcggcga | ccgagtgtct | 5860 |
| cctgcccggc | gtcaatacgg | gataataccg | cgccacatag | cagaacttta | aaagtgtca | 5920 |
| tcattggaaa | acgttctctg | gggcgaaaac | tctcaaggat | cttaccgctg | ttgagatcca | 5980 |
| gttcgatgta | accactcgt | gcacccaact | gatcttcagc | atcttttact | ttcaccagcg | 6040 |
| tttctgggtg | agcaaaaaa | ggaaggcaaa | atgccgcaaa | aaagggata | agggcgacac | 6100 |
| ggaaatggtg | aatactcata | ctcttccttt | ttcaatatta | ttgaagcatt | ttcagggtt | 6160 |
| attgtctcat | gagcggatac | atatttgaat | gtatttagaa | aaataaacia | ataggggttc | 6220 |
| cgcgcacatt | tccccgaaaa | gtgccacctg | acgtctaaga | aaccattatt | atcatgacat | 6280 |
| taacctataa | aaatagggct | atcacgaggc | cctgatggct | ctttgcggca | cccactcgtc | 6340 |
| gtaatgttcc | gtggcaccca | ggacaacctt | caagagaaaa | tgtaatcaca | ctgggaagtt | 6400 |
| cctattccga | agttcctatt | cttcaaaagg | tataggaact | tcctgcagtg | aataataaaa | 6460 |
| tgtgtgtttg | tccgaaatac | gvcgctactg | agtcattagg | gactttccaa | tgggttttgc | 6520 |
| ccagtaacata | aggtcaatag | gggtgaatca | acaggaaagt | cccattggag | ccaagtacac | 6580 |
| tgagtcaata | gggactttcc | attgggtttt | gcccagtaca | aaaggtcaat | agggggtgag | 6640 |
| tcaatgggtt | tttccatta | ttggcagcta | cataaggctc | ataggggtga | gtcattgggt | 6700 |
| ttttccagcc | aatttaatta | aaacgccatg | tactttccca | ccattgacgt | caatgggcta | 6760 |
| ttgaaactaa | tgcaacgtga | cctttaaacg | gtactttccc | atagctgatt | aatgggaaaag | 6820 |
| taccggttac | gagccaatac | acgtcaatgg | gaagtgaag | ggcagccaaa | acgtaacacc | 6880 |
| gcccgggttt | tcccctggaa | attccatatt | ggcacgcatt | ctattggctg | agctgcggtc | 6940 |
| tacgtgggta | taagaggcgc | gaccagcgtc | ggtaccgctc | cagtcttcgg | tctgaccacc | 7000 |
| gtagaacgca | gcctcaggac | ctccatagaa | gacaccggga | ccgatccagc | ctccgcggcc | 7060 |
| gggaacgggt | cattggaacg | cggtattcccc | gtgccaagag | tgacgtaagt | accgcctata | 7120 |
| gagtcctatag | gcccaccccc | ttggcttctt | atgcatgcta | tactgttttt | ggcttggggt | 7180 |
| ctatacacc | ccgcttctct | atgttatagg | tgatgggata | gcttagccta | taggtgtggg | 7240 |
| ttattgacca | ttattgacca | ctccccctatt | ggtgacgata | ctttccatta | ctaatccata | 7300 |
| acatggctct | ttgccacaac | tctctttatt | ggctatatgc | caatacactg | tccttcagag | 7360 |
| actgacacgg | actctgtatt | ttacaggat | ggggtctcat | ttattattta | caaattcaca | 7420 |
| tatacaaac | caccgtcccc | agtgcccgca | gtttttatta | aacataacgt | gggatctcca | 7480 |
| cgcaaatctc | gggtacgtgt | tccggacatg | ggctcttctc | cggtagcggc | ggagcttcta | 7540 |
| catccgagcc | ctgctcccat | gcctccagcg | actcatggtc | gctcggcagc | tccttgctcc | 7600 |
| taacagtgga | ggccagactt | aggcacagca | cgatgcccac | caccaccagt | gtgccgcaca | 7660 |
| aggccgtggc | ggtaggggat | gtgtctgaaa | atgagctcgg | ggagcgggct | tgaccgcctg | 7720 |
| acgcatttgg | aagacttaag | gcagcggcag | aagaagatgc | aggcagctga | gttgttgtgt | 7780 |
| tctgataaga | gtcagaggta | actcccgttg | cggtgctgtt | aacgggtggag | ggcaggtgag | 7840 |
| tctgagcagt | actcgttgct | gccgcgcgcg | ccaccagaca | taatagctga | cagactaaca | 7900 |
| gactgttctc | ttccatgggt | cttttctgca | gtcaccgtcc | ttgacacg | | 7948 |

<210>3
 <211> 492
 5 <212> ADN
 <213> Secuencia artificial

<220>
 <223> fragmento de la secuencia promotora de mCMV IE1

10

ES 2 540 753 T3

| | | | | | | | |
|----|---|------------|------------|------------|------------|------------|-----|
| | <400>3 | | | | | | |
| | tactgagtc | ttagggactt | tccaatgggt | ttgcccagt | acataaggtc | aataggggtg | 60 |
| | aatcaacagg | aaagtcccat | tggagccaag | tacactgagt | caatagggac | tttccattgg | 120 |
| | gttttgccc | gtacaaaagg | tcaatagggg | gtgagtcaat | gggtttttcc | cattattggc | 180 |
| | acgtacataa | ggtcaatagg | ggtgagtc | tgggtttttc | cagccaattt | aattaaaacg | 240 |
| | ccatgtactt | tcccaccatt | gacgtcaatg | ggctattgaa | actaatgcaa | cgtgaccttt | 300 |
| | aaacgggtact | ttcccatagc | tgattaatgg | gaaagtaccg | ttctcgagcc | aatacacgtc | 360 |
| | aatgggaagt | gaaagggcag | ccaaaacgta | acaccgcccc | ggttttcccc | tggaaattcc | 420 |
| | atattggcac | gcattctatt | ggctgagctg | cgttctacgt | gggtataaga | ggcgcgacca | 480 |
| | gcgtcggtag | cg | | | | | 492 |
| | <210>4 | | | | | | |
| 5 | <211> 102 | | | | | | |
| | <212> ADN | | | | | | |
| | <213> Secuencia artificial | | | | | | |
| | <220> | | | | | | |
| 10 | <223> fragmento de la secuencia promotora de mCMV IE1 ("promotor núcleo") | | | | | | |
| | <400> 4 | | | | | | |
| | acaccgcccc | ggttttcccc | tggaaattcc | atattggcac | gcattctatt | ggctgagctg | 60 |
| | cgttctacgt | gggtataaga | ggcgcgacca | gcgtcggtag | cg | | 102 |
| | <210>5 | | | | | | |
| 15 | <211> 452 | | | | | | |
| | <212> ADN | | | | | | |
| | <213> Secuencia artificial | | | | | | |
| | <220> | | | | | | |
| 20 | <223> fragmento de la secuencia potenciadora de sCMV IE1 | | | | | | |
| | <400> 5 | | | | | | |
| | gaccatagcc | aattcaatat | ggcgtatatg | gactcatgcc | aattcaatat | ggtggatctg | 60 |
| | gacctgtgcc | aattcaatat | ggcgtatatg | gactcgtgcc | aattcaatat | ggtggatctg | 120 |
| | gaccccagcc | aattcaatat | ggcggacttg | gcaccatgcc | aattcaatat | ggcggacctg | 180 |
| | gcactgtgcc | aactggggag | gggtctactt | ggcacggtgc | caagtttgag | gaggggtctt | 240 |
| | ggccctgtgc | caagtccgcc | atattgaatt | ggcatggtgc | caataatggc | ggccatattg | 300 |
| | gctatatgcc | aggatcaata | tataggcaat | atccaatatg | gccctatgcc | aatatggcta | 360 |
| | ttggccaggt | tcaatactat | gtattggccc | tatgccatat | agtattccat | atatggggtt | 420 |
| 25 | tcctattgac | gtagatagcc | cctcccaatg | gg | | | 452 |
| | <210>6 | | | | | | |
| | <211> 1074 | | | | | | |
| | <212> ADN | | | | | | |
| 30 | <213> Secuencia artificial | | | | | | |
| | <220> | | | | | | |
| | <223> secuencia reguladora 5' no traducida quimérica ("construcción 1") | | | | | | |
| | <220> | | | | | | |
| 35 | <221> potenciador | | | | | | |
| | <222> (1)...(582) | | | | | | |
| | <223> fragmento de secuencia potenciadora de hCMV IE1 | | | | | | |
| | <220> | | | | | | |
| 40 | <221> promotor | | | | | | |
| | <222> (583)...(1074) | | | | | | |
| | <223> fragmento de secuencia promotora de mCMV IE1 | | | | | | |
| | <400>6 | | | | | | |
| 45 | | | | | | | |

ES 2 540 753 T3

ctgcagtgaa taataaaatg tgtgtttgtc cgaaatacgc gttttgagat ttctgtcgcc 60
 gactaaattc atgtcgcgcg atagtgggtg ttatcgccga tagagatggc gatattggaa 120

aaatcगतat ttgaaaatat ggcatattga aaatgtcgcg gatgtgagtt tctgtgtaac 180
 tgataatcgcc atttttccaa aagtgatttt tgggcatacg cgataatctgg cgatagcgt 240
 tataatcgttt acgggggatg gcgatagacg actttgggtga cttgggcat tctgtgtgtc 300
 gcaaatatcg cagtttcgat atagggtgaca gacgatatga ggctatatcg ccgatagagg 360
 cgacatcaag ctggcacatg gccaatgcat atcgatctat acattgaatc aatattggcc 420
 attagccata ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca 480
 tacgttgat ccatatcata atatgtacat ttatattggc tcatgtccaa cattaccgcc 540
 atgttgacat tgattattga ctagtattta atagtaatca attactgagt cattagggac 600
 ttccaatgg gttttgccc gtacataagg tcaatagggg tgaatcaaca ggaaagtccc 660
 attggagcca agtacctga gtcaataggg actttccatt gggtttgccc cagtacaaaa 720
 ggtcaatagg ggggtgagtc atgggttttt cccattattg gcacgtacat aaggtaata 780
 ggggtgagtc attgggtttt tccagccaat ttaattaaaa cgccatgtac tttcccacca 840
 ttgacgtcaa tgggctattg aaactaatgc aacgtgacct ttaaaccgta cttcccata 900
 gctgattaat gggaaagtac cgttctcgag ccaatacagc tcaatgggaa gtgaaagggc 960
 agcctaaacg taacaccgcc ccggttttcc cctggaaatt ccatattggc acgattcta 1020
 ttggctgagc tgcgttctac gtgggtataa gaggcgcgac cagcgtcggg accg 1074

<210>7

<211> 1128

5 <212> ADN

<213> Secuencia artificial

<220>

<223> secuencia reguladora 5' no traducida química ("construcción 2")

10

<220>

<221> potenciador

<222> (1)...(1026)

<223> fragmento de secuencia potenciadora de hCMV IE1

15

<220>

<221> promotor

<222> (1027)...(1128)

<223> fragmento de secuencia promotora de mCMV IE1 ("promotor núcleo")

20

<400> 7

ctgcagtgaa taataaaatg tgtgtttgtc cgaaatacgc gttttgagat ttctgtcgcc 60
 gactaaattc atgtcgcgcg atagtgggtg ttatcgccga tagagatggc gatattggaa 120
 aaatcगतat ttgaaaatat ggcatattga aaatgtcgcg gatgtgagtt tctgtgtaac 180
 tgataatcgcc atttttccaa aagtgatttt tgggcatacg cgataatctgg cgatagcgt 240
 tataatcgttt acgggggatg gcgatagacg actttgggtga cttgggcat tctgtgtgtc 300
 gcaaatatcg cagtttcgat atagggtgaca gacgatatga ggctatatcg ccgatagagg 360
 cgacatcaag ctggcacatg gccaatgcat atcgatctat acattgaatc aatattggcc 420
 attagccata ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca 480
 tacgttgat ccatatcata atatgtacat ttatattggc tcatgtccaa cattaccgcc 540
 atgttgacat tgattattga ctagtattta atagtaatca attacggggt cattagttca 600
 tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 660
 gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 720
 agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 780
 acatcaagtg tatcatafgc caagtacgcc ccctattgac gtcaatgacg gtaaatggcc 840
 cgcttgcat tatgcccagt acatgacctt atgggacttt cctacttggc agtacatcta 900
 cgtattagtc atcgctatta ccatgggtgat gcggttttgg cagtacatca atgggctgg 960
 atagcggttt gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt 1020
 gtttgacac cgccccggtt tccccctgga aattccatat tggcacgcac tctattggct 1080
 gagctgcggt ctacgtgggt ataagaggcg cgaccagcgt cgggtaccg 1128

25 <210> 8

<211> 509

<212> ADN

<213> Secuencia artificial

30 <220>

ES 2 540 753 T3

<223> secuencia reguladora 5' no traducida quimérica ("construcción 3")

<220>

<221> potenciador

5 <222> (1)...(407)

<223> fragmento de secuencia potenciadora de hCMV IE1

<220>

<221> promotor

10 <222> (408)...(509)

<223> fragmento de secuencia promotora de mCMV IE1 ("promotor núcleo")

<400> 8

| | | | | | | | |
|----|-------------|-------------|------------|------------|------------|------------|-----|
| | cgcgttacat | aacttacggt | aaatggcccg | cctggctgac | cgcccaacga | ccccgccc | 60 |
| | ttgacgtcaa | taatgacgta | tgttcccata | gtaacgccaa | tagggacttt | ccattgacgt | 120 |
| | caatgggtgg | agtattttacg | gtaaactgcc | cacttggcag | tacatcaagt | gtatcatatg | 180 |
| | ccaagtacgc | cccctattga | cgccaatgac | ggtaaatggc | ccgcctggca | ttatgccag | 240 |
| | tacatgacct | tatgggacct | tcctacttgg | cagtacatct | acgtattagt | catcgctatt | 300 |
| | accatgggtga | tgcggttttg | gcagtacatc | aatgggcgtg | gatagcggtt | tgactcacgg | 360 |
| | ggatttccaa | gtctccacc | cattgacgtc | aatgggagtt | tgttttgaca | ccgccccggt | 420 |
| | ttccccctgg | aaattccata | ttggcacgca | ttctattggc | tgagctgcgt | tctacgtggg | 480 |
| 15 | tataagaggc | gcgaccagcg | tcggtaccg | | | | 509 |

<210> 9

<211> 961

<212> ADN

20 <213> Secuencia artificial

<220>

<223> secuencia reguladora 5' no traducida quimérica ("construcción 4")

25 <220>

<221> potenciador

<222> (1)...(452)

<223> fragmento de secuencia potenciadora de sCMV IE1

30 <220>

<221> potenciador

<222> (453)...(859)

<223> fragmento de secuencia potenciadora de hCMV IE1

35 <220>

<221> > promotor

<222> (860)...(961)

<223> fragmento de secuencia promotora de mCMV IE1 ("promotor núcleo")

40 <400> 9

| | | | | | | | |
|--|------------|------------|------------|------------|------------|-------------|-----|
| | gaccatagcc | aattcaatat | ggcgtatatg | gactcatgcc | aattcaatat | ggtggatctg | 60 |
| | gacctgtgcc | aattcaatat | ggcgtatatg | gactcgtgcc | aattcaatat | ggtggatctg | 120 |
| | gacccagcc | aattcaatat | ggcggacttg | gcaccatgcc | aattcaatat | ggcggacctg | 180 |
| | gcactgtgcc | aactggggag | gggtctactt | ggcacgggtg | caagttagag | gaggggtctt | 240 |
| | ggccctgtgc | caagtccgcc | atattgaatt | ggcatgggtg | caataatggc | ggccatattg | 300 |
| | gctatatgcc | aggatcaata | tataggcaat | atccaatatg | gccctatgcc | aatatggcta | 360 |
| | ttggccaggt | tcaatactat | gtattggccc | tatgccatat | agtattccat | atatgggttt | 420 |
| | tcctattgac | gtagatagcc | cctccaatg | ggcgcgttac | ataacttacg | gtaaattggcc | 480 |
| | cgctggctg | accgcccac | gacccccgcc | cattgacgtc | aataatgacg | tatgttccca | 540 |
| | tagtaacgcc | aatagggact | ttccattgac | gtcaatgggt | ggagtattta | cggtaaactg | 600 |
| | cccacttggc | agtacatcaa | gtgtatcata | tgccaagtac | gccccctatt | gacgtcaatg | 660 |
| | acggtaaatg | gcccgcctgg | cattatgccc | agtacatgac | cttatgggac | tttctactt | 720 |
| | ggcagtacat | ctacgtatta | gtcatcgcta | ttaccatggg | gatgcggttt | tggcagtaca | 780 |
| | tcaatggcg | tggatagcgg | tttgactcac | ggggatttcc | aagtcctcac | cccattgacg | 840 |
| | tcaatgggag | tttgttttga | caccgccccg | gttttcccct | ggaaattcca | tattggcacg | 900 |
| | cattctattg | gctgagctgc | gttctacgtg | ggtataagag | gcgcgaccag | cgtcgggtacc | 960 |
| | g | | | | | | 961 |

ES 2 540 753 T3

<210> 10
 <211> 909
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> secuencia reguladora 5' no traducida quimérica ("construcción 5")
 <220>
 <221> potenciador
 <222> (1)...(807)
 <223> fragmento de secuencia potenciadora de sCMV IE1
 10
 <220>
 <221 > promotor
 <222> (808)...(909)
 <223> fragmento de secuencia promotora de mCMV IE1 ("promotor núcleo")
 15
 <400> 10
 gaccatagcc aattcaatat ggcgtatatg gactcatgcc aattcaatat ggtggatctg 60
 gacctgtgcc aattcaatat ggcgtatatg gactcgtgcc aattcaatat ggtggatctg 120
 gaccccagcc aattcaatat ggcggacttg gcaccatgcc aattcaatat ggcggacctg 180
 gcactgtgcc aactggggag gggcttactt ggcacggtgc caagtttgag gaggggtctt 240
 ggccctgtgc caagtccgcc atattgaatt ggcatggtgc caataatggc ggccatattg 300
 gctatatgcc aggatcaata tataggcaat atccaatatg gccctatgcc aatatggcta 360
 ttggccaggt tcaatactat gtattggccc tatgccatat agtattccat atatggggtt 420
 tcctattgac gtagatagcc cctcccaatg ggcggtccca tataccatat atggggcttc 480
 ctaataccgc ccatagccac tccccattg acgtcaatgg tctctatata tggctttcc 540
 tattgacgtc atatgggagg tcctattgac gtatatggcg cctcccccat tgacgtcaat 600
 tacggtaaat ggcccgcctg gctcaatgcc cattgacgtc aataggacca cccaccattg 660
 acgtcaatgg gatggctcat tgcccattca tatccgttct cacgccccct attgacgtca 720
 atgacggtaa atggcccact tggcagtaca tcaatatcta ttaatagtaa cttggcaagt 780
 acattactat tgggaagtacg ccagggtaca ccgccccggt tttcccctgg aaattccata 840
 ttggcacgca ttctattggc tgagctgcgt tctacgtggg tataagaggc gcgaccagcg 900
 20 tcggtaccg 909
 <210> 11
 <211> 214
 <212> PRT
 <213> Secuencia artificial
 25
 <220>
 <223> cadena ligera kappa de cB72.3 optimizada
 <220>
 <400> 11
 30

ES 2 540 753 T3

Asp Ile Gln Met Thr Gln Ser Pro Ala Ser Leu Ser Val Ser Val Gly
 1 5 10 15
 Asp Thr Val Thr Ile Thr Cys Arg Ala Ser Glu Asn Ile Tyr Ser Asn
 20 25 30
 Leu Ala Trp Tyr Gln Gln Lys Gln Gly Lys Ser Pro Gln Leu Leu Val
 35 40 45
 Tyr Ala Ala Thr Asn Leu Ala Asp Gly Val Pro Ser Arg Phe Ser Gly
 50 55 60
 Ser Gly Ser Gly Thr Gln Tyr Ser Leu Lys Ile Asn Ser Leu Gln Ser
 65 70 75 80
 Glu Asp Phe Gly Ser Tyr Tyr Cys Gln His Phe Trp Gly Thr Pro Tyr
 85 90 95
 Thr Phe Gly Gly Thr Arg Leu Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110
 Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125
 Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140
 Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160
 Asp Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175
 Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190
 Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205
 Phe Asn Arg Gly Glu Cys
 210

<210> 12
 <211> 434
 5 <212> PRT
 <213> Secuencia artificial

<220>
 <223> cadena pesada gamma-4 de cB72.3 optimizada
 10
 <400> 12

Gln Val Gln Leu Gln Gln Ser Asp Ala Glu Leu Val Lys Pro Gly Ala
 1 Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asp His
 20 Ala Ile His Trp Ala Lys Gln Lys pro Glu Gln Gly Leu Glu Trp Ile
 35 Gly Tyr Ile Ser Pro Gly Asn Asp Asp Ile Lys Tyr Asn Glu Lys Phe
 50 Lys Gly Lys Ala Thr Leu Thr Ala Asp Lys Ser Ser Ser Thr Ala Tyr
 65 Met Gln Leu Asn Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
 85 Lys Arg Ser Tyr Tyr Gly His Trp Gly Gln Gly Thr Thr Leu Thr Val
 100 Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys
 115 Ser Arg Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu Val Lys
 130 Glu Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu
 145 Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu
 160 Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr
 175 Lys Thr Tyr Thr cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val
 180 Asp Lys Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser Cys Pro
 195 Ala Pro Glu Glu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
 210 Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
 225 Val Val Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr
 240 Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
 255 Gln Phe Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
 270 Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
 285 Gly Leu Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
 300 Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met
 315 Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
 330 Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
 345 Tyr Lys Thr Thr pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
 360 Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val
 375 Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn Pro Thr Pro Arg
 385 Arg Ala

REIVINDICACIONES

- 5 1. Un vector de expresión para la expresión heteróloga de una secuencia de ácido nucleico de interés en células de mamífero, comprendiendo el vector una primera secuencia quimérica reguladora de promotor que está unida operativamente a una primera secuencia de ácido nucleico que se va a expresar, en donde la secuencia quimérica reguladora de promotor comprende:
- 10 (i) una secuencia promotora del promotor IE1 de citomegalovirus murino y que está unida operativamente al sitio de inicio de la transcripción de la secuencia de ácido nucleico que se va a expresar; y
- 10 (ii) una secuencia amplificadora de la región IE1 de citomegalovirus humano y/o de simio, estando localizada la secuencia amplificadora 5' y unida operativamente a la secuencia promotora IE1 de citomegalovirus murino; y
- 15 en donde la secuencia quimérica reguladora de promotor comprende una secuencia de nucleótidos que se selecciona de entre el grupo que consiste en la SEC ID N° 7, SEC ID N° 9 y SEC ID N° 10.
- 20 2. El vector de expresión de la reivindicación 1, que comprende además una segunda secuencia quimérica reguladora de promotor que está unida operativamente a una segunda secuencia de ácido nucleico que se va a expresar, en donde la segunda secuencia quimérica reguladora de promotor es idéntica a la primera secuencia quimérica reguladora de promotor.
- 25 3. El vector de expresión de la reivindicación 1, que comprende además una segunda secuencia quimérica reguladora de promotor que está unida operativamente a una segunda secuencia de ácido nucleico que se va a expresar, en donde la segunda secuencia quimérica reguladora de promotor es diferente de la primera secuencia quimérica reguladora de promotor.
- 30 4. El vector de expresión de las reivindicaciones 2 o 3, en donde la primera y la segunda secuencias de ácido nucleico que se van a expresar codifican polipéptidos diferentes.
- 35 5. El vector de expresión de la reivindicación 4, en donde los polipéptidos diferentes representan subunidades de una proteína dimérica o multimérica.
- 40 6. El vector de expresión de la reivindicación 5, en donde la proteína dimérica o multimérica es una molécula de anticuerpo.
- 45 7. Una célula huésped de mamífero que se transfecta con un vector de expresión como se define en una cualquiera de las reivindicaciones 1 a 6.
8. La célula huésped de mamífero de la reivindicación 7, en donde la célula huésped es una célula CHO.
9. Un método para la expresión heteróloga de una secuencia de ácido nucleico de interés en una célula huésped de mamífero, que comprende:
- 45 (i) transfectar la célula huésped de mamífero con un vector de expresión como se define en una cualquiera de las reivindicaciones 1 a 6; y
- 45 (ii) cultivar la célula huésped de mamífero transfectada bajo condiciones que permitan la expresión de la secuencia de ácido nucleico de interés.
- 50 10. El método de la reivindicación 9, en donde la transfección es una transfección estable.
- 50 11. Uso de un vector de expresión como se define en una cualquiera de las reivindicaciones 1 a 6 para la expresión heteróloga *in vitro* de una secuencia de ácido nucleico de interés en una célula huésped de mamífero.

FIGURA 1

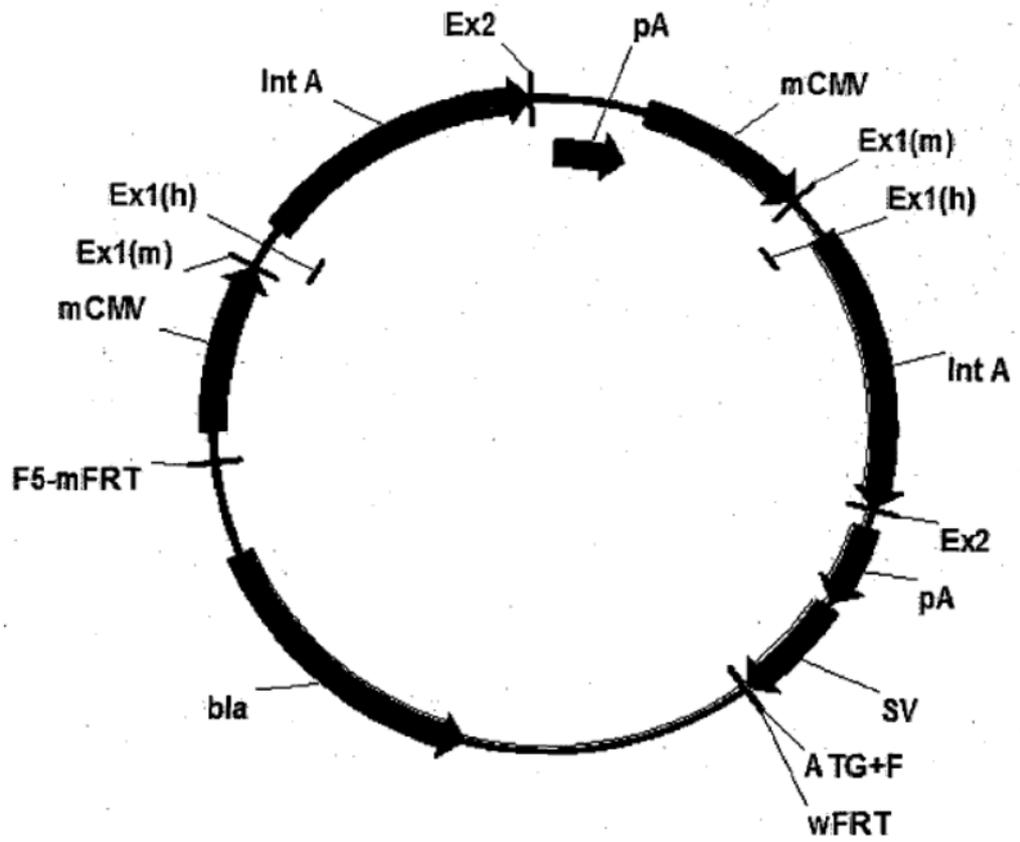


FIGURA 2

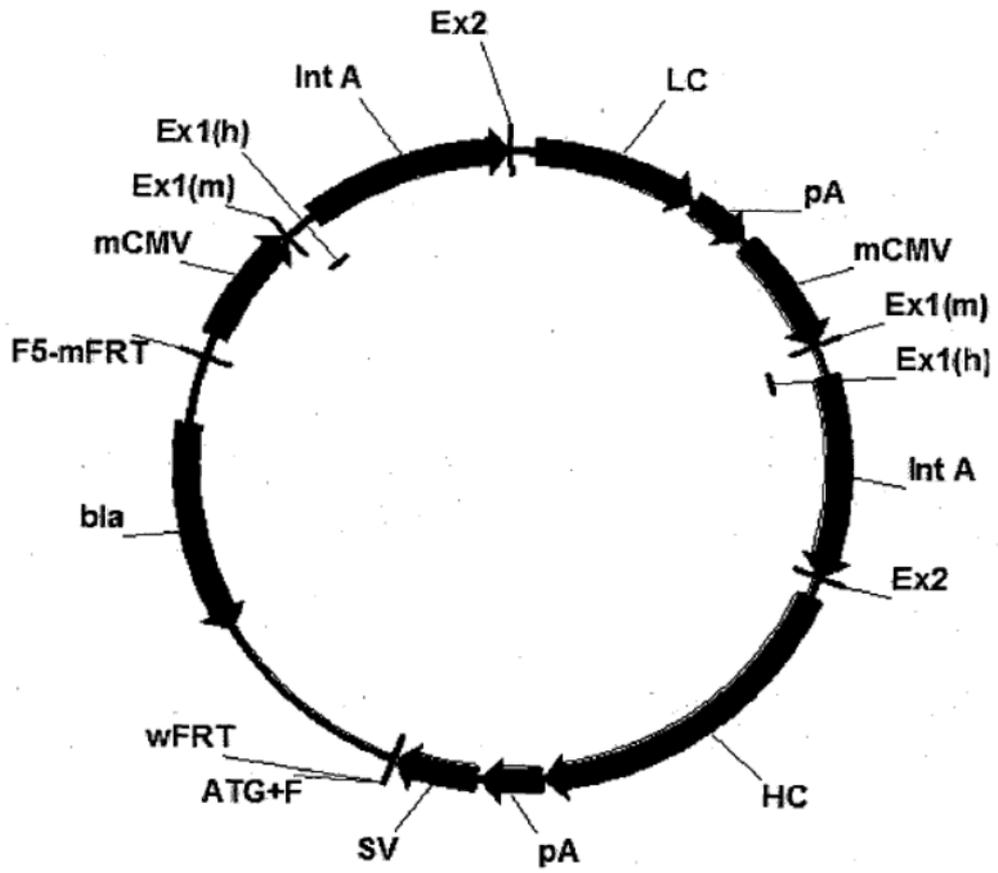


FIGURA 3

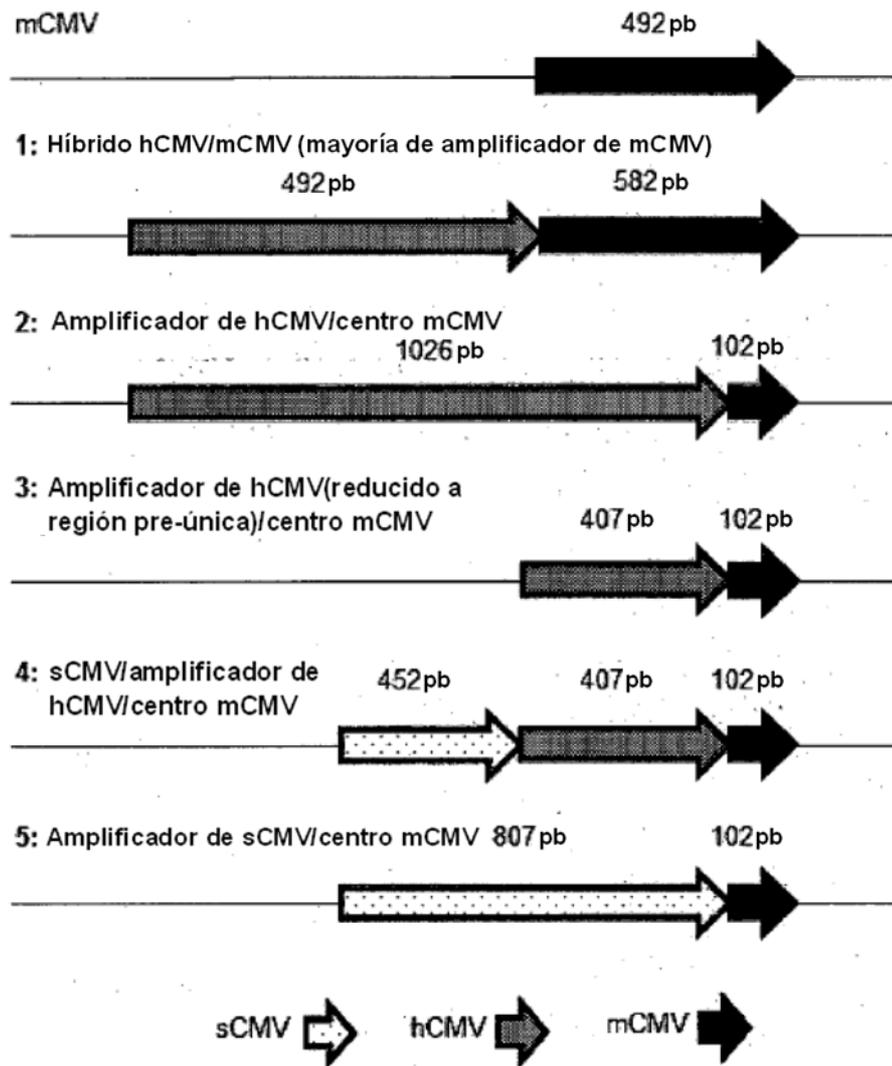


FIGURA 4