

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 706 899**

51 Int. Cl.:

C12N 15/52 (2006.01)
C12N 5/10 (2006.01)
C12N 15/867 (2006.01)
C07K 14/39 (2006.01)
C12N 9/10 (2006.01)
C07K 14/55 (2006.01)
C07K 14/57 (2006.01)
C12N 9/78 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.09.2009** **PCT/US2009/058510**

87 Fecha y número de publicación internacional: **22.04.2010** **WO10045002**

96 Fecha de presentación y número de la solicitud europea: **26.09.2009** **E 09820986 (9)**

97 Fecha y número de publicación de la concesión europea: **07.11.2018** **EP 2346995**

54 Título: **Vectores de terapia génica y citosina desaminasas**

30 Prioridad:

13.06.2009 US 186823 P
26.09.2008 US 100666 P
08.12.2008 US 120618 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
01.04.2019

73 Titular/es:

TOCAGEN INC. (100.0%)
3030 Bunker Hill Street Suite 230
San Diego, CA 92109, US

72 Inventor/es:

GRUBER, HARRY, E.;
JOLLY, DOUGLAS;
PEREZ, OMAR y
LOGG, CHRISTOPHER R.

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

ES 2 706 899 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 terapia génica y citosina desaminasas

5 **Campo técnico**

La presente divulgación se refiere a citosina desaminasas modificadas (CD). La divulgación se refiere además a las células que expresan tales CD modificadas y a los procedimientos para usar dichas CD modificadas en el tratamiento de enfermedades y trastornos.

10 **Antecedentes**

La citosina desaminasa de levadura o bacteriana convierte el inocuo profármaco antibiótico 5-FC en el agente quimioterapéutico citotóxico 5-fluorouracilo (5-FU). No se sabe que los seres humanos (y los mamíferos en general) posean un gen natural que codifica una enzima con una actividad citosina desaminasa significativa. La citosina desaminasa bacteriana y de levadura han ganado reconocimiento en el tratamiento de los cánceres mediante la liberación de genes y vectores virales para el suministro de la enzima seguido del tratamiento con 5-FC, que luego se convierte mediante la enzima en un fármaco citotóxico (Miller et al., Can Res 62:773-780 2002; Kievit et al., Can Res 59:1417-1421 1999).

Korkegian et al, (2005), Nature, vol. 308, pp. 857-860 describen la termoestabilización computacional de una enzima. Stolworthy et al. (2008), J. Mol Bio, vol. 377, pp. 854-869 indican que los mutantes de la citosina desaminasa de levadura con mayor termoestabilidad imparten sensibilidad a la 5-fluorocitosina. Khatri et al (2006), J. Gene Med vol. 8, pp. 1086-1096 indican que la combinación de citosina desaminasa con uracil fosforribosil transferasa conduce a efectos secundarios locales y distantes contra el cáncer de próstata RM₁ en ratones.

El documento WO2006 / 05831 desvela un virus competente para la replicación que codifica plásmidos para su uso en terapia, más particularmente para su uso en el tratamiento de una enfermedad de proliferación celular, una enfermedad inmunológica, una infección adquirida e inflamación.

Kei, H. et al (Cancer Research, vol. 67, no. 11, 5345-5353) desvela la eficacia terapéutica de la terapia génica suicida mediada por vector de retrovirus con replicación competente en un modelo de metástasis de cáncer colorrectal multifocal.

Xing, L. et al. (International Journal of Radiation, vol. 69, n.º 3, S617-S618) desvela que una combinación de citosina desaminasa con uracil fosforribosil transferasa con 5-fluorocitosina radiosensibiliza las células de cáncer de próstata de rata. El documento WO2005086922 desvela composiciones y procedimientos para tratar el cáncer usando un adenovirus mutante que comprende un polinucleótido que codifica un polipéptido terapéutico que está dirigido a células con una ruta de retinoblastoma mutante. El adenovirus mutante es capaz de matar las células tumorales sin dañar las células con una ruta de retinoblastoma de tipo salvaje.

Sumario

En un aspecto, la presente invención proporciona un vector según la reivindicación 1. En otro aspecto, la presente invención proporciona un polinucleótido aislado según la reivindicación 8. En un aspecto adicional más, la presente invención proporciona un polinucleótido aislado según la reivindicación 14. En el presente documento se proporcionan polipéptidos que convierten 5-FC en 5-FU. También se proporcionan moléculas de ácido nucleico que codifican tales polipéptidos, células que expresan tales polipéptidos, vectores que contienen dichos polinucleótidos y polipéptidos y procedimientos para sintetizar 5-FU o derivados de los mismos a partir de un uso adecuado de tales polipéptidos. Por consiguiente, en diversos aspectos de la divulgación, se proporcionan polipéptidos aislados o recombinantes que comprenden una citosina desaminasa que tiene una secuencia como se indica en la SEQ ID NO: 2. En otros aspectos de la divulgación, los polipéptidos comprenden la SEQ ID NO: 2 con una mutación seleccionada del grupo que consiste en A23L, V108I, I140L y cualquier combinación de los mismos. En aún otro aspecto de la divulgación, el polipéptido comprende la secuencia que se establece en la SEQ ID NO: 4 e incluye hasta 50, 25, 10 o 5 sustituciones conservadoras de aminoácidos, excluyendo los restos: (a) 23, 108 y 140. En general, Los polipéptidos proporcionados en el presente documento muestran la actividad de citosina desaminasa útil para convertir 5-FC en 5-FU.

La divulgación también proporciona un polipéptido que comprende una secuencia que es al menos un 80 %, 90 %, 95 %, 98 % o 99 % idéntico a la SEQ ID NO: 4, donde el polipéptido tiene una leucina en la posición 23, una isoleucina en la posición 108 y una leucina en la posición 140, y donde el polipéptido tiene actividad citosina desaminasa. En aún otro aspecto de la divulgación, el polipéptido comprende la secuencia expuesta en la SEQ ID NO: 4.

La divulgación proporciona además construcciones de fusión que comprenden uno cualquiera de los polipéptidos anteriores unidos operativamente a una uracil fosforribosiltransferasa (UPRT) o una orotato fosforribosiltransferasa (OPRT). En un aspecto de la divulgación, la construcción de fusión comprende un primer polipéptido que comprende la SEQ ID NO: 2 con una mutación seleccionada del grupo que consiste en A23L, V108I, I140L y cualquier combinación de las mismas unidas a un segundo polipéptido que tiene actividad UPRT u OPRT. En aún otro aspecto de la divulgación, la construcción de fusión comprende un primer polipéptido que tiene la secuencia como se expone

en la SEQ ID NO: 4 e incluye hasta 50, 25, 10 o 5 sustituciones conservadoras de aminoácidos, excluyendo los restos: 23, 108 y 140, donde el primer polipéptido está unido a un segundo polipéptido que tiene actividad UPRT u OPRT. La divulgación también proporciona una construcción de fusión que comprende un primer polipéptido que tiene actividad citosina desaminasa y que comprende una secuencia que es al menos un 80 %, 90 %, 95 %, 98 % o 99 % idéntico a la SEQ ID NO: 4, donde el polipéptido tiene una leucina en la posición 23, una isoleucina en la posición 108 y una leucina en la posición 140, y donde el primer polipéptido está unido a un segundo polipéptido que comprende actividad UPRT o OPRT. En un aspecto de la divulgación, los polipéptidos que tienen actividad UPRT u OPRT comprenden una secuencia como se expone en las SEQ ID NO: 8 y 10, respectivamente, o variantes de los mismos. En aún otro aspecto de la divulgación, un primer polipéptido que comprende actividad citosina desaminasa está unido a un polipéptido que tiene actividad UPRT u OPRT, donde los polipéptidos están separados por un enlazador peptídico. En otro aspecto de la divulgación, la construcción de fusión comprende una secuencia seleccionada del grupo que consiste en las SEQ ID NO: 12, 14, 16 o 18.

La divulgación proporciona además polinucleótidos que codifican cualquiera de los polipéptidos anteriores. Por ejemplo, la divulgación proporciona un polinucleótido que codifica un polipéptido que comprende la SEQ ID NO: 2 con una mutación seleccionada del grupo que consiste en A23L, V108I, I140L y cualquier combinación de los mismos. En aún otro aspecto de la divulgación, El polinucleótido codifica un polipéptido que comprende una secuencia que se expone en la SEQ ID NO: 4 e incluye hasta 50, 25, 10 o 5 sustituciones conservadoras de aminoácidos, excluyendo los restos: (a) 23, 108 y 140. En un aspecto adicional de la divulgación, la divulgación proporciona un polinucleótido que codifica un polipéptido que comprende una secuencia que es al menos un 80 %, 90 %, 95 %, 98 % o 99 % idéntico a la SEQ ID NO: 4, donde el polipéptido tiene una leucina en la posición 23, una isoleucina en la posición 108 y una leucina en la posición 140, y donde el polipéptido tiene actividad citosina desaminasa. En aún otro aspecto de la divulgación, el polinucleótido codifica un polipéptido que comprende la SEQ ID NO: 4. En un aspecto adicional de la divulgación, el polinucleótido codifica un polipéptido que comprende una secuencia como se expone en las SEQ ID NO: 4, 6, 8, 10, 11, 12 o 13.

La divulgación proporciona un polinucleótido optimizado por codones humano que codifica una citosina desaminasa. En un aspecto de la divulgación, el polinucleótido optimizado por codones humano comprende una secuencia como se expone en la SEQ ID NO: 3. En aún otro aspecto de la divulgación, el polinucleótido comprende la secuencia expuesta en la SEQ ID NO: 5, 7 o 9.

En otros aspectos de la divulgación, una citosina desaminasa o construcción de fusión de la divulgación se libera usando un sistema de liberación de genes (GDS). En otro aspecto, el polinucleótido que codifica una citosina desaminasa se libera con un GDS que es un vector viral o derivado de virus. El vector viral puede ser replicante o no replicante, y puede ser un vector adenoviral, un vector de sarampión, un vector de herpes, un vector retroviral (incluido un vector lentiviral), un vector rabinoviral, tal como un vector viral de estomatitis vesicular, un vector de reovirus, un vector del virus del valle de Séneca, un vector de poxvirus (incluyendo vectores derivados de la viruela animal o vaccinia), un vector de parvovirus (incluyendo un vector de AAV), un vector de alfavirus u otro vector viral conocido por un experto en la técnica. En un aspecto de la divulgación, el vector viral puede ser un vector retroviral competente para la replicación capaz de infectar células de mamíferos en replicación. El vector retroviral competente para la replicación puede comprender un ortoretrovirus o más típicamente un vector de retrovirus gamma. En un aspecto, un vector retroviral competente para la replicación comprende un sitio interno de entrada al ribosoma interno (IRES) en 5' al polinucleótido que codifica una citosina desaminasa de la divulgación. En un aspecto de la divulgación, el polinucleótido que codifica una citosina desaminasa está en 3' a un polinucleótido ENV de un vector retroviral.

En otros aspectos de la divulgación, se proporcionan células huésped transfectadas con una citosina desaminasa o construcción de fusión de la divulgación, o un vector que incluye un polinucleótido o construcción de fusión de la divulgación. Las células huésped incluyen células eucariotas, tales como células de levadura, células de insecto o células de animales. Las células huésped también incluyen células procariotas, tales como células bacterianas.

La divulgación también proporciona procedimientos para tratar un trastorno de proliferación celular que incluye un cáncer que comprende administrar un polinucleótido o polipéptido de la divulgación a un sujeto y ponerlo en contacto con un fármaco citotóxico que comprende 5-fluorocitosina (5-FC).

La divulgación proporciona un retrovirus competente para la replicación recombinante (RCR) que comprende: una proteína GAG retroviral; una proteína POL retroviral; una cubierta retroviral; un polinucleótido retroviral que comprende secuencias de repetición terminales largas (LTR) en el extremo 3' de la secuencia de polinucleótido retroviral, una secuencia promotora en el extremo 5' del polinucleótido retroviral, siendo dicho promotor adecuado para la expresión en una célula de mamífero, un dominio de ácido nucleico gag, un dominio de ácido nucleico pol y un dominio de ácido nucleico env; un casete que comprende un sitio interno de entrada al ribosoma (IRES) unido operativamente a un polinucleótido heterólogo, donde el casete está posicionado en 5' a 3' LTR y 3' al dominio de ácido nucleico env que codifica la cubierta retroviral; y secuencias que actúan en cis necesarias para la transcripción inversa, empaquetamiento e integración en una célula objetivo, donde el RCR mantiene una mayor competencia de replicación después de 6 pases en comparación con un vector pACE (SEQ ID NO: 21). En un aspecto de la divulgación, la secuencia polinucleotídica retroviral deriva del virus de la leucemia murina (MLV), el virus de la

leucemia murina de Moloney (MoMLV), el virus de la leucemia felina o virus de la leucemia del mono de Gibbon (GALV). En otro aspecto de la divulgación, el MLV es un MLV anfitriónico. En aún otro aspecto de la divulgación, el retrovirus es un oncoretrovirus. En aún otro aspecto de la divulgación, la célula objetivo es una célula que tiene un trastorno proliferativo celular. El trastorno de proliferación celular puede seleccionarse del grupo que consiste en, aunque no de forma limitativa, cáncer de pulmón, cáncer de colon y recto, cáncer de mama, cáncer de próstata, cáncer del tracto urinario, cáncer de útero, cáncer cerebral, cáncer de cabeza y cuello, cáncer pancreático, melanoma, cáncer de estómago y cáncer de ovarios, artritis reumatoide y otras enfermedades autoinmunes. En un aspecto de la divulgación, el promotor puede comprender un promotor de CMV que tiene una secuencia como se expone en la SEQ ID NO:19, 20 o 22 desde el nucleótido 1 a aproximadamente el nucleótido 582 y puede incluir una modificación en una o más bases de ácidos nucleicos y que es capaz de dirigir e iniciar la transcripción. En aún un aspecto adicional de la divulgación, el promotor comprende una secuencia como se expone en las SEQ ID NO: 19, 20 o 22 desde el nucleótido 1 hasta aproximadamente el nucleótido 582. En un aspecto adicional de la divulgación, el promotor comprende un polinucleótido de dominio CMV-R-U5. En un aspecto de la divulgación, el dominio CMV-R-U5 comprende el promotor inmediatamente temprano de citomegalovirus humano unido a una región MLV R-U5. En aún otro aspecto de la divulgación, el polinucleótido del dominio CMV-R-U5 comprende una secuencia como se expone en las SEQ ID NO: 19, 20, o 22 desde aproximadamente el nucleótido 1 a aproximadamente el nucleótido 1202 o las secuencias que son al menos un 95 % idénticas a una secuencia como se expone en las SEQ ID NO: 19, 20 o 22, donde el polinucleótido promueve la transcripción de una molécula de ácido nucleico unida operativamente a la anterior. En otro aspecto de la divulgación, el gag y pol del polinucleótido derivan de un oncoretrovirus. El dominio gag de ácido nucleico puede comprender una secuencia desde aproximadamente el nucleótido número 1203 a aproximadamente el nucleótido 2819 de la SEQ ID NO: 19 o una secuencia que tiene al menos un 95 %, 98 %, 99 % o 99,8 % de identidad con la anterior. El dominio pol comprende una secuencia desde aproximadamente el nucleótido número 2820 a aproximadamente el nucleótido 6358 de la SEQ ID NO:19 o una secuencia que tiene al menos un 95 %, 98 %, 99 % o 99,9 % de identidad con la anterior. En un aspecto de la divulgación, el dominio env codifica una proteína env anfótera. El dominio env comprende una secuencia desde aproximadamente el número 6359 a aproximadamente el nucleótido 8323 de las SEQ ID NO: 19 o una secuencia que tiene al menos un 95 %, 98 %, 99 % o 99,8 % de identidad con la anterior. El dominio IRES del vector puede ser cualquier IRES, sin embargo, en un aspecto de la divulgación, el IRES deriva de un virus de encefalomiocarditis. En un aspecto adicional de la divulgación, el IRES comprende una secuencia desde aproximadamente el número 8327 a aproximadamente el nucleótido 8876 de la SEQ ID NO: 19 o una secuencia que tiene al menos un 95 %, 98 % o 99% de identidad con la misma. En aún un aspecto adicional de la divulgación, el polinucleótido heterólogo comprende un polinucleótido que tiene una secuencia como se expone en la SEQ ID NO: 3, 5, 11, 13, 15 o 17. En un aspecto adicional de la divulgación, la secuencia heteróloga codifica un polipéptido que comprende una secuencia como se expone en la SEQ ID NO:4. El ácido nucleico heterólogo es un codón humano optimizado y codifica un polipéptido como se expone en la SEQ ID NO: 4. En un aspecto adicional de la divulgación, el ácido nucleico heterólogo comprende una secuencia como se expone en la SEQ ID NO: 19, desde aproximadamente el nucleótido 8877 hasta aproximadamente el nucleótido 9353. En un aspecto de la divulgación, la LTR en 3' deriva de un oncoretrovirus. En un aspecto adicional de la divulgación, la 3' LTR comprende un dominio U3-R-U5. En aún un aspecto adicional de la divulgación, la LTR en 3' comprende una secuencia como se expone en la SEQ ID NO: 19 desde aproximadamente el nucleótido 9405 a aproximadamente el nucleótido 9998 o una secuencia que tiene al menos un 95 %, 98 % o 99,5 % de identidad con la misma.

La divulgación proporciona un polinucleótido que comprende una secuencia como se expone en las SEQ ID NO:19, 20 o 22.

La divulgación proporciona un polinucleótido aislado que comprende de 5' a 3': una fusión CMV-R-U5 del promotor temprano inmediato del citomegalovirus humano en una región R-U5 del VLM; un PBS, un sitio de unión al cebador para la transcriptasa inversa; un sitio 5' de corte y empalme; una señal de empaquetamiento ψ ; una secuencia de codificación gag para el antígeno específico del grupo del VLM; una secuencia de codificación pol para la poliproteína de la polimerasa de VLM; un sitio 3' de corte y empalme; una secuencia de codificación de env 4070A para la proteína de la cubierta de la cepa 4070A del MLV; un sitio interno de entrada al ribosoma (IRES) del virus de la encefalomiocarditis; una secuencia codificante de citosina desaminasa modificada; un tracto de polipurina; y una repetición terminal larga U3-R-U5 del MLV.

La divulgación proporciona un procedimiento para tratar a un sujeto con un trastorno de proliferación celular que comprende poner en contacto al sujeto con un polinucleótido, un polipéptido de la divulgación que tiene actividad citosina desaminasa en condiciones tales que el polinucleótido se expresa y poner en contacto al sujeto con 5-fluorocitosina.

La divulgación también proporciona un procedimiento para tratar un trastorno de proliferación celular en un sujeto que comprende poner en contacto al sujeto con un retrovirus de la divulgación, donde la secuencia de ácido nucleico heteróloga codifica una proteína terapéutica que inhibe la proliferación de una célula neoplásica. En una realización, el retrovirus comprende un polinucleótido que codifica un polipéptido que tiene una secuencia como se expone en las SEQ ID NO: 4, 12, 14, 16 o 18.

Los detalles de una o más realizaciones de la divulgación se exponen en los dibujos adjuntos y en la descripción a

continuación. Otras características, objetos y ventajas serán evidentes a partir de la descripción y los dibujos, y de las reivindicaciones.

Breve descripción de los dibujos

La **Figura 1A-D** muestra (a) un esquema del vector retroviral recombinante de la divulgación; (b) un mapa plasmídico de un polinucleótido de la divulgación; (c y d) una secuencia de un polinucleótido de la divulgación (SEQ ID NO: 19).

La **Figura 2A-D** muestra esquemas para la generación de varias realizaciones de la divulgación que comprenden polipéptidos con actividad CD, OPRT y UPRT.

La **Figura 3** muestra que se observan niveles más altos de proteína yCD2 en comparación con la proteína yCD de tipo salvaje en las células U-87 infectadas.

La **Figura 4** muestra la estabilidad de un vector que comprende un polipéptido CD de la divulgación y la comparación con otros vectores de la divulgación.

La **Figura 5A-B** muestra que la actividad de citosina desaminasa y el vector de la divulgación proporcionan una expresión comparable o mejor, y por lo tanto, la destrucción, de células RG2 (5A) o U87 de rata infectadas (5B) en comparación con la actividad de yCD de tipo salvaje (T5.0007), cuando las células infectadas se exponen a niveles crecientes de 5-FC.

La **Figura 6** muestra que la actividad específica en células U87 infectadas de T5.0002 (yCD2) es mayor que T5.0001 (yCD parcialmente optimizada) que es mayor que T5.0007 (yCD salvaje).

Símbolos de referencia similares en los diversos dibujos indican elementos similares.

Descripción detallada

Tal como se usa en el presente documento y en las reivindicaciones adjuntas, las formas en singular "un/a", "y", y "el/la" incluyen las referencias plurales salvo que el contexto indique claramente otra cosa. Por lo tanto, por ejemplo, la referencia a "una célula" incluye una pluralidad de dichas células y la referencia a "el agente" incluye una referencia a uno o más agentes conocidos por los expertos en la técnica, y así sucesivamente.

Asimismo, el uso de "o" significa "y/o" a no ser que se indique otra cosa. De forma similar, "comprende", "comprenden", "que comprende" "incluye", "incluyen" y "que incluye", es indistinto y no se pretende que sea limitante.

Debe entenderse además que, cuando las descripciones de diversas realizaciones utilizan el término "que comprende", los expertos en la materia entenderán que, en algunos casos específicos, una realización puede describirse alternativamente utilizando el lenguaje "consiste esencialmente en" o "consiste en".

A no ser que se defina de otra forma, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que entiende comúnmente una persona normalmente experta en la materia a la cual pertenece esta divulgación. Aunque los procedimientos y materiales similares o equivalentes a los descritos en el presente documento se pueden usar en la práctica de los procedimientos y composiciones descritos, en el presente documento se describen los procedimientos, dispositivos y materiales de ejemplo.

Las publicaciones tratadas anteriormente y a lo largo del texto se proporcionan ricamente para su divulgación antes de la fecha de presentación de la presente solicitud.

Una citosina desaminasa (EC 3.5.4.1) es una enzima que cataliza la



Por lo tanto, los dos sustratos de esta enzima son citosina y H₂O, mientras que sus dos productos son uracilo y NH₃. Esta enzima pertenece a la familia de las hidrolasas, las que actúan sobre enlaces carbono-nitrógeno distintos de los enlaces peptídicos, específicamente en amidinas cíclicas. El nombre sistemático de esta clase de enzima es citosina aminohidrolasa. Esta enzima también se llama isocitosina desaminasa. Esta enzima participa en el metabolismo de las pirimidinas.

Más particularmente, la citosina desaminasa es una enzima involucrada en la ruta metabólica de las pirimidinas, a través de la cual se transforma la citosina exógena, mediante desaminación hidrolítica, en uracilo. Se han demostrado actividades de la citosina desaminasa (CDasa o CD) en procariotas y eucariotas inferiores, pero están ausentes en mamíferos (Koechlin et al., 1966, Biochem. Pharmacol. 15, 435-446; Polak et al., 1976, Chemotherapy 22, 137-153). El gen *FCY1* de *Saccharomyces cerevisiae* (*S. cerevisiae*) y el gen *coda* de *E. coli*, que codifican, respectivamente, la CDasa de estos dos organismos, se conocen y sus secuencias están publicadas (documento EP 402 108; Erbs et al., 1997, Curr. Genet. 31, 1-6; documento WO 93/01281). La CDasa también desamina un análogo de la citosina, 5-fluorocitosina (5-FC) a 5-fluorouracilo (5-FU), que es un compuesto altamente citotóxico, en particular cuando se convierte a 5-fluoro-UMP (5-FUMP). Las células que carecen de actividad CDasa, debido a una

mutación inactivadora del gen que codifica la enzima o a su deficiencia natural para esta enzima (por ejemplo, células de mamífero) son resistentes al 5-FC (Jund y Lacroute, 1970, J. Bacteriol. 102, 607-615; Kilstrup et al., 1989, J. Bacteriol., 171, 2124-2127). Por otro lado, se ha demostrado que es posible transmitir la sensibilidad a 5-FC a células de mamífero a las que se ha transferido la secuencia que codifica una actividad de CDasa (Huber et al., 1993, Cancer Res. 53, 4619-4626; Mullen et al., 1992, Proc. Natl. Acad. Sci. USA 89, 33-37; documento WO 93/01281). Por consiguiente, el uso de CD es ventajoso en el contexto de la terapia génica, en particular, de la terapia génica contra el cáncer.

Sin embargo, la sensibilidad al 5-FC varía mucho dependiendo de las líneas celulares. Se observa baja sensibilidad, por ejemplo, en las líneas tumorales humanas PANC-1 (carcinoma de páncreas) y SK-BR-3 (adenocarcinoma de mama) transducidas con un retrovirus que expresa el gen *coda* de *E. coli* (Harris et al., 1994, Gene Therapy 1, 170-175). Este fenómeno se explica por la ausencia o mala conversión endógena del 5-FU formado por la acción enzimática de la CDasa, al 5-FUMP citotóxico. Esta etapa, que normalmente se lleva a cabo en células de mamíferos mediante orotato fosforibosiltransferasa (OPRTasa), puede estar ausente en ciertos tumores y, por lo tanto, hacer que la terapia génica basada en la CDasa sea inefectiva. En procariotas y eucariotas inferiores, el uracilo se transforma en UMP a través de la acción de la uracil fosforibosiltransferasa (actividad UPRTasa). Esta enzima también convierte 5-FU en 5-FUMP. Es importante destacar que, la uracil fosforibosiltransferasa bacteriana (UPRT) es funcionalmente equivalente a la orotato fosforibosiltransferasa (OPRT) o uridina-5' -monofosfato sintasa de células de mamífero. Estas enzimas median la conversión de 5-fluorouracilo (5-FU) (un análogo fluorado de uracilo) en monofosfato de 5-fluorouridina 5' (5-FUMP). El 5-fluorouridina 5' monofosfato se convierte posteriormente en 5-FdUDP y FdUTP a través de la ruta de las pirimidinas *de novo* de los mamíferos. Cada 5-FdUTP es un inhibidor irreversible de la timidilato sintasa (Thy-A) y da como resultado la inanición de dTTP. Está ampliamente aceptado que esta conversión es una de las vías necesarias para lograr los efectos citotóxicos del 5-fluorouracilo y que la uracil fosforibosiltransferasa bacteriana de origen bacteriano es capaz de convertir el 5-fluorouracilo en el mismo metabolito activo que la orotato fosforibosiltransferasa. En ausencia de actividad UPRTasa o actividad OPRTasa, el 5-FU, originado de la desaminación del 5-FC por la CDasa, no se transforma en 5-FUMP citotóxico (Jund y Lacroute, 1970, J. Bacteriol. 102, 607-615).

Como se describe en el presente documento, la divulgación proporciona polinucleótidos que codifican polipéptidos y polipéptidos que comprenden una fusión de un primer polipéptido que tiene actividad citosina desaminasa y un segundo polipéptido que comprende actividad UPRT o OPRT. Tales construcciones de fusión son útiles para convertir uracilo, o un derivado del mismo, en un análogo monofosfato, y en particular 5-FU en 5-FUMP.

Como se describirá con más detalle a continuación, la divulgación se basa, al menos en parte, en la generación y expresión de nuevos polipéptidos que catalizan la conversión 5-FC a 5-FU. En un aspecto de la divulgación, se proporcionan nuevos polipéptidos que han sido diseñados para tener propiedades catalíticas mejoradas para la conversión de 5-FC a 5-FU. Tales polipéptidos incluyen variantes que se han alterado para incluir sustituciones de aminoácidos en residuos específicos. Si bien estas variantes se describirán con más detalle a continuación, se entiende que los polipéptidos de la divulgación pueden contener uno o más aminoácidos modificados. La presencia de aminoácidos modificados puede ser ventajosa en, por ejemplo, (a) aumentar la semivida *in vivo* del polipéptido, (b) reducir o aumentar la antigenicidad del polipéptido, y (c) aumentar la estabilidad de almacenamiento del polipéptido. Los aminoácidos son modificados, por ejemplo, co-traduccionalmente o post-traduccionalmente durante la producción recombinante (por ejemplo, glicosilación ligada a N en los motivos N - X - S / T durante la expresión en células de mamífero) o modificada por medios sintéticos. Por consiguiente, una proteína, enzima, polinucleótido, gen, o célula "mutante", "variante" o "modificada" significa una proteína, enzima, polinucleótido, gen, o célula, que ha sido alterado o derivado, o que es de alguna manera diferente o ha cambiado, respecto de una proteína, enzima, polinucleótido, gen, o célula original. Una proteína o enzima mutante o modificada suele, aunque no necesariamente, expresarse a partir de un polinucleótido o gen mutante.

Una "mutación" significa cualquier proceso o mecanismo que resulta en una proteína, enzima, polinucleótido, gen, o célula mutante. Esto incluye cualquier mutación en la que una proteína, enzima, polinucleótido o secuencia génica está alterada y cualquier cambio detectable en una célula que surja de tal mutación. Típicamente, una mutación se produce en un polinucleótido o secuencia génica, mediante mutaciones puntuales, deleciones o inserciones de residuos de nucleótidos simples o múltiples. Una mutación incluye alteraciones polinucleotídicas que surgen dentro de una región codificadora de proteínas de un gen, así como alteraciones en regiones fuera de una secuencia codificadora de proteínas, tal como, pero sin limitación, secuencias reguladoras o promotoras. Una mutación en un gen puede ser "silenciosa", es decir, no se refleja en una alteración de aminoácidos en la expresión, lo que conduce a una variante "conservadora de secuencia" del gen. Esto generalmente surge cuando un aminoácido corresponde a más de un codón.

Los ejemplos no limitantes de un aminoácido modificado incluyen un aminoácido glicosilado, un aminoácido sulfatado, un aminoácido prenilado (por ejemplo, aminoácido farnesilado, geranilgeranilado), un aminoácido acetilado, un aminoácido acilado, un aminoácido pegilado, un aminoácido biotilado, un aminoácido carboxilado, un aminoácido fosforilado, y similares. Las referencias adecuadas para guiar a un experto en la modificación de aminoácidos están repletas en toda la literatura. Ejemplos de protocolos se encuentran en Walker (1998) Protein Protocols en CD-ROM (Humana Press, Towata, N.J.).

En el presente documento se describen procedimientos recombinantes para la producción, el aislamiento y el uso de los polipéptidos y polinucleótidos de CD modificados de la divulgación. Además de la producción recombinante, los polipéptidos pueden producirse por síntesis peptídica directa utilizando técnicas en fase sólida (por ejemplo, Stewart et al. (1969) Solid-Phase Peptide Synthesis (WH Freeman Co, San Francisco); y Merrifield (1963) J. Am. Chem. Soc. 85: 2149-2154). La síntesis de péptidos se puede realizar utilizando técnicas manuales o por automatización. Se puede lograr la síntesis automatizada, por ejemplo, utilizando el sintetizador de péptidos Applied Biosystems 431A (Perkin Elmer, Foster City, Calif.) de acuerdo con las instrucciones proporcionadas por el fabricante.

A modo de ilustración, las secuencias de ácido nucleico que codifican las UPRTasas de *E. coli* (Anderson et al., 1992, Eur. J. Biochem 204, 51-56), de *Lactococcus lactis* (Martinussen y Hammer, 1994, J. Bacteriol. 176, 6457-6463), de *Mycobacterium bovis* (Kim et al., 1997, Biochem Mol. Biol. Int 41, 1117-1124) y de *Bacillus subtilis* (Martinussen et al., 1995, J. Bacteriol. 177, 271-274) se puede utilizar en el contexto de la divulgación. Sin embargo, se prefiere el uso de una UPRTasa de levadura y, en particular, la codificada por el gen FUR1 de *S. cerevisiae*, cuya secuencia se desvela en Kern et al. (1990, Gene 88, 149-157). A modo de indicación, las secuencias de los genes y las de las UPRTasas correspondientes se pueden encontrar en la literatura y en los bancos de datos especializados (SWISSPROT, EMBL, Genbank Medline, etc.).

Una "proteína" o "polipéptido", términos que se usan indistintamente en el presente documento, comprende una o más cadenas de bloques componentes químicos denominados aminoácidos que están unidos entre sí por enlaces químicos llamados enlaces peptídicos. Una "enzima" significa cualquier sustancia, preferentemente compuesta en su totalidad o en gran parte por proteínas, que cataliza o promueve, más o menos específicamente, una o más reacciones químicas o bioquímicas. El término "enzima" también puede referirse a un polinucleótido catalítico (por ejemplo, ARN o ADN). Una proteína, enzima, polinucleótido, gen, o célula "nativa" o "tipo salvaje", significa una proteína, enzima, polinucleótido, gen, o célula, significa una proteína, enzima, polinucleótido, gen, o célula que se produce en la naturaleza.

Por consiguiente, se proporcionan polipéptidos aislados o recombinantes que comprenden la SEQ ID NO: 2 con una mutación seleccionada del grupo que consiste en A23L, V108I, I140L y cualquier combinación de las mismas. En aún otro aspecto de la divulgación, el polipéptido comprende la secuencia que se expone en la SEQ ID NO: 4 e incluye hasta 50, 25, 10 o 5 sustituciones conservadoras de aminoácidos, excluyendo los restos: (a) 23, 108 y 140. En general, Los polipéptidos proporcionados en el presente documento muestran la actividad de citosina desaminasa útil para convertir 5-FC en 5-FU.

La divulgación también proporciona un polipéptido que comprende una secuencia que es al menos un 80 %, 90 %, 95 %, 98 % o 99 % idéntico a la SEQ ID NO: 4, donde el polipéptido tiene una leucina en la posición 23, una isoleucina en la posición 108 y una leucina en la posición 140, y donde el polipéptido tiene actividad citosina desaminasa. En aún otro aspecto de la divulgación, el polipéptido comprende la secuencia expuesta en la SEQ ID NO: 4.

La divulgación proporciona además construcciones de fusión que comprenden uno cualquiera de los polipéptidos anteriores unidos operativamente a una uracil fosforribosiltransferasa (UPRT) o una orotato fosforribosiltransferasa (OPRT). En un aspecto de la divulgación, la construcción de fusión comprende un primer polipéptido que comprende la SEQ ID NO: 2 con una mutación seleccionada del grupo que consiste en A23L, V108I, I140L y cualquier combinación de las mismas unidas a un segundo polipéptido que tiene actividad UPRT u OPRT. En aún otro aspecto de la divulgación, la construcción de fusión comprende un primer polipéptido que tiene la secuencia como se expone en la SEQ ID NO: 4 e incluye hasta 50, 25, 10 o 5 sustituciones conservadoras de aminoácidos, excluyendo los restos: 23, 108 y 140, donde el primer polipéptido está unido a un segundo polipéptido que tiene actividad UPRT u OPRT. La divulgación también proporciona una construcción de fusión que comprende un primer polipéptido que tiene actividad citosina desaminasa y que comprende una secuencia que es al menos un 80 %, 90 %, 95 %, 98 % o 99 % idéntico a la SEQ ID NO: 4, donde el polipéptido tiene una leucina en la posición 23, una isoleucina en la posición 108 y una leucina en la posición 140, y donde el primer polipéptido está unido a un segundo polipéptido que comprende actividad UPRT o OPRT. En un aspecto de la divulgación, los polipéptidos que tienen actividad UPRT u OPRT comprenden una secuencia como se expone en las SEQ ID NO: 8 y 10, respectivamente, o variantes de los mismos. En aún otro aspecto de la divulgación, un primer polipéptido que comprende actividad citosina desaminasa está unido a un polipéptido que tiene actividad UPRT u OPRT, donde los polipéptidos están separados por un enlazador peptídico. En otro aspecto de la divulgación, la construcción de fusión comprende una secuencia seleccionada del grupo que consiste en las SEQ ID NO: 11, 12 o 13.

La expresión "unido operativamente" o "asociado operativamente" se refiere a un enlace o asociación funcional entre una secuencia reguladora y el polinucleótido regulado por la secuencia reguladora o entre dos polipéptidos o polinucleótidos diferentes que codifican tales polipéptidos.

Una construcción de fusión que comprende un polipéptido que tiene actividad de CD y un polipéptido que comprende actividad UPRT u OPRT puede modificarse para que contenga un sitio de escisión para ayudar en la recuperación de proteínas u otro resto enlazador. Típicamente, un enlazador será un resto enlazador peptídico. La longitud del resto enlazador se elige para optimizar la actividad biológica del polipéptido CD y el polipéptido UPRT u

OPRT y puede determinarse empíricamente sin experimentación excesiva. El resto enlazador debe ser lo suficientemente largo y flexible para permitir que un sustrato interaccione con el polipéptido CD y un sustrato con el polipéptido UPRT o OPRT. Un resto enlazador es un péptido de entre aproximadamente uno y 30 residuos de aminoácidos de longitud, típicamente entre aproximadamente dos y 15 residuos de aminoácidos. Ejemplos de grupos enlazadores son --Gly - Gly--, GGGGS (SEQ ID NO:22), (GGGGS)_N (SEQ ID NO:23), (SGGGG)_N (SEQ ID NO: 24), GKSSGSGSESKS (SEQ ID NO:25), GSTSGSGKSSEGKG (SEQ ID NO:26), GSTSGSGKSSEGSGSTKG (SEQ ID NO: 27), GSTSGSGKPGSGEGSTKG (SEQ ID NO:28) o EGKSSGSGSESKEF (SEQ ID NO:29). Los restos de unión se describen, por ejemplo, en Huston et al., Proc. Nat'l Acad. Sci 85:5879, 1988; Whitlow et al., Protein Engineering 6: 989, 1993; y Newton et al., Biochemistry 35:545, 1996. Otros enlazadores peptídicos adecuados son los descritos en las patentes de Estados Unidos n.º 4.751.180 y 4.935.233. Una secuencia de ADN que codifica un enlazador peptídico deseado puede insertarse entre, y en el mismo marco de lectura que un polinucleótido que codifica un polipéptido CD seguido de un polipéptido UPRT o OPRT, utilizando cualquier técnica convencional adecuada. Por ejemplo, un oligonucleótido sintetizado químicamente que codifica el enlazador puede ligarse entre dos a cuatro dominios separados (por ejemplo, un dominio CD y un UPRT u OPRT) están separados por enlazadores peptídicos.

"Sustitución conservadora de aminoácidos" o, simplemente, "variaciones conservadoras" de una secuencia particular se refiere al reemplazo de un aminoácido, o serie de aminoácidos, con secuencias de aminoácidos esencialmente idénticas. Un experto reconocerá que las sustituciones, deleciones o adiciones individuales que alteran, añaden o eliminan un solo aminoácido o un porcentaje de aminoácidos en una secuencia codificada dan lugar a "variaciones conservadoras" donde las alteraciones dan como resultado la deleción de un aminoácido, la adición de un aminoácido o la sustitución de un aminoácido con un aminoácido químicamente similar.

Las tablas de sustituciones conservadoras que proporcionan aminoácidos funcionalmente similares son bien conocidas en la técnica. Por ejemplo, un grupo de sustitución conservadora incluye alanina (A), serina (S) y treonina (T). Otro grupo de sustitución conservadora incluye ácido aspártico (D) y ácido glutámico (E). Otro grupo de sustitución conservadora incluye asparagina (N) y glutamina (Q). Aún otro grupo de sustitución conservadora incluye arginina (R) y lisina (K). Aún otro grupo de sustitución conservadora incluye isoleucina, (I), leucina (L), metionina (M) y valina (V). Otro grupo de sustitución conservadora incluye fenilalanina (F), tirosina (Y) y triptófano (W).

Por lo tanto, "sustituciones de aminoácidos conservadoras" de una secuencia polipeptídica enumerada (por ejemplo, SEQ ID NO: 2, 4, 6, 8, 10, 11, 12, 13, etc.) incluyen sustituciones de un porcentaje, típicamente menos del 10 %, de los aminoácidos de la secuencia polipeptídica, con un aminoácido seleccionado de forma conservadora del mismo grupo de sustitución conservadora. Por consiguiente, una variación sustituida de forma conservadora de un polipéptido de la divulgación puede contener 100, 75, 50, 25 o 10 sustituciones con una variación sustituida de forma conservadora del mismo grupo de sustitución conservadora.

La "actividad" de una enzima es una medida de su capacidad para catalizar una reacción, es decir, para "funcionar", y puede expresarse como la velocidad a la que se produce el producto de la reacción. Por ejemplo, la actividad enzimática se puede representar como la cantidad de producto producido por unidad de tiempo o por unidad de enzima (por ejemplo, concentración o peso), o en términos de afinidad o constantes de disociación. Como se usa indistintamente en el presente documento, una "actividad citosina desaminasa", "actividad biológica de la citosina desaminasa" o "actividad funcional de una citosina desaminasa", se refiere a una actividad ejercida por una proteína citosina desaminasa, o polipéptido de la divulgación, en un sustrato de citosina desaminasa, según lo determinado *in vivo* o *in vitro*, de acuerdo con las técnicas estándar. Los ensayos para medir la actividad de la citosina desaminasa son conocidos en la técnica. Por ejemplo, una actividad de citosina desaminasa puede medirse determinando la tasa de conversión de 5-FC en 5-FU o de citosina en uracilo. La detección de 5-FC, 5-FU, citosina y uracilo se puede realizar por cromatografía y otros procedimientos conocidos en la técnica.

Un experto en la técnica apreciará que muchas variaciones conservadoras de las construcciones de ácido nucleico que se describen producen una construcción funcionalmente idéntica. Por ejemplo, tal como se ha tratado anteriormente, debido a la degeneración del código genético, "sustituciones silenciosas" (es decir, sustituciones en una secuencia de ácido nucleico que no producen una alteración en un polipéptido codificado) son una característica implícita de cada secuencia de ácido nucleico que codifica un aminoácido. Los expertos en la técnica apreciarán que, debido a la degeneración del código genético, se puede producir una multitud de secuencias de nucleótidos que codifican polipéptidos de citosina desaminasa modificada de la divulgación, algunas de las cuales tienen una identidad sustancial con las secuencias de ácido nucleico desveladas explícitamente en el presente documento. Por ejemplo, los codones AGA, AGG, CGA, CGC, CGG y CGU codifican todos el aminoácido arginina. Por lo tanto, en cada posición en los ácidos nucleicos de la divulgación, donde una arginina está especificada por un codón, el codón puede alterarse a cualquiera de los codones correspondientes descritos anteriormente sin alterar el polipéptido codificado. Se entiende que U en una secuencia de ARN corresponde a T en una secuencia de ADN.

De forma similar, "sustituciones conservadoras de aminoácidos" en uno o unos pocos aminoácidos en una secuencia de aminoácidos están sustituidos con diferentes aminoácidos con propiedades altamente similares, también se identifican fácilmente como muy similares a una construcción desvelada. Dichas variaciones

conservadoras de cada secuencia desvelada son una característica de los polipéptidos proporcionados en el presente documento.

Las "variantes conservadoras" son proteínas o enzimas en las que un residuo de aminoácido dado se ha modificado sin alterar la conformación y función general de la proteína o enzima, incluyendo, pero sin limitación, la sustitución de un aminoácido por uno con propiedades similares, incluyendo caracteres polares o no polares, tamaño, forma y carga. Los aminoácidos distintos de los indicados como conservados pueden diferir en una proteína o enzima, de modo que el porcentaje de similitud de la secuencia de la proteína o aminoácido entre cualquiera de las dos proteínas de función similar puede variar y puede ser, por ejemplo, al menos el 30 %, al menos el 50 %, al menos el 70 %, al menos el 80 %, o al menos el 90 %, según lo determinado de acuerdo con un esquema de alineación. Como se cita en el presente documento, "similitud de secuencia" significa el grado en que las secuencias de nucleótidos o de proteínas están relacionadas. El grado de similitud entre dos secuencias puede basarse en el porcentaje de identidad de secuencia y / o conservación. La "identidad de secuencia" en el presente documento significa el grado en que dos secuencias de nucleótidos o aminoácidos son invariables. "Alineación de secuencias" significa el proceso de alinear dos o más secuencias para lograr niveles máximos de identidad (y, en el caso de secuencias de aminoácidos, conservación) con el fin de evaluar el grado de similitud. Se conocen numerosos procedimientos para alinear secuencias y evaluar la similitud / identidad en la técnica, tales como, por ejemplo, el procedimiento de Cluster, donde la similitud se basa en el algoritmo MEGALIGN, así como BLASTN, BLASTP y FASTA (Lipman y Pearson, 1985; Pearson y Lipman, 1988). Al usar todos estos programas, las configuraciones preferidas son aquellas que resultan en la mayor similitud de secuencia. Por ejemplo, la "identidad" o "porcentaje de identidad" con respecto a un par particular de secuencias de aminoácidos alineadas puede referirse al porcentaje de identidad de secuencia de aminoácidos que se obtiene mediante análisis ClustalW (versión W 1.8 disponible en el Instituto Europeo de Bioinformática, Cambridge, Reino Unido), contando el número de coincidencias idénticas en la alineación y dividiendo dicho número de coincidencias idénticas por la mayor de (i) la longitud de las secuencias alineadas, y (ii) 96, y utilizando los siguientes parámetros ClustalW predeterminados para lograr alineaciones pareadas lentas / precisas: Gap Penalización abierta: 10; penalización por extensión huecos: 0,10; matriz del peso de la proteína: serie Gonnet; matriz del peso del ADN: IUB; alternar alineaciones pareadas lentas / rápidas = alineación LENTA O COMPLETA.

Dos secuencias están "alineadas de manera óptima" cuando están alineadas para una puntuación de similitud utilizando una matriz de sustitución de aminoácidos definida (por ejemplo, BLOSUM62), la penalización de existencia de brecha y la penalización de extensión de brecha para llegar a la puntuación más alta posible para ese par de secuencias. Las matrices de sustitución de aminoácidos y su uso para cuantificar la similitud entre dos secuencias son bien conocidas en la técnica y se describen, por ejemplo, en Dayhoff et al. (1978) "A model of evolutionary change in proteins" en "Atlas of Protein Sequence and Structure," Vol. 5, Suppl. 3 (ed. M. O. Dayhoff), pp. 345-352. Natl Biomed. Res. Found., Washington, D.C. y Henikoff et al. (1992) Proc. Nat'l. Acad. Sci. USA 89: 10915-10919. La matriz BLOSUM62 (FIG. 10) se utiliza a menudo como una matriz de sustitución de puntuación predeterminada en los protocolos de alineación de secuencias como Gapped BLAST 2.0. La penalización por existencia de brecha se impone por la introducción de una brecha de un solo aminoácido en una de las secuencias alineadas, y la penalización por extensión de brecha se impone por cada posición de aminoácido vacía adicional insertada en una brecha ya abierta. La alineación se define por las posiciones de los aminoácidos de cada secuencia donde la alineación comienza y termina, y, opcionalmente, por la inserción de una brecha o múltiples brechas en una o ambas secuencias para llegar a la puntuación más alta posible. Si bien la alineación y puntuación óptimas se pueden lograr manualmente, el proceso se facilita mediante el uso de un algoritmo de alineación implementado por ordenador, por ejemplo, BLAST 2.0 con huecos, descrito en Altschul et al. (1997) Nucl. Acids Res. 25: 3389-3402, y se puso a disposición del público en el sitio web del Centro Nacional de Información Biotecnológica (NCBI) (www.ncbi.nlm.nih.gov). Las alineaciones óptimas, incluyendo múltiples alineaciones, se puede preparar utilizando, por ejemplo, PSI-BLAST, disponible a través del sitio web de NCBI y descrito por Altschul et al. (1997) Nucl. Acids Res. 25:3389-3402.

Con respecto a una secuencia de aminoácidos que está alineada de manera óptima con una secuencia de referencia, un residuo de aminoácido "corresponde a" la posición en la secuencia de referencia con la cual el residuo está emparejado en la alineación. La "posición" se denota mediante un número que identifica secuencialmente cada aminoácido en la secuencia de referencia en función de su posición en relación con el extremo N-terminal. Debido a las delecciones, inserciones, truncamientos, fusiones, etc., que deben tenerse en cuenta al determinar una alineación óptima, en general, el número de residuos de aminoácidos en una secuencia de prueba como se determina simplemente contando desde el N-terminal no será necesariamente el mismo que el número de su posición correspondiente en la secuencia de referencia. Por ejemplo, en un caso en el que hay una delección en una secuencia de prueba alineada, no habrá ningún aminoácido que corresponda a una posición en la secuencia de referencia en el sitio de la delección. Cuando hay una inserción en una secuencia de referencia alineada, dicha inserción no corresponderá a ninguna posición de aminoácido en la secuencia de referencia. En el caso de truncamientos o fusiones, puede haber tramos de aminoácidos en la secuencia de referencia o alineada que no corresponden a ningún aminoácido en la secuencia correspondiente.

Las modificaciones no conservadoras de un polipéptido particular son aquellas que sustituyen cualquier aminoácido no caracterizado como una sustitución conservadora. Por ejemplo, cualquier sustitución que cruce los límites de los

seis grupos establecidos anteriormente. Estos incluyen sustituciones de aminoácidos básicos o ácidos por aminoácidos neutros (por ejemplo, Asp, Glu, Asn o Gln por Val, Ile, Leu o Met), aminoácidos aromáticos por aminoácidos básicos o ácidos (por ejemplo, Phe, Tyr o Trp por Asp, Asn, Glu o Gln) o cualquier otra sustitución que no reemplace un aminoácido con un aminoácido similar. Las cadenas laterales básicas incluyen lisina (K), arginina (R), histidina (H); las cadenas laterales ácidas incluyen ácido aspártico (D), ácido glutámico (E); las cadenas laterales polares sin carga incluyen glicina (G), asparagina (N), glutamina (Q), serina (S), treonina (T), tirosina (Y), cisteína (C); las cadenas laterales no polares incluyen alanina (A), valina (V), leucina (L), isoleucina (I), prolina (P), fenilalanina (F), metionina (M), triptófano (W); las cadenas laterales beta-ramificadas incluyen treonina (T), valina (V), isoleucina (I); las cadenas laterales aromáticas incluyen tirosina (Y), fenilalanina (F), triptófano (W), histidina (H).

Por consiguiente, algunos residuos de aminoácidos en posiciones específicas en un polipéptido están "excluidos" de las sustituciones conservadoras de aminoácidos. Por ejemplo, la divulgación proporciona un polipéptido que comprende una secuencia como se expone en la SEQ ID NO:4. En algunos aspectos de la divulgación, el polipéptido puede alterarse con sustituciones conservadoras de aminoácidos como se ha descrito anteriormente, sin embargo, es deseable que ciertos residuos queden sin sustituir, como los residuos en las posiciones 23, 108 y 140.

Una proteína, enzima, polinucleótido, gen o célula original es cualquier proteína, enzima, polinucleótido, gen, o célula, de la cual cualquier otra proteína, enzima, polinucleótido, gen, o célula, deriva o se fabrica, utilizando cualquier procedimiento, herramienta o técnica, y si el padre es o no nativo o mutante. Un gen o polinucleótido original codifica una proteína o enzima original.

Un polinucleótido, polipéptido u otro componente está "aislado" cuando está parcial o completamente separado de los componentes con los que normalmente está asociado (otras proteínas, ácidos nucleicos, células, reactivos sintéticos, etc.). Un ácido nucleico o polipéptido es "recombinante" cuando es artificial o está modificado, o deriva de una proteína o ácido nucleico artificial o modificado. Por ejemplo, un polinucleótido que se inserta en un vector o en cualquier otra ubicación heteróloga, por ejemplo, en un genoma de un organismo recombinante, de modo que no se asocia con secuencias de nucleótidos que normalmente flanquean al polinucleótido, ya que se encuentra en la naturaleza, es un polinucleótido recombinante. Una proteína expresada *in vitro* o *in vivo* a partir de un polinucleótido recombinante es un ejemplo de un polipéptido recombinante. Del mismo modo, una secuencia de polinucleótidos que no aparece en la naturaleza, por ejemplo, una variante de un gen de origen natural, es recombinante.

En otros aspectos de la divulgación, se proporcionan polinucleótidos aislados que codifican un polipéptido de citosina desaminasa o construcción de fusión de la divulgación. En un aspecto, la divulgación proporciona un polinucleótido aislado o recombinante referido en el presente documento como "polinucleótidos optimizados para CD", "polinucleótidos CD" o "polinucleótidos de fusión CD". Los términos "polinucleótido", "secuencia de nucleótidos" y "molécula de ácido nucleico" se utilizan para referirse a un polímero de nucleótidos (A, C, T, U, G, etc. o análogos de nucleótidos artificiales o naturales), por ejemplo, ADN o ARN, o una representación de los mismos, por ejemplo, una cadena de caracteres, etc., dependiendo del contexto relevante. Un polinucleótido dado o un polinucleótido complementario se puede determinar a partir de cualquier secuencia de nucleótidos especificada.

Un aspecto de la divulgación se refiere a polinucleótidos aislados que codifican una citosina desaminasa o un polipéptido de citosina desaminasa mutante o partes biológicamente activas del mismo. Tal como se utiliza en el presente documento, el término "molécula de ácido nucleico" o "polinucleótido" pretende incluir moléculas de ADN (por ejemplo, ADNc o ADN genómico) y moléculas de ARN (por ejemplo, ARNm) y análogos del ADN o ARN generados usando análogos de nucleótidos. La molécula de ácido nucleico puede ser monocatenaria o bicatenaria.

En un aspecto de la divulgación, un polinucleótido CD optimizado o polinucleótido CD comprende formas recombinantes, modificadas o aisladas de ácidos nucleicos de origen natural aislados de un organismo, por ejemplo, una cepa bacteriana o de levadura. Los polinucleótidos de CD de ejemplo incluyen aquellos que codifican los polipéptidos de tipo salvaje expuestos en la SEQ ID NO: 2. En otro aspecto de la divulgación, los polinucleótidos de CD se producen mediante diversificación, por ejemplo, recombinación y / o mutación de uno o más polinucleótidos de CD de origen natural, aislados o recombinantes. Como se describe con más detalle en otra parte del presente documento, es posible generar polinucleótidos de CD diversificados que codifican polipéptidos de CD con características funcionales superiores, por ejemplo, aumento de la función catalítica, mayor estabilidad o mayor nivel de expresión, que un polinucleótido de CD utilizado como sustrato o padre en el proceso de diversificación. Los polinucleótidos de ejemplo incluyen las SEQ ID NO : 3 y 5 y aquellos que codifican los polipéptidos variantes de CD de la divulgación.

Los polinucleótidos de la divulgación tienen diversos usos en, por ejemplo, la producción recombinante (es decir, expresión) de los polipéptidos de CD de la divulgación y como sustratos para una mayor generación de diversidad, por ejemplo, reacciones de recombinación o reacciones de mutación para producir homólogos de CD nuevos y / o mejorados, y similares.

Es importante tener en cuenta que ciertas utilidades específicas, sustanciales y creíbles de los polinucleótidos de CD no requieren que el polinucleótido codifique un polipéptido con actividad de CD sustancial o incluso actividad de CD variante. Por ejemplo, los polinucleótidos de CD que no codifican enzimas activas pueden ser fuentes valiosas de

polinucleótidos originales para su uso en procedimientos de diversificación para llegar a variantes de polinucleótidos de CD, o polinucleótidos que no son de CD, con propiedades funcionales deseables (por ejemplo, k_{cat} alta o k_{cat}/K_m , K_m baja, alta estabilidad frente al calor u otros factores ambientales, altas tasas de transcripción o traducción, resistencia a la escisión proteolítica, etc.).

Los polinucleótidos de CD, incluyendo las secuencias de nucleótidos que codifican los polipéptidos de CD y variantes de los mismos, fragmentos de polipéptidos de CD, proteínas de fusión relacionadas o equivalentes funcionales de los mismos, se utilizan en moléculas de ADN recombinante que dirigen la expresión de los polipéptidos de CD en las células huésped apropiadas, tales como células bacterianas. Debido a la degeneración inherente del código genético, También se pueden usar otras secuencias de ácido nucleico que codifican sustancialmente la misma o una secuencia de aminoácidos funcionalmente equivalente para clonar y expresar los polinucleótidos de CD. La expresión "célula huésped", como se usa en el presente documento, incluye cualquier tipo de célula que sea susceptible de transformación con una construcción de ácido nucleico. El término "transformación" significa la introducción de un gen (es decir, extrínseco o extracelular), secuencia de ADN o ARN extraños en una célula huésped, de modo que la célula huésped expresará el gen o secuencia introducidos para producir una sustancia deseada, típicamente una proteína o enzima codificada por el gen o secuencia introducidos. El gen o secuencia introducidos pueden incluir secuencias reguladoras o de control, tales como las secuencias de inicio, terminación, promotor, señal, secreción u otras secuencias utilizadas por la maquinaria genética de la célula. Una célula huésped que recibe y expresa el ADN o ARN introducido se ha "transformado" y es un "transformante" o un "clon". El ADN o ARN introducido en una célula huésped puede provenir de cualquier fuente, incluidas las células del mismo género o especie que la célula huésped, o las células de un género o especie diferente.

Como apreciarán los expertos en la materia, puede ser ventajoso modificar una secuencia de codificación para mejorar su expresión en un huésped particular. El código genético es redundante con 64 codones posibles, pero la mayoría de los organismos utilizan preferentemente un subconjunto de estos codones. Los codones que se utilizan con mayor frecuencia en una especie se llaman codones óptimos y los que no se utilizan muy a menudo se clasifican como codones raros o de bajo uso (consulte, por ejemplo, Zhang et al. (1991) Gene 105:61-72). Los codones se pueden sustituir para reflejar el uso de codones preferido del huésped, un proceso a veces llamado "optimización de codones" o "control del sesgo de codones de especies".

Se pueden preparar secuencias codificantes optimizadas que contienen codones preferidos por un huésped procariótico o eucariótico particular (véase también Murray et al. (1989) Nucl. Acids Res. 17: 477-508), por ejemplo, para aumentar la velocidad de traducción o para producir transcritos de ARN recombinante que tengan propiedades deseables, como una vida media más larga, en comparación con los transcritos producidos a partir de una secuencia no optimizada. Los codones de terminación de la traducción también pueden modificarse para reflejar las preferencias del huésped. Por ejemplo, los codones de terminación preferidos para *S. cerevisiae* y mamíferos son UAA y UGA, respectivamente. El codón de terminación preferido para plantas monocotiledóneas es UGA, mientras que los insectos y *E. coli* prefieren usar UAA como el codón de terminación (Dalphin et al. (1996) Nucl. Acids Res. 24: 216-218).

El sesgo del uso del codón se refiere a las diferencias entre los organismos en la frecuencia de aparición de los codones en las secuencias de ADN de codificación genética (genes). Un codón es una serie de tres nucleótidos (tripletes) que codifica un residuo de aminoácido específico en una cadena polipeptídica. Porque hay cuatro nucleótidos en el ADN, adenina (A), guanina (G), citosina (C) y timina (T), hay 64 posibles tripletes que codifican 20 aminoácidos y tres codones de terminación de la traducción (sin sentido). Debido a esta redundancia, todos menos dos aminoácidos están codificados por más de un triplete. Los diferentes organismos a menudo muestran preferencias particulares para uno de los varios codones que codifican el mismo aminoácido. Cómo surgen estas preferencias es un área muy debatida de la evolución molecular.

En general, se reconoce que las preferencias de codón reflejan un equilibrio entre los sesgos mutacionales y la selección natural para la optimización de la traducción. Los codones óptimos en microorganismos de rápido crecimiento, como *Escherichia coli* o *Saccharomyces cerevisiae* (levadura de panadero), reflejan la composición de su respectivo conjunto de ARNt genómico. Se cree que los codones óptimos ayudan a lograr velocidades de traducción más rápidas y una alta precisión. Como resultado de estos factores, se espera que la selección traduccional sea más fuerte en genes altamente expresados, como es de hecho el caso de los organismos antes mencionados. En otros organismos que no muestran altas tasas de crecimiento o que presentan genomas pequeños, la optimización del uso de codones normalmente está ausente y las preferencias de codones están determinadas por los sesgos mutacionales característicos observados en ese genoma en particular. Ejemplos de esto son *Homo sapiens* (humano) y *Helicobacter pylori*. Los organismos que muestran un nivel intermedio de optimización del uso de codones incluyen *Drosophila melanogaster* (mosca de la fruta), *Caenorhabditis elegans* (gusano nematodo) o *Arabidopsis thaliana* (berro de pared).

El término "secuencias optimizadas de codones" generalmente se refiere a secuencias de nucleótidos que se han optimizado para una especie de huésped en particular al reemplazar cualquier codón que tenga una frecuencia de uso de menos de aproximadamente el 20%. Las secuencias de nucleótidos que se han optimizado para la expresión en una especie huésped dada mediante la eliminación de secuencias de poliadenilación espurias, eliminación de

señales de empalme de exón / intrón, eliminación de repeticiones similares a transposones y / u optimización del contenido de GC además de la optimización de codones se denominan en el presente documento "secuencias mejoradas de expresión".

Tabla 1: El uso de codones humanos y la preferencia de codones. Para cada codón, la tabla muestra la frecuencia de uso de cada codón (por mil) en las regiones de codificación humana (primera columna) y la frecuencia relativa de cada codón entre los codones sinónimos (segunda columna).

La tabla de uso de codones humanos															
Gly	GGG	17,08	0,23	Arg	AGG	12,09	0,22	Trp	TGG	14,74	1,00	Arg	CGG	10,40	0,19
Gly	GGA	19,31	0,26	Arg	AGA	11,73	0,21	Final	TGA	2,64	0,61	Arg	CGA	5,63	0,10
Gly	GGT	13,66	0,18	Ser	AGT	10,18	0,14	Cys	TGT	9,99	0,42	Arg	CGT	5,16	0,09
Gly	GGC	24,94	0,33	Ser	AGC	18,54	0,25	Cys	TGC	13,86	0,58	Arg	CGC	10,82	0,19
Glu	GAG	38,82	0,59	Lys	AAG	33,79	0,60	Final	TAG	0,73	0,17	Gln	CAG	32,95	0,73
Glu	GAA	27,51	0,41	Lys	AAA	22,32	0,40	Final	TAA	0,95	0,22	Gln	CAA	11,94	0,27
Asp	GAT	21,45	0,44	Asn	AAT	16,43	0,44	Tyr	TAT	11,80	0,42	His	CAT	9,56	0,41
Asp	GAC	27,06	0,56	Asn	AAC	21,30	0,56	Tyr	TAC	16,48	0,58	His	CAC	14,00	0,59
Val	GTG	28,60	0,48	Met	ATG	21,86	1,00	Leu	TTG	11,43	0,12	Leu	CTG	39,93	0,43
Val	GTA	6,09	0,10	Ile	ATA	6,05	0,14	Leu	TTA	5,55	0,06	Leu	CTA	6,42	0,07
Val	GTT	10,30	0,17	Ile	ATT	15,03	0,35	Phe	TTT	15,36	0,43	Leu	CTT	11,24	0,12
Val	GTC	15,01	0,25	Ile	ATC	22,47	0,52	Phe	TTC	20,72	0,57	Leu	CTC	19,14	0,20
Ala	GCG	7,27	0,10	Thr	ACG	6,80	0,12	Ser	TCG	4,38	0,06	Pro	CCG	7,02	0,11
Ala	GCA	15,50	0,22	Thr	ACA	15,04	0,27	Ser	TCA	10,96	0,15	Pro	CCA	17,11	0,27
Ala	GCT	20,23	0,28	Thr	ACT	13,24	0,23	Ser	TCT	13,51	0,18	Pro	CCT	18,03	0,29
Ala	GCC	28,43	0,40	Thr	ACC	21,52	0,38	Ser	TCC	17,37	0,23	Pro	CCC	20,51	0,33

Por consiguiente, en algunos aspectos de la divulgación, un polinucleótido comprende un codón de molécula optimizado para la traducción en una célula humana. Por ejemplo, las SEQ ID NO: 3 y 5 comprenden secuencias que se han optimizado para producir citosina desaminasa en una célula huésped humana.

- 5 La divulgación proporciona así un polinucleótido optimizado en codón humano que codifica un polipéptido que tiene actividad citosina desaminasa. Un polinucleótido optimizado por codones humanos de la divulgación (por ejemplo, la SEQ ID NO: 5) puede incluir además mutaciones adicionales que dan como resultado una sustitución conservadora de aminoácidos y / o una actividad o estabilidad mejoradas del polipéptido codificado. Por ejemplo, las mutaciones en las posiciones 23, 108 y 140 del polipéptido que comprende la SEQ ID NO:6 proporcionan una mayor
- 10 termoestabilidad. Por consiguiente, la divulgación proporciona un polinucleótido optimizado en codón humano que codifica un polipéptido que tiene actividad citosina desaminasa y una mayor termoestabilidad. En un aspecto de la divulgación, el polinucleótido optimizado en codón humano comprende una secuencia que codifica un polipéptido de la SEQ ID NO: 4, donde los codones que codifican los aminoácidos del polipéptido se han optimizado para la expresión en una célula humana. En otro, un polinucleótido de la divulgación comprende las SEC ID NO: 3 o 5.

- 15 La divulgación proporciona un polinucleótido que comprende un polinucleótido de citosina desaminasa o un polinucleótido o mutante optimizado de codón unido a un polinucleótido heterólogo que codifica un polipéptido que tiene actividad UPRT u OPRT.

- 20 Un polinucleótido de la divulgación que comprende, por ejemplo, una secuencia que codifica el polipéptido de las SEQ ID NO: 4, 8, 10, 12, 14, 16 o 18; o que tiene la secuencia de nucleótidos establecida en cualquiera de las SEQ ID NO: 3, 5, 7, 9, 11, 13, 15 o 17, o una porción de los mismos, se puede aislar y generar utilizando técnicas estándar de biología molecular y la información de secuencia proporcionada en el presente documento.

- 25 Un polinucleótido de la divulgación puede amplificarse utilizando ADNc, ARNm o, como alternativa, ADN genómico, como molde y cebadores de oligonucleótidos apropiados de acuerdo con las técnicas de amplificación por PCR estándar. El ácido nucleico así amplificado puede clonarse en un vector apropiado y caracterizarse por análisis de secuencia de ADN. Adicionalmente, los oligonucleótidos correspondientes a secuencias de nucleótidos pueden prepararse mediante técnicas sintéticas estándar, por ejemplo, usando un sintetizador de ADN automático. En
- 30 algunos aspectos de la divulgación, una molécula de ácido nucleico aislada de la divulgación comprende una molécula de ácido nucleico que es complementaria de una secuencia nucleotídica que codifica un polipéptido expuesto en cualquiera de las SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, o 18, o que tiene la secuencia nucleotídica expuesta en cualquiera de las SEQ ID NO :1, 3, 5, 7, 9, 11, 13, 15, o 17.

- 35 En otro aspecto de la divulgación, un polinucleótido aislado de la divulgación comprende una molécula de ácido nucleico que hibrida en condiciones rigurosas a una molécula de ácido nucleico que consiste en la secuencia de nucleótidos que codifica un polipéptido expuesto en cualquiera de las SEQ NO: 4, 12, 14, 16 o 18, o que consiste en la secuencia de nucleótidos expuesta en cualquiera de las SEQ ID NO:3, 5, 11, 13, 15 o 17. En otros aspectos de la divulgación, el ácido nucleico tiene al menos 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550 o 600
- 40 nucleótidos de longitud. Las moléculas de ácido nucleico son "hibridables" entre sí cuando al menos una hebra de un polinucleótido puede unirse a otro polinucleótido en condiciones de rigurosidad definidas. Se determina la rigurosidad de la hibridación, por ejemplo, por (a) la temperatura a la que se realiza la hibridación y / o lavado, y (b) la fuerza iónica y la polaridad (por ejemplo, formamida) de las soluciones de hibridación y lavado, así como otros parámetros. La hibridación requiere que los dos polinucleótidos contienen secuencias sustancialmente
- 45 complementarias; dependiendo de la rigurosidad de la hibridación, sin embargo, se pueden tolerar emparejamientos erróneos. Típicamente, la hibridación de dos secuencias en alta rigurosidad (como, por ejemplo, por ejemplo, en una solución acuosa de 0,5 X SSC a 65 °C) se requiere que las secuencias exhiban un alto grado de complementariedad en toda su secuencia. Las condiciones de rigurosidad intermedia (tales como, por ejemplo, una solución acuosa de 2 X SSC a 65 ° C) y baja rigurosidad (como, por ejemplo, una solución acuosa de 2 X SSC a 55 °C), requieren
- 50 correspondientemente menos complementariedad global entre las secuencias de hibridación (1 X SSC es NaCl 0,15 M, citrato Na 0,015 M). Las moléculas de ácido nucleico que hibridan incluyen aquellas que hibridan en condiciones de rigurosidad adecuadas y que codifican polipéptidos o enzimas que tienen la misma función, tal como la capacidad de catalizar la conversión de 5-fluorocitosina (5-FC) en 5-fluorouracilo (5-FU), de la divulgación. Además, la expresión "hibrida en condiciones rigurosas" pretende describir las condiciones para la hibridación y el lavado bajo las cuales
- 55 las secuencias de nucleótidos al menos el 30 %, 40 %, 50%, o el 60% de homólogos entre sí, típicamente permanecen hibridados entre sí. Preferiblemente, las condiciones son tales que las secuencias al menos aproximadamente el 70 %, más preferentemente al menos aproximadamente 80 %, incluso más preferentemente, al menos aproximadamente el 85 % o 90 % de los homólogos entre sí, típicamente permanecen hibridados entre sí. En algunos casos, una molécula de ácido nucleico aislada de la divulgación hibrida en condiciones rigurosas con una
- 60 secuencia de ácido nucleico que codifica un polipéptido expuesto en cualquiera de las SEC ID NO: 4, 12, 14, 16 o 18, o que tiene la secuencia de nucleótidos expuesta en cualquiera de las SEQ ID NO: 3, 5, 11, 13, 15 o 17. Tal como se utiliza en el presente documento, una molécula de ácido nucleico "de origen natural" se refiere a una molécula de ARN o ADN que tiene una secuencia de nucleótidos que ocurre en la naturaleza (por ejemplo, codifica una proteína natural).

- 65 El experto en la materia apreciará que pueden introducirse cambios por mutación en las secuencias de nucleótidos

de cualquier polinucleótido que codifica un polipéptido expuesto en cualquiera de las SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16 o 18, o que tienen la secuencia de nucleótidos expuesta en cualquiera de las SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15 o 17, conduciendo así a cambios en la secuencia de aminoácidos de las proteínas codificadas. En algunos casos, la alteración conducirá a una función alterada del polipéptido. En otros casos, el cambio no alterará la capacidad funcional del polipéptido codificado. En general, Las sustituciones que no alteran la función de un polipéptido incluyen sustituciones de nucleótidos que conducen a sustituciones de aminoácidos en restos de aminoácidos "no esenciales". Un resto de aminoácido "no esencial" es un resto que puede alterarse de la secuencia original sin alterar la actividad biológica del polipéptido resultante, por ejemplo, catalizando la conversión de 5-FC en 5-FU.

También se contemplan las situaciones en las que es deseable alterar la actividad de un polipéptido original de manera que el polipéptido tenga una actividad nueva o aumentada sobre un sustrato particular o una mayor estabilidad o una menor degradación. Se entiende que estas sustituciones de aminoácidos generalmente no constituirán sustituciones "conservadoras". En cambio, estas sustituciones constituyen sustituciones no conservadoras introducidas en una secuencia para obtener una actividad nueva o mejorada. Por ejemplo, la SEQ ID NO:1 proporciona la secuencia de ácido nucleico original para la citosina desaminasa de *S. cerevisiae* (la SEC ID NO: 2 proporciona el polipéptido correspondiente). La SEQ ID NO: 3 proporciona la secuencia de ácido nucleico de una secuencia mutante que incluye sustituciones de aminoácidos que imparten mayor estabilidad al polipéptido. Por consiguiente, la molécula de ácido nucleico que codifica la secuencia de aminoácidos de SEQ ID NO: 2 proporciona una molécula "original" de ácido nucleico a partir de la cual se pueden hacer mutaciones para obtener una molécula de ácido nucleico que codifica un polipéptido modificado que incluye sustituciones de aminoácidos.

También se entiende que un polipéptido modificado puede constituir un polipéptido "original" a partir del cual se pueden realizar sustituciones adicionales. Por consiguiente, un polipéptido original y una molécula de ácido nucleico que codifica un polipéptido original, incluye polipéptidos modificados y no solo secuencias "de tipo salvaje". Por ejemplo, el polinucleótido de la SEQ ID NO: 5 es un polinucleótido modificado con respecto a la SEQ ID NO: 1 (es decir, el polinucleótido "original"). De manera similar, el polinucleótido de la SEQ ID NO: 3 es un polinucleótido modificado con respecto a la SEQ ID NO: 5. Por consiguiente, la SEQ ID NO: 5 es la secuencia original de SEQ ID NO: 3.

Los métodos mutacionales de generar diversidad incluyen, por ejemplo, mutagénesis dirigida al sitio (Ling et al. (1997) "Approaches to DNA mutagenesis: an overview" *Anal Biochem.* 254(2): 157-178; Dale et al. (1996) "Oligonucleotide-directed random mutagenesis using the phosphorothioate method" *Methods Mol. Biol.* 57:369-374; Smith (1985) "*In vitro* mutagenesis" *Ann. Rev. Genet.* 19:423-462; Botstein y Shortle (1985) "Strategies and applications of *in vitro* mutagenesis" *Science* 229:1193-1201; Carter (1986) "Site-directed mutagenesis" *Biochem. J.* 237:1-7; y Kunkel (1987) "The efficiency of oligonucleotide directed mutagenesis" in *Nucleic Acids & Molecular Biology* (Eckstein, F. y Lilley, D. M. J. eds., Springer Verlag, Berlín)); mutagénesis utilizando moldes que contienen uracilo (Kunkel (1985) "Rapid and efficient site-specific mutagenesis without phenotypic selection" *Proc. Natl. Acad. Sci. USA* 82:488-492; Kunkel et al. (1987) "Rapid and efficient site-specific mutagenesis without phenotypic selection" *Methods in Enzymol.* 154, 367-382; y Bass et al. (1988) "Mutant Trp repressors with new DNA-binding specificities" *Science* 242:240-245); oligonucleotide-directed mutagenesis (*Methods in Enzymol.* 100: 468-500 (1983); *Methods in Enzymol.* 154: 329-350 (1987); Zoller y Smith (1982) "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment" *Nucleic Acids Res.* 10:6487-6500; Zoller y Smith (1983) "Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors" *Methods in Enzymol.* 100:468-500; y Zoller & Smith (1987) "Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template" *Methods in Enzymol.* 154:329-350); mutagénesis de ADN modificado con fosforotioato (Taylor et al. (1985) "The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA" *Nucl. Acids Res.* 13: 8749-8764; Taylor et al. (1985) "The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA" *Nucl. Acids Res.* 13: 8765-8787; Nakamaye y Eckstein (1986) "Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis" *Nucl. Acids Res.* 14: 9679-9698; Sayers et al. (1988) "Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis" *Nucl. Acids Res.* 16:791-802; y Sayers et al. (1988) "Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide" *Nucl. Acids Res.* 16: 803-814); mutagénesis utilizando ADN dúplex con huecos (Kramer et al. (1984) "The gapped duplex DNA approach to oligonucleotide-directed mutation construction" *Nucl. Acids Res.* 12: 9441-9456; Kramer y Fritz (1987) *Methods in Enzymol.* "Oligonucleotide-directed construction of mutations via gapped duplex DNA" 154:350-367; Kramer et al. (1988) "Improved enzymatic *in vitro* reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations" *Nucl. Acids Res.* 16: 7207; y Fritz et al. (1988) "Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions *in vitro*" *Nucl. Acids Res.* 16: 6987-6999).

Métodos adecuados adicionales incluyen reparación de desajuste de puntos (Kramer et al. (1984) "Point Mismatch Repair" *Cell* 38:879-887), mutagénesis usando cepas de hospedador deficientes en la reparación (Carter et al. (1985) "Improved oligonucleotide site-directed mutagenesis using M13 vectors" *Nucl. Acids Res.* 13: 4431-4443; y Carter (1987) "Improved oligonucleotide-directed mutagenesis using M13 vectors" *Methods in Enzymol.* 154: 382-403), mutagénesis por delección (Eghedarzadeh & Henikoff (1986) "Use of oligonucleotides to generate large

deletions" Nucl. Acids Res. 14: 5115), selección de restricción y purificación de restricción (Wells et al. (1986) "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin" Phil. Trans. R. Soc. Lond. A 317: 415-423), mutagénesis por síntesis génica total (Nambiar et al. (1984) "Total synthesis and cloning of a gene coding for the ribonuclease S protein" Science 223: 1299-1301; Sakamar y Khorana (1988) "Total synthesis and expression of a gene for the α -subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)" Nucl. Acids Res. 14: 6361-6372; Wells et al. (1985) "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites" Gene 34:315-323; y Grundstrom et al. (1985) "Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis" Nucl. Acids Res. 13: 3305-3316); reparación de rotura de doble hebra (Mandecki (1986); Arnold (1993) "Protein engineering for unusual environments" Current Opinion in Biotechnology 4:450-455; y "Oligonucleotide-directed double-strand break repair in plasmids of *Escherichia coli*: a method for site-specific mutagenesis" Proc. Natl. Acad. Sci. USA, 83:7177-7181). Se pueden encontrar detalles adicionales sobre muchos de los métodos anteriores en Methods in Enzymology Volume 154, que también describe controles útiles para solucionar problemas con varios métodos de mutagénesis.

Se pueden encontrar detalles adicionales sobre varios métodos de generación de diversidad en las siguientes patentes de Estados Unidos, publicaciones del PCT, y publicaciones de la EPO: patente de Estados Unidos n.º 5.605.793 concedida a Stemmer (Feb. 25, 1997), "Methods for *In vitro* Recombination;" patente de Estados Unidos n.º 5.811.238 concedida a Stemmer et al. (Sep. 22, 1998) "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" patente de Estados Unidos n.º 5.830.721 concedida a Stemmer et al. (Nov. 3, 1998), "DNA Mutagenesis by Random Fragmentation and Reassembly;" patente de Estados Unidos n.º 5.834.252 concedida a Stemmer et al. (Nov. 10, 1998) "End-Complementary Polymerase Reaction;" patente de Estados Unidos n.º 5.837.458 concedida a Minshull, et al. (Nov. 17, 1998), "Methods and Compositions for Cellular and Metabolic Engineering;" WO 95/22625, Stemmer y Cramer, "Mutagenesis by Random Fragmentation and Reassembly;" el documento WO 96/33207 de Stemmer y Lipschutz "End Complementary Polymerase Chain Reaction;" el documento WO 97/20078 de Stemmer and Cramer "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" documento WO 97/35966 de Minshull y Stemmer, "Methods and Compositions for Cellular and Metabolic Engineering;" documento WO 99/41402 de Punnonen et al. "Targeting of Genetic Vaccine Vectors;" el documento WO 99/41383 de Punnonen et al. "Antigen Library Immunization;" documento WO 99/41369 de Punnonen et al. "Genetic Vaccine Vector Engineering;" documento WO 99/41368 de Punnonen et al. "Optimization of Immunomodulatory Properties of Genetic Vaccines;" documento EP 752008 de Stemmer and Cramer, "DNA Mutagenesis by Random Fragmentation and Reassembly;" documento EP 0932670 de Stemmer "Evolving Cellular DNA Uptake by Recursive Sequence Recombination;" el documento WO 99/23107 de Stemmer et al., "Modification of Virus Tropism and Host Range by Viral Genome Shuffling;" el documento WO 99/21979 de Apt et al., "Human Papillomavirus Vectors;" el documento WO 98/31837 de del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination;" el documento WO 98/27230 de Patten y Stemmer, "Methods and Compositions for Polypeptide Engineering;" el documento WO 98/13487 de Stemmer et al., "Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection;" el documento WO 00/00632, "Methods for Generating Highly Diverse Libraries;" el documento WO 00/09679, "Methods for Obtaining *in vitro* Recombined Polynucleotide Sequence Banks and Resulting Sequences;" el documento WO 98/42832 de Arnold et al., "Recombination of Polynucleotide Sequences Using Random or Defined Primers;" el documento WO 99/29902 de Arnold et al., "Method for Creating Polynucleotide and Polypeptide Sequences;" documento WO 98/41653 de Vind, "An *in vitro* Method for Construction of a DNA Library;" documento WO 98/41622 de Borchert et al., "Method for Constructing a Library Using DNA Shuffling;" documento WO 98/42727 de Pati y Zarling, "Sequence Alterations using Homologous Recombination;" documento WO 00/18906 de Patten et al., "Shuffling of Codon-Altered Genes;" el documento WO 00/04190 de del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Recombination;" el documento WO 00/42561 de Cramer et al., "Oligonucleotide Mediated Nucleic Acid Recombination;" el documento WO 00/42559 de Selifonov and Stemmer "Methods of Populating Data Structures for Use in Evolutionary Simulations;" documento WO 00/42560 de Selifonov et al., "Methods for Making Character Strings, Polynucleotides & Polypeptides Having Desired Characteristics;" el documento WO 01/23401 de Welch et al., "Use of Codon-Variied Oligonucleotide Synthesis for Synthetic Shuffling;" y documento WO 01/64864 "Single-Stranded Nucleic Acid Template-Mediated Recombination and Nucleic Acid Fragment Isolation" de Affholter.

También se proporcionan construcciones recombinantes que comprenden una o más de las secuencias de ácido nucleico como se describe ampliamente en lo que antecede. Las construcciones comprenden un vector, tal como, un plásmido, un cósmido, un fago, virus, un cromosoma bacteriano artificial (CBA), un cromosoma de levadura artificial (CBA), un vector viral o similar, en el que se ha insertado un polinucleótido de la divulgación, en una orientación hacia adelante o hacia atrás. Los expertos en la técnica conocen un número elevado de vectores y promotores adecuados y comercialmente disponibles. En un aspecto de la divulgación, el vector viral es un vector retroviral.

Los términos "vector", "construcción de vector" y "vector de expresión" significan el vehículo por el cual una secuencia de ADN o ARN (por ejemplo, un gen extraño) se puede introducir en una célula huésped, con el fin de transformar el huésped y promover la expresión (por ejemplo, la transcripción y la traducción) de la secuencia introducida. Los vectores comprenden normalmente el ADN de un agente transmisible, donde el ADN extraño que codifica una proteína se inserta mediante la tecnología de la enzima de restricción. Un tipo común de vector es un "plásmido", que generalmente es una molécula de ADN bicatenario autocontenida que puede aceptar fácilmente el

ADN (extraño) adicional y que se puede introducir fácilmente en una célula hospedadora adecuada. Se ha descrito un gran número de vectores, incluyendo vectores plasmídicos y fúngicos, para la replicación y / o expresión en una variedad de huéspedes eucarióticos y procarióticos. Los ejemplos no limitantes incluyen plásmidos pKK (Clontech), plásmidos pUC, plásmidos pET (Novagen, Inc., Madison, Wis.), plásmidos pRSET o pREP (Invitrogen, San Diego, Calif.), o plásmidos pMAL (New England Biolabs, Beverly, Mass.), y muchas células huésped adecuadas, utilizando los métodos divulgados o citados en el presente documento o conocidos de otra manera por los expertos en la materia en la técnica relevante. Los vectores de clonación recombinantes incluirán a menudo uno o más sistemas de replicación para la clonación o la expresión, uno o más marcadores para la selección en el hospedador, por ejemplo, resistencia a antibióticos, y uno o más casetes de expresión.

Los términos "expresa" y "expresión" significan permitir u ocasionar que la información en un gen o secuencia de ADN llegue a manifestarse, por ejemplo, producir una proteína activando las funciones celulares implicadas en la transcripción y la traducción de un gen o secuencia de ADN correspondiente. Una secuencia de ADN se expresa en o mediante una célula para formar un "producto de expresión" tal como una proteína. El producto de expresión por sí mismo, por ejemplo, la proteína resultante, también puede decirse que se "expresa" por la célula. Un polinucleótido o polipéptido se expresa de forma recombinante, por ejemplo, cuando se expresa o produce en una célula hospedadora extraña bajo el control de un promotor extraño o nativo, o en una célula hospedadora nativa bajo el control de un promotor extraño.

Los polinucleótidos proporcionados en el presente documento pueden incorporarse en uno cualquiera de diversos vectores de expresión adecuados para expresar un polipéptido. Los vectores adecuados incluyen secuencias de ADN cromosómicas, no cromosómicas y sintéticas, por ejemplo, derivados de SV40; plásmidos bacterianos; ADN de fago; baculovirus; plásmidos de levadura; vectores derivados de combinaciones de plásmidos y ADN de fago, ADN viral, tal como vaccinia, adenovirus, virus de la viruela aviar, pseudorabia, adenovirus, virus adenoasociados, retrovirus y muchos otros. Cualquier vector que transduce material genético en una célula y, si se desea replicación, que es replicable y viable en el huésped relevante puede usarse. Estos vectores de ácido nucleico pueden liberarse utilizando un sistema de administración no viral tal como se describe en et al. Brit J Pharm 157: 166-178 2009 o Kamimura y Liu J AAPS 10:589-595 (2008) o un sistema vírico como se describe a continuación.

En otro aspecto, el polinucleótido que codifica una citosina desaminasa o mutante de la misma se libera con un sistema de administración de genes que es un vector vírico derivado de virus. El vector viral puede ser replicante o no replicante, y puede ser un vector adenoviral, un vector de sarampión, un vector de herpes, un vector retroviral (incluido un vector lentiviral), un vector rabdoviral, tal como un vector viral de estomatitis vesicular, un vector de reovirus, un vector del virus del valle de Séneca, un vector de poxvirus (incluyendo vectores derivados de la viruela animal o vaccinia), un vector de parvovirus (incluyendo un vector de AAV), un vector de alfavirus u otro vector vírico conocido por un experto en la técnica (véase también, por ejemplo, Concepts in Genetic Medicine, ed. Boro Dropulic and Barrie Carter, Wiley, 2008, Hoboken, NJ.; The Development of Human Gene Therapy, ed. Theodore Friedmann, Cold Springs Harbor Laboratory Press, Cold springs Harbor, Nueva York, 1999; Gene and Cell Therapy, ed. Nancy Smyth Templeton, Marcel Dekker Inc., Nueva York, Nueva York, 2000 y Gene Therapy: Therapeutic Mechanism and Strategies, Nancy Smyth Templeton y Danilo D Lasic, Marcel Dekker, Inc., Nueva York, Nueva York, 2000).

En un aspecto de la divulgación, el vector vírico puede ser un vector retroviral competente para la replicación capaz de infectar solo células de mamífero en replicación. Los retrovirus se han clasificado de diversas maneras, pero la nomenclatura se ha normalizado en la última década (véase ICTVdB - The Universal Virus Database, v 4 en la World Wide Web (www.ncbi.nlm.nih.gov/ICTVdb/ICTVdb/) y el libro de texto "Retroviruses" Eds Coffin, Hughs y Varmus, Cold Spring Harbor Press 1997. El vector retroviral competente para la replicación puede comprender un ortoretrovirus o más típicamente un vector de retrovirus gamma. En un aspecto, un vector retroviral competente para la replicación comprende un sitio interno de entrada al ribosoma (IRES) en 5' al polinucleótido que codifica una citosina desaminasa. En un aspecto de la divulgación, el polinucleótido que codifica una citosina desaminasa se encuentra en 3' con respecto a un polinucleótido env de un vector retroviral.

Por consiguiente, en otros aspectos de la divulgación, se proporcionan vectores que incluyen un polinucleótido de la divulgación. En otros aspectos de la divulgación, se proporcionan células huésped transfectadas con una molécula de ácido nucleico de la divulgación, o un vector que incluye un polinucleótido de la divulgación. Las células huésped incluyen células eucarióticas tales como células de levadura, células de insecto o células de animales. Las células huésped también incluyen células procariotas tales como células bacterianas.

Como se ha tratado anteriormente, textos generales que describen técnicas de biología molecular útiles en el presente documento, incluyendo el uso de vectores, promotores y muchos otros temas relevantes, incluyen a Berger y Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology Volume 152, (Academic Press, Inc., San Diego, Calif.) ("Berger"); Sambrook et al., Molecular Cloning--A Laboratory Manual, 2ª ed., Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y., 1989 ("Sambrook") y Current Protocols in Molecular Biology, F. M. Ausubel et al., eds., Current Protocols, una empresa conjunta entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (complementado hasta 1999) ("Ausubel"). Ejemplos de protocolos suficientes para dirigir a personas expertas a través de procedimientos de amplificación *in vitro*, incluyendo la reacción en cadena de la polimerasa (PCR), la reacción en cadena de la ligasa (LCR), la amplificación de la Q β -replicasa y otras técnicas mediadas por la

ARN polimerasa (por ejemplo, NASBA), por ejemplo, para la producción de los ácidos nucleicos homólogos de la divulgación se encuentran en Berger, Sambrook y Ausubel, así como en Mullis et al. (1987) patente de Estados Unidos n.º 4.683.202; Innis et al., eds. 1990) PCR Protocols: A Guide to Methods and Applications (Academic Press Inc. San Diego, Calif.) ("Innis"); Arnheim y Levinson (1 de octubre de 1990) C&EN 36-47; The Journal Of NIH Research (1991) 3: 81-94; Kwok et al. (1989) Proc. Natl. Acad. Sci. USA 86: 1173; Guatelli et al. (1990) Proc. Nat'l. Acad. Sci. USA 87: 1874; Lomell et al. (1989) J. Clin. Chem 35: 1826; Landegren et al. (1988) Science 241: 1077-1080; Van Brunt (1990) Biotechnology 8: 291-294; Wu y Wallace (1989) Gene 4:560; Barringer et al. (1990) Gene 89:117; y Sooknanan y Malek (1995) Biotechnology 13: 563-564. Los procedimientos mejorados para la clonación de ácidos nucleicos amplificados *in vitro* se describen en Wallace et al., la patente de Estados Unidos n.º 5.426.039. Los procedimientos mejorados para amplificar ácidos nucleicos grandes mediante PCR se resumen en Cheng et al. (1994) Nature 369: 684-685 y en las referencias citadas en el mismo, donde se generan amplicones de PCR de hasta 40 kb. Un experto apreciará que esencialmente cualquier ARN puede convertirse en un ADN de doble cadena adecuado para la digestión de restricción, expansión de PCR y secuenciación utilizando transcriptasa inversa y una polimerasa. Véase, por ejemplo, Ausubel, Sambrook y Berger, todos citados anteriormente.

También se proporcionan células huésped modificadas que se transducen (transforman o transfectan) con un vector proporcionado en el presente documento (por ejemplo, un vector de clonación o un vector de expresión), así como la producción de polipéptidos de la divulgación mediante técnicas recombinantes. el vector puede ser, por ejemplo, un plásmido, una partícula viral, un fago, etc. Las células huésped modificadas pueden cultivarse en medios nutritivos convencionales modificados según sea apropiado para activar los promotores, seleccionar transformantes, o amplificar un polinucleótido codificante. Las condiciones de cultivo, tales como la temperatura, el pH y similares, son aquellas utilizadas anteriormente con la célula hospedadora seleccionada para la expresión, y serán evidentes para una persona normalmente experta en la técnica y en las referencias citadas en el presente documento, que incluyen, por ejemplo, Sambrook, Ausubel y Berger, así como, por ejemplo, Freshney (1994) Culture of Animal Cells: A Manual of Basic Technique, 3ª ed. (Wiley-Liss, Nueva York) y las referencias citadas en los anteriores.

Pueden emplearse vectores para transformar un huésped apropiado para permitir que el huésped exprese una proteína o polipéptido. Ejemplos de huéspedes de expresión apropiados incluyen: células bacterianas, como *E. coli*, *B. subtilis*, estreptomicos y *Salmonella typhimurium*; células fúngicas, tales como *Saccharomyces cerevisiae*, *Pichia pastoris* y *Neurospora crassa*; células de insectos tales como *Drosophila* y *Spodoptera frugiperda*; células de mamífero, tales como CHO, células COS, BHK, HEK 293 o melanoma de Bowes; o células vegetales o explantes, etc.

En los sistemas bacterianos, puede seleccionarse una serie de vectores de expresión dependiendo del uso del polipéptido de citosina desaminasa previsto. Por ejemplo, cuando se necesitan grandes cantidades de polipéptido de CD o fragmentos del mismo para la producción comercial o para la inducción de anticuerpos, pueden ser deseables vectores que dirijan la expresión de alto nivel de proteínas de fusión que se purifican fácilmente. Dichos vectores incluyen, pero sin limitaciones, vectores multifuncionales de clonación y expresión de *E. coli* como BLUESCRIPT (Stratagene), donde la secuencia codificante del polipéptido de CD se puede ligar en el vector en el marco con secuencias para la Met amino terminal y los 7 residuos posteriores de beta-galactosidasa de modo que se produzca una proteína híbrida; vectores pIN (Van Heeke y Schuster (1989) J. Biol. Chem. 264: 5503-5509); vectores pET (Novagen, Madison Wis.); y similares.

De forma similar, en la levadura *Saccharomyces cerevisiae*, una serie de vectores que contienen promotores constitutivos o inducibles, como el factor alfa, la alcohol oxidasa y la PGH se pueden usar para la producción de los polipéptidos deCD de la divulgación. Para una revisión, véase Ausubel (citado anteriormente) y Grant et al. (1987) Methods in Enzymology 153:516-544.

La divulgación también proporciona vectores retrovirales competentes para la replicación que tienen estabilidad incrementada en relación con vectores retrovirales anteriores. Este aumento de la estabilidad durante la infección y la replicación es importante para el tratamiento de trastornos de proliferación celular. La combinación de la eficacia de transducción, la estabilidad del transgén y la selectividad diana se proporciona por el retrovirus competente para la replicación. Las composiciones y los métodos proporcionan una estabilidad de inserción y mantiene la actividad de la transcripción del transgén y la viabilidad de la traducción del polipéptido codificado.

Los retrovirus pueden transmitirse horizontal y verticalmente. La transmisión infecciosa eficaz de los retrovirus requiere la expresión sobre la célula diana de receptores que reconocen específicamente las proteínas de la envoltura vírica, aunque los virus pueden usar receptores independientes, rutas de entrada no específicas de baja eficiencia. Además, el tipo de célula diana debe ser capaz de soportar todas las etapas del ciclo de replicación una vez que el virus se ha unido y penetrado. La transmisión vertical se produce cuando el genoma vírico queda integrado en la línea germinal del hospedador. A continuación, el provirus pasará de una generación a otra como si fuera un gen celular. Por tanto, se establecen provirus endógenos que a menudo se encuentran latentes, pero que pueden activarse cuando el hospedador se expone a los agentes adecuados.

Como se ha mencionado anteriormente, el ADN intermedio integrado se denomina provirus. La terapia génica o los sistemas de administración de genes anteriores utilizan procedimientos y retrovirus que requieren la transcripción del

provirus y el ensamblaje en virus infecciosos en presencia de un virus auxiliar apropiado o en una línea celular que contiene secuencias apropiadas que permiten la encapsidación sin la producción coincidente de un virus auxiliar contaminante. Como se describe a continuación, no se requiere un virus auxiliar para la producción del retrovirus recombinante de la divulgación, ya que se proporcionan secuencias para la encapsidación en el genoma proporcionando de esta forma un vector retroviral competente para la replicación para la administración o el tratamiento génico.

Un genoma retroviral útil en los procedimientos y composiciones de la divulgación comprende un ADN proviral que tiene al menos tres genes: el gag, el pol, y el env, estos genes pueden estar flanqueados por una o dos repeticiones terminales largas (LTR) o, en el provirus, están flanqueados por dos repeticiones terminales largas (LTR) y las secuencias que contienen las secuencias que actúan en cis tales como psi. El gen gag codifica las proteínas estructurales internas (matriz, cápside, y nucleocápside); el gen pol codifica la ADN polimerasa dirigida al ARN (transcriptasa inversa), la proteasa y la integrasa; y el gen env codifica las glicoproteínas de la envoltura vírica. Las LTR 5' y/o 3' promueven la transcripción y la poliadenilación del ARN de los viriones. La LTR contiene el resto de secuencias que actúan en cis necesarias para la replicación vírica. Los lentivirus tienen genes adicionales que incluyen *vif*, *vpr*, *tat*, *rev*, *vpu*, *nef* y *vpx* (en HIV-1, HIV-2 y/o SIV).

Adyacentes a las LTR 5' se encuentran las secuencias necesarias para la transcripción inversa del genoma (el sitio de unión al cebador del ARNt) y para la encapsidación eficaz del ARN vírico en partículas (el sitio Psi). Si las secuencias necesarias para la encapsidación (o el empaquetamiento del ARN retroviral en el virión infeccioso) desaparecen del genoma vírico, el resultado es un defecto en cis que evita la encapsidación del ARN vírico genómico. Este tipo de vector modificado es lo que se ha utilizado típicamente en sistemas de administración de genes anteriores (es decir, sistemas que carecen de elementos que son necesarios para la encapsidación del virión).

En un aspecto de la divulgación, la divulgación proporciona un retrovirus recombinante capaz de infectar una célula que no está en división, una célula en división, o una célula que tiene un trastorno de proliferación celular. El retrovirus competente para la replicación recombinante de la divulgación comprende una secuencia de polinucleótido que codifica un GAG vírico, un POL vírico, un ENV viral, un polinucleótido heterólogo precedido por un sitio interno de entrada al ribosoma (IRES) encapsulado dentro de un virión.

La secuencia de ácido nucleico heterólogo está unida operativamente a un IRES. Tal como se utiliza en el presente documento, el término "heterólogo" en la secuencia de ácido nucleico o transgén se refiere a (i) una secuencia que no existe normalmente en un retrovirus natural, (ii) una secuencia que se origina de una especie extraña, o (iii) si procede de la misma especie, puede modificarse sustancialmente a partir de su forma original. Como alternativa, una secuencia de ácido nucleico sin cambiar que no se expresa normalmente en una célula es una secuencia de ácido nucleico heterólogo. En una realización específica, el polinucleótido heterólogo es un polipéptido de la divulgación que tiene actividad citosina desaminasa.

La frase célula "que no se dividen" se refiere a una célula que no sufre mitosis. Las células que no se dividen pueden bloquearse en cualquier punto del ciclo celular, (por ejemplo, G₀/G₁, G_{1/s}, G_{2/M}), siempre que la célula no se esté dividiendo activamente. Para la infección *ex vivo*, una célula en división puede tratarse para bloquear la división celular mediante técnicas estándar utilizadas por los expertos en la técnica, que incluyen, irradiación, tratamiento con afidocolina, inanición sérica e inhibición por contacto. Sin embargo, debe entenderse que la infección *ex vivo* a menudo se realiza sin bloquear las células ya que muchas células ya están detenidas (por ejemplo, células madre). Por ejemplo, un vector de lentivirus recombinante es capaz de infectar cualquier célula que no se divide, independientemente del mecanismo utilizado para bloquear la división celular o el punto en el ciclo celular donde se bloquea la célula. Los ejemplos de células que no se dividen preexistentes en el cuerpo incluyen células neuronales, musculares, hepáticas, de la piel, cardíacas, pulmonares y de la médula ósea, y sus derivados. Para las células que se dividen se pueden utilizar vectores onco-retrovirales.

Por células "que se dividen" se entiende una célula que sufre mitosis activa o meiosis. Tales células en división incluyen células madre, células de la piel (por ejemplo, fibroblastos y queratinocitos), gametos y otras células que se dividen conocidas en la técnica. De particular interés y abarcadas por el término célula en división están las células que tienen trastornos de proliferación celular, como las células neoplásicas. La expresión "trastorno de proliferación celular" se refiere a una afección caracterizada por un número anormal de células. La afección puede incluir tanto el crecimiento celular hipertrófico (la multiplicación continua de células que resulta en un crecimiento excesivo de una población celular dentro de un tejido) como el hipotrófico (una falta o deficiencia de células dentro de un tejido) o un influjo excesivo o migración de células a un área de un cuerpo. Las poblaciones celulares no son necesariamente células tumorigénicas o malignas transformadas, pero también puede incluir células normales. Los trastornos de proliferación celular incluyen trastornos asociados con un crecimiento excesivo de tejidos conectivos, tales como diversas afecciones fibróticas, incluyendo esclerodermia, artritis y cirrosis hepática. Los trastornos de proliferación celular incluyen trastornos neoplásicos, tales como carcinomas de cabeza y cuello. Los carcinomas de cabeza y cuello incluyen, por ejemplo, carcinoma de la boca, esófago, garganta, laringe, glándula tiroides, lengua, labios, glándulas salivales, nariz, senos paranasales, nasofaringe, tumores de la bóveda nasal superior y sinusales, etesioneuroblastoma, cáncer de células escamosas, melanoma maligno, carcinoma sinonasal indiferenciado (SNUC), neoplasia de cerebro (incluyendo glioblastomas) o sanguínea. También se incluyen los carcinomas de los

ganglios linfáticos regionales, incluidos los ganglios linfáticos cervicales, ganglios linfáticos prelaríngeos, ganglios linfáticos yuxtaesofágicos pulmonares y ganglios linfáticos submandibulares (Harrison's Principles of Internal Medicine (eds., Isselbacher, et al., McGraw-Hill, Inc., 13ª edición, pp 1850-1853, 1994). Otros tipos de cáncer incluyen, pero sin limitaciones, cáncer de pulmón, cáncer de colon y recto, cáncer de mama, cáncer de próstata, 5 cáncer del tracto urinario, linfoma de cáncer uterino, cáncer oral, cáncer pancreático, leucemia, melanoma, cáncer de estómago, cáncer de piel y cáncer de ovario.

En un aspecto de la divulgación, el polinucleótido heterólogo dentro del vector comprende una citosina desaminasa que se ha optimizado para la expresión en una célula humana. En un aspecto adicional de la divulgación, la citosina 10 desaminasa comprende una secuencia que se ha optimizado por codones humanos y comprende mutaciones que aumentan la estabilidad de la citosina desaminasa (por ejemplo, degradación reducida o termoestabilidad aumentada) en comparación con una citosina desaminasa de tipo salvaje. En aún otro aspecto de la divulgación, el polinucleótido heterólogo codifica una construcción de fusión que comprende una citosina desaminasa (codón humano optimizado o no optimizado, ya sea mutado o no mutado) unida operativamente a un polinucleótido que 15 codifica un polipéptido que tiene actividad UPRT u OPRT. En otro aspecto de la divulgación, el polinucleótido heterólogo comprende un polinucleótido de CD o una construcción de fusión de la divulgación (por ejemplo, SEQ ID NO: 3, 5, 11, 13, 15 o 17).

En otro aspecto de la divulgación, el vector retroviral competente para la replicación puede comprender un polinucleótido heterólogo que codifica un polipéptido que comprende una citosina desaminasa (como se describe en 20 el presente documento) y puede comprender además un polinucleótido que comprende una molécula de ARNip o ARNip unida a un promotor específico de tipo celular o DE tejido.

La expresión "secuencia reguladora de ácido nucleico" se refiere colectivamente a secuencias promotoras, señales 25 de poliadenilación, secuencias de terminación de la transcripción, dominios reguladores aguas arriba, orígenes de replicación, potenciadores y similares, que colectivamente proporcionan la replicación, transcripción y traducción de una secuencia de codificación en una célula receptora. No es necesario que todas estas secuencias de control estén presentes siempre que la secuencia de codificación seleccionada sea capaz de replicarse, transcribirse y traducirse en una célula huésped apropiada. Un experto en la materia puede identificar fácilmente la secuencia reguladora de 30 ácido nucleico a partir de bases de datos y materiales públicos. Adicionalmente, un experto en la técnica puede identificar una secuencia reguladora que sea aplicable para el uso previsto, por ejemplo, *in vivo*, *ex vivo* o *in vitro*.

La expresión "región promotora" se usa en el presente documento en su sentido ordinario para referirse a una región 35 nucleotídica que comprende una secuencia reguladora de ADN, donde la secuencia reguladora deriva de un gen que es capaz de unirse a la ARN polimerasa e iniciar la transcripción de una secuencia codificante aguas abajo (dirección 3'). La secuencia reguladora puede ser homóloga o heteróloga a la secuencia génica deseada. Por ejemplo, se puede utilizar una amplia gama de promotores, incluyendo promotor viral o de mamífero como se ha descrito anteriormente.

Un sitio interno de entrada al ribosoma ("IRES") se refiere a un segmento de ácido nucleico que promueve la entrada 40 o retención de un ribosoma durante la traducción de una secuencia de codificación generalmente 3' al IRES. En algunos aspectos de la divulgación, el IRES puede comprender un sitio de aceptor / donante de corte y empalme, sin embargo, los IRES preferidos carecen de un sitio aceptor / donante de corte y empalme. Normalmente, la entrada de ribosomas en el ARN mensajero se realiza a través de la caperuza ubicada en el extremo 5' de todos los ARNm eucarióticos. Sin embargo, hay excepciones a esta regla universal. La ausencia de una caperuza en algunos ARNm 45 virales sugiere la existencia de estructuras alternativas que permiten la entrada de ribosomas en un sitio interno de estos ARN. Hasta la fecha, una serie de estas estructuras, designadas IRES por su función, se han identificado en la región no codificadora 5' de los ARNm virales sin protección, tal como, en particular, de picornavirus como el virus de la poliomieltis (Pelletier et al., 1988, Mol. Cell. Biol., 8, 1103-1112) y el virus EMCV (virus de la encefalo-miocarditis (Jang et al., J. Virol., 1988, 62, 2636-2643). La divulgación proporciona el uso de un IRES en el contexto de un 50 vector retroviral competente para la replicación.

La secuencia heteróloga de ácido nucleico está típicamente bajo el control de las señales potenciadoras del promotor LTR viral o de un promotor interno, y las señales retenidas dentro de la LTR retroviral todavía pueden 55 lograr una integración eficiente del vector en el genoma de la célula huésped. Por consiguiente, los vectores retrovirales recombinantes de la divulgación, las secuencias deseadas, los genes y / o fragmentos de genes pueden insertarse en varios sitios y bajo diferentes secuencias reguladoras. Por ejemplo, un sitio para la inserción puede ser el sitio proximal del potenciador / promotor viral (es decir, el locus del gen dirigido por LTR en 5'). Como alternativa, las secuencias deseadas pueden insertarse en un sitio distal (por ejemplo, la secuencia IRES 3' al gen env) o 60 cuando están presentes dos o más secuencias heterólogas, una secuencia heteróloga puede estar bajo el control de una primera región reguladora y una segunda secuencia heteróloga bajo el control de una segunda región reguladora. Otros sitios distales incluyen secuencias promotoras virales, donde la expresión de la secuencia o secuencias deseadas es a través del corte y empalme del cistron proximal del promotor, se puede usar un promotor heterólogo interno como SV40 o CMV, o un sitio interno de entrada al ribosoma (IRES).

En un aspecto de la divulgación, el genoma retroviral de la divulgación contiene un IRES que comprende un sitio de 65

clonación para la inserción de una secuencia de polinucleótidos deseada. En un aspecto de la divulgación, el IRES está ubicado 3' al gen env en el vector retroviral, pero 5' al deseado ácido nucleico heterólogo. Por consiguiente, una secuencia polinucleotídica heteróloga que codifica un polipéptido deseado puede unirse operativamente al IRES.

5 En otro aspecto de la divulgación, se incluye una secuencia de polinucleótidos dirigidos como parte del vector retroviral recombinante de la divulgación. La secuencia del polinucleótido de direccionamiento es un ligando de direccionamiento (por ejemplo, hormonas peptídicas como la heregulina, un anticuerpo de cadena única, un receptor o un ligando para un receptor), un elemento regulador específico de tejido o de tipo celular (por ejemplo, un promotor o potenciador específico de tejido o específico de tipo celular), o una combinación de un ligando dirigido y un
10 elemento regulador específico de tejido / específico de tipo celular. Preferiblemente, el ligando de direccionamiento está unido operativamente a la proteína env del retrovirus, creando una proteína env retroviral quimérica. Las proteínas virales GAG, POL y ENV pueden derivarse de cualquier retrovirus adecuado (por ejemplo, derivado de MLV o lentivirus). En otro aspecto de la divulgación, la proteína ENV viral no deriva de retrovirus (por ejemplo, CMV o VSV).

15 El retrovirus recombinante de la divulgación se modifica genéticamente de tal manera que el virus se dirige a un tipo de célula en particular (por ejemplo, células del músculo liso, células hepáticas, células renales, fibroblastos, queratinocitos, células madre mesenquimales, células de médula ósea, condrocitos, células epiteliales, células intestinales, células neoplásicas, células del glioma, células neuronales y otras conocidas en la técnica) de manera
20 que el genoma del ácido nucleico se libera a un objetivo que no se divide, una célula que se divide en un objetivo, o una célula diana que tiene un trastorno de proliferación celular. El direccionamiento se puede lograr de dos maneras. La primera forma dirige el retrovirus a una célula diana uniéndose a las células que tienen una molécula en la superficie externa de la célula. Este procedimiento para atacar al retrovirus utiliza la expresión de un ligando de direccionamiento en la capa del retrovirus para ayudar a dirigir el virus a las células o tejidos que tienen un receptor o molécula de unión que interacciona con el ligando de direccionamiento en la superficie del retrovirus. Después de
25 la infección de una célula por el virus, el virus inyecta su ácido nucleico en la célula y el material genético del retrovirus puede integrarse en el genoma de la célula huésped. El segundo procedimiento de direccionamiento utiliza elementos reguladores específicos de células o tejidos para promover la expresión y la transcripción del genoma viral en una célula específica que utiliza activamente los elementos reguladores, como se describe más
30 completamente a continuación. El material genético retrovirus transferido luego se transcribe y se traduce en proteínas dentro de la célula huésped. El elemento regulador diana está típicamente vinculado a la LTR en 5' y / o 3', creando una LTR quimérica.

35 Al insertar una secuencia de ácido nucleico heterólogo de interés en el vector viral de la divulgación, junto con otro gen que codifica, por ejemplo, el ligando para un receptor en una célula diana específica, el vector ahora es objetivo específico. Los vectores virales se pueden hacer específicos del objetivo mediante la unión, por ejemplo, de un azúcar, un glicolípido, o una proteína. El direccionamiento se puede lograr usando un anticuerpo para dirigirse al vector viral. Los expertos en la técnica sabrán, o pueden determinar fácilmente, secuencias polinucleotídicas específicas que pueden insertarse en el genoma viral o proteínas que pueden unirse a una envoltura viral para
40 permitir la administración específica del objetivo del vector viral que contiene la secuencia de ácido nucleico de interés.

Por lo tanto, la divulgación incluye una proteína env quimérica que comprende una proteína env retroviral unida operativamente a un polipéptido de direccionamiento. El polipéptido de direccionamiento puede ser una molécula
45 receptora específica de la célula, un ligando para un receptor celular específico, un anticuerpo o fragmento de anticuerpo para un epítipo antigénico específico de célula o cualquier otro ligando fácilmente identificable en la técnica que sea capaz de unirse o interactuar con una célula diana. Los ejemplos de polipéptidos o moléculas dirigidas a la enfermedad incluyen anticuerpos bivalentes que usan biotina-estreptavidina como enlazadores (Etienne-Julan et al., J. Of General Virol., 73: 3251-3255 (1992); Roux et al., Proc. Natl. Acad. Sci. USA 86:9079-9083 (1989)), virus recombinante que contiene en su envoltura una secuencia que codifica una región variable de anticuerpo de cadena única contra un hapteno (Russell et al., Nucleic Acids Research, 21,1081-1085 (1993)),
50 clonación de ligandos de hormonas peptídicas en la envoltura del retrovirus (Kasahara et al., Science, 266, 1373-1376 (1994)), construcciones de EPO / env quiméricas (Kasahara et al., 1994), anticuerpo de cadena única contra el receptor de lipoproteínas de baja densidad (LDL) en la envoltura MLV ecotrópica, dando como resultado una infección específica de células HeLa que expresan el receptor de LDL (Somia et al., Proc. Natl. Acad. Sci USA, 92, 7570-7574 (1995)), de manera similar, el rango de ALV del huésped se puede alterar mediante la incorporación de un ligando de integrina, permitiendo que el virus ahora cruce especies para infectar específicamente células de glioblastoma de rata (Valsesia-Wittmann et al., J. Virol. 68, 4609-4619 (1994)), y Dornberg y colaboradores (Chu y Dornburg, J. Virol 69, 2659-2663 (1995)) han informado sobre el objetivo específico del tejido del virus de la necrosis del bazo (SNV), un retrovirus aviar, utilizando cubiertas que contienen anticuerpos de cadena única dirigidos contra
60 marcadores tumorales.

La divulgación proporciona un procedimiento para producir un retrovirus recombinante capaz de infectar una célula diana que comprende transfectar una célula huésped adecuada con lo siguiente: un vector que comprende
65 una secuencia de polinucleótido que codifica una gag viral, una pol viral y una env viral, donde el vector contiene un sitio de clonación para la introducción de un gen heterólogo, ligado operativamente a una secuencia reguladora de ácido

nucleico, y recuperando el virus recombinante.

Los retrovirus y métodos de la divulgación proporcionan un retrovirus competente para la replicación que no requiere un virus auxiliar o una secuencia de ácido nucleico o proteínas adicionales a fin de propagar y producir el virión. Por ejemplo, las secuencias de ácido nucleico del retrovirus de la divulgación codifican, por ejemplo, un antígeno específico de grupo y una transcriptasa inversa, (y las enzimas integrasa y proteasa necesarias para la maduración y la transcripción inversa), respectivamente, como se ha descrito anteriormente. Los gag y pol víricos pueden derivarse de un lentivirus, como el VIH o un oncovirus como MoMLV. Además, el genoma de ácido nucleico del retrovirus de la divulgación incluye una secuencia que codifica una proteína de la envoltura vírica (ENV). El gen env puede derivarse de cualesquiera retrovirus. La env puede ser una proteína de envoltura anfotrópica que permite la transducción de células de humanos y otras especies, o puede ser una proteína de envoltura ecotrópica, que es capaz de transducir solo ratas de ratas y ratones. Además, puede ser deseable apuntar al virus recombinante mediante la unión de la proteína de la envoltura con un anticuerpo o un ligando particular para apuntar a un receptor de un tipo de célula particular. Como se ha mencionado anteriormente, los vectores retrovirales se pueden hacer específicos del objetivo insertando, por ejemplo, un glicolípidio, o una proteína. A menudo, el direccionamiento se realiza mediante el uso de un anticuerpo para dirigir el vector retroviral a un antígeno en un tipo de célula particular (por ejemplo, un tipo de célula que se encuentra en un determinado tejido o un tipo de célula cancerosa). Los expertos en la técnica sabrán, o pueden determinar fácilmente sin experimentación indebida, procedimientos específicos para lograr la liberación de un vector retroviral a un objetivo específico. En un aspecto de la divulgación, el gen env deriva de un no retrovirus (por ejemplo, CMV o VSV). Ejemplos de genes env derivados de retrovirus incluyen, aunque no de forma limitativa: virus de la leucemia murina de Moloney (MoMuLV), virus del sarcoma murino de Harvey (HaMuSV), virus del tumor de mama de murino (MuMTV), virus de la leucemia del mono gibbon (GaLV), virus de inmunodeficiencia humana (VIH) y virus del sarcoma de Rous (RSV). También se pueden usar otros genes env, como el virus de la estomatitis vesicular (VSV) (proteína G), o la envoltura de citomegalovirus (CMV) o el virus de la gripe con hemaglutinina (HA).

En un aspecto de la divulgación, el genoma retroviral deriva de un onco-retrovirus o gammaretrovirus, y más particularmente un onco-retrovirus de mamíferos o gamma retrovirus. Por "derivado" se entiende que la secuencia polinucleotídica principal es un oncovirus de tipo salvaje que se ha modificado mediante la inserción o eliminación de secuencias naturales (por ejemplo, la inserción de un IRES, inserción de un polinucleótido heterólogo que codifica, por ejemplo, un polipéptido que tiene actividad citosina desaminasa de la divulgación y similares).

En otro aspecto de la divulgación, la divulgación proporciona vectores retrovirales que se dirigen utilizando secuencias reguladoras. Se pueden utilizar secuencias reguladoras específicas de células o tejidos (por ejemplo, promotores) para dirigir la expresión de secuencias de genes en poblaciones de células específicas. Los promotores de mamíferos y virales adecuados para la divulgación se describen en otra parte del presente documento. Por consiguiente, en un aspecto de la divulgación, la divulgación proporciona un retrovirus que tiene elementos promotores específicos de tejido en el extremo 5' del genoma retroviral. Preferiblemente, los elementos / secuencias reguladores específicos del tejido se encuentran en la región U3 de la LTR del genoma retroviral, incluyendo, por ejemplo, promotores y potenciadores específicos de células o tejidos para células neoplásicas (por ejemplo, potenciadores y promotores específicos de células tumorales), y promotores inducibles (por ejemplo, tetraciclina).

En algunas circunstancias, puede ser conveniente regular la expresión. Por ejemplo, se pueden utilizar diferentes promotores virales con diferentes fuerzas de actividad dependiendo del nivel de expresión deseado. En células de mamíferos, el promotor temprano inmediato de CMV se usa a menudo para proporcionar una fuerte activación transcripcional. También se han utilizado versiones modificadas del promotor de CMV que son menos potentes cuando se desean niveles reducidos de expresión del transgén. Cuando se desea la expresión de un transgén en células hematopoyéticas, se pueden usar promotores retrovirales, como los LTR de MLV o MMTV. Otros promotores virales que pueden usarse incluyen SV40, RSV LTR, HIV-1 y HIV-2 LTR, promotores de adenovirus tales como de la E1A, E2A o región MLP, AAV LTR, virus del mosaico de la coliflor, HSV-TK, y el virus del sarcoma aviar.

De manera similar, se pueden usar promotores selectivos o específicos de tejidos para efectuar la transcripción en tejidos o células específicos a fin de reducir la toxicidad potencial o los efectos indeseables para los tejidos no dirigidos. Por ejemplo, los promotores como el PSA, probasina, fosfatasa ácida prostática o calicreína glandular específica de la próstata (hK2) se pueden usar para dirigir la expresión génica en la próstata. Otros promotores / dominios reguladores que pueden usarse se establecen en la Tabla 1.

En ciertas indicaciones, puede ser deseable activar la transcripción en momentos específicos después de la administración del vector de terapia génica. Esto se puede hacer con promotores como los que son hormonales o citocinas regulables. Por ejemplo, en aplicaciones terapéuticas donde la indicación es un tejido gonadal donde se producen o envían esteroides específicos, el uso de andrógenos o estrógenos regulados promotores puede ser ventajoso. Tales promotores que son regulables por hormonas incluyen MMTV, MT-1, ecdisona y RuBisco. Otros promotores regulados por hormonas, como los que responden a la tiroides, se pueden usar hormonas hipofisarias y suprarrenales. Citocinas y promotores sensibles a proteínas inflamatorias que podrían usarse incluyen K y T Kininogen (Kageyama et al., 1987), c-fos, TNF-alfa, proteína C reactiva (Arcone et al., 1988), haptoglobina (Oliviero et al., 1987), suero amiloide A2, C / EBP alfa, IL-1, IL-6 (Poli y Cortese, 1989), complemento C3 (Wilson et al., 1990),

IL-8, glicoproteína ácida alfa-1 (Prowse y Baumann, 1988), antitripsina alfa-1, lipoproteína lipasa (Zechner et al., 1988), angiotensinógeno (Ron et al., 1990), fibrinógeno, c-jun (inducible por los ésteres de forbol, TNF-alfa, radiación UV, ácido retinoico y peróxido de hidrógeno), collagenasa (inducida por los ésteres de forbol y el ácido retinoico), metalotioneína (inducible por metales pesados y glucocorticoides), estromelina (inducible por éster de forbol, interleucina-1 y EGF), alfa-2 macroglobulina y la alfa-1 antihimotripsina. Promotores específicos de tumores como la osteocalcina, elemento sensible a la hipoxia (HRE), MAGE-4, CEA, alfa-fetoproteína, GRP78 / BiP y la tirosinasa también se pueden usar para regular la expresión génica en células tumorales.

Además, esta lista de promotores no debe interpretarse como exhaustiva o limitante, los expertos en la técnica conocerán otros promotores que pueden usarse junto con los promotores y procedimientos descritos en el presente documento.

TABLA 1 PROMOTORES ESPECÍFICOS DE TEJIDOS

Tejido	Promotor
Páncreas	Insulina Elastina Amilasa pdr-1 pdx-1 glucoquinasa
Hígado	Albúmina PEPCCK Potenciador de VHB α fetoproteína apolipoproteína C α -1 antitripsina vitelogenina, NF-AB Transtiretina
Músculo esquelético	Cadena h de miosina creatina quinasa de músculo Distrofina Calpaína p94 alfa-actina troponina 1 rápida de músculo esquelético
Piel	Queratina K6 Queratina K1
Pulmón	CFTR citoqueratina humana 18 (K18) proteínas surfactantes pulmonares A, B y C CC-10 P1
Músculo liso	sm22 α SM-alfa-actina
Endotelio	Endotelina-1 E-selectina Factor de von Willebrand TIE, KDR / flk-1 Melanocitos Tirosinasa
Tejido adiposo	Lipoproteína lipasa (Zechner et al., 1988) Adipsina (Spiegelman et al., 1989) acetil-CoA carboxilasa (Pape and Kim, 1989) glicerofosfato deshidrogenasa (Dani et al., 1989) adipocito P2 (Hunt et al., 1986)
Mama	Proteína ácida del suero de leche (WAP) (Andres et al. PNAS 84:1299-1303 1987)
Sangre	β -globina

Los "elementos reguladores específicos de tejido" son elementos reguladores (por ejemplo, promotores) que son capaces de dirigir la transcripción de un gen en un tejido mientras permanecen en gran parte "silenciosos" en otros tipos de tejidos. Se entenderá, sin embargo, que los promotores específicos de tejido pueden tener una cantidad detectable de actividad "de fondo" o "base" en los tejidos donde están en silencio. El grado en que un promotor se activa selectivamente en un tejido diana puede expresarse como una relación de selectividad (actividad en un tejido diana / actividad en un tejido de control). A este respecto, un promotor específico de tejido útil en la práctica de la divulgación tiene, normalmente, una relación de selectividad mayor que aproximadamente 5. Preferiblemente, la relación de selectividad es mayor que aproximadamente 15.

Se entenderá además que ciertos promotores, si bien no tiene actividad restringida a un solo tipo de tejido, Sin embargo, pueden mostrar selectividad porque pueden estar activos en un grupo de tejidos y menos activos o silenciosos en otro grupo. Dichos promotores también se denominan "específicos de tejido" y se contemplan para su uso con la divulgación. Por consiguiente, los elementos reguladores específicos del tejido utilizados en la divulgación, tienen aplicabilidad a la regulación de las proteínas heterólogas, tales como los polipéptidos que tienen actividad citosina desaminasa de la divulgación, así como una aplicabilidad como una secuencia polinucleotídica de direccionamiento en vectores retrovirales.

Los vectores retrovirales y los polinucleótidos y polipéptidos de CD de la divulgación se pueden usar para tratar una amplia gama de enfermedades y trastornos que incluyen varias enfermedades y trastornos de proliferación celular (véase, por ejemplo, las Patentes de Estados Unidos N.º 4.405.712 y 4.650.764; Friedmann, 1989, Science, 244:1275-1281; Mulligan, 1993, Science, 260: 926-932, R. Crystal, 1995, Science 270:404-410, véase también, The Development of Human Gene Therapy, Theodore Friedmann, Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1999. ISBN 0-87969-528-5.

La divulgación también proporciona terapia génica para el tratamiento de trastornos proliferativos celulares. Dicha terapia lograría su efecto terapéutico mediante la introducción de una secuencia polinucleotídica terapéutica apropiada (por ejemplo, un polipéptido de la divulgación que tiene actividad citosina desaminasa), en las células del sujeto que tiene el trastorno proliferativo. El suministro de construcciones polinucleotídicas se puede lograr usando el vector retroviral recombinante de la divulgación.

Además, los procedimientos terapéuticos (por ejemplo, la terapia génica o los procedimientos de administración de genes) tal como se describen en el presente documento pueden realizarse in vivo o ex vivo. Por ejemplo, en los procedimientos para el tratamiento de enfermedades o trastornos de proliferación celular puede ser útil eliminar la

mayoría de un tumor antes de la terapia génica, por ejemplo quirúrgicamente o por radiación. En algunos aspectos de la divulgación, la terapia retroviral puede ser precedida o seguida por quimioterapia.

Por lo tanto, la divulgación proporciona un retrovirus recombinante capaz de infectar una célula que no está en división, una célula en división o una célula neoplásica, donde el retrovirus recombinante comprende un GAG viral; un POL viral; un ENV viral; un ácido nucleico heterólogo (por ejemplo, que comprende un polipéptido de la divulgación que tiene actividad citosina desaminasa) unido operativamente a un IRES; y las secuencias de ácido nucleico que actúan en cis necesarias para el empaquetamiento, la transcripción inversa e integración. El retrovirus recombinante puede ser un lentivirus, como el VIH, o puede ser un oncovirus. Como se ha descrito anteriormente para el procedimiento de producción de un retrovirus recombinante, el retrovirus recombinante de la divulgación puede incluir además al menos uno de VPR, VIF, NEF, VPX, TAT, REV, y la proteína VPU. Si bien no queriendo estar limitado por una teoría particular, se cree que uno o más de estos productos de genes / proteínas son importantes para aumentar el título viral del retrovirus recombinante producido (por ejemplo, NEF) o pueden ser necesarios para la infección y el empaquetamiento del virión.

La divulgación también proporciona un procedimiento de transferencia de ácido nucleico a una célula diana para proporcionar la expresión de un ácido nucleico particular (por ejemplo, una secuencia heteróloga tal como un polinucleótido que codifica un polipéptido que tiene actividad citosina desaminasa). Por lo tanto, en otro aspecto de la divulgación, la divulgación proporciona un procedimiento para la introducción y expresión de un ácido nucleico heterólogo en una célula diana que comprende infectar la célula diana con el virus recombinante de la divulgación y expresar el ácido nucleico heterólogo en la célula diana. Como se ha mencionado anteriormente, la célula diana puede ser cualquier tipo de célula, incluyendo células que se dividen, que no se dividen, neoplásicas, inmortalizadas, modificadas y de otros tipos reconocidos por los expertos en la técnica, siempre que sean capaces de infectarse por un retrovirus.

La divulgación también proporciona terapia génica para el tratamiento de trastornos de proliferación celular o inmunológicos. En un aspecto de la divulgación, un trastorno de proliferación celular se trata introduciendo un polinucleótido de CD de la divulgación, expresando el polinucleótido para producir un polipéptido que comprende actividad citosina desaminasa y poner en contacto la célula con 5-fluorocitosina en una cantidad y durante un período de tiempo para producir una cantidad citotóxica de 5-FU.

Además, la divulgación proporciona una secuencia de polinucleótidos que codifica un vector retroviral recombinante de la divulgación. La secuencia del polinucleótido puede incorporarse en diversas partículas víricas. Por ejemplo, varios vectores virales que se pueden usar para terapia génica incluyen adenovirus, virus del herpes, vaccinia o, preferentemente, un virus de ARN, tal como un retrovirus y, más particularmente, un oncovirus de mamíferos. El vector retroviral puede ser un derivado de un retrovirus murino, de simio o humano. Los ejemplos de vectores retrovirales en los que se puede insertar un gen extraño (por ejemplo, una secuencia polinucleotídica heteróloga) incluyen, aunque no de forma limitativa: derivados del virus de la leucemia murina Moloney (MoMuLV), del virus del sarcoma murino de Harvey (HaMuSV), del virus del tumor mamario murino (MuMTV) y del virus del sarcoma de Rous (RSV). Todos estos vectores pueden transferir o incorporar un gen para un marcador seleccionable para que las células transducidas puedan identificarse y generarse. En aún otro aspecto de la divulgación, la divulgación proporciona plásmidos que comprenden una construcción derivada de un retrovirus recombinante. El plásmido se puede introducir directamente en una célula diana o en un cultivo celular, tal como NIH 3T3, HT1080 (humano), CF2 (perro) u otras células de cultivo de tejidos. Las células resultantes liberan el vector retroviral en el medio de cultivo.

La divulgación proporciona una construcción de polinucleótido que comprende, desde 5' a 3': un promotor o región reguladora útil para iniciar la transcripción; una señal de empaquetamiento psi; una secuencia de ácido nucleico que codifica gag, una secuencia de ácido nucleico que codifica pol; una secuencia de ácido nucleico que codifica env; una secuencia de ácido nucleico interna de entrada al ribosoma; un polinucleótido heterólogo que codifica un marcador, terapéutico (por ejemplo, un polipéptido que tiene actividad citosina desaminasa) o polipéptido de diagnóstico; y una secuencia de ácido nucleico de LTR. En aspectos específicos de la divulgación, un polinucleótido heterólogo que codifica un polipéptido que tiene actividad citosina desaminasa puede comprender además un dominio que codifica un polipéptido que comprende actividad UPRT o OPRT.

Como se describe en otra parte en el presente documento y como sigue, los diversos segmentos de la construcción polinucleotídica de la divulgación (por ejemplo, un polinucleótido retroviral competente para la replicación recombinante) se modifican dependiendo en parte de la célula huésped deseada, el momento o la cantidad de expresión y el polinucleótido heterólogo. Una construcción retroviral competente para replicación de la divulgación se puede dividir en una serie de dominios que pueden modificarse individualmente por los expertos en la técnica.

Por ejemplo, el promotor puede comprender un promotor del CMV que tiene una secuencia como se expone en las SEQ ID NO: 19, 20, o 22 desde el nucleótido 1 a aproximadamente el nucleótido 582 y puede incluir la modificación de una o más (por ejemplo, 2-5, 5-10, 10-20, 20-30, 30-50, 50-100 o más bases de ácidos nucleicos) siempre que el promotor modificado sea capaz de dirigir e iniciar la transcripción. En un aspecto de la divulgación, el promotor o la región reguladora comprende un polinucleótido del dominio CMV-R-U5. El dominio CMV-R-U5 comprende el promotor temprano inmediato del citomegalovirus humano en una región R-U5 del VLM. En un aspecto de la

divulgación, el polinucleótido del dominio CMV-R-U5 comprende una secuencia como se expone en las SEQ ID NO: 19, 20, o 22 desde aproximadamente el nucleótido 1 a aproximadamente el nucleótido 1202 o las secuencias que son al menos un 95 % idénticas a una secuencia como se expone en las SEQ ID NO: 19, 20 o 22, donde el polinucleótido promueve la transcripción de una molécula de ácido nucleico unida operativamente a la anterior. El dominio gag del polinucleótido puede derivarse de cualquier número de retrovirus, pero típicamente se derivará de un oncoretrovirus y, más particularmente, de un oncoretrovirus de mamífero. En un aspecto de la divulgación, el dominio gag comprende una secuencia desde aproximadamente el nucleótido número 1203 a aproximadamente el nucleótido 2819 o una secuencia que tiene al menos 95 %, 98 %, 99 % o 99,8 % (redondeado a la décima más cercana) de identidad con la anterior. El dominio pol del polinucleótido puede derivarse de cualquier número de retrovirus, pero típicamente se derivará de un oncoretrovirus y, más particularmente, de un oncoretrovirus de mamífero. En un aspecto de la divulgación, el dominio pol comprende una secuencia desde aproximadamente el nucleótido número 2820 a aproximadamente el nucleótido 6358 o una secuencia que tiene al menos un 95 %, 98 %, 99 % o 99,9 % (redondeado a la décima más cercana) de identidad con la anterior. El dominio env del polinucleótido puede derivarse de cualquier número de retrovirus, pero típicamente se derivará de un oncoretrovirus y, más particularmente, de un oncoretrovirus de mamífero. En algunos aspectos de la divulgación, el dominio de codificación env comprende un dominio env anfotrópico. En un aspecto de la divulgación, el dominio env comprende una secuencia desde aproximadamente el número de nucleótidos 6359 hasta aproximadamente el nucleótido 8323 o una secuencia que tiene al menos un 95 %, 98 %, 99 % o 99,8 % (redondeado a la décima más cercana) de identidad con la anterior. El dominio IRES del polinucleótido se puede obtener a partir de cualquier número de sitios internos de entrada al ribosoma. En un aspecto de la divulgación, el IRES deriva de un virus de la encefalomiocarditis. En un aspecto de la divulgación, el dominio IRES comprende una secuencia desde aproximadamente el número de nucleótido 8327 hasta aproximadamente el nucleótido 8876 o una secuencia que tiene al menos un 95 %, 98 % o 99 % (redondeado a la décima más próxima) de identidad, siempre y cuando el dominio permita la entrada de un ribosoma. El dominio heterólogo puede comprender una citosina desaminasa de la divulgación. En un aspecto de la divulgación, el polinucleótido de CD comprende una secuencia optimizada de codones humanos. En aún otro aspecto de la divulgación, el polinucleótido de CD codifica un polipéptido mutante que tiene citosina desaminasa, donde las mutaciones confieren un aumento de la estabilización térmica que aumenta la temperatura de fusión (T_m) en 10 °C, lo que permite una actividad cinética sostenida en un intervalo de temperatura más amplio y mayores niveles acumulados de proteína. En un aspecto de la divulgación, la citosina desaminasa comprende una secuencia como se expone en la SEQ ID NO: 19 desde aproximadamente el nucleótido número 8877 hasta aproximadamente 9353. El dominio heterólogo puede ir seguido de un dominio rico en polipurina. La 3' LTR puede derivarse de cualquier número de retrovirus, normalmente, un oncoretrovirus y, preferentemente, un oncoretrovirus de mamífero. En un aspecto de la divulgación, la 3' LTR comprende un dominio U3-R-U5. En aún otro aspecto de la divulgación, la LTR comprende una secuencia como se expone en la SEQ ID NO: 19 desde aproximadamente el nucleótido 9405 hasta aproximadamente 9998 o una secuencia que es al menos un 95 %, 98 % o 99,5 % (redondeado a la más cercana 10⁰) de identidad con la anterior.

La divulgación también proporciona un vector retroviral recombinante que comprende de 5' a 3' un CMV-R-U5, fusión del promotor temprano inmediato del citomegalovirus humano a la región R-U5 del MLV; un PBS, un sitio de unión al cebador para la transcriptasa inversa; un sitio 5' de corte y empalme; una señal de empaquetamiento ψ ; un gag, un antígeno específico del grupo de ORF del VLM; un pol, un poliproteína del ORF de la polimerasa del VLM; un sitio 3' de corte y empalme; un 4070A env, un ORF de la proteína de la envoltura de la cepa 4070A del VLM; un IRES, sitio interno de entrada al ribosoma del virus de la encefalomiocarditis; una citosina desaminasa modificada (termoestable y de codones optimizados); un PPT, un tracto de polipurina; y U3-R-U5, una repetición terminal larga del VLM. Esta estructura se representa gráficamente además en la Figura 1.

La divulgación también proporciona un vector retroviral que comprende una secuencia como se expone en la SEC ID NO:19, 20 o 22.

En otro aspecto de la divulgación, la divulgación proporciona un procedimiento para tratar a un sujeto que tiene un trastorno de proliferación celular. El sujeto puede ser cualquier mamífero y es, preferentemente, un ser humano. El sujeto se pone en contacto con un vector retroviral competente para replicación recombinante de la divulgación. El contacto puede ser *in vivo* o *ex vivo*. Los procedimientos de administración del vector retroviral de la divulgación son conocidos en la técnica e incluyen, por ejemplo, administración sistémica, administración tópica, administración intraperitoneal, administración intramuscular, intracraneal, cerebroespinal, así como la administración directamente en el sitio de un tumor o trastorno de proliferación celular. Otras vías de administración conocidas en la técnica.

Por lo tanto, la divulgación incluye diversas composiciones farmacéuticas útiles para tratar un trastorno de proliferación celular. Las composiciones farmacéuticas según la divulgación se preparan llevando un vector retroviral que contiene una secuencia polinucleotídica heteróloga útil para tratar o modular un trastorno de proliferación celular según la divulgación a una forma adecuada para la administración a un sujeto usando vehículos, excipientes y aditivos o auxiliares. Los vehículos o auxiliares de uso frecuente incluyen carbonato de magnesio, dióxido de titanio, lactosa, manitol y otros azúcares, talco, proteínas de la leche, gelatina, almidón, vitaminas, celulosa y sus derivados, aceites animales y vegetales, polietilenglicoles y disolventes, tales como agua estéril, alcoholes, glicerol y alcoholes polihídricos. Los vehículos intravenosos incluyen reponedores de líquidos y nutrientes. Los conservantes incluyen antimicrobianos, antioxidantes, agentes quelantes y gases inertes. Otros vehículos farmacéuticamente aceptables

incluyen soluciones acuosas, excipientes no tóxicos, incluyendo sales, conservantes, tampones y similares, como se describe, por ejemplo, en Remington's Pharmaceutical Sciences, 15ª ed. Easton: Mack Publishing Co., 1405-1412, 1461-1487 (1975) y el Formulario Nacional XIV., 14ª ed. Washington: American Pharmaceutical Association (1975). El pH y la concentración exacta de los diversos componentes de la composición farmacéutica se ajustan de acuerdo con las habilidades de rutina en la técnica. Véase Goodman y Gilman en The Pharmacological Basis for Therapeutics (7ª ed.).

Por ejemplo, y no a modo de limitación, un vector retroviral útil en el tratamiento de un trastorno de proliferación celular incluirá una proteína ENV anfotrópica, las proteínas GAG y POL, una secuencia promotora en el genoma retroviral de la región U3, y todas las secuencias que actúan en cis necesarias para la replicación, empaquetamiento e integración del genoma retroviral en la célula diana.

Los siguientes ejemplos pretenden ilustrar, pero no limitar, la divulgación. Si bien estos ejemplos son típicos de los que podrían usarse, como alternativa, pueden utilizarse otros procedimientos conocidos por los expertos en la técnica.

Ejemplos

Ejemplo 1 Construcción de genes de CD modificados e inserción en vectores plasmídicos.

Se han hecho mejoras genéticas al gen de la citosina desaminasa de levadura de tipo salvaje para incluir: (1) tres mutaciones posicionales que cambian tres aminoácidos (A23L, I140L y V108I) para aumentar la estabilidad térmica de la proteína de la citosina desaminasa de levadura y (2) modificaciones en la secuencia génica adicional para mejorar las secuencias de uso de codones humanos para mejorar la eficiencia de la traducción de proteínas en células humanas sin más cambios en la secuencia de aminoácidos.

Diseño de secuencia para incluir CD de codones optimizados, CD-UPRT (+ / enlazador) y CD-OPRTasa (+/- enlazador). La secuencia de codificación final de la citosina desaminasa puede comprender en el extremo 5' un sitio PSI1 (de longitud completa) y en el extremo 3' un sitio Not1 más una cola poli A para la estrategia basada en casete PSI1 / Not1.

La siguiente secuencia que comprende una citosina desaminasa de levadura se usó para la clonación, optimización y mutación (los ácidos nucleicos en recuadro comprenden los sitios de restricción utilizados en los procedimientos posteriores para la clonación:

```
AACACGATTATAAATGGTGACAGGGGGAATGGCAAGCAAGTGGGATCAGAAGGGTATGGACA
TTGCCTATGAGGAGGCGGCCTTAGGTTACAAAGAGGGTGGTGTTCCTATTGGCGGATGTCTT
ATCAATAACAAAGACGGAAGTGTTCTCGGTCGTGGTCACAACATGAGATTTCAAAGGGATC
CGCCACACTACATGGTGAGATCTCCACTTTGGAAAACGTGGGAGATTAGAGGGCAAAGTGT
ACAAAGATACCACTTTGTATACGACGCTGTCTCCATGCGACATGTGTACAGGTGCCATCATC
ATGTATGGTATTCCACGCTGTGTTGTGCGTGAGAACGTTAATTTCAAAGTAAGGGCGAGAA
ATATTTACAACTAGAGGTCACGAGGTTGTTGTTGTTGACGATGAGAGGTGTAAGATCA
TGAAACAATTTATCGATGAAAGACCTCAGGATTGGTTTGAAGATATTGGTGAGTAGGCGGCC
GCCCATAGATAAAAATAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGG (SEQ ID
NO: 31)
```

La siguiente tabla resume los genes y los vectores plasmídicos resultantes que se fabricaron y sus nombres.

Tabla: Construcciones de vectores y nombres

Código Tocagen	Nombre de referencia	Nombre original	5'LTR Prom	Cubierta	Vector	IRES	Transgén	3'LTR
T5.0000	pACE-yCD	pACE-CD (Tai et al. 2005)	CMV	Ampho (4070A)	pACE	EMCV	CD de levadura Wt	U3 de MLV
T5.0001	pAC3-yCD1	Secuencia CDopt	CMV	Ampho (4070A)	pAC3	EMCV	CD modificada	U3 de MLV

Código Tocagen	Nombre de referencia	Nombre original	5'LTR Prom	Cubierta	Vector	IRES	Transgén	3'LTR
T5.0002	pAC3-yCD2	CDopt+3pt	CMV	Ampho (4070A)	pAC3	EMCV	CD modificada	U3 de MLV
T5.0003	pAC3-yCD2-U	Cdopt+3pt-UPRT	CMV	Ampho (4070A)	pAC3	EMCV	CD2-UPRT	U3 de MLV
T5.0004	pAC3-yCD2-O	CDopt+3pt-OPRT	CMV	Ampho (4070A)	pAC3	EMCV	CD2-OPRT	U3 de MLV
T5.0005	pAC3-yCD2-LO	CDopt+3pt-LINK-OPRT	CMV	Ampho (4070A)	pAC3	EMCV	CD2-L-OPRT	U3 de MLV
T5.0006	pAC3-eGFP	pAC3-emd, pAC3GFP	CMV	Ampho (4070A)	pAC3	EMCV	Emera Id GFP	U3 de MLV
T5.0007	pAC3-yCD	pAC3-yCD	CMV	Ampho (4070A)	pAC3	EMCV	CD de levadura Wt	U3 de MLV

El vector retroviral de replicación competente descrito por Kasahara et al. Se usó pACE-CD (patente de Estados Unidos N° 6.899.871) como base para modificaciones adicionales. Se modificó un vector (pAC3-yCD) para expresar un gen de la citosina desaminasa de levadura modificado como se describe en el presente documento y se usó en las construcciones. Consulte 1A a continuación para ver un diagrama de la construcción del vector para el retrovirus competente para la replicación transfectado inicial. El CMV es el promotor temprano inmediato del CMV humano, U3, R y U5 son las regiones correspondientes de la repetición terminal larga viral (LTR). *Gag*, *pol* y *env* son las regiones codificantes de proteínas virales. 1B y 1D muestran la estructura del plásmido y una secuencia de la divulgación. Después de sintetizar los genes por contrato (Bio Basic Inc., Markham, Ontario, Canadá) se insertaron en el sitio Psi1 -Not1 del esqueleto del vector pAC3 (Figura 1). El esqueleto del plásmido se generó normalmente cortando el plásmido pAC3-eGFP con Psi1 y Not1 y purificando el fragmento grande (aproximadamente 11 kb) que codifica el plásmido y el esqueleto retroviral)

A. Gen de CD optimizado con codones humanizados (CDopt, también conocido como CD1, T5.0001) Una comparación de una citosina desaminasa optimizada con codones humano de Conrad et al. y PCT WO 99/60008 indica 91 codones totales optimizados en ambos, 36 codones idénticos, 47 codones tenían cambios en el tercer par de bases (todos codifican el mismo aminoácido) y 9 codones fueron diferentes (sin embargo, codificaban el mismo aminoácido). De los 9 codones que diferían:

AGC (Ser) a TCC (Ser)

CGT (Arg) a AGG (Arg)

CCA (Pro) a CCT (Pro)

Todos tienen un contenido de GC equivalente y codifican el mismo aminoácido.

La secuencia del gen de levadura nativa anterior se optimizó por codones por separado para proporcionar el siguiente gen de CD (CD1) y se denominó T5.0001 cuando se insertó en el vector plasmídico pAC3 que codifica el retrovirus competente para la replicación (RCR) con IRES.

TTATAAATGGTGACCGCGGCATGGCCTCCAAGTGGGATCAAAGGGCATGGATATCGCTTA
CGAGGAGGCCGCCCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACA
ACAAGGACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACC
CTGCACGGCGAGATCTCCACCCTGGAGAAGTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGA
CACCACCCTGTACACCACCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTACG
GCATCCCTAGGTGTGTGGTGGGCGGAGAACGTGAACCTTCAAGTCCAAGGGCGAGAAGTACCTG
CAAACCAGGGGCCACGAGGTGGTGGTTGTTGACGATGAGAGGTGTAAGAAGATCATGAAGCA
GTTTCATCGACGAGAGGCCCTCAGGACTGGTTCGAGGATATCGGCGAGTGATAAGCGGGCCGAG
ATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGG. (SEQ ID NO:32)

B. Gen de CD estabilizado con calor. Se realizaron modificaciones adicionales para mejorar la estabilidad de la citosina desaminasa. Se realizaron mejoras genéticas en el gen de la citosina desaminasa de levadura de tipo salvaje para incluir tres mutaciones posicionales que cambian tres aminoácidos (A23L, I140L y V108I) para aumentar la estabilidad térmica de la proteína citosina desaminasa de levadura.

Los siguientes pares de cebadores se utilizaron en la generación del gen para el polipéptido de la citosina desaminasa de la divulgación: sentido: 5'-tcgaggatcgccgagtgaaaccggtattcttttggc-3' (SEQ ID NO:25) antisentido: 5'-gccaaaagaataacgggttccactgcgcgatctctga-3' (SEQ ID NO:26)

sentido: 5'tcggcgcagtgatccggcggcgccgctccggcggcgccgctccggcggcgccgctccggcggcgcccaaccggtatt-3'(SEQ ID NO:27)

antisentido: 5'-aataacgggttgccgcccgcggaggcgccgcccgcggaggcgccgcccgcggaggcgccgcccgcggatcactcgccga-3'(SEQ ID NO:28)

Para aumentar la estabilidad de la CD de levadura nativa, se realizaron tres sustituciones de aminoácidos en la proteína. Estas sustituciones se realizaron solas o en combinación con optimización de codones humanos.

Las tres sustituciones de aminoácidos son: A23L, V108I, I140L. A continuación se muestra una secuencia que codifica estas sustituciones.

ATGGTGACAGGGGGAATGGCAAGCAAGTGGGATCAGAAGGGTATGGACATTGCCTATGAGGA
GGCG**TTA**TTAGGTTACAAAGAGGGTGGTGTTCCTATTGGCGGATGTCTTATCAATAACAAAG
ACGGAAGTGTCTCGGTCGTGGTCACAACATGAGATTTCAAAGGGATCCGCCACACTACAT
GGTGAGATCTCCACTTTGGAAAAGTGTGGGAGATTAGAGGGCAAAGTGTACAAAGATAACAC
TTTGTATACGACGCTGTCTCCATGCGACATGTGTACAGGTGCCATCATCATGTATGGTATTC
CACGCTGTGT**CATC**GGTGAGAACGTTAATTTCAAAGTAAGGGCGAGAAATATTTACAAAGT
AGAGGTCACGAGGTTGTTGTTGTTGACGATGAGAGGTGTAAAAAG**TTA**ATGAAACAATTTAT
CGATGAAAGACCTCAGGATTGGTTTGAAGATATTGGTGAGTAG**GCGGCCGC**GCCATAGATAA
AATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGG (SEQ ID NO:33)

El polipéptido codificado comprende la siguiente secuencia (aminoácidos sustituidos en negrita subrayados):

1 MVTGGMASKWDQKGMDIAYEEA**L**LGYKEGGVPIGGCLINNKGSVLGRGHNMRFQKGSAT
61 LHGEISTLENCGRLEGKVYKDTTLYTTLSPCDMCTGAIIMYGIPRCV**I**GENVNFKSKGEK
121 YLQTRGHEVVVVDDERCK**L**MKQFIDERPDWFEDIGE-

El diseño final de la construcción que integra 3 sustituciones de aminoácidos A23L / V108I / I140L utilizando codones preferidos y utiliza el uso de codones humanos preferidos para la secuencia completa (este gen se llama CDopt + 3pt [también conocido como CD2] y T5.0002 cuando se inserta en el vector plasmídico pAC3, que codifica el RCR con IRES) (SEQ ID NO: 34):

1 ATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAGGGCATGGATATCGCTTACGAG
61 GAGGCC**CTG**CTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAAC
121 AAGGACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACC
181 CTGCACGGCGAGATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAG
241 GACACCACCCTGTACACCACCCTGTCCCTTGTGACATGTGTACCGGCGCTATCATCATG
301 TACGGCATCCCTAGGTGTGTGAT**CGGCG**GAGAACGTGAACTTCAAGTCCAAGGGCGAGAAG
361 TACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGTTGACGATGAGAGGTGTAAGAAGCTG
421 ATGAAGCAGTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGTGATAA

Los codones subrayados denotan codones preferidos para sustituciones de aminoácidos.

Diseño de secuencia CD optimizada (preferencia de codones humanos + 3 sustituciones de aminoácidos)

AACACGA**TTATAA**ATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAGGGCATGGATA
TCGCTTACGAGGAGGCCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTG
ATCAACAACAAGGACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTC
CGCCACCCTGCACGGCGAGATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGT
ACAAGGACACCACCCTGTACACCACCCTGTCCCTTGTGACATGTGTACCGGCGCTATCATC
ATGTACGGCATCCCTAGGTGTGTGAT**CGGCG**GAGAACGTGAACTTCAAGTCCAAGGGCGAGAA
GTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGTTGACGATGAGAGGTGTAAGAAGCTGA
TGAAGCAGTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGGATATCGGCGAGTAA**GCGGCC**
GCGCCATAGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGG (SEQ ID
NO: 35)

C. Construcción del gen de fusión CD-UPRT (CDopt + 3pt-UPRT, [también conocido como CDopt-UPRT y CD2-UPRT], T5.0003 en el vector de RCR del plásmido pAC3) También se desarrolló una construcción de fusión que comprende un polipéptido de CD como se ha descrito anteriormente unido a un polipéptido UPRT para generar una secuencia UPRT CD optimizada utilizando el Esquema I como se muestra en la Figura 2A. Los siguientes cebadores se utilizaron para delecionar el inicio-terminación entre CD y UPRT.

5

Secuencia del cebador:

Nombre del cebador	Secuencia de cebador (5' a 3') (SEC ID NO:)			
del118-123	5'-tcgaggatatcggcgagtgaaacccgttattcttttggc-3'(36)			
del118-123-antisentido	5'-gccaaaaagaataacgggttcactcgccgatcctcga-3 '(37)			
Nombre del cebador	Longitud (nt.)	Tm	Energía dúplex a 68 °C	Coste energético de los apareamientos erróneos
del118-123	40	79,06 °C	-44,37 kcal / mol	21,1 %
del118-123-antisentido	40	79,06 °C	-47,95 kcal / mol	20,3 %
Nombre del cebador	Dúplex Cebador-molde			
del118-123	5'-tcgaggatatcggcgagtg-----aaccggttattcttttggc-3' ccaagctcctatagccgctcactatctacttgggcaataagaaaaaccgaag			
del118-123-antisentido	ggttcgaggatatcggcgagtgatagatgaacccgttattcttttggcttc 3'-agctcctatagccgctcact-----ttgggcaataagaaaaaccg			

El polinucleótido de fusión resultante comprende 1296 pb y la secuencia expuesta inmediatamente a continuación:

10

AACACGA**TTATAA**ATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTT
 ACGAGGAGGCCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAAG
 GACGGCAGTGTGCTGGGCGAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCTGCACGGCGA
 GATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACCA
 CCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTACGGCATCCCTAGGTGTGTGATCGGC
 GAGAACGTGAACCTCAAGTCCAAGGGCGAGAAGTACCTGCAAACCAGGGGCCACGAGGTGGTGGTTGT
 TGACGATGAGAGGTGTAAGAAGCTGATGAAGCAGTTTCATCGACGAGAGGCCTCAGGACTGGTTTCGAGG
 ATATCGGCGAGAACCCGTTATTCTTTTTGGCTTCTCCATTCTTGTACCTTACATATCTTATATATTAT
 CCAAACAAAGGGTCTTTTCGTTAGCAAACCTAGAAATCTGCAAAAAATGTCTTCGGAACCATTTAAGAA
 CGTCTACTTGCTACCTCAAACAAACCAATTGCTGGGTTTGTACACCATCATCAGAAATAAGAATACAA
 CTAGACCTGATTTTCAATTTCTACTCCGATAGAATCATCAGATTGTTGGTTGAAGAAGGTTTGAACCAT
 CTACCTGTGCAAAAGCAAATTTGTGGAACTGACACCAACGAAAACCTCGAAGGTGTCTCATTCATGGG
 TAAAATCTGTGGTGTTCATTGTGAGAGCTGGTGAATCGATGGAGCAAGGATTAAGAGACTGTTGTA
 GGTCTGTGCGTATCGGTAAAATTTTAATTCAAAGGGACGAGGAGACTGCTTTACCAAAGTTATTCTAC
 GAAAAATTACCAGAGGATATATCTGAAAGGTATGTCTTCTATTAGACCCAATGCTGGCCACCGGTGG
 TAGTGCTATCATGGCTACAGAAGTCTTGATTAAGAGAGGTGTTAAGCCAGAGAGAATTTACTTCTTAA
 ACCTAATCTGTAGTAAGGAAGGGATTGAAAAATACCATGCCGCCTTCCAGAGGTGAGAATTGTTACT
 GGTGCCCTCGACAGAGGTCTAGATGAAAACAAGTATCTAGTTCCAGGGTTGGGTGACTTTGGTGACAG
 ATACTACTGTGTTTAA**GCGGCCGC**GCCATAGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGG
 GGGG (SEQ ID NO:38)

D. Construcción del gen de fusión de CD-enlazador UPRT (CDopt + 3pt-LINK-UPRT [también conocido como CDopt-ENLAZADOR-UPRT y CD2-L-UPRT])._También se desarrolló una construcción de fusión mediante la

clonación de un dominio enlazador (Ser-Gly-Gly-Gly-Gly)₄ dentro y en el marco con el polipéptido de CD y el polipéptido de UPRT para generar una secuencia de UPRT enlazador CDOptimizado utilizando el Esquema II como se representa en 2B. Los siguientes cebadores se utilizaron para insertar el enlazador.

Nombre del cebador	Secuencia de cebador (5' a 3') (SEC ID NO:)			
ins_60nt_after_477	5' - tcggcgagtgatccggcgggcgggcgccctccggcgggcgccctccggcg gcgggcgccctccggcgggcgggcgccaaccggttatt-3' (39)			
ins_60nt_after_477-antisentido	5' - aataacgggttggcgccgcccgcggaggcgccgcccggaggcgcc gccgcccggaggcgcccgcggccgcatcactcgccga-3' (40)			
Nombre del cebador	Longitud (nt.)	Tm	Energía dúplex a 68 °C	Coste energético de los apareamientos erróneos
ins_60nt_after_477	82	79,77 °C	-30,19 kcal / mol	83,3 %
ins_60nt_after_477-antisentido	82	79,77 °C	-32,31 kcal / mol	82,2 %

5

La construcción resultante tiene el tamaño: 1356 pb y la secuencia inmediatamente a continuación:

AACACGA **TTATAA** ATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTT
ACGAGGAGGGCCCTGCTGGGCTACAAGGAGGGCGGGCTGCCTATCGGCGGCTGTCTGATCAACAACAAG
GACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCTGCACGGCGA
GATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACCA
CCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTACGGCATCCCTAGGTGTGTGATCGGC
GAGAACGTGAACCTCAAGTCCAAGGGCGAGAAGTACCTGCAAACCAGGGGGCCACGAGGTGGTGGTTGT
TGACGATGAGAGGTGTAAGAAGCTGATGAAGCAGTTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGG
ATATCGGCGAGTCCGGCGGGCGGCCTCCGGCGGGCGGCCTCCGGCGGGCGGCCTCCGGCGGGCGGC
GCCAACCCGTTATTCTTTTGGCTTCTCCATTCTTGTACCTTACATATCTTATATATTATCCAAACAA
AGGGTCTTTTCGTTAGCAAACCTAGAAATCTGCAAAAAATGTCTTCGGAACCATTTAAGAACGTCTACT
TGCTACCTCAAACAACCAATTGCTGGGTTTGTACACCATCATCAGAAATAAGAATACAACCTAGACCT
GATTTCATTTTCTACTCCGATAGAATCATCAGATTGTTGGTTGAAGAAGGTTTGAACCATCTACCTGT
GCAAAAGCAAATTGTGGAACTGACACCAACGAAACTTCGAAGGTGTCTCATTTCATGGGTAAAAATCT
GTGGTGTTCATTGTGAGAGCTGGTGAATCGATGGAGCAAGGATTAAGAGACTGTTGTAGGTCTGTG
CGTATCGGTAAAATTTAATTCAAAGGGACGAGGAGACTGCTTTACCAAAGTTATTCTACGAAAAATT
ACCAGAGGATATATCTGAAAGGTATGTCTTCCTATTAGACCCAATGCTGGCCACCGGTGGTAGTGCTA
TCATGGCTACAGAAGTCTTGATTAAGAGAGGTGTTAAGCCAGAGAGAATTTACTTCTTAAACCTAATC
TGTAAGTAAGGAAGGGATTGAAAAATACCATGCCGCCTTCCAGAGGTCAGAATTGTTACTGGTGCCCT
CGACAGAGGTCTAGATGAAAACAAGTATCTAGTTCCAGGGTTGGGTGACTTTGGTGACAGATACTACT
GTGTTTAA **GCGGCGCG** GCCATAGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGG
(SEQ ID NO: 41)

10 E. Construcción del gen de fusión CD-OPRT (CDopt + 3pt-OPRT [también conocido como CDopt-OPRT y CD2-OPRT], T5.0004 cuando se inserta en el vector RCR del plásmido pAC3) También se desarrolló una construcción de fusión que comprende un polipéptido de CD como se ha descrito anteriormente unido a un polipéptido OPRT para generar una OPRTasa de CD optimizada (CD humanizada + mutación 3pt + OPRTasa dominio funcional humano) usando el Esquema III como se muestra en la Figura 2C.

15 La construcción resultante comprende un tamaño de 1269 pb y la secuencia inmediatamente a continuación:

AACACGA**TTATAA**ATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTT
 ACGAGGAGGCCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAAG
 GACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCTGCACGGCGA
 GATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACCA
 CCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTACGGCATCCCTAGGTGTGTGATCGGC
 GAGAACGTGAACCTCAAGTCCAAGGGCGAGAAGTACCTGCAAACCAGGGGGCCACGAGGTGGTGGTTGT
 TGACGATGAGAGGTGTAAGAAGCTGATGAAGCAGTTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGG
 ATATCGGCGAG**GCGGTCGCTCGTG**cagcttttggggccatttggtgacgggtctgtacgacgtgcaggct
 ttcaagtttggggacttcgtgctgaagagcgggctttcctcccccatctacatcgatctgcggggcat

cgtgtctcgaccgctcttctgagtcaggttgacagatatatttattccaaactgccccaaatgcaggca
 tcagttttgacaccgtgtgtggagtgccttatacagctttgccattggctacagttatctgttcaacc
 aatcaaattccaatgcttattagaaggaagaaacaaaggattatggaactaagcgtctttagtagaagg
 aactattaatccaggagaaacctgtttaatcattgaagatggtgtcaccagtggtatctagtgttttgg
 aaactgttgaggttcttcagaaggagggcttgaaggtcactgatgccatagtgtgttggacagagag
 caggagggcaaggacaagttgcaggcgcacgggatccgcctccactcagtggtgtacattgtccaaaat
 gctggagattctcgagcagcagaaaaaagttgatgctgagacagttgggagagtgaagaggtttattc
 aggagaatgtcttctgtggcagcgaatcataatggttctccccctttctataaaggaagcaccacaaagaa
 ctcaGCTTCGGTGCACGTGCAGAGCTGCCCAGGATCCACCCAGTTGCATCGAAGTAA**GCGGCCGC**GCC
 ATAGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGGG (SEQ ID NO:42)

- 5 **F. Construcción del gen de fusión CD-enlazador-OPRT (CDopt + 3pt-LINK-OPRT, [también conocido como CDopt-ENLAZADOR-OPRT y CD2-L-OPRT]. T5.0005 en el vector RCR plasmídico pAC3)** También se desarrolló una construcción de fusión mediante la clonación de un dominio enlazador (Ser-Gly-Gly-Gly-Gly)4) entre y en el marco con el polipeptídico de CD y el polipéptido OPRT para generar una secuencia de CDoptimizada -enlazador-OPRT utilizando el Esquema IV como se muestra en la Figura 2D.
 La construcción resultante comprende un tamaño de 1329 pb y la secuencia inmediatamente a continuación:

AACACGA**TTATAA**ATGGTGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTT
 ACGAGGAGGCCCTGCTGGGCTACAAGGAGGGCGGCGTGCCTATCGGCGGCTGTCTGATCAACAACAAG
 GACGGCAGTGTGCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCTGCACGGCGA
 GATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACACCACCCTGTACACCA
 CCCTGTCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTACGGCATCCCTAGGTGTGTGATCGGC
 GAGAACGTGAACCTCAAGTCCAAGGGCGAGAAGTACCTGCAAACCAGGGGGCCACGAGGTGGTGGTTGT
 TGACGATGAGAGGTGTAAGAAGCTGATGAAGCAGTTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGG
 ATATCGGCGAGT**CCGGCGGCGGCGCCT**CCGGCGGCGGCGCTCCGGCGGCGGCGCCTCCGGCGGCGGCGC
 GCCGCGGTTCGTCGTCagcttttggggccatttggtgacgggtctgtacgacgtgcaggctttcaagtt
 tggggacttcgtgctgaagagcgggctttcctcccccatctacatcgatctgcggggcagctgctgtc
 gaccgctcttctgagtcaggttgacagatatatttattccaaactgccccaaatgcaggcatcagtttt
 gacaccgtgtgtggagtgccttatacagctttgccattggctacagttatctgttcaaccaatcaa
 tccaatgcttattagaaggaagaaacaaaggattatggaactaagcgtctttagtagaaggaactatta
 atccaggagaaacctgtttaatcattgaagatggtgtcaccagtggtatctagtgttttggaaactgtt
 gaggttcttcagaaggagggcttgaaggtcactgatgccatagtgtgttggacagagagcaggagg
 caaggacaagttgcaggcgcacgggatccgcctccactcagtggtgtacattgtccaaaatgctggaga
 ttctcgagcagcagaaaaaagttgatgctgagacagttgggagagtgaagaggtttattcaggagaat
 gtcttctgtggcagcgaatcataatggttctccccctttctataaaggaagcaccacaaagaactcaGCTT
 CGGTGCACGTGCAGAGCTGCCCAGGATCCACCCAGTTGCATCGAAGTAA**GCGGCCGC**GCCATAGATAA
 AATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGGG. (SEQ ID NO:43)

Ejemplo 2 Producción de vectores infecciosos

- 15 El vector se puede producir de varias maneras, pero la primera etapa es introducir el vector de ADN plasmídico en las células para permitir la producción de partículas infecciosas, que luego se pueden recoger del sobrenadante celular. Una vez que se han generado partículas infecciosas, los expertos en la técnica pueden implementar otros procedimientos de producción. Las partículas del vector se generaron por transfección transitoria de células 293T Pear et al. Proc Natl Acad Sci U S A. 90:8392-8396 1993). Las células 293T se descongelaron y se pusieron en cultivo, luego se pasaron dos veces en matraces T-75 que contenían 15 ml del medio DMEM que se preparó

mezclando medio DMEM rico en glucosa (Hyclone # 30081, 500 ml) con FBS (Hyclone # SH30070, 50 ml), L-Glutamina (Cellgro # 25-005-Cl, 5 ml), NEAA (Hyclone # SH30238, 5 ml) y Penicilina-estrep (Cellgro # 30-002-Cl, 5 ml). Los matraces se incubaron a 37 °C y 5 % de CO₂. Después del tercer paso, las células se sembraron en 6 T-25, conteniendo cada 5 ml del medio, a una densidad celular de $1,8 \times 10^6$ células / T-25 (o $7,2 \times 10^4$ células / cm²). Un día después de sembrar los T-25, las células se transfectaron con el plásmido T5.0002 que expresaba el vector viral utilizando el kit de transfección de fosfato de calcio de Promega (N.º de cat. E1200). Dieciocho horas después de la transfección, los medios en un conjunto de matraces (3 matraces cada conjunto) se reemplazaron con medio fresco que contenía NaB 10 mM. Los medios en el segundo conjunto de los matraces no fueron reemplazados, sirvieron como control (NaB cero). Ocho horas después del tratamiento con NaB, los medios en todos los matraces se reemplazaron con el medio fresco que no contenía NaB. Se permitió que la expresión continuara para ambos conjuntos de matraces hasta el día siguiente (22 horas de duración). Los sobrenadantes de ambos conjuntos de matraces se recolectaron y analizaron para determinar sus títulos mediante qPCR expresados en Unidades de Transducción (UT) / ml (véase el ejemplo 3).

Los resultados del título se muestran en la siguiente tabla.

Condición	Primer título	Segundo título (después de almacenar a -80 °C durante 68 días)
Sin NaB	$1,5 (\pm 0,05) \times 10^6$ UT / ml	$1,2 (\pm 0,2) \times 10^6$ UT / ml
NaB 10 mM	$1,4 (\pm 0,3) \times 10^6$ UT / ml	$7,0 (\pm 0,14) \times 10^5$ UT / ml

Las preparaciones de vector posteriores se produjeron de esta manera, sin butirato de sodio. Los otros plásmidos vectores se han utilizado de la misma manera para generar preparaciones de vectores con títulos entre 10^5 UT / ml y 10^7 UT / ml. Tal material puede purificarse adicionalmente y concentrarse, si se desea, como se describe a continuación y véase también: documento US 5792643; T. Rodríguez et al. J Gene Med 9:233 2007; solicitud de Estados Unidos 61218063. En ciertas realizaciones de la divulgación, la dosificación se calculó por gramos de peso de cerebro. En dichas realizaciones, la dosificación de un vector retroviral competente para la replicación de la divulgación útil en los procedimientos para el tratamiento puede variar de 10^3 a 10^7 UT por gramo de peso de cerebro.

Ejemplo 3 Ensayo cuantitativo de titulación por PCR

La concentración del vector funcional, o título, se determinó utilizando un método basado en la PCR cuantitativa (qPCR). En este método, el vector se titula infectando una línea de células huésped transducible (por ejemplo, células PC-3 de carcinoma prostático humano, ATCC n.º de cat. CRL-1435) con un volumen normalizado de vector y midiendo la cantidad resultante de provirus presente en las células huésped tras la transducción. Las células y el vector se incubaron en condiciones de cultivo normalizadas (37 °C, 5 % de CO₂) durante 24 horas para permitir una infección completa antes de la adición del AZT antirretroviral para detener la replicación del vector.

A continuación, las células se recogieron de la placa de cultivo y el ADN genómico (ADNg) se purificó utilizando un kit de purificación del ADNg Purelink de Invitrogen y se eluyeron de la columna de purificación con agua estéril exenta de ARNasa/ADNasa. La relación de absorbancia A260 / A280 se mide en un espectrofotómetro para determinar la concentración y la pureza relativa de la muestra. Las concentraciones de ADNg se normalizaron con agua exenta de ARNasa/ADNasa adicional hasta la concentración más baja de cualquier conjunto dado de preparaciones de ADNg de tal manera que la entrada de ADN para la qPCR es constante para todas las muestras analizadas. Se evaluó además la pureza del ADN genómico mediante la electroforesis de una alícuota de cada muestra en un gel de agarosa al 0,8 % teñido con bromuro de etidio. Si la muestra pasa un intervalo de absorbancia A260 / A280 de 1,8-2,0 y muestra una sola banda de ADNg, a continuación, la muestra está lista para el análisis de la qPCR del número de copias províricas del vector. Utilizando cebadores que interrogan la región LTR del provirus (vector de ADN transcrito de forma inversa y vector de ADN que se integra en el ADNg hospedador), se llevó a cabo la qPCR para estimar el número total de eventos de transducción producidos cuando el volumen conocido del vector se utilizó para transducir el número de células conocido. El número de eventos de transducción por reacción se calcula a partir de una curva estándar que utiliza un plásmido portador de diana del número de copia conocido que se diluye en serie de 10^7 a 10^1 copias y se mide en condiciones de qPCR idénticas a las muestras. A partir del conocimiento de cuántos equivalentes genómicos se utilizaron para cada reacción de la qPCR (a partir de la concentración previamente determinada) y cuántos eventos de transducción se produjeron por reacción, los inventores determinaron el número total de eventos de transducción que se produjeron basándose en el número total de células que estuvieron presentes en el momento de la transducción. Este valor es el título del vector tras la dilución en el medio que contiene las células durante la transducción inicial. Para calcular el valor del título corregido, la dilución se corrige multiplicando por el rendimiento del volumen del cultivo y el volumen del título dividido por el volumen del título. Estos experimentos se llevaron a cabo en cultivos replicados y se analizaron mediante la qPCR utilizando mediciones por triplicado para cada condición para determinar un título promedio y con su desviación estándar y el coeficiente de la varianza asociada.

Ejemplo 4 Niveles de expresión medidos por transferencia de tipo Western

La Figura 3 demuestra que se observan niveles más altos del codón humano optimizado con las tres mutaciones para una mayor estabilidad en comparación con la proteína yCD de tipo salvaje en un análisis de transferencia Western de células U-87 infectadas con virus que codifican los genes de citosina desaminasa de tipo salvaje (ACE-yCD) o completamente optimizados (AC3-yCD2).

Ejemplo 5 Estabilidad genética de vectores virales.

Se reconoce que después de la transcripción inversa y el primer evento de integración en las células tratadas, el provirus de ADN y cualquier retrovirus de progenie posterior tiene una estructura de LTR convencional de MLV en cada extremo. Se ha demostrado que esta configuración es estable después de múltiples ciclos de infección (consulte la Figura 4 a continuación).

Aproximadamente 10^6 células U-87 intactas se infectaron inicialmente con el vector viral a una MOI de 0,01 y se cultivaron hasta que se infectaron completamente para completar un solo ciclo de infección. El sobrenadante se recalienta en células no infectadas y se repite el ciclo. La estabilidad genómica de la secuencia de yCD2 se evaluó mediante amplificación por PCR del provirus integrado de las células infectadas utilizando cebadores específicos de MLV que flanquean el sitio de inserción del transgén. Para cada conjunto de infecciones, También se realizó amplificación del plásmido vector (pAC3-yCD2 y el vector pACE-CD de Kasahara et al.) para rastrear el tamaño del amplicón de longitud completa en el gel. La aparición de cualquier banda más pequeña que un amplicón de longitud completa sería un indicador de la inestabilidad del vector. Dichos experimentos demostraron que un vector de la divulgación (T5.0002, que comprende el vector modificado y el polinucleótido heterólogo CDopt + 3pt (CD2) mantuvo la estabilidad de más pasajes que pACE-CD o T5.007, los cuales llevan la levadura de tipo salvaje.

Ejemplo 6 Experimentos de destrucción celular

En experimentos de cultivo celular *in vitro*, los experimentos demuestran que la citosina desaminasa en las células que expresan la proteína yCD2 es al menos tan activa como la de las células que expresan la proteína yCD de tipo salvaje, realizando valoraciones de 5-FC en células de rata RG2 (Figura 5A) o células U-87 (Figura 5B) infectadas, ya sea con virus creado a partir de pAC3-yCD2 / T5.0002 [AC3-yCD2 (V)] a partir de pAC3-yCD / T5.0007 [AC3-yCD (V)] o los otros vectores. En resumen, para células U-87, 5 días después de la infección a una multiplicidad de infección de 0,1 (es decir, 100 % infectada) con el vector AC3-yCD (CD de tipo salvaje) o el vector AC3-yCD2 (termoestabilizado y codón optimizado) se sometieron a cantidades crecientes de 5-FC o 0,1 mM de 5-FU como control positivo durante 8 días. El día 8 del tratamiento 5-FC, se evaluó la viabilidad de los cultivos celulares utilizando un ensayo MTS (Promega CellTiter 96 AQUEOUS One Solution Proliferation Assay). Los datos muestran una destrucción comparable entre los dos vectores retrovirales a dosis crecientes de tratamiento con 5-FC. Los cultivos de RG2 se trataron de manera similar y también muestran, en todo caso, un ligero cambio en la curva de destrucción hacia la concentración más baja de 5-FC para el virus del ensayo de expresión T5.0002CD. Las células U87 se transdujeron a una multiplicidad de infección (MOI) de 0,1, se cultivaron durante 5 días para permitir la propagación viral y se recolectaron las células a partir del día 5 posterior a la transducción. Las células se recogieron luego por centrifugación a 800 x g durante 5 minutos. El sobrenadante se aspiró del sedimento celular y se lavó con 5 ml de solución salina tamponada con fosfato (PBS) y se centrifugó nuevamente a 800 x g durante 5 minutos. El sedimento celular resultante se recogió en 1,5 ml de PBS, se resuspendió mediante el paso a través de una punta de pipeta y se coloca en un congelador a -20 °C. Las células se lisaron mediante un procedimiento de congelación / descongelación. Las células previamente resuspendidas se dejaron descongelar a temperatura ambiente, se pasaron a través de una punta de pipeta, se mezclaron con un cóctel inhibidor de proteasa y nuevamente se congelaron a -20 °C. Previo al ensayo enzimático, la muestra se descongeló de nuevo a temperatura ambiente y se pasó a través de una punta de pipeta. A continuación, la suspensión se centrifugó a 14.000 rpm en una centrífuga de mesa durante 5 minutos. El sobrenadante se decantó del sedimento y se colocó en un tubo eppendorf nuevo y se colocó en hielo. La actividad de la enzima yCD se evaluó utilizando un ensayo de HPLC.

El ensayo de HPLC se realizó en una unidad Shimadzu LC20AT conectada en serie con un detector de fotomatrix y un autoinyector. La fase sólida fue una columna de HPLC Hypersil BDS C18 con un tamaño de esfera de 5 μ m y dimensiones de columna de 4,0 x 250 mm. La fase móvil fue fosfato de amonio 50 mM, pH 2,1, que contiene 0,01 % de perclorato de terc-butilamonio y 5 % de metanol; el sistema se equilibró a 22 °C. Todos los reactivos fueron de calidad ACS y los disolventes fueron de calidad HPLC. Se hizo una mezcla de reacción que consistía en 800 μ l con una concentración final de 0,125 mg / ml de 5FC (1 mM) en PBS y se colocó en un vial de automuestreador de 1,5 ml. A continuación, la reacción se inició añadiendo 200 μ l de cada lisado celular. Los viales de reacción / automuestreador se colocaron en el automuestreador y se inyectaron 5 μ l de la mezcla de reacción. Los puntos de tiempo se tomaron periódicamente recuperando una alícuota de 5 μ l de cada vial de reacción y analizando en la columna de HPLC. Las tasas de conversión de 5FC en 5FU se calcularon comparando las áreas de los picos con cantidades conocidas de una curva estándar generada previamente de 5FU. La tasa de conversión de 5FC en 5FU se derivó trazando la cantidad de 5FU (en nmol) generada contra su intervalo de tiempo correspondiente. Se derivó la concentración de proteína para la muestra de células y la actividad específica de las muestras de lisado celular se calculó dividiendo la tasa de conversión (nmol / min) por la cantidad de proteína utilizada en el ensayo en mg. La

Figura 6 muestra la actividad específica de varios vectores después de 5 días en la transducción a una MOI de 0,1. Los datos demuestran que pACE-yCD (T5.0000) <pAC3-yCD1 (T5.0001) <pAC3-CD2 (T5.0002) en términos de la actividad específica de la citosina desaminasa en células de cultivo tisular.

5 Ejemplo 7. Los tumores tratados con el gen de CD completamente modificado (yCD2) se eliminan más eficientemente que los tumores tratados con el gen de CD de levadura no modificado y no recurren (PR-01-08-001).

10 Para determinar qué construcción vectorial da lugar al vector más efectivo en un modelo de xenoinjerto de ratón / humano subcutáneo. Se evaluaron tres construcciones diferentes: T5.0001 (CD parcialmente modificada); T5.0007 (gen CD de levadura no modificado); T5.0002 (gen yCD2 totalmente modificado). El crecimiento tumoral, la supervivencia y la regresión del tumor después de la administración de profármacos de 5-FC se evaluaron en un modelo subcutáneo de glioma humano (U87) en ratones inmunodeficientes. Se evaluaron dos concentraciones diferentes de 5-FC para determinar una respuesta a la dosis observada entre 5-FC y las construcciones de vector tumoral.

Se estudió un total de 12 grupos que consistían en 9-11 ratones hembra por grupo. Todos los ratones se sometieron a la implantación en el flanco dorsal derecho el día 0, ya sea con: una mezcla del 98 % de una línea celular de tumor U-87 no infectada y una mezcla del 2 % de una línea celular U-87 que se ha infectado con una de tres construcciones de vectores TOCA 511, T5.0002, T5.0001, T5.0007; o una línea celular U-87 de control no infectada (100 %). Se inoculó a los ratones 2×10^6 células / ratón. Los ratones en los grupos 1-3 tenían células U87 no infectadas, los grupos 4-6 tenían una mezcla de células U87 que contenía T5.0002 U87 transducida, los grupos 7-9 tenían una mezcla de células U87 que contenía transducida con T5.0001, y los grupos 10-12 tenían una mezcla de células U87 que contenía transducida con T5.0007. Los tumores se dejaron crecer durante 6 días, hasta que el tamaño del tumor fue de aproximadamente 100 mm^3 . Cada grupo de ratones con dosis de vectores se aleatorizó para recibir una de las dos dosis de 5-FC (200 o 500 mg / kg / día), administradas como una sola inyección IP, comenzando el día 6, o no 5-FC. La administración de 5-FC continuó diariamente durante 28 días consecutivos. Los ratones que sobrevivieron después del día 29 se evaluaron para determinar el tamaño del tumor y luego se sacrificaron en días variables y los tumores, si estaban presentes, se extrajeron para su análisis. El día 29, los ratones del grupo 4-12 se volvieron a aleatorizar y los ratones se subdividieron en 2 subgrupos (los que continuaron con 5-FC para controlar la regresión del tumor o los que suspendieron el tratamiento con 5-FC para controlar el recrecimiento del tumor).

Resultados: Los tumores tratados con T5.0002, T5.0000 y T5.0007 demostraron regresión del tumor a niveles indetectables a dosis de 200 y 500 mg / kg en el día 29. Todos los animales tratados con T5.0002 no reformaron los tumores cuando se suspendió el tratamiento con 5FC, en 39 días. En contraste, algunos tumores tratados con T5.0001 y T5.0007 recurrieron en el punto de tiempo de 39 días. El experimento finalizó el día 39.

Conclusión: T5.0002 más 5-FC es una terapia antitumoral más efectiva que T5.0001 o T5.0007 más 5-FC.

LISTADO DE SECUENCIAS

40 <110> Tocagen Inc.
Gruber, Harry E.
Jolly, Douglas
Perez, Omar

45 <120> VECTORES DE TERAPIA GÉNICA Y CITOSINA DESAMINASAS

<130> 00014-007WO1

50 <140> Pendiente de conocer
<141> 26/09/2009

<150> US 61/100.666
<151> 26/09/2008

55 <150> US 61/120.618
<151> 08/12/2008

60 <150> US 61/186.823
<151> 13/06/2009

<160> 22

<170> PatentIn versión 3.5

65 <210> 1

ES 2 706 899 T3

<211> 477
 <212> ADN
 <213> *Saccharomyces cerevisiae*

5 <220>
 <221> CDS
 <222> (1)..(477)

<400> 1

10
 atg gtg aca ggg gga atg gca agc aag tgg gat cag aag ggt atg gac 48
 Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15
 att gcc tat gag gag gcg gcc tta ggt tac aaa gag ggt ggt gtt cct 96
 Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30
 att ggc gga tgt ctt atc aat aac aaa gac gga agt gtt ctc ggt cgt 144
 Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45
 ggt cac aac atg aga ttt caa aag gga tcc gcc aca cta cat ggt gag 192
 Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60
 atc tcc act ttg gaa aac tgt ggg aga tta gag ggc aaa gtg tac aaa 240
 Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80
 gat acc act ttg tat acg acg ctg tct cca tgc gac atg tgt aca ggt 288
 Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95
 gcc atc atc atg tat ggt att cca cgc tgt gtt gtc ggt gag aac gtt 336
 Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
 100 105 110
 aat ttc aaa agt aag ggc gag aaa tat tta caa act aga ggt cac gag 384
 Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125
 gtt gtt gtt gtt gac gat gag agg tgt aaa aag atc atg aaa caa ttt 432
 Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
 130 135 140
 atc gat gaa aga cct cag gat tgg ttt gaa gat att ggt gag tag 477
 Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
 145 150 155

<210> 2
 <211> 158
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 2

15

ES 2 706 899 T3

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

<210> 3
 <211> 477
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Citosina desaminasa modificada

<220>
 <221> CDS
 <222> (1)..(477)

<400> 3

ES 2 706 899 T3

atg gtg aca ggg gga atg gca agc aag tgg gat cag aag ggt atg gac	48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp	
1 5 10 15	
att gcc tat gag gag gcg tta tta ggt tac aaa gag ggt ggt gtt cct	96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro	
20 25 30	
att ggc gga tgt ctt atc aat aac aaa gac gga agt gtt ctc ggt cgt	144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg	
35 40 45	
ggg cac aac atg aga ttt caa aag gga tcc gcc aca cta cat ggt gag	192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu	
50 55 60	
atc tcc act ttg gaa aac tgt ggg aga tta gag ggc aaa gtg tac aaa	240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys	
65 70 75 80	
gat acc act ttg tat acg acg ctg tct cca tgc gac atg tgt aca ggt	288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly	
85 90 95	
gcc atc atc atg tat ggt att cca cgc tgt gtc atc ggt gag aac gtt	336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val	
100 105 110	
aat ttc aaa agt aag ggc gag aaa tat tta caa act aga ggt cac gag	384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu	
115 120 125	
gtt gtt gtt gtt gac gat gag agg tgt aaa aag tta atg aaa caa ttt	432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe	
130 135 140	
atc gat gaa aga cct cag gat tgg ttt gaa gat att ggt gag tag	477
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu	
145 150 155	

<210> 4
 <211> 158
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Construcción sintética

<400> 4

ES 2 706 899 T3

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 5

<211> 480

<212> ADN

<213> Secuencia artificial

<220>

<223> Citosina desaminada optimizada por codones humanos

<220>

<221> CDS

<222> (1)..(480)

<400> 5

ES 2 706 899 T3

atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat	48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp	
1 5 10 15	
atc gct tac gag gag gcc gca ctg ggc tac aag gag ggc ggc gtg cct	96
Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro	
20 25 30	
atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg	144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg	
35 40 45	
ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag	192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu	
50 55 60	
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag	240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys	
65 70 75 80	
gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc	288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly	
85 90 95	
gct atc atc atg tac ggc atc cct agg tgt gtg gtc ggc gag aac gtg	336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val	
100 105 110	
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag	384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu	
115 120 125	
gtg gtg gtt gtt gac gat gag agg tgt aag aag atc atg aag cag ttc	432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe	
130 135 140	
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tga taa	480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu	
145 150 155	

<210> 6
 <211> 158
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Construcción sintética

<400> 6

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp	
1 5 10 15	
Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro	
20 25 30	
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg	

ES 2 706 899 T3

35

40

45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 7
<211> 756
<212> ADN
<213> *Saccharomyces cerevisiae*

<220>
<221> CDS
<222> (1)..(756)

<400> 7

atg aac ccg tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca tat 48
Met Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr
1 5 10 15

ctt ata tat tat cca aac aaa ggg tct ttc gtt agc aaa cct aga aat 96
Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn
20 25 30

ctg caa aaa atg tct tcg gaa cca ttt aag aac gtc tac ttg cta cct 144
Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro
35 40 45

caa aca aac caa ttg ctg ggt ttg tac acc atc atc aga aat aag aat 192
Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn
50 55 60

aca act aga cct gat ttc att ttc tac tcc gat aga atc atc aga ttg 240
Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu
65 70 75 80

15

ES 2 706 899 T3

```

ttg gtt gaa gaa ggt ttg aac cat cta cct gtg caa aag caa att gtg      288
Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val
      85                      90                      95

gaa act gac acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt aaa      336
Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys
      100                    105                    110

atc tgt ggt gtt tcc att gtc aga gct ggt gaa tcg atg gag caa gga      384
Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly
      115                    120                    125

tta aga gac tgt tgt agg tct gtg cgt atc ggt aaa att tta att caa      432
Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln
      130                    135                    140

agg gac gag gag act gct tta cca aag tta ttc tac gaa aaa tta cca      480
Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro
      145                    150                    155                    160

gag gat ata tct gaa agg tat gtc ttc cta tta gac cca atg ctg gcc      528
Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala
      165                    170                    175

acc ggt ggt agt gct atc atg gct aca gaa gtc ttg att aag aga ggt      576
Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly
      180                    185                    190

gtt aag cca gag aga att tac ttc tta aac cta atc tgt agt aag gaa      624
Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu
      195                    200                    205

ggg att gaa aaa tac cat gcc gcc ttc cca gag gtc aga att gtt act      672
Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr
      210                    215                    220

ggt gcc ctc gac aga ggt cta gat gaa aac aag tat cta gtt cca ggg      720
Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly
      225                    230                    235                    240

ttg ggt gac ttt ggt gac aga tac tac tgt gtt taa      756
Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val
      245                    250

```

<210> 8
 <211> 251
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 8

```

Met Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr
1          5          10          15

Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn
20          25          30

```

5

10

ES 2 706 899 T3

Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro
 35 40 45
 Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn
 50 55 60
 Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu
 65 70 75 80
 Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val
 85 90 95
 Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys
 100 105 110
 Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly
 115 120 125
 Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln
 130 135 140
 Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro
 145 150 155 160
 Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala
 165 170 175
 Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly
 180 185 190
 Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu
 195 200 205
 Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr
 210 215 220
 Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly
 225 230 235 240
 Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val
 245 250

<210> 9
 <211> 1443
 <212> ADN
 <213> homo sapiens

<220>
 <221> CDS
 <222> (1)..(1443)

ES 2 706 899 T3

<400> 9

atg gct gtt gct cgt gct gct ctt ggt cct ctt gtt act ggt ctt tat	48
Met Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr	
1 5 10 15	
gat gtt caa gct ttt aaa ttt ggt gat ttt gtt ctt aaa tct ggt ctt	96
Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu	
20 25 30	
tct tct cct att tat att gat ctt cgt ggt att gtt tct cgt cct cgt	144
Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg	
35 40 45	
ctt ctt tct caa gtt gct gat att ctt ttt caa act gct caa aat gct	192
Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala	
50 55 60	
ggt att tct ttt gat act gtt tgt ggt gtt cct tat act gct ctt cct	240
Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro	
65 70 75 80	
cttgct act gtt att tgt tct act aat caa att cct atg ctt att cgt	288
Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg	
85 90 95	
cgt aaa gaa act aaa gat tat ggt act aaa cgt ctt gtt gaa ggt act	336
Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr	
100 105 110	
att aat cct ggt gaa act tgt ctt att att gaa gat gtt gtt act tct	384
Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser	
115 120 125	
ggt tct tct gtt ctt gaa act gtt gaa gtt ctt caa aaa gaa ggt ctt	432
Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu	
130 135 140	
aaa gtt act gat gct att gtt ctt ctt gat cgt gaa caa ggt ggt aaa	480
Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys	
145 150 155 160	
gat aaa ctt caa gct cat ggt att cgt ctt cat tct gtt tgt act ctt	528
Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu	
165 170 175	
tct aaa atg ctt gaa att ctt gaa caa caa aaa aaa gtt gat gct gaa	576
Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu	
180 185 190	
act gtt ggt cgt gtt aaa cgt ttt att caa gaa aat gtt ttt gtt gct	624
Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala	
195 200 205	
gct aat cat aat ggt tct cct ctt tct att aaa gaa gct cct aaa gaa	672
Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu	
210 215 220	

ES 2 706 899 T3

ctt tct ttt ggt gct cgt gct gaa ctt cct cgt att cat cct gtt gct Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala 225 230 235 240	720
tct aaa ctt ctt cgt ctt atg caa aaa aaa gaa act aat ctt tgt ctt Ser Lys Leu Leu Arg Leu Met Gln Lys Lys Glu Thr Asn Leu Cys Leu 245 250 255	768
tct gct gat gtt tct ctt gct cgt gaa ctt ctt caa ctt gct gat gct Ser Ala Asp Val Ser Leu Ala Arg Glu Leu Leu Gln Leu Ala Asp Ala 260 265 270	816
ctt ggt cct tct att tgt atg ctt aaa act cat gtt gat att ctt aat Leu Gly Pro Ser Ile Cys Met Leu Lys Thr His Val Asp Ile Leu Asn 275 280 285	864
gat ttt act ctt gat gtt atg aaa gaa ctt att act ctt gct aaa tgt Asp Phe Thr Leu Asp Val Met Lys Glu Leu Ile Thr Leu Ala Lys Cys 290 295 300	912
cat gaa ttt ctt att ttt gaa gat cgt aaa ttt gct gat att ggt aat His Glu Phe Leu Ile Phe Glu Asp Arg Lys Phe Ala Asp Ile Gly Asn 305 310 315 320	960
act gtt aaa aaa caa tat gaa ggt ggt att ttt aaa att gct tct tgg Thr Val Lys Lys Gln Tyr Glu Gly Gly Ile Phe Lys Ile Ala Ser Trp 325 330 335	1008
gct gat ctt gtt aat gct cat gtt gtt cct ggt tct ggt gtt gtt aaa Ala Asp Leu Val Asn Ala His Val Val Pro Gly Ser Gly Val Val Lys 340 345 350	1056
ggt ctt caa gaa gtt ggt ctt cct ctt cat cgt ggt tgt ctt ctt att Gly Leu Gln Glu Val Gly Leu Pro Leu His Arg Gly Cys Leu Leu Ile 355 360 365	1104
gct gaa atg tct tct act ggt tct ctt gct act ggt gat tat act cgt Ala Glu Met Ser Ser Thr Gly Ser Leu Ala Thr Gly Asp Tyr Thr Arg 370 375 380	1152
gct gct gtt cgt atg gct gaa gaa cat tct gaa ttt gtt gtt ggt ttt Ala Ala Val Arg Met Ala Glu Glu His Ser Glu Phe Val Val Gly Phe 385 390 395 400	1200
att tct ggt tct cgt gtt tct atg aaa cct gaa ttt ctt cat ctt act Ile Ser Gly Ser Arg Val Ser Met Lys Pro Glu Phe Leu His Leu Thr 405 410 415	1248
cct ggt gtt caa ctt gaa gct ggt ggt gat aat ctt ggt caa caa tat Pro Gly Val Gln Leu Glu Ala Gly Gly Asp Asn Leu Gly Gln Gln Tyr 420 425 430	1296
aat tct cct caa gaa gtt att ggt aaa cgt ggt tct gat att att att Asn Ser Pro Gln Glu Val Ile Gly Lys Arg Gly Ser Asp Ile Ile Ile 435 440 445	1344
gtt ggt cgt ggt att att tct gct gct gat cgt ctt gaa gct gct gaa Val Gly Arg Gly Ile Ile Ser Ala Ala Asp Arg Leu Glu Ala Ala Glu 450 455 460	1392

ES 2 706 899 T3

atg tat cgt aaa gct gct tgg gaa gct tat ctt tct cgt ctt ggt gtt 1440
 Met Tyr Arg Lys Ala Ala Trp Glu Ala Tyr Leu Ser Arg Leu Gly Val
 465 470 475 480

taa 1443

5
 <210> 10
 <211> 480
 <212> PRT
 <213> homo sapiens
 <400> 10

Met Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr
 1 5 10 15

Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu
 20 25 30

Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg
 35 40 45

Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala
 50 55 60

Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro
 65 70 75 80

Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg
 85 90 95

Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr
 100 105 110

Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser
 115 120 125

Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu
 130 135 140

Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys
 145 150 155 160

Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu
 165 170 175

Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu
 180 185 190

ES 2 706 899 T3

Thr	Val	Gly	Arg	Val	Lys	Arg	Phe	Ile	Gln	Glu	Asn	Val	Phe	Val	Ala		
		195					200					205					
Ala	Asn	His	Asn	Gly	Ser	Pro	Leu	Ser	Ile	Lys	Glu	Ala	Pro	Lys	Glu		
	210					215					220						
Leu	Ser	Phe	Gly	Ala	Arg	Ala	Glu	Leu	Pro	Arg	Ile	His	Pro	Val	Ala		
225					230					235					240		
Ser	Lys	Leu	Leu	Arg	Leu	Met	Gln	Lys	Lys	Glu	Thr	Asn	Leu	Cys	Leu		
				245					250					255			
Ser	Ala	Asp	Val	Ser	Leu	Ala	Arg	Glu	Leu	Leu	Gln	Leu	Ala	Asp	Ala		
			260					265						270			
Leu	Gly	Pro	Ser	Ile	Cys	Met	Leu	Lys	Thr	His	Val	Asp	Ile	Leu	Asn		
		275					280					285					
Asp	Phe	Thr	Leu	Asp	Val	Met	Lys	Glu	Leu	Ile	Thr	Leu	Ala	Lys	Cys		
	290					295					300						
His	Glu	Phe	Leu	Ile	Phe	Glu	Asp	Arg	Lys	Phe	Ala	Asp	Ile	Gly	Asn		
305					310					315					320		
Thr	Val	Lys	Lys	Gln	Tyr	Glu	Gly	Gly	Ile	Phe	Lys	Ile	Ala	Ser	Trp		
				325					330					335			
Ala	Asp	Leu	Val	Asn	Ala	His	Val	Val	Pro	Gly	Ser	Gly	Val	Val	Lys		
			340					345					350				
Gly	Leu	Gln	Glu	Val	Gly	Leu	Pro	Leu	His	Arg	Gly	Cys	Leu	Leu	Ile		
		355					360					365					
Ala	Glu	Met	Ser	Ser	Thr	Gly	Ser	Leu	Ala	Thr	Gly	Asp	Tyr	Thr	Arg		
	370					375					380						
Ala	Ala	Val	Arg	Met	Ala	Glu	Glu	His	Ser	Glu	Phe	Val	Val	Gly	Phe		
385					390					395					400		
Ile	Ser	Gly	Ser	Arg	Val	Ser	Met	Lys	Pro	Glu	Phe	Leu	His	Leu	Thr		
				405					410					415			
Pro	Gly	Val	Gln	Leu	Glu	Ala	Gly	Gly	Asp	Asn	Leu	Gly	Gln	Gln	Tyr		
			420					425					430				

ES 2 706 899 T3

Asn Ser Pro Gln Glu Val Ile Gly Lys Arg Gly Ser Asp Ile Ile Ile
435 440 445

Val Gly Arg Gly Ile Ile Ser Ala Ala Asp Arg Leu Glu Ala Ala Glu
450 455 460

Met Tyr Arg Lys Ala Ala Trp Glu Ala Tyr Leu Ser Arg Leu Gly Val
465 470 475 480

<210> 11
<211> 1227
<212> ADN
<213> Secuencia artificial

<220>
<223> Construcción de fusión CDopt-UPRT

<220>
<221> CDS
<222> (1)..(1227)

<400> 11

atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat	48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp	
1 5 10 15	
atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct	96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro	
20 25 30	
atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg	144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg	
35 40 45	
ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag	192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu	
50 55 60	
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag	240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys	
65 70 75 80	
gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc	288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly	
85 90 95	
gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg	336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val	
100 105 110	
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag	384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu	
115 120 125	
gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc	432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe	
130 135 140	

ES 2 706 899 T3

atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag aac ccg Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Asn Pro 145 150 155 160	480
tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca tat ctt ata tat Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr Leu Ile Tyr 165 170 175	528
tat cca aac aaa ggg tct ttc gtt agc aaa cct aga aat ctg caa aaa Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn Leu Gln Lys 180 185 190	576
atg tct tcg gaa cca ttt aag aac gtc tac ttg cta cct caa aca aac Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn 195 200 205	624
caa ttg ctg ggt ttg tac acc atc atc aga aat aag aat aca act aga Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn Thr Thr Arg 210 215 220	672
cct gat ttc att ttc tac tcc gat aga atc atc aga ttg ttg gtt gaa Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu 225 230 235 240	720
gaa ggt ttg aac cat cta cct gtg caa aag caa att gtg gaa act gac Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp 245 250 255	768
acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt aaa atc tgt ggt Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly 260 265 270	816
gtt tcc att gtc aga gct ggt gaa tcg atg gag caa gga tta aga gac Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp 275 280 285	864
tgt tgt agg tct gtg cgt atc ggt aaa att tta att caa agg gac gag Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu 290 295 300	912
gag act gct tta cca aag tta ttc tac gaa aaa tta cca gag gat ata Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile 305 310 315 320	960
tct gaa agg tat gtc ttc cta tta gac cca atg ctg gcc acc ggt ggt Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly 325 330 335	1008
agt gct atc atg gct aca gaa gtc ttg att aag aga ggt gtt aag cca Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro 340 345 350	1056
gag aga att tac ttc tta aac cta atc tgt agt aag gaa ggg att gaa Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu 355 360 365	1104
aaa tac cat gcc gcc ttc cca gag gtc aga att gtt act ggt gcc ctc Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu 370 375 380	1152

ES 2 706 899 T3

gac aga ggt cta gat gaa aac aag tat cta gtt cca ggg ttg ggt gac 1200
 Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp
 385 390 395 400

ttt ggt gac aga tac tac tgt gtt taa 1227
 Phe Gly Asp Arg Tyr Tyr Cys Val
 405

<210> 12

<211> 408

<212> PRT

<213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 12

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
 100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
 130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Asn Pro
 145 150 155 160

Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr Leu Ile Tyr
 165 170 175

ES 2 706 899 T3

Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn Leu Gln Lys
 180 185 190
 Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn
 195 200 205
 Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn Thr Thr Arg
 210 215 220
 Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu
 225 230 235 240
 Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp
 245 250 255
 Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly
 260 265 270
 Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp
 275 280 285
 Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu
 290 295 300
 Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile
 305 310 315 320
 Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly
 325 330 335
 Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro
 340 345 350
 Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu
 355 360 365
 Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu
 370 375 380
 Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp
 385 390 395 400
 Phe Gly Asp Arg Tyr Tyr Cys Val
 405

<210> 13
 <211> 1287
 <212> ADN
 <213> Secuencia artificial

<220>

<223> Construcción de fusión - CDopt - enlazador - UPRT

<220>

<221> CDS

<222> (1)..(1287)

<400> 13

atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat	48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp	
1 5 10 15	
atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct	96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro	
20 25 30	
atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg	144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg	
35 40 45	
ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag	192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu	
50 55 60	
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag	240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys	
65 70 75 80	
gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc	288
Asp Thr Thr Leu Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly	
85 90 95	
gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg	336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val	
100 105 110	
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag	384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu	
115 120 125	
gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc	432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe	
130 135 140	
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tcc ggc	480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly	
145 150 155 160	
ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc	528
Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly	
165 170 175	
ggc gcc aac ccg tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca	576
Gly Ala Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr	

ES 2 706 899 T3

	180		185		190	
tat ctt ata tat tat cca aac aaa ggg tct ttc gtt agc aaa cct aga						624
Tyr Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg						
195			200		205	
aat ctg caa aaa atg tct tcg gaa cca ttt aag aac gtc tac ttg cta						672
Asn Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu						
210			215		220	
cct caa aca aac caa ttg ctg ggt ttg tac acc atc atc aga aat aag						720
Pro Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys						
225			230		235	240
aat aca act aga cct gat ttc att ttc tac tcc gat aga atc atc aga						768
Asn Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg						
245			250		255	
ttg ttg gtt gaa gaa ggt ttg aac cat cta cct gtg caa aag caa att						816
Leu Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile						
260			265		270	
gtg gaa act gac acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt						864
Val Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly						
275			280		285	
aaa atc tgt ggt gtt tcc att gtc aga gct ggt gaa tcg atg gag caa						912
Lys Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln						
290			295		300	
gga tta aga gac tgt tgt agg tct gtg cgt atc ggt aaa att tta att						960
Gly Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile						
305			310		315	320
caa agg gac gag gag act gct tta cca aag tta ttc tac gaa aaa tta						1008
Gln Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu						
325			330		335	
cca gag gat ata tct gaa agg tat gtc ttc cta tta gac cca atg ctg						1056
Pro Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu						
340			345		350	
gcc acc ggt ggt agt gct atc atg gct aca gaa gtc ttg att aag aga						1104
Ala Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg						
355			360		365	
ggt gtt aag cca gag aga att tac ttc tta aac cta atc tgt agt aag						1152
Gly Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys						
370			375		380	
gaa ggg att gaa aaa tac cat gcc gcc ttc cca gag gtc aga att gtt						1200
Glu Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val						
385			390		395	400
act ggt gcc ctc gac aga ggt cta gat gaa aac aag tat cta gtt cca						1248
Thr Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro						
405			410		415	
ggg ttg ggt gac ttt ggt gac aga tac tac tgt gtt taa						1287
Gly Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val						
420			425			

<211> 428
 <212> PRT
 <213> Secuencia artificial

5 <220>
 <223> Construcción sintética

<400> 14

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
 100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
 130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly
 145 150 155 160

Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly
 165 170 175

Gly Ala Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr
 180 185 190

10 Tyr Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg

ES 2 706 899 T3

195	200	205
Asn Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu 210 215 220		
Pro Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys 225 230 235 240		
Asn Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg 245 250 255		
Leu Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile 260 265 270		
Val Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly 275 280 285		
Lys Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln 290 295 300		
Gly Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile 305 310 315 320		
Gln Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu 325 330 335		
Pro Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu 340 345 350		
Ala Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg 355 360 365		
Gly Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys 370 375 380		
Glu Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val 385 390 395 400		
Thr Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro 405 410 415		
Gly Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val 420 425		

<210> 15
 <211> 1200
 <212> ADN
 <213> Secuencia artificial

<220>

<223> Construcción de fusión - CDopt3 - OPRT

<220>

<221> CDS

<222> (1)..(1200)

<400> 15

atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat	48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp	
1 5 10 15	
atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct	96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro	
20 25 30	
atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg	144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg	
35 40 45	
ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag	192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu	
50 55 60	
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag	240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys	
65 70 75 80	
gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc	288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly	
85 90 95	
gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg	336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val	
100 105 110	
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag	384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu	
115 120 125	
gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc	432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe	
130 135 140	
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag gcg gtc	480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ala Val	
145 150 155 160	
gct cgt gca gct ttg ggg cca ttg gtg acg ggt ctg tac gac gtg cag	528
Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr Asp Val Gln	
165 170 175	
gct ttc aag ttt ggg gac ttc gtg ctg aag agc ggg ctt tcc tcc ccc	576
Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu Ser Ser Pro	
180 185 190	
atc tac atc gat ctg cgg ggc atc gtg tct cga ccg cgt ctt ctg agt	624

ES 2 706 899 T3

Ile	Tyr	Ile	Asp	Leu	Arg	Gly	Ile	Val	Ser	Arg	Pro	Arg	Leu	Leu	Ser		
		195					200					205					
cag	gtt	gca	gat	att	tta	ttc	caa	act	gcc	caa	aat	gca	ggc	atc	agt	672	
Gln	Val	Ala	Asp	Ile	Leu	Phe	Gln	Thr	Ala	Gln	Asn	Ala	Gly	Ile	Ser		
	210					215					220						
ttt	gac	acc	gtg	tgt	gga	gtg	cct	tat	aca	gct	ttg	cca	ttg	gct	aca	720	
Phe	Asp	Thr	Val	Cys	Gly	Val	Pro	Tyr	Thr	Ala	Leu	Pro	Leu	Ala	Thr		
225					230					235					240		
gtt	atc	tgt	tca	acc	aat	caa	att	cca	atg	ctt	att	aga	agg	aaa	gaa	768	
Val	Ile	Cys	Ser	Thr	Asn	Gln	Ile	Pro	Met	Leu	Ile	Arg	Arg	Lys	Glu		
				245					250					255			
aca	aag	gat	tat	gga	act	aag	cgt	ctt	gta	gaa	gga	act	att	aat	cca	816	
Thr	Lys	Asp	Tyr	Gly	Thr	Lys	Arg	Leu	Val	Glu	Gly	Thr	Ile	Asn	Pro		
			260					265					270				
gga	gaa	acc	tgt	tta	atc	att	gaa	gat	gtt	gtc	acc	agt	gga	tct	agt	864	
Gly	Glu	Thr	Cys	Leu	Ile	Ile	Glu	Asp	Val	Val	Thr	Ser	Gly	Ser	Ser		
	275						280					285					
gtt	ttg	gaa	act	gtt	gag	gtt	ctt	cag	aag	gag	ggc	ttg	aag	gtc	act	912	
Val	Leu	Glu	Thr	Val	Glu	Val	Leu	Gln	Lys	Glu	Gly	Leu	Lys	Val	Thr		
	290					295					300						
gat	gcc	ata	gtg	ctg	ttg	gac	aga	gag	cag	gga	ggc	aag	gac	aag	ttg	960	
Asp	Ala	Ile	Val	Leu	Leu	Asp	Arg	Glu	Gln	Gly	Gly	Lys	Asp	Lys	Leu		
305					310					315					320		
cag	gcg	cac	ggg	atc	cgc	ctc	cac	tca	gtg	tgt	aca	ttg	tcc	aaa	atg	1008	
Gln	Ala	His	Gly	Ile	Arg	Leu	His	Ser	Val	Cys	Thr	Leu	Ser	Lys	Met		
				325					330					335			
ctg	gag	att	ctc	gag	cag	cag	aaa	aaa	gtt	gat	gct	gag	aca	gtt	ggg	1056	
Leu	Glu	Ile	Leu	Glu	Gln	Gln	Lys	Lys	Val	Asp	Ala	Glu	Thr	Val	Gly		
			340					345					350				
aga	gtg	aag	agg	ttt	att	cag	gag	aat	gtc	ttt	gtg	gca	gcg	aat	cat	1104	
Arg	Val	Lys	Arg	Phe	Ile	Gln	Glu	Asn	Val	Phe	Val	Ala	Ala	Asn	His		
		355					360					365					
aat	ggt	tct	ccc	ctt	tct	ata	aag	gaa	gca	ccc	aaa	gaa	ctc	agc	ttc	1152	
Asn	Gly	Ser	Pro	Leu	Ser	Ile	Lys	Glu	Ala	Pro	Lys	Glu	Leu	Ser	Phe		
	370					375					380						
ggt	gca	cgt	gca	gag	ctg	ccc	agg	atc	cac	cca	gtt	gca	tcg	aag	taa	1200	
Gly	Ala	Arg	Ala	Glu	Leu	Pro	Arg	Ile	His	Pro	Val	Ala	Ser	Lys			
	385				390					395							

<210> 16
 <211> 399
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Construcción sintética

<400> 16

ES 2 706 899 T3

Met	Val	Thr	Gly	Gly	Met	Ala	Ser	Lys	Trp	Asp	Gln	Lys	Gly	Met	Asp	1	5	10	15
Ile	Ala	Tyr	Glu	Glu	Ala	Leu	Leu	Gly	Tyr	Lys	Glu	Gly	Gly	Val	Pro	20	25	30	
Ile	Gly	Gly	Cys	Leu	Ile	Asn	Asn	Lys	Asp	Gly	Ser	Val	Leu	Gly	Arg	35	40	45	
Gly	His	Asn	Met	Arg	Phe	Gln	Lys	Gly	Ser	Ala	Thr	Leu	His	Gly	Glu	50	55	60	
Ile	Ser	Thr	Leu	Glu	Asn	Cys	Gly	Arg	Leu	Glu	Gly	Lys	Val	Tyr	Lys	65	70	75	80
Asp	Thr	Thr	Leu	Tyr	Thr	Thr	Leu	Ser	Pro	Cys	Asp	Met	Cys	Thr	Gly	85	90	95	
Ala	Ile	Ile	Met	Tyr	Gly	Ile	Pro	Arg	Cys	Val	Ile	Gly	Glu	Asn	Val	100	105	110	
Asn	Phe	Lys	Ser	Lys	Gly	Glu	Lys	Tyr	Leu	Gln	Thr	Arg	Gly	His	Glu	115	120	125	
Val	Val	Val	Val	Asp	Asp	Glu	Arg	Cys	Lys	Lys	Leu	Met	Lys	Gln	Phe	130	135	140	
Ile	Asp	Glu	Arg	Pro	Gln	Asp	Trp	Phe	Glu	Asp	Ile	Gly	Glu	Ala	Val	145	150	155	160
Ala	Arg	Ala	Ala	Leu	Gly	Pro	Leu	Val	Thr	Gly	Leu	Tyr	Asp	Val	Gln	165	170	175	
Ala	Phe	Lys	Phe	Gly	Asp	Phe	Val	Leu	Lys	Ser	Gly	Leu	Ser	Ser	Pro	180	185	190	
Ile	Tyr	Ile	Asp	Leu	Arg	Gly	Ile	Val	Ser	Arg	Pro	Arg	Leu	Leu	Ser	195	200	205	
Gln	Val	Ala	Asp	Ile	Leu	Phe	Gln	Thr	Ala	Gln	Asn	Ala	Gly	Ile	Ser	210	215	220	
Phe	Asp	Thr	Val	Cys	Gly	Val	Pro	Tyr	Thr	Ala	Leu	Pro	Leu	Ala	Thr	225	230	235	240

ES 2 706 899 T3

Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg Arg Lys Glu
245 250 255

Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr Ile Asn Pro
260 265 270

Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser Gly Ser Ser
275 280 285

Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu Lys Val Thr
290 295 300

Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys Asp Lys Leu
305 310 315 320

Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu Ser Lys Met
325 330 335

Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu Thr Val Gly
340 345 350

Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala Ala Asn His
355 360 365

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

Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala Ser Lys
385 390 395

<210> 17

<211> 1260

<212> ADN

<213> Secuencia artificial

<220>

<223> Construcción de fusión - CDopt3 - enlazador - OPRT

<220>

<221> CDS

<222> (1)..(1260)

<400> 17

atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat 48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct 96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

ES 2 706 899 T3

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

ES 2 706 899 T3

aga agg aaa gaa aca aag gat tat gga act aag cgt ctt gta gaa gga	864
Arg Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly	
275 280 285	
act att aat cca gga gaa acc tgt tta atc att gaa gat gtt gtc acc	912
Thr Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr	
290 295 300	
agt gga tct agt gtt ttg gaa act gtt gag gtt ctt cag aag gag ggc	960
Ser Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly	
305 310 315 320	
ttg aag gtc act gat gcc ata gtg ctg ttg gac aga gag cag gga ggc	1008
Leu Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly	
325 330 335	
aag gac aag ttg cag gcg cac ggg atc cgc ctc cac tca gtg tgt aca	1056
Lys Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr	
340 345 350	
ttg tcc aaa atg ctg gag att ctc gag cag cag aaa aaa gtt gat gct	1104
Leu Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala	
355 360 365	
gag aca gtt ggg aga gtg aag agg ttt att cag gag aat gtc ttt gtg	1152
Glu Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val	
370 375 380	
gca gcg aat cat aat ggt tct ccc ctt tct ata aag gaa gca ccc aaa	1200
Ala Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys	
385 390 395 400	
gaa ctc agc ttc ggt gca cgt gca gag ctg ccc agg atc cac cca gtt	1248
Glu Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val	
405 410 415	
gca tcg aag taa	1260
Ala Ser Lys	

<210> 18

<211> 419

<212> PRT

<213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 18

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

ES 2 706 899 T3

Gly	His	Asn	Met	Arg	Phe	Gln	Lys	Gly	Ser	Ala	Thr	Leu	His	Gly	Glu
50						55					60				
Ile	Ser	Thr	Leu	Glu	Asn	Cys	Gly	Arg	Leu	Glu	Gly	Lys	Val	Tyr	Lys
65					70					75					80
Asp	Thr	Thr	Leu	Tyr	Thr	Thr	Leu	Ser	Pro	Cys	Asp	Met	Cys	Thr	Gly
				85					90					95	
Ala	Ile	Ile	Met	Tyr	Gly	Ile	Pro	Arg	Cys	Val	Ile	Gly	Glu	Asn	Val
			100					105					110		
Asn	Phe	Lys	Ser	Lys	Gly	Glu	Lys	Tyr	Leu	Gln	Thr	Arg	Gly	His	Glu
		115					120					125			
Val	Val	Val	Val	Asp	Asp	Glu	Arg	Cys	Lys	Lys	Leu	Met	Lys	Gln	Phe
	130					135					140				
Ile	Asp	Glu	Arg	Pro	Gln	Asp	Trp	Phe	Glu	Asp	Ile	Gly	Glu	Ser	Gly
145					150					155					160
Gly	Gly	Ala	Ser	Gly	Gly	Gly	Ala	Ser	Gly	Gly	Gly	Ala	Ser	Gly	Gly
				165					170					175	
Gly	Ala	Ala	Val	Ala	Arg	Ala	Ala	Leu	Gly	Pro	Leu	Val	Thr	Gly	Leu
			180					185					190		
Tyr	Asp	Val	Gln	Ala	Phe	Lys	Phe	Gly	Asp	Phe	Val	Leu	Lys	Ser	Gly
		195					200					205			
Leu	Ser	Ser	Pro	Ile	Tyr	Ile	Asp	Leu	Arg	Gly	Ile	Val	Ser	Arg	Pro
	210					215					220				
Arg	Leu	Leu	Ser	Gln	Val	Ala	Asp	Ile	Leu	Phe	Gln	Thr	Ala	Gln	Asn
225					230					235					240
Ala	Gly	Ile	Ser	Phe	Asp	Thr	Val	Cys	Gly	Val	Pro	Tyr	Thr	Ala	Leu
				245					250					255	
Pro	Leu	Ala	Thr	Val	Ile	Cys	Ser	Thr	Asn	Gln	Ile	Pro	Met	Leu	Ile
			260					265					270		
Arg	Arg	Lys	Glu	Thr	Lys	Asp	Tyr	Gly	Thr	Lys	Arg	Leu	Val	Glu	Gly
		275					280					285			

ES 2 706 899 T3

Thr Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr
 290 295 300

Ser Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly
 305 310 315 320

Leu Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly
 325 330 335

Lys Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr
 340 345 350

Leu Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala
 355 360 365

Glu Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val
 370 375 380

Ala Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys
 385 390 395 400

Glu Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val
 405 410 415

Ala Ser Lys

<210> 19
 <211> 11892
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Vector RCR - pAC3-yCD2

<400> 19

tagttattaa tagtaatcaa ttacgggggtc attagttcat agcccatata tggagttccg	60
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt	120
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca	180
atgggtggag tatttacggt aaactgccc cttggcagta catcaagtgt atcatatgcc	240
aagtagccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta	300
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	420
atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480

ES 2 706 899 T3

ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaatgggcg	gtaggcgtgt	540
acggtgggag	gtctatataa	gcagagctgg	tttagtgaac	cggcgccagt	cctccgattg	600
actgagtcgc	ccgggtaccc	gtgtatccaa	taaacctct	tgcagttgca	tccgacttgt	660
ggctctcgctg	ttccttgga	gggtctcctc	tgagtgattg	actaccgcgc	agcggggggtc	720
tttcatttgg	gggctcgtec	gggatcgga	gacccctgcc	cagggaccac	cgacccacca	780
ccgggaggta	agctggccag	caacttatct	gtgtctgtcc	gattgtctag	tgtctatgac	840
tgattttatg	cgctgcgtc	ggtactagtt	agctaactag	ctctgtatct	ggcggaccgc	900
tgggtggaact	gacgagttcg	gaacaccgcg	ccgcaaccct	gggagacgtc	ccagggactt	960
cggggggccgt	ttttgtggcc	cgacctgagt	ccaaaaatcc	cgatcgtttt	ggactctttg	1020
gtgcaccccc	cttagaggag	ggatatgtgg	ttctggtagg	agacgagaac	ctaaaacagt	1080
tccgcctcc	gtctgaattt	ttgctttcgg	tttgggaccg	aagccgcgcc	gcgcgtcttg	1140
tctgctgcag	catcgttctg	tgttgtctct	gtctgactgt	gtttctgtat	ttgtctgaga	1200
atatgggcca	gactgttacc	actcccttaa	gtttgacctt	aggtcactgg	aaagatgtcg	1260
agcggatcgc	tcacaaccag	tcggtagatg	tcaagaagag	acgttgggtt	accttctgct	1320
ctgcagaatg	gccaaccttt	aacgtcggat	ggccgcgaga	cggcaccttt	aaccgagacc	1380
tcatacccca	ggttaagatc	aaggctcttt	cacctggccc	gcatggacac	ccagaccagg	1440
tcccctacat	cgtgacctgg	gaagccttgg	cttttgaccc	ccctccctgg	gtcaagccct	1500
ttgtacaccc	taagcctccg	cctcctcttc	ctccatccgc	cccgtctctc	ccccttgaac	1560
ctcctcgttc	gaccccgcc	cgatcctccc	tttatccagc	cctcactcct	tctctaggcg	1620
ccaaacctaa	acctcaagtt	ctttctgaca	gtgggggggc	gctcatcgac	ctacttacag	1680
aagacccccc	gccttatagg	gacccaagac	caccccttcc	cgacagggac	ggaaatggtg	1740
gagaagcgac	ccctgcggga	gaggcaccgg	acccctcccc	aatggcatct	cgcctacgtg	1800
ggagacggga	gccccctgtg	gccgactcca	ctacctcgca	ggcattcccc	ctccgcgcag	1860
gaggaaacgg	acagcttcaa	tactggccgt	tctcctcttc	tgacctttac	aactggaaaa	1920
ataataaccc	ttctttttct	gaagatccag	gtaaactgac	agctctgata	gagctctgtc	1980
tcatacccca	tcagcccacc	tgggacgact	gtcagcagct	gttggggact	ctgctgaccg	2040
gagaagaaaa	acaacgggtg	ctcttagagg	ctagaaaggc	ggtgcggggc	gatgatgggc	2100
gccccactca	actgccaat	gaagtcgatg	ccgcttttcc	cctcgagcgc	ccagactggg	2160
attacaccac	ccaggcaggt	aggaaccacc	tagtccacta	tcgccagttg	ctcctagcgg	2220
gtctccaaaa	cgcgggcaga	agccccacca	atttggccaa	ggtaaaagga	ataacacaag	2280

ES 2 706 899 T3

ggcccaatga	gtctccctcg	gccttcctag	agagacttaa	ggaagcctat	cgcaggtaca	2340
ctccttatga	ccctgaggac	ccagggcaag	aaactaatgt	gtctatgtct	ttcatttggc	2400
agtctgcccc	agacattggg	agaaagttag	agaggttaga	agatttaaaa	aacaagacgc	2460
ttggagattt	ggttagagag	gcagaaaaga	tctttaataa	acgagaaacc	ccggaagaaa	2520
gagaggaacg	tatcaggaga	gaaacagagg	aaaaagaaga	acgccgtagg	acagaggatg	2580
agcagaaaga	gaaagaaaga	gacgtagga	gacatagaga	gatgagcaag	ctattggcca	2640
ctgtcgtagg	tggacagaaa	caggatagac	agggaggaga	acgaaggagg	tcccaactcg	2700
atcgcgacca	gtgtgcctac	tgcaaagaaa	aggggcactg	ggctaaagat	tgtcccaaga	2760
aaccacgagg	acctcgggga	ccaagacccc	agacctccct	cctgacccta	gatgactagg	2820
gaggtcaggg	tcaggagccc	ccccctgaac	ccaggataac	cctcaaagtc	ggggggcaac	2880
ccgtcacctt	cctggtagat	actggggccc	aacctccgt	gctgacccaa	aatcctggac	2940
ccctaagtga	taagtctgcc	tgggtccaag	gggctactgg	aggaaagcgg	tatcgctgga	3000
ccacggatcg	caaagtacat	ctagctaccg	gtaaggtcac	ccactctttc	ctccatgtac	3060
cagactgtcc	ctatcctctg	ttaggaagag	atttgctgac	taaactaaaa	gcccaaattc	3120
actttgaggg	atcaggagcc	caggttatgg	gaccaatggg	gcagcccctg	caagtgttga	3180
ccctaaatat	agaagatgag	catcggttac	atgagacctc	aaaagagcca	gatgtttctc	3240
taggggtccac	atggctgtct	gattttcctc	aggcctgggc	ggaaaccggg	ggcatgggac	3300
tggcagttcg	ccaagctcct	ctgatcatac	ctctgaaagc	aacctctacc	cccgtgtcca	3360
taaaacaata	ccccatgtca	caagaagcca	gactggggat	caagccccac	atacagagac	3420
tgttggaacca	gggaataactg	gtaccctgcc	agtccccctg	gaacaagccc	ctgctaccgg	3480
ttaagaaacc	agggactaat	gattataggc	ctgtccagga	tctgagagaa	gtcaacaagc	3540
gggtggaaga	catccacccc	accgtgccc	acccttaca	cctcttgagc	gggctccac	3600
cgtcccacca	gtggtacact	gtgcttgatt	taaaggatgc	ctttttctgc	ctgagactcc	3660
accccaccag	tcagcctctc	ttcgcccttg	agtggagaga	tccagagatg	ggaatctcag	3720
gacaattgac	ctggaccaga	ctccacagg	gtttcaaaaa	cagtcccacc	ctgtttgatg	3780
aggcactgca	cagagacctc	gcagacttcc	ggatccagca	cccagacttg	atcctgctac	3840
agtacgtgga	tgacttactg	ctggccgcca	cttctgagct	agactgcca	caaggctactc	3900
gggccctggt	acaaacccta	gggaacctcg	ggtatcgggc	ctcggccaag	aaagcccaaa	3960
tttgccagaa	acaggtcaag	tatctgggg	atcttctaaa	agagggtcag	agatggctga	4020
ctgaggccag	aaaagagact	gtgatggggc	agcctactcc	gaagaccctc	cgacaactaa	4080
gggagttcct	agggacggca	ggcttctgtc	gcctctggat	ccctgggttt	gcagaaatgg	4140

ES 2 706 899 T3

cagccccctt	gtaccctctc	accaaaacgg	ggactctgtt	taattggggc	ccagaccaac	4200
aaaaggccta	tcaagaaatc	aagcaagctc	ttctaactgc	cccagccctg	gggttgccag	4260
atttgactaa	gccctttgaa	ctctttgtcg	acgagaagca	gggctacgcc	aaaggtgtcc	4320
taacgcaaaa	actgggacct	tggcgtcggc	cgggtggccta	cctgtccaaa	aagctagacc	4380
cagtagcagc	tgggtggccc	ccttgcctac	ggatggtagc	agccattgcc	gtactgacaa	4440
aggatgcagg	caagctaacc	atgggacagc	cactagtcat	tctggcccc	catgcagtag	4500
aggcactagt	caaacaaccc	cccgaccgct	ggctttccaa	cgcccggatg	actcactatc	4560
aggccttgct	tttggacacg	gaccgggtcc	agttcggacc	ggtggtagcc	ctgaacccgg	4620
ctacgtgct	cccactgcct	gaggaagggc	tgcaacacaa	ctgccttgat	atcctggccg	4680
aagcccacgg	aacccgaccc	gacctaacgg	accagccgct	cccagacgcc	gaccacacct	4740
ggtacacgga	tggaagcagt	ctcttacaag	agggacagcg	taaggcggga	gctgcggtga	4800
ccaccgagac	cgaggtaatc	tgggctaaag	ccctgccagc	cgggacatcc	gctcagcggg	4860
ctgaactgat	agcactcacc	caggccctaa	agatggcaga	aggtaagaag	ctaaatgttt	4920
atactgatag	ccgttatgct	tttgctactg	cccatatcca	tggagaaata	tacagaaggc	4980
gtgggttgct	cacatcagaa	ggcaaagaga	tcaaaaataa	agacgagatc	ttggccctac	5040
taaaagccct	ctttctgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaagg	5100
gacacagcgc	cgaggctaga	ggcaaccgga	tggctgacca	agcggcccgga	aaggcagcca	5160
tcacagagac	tccagacacc	tctaccctcc	tcatagaaaa	ttcatcacc	tacacctcag	5220
aacattttca	ttacacagtg	actgatataa	aggacctaac	caagttgggg	gccatttatg	5280
ataaaacaaa	gaagtattgg	gtctaccaag	gaaaacctgt	gatgcctgac	cagtttactt	5340
ttgaattatt	agactttctt	catcagctga	ctcacctcag	cttctcaaaa	atgaaggctc	5400
tcctagagag	aagccacagt	ccctactaca	tgctgaaccg	ggatcgaaca	ctcaaaaata	5460
tcactgagac	ctgcaaagct	tgtgcacaag	tcaacgccag	caagtctgcc	gttaaacagg	5520
gaactagggg	ccgcgggcat	cggcccggca	ctcattggga	gatcgatttc	accgagataa	5580
agcccggatt	gtatggctat	aaatatcttc	tagtttttat	agataccttt	tctggctgga	5640
tagaagcctt	cccaaccaag	aaagaaaccg	ccaaggtcgt	aaccaagaag	ctactagagg	5700
agatcttccc	caggttcggc	atgcctcagg	tattgggaac	tgacaatggg	cctgccttcg	5760
tctccaagg	gagtcagaca	gtggccgatc	tgttggggat	tgattggaaa	ttacattgtg	5820
catacagacc	ccaaagctca	ggccaggtag	aaagaatgaa	tagaaccatc	aaggagactt	5880
taactaaatt	aacgcttgca	actggctcta	gagactgggt	gctcctactc	cccttagccc	5940

ES 2 706 899 T3

tgtaccgagc	ccgcaacacg	ccgggcccc	atggcctcac	cccatatgag	atcttatatg	6000
gggcaccccc	gccccttgta	aacttccttg	accctgacat	gacaagagtt	actaacagcc	6060
cctctctcca	agctcactta	caggctctct	acttagtcca	gcacgaagtc	tggagacctc	6120
tggcggcagc	ctaccaagaa	caactggacc	gaccggtggg	acctcaccct	taccgagtcg	6180
gcgacacagt	gtgggtccgc	cgacaccaga	ctaagaacct	agaacctcgc	tggaaaggac	6240
cttacacagt	cctgctgacc	acccccaccg	ccctcaaagt	agacggcatc	gcagcttgga	6300
tacacgccgc	ccacgtgaag	gctgccgacc	ccgggggtgg	accatcctct	agactgacat	6360
ggcgcgttca	acgctctcaa	aacccccctc	agataagatt	aacccgtgga	agcccttaat	6420
agtcatggga	gtcctgttag	gagtagggat	ggcagagagc	ccccatcagg	tctttaatgt	6480
aacctggaga	gtcaccaacc	tgatgactgg	gcgtaccgcc	aatgccacct	ccctcctggg	6540
aactgtacaa	gatgccttcc	caaaattata	ttttgatcta	tgtgatctgg	tggagagga	6600
gtgggaccct	tcagaccagg	aaccgtatgt	cgggtatggc	tgcaagtacc	ccgcagggag	6660
acagcggacc	cggacttttg	acttttacgt	gtgccctggg	cataccgtaa	agtcgggggtg	6720
tgggggacca	ggagagggct	actgtggtaa	atgggggtgt	gaaaccaccg	gacaggctta	6780
ctggaagccc	acatcatcgt	gggacctaat	ctcccttaag	cgcggttaaca	ccccctggga	6840
cacgggatgc	tctaaagttg	cctgtggccc	ctgctacgac	ctctccaaag	tatccaattc	6900
cttccaaggg	gctactcgag	ggggcagatg	caaccctcta	gtcctagaat	tactgatgc	6960
aggaaaaaag	gctaactggg	acggggccaa	atcgtgggga	ctgagactgt	accggacagg	7020
aacagatcct	attaccatgt	tctccctgac	ccggcaggtc	cttaatgtgg	gaccccgagt	7080
ccccataggg	cccaaccag	tattaccgga	caaagactc	ccttcctcac	caatagagat	7140
tgtaccggct	ccacagccac	ctagccccct	caataccagt	tacccccctt	ccactaccag	7200
tacacctca	acctccccta	caagtccaag	tgtcccacag	ccacccccag	gaactggaga	7260
tagactacta	gctctagtca	aaggagccta	tcaggcgctt	aacctacca	atcccgacaa	7320
gaccaagaa	tgttggctgt	gcttagtgtc	gggacctcct	tattacgaag	gagtagcggg	7380
cgtgggcact	tataccaatc	attccaccgc	tccggccaac	tgtacggcca	cttcccaaca	7440
taagcttacc	ctatctgaag	tgacaggaca	gggcctatgc	atgggggcag	tacctaaaac	7500
tcaccaggcc	ttatgtaaca	ccacccaaag	cgccggctca	ggatcctact	accttgacgc	7560
acccgccgga	acaatgtggg	cttgacgac	tggattgact	ccctgcttgt	ccaccacggg	7620
gctcaatcta	accacagatt	attgtgtatt	agttgaactc	tggcccagag	taattttacca	7680
ctccccgat	tatatgtatg	gtcagcttga	acagcgtacc	aaatataaaa	gagagccagt	7740
atcattgacc	ctggcccttc	tactaggagg	attaacctatg	ggagggattg	cagctggaat	7800

ES 2 706 899 T3

agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccgctat	7860
ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc	7920
gttgtctgaa gtagtcctac agaaccgcag aggccatagat ttgctattcc taaaggaggg	7980
aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt	8040
gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac	8100
aggccaagga tggttcgaag ggctgtttta tagatcccc tggtttacca ccttaatctc	8160
caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct	8220
caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac	8280
tcagcaatat caccagctaa aacccataga gtacgagcca tgaacgcgtt actggccgaa	8340
gccgcttgga ataaggccgg tgtgcgtttg tctatatgtt attttccacc atattgccgt	8400
cttttgga tgtgagggcc cggaaacctg gccctgtctt cttgacgagc attcctaggg	8460
gtctttcccc tctgcgcaaa ggaatgcaag gtctgttgaa tgtcgtgaag gaagcagttc	8520
ctctggaagc ttctgaaga caaacaacgt ctgtagcgac cctttgcagg cagcggaacc	8580
ccccacctgg cgacaggtgc ctctgcggcc aaaagccacg tgtataagat acacctgcaa	8640
aggcggcaca accccagtgc cacgttgatg gttggatagt tgtggaaaga gtcaaatggc	8700
tctcctcaag cgtattcaac aaggggctga aggatgcccc gaagggtacc cattgtatgg	8760
gatctgatct ggggcctcgg tgcacatgct ttacatgtgt ttagtcgagg ttaaaaaaac	8820
gtctaggccc cccgaaccac ggggacgtgg ttttcctttg aaaaacacga ttataaatgg	8880
tgaccggcgg catggcctcc aagtgggatc aaaagggcat ggatatcgct tacgaggagg	8940
ccctgctggg ctacaaggag ggcggcgtgc ctatcggcgg ctgtctgatc aacaacaagg	9000
acggcagtgt gctgggcagg ggccacaaca tgagggtcca gaagggtcc gccaccctgc	9060
acggcgagat ctccacctg gagaactgtg gcaggctgga gggcaagggtg tacaaggaca	9120
ccacctgta caccacctg tccccttggtg acatgtgtac cggcgctatc atcatgtacg	9180
gcatccctag gtgtgtgatc ggcgagaacg tgaacttcaa gtccaagggtc gagaagtacc	9240
tgcaaacag gggccacgag gtggtggttg ttgacgatga gaggtgtaag aagctgatga	9300
agcagttcat cgacgagagg cctcaggact gggtcgagga tatcggcgag taagcggccg	9360
cagataaaat aaaagatttt atttagtctc cagaaaaagg ggggaatgaa agaccccacc	9420
tgtaggtttg gcaagctagc ttaagtaacg ccattttgca aggcattggaa aaatacataa	9480
ctgagaatag agaagttcag atcaaggatc ggaacagatg gaacagctga atatgggcca	9540
aacaggatat ctgtggtaag cagttcctgc ccggtctcag ggccaagaac agatggaaca	9600

ES 2 706 899 T3

gctgaatatg	ggccaaacag	gatatctgtg	gtaagcagtt	cctgccccgg	ctcagggcca	9660
agaacagatg	gtccccagat	gcggtccagc	cctcagcagt	ttctagagaa	ccatcagatg	9720
tttccagggg	gccccaaagga	cctgaaatga	ccctgtgcct	tatttgaact	aaccaatcag	9780
ttcgcttctc	gcttctgttc	gcgcgcttct	gtccccgag	ctcaataaaa	gagcccacaa	9840
cccctcactc	ggggcgccag	tcctccgatt	gactgagtcg	cccgggtacc	cgtgtatcca	9900
ataaaccttc	ttgcagttgc	atccgacttg	tggtctcgct	gttccttggg	agggctctct	9960
ctgagtgatt	gactaccggt	cagcgggggt	ctttcattac	atgtgagcaa	aaggccagca	10020
aaaggccagg	aaccgtaaaa	aggccgcgtt	gctggcgttt	ttccataggc	tccgcccccc	10080
tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaaccgga	caggactata	10140
aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	cgaccctgcc	10200
gcttaccgga	tacctgtccg	cctttctccc	ttcgggaagc	gtggcgcttt	ctcaatgctc	10260
acgctgtagg	tatctcagtt	cggtgtaggt	cgttcgctcc	aagctggggt	gtgtgcacga	10320
accccccggt	cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	agtcacaacc	10380
ggtaagacac	gacttatcgc	cactggcagc	agccactggg	aacaggatta	gcagagcgag	10440
gtatgtaggc	ggtgctacag	agttcttgaa	gtggtggcct	aactacggct	acactagaag	10500
gacagtatct	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	10560
ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcgggtgg	ttttttgttt	gcaagcagca	10620
gattacgcgc	agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	cgggggtctga	10680
cgctcagtg	aacgaaaact	cacgttaagg	gattttgggt	atgagattat	caaaaaggat	10740
cttcacctag	atccttttta	attaaaaatg	aagtttttaa	tcaatctaaa	gtatatatga	10800
gtaaacttgg	tctgacagtt	accaatgctt	aatcagtgag	gcacctatct	cagcgatctg	10860
tctatttcgt	tcattccatg	ttgcctgact	ccccgtcgtg	tagataacta	cgatacggga	10920
gggcttacca	tctggcccca	gtgctgcaat	gataccgcga	gaccacgct	caccggctcc	10980
agatttatca	gcaataaacc	agccagccgg	aaggggccgag	cgcagaagtg	gtcctgcaac	11040
tttatccgcc	tccatccagt	ctattaattg	ttgccgggaa	gctagagtaa	gtagtctgcc	11100
agttaatagt	ttgcgcaacg	ttgttgccat	tgctgcaggc	atcgtggtgt	cacgctcgtc	11160
gtttggtatg	gcttcattca	gctccggttc	ccaacgatca	aggcgagtta	catgatcccc	11220
catgttgtgc	aaaaaagcgg	ttagctcctt	cggctcctccg	atcgttgtca	gaagtaagtt	11280
ggcgcagtg	ttatcaactc	tggttatggc	agcaactgcat	aattctctta	ctgtcatgcc	11340
atccgtaaga	tgcttttctg	tgactggtga	gtactcaacc	aagtcattct	gagaatagtg	11400
tatgcggcga	ccgagttgct	cttgcccggc	gtcaacacgg	gataataccg	cgccacatag	11460

ES 2 706 899 T3

cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat 11520
 cttaccgctg ttgagatcca gttcgatgta acccactcgt gcaccaact gatcttcagc 11580
 atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa 11640
 aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta 11700
 ttgaagcatt tatcaggggtt attgtctcat gagcggatac atatttgaat gtatttagaa 11760
 aaataaacia ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga 11820
 aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtct 11880
 tcaagaattc at 11892

<210> 20
 <211> 11892
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Vector RCR - pAC3-yCD

<400> 20

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg 60
 cgttacataa cttacggtaa atggcccgcc tggtgaccg cccaacgacc ccgcgccatt 120
 gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 180
 atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt atcatatgcc 240
 aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta 300
 catgacctta tgggactttc ctacttgga gtacatctac gtattagtca tcgctattac 360
 catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 420
 atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 480
 ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt 540
 acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg 600
 actgagtcgc ccgggtaccc gtgtatccaa taaacctct tgcagttgca tccgacttgt 660
 ggtctcgtg ttccttggga gggctctctc tgagtgattg actaccgctc agcgggggtc 720
 tttcatttgg gggctcgtcc gggatcgga gaccctgcc cagggaccac cgaccacca 780
 ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac 840
 tgattttatg cgctgcgtc ggtactagtt agctaactag ctctgtatct ggcggaccgc 900
 tgggtggaact gacgagttcg gaacaccgcg ccgcaacct gggagacgtc ccagggactt 960
 cgggggcccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg 1020

ES 2 706 899 T3

gtgcaccccc	cttagaggag	ggatatgtgg	ttctggtagg	agacgagaac	ctaaaacagt	1080
tcccgcctcc	gtctgaattt	ttgctttcgg	tttgggaccg	aagccgcgcc	gcgcgtcttg	1140
tctgctgcag	catcgttctg	tgttgctctt	gtctgactgt	gtttctgtat	ttgtctgaga	1200
atatgggcca	gactgttacc	actcccttaa	gtttgacctt	aggtcactgg	aaagatgtcg	1260
agcggatcgc	tcacaaccag	tcggtagatg	tcaagaagag	acgttggggt	accttctgct	1320
ctgcagaatg	gccaaccttt	aacgtcggat	ggccgcgaga	cggcaccttt	aaccgagacc	1380
tcatacccca	ggtaagatc	aaggtctttt	caactggccc	gcatggacac	ccagaccagg	1440
tcccctacat	cgtgacctgg	gaagccttgg	cttttgaccc	ccctccctgg	gtcaagccct	1500
ttgtacaccc	taagcctccg	cctcctcttc	ctccatccgc	cccgtctctc	ccccttgaac	1560
ctcctcgttc	gaccccgctt	cgatcctccc	tttatccagc	cctcactcct	tctctaggcg	1620
ccaaacctaa	acctcaagtt	ctttctgaca	gtgggggggc	gctcatcgac	ctacttacag	1680
aagacccccc	gccttatagg	gacccaagac	caccccttcc	cgacagggac	ggaaatggtg	1740
gagaagcgac	ccctgcggga	gaggcaccgg	acccctcccc	aatggcatct	cgcctacgty	1800
ggagacggga	gccccctgtg	gccgactcca	ctacctcgca	ggcattcccc	ctccgcgcag	1860
gaggaaacgg	acagcttcaa	tactggccgt	tctcctcttc	tgacctttac	aactggaaaa	1920
ataataaccc	ttctttttct	gaagatccag	gtaaaactgac	agctctgata	gagtctgttc	1980
tcatacccca	tcagcccacc	tgggacgact	gtcagcagct	gttggggact	ctgctgaccg	2040
gagaagaaaa	acaacgggtg	ctcttagagg	ctagaaaggc	ggtgcggggc	gatgatgggc	2100
gccccactca	actgcccatt	gaagtcgatg	ccgcttttcc	cctcgagcgc	ccagactggg	2160
attacaccac	ccaggcaggt	aggaaccacc	tagtccacta	tcgccagttg	ctcctagcgg	2220
gtctccaaaa	cgcgggcgaga	agccccacca	atttggccaa	ggtaaaagga	ataacacaag	2280
ggcccaatga	gtctccctcg	gccttcctag	agagacttaa	ggaagcctat	cgcaggtaca	2340
ctccttatga	ccctgaggac	ccagggcaag	aaactaatgt	gtctatgtct	ttcatttggc	2400
agtctgcccc	agacattggg	agaaagttag	agagggttaga	agatttaaaa	aacaagacgc	2460
ttggagattt	ggttagagag	gcagaaaaga	tctttaataa	acgagaaaacc	ccggaagaaa	2520
gagaggaacg	tatcaggaga	gaaacagagg	aaaaagaaga	acgccgtagg	acagaggatg	2580
agcagaaaga	gaaagaaaga	gatcgtagga	gacatagaga	gatgagcaag	ctattggcca	2640
ctgtcgttag	tggacagaaa	caggatagac	agggaggaga	acgaaggagg	tcccaactcg	2700
atcgcgacca	gtgtgcctac	tgcaaagaaa	aggggcaactg	ggctaaagat	tgtcccaaga	2760
aaccacgagg	acctcgggga	ccaagacccc	agacctccct	cctgacccta	gatgactagg	2820
gaggtcaggg	tcaggagccc	ccccctgaac	ccaggataac	cctcaaagtc	ggggggcaac	2880

ES 2 706 899 T3

ccgtcacctt cctggtagat actggggccc aacactccgt gctgacccaa aatcctggac 2940
 ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgttgga 3000
 ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactctttc ctccatgtac 3060
 cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaaatcc 3120
 actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga 3180
 ccctaaatat agaagatgag catcggctac atgagacctc aaaagagcca gatgtttctc 3240
 taggggtccac atggctgtct gattttcttc aggcctgggc ggaaaccggg ggcatgggac 3300
 tggcagttcg ccaagtcctt ctgatcatac ctctgaaagc aacctctacc cccgtgtcca 3360
 taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac 3420
 tgttggaacca gggaatactg gtaccctgcc agtccccctg gaacacgccc ctgctaccog 3480
 ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc 3540
 ggggtggaaga catccacccc accgtgccca acccttaciaa cctcttgagc gggctccac 3600
 cgtcccacca gtggtacact gtgcttgatt taaaggatgc ctttttctgc ctgagactcc 3660
 accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag 3720
 gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg 3780
 aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac 3840
 agtacgtgga tgacttactg ctggccgcca cttctgagct agactgcaa caaggtactc 3900
 gggccctgtt acaaacccta gggaacctcg ggtatcgggc ctcgccaag aaagcccaaa 3960
 tttgccagaa acaggtcaag tatctggggt atcttctaaa agagggtcag agatggctga 4020
 ctgaggccag aaaagagact gtgatggggc agcctactcc gaagaccctc cgacaactaa 4080
 gggagttcct agggacggca ggcttctgtc gcctctggat ccctggggtt gcagaaatgg 4140
 cagccccctt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac 4200
 aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccag 4260
 atttgactaa gccctttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc 4320
 taacgcaaaa actgggacct tggcgtcggc cggtggccta cctgtccaaa aagctagacc 4380
 cagtagcagc tgggtggccc ccttgccctac ggatggtagc agccattgcc gtactgacaa 4440
 aggatgcagg caagctaacc atgggacagc cactagtcac tctggcccc catgcagtag 4500
 aggcactagt caaacaaccc cccgaccgct ggctttccaa cgcccggatg actcactatc 4560
 aggccttgct tttggacacg gaccgggtcc agttcggacc ggtggtagcc ctgaaccggg 4620
 ctacgtgct cccactgcct gaggaagggc tgcaacacaa ctgccttgat atcctggccg 4680

ES 2 706 899 T3

aagcccacgg	aacccgaccc	gacctaacgg	accagccgct	cccagacgcc	gaccacacct	4740
ggtacacgga	tggaagcagt	ctcttacaag	agggacagcg	taaggcggga	gctgcggtga	4800
ccaccgagac	cgaggtaatc	tgggctaaag	ccctgccagc	cgggacatcc	gctcagcggg	4860
ctgaactgat	agcactcacc	caggccctaa	agatggcaga	aggtaagaag	ctaaatgttt	4920
atactgatag	ccgttatgct	tttgctactg	cccatatcca	tggagaaata	tacagaaggc	4980
gtggggttgct	cacatcagaa	ggcaaagaga	tcaaaaataa	agacgagatc	ttggccctac	5040
taaaagccct	ctttctgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaagg	5100
gacacagcgc	cgaggctaga	ggcaaccgga	tggctgacca	agcggcccca	aaggcagcca	5160
tcacagagac	tccagacacc	tctaccctcc	tcatagaaaa	ttcatcacc	tacacctcag	5220
aacattttca	ttacacagtg	actgatataa	aggacctaac	caagttgggg	gccatttatg	5280
ataaaacaaa	gaagtattgg	gtctaccaag	gaaaacctgt	gatgcctgac	cagtttactt	5340
ttgaattatt	agactttctt	catcagctga	ctcacctcag	cttctcaaaa	atgaaggctc	5400
tcctagagag	aagccacagt	ccctactaca	tgctgaaccg	ggatcgaaca	ctcaaaaata	5460
tcactgagac	ctgcaaagct	tgtgcacaag	tcaacgccag	caagtctgcc	gttaaacagg	5520
gaactagggg	ccgcgggcat	cggcccggca	ctcattggga	gatcgatttc	accgagataa	5580
agcccggatt	gtatggctat	aaatatcttc	tagtttttat	agataccttt	tctggctgga	5640
tagaagcctt	cccaaccaag	aaagaaaccg	ccaaggtcgt	aaccaagaag	ctactagagg	5700
agatcttccc	caggttcggc	atgcctcagg	tattgggaac	tgacaatggg	cctgccttcg	5760
tctccaaggt	gagtcagaca	gtggccgac	tgttggggat	tgattggaaa	ttacattgtg	5820
catacagacc	ccaaagctca	ggccaggtag	aaagaatgaa	tagaaccatc	aaggagactt	5880
taactaaatt	aacgcttgca	actggctcta	gagactgggt	gctcctactc	cccttagccc	5940
tgtaccgagc	ccgcaacacg	ccgggcccc	atggcctcac	cccatatgag	atcttatatg	6000
gggcaccccc	gccccttgta	aacttccttg	accctgacat	gacaagagtt	actaacagcc	6060
cctctctcca	agctcactta	caggctctct	acttagtcca	gcacgaagtc	tggagacctc	6120
tggcggcgagc	ctaccaagaa	caactggacc	gaccggtggt	acctcaccct	taccgagtcg	6180
gcgacacagt	gtgggtccgc	cgacaccaga	ctaagaacct	agaacctcgc	tggaaaggac	6240
cttacacagt	cctgctgacc	acccccaccg	ccctcaaagt	agacggcatc	gcagcttgga	6300
tacacgccgc	ccacgtgaag	gctgccgacc	ccgggggtgg	accatcctct	agactgacat	6360
ggcgcgttca	acgctctcaa	aacccctca	agataagatt	aaccgtgga	agcccttaat	6420
agtcatggga	gtcctgttag	gagtagggat	ggcagagagc	ccccatcagg	tctttaatgt	6480
aacctggaga	gtcaccaacc	tgatgactgg	gcgtaccgcc	aatgccacct	ccctcctggg	6540

ES 2 706 899 T3

aactgtacaa gatgccttcc caaaattata ttttgatcta tgtgatctgg tcggagagga	6600
gtgggaccct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcagggag	6660
acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcgggggtg	6720
tgggggacca ggagaggggt actgtggtaa atgggggtgt gaaaccaccg gacaggctta	6780
ctggaagccc acatcatcgt gggacctaata ctcccttaag cgcgtaaca cccctggga	6840
cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaattc	6900
cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tcaactgatgc	6960
aggaaaaaag gctaactggg acgggcccac atcgtgggga ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt	7080
ccccataggg cccaaccagc tattaccgca ccaagactc ccttcctcac caatagagat	7140
tgtaccggct ccacagccac ctagccccct caataccagt taccgccctt ccactaccag	7200
tacaccctca acctccccta caagtccaag tgtccacag ccaccccag gaactggaga	7260
tagactacta gctctagtca aaggagccta tcaggcgctt aacctacca atcccgacaa	7320
gaccaagaa tgttggtgtg gcttagtgtc gggacctcct tattacgaag gagtagcggc	7380
cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca cttcccaaca	7440
taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac	7500
tcaccaggcc ttatgtaaca ccacccaaag cgcgggtca ggatcctact accttgagc	7560
acccgccgga acaatgtggg cttgcagcac tggattgact cctgcttgt ccaccacggc	7620
gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca	7680
ctccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt	7740
atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattg cagctggaat	7800
agggacgggg accactgcct taattaaaac ccagcagttt gagcagcttc atgccgctat	7860
ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgacctc	7920
gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgctattcc taaaggaggg	7980
aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt	8040
gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac	8100
aggccaagga tggttcgaag ggctgtttaa tagatcccc tggtttacca ccttaatctc	8160
caccatcatg ggacctctaa tagtactctt actgatctta ctctttggac cttgcattct	8220
caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac	8280
tcagcaatat caccagctaa aaccataga gtacgagcca tgaacgcgtt actggccgaa	8340

ES 2 706 899 T3

gccgcttgga ataaggccgg tgtgcgtttg tctatatgtt attttccacc atattgccgt	8400
cttttgga tgtgagggcc cggaacctg gccctgtctt cttgacgagc attcctaggg	8460
gtctttcccc tctcgccaaa ggaatgcaag gtctgttgaa tgcgtgaag gaagcagttc	8520
ctctggaagc ttcttgaaga caaacaacgt ctgtagcgac cctttgcagg cagcggaacc	8580
ccccacctgg cgacaggtgc ctctgcggcc aaaagccacg tgtataagat acacctgcaa	8640
aggcggcaca accccagtgc cacgttgtga gttggatagt tgtggaaaga gtcaaatggc	8700
tctcctcaag cgtattcaac aaggggctga aggatgccca gaaggtaccc cattgtatgg	8760
gatctgatct ggggcctcgg tgcacatgct ttacatgtgt ttagtcgagg ttaaaaaaac	8820
gtctaggccc ccgaaccac ggggacgtgg ttttcctttg aaaaacacga ttataaatgg	8880
tgacaggggg aatggcaagc aagtgggatc agaagggat ggacattgcc tatgaggagg	8940
cggccttagg ttacaaagag ggtggtgttc ctattggcgg atgtcttata aataacaaag	9000
acggaagtgt tctcggtcgt ggtcacaaca tgagatttca aaagggatcc gccacactac	9060
atggtgagat ctccactttg gaaaactgtg ggagattaga gggcaaagtg tacaagata	9120
ccactttgta tacgacgctg tctccatgcg acatgtgtac aggtgccatc atcatgtatg	9180
gtattccacg ctgtgttgtc ggtgagaacg ttaatttcaa aagtaagggc gagaaatatt	9240
tacaaactag aggtcacgag gttgttgttg ttgacgatga gaggtgtaaa aagatcatga	9300
aacaatttat cgatgaaaga cctcaggatt ggtttgaaga tattggtgag taggcggccg	9360
cagataaaat aaaagatttt atttagtctc cagaaaaagg ggggaatgaa agaccccacc	9420
tgtaggtttg gcaagctagc ttaagtaacg ccattttgca aggcattggaa aaatacataa	9480
ctgagaatag agaagttcag atcaaggtca ggaacagatg gaacagctga atatgggcca	9540
aacaggatat ctgtggtaag cagttcctgc cccggctcag ggccaagaac agatggaaca	9600
gctgaatatg ggccaaacag gatattctgtg gtaagcagtt cctgccccgg ctgaggcca	9660
agaacagatg gtccccagat gcggtccagc cctcagcagt ttctagagaa ccatcagatg	9720
tttccagggt gccccaaagga cctgaaatga ccctgtgcct tatttgaact aaccaatcag	9780
ttcgcttctc gcttctgttc gcgcgttct gctccccgag ctcaataaaa gagcccacaa	9840
cccctcactc ggggcgccag tcctccgatt gactgagtcg cccgggtacc cgtgtatcca	9900
ataaacctc ttgcagttgc atccgacttg tggctctgct gttccttggg agggctctct	9960
ctgagtgatt gactaccctg cagcgggggt ctttcattac atgtgagcaa aaggccagca	10020
aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgcccccc	10080
tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaaccga caggactata	10140
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc	10200

ES 2 706 899 T3

gcttaccgga tacctgtccg cttttctccc ttcgggaagc gtggcgcttt ctcaatgctc 10260
acgctgtagg tatctcagtt cgggtgtaggt cgttcgctcc aagctgggct gtgtgcacga 10320
acccccggtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc 10380
ggtaagacac gacttatcgc cactggcagc agccactggg aacaggatta gcagagcgag 10440
gtatgtaggc ggtgctacag agttcttgaa gtgggtggcct aactacggct acactagaag 10500
gacagtatth ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 10560
ctcttgatcc ggcaaaaaa ccaccgctgg tagcgggtgg ttttttgtht gcaagcagca 10620
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggctctga 10680
cgctcagtg aacgaaaaact cacgttaagg gattttggtc atgagattat caaaaaggat 10740
cttcacctag atccttttaa attaaaaatg aagtttttaa tcaatctaaa gtatatatga 10800
gtaaaacttg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 10860
tctatttctg tcatccatag ttgcctgact ccccgctcgt tagataacta cgatacggga 10920
gggcttacca tctggcccca gtgctgcaat gataccgcga gaccacgct caccggctcc 10980
agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg gtccgtcaac 11040
tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc 11100
agttaatagt ttgcgcaacg ttgttgccat tgctgcaggc atcgtgggtg cagcctcgtc 11160
gtttggtatg gcttcattca gctccggtc ccaacgatca aggcgagtta catgatcccc 11220
catgttgtgc aaaaaagcgg ttagctcctt cggctcctcc atcgttgtca gaagtaagtt 11280
ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc 11340
atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg 11400
tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataacc cgccacatag 11460
cagaacttta aaagtgtca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat 11520
cttaccgctg ttgagatcca gtctgatgta acccactcgt gcacccaact gatcttcagc 11580
atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa 11640
aaaggaata agggcgacac ggaaatgttg aatactcata ctcttctttt ttcaatatta 11700
ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa 11760
aaataaacia ataggggttc cgcgcacatt tccccgaaa gtgccacctg acgtctaaga 11820
aaccattatt atcatgacat taacctataa aaataggcgt atcacgagge cttttcgtct 11880
tcaagaattc at 11892

<210> 21
<211> 12007
<212> ADN
<213> Secuencia artificial

ES 2 706 899 T3

<220>

<223> Vector RCR - pACE-CD

<400> 21

5

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg	60
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt	120
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca	180
atgggtggag tatttacggg aaactgccc cttggcagta catcaagtgt atcatatgcc	240
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	300
catgacctta tgggactttc ctacttgga gtacatctac gtattagtea tgcctattac	360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	420
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480
ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggag gtaggcgtgt	540
acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg	600
actgagtcgc cggggtaccc gtgtatccaa taaacctct tgcagttgca tccgacttgt	660
ggtctcgctg ttccctggga ggtctctc tgagtgttg actaccgctc agcgggggtc	720
tttcatttg gggctcgctc gggatcggga gaccctgcc cagggaccac cgaccacca	780
ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac	840
tgattttatg cgctgcgctc ggtactagtt agctaactag ctctgtatct ggcggacccg	900
tgggtggaact gacgagttcg gaacaccccg ccgcaaccct gggagacgct ccagggactt	960
cggggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg	1020
gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagAAC ctaaaacagt	1080
tcccgctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgctcttg	1140
tctgctgcag catcgttctg tgtgtctct gtctgactgt gtttctgtat ttgtctgaga	1200
atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg	1260
agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct	1320
ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc	1380
tcatcaccca ggttaagatc aaggctttt cacctggccc gcatggacac ccagaccagg	1440
tcccctacat cgtgacctgg gaagccttgg cttttgaccc cctccctgg gtcaagccct	1500
ttgtacaccc taagcctccg cctcctcttc ctccatccgc cccgtctctc ccccttgaac	1560
ctcctcgttc gacccgcct cgatcctccc tttatccagc cctcactcct tctctaggcg	1620

ES 2 706 899 T3

ccaaacctaa acctcaagtt ctttctgaca gtgggggggccc gctcatcgac ctacttacag	1680
aagaccccc gccttatagg gacccaagac caccaccttc cgacagggac ggaaatggtg	1740
gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cgcctacgtg	1800
ggagacggga gccccctgtg gccgactcca ctacctcgca ggcatctccc ctccgcgcag	1860
gaggaaacgg acagcttcaa tactggccgt tctcctcttc tgacctttac aactggaaaa	1920
ataataaccc ttctttttct gaagatccag gtaaactgac agctctgatc gagtctgttc	1980
tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg	2040
gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc	2100
gccccactca actgccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg	2160
attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagttg ctccatagcgg	2220
gtctccaaaa cgcgggcaga agccccacca atttgcccaa ggtaaaagga ataacacaag	2280
ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggatca	2340
ctccttatga ccctgaggac ccagggaag aaactaatgt gtctatgtct ttcatttggc	2400
agtctgcccc agacattggg agaaagttag agaggttaga agatttaaaa aacaagacgc	2460
ttggagattt ggtagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa	2520
gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg	2580
agcagaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca	2640
ctgtcgtttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg	2700
atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtcccaaga	2760
aaccacgagg acctcgggga ccaagacccc agacctccct cctgacccta gatgactagg	2820
gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac	2880
ccgtcacctt cctggtagat actggggccc aacactccgt gctgacccaa aatcctggac	2940
ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgctgga	3000
ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactcttct ctccatgtac	3060
cagactgtcc ctatcctctg ttaggaagag atttgctgac taaactaaaa gcccaaatec	3120
actttgaggg atcaggagcc cagggttatgg gaccaatggg gcagcccctg caagtgttga	3180
ccctaaatat agaagatgag catcggtac atgagacctc aaaagagcca gatgtttctc	3240
taggggtccac atggctgtct gattttctct aggcctgggc ggaaaccggg ggcatgggac	3300
tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca	3360
taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac	3420

ES 2 706 899 T3

tgttggacca	gggaataactg	gtaccctgcc	agtccccctg	gaacacgccc	ctgctacccg	3480
ttaagaaacc	agggactaat	gattataggc	ctgtccagga	tctgagagaa	gtcaacaagc	3540
gggtggaaga	catccacccc	accgtgccca	acccttacia	cctcttgagc	gggctccac	3600
cgtcccacca	gtggtacact	gtgcttgatt	taaaggatgc	ctttttctgc	ctgagactcc	3660
accccaccag	tcagcctctc	ttcgctttg	agtggagaga	tccagagatg	ggaatctcag	3720
gacaattgac	ctggaccaga	ctcccacagg	gtttcaaaaa	cagtcccacc	ctgtttgatg	3780
aggcactgca	cagagaccta	gcagacttcc	ggatccagca	cccagacttg	atcctgctac	3840
agtacgtgga	tgacttactg	ctggccgcca	cttctgagct	agactgcca	caaggtactc	3900
gggccctgtt	acaaacccta	gggaacctcg	ggtatcgggc	ctcggccaag	aaagcccaaa	3960
tttgccagaa	acaggtcaag	tatctgggg	atcttctaaa	agagggtcag	agatggctga	4020
ctgaggccag	aaaagagact	gtgatggggc	agcctactcc	gaagaccct	cgacaactaa	4080
gggagttcct	agggacggca	ggcttctgtc	gcctctggat	ccctgggttt	gcagaaatgg	4140
cagccccctt	gtaccctctc	acaaaaacgg	ggactctgtt	taattggggc	ccagaccaac	4200
aaaaggccta	tcaagaaatc	aagcaagctc	ttctaactgc	cccagccctg	gggttgccag	4260
atttgactaa	gccctttgaa	ctctttgtcg	acgagaagca	gggctacgcc	aaaggtgtcc	4320
taacgcaaaa	actgggacct	tggcgtcggc	cggtggccta	cctgtccaaa	aagctagacc	4380
cagtagcagc	tgggtggccc	ccttgccctac	ggatggtagc	agccattgcc	gtactgacaa	4440
aggatgcagg	caagctaacc	atgggacagc	cactagtcac	tctggccccc	catgcagtag	4500
aggcactagt	caaacacccc	cccgaccgct	ggctttccaa	cggccggatg	actcactatc	4560
aggccttgct	tttggaacag	gaccgggtcc	agttcggacc	ggtggtagcc	ctgaaccggg	4620
ctacgtgct	cccactgcct	gaggaagggc	tgcaacacaa	ctgccttgat	atcctggccg	4680
aagcccacgg	aaccgacccc	gacctaacgg	accagccgct	cccagacgcc	gaccacacct	4740
ggtacacgga	tggaagcagt	ctcttacaag	agggacagcg	taaggcggga	gctgcgggtga	4800
ccaccgagac	cgaggtaatc	tgggctaaag	ccctgccagc	cgggacatcc	gctcagcggg	4860
ctgaactgat	agcactcacc	caggccctaa	agatggcaga	aggtaagaag	ctaaatgttt	4920
atactgatag	cgttatgct	tttgctactg	cccatatcca	tggagaaata	tacagaaggc	4980
gtgggttgct	cacatcagaa	ggcaaagaga	tcaaaaataa	agacgagatc	ttggccctac	5040
taaaagccct	ctttctgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaagg	5100
gacacagcgc	cgaggctaga	ggcaaccgga	tggctgacca	agcggcccca	aaggcagcca	5160
tcacagagac	tccagacacc	tctaccctcc	tcatagaaaa	ttcatcacc	tacacctcag	5220
aacattttca	ttacacagtg	actgatataa	aggacctaac	caagttgggg	gccatttatg	5280

ES 2 706 899 T3

ataaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt	5340
ttgaattatt agactttctt catcagctga ctcacctcag cttctcaaaa atgaaggctc	5400
tcctagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata	5460
tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg	5520
gaactagggg cgcggggcat cggcccgga ctcattggga gatcgatttc accgagataa	5580
agcccggtt gtatggctat aaatatcttc tagtttttat agataccttt tctggctgga	5640
tagaagcctt cccaaccaag aaagaaaccg ccaaggctgt aaccaagaag ctactagagg	5700
agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg	5760
tctccaaggt gagtcagaca gtggccgacg tgttggggat tgattggaaa ttacattgtg	5820
catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaaccatc aaggagactt	5880
taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc	5940
tgtaccgagc ccgcaacacg ccggggcccc atggcctcac cccatatgag atcttatatg	6000
gggcaccccc gcccttgta aacttcctg accctgacat gacaagagtt actaacagcc	6060
cctctctcca agctactta caggctctct acttagtcca gcacgaagtc tggagacctc	6120
tggcggcagc ctaccaagaa caactggacc gaccgggtgt acctaccct taccgagtcg	6180
gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcg tggaaaggac	6240
cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttgga	6300
tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat	6360
ggcgcggttca acgctctcaa aacccccctca agataagatt aaccctgga agcccttaat	6420
agtcattgga gtcctgttag gagtagggat ggcagagagc ccccatcagg tctttaatgt	6480
aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct ccctcctggg	6540
aactgtacaa gatgccttcc caaaattata ttttgatcta tgtgatctgg tcggagagga	6600
gtgggaccct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcagggag	6660
acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcgggggtg	6720
tgggggacca ggagagggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta	6780
ctggaagccc acatcatcgt gggacctaata ctccttaag cgcggtaaca cccctggga	6840
cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaattc	6900
cttccaaggg gctactcgag ggggcagatg caaccctcta gtcctagaat tcaactgatgc	6960
aggaaaaaag gctaactggg acgggcccac atcgtgggga ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt	7080

ES 2 706 899 T3

cccataggg	cccaaccag	tattaccga	caaagaactc	ccttcctcac	caatagagat	7140
tgtaccggct	ccacagccac	ctagccccct	caataccagt	tacccccctt	ccactaccag	7200
tacaccctca	acctccccta	caagtccaag	tgtcccacag	ccacccccag	gaactggaga	7260
tagactacta	gctctagtca	aaggagccta	tcaggcgctt	aacctcacca	atccccgaaa	7320
gaccaagaa	tggtggctgt	gcttagtgtc	gggacctcct	tattacgaag	gagtagcggt	7380
cgtgggcact	tataccaatc	attccaccgc	tccggccaac	tgtacggcca	cttcccaca	7440
taagcttacc	ctatctgaag	tgacaggaca	gggcctatgc	atgggggcag	tacctaaaac	7500
tcaccaggcc	ttatgtaaca	ccacccaaag	cgccggctca	ggatcctact	accttgacgc	7560
acccgccgga	acaatgtggg	cttgacgac	tggattgact	ccctgcttgt	ccaccacggt	7620
gctcaatcta	accacagatt	attgtgtatt	agttgaactc	tggcccagag	taattttacca	7680
ctccccgat	tatatgtatg	gtcagcttga	acagcgtacc	aaatataaaa	gagagccagt	7740
atcattgacc	ctggcccttc	tactaggagg	attaaccatg	ggagggattg	cagctggaat	7800
agggacgggg	accactgcct	taattaaaac	ccagcagttt	gagcagcttc	atgccgctat	7860
ccagacagac	ctcaacgaag	tcgaaaagtc	aattaccaac	ctagaaaagt	cactgacctc	7920
gttgtctgaa	gtagtcctac	agaaccgcag	aggcctagat	ttgctattcc	taaaggaggg	7980
aggtctctgc	gcagccctaa	aagaagaatg	ttgtttttat	gcagaccaca	cggggcctagt	8040
gagagacagc	atggccaaat	taagagaaag	gcttaatcag	agacaaaaac	tatttgagac	8100
aggccaagga	tggttcgaag	ggctgtttta	tagatcccc	tggtttacca	ccttaatctc	8160
caccatcatg	ggacctctaa	tagtactctt	actgatctta	ctctttggac	cttgcaattct	8220
caatcgatta	gtccaatttg	ttaaagacag	gatatcagtg	gtccaggctc	tagttttgac	8280
tcaacaatat	caccagctga	agcctataga	gtacgagcca	tgacgtacgt	tactggccga	8340
agccgcttg	aataaggccg	gtgtgcgttt	gtctatatgt	tattttccac	catattgccg	8400
tcttttgga	atgtgagggc	cgggaaacct	ggccctgtct	tcttgacgag	cattcctagg	8460
ggtctttccc	ctctcgccaa	aggaatgcaa	ggtctgttga	atgtcgtgaa	ggaagcagtt	8520
cctctggaag	cttcttgaag	acaaacaacg	tctgtagcga	ccctttgcag	gcagcggaac	8580
ccccacctg	gcgacaggtg	cctctgcggc	caaagccac	gtgtataaga	tacacctgca	8640
aaggcggcac	aacccacagt	ccacgttgtg	agttggatag	ttgtggaaag	agtcaaattg	8700
ctctcctcaa	gcgtattcaa	caaggggctg	aaggatgccc	agaaggtacc	ccattgtatg	8760
ggatctgatc	tggggcctcg	gtgcacatgc	tttacctgtg	tttagtcgag	gttaaaaaaa	8820
cgtctaggcc	ccccgaacca	cggggacgtg	gttttccttt	gaaaaacacg	ataataccat	8880
ggtgacaggg	ggaatggcaa	gcaagtggga	tcagaagggg	atggacattg	cctatgagga	8940

ES 2 706 899 T3

ggcggcctta ggttacaaag aggggtggtgt tcctattggc ggatgtctta tcaataacaa	9000
agacggaagt gttctcggtc gtggtcacaa catgagattt caaaagggat ccgccacact	9060
acatggtgag atctccactt tggaaaactg tgggagatta gagggcaaag tgtacaaaga	9120
taccactttg tatacgacgc tgtctccatg cgacatgtgt acagggtgcca tcatcatgta	9180
tggatattcca cgctgtgttg tcggtgagaa cgtaatttc aaaagtaagg gcgagaaata	9240
tttacaaact agaggtcacg aggttggtgt tggtgacgat gagagggtgta aaaagatcat	9300
gaaacaattt atcgatgaaa gacctcagga ttgggttgaa gatattgggtg agtaggcggc	9360
cgcgccatag ataaaataaa agattttatt tagtctccag aaaaaggggg gaatgaaaga	9420
ccccacctgt aggtttggca agctagctta agtaacgcca ttttgcaagg catggaaaaa	9480
tacataactg agaatagaga agttcagatc aaggtcagga acagatggaa cagctgaata	9540
tgggccaac aggatatctg tggtaagcag ttcttgcccc ggctcagggc caagaacaga	9600
tggaacagct gaatatgggc caaacaggat atctgtggta agcagttcct gccccggctc	9660
agggccaaga acagatggtc ccagatgcg gtccagccct cagcagtttc tagagaacca	9720
tcagatgttt ccagggtgcc ccaaggacct gaaatgacct tgtgccttgt ttaactaac	9780
caatcagttc gcttctcgct tctgttcgcg cgcttctgct ccccgagctc aataaaagag	9840
cccacaaccc ctcaactcggg gcgccagtcc tcgattgac tgagtcgccc gggtagccgt	9900
gtatccaata aaccctcttg cagttgcac cagcttggtg tctcgctgtt ccttgggagg	9960
gtctcctctg agtgattgac taccgctcag cgggggtctt tcatttgggg gctcgtccgg	10020
gatcgggaga cccctgcca gggaccaccg acccaccacc gggaggtaag ctggctgcct	10080
cgcgctttc ggtgatgacg gtgaaaacct ctgacatgtg agcaaaaggc cagcaaaagg	10140
ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg	10200
agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat	10260
accaggcggt tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta	10320
ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcaa tgctcacgct	10380
gtaggtatct cagttcgggtg taggtcgttc gtcceaagct gggctgtgtg cacgaacccc	10440
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa	10500
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtagt	10560
taggcggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact agaaggacag	10620
tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt	10680
gatccggcaa acaaaccacc gctggtagcg gtgggttttt tgtttgcaag cagcagatta	10740

ES 2 706 899 T3

```

cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 10800
agtggaacga aaactcacgt taagggatct tggatcatgag attatcaaaa aggatcttca 10860
cctagatcct tttaaattaa aaatgaagtt ttaaataaat cttaaagtata tatgagtaaa 10920
cttgggtctga cagttaccaa tgcttaataca gtgaggcacc tatctcagcg atctgtctat 10980
ttcggttcac catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct 11040
taccatctgg cccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt 11100
tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggctct gcaactttat 11160
ccgcctccat ccagtctatt aattgttgcc ggaagctag agtaagtagt tcgccagtta 11220
atagtttgcg caacgttggt gccattgctg caggcatcgt ggtgtcacgc tcgtcgtttg 11280
gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt 11340
tgtgcaaaaa agcggtagc tccttcgggc ctccgatcgt tgtcagaagt aagtggccg 11400
cagtgttacc actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg 11460
taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc 11520
ggcgaccgag ttgctcttgc ccggcgtaa cacgggataa taccgcgcca catagcagaa 11580
ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac 11640
cgctgttgag atccagttcg atgtaaccca ctctgcacc caactgatct tcagcatctt 11700
ttactttcac cagcgtttct gggtagcaaa aaacaggaag gcaaaatgcc gcaaaaagg 11760
gaataagggc gacacgaaa tgttgaatac tcatactctt cctttttcaa tattattgaa 11820
gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 11880
aacaatatag ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca 11940
ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtcttcaag 12000
aatcat 12007

```

<210> 22
 <211> 11893
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Vector RCR - pAC3-yCD2

<400> 22

```

tagttattaa tagtaatcaa ttacgggggc attagttcat agcccatata tggagttccg 60
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgccatt 120
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 180
atgggtggag tatttacggt aaactgcca cttggcagta catcaagtg atcatatgcc 240

```

ES 2 706 899 T3

aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	300
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	360
catggtgatg cggtttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	420
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480
ggactttcca aaatgtcgtg acaactccgc cccattgacg caaatgggcg gtaggcgtgt	540
acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg	600
actgagtcgc ccgggtaccc gtgtatccaa taaacctctt tgcagttgca tccgacttgt	660
ggtctcgctg ttccttggga gggctcctc tgagtgattg actaccctgc agcgggggtc	720
tttcatttgg gggctcgtcc gggatcggga gaccctgcc cagggaccac cgaccacca	780
ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac	840
tgattttatg cgctgcgtc ggtactagtt agctaactag ctctgtatct ggcggaccgc	900
tgggtggaact gacgagttcg gaacacccgg ccgcaaccct gggagacgtc ccagggactt	960
cggggggcgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg	1020
gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt	1080
tccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg	1140
tctgctgcag catcgttctg tgttgtctct gtctgactgt gtttctgtat ttgtctgaaa	1200
atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg	1260
agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct	1320
ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcaccttt aaccgagacc	1380
tcatcaccca ggttaagatc aaggctcttt cacctggccc gcatggacac ccagaccagg	1440
tcccctacat cgtgacctgg gaagccttgg cttttgacct cctccctgg gtcaagcct	1500
ttgtacacct taagcctccg cctcctcttc ctccatccgc cccgtctctc ccccttgaac	1560
ctcctcgttc gacccgcct cgatcctccc tttatccagc cctcactcct tctctaggcg	1620
ccaaacctaa acctcaagtt ctttctgaca gtggggggcc gctcatcgac ctacttacag	1680
aagacccccc gccttatagg gacccaagac caccctcttc cgacagggac ggaaatggtg	1740
gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cgcctacgtg	1800
ggagacggga gccccctgtg gccgactcca ctacctcgca ggcattcccc ctccgcgcag	1860
gaggaaacgg acagcttcaa tactggccgt tctcctcttc tgacctttac aactggaaaa	1920
ataataaccc ttctttttct gaagatccag gtaaactgac agctctgatc gagtctgtcc	1980
tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg	2040

ES 2 706 899 T3

gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatgggc	2100
gccccactca actgccaat gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg	2160
attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagttg ctccctagcgg	2220
gtctccaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag	2280
ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaggtaca	2340
ctccttatga ccctgaggac ccagggcaag aaactaatgt gtctatgtct ttcatttggc	2400
agtctgcccc agacattggg agaaagttag agagggttaga agatttaaaa aacaagacgc	2460
ttggagattt ggtagagag gcagaaaaga tctttaataa acgagaaacc ccggaagaaa	2520
gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccgtagg acagaggatg	2580
agcagaaaaga gaaagaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca	2640
ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg	2700
atcgcgacca gtgtgcctac tgcaaagaaa aggggactg ggctaaagat tgtccaaga	2760
aaccacgagg acctcgggga ccaagacccc agacctcct cctgacccta gatgactagg	2820
gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac	2880
ccgtcacctt cctggtagat actggggccc aacctccgt gctgaccaa aatcctggac	2940
ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgtgga	3000
ccacggatcg caaagtacat ctagctaccg gtaaggtcac ccactcttcc ctccatgtac	3060
cagactgtcc ctatcctctg ttaggaagag atttgtgac taaactaaaa gcccaaatcc	3120
actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga	3180
ccctaaatat agaagatgag tatcggctac atgagacctc aaaagagcca gatgtttctc	3240
taggggccac atggctgtct gattttcctc aggcctgggc ggaaaccggg ggcattggac	3300
tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc ccctgtcca	3360
taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac	3420
tgttggaaca gggaatactg gtaccctgcc agtccccctg gaacacgccc ctgctacccg	3480
ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc	3540
gggtggaaga catccacccc accgtgccca acccttaca cctcttgagc gggctccac	3600
cgtcccacca gtggtacact gtgcttgatt taaaggatgc ctttttctgc ctgagactcc	3660
accccaccag tcagcctctc ttgcctttg agtggagaga tccagagatg ggaatctcag	3720
gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg	3780
aggcactgca cagagacctc gcagacttcc ggatccagca ccagacttg atcctgctac	3840
agtacgtgga tgacttactg ctggccgcca cttctgagct agactgcca caaggtactc	3900

ES 2 706 899 T3

gggcctgtt	acaaacccta	gggaacctcg	ggtatcgggc	ctcggccaag	aaagcccaaa	3960
tttgccagaa	acaggtcaag	tatctgggggt	atcttctaaa	agagggtcag	agatggctga	4020
ctgaggccag	aaaagagact	gtgatggggc	agcctactcc	gaagaccctc	cgacaactaa	4080
gggagttcct	agggacggca	ggcttctgtc	gcctctggat	ccctgggttt	gcagaaatgg	4140
cagccccctt	gtaccctctc	acaaaaacgg	ggactctgtt	taattggggc	ccagaccaac	4200
aaaaggccta	tcaagaaatc	aagcaagctc	ttctaactgc	cccagccctg	gggttgccag	4260
atttgactaa	gccctttgaa	ctctttgtcg	acgagaagca	gggctacgcc	aaaggtgtcc	4320
taacgcaaaa	actgggacct	tggcgtcggc	cggtggccta	cctgtccaaa	aagctagacc	4380
cagtagcagc	tgggtggccc	ccttgccctac	ggatggtagc	agccattgcc	gtactgacaa	4440
aggatgcagg	caagctaacc	atgggacagc	cactagtcat	tctggccccc	catgcagtag	4500
aggcactagt	caaacaacct	cccagaccgt	ggctttccaa	cgcccggatg	actcactatc	4560
aggccttgct	tttggaacag	gaccgggtcc	agttcggacc	ggtggtagcc	ctgaaccggy	4620
ctacgtgct	cccactgcct	gaggaagggc	tgcaacacaa	ctgccttgat	atcctggccg	4680
aagcccacgg	aacccgacct	gacctaacgg	accagccgct	cccagacgcc	gaccacacct	4740
ggtacacgga	tgaagcagt	ctcttacaag	agggacagcg	taaggcgga	gctgcggtga	4800
ccaccgagac	cgaggtaatc	tgggctaaag	ccctgccagc	cgggacatcc	gctcagcggg	4860
ctgaactgat	agcactcacc	caggccctaa	agatggcaga	aggtaagaag	ctaaatgttt	4920
atactgatag	ccgttatgct	tttgctactg	cccatatcca	tggagaaata	tacagaaggc	4980
gtgggttgct	cacatcagaa	ggcaaagaga	tcaaaaataa	agacgagatc	ttggccctac	5040
taaaagccct	ctttctgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaagg	5100
gacacagcgc	cgaggctaga	ggcaaccgga	tggtgacca	agcggcccga	aaggcagcca	5160
tcacagagac	tccagacacc	tctaccctcc	tcatagaaaa	ttcatcacc	tacacctcag	5220
aacattttca	ttacacagtg	actgatataa	aggacctaac	caagttgggg	gccattttatg	5280
ataaaacaaa	gaagtattgg	gtctaccaag	gaaaacctgt	gatgcctgac	cagtttactt	5340
ttgaattatt	agactttctt	catcagctga	ctcacctcag	cttctcaaaa	atgaaggctc	5400
tcctagagag	aagccacagt	ccctactaca	tgctgaaccg	ggatcgaaca	ctcaaaaata	5460
tcactgagac	ctgcaaagct	tgtgcacaag	tcaacgccag	caagtctgcc	gttaaacagg	5520
gaactagggg	ccgcgggcat	cggcccggca	ctcattggga	gatcgatttc	accgagataa	5580
agcccggatt	gtatggctat	aaatatcttc	tagtttttat	agataccttt	tctggctgga	5640
tagaagcctt	cccaaccaag	aaagaaaccg	ccaaggctcg	aaccaagaag	ctactagagg	5700

ES 2 706 899 T3

agatcttccc caggttcggc atgcctcagg tattgggaac tgacaatggg cctgccttcg	5760
tctccaaggt gagtcagaca gtggccgacg tgttggggat tgattggaaa ttacattgtg	5820
catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaaccatc aaggagactt	5880
taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc	5940
tgtaccgagc ccgcaacacg ccggggccccc atggcctcac cccatatgag atcttatatg	6000
gggcaccccc gcccttgta aacttccctg accctgacat gacaagagtt actaacagcc	6060
cctctctcca agctcaacta caggctctct acttagtcca gcacgaagtc tggagacctc	6120
tggcggcagc ctaccaagaa caactggacc gaccggtggt acctcaccct taccgagtcg	6180
gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac	6240
cttacacagt cctgctgacc acccccaccg ccctcaaagt agacggcatc gcagcttgga	6300
tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat	6360
ggcgcgttca acgctctcaa aacccccctc agataagatt aacccttgga agcccttaat	6420
agtcattgga gtccctgttag gagtagggat ggcagagagc ccccatcagg tctttaatgt	6480
aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct cctcctggg	6540
aactgtacaa gatgccttcc caaaattata ttttgatcta tgtgatctgg tcggagagga	6600
gtgggacctc tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcaggag	6660
acagcggacc cggacttttg acttttacgt gtgccctggg cataccgtaa agtcgggggtg	6720
tgggggacca ggagagggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta	6780
ctggaagccc acatcatcgt gggacctaact ctcccttaag cgcgtaaca cccctggga	6840
cacgggatgc tctaaagtgt cctgtggccc ctgctacgac ctctccaaag tatccaattc	6900
cttccaaggg gctactcgag ggggcagatg caaccctcta gtccctagaat tcaactgatgc	6960
aggaaaaaag gctaactggg acgggcccac atcgtgggga ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt	7080
cccataggg cccaaccag tattaccga ccaaagactc ccttcctcac caatagagat	7140
tgtaccggct ccacagccac ctagccccct caataccagt taccctccct ccactaccag	7200
tacaccctca acctccccta caagtccaag tgtccacag ccaccccccag gaactggaga	7260
tagactacta gctctagtca aaggagccta tcaggcgctt aacctacca atcccgacaa	7320
gaccaagaa tgttggtgtg gcttagtgtc gggacctcct tattacgaag gagtagcgg	7380
cgtgggcact tataccaatc attccaccgc tccggccaac tgtacggcca ctcccaaca	7440
taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac	7500
tcaccaggcc ttatgtaaca ccacccaaag cgccggctca ggatcctact accttgacg	7560

ES 2 706 899 T3

acccgccgga	acaatgtggg	cttgacgac	tggttgact	ccctgcttgt	ccaccacggt	7620
gctcaatcta	accacagatt	attgtgtatt	agttgaactc	tggtccagag	taatttacca	7680
ctccccgat	tatatgtatg	gtcagcttga	acagcgtacc	aaatataaaa	gagagccagt	7740
atcattgacc	ctggcccttc	tactaggagg	attaaccatg	ggagggttg	cagctggaat	7800
agggacgggg	accactgcct	taattaaaac	ccagcagttt	gagcagcttc	atgccgctat	7860
ccagacagac	ctcaacgaag	tcgaaaagtc	aattaccaac	ctagaaaagt	cactgacctc	7920
gttgtctgaa	gtagtcctac	agaaccgcag	aggcctagat	ttgctattcc	taaaggaggg	7980
aggtctctgc	gcagccctaa	aagaagaatg	ttgtttttat	gcagaccaca	cggggctagt	8040
gagagacagc	atggccaaat	taagagaaag	gcttaatcag	agacaaaaac	tatttgagac	8100
aggccaagga	tggttcgaag	ggctgtttaa	tagatcccc	tggtttacca	ccttaatctc	8160
caccatcatg	ggacctctaa	tagtactctt	actgatctta	ctctttggac	cttgcatctt	8220
caatcgattg	gtccaatttg	ttaaagacag	gatctcagtg	gtccaggctc	tggttttgac	8280
tcagcaatat	caccagctaa	aacctataga	gtacgagcca	tgaacgcgtt	actggccgaa	8340
gccgcttga	ataaggccgg	tgtgcgtttg	tctatatgtt	attttccacc	atattgccgt	8400
cttttgcaa	tgtgagggcc	cggaaacctg	gccctgtctt	cttgacgagc	attcctaggg	8460
gtctttcccc	tctcgccaaa	ggaatgcaag	gtctgttgaa	tgtcgtgaag	gaagcagttc	8520
ctctggaagc	ttcttgaaga	caaacaacgt	ctgtagcgac	cctttgcagg	cagcggaacc	8580
ccccacctgg	cgacaggtgc	ctctgcggcc	aaaagccacg	tgtataagat	acacctgcaa	8640
aggcgcaca	acccagtg	cacgttgtga	gttgatagtg	tgtggaaaga	gtcaaatggc	8700
tctcctcaag	cgtattcaac	aaggggctga	aggatgccca	gaaggtaccc	cattgtatgg	8760
gatctgatct	ggggcctcgg	tgcacatgct	ttacatgtgt	ttagtcgagg	ttaaaaaac	8820
gtctaggccc	cccgaaccac	ggggacgtgg	ttttcctttg	aaaaacacga	ttataaatgg	8880
tgaccggcgg	catggcctcc	aagtgggatc	aaaagggcat	ggatatcgct	tacgaggagg	8940
ccctgctggg	ctacaaggag	ggcggcgtgc	ctatcggcgg	ctgtctgatc	aacaacaagg	9000
acggcagtg	gctgggcagg	ggccacaaca	tgagggtcca	gaagggtcc	gccaccctgc	9060
acggcgagat	ctccaccctg	gagaactgtg	gcagggtgga	gggcaagggtg	tacaaggaca	9120
ccaccctgta	caccaccctg	tccccttggtg	acatgtgtac	cggcgctatc	atcatgtacg	9180
gcatccctag	gtgtgtgatc	ggcgagaacg	tgaacttcaa	gtccaagggc	gagaagtacc	9240
tgcaaaccag	gggccacgag	gtggtggttg	ttgacgatga	gaggtgtaag	aagctgatga	9300
agcagttcat	cgacgagagg	cctcaggact	ggttcgagga	tatcggcgag	taagcggccg	9360

ES 2 706 899 T3

cagataaaat aaaagatttt atttagtctc cagaaaaagg ggggaatgaa agaccccacc	9420
tgtaggtttg gcaagctagc ttaagtaacg ccattttgca aggcattggaa aaatacataa	9480
ctgagaatag agaagttcag atcaagggtca ggaacagatg gaacagctga atatgggcca	9540
aacaggatat ctgtggtaag cagttcctgc cccggctcag ggccaagaac agatggaaca	9600
gctgaatatg ggccaaacag gatattctgtg gtaagcagtt cctgccccgg ctcagggcca	9660
agaacagatg gtccccagat gcggtccagc cctcagcagt ttctagagaa ccatcagatg	9720
tttccagggt gcccgaagga cctgaaatga ccctgtgcct tatttgaact aaccaatcag	9780
ttcgcttctc gcttctgttc gcgcgcttct gctccccgag ctcaataaaa gagcccacaa	9840
ccctcactc ggggcgccag tctccgatt gactgagtcg cccgggtacc cgtgtatcca	9900
ataaacctc ttgcagttgc atccgacttg tggctctcgt gtcccttggg agggctctct	9960
ctgagtgatt gactaccgt cagcgggggt ctttcattac atgtgagcaa aaggccagca	10020
aaaggccagg aaccgtaaaa aggcgcggtt gctggcgttt ttccataggc tccgcccccc	10080
tgacgagcat cacaaaaatc gacgctcaag tcagagggtg cgaaaccga caggactata	10140
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc	10200
gcttaccgga tacctgtccg cttttctccc ttcggaagc gtggcgcttt ctcatagctc	10260
acgctgtagg tatctcagtt cgggtgtagt cgttcgtcc aagctgggct gtgtgcacga	10320
acccccgtt cagcccagc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc	10380
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag	10440
gtatgtaggc ggtgctacag agttcttgaa gtgggtggcct aactacggct acactagaag	10500
gacagtattt ggtatctgcg ctctgtgaa gccagttacc ttcggaaaaa gagttggtag	10560
ctcttgatcc ggcaaaaaa ccaccgctgg tagcgggtgt tttttgttt gcaagcagca	10620
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga	10680
cgctcagtg aacgaaaact cacgttaagg gattttggct atgagattat caaaaaggat	10740
cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga	10800
gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg	10860
tctatttcgt tcatccatag ttgcctgact ccccgctcgt tagataacta cgatacggga	10920
gggcttacca tctggcccca gtgctgcaat gataccgcga gacccacgct caccggctcc	10980
agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcctgcaac	11040
tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc	11100
agttaatagt ttgcgcaacg ttgttgccat tgctgcaggc atcgtgggtg cacgctcgtc	11160
gtttggtagt gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc	11220

ES 2 706 899 T3

catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt	11280
ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc	11340
atccgtaaga tgcttttctg tgactgggga gtactcaacc aagtcattct gagaatagtg	11400
tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataccg cgccacatag	11460
cagaacttta aaagtgtca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat	11520
cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc	11580
atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa	11640
aaaggaata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta	11700
ttgaagcatt tatcaggggt attgtctcat gagcggatac atatttgaat gtatttagaa	11760
aaataaacia ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga	11820
aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtct	11880
tcaagaattc cat	11893

REIVINDICACIONES

1. Un vector que comprende un polinucleótido que comprende los pares de bases 8877 a 9353 de la SEQ ID NO: 19, que codifica un polipéptido que tiene actividad citosina desaminasa (CD).

2. El vector de la reivindicación 1, donde el polinucleótido comprende además un polinucleótido UPRT u OPRT que codifica un polipéptido que tiene actividad UPRT u OPRT unida operativamente al polipéptido que tiene actividad CD, preferentemente donde el polinucleótido UPRT u OPRT está optimizado por codones y / o separado del polinucleótido que codifica el polipéptido que tiene actividad CD por un enlazador.

3. Una célula huésped que contiene el vector de la reivindicación 1 o la reivindicación 2, donde la célula huésped es una célula humana, preferentemente una célula humana transformada de manera estable y / o donde la célula huésped es una célula que comprende un trastorno de proliferación celular, preferentemente glioblastoma multiforme.

4. El vector de la reivindicación 1 o la reivindicación 2, donde el vector es un vector de expresión y / o en donde el vector es un plásmido o un vector viral, preferentemente donde el vector viral es un retrovirus competente para la replicación.

5. El vector de la reivindicación 4, donde el retrovirus competente para la replicación (RCR) comprende:

una proteína GAG retroviral;

una proteína POL retroviral;

una cubierta retroviral;

un polinucleótido retroviral que comprende secuencias de repetición terminales largas (LTR) en el extremo 3' de la secuencia de polinucleótido retroviral, una secuencia promotora en el extremo 5' del polinucleótido retroviral, siendo dicho promotor adecuado para la expresión en una célula de mamífero, un dominio de ácido nucleico gag, un dominio de ácido nucleico pol y un dominio de ácido nucleico env;

un casete que comprende un sitio interno de entrada al ribosoma (IRES) unido operativamente a un ácido nucleico heterólogo que comprende un polinucleótido que comprende los pares de bases 8877 a 9353 de la SEQ ID NO: 19, donde T es U, donde el casete está posicionado en 5' a 3' LTR y 3' al dominio de ácido nucleico env que codifica la cubierta retroviral; y secuencias que actúan en cis necesarias para la transcripción inversa, empaquetamiento e integración en una célula objetivo, opcionalmente, donde la célula diana es una célula neoplásica y / o una célula que tiene un trastorno de proliferación celular, preferentemente, donde el trastorno de proliferación celular se selecciona de cáncer de pulmón, cáncer de colon y recto, cáncer de mama, cáncer de próstata, cáncer del tracto urinario, cáncer de útero, cáncer cerebral, cáncer de cabeza y cuello, cáncer pancreático, melanoma, cáncer de estómago y cáncer de ovarios, Artritis reumatoide u otra enfermedad autoinmune.

6. El vector de la reivindicación 5, donde el retrovirus tiene una secuencia polinucleotídica derivada del virus de la leucemia murina (MLV), opcionalmente donde el MLV es un MLV anfitriónico, el virus de la leucemia murina de Moloney (MoMLV), virus de la leucemia felina o virus de la leucemia del simio de Gibbon (GALV), y / o donde el retrovirus es un gammaretrovirus, y / o donde la secuencia promotora está asociada con un gen regulador del crecimiento, y / o donde la secuencia promotora comprende:

a) una secuencia promotora específica de tejido, preferentemente donde la secuencia promotora específica de tejido comprende al menos un elemento de respuesta a andrógenos (ARE), preferentemente donde el elemento de respuesta a andrógenos deriva del promotor de probasina y, opcionalmente, donde la secuencia promotora específica de tejido es un promotor de probasina; o

b) comprende un promotor del CMV que tiene una secuencia como se expone en la SEQ ID NO: 19, 20 o 22 desde el nucleótido 1 hasta el nucleótido 582, donde T es U, y que es capaz de dirigir e iniciar la transcripción; o

c) comprende un polinucleótido del dominio CMV-R-U5, preferentemente donde el dominio CMV-R-U5 comprende el promotor inmediatamente temprano de citomegalovirus humano unido a una región R-U5 del MLV, preferentemente, donde el polinucleótido del dominio CMV-R-U5 comprende una secuencia como se expone en la SEQ ID NO: 19, 20 o 22 desde el nucleótido 1 al nucleótido 1202 o secuencias que son al menos un 95 % idénticas a una secuencia como se expone en la SEQ. ID NO: 19, 20 o 22, donde el polinucleótido promueve la transcripción de una molécula de ácido nucleico unida operativamente a la anterior.

7. El vector de acuerdo con la reivindicación 6, donde el gag del polinucleótido deriva de un gammaretrovirus, preferentemente, donde el dominio de ácido nucleico gag comprende una secuencia desde el número de nucleótido 1203 hasta el nucleótido 2819 de la SEQ ID NO: 19, donde T es U, o una secuencia que tiene al menos un 95 %, 98 %, 99 % o 99,8 % de identidad con la anterior, y / o donde el dominio pol del polinucleótido se deriva de un gammaretrovirus, preferentemente en donde el dominio pol comprende una secuencia desde el número de nucleótido 2820 hasta el nucleótido 6358 de la SEQ ID NO: 19, donde T es U, o una secuencia que tiene al menos un 95 %, 98 %, 99 % o 99,9 % de identidad con la anterior,

y / o donde el dominio env comprende una secuencia desde el número de nucleótido 6359 hasta el nucleótido 8323 de la SEQ ID NO: 19, donde T es U, o una secuencia que tiene al menos un 95 %, 98 %, 99 % o 99,8 % de identidad, y / o donde el IRES deriva de un virus de encefalomiocarditis, preferentemente donde el IRES comprende una secuencia desde el número de nucleótido 8327 hasta aproximadamente el nucleótido 8876 de la SEQ ID NO: 19, donde T es U, o una secuencia que tiene al menos un 95 %, 98 % o 99 % de identidad con el mismo, y / o donde el LTR en 3' deriva de un gammaretrovirus, preferentemente, donde la LTR en 3' comprende un dominio U3-R-U5, o comprende una secuencia como se expone en la SEQ ID NO: 19 desde el nucleótido 9405 hasta el 9998 o una secuencia que es al menos un 95 %, 98 % o 99,5 % idéntico al mismo, y / o donde el polinucleótido retroviral comprende una secuencia como se expone en la SEQ ID NO: 19, 20 o 22, donde T es U.

8. Un polinucleótido aislado que comprende de 5' a 3':

una fusión CMV-R-U5 del promotor temprano inmediato del citomegalovirus humano en una región R-U5 del VLM; un PBS, un sitio de unión al cebador para la transcriptasa inversa;
un sitio 5' de corte y empalme;
una señal de empaquetamiento ψ ;
una secuencia de codificación gag para el antígeno específico del grupo del VLM;
una secuencia de codificación pol para la poliproteína de la polimerasa de VLM;
un sitio 3' de corte y empalme;
una secuencia de codificación de env 4070A para la proteína de la cubierta de la cepa 4070A del MLV;
un sitio interno de entrada al ribosoma (IRES) del virus de la encefalomiocarditis;
un polinucleótido que comprende los pares de bases 8877 a 9353 de la SEQ ID NO:19, que codifica un polipéptido que tiene actividad citosina desaminasa; un tracto de polipurina; y
una repetición terminal larga U3-R-U5 del MLV.

9. El polinucleótido de la reivindicación 8, donde el polinucleótido retroviral comprende una secuencia como se expone en la SEQ ID NO: 19 o 22,

y / o donde la secuencia de polinucleótidos retroviral deriva del virus de la leucemia murina (MLV), el virus de la leucemia murina de Moloney (MoMLV), el virus de la leucemia felina o el virus de la leucemia de simio de Gibbon (GALV), preferentemente donde el MoMLV es un MoMLV anfotrópico.

10. Un polinucleótido como se define en la reivindicación 8 o la reivindicación 9 para su uso en el tratamiento de un sujeto que tiene un trastorno de proliferación celular en combinación con 5-fluorocitosina en condiciones tales que se expresa el polinucleótido.

11. El polinucleótido para su uso de acuerdo con la reivindicación 10, donde las condiciones son tales que el polinucleótido está integrado en la célula del sujeto, y / o donde el polinucleótido está comprendido en un vector retroviral, preferentemente donde el vector retroviral comprende una secuencia como se expone en la SEQ ID NO: 19 o 22, y / o donde el trastorno de proliferación celular se selecciona de cáncer de pulmón, cáncer de colon y recto, cáncer de mama, cáncer de próstata, cáncer del tracto urinario, cáncer de útero, cáncer cerebral, cáncer de cabeza y cuello, cáncer pancreático, melanoma, cáncer de estómago y cáncer de ovarios, artritis reumatoide u otra enfermedad autoinmune, opcionalmente, donde el trastorno de proliferación celular es glioblastoma multiforme.

12. Un vector como se define en la reivindicación 5, para su uso en el tratamiento de un sujeto que tiene un trastorno de proliferación celular en combinación con 5-fluorocitosina.

13. El vector para su uso de acuerdo con la reivindicación 12, donde el ácido nucleico heterólogo codifica un polipéptido que comprende una secuencia como se expone en la SEQ ID NO: 4, 12, 14, 16, o 18, y / o donde el polinucleótido retroviral comprende una secuencia como se expone en la SEQ ID NO: 19 o 22.

14. Un polinucleótido aislado que comprende:

una primera secuencia de codificación que comprende los pares de bases 8877 a 9353 de la SEQ ID NO: 19 que codifica un polipéptido que tiene actividad citosina desaminasa; y
una segunda secuencia de codificación que codifica un polipéptido que tiene actividad UPRT u OPRT, donde la primera secuencia de codificación está unida operativamente a la segunda secuencia de codificación.

15. El polinucleótido aislado de la reivindicación 14, donde la primera y la segunda secuencia de codificación están separadas por una secuencia que codifica un enlazador peptídico, y / o donde el polinucleótido codifica un polipéptido que comprende una secuencia como se expone en la SEQ ID NO: 12, 14 o 16.

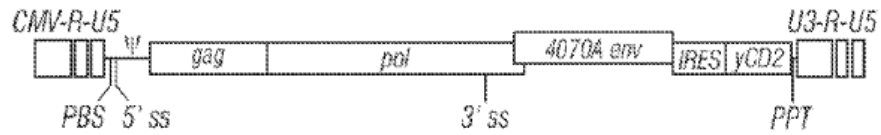


FIGURA 1A

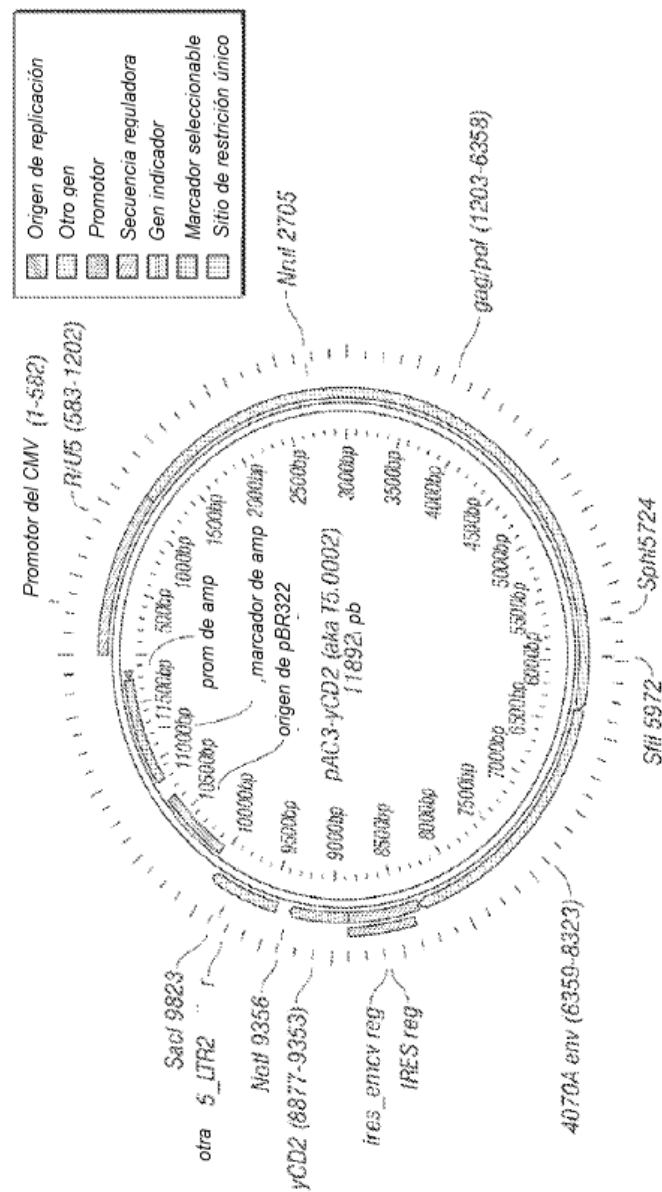


FIGURA 1B

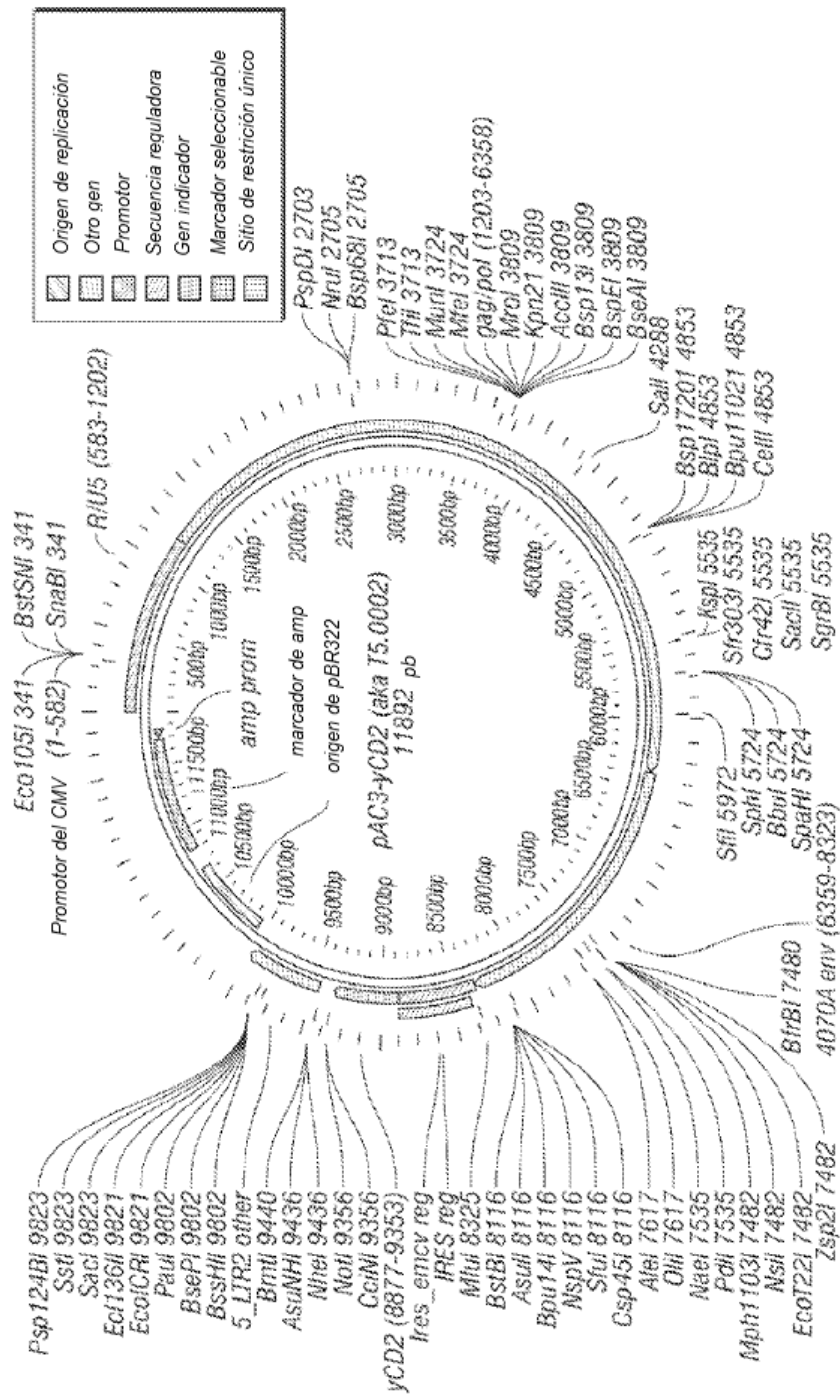


FIGURA 1C

222

C C A G A C C T C C C T C C T G A C C C T A G A T G A C T A G G A G G T C A G G G T C A G G A C C C C C C C C T G A A C C C A G G N T A A C C C T C A A A G T C
 G G G G G C A A C C C G T C A C C T T C C T G G T A G A T A C T G G G G C C A A C A C A C T C C G T G C T G A C C C A A A A T C C T G G A C C C C T A A G T G A T A
 A G T C T G C C T G G G T C C A A G G G G C T A C T G G A G G A A G G G G T A T C G C T G G A C C A G G G A T C G C A A A G T A C A T C T A G C T A C G G T A A
 G G T C A C C C A C T C T T T C C C A T G T A C C A G A C T G T C C C T A T C C T C T G T T A G G A G A G A T T T G C T G A C T A A A C T A A A G C C C A A
 A T C C A C T T T T G A G G G A T C A G G A G C C A G G T T A T G G G A C C A A T G G G C A G C C C C T G C A A C T G T T G A C C C T A A A T A T A G A A G A T G
 A G C A T C G G C T A C A T G A G A C C T C A A A G A G C C A G A T G T T T C T C T A G G G T C C A C A T G G C T G T C T G A T T T T C C T C A G G C C T G G G C
 G G A A C C G G G G C A T G G G A C T G G C A G T T C G C C A A G C T C C T C T G A T C A T A C C T C T G A A A G C A A C C T C T A C C C C C T G T C C A T A
 A A A C A A T C C C C A T G T C A C A A G A B C C A G A C T G G G G A T C A A G C C C A C A T A C A G A G A C T G T T G A C C A G G E A N T A C T G G T A C
 C C T G C C A G T C C C C C T G G A A C A C C C C C T G T A C C C G T T A A G A A A C C A G G G A C T A A T E A T T A T A G G C T G T C A G E A T C T G A G
 A G A A G T C A A C A A G C G G G T G G A A G A C A T C C A C C C A C C G T G C C A A C C C T A C A C C T C T T G A G C G G G T C C C A C C G T C C A C
 C A G T G C T A C A C T G T G C T T G A T T T A A G G A T G C C T T T T C T G C C T G A G A C T C C A C C C A C C A G T C A G C C T C T C T T C G C C T T T G
 A G T G A G A G A T C C A G A G A T G G G A T C T C A G A C A A T T G A C C T G G A C C A G A C T C C A C A G G C T T T C A A A A C A G T C C C A C C T
 G T T T G A T G A G C C A C T G C A C A G A C C T A G C A G A C T T C C G G A T C C A G C A C C C A G A C T T G A T C C T G C T A C A G T A C G T G G A T G A C
 T T A C T G C T G C C C G C A C T T C T G A G C T A G A C T G C C A A C A G G T A C T C G G G C C C T G T T A C A A C C C T A G G G A A C C T C G G G T A T C
 G G G C C T C G G C C A A G A A G C C C A A A T T T G C C A G A A C A G G T C A A G T A T C T G G G T A T C T T C T A A A G A G G G T C A G A G A T G C C T
 G A C T G A G G C C A G A A A G A G A C T G T G A T G G G G C A G C C T A C T C C G A A G A C C C T G A C A A C T A A G G G A G T T C C T A G G A C G G C A
 G G C T T C T G T C G C C T C T G G A T C C C T G G T T T G C A G A A A T G C C A G C C C C C T T G T A C C C T C T C A C C A A A A C G G G A C T C T G T T T A
 A T T G G G G C C A G A C C A A C A A A G G C C T A T C A G A A A T C A A G C A A G C C T C T T A A C T G C C C A G C C T C G G G T T G C C A G A T T T
 G A C T A A G C C C T T T G A C T C T T T G T C G A C G A A G C A G G G C T A C G C C A A G G T G T C C T A A C C A A A A A C T G G G A C C T T G C G T
 C E G C G G T G G C C T A C C T G T C C A A A A C T A G A C C A G T A G C A G C T G G G T G B C C C C T T G C C T A C G G A T G G T A G C A G C A T T G
 C C G T A C T G A C A A A G A T G C A G G C A A G C T A A C C A T G G G A C A G C C A C T A G T C A T T C T G C C C C C A T G C A G T A G A G C C A C T A G T
 C A A A C A A C C C C C G A C C G C T G C C T T C C A A G C C C G G T G A C T C A C T A T C A G G C C T T G C C T T T G A C A C C G A C C G G G T C C A G
 T T C G A C C G G T G G T A G C C C T G A A C C C G G T A C G C T G C T C C A C T G C C T H A G G A A G G C C T G C A A C A C A A C T G C C T T G A T A T C C
 T G C C G A A G C C C C A C G G A C C C G A C C C A C C T A A C G A C C A G C C C G C T C C A G A C C C A C C A C A C T T G G T A C A C G A T G E A A G
 C A G T C T C T A C A A G A G G A C A G C G T A A G G C G G G A C T G C G G T G A C C A C C G A G A C C G A G G T A A T C T G G G C T A A A G C C C T G C C A
 G C C G G A C A T C C C C T C A G C G G G C T G A C T G A T A G C A C T C A C C C A G G C C C T A A A G A T G G C A G A G G T A A G A G C T A A A T G T T T
 A T A C T G A T A G C C G T T A T G C T T T T G C T A C T G C C A T A T C C A T G G A G A A A T A T A C A G A A G G C G T G G G T T G C T A C A T C A G A A G G
 C A A A G A G A T C A A A A T A A A G A C G A G A T C T T G C C C T A C T A A A A G C C C T C T T T C T G C C A A A A G A C T T A G C A T A T C C A T T G T
 C C A G A C A T C A A A A G G A C A C A G C C G A G C T A G A G C A A C C G A T G C C T G A C C A G C G G C C G A A A G C A G C C A T C A C A G
 A G A C C C A G A C C C T C A C C C T C C A T A G A A A T T C A T C A C C C T A C A C C C T A G A C A T T T T C A T T A C A C A G T G A C T G A T A T

[pol]

FIGURA 1D-2

[illegible]**FIGURA 1D-3**

AGTAAAGCGCCGCGAGATAAAATAAAGATTTTATTTAGTCTCCAGAAAAAGCGCGAATGAAGGACCCACCTGTAGGTTTG
 CCAAGCTAGCTTAAGTAAGCCCATTTTCAAGGCATGCAAAATATCTAACTGAGAATAGAGAACTTCAGATCAAGGTCAGG
 AACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCCCCGGGTCAGGGCCAGAACAGAT
 GGAACAGCTGAATATGGGCCAAACAGGATATCTGTGCTAAGCAGTTCCTGCCCCGGGTCAGGGCCAGAACAGATGCTCCCC
 AGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCAGGGTGCCCCAGGACCTGAATGACCTGTGGC
 CTATATTGAACTAACCAATCAGTTGCTTCTCGGCTTCTGTTCGCGGCTTCCTGCTCCCGGAGCTCAATATAAAGAGCCCAAA
 CCCCCTCAGTCGGGCGCCAGTCTCTCGATTGACTEAGTGGCCCGGTACCCGCTATCCATAAACCCTCTTGCAGTTGCT
 CCGACTTGTGCTCTGCTGCTCTTGGAGGGTCTCTCTGAGTGATTGACTACCCGCTCAGCGGGGTCTTTCAATTACATGT
 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAABAAGGCCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCCCTGACGA
 GCATCACAAAAATCGAGCTCAAGTCAGAGTGCGCAACCCGACAGACTATAAAGATACCAAGCGTTTCCCCCTGGAGC
 TCCCTCGTGCGCTCTCTCTGTCGACCTGCGCGTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGGCTGGCGCTT
 CTCATGCTCAGCGCTGATGATCTCAGTTGCGTGATGCTGCTCCGCTCAAGCTGGGCTGTTGCAAGAACCCCGCTCA
 GCGCGAGCTGCTGCGCTTATCCGTAATCTATCTCTGAGTCACACCGGTAAGACAGCACTTATGCGCACTGGCAGCAGCC
 ACTGTAACAGGATTAGCAGAGGAGGATATGTAAGCGGTGCTACAGAGTTCTGAACTGGTGGCTAACTACGGCTACACTA
 GAAGGACAGTATTGCTATCTGCGCTCTGCTGAAGCCAGTTACCTTGGAAAAAGAGTTGTAAGCTCTTGATCCGCAACA
 AAGCCCGCTGGTACCGGTGCTTTTCTGTTTGCAGCAGTACATACCGCGCAAAAAAGGATCTCAAGAGATCTTTTG
 ATCTTTTCTACCGGCTCTGCGCTCAGTGGACGAAACTCAAGTTAAGGATTTTGGTCAAGGATATCAAAAGGATCT
 TCACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCAAAAGTATATATGAGTAACTTGGTCTGACAGTTTACCA
 ATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGCTATTTCTGTTTATCTAGTTTGGTGGCTCCCGCTGCTGATTA
 ACTACGATACCGGAGGGCTTACCATCTGGGCCAGTGGCTGCAATGATACCCGAGAGCCACGCTCAGCGCTCCAGTTTAT
 CAGCAATATAACAGCCAGCCGGAAGGCGGAGCCGCAAGTGGTCTGCACTTTATCCGCTCCATCCAGTCTATTAATG
 TTGCGCGGAGCTAGAGTAAGTAGTTGCGCAGTTAATAGTTTGGCAAGCTTGTGGCATTGCTGCAAGCATCGTGGTGTCA
 CGCTGCTGCTTGGTATGCTTCATTGAGTCCGCTTCCCAAGCATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAA
 AAGCGGTAGCTCTCTCGGCTCTCGGATCTGTGTCAGAAAGTAAGTTGGCCGAGTGTATCACTCATGCTTATGGCAGCACT
 GCAATATCTCTTACGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTCACTACTCAACCAAGTCATTCTGAGAAATG
 TGTATGCGCGGAGCGAGTTGCTCTTGGCCGGCTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTCTCTCA
 TCATTGGAAAACGTTCTTGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTGATGTAACCCACTGCTGC
 ACCCAACTGATCTTCAGCATCTTTACTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAAAG
 GGAATAAGGCGACACGGAAATGTTGAATACTCTATCTCTCTCTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTC
 TCATGAGCGGATACATATTTGAATGATTTAGAAAAATACAAATAGGGGTTCGCGGCACATTTCCCGAAGGTGCCACC
 TGACGCTCAGAGAACCATTTATCATGACATTAACCTATAAAATAGGGGTATCAGGAGCGCTTTTGGTCTTCAAGATTC
 AT

NotI

U3

U

U5

AseI

Amp^r ORF

ScaI

FIGURA 1D-4

Promotor del CMV (1-582)>>>

```

1   | TAGTATTAATACTAATCAATTAACGGGGTCATTAGTTTCATAGCCCATATATGGAGTTCCG 60
    | ATCAATAAATTATCATTAGTTAATGCCCCAGTAATCAAGTATCGGGTATATACCTCAAGGC
61  CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCATT 120
    | GCAATGTATTGAATGCCATTTACCGGGCGGACCGACTGGCGGGTTGCTGGGGCGGGTAA
121 GACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCA 180
    | CTGCAGTTATTACTGCATACAAGGGTATCATTGCGGGTATCCCTGAAAGGTAACGTGCAGT
181 ATGGGTGGAGTATTTACGGTAAACTGCCCCACTTGGCAGTACATCAAGTGTATCATATGCC 240
    | TACCCACCTCATAAATGCCATTTGACGGGTGAACCGTCATGTAGTTTACATAGTATACGG
241 AAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTA 300
    | TTCATCGGGGGGATAACTGCAGTTACTGCCATTTACCGGCGGACCGTAATACGGGTGAT

```

Eco105I

|

SnaBI

|

BstSNI

|

```

301 CATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC 360
    | GTACTGGAATACCTGAAAGGATGAACCGTCATGTAGATGCATAATCAGTAGCGATAATG
361 CATGGTGATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGAATCACGGGG 420
    | GTACCACTACGCCAAAACCGTCATGTAGTTACCCGCACCTATCGCCAAAACGAGTGCCCC
421 ATTTCCAAGTCTCCACCCCATGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG 480
    | TAAAGGTTTACAGAGGTGGGGTAACTGCAGTTACCCCTCAAACAAAACCGTGGTTTTAGTTGC
481 GGACTTTCCAAAATGTCGTAACAACCTCCGCCCATTTGACGCAATGGGCGGTAGGCGTGT 540
    | CCTGAAAGGTTTTACAGCATTGTTGAGGCGGGGTAACTGCCTTTACCCGCCATCCGCACA

```

Región R (583-650)>>>

|

```

541 ACCGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGGCGCCAGTCTCCGATTG 600
    | TGCCACCCCTCCAGATATATTCTCTCGACCAAATCACTTGGCCGCGGTACAGGAGGCTAAC

```

FIGURA 1D-4

Región U5 (851-1202)>>>

|

601 ACTGAGTCGCCCCGGGTACCCGTGTATCCAATAAACCCCTCTTGCAGTTGCATCCGACTTGT 660
 TGACTCAGCGGGCCCATGGGCACATAGGTTATTTGGGAGAACCTCAACGTAGGCTGAACA

661 GGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGATTGACTACCCGTCAGCGGGGGTC 720
 CCAGAGCGACAAGGAACCCCTCCAGAGGAGACTCACTAACTGATGGGCAGTCGCCCCCAG

721 TTTCATTTGGGGGCTCGTCCGGGATCGGGAGACCCCTGCCAGGGACCAACGACCCACCA 780
 AAGCTAAACCCCGAGCAGGCCCTAGCCCTCTGGGGACGGGTCCCTGGTGGCTGGGTGGT

5'SS (780)

|

781 CCGGGAGGTAAGCTGGCCAGCAACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGAC 840
 GGCCCTCCATTTCGACCGGTCTGTGAATAGACACAGACAGGCTAACAGATCACAGATACTG

841 TGAATTTTATGGCCCTGGGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCG 900

FIGURA 1D-5

ACTAAAATACGCGGACGCGGCCATGATCAATCGATTGATCGAGACATAGACCGCCTGGGC

901 TGGTGGAACTGACGAGTTGCGGAACACCGGCGCAACCCCTGGGAGACGTCACAGGGACTT 960
ACCACCTTGACTGCTCAAGCCTTGTGGGCGGGCTTGGGACCCTCTGCAGGGTCCCTGAA

961 CGGGGGCGGTTTTTGTGGCCGACCTGAGTCCAAAAATCCCGATCGTTTTGGAATCTTTG 1020
GCCCCCGGCAAAAACACCGGGCTGGACTCAGGTTTTTAGGGCTAGCAAAACCTGAGAAAC

1021 GTGCACCCCCCTTAGAGGAGGGATATGTGGTTCTGGTAGGAGACGAGAACCCTAAACAGT 1080
CACGTGGGGGGAATCTCCTCCCTATACACCAAGACCATCCTCTGCTCTTGATTTTGTCA

1081 TCCCGCCTCCGTCTGAATTTTGTCTTTCGGTTTGGGACCGAAGCCGCGCGCGCGCTCTTG 1140
AGGGCGGAGGCAGACTTAAAAACGAAGCCAAACCCCTGGCTTCGGCGCGCGCGCGCAGAAC

1141 TCTGCTGCAGCATCGTTCTGTGTGTCTCTGTCTGACTGTGTTTCTGTATTTGTCTGAGA 1200
AGACGACGCTCGTAGCAAGACACAACAGAGACAGACTGACACAAAGACATAAACAGACTCT

gag (1203, 2819) >>>

1201 ATATGGGGCCAGACTGTTACCACTCCCTTAAGTTTGACCTTAGGTCACTGGAAAGATGTCTG 1260
TATACCCGGTCTGACAATGGTGAGGGAATTCAAACGGAATCCAGTGACCTTTCTACAGC

1261 AGCGGATCGCTCACAACCAAGTCGGTAGATGTCAAGAAGAGACGTTGGGTTACCTTCTGCT 1320
TCGCCTAGCGAGTGTGGTCAGCCATCTACAGTTCTTCTCTGCAACCAATGGAAGACGA

1321 CTGCAGAATGGCCAACCTTTAAGCTCGGATGGCGCGGAGACGGCACCTTTAACCAGAGACC 1380
GACGTCTTACCGGTTGGAATTCAGGCTACCGGCGCTCTGCCGTGGAAATTTGGCTCTGG

1381 TCATCACCCAGGTTAAGATCAAGGTCTTTTCACTTGGCCCGCATGGACACCCAGACAGG 1440
AGTAGTGGGTCCAATTCTAGTTCCAGAAAAGTGGACCGGGCGTACCTGTGGGTCTGGTCC

1441 TCCCTACATCGTGACCTGGGAAGCCTTGGCTTTTGACCCCCCTCCCTGGGTCAAGCCCT 1500
AGGGGATGTAGCACTGGACCCCTTCGGAACCGAAAACCTGGGGGGAGGGACCCAGTTCGGGA

1501 TTGTACACCCCTAAGCCTCCGCTCCTCTTCTCCATCCGCCCCGTCTCTCCCCCTTGAAC 1560
AACATGTGGGATTCGGAGCGGAGGAGAAGGAGGTAGGCGGGGAGAGAGGGGGAACCTG

1561 CTCTCGTTTCGACCCCGCTCGATCTCCCTTTATCCAGCCCTCACTCCTTCTCTAGGCG 1620
GAGGAGCAAGCTGGGCGGAGCTAGGAGGGAATAGGTCGGGAGTGAGGAAGAGATCCGC

1621 CCAAACTTAAACCTCAAGTTCTTTCTGACAGTGGGGGGCGCTCATCGACCTACTTACAG 1680
GGTTTGGATTTGGAGTTCAAGAAAGACTGTACCCCCCGGCGAGTAGCTGGATGAATGTC

1681 AAGACCCCCCGCCTTATAGGGACCCAAGACCACCCCTTCCGACAGGGACGGAAATGGTG 1740
TTCTGGGGGGCGGAATATCCCTGGGTTCTGGTGGGGGAAGGCTGTCCCTGCCTTTACCAC

FIGURA 1D-6

1741 GAGAAGCGACCCCTGCGGGAGAGGCACCGGACCCCTCCCCAATGGCATCTCGCCTACGTG 1800
CTCTTCGCTGGGGACGCCCTCTCCGTGGCTTGGGGAGGGGTTACCGTAGAGCGGATGCAC

1801 GGAGACGGGAGCCCCCTGTGGCGGACTCCACTACCTCGCAGGCATTCCCCCTCCGCGCAG 1860
CCTCTGCCCTCGGGGGACACCGGCTGAGGTGATGGAGCGTCCGTAAGGGGGAGGCGCGTC

1861 GAGGAAACGGACAGCTTCAATACTGGCCGTTCTCCTCTTCTGACCTTTACAACCTGGAAAA 1920
CTCCTTTGCCTGTGGAAGTTATGACCGGCAAGAGGAGAAGACTGGAAATGTTGACCTTTT

1921 ATAATAACCCCTTCTTTTCTGAAGATCCAGGTAAACTGACAGCTCTGATCGAGTCTGTTT 1980
TATTATTGGGAAGAAAAGACTTCTAGGTCCATTGACTGTGAGACTAGCTCAGACAAG

1981 TCATCACCCATCAGCCACCTGGGACGACTGTCAGCAGCTGTTGGGGACTCTGCTGACCG 2040
AGTAGTGGGTAGTCGGGTGGACCCCTGCTGACAGTCGTGACACACCCCTGAGACGACTGGC

3121 ACTTTGAGGGATCAGGAGCCCAGGTTATGGGACCAATGGGGCAGCCCTGCAAGTGTTGA 3180
TGAAACTCCCTAGTCCCTCGGGTCCAATACCCTGGTTACCCCGTUGGGGACGTTCAAACT

3181 CCCTAAATATAGAAGATGAGCATCGGCTACATGAGACCTCAAAAGAGCCAGATGTTTCTC 3240
GGGATTTATATCTTCTACTCGTAGCCGATGTACTCTGCAGTTTTCTCGGTCTACAAAGAG

3241 TAGGGTCCACATGGCTGTCTGATTTTCTCAGGCCGCGGCGAAACCGGGGGCATGGGAC 3300
ATCCCAGGTGTACCGACAGACTAAAAGGAGTCCGGACCCGCCCTTGGCCCCCGTACCCCTG

3'SS (3314)

3301 TGGCAGTTCCGCAAGCTCCTCTGATCATACTCTGAAAGCAACCTCTACCCCGTGTCCA 3360
ACCGTCAAGCGGTTCGAGGAGACTAGTATGGAGACTTTCGTTGGAGATGGGGGCACAGGT

3361 TAAACAATACCCCATGTCACAAGAAGCCAGACTGGGGATCAAGCCCCACATACAGAGAC 3420
ATTTTGTATGGGTACAGTGTCTTCGGTCTGACCCCTAGTTGCGGGGTGTATGTCTCTG

3421 TGTTGGACCCAGGGAATACTGGTACCCTGCCAGTCCCCCTGGAACACGCCCCCTGCTACCCG 3480
ACAACCTGGTCCCTTATGACCATGGGACGGTCAGGGGGACCTTGTGCGGGGACGATGGGC

3481 TTAAGAAACCAGGGACTAATGATTATAGGCCTGTCCAGGATCTGAGAGAACTCAACAAGC 3540
AATTCTTTGGTCCCTGATTACTAATATCCGGACAGGTCTAGACTCTCTTCAGTTGTTCTG

3541 GGGTGAAGACATCCACCCACCGTGCCCAACCCCTTACAACCTCTTGAGCGGGCTCCAC 3600
CCCACCTCTGTAGGTGGGTGGCACGGGTGGGAATGTTGGAGAACTCGCCCGAGGGTG

3601 CGTCCACCACTGGTACACTGTGCTTGATTTAAAGGATGCCCTTTTCTGCCTGAGACTCC 3660
GCAGGGTGGTCACCATGTGACACGAATAAATTTCTACGGAAAAAGACGGACTCTGAGG

PfeI
|
TfiI
|

3661 ACCCCACCACTCAGCCTCTCTTCGCCCTTTGAGTGGAGAGATCCAGAGATGGGAATCTCAG 3720
TGGGGTGGTCAGTCGGAGAGAAGCGGAAACTCACCTCTCTAGGTCTCTACCCCTAGAGTC

FIGURA 1D-7

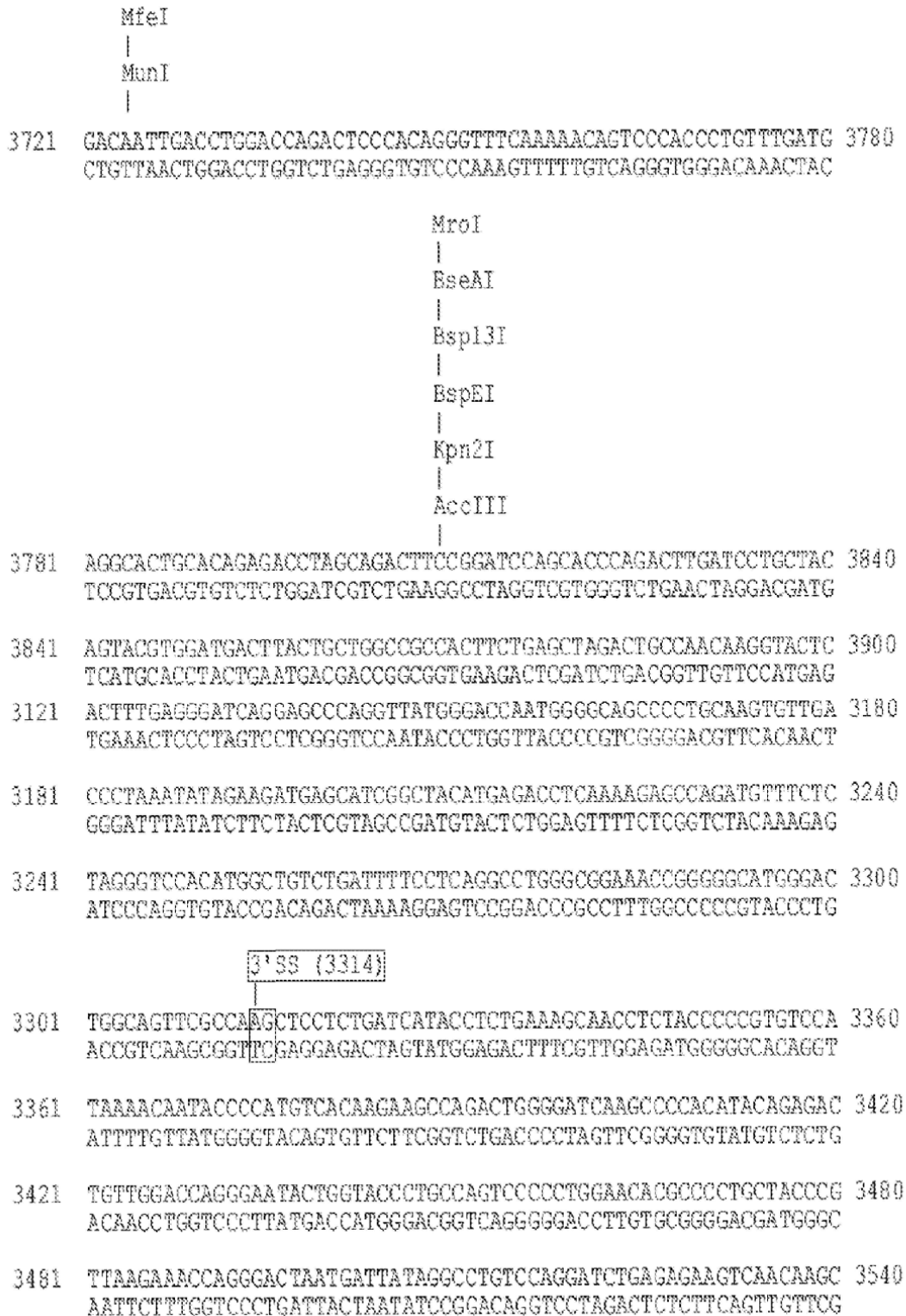


FIGURA 1D-8

```

3541 GGGTGGGAAGACATCCACCCCAACCGTGCCCAACCCTTACAACTCTTGAGCGGGCTCCAC 3600
      CCCAOCCTTCTGTAGGTGGGGTGGCACGGGTGGGAATGTTGGAGAACTCGCCCGAGGGTG

3601 CGTCCCAACAGTGGTACACTGTGCTTGATTTAAAGGATGCCTTTTCTGCCTGAGACTCC 3660
      GCAGGGTGGTCAACATGTGACACGAACTAAATTTCTACGGAAAAAGACGGACTCTGAGG

                                     PfeI
                                     |
                                     TfiI
                                     |
3661 ACCCCACCAAGTCAGCCTCTCTTCGCCTTTGAGTGGAGAGATCCAGAGATGGGAATCTCAG 3720
      TGGGGTGGTCAGTCGGAGAGAAAGCGGAAACTCACCTCTCTAGGTCTCTACCCCTTAGAGTC

      MfeI
      |
      MunI
      |
3721 GACAATTGACCTGGACCAAGACTCCACAGGGTTTCAAAAACAGTCCACCCCTGTTTGATG 3780
      CTGTTAACTGGACCTGGTCTGAGGGTGTCCCAAAGTTTTTGTGAGGGTGGGACAAACTAC

                                     MroI
                                     |
                                     EaeAI
                                     |
                                     Bsp13I
                                     |
                                     BspEI
                                     |
                                     Kpn2I
                                     |
                                     AccIII
                                     |
3781 AGGCACTGCACAGAGACCTAGCAGACTTCCGGATCCAGCAACCAGACTTGATCCTGCTAC 3840
      TCCGTGACGTCTCTCTGGATCGTCTGAAGGCCTAGGTCTGGTCTGAACTAGGACGATG

3841 AGTACGTGGATGACTTACTGCTGCCCCCACTTCTGAGCTAGACTGCCAACAAAGTACTC 3900
      TCATGCACCTACTGAATGACGACCGCGGTGAAGACTCGATCTGACGGTTGTTCCATGAG

```

FIGURA 1D-9

3901 GGGCCCTGTTACAAACCCCTAGGGAACCTCGGGTATCGGGCCTCGGCCAAGAAAGCCCAAA 3960
 CCCGGGACAATGTTTGGGATCCCTTGGAGCCCATAGCCCAGGCGGTTCTTTCGGGTTT
 3961 TTTGCCAGAAACAGGTCAAGTATCTGGGGTATCTTCTAAAAGAGGGTCAGAGATGGCTGA 4020
 AAACGGTCTTTGTCCAGTTCATAGACCCCATAGAAGATTTTCTCCAGTCTCTACCGACT
 4021 CTGAGGCCAGAAAAGAGACTGTGATGGGGCAGCCTACTCCGAAGACCCCTCGACAACATA 4080
 GACTCCGGTCTTTTCTCTGACACTACCCCGTCGGATGAGGCTTCTGGGGAGCTGTTGATT
 4081 GGGAGTTCCTAGGGACGGCAGGCTTCTGTGCGCTCTGGATCCCTGGGTTTGCAGAAATGG 4140
 CCTCAAGGATCCCTGCCGTCCGAAGACAGCGGAGACCTAGGGACCCAAACGTCCTTACC
 4141 CAGCCCCCTTGTACCTCTCAACAAAACGGGGACTCTGTTTAATTGGGGCCAGACCAAC 4200
 GTCGGGGGAACATGGGAGAGTGCTTTTGCCCTGAGACAAATTAACCCCGGGTCTGGTTG
 4201 AAAAGGCCTATCAAGAAATCAAGCAAGCTCTTCTAACTGCCCCAGCCCTGGGGTTGCCAG 4260
 TTTTCCGATAGTTCTTTAGTTCTGTCGAGAAGATTGACGGGGTCGGGACCCCAACGGTC

 SalI
 |
 4261 ATTTGACTAAGCCCTTTGAACTCTTTGTGACGAGAAGCAGGGCTACGCCAAAGGTGTCC 4320
 TAAACTGATTTCGGGAACTTGAGAAACAGCTGCTCTTCGTCCCGATGCGGTTTCCACAGG
 4321 TAACGCAAAAACTGGGACCTTGGCGTCGGCCGGTGGCCTACCTGTCCAAAAGCTAGACC 4380
 ATTGCGTTTTTGAACCTGGAAACCGCAGCCGGCCACCGGATGGACAGGTTTTTCGATCTGG
 4381 CAGTAGCAGCTGGGTGGCCCCCTTGCCCTACGGATGGTAGCAGCCATTGCCGTACTGACAA 4440
 GTCATCGTCGACCCACCGGGGGAACGGATGCCCTACCATCGTCGGTAACGGCATGACTGTT
 4441 AGGATGCAGGCAAGCTAACCATGGGACAGCCACTAGTCATTCTGGCCCCCATGCAGTAG 4500
 TCCTACGTCCGTTTCGATTGGTACCCTGTCGGTGATCAGTAAGACCGGGGGTACGTCATC
 4501 AGGCACTAGTCAAACAACCCCGGACCGCTGGCTTTCCAAACGCCCGGATGACTCACTATC 4560
 TCCGTGATCAGTTTGTGGGGGGCTGGCGACCGAAAGGTTGCGGGCCTACTGAGTGATAG
 4561 AGGCCTTGCTTTTGGACACGGACCGGGTCCAGTTCGGACCGGTGGTAGCCCTGAACCCGG 4620
 TCCGGAACGAAACCTGTGCTGGCCAGGTCAAGCCTGGCCACCATCGGGACTTGGGCC
 4621 CTACGCTGCTCCCACTGCCCTGAGGAAGGGCTGCAACACAACCTGCCTTGATATCCTGGCCG 4680
 GATGCGACGAGGGTGACGGACTCCTTCCCGACGTTGTGTTGACGGAACATATAGGACCGGC
 4681 AAGCCACGGAACCCGACCCGACCTAACGGACAGCCGCTCCAGACGCCGACCCACACCT 4740
 TTCGGGTGCCTTGGGCTGGGCTGGATTGCCTGGTCGGCGAGGGTCTGCGGCTGGTGTGGA
 4741 GGTACACGGATGGAAGCAGTCTCTTACAAGAGGGACAGCGTAAGGCGGGAGCTGCGGTGA 4800
 CCATGTGCCTACCTTCGTCAGAGAATGTTCTCCCTGTCGATTCGCGCCTCGACGCCACT

FIGURA 1D-10

BspI
 |
 CclII
 |
 Bpu1102I
 |
 Bsp1720I
 |

4801 CCACCGAGACCGAGGTAATCTGGGGCTAAAGCCCTGCCAGCCGGGACATCCGCTCAGCGGG 4860
 GGTGGCTCTGGCTCCATTAGACCCGATTTCGGGACGGTCGGCCCTGTAGCGGAGTCGCCC
 4861 CTGAACTGATAGCACTCAGCCAGGCCCTAAAGATGGCAGAAGGTAAGAAGCTAAATGTTT 4920
 GACTPGACTATCGTGAGTGGGTCCGGGATTTCTACCGTCTTCCATTCTTCGATTTACAAA
 4921 ATACTGATAGCCGTTATGCTTTTGCTACTGCCCATATCCATGGAGAAATATACAGAAGGC 4980
 TATGACTATCGGCAATACGAAAACGATGACGGGTATAGGTACCTCTTTATATGTCTTCCG
 4981 GTGGGTTGCTCAGATCAGAAAGCAAGAGATCAAAAATAAAGACGAGATCTTGGCCCTAC 5040
 CACCCAAACGAGTGTAGTCTTCCGTTTCTCTAGTTTTTATTTCTGCTCTAGAACCAGGATG
 5041 TAAAAGCCCTCTTTCTGCCCCAAAAGACTTAGCATAATCCATTGTCCAGGACATCAAAAGG 5100
 ATTTTCGGGAGAAAGACGGGTTTTCTGAATCGTATTAGGTAAACAGGTCCGTAGTTTTCC
 5101 GACACAGCGCCGAGGCTAGAGGCAACCGGATGGCTGACCAAGCGGCCCGAAAGGCAGCCA 5160
 CTGTGTGCGGGCTCCGATCTCCGTTGGCCTACCGACTGGTTCGCCGGGCTTTCCGTGGT
 5161 TCACAGAGACTCCAGACACCTCTACCCCTCCTCATAGAAAATTCATCACCTACACCTCAG 5220
 AGTGTCTCTGAGGTCTGTGGAGATGGGAGGAGTATCTTTTAAGTAGTGGGATGTGGAGTC
 5221 AACATTTTCATTACACAGTGACTGATATAAAGGACCTAACCAAGTTGGGGGCCATTTATG 5280
 TTGTAAAAGTAATGTGTCAGTACTATATTTCTGGATTGGTTCAACCCCGGTAAATAC
 5281 ATAAAACAAAGAAGTATTGGGTCTACCAAGGAAAACCTGTGATGCCCTGACCAGTTTACTT 5340
 TATTTTGTTCCTTCATAAACCAGATGGTTCCTTTGGACACTACGGACTGGTCAAATGAA
 5341 TTGAATTATTAGACTTTCTTCATCAGCTGACTCACCTCAGCTTCTCAAAAATGAAGGCTC 5400
 AACTTAATAATCTGAAAGAAGTAGTCGACTGAGTGGAGTCGAAGAGTTTTTACTTCCGAG
 5401 TCCTAGAGAGAAGCCACAGTCCCTACTACATGCTGAACCGGGATCGAAGACTCAAAAATA 5460
 AGGATCTCTCTTCGGTGTGAGGATGATGTACGACTTGGCCCTAGCTTGTGAGTTTTTAT

FIGURA 1D-11

5461 TCACTGAGACCTGCAAAGCTTGTGCACAAGTCAACGCCAGCAAGTCTGCCGTTAAACAGG 5520
AGTGACTCTGGACGTTTCGAACACGTGTTTCAGTTGCCGTCGTTTCAGACGGCAATTTGTCC

SacII
|
SgrBI
|
Cfr42I
|
Sfr303I
|
KspI
|

5521 GAACTAGGGTCCCGGGGCATCGGCCGGGCACTCATTGGGAGATCGATTTCACCGAGATRA 5580
CTTGATCCCAGGCGCCCGTAGCCGGGCGGTGAGTAACCCCTCTAGCTAAAGTGGCTCTATT

5581 AGCCCGGATTGTATGGCTATAAATATCTTCTAGTTTTTATAGATACCTTTTCTGGCTGGA 5640
TCGGGCTTAAACATACCGATATTTATAGAAGATCAAAAATATCTATGGAAAAGACCGACCT

5641 TAGAAGCCTTCCCAACCAAGAAGAAACCGCCAAGGTCGTAACCAAGAAGCTACTAGAGG 5700
ATCTTCGGAAGGGTTGGTTCTTTCTTTGGCGGTTCCAGCATTGGTTCTTCGATGATCTCC

PaeI
|
BbuI
|
SpaHI
|
SphI
|

5701 AGATCTTCCCCAGGTTTCGGCATGCCTCAGGTATTGGGAACTGACAATGGGGCTGCCTTCG 5760
TCTAGAAGGGGTCCAAGCCGTACGGAGTCCATAACCCCTTGACTGTTACCCGGACGGAAGC

6901 CTTCCAAGGGGCTACTCGAGGGGGCAGATGCAACCCCTCTAGTCTTAGAATTCAGTGATGC 6960
GAAGGTTCCCGATGAGCTCCCCGCTACGTTGGGAGATCAGGATCTTAAGTGACTACG

6961 AGGAAAAAAGGCTAACTGGGACGGGCCAAATCGTGGGGAAGTGAAGTGTACCGGACAGG 7020
TCCTTTTTTTCGATTGACCCCTGCCCCGGGTTTAGCACCCCTGACTCTGACATGGCCTGTCC

7021 AACAGATCCTATTACCATGTTCTCCCTGACCCGGCAGGTCCCTTAATGTGGGACCCCGAGT 7080
TTGCTC TAGGATAATGGTACAAGAGGGACTGGGCCGTCCAGGAATTACACCCCTGGGGCTCA

7081 CCCCATAGGGCCCAACCCAGTATTACCCGACCAAGAGCTCCCTTCCTCACC AATAGAGAT 7140
GGGGTATCCCGGTTGGGTCATAATCGGCTGGTTTCTGAGGGAAGGAGTGGTTATCTCTA

7141 TGTACCGGCTCCACAGCCACCTAGCCCCCTCAATACCAGTTACCCCCCTTCCACTACCAG 7200
ACATGGCCGAGGTGTCGGTGGATCGGGGAGTTATGGTCAATGGGGGAAGGTGATGGTC

7201 TACACCCCTCAACCTCCCTACAAGTCCAAGTGTCCACAGCCACCCCGAGGAAGTGGAGA 7260
ATGTGGGAGTTGGAGGGGATGTTTCAGGTTACAGGGTGTGGTGGGGGTCTTGACCTCT

7261 TAGACTACTAGCTCTAGTCAAAGGAGCCTATCAGGCGCTTAACCTCACC AATCCCAGCAA 7320
ATCTGATGATCGAGATCAGTTTCCTCGGATAGTCCGCGAATTGGAGTGGTTAGGGCTGTT

FIGURA 1D-12

```

7321  GACCCAAGAANTGTTGGCTGTGCTTAGTGTCGGGACCTCCTTATTACGAAGGAGTAGCGGT 7380
      CTGGGTTCTTACAACCGACACGAATCACAGCCCTGGAGGAATAATGTTCTCATCGCCA

7381  CGTGGGCACCTTATACCAATCATTCACCCGCTCCGGCCAACTGTACGGCCACTTCCCAACA 7440
      GCACCCGTGAATATGGTTAGTAAGGTGGGAGGCCCGGTTGACATGCCGGTGAAGGGTTGT

                                     EcoT22I
                                     |
                                     NsiI
                                     |
                                     Mph1103I
                                     |
                                     Zsp2I
                                     |
                                     BfrBI
                                     ||
7441  TAAGCTTACCCTATCTGAAGTGACAGGACAGGGCCTATGCATGGGGGACGTACCTAAAAC 7500
      ATTCGAATGGGATAGACTTCACTGTCTGTCCCGGATACGTACCCCGTCATGGATTTTG

                                     NaeI
                                     |
                                     PdiI
                                     |
                                     MroNI
                                     ||
                                     NgoMIV
                                     ||
7501  TCACCAGGCCCTTATGTAACACCACCCAAAGCGCCGGCTCAGGATCCTACTACCTTGCAGC 7560
      AGTGGTCCGGAATACATTGTGGTGGGTTTTCCGGGCCGAGTCTAGGATGATGGAACGTCC

                                     AleI
                                     |
                                     oliI
                                     |
7561  ACCCGCCGGAACAATGTGGGCTTGCAGCACTGGATTGACTCCCTGCTTGTCCACCACGGT 7620
      TGGGCGGCCCTTGTACACCCGAACGTGCTGACCTAACTGAGGGACGAACAGGTGGTGCCA

7621  GCTCAATCTAACCACAGATTATTGTGTATTAGTTGAACTCTGGCCCAGAGTAATTTACCA 7680
      CGAGTTAGATTGGTGTCTAATAACACATAATCAACTTGAGACCGGGTCTCATTAAATGGT

```

FIGURA 1D-13

6901 CTTCCAAGGGGCTACTCGAGGGGGCAGATGCAACCCCTCTAGTCCTAGAAATTCAGTATGC 6960
GAAGGTTCCCCGATGAGCTCCCCCGTCTACGTTGGGAGATCAGGATCTTAAGTGACTACG

6961 AGGAAAAAGGCTAACTGGGACGGGCCCCAANTCGTGGGGACTGAGACTGTACCGGACAGG 7020
TCCTTTTTTCGATTGACCCGCCCCGGGTTTAGCACCCCTGACTCTGACATGGCCTGTCC

7021 AACAGATCCTATTACCATGTTCCTCCCTGACCCGGCAGGTCCTTAATGTGGGACCCCGAGT 7080
TTGTCTAGGATAATGGTACAAGAGGCACTGGGCCGTCCAGGAATTACACCTGGGGCTCA

7081 CCCCATAGGGCCCAACCCAGTATTACCCGACCAAGACTCCCTTCCTCACC AATAGAGAT 7140
GGGGTATCCCGGGTTGGGTCATAATGGGCTGGTTTCTGAGGGAAGGAGTGGTTATCTCTA

7141 TGTACGGGCTCCACAGCCACCTAGCCCCCTCAATACCAGTTACCCCCCTTCCACTACCAG 7200
ACATGGCCGAGGTGTGGTGGATCGGGGAGTTATGGTCAATGGGGGGAAGGTGATGGTC

7201 TACACCTCAACCTCCCCCTACAAGTCCAAGTGTCCACAGCCACCCCAAGGAACTGGAGA 7260
ATGTGGGAGTTGGAGGGGATGTTCAAGTTACAGGGTGTGGTGGGGTCCCTGACCTCT

7261 TAGACTACTAGCTCTAGTCAAGGAGCCTATCAGGCGCTTAACCTCACCAATCCCGACAA 7320
ATCTGATGATCGAGATCAGTTTCCTCGGATAGTCCCGGAATTGGAGTGGTTAGGGCTGTT

7321 GACCCAAGAATGTTGGCTGTGCTTAGTGTGGGACCTCCTTATTACGAAGGAGTAGCGGT 7380
CTGGGTTCCTACAACCGACACGAATCACAGCCCTGGAGGAATAATGCTTCCTCATCGCCA

7381 CGTGGGCACCTTATACCAATCATTCACCCGCTCCGGCCAACGTGTACGGCCACTTCCCAACA 7440
GCACCCGTGAATATGGTTAGTAAGGTGGCGAGGCCGGTTGACATGCCGGTGAAGGGTTGT

EcoT22I
|
NsiI
|
Mph1103I
|
Zsp2I
|
BfrBI
| |

7441 TAAGCTTACCTATCTGAAGTGACAGGACAGGGCCTATGCATGGGGGCAGTACCTAAAAC 7500
ATTCGAATGGGATAGACTTCACTGTCTGTCCCGGATACGTACCCCGTCAATGGATTTTG

NaeI
|
PdiI
|
MroNI
| |
NgoMIV
| |

7501 TCACCAGGCCTTATGTAACACCACCCAAAGCGCCGGCTCAGGATCCTACTACCTTGCAGC 7560
AGTGGTCCCGAATACATTGTGGTGGGTTTCGCGGCCGAGTCTAGGATGATGCAACGTCC

FIGURA 1D-14

AleI
 |
 OllI
 |

7561 ACCCGCCGGAACAATGTGGGCTTGCAGCACTGGATTGACTCCCTGCTTGTCCACCACGGT 7620
 TGGGCGGCCTTGTTACACCCGAACGTCGTGACCTAACTGAGGGACGAACAGGTGGTGCCA
 7621 GCTCAATCTAACCACAGATTATTGTGTATTAGTTGAACTCTGGCCCAGAGTAATTTACCA 7680
 CGAGTTAGATTGGTGTCTAATAACACATAATCAACTTGAGACCGGGTCTCATTAAATGGT
 7681 CTCCCCCGATTATATGTATGGTCAGCTTGAACAGCGTACCAAATATAAAAGAGAGCCAGT 7740
 GAGGGGGCTAATATACATACCAGTCGAACCTTGTGCGATGGTTTATATTTTCTCTCGGTCA
 7741 ATCATTGACCCTGCCCCCTTCTACTAGGAGGATTAACCATGGGAGGGATTGCAGCTGGAAT 7800
 TAGTAACTGGGACCGGGAAGATGATCCTCCTAATTGGTACCCTCCCTAAGTCGACCTTA
 7801 AGGGACGGGGACCACTECCTTAATTAAAAACCCAGCAGTTTGAGCAGCTTCATGCCGCTAT 7860
 TCCCTGCCCTGGTGACGGAATTAATTTTGGGTCGTCAAACCTCGTCGAAGTACGGCGATA
 7861 CCAGACAGACCTCAACGAAGTCGAAAAGTCAATTACCAACCTAGAAAAGTCACTGACCTC 7920
 GGTCTGTCTGGAGTTGCTTCAGCTTTTCAGTTAATGGTTGGATCTTTTCAGTGACTGGAG
 7921 GTTGTCTGAAGTAGTCCTACAGAACCGLAGAGGCCTAGATTGCTATTCTAAAGGAGGG 7980
 CAACAGACTTCATCAGEATGTCTTGGCGTCTCCGGATCTAAACGATAAGGATTTCTCTCC
 7981 AGGTCTCTGCGCAGCCCTAAAAGAAGAATGTTGTTTTTATGCAGACCACACGGGGCTAGT 8040
 TCCAGAGACGGCTCGGGATTTTCTTCTTACAACAAAAATACCTCTGGTGTGCCCCGATCA
 8041 GAGAGACAGCATGGCCAAATTAAGAGAAAGGCTTAATCAGAGACAAAAACTATTTGAGAC 8100
 CTCTCTGTGCTACCGGTTTAATTCCTTTCCGAATTAGTCTCTGTTTTTGATAAACTCTG

NspV
 |
 BstBI
 |
 Bsp119I
 |
 AsuII
 |
 Csp45I
 |
 SfuI
 |
 Bpu14I
 |
 BspT104I
 |

8101 AGGCCAAGGATGGTTTCGAAGGGCTGTTTAATAGATCCCCCTGGTTTACCACCTTAATCTC 8160
 TCCGGTTCTACCAAGCTTCCCGACAAATTATCTAGGGGGACCAATGGTGAATTAGAG

FIGURA 1D-15

```

8161 CACCATCATGGGACCTCTAATAGTACTCTTACTGATCTTACTCTTTGGACCTTGCATTCT 8220
      GTGGTAGTACCTGGAGATTATCATGAGAATGACTAGAATGAGAAACCTGGAACGTAAGA

8221 CAATCGATTGGTCCAATTTGTTAAAGACAGGATCTCAGTGGTCCAGGCTCTGCTTTTGAC 8280
      CTTAGCTAACCCAGGTTAAACAATTTCTGTCTAGAGTCACCAGGTCCGAGACCAAACTG
                                     IRES reg(8327,8876)>>>
                                     |
                                     MluI(8325)
                                     | |
8281 TCAGCAATATCACCAGCTAAAACCCATAGAGTACGAGCCATGAACGCGTTACTGGCCGAA 8340
      AGTCGTTATAGTGGTCGATTTTGGGTATCTCATGCTCGGTACTTGCGCAATGACCGGCTT

                                     ires_emcv reg(8378,8876)>>>
                                     |
8341 GCCGCTTGGAAATAAGGCCGGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGT 8400
      CGGCGAACCTTATTCCGGCCACACGCAACAGATATACAATAAAAGGTGGTATAACGGCA

8401 CTTTTGGCAATCTGAGGGCCCCGAAACCTGCCCTGTCTTCTTGACGAGCATTCCTAGGG 8460
      GAAACCGTTACACTCCCGGCCCTTTGGACCGGGACAGAGAACTGCTCGTAAGGATCCC

8461 GTCTTTCCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGTCGTGAAGGAAGCAGTTC 8520
      CAGAAAGGGGAGAGCGGTTTCCTTACGTTCCAGACAACTTACAGCACTTCCTTCGTCAAG

8521 CTCTGGAAGCTTCTTGAAGACAAACAAGTCTGTAGCGACCCCTTTGCAGGCAGCGGAACC 8580
      GAGACCTTCGAAGAACTTCTGTTTGTGTCAGACATCGCTGGGAAACCTCCGTCGCCCTTGG

8581 CCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAA 8640
      GGGGTGGACCGCTGTCCACGGAGACGCGGTTTTTCGGTGCACATATTCTATGTGGACGTT

8641 AGCGGGCACAACCCCACTGCCACGTTGTGAGTTGGATAGTTGTGGAAGAGTCAAAATGGC 8700
      TCCGCCGTGTTGGGGTCACGCTCAACACTCAACCTATCAACACCTTTCTCAGTTTACCG

8701 TCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTTGTATGG 8760
      AGAGGAGTTGCATAAGTTGTTCCCGACTTCCTACGGGTCTTCCATGGGGTAACATACC

8761 GATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAC 8820
      CTAGACTAGACCCCGGAGCCACGTGTACGAAATGTACACAAATCAGCTCCAATTTTTTTG

```

FIGURA 1D-16

yCD2 (8877, 9353)>>>

|

PstI (8874)

|

8821 GTCTAGGCCCCCGAACCACGGGGACGTGGTTTTTCCTTTGAAAAACACGATTATAAATGG 8880
CAGATCCGGGGGGGCTTGGTCCCCCTGCACCAAAAGGAAACTTTTTCTGCTAATATTTACC

8881 TGACCGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTTACGAGGAGG 8940
ACTGGCGCGCGTACCGGAGGTTTACCCCTAGTPTTCCCGTACCTATAGCGAATGCTCCTCC

8941 CCCTGCTGGGCTACAAGGAGGGCGGGCTGCTATCGGCGGCTGCTGATCAACAACAAGG 9000
GGGACGACCCGATGTTCCCTCCCGCGCACGGATAGCCGCGGACAGACTAGTTGTTGTTCC

9001 ACGGCAGTGTGCTGGGCAGGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCCTGC 9060
TGCCGTCACACGACCCGTCCCCGGTGTGTGTAATCCAGGCTCTCCCGAGGCGGTGGGACG

9061 ACGGCGACATCTCCACCCTCGAGAAGCTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACN 9120
TGCCGCTCTAGAGGTGGGACCTCTTGACACCGTCCGACCTCCCGTTCCACATGTTCTCTGT

9121 CCACCCCTGTACACCAACCTGTCCCCCTTGTGACATGTGTACCGGCGCTATCATCATGTAGG 9180
GGTGGGACATGTGCTGGGACAGGGGAACACTGTACACATGGCGGCGATAGTAGTACATGC

9181 GCATCCCTAGGTGTGTGATCGGCGAGAACGTGAAGTTCAAGTCCAGGGCGAGAACTACC 9240
CCTAGGGATCCACACACTAGCGGCTCTTGCACTTGAAATTGAGTTCCCGCTCTTCATGG

9241 TGCAAAACAGGGGCCACGAGGTGGTGGTTGTTGACCATGACAGGTGTAAAGAGCTGATGA 9300
ACGTTTGGTCCCCGGTGTCCACCACCAACAACCTGCTACTCTCCACATTCTTCCACTACT

NotI (9356)

|

CciNI

|

9301 AGCAGTTTCATCGACGAGAGGCCCTCAGGACTGGTTCCAGGATATCGCGAGTAAGCGGCGG 9360
TCGTCAAGTAGCTGCTCTCCGGAGTCTGACCAAGCTCCTATAGCCGCTCATTCGCGGCG

Región U3 (9405, 9854)>>>

|

9361 CAGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGGAATGAAAGACCCACC 9420
GTCTATTTTATTTTCTAAATAAATCAGAGGTCTTTTCCCCCCTTACTTTCTGGGGTGG

FIGURA 1D-17

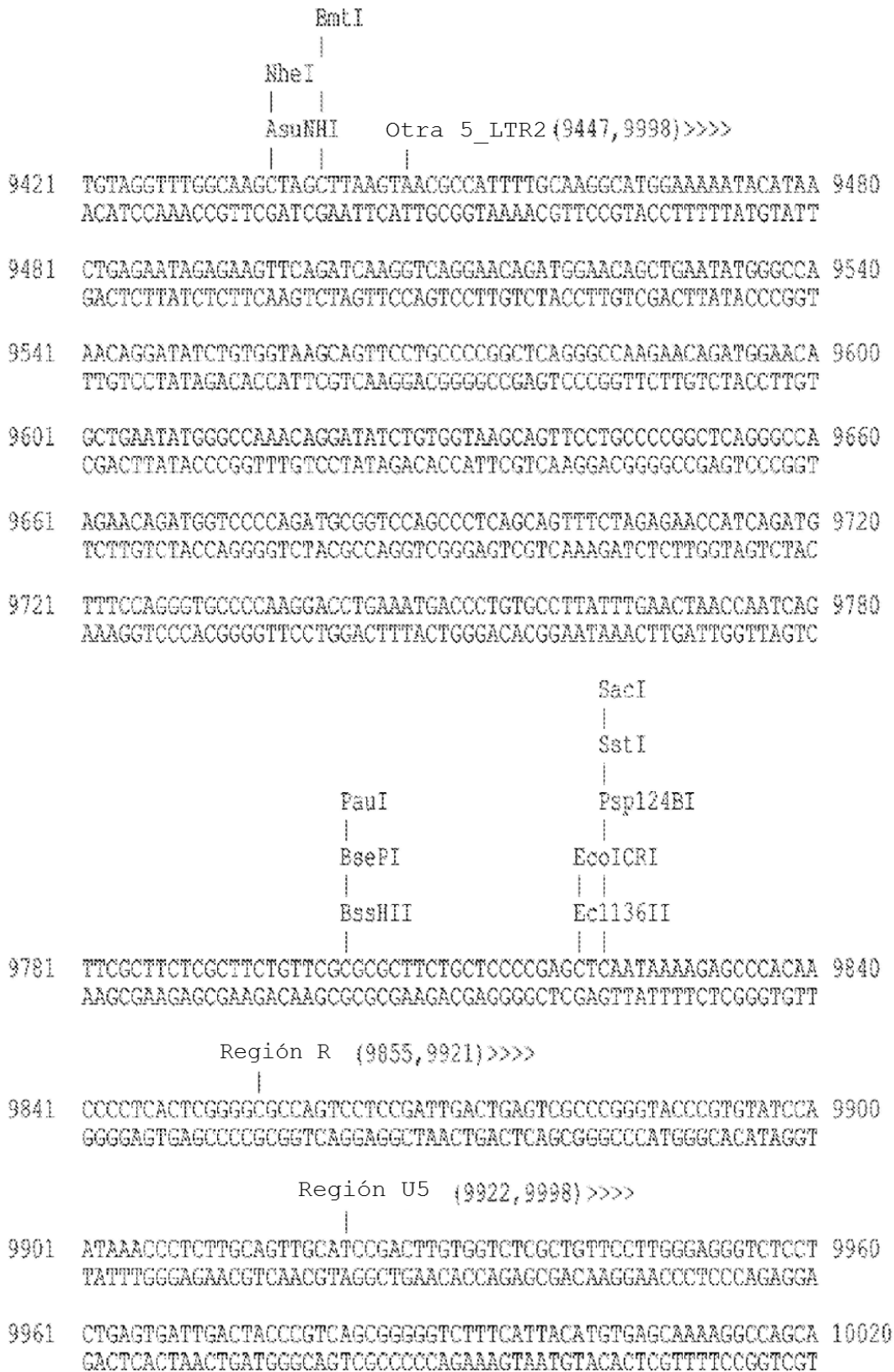


FIGURA 1D-18

Origen pBR322 {10045,10664}<<<
|

```

10021 AAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC 10080
      TTTCCGGTCCTTGGCATTTTTCCGGCGCAACGACCGCAAAAAGGTATCCGAGCGGGGGG

10081 TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA 10140
      ACTGCTCGTAGTGTTTTAGCTGCGAGTTCAGTCTCCACCGCTTTGGGCTGTCTGATAT

10141 AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGGCGCTCTCCTGTTCCGACCCCTGCC 10200
      TTCTATGGTCCGCAAGGGGGACCTTCGAGGGAGCACCGGAGAGGACAAGGCTGGGACGG

10201 GCTTACCGGATACCTGTCCGCCCTTCTCCCTTCGGGAAGCGTGGCGCTTTCCTCAATGCTC 10260
      CGAATGGCCTATGGACAGCGGGAAGAGGGAGCCCTTCGCACCGCGAAAGAGTTACGAG

10261 ACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCCCTCCAAGCTGGGCTGTGTGCACGA 10320
      TGGACATCCATAGAGTCAAGCCACATCCAGCAAGCGAGGTTCCGACCGACACACGTGCT

10321 ACCCCCGCTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC 10380
      TGGGGGGCAAGTCGGCTGCGGACGCGGAATAGGCCATTGATAGCAGAACTCAGGTTGGG

10381 GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 10440
      CCATTCTGTGCTGAATAGCGGTGACCGTCTGCGTGACCATTTGTCCTAATCGTCTCGCTC

10441 GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG 10500
      CATACATCCGCCACGATGTCTCAAGAACTTCACCACCGGATTGATGCCGATGTGATCTTC

10501 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 10560
      CTGTCATAAACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATC

10561 CTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCA 10620
      GAGAACTAGGCGGTTTTGTTTGGTGGCGACCATCGCCACCAAAAAACAAACGTTCTGTCGT

10621 GATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGCTCTGA 10680
      CTAATGCGCGTCTTTTTTCTTAGAGTTCTTCTAGGAACTAGAAAAGATGCCCCAGACT

10681 CGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGAT 10740
      GCGAGTCACCTTGCTTTTGAGTGCAATTCCTAAAACAGTACTCTAATAGTTTTTCCTA

10741 CTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTAAATCAATCTAAAGTATATATGA 10800
      GAAGTGGATCTAGGAAAATTTAATTTTACTTCAAAATTTAGTTAGATTTCTATATATACT
  
```

FIGURA 1D-19

Marcador de amp{10819,11679}<<<
|

10801 GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG 10860
CATTGGAACAGACTGTCAATGGTTACGAATTAGTCACTCCGTGGATAGAGTCGCTAGAC

10861 TCTATTTCTGTTTCATCCATAGTTGCCCTGACTCCCCGTCTGTAGATAACTACGATACGGGA 10920
AGATAAAGCAAGTAGGTATCAACGGACTGAGGGGCAGCACATCTATTGATGCTATGCCCT

10921 GGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACCGGCTCC 10980
CCCGAATGGTAGACCGGGGTCAAGACGTTACTATGGCGCTCTGGGTGCGAGTGGCGGAGG

10981 AGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCCTGCAAC 11040
TCTAAATAGTCGTTATTTGGTGGTCCGGCCTTCCCGGCTCGCGTCTTCACCAGGACGTTG

11041 TTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGC 11100
AAATAGGCGGAGGTAGGTACAGATAATTAACAACGGCCCTTCGATCTCATTTCATCAAGCGG

11101 AGTTAATAGTTTGGCCAAAGTTGTTGCCATTGCTGCAGGCATCGTGGTGTACGCTCGTC 11160
TCAATTATCAAACGCGTTGCAACAACGGTAACGACGTCGGTAGCACCACAGTGGCAGCAG

11161 GTTGTGTATGGCTTCATTACGCTCCGCTTCCCAACGATCAAGCCGAGTTACATGATCCCC 11220
CAAACCATACCGAAGTAAGTCGAGGCCAAGGGTTGCTAGTTCCGCTCAATGTACTAGGGG

11221 CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAGTAAGTT 11280
GTACAACACGTTTTTTTCCCAATCGAGGAAGCCAGGAGGCTAGCAACAGTCTTCATTCAA

11281 GGCCGCAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCTATGCC 11340
CCGGCGTCACAATAGTGAGTACCAATACCGTCTGTGACGTATTAAGAGAATGACAGTACGG

11341 ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 11400
TAGGCATTCTACGAAAGACACTGACCACCTCATGAGTTGGTTCAGTAAGACTCTTATCAC

11401 TATGCGGCGACCGAGTTGCTCTTGGCCGGCGTCAACACGGGATAATACCGCGCCACATAG 11460
ATACGCGCGCTGGCTCAACGAGAACGGGCGCAGTTGTGCCCTATTATGGCGCGGTGTATC

11461 CAGAACTTTAAAAGTGTCTCATCATTGGAAGAGTTCTTCGGGCGGAAACTCTCAAGGAT 11520
GTCTTGAAATTTTACGAGTAGTAACCTTTTGCAGAAGCCCCGCTTTTGAGAGTTCTTA

11521 CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAAGTATCTTCAGC 11580
GAATGGCGACAACCTCTAGGTCAAGCTACATTGGGTGAGCACGTGGGTGACTAGAAAGTGG

11581 ATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGAAA 11640
TAGAAAATGAAAGTGGTCCGAAAGACCCACTCGTTTTTGTCTTCCGTTTTACGGCGTTT

11641 AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATCTCTTCTTTTCAATATTA 11700
TTTCCCTTATTTCCGCTGTGCCTTTACAACCTATGAGTATGAGAAGGAAAAAGTTATAAT

FIGURA 1D-20

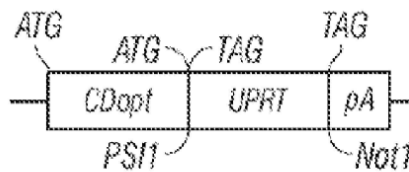
Prom.de amp |{11721,11749}<<<
 |
 11701 TTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAA 11760
 AACTTCGTAAATAGTCCCAATAACAGAGTACTCGCTATGTATAAACTTACATAAATCTT
 11761 AAATAAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTCTAAGA 11820
 TTTATTTGTTTATCCCCAAGGCGCGTGTAAGGGGCTTTTCACGGTGGACTGCAGATTCT
 11821 AACCATTATTATCATGACATTAACTATAAAAAATAGGCGTATCAGAGGCCCTTTTCGTCT 11880
 TTGGTAATAATAGTACTGTAATTGCATATTTTATCCGCATAGTGCTCCGGGAAAGCAGA
 11881 TCAAGAATTCAT 11892
 AGTTCTTAAGTA

FIGURA 1D-21

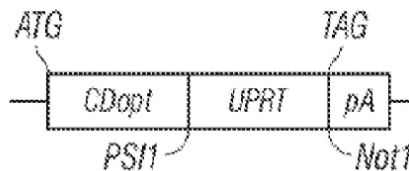
1. Análisis separado de los genes de CDopt y UPRT con sitios PSI1/Not1



2. Digerir con PSI1 y realizar ligado de extremos romos



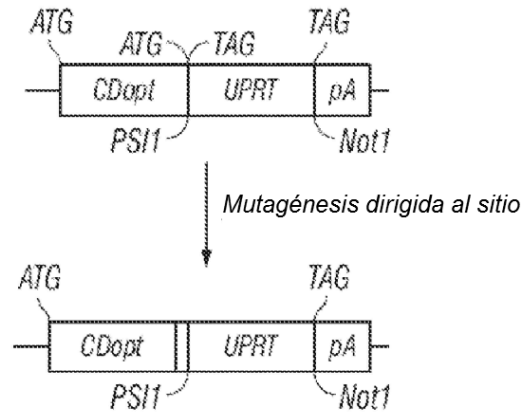
3. Delección del gen codón de TERMINACIÓN del gen de CDopt y el codón de INICIACIÓN de UPRT mediante mutagénesis dirigida a sitio



4. Tamaño del gen de CDopt-UPRT: 1,3 kb

FIGURA 2A

1. Síntesis de CDopt-ENLAZADOR-UPRT: Mutagénesis dirigida a sitio del enlazador



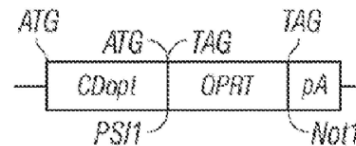
2. Tamaño del gen de UPRT .enlazador CDopt: 1,36 kb

FIGURA 2B

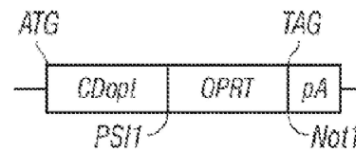
1. Síntesis separada del gen CDopt y el dominio OPRT con sitios PSI1/Not1



2. Digerir con PSI1 y realizar ligado de los extremos romos



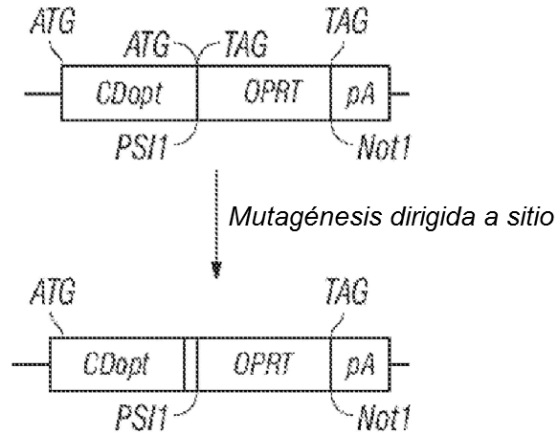
3. Delección del gen de terminación de CDopt y el codón de INICIACIÓN OPRT por mutagénesis dirigida a sitio



4. Tamaño del gen CDopt-OPRT: 1,3 kb

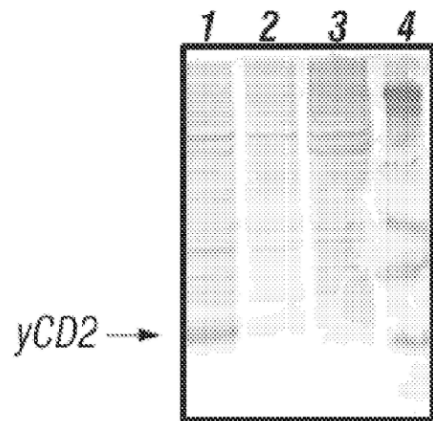
FIGURA 2C

1. Síntesis separada de CDopt-ENLAZADOR-OPRT: Inserción por mutagénesis dirigida a sitio del enlazador



2. Tamaño del gen CD-OPRT: 1,33 kb

FIGURA 2D



1 - U87 + AC3-yCD2 (V)
 2 - U87 + ACE-yCD (V)
 3 - U87 (sin infectar)
 4 - Patrones del peso molecular

FIGURA 3

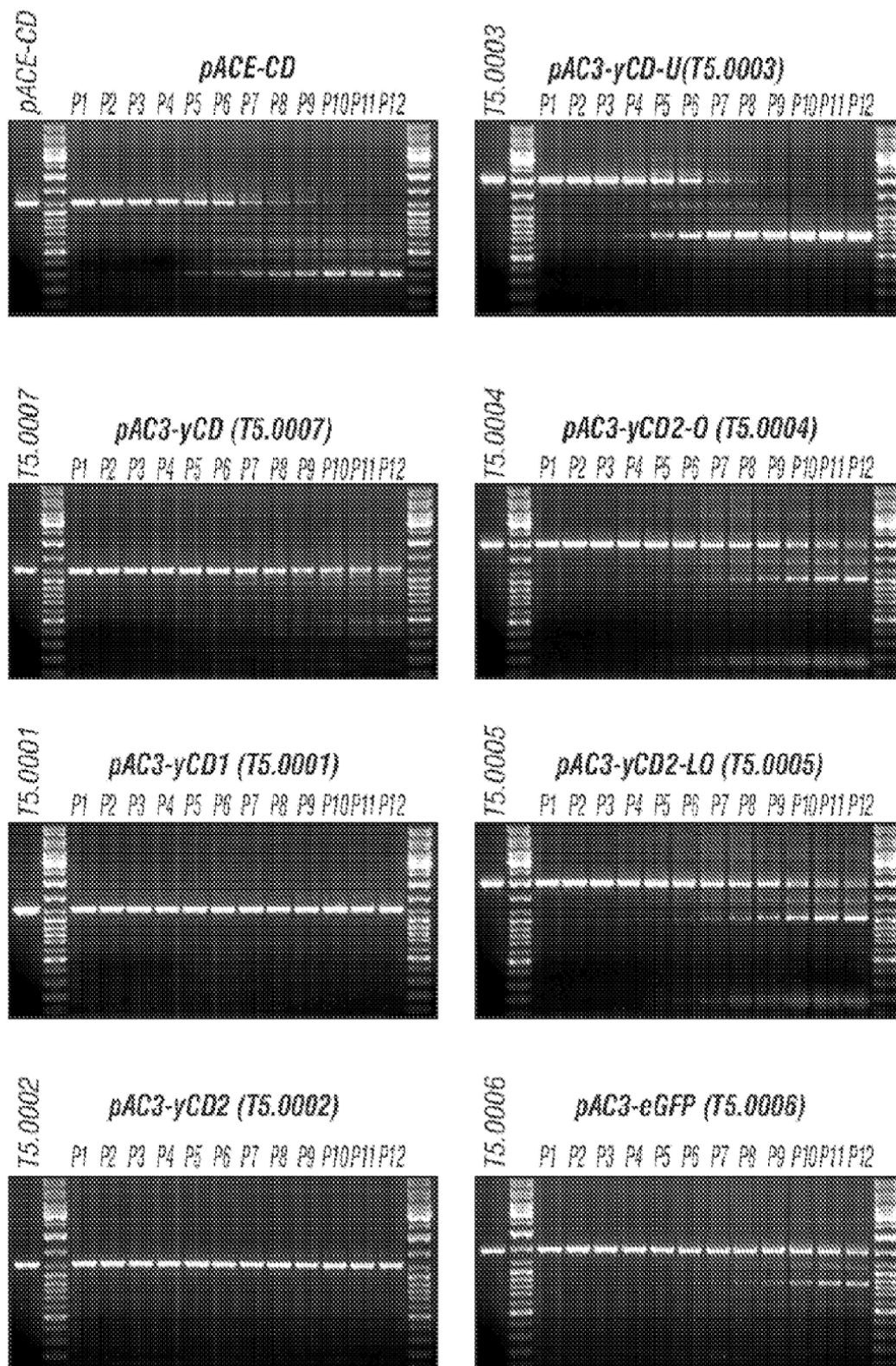


FIGURA 4

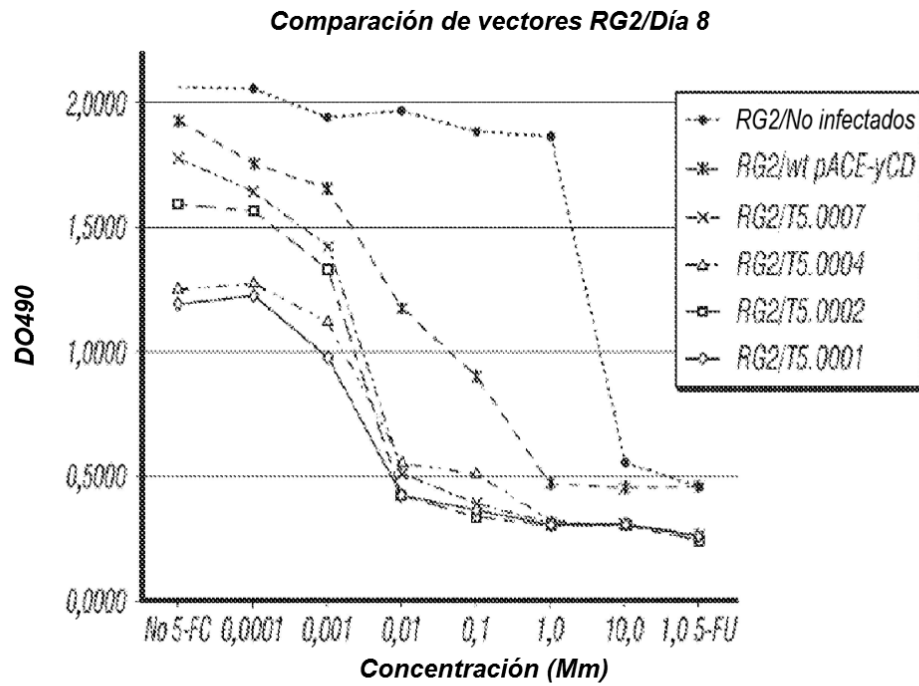


FIGURA 5A

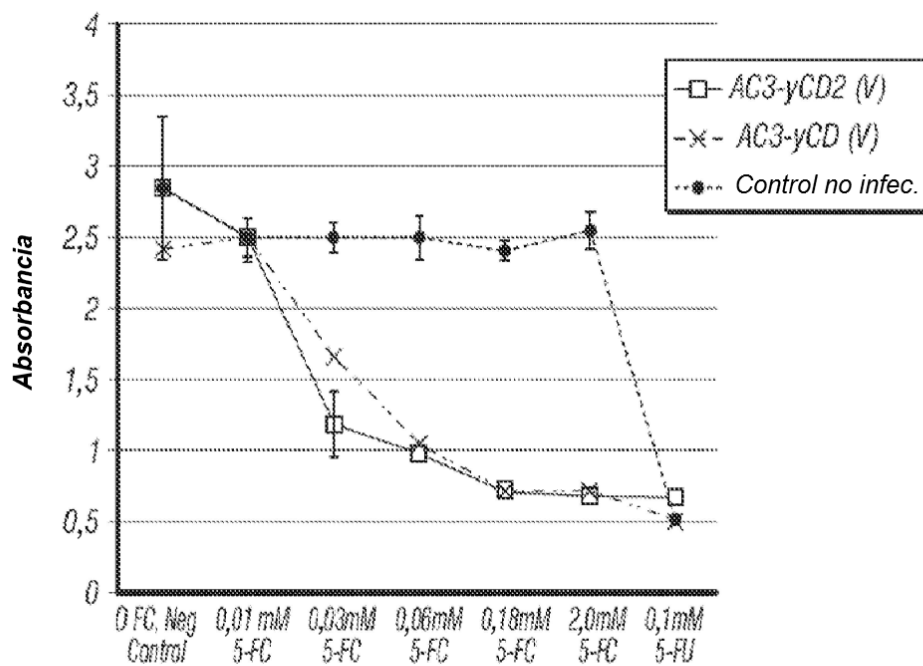


FIGURA 5B

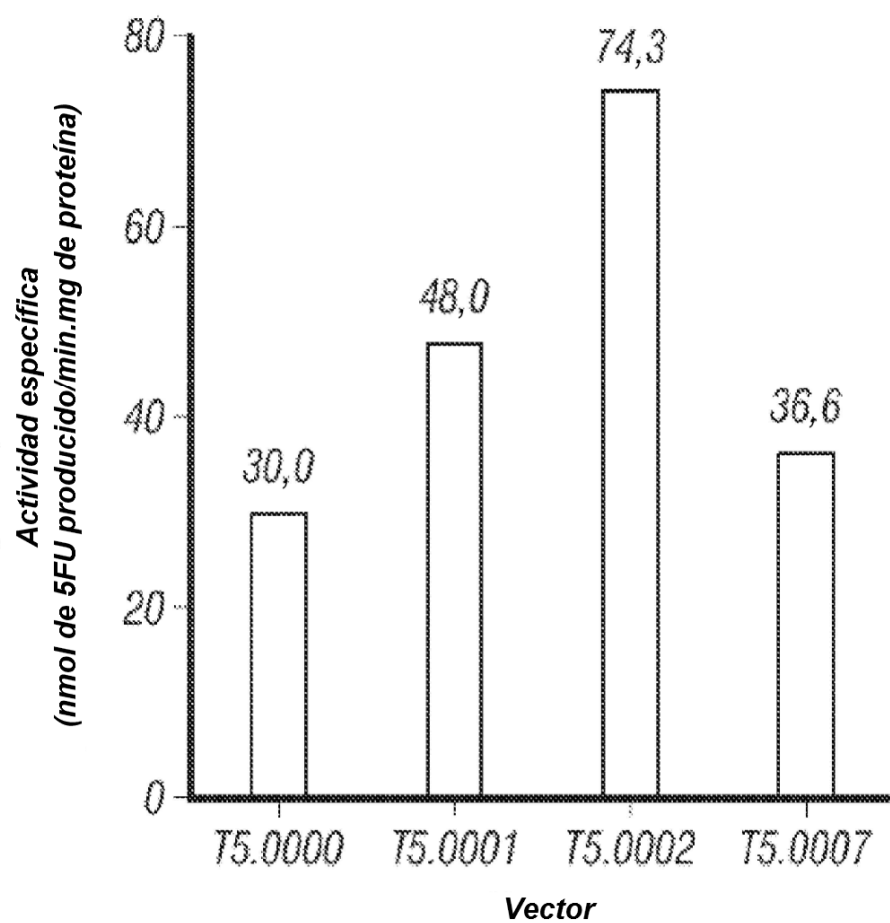


FIGURA 6