



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 358 680**

51 Int. Cl.:
C12N 15/85 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06727026 .4**

96 Fecha de presentación : **09.05.2006**

97 Número de publicación de la solicitud: **1891223**

97 Fecha de publicación de la solicitud: **27.02.2008**

54 Título: **Elementos de expresión mejorados.**

30 Prioridad: **17.05.2005 GB 0509965**
18.05.2005 US 682277 P

45 Fecha de publicación de la mención BOPI:
12.05.2011

45 Fecha de la publicación del folleto de la patente:
12.05.2011

73 Titular/es: **MILLIPORE CORPORATION**
290 Concord Road
Billerica, Massachusetts 01821, US

72 Inventor/es: **Simpson, David John;**
Williams, Steven Geraint y
Irvine, Alistair Simpson

74 Agente: **Polo Flores, Carlos**

ES 2 358 680 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

Elementos de expresión mejorados

ANTECEDENTES

La presente invención se refiere a polinucleótidos
5 que comprenden elementos que confieren una expresión
mejorada a unidades de transcripción unidas de manera
operable. Estos elementos están asociados de manera
natural a las regiones promotoras de genes de proteínas
ribosómicas y, en construcciones de ADN recombinante,
10 confieren niveles de expresión génica elevados y
reproducibles. La presente invención también se refiere a
vectores que comprenden dichas secuencias de
polinucleótidos, a células hospedadoras que comprenden
dichos vectores y al uso de dichos polinucleótidos,
15 vectores o células hospedadoras en terapia, para la
producción de proteínas recombinantes en cultivos
celulares y otras aplicaciones biotecnológicas.

El modelo actual de la estructura de la cromatina en
eucariotas superiores postula que los genes están
20 organizados en "dominios" (Dillon, N. & Grosveld, F.
Chromatin domains as potential units of eukaryotic gene
function. Curr. Opin. Genet. Dev. 4, 260-264 (1994);
Higgs, D.R. Do LCRs open chromatin domains? Cell 95, 299-
302 (1998)). Se contempla que los dominios de la cromatina
25 se encuentren en un estado condensado, "cerrado",
transcripcionalmente silente, o en una configuración
descondensada, "abierta" y transcripcionalmente
competente. El establecimiento de una estructura abierta
de la cromatina caracterizada por una sensibilidad a la
30 DNasa I, una hipometilación del ADN y una hiperacetilación
de las histonas incrementadas se considera un requisito
previo para el comienzo de la expresión génica.

La naturaleza abierta y cerrada de regiones de la
cromatina está reflejada en el comportamiento de
35 transgenes que se integran aleatoriamente en el genoma de
la célula hospedadora. Construcciones idénticas

proporcionan diferentes patrones de expresión específica de tejido y de expresión específica de la fase de desarrollo cuando se integran en localizaciones diferentes del genoma de ratón (Palmiter, R.D. & Brinster, R.L. Ann. Ref. Genet. 20, 465-499 (1986); Allen, N.D. y col. Nature 333, 852-855 (1988); Bonnerot, C., Grimber, G., Briand, P. & Nicolas, J.F. Proc. Natl. Acad. Sci. USA 87:6331-6335 (1990)).

El modelo de organización génica del dominio de la cromatina sugiere que los elementos de control genético que son capaces de establecer y mantener una estructura abierta de la cromatina transcripcionalmente competente deben estar asociados a regiones activas del genoma.

Las Regiones para el control del locus (LCRs) son una clase de elementos reguladores transcripcionales con una capacidad de reestructuración de la cromatina de largo alcance. Las LCR se definen funcionalmente en ratones transgénicos por su capacidad para conferir niveles de expresión fisiológicos que dependen del número de copias del transgen y que es independiente del sitio de integración sobre un gen unido en *cis*, especialmente en transgenes de una sola copia (Fraser, P. & Grosveld, F. Curr. Opin. Cell Biol 10, 361-365 (1998); Li, Q., Harju, S. & Peterson, K.R. Trends Genet. 15: 403-408 (1999). De manera crucial, dicha expresión es específica de tejido. Las LCRs son capaces de impedir la dispersión de la heterocromatina, evitar la PEV (Kloussis, D. & Festenstein, R. Curr. Opin. Genet. Dev. 7, 614-619 (1997)) y constan de una serie de sitios hipersensibles (HS) a la DNasa I que pueden estar localizados en 5' o 3' de los genes que regulan (Li, Q., Harju, S. & Peterson, K.R. Trends Genet. 15: 403-408 (1999)).

La generación de líneas celulares de mamífero en cultivo que producen niveles elevados de un producto proteico terapéutico es una gran industria en desarrollo. Los efectos de la posición de la cromatina lo convierten

en un proceso difícil, que requiere tiempo y caro. El enfoque más habitual usado para la producción de dichas "factorías celulares" de mamífero depende de la amplificación génica inducida por una combinación de un gen de resistencia a fármacos (por ejemplo, DHFR, glutamina sintetasa (Kaufman RJ. *Methods Enzymol* 185, 537-566 (1990)), y el mantenimiento de una presión rigurosa y selectiva. El uso de vectores que contienen LCRs procedentes de dominios génicos muy expresados, usando células derivadas del tejido adecuado, simplifica enormemente el procedimiento, dando una elevada proporción de líneas celulares clonales que presentan unos niveles de expresión elevados y estables (Needham M, Gooding C, Hudson K, Antoniou M, Grosveld F y Hollis M. *Nucleic Acids Res* 20, 997-1003 (1992); Needham M, Egerton M, Millett A, Evans S, Popplewell M, Cerillo G, McPheat J, Monk A, Jack A, Johnstone D y Hollis M. *Protein Expr Purif* 6,124-131 (1995).

No obstante, la especificidad de tejido de las LCR, aunque es útil en algunas circunstancias, también es una limitación importante para muchas aplicaciones, por ejemplo, cuando no se conocen LCR para el tejido en el que se requiere la expresión, o cuando se requiere la expresión en muchos, o en todos los tejidos.

El documento US 6.689.606 y la solicitud de patente copendiente WO 00/0539, incorporadas en el presente documento por referencia, describen elementos que son responsables, en su contexto cromosómico natural, del establecimiento de una estructura de cromatina abierta a lo largo de un locus que consta exclusivamente de genes domésticos que se expresan de manera ubicua. Estos elementos no proceden de una LCR y comprenden islas CpG extendidas exentas de metilación.

En ADN de mamífero, el dinucleótido CpG es reconocido por una enzima ADN-metiltransferasa que metila la citosina a 5-metilcitosina. No obstante, la 5-metilcitosina es

inestable y se convierte en timina. Como resultado, los dinucleótidos CpG se producen mucho menos frecuentemente de lo que cabría esperar por azar. Sin embargo, algunas secciones de ADN genómico presentan una frecuencia de CpG
 5 cercana a la esperada, y estas secuencias se conocen como "islas CpG". Como se usa en el presente documento una "isla CpG" se define como una secuencia de ADN, de al menos 200 pb, que tiene un contenido en GC de al menos el 50% y una relación de contenido en CpG observada/esperada
 10 de al menos 0,6 (es decir, un contenido en dinucleótidos CpG de al menos el 60% de lo que cabría esperar por azar) (Gardiner-Green M y Frommer M. J Mol Biol 196, 261-282 (1987); Rice P, Longden I y Bleasby A Trends Genet 16, 276-277 (2000)).

15 Las islas CpG exentas de metilación son muy conocidas en la materia (Bird y col. (1985) Cell 40: 91-99, Tazi y Bird (1990) Cell 60: 909-920) y se pueden definir como islas CpG en las que una proporción sustancial de los restos de citosina no están metilados y que normalmente se
 20 extienden más allá de los extremos 5' de dos genes transcritos de manera divergente muy próximos en el espacio (0,1-3 kb). Se ha informado de que estas regiones de ADN permanecen hipometiladas en todos los tejidos durante el desarrollo (Wise y Pravtcheva (1999) Genomics
 25 60: 258-271). A menudo están asociadas a los extremos 5' de genes expresados de manera ubicua, así como en un 40% de los genes que presentan un perfil de expresión restringido a determinados tejidos (Antequera, F. & Bird, A. Proc. Natl. Acad. Sci. USA 90, 1195-11999 (1993);
 30 Cross, S.H. & Bird, A.P. Curr. Opin, Genet. Dev. 5, 309-314 (1995)) y son conocidas por tener regiones localizadas de cromatina activa (Tazi, J. & Bird, A. Cell 60, 909-920 (1990)).

Una isla CpG "extendida" exenta de metilación es una
 35 isla CpG exenta de metilación que se extiende a través de una región que engloba más de un sitio de inicio de la

transcripción y/o que se extiende más de 300 pb y preferentemente más de 500 pb. Los límites de la isla CpG extendida exenta de metilación se definen funcionalmente con el uso de la PCR sobre la región, en combinación con
 5 enzimas endonucleasas de restricción, cuya capacidad para digerir (cortar) el ADN en su secuencia de reconocimiento es sensible al estado de metilación de cualquier resto de CpG que haya presente. Una de esas enzimas es la HpaII, que reconoce y digiere el sitio CCGG, que se encuentra
 10 habitualmente dentro de las islas CpG, pero sólo si los restos de CG centrales no están metilados. Por tanto, la PCR llevada a cabo con ADN digerido con HpaII y sobre una región que alberga sitios HpaII, no proporciona un producto de amplificación debido a la digestión de HpaII
 15 si el ADN está sin metilar. La PCR sólo proporcionará un producto amplificado si el ADN está metilado. Por tanto, más allá de la región exenta de metilación el HpaII no digerirá el ADN y se observará un producto amplificado por la PCR, definiendo así los límites de la "isla CpG
 20 extendida exenta de metilación".

Se ha demostrado (documento WO 00/05393) que regiones que se extienden sobre islas CpG exentas de metilación que engloban promotores dobles transcritos de manera divergente desde los loci de los genes de la proteína de
 25 unión a la caja TATA humana (TBP)/proteosoma componente-B1 (PSMBI) y la ribonucleoproteína nuclear heterogénea A2/B1 (hnRNPA2)/heterocromatina proteína 1H• (HP1^{HS}) proporcionan niveles fisiológicos reproducibles de expresión génica y que son capaces de prevenir un patrón de expresión veteadado
 30 y el silenciamiento que se produce normalmente con la integración de transgenes dentro de la heterocromatina centromérica.

Se sabe que las islas CpG exentas de metilación asociadas a promotores que se transcriben activamente
 35 poseen la capacidad de reestructurar la cromatina y así, se cree que son un determinante primordial en el

establecimiento y el mantenimiento de un dominio abierto en loci de genes domésticos (documento WO 00/05393) y que esos elementos confieren un aumento en la proporción de eventos de consecución de genes productivos con mejoras en el nivel y la estabilidad de la expresión de transgenes.

Los ribosomas son grandes complejos de ARN y proteínas responsables de la traducción del ARNm en polipéptidos. Cada ribosoma está constituido de 4 moléculas de ARN ribosómico (ARNr) y un gran número de proteínas ribosómicas (actualmente se piensa que son 79 en las células de mamífero). Las proteínas ribosómicas tienen funciones que incluyen el facilitar el plegamiento del ARNr, la protección frente a las ribonucleasas celulares, y la coordinación de la síntesis de proteínas. Algunas proteínas ribosómicas presentan funciones extrarribosómicas adicionales (Wool, 1996, TIBS 21: 164-165). Dadas las similitudes estructurales y funcionales de los ribosomas entre especies, no sorprende que la conservación de la secuencia de aminoácidos de las proteínas ribosómicas sea elevada, y entre los mamíferos, las secuencias de la mayoría de proteínas ribosómicas son casi idénticas (Wool y col., 1995, Biochem Cell Biol 73: 933-947).

Hay dos proteínas ribosómicas que parecen atípicas, puesto que se expresan en forma de propéptidos (proteínas con extensión carboxi) fusionados a ubiquitina. La ubiquitina es un polipéptido de 76 aminoácidos muy conservado involucrado en una variedad de funciones celulares, incluyendo la regulación de la descomposición intracelular de proteínas, la regulación del ciclo celular y la respuesta al estrés (Hershko & Ciechanover, 1992, Annu Rev Biochem 61: 761-807; Coux y col., 1996, Annu Rev Biochem 65: 801-847).

La ubiquitina está codificada por dos clases de genes diferentes. Uno es un gen de poli-ubiquitina que codifica un polímero lineal de repeticiones de ubiquitina. El otro

comprende genes que codifican proteínas de fusión naturales en las que una sola molécula de ubiquitina está unida a la proteína ribosómica rps27A o rpL40 (Finley y col., 1989, Nature 338: 394-401; Chan y col., 1995, Biochem Biophys Res Commun 215: 682-690; Redman & Burris, 1996, Biochem J 315: 315-321).

Las características estructurales comunes de los promotores de proteínas ribosómicas han sido descritas por Perry (2005, BMC Evolutionary Biology 5: 15). Los promotores se pueden clasificar según la naturaleza de los motivos de la caja TATA, el número y el tipo de sitios de unión a factores de transcripción y la localización de los codones de iniciación AUG. No obstante, esa clasificación no parece predecir la fuerza del promotor y hay evidencias que sugieren que algunos de esos promotores probados presentan una actividad transcripcional equivalente a la medida para la expresión de un gen informador unido (Hariharan y col., 1989, Genes Dev 3: 1789-800).

La patente de EE.UU. 6.063.598 describe el promotor de la ubiquitina de hámster/S27a y su uso para conseguir un nivel de producción elevado de proteínas recombinantes. No obstante, no hace ninguna sugerencia en cuanto a su uso en la mejora de la expresión de un gen transcrito principalmente a partir de otro promotor (es decir, uno distinto del promotor de la ubiquitina de hámster/S27a).

La solicitud de EE.UU. US 2004/0148647 describe un ensayo informador usando un vector de expresión que comprende un promotor de la ubiquitina de hámster/S27a unido de manera funcional a un gen para un producto de interés y una proteína informadora fluorescente. De nuevo, la aplicación sólo describe construcciones en las que la transcripción del gen de interés se produce a partir del propio promotor de la ubiquitina de hámster/S27a.

Sigue siendo un objetivo en materia de expresión de genes recombinantes la obtención de niveles de expresión más elevados y más exactos, particularmente para

aplicaciones terapéuticas *in vivo* y *ex vivo* y para la producción de proteínas recombinantes *in vitro*.

RESUMEN DE LA INVENCIÓN

A lo largo de la descripción y las reivindicaciones de esta solicitud, las palabras "comprenden" y "contienen" y sus variaciones, por ejemplo, "que comprende" y "comprende", significa "que incluye, pero no limitado a", y no se pretende que se excluyan (y no se hace) otros restos, aditivos, componentes, pasos o etapas.

10 Definiciones

Como se usa en el presente documento, una región promotora se define como una secuencia genómica de nucleótidos que consta de un promotor y un sitio de inicio de la transcripción junto con 5 kb de la secuencia en 5' 15 aguas arriba del sitio de inicio de la transcripción y 500 pb de la secuencia en 3' aguas abajo del extremo distal del primer exón.

Una región en 5' sin traducir significa una región en 5' del inicio de la traducción codificada en la secuencia 20 genómica o de ADNc. Se toma para que incluya todos los elementos reguladores en la dirección 5'. Una secuencia en 5' aguas arriba se usa para referirse a una secuencia en 5' al inicio de la transcripción codificado en la secuencia genómica.

25 Como se usa en el presente documento, "ácido nucleico transcribible" significa un ácido nucleico que, cuando está unido de manera operable a un promotor funcional u otras secuencias reguladoras, puede ser transcrito para dar una molécula de ARN funcional, tal como ARNm. Estas 30 secuencias pueden comprender marcos de lectura abiertos, que codifican secuencias de polipéptidos traducibles. Alternativamente, el ARN funcional puede tener otra función, tal como ARN ribosómico, ribozimas o ARN no codificante.

35 "Gen" normalmente significa la combinación de la región codificante del ácido nucleico transcribible, junto

con el promotor a partir del cual se transcribe y otras
secuencias reguladoras tales como potenciadores y señales
de poliadenilación en 3'. Los genes de ADN genómico
también contienen intrones. "Unidad de transcripción" a
5 veces se usa para describir una combinación funcional de
al menos un promotor y secuencias reguladoras mínimas con
un ácido nucleico transcribible, a menudo derivado de ADNc
a partir del cual se han procesado los intrones. "Cistrón"
se define como un ácido nucleico que codifica un único
10 polipéptido junto con las señales funcionales de
iniciación y terminación. "Transgen" implica un gel que ha
sido transferido de un genoma a otro, aunque el término se
puede aplicar más libremente a cualquier gen o incluso a
un ácido nucleico transcribible comprendido en una
15 construcción de ADN recombinante tal como un vector.

Promotor y potenciador son términos muy conocidos en
la materia e incluyen las siguientes características que
se proporcionan sólo a modo de ejemplo, y no a modo de
limitación. Los promotores son secuencias reguladoras en
20 5' que actúan en *cis* unidas directamente al sitio de
inicio de la transcripción. Los elementos promotores
incluyen las secuencias denominadas caja TATA y la
selección de iniciación de la ARN polimerasa (RIS) cuya
función es seleccionar un sitio de inicio de la
25 transcripción. Estas secuencias también se unen a
polipéptidos cuya función, entre otras, es facilitar la
selección del inicio de la transcripción por parte de la
ARN polimerasa.

En términos sencillos, los promotores son elementos
30 direccionales que actúan para iniciar la transcripción de
secuencias situadas a menos de 100 (y normalmente a menos
de 50) pares de bases (pb) de nucleótidos en dirección 3'.
Contienen una serie de secuencias cortas de nucleótidos
consenso que actúan como sitios de unión para diversas
35 proteínas que participan en el inicio de la transcripción
y en el ensamblaje de un complejo multi-subunidad conocido

como complejo de pre-iniciación (McKnight y Tjian, 1987, Cell 46: 795-805). En la mayoría de los genes, esto se produce en una secuencia muy ampliamente conservada conocida como caja TATA (TATAAA) a la cual se une la
 5 proteína de unión a la caja TATA (TBP, una subunidad del factor de transcripción general TFIID). A continuación se produce el ensamblaje ordenado de más de 10 factores de transcripción adicionales para formar en último término el complejo de la holoenzima Pol II. La transcripción del ARN
 10 en realidad comienza en el sitio de iniciación aproximadamente 25-30 bases en dirección 3' (Breathnach y Chambon, 1981, Annu Rev Biochem 50: 349-393) al cual también se une la TBP.

La mayoría de promotores funcionales contienen
 15 elementos promotores adicionales en dirección 5' (UPEs), de los cuales los más conservados son la caja CAAT (CCAAT, el sitio de unión para los factores de transcripción CBF, C/EBP y NF-1), aproximadamente 70-200 pb en dirección 5', y la caja GC (GGGCGG, sitio de unión para el factor de
 20 transcripción general Sp-1) a una distancia similar en dirección 5'. Aunque se producen niveles basales de transcripción con la caja TATA sola, para la mayoría de promotores son necesarias al menos las cajas CAAT y GC para unos niveles óptimos de transcripción.

Los potenciadores son secuencias que actúan no
 25 direccionalmente para incrementar la transcripción de promotores situados localmente, pero no de manera necesaria inmediatamente adyacentes (hasta varias kilobases de distancia (Kadonaga (2004) Cell 116: 247-
 30 257)). Los potenciadores contienen secuencias consenso cortas (8-12 pb) que representan los sitios de unión para una amplia variedad de proteínas activadoras de la transcripción (Ondek y col., 1988, Science 236: 1237-1244) incluyendo algunas, tales como NF-1 y SP-1 que también
 35 están asociadas a elementos promotores. A menudo estas secuencias están duplicadas en tándem o en repeticiones

invertidas.

En algunas unidades de transcripción naturales, incluyendo las unidades de transcripción génica inmediata/temprana muy activas de muchos virus de ADN
 5 tales como citomegalovirus, los elementos potenciadores y promotores se pueden combinar funcionalmente en lo que es efectivamente un elemento extendido en dirección 5'.

Los promotores pueden estar regulados, que responden al tipo celular, a la temperatura, a iones metálicos u
 10 otros factores; o ser constitutivos, que producen una transcripción que no responde a esos factores. Para muchos propósitos es muy ventajoso un promotor constitutivo que dé niveles de transcripción elevados y consistente en muchos, si no en todos, los tipos celulares. Durante
 15 muchos años, el elemento potenciador/promotor que conduce la expresión génica inmediata/temprana en citomegalovirus humano se ha usado muy ampliamente para conducir la expresión de genes heterólogos en vectores de expresión eucariotas (Foecking & Hoffstetter, 1986, Gene 45: 101-
 20 105).

Se propuso la hipótesis de que regiones promotoras de genes de proteínas ribosómicas podrían tener una actividad útil en la potenciación y estabilización de la expresión de transgenes unidos y que las regiones reguladoras
 25 procedentes de genes muy expresados muy probablemente podrían contener elementos que son muy eficaces en el mantenimiento de la cromatina en una conformación transcripcionalmente activa. La unión de dichos elementos a un promotor heterólogo podría generar por tanto un
 30 entorno de la cromatina más abierto en los alrededores del promotor, lo que daría como resultado una expresión incrementada. Alguien experto en la materia entenderá que los promotores derivados de genes de proteínas ribosómicas se deben distinguir de los promotores ribosómicos, que son
 35 promotores que dependen de la ARN-polimerasa de tipo I a partir de la cual se transcribe el ARNr.

Para comprobar esta hipótesis, se obtuvo ARN
 procedente de líneas celulares CHO-K1 y NS0 en la fase de
 crecimiento exponencial y se llevó a cabo un análisis de
 micromatrices frente a 13.443 genes murinos. Limitamos
 5 nuestro análisis a elementos que tienen un contenido
 elevado en islas CpG y una probabilidad de promotores
 bidireccionales. Usando criterios basados en la secuencia
 efectiva mínima de la región reguladora hnRNPA2, se
 amplificaron por PCR aproximadamente 3 kb de ADN
 10 procedente de una selección de estos genes a partir de ADN
 genómico de NS0. A continuación, estas secuencias se
 clonaron en vectores de expresión EGFP y se transfectaron
 a células CHO-K1 junto con las versiones control de
 hnRNPA2 del mismo vector.

15 Se encontró que las secuencias derivadas de las
 regiones promotoras de dos proteínas ribosómicas
 proporcionaban de manera consistente niveles de expresión
 elevados de la secuencia informadora heteróloga en el
 ensayo utilizado. En cada caso, la región promotora
 20 comprendía una secuencia rica en GC que se extiende desde
 la región en 5' aguas arriba de los elementos promotores
 reales hasta bien entrado en el primer exón y, de hecho,
 en el primer intrón. Esta secuencia rica en GC cumplía los
 criterios de ser una isla CpG extendida, como se define en
 25 el presente documento puesto que se extiende más de 300
 pb.

Por consiguiente, la invención proporciona un
 polinucleótido aislado que comprende:

- a. una isla CpG extendida exenta de metilación que
 30 comprende más de 3000 nucleótidos contiguos desde la
 región promotora de un gen rps3 de mamífero;
- b. un promotor heterólogo; y
- c. una secuencia de ácidos nucleicos transcribible
 adyacente al promotor heterólogo,
- 35 en el que la transcripción de dicha secuencia de ácidos
 nucleicos transcribible desde dicho promotor heterólogo se

mejora en presencia de dicha isla CpG extendida exenta de metilación con respecto a la transcripción en ausencia de dicha isla CpG extendida exenta de metilación.

Los nucleótidos contiguos se seleccionan del área del
 5 promotor, que se extiende desde un punto 5 kb aguas arriba
 (en dirección 5' con respecto a la hebra codificante) del
 sitio de inicio de la transcripción hasta un punto 500 pb
 aguas abajo (en dirección 3' con respecto a la hebra
 codificante) del extremo distal (3') del primer exón.
 10 También se prefiere que dicha isla CpG comprenda el
 promotor del gen rps3 y desde el cual se inicia
 naturalmente la transcripción del gen rps3. Ese promotor,
 a menudo se denomina promotor endógeno. En una forma de
 realización preferida, la isla CpG comprende
 15 adicionalmente uno o varios exones de dicho gen rps3.

El gen rps3 es un gen de mamífero, aunque esos genes
 y sus promotores y las secuencias 5' aguas arriba están
 muy conservadas entre especies, y alternativamente también
 pueden estar descritos genes de insectos, nematodos y
 20 levaduras. No obstante, preferentemente es un gen humano o
 de roedor y lo más preferentemente es un gen de ratón.

En una forma de realización muy preferida, el
 polinucleótido aislado de la invención comprende los
 nucleótidos 38 a 3154 de la secuencia de nucleótidos rps3
 25 de ratón representada en la SEQ ID NO: 1. También se
 describen los nucleótidos 12 a 3032 de la secuencia de
 nucleótidos rps11 de ratón como se presenta en la SEQ ID
 NO: 2.

En un aspecto, además del elemento descrito, el
 30 polinucleótido comprende adicionalmente un promotor no
 asociado de manera natural a dicho elemento procedente de
 un gen de una proteína ribosómica. En esta forma de
 realización, un promotor heterólogo (distinto del promotor
 endógeno que puede o puede no estar presente en el primer
 35 elemento) está situado en una posición adyacente y unido
 de manera operable en dirección 3' del elemento que

contiene la secuencia en 5' derivada del gen de la proteína ribosómica. En esta disposición, la expresión dirigida por este promotor heterólogo se mejora por el efecto del elemento del gen de la proteína ribosómica.

5 En una forma de realización, dicho promotor es un promotor constitutivo, más preferentemente seleccionado de la lista constituida por el promotor temprano/inmediato de citomegalovirus, SV40, EF-1 α , las LTR del virus del sarcoma de Rous (RSV), o las LTR del VIH2 o combinaciones
10 de secuencias derivadas de ellos. Más preferentemente, el promotor es un promotor inmediato/temprano de CMV. Más preferentemente, es el promotor inmediato/temprano de CMV en ratón o conejillo de indias.

Alternativamente, dicho promotor puede ser un
15 promotor específico de tejido, que dirige la expresión en un espectro limitado de tejidos. Dichos promotores son muy conocidos en la materia e incluyen aquellos derivados de •-globina, las cadenas ligeras de inmunoglobulinas • y •, la cadena pesada de inmunoglobulinas, desmina, tirosinasa,
20 CD2, IL-3, la cadena ligera de la miosina, el promotor del gen de la actividad inhibitoria del melanoma humano y queratinas. En una forma de realización particularmente preferida el promotor es un promotor selectivo de tumores, que dirige preferentemente la expresión de uno o más tipos
25 tumorales. Ejemplos de dichos promotores incluyen los promotores del antígeno carcinoembrionario (CEA), antígeno prostático específico (PSA), ciclooxigenasa-2 (COX-2), alfa-fetoproteína (AFP), tirosinasa, y los promotores basados en los factores 1-4 de linfocitos T (TCF).

30 El ácido nucleico transcribible puede codificar cualquier polipéptido útil para la expresión *in vitro* y preferentemente se selecciona de la lista constituida por un anticuerpo, un fragmento funcional de un anticuerpo que se une a un epítipo, un factor de crecimiento, una
35 citoquina, una proteína quinasa, un receptor soluble, un receptor unido a membrana, o un factor de coagulación

sanguíneo. Alternativamente, el ácido nucleico transcribible puede codificar un gen terapéutico para su uso en terapia génica *in vivo* o *ex vivo*. Dicho ácido nucleico terapéutico puede actuar sustituyendo o
5 suplementando la función de un gen defectuoso que provoca una enfermedad como la fibrosis quística, talasemia, anemia falciforme, anemia de Fanconi, hemofilia, inmunodeficiencia combinada severa (SCID), fenilcetonuria (PKU), deficiencia en antitripsina alfa-1, distrofia
10 muscular de Duchenne, deficiencia en ornitina transcarbamilasa u osteogénesis imperfecta. Alternativamente, puede codificar un agente citotóxico o una enzima convertidora de un profármaco expresada selectivamente en una célula diana, como una célula de
15 cáncer maligno, con el fin de acabar con él. Estas aplicaciones, y muchas otras, son muy conocidas por aquellos expertos en la materia y la importancia de la presente invención en la mejora de la expresión de los ácidos nucleicos terapéuticos será evidente para dichos
20 profesionales cualificados.

En otro aspecto, la invención proporciona un vector que comprende el polinucleótido de la invención como se describe anteriormente. Preferiblemente, dicho vector es un vector de expresión adaptado para la expresión de genes
25 eucariotas.

Normalmente, dicha adaptación incluye, a modo de ejemplo y no de limitación, la provisión de secuencias para el control de la transcripción (secuencias promotoras) que median en la expresión específica de
30 células/tejido. Las adaptaciones también incluyen la provisión de marcadores seleccionables y secuencias de replicación autónomas que facilitan el mantenimiento de dicho vector en la célula eucariota o en el hospedador procariota. Los vectores que se mantienen de manera
35 autónoma se denominan vectores episomales. Se prefieren los vectores episomales puesto que son auto-replicativos y

de esta forma persisten sin necesidad de integración. Vectores episomales de este tipo se describen en el documento WO 98/07876.

Las adaptaciones que facilitan la expresión de genes
 5 codificados en vectores incluye la provisión de secuencias de terminación de la transcripción/poliadenilación. Esto también incluye la provisión de sitios internos de entrada de ribosomas (IRES) cuya función es maximizar la expresión de los genes codificados en los vectores dispuestos en
 10 casetes de expresión bicistrónicos o multi-cistrónicos.

Estas adaptaciones son muy conocidas en la materia. Hay una cantidad importante de bibliografía publicada con respecto a la construcción de vectores de expresión y a las técnicas de ADN recombinante, en general. Por favor,
 15 véase, Sambrook y col. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY y las referencias allí citadas; Marston, F (1987) DNA Cloning Techniques: A Practical Approach Vol III IRL Press, Oxford UK; DNA Cloning: F M
 20 Ausubel y col., Current Protocols In Molecular Biology, John Wiley & Sons, Inc.(1994).

El vector puede ser un vector episomal o un vector de integración. Preferentemente, el vector es un plásmido. Alternativamente, el vector puede ser un virus, tal como
 25 un adenovirus, virus asociado a adenovirus, un herpesvirus, virus de la vacuna, un lentivirus, u otro retrovirus.

Alternativamente, ese vector puede comprender:

a. un elemento que comprende más de 3000 nucleótidos
 30 contiguos desde la región promotora de un gen rps3 de mamífero;

b. un promotor heterólogo; y

c. un sitio de clonación múltiple,
 en el que una secuencia de ácidos nucleicos transcribible
 35 insertada en el sitio de clonación múltiple puede ser transcrita desde el promotor heterólogo y el nivel de

transcripción se mejora mediante el elemento.

En un aspecto adicional, la invención proporciona una célula hospedadora que comprende el polinucleótido o el vector aislado como se describe en el presente documento.

5 Preferentemente, dicha célula hospedadora es una célula de mamífero, más preferentemente seleccionada de la lista constituida por células CHO, NS0, BHK, HeLa, HepG2.

Además, la invención proporciona un procedimiento de expresión de un polipéptido que comprende la inserción de un vector de expresión que contiene el polinucleótido de la invención en una célula hospedadora apropiada como se describe en el presente documento y el cultivo de dicha célula hospedadora en condiciones adecuadas para permitir su expresión. Preferentemente, dicho polipéptido es un polipéptido terapéuticamente útil.

En un aspecto adicional la invención proporciona una preparación farmacéutica que comprende el polinucleótido, el vector o la célula hospedadora como se describe en el presente documento y un vehículo, excipiente, tampón o medio farmacéuticamente aceptable.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La **Figura 1** muestra un plásmido con el mapa del vector rps3-1005-EGFP (véase Ejemplo 1)

La **Figura 2** muestra un plásmido con el mapa del vector rps11-1005-EGFP (véase Ejemplo 2)

25 La **Figura 3** muestra la expresión del informador EGFP como se expresa por diversas construcciones de rps3 en células CHO-K1 analizadas mediante análisis FACS, 8 días después de la transfección. La Figura A muestra la fluorescencia media, y la Figura B indica el porcentaje de células que expresan el gen informador a un nivel detectable (% de células positivas). Véase Ejemplo 1.

30 La **Figura 4** muestra la expresión del gen informador de construcciones de rps11 en células CHO-K1,

7 días después de la transfección (analizadas por FACS). A y C son las cuentas totales, de B a E son los resultados basados sólo en las células que expresan dentro de la población. La Figura A muestra la fluorescencia media de células en el fondo de reserva seleccionado de manera estable, y la Figura B muestra el porcentaje (%) de células positivas. Véase Ejemplo 2.

La Figura 5

muestra el nivel de expresión de un gen informador en células NS0 transfectadas de manera estable cuando se lleva a cabo por el promotor hCMV sin elementos adicionales presentes o cuando se coloca un elemento hnRNPA2 de 8 kb o un elemento RPS3 de 3 kb inmediatamente en 5' al promotor de hCMV. La Figura A muestra la intensidad de fluorescencia media de los fondos de reserva estables a los 28 días y la Figura B muestra el porcentaje (%) de células positivas.

La Figura 6

muestra datos similares a la Figura 5 para construcciones de rps11. La Figura A muestra los valores de la intensidad de fluorescencia media para fondos de reserva estables generados con construcciones de hCMV o construcciones idénticas con un elemento hnRNPA2 de 8 kb o un elemento RPS3 de 3 kb inmediatamente en 5' al promotor. La Figura B muestra el porcentaje de células positivas dentro del fondo de reserva que expresan el gen informador (porcentaje (%) de células positivas).

DESCRIPCIÓN DETALLADA**Materiales y procedimientos****Análisis de micromatrices**

Se extrajo el ARN total de células CHO-K1
 5 confluentes al ~80% usando el kit de extracción RNeasy
 RNA (Qiagen, Crawley, RU) según el protocolo del
 fabricante. El ARN total (2 µg/µl) se sometió a análisis
 de expresión en micromatrices usando un banco de
 oligonucleótidos de ratón de oligómeros de 70 unidades
 10 (Operon V.1) que representa 13.443 transcritos conocidos.
 El laboratorio de Genómica y micromatrices de la
 Universidad de Cincinnati realizó el análisis de
 micromatrices según los protocolos de referencia
 (<http://microarray.uc.edu>).

15 Las secuencias de los transcritos génicos se
 clasificaron según su fluorescencia creciente. Puesto que
 nuestros estudios previos detallaban los loci
 HNRPA2B1/CBX3 como un elemento de reestructuración de la
 cromatina y que confiere beneficios al hCMV, el
 20 transcrito HNRPA2 fue identificado como el nivel de
 expresión basal. No obstante, usando el análisis de
 micromatrices disponible, el transcrito HNRPA2 apenas fue
 detectable. Puesto que el nivel de expresión del
 transcrito HNRPA2 era mínimo, usando HNRPA2 como
 25 referencia identificamos 3829 secuencias para su
 potencial análisis. Por tanto, se identificaron 7
 secuencias del 2% superior (76 secuencias) de los
 transcritos ordenados y expresados según los criterios de
 inclusión de una isla CpG y uno o más sitios de inicio de
 30 la transcripción conocidos/putativos (véase Tabla 1). La
 posición de las islas CpG, el tamaño y las relaciones
 GC:CG se verificaron usando GrailEXP
 (<http://compbio.ornl.gov>). Los sitios de inicio de la
 transcripción conocidos/putativos se identificaron a
 35 partir del análisis NIX blast (<http://www.hamp.mrc.ac.uk>)
 y las bases de datos de Ensembl (<http://www.ensembl.org>).

Amplificación por PCR de fragmentos que contienen islas CpG

Los oligonucleótidos de la PCR estaban diseñados para amplificar fragmentos de 3 kb aproximadamente que contienen la isla CpG completa embebida en la región promotora y que incluyen 500 pb aproximadamente de secuencia codificante según la estructura de la secuencia codificante conocida o predicha (véase Tabla 2).

Las reacciones de PCR contenían grupos de oligonucleótidos específicos de cada fragmento genómico (2 pmol de cada cebador; Tabla 2). La amplificación por PCR se consiguió usando las premezclas Failsafe™ PCR A-F (Cambio, RU), 1 unidad de Taq ADN polimerasa (Promega, RU) y 200 ng de ADN no codificante. La desnaturalización inicial se realizó a 96°C durante 2 minutos, mientras que la amplificación por PCR se llevó a cabo durante 35 ciclos (94°C durante 1 min, 55-60°C durante 1 min, 72°C durante 5 min). Se incluyó una etapa de extensión final (72°C durante 10 min).

Los productos de la PCR se purificaron en gel, usando columnas de purificación GFX DNA (Amersham, RU) según el protocolo del fabricante, y se sometieron al proceso de clonación TOPO TA cloning® según el protocolo del fabricante (TOPO; Invitrogen, RU). Se obtuvieron las orientaciones en sentido directo y en sentido contrario para cada fragmento que contiene islas CpG clonadas en vectores TOPO (Invitrogen, RU).

Construcción del vector de expresión

Se construyó un vector de expresión control (denominado CET1005EGFP, SEQ ID NO: 20) con la inserción de un hCMVIEGFP/sv40 pA (sitio de clonación múltiple NheI/AgeI borrado) procedente de pEGFP-N1 en CET 900, seguido de la inserción del casete AscI de este vector en el sitio AscI de CET 1005.

Todos los fragmentos con islas CpG se eliminaron de TOP02.1 (Invitrogen, RU) a menos que se indique otra

cosa. El fragmento Terf2ip Acc65I/EcoRV se insertó en
 Acc65I/SwaI de 1005. GAPDH SpeI/SnaBI se insertó en
 PmeI/XbaI de 1005. El fragmento RPS3 XbaI/SpeI se insertó
 en XbaI de 1005. Los fragmentos romos RPS11 y TUBA1 EcoRI
 5 se eliminaron de TOPO4.0 y TOPO2.1, respectivamente
 (Invitrogen, RU) y se insertaron en PmeI de 1005. Por
 último, los fragmentos A430106P18Rik (EcoRV) y
 2510006D16Rik (BstXI) también se insertaron en PmeI de
 1005. Todos los fragmentos que contienen islas CpG se
 10 insertaron en ambas orientaciones en sentido directo y en
 sentido contrario inmediatamente en 5' del promotor hCMV.

Líneas celulares y transfecciones.

Las células CHO-K1 se crecieron en HAMS F12
 (Invitrogen, Paisley, RU) más 4500 mg/l de L-ananil-L-
 15 glutamina, 10 µg/ml de cada una de penicilina y
 estreptomycin, y el 10% (v/v) de suero fetal bovino
 inactivado por calor (FCS; Invitrogen, Paisley, RU). La
 transfección se llevó a cabo por electroporación usando
 aproximadamente 10^7 células procedentes de cultivos
 20 confluentes al 80% y un set Gene Pulser II™ de BioRad
 para proporcionar un único pulso de 975 µF a 250 V. Las
 transfecciones requirieron 2 µg de plásmido CET1005EGFP
 linealizado y cantidades molares equivalentes para
 vectores de expresión de diferentes tamaños. Las células
 25 transfectadas de manera estable se seleccionaron y se
 mantuvieron en medio de crecimiento que contiene 12,58
 µ/ml de sulfato de puromicina (Sigma, RU).

Cuantificación de la expresión transgénica

El análisis de las células transfectadas con las
 30 construcciones informadoras EGFP se realizó con un
 Becton-Dickinson.

El FACScan se realizó usando la línea celular CHO-K1
 parental como control de fondo autofluorescente.

Tabla 1. Secuencias analizadas

Locus	Acc. # ^a	Descripción ^b	Isla CpG ^c		
			pb	% CG	GC/CG
<i>Tert2ip</i>	AB041557	Factor 2 que se une a repeticiones teloméricas que interacciona con la proteína 1 (proteína telomérica Rap1 que interacciona con TRF2) .	968	64,56	0,90
<i>Gapd</i> ^d	M32599	Glicerinaldehído-3-fosfato deshidrogenasa	1187	60,50	0,84
<i>RPS3</i>	NM012052	RPS3 - proteína ribosómica S3 de 40S.	419	60,39	0,87
<i>TUBA 1</i>	M13445	Cadena de la tubulina alfa-1 (Alfa-tubulina 1) .	850	66,30	0,91
<i>RPS11</i>	AK011207	RPS11 - proteína ribosómica S11 de 40S.	957	59,71	0,95
<i>A430106Pl8Rik</i>	AK020778	Marcaje de la secuencia expresada	982	63,62	0,93
<i>2510006D16Rik</i>	AK010915	Marcaje de la secuencia expresada	679	67,69	0,75

^aNº de acceso del Genbank^bDescripción del Ensembl (<http://www.ensembl.org/>)^cGrailexp (<http://compbio.ornl.gov/orailexp>)^dGapd - derivado de la secuencia humana

Tabla 2. Oligonucleótidos de la PCR y tamaños de los amplicones

Locus	Sentido directo	Sentido contrario	Amplicón
<i>Terf2ip</i>	gtagtttctgacttggaat gt (SEQ ID NO: 3)	aactgacctgccatgcca ttc (SEQ ID NO: 4)	2995 pb
<i>Gapd</i>	gagcagtccggtgtcacta (SEQ ID NO: 5)	gcagagaagcagacagtt atg (SEQ ID NO: 6)	3096 pb
<i>RPS3</i>	cagagcatcaagtacctgtg a (SEQ ID NO: 7)	taaccactaagccatctc tcc (SEQ ID NO: 8)	3056 pb
<i>TUBA1</i>	caagaacaaggaagctggcc (SEQ ID NO: 9)	taaaacccacagcactgt aggg (SEQ ID NO: 10)	3049 pb
<i>RPS11</i>	aagactgtttgcctcatgcc (SEQ ID NO: 11)	ggatgacaatggtcctct gc (SEQ ID NO: 12)	3020 pb
<i>A430106P18 Rik</i>	atggtttagg ttcacgtcc (SEQ ID NO: 13)	atccctcacattgccaag cc (SEQ ID NO: 14)	3128 pb
<i>2510006D16 Rik</i>	acttaagacct gatgcctcc (SEQ ID NO: 15)	gctagcttacataggcag cc (SEQ ID NO: 16)	2997 bp

Ejemplo 1: Expresión conducida por el elemento rps3

La SEQ ID NO: 1 muestra la secuencia clonada RPS3 (nucleótidos 38 a 3154); la SEQ ID NO: 17 muestra la secuencia completa del plásmido pRPS3-1005-EGFP; la SEQ ID NO: 18 muestra la secuencia completa del plásmido pCET1015-EGFP.

Se investigaron los niveles de expresión de EGFP, 8 días después de la transfección, en fondos de reserva de CHO-K1 que contienen hCMV solo (construcción control; plásmido pCET1005-EGFP, linealizado con PmeI antes de la transfección), construcciones que contienen un fragmento RNPA2 de 8 kb (plásmido pCET1015-EGFP, linealizado con PmeI antes de la transfección) y Rps3 (plásmido pRPS3-1005-EGFP; linealizado con PmeI antes de la transfección).

Los fondos de reserva generados con construcciones que contienen Rps3 presentan un incremento significativo en los niveles de expresión de EGFP comparados con las construcciones control. La adición de la secuencia Rps3 en 5' del promotor hCMV dio como resultado un incremento

de 5,5 ó 1,5 veces en la intensidad de fluorescencia media con respecto a construcciones que contienen el elemento de control o el elemento hnRNPA2, respectivamente (Figura 3A).

5 Se investigó la actividad de las construcciones en células NS0. El incremento en la intensidad de fluorescencia media en fondos de reserva estables cuando el elemento RPS3 o los elementos hnRNPA2 están incluidos en las construcciones, comparado con el promotor hCMV
10 solo, fue de 28 ó 18 veces, respectivamente (Figura 5A). En ambos tipos de células CHO-K1 y NS0, el porcentaje de células positivas se incrementó significativamente con el elemento hnRNPA2, pero este incremento fue superior con el elemento RPS3 (Figuras 3B y 5B).

15 **Ejemplo 2: Expresión conducida por el elemento rps11**

La SEQ ID NO: 2 muestra la secuencia clonada RPS11 (nucleótidos 12 a 3032); la SEQ ID NO: 19 muestra la secuencia completa de pRPS11-1005-EGFP.

Los vectores que contienen Rps11 y el control (Pmel linealizado) se transfectaron en líneas celulares CHO-K1
20 y NS0 y se generaron fondos de reserva estables mediante la selección con puromicina. Los niveles de expresión media de EGFP se valoraron mediante análisis FACScan.

La adición del elemento Rps11 en dirección 5' del hCMV dio como resultado un incremento de 1,2 veces en los
25 niveles de expresión media de EGFP, en fondos de reserva de CHO-K1, comparados con una construcción que contiene el fragmento RNPA2 descrito previamente (Fig 4A).

Líneas celulares NS0 transfectadas de manera estable con construcciones que contienen Rps11 mostraron un
30 incremento de 1,8 y 1,5 veces (respectivamente), en los niveles de expresión media de EGFP comparados con las construcciones de hCMV y RNPA2 (Fig 6A).

Se observó un incremento en el porcentaje de células
35 positivas para líneas celulares CHO-K1 transfectadas con construcciones Rps11 comparadas con construcciones RNPA2

(Fig 4B). Además, se observó un incremento en el porcentaje de células positivas en fondos de reserva de NS0 transfectadas con construcciones Rps11 comparadas tanto con hCMV como con RNPA2 (Fig 6B).

5 **LISTADO DE SECUENCIAS**

<110> serologicals Investment company
 <120> Elementos de expresión mejorados
 <130> P108872WO
 <160> 20
 10 <170> PatentIn version 3.2
 <210> 1
 <211> 3145
 <212> ADN
 <213> Mus musculus rps3
 15 <400> 1

ctagtaacgg ccgccagtgt gctggaattc gcccttataa ccactgagcc atctctccag	60
ccctgagtca tgatttttagt gtgagaggca tcattgaatt ttctgagcac ggccatcagg	120
gtagctggca caggtcttca gatacaagga gatagttata agaaggcagc catggctgtg	180
gtgcactaga aatggagaaa cagcttcacg aggtgacaga ccagtctgac tctgtcccat	240
gattagaagc catcttggtt caaggtcaaa ataagttcat tcctgttttc tgtaacactt	300
gggtttgatc ctgtcgtcaa cccattttct ggaatttgac atgttccata ctccattata	360
ccctgacttc caccctgata agatgttctg ccaagttcct gtgtagccaa cattccccctg	420
gaaatctctc ttcccttgga aaccacctag tcttagaaat tttagattat ataaattcca	480
cttctatggt tgatgctatt ctttaaaact ccactttagg gagatagccc tgtctgatag	540
aaaataaaac ttgcttaatt tgtctaaaag agtttaagta atagttttta cttttgttcc	600
gtgggattaa tacagggtga aacagactcc cgtgtttcca gtgtgaagtg agccacacac	660
tgcagtacaa gttatatcag caggttctgc ctctgggcaa tgaacttttg cttgtgtgga	720
catcagggtc tgtgtgaagg gaaggtccta tggcctagat ttatactatt caacagtctg	780
tccccgaagc cctgggtgctt tattattttg acaagcccct gctgctggta ttccaccctg	840
ctgcgagtca aaaaagttcc tgtctcgga aaacaaaaca aaacaaaaca accaaaaaat	900
aaaatttttt tttccacag gttctagtgg aggtgctcac taccagaaat cctacaaata	960
agcccatctc atggatcagg gtttaccttt gtaataatat taaatctgtg tgcattgtcg	1020
cacgcatgtg ttttatgctt gcatatatgt atacgcagcc atggttttct actgtcccac	1080
tcactctgta acttactgag ccattccagct ggtcctctaa atacatttca atgaaagttt	1140
tcattagcgt gaacgtgaag gtggtaaaat ctgttagtgt gtgcttatgc ctgtgggttg	1200
cacctctagt ctgaagggtt ctcttttcaa attttttatt tatttacgtt ttacttttg	1260
agtcagaaac tcataaaggc caagctggcc tcgaattcgc tatgtagtca atgatgacct	1320
taaacttgtg accctctact tcgttagtgc tggaaaccca agcttgctga gtacagagca	1380
ctttcagacc ggaactagat gtctacttcc tgttccgcct acattacagg ttgctagggt	1440
acaccccccc tacgccgttt tagacgcaaa acttcatttc ccatgcaaaa cttcatttcc	1500

```

catgaacact tgcaagggtc gccgcgctgc gcggcgctcat tgctcccgcc ctatatacct 1560
acttccgccc gcgagccact tcctttcctt tcagcggcgc gcggctgcaa gatggcggtg 1620
cagatttcca agaagaggaa ggtaagcgtc tgggcccggg tcgggagtcg gccgcgggtt 1680
ctacaagtgc cagggaggcc tgtggctccg tgatcagtcg tgtggagcgt ctggggccgc 1740
ctgccgtctc ttcgagcctc ggatggccgt agattgtgta ttgggccgga gccgggagag 1800
tgctgtgtgc ctgggcaagg gagggacaaa ctctcagagt tctggaccga ctcgaacacc 1860
gggcgccctc agttccggac tagacacctt tgagcgcttc ttggtctcca taatagtaat 1920
cctgtggcac agttagaggg cgtgtgccat cagatctagt ccagtttctt tagtaagtga 1980
agtttagcag tcccttctct tagtcgctg atcctgcaag tggccatagt tgaaagccta 2040
cttactgact gctgccgtgt tctactggga cccggagctg cagcgctcct gtggttatca 2100
tttcatgggg gaaaagtgtg caggttgcca ggtttagaaa tagatggtct gtcgtttgtg 2160
cttatgcaca cagatgataa acctgttttg agtcaggatt cctctcctat ccgaggatca 2220
acttacagtc ccagctgtac atgtgctact tggagacaga tttttctttg tctcttgggt 2280
gtagattatg ccgtagagcc cttcgatgaa gaggtgatga cgagtctgag taggaagtgt 2340
tgtctttgtc caagatgcct cactatgctg cgttctgtgg cacagctgaa agcactgtgg 2400
tcaaaagaaa cttcctaaag atgaccaaga ggcatttgtc tgagaagggt tgctgctttt 2460
ctgtagggcc attgggcttg ctctgactaa cctgtcttc acctcagagg taacttgttt 2520
cctttggttc agttttagc tgatggcatc ttcaaagctg agctgaatga atttctcact 2580
cgggagctgg ctgaagatgg ctactctgga gttgaagtcc gagttacacc aaccaggaca 2640
gaaatcatta ttttagccac caggtagaaa taccattgat tgtcacctgt aaatactgtg 2700
tgtactgaga tgctgtgtaa acttgggcca accaagcagt aaatctggcc tcagtgggtg 2760
taactgcttt gttagaactg catttgggaa gaacttacct tccatttaac gtgtgtgctg 2820
gcgttgtggt gggcggcagg tgggatcttg agtaaagtgt tgcgcttccc ctctacagga 2880
cacagaatgt tcttggggag aagggtcgct ggatcagaga gttgaccgca gttgtccaga 2940
agcgcttttg cttccctgaa ggcagcgtag aggtgagttc ctctgcttta tctcccggg 3000
gttttagact gagttgggat gtggcttctg ctatagaatt gtacttctga aaacctgaca 3060
tggccagtga cagtcacagg tacttgatgc tctgagggcg aattctgcag atatccatca 3120
cactggcggc cgctcgagca tgcatt 3145

```

<210> 2

<211> 3039

<212> ADN

5 <213> Mus musculus rps11

<400> 2 60

```

aattcgccct taagactgtt tgcctcatgc ctgcctggcc tgccttcct ccgccgcaa 60
ctaggaagt ggggacaaa gggtccttag gcactgctcc tgtgggtaga ggggacatta 120
gagagctgac agcgaccac ctgcatgagt ttttattaaa gtgcaaacca tgggatgaat 180

```

cagttgagct tcagtgttga aaatgagtag cagggctgcc ccacccacct gaccaagtac 240
 cctattctgc agctatgaaa atgagatctg cacatgagct ggggttcaca agtgcacact 300
 tggagcactg ccttgctcct tcccagcaga ccacaaagca gtatTTTTct ggaggatttt 360
 atgtgctaataa aaattatttg acttaagtgt gtacgatgtg tgctgtgcag agaggggcag 420
 agggcaccag caggtcatct gcatgggggg cccctttggg tgaatccttg ctcacgggat 480
 aggctttgtt gctcaaaagt tgcagatata catcttgggt cctgtcctag atggtgttac 540
 tgtaagtcag caccaagata caagagctgg tacctggact gtaggagggtc aggccatgac 600
 acaaaggctg ggactaaagg catTTaccac gcctgagtct tctggttctt taaacatcaa 660
 atccttccgg gggctggcga gatggctcag tggTTaagag cacagactgc tcttacgaag 720
 gatccgagtt caaatcccag caaccaaagt gtgcctaaca actatccata atgaaatctg 780
 atgccctctt ctggagtatc tgagaacagc tacagtgtac ttacatataa tcttaaaaaat 840
 gcttcccatg ttaaccacca ctagagtttt tattacagct agctgacctg gaagccaagt 900
 ccttatgcct ccgtgagtgc tggggTTaaa aagatccagc accactcaaa atgtcaatct 960
 atTTTgaaaa tatgctttat actgttctag cccatctgtg cagggctaga acggtgaata 1020
 cgagaaactg acacaagctt ttgccacctg gctaaatggg tcctctatta cctgggggtg 1080
 tcacctaagg ttagacactc atccacgagt agtcaggaca taaacccatc aaagtgtggg 1140
 tagacgcgca gcctgagata ctgtcaacaa aggacatgcg accttggtga cgtcggcctt 1200
 taataaaagg aagaaagggt gactattcgg tcgacgctgg ctgctcctga catcgtatgg 1260
 cagatactct gctgtaaagc ggTtcccc tttcttgaga cccgctctgc acggccgctt 1320
 ctctctggaa actgaatccc agcacgtgtt tcccaacccg tacggcacgc ctctccgcc 1380
 ctaagcctcg ccgtaccaca tgatgcacgt ttcctccaca tcgtgctcct gaaatctcgc 1440
 gagatgatag gatcttcccc ccccttagtc ctccccgc atggcggcgt acggacagtc 1500
 ccaggaacgc gggctctcgc cggaagtacc tccacctcc gtgaggataa ccccgcgta 1560
 ctccgcccc gacctcgcgt ggtgaataag gaagccggga gcggccctgc ctctcccttt 1620
 ctccggcggc cgggaagatg gcggacattc aggttcgagc gtttagttgc tttccccga 1680
 cgcttcggtg tggagcgtat cccttggcgt cctcgttgtc ttacgcatta gctgaagcga 1740
 ggatgcctgc gaatgccttc gtctcaggcg gctcggaaat ccgggctcta cgcagtaatg 1800
 gggTccctgg cgcttcggga gttggttctt aaagctcaga gcttaacggg tgagggattg 1860
 tggcgggagg agggcatcct gcggcgcggg agtcctgcgg cggcagagcc ggggacactg 1920
 ggtaaagcag gttttttccc ctgatggag actgaggccc ggacctcgtg cgctctacgg 1980
 cagggctgcg gtcccacact cgctgtagtt ttcagtgtga gcgcagctct ggcctcgatg 2040
 agcttaggct tgtcttaaac ttgccatcct gcctcaacct caaccgggat gacagatccg 2100
 gccaccagg ctcggctacg tggacataag cttgaatccc gaatgagtggt atttgtatgt 2160
 tttggagggtc cagtctggct gaaaagctct ttttgatctc agccgtgagt tctgcaggct 2220

gtggaagggtgt tagatgggac gcagtgtgtg agctaaacta gacttgggggt ggttggagag 2280
ccctgaccag ccggtttttg cgattggggc aaataagggtt gaaggtagga aggaagaaat 2340
attgtctctg atttccttga actttacctg caacctcacc aaattctcat ccctacagac 2400
ggagcgtgct taccaaaagc agcctacgat ctttcaaac aagaagcggg ttctgctggg 2460
agaaaccggc aaggaaaaac tccctcggtg ctacaagaat atcgggtctag gcttcaagac 2520
gcctaaagag gtacaggacc ctccagcaga tgagatccct gctgccctgc acgtgtggga 2580
gcacagccac cccgccccct tcacagtggc ttcccatggg cccctgggaa ttgtagtatg 2640
ggccctgagg cgtcatcctt ggttctgttt aggaagtggg aatctaaacc ccactttctt 2700
aactttgcag gctattgagg gtacctacat agacaagaaa tgccccctca ctggtaacgt 2760
ctccatccga ggtcggatcc tgtctggtga gtgggatggt ggaaggggtg ttctaggttc 2820
ctgcgtccag gggcgctggc aagtgatgtc tgttctcacg atggtcttca gatgtcctct 2880
agggcactgc tgagacagcc agttgacaaa gctgatgcca taaatggagc ttcttgggag 2940
ccccgttcaa ctgactccta cctgctaaca cctttctgtt actctcccag gtgtcgtgac 3000
gaagatgaag atgcagagga ccattgtcat ccaagggcg 3039

<210> 3

<211> 22

5 <212> ADN

<213> Artificial

<220>

<223> Oligonucleótido sintético de la PCR

<400> 3 gtagttctg acttgaaat gt 22

10 <210> 4

<211> 21

<212> ADN

<213> Artificial

<220>

15 <223> Oligonucleótido sintético de la PCR

<400> 4 aactgacctg ccatgccatt c 21

<210> 5

<211> 19

<212> ADN

20 <213> Artificial

<220>

<223> Oligonucleótido sintético de la PCR

<400> 5 gagcagtcg gtgtcacta 19

<210> 6
 <211> 21
 <212> ADN
 <213> Artificial
 5 <220>
 <223> Oligonucleótido sintético de la PCR
 <400> 6 gcagagaagc agacagttat g 21
 <210> 7
 <211> 21
 10 <212> ADN
 <213> Artificial
 <220>
 <223> Oligonucleótido sintético de la PCR
 <400> 7
 15 cagagcatca agtacctgtg a 21
 <210> 8
 <211> 21
 <212> ADN
 <213> Artificial
 20 <220>
 <223> Oligonucleótido sintético de la PCR
 <400> 8
 taaccactaa gccatctctc c 21
 <210> 9
 25 <211> 20
 <212> ADN
 <213> Artificial
 <220>
 <223> Oligonucleótido sintético de la PCR
 30 <400> 9
 caagaacaag gaagctggcc 20
 <210> 10
 <211> 22
 <212> ADN
 35 <213> Artificial
 <220>

<223> Oligonucleótido sintético de la PCR
 <400> 10
 taaaaccac agcactgtag gg 22
 <210> 11
 5 <211> 20
 <212> ADN
 <213> Artificial
 <220>
 <223> Oligonucleótido sintético de la PCR
 10 <400> 11
 aagactgttt gcctcatgcc 20
 <210> 12
 <211> 20
 <212> ADN
 15 <213> Artificial
 <220>
 <223> Oligonucleótido sintético de la PCR
 <400> 12
 ggatgacaat ggtcctctgc 20
 20 <210> 13
 <211> 20
 <212> ADN
 <213> Artificial
 <220>
 25 <223> Oligonucleótido sintético de la PCR
 <400> 13
 gtggtgtag gttcacgtcc 20
 <210> 14
 <211> 20
 30 <212> ADN
 <213> Artificial
 <220>
 <223> Oligonucleótido sintético de la PCR
 <400> 14
 35 atccctcaca ttgccaagcc 20
 <210> 15

```

<211> 20
<212> ADN
<213> Artificial
<220>
5  <223> Oligonucleótido sintético de la PCR
   <400> 15
   acttaagacc tgatgcctcc      20
   <210> 16
   <211> 20
10 <212> ADN
   <213> Artificial
   <220>
   <223> Oligonucleótido sintético de la PCR
   <400> 16
15 gctagcttac ataggcagcc      20
   <210> 17
   <211> 8691
   <212> ADN
   <213> Artificial
20 <220>
   <223> Vector pRPS3 1005 EGFP
   <400> 17

```

cgttgtaaaa	cgacggccag	tgaattgtaa	tacgactcac	tatagggcga	attgggtacc	60
gggccccccc	tcgaagttaa	aacattttaa	tctagtaacg	gccgccagtg	tgtgtgaatt	120
cgcctttata	accactgagc	catctctcca	gccctgagtc	atgattttag	tgtgagaggc	180
atcattgaat	tttctgagca	cggccatcag	ggtagctggc	acaggctctc	agatacaagg	240
agatagttaa	aagaaggcag	ccatggctgt	ggtgcactag	aaatggagaa	acagcttcat	300
caggtgacag	accagtctga	ctctgtccca	tgattagaag	ccatcttggt	acaaggtaaa	360
aataagttca	ttcctgtttt	ctgtaacact	tgggtttgat	cctgtcgtca	acccattttc	420
tggaatttga	catgttccat	actccattat	accctgactt	ccaccctgat	aagatgttct	480
gccaagttcc	tgtgtagcca	acattcccct	ggaaatctct	cttcccttgg	aaaccaccta	540
gtcttagaaa	ttttgagtta	tataaattcc	acttctatgt	ttgatgctat	tctttaaaaa	600
tccactttag	ggagatagcc	ctgtctgata	gaaaataaaa	cttgcttaat	ttgtctaaaa	660
gagtttaagt	aatagttttt	acttttgttc	cgtgggatta	atacagggtg	aaacagactc	720
ccgtgtttcc	agtggtgaag	gagccacaca	ctgcagtaca	agttatatca	gcaggttctg	780
cctctgggca	atgaactttt	gcttgtgtgg	acatcagggg	ctgtgtgaag	ggaaggctct	840
atggcctaga	tttatactat	tcaacagtct	gtccccgaag	ccctgggtgct	ttattatttt	900
gacaagcccc	tgctgtggtt	attccaccct	gctgcgagtc	aaaaaagttc	ctgtctcgga	960
aaaacaaaaa	aaaacaaaaa	aaccaaaaaa	taaaattttt	ttttcccaca	ggttctagt	1020
gaggtgtctc	ctaccagaaa	tctacaaaat	aagcccatct	catggatcag	ggtttacctt	1080
tgtaataata	ttaaatctgt	gtgcatgtgc	gcacgcatgt	gttttatgct	tgcatatatg	1140
tatacgagc	catggttttc	tactgtccca	ctcactctgt	aacttactga	gccatccagc	1200
tggctctcta	aatacatttc	aatgaaagtt	ttcattagcg	tgaacgtgaa	ggtggtaaaa	1260
tctgttagtg	tgtgcttatg	cctgtgggtt	gcacctctag	tctgaagggt	gctcttttca	1320
aattttttat	ttatttacgt	ttttactttt	gagtcagaaa	ctcataaagg	ccaagctggc	1380
ctcgaattcg	ctatgtagtc	aatgatgacc	ttaaacttgt	gaccctctac	ttcgtttagtg	1440
ctggaacccc	aagcttgctg	agtacagagc	actttcagac	cggaaactaga	tgtctacttc	1500
ctgttcgcc	tacattacag	gttgctaggt	tacacccccc	ctacgccggt	ttagacgcaa	1560
aacttcattt	cccatgcaaa	acttcatttc	ccatgaacac	ttgcaagggt	cgccgcgctg	1620
cgcggcgta	ttgtcccg	cctatatacc	tacttccg	cgcgagccac	ttcctttcct	1680
ttcagcggcg	cgcggctgca	agatggcggg	gcagatttcc	aagaagagga	aggtaagcgt	1740
ctgggcccgg	ttcgggagtc	cgcgcgggt	tctacaagtg	ccaggagggc	ctgtggctcc	1800
gtgatcagtc	ctgtggagcg	tctggggccg	cctgccgtct	cttcgagcct	cggatggccg	1860
tagatttgtg	attgggcccg	agccgggcca	gtgctgtgtg	cctgggcaag	ggagggacaa	1920
actcctcgag	ttctggaccg	actcgaacac	cgggcgcctc	cagttccgga	ctagacacct	1980

ttgagcggtt	cttgggtctcc	ataatagtaa	tcctgtggca	cagttagagg	gcgtgtgcca	2040
tcagatctag	tccagtttct	ttagtaagt	aagtttagca	gtcccttctc	ttagtcgcgt	2100
gatcctgcaa	gtggccatag	ttgaaagcct	acttactgac	tgctgccgtg	ttcactcggg	2160
acccggagct	gcagcgctcc	tgtggttatc	atttcattgg	ggaaaagtgt	gcaggttgcc	2220
aggttttagaa	atagatggtc	tgtcgtttgt	gcttatgcac	acagatgata	aacctgtttt	2280
gagtcaggat	tcctctccta	tccgaggtac	aacttacagt	cccagctgta	catgtgctac	2340
ttggagacag	atcttctctt	gtctcttggg	tgtagattat	gccgtagagc	ccttcgatga	2400
agaggtgatg	acgagtcctg	gtaggaagt	ttgtctttgt	ccaagatgcc	tcactatgct	2460
gcgttctgtg	gcacagctga	aagcactgtg	gtcaaaagaa	acttcctaaa	gatgaccaag	2520
aggcatttgt	ctgagaagg	ttgctgcttt	tctgtagggc	cattgggctt	gctctgacta	2580
acctgtctt	cacctcagag	gtaacttggt	tcctttgggt	cagttttag	ctgatggcat	2640
cttcaaagct	gagctgaatg	aatttctcac	tcgggagctg	gctgaagatg	gctactctgg	2700
agttgaagtc	cgagttacac	caaccaggac	agaaatcatt	atcttagcca	ccaggtagaa	2760
ataccattga	ttgtcacctg	taaatactgt	gtgtactgag	atgctgtgta	aacttgggcc	2820
aaccaagcag	taaatctggc	ctcagtggtg	gtaactgctt	tgtagaact	gcatttggga	2880
agaacttacc	ttccatttaa	cgtgtgtgct	ggcgttgtgg	tgggcggcag	gtgggatctt	2940
gagtaaagtg	ttgcgcttcc	cctctacagg	acacagaatg	ttcttgggga	gaagggtcgt	3000
cggatcagag	agttgaccgc	agttgtccag	aagcgctttg	gcttccctga	aggcagcgta	3060
gaggtgagtt	cctctgcttt	atctcccggg	ggtttttagac	tgagttggga	tgtggcttct	3120
gctatagaat	tgtacttctg	aaaacctgac	atggccagtg	acagtcacag	gtacttgatg	3180
ctctgagggc	gaattctgca	gatatccatc	acactggcgg	ccgctcgagc	atgcatctag	3240
aagcttatcg	ataccggtgg	cgcgccatt	gaattaagat	ctggcccaat	gggccgtacg	3300
aattcgagct	cggtagccgg	ggatcctgat	ctaatagtaa	tcaattacgg	ggtcattagt	3360
tcatagccca	tatatggagt	tccgcgttac	ataacttacg	gtaaatggcc	cgcctggctg	3420
accgccaac	gacccccgcc	cattgacgtc	aataatgacg	tatgttccca	tagtaacgcc	3480
aatagggact	ttccattgac	gtcaatgggt	ggagtattta	cggtaaactg	cccacttggc	3540
agtacatcaa	gtgtatcata	tgccaagtac	gccccctatt	gacgtcaatg	acggtaaagt	3600
gccccgctgg	cattatgccc	agtacatgac	cttatgggac	tttctactt	ggcagtagat	3660
ctacgtatta	gtcatcgcta	ttaccatggt	gatgcggttt	tggcagtaca	tcaatgggag	3720
tggatagcgg	tttgactcac	ggggatttcc	aagtctccac	cccattgacg	tcaatgggag	3780
tttgttttgg	cacaaaaatc	aacgggactt	tcaaaaatgt	cgtaacaact	ccgccccatt	3840
gacgcaaagt	ggcggtaggc	gtgtacgggt	ggaggtctat	ataagcagag	ctggtttagt	3900
gaaccgtcag	atccgtcgcc	accatgggtg	gcaagggcga	ggagctgttc	accgggggtg	3960
tgcccatcct	ggtcgagctg	gacggcgacg	taaacggcca	caagttcagc	gtgtccggcg	4020

agggcgaggg cgatgccacc tacggcaagc tgaccctgaa gttcatctgc accaccggca 4080
 agctgcccgt gccctggccc accctcgtga ccaccctgac ctacggcgtg cagtgccttca 4140
 gccgctaccc cgaccacatg aagcagcacg acttcttcaa gtccgccatg cccgaaggct 4200
 acgtccagga ggcgaccatc ttcttcaagg acgacggcaa ctacaagacc cgcgccgagg 4260
 tgaagttcga gggcgacacc ctggtgaacc gcatcgagct gaagggcacg gacttcaagg 4320
 aggacggcaa catcctgggg cacaagctgg agtacaacta caacagccac aacgtctata 4380
 tcatggccga caagcagaag aacggcatca aggtgaactt caagatccgc cacaacatcg 4440
 aggacggcag cgtgcagctc gccgaccact accagcagaa ccccccatc ggcgacggcc 4500
 ccgtgctgct gcccgacaac cactacctga gcaccagtc cgccctgagc aaagacccca 4560
 acgagaagcg cgatcacatg gtcctgctgg agttcgtgac cgccgccggg atcactctcg 4620
 gcatggacga gctgtacaag taaagcggcc gcgactctag atcataatca gccataccac 4680
 atttgtagag gttttacttg ctttaaaaaa cctcccacac ctccccctga acctgaaaca 4740
 taaaatgaat gcaattgttg ttgttaactt gtttattgca gcttataatg gttacaaata 4800
 aagcaatagc atcacaaatt tcacaaataa agcatttttt tctactgcatt ctagtgtggt 4860
 tttgtccaaa ctcacatg tatcttaact agagtcgacc tgcaggcatg caagcttacc 4920
 ggtggcgcg cgccaattg ttaattaaga tctggcccaa tgggcccgtac gaattcctta 4980
 ggctaccggg taggggaggc gcttttccca aggcagtctg gagcatgcgc ttttagcagcc 5040
 ccgctgggca cttggcgcta cacaagtggc ctctggcctc gcacacattc cacatccacc 5100
 ggccggtagg cgccaaccgg ctccgttctt tgggtggccc ttcgcccac cttctactcc 5160
 tcccctagtc aggaagtcc ccccgcccc gcagctcgcg tcgtgcagga cgtgacaaat 5220
 ggaagtagca cgtctcacta gtctcgtgca gatggacagc accgctgagc aatggaagcg 5280
 ggtaggcctt tggggcagcg gccaatagca gctttgctcc ttcgctttct gggctcagag 5340
 gctgggaagg ggtgggtccg ggggcgggct caggggcggg ctcaggggcg gggcgggcg 5400
 ccgaaggctc tcgggaggcc cggcattctg cacgcttcaa aagcgacgt ctgccgcgt 5460
 gttctcctct tcctcatctc cgggcctttc gaccagctta ccatgaccga gtacaagccc 5520
 acggtgcgcc tcgccaccg cgacgacgtc ccagggccg tacgcacct cgccgcccg 5580
 ttcgccgact accccgccac gcgccacacc gtcgatccgg accgccacat cgagcgggtc 5640
 accgagctgc aagaactctt cctcacgcgc gtcgggctcg acatcggcaa ggtgtgggtc 5700
 gcggacgacg gcgccgcggt ggcgggtctgg accacgccgg agagcgtcga agcggggcg 5760
 gtgttcgccg agatcgggcc gcgcatggcc gagttgagcg gttcccggt ggccgcgag 5820
 caacagatgg aaggcctcct ggcgccgcac cggcccaagg agcccgcgtg gttcctggcc 5880
 accgtcggcg tctcgcccga ccaccagggc aagggctctg gcagcgccgt cgtgctcccc 5940
 ggagtggagg cgccgagcg cgccggggtg ccgccttcc tggagacct cgcccccgc 6000
 aacctcccct tctacgagcg gctcggcttc accgtcaccg ccgacgtcga ggtgcccga 6060

ggaccgcgca cctggtgcat gacccgcaag cccggtgcct gacgcccgcc ccacgacccg 6120
 cagcgcccga ccgaaaggag cgcacgaccc catgcatcgt agacgaaatg accgaccaag 6180
 cgacgcccga cctgccatca cgagatttcg attccaccgc cgctttctat gaaagggttg 6240
 gcttcggaat cgttttccgg gacgcccggc ggatgatcct ccagcgcggg gatctcatgc 6300
 tggagttctt cgcccaccct agggggaggc taactgaaac acggaaggag acaataaccg 6360
 aaggaacccg cgctatgacg gcaataaaaa gacagaataa aacgcacggg gttgggtcgt 6420
 ttgttcataa acgcgggggt cgggtcccagg gctggcactc tgtcgatacc ccaccgagac 6480
 cccattgggg ccaatacgcc cgcttttctt ctttttcccc accccacccc ccaagttcgg 6540
 gtgaaggccc agggctcgca gccaacgtcg gggcggcagg ccccgagctt ttgttccctt 6600
 tagtgagggt taatttcgag cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat 6660
 tgttatccgc tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg 6720
 ggtgcctaata gagtgagcta actcacatta attgctgtgc gctcactgcc cgctttccag 6780
 tcgggaaacc tgtcgtgcca gcatcgcgag cacttttcgg ggaaatgtgc gcggaacccc 6840
 tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcattgagac aataaccctg 6900
 ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc 6960
 ccttattccc ttttttgagg cattttgcct tcctgttttt gctcaccag aaacgctggg 7020
 gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg gggtacatcg aactggatct 7080
 caacagcggg aagatccttg agagttttcg ccccgagaa cgttttccaa tgatgagcac 7140
 ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt gacgcccggc aagagcaact 7200
 cggtcgccgc atacactatt ctcaaatga cttggttgag tactcaccag tcacagaaaa 7260
 gcatcttacg gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga 7320
 taacactgcg gccaaacttac ttctgacaac gatcggagga ccgaaggagc taaccgcttt 7380
 tttgcacaac atgggggatc atgtaactcg ccttgatcgt tgggaaccgg agctgaatga 7440
 agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa caacgttgcg 7500
 caaactatta actggcgaac tacttactct agcttcccgg caacaattaa tagactggat 7560
 ggaggcggat aaagttgcag gaccacttct gcgctcggcc cttccggctg gctggtttat 7620
 tgctgataaa tctggagccg gtgagcgtgg gtctcgcggg atcattgcag cactggggcc 7680
 agatggtaag ccctcccgtg tcgtagttat ctacacgacg gggagtcagg caactatgga 7740
 tgaacgaaat agacagatcg ctgagatagg tgcctcactg attaagcatt ggtaactgtc 7800
 agactcgca cactgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt 7860
 gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggctcgttcg gctgcggcga 7920
 gcggtatcag ctactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 7980
 ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 8040
 ctggcgtttt tccataggct ccgccccct gacgagcatc aaaaaatcg acgctcaagt 8100

```

cagaggtggc gaaacccgac aggactataa agataccagg cgtttcccc tggaagctcc 8160
ctcgtgcgct ctctgttcc gacctgccc cttaccggat acctgtccgc ttttctccct 8220
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 8280
gttcgctcca agctgggctg tgtgcacgaa cccccggtc agcccgaccg ctgcgctta 8340
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 8400
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 8460
tgggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 8520
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaac caccgctggt 8580
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaagg atctcaagaa 8640
gatcctttga tcttttctac ggggtctgac gctcagtga acgaaaactc a 8691
<210> 18
<211> 13827
<212> ADN
<213> Artificial
<220>
<223> Vector pCET 1015 EGFP
<400> 18
  cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga attgggtacc 60
  gggccccccc tcgaagttaa aacatttaaa tctagaagct tcaatgtttt tagcacccctc 120
  tgtgtggagg aaaataatgc agattattct aattagtgtg atatctaacc acattaaaaat 180
  atattacata gtaaaactaca ctccataatt ttataaattt gactccccag ggtaataaac 240
  tagtctctag tctgctcacc ttcaactgta caataaagtc ttggttcttt tgaaatagac 300
  ctcaaatgag acacctaaaa ttcaaagtgt ctttacattt aaagacacct acaggaaagc 360
  aggtaaaaga gccagggtta aaacaaattc taaaaccact tagctgcagt taaacatata 420
  gtaaagatgc actaaagttt ctactctgt aaatccctc cacttcagga aatattccac 480
  tttcccattc actacacgtc gatctagtag tttttccacg acaaattctt caggctctgc 540
  ctcttcaact tttttactct ttccattctg tttttttccc attttttgct aaaataaaac 600
  aaaagagaaa ttaagaaata ttctcttga attttgagca ctttttcaag gctcaattgc 660
  ttatattatt atcacattcg acataaattt ttacttctat atcccagggc agacaccttc 720
  tggaaagatt aaaagtcaac agacaataaa ataaaagaat gctttatctt gttcatttag 780
  ttcaaactta caaccacca ccaaaataat acaataaaaa aacactatct ggaaacagtt 840
  atttttttcc agtctttttt ttigagacag ggtctcacac tcttgctgcc caggctggag 900
  tgcagtggcg tgatctcagc tactgcaac ctccgcctcc ccaggttcaa gcagttctca 960
  tgcctcagcc tccagagtag ctgggattat aggcggatgc caccatgccg ggctaatttt 1020
  ttttgtgttt ttattagaaa cagggtttca ccatgttgac caggctggtc tcaaactcct 1080
  gacctgaagt gattcaccag cctgggcctc ccaaagtgtt ggcattacag gcgtgagcca 1140

```

ctgcgcccgg	ccctgtagtc	ttaaaagacc	aagtttacta	atthttcactc	attttaacaa	1200
cactgcaaca	aacaactatg	caggaagtag	ctaaagggtg	atccagagaa	gcaagtagta	1260
gtgacaggtc	ttaggtgaac	ctatgacaga	ccttgtagcc	accccagat	ggtaaaagcc	1320
ccagccccct	tetcaattca	aataattaatg	tcaaaagcat	caatgatata	gagaaaagat	1380
aaatgcagaa	tgaaaacatg	gttcaaaatc	ctgataccaa	ctgcagggtc	aactatagag	1440
accactagga	ggttcaatta	aaggacaaga	ttatthttcc	ataatctctg	tagataatat	1500
ttcctaccac	ttagaacaaa	actataaagc	tatcacttca	agagaccaac	attacaaatt	1560
tatthttaatt	ccctaagggtg	aaaaaaatcc	ttccttccctg	gtttctcaag	agaaagtcta	1620
tactggtaac	caaattcact	ttaaacaggc	atthttctttg	gtatgacact	atttaagaga	1680
agcaggaaac	caacgtgaac	cagctctttc	caatggctca	agatthtccta	tgagaggact	1740
aaaaatgggg	aaaatthtta	tgagaggatt	aaaaatgggg	gaaaaaaaac	cctgaaatgg	1800
ttaatcagaa	gacccatagg	gctgagaagg	aatccatctt	aacatthcat	cttaaagcaa	1860
atgctattgc	cgggggcagt	ggctcatgcc	tgtaatccca	gcactthggg	aggccgaggt	1920
gggcagatca	tctgaggcca	ggagthtgag	accagcccta	ccaacatgga	gaaaccccg	1980
ttctactaaa	aatacaaaat	tagccaggca	tagtggtgca	tgccctgtaat	cccagctact	2040
tgggaggctg	aggcaggaga	actgcttgaa	cccaggaggc	ttaagthtgcg	gtgagccaag	2100
atcacgccat	tgactcttag	cctggacaac	aagagaaaaa	ctctgtctca	aaaaaacaca	2160
aaaacaaaaa	acccaataac	tatttaaaaa	agataaacct	taattgctca	atcattaaag	2220
ccatcccaca	agtaaagcag	caagcagaaa	aaagthtaaga	acacctcaag	gctacagaag	2280
gacatthcaa	gctatgcagg	catatgaagt	gtgcagacag	atatgtaaga	aaggccctcaa	2340
gactgcaaaa	gggcatttca	agctatgcaa	gcataatagg	aacacataca	cacacacaaa	2400
ataaaatccc	ctgaaatata	aaaacatgca	gcaaacacct	gacgthttttg	gataccatth	2460
ctaagtcagg	tgthtatgatt	ctcattagtc	aagatacttg	agtactgggc	ccaaacagct	2520
ttctgccact	gtacagtata	agaaggtagg	aataatgggtg	ggaggagcaa	agacaaactg	2580
taatagacag	aagtgtatca	gataccatata	ctacatgaaa	aacaaaacag	ctactgccac	2640
aaaggagaaa	ggctaacaaa	ataaagtcaa	caataaatac	agaaaatgaa	aaggatacac	2700
actaaggtht	acaaaaaaa	aaaggcagac	aaaatgccat	acagtatthca	ttcactacta	2760
tggcattcat	aagctagtht	caaatgctca	ctatthttctt	ttatagtata	tatttgccct	2820
aaccagcac	thttttccaa	aagtggatga	gtcaaaataa	atthcccat	atttaagtga	2880
aattaacagc	acacatatct	cacaacacta	atgaatthtt	aaaatggaaa	gttaagaact	2940
thtaagtgg	ccaacctgtg	atccttcaca	aaataaacta	aatacaataa	cagaccccaa	3000
aggctatcaa	ttgcgtgcaa	aaacaacttc	tgthttccag	ggtaaacaga	atctaattgca	3060
gaatctaattg	cagggtaaac	agacttaattg	cagaatctaa	tgatggcaca	aattaaaaat	3120
cactaacgtg	ccctthtttag	tgtgaaaccc	agagagagca	catacaagcc	aaaaacaaat	3180

gctttatttt	acctaggaga	cattaacatt	cacctttacg	tgtttaagat	taatgcaatg	3240
ttaaataattg	tgaaaactgt	aactttgaat	ttcatgattt	ttatgtgaat	attccagggt	3300
ttaaaaaaac	ttgtaacatg	acatggctga	ataagataaa	aaaaaaatct	agccttttct	3360
cccttctggc	tcataattgc	gatttcgac	atthttgtta	aaaaacaaaa	cactgcaatg	3420
aattaaactt	aatattcttc	tatgttttag	agtaagttaa	aacaagataa	agtgaccaa	3480
gtaatttgaa	agattcaatg	acttttgctc	caacctagg	gcacaaggta	ccttgttctt	3540
taaattgggc	tttaatgaaa	atacttctcc	agaattctgg	ggatttaaga	aaaattatgc	3600
caaccaacaa	gggctttacc	atthttatgta	acatthttca	acgtgcaaa	aatgtgtgta	3660
tttctatttg	aagataaaaa	tcctcagcaa	aatccacatt	gcactgtcct	tcaaagatta	3720
gccttctttg	aactagttaa	gacactatta	agccaagcca	gtatctccct	gtaatgaatt	3780
cgthtttctc	ttaatthttcc	cctgtaattt	acactgggag	agctgggaaa	tatgtggatg	3840
taaatttctc	agccacagag	atgcaaagtt	atactgtggg	gaaaaaaac	ttgagttaaa	3900
tccttacata	ttttaggtht	tcattaactt	accaatgtag	ttttgttgga	ggccatthtt	3960
tttatgtcag	acttgaagag	ctattactag	aaaaatgcat	gacagttaa	gtaagthtgc	4020
atgacacaaa	aaaggtaact	aaatacaaat	tctgtthtga	ttccaacccc	caagtagaga	4080
gcgcacactt	tcaaacgtga	atacaaatcc	agagtagatc	tgcgctccta	cctacattgc	4140
ttatgatgta	cttaagtacg	tgctctaacc	atgtgagtct	agaaagactt	tactggggat	4200
cctggtacct	aaaacagctt	cacatggctt	aaaatagggg	accaatgtct	tttccaatct	4260
aagthccatt	tataataaag	tccatgtthc	atthtttaa	gacaatcctt	tcggthttaa	4320
accaggcacg	attacccaaa	caactcacia	cggtaaagca	ctgtgaatct	tctctgtthc	4380
gcaatcccaa	cttggtthct	gctcagaaac	cctccctctt	tccaatcgg	aattaaataa	4440
caaaaggaaa	aaacttaaga	tgcttcaacc	ccgtthcgtg	acactthgaa	aaaagaatca	4500
cctcttgcaa	acacccgctc	ccgaccccg	ccgtgaagc	ccggcgcca	gaggcctaag	4560
cgcggtgccc	cgccccacc	cgggagcgcg	ggcctcgtgg	tcagcgcatc	cgcggggaga	4620
aacaaaggcc	gcggcacggg	ggctcaagg	cactgcgcca	caccgcacgc	gcctaccccc	4680
gcgcggccac	gttaactggc	ggtcgcccga	gcctcgggac	agccggccgc	gcgcggccag	4740
gctcgcggac	gcgggaccac	gcgcgcccct	ccgggaggcc	caagtctcga	cccagccccg	4800
cgthggcgctg	ggggaggggg	cgctccgcc	ggaacgcggg	tgggggagg	gagggggaaa	4860
tgcgctthgt	ctcgaaatgg	ggcaaccgtc	gccacagctc	cctacccct	cgagggcaga	4920
gcagthcccc	cactaactac	cggtctggcc	gcgcgccagg	ccagccgcga	ggccaccgcc	4980
cgaccctcca	ctccttccc	cagctcccgg	cgcggggtcc	ggcgagaagg	ggaggggagg	5040
ggagcgggaga	accgggcccc	cgggacgcgt	gtggcatctg	aagcaccacc	agcgagcgag	5100
agctagagag	aaggaaagcc	accgacttca	ccgcctccga	gctgctccgg	gtcgcgggtc	5160
tgacgcgtct	ccggccctcc	gcgcctacag	ctcaagccac	atccgaagg	ggagggagcc	5220

gggagctgcg	cgcggggccg	ccggggggag	gggtggcacc	gcccacgccg	ggcggccacg	5280
aagggcgggg	cagcgggccc	gcgcgcggcg	gggggagggg	ccggcgccgc	gcccgcctggg	5340
aattggggcc	ctagggggag	ggcggaggcg	ccgacgaccg	cggcacttac	cgttcgcggc	5400
gtggcgcccc	gtggtcccca	aggggagggg	agggggaggc	ggggcgagga	cagtgaaccg	5460
agtctcctca	gcggtggcct	ttctgcttgg	cagcctcagc	ggctggcgcc	aaaaccggac	5520
tccgcccact	tcctcgcccc	ccggtgagag	ggtgtggaat	cctccagacg	ctgggggagg	5580
gggagttggg	agcttaaaaa	ctagtacccc	tttgggacca	ctttcagcag	cgaactctcc	5640
tgtacaccag	gggtcagttc	cacagacgcg	ggccaggggt	gggtcattgc	ggcgtgaaca	5700
ataatttgac	tagaagttga	ttcgggtgtt	tccggaaggg	gccgagtcaa	tccgccgagt	5760
tggggcacgg	aaaacaaaaa	gggaaggcta	ctaagatttt	tctggcgggg	gttatcattg	5820
gcgtaactgc	agggaccacc	tcccgggttg	agggggctgg	atctccaggc	tgcggaattaa	5880
gcccctcccc	tcggcggtta	tttcaaactg	cgcgacgttt	ctcacctgcc	ttcgccaagg	5940
caggggcccg	gacctatttc	caagaggtag	taactagcag	gactctagcc	ttccgcaatt	6000
cattgagcgc	atttacggaa	gtaacgtcgg	gtactgtctc	tggccgcaag	ggtgggagga	6060
gtacgcattt	ggcgtaaggt	ggggcgtaga	gccttccccg	cattggcggc	ggatagggcg	6120
tttacgcgac	ggcctgacgt	agcggaagac	gcgttagtg	gggggaaggt	tctagaaaag	6180
cggcggcagc	ggctctagcg	gcagtagcag	cagcgccggg	tcccgtgcgg	aggtgctcct	6240
cgcagagttg	tttctcgagc	agcggcagtt	ctcactacag	cgccaggacg	agtccggttc	6300
gtgttctgtc	gcggagatct	ctctcatctc	gctcggctgc	gggaaatcgg	gctgaagcga	6360
ctgagtccgc	gatggaggta	acgggtttga	aatcaatgag	ttattgaaaa	gggcatggcg	6420
aggccgttgg	cgcctcagtg	gaagtcggcc	agccgcctcc	gtgggagaga	ggcaggaaat	6480
cggaccaatt	cagtagcagt	ggggcttaag	gtttatgaac	ggggtcttga	gcggaggcct	6540
gagcgtacaa	acagcttccc	cacctcagc	ctcccggcgc	catttccctt	cactgggggt	6600
gggggatggg	gagctttcac	atggcggacg	ctgccccgct	ggggtgaaag	tggggcgcg	6660
aggcggaat	tcttattccc	tttctaaagc	acgctgcttc	gggggccacg	gcgtctcctc	6720
ggcgagcggt	tcggcgggca	gcaggtcctc	gtgagcgagg	ctgcggagct	tcccctcccc	6780
ctctctcccc	ggaaccgatt	tggcggccgc	cattttcatg	gctcgccttc	ctctcagcgt	6840
tttctttata	actcttttat	tttcttagtg	tgttttctct	atcaagaagt	agaagtgggt	6900
aactattttt	tttttcttct	cgggctgttt	tcatatcggt	tcgaggtgga	tttgaggtgt	6960
tttgtgagct	tggatcttta	gagtcctg	cacctcatta	aaggcgctca	gccttcccct	7020
cgatgaaatg	gcgccattgc	gttcggaagc	cacaccgaag	agcggggagg	gggggtgctc	7080
cgggttttgc	ggcccgggtt	cagagaagat	atcaccaccc	agggcgctcg	gccgggttca	7140
atgcgagccg	taggacaaag	aaaccatttt	atgtttttcc	tgtctttttt	ttcctttgag	7200
taacggtttt	atctgggtct	gcagtcagta	aaacgacaga	tgaaccgcgg	caaaataaac	7260

ataaattgga agccatcggc cacgaggggc agggacgaag gtggttttct gggcggggga	7320
gggatattcg cgtcagaatc ctttactggt ctttaaggatt ccgtttaagt tgtagagctg	7380
actcatttta agtaatgttg ttactgagaa gtttaaccct tacgggacag atccatggac	7440
ctttatagat gattacgagg aaagtgaat aacgattttg tccttagtta tacttcgatt	7500
aaaacatggc ttcagaggct ctttcctgta atgcgtatgg attgatgtgc aaaactgttt	7560
tgggcctggg ccgctctgta tttgaacttt gttacttttc tcattttgtt tgcaatcttg	7620
gttgaacatt acattgataa gcataaggtc tcaagcgaag ggggtctacc tggttatitt	7680
tctttgaccc taagcacgtt tataaaataa cattgtttaa aatcgatagt ggacatcggg	7740
taagtttgga taaattgtga ggtaagtaat gagtttttgc tttttgtag tgatttghaa	7800
aacttgttat aaatgtacat tatccgtaat ttcagtttag agataacctg tgtgctgacg	7860
acaattaaga ataaaaacta gctgaaaaaa tgaaaataac tatcgtgaca agtaaccatt	7920
tcaaaagact gctttgtgtc tcataggagc tagtttgatc atttcagtta attttttctt	7980
taatttttac gagtcatgaa aactacagga aaaaaaatct gaactgggtt ttaccactac	8040
tttttaggag ttgggagcat gcgaatggag ggagagctcc gtagaactgg gatgagagca	8100
gcaattaatg ctgcttgcta ggaacaaaaa ataattgatt gaaaattacg tgtgactttt	8160
tagtttgcat tatgcgtttg tagcagttgg tcctggatat cactttctct cgtttgaggt	8220
tttttaacct agttaacttt taagacaggt ttccttaaca ttcataagt cccagaatac	8280
agctgtgtag tacagcatat aaagatttca gctctgaggt ttttcctatt gacttgga	8340
attgttttgt gcctgtcgct tgccacatgg ccaatcaagt aagcttatcg ataccggtgg	8400
cgcgcccaatt gaattaagat ctggcccaat gggccgtacg aattcgagct cggtagccgg	8460
ggatcctgat ctaatagtaa tcaattacgg ggtcattagt tcatagccca tatatggagt	8520
tccgcgttac ataaacttac gtaaatggcc cgcctggctg accgcccaac gacccccgcc	8580
cattgacgtc aataatgacg tatgttccca tagtaacgcc aatagggact ttccattgac	8640
gtcaatgggt ggagtattta cggtaaactg cccacttggc agtacatcaa gtgtatcata	8700
tgccaagtac gccccctatt gacgtcaatg acggtaaatg gccgcctgg cattatgccc	8760
agtacatgac cttatgggac tttcctactt ggcagtacat ctacgtatta gtcacgcta	8820
ttaccatggg gatgcggttt tggcagtaca tcaatgggag tggatagcgg tttgactcac	8880
ggggatttcc aagtctccac cccattgacg tcaatgggag tttgttttgg caccaaatc	8940
aacgggactt tccaaaatgt cgtaacaact ccgccccatt gacgcaaatg ggcggtaggc	9000
gtgtacggtg ggaggtctat ataagcagag ctggttttagt gaaccgtcag atccgtcgcc	9060
accatggtga gcaagggcga ggagctgttc accggggtgg tgccatcct ggtcgagctg	9120
gacggcgacg taaacggcca caagttcagc gtgtccggcg agggcgaggg cgatgccacc	9180
tacggcaagc tgaccctgaa gttcatctgc accaccggca agctgcccgt gccctggccc	9240
accctcgtga ccaccctgac ctacggcgtg cagtgccttca gccgctaccc cgaccacatg	9300

aagcagcacg acttcttcaa gtccgccatg cccgaaggct acgtccagga gcgcaccatc 9360
ttcttcaagg acgacggcaa ctacaagacc cgcgccgagg tgaagtctga gggcgacacc 9420
ctggtgaacc gcatcgagct gaagggcatc gacttcaagg aggacggcaa catcctgggg 9480
cacaagctgg agtacaacta caacagccac aacgtctata tcatggccga caagcagaag 9540
aacggcatca aggtgaactt caagatccgc cacaacatcg aggacggcag cgtgcagctc 9600
gccgaccact accagcagaa ccccccatc ggcgacggcc ccgtgctgct gcccagacaac 9660
cactacctga gcaccagtc cgccctgagc aaagacccca acgagaagcg cgatcacatg 9720
gtcctgctgg agttctgtgac cgccgcccggg atcactctcg gcatggacga gctgtacaag 9780
taaagcggcc gcgactctag atcataatca gccataccac atttgtagag gttttacttg 9840
ctttaaaaaa cctccacac ctccccctga acctgaaaca taaaatgaat gcaattgttg 9900
ttgttaactt gtttattgca gcttataatg gttacaaata aagcaatagc atcacaaatt 9960
tcacaaataa agcatttttt tcaactgcatt ctagttgtgg tttgtccaaa ctcatcaatg 10020
tatcttaact agagtcgacc tgaggcatg caagcttacc ggtggcgcg ccaattgttaa 10080
ttaagatctg gcccaatggg ccgtacgaat tccttaggct accgggtagg ggaggcgctt 10140
ttccaaggc agtctggagc atgcgcttta gcagccccgc tgggcaactg gcgctacaca 10200
agtggcctct ggcctcgac acattccaca tccaccggcc ggtaggcgcc aaccggctcc 10260
gttcttttgt gggcccttcg cgccaccttc tactcctccc ctagtcagga agttccccc 10320
cgccccgag ctgcgctgt gcaggacgtg acaaatggaa gtagcacgtc tcaactagtct 10380
cgtgcagatg gacagcaccg ctgagcaatg gaagcgggta ggcctttggg gcagcggcca 10440
atagcagctt tgctccttcg ctttctgggc tcagaggctg ggaaggggtg ggtccggggg 10500
cgggctcagg ggcgggctca ggggcggggc gggcgcccga aggtcctccg gaggccgggc 10560
attctgcacg cttcaaaagc gcacgtctgc cgcgctgttc tcctcttcct catctccggg 10620
cctttcgacc agcttaccat gaccgagtag aagcccacgg tgcgcctcgc caccgcgac 10680
gacgtcccca gggccgtacg caccctcgcc gccgcgttcg ccgactaccc cgccacgcgc 10740
cacaccgtcg atccggaccg ccacatcgag cgggtcaccg agctgcaaga actcttcctc 10800
acgcgcgtcg ggctcgacat cggcaagggtg tgggtcgcgg acgacggcg cgcggtggcg 10860
gtctggacca cgccggagag cgtcgaagcg ggggcgggtg tcgccgagat cggcccgcgc 10920
atggccgagt tgagcggttc ccggtggcc gcgcagcaac agatggaagg ctccttggcg 10980
ccgcaccggc ccaaggagcc cgcggtgttc ctggccaccg tcggcgcttc gcccgaccac 11040
cagggaagg gtctgggcag cgccgtcgtg tccccggag tggaggcggc cgagcgcgcc 11100
ggggtgcccg ccttcctgga gacctcgcg ccccgcaacc tccccctcta cgagcggctc 11160
ggcttcaccg tcaccgccga cgtcgagggtg cccgaaggac cgcgcacctg gtgcatgacc 11220
cgcaagcccg gtgcctgacg cccgccccac gaccgcagc gcccgaccga aaggagcgca 11280
cgaccccatg catcgtagac gaaatgaccg accaagcgac gcccaacctg ccatcacgag 11340

atttcgattc caccgcccgc ttctatgaaa ggttgggctt cggaatcgtt ttccgggacg 11400
 ccggctggat gatcctccag cgcggggatc tcatgctgga gttcttcgcc caccctaggg 11460
 ggaggctaac tgaaacacgg aaggagacaa taccggaagg aacccgcgct atgacggcaa 11520
 taaaaagaca gaataaaacg cacggtgttg ggtcgtttgt tcataaacgc ggggttcggt 11580
 cccagggctg gactctgtc gatacccccac cgagacccca ttggggccaa tacgcccgcg 11640
 tttcttcctt tccccaccc cccccccaa gttcgggtga agggccaggg ctgcagcca 11700
 acgtcggggc ggcaggcccc cagcttttgt tcccttagt gagggttaat ttcgagcttg 11760
 gcgtaatcat ggcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac 11820
 aacatacgag ccggaagcat aaagtgtaaa gcctggggtg cctaatagag gagctaactc 11880
 acattaattg cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc gtgccagcat 11940
 cgcgagcact tttcggggaa atgtgcgcgg aaccctatt tgtttatatt tctaaatata 12000
 ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 12060
 aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt 12120
 ttgccttcct gtttttctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 12180
 gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 12240
 ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagtctctg tatgtggcgc 12300
 ggtattatcc cgtattgacg ccgggcaaga gcaactcggg cgccgcatc actattctca 12360
 gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 12420
 aagagaatta tgcagtgtg ccataaccat gagtgataac actgcggcca acttacttct 12480
 gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt 12540
 aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga 12600
 caccacgatg cctgtagcaa tggcaacaac gttgcgcaa ctattaactg gcgaactact 12660
 tactctagct tcccggaac aattaataga ctggatggag gcggataaag ttgcaggacc 12720
 acttctgctc tcggcccttc cggttggtg gtttattgct gataaatctg gagccggtga 12780
 gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct ccggtatcgt 12840
 agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgtga 12900
 gataggtgcc tactgatta agcattggtg actgtcagac tcgcgacact gcattaatga 12960
 atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttcgc ttccctcgtc 13020
 actgactcgc tgcgctcggg cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg 13080
 gtaatacggg tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc 13140
 cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc 13200
 cccctgacg agcatcaca aaatcgacgc tcaagtcaga ggtggcgaaa ccgacagga 13260
 ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc gttccgacc 13320
 ctgcccgtta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat 13380

```

agctcacgct gtaggtatct cagttcggtg taggtcggtc gctccaagct gggctgtgtg 13440
cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc 13500
aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga 13560
gcgaggtagt taggcggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact 13620
agaaggacag ttttgggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt 13680
ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag 13740
cagcagatta cgcgagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg 13800
tctgacgctc agtgaacga aaactca 13827

```

<210> 19

<211> 8585

<212> ADN

5 <213> Artificial

<220>

<223> Vector pRPS11 1005 EGFP

<400> 19

```

cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga attgggtacc 60
gggccccccc tcgaagttaa attcgccctt aagactgttt gcctcatgcc tgcttggcct 120
gcccttcctc cgccccaac taggggaagt gggaccaaag gttccttagg cactgctcct 180
gtgggtagag gggacattag agagctgaca gcgcaccacc tgcattgagt tttattaaag 240
tgcaaacatc gggatgaatc agttgagctt cagtgttgaa aatgagtagc agggctgccc 300
caccacctg accaagtacc ctattctgca gctatgaaaa tgagatctgc acatgagctg 360
gggttcacaa gtgcacactt ggagcactgc cttgctcctt cccagcagac cacaagcag 420
tatttttctg gaggatttta tgtgctaata aattatttga cttaagtgtg tacgatgtgt 480
gctgtgcaga gaggggcaga gggcaccagc aggtcatctg catggggggc ccctttgggt 540
gaatccttgc tcacgggata ggctttgttg ctcaaaagt ggcagatatac atcttgggtc 600
ctgtcctaga tgggtgttact gtaagtcagc accaagatac aagagctggg acctggactg 660
taggaggtca ggccatgaca caaaggctgg gactaaaggc atttaccacg cctgagtctt 720
ctggttcttt aaacatcaaa tccttccggg ggctggcgag atggctcagt ggtaagagc 780
acagactgct cttacgaagg atccgagttc aaatcccagc aaccaaattg tgcctaacaa 840
ctatccataa tgaaatctga tgccctcttc tggagtatct gagaacagct acagtgtact 900
tacatataat cttaaaaatg cttcccatgt taaccaccac tagagttttt attacagcta 960
gctgacctgg aagccaagtc cttatgcctc cgtgagtgct ggggttaaaa agatccagca 1020
ccactcaaaa tgtcaatcta ttttgaaaat atgctttata ctgttctagc ccatctgtgc 1080
agggctagaa cggtgaatac gagaaactga cacaagcttt tgccacctgg ctaaattggt 1140
cctctattac ctggggtggg cacctaaggc tagacactca tccacgagta gtcaggacat 1200
aaacccatca aagtgtgggt agacgcgcag cctgagatac tgtcaacaaa ggacatgcga 1260

```

ccttggtgac	gtcggccttt	aataaaagga	agaaagggtt	actattcggt	cgacgtggc	1320
tgctcctgac	atcgtatggc	agatactctg	ctgtaaagcg	gttcacccct	ttcttgagac	1380
ccgctctgca	cggccgcttc	tctctgga	ctgaatccca	gcacgtgttt	cccaaccctg	1440
acggcacgcc	ttctccgccc	taagcctcgc	cgtaccacat	gatgcacgtt	tcctccacat	1500
cgtgctcctg	aaatctcgcg	agatgatagg	atcttcccg	cccttagtcc	tccccgtca	1560
tggcggcgta	cggacagtcc	caggaacgcg	ggctctcgcc	ggaagtacct	cccacctccg	1620
tgaggataac	cccgcgtcac	ttccgccccg	acctcgcgtg	gtgaataagg	aagccgggag	1680
cggccctgcc	tctccctttc	tccggcggcc	gggaagatgg	cggacattca	ggttcgagcg	1740
tttagttgct	ttccccgac	gcttcggtgt	ggagcgtatc	ccttggcgtc	ctcgttgtct	1800
tacgcattag	ctgaagcgag	gatgcctgcg	aatgccttcg	tctcaggcgg	ctcggaaatc	1860
cgggctctac	gcagtaatgg	ggccctggc	gcttcgggag	ttggttctta	aagctcagag	1920
cttaacgggt	gagggattgt	ggcgggagga	gggcacccct	cggcgcggga	gtcctgcggc	1980
ggcagagccg	gggacactgg	gtaaagcagg	ttttttcccc	ttgatggaga	ctgaggcccc	2040
gacctcgtgc	gctctacggc	agggctgcgg	tcccgacctc	gctgtagttt	tcagtgtgag	2100
cgcagctctg	gcctcgatga	gcttaggctt	gtcttaaaact	tgccatccct	cctcaacctc	2160
aaccgggatg	acagatccgg	cccaccaggc	tcggctacgt	ggacataagc	ttgaatcccc	2220
aatgagtggg	tttgtatgtt	ttggagggtc	agtctggctg	aaaagctctt	ttgatctca	2280
gccgtgagtt	ctgcaggctg	tggagggtgt	agatgggacg	cagtgtgtga	gctaaactag	2340
acttgggggtg	gttggagagc	cctgaccagc	cggttttggc	gattggggca	aataagggtt	2400
aaggtaggaa	ggaagaaata	ttgtctctga	tttccttgaa	ctttacctgc	aacctcacca	2460
aattctcatc	cctacagacg	gagcgtgctt	acccaaagca	gcctacgata	tttcaaaaca	2520
agaagcgggt	tctgctggga	gaaaccggca	aggaaaaact	ccctcggtac	tacaagaata	2580
tcggtctagg	cttcaagacg	cctaaagagg	tacaggacct	tccagcagat	gagatccctg	2640
ctgccctgca	cgtgtgggag	cacagccacc	ccgccccctt	cacagtggct	tcccatgggc	2700
ccctgggaat	tgtagtatgg	gccctgaggc	gtcatccttg	gttctgttta	ggaagtggta	2760
atctaaacct	cactttctta	actttgcagg	ctattgaggg	tacctacata	gacaagaaat	2820
gccccctcac	tggtaacgtc	tccatccgag	gtcggatcct	gtctgggtgag	tgggatgttg	2880
gaagggtggt	tctaggttcc	tgcgccagg	ggcgctggca	agtgatgtct	gttctcacga	2940
tggctcttcag	atgtcctcta	gggcactgct	gagacagcca	gttgacaaag	ctgatgccat	3000
aaatggagct	tcttgggagc	cccgttcaac	tgactcctac	ctgctaacac	ctttctgtta	3060
ctctcccagg	tgtcgtgacg	aagatgaaga	tgcagaggac	cattgtcatc	caagggcgaa	3120
acatttaaat	ctagaagctt	atcgataccg	gtggcgcgcc	aattgaatta	agatctggcc	3180
caatgggccg	tacgaattcg	agctcggtag	ccggggatcc	tgatctaata	gtaatcaatt	3240
acggggtcat	tagttcatag	cccatatatg	gagttccgcg	ttacataact	tacggtaaat	3300

```

ggcccgcctg gctgaccgcc caacgacccc cgcccattga cgtcaataat gacgtatggt 3360
cccatagtaa cgccaatagg gactttccat tgacgtcaat gggtaggagta tttagcgtaa 3420
actgcccact tggcagtaca tcaagtgtat catatgccaa gtacgcccc tattgacgtc 3480
aatgacggta aatggcccgc ctggcattat gcccagtaca tgaccttatg ggactttcct 3540
acttggcagt acatctacgt attagtcacg gctattacca tggtagatgcg gttttggcag 3600
tacaatcaatg ggcgtggata gcggtttgac tcacggggat ttccaagtct ccacccatt 3660
gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg actttccaaa atgtcgtaac 3720
aactccgccc cattgacgca aatgggcggg aggcgtgtac ggtgggaggt ctatataagc 3780
agagctggtt tagtgaaccg tcagatccgt cgccaccatg gtgagcaagg gcgaggagct 3840
gttcaccggg gtggtgccc tcctggtcga gctggacggc gacgtaaagc gccacaagtt 3900
cagcgtgtcc ggcgagggcg agggcgatgc cacctacggc aagctgacct tgaagttcat 3960
ctgcaccacc ggcaagctgc ccgtgccctg gccaccctc gtgaccacc tgacctacgg 4020
cgtgcagtgc ttcagccgct accccgacca catgaagcag cacgacttct tcaagtccgc 4080
catgcccga ggtacgtcc aggagcgac catcttctt aaggacgacg gcaactacaa 4140
gacccgcgcc gaggtgaagt tcgagggcga caccctggtg aaccgcatcg agctgaagg 4200
catcgacttc aaggaggacg gcaacatcct ggggcacaag ctggagtaca actacaacag 4260
ccacaacgct tatatcatgg ccgacaagca gaagaacggc atcaaggtag acttcaagat 4320
ccgccacaac atcgaggacg gcagcgtgca gctcgccgac cactaccagc agaacacccc 4380
catcggcgac ggccccgtgc tgctgcccga caaccactac ctgagcacc agtccgccct 4440
gagcaaagac cccaacgaga agcgcgatca catggtcctg ctggagtctg tgaccgccgc 4500
cgggatcact ctcggcatgg acgagctgta caagtaaagc ggccgcgact ctagatcata 4560
atcagccata ccacatttgt agaggtttta cttgcttta aaaacctccc acacctccc 4620
ctgaacctga aacataaaat gaatgcaatt gttgtgtta acttgtttat tgcagcttat 4680
aatggttaca aataaagcaa tagcatcaca aatttcacaa ataaagcatt tttttactg 4740
cattctagtt gtggtttgtc caaactcatc aatgtatctt aactagagtc gacctgcagg 4800
catgcaagct taccggtggc gcgcgcgcca attgttaatt aagatctggc ccaatgggccc 4860
gtacgaattc cttaggctac cgggtagggg aggcgctttt cccaaggcag tctggagcat 4920
gcgcttttagc agccccgctg ggcacttggc gctacacaag tggcctctgg cctcgcacac 4980
attccacatc caccggccgg taggcgcca ccggctccgt tctttggtgg ccccttcgcg 5040
ccaccttcta ctctcccct agtcaggaag ttccccccg cccgcagct cgcgtcgtgc 5100
aggacgtgac aaatggaagt agcacgtctc actagtctcg tgcagatgga cagcaccgct 5160
gagcaatgga agcgggtagg ctttgggggc agcggccaat agcagctttg ctcttcgct 5220
ttctgggctc agaggctggg aaggggtggg tccggggcg ggctcaggg cgggctcagg 5280
ggcggggcg gcgcccgaag gtcctccgga gggcgccat tctgcacgct tcaaaagcgc 5340

```

acgtctgccg	cgctgttctc	ctcttcctca	tctccgggcc	tttcgaccag	cttaccatga	5400
ccgagtacaa	gcccacggtg	cgcttcgcca	cccgcgacga	cgccccagg	gccgtacgca	5460
ccctcgccgc	cgcttcgcc	gactaccccg	ccacgcgcca	caccgtcgat	ccggaccgcc	5520
acatcgagcg	ggtcaccgag	ctgcaagaac	tcttcctcac	gcgcgtcggg	ctcgacatcg	5580
gcaagggtgtg	ggtcgcgga	gacggcgccg	cggtggcggg	ctggaccacg	ccggagagcg	5640
tcgaagcggg	ggcgggtgtc	gccgagatcg	gcccgcgcat	ggccgagttg	agcggttccc	5700
ggctggccgc	gcagcaacag	atggaaggcc	tcctggcgcc	gcaccggccc	aaggagcccg	5760
cggtggttct	ggccaccgtc	ggcgtctcgc	ccgaccacca	gggcaagggt	ctgggcagcg	5820
ccgtcggtct	ccccggagtg	gaggcgggcg	agcgcgcccg	ggtgcccgcc	ttcctggaga	5880
cctccgcgcc	ccgcaacctc	cccttctacg	agcggctcgg	cttcaccgtc	accgccgacg	5940
tcgaggtgcc	cgaaggaccg	cgcacctggt	gcatgaccgc	caagcccggg	gcctgacgcc	6000
cgccccacga	cccgcagcgc	ccgaccgaaa	ggagcgcacg	accccatgca	tcgtagacga	6060
aatgaccgac	caagcgacgc	ccaacctgcc	atcacgagat	ttcgattcca	ccgccgcctt	6120
ctatgaaaagg	ttgggtcttcg	gaatcgtttt	ccgggacgcc	ggctggatga	tcctccagcg	6180
cggggatctc	atgctggagt	tcttcgcccc	ccctaggggg	aggctaactg	aaacacggaa	6240
ggagacaata	ccggaaggaa	cccgcgctat	gacggcaata	aaaagacaga	ataaaacgca	6300
cggtgttggg	tcgtttgttc	ataaacgcgg	ggttcgggtc	cagggctggc	actctgtcga	6360
tacccccaccg	agacccatt	ggggccaata	cgcccgcggt	tcttcctttt	ccccacccca	6420
cccccaagt	tcgggtgaag	gcccgaggct	cgagccaac	gtcggggcgg	caggccccca	6480
gcttttgttc	cctttagtga	gggttaattt	cgagcttggc	gtaatcatgg	tcatagtgt	6540
ttcctgtgtg	aaattgttat	ccgctcacia	ttccacacia	catacgagcc	ggaagcataa	6600
agtgtaaaagc	ctgggggtgcc	taatgagtga	gctaactcac	attaattgcg	ttgcgctcac	6660
tgcccgcctt	ccagtcggga	aacctgtcgt	gccagcatcg	cgagcacttt	tcgggggaaat	6720
gtgcgcggaa	cccctatttg	tttatttttc	taaatacatt	caaatatgta	tccgctcatg	6780
agacaataac	cctgataaat	gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	6840
catttcgctg	tcgcccttat	tccctttttt	gcggcatttt	gccttcctgt	ttttgctcac	6900
ccagaaacgc	tggtgaaagt	aaaagatgct	gaagatcagt	tggtgacacg	agtgggttac	6960
atcgaactgg	atctcaacag	cggtaaagatc	cttgagagtt	ttcgccccga	agaacgtttt	7020
ccaatgatga	gcacttttaa	agttctgcta	tgtggcgcg	tattatcccc	tattgacgcc	7080
gggcaagagc	aactcgggtc	ccgcatacac	tattctcaga	atgacttggg	tgagtactca	7140
ccagtcacag	aaaagcatct	tacggatggc	atgacagtaa	gagaattatg	cagtgtgtcc	7200
ataacatga	gtgataacac	tgcgccaac	ttacttctga	caacgatcgg	aggaccgaag	7260
gagctaaccg	cttttttgca	caacatgggg	gatcatgtaa	ctcgccttga	tcgttgggaa	7320
ccggagctga	atgaagccat	accaaacgac	gagcgtgaca	ccacgatgcc	tgtagcaatg	7380

gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc ccggcaacaa	7440
ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg	7500
gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg cggatatcatt	7560
gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt	7620
caggcaacta tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag	7680
cattggtaac tgcagactc gcgacactgc attaatgaat cggccaacgc gcggggagag	7740
gcggtttgcg tattgggagc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg	7800
ttcggctgcg gcgagcggta tcagctcact caaaggcggg aatacgggta tccacagaat	7860
caggggataa cgaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta	7920
aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa	7980
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc	8040
cccttggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt	8100
ccgcctttct cccttcggga agcgtggcg tttctcatag ctacgctgt aggtatctca	8160
gttcggtgta ggtcgctgc tccaagctgg gctgtgtgca cgaaccccc gttcagccc	8220
accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaga cagacttat	8280
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta	8340
cagagtcttt gaagtgggtg cctaactacg gctacactag aaggacagta tttggtatct	8400
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac	8460
aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaaa	8520
aaggatctca agaagatcct ttgatctttt ctacgggggc tgacgctcag tggaacgaaa	8580
actca	8585
<210> 20	
<211> 5546	
<212> ADN	
<213> Artificial	
<220>	
<223> Sector pCET 1005 EGFP	
<400> 20	

cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga attgggtacc	60
gggccccccc tcgaagttaa aacatttaaa tctagaagct tatcgatacc ggtggcgcgc	120
caattgaatt aagatctggc ccaatgggccc gtacgaattc gagctcggta cccggggatc	180
ctgatctaata agtaataaat tacgggggtca ttagttcata gcccatatat ggagttccgc	240
gttacataac ttacggtaaa tggcccgctt ggctgaccgc ccaacgacct ccgcccattg	300
acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca ttgacgtcaa	360
tgggtggagt atttacggta aactgcccac ttggcagtac atcaagtgt tcatatgcc	420
agtacgcccc ctattgacgt caatgacggg aaatggcccc cctggcatta tgcccagtac	480

atgaccttat	gggactttcc	tacttggcag	tacatctacg	tattagtcac	cgctattacc	540
atggtgatgc	ggttttggca	gtacatcaat	gggcgtggat	agcggtttga	ctcacgggga	600
tttccaagtc	tccaccccat	tgacgtcaat	gggagtttgt	tttggcacca	aaatcaacgg	660
gactttccaa	aatgtcgtaa	caactccgcc	ccattgacgc	aaatgggcgg	taggcgtgta	720
cggtgggagg	tctatataag	cagagctggt	ttagtgaacc	gtcagatccg	tcgccaccat	780
ggtgagcaag	ggcgaggagc	tgttcaccgg	ggtggtgccc	atcctggtcg	agctggacgg	840
cgacgtaaac	ggccacaagt	tcagcgtgtc	cggcgagggc	gagggcgatg	ccacctacgg	900
caagctgacc	ctgaagttca	tctgcaccac	cggcaagctg	cccgtgccct	ggccccacct	960
cgtgaccacc	ctgacctacg	gcgtgcagtg	cttcagccgc	taccccgacc	acatgaagca	1020
gcacgacttc	ttcaagtccg	ccatgcccga	aggctacgtc	caggagcgca	ccatcttctt	1080
caaggacgac	ggcaactaca	agacccgcgc	cgaggtgaag	ttcgagggcg	acaccctggt	1140
gaaccgcatc	gagctgaagg	gcatcgactt	caaggaggac	ggcaacatcc	tggggcacia	1200
gctggagtac	aactacaaca	gccacaacgt	ctatatcatg	gccgacaagc	agaagaacgg	1260
catcaaggty	aacttcaaga	tccgccacaa	catcgaggac	ggcagcgtgc	agctcgccga	1320
ccactaccag	cagaacaccc	ccatcggcga	cggccccgtg	ctgctgcccg	acaaccacta	1380
cctgagcacc	cagtccgccc	tgagcaaaga	ccccaacgag	aagcgcgac	acatggctct	1440
gctggagttc	gtgaccgccg	ccgggatcac	tctcggcatg	gacgagctgt	acaagtaaag	1500
cggccgcgac	tctagatcat	aatcagccat	accacatttg	tagaggtttt	acttgcttta	1560
aaaaacctcc	cacacctccc	cctgaacctg	aaacataaaa	tgaatgcaat	tgttgttggt	1620
aacttgttta	ttgcagctta	taatggttac	aaataaagca	atagcatcac	aaatttcaca	1680
aataaagcat	ttttttcact	gcattctagt	tgtggtttgt	ccaaactcat	caatgtatct	1740
taactagagt	cgacctgcag	gcatgcaagc	ttaccggtgg	cgcgcgccg	aattgttaat	1800
taagatctgg	cccaatgggc	cgtacgaatt	ccttaggcta	ccgggtaggg	gaggcgcttt	1860
tccaaggca	gtctggagca	tgcgctttag	cagccccgct	gggcacttgg	cgctacacaa	1920
gtggcctctg	gcctcgaca	cattccacat	ccaccggccg	gtaggcgcca	accggctccg	1980
ttctttggtg	gccccttcgc	gccaccttct	actcctcccc	tagtcaggaa	gttccccccc	2040
gccccgcagc	tcgcgtcgtg	caggacgtga	caaatggaag	tagcacgtct	cactagtctc	2100
gtgcagatgg	acagcaccgc	tgagcaatgg	aagcgggtag	gcctttgggg	cagcggccaa	2160
tagcagcttt	gtccttcgc	tttctgggct	cagaggctgg	gaaggggtgg	gtccgggggc	2220
gggctcaggg	gcgggctcag	gggcggggcg	ggcgcccgaa	ggctctccgg	aggcccgga	2280
ttctgcacgc	ttcaaaagcg	cacgtctgcc	gcgtgttct	cctcttcctc	atctccgggc	2340
ctttcgacca	gcttaccatg	accgagtaca	agcccacggt	gcgcctcgcc	accgcgacg	2400
acgtccccag	ggccgtacgc	accctcgccg	ccgcgttcgc	cgactacccc	gccacgcgcc	2460
acaccgtcga	tccggaccgc	cacatcgagc	gggtcaccga	gctgcaagaa	ctcttcctca	2520

cgcgcgtcgg gctcgacatc ggcaagggtgt gggtcgcgga cgacggcgcc gcggtggcgg	2580
tctggaccac gccggagagc gtcgaagcgg gggcgggtgtt cgccgagatc ggcccgcgca	2640
tggccgagtt gagcggttcc cggctggccg cgcagcaaca gatggaaggc ctcttggcgc	2700
cgcaccggcc caaggagccc gcgtggttcc tggccaccgt cggcgtctcg cccgaccacc	2760
agggcaaggg tctgggcagc gccgtcgtgc tccccggagt ggaggcggcc gagcgcgccg	2820
gggtgcccgc cttcctggag acctccgcgc cccgcaacct ccccttctac gagcggctcg	2880
gcttcaccgt caccgccgac gtcgaggtgc ccgaaggacc gcgcacctgg tgcattgacc	2940
gcaagcccgg tgcttgacgc ccgcccacg acccgagcg cccgaccgaa aggagcgac	3000
gaccccatgc atcgtagacg aaatgaccga ccaagcgacg cccaacctgc catcacgaga	3060
tttcgattcc accgccgcct tctatgaaag gttgggcttc ggaatcgttt tccgggacgc	3120
cggctggatg atcctccagc gcggggatct catgctggag ttcttcgccc accctagggg	3180
gaggctaact gaaacacgga aggagacaat accggaagga acccgcgcta tgacggcaat	3240
aaaaagacag aataaacgc acggtgttgg gtcgtttgtt cataaacgcg gggttcggtc	3300
ccagggttgg cactctgtcg atacccacc gagacccat tggggccaat acgcccgcgt	3360
ttcttccttt tccccacccc acccccaag ttcgggtgaa ggcccagggc tcgcagccaa	3420
cgtcggggcg gcaggccccc agcttttgtt cccttagtg agggttaatt tcgagcttgg	3480
cgtaatcatg gtcattagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca	3540
acatacgagc cggaagcata aagtgtaaag cctggggtgc ctaatgagt agctaactca	3600
cattaattgc gttgcgctca ctgccgctt tccagtcggg aaacctgtcg tgccagcatc	3660
gcgagcactt ttcggggaaa tgtgcgcgga acccctattt gtttttttt ctaaatacat	3720
tcaaataatgt atccgctcat gagacaataa cctgataaa tgcttcaata atattgaaaa	3780
aggaagagta tgagtattca acatttccgt gtcgccctta ttcccttttt tgccgcattt	3840
tgctttcctg tttttgctca ccagaaaacg ctggtgaaag taaaagatgc tgaagatcag	3900
ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaatg ccttgagagt	3960
tttcgccccg aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg	4020
gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca ctatttctcag	4080
aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta	4140
agagaattat gcagtgtctg cataaccatg agtgataaca ctgcggccaa cttacttctg	4200
acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta	4260
actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac	4320
accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt	4380
actctagctt cccggcaaca attaatagac tggatggagg cgataaagt tgcaggacca	4440
cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg agccgggtgag	4500
cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta	4560

gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag	4620
ataggtgcct cactgattaa gcattggttaa ctgtcagact cgcgacactg cattaatgaa	4680
tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca	4740
ctgactcgct gcgctcggtc gttcggctgc ggcgagcggc atcagctcac tcaaaggcgg	4800
taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc	4860
agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc	4920
cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac	4980
tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgacct	5040
tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata	5100
gctcacgctg taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc	5160
acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca	5220
acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag	5280
cgaggtatgt aggcggtgct acagagttct tgaagtggc gcctaactac ggctacacta	5340
gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg	5400
gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc	5460
agcagattac gcgcagaaaa aaaggatctc aagaagatcc ttgatcttt tctacggggt	5520
ctgacgctca gtggaacgaa aactca	5546

REIVINDICACIONES

1. Un polinucleótido aislado que comprende:
 - a. una isla CpG extendida exenta de metilación que comprende más de 3000 nucleótidos contiguos desde la
 - 5 región promotora de un gen rps3 de mamífero;
 - b. un promotor heterólogo; y
 - c. una secuencia de ácidos nucleicos transcribible adyacente al promotor heterólogo,
 en el que la transcripción de dicha secuencia de ácidos
 - 10 nucleicos transcribible desde dicho promotor heterólogo se mejora en presencia de dicha isla CpG extendida exenta de metilación con respecto a la transcripción en ausencia de dicha isla CpG extendida exenta de metilación.
2. El polinucleótido según la reivindicación 1, en el
 - 15 que la isla CpG extendida exenta de metilación comprende adicionalmente uno o varios exones del gen rps3 de mamífero.
3. El polinucleótido según la reivindicación 1, en el que el gen rps3 de mamífero es un gen rps3 humano.
 - 20 4. El polinucleótido según la reivindicación 1, en el que el gen rps3 de mamífero es un gen rps3 de roedor.
5. El polinucleótido según la reivindicación 1, en el que el gen de roedor es un gen rps3 de ratón.
 6. El polinucleótido según la reivindicación 5, que
 - 25 comprende la secuencia de nucleótidos de SEQ ID NO: 1.
7. El polinucleótido según la reivindicación 1, en el que el promotor heterólogo es un promotor constitutivo.
 8. El polinucleótido según la reivindicación 7, en el que el promotor constitutivo se selecciona del grupo
 - 30 constituido por el promotor temprano/inmediato de citomegalovirus, SV40, EF-1 α , las LTR del virus del sarcoma de Rous (RSV), o las LTR del VIH2.
9. El polinucleótido según la reivindicación 10, en el que el promotor constitutivo es el promotor
 - 35 temprano/inmediato de citomegalovirus.
10. El polinucleótido según la reivindicación 9, en el

que el promotor constitutivo es un promotor temprano/inmediato de citomegalovirus en conejillo de indias.

11. El polinucleótido según la reivindicación 9, en el
5 que el promotor constitutivo es un promotor temprano/inmediato de citomegalovirus en ratón.

12. El polinucleótido según la reivindicación 1, en el que el promotor heterólogo es un promotor específico de tejido.

10 13. El polinucleótido según la reivindicación 12, en el que el promotor heterólogo es un promotor selectivo de tumores.

14. El polinucleótido según la reivindicación 13, en el que el promotor se selecciona del grupo constituido por
15 promotores del antígeno carcinoembrionario (CEA), antígeno prostático específico (PSA), ciclooxigenasa-2 (COX-2), alfa-fetoproteína (AFP), tirosinasa, y los promotores basados en los factores 1-4 de linfocitos T (TCF).

15. El polinucleótido según la reivindicación 1, en el
20 que el ácido nucleico transcribible codifica un polipéptido seleccionado del grupo constituido por un anticuerpo, un fragmento funcional de un anticuerpo que se une a un epítipo, un factor de crecimiento, una citoquina, una proteína quinasa, un receptor soluble, un receptor
25 unido a membrana y un factor de coagulación sanguíneo.

16. Un vector que comprende el polinucleótido de una cualquiera de las reivindicaciones 1 a 15.

17. El vector según la reivindicación 16, en el que el vector es un vector de expresión en eucariotas.

30 18. Un vector de expresión en eucariotas que comprende:
a. un elemento que comprende más de 3000 nucleótidos contiguos desde la región promotora de un gen rps3 de mamífero;

b. un promotor heterólogo; y
35 c. un sitio de clonación múltiple,
en el que una secuencia de ácidos nucleicos transcribible

insertada en el sitio de clonación múltiple puede ser transcrita desde el promotor heterólogo y el nivel de transcripción se mejora con el elemento.

5 19. Una célula hospedadora que comprende el polinucleótido de una cualquiera de las reivindicaciones 1 a 15 o el vector de la reivindicación 16, 17 ó 18.

20. La célula hospedadora según la reivindicación 19, en la que las células se seleccionan del grupo constituido por CHO, NS0, BHK, HeLa, HepG2.

10 21. Un procedimiento de expresión de un polipéptido codificado por un ácido nucleico transcribible que comprende la inserción de un vector de expresión según cualquiera de las reivindicaciones 16 a 18 en una célula hospedadora adecuada y el cultivo de la célula hospedadora en condiciones que permitan la expresión del polipéptido.

15 22. Una composición farmacéutica que comprende el polinucleótido según una cualquiera de las reivindicaciones 1 a 15, el vector según una cualquiera de las reivindicaciones 16 a 18 o la célula hospedadora según una cualquiera de las reivindicaciones 19 ó 20, y un
20 vehículo farmacéuticamente aceptable.

23. Un polinucleótido según una cualquiera de las reivindicaciones 1 a 15, el vector según una cualquiera de las reivindicaciones 16 a 18 o la célula hospedadora según una cualquiera de las reivindicaciones 19 ó 20 para su uso
25 como medicamento.

Figura 1.

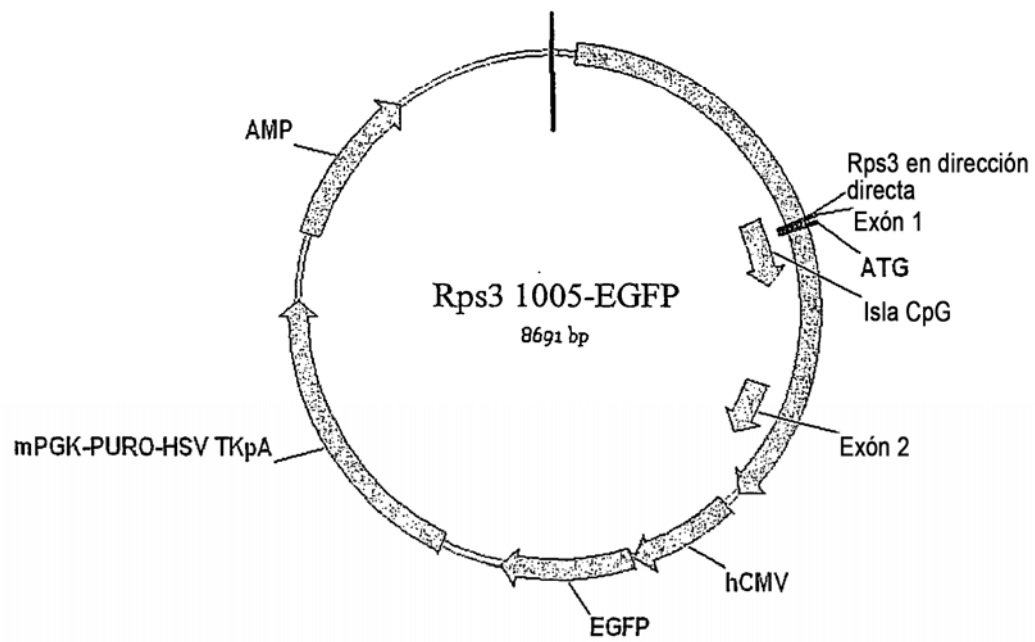


Figura 2

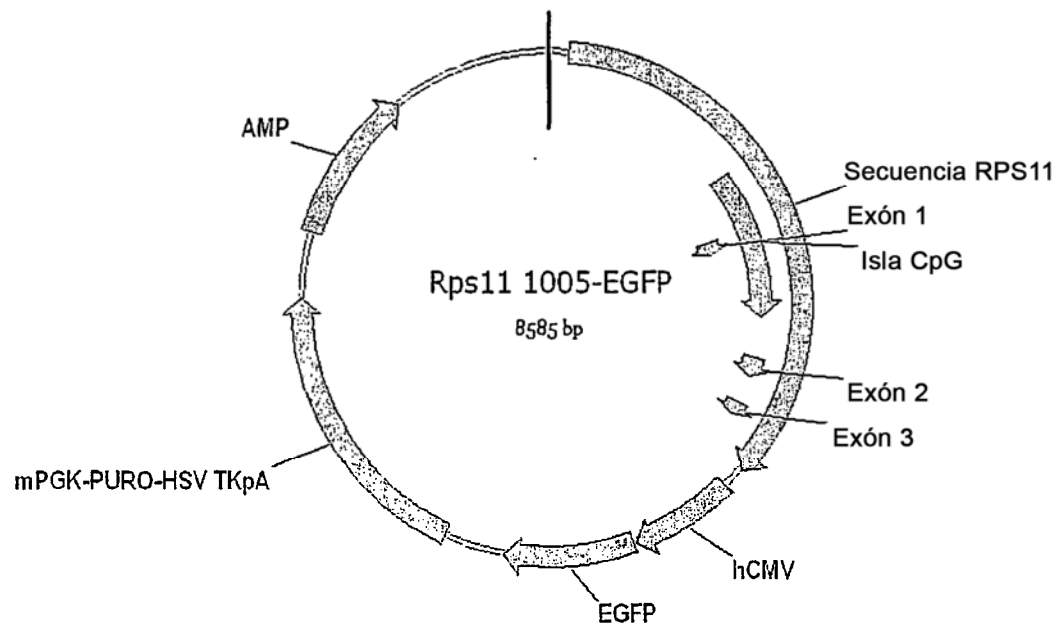
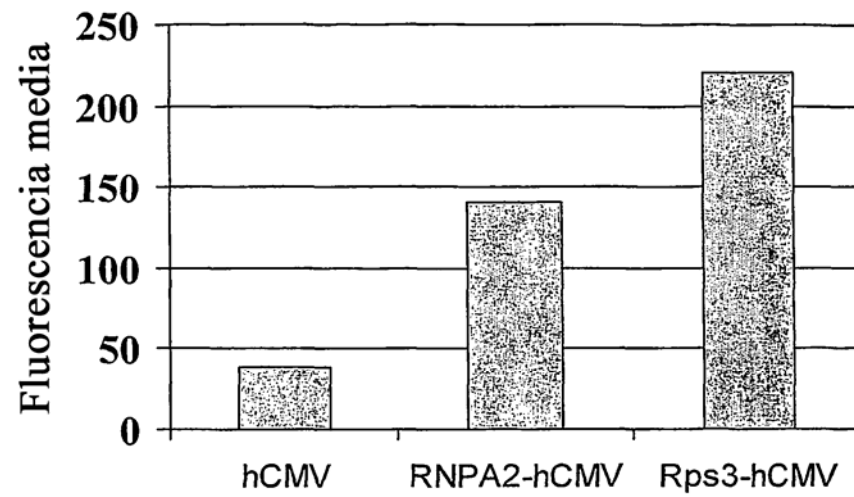


Figura 3.

A



B

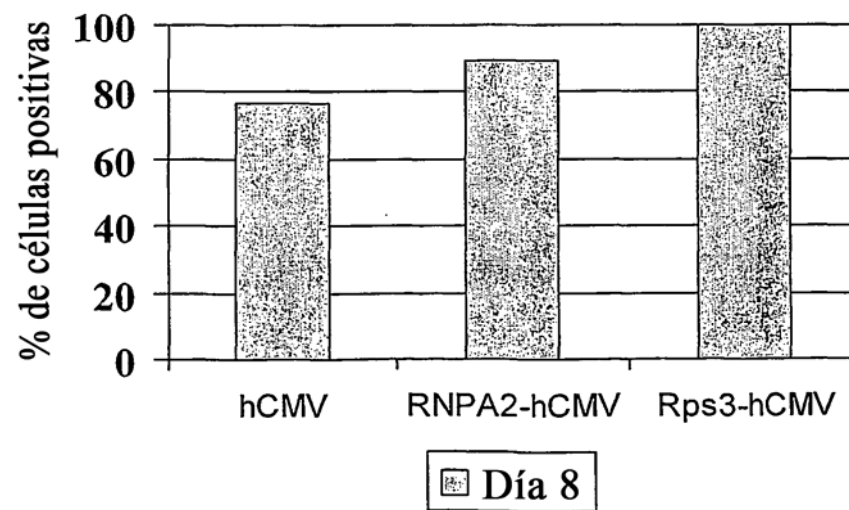
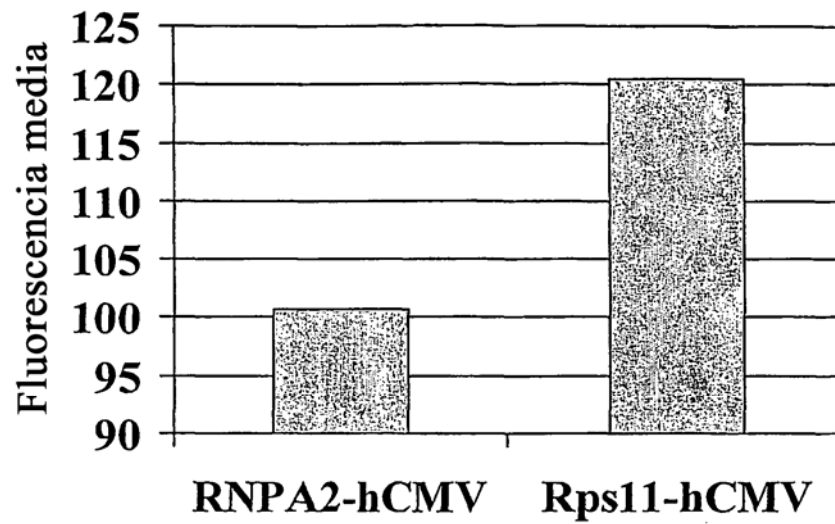


Figura 4.

A



B

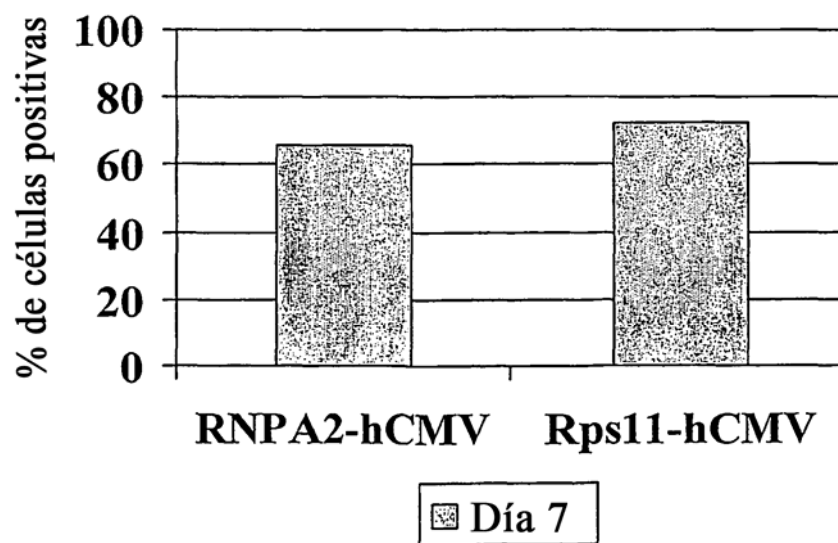


Figura 5.

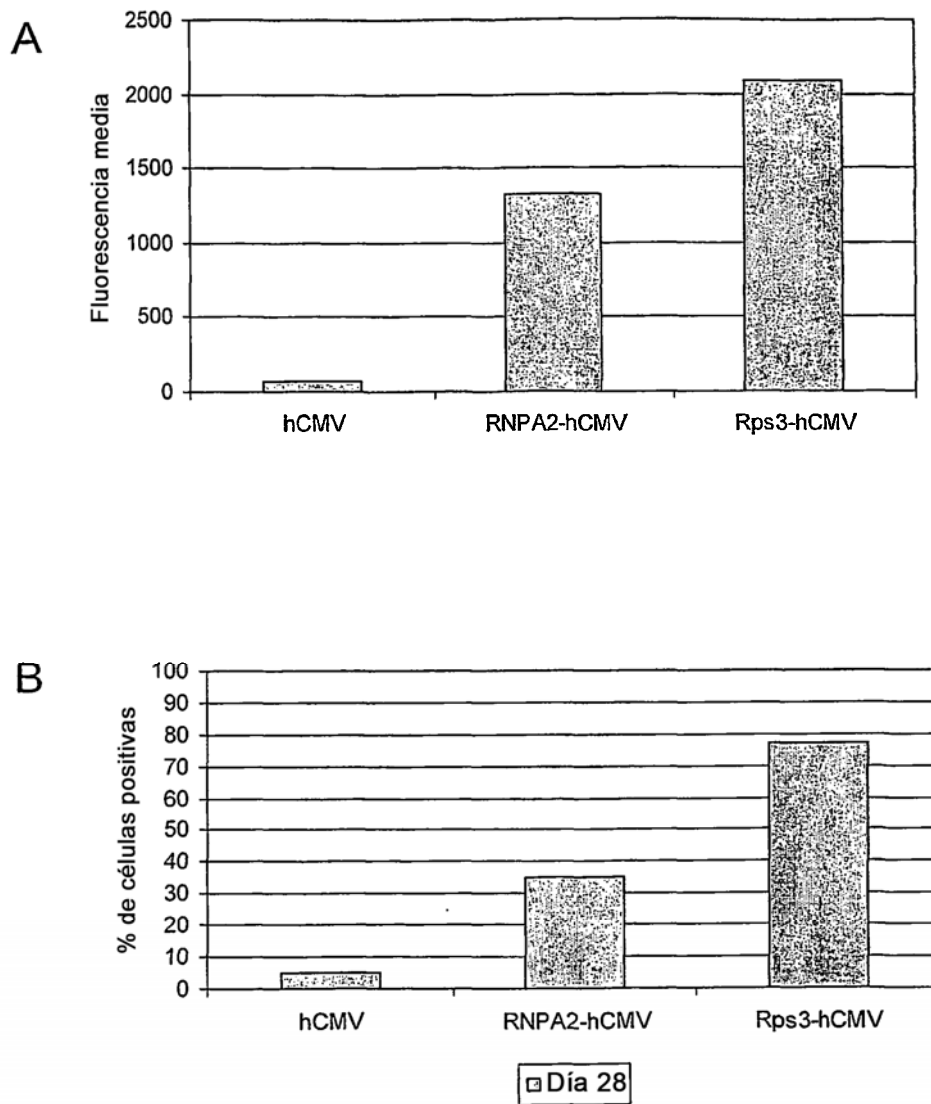
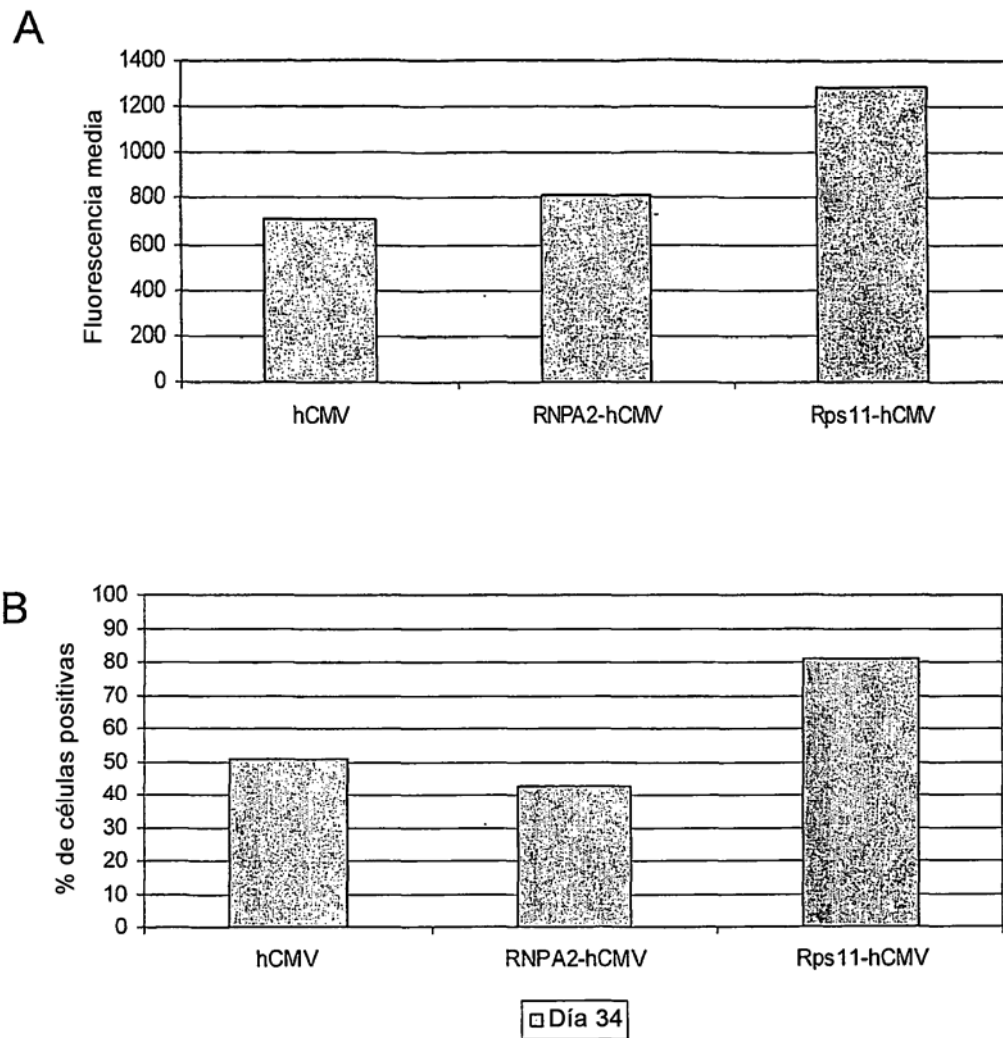


Figura 6.



REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Esta lista de referencias citadas por el solicitante pretende únicamente servir de ayuda al lector y no forma parte del documento de patente europeo. Aun cuando se ha puesto el máximo esmero en su elaboración, no pueden excluirse errores u omisiones y la EPO declina toda responsabilidad a este respecto.

Documentos de patente citados en la descripción

- US 6689606 B [0008]
- WO 000539 A [0008]
- WO 0005393 A [0012] [0013]
- US 6063598 A [0018]
- US 20040148647 A [0019]
- WO 9807876 A [0044]

Bibliografía distinta de patentes citada en la descripción

- **Dillon, N. ; Grosveld, F.** Chromatin domains as potential units of eukaryotic gene function. Curr. Opin. Genet. Dev., 1994, vol. 4, 260-264 [0002]
- **Higgs, D.R.** Do LCRs open chromatin domains? Cell, 1998, vol. 95, 299-302 [0002]
- **Palmiter, R.D. ; Brinster, R.L.** Ann. Ref. Genet., 1986, vol. 20, 465-499 [0003]
- **Allen, N.D. et al.** Nature, 1988, vol. 333, 852-855 [0003]
- **Bonnerot, C. ; Grimber, G. ; Briand, P. ; Nicolas, J.F.** Proc. Natl. Acad. Sci. USA, 1990, vol. 87, 6331-6335 [0003]
- **Fraser, P. ; Grosveld, F.** Curr. Opin. Cell Biol, 1998, vol. 10, 361-365 [0005]
- **Li, Q. ; Harju, S. ; Peterson, K.R.** Trends Genet., 1999, vol. 15, 403-408 [0005]
- **Kloussis, D. ; Festenstein, R.** Curr. Opin. Genet. Dev., 1997, vol. 7, 614-619 [0005]
- **LI, Q. ; Harju, S. ; Peterson, K.R.** Trends Genet, 1999, vol. 15, 403-408 [0005]
- **Kaufman RJ.** Methods Enzymol, 1990, vol. 185, 537-566 [0006]
- **Needham M ; Gooding C ; Hudson K ; Antoniou M ; Grosveld F ; Hollis M.** Nucleic Acids Res, 1992,

vol. 20, 997-1003 [0006]
 • **Needham M ; Egerton M ; Millest A ; Evans S ; Popplewell M ; Cerillo G ; McPheat J ; Monk A ; Jack A ; Johnstone D.** Protein Expr Purif, 1995, vol. 6, 124-131 [0006]
 • **Gardiner-Green M ; Frommer M.** J Mol Biol, 1987, vol. 196, 261-282 [0009]
 • **Rice P ; Longden I ; Bleasby A.** Trends Genet, 2000, vol. 16, 276-277 [0009]
 • **Bird y col.** Cell, 1985, vol. 40, 91-99 [0010]
 • **Tazi; Bird.** Cell, 1990, vol. 60, 909-920 [0010]
 • **Wise; Pravtcheva.** Genomics, 1999, vol. 60, 258-271 [0010]
 • **Antequera, F. ; Bird. A.** Proc. Natl. Acad. Sci. USA, 1993, vol. 90, 1195-11999 [0010]
 • **Cross, S.H. ; Bird, A.P.** Curr. Opin, Genet. Dev., 1995, vol. 5, 309-314 [0010]
 • **Tazi, J. ; Bird, A.** Cell, 1990, vol. 60, 909-920 [0010]
 • **Wool.** TIBS, 1996, vol.

21, 164-165 [0014]
 • **Wool y col.** Biochem Cell Biol, 1995, vol. 73, 933-947 [0014]
 • **Hershko ; Ciechanover.** Annu Rev Biochem, 1992, vol. 61, 761-807 [0015]
 • **Coux y col.** Annu Rev Biochem, 1996, vol. 65, 801-847 [0015]
 • **Finley y col.** Nature, 1989, vol. 338, 394-401 [0016]
 • **Chan y col.** Biochem Biophys Res Commun, 1995, vol. 215, 682-690 [0016]
 • **Redman ; Burris.** Biochem J, 1996, vol. 315, 315-321 [0016]
 • **Perry.** BMC Evolutionary Biology, 2005, vol. 5, 15 [0017]
 • **Hariharan y col.** Genes Dev, 1989, vol. 3, 1789-800 [0017]
 • **McKnight ; Tjian.** Cell, 1987, vol. 46, 795-805 [0027]
 • **Breathnach ; Chambon.** Annu Rev Biochem, 1981, vol. 50, 349-393 [0027]
 • **Kadonaga.** Cell, 2004, vol. 116, 247-257 [0029]
 • **Ondek y col.** Science, 1988, vol. 236, 1237-1244

[0029]

• **Foecking ; Hoffstetter.**
Gene, 1986, vol. 45, 101-
105 **[0031]**

• **Sambrook y col.**
Molecular Cloning: A
Laboratory Manual. Cold
Spring Harbour Laboratory,
1989 **[0046]**

• **Marston, F.** DNA Cloning
Techniques: A Practical
Approach. RL Press, 1987,
vol. III 1 **[0046]**

• Current Protocols. **F M**
Ausubel y col. Molecular
Biology. John Wiley &
Sons, Inc, 1994 **[0046]**