(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 345 335**

(51) Int. Cl.:
C12N 15/85 (2006.01)
C12N 5/10 (2006.01)
C12Q 1/68 (2006.01)
A61K 48/00 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Número de solicitud europea: **04707256 .6**
(96) Fecha de presentación : **02.02.2004**
(97) Número de publicación de la solicitud: **1587934**
(97) Fecha de publicación de la solicitud: **26.10.2005**

(54) Título: **Elementos genéticos mejorados que proporcionan altos niveles de expresión.**

(30) Prioridad: **01.02.2003 GB 0302330**
12.04.2003 GB 0308503
05.09.2003 GB 0320824

(45) Fecha de publicación de la mención BOPI:
21.09.2010

(45) Fecha de la publicación del folleto de la patente:
21.09.2010

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

(72) Inventor/es: **Williams, Steven Geraint;**
Crombie, Robert Lachlan;
Lipinski, Kai Stefan y
Irvine, Alistair Simpson

(74) Agente: **Polo Flores, Carlos**

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).

Venta de fascículos: Oficina Española de Patentes y Marcas. Pº de la Castellana, 75 – 28071 Madrid

DESCRIPCIÓN

Elementos genéticos mejorados que proporcionan altos niveles de expresión.

Campo de la invención

La presente invención se refiere a un polinucleótido que comprende un elemento de apertura de cromatina ubicuo (UCOE) más pequeño mejorado. Cuando está operativamente ligado a una secuencia de ácidos nucleicos expresable, este elemento proporciona niveles altos y reproducibles de expresión de genes. La presente invención también se refiere a un vector que comprende la secuencia de polinucleótidos, una célula huésped que comprende el vector y al uso del polinucleótido, vector o célula huésped en terapia, o para aplicaciones que implican la expresión de proteínas en cultivo celular.

Antecedentes de la invención

El presente modelo de estructura de cromatina en eucariotas superiores postula que los genes están 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 prevé que los dominios de cromatina existan bien un estado condensado "cerrado" transcripcionalmente silencioso o en una configuración descondensada "abierta" y transcripcionalmente competente. El establecimiento de una estructura de cromatina abierta caracterizada por un aumento de la sensibilidad a DNasaI, hipometilación de ADN e hiperacetilación de histonas se considera un requisito previo para el comienzo de la expresión de genes.

La naturaleza abierta y cerrada de las regiones de cromatina se refleja en el comportamiento de los transgenes que están aleatoriamente integrados en el genoma de células huésped. Constructos idénticos dan diferentes modelos de expresión específica de tejido y específica de fase de desarrollo cuando se integran en diferentes localizaciones en el 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)).

Frecuentemente también se observa un modelo de expresión variegado dentro de un tejido de ratón transgénico dado conocido como variegación por efecto de posición (PEV) (Kioussis, D. & Festenstein, R. Curr. Opin. Genet. Dev. 7, 614-619 (1997)). Cuando los genes exógenos están integrados en el cromosoma de cultivos de células de mamífero *in vitro,* muchos de los acontecimientos de integración dan como resultado el rápido silenciamiento del transgén y el resto da una gran variabilidad en los niveles de expresión (Pikaart, M.J., Recillas-Targa, F. & Felsenfield, G. Genes Dev. 12, 2852-2862 (1998); Fussenegger, M., Bailey, J.E., Hauser, H. & Mueller, P.P Trends Biotech. 17, 35-42 (1999)). Estos efectos de posición hacen ineficaz la expresión de transgenes con implicaciones en aplicaciones tanto de investigación básica como de biotecnología.

El modelo de dominios de cromatina para la organización de genes sugiere que los elementos de control genético que pueden establecer y mantener una estructura de cromatina abierta transcripcionalmente competente deben estar asociados a regiones activas del genoma.

Las regiones de control de locus (LCR) son una clase de elementos reguladores de la transcripción con capacidad de remodelación de la cromatina de largo alcance. Las LCR se definen funcionalmente en ratones transgénicos por su capacidad para conferir niveles fisiológicos independientes del sitio de integración, dependientes del número de copias de transgenes de expresión en un gen ligado en *cis*, especialmente transgenes de copia única, 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 forma crucial, tal expresión es específica de tejido. Las LCR pueden obstruir la propagación de heterocromatina, evitar el PEV (Kioussis, D. & Festenstein, R. Curr. Opin. Genet. Dev. 7, 614-619 (1997)) y estar constituidas por una serie de sitios hipersensibles (HS) a DNasaI que puede localizarse bien en 5' ó 3' de los genes que regulan (Li, Q., Harju, S. & Peterson, K.R. Trends Genet. 15: 403-408 (1999)).

Las LCR parecen estar constituidas por dos componentes separados, aunque no necesariamente independientes. Primero, el establecimiento de un "dominio de cromatina abierto" y segundo una capacidad de activación transcripcional dominante para conferir una expresión dependiente del número de copias de transgenes (Fraser, P. & Grosveld, F. Curr. Opin. Cell Biol. 10, 361-365 (1998). Los mecanismos moleculares mediante los que las LCR ejercen su función sigue siendo un punto de discrepancia (Higgs, D.R. Cell 95, 299-302 (1998); Bulger, M. & Groudine, M. Genes Dev. 13, 2465-2477 (1999); Grosveld, F. Curr. Opin. Genet. Dev. 9 152-157 (1999); Bender, M.A., Bulger, M., Close, J. & Groudine, M. Mol. Cell 5, 387-393 (2000).

La generación de líneas celulares de mamífero cultivadas que producen altos niveles de un producto proteínico terapéutico es una aplicación en desarrollo importante. Los efectos de posición de la cromatina hacen que sea un procedimiento difícil, que requiere mucho tiempo y caro. La solución más comúnmente usada para la producción de tales "fábricas de células" de mamífero se basa en 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 presión selectiva reducida. El uso de vectores que contienen LCR de dominios de genes sumamente expresados usando células derivadas del tejido apropiado simplifica mucho el procedimiento dando una

gran proporción de líneas celulares clónicas que muestran niveles de expresión altos estables (Needham M, Gooding C, Hudson K, Antoniou M, Grosfeld F y Hollis M. Nucleic Acids Res 20, 997-1003 (1992); Needham M, Egerton M, Millest 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).

Sin embargo, la especificidad de tejido de LCR, aunque es útil en algunas circunstancias, también es una gran limitación para muchas aplicaciones, por ejemplo, cuando no se conoce ninguna LCR para el tejido en el que se requiere la expresión, o cuando se requiere la expresión en muchos o todos los tejidos.

Las solicitudes de patente en tramitación junto con la presente PCT/GB99/02357 (documento WO 00/05393), US 09/358082, GB 0022995.5 y US 60/252.048 de los inventores incorporadas por referencia al presente documento describen elementos que son responsables, en su contexto cromosómico natural, de establecer una estructura de cromatina abierta a través de un locus que está constituido exclusivamente por genes de mantenimiento expresados ubicuamente. Estos elementos no se derivan de una LCR y comprenden islas CpG libres de metilación extendidas. Los inventores han usado el término elemento de apertura de cromatina ubicuo (UCOE) para describir tales elementos.

En ADN de mamífero, el dinucleótido CpG es reconocido por una enzima ADN metiltransferasa que metila la citosina en 5-metilcitosina. Sin embargo, la 5-metilcitosina es inestable y se convierte en timina. Como resultado, los dinucleótidos CpG se producen mucho menos frecuentemente que lo que se esperaría por casualidad. Sin embargo, algunas secciones de ADN genómico tienen una frecuencia de CpG que es más próxima a la esperada y estas secuencias se conocen como "islas CpG". Como se usa en este documento, una "isla CpG" se define como una secuencia de ADN de al menos 200 pb que tiene un contenido de GC de al menos el 50% y una relación de contenido de CpG observada/esperada de al menos 0,6 (es decir, un contenido de dinucleótido CpG de al menos el 60% del que se esperaría por casualidad) (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).

Las islas CpG libres de metilación son muy conocidas en la técnica (Bird y col. (1985) Cell 40: 91-99, Tazi y Bird (1990) Cell 60: 909-920) y pueden definirse como islas CpG en las que una proporción sustancial de los residuos de citosina no están metilados y que normalmente se extienden por los extremos 5' de dos genes divergentemente transcritos poco espaciados (0,1-3 kb). Se presenta que estas regiones de ADN permanecen hipometiladas en todos los tejidos durante todo el desarrollo (Wise y Pravtcheva (1999) Genomics 60: 258-271). Frecuentemente están asociadas a los extremos 5' de genes ubicuamente expresados, mostrando además un 40% estimado de los genes un perfil de expresión restringido de tejido (Antequera, F. & Bird, A. Proc. Natl. Acad. Sci. USA 90, 1195-11999 (1993); Cross, S.H. & Bird, A.P. Curr. Opin, Genet. Dev. 5, 309-314 (1995) y son conocidos por ser regiones localizadas de cromatina activa (Tazi, J. & Bird, A. Cell 60, 909-920 (1990).

Una isla CpG libre de metilación "extendida" es una isla CpG libre de metilación que se extiende a lo largo de una región que engloba más de un sitio de inicio de la transcripción y/o se extiende para más de 300 pb y preferentemente más de 500 pb. Las fronteras de la isla CpG libre de metilación extendida se definen funcionalmente por el uso de PCR a lo largo de la región en combinación con enzimas endonucleasas de restricción cuya capacidad para digerir (cortar) ADN en su secuencia de reconocimiento es sensible al estado de metilación de cualquier residuo de CpG que esté presente. Una enzima tal es *Hpa*II, que reconoce y digiere en el sitio CCGG, que comúnmente se encuentra dentro de las islas CpG, pero sólo si los residuos CG centrales *no* están metilados. Por tanto, la PCR realizada con ADN digerido con *Hpa*II y a lo largo de una región que alberga sitios *Hpa*II *no* da un producto de amplificación debido a la digestión de *Hpa*II si el ADN está *sin metilar.* La PCR *solo* dará un producto amplificado si el ADN está *metilado.* Por tanto, más allá de la región libre de metilación, *Hpa*II no digerirá el ADN, se observará un producto amplificado por PCR definiéndose así los límites de la "isla CpG libre de metilación extendida".

Los inventores han demostrado (documento WO 00/05393) que las regiones que abarcan las islas CpG libres de metilación que engloban promotores divergentemente transcritos duales de la proteína de unión a TATA humana (*TBP*)/componente B1 de proteosoma (*PSMBI*) y loci de genes A2/B1 de ribonucleoproteína nuclear heterogénea (*hnRNPA2*)/proteína 1Hsγ de heterocromatina ($HP1^{Hs\gamma}$) dan niveles de expresión de genes fisiológicos reproducibles y que pueden prevenir un modelo de expresión variegado y el silenciamiento que normalmente se produce con la integración de transgenes dentro de la heterocromatina centromérica.

Como se usa en este documento, el término "expresión reproducible" significa que el polinucleótido de la invención dirigirá la expresión del gen expresable a sustancialmente el mismo nivel de expresión independientemente de su entorno de cromatina y preferentemente independientemente del tipo de célula o tipo de tejido en el que puede estar el polinucleótido de la invención. Aquellos expertos en la materia reconocerán que sustancialmente se logra el mismo nivel de expresión del gen expresable operativamente ligado, independientemente del entorno de cromatina del polinucleótido reivindicado y preferentemente independientemente del tipo de célula, asumiendo que la célula puede activar la expresión de genes.

Los inventores han mostrado (documento WO 00/05393) que las islas CpG libres de metilación asociadas a promotores activamente transcriptores poseen la capacidad de remodelar la cromatina y, por tanto, se cree que son un determinante principal en el establecimiento y el mantenimiento de un dominio abierto en los loci de genes de mantenimiento.

Los UCOE confieren un aumento de la proporción de acontecimientos de liberación de genes productores con mejoras en el nivel y la estabilidad de la expresión de transgenes. Esto tiene importantes aplicaciones en investigación y biotecnológicas que incluyen la generación de animales transgénicos y productos de proteínas recombinantes en células cultivadas. Los inventores han mostrado (documento WO 00/05393) efectos beneficiosos de los UCOE en la expresión del constructo indicador CMV-EGFP y con la proteína secretada farmacéuticamente valiosa eritropoyetina. Las propiedades de los UCOE también sugieren utilidad en la terapia génica cuya eficacia está frecuentemente limitada por una baja frecuencia de acontecimientos de liberación de genes productores y un nivel y duración de la expresión inadecuados (Verma, I.M. & Somia, N. Nature 389: 239-242 (1997).

La solicitud anteriormente mencionada PCT/GB99/02357 (documento WO 00/05393) de los inventores desvela fragmentos de UCOE funcionales de aproximadamente 4,0 kb, en particular el fragmento "5,5 RNP" definido por los nucleótidos 4102 a 8286 de la Figura 21 (como se desvela en la pág. 11, líneas 6 y 7). La misma solicitud desvela un fragmento "1,5 kb RNP" (Figs 22 y 29, derivación descrita en la pág. 51, líneas 1 a 5). Sin embargo, este fragmento es en realidad un fragmento *Bam*HI - *Tth*111I de 2165 pb del fragmento "5,5 RNP" descrito anteriormente que está constituido por los nucleótidos 4102 a 6267 de la Figura 21 de esa solicitud.

En otra solicitud (documento WO 02/24930), los inventores desvelan UCOE artificialmente construidos compuestos por fragmentos de islas CpG que se producen naturalmente. Los fragmentos desvelados son mayores que los reivindicados en la presente solicitud y en ese momento no se consideró posible usar individualmente fragmentos pequeños en vez de cómo meros componentes de constructos de UCOE sintéticos o "híbridos".

Dadas estas implicaciones significativas y aplicaciones generales, existe el deseo de optimizar adicionalmente los niveles de expresión de transgenes. Existe la necesidad de optimizar adicionalmente los niveles de expresión de transgenes, particularmente en los campos de terapia génica *in vivo* y para la producción *in vitro* de proteínas recombinantes.

Una necesidad particular es reducir el tamaño de elementos usados para potenciar la expresión génica. Con esto pueden producirse vectores más pequeños, o vectores con una mayor capacidad en términos del tamaño del inserto que pueden contener y expresar establemente.

Exposición de la invención

Es un objeto de la invención proporcionar elementos de apertura de cromatina más pequeños para potenciar el nivel y la reproducibilidad de transgenes operativamente ligados.

Según la presente invención se proporcionan polinucleótidos que comprenden fragmentos funcionales pequeños de UCOE. Tales polinucleótidos comprenden islas CpG libres de metilación de no más de aproximadamente 2 kb, o fragmentos mayores de tales islas de no más de aproximadamente 2 kb.

Aunque se conocen polinucleótidos mayores que comprenden islas CpG libres de metilación extendidas en la técnica (véase la solicitud previa de los propios solicitantes WO 00/05393), previamente no se ha establecido si es posible usar fragmentos significativamente más pequeños y que todavía mantengan la potenciación en niveles y conformidad de expresión obtenida con UCOE mayores/islas CpG libres de metilación.

Como se usa en este documento, el término "operativamente ligado" se refiere a una relación de operabilidad entre elementos en los polinucleótidos de la invención. "Operativamente ligado" es un término muy conocidos para aquellos expertos en la materia que describe una relación funcional entre secuencias de ADN que actúan en cis. La exacta relación estructural puede o puede no ser relevante y es diferente para diferentes tipos de elementos. Para un promotor implica una posición 5' esencialmente adyacente (normalmente dentro de menos de 100 pb) respecto al marco de lectura abierto que acciona. En el caso de islas CpG libres de metilación extendidas parece que un efecto regional sobre la estructura de la cromatina es responsable de aumentar el nivel y la conformidad de la expresión de genes. A modo de ejemplo, el elemento que comprende una isla CpG libre de metilación extendida está situado inmediatamente en 5' del promotor que controla la transcripción del gen expresable. Sin embargo, "operativamente ligado" engloba la posibilidad de estar situado en cualquier sitio con tal que pueda demostrarse un claro efecto funcional.

La presente invención proporciona un polinucleótido aislado que comprende

a. una isla CpG libre de metilación extendida;

b. un marco de lectura abierto expresable operativamente ligado a dicha isla CpG libre de metilación extendida;

c. un promotor operativamente ligado a dicho marco de lectura abierto, en el que dicho promotor no está naturalmente ligado a dicha isla CpG;

caracterizado porque dicha isla CpG comprende un fragmento de un locus HP-1/hnRNP A2 humano de no más de 1,6 kb y en el que se obtiene expresión reproducible de dicho marco de lectura abierto en dos o más tipos de tejido.

Alternativamente, el polinucleótido comprende un fragmento de restricción *Esp*3I, más preferentemente la secuencia de la Figura 2, nucleótidos 977 a 2522 (SEQ ID NO: 1), o un homólogo funcional del mismo. Preferentemente, dicho fragmento está orientado en orientación directa.

Por "homólogo funcional" se indica una secuencia de polinucleótidos que puede hibridarse, bajo condiciones restrictivas, con la secuencia desvelada y que tiene propiedades similares de conferir expresión reproducible de marcos de lectura abiertos expresables operativamente ligados en dos o más tejidos. Las condiciones de hibridación/lavado restrictivas son muy conocidas en la técnica. Por ejemplo, híbridos de ácidos nucleicos que son estables después de lavar en 0,1 x SSC, SDS al 0,1% a 60ºC. Es muy conocido en la técnica que las condiciones de hibridación óptimas pueden calcularse si la secuencia del ácido nucleico es conocida. Por ejemplo, las condiciones de hibridación pueden determinarse por el contenido de GC del ácido nucleico sometido a hibridación. Véase Sambrook y col. (1989), Molecular Cloning; A Laboratory Approach. Una fórmula común para calcular las condiciones de rigurosidad requeridas para lograr la hibridación entre moléculas de ácidos nucleicos de una homología especificada es:

$$T_m = 81{,}5^oC + 16{,}6\log[Na^+] + 0{,}41[\%\text{ de G } +\text{ C}] - 0{,}63\ (\%\text{ de formamida})$$

Preferentemente, el polinucleótido de la presente invención facilita la expresión reproducible de un gen operativamente ligado específicamente en no tejido.

En otra realización preferida, el polinucleótido comprende un fragmento del gen hnRNP A2 humano de no más de aproximadamente 1 kb, preferentemente comprende un fragmento de restricción *Bsp*E1 - *Esp*3I de 987 pb, más preferentemente comprende la secuencia de la Figura 2, nucleótidos 1536 a 2522 (SEQ ID NO: 2).

Preferentemente, el polinucleótido de la presente invención comprende uno o más promotores que se producen naturalmente.

Preferentemente, el polinucleótido de la presente invención comprende promotores duales o bidireccionales que se transcriben divergentemente.

También se describe un fragmento de la isla CpG/región promotora de β-actina, preferentemente de origen humano. Más preferentemente, el fragmento comprende un fragmento de ADN dentro del intervalo de 100 pb a 2 kb que abarca la isla CpG/región promotora de β-actina humana.

También se describe un fragmento de la isla CpG/región promotora de PDCD2, preferentemente de origen humano. Más preferentemente, el fragmento comprende un fragmento de ADN dentro del intervalo de 100 pb a 2 kb que abarca la isla CpG/región promotora de PDCD2 humana.

También se describe un fragmento de ADN dentro del intervalo de 100 pb a 1,9 kb que abarca la isla CpG/región promotora de β-actina humana y un fragmento de ADN dentro del intervalo de 100 pb a 2 kb que abarca la isla CpG/región promotora de PDCD2 humana. Preferentemente, dicho fragmentos son directamente adyacentes a sus promotores orientados divergentemente.

En otro aspecto, la invención proporciona un vector que comprende el polinucleótido de la invención como se desvela anteriormente. Preferentemente, dicho vector es un vector de expresión adaptado para la expresión génica eucariota.

Normalmente, dicha adaptación incluye, por ejemplo y no a modo de limitación, la provisión de secuencias del control de la transcripción (secuencias promotoras) que median en la expresión específica de célula/tejido.

Promotor y potenciador son términos muy conocidos en la técnica e incluyen las siguientes características que se proporcionan sólo por ejemplo y no a modo de limitación.

Los promotores son secuencias reguladoras que actúan en cis de 5' ligadas directamente al inicio de la transcripción. Los elementos promotores incluyen la denominada caja TATA y secuencias de selección del inicio de ARN polimerasa (RIS) que actúan seleccionando un sitio de inicio de la transcripción. Estas secuencias también se unen a polipéptidos que actúan, entre otras cosas, facilitando la selección del inicio de la transcripción por ARN polimerasa.

Preferentemente, el promotor se selecciona de CMV, EF-1α, LTR del virus del sarcoma de Rous (RSV) o LTR de VIH2 o combinaciones de secuencias derivadas de los mismos. Más preferentemente, el promotor es un promotor inmediato/temprano de CMV. Lo más preferentemente, es el promotor inmediato/temprano de CMV de ratón.

En una realización preferida de vector, la isla CpG de la invención está situada adyacente y 5' al promotor operativo que controla la expresión del marco de lectura abierto expresable.

Los elementos potenciadores son secuencias de ácidos nucleicos que actúan en cis encontradas frecuentemente en 5' respecto al sitio de inicio de la transcripción de un gen (también pueden encontrarse potenciadores en 3' respecto a una secuencia de gen o incluso localizarse en secuencias intrónicas y, por tanto, es independiente de la posición).

Los potenciadores actúan aumentando la velocidad de transcripción del gen al que está ligado el potenciador. La actividad de los potenciadores es responsable de los factores de transcripción que actúan en *trans* (polipéptidos) que se ha mostrado que se unen específicamente a elementos potenciadores. La unión/actividad de los factores de transcripción es responsable de varias claves medioambientales que incluyen, a modo de ejemplo y no a modo de limitación, metabolitos intermedios (por ejemplo, glucosa), efectores medioambientales (por ejemplo, calor). (Véase Eukaryotic Transcription Factors, de David S Latchman, Academic Press Ltd, San Diego).

Las adaptaciones también incluyen la provisión de marcadores de selección y secuencias de replicación autónomas que facilitan ambas el mantenimiento de dicho vector en o bien la célula eucariota o en el huésped procariota. Los vectores que se mantienen autónomamente en células eucariotas se denominan en lo sucesivo vectores episomales. Los vectores episomales son deseables ya que son autorreplicantes y así persisten sin necesidad de integración. Los vectores episomales de este tipo se describen en el documento WO98/07876.

Las adaptaciones que facilitan la expresión de genes codificados por vectores incluyen la provisión de secuencias de terminación de la transcripción/poliadenilación. Esto también incluye la provisión de sitios internos de entrada al ribosoma (IRES) que actúan maximizando la expresión de genes codificados por vectores dispuestos en casetes de expresión bicistrónicos o multi-cistrónicos.

Estas adaptaciones son muy conocidas en la técnica. Hay una cantidad significativa de bibliografía publicada con respecto a la construcción de vectores de expresión y a técnicas de ADN recombinante en general. Por favor, véanse Sambrook y col. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY y referencias allí dentro; Marston, F (1987) DNA Cloning Techniques: A Practical Approach Vol. III IRL Press, Oxford UK; DNA Cloning: F M 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 un adenovirus, virus adenoasociado, un herpesvirus, virus de la variolovacuna, lentivirus u otros retrovirus.

Preferentemente, el vector comprende un gen operativamente ligado que es un ácido nucleico terapéutico. Un ácido nucleico terapéutico tal puede actuar reemplazando o complementando la función de un gen defectuoso que produce una enfermedad tal como fibrosis quística, talasemia, anemia de células falciformes, anemia de Fanconi, hemofilia, inmunodeficiencia combinada grave (IDCG), fenilcetonuria (FCN), deficiencia de alfa-1 antitripsina, distrofia muscular de Duchenne, deficiencia de ornitina transcarbamilasa u osteogénesis imperfecta. Alternativamente, puede codificar un agente citotóxico o enzima conversora de profármaco selectivamente expresada en una célula diana tal como una célula de cáncer maligno con el fin de matarla. Tales aplicaciones y muchas otras son muy conocidas para aquellos expertos en la materia y la relevancia de la presente invención en potenciar la expresión de ácidos nucleicos terapéuticos será evidente para tales médicos expertos.

Lo más preferentemente, el vector comprende una cualquiera de SEQ ID Nos 1 ó 2, un promotor de CMV, un sitio de clonación múltiple, una secuencia de poliadenilación y genes que codifican marcadores de selección bajo elementos de control adecuados.

También se proporciona una célula huésped transfectada con el vector. Preferentemente, ésta es una célula eucariota, más preferentemente una célula de mamífero, lo más preferentemente una célula humana o de roedor.

Alternativamente puede ser una célula vegetal.

Otro aspecto de la invención es el uso de cualquiera de los polinucleótidos aislados, islas CpG extendidas libres de metilación (en particular como se desvelan por SEQ ID Nos 1 ó 2), vectores o células huésped desveladas anteriormente en terapia, en particular, su uso en terapia génica.

La presente invención también proporciona el uso del polinucleótido, vector o célula huésped en la preparación de un medicamento o composición para uso en terapia génica.

La presente invención también proporciona un procedimiento de tratamiento que comprende administrar a un paciente en necesidad de tal tratamiento una dosis eficaz del polinucleótido, vector o célula huésped de la presente invención. Preferentemente, el paciente padece una enfermedad tratable por terapia génica.

La presente invención también proporciona una composición farmacéutica que comprende el polinucleótido y/o el vector y/o la célula huésped, opcionalmente en mezcla con un vehículo o diluyente farmacéuticamente aceptable, para terapia para tratar una enfermedad o para proveer las células de un tejido particular de una proteína o función ventajosa.

El polinucleótido, vector o célula huésped de la invención o la composición farmacéutica pueden administrarse por una vía que incluye sistémica, intramuscular, intravenosa, aerosol, oral (forma sólida o líquida), tópica, ocular, rectal, intraperitoneal y/o intratecal e inyección directa local.

Por supuesto, la pauta de dosificación exacta necesitará ser determinada por médicos individuales para pacientes individuales y esto se controlará a su vez por la naturaleza exacta de la proteína expresada por el gen de interés y el tipo de tejido que está eligiéndose como diana para el tratamiento.

La dosificación también dependerá de la indicación de enfermedad y la vía de administración. El número de dosis dependerá de la enfermedad, y los datos de eficacia de ensayos clínicos.

La cantidad de ADN de polinucleótido o vector administrado para la terapia génica eficaz según la invención estará preferentemente en el intervalo de entre 50 ng - 1000 μg de ADN de vector/kg de peso corporal; y más preferentemente en el intervalo de entre aproximadamente 1-100 μg de ADN de vector/kg.

Aunque según la invención se prefiere administrar el polinucleótido, vector o célula huésped a un mamífero para la captación celular *in vivo*, puede utilizarse una solución *ex vivo* por la cual las células se extirpan de un animal, se transducen con el polinucleótido o vector y luego se reimplantan en el animal. Por ejemplo, puede accederse al hígado mediante una vía de acceso *ex vivo* extirpando hepatocitos de un animal, transduciendo los hepatocitos *in vitro* y reimplantando los hepatocitos transducidos en el animal (por ejemplo, como se describe para conejos por Chowdhury y col., Science 254:1802-1805, 1991, o en seres humanos por Wilson, Hum. Gene Ther. 3:179-222, 1992). Tales procedimientos también pueden ser eficaces para la administración a diversas poblaciones de células en los sistemas circulatorio o linfático tales como eritrocitos, células T, células B y células madre hematopoyéticas.

En otro aspecto, la invención proporciona el uso de polinucleótidos, vectores o células huésped de la invención en un sistema de cultivo celular con el fin de obtener un producto génico deseado. Preferentemente, el ácido nucleico expresable codifica una proteína recombinante para la expresión en un sistema de cultivo celular *in vitro*.

Los sistemas de cultivo celular adecuados son muy conocidos en la técnica y se describen completamente en la abundante bibliografía conocida para aquellos expertos en la materia. Se proporciona un procedimiento para la producción de un polipéptido según la invención que comprende:

i) proporcionar una célula transformada/transfectada con un polinucleótido según la invención;

ii) cultivar dicha célula en condiciones propicias para la preparación de dicho polipéptido; y

iii) purificar dicho polipéptido a partir de dicha célula, o su entorno de crecimiento.

Alternativamente, el gen expresable codifica un producto de no polipéptido tal como ARN. Tal ARN puede ser un ARN antisentido que puede inhibir la expresión de un gen particular a un nivel post-transcripcional, o puede tener una función enzimática (de ribozima) u otra función tal como un ARN ribosómico.

También se proporciona el uso de un polinucleótido de isla CpG libre de metilación extendida según cualquiera de las SEQ ID NOs: 1 ó 2 para aumentar la expresión de un gen endógeno que comprende insertar el polinucleótido en el genoma de una célula en una posición operativamente asociada al gen endógeno, aumentando así el nivel de expresión del gen.

En otra realización de la invención se proporciona un animal transgénico no humano que comprende cualquiera de los polinucleótidos o vectores de la invención o cualquiera de las islas CpG libres de metilación extendidas de la invención en el que dicha isla CpG se ha introducido artificialmente. Los procedimientos para hacer ratones transgénicos (Gordon y col., Proc. Natl. Acad. Sci. USA 77:7380 (1980); Harbers y col., Nature 293:540 (1981); Wagner y col., Proc. Natl. Acad. Sci. USA 78:5016 (1981); y Wagner y col., Proc. Natl. Acad. Sci. USA 78:6376 (1981), ovejas, cerdos, pollos (véase Hammer y col., Nature 315:680 (1985)), etc., son muy conocidos en la técnica y se contemplan para uso según la invención.

Tales animales transgénicos que contienen el polinucleótido de la invención también pueden usarse para la producción a largo plazo de una proteína de interés.

También se proporciona un modelo de mamífero para determinar la eficacia de la terapia génica usando el polinucleótido, vector o célula huésped de la invención. El modelo de mamífero comprende un animal transgénico cuyas células contienen el vector de la presente invención. Tales animales permiten hacer pruebas antes de ensayos clínicos en seres humanos.

La presente invención también proporciona el uso de los polinucleótidos e islas CpG libres de metilación extendidas de la presente invención en la producción de plantas transgénicas.

La generación de plantas transgénicas que tienen una elevada producción o una elevada resistencia a enfermedad, plagas, sequía o sal son muy conocidas para aquellos expertos en la materia. La presente invención también proporciona plantas transgénicas que contienen células que contienen el polinucleótido de la presente invención. Algunas o todas las células que comprenden el polinucleótido de la invención pueden originarse en plantas.

La presente invención también se refiere al uso de polinucleótidos, vectores o islas CpG libres de metilación extendidas de la presente invención en aplicaciones de genómica funcional. La genómica funcional se refiere principalmente a la identificación de genes específicamente expresados en tipos de células o estados de enfermedad particulares y ahora proporciona miles de secuencias de genes novedosas de posible interés para fines del descubrimiento de fármacos o de terapia génica. El problema principal de usar esta información para el desarrollo de terapias novedosas radica en cómo se determinan las funciones de estos genes. Los polipéptidos de la invención pueden usarse en varias aplicaciones de genómica funcional con el fin de determinar la función de secuencias de genes. Las aplicaciones de genómica funcional de la presente invención incluyen, pero no se limitan a:

(1) Usar polinucleótidos de la presente invención para lograr la expresión sostenida de versiones antisentido de secuencias de genes o bibliotecas de silenciamiento de ribozimas, determinándose así los efectos de la inactivación del gen sobre el fenotipo de células.

(2) Usar polinucleótidos de la presente invención para preparar bibliotecas de expresión para las secuencias de genes de forma que la administración a células dará como resultado la expresión sostenida reproducible fidedigna de las secuencias de genes. Las células resultantes que expresan las secuencias de genes pueden usarse en una variedad de soluciones para la determinación de la función y el descubrimiento de fármacos. Por ejemplo, producción de anticuerpos neutralizantes contra el producto génico; rápida purificación del producto de proteína del propio gen para uso en estudios de barrido estructurales, funcionales o de fármacos; o en el barrido de fármacos basados en células.

(3) Usar polinucleótidos de la presente invención en soluciones que implican células madre embriónicas (EM) de ratón y ratones transgénicos. Una de las soluciones de la genómica funcional más poderosas implica la inserción al azar en genes en células ES de ratón de constructos que sólo permiten la selección de fármacos tras la inserción en genes expresados y que fácilmente pueden rescatarse para la secuenciación (G. Hicks y col., Nature Genetics, 16, 338-334). Entonces, los ratones transgénicos con mutaciones inactivadas en genes con secuencias novedosas pueden prepararse fácilmente para investigar su función. En este momento, esta tecnología funciona bien para el 10% de los genes de ratón que están bien expresados en células ES de ratón. La incorporación de los polinucleótidos de la presente invención en los constructos integrantes permitirá que esta técnica se extienda para identificar todos los genes expresados en ratones.

Descripción detallada de la invención

La invención se describirá ahora sólo modo de ejemplo y con referencia a las figuras adjuntas en las que:

La Figura 1 muestra un mapa del locus HP-1/hnRNPA2 que indica sitios de restricción *Bam*HI, *Hind*III, *Esp*3I, *Tth*111I y *Bsp*EI que definen diversos fragmentos de islas CpG de 4, 2, 1,5 y 1 kb.

La Figura 2 muestra la secuencia de nucleótidos del locus HP-1/hnRNPA2 que indica sitios de restricción *Bam*HI, *Esp*3I, *Bsp*EI y *Tth*111I.

La Figura 3 muestra la expresión de un gen indicador EGFP operativamente ligado a los fragmentos de islas CpG de 4 kb y 1,5 kb (en ambas orientaciones). Los datos de FACScan se expresaron como mediana de la fluorescencia durante 71 días.

La Figura 4 muestra la expresión de un gen indicador EGFP operativamente ligado a los fragmentos de islas CpG de 4 kb y 1,5 kb (en ambas orientaciones). Los datos de FACScan se expresaron como células positivas en % durante 71 días.

La Figura 5 muestra la expresión de un gen indicador EGFP operativamente ligado a los fragmentos de islas CpG de 4 kb, 1,5 kb y 1 kb. Los datos de FACScan se expresaron como mediana de la fluorescencia durante 68 días.

La Figura 6 muestra la expresión de un gen indicador EGFP operativamente ligado a los fragmentos de islas CpG de 4 kb, 1,5 kb y 1 kb. Los datos de FACScan se expresaron como células positivas en % durante 68 días.

La Figura 7 muestra la estructura de los dos constructos adenovirales Ad.CMV-Luc-SV40p(A) y Ad.1,5kb(F) UCOE-CMV-Luc-SV40p(A) que se basan en el serotipo 5 de adenovirus. Los casetes de expresión de luciferasa se insertaron en la región E1 en una orientación izquierda a derecha. Las regiones E1 y E3 se delecionan de los virus. Debido a la deleción de E1, los virus son defectuosos en la replicación. El CMV (citomegalovirus) es el potenciador/promotor de CMV humano. SV40p(A) es la señal poliadenilación tardía del virus SV40 de pGL3-Basic (Promega).

La Figura 8 muestra que un UCOE de 1,5 kb refuerza la expresión génica en células HeLa cuando se libera por medio de un vector adenoviral. Las células HeLa se infectaron a una MOI de 50 durante 2-3 horas en 200 μl de medio de infección (medio HeLa normal que sólo contenía FCS al 1%) con los virus Ad.CMV-Luc-SV40p(A) y Ad.1,5kb (F)UCOE-CMV-Luc-SV40p(A). Después de la incubación se añadieron 2 ml de medio completo y las células se sembraron en placas de 6 pocillos. La actividad de la luciferasa se analizó 2 días después de la infección. Se muestran

los resultados de tres experimentos independientes. Los experimentos se hicieron por triplicado. Se muestra la media +/- la desviación estándar.

La Figura 9 muestra que el efecto de UCOE de 1,5 kb es independiente de la preparación viral. Las células HeLa se infectaron a una MOI de 50 durante 2-3 horas en 200 μl de medio de infección (medio HeLa normal que sólo contenía FCS al 1%) con los virus Ad.CMV-Luc-SV40p(A) y Ad.1,5kb(F)UCOE-CMV-Luc-SV40p(A) de dos preparaciones virales independientes, cada una. Después de la incubación se añadieron 2 ml de medio completo y las células se sembraron en placas de 6 pocillos. La actividad de la luciferasa se analizó 2 días después de la infección. Se muestran los resultados de un experimento representativo. Los experimentos se hicieron por triplicado.

La Figura 10 muestra que el UCOE de 1,5 kb aumenta el nivel de expresión transgénica y la estabilidad en conjuntos de CHO-K1 retroviralmente transducidos. (A) Valores de GFP medios con el tiempo para conjuntos de CHO-K1 clasificados con FACS (36.000 células originalmente clasificadas con FACS) después de la transducción retroviral para CMV y constructos de 1,5UCOE-CMV. Día 1 se refiere al 1º día de medición (8 días después de la clasificación con FACS). (B) Histogramas de los conjuntos de CHO-K1 clasificados con FACS para cada momento de medición.

La Figura 11 muestra la expresión de GFP de CHO-K1 transducido (conjunto I) después de infección a baja MOI. (A) Valores de GFP medios para poblaciones de CHO-K1 positivas para GFP (M1) después de la transducción retroviral para cada momento. La población de CMV empezó con 3,19% de células positivas para GFP y la población de 1,5UCOE-CMV con 4,99% de células positivas para GFP. (B) Histogramas para las poblaciones en cada momento.

La Figura 12 muestra la expresión de GFP de CHO-K1 transducido (conjunto II) después de infección a MOI intermedia. (A) Valores de GFP medios para poblaciones de CHO-K1 positivas para GFP (M1) después de la transducción retroviral para cada momento. La población de CMV empezó con 25,2% de células positivas para GFP y la población de 1,5UCOE-CMV con 14,35% de células positivas para GFP. (B) Histogramas de las poblaciones para cada momento.

La Figura 13 muestra que el UCOE de 1,5 kb aumenta el nivel de expresión transgénica y la estabilidad en conjuntos de HeLa retroviralmente transducidos. (A) Valores de GFP medios con el tiempo para conjuntos de HeLa clasificados con FACS (10.000 células originalmente clasificadas con FACS) después de la transducción retroviral para CMV y constructos 1,5UCOE-CMV. Día 1 se refiere al 1º día de medición (5 días después de la clasificación con FACS). (B) Histogramas de los conjuntos de HeLa clasificados con FACS para cada momento de medición.

La Figura 14 muestra que el UCOE de 1,5 kb aumenta la conformidad de la expresión de GFP tanto con como entre clones de células HeLa retroviralmente transducidas. Los clones de HeLa clasificados con FACS se expandieron al tamaño de 6 pocillos y se analizó la expresión de GFP. (A-C) Histogramas para todos los clones de CMV. (D-E) Histogramas para todos los clones de 1,5UCOE-CMV.

La Figura 15 muestra que el UCOE de 1,5 kb disminuye el coeficiente de variación (CV) de la expresión de GFP en clones de células HeLa. Se muestra el valor CV promedio (+/-DE) de las medias para todos los clones de HeLa retroviralmente transducidos y clasificados con FACS de la Figura 5. Tanto el CV como la desviación estándar (DE) de los clones de 1,5UCOE-CMV disminuyen significativamente en comparación con los clones de CMV.

Ejemplos

Ejemplo 1

Un UCOE HP-1/hnRNP A2 de 1,5 kb potencia la expresión

Materiales y procedimientos

Construcción de vectores

Como se trata en la solicitud anterior WO 00/05393 de los inventores, el vector CET20 se generó clonando un fragmento *Hind*III de 8,3 kb del locus HP-1/hnRNP A2 humano (que contiene los promotores HP1/RNP y la isla CpG extendida) en pBluescript (Stratagene).

Entonces, un fragmento de 4186 pb (denominado en lo sucesivo 4 kb) se eliminó del inserto por digestión con *Bam*HI y *Hind*III. Este fragmento se llenó en el extremo usando T4 ADN polimerasa y se ligó en pEGFPN-1 (Clontech) que había sido digerido con *Ase*I y de nuevo se llenó en el extremo usando T4 ADN polimerasa. Entonces, los clones se aislaron con el fragmento en ambas orientaciones.

El fragmento *Esp*3I de 1546 pb (isoesquizómero de *Bsm*BI) (denominado en lo sucesivo el fragmento de 1,5 kb) se aisló de nuevo de CET20 por digestión con *Esp*3I (*Bsm*BI) seguido por llenado del extremo, éste se ligó de nuevo en el sitio AseI de pEGFPN-1 como se describe anteriormente y los clones se identificaron con el fragmento en ambas orientaciones (estando la orientación “directa” con el promotor RNPA2 orientada en la misma dirección 5’ a 3’ que el promotor de CMV operativo aguas abajo adyacente a partir del cual se transcribió el transgén de EGFP).

Transfección

Se transfectaron células CHO-K1 y se seleccionaron en G418 según procedimientos estándar y como se describen en las solicitudes de patente en tramitación junto con la presente incorporadas por referencia.

Análisis de la expresión de GFP

Las células transfectadas se mantuvieron en selección con G418 a 600 μg/ml. Las células se lavaron con solución salina tamponada con fosfato patrón y se sometieron a arrastre del sustrato con tripsina/EDTA según procedimientos estándar. Se añadió un exceso de medio Nutrient mixture F12 (HAM) (Gibco) y las células se transfirieron a tubos de poliestireno de fondo redondo de 5 ml para análisis por FACScan de Becton Dickinson. La fluorescencia de GFP se detectó y se comparó con la autofluorescencia de la población de células parental. La expresión en la población de células se expresó tanto como la mediana de la fluorescencia expresada como unidades arbitrarias ajustadas respecto al control (según procedimientos estándar) en una escala lineal en las figuras, y en términos de % de células juzgadas como expresores positivos.

Resultados

Como se muestra en la Figura 3, el fragmento de 1,5 kb dio una potenciación significativa de la expresión en comparación con el control (aproximadamente 10 veces durante al menos 70 días) en términos de mediana de la fluorescencia y aproximadamente el 60% de la potenciación observada con el fragmento de 4 kb cuando se insertó en la orientación directa. En el experimento mostrado en la Figura 5, la expresión del fragmento de 1,5 kb directo es comparable con la obtenida con el fragmento de 4 kb. Sin embargo, parece que la orientación inversa es bastante menos eficaz en términos de mediana de la fluorescencia, como se muestra en la Figura 3. La Figura 4 muestra ambas orientaciones para que sean comparables en términos de % de células positivas. Sin embargo, estos datos sugieren que este fragmento puede tener un grado de direccionalidad y que, para la mayor parte de los fines, la orientación directa es preferible.

Ejemplo 2

Un UCOE HP-1/hnRNP A2 de 1 kb potencia la expresión

Materiales y procedimientos

Construcción de vectores

El vector que contiene el UCOE de 1 kb se construyó digiriendo el vector pEGFPN-1 con el fragmento *Esp*3I de 1,5 kb en la orientación directa con *Pci*I y *Bsp*EI para eliminar los 500 pb de 5', seguido por llenado del extremo y re-ligación. Esto generó un vector con el fragmento BspEI-*Esp*3I de 987 pb en sólo una orientación.

Resultados

El fragmento de 1 kb (987 pb) en la orientación directa parece comparable al fragmento de 1,5 kb en la orientación directa en términos de tanto la mediana de la fluorescencia (Figura 5) como el % de células positivas (Figura 6).

Ejemplo 3

Un UCOE HP-1/hnRNP A2 de 1,5 kb potencia la expresión en un constructo adenoviralmente codificado

Material y procedimientos

Cultivo celular

Se obtuvieron HeLa de ATCC (Manassas, Virginia). Se obtuvieron PER.C6 de Crucell (Leiden, Los Países Bajos). Todas las líneas celulares compradas se cultivaron como recomendó el proveedor. El Prof. L. S. Young (Cancer Research UK Institute for Cancer Studies, Universidad de Birmingham, Birmingham, RU) proporcionó amablemente células 911 y se cultivaron en antibióticos que contenían DMEM/FCS al 10%.

Construcción de plásmidos

Se obtuvo PGL3-Basic de Promega (Madison, WI, EE.UU.) y contiene el casete de luciferase-SV40p(A) aguas abajo del sitio de clonación múltiple. El potenciador/promotor de CMV humano (0,9 kb) se clonó en pGL3-Basic digerido con SmaI para generar pGL3/CMV-Luc-SV40p(A). Para generar pGL3/1,5kb(F)UCOE-CMV-SV40p(A), el fragmento *Esp*3I de UCOE de 1,5 kb (véase la Figura 1 y 2) se hizo romo con T4 ADN polimerasa (NEB, Beverly, MA, EE.UU.) y luego se ligó en pGL3-Basic/CMV-Luc digerido con *Nhe*I y hecho romo con T4.

Construcción de vectores virales

Para construir pPS1128/CMV-Luc-SV40p(A) y pPS1128/1,5kb(F)UCOE-CMV-Luc-SV40p(A), los casetes de expresión se cortaron de pGL3/CMV-Luc-SV40p(A) por el digesto *Pvu*I/*Nhe*I/*Bam*HI y de pGL3/1,5kb(F)UCOE-CMV-Luc-SV40p(A) por el digesto *Kpn*I/*Bam*HI, luego ambos casetes se hicieron romos y se clonaron en pPS1128 digerido con SpeI y hecho romo (Djeha y col., Cancer Gene Therapy 2000: 721-731). Los plásmidos se analizaron por digestos de enzimas de restricción.

Los virus Ad.CMV-Luc-SV40p(A) y Ad.1,5kb(F)UCOE-CMV-Luc-SV40p(A) se construyeron por recombinación homóloga en PER.C6 usando pPS1128/CMV-Luc-SV40p(A) y pPS1128/1,5kb(F)UCOE-CMV-Luc-SV40p(A), respectivamente, y el vector del esqueleto adenoviral de solapamiento pPS1160 (Djeha y col., Cancer Gene Therapy 2000: 721-731) se amplió y se purificó con CsCl y se valoró como se describe en otra parte (Lipinski y col., Gene Therapy, 8, 2001: 274-281).

Infecciones por virus y ensayo de la actividad de la luciferasa

Se infectaron células HeLa (5,0 x 10^4 por 6 pocillos) en suspensión con los virus respectivos en 200 μl de medio de infección (antibióticos que contienen EMEM/FCS al 1%) durante 2-3 horas. Entonces se añadió 1 ml de medio completo y las células se sembraron en placas de 6 pocillos. Se prepararon extractos de células completas lisando las células en 200 μl de tampón de lisis (fosfato de sodio 10 mM, pH 7,8, $MgCl_2$ 8 mM, EDTA 1 mM, Triton-X-100 al 1%, glicerina al 15%) y eliminando el lisado por centrifugación (1 min 13.000 x g, TA). Una alícuota de cada sobrenadante se ensayó para la actividad de la luciferasa en el intervalo lineal usando un luminómetro (Lumat LB 9501, Berthold, Wildbad, Alemania) a los momentos indicados.

Resultados

Para analizar si el UCOE puede reforzar la expresión génica en un vector adenoviral, el UCOE de 1,5 kb se clonó en una orientación directa aguas arriba (el promotor de RNP tiene la misma orientación que el promotor de CMV) del potenciador/promotor de CMV humano que acciona un gen indicador de luciferasa. La Figura 7 muestra esquemáticamente la estructura de los dos virus que se compararon para su la actividad de la luciferasa. Las células HeLa se infectaron a una MOI de 50 con los virus respectivos y la actividad de la luciferasa se analizó 2 días después de la infección. La Figura 8 muestra claramente que el fragmento de UCOE de 1,5 kb puede aumentar espectacularmente el nivel de expresión de genes indicadores en células HeLa.

Para excluir un efecto específico asociado a las preparaciones virales, los inventores prepararon nuevas preparaciones de ambos virus y repitieron los experimentos. La Figura 9 muestra que el efecto de UCOE no dependía de la preparación viral y era completamente reproducible con las nuevas preparaciones independientes.

Ejemplo 4

Un UCOE HP-1/hnRNP A2 de 1,5 kb potencia la expresión en un constructo retroviralmente codificado

Material y procedimientos

Cultivo celular

Se obtuvieron HeLa, CHO-K1 y 293 de ATCC (Manassas, Virginia). HeLa y 293 se cultivaron ambas en DMEM, FCS al 10% añadiendo NEAA al 1% y conteniendo antibióticos. CHO-K1 se cultivaron en medio Nutrient mix F12 (HAM) que contenía FCS al 10% y antibióticos.

Construcción de plásmidos

Los plásmidos pVPack-GP y pVPack-VSV-G se obtuvieron de Stratagene (LaJolla, CA, USA) y expresan proteínas gag-pol retrovirales y VSV-G de la envuelta, respectivamente. El vector retroviral pQCXIX se obtuvo de BD Bioscience y contiene el promotor de CMV humano para la expresión transgénica (Palo Alto, CA, EE.UU.). El plásmido phrGFP-1 se obtuvo de Stratagene y se usó como fuente para un ADNc de GFP modificado. Para clonar pQCXIX-CMV-hrGFP, el ADNc de hrGFP se cortó de phr-GFP-1 por el digesto *Bam*HI/*Eco*RV y se clonó en pQCXIX digerido con *Bam*HI/*Eco*RV. Para generar pQCXIX-1,5UCOE-CMV-hrGFP, el fragmento de UCOE de 1,5 kb (fragmento *Esp*3I/*Bsm*BI) se clonó como un fragmento romo en pQCXIX-CMV-hrGFP digerido con XbaI y hecho romo.

Producción de partículas de retrovirus anfotrópicos envueltas por VSV-G

Todos los virus se pseudotiparon con la envuelta de VSV-G (glicoproteína del virus de la estomatitis vesicular) anfotrópica permitiendo una transducción altamente eficaz y de amplio intervalo de huéspedes. El vector retroviral usado (pQCXIX) genera virus autoinactivantes debido a una deleción en la región U3 de la LTR en 3' que se copia a la LTR en 5' después de la transcripción inversa y, por tanto, inhibe cualquier transcripción posterior de la LTR en 5'.

Se sembraron 4,5x10^6 células 293 (al día siguiente las células deberán ser el 80-90% de confluentes) en placas de 8,4 cm de diámetro (55,6 cm^2) recubiertas con colagenasa tipo I el día antes de la transfección transitoria. Al día siguiente se mezclaron 114 μl de lipofectamina 2000 (Invitrogen, Groenningen, Los Países Bajos) con 286 μl de OptiMEM (Invitrogen) y se incubaron durante 5 min a temperatura ambiente (TA). En paralelo, 1,5 μg de cada plásmido pVPack-GP, pVPack-VSV-G y vector retroviral se mezclaron con OptiMEM para dar un volumen final de 400 μl. La mezcla de ADN se añadió a la disolución de lipofectamina 2000 y se incubó durante 20-30 min a TA. Durante esta incubación, las células 293 se lavaron una vez con PBS y se añadieron 6 ml de OptiMEM/FCS al 10% a las células. Finalmente, la mezcla de ADN/lipofectamina se añadió gota a gota a las células y las células se incubaron en un incubador con 5% de CO_2 a 37ºC durante 5 horas. Entonces, el medio se sustituyó por 8 ml de medio de cultivo de 293 fresco y las células se incubaron durante otras 24-36 horas. Finalmente, el sobrenadante que contenía las partículas virales se recogió de las células, se centrifugó durante 5 min a 1000 rpm para sedimentar células y residuos de células, y el sobrenadante se esterilizó por filtración usando filtros de 0,45 μm de baja unión a proteínas de PVDF Millex-HV (Millipore, Molsheim, Francia). El sobrenadante se tomó en alícuotas, se congeló criogénicamente con nitrógeno líquido y se guardó a -80ºC.

Transducción retroviral de líneas de células diana

Se sembraron células diana (HeLa, CHO-K1) a 1x10^5 células/pocillo en placas de 6 pocillos el día antes de la infección. Para la transducción viral, diferentes cantidades de sobrenadante que contenía virus se añadieron en presencia de 8,0 μg/ml de polibreno (Sigma, St. Louis, Missouri, EE.UU.). Las células se incubaron con el virus durante 24 horas y el medio se sustituyó luego por medio fresco. La expresión génica puede observarse a partir de dos días post-transducción. Para evitar una copia múltiple por efectos celulares, para todos los análisis se infectaron células diana con una cantidad de virus que dio como resultado una eficiencia de transducción muy inferior al 100% (normalmente da el 1-20% de células positivas para GFP). Dentro de este intervalo de eficiencia de transducción, el volumen de sobrenadante usado para la infección guarda una relación lineal con el número de células positivas, mientras que el nivel de expresión medio muestra un aumento mucho menor.

Análisis con FACS

Para los análisis con FACS, células diana se lavaron con PBS, se tripsinaron y se resuspendieron en medio completo y la expresión de hrGFP (proteína verde fluorescente) se analizó con un FACScan de Becton Dickinson (BD, Franklin Lakes, NJ, EE.UU.) usando parámetros de instrumento constantes para cada estudio. Los datos se analizaron con el software CellQuest para Apple McIntosh.

Clasificación con FACS

Para clasificar células HeLa y CHO-K1 que expresan GFP, las células se prepararon como para el análisis con FACS y luego se clasificaron con un clasificador FACS de BD Bioscience (proporcionado amablemente por The Institute for Cancer Studies, Universidad de Birmingham, RU). Los clones individuales y los conjuntos clasificados se expandieron luego y la expresión de GFP se siguió con el tiempo.

Resultados

En un primer estudio, células CHO-K1 se transdujeron con sobrenadantes celulares que contenían los retrovirus CMV-hrGFP (CMV) y 1,5UCOE-CMV-hrGFP (1,5UCOE-CMV). Las células CHO-K1 se transdujeron como se describe en Material y procedimientos y tres días después de la infección las células positivas para GFP se clasificaron con FACS como un conjunto de poblaciones (36.000 células para cada constructo viral) y luego se expandieron. La expresión de GFP y el modelo de histogramas se siguieron con el tiempo y se ilustran en la Figura 10A y B, respectivamente. En el día 17 de análisis, el valor medio para el conjunto de 1,5UCOE-CMV es 2,6 veces superior a la media del conjunto de CMV. Adicionalmente, como muestran los histogramas en la Figura 10B, el conjunto de 1,5UCOE-CMV da un pico de la expresión de GFP más estrecho que el conjunto de CMV. El valor de CV promedio (coeficiente de variación; éste es un marcador de la homogeneidad de la expresión de GFP) para las células positivas para GFP en el marcador 1 es 146,0 para el conjunto de CMV y 83,0 para el conjunto de 1,5UCOE-CMV.

Comparable a este estudio, se transdujeron CHO-K1 y la población se siguió con el tiempo sin clasificación con FACS. La MOI se eligió para dar un porcentaje relativamente bajo de células positivas para excluir la copia múltiple por efectos celulares. Como se muestra en la Figura 11A (conjunto I con aproximadamente 5% de células positivas para GFP), la media de las células positivas para la población de 1,5UCOE-CMV en el último momento de medición es 2,8 veces superior a la media de la población de CMV. Para el conjunto II (aproximadamente 15-25% de células positivas para GFP; Figura 12A), la media para la población de 1,5UCOE-CMV en el último momento de medición es 2,3 veces superior a la media para la población de CMV. Comparable a la población clasificada con FACS, las poblaciones de 1,5UCOE-CMV dan un pico de expresión claramente más estrecho en comparación con las poblaciones de CMV en todos los momentos analizados (Figura 11B y Figura 12B).

Como para CHO-K1, los conjuntos de 10.000 células HeLa establemente transducidos con los constructos retrovirales se clasificaron con FACS. En la Figura 13A se muestra el valor medio de GFP incluyendo el valor medio en el día de la clasificación con FACS (día -5). En el último momento de medición, la media del conjunto de 1,5UCOE-CMV es 1,6 veces superior a la media del conjunto de CMV. Como se ilustra por los histogramas en la Figura 13B,

muchas de las células de CMV pierden su alto nivel de expresión de GFP con el tiempo; a diferencia, la mayoría de las células de 1,5UCOE-CMV conservan su alta expresión de GFP.

Los clones de células individuales HeLa también se seleccionaron con FACS y se expandieron en placas de 6 pocillos. Los clones de CMV tienden a dar una expresión pico ligeramente superior que los clones de 1,5UCOE-CMV; sin embargo, la conformidad del nivel de expresión, tanto dentro de un clon como entre clones, aumenta significativamente por la presencia del UCOE como se indica por los histogramas para todos los clones analizados (Figura 14 A-E). La Figura 15 muestra el coeficiente promedio de la variación de la expresión de GFP (CV) y su variación estándar (DE). El valor de CV+/-DE para los clones de CMV es 135,0 +/-165,3 y para los clones de 1,5UCOE-CMV 49,4+/-10,9.

Antecedentes citados en la descripción

Esta lista de antecedentes citados por el solicitante es sólo por conveniencia del lector. No forma parte del documento de patente europea. Aun cuando se ha tenido mucho cuidado al compilar los antecedentes, no pueden excluirse errores u omisiones y la Oficina Europea de Patentes declina toda responsabilidad a este respecto.

Documentos de patente citados en la descripción

- GB 9902357 W [0010] [0018]
- GB 0022995 A [0010]
- WO 0005393 A [0010] [0014] [0016] [0017] [0018] [0023] [0069]
- US 60252048 B [0010]
- WO 0224930 A [0019]
- WO 9807876 A [0041]
- US 09358082 B [0010]

Bibliografía no 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. *y col. 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]
- **Kioussis**, D.; **Festenstein**, R. *Curr. Opin. Genet. Dev.,* 1997, vol. 7, 614-619 [0004] [0006]
- **Pikaart**, M. J.; **Recillas-Targa**, F.; **Felsenfield**, G. *Genes Dev.,* 1998, vol. 12, 2852-2862 [0004]
- **Fussenegger**, M.; **Bailey**, J. E.; **Hauser**, H.; **Mueller**, P. P. *Trends Biotech.,* 1999, vol. 17, 35-42 [0004]
- **Fraser**, P.; **Grosveld**, F. *Curr. Opin. Cell Biol.,* 1998, vol. 10, 361-365 [0006] [0007]
- **Li**, Q.; **Harju**, S.; **Peterson**, K. R. *Trends Genet.,* 1999, vol. 15, 403-408 [0006]
- **Higgs**, D. R. *Cell,* 1998, vol. 95, 299-302 [0007]
- **Bulger**, M.; **Groudine**, M. *Genes Dev.,* 1999, vol. 13, 2465-2477 [0007]
- **Grosveld**, F. *Curr. Opin. Genet. Dev.,* 1999, vol. 9, 152-157 [0007]
- **Bender**, M. A.; **Bulger**, M.; **Close**, J.; **Groudine**, M. *Mol. Cell,* 2000, vol. 5, 387-393 [0007]
- **Kaufman** R J. *Methods Enzymol,* 1990, vol. 185, 537-566 [0008]
- **Needham** M; **Gooding** C; **Hudson** K; **Antoniou** M; **Grosfeld** F; **Hollis** M. *Nucleic Acids Res,* 1992, vol. 20, 997-1003 [0008]
- **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 [0008]

- **Gardiner-Green** M; **Frommer** M. *J Mol Biol,* 1987, vol. 196, 261-282 [0011]
- **Lipinski** y col. *Gene Therapy,* 2001, vol. 8, 274-281 [0080]
- **Rice** P; **Longden** I; **Bleasby** A. *Trends Genet,* 2000, vol. 16, 276-277 [0011]
- **Bird** *y col*. *Cell,* 1985, vol. 40, 91-99 [0012]
- **Tazi**; **Bird**. *Cell,* 1990, vol. 60, 909-920 [0012]
- **Wise**; **Pravtcheva**. *Genomics,* 1999, vol. 60, 258-271 [0012]
- **Antequera**, F.; **Bird**, A. *Proc. Natl. Acad. Sci. USA,* 1993, vol. 90, 1195-11999 [0012]
- **Cross**, S. H.; **Bird**, A. P. *Curr. Opin, Genet. Dev.,* 1995, vol. 5, 309-314 [0012]
- **Tazi**, J.; **Bird**, A. *Cell,* 1990, vol. 60, 909-920 [0012]
- **Verma**, I. M.; **Somia**, N. *Nature,* 1997, vol. 389, 239-242 [0017]
- **Sambrook** *y col*. *Molecular Cloning; A Laboratory Approach*. 1989 [0027]
- **David** S **Latchman**. Eukaryotic Transcription Factors. *Academic Press Ltd* [0040]
- **Sambrook** *y col*. Molecular Cloning: A Laboratory Manual. *Cold Spring Harbour Laboratory* [0043]
- **Marston**, F. DNA Cloning Techniques: A Practical Approach. *IRL Press*, 1987, vol. III [0043]
- DNA Cloning. F M **Ausubel** *y col*. Current Protocols in Molecular Biology. *John Wiley & Sons, Inc*, 1994 [0043]
- **Chowdhury** *y col*. *Science,* 1991, vol. 254, 1802-1805 [0057]
- **Wilson**. *Hum. Gene Ther.,* 1992, vol. 3, 179-222 [0057]
- **Gordon** *y col*. *Proc. Natl. Acad. Sci. USA,* 1980, vol. 77, 7380 [0062]
- **Harbers** *y col*. *Nature,* 1981, vol. 293, 540 [0062]
- **Wagner** *y col*. *Proc. Natl. Acad. Sci. USA,* 1981, vol. 78, 5016 [0062]
- **Wagner** *y col*. *Proc. Natl. Acad. Sci. USA,* 1981, vol. 78, 6376 [0062]
- **Hammer** *y col*. *Nature,* 1985, vol. 315, 680 [0062]
- G. **Hicks** *y col*. *Nature Genetics,* vol. 16, 338-334 [0067]
- **Djeha** *y col*. *Cancer Gene Therapy,* 2000, 721-731 [0079] [0080]

REIVINDICACIONES

1. Un polinucleótido aislado que comprende

a. una isla CpG libre de metilación extendida;

b. un marco de lectura abierto expresable operativamente ligado a dicha isla CpG libre de metilación extendida;

c. un promotor operativamente ligado a dicho marco de lectura abierto, en el que dicho promotor no está naturalmente ligado a dicha isla CpG;

caracterizado porque dicha isla CpG comprende un fragmento de un locus HP-1/hnRNP A2 humano de no más de 1,6 kb y en el que la expresión reproducible de dicho marco de lectura abierto se obtiene en dos o más tipos de tejido.

2. Un polinucleótido según la reivindicación 1 que comprende un fragmento de restricción *Esp*3I.

3. Un polinucleótido según la reivindicación 2 que comprende la secuencia de la Figura 2, nucleótidos 977 a 2522 (SEQ ID NO: 1), o un homólogo funcional del mismo.

4. Un polinucleótido según cualquiera de las reivindicaciones 1 a 3 que comprende un fragmento de un locus HP-1/hnRNP A2 de no más de 1 kb.

5. El polinucleótido según la reivindicación 4 que comprende un fragmento de restricción *Bsp*E1-*Esp*3I.

6. Un polinucleótido según la reivindicación 5 que comprende la secuencia de la Figura 2, nucleótidos 1536 a 2522 (SEQ ID NO: 2), o un homólogo funcional del mismo.

7. Un vector que comprende el polinucleótido de cualquiera de las reivindicaciones 1 a 6.

8. Un vector según la reivindicación 7, en el que el vector es un vector episomal.

9. Un vector según la reivindicación 8, en el que el vector es un vector de integración.

10. Un vector según cualquiera de las reivindicaciones 7 a 9, en el que el vector es un plásmido.

11. Un vector según cualquiera de las reivindicaciones 7 a 10, en el que el gen operativamente ligado es un ácido nucleico terapéutico.

12. Un vector según cualquiera de las reivindicaciones 7 a 11 que comprende cualquiera de SEQ ID NOs 1 ó 2, un promotor de CMV, un sitio de clonación múltiple, una secuencia de poliadenilación y genes que codifican marcadores de selección bajo elementos de control adecuados.

13. Una célula huésped que comprende el polinucleótido de cualquier de las reivindicaciones 1 a 6, o el vector de cualquiera de las reivindicaciones 7 a 12.

14. Un polinucleótido aislado según cualquiera de las reivindicaciones 1 a 6, un vector según cualquiera de las reivindicaciones 7 a 12 o la célula huésped de la reivindicación 13 para uso en terapia génica.

15. Uso del polinucleótido según cualquiera de las reivindicaciones 1 a 6, un vector según cualquiera de las reivindicaciones 7 a 12 o la célula huésped de la reivindicación 13 para la preparación de un medicamento para tratar inmunodeficiencia combinada grave (IDCG).

16. Un polinucleótido según cualquiera de las reivindicaciones 1 a 6, un vector según cualquiera de las reivindicaciones 7 a 12 o la célula huésped de la reivindicación 13 para uso en el tratamiento de inmunodeficiencia combinada grave (IDCG).

17. Una composición farmacéutica del polinucleótido según cualquiera de las reivindicaciones 1 a 6, un vector según cualquiera de las reivindicaciones 7 a 12 o la célula huésped de la reivindicación 13 en combinación con un excipiente farmacéuticamente aceptable.

18. Uso del polinucleótido según cualquiera de las reivindicaciones 1 a 6, un vector según cualquiera de las reivindicaciones 7 a 12 o la célula huésped de la reivindicación 13 en un sistema de cultivo celular con el fin de obtener un producto génico deseado.

19. Uso *ex vivo* del polinucleótido según cualquiera de SEQ ID NOs 1 ó 2 para aumentar la expresión de un gen endógeno que comprende insertar el polinucleótido en el genoma de una célula en una posición operativamente asociada al gen endógeno, aumentando así el nivel de expresión del gen.

20. Un animal no humano transgénico que contiene células que comprenden un polinucleótido que comprende el polinucleótido de cualquiera de las reivindicaciones 1 a 6, o un vector según cualquiera de las reivindicaciones 7 a 12, en el que dicho polinucleótido se ha introducido artificialmente.

21. Uso *ex vivo* del polinucleótido según cualquiera de las reivindicaciones 1 a 6, un vector según cualquiera de las reivindicaciones 7 a 12 o la célula huésped de la reivindicación 13 para obtener expresión de una secuencia génica antisentido para inactivar la expresión de la secuencia génica correspondiente.

22. Uso del polinucleótido según cualquiera del polinucleótido según cualquiera de las reivindicaciones 1 a 6, un vector según cualquiera de las reivindicaciones 7 a 12 en la preparación de una biblioteca de expresión.

23. Uso del polinucleótido según cualquier de SEQ ID NOs 1 ó 2 en un procedimiento para identificar genes expresables en un animal no humano que comprende insertar un constructo que comprende el polinucleótido en células madre embriónicas del animal no humano, en el que el constructo permite la selección de fármacos tras la inserción en genes expresados.

Figura 1

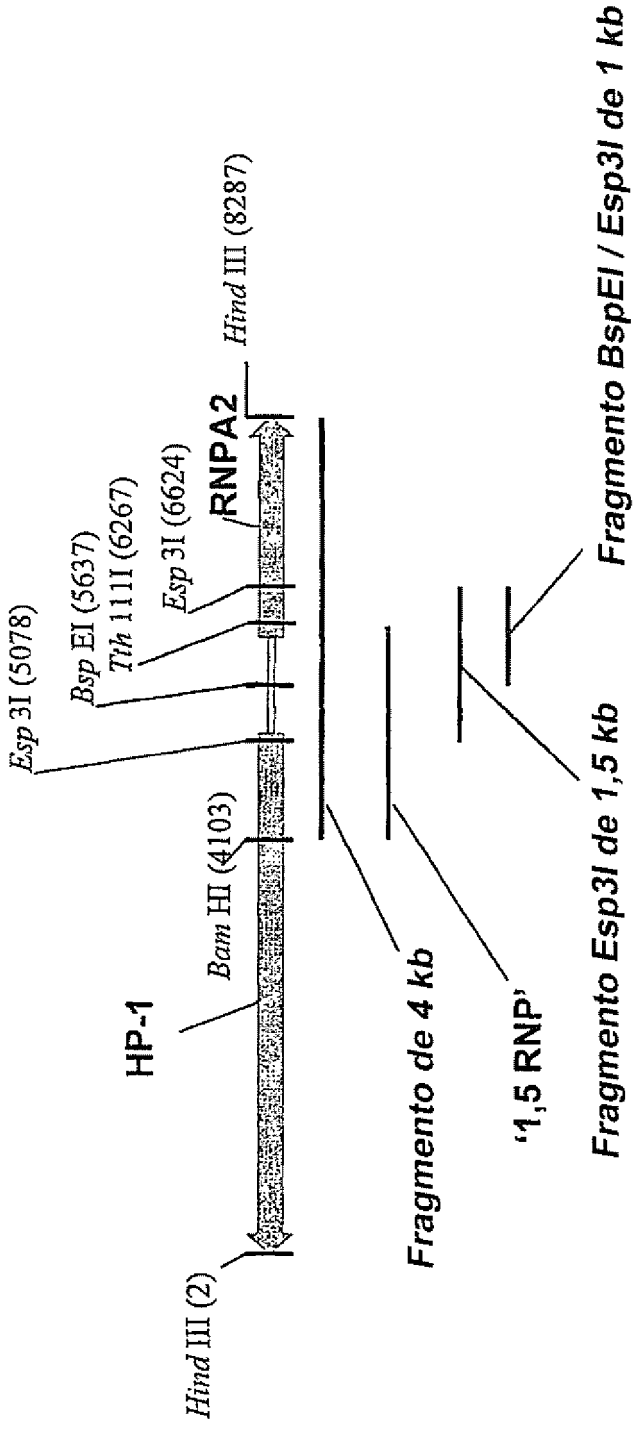

Esp 3I (5078)
Bsp EI (5637)
Tth 111I (6267)
Esp 3I (6624)
Hind III (8287)
RNPA2
HP-1
Bam HI (4103)
Hind III (2)
Fragmento de 4 kb
'1,5 RNP'
Fragmento Esp3I de 1,5 kb
Fragmento BspEI / Esp3I de 1 kb

Figura 2

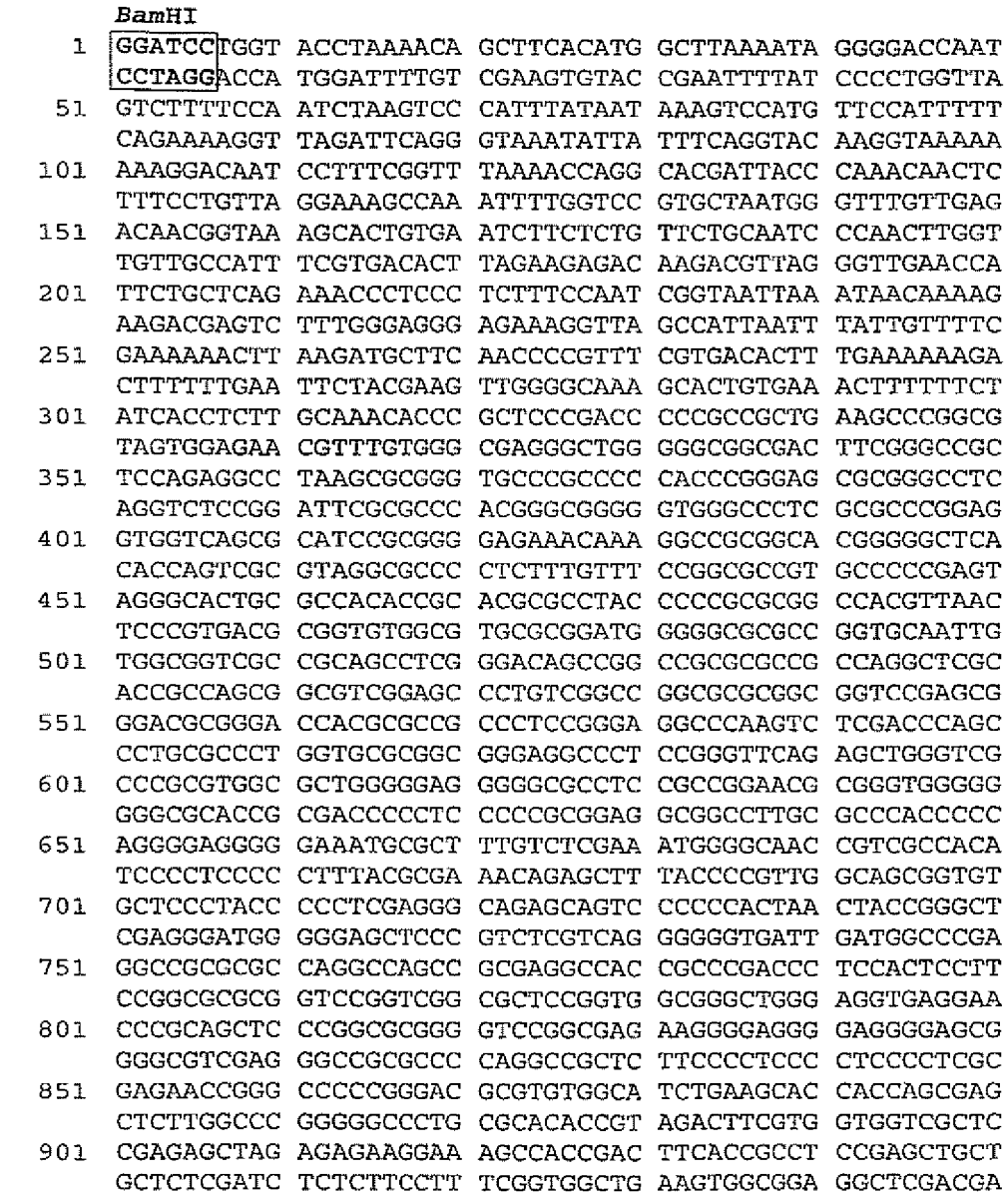

BamHI
1 GGATCCTGGT ACCTAAAACA GCTTCACATG GCTTAAAATA GGGGACCAAT
CCTAGGACCA TGGATTTTGT CGAAGTGTAC CGAATTTTAT CCCCTGGTTA
51 GTCTTTTCCA ATCTAAGTCC CATTTATAAT AAAGTCCATG TTCCATTTTT
CAGAAAAGGT TAGATTCAGG GTAAATATTA TTTCAGGTAC AAGGTAAAAA
101 AAAGGACAAT CCTTTCGGTT TAAAACCAGG CACGATTACC CAAACAACTC
TTTCCTGTTA GGAAAGCCAA ATTTTGGTCC GTGCTAATGG GTTTGTTGAG
151 ACAACGGTAA AGCACTGTGA ATCTTCTCTG TTCTGCAATC CCAACTTGGT
TGTTGCCATT TCGTGACACT TAGAAGAGAC AAGACGTTAG GGTTGAACCA
201 TTCTGCTCAG AAACCCTCCC TCTTTCCAAT CGGTAATTAA ATAACAAAAG
AAGACGAGTC TTTGGGAGGG AGAAAGGTTA GCCATTAATT TATTGTTTTC
251 GAAAAAACTT AAGATGCTTC AACCCCGTTT CGTGACACTT TGAAAAAAGA
CTTTTTTGAA TTCTACGAAG TTGGGGCAAA GCACTGTGAA ACTTTTTTCT
301 ATCACCTCTT GCAAACACCC GCTCCCGACC CCCGCCGCTG AAGCCCGGCG
TAGTGGAGAA CGTTTGTGGG CGAGGGCTGG GGGCGGCGAC TTCGGGCCGC
351 TCCAGAGGCC TAAGCGCGGG TGCCCGCCCC CACCCGGGAG CGCGGGCCTC
AGGTCTCCGG ATTCGCGCCC ACGGGCGGGG GTGGGCCCTC GCGCCCGGAG
401 GTGGTCAGCG CATCCGCGGG GAGAAACAAA GGCCGCGGCA CGGGGGCTCA
CACCAGTCGC GTAGGCGCCC CTCTTTGTTT CCGGCGCCGT GCCCCCGAGT
451 AGGGCACTGC GCCACACCGC ACGCGCCTAC CCCCGCGCGG CCACGTTAAC
TCCCGTGACG CGGTGTGGCG TGCGCGGATG GGGGCGCGCC GGTGCAATTG
501 TGGCGGTCGC CGCAGCCTCG GGACAGCCGG CCGCGCGCCG CCAGGCTCGC
ACCGCCAGCG GCGTCGGAGC CCTGTCGGCC GGCGCGCGGC GGTCCGAGCG
551 GGACGCGGGA CCACGCGCCG CCCTCCGGGA GGCCCAAGTC TCGACCCAGC
CCTGCGCCCT GGTGCGCGGC GGGAGGCCCT CCGGGTTCAG AGCTGGGTCG
601 CCCGCGTGGC GCTGGGGGAG GGGGCGCCTC CGCCGGAACG CGGGTGGGGG
GGGCGCACCG CGACCCCCTC CCCCGCGGAG GCGGCCTTGC GCCCACCCCC
651 AGGGGAGGGG GAAATGCGCT TTGTCTCGAA ATGGGGCAAC CGTCGCCACA
TCCCCTCCCC CTTTACGCGA AACAGAGCTT TACCCCGTTG GCAGCGGTGT
701 GCTCCCTACC CCCTCGAGGG CAGAGCAGTC CCCCCACTAA CTACCGGGCT
CGAGGGATGG GGGAGCTCCC GTCTCGTCAG GGGGGTGATT GATGGCCCGA
751 GGCCGCGCGC CAGGCCAGCC GCGAGGCCAC CGCCCGACCC TCCACTCCTT
CCGGCGCGCG GTCCGGTCGG CGCTCCGGTG GCGGGCTGGG AGGTGAGGAA
801 CCCGCAGCTC CCGGCGCGGG GTCCGGCGAG AAGGGGAGGG GAGGGGAGCG
GGGCGTCGAG GGCCGCGCCC CAGGCCGCTC TTCCCCTCCC CTCCCCTCGC
851 GAGAACCGGG CCCCCGGGAC GCGTGTGGCA TCTGAAGCAC CACCAGCGAG
CTCTTGGCCC GGGGGCCCTG CGCACACCGT AGACTTCGTG GTGGTCGCTC
901 CGAGAGCTAG AGAGAAGGAA AGCCACCGAC TTCACCGCCT CCGAGCTGCT
GCTCTCGATC TCTCTTCCTT TCGGTGGCTG AAGTGGCGGA GGCTCGACGA

Figura 2 (cont.)

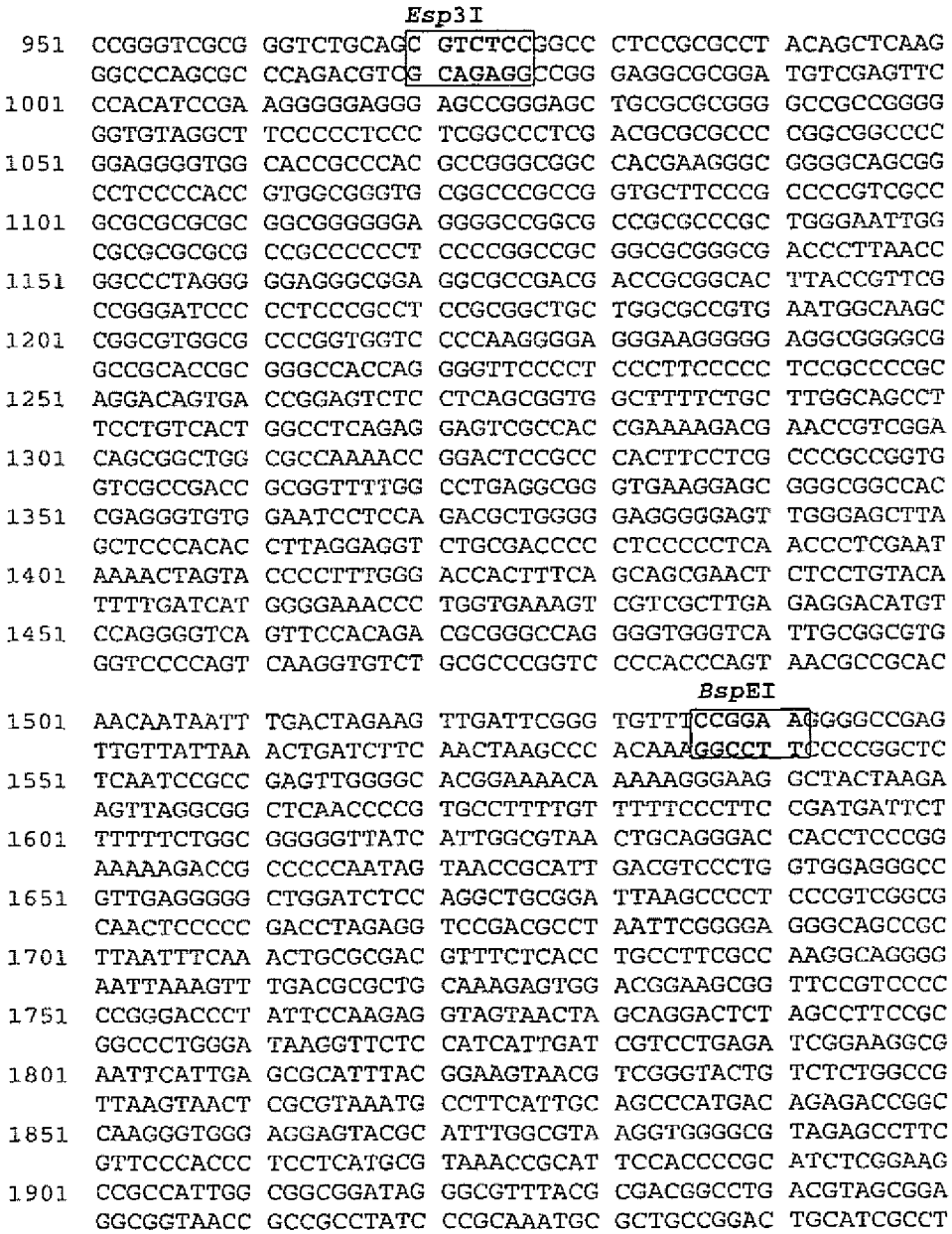


Figura 2 (cont.)

```
1951  AGACGCGTTA GTGGGGGGGA AGGTTCTAGA AAAGCGGCGG CAGCGGCTCT
      TCTGCGCAAT CACCCCCCCT TCCAAGATCT TTTCGCCGCC GTCGCCGAGA
2001  AGCGGCAGTA GCAGCAGCGC CGGGTCCCGT GCGGAGGTGC TCCTCGCAGA
      TCGCCGTCAT CGTCGTCGCG GCCCAGGGCA CGCCTCCACG AGGAGCGTCT
2051  GTTGTTTCTC GAGCAGCGGC AGTTCTCACT ACAGCGCCAG GACGAGTCCG
      CAACAAAGAG CTCGTCGCCG TCAAGAGTGA TGTCGCGGTC CTGCTCAGGC
2101  GTTCGTGTTC GTCCGCGGAG ATCTCTCTCA TCTCGCTCGG CTGCGGGAAA
      CAAGCACAAG CAGGCGCCTC TAGAGAGAGT AGAGCGAGCC GACGCCCTTT
                 Tth111I
2151  TCGGGCTGAA GCGACTGAGT CCGCGATGGA GGTAACGGGT TTGAAATCAA
      AGCCCGACTT CGCTGACTCA GGCGCTACCT CCATTGCCCA AACTTTAGTT
2201  TGAGTTATTG AAAAGGGCAT GGCGAGGCCG TTGGCGCCTC AGTGGAAGTC
      ACTCAATAAC TTTTCCCGTA CCGCTCCGGC AACCGCGGAG TCACCTTCAG
2251  GGCCAGCCGC CTCCGTGGGA GAGAGGCAGG AAATCGGACC AATTCAGTAG
      CCGGTCGGCG GAGGCACCCT CTCTCCGTCC TTTAGCCTGG TTAAGTCATC
2301  CAGTGGGGCT TAAGGTTTAT GAACGGGGTC TTGAGCGGAG GCCTGAGCGT
      GTCACCCCGA ATTCCAAATA CTTGCCCCAG AACTCGCCTC CGGACTCGCA
2351  ACAAACAGCT TCCCCACCCT CAGCCTCCCG GCGCCATTTC CCTTCACTGG
      TGTTTGTCGA AGGGGTGGGA GTCGGAGGGC CGCGGTAAAG GGAAGTGACC
2401  GGGTGGGGGA TGGGGAGCTT TCACATGGCG GACGCTGCCC CGCTGGGGTG
      CCCACCCCCT ACCCCTCGAA AGTGTACCGC CTGCGACGGG GCGACCCCAC
2451  AAAGTGGGGC GCGGAGGCGG GAATTCTTAT TCCCTTTCTA AAGCACGCTG
      TTTCACCCCG CGCCTCCGCC CTTAAGAATA AGGGAAAGAT TTCGTGCGAC
                       Esp3I
2501  CTTCGGGGGC CACGGCGTCT CCTCGG
      GAAGCCCCCG GTGCCGCAGA GGAGCC
```

Figura 3

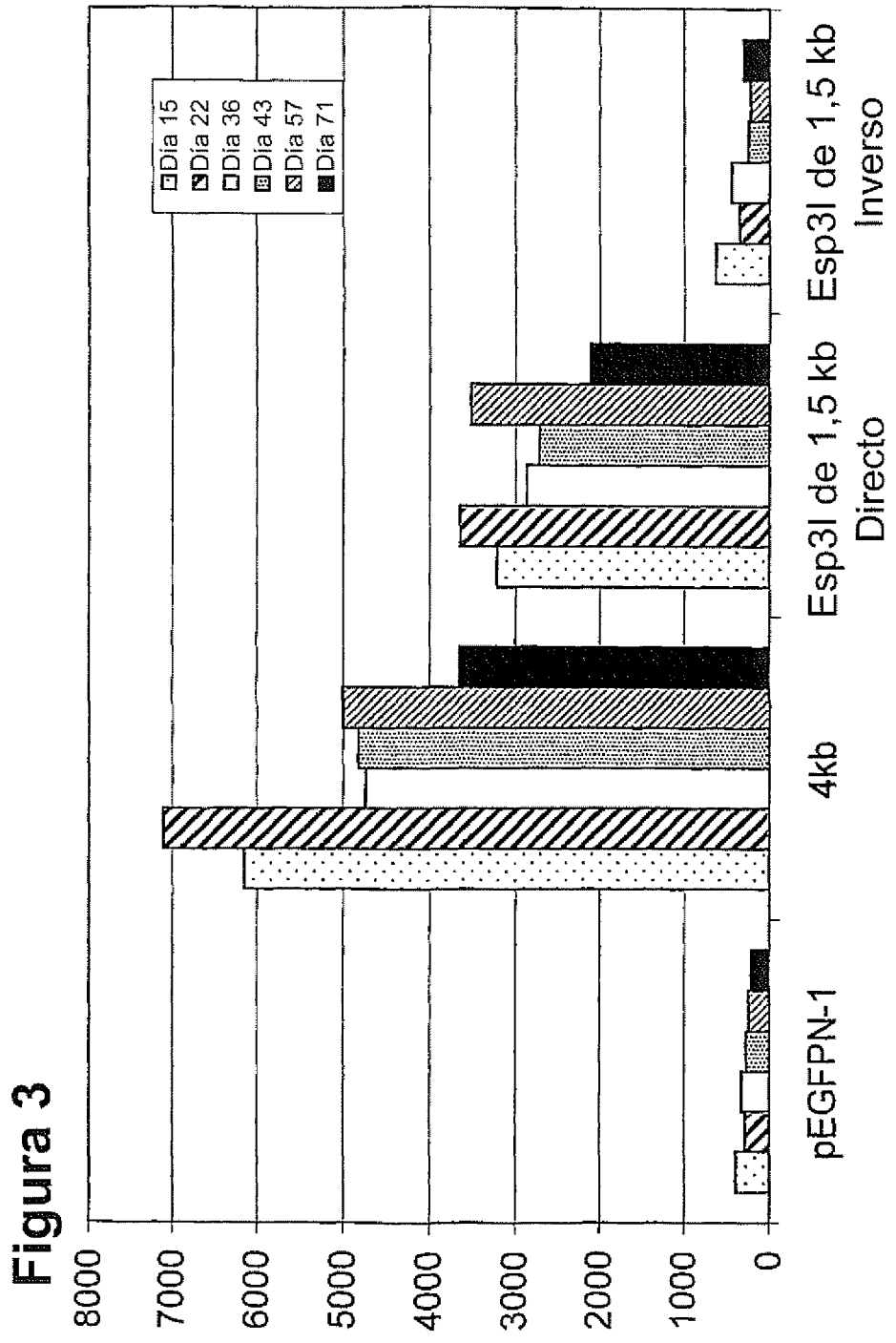

Dia 15
Dia 22
Dia 36
Dia 43
Dia 57
Dia 71
8000
7000
6000
5000
4000
3000
2000
1000
0
pEGFPN-1
4kb
Esp3I de 1,5 kb Directo
Esp3I de 1,5 kb Inverso

Figura 4

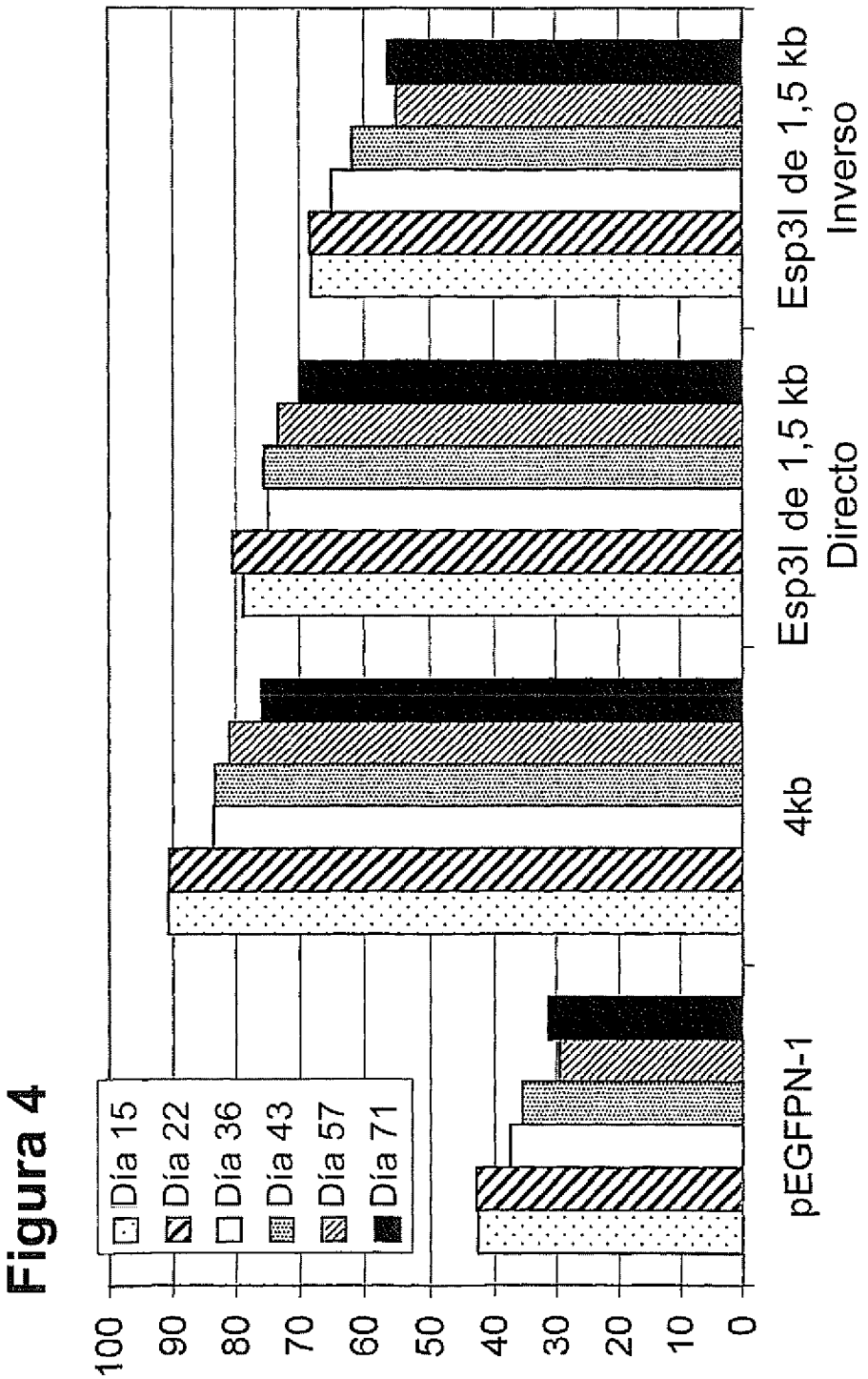

Día 15
Día 22
Día 36
Día 43
Día 57
Día 71
100
90
80
70
60
50
40
30
20
10
0
pEGFPN-1
4kb
Esp3I de 1,5 kb
Directo
Esp3I de 1,5 kb
Inverso

Figura 5

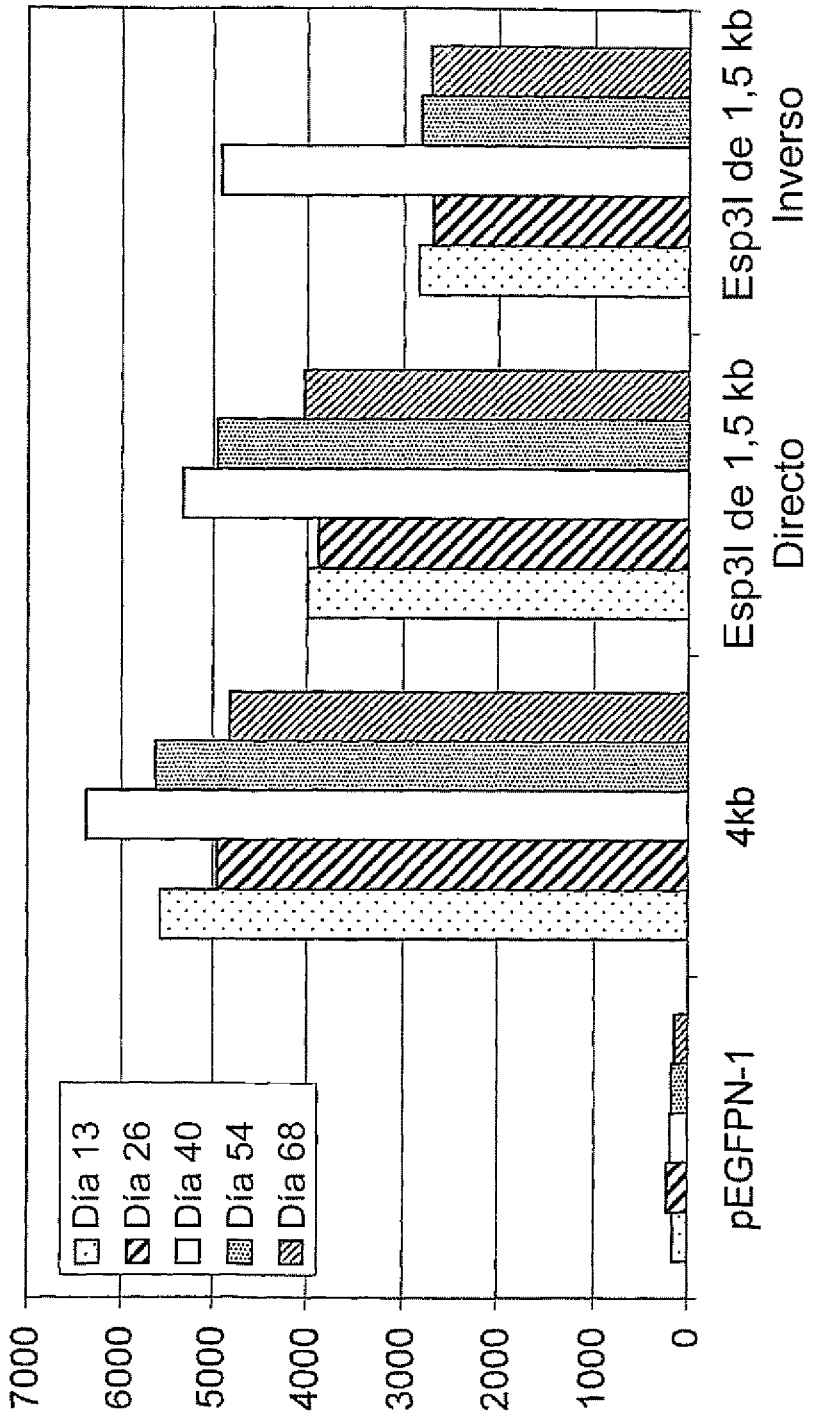

7000
6000
5000
4000
3000
2000
1000
0
Día 13
Día 26
Día 40
Día 54
Día 68
pEGFPN-1
4kb
Esp3I de 1,5 kb
Directo
Esp3I de 1,5 kb
Inverso

Figura 6

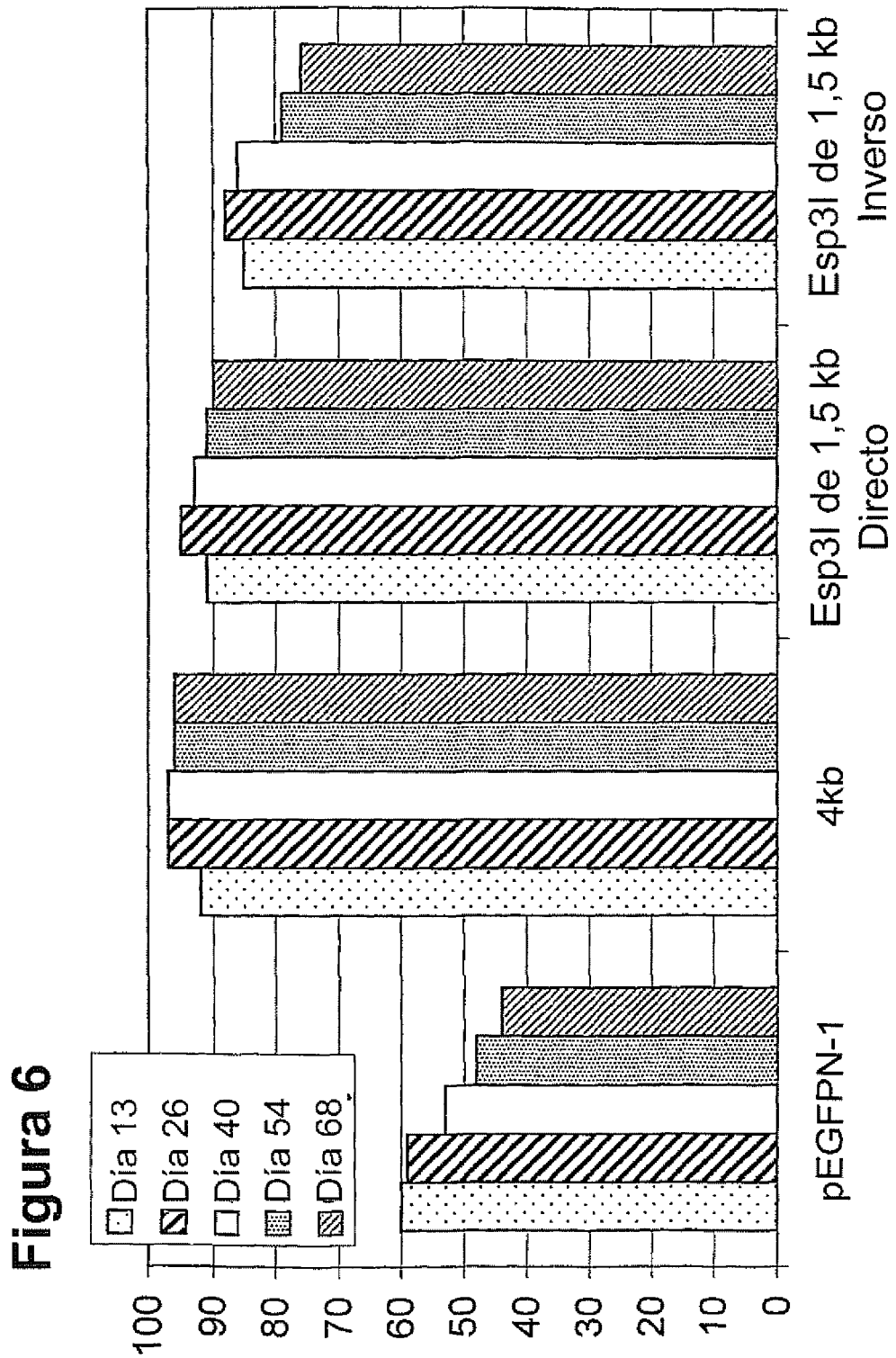

Día 13
Día 26
Día 40
Día 54
Día 68
100
90
80
70
60
50
40
30
20
10
0
pEGFPN-1
4kb
Esp3I de 1,5 kb
Directo
Esp3I de 1,5 kb
Inverso

Figura 7

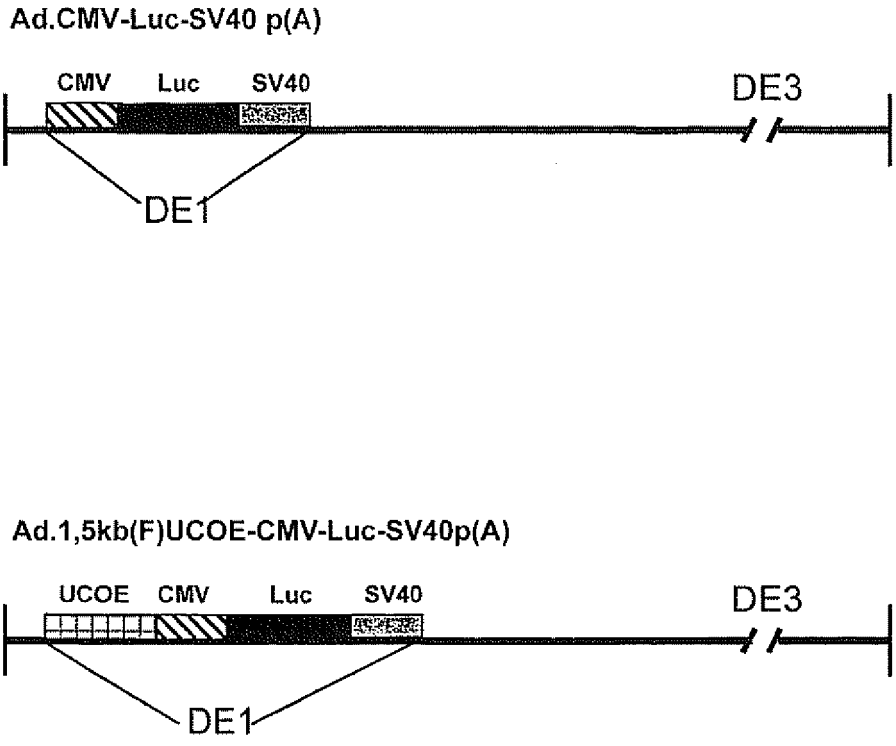

Ad.CMV-Luc-SV40 p(A)
CMV
Luc
SV40
DE3
DE1
Ad.1,5kb(F)UCOE-CMV-Luc-SV40p(A)
UCOE
CMV
Luc
SV40
DE3
DE1

Figura 8

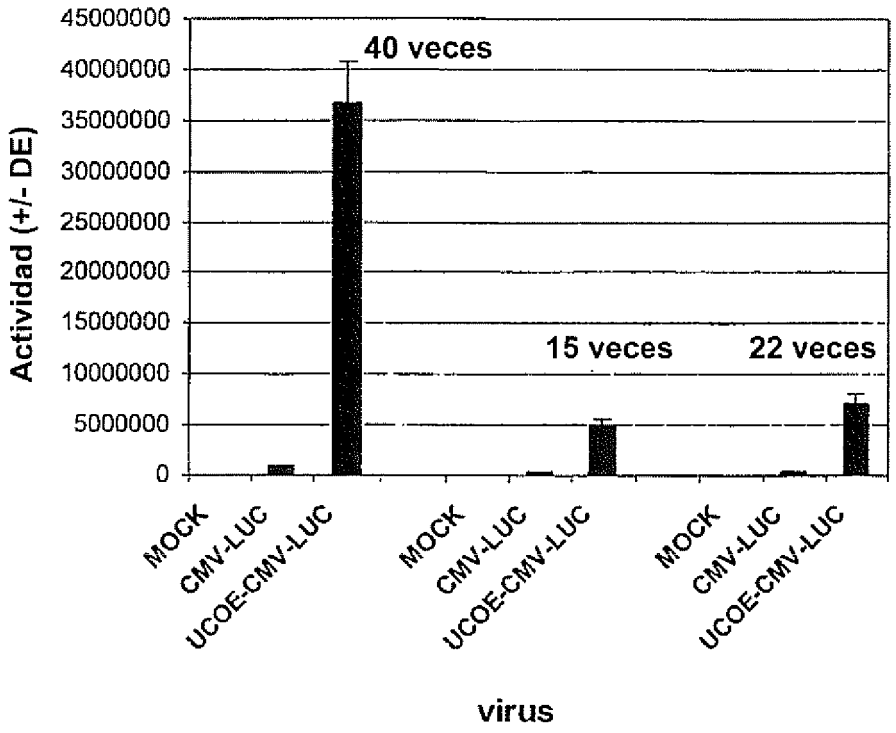

45000000
40000000
35000000
30000000
25000000
20000000
15000000
10000000
5000000
0
Actividad (+/- DE)
40 veces
15 veces
22 veces
MOCK
CMV-LUC
UCOE-CMV-LUC
MOCK
CMV-LUC
UCOE-CMV-LUC
MOCK
CMV-LUC
UCOE-CMV-LUC
virus

Figura 9

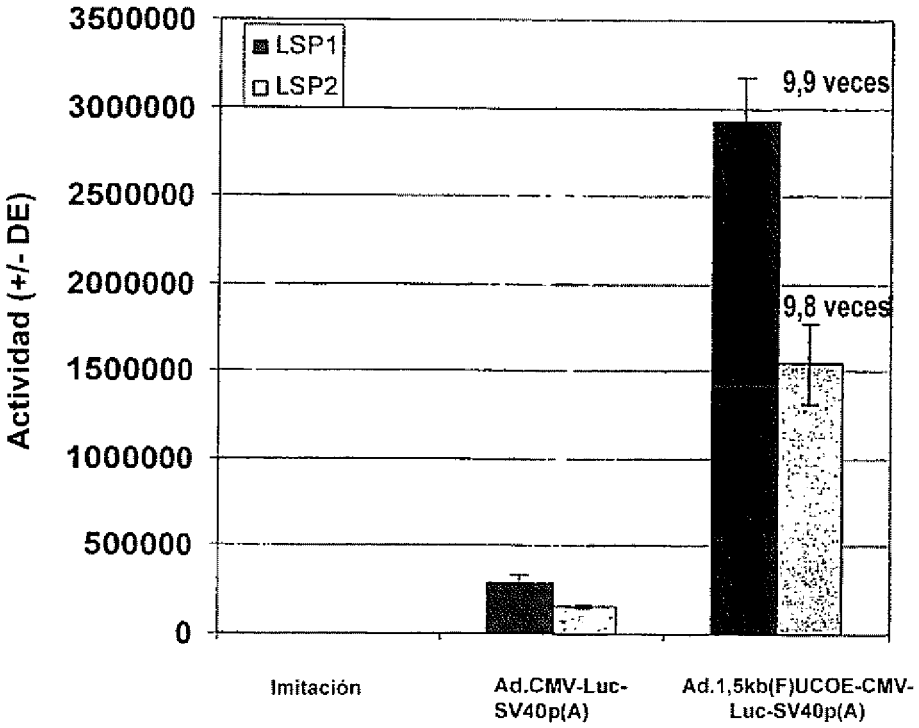

3500000
3000000
2500000
2000000
1500000
1000000
500000
0
Actividad (+/- DE)
LSP1
LSP2
9,9 veces
9,8 veces
Imitación
Ad.CMV-Luc-SV40p(A)
Ad.1,5kb(F)UCOE-CMV-Luc-SV40p(A)

Figura 10

A

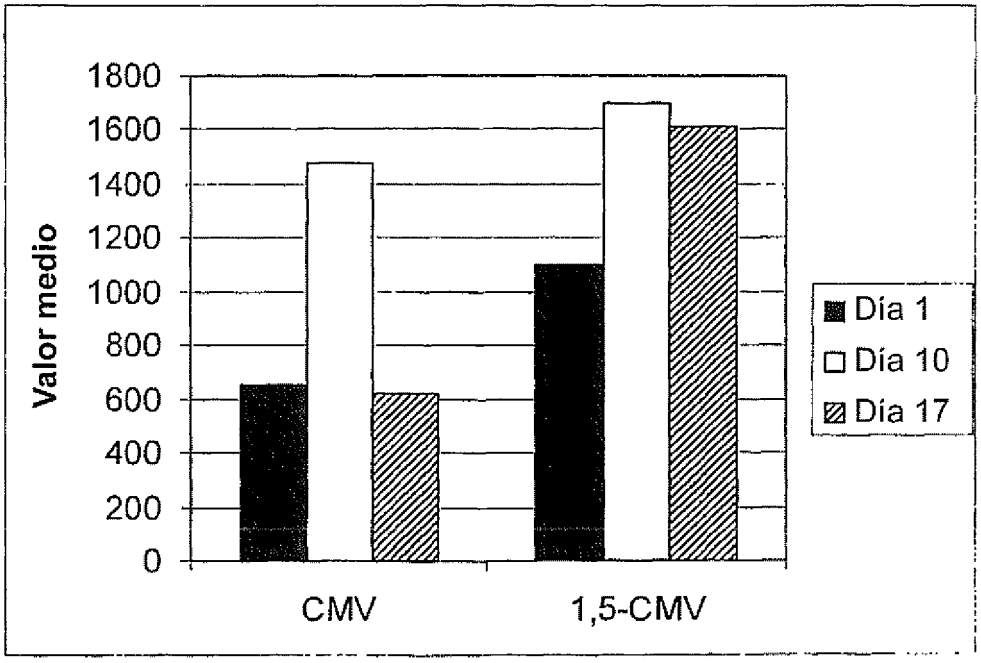

1800
1600
1400
1200
1000
800
600
400
200
0
Valor medio
CMV
1,5-CMV
Día 1
Día 10
Día 17

Figura 10 (cont.)

B

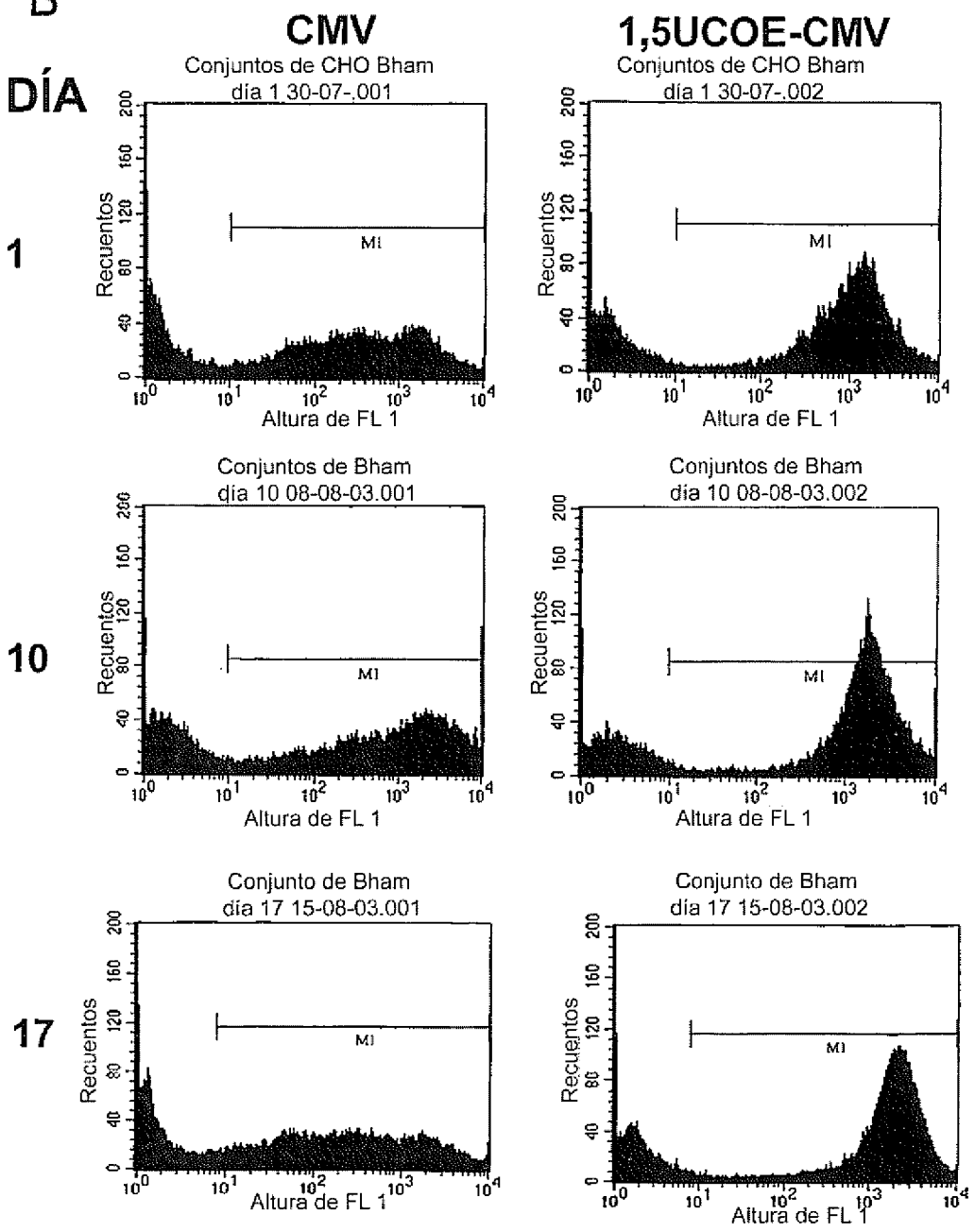


Figura 11

A

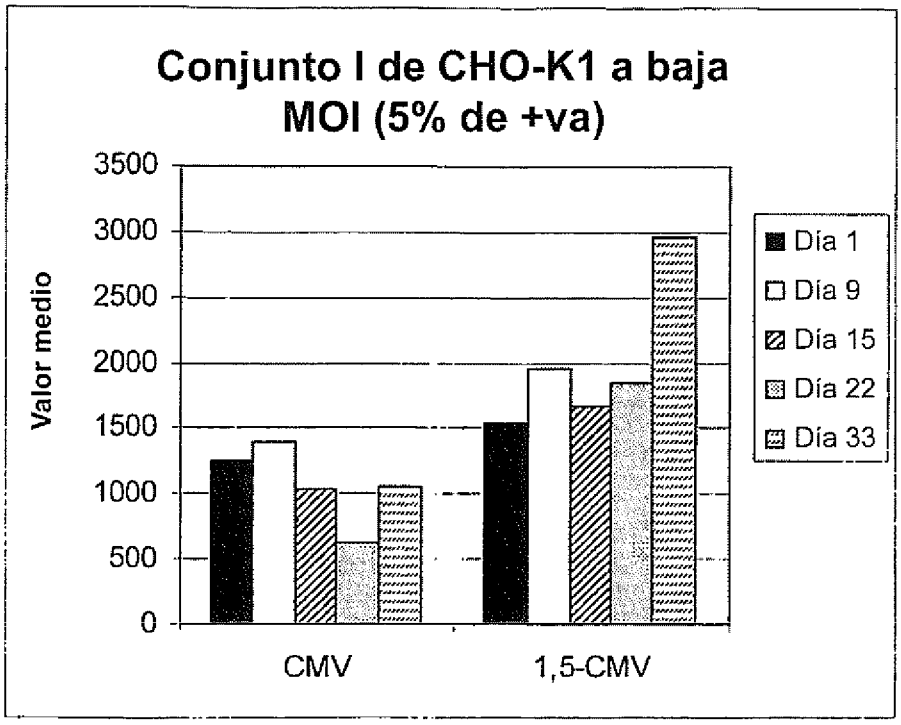

Conjunto I de CHO-K1 a baja MOI (5% de +va)
Valor medio
3500
3000
2500
2000
1500
1000
500
0
CMV
1,5-CMV
Día 1
Día 9
Día 15
Día 22
Día 33

Figura 11 (cont.)

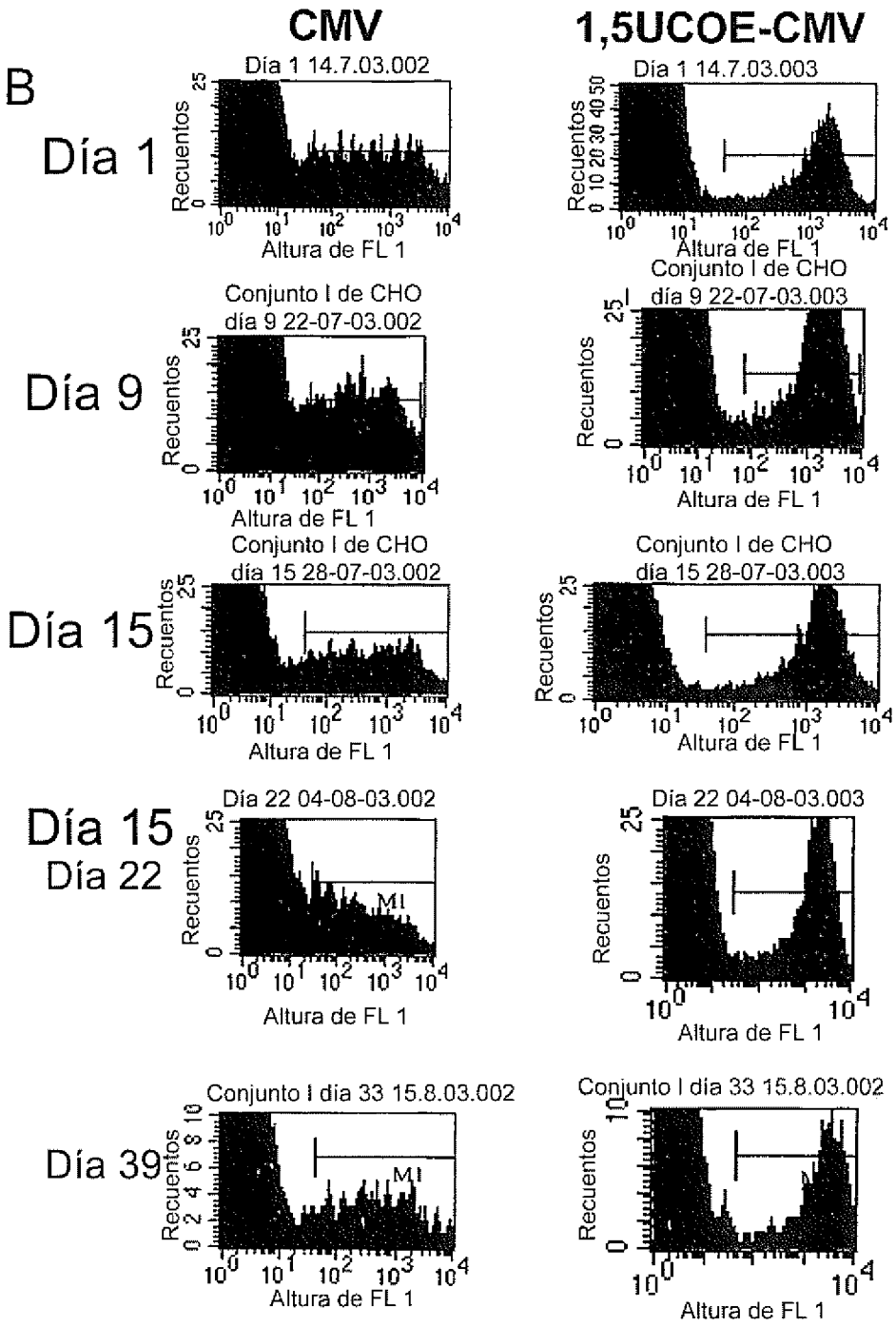


Figura 12

A

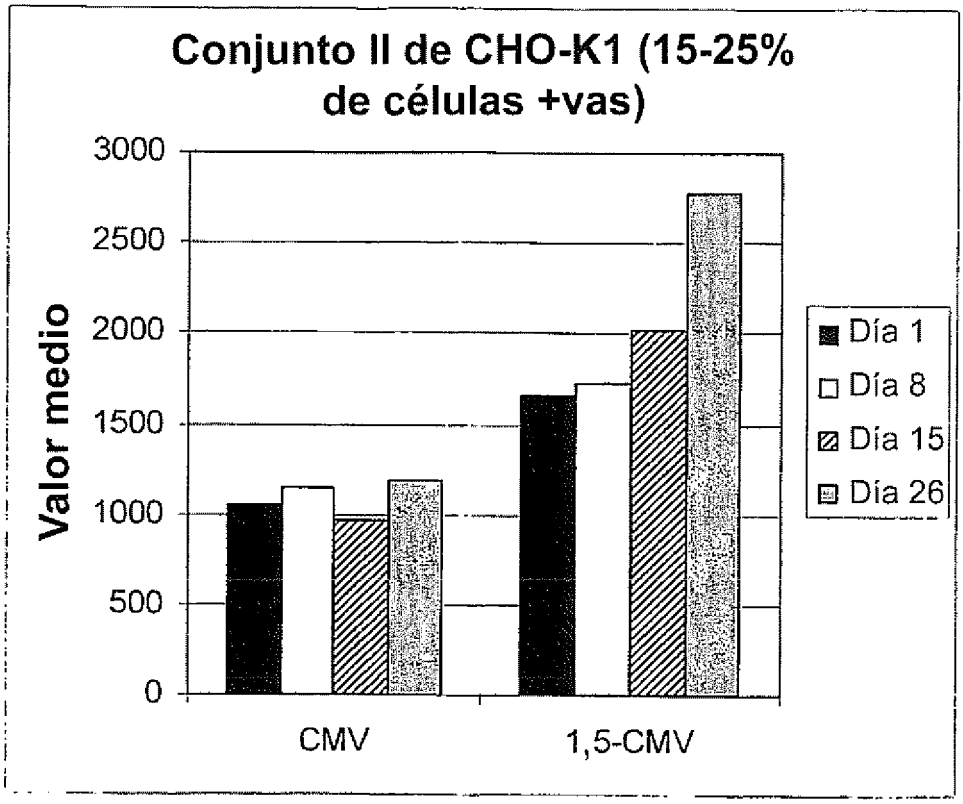

Conjunto II de CHO-K1 (15-25% de células +vas)
Valor medio
3000
2500
2000
1500
1000
500
0
CMV
1,5-CMV
Día 1
Día 8
Día 15
Día 26

Figura 12 (cont.)

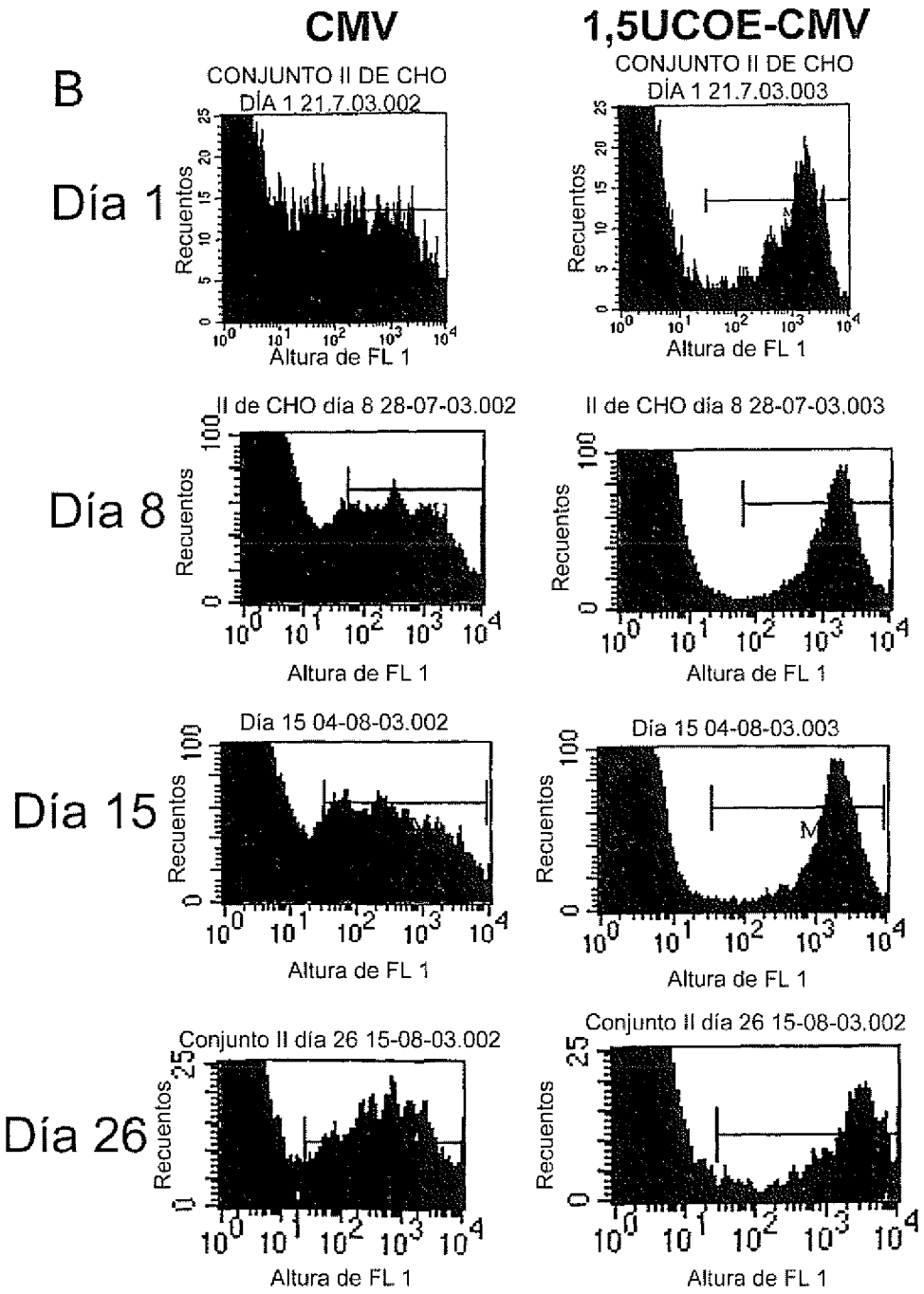


Figura 13

A

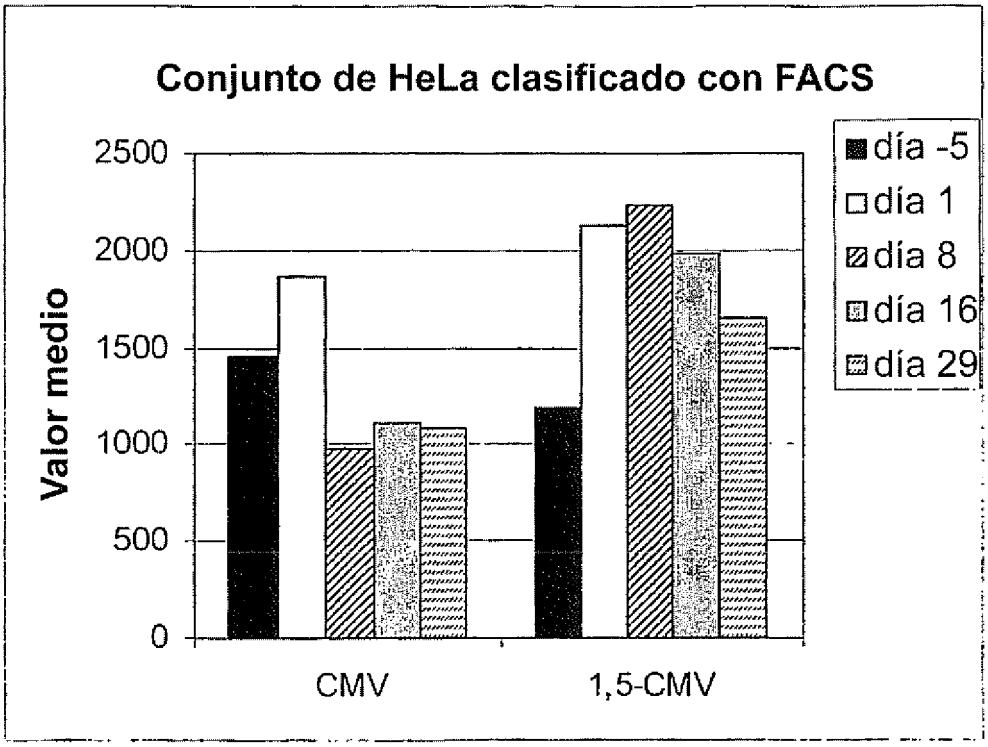

Conjunto de HeLa clasificado con FACS
Valor medio
2500
2000
1500
1000
500
0
CMV
1,5-CMV
día -5
día 1
día 8
día 16
día 29

Figura 13 (cont.)

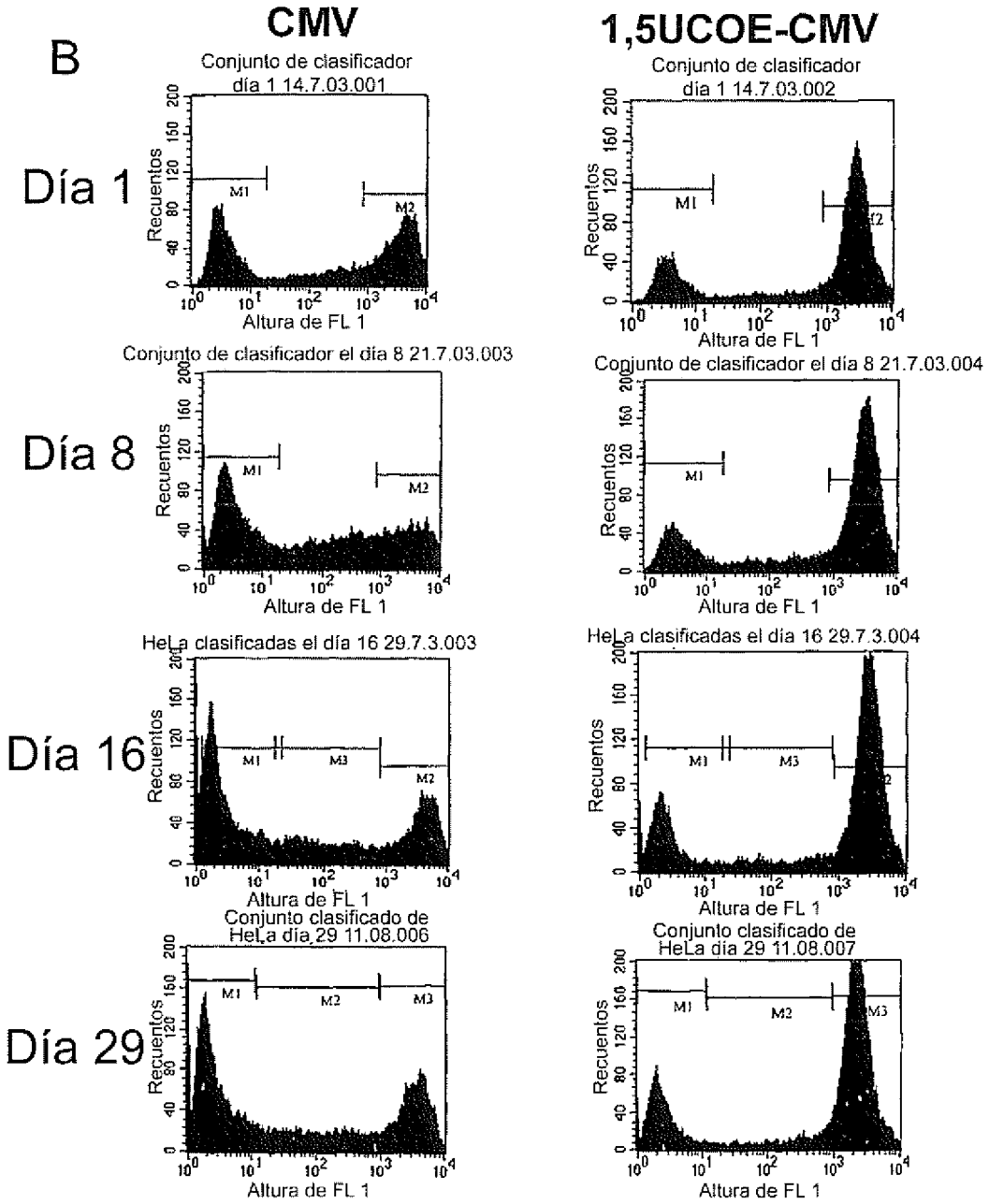


Figura 14

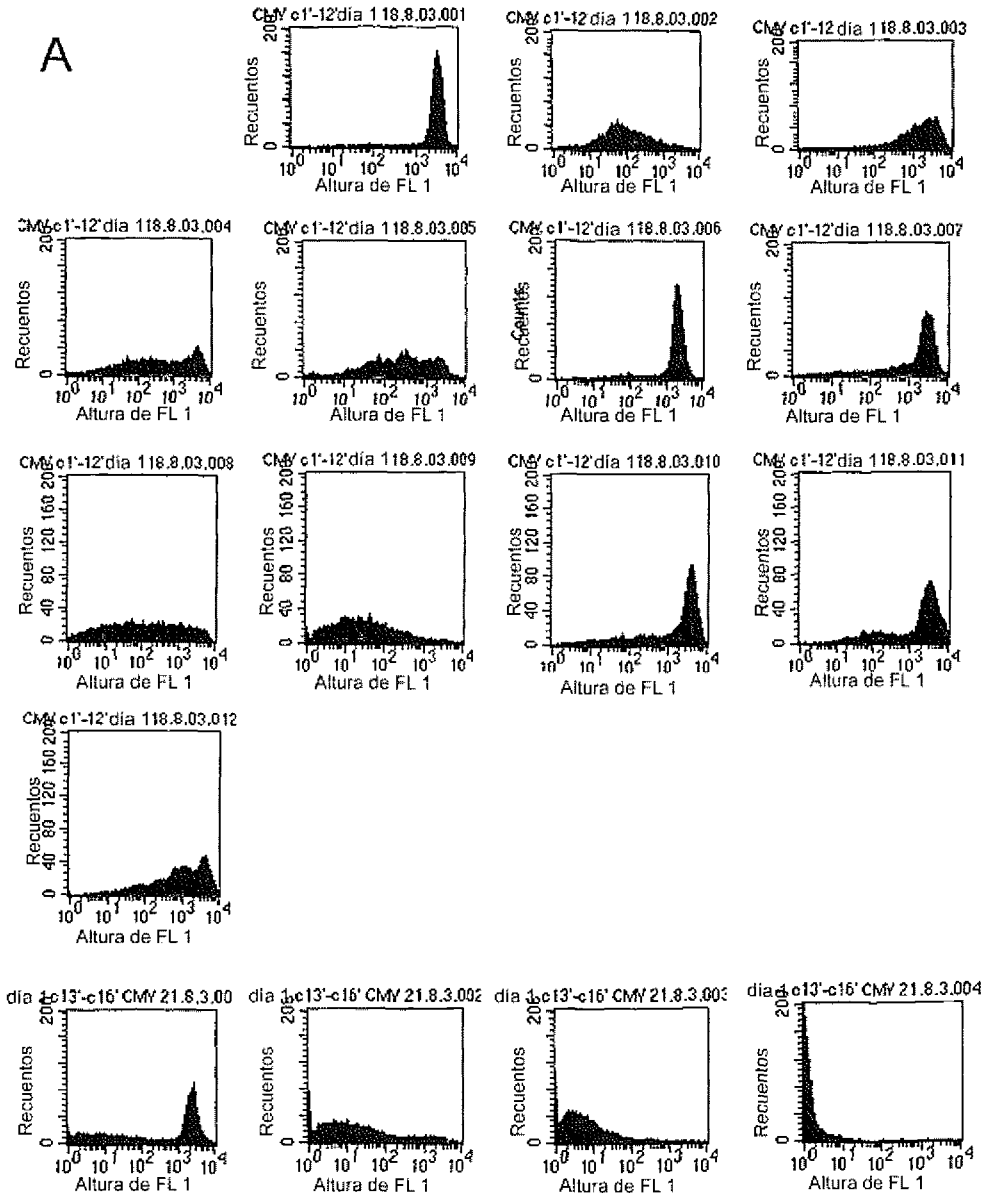

A
CMV c1'-12'día 118.8.03.001
CMV c1'-12'día 118.8.03.002
CMV c1'-12'día 118.8.03.003
CMV c1'-12'día 118.8.03.004
CMV c1'-12'día 118.8.03.005
CMV c1'-12'día 118.8.03.006
CMV c1'-12'día 118.8.03.007
CMV c1'-12'día 118.8.03.008
CMV c1'-12'día 118.8.03.009
CMV c1'-12'día 118.8.03.010
CMV c1'-12'día 118.8.03.011
CMV c1'-12'día 118.8.03.012
día 1 c13'-c16' CMV 21.8.3.001
día 1 c13'-c16' CMV 21.8.3.002
día 1 c13'-c16' CMV 21.8.3.003
día 1 c13'-c16' CMV 21.8.3.004
Recuentos
Altura de FL 1

Figura 14 (cont.)

B

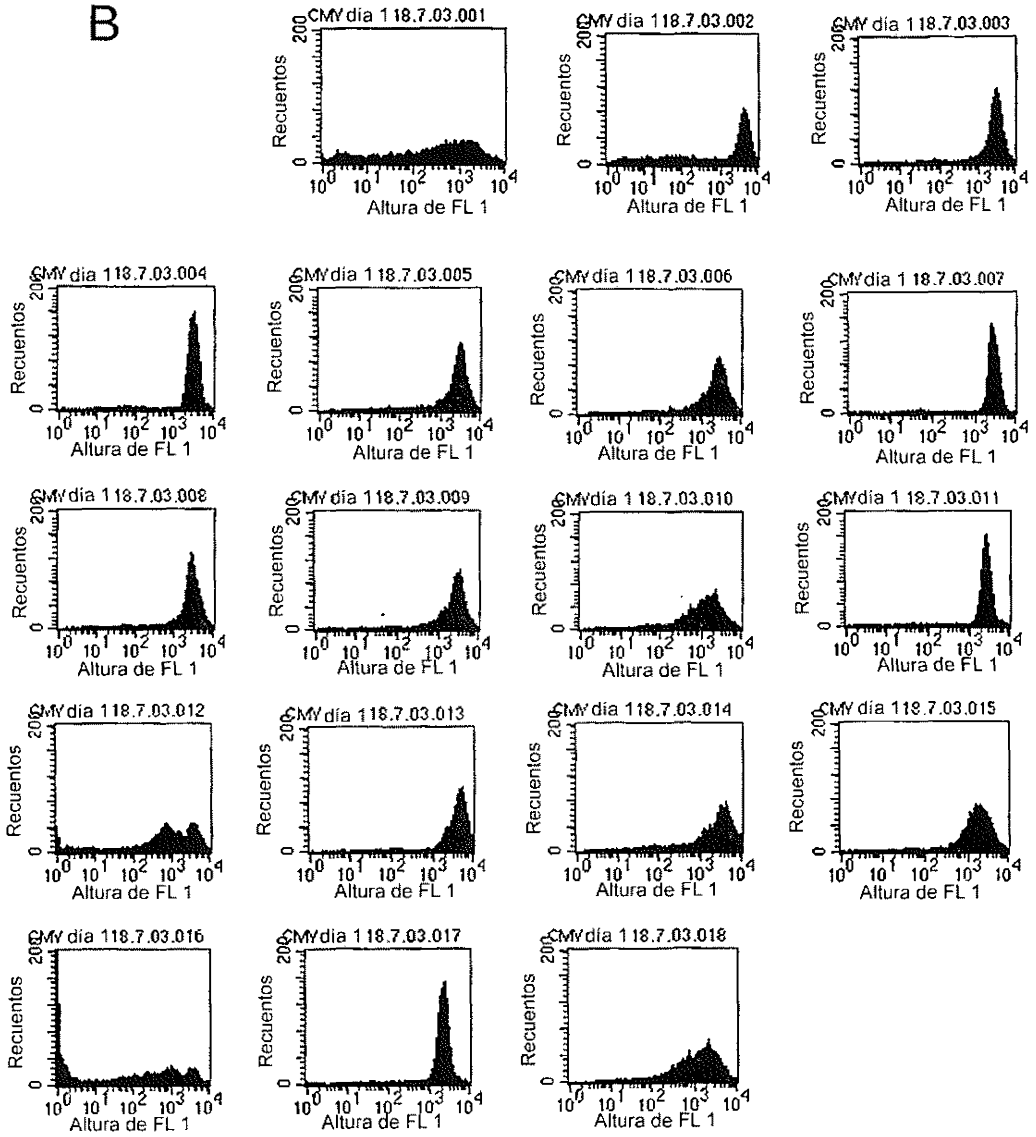


Figura 14 (cont.)

C

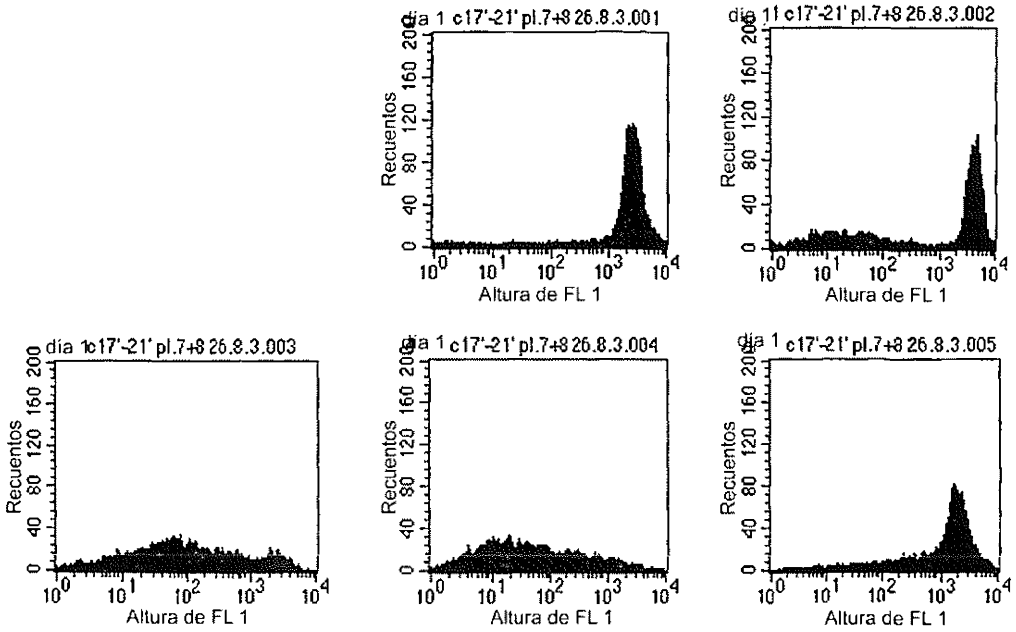


Figura 14 (cont.)

D

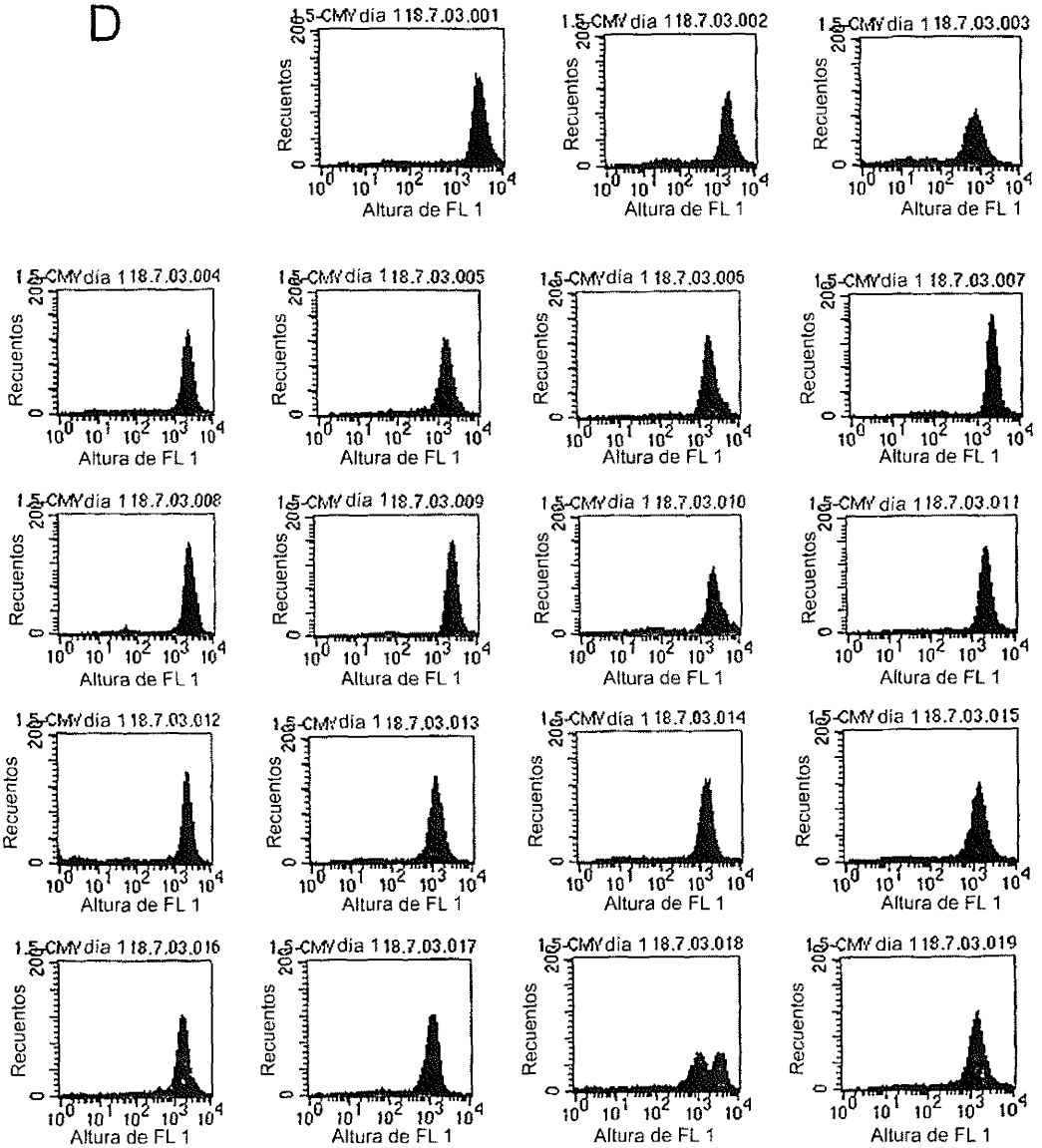


Figura 14 (cont.)

E

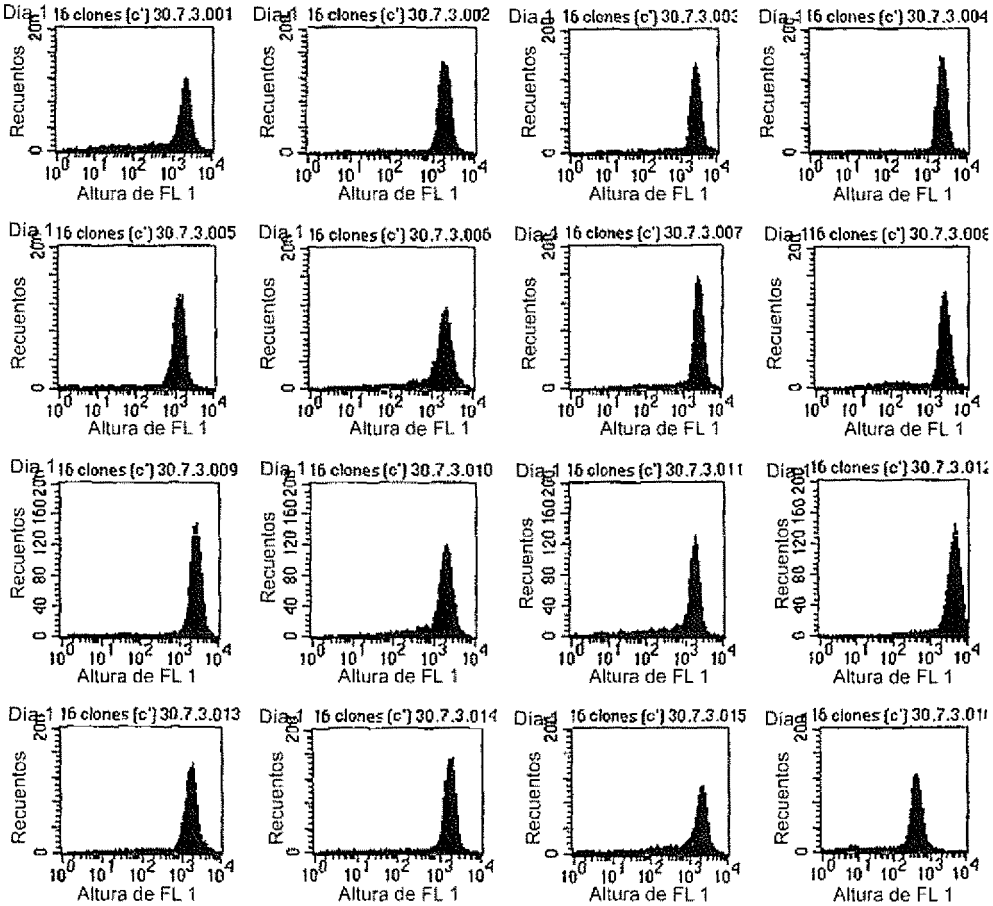
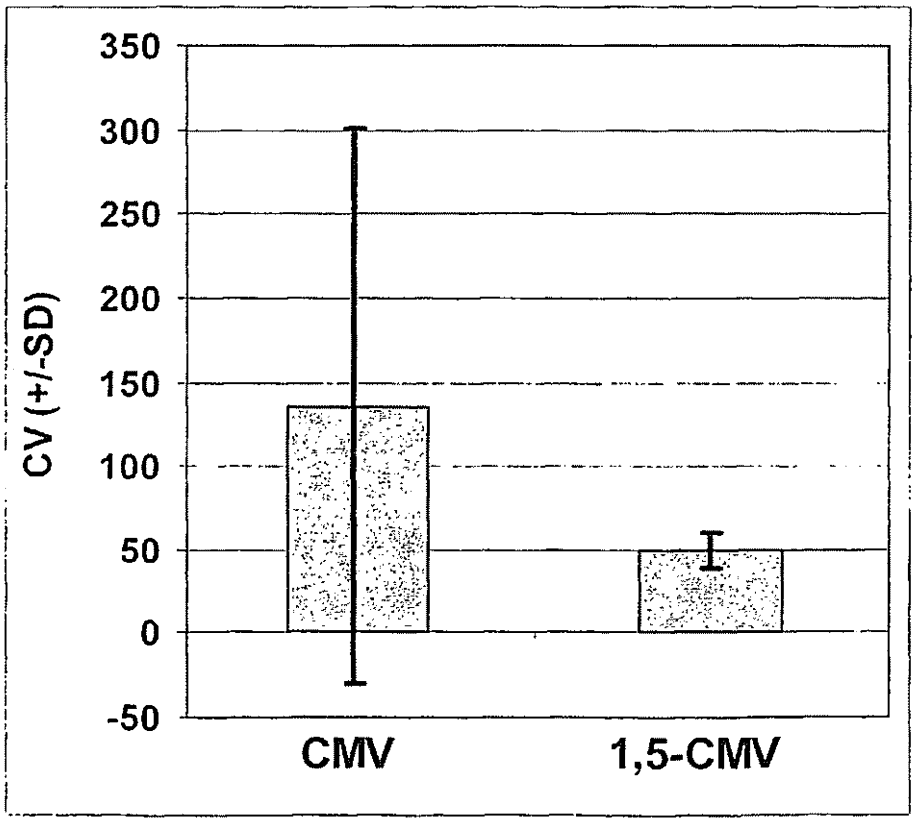


Figura 15



LISTA DE SECUENCIAS

<110> ML Laboratories PLC williams, Steven G Crombie, Robert L
<120> Elementos de expresión mejorados
<130> P101405
<160> 2
<170> PatentIn versión 3.1
<210> 1
<211> 1546
<212> ADN
<213> *Homo sapiens*

<400> 1

ggccctccgc gcctacagct caagccacat ccgaaggggg agggagccgg gagctgcgcg 60
cggggccgcc gggggggaggg gtggcaccgc ccacgccggg cggccacgaa gggcggggca 120
gcgggcgcgc gcgcggcggg gggaggggcc ggcgccgcgc ccgctgggaa ttggggccct 180
agggggaggg cggaggcgcc gacgaccgcg gcacttaccg ttcgcggcgt ggcgcccggt 240
ggtccccaag gggagggaag ggggaggcgg ggcgaggaca gtgaccggag tctcctcagc 300
ggtggctttt ctgcttggca gcctcagcgg ctggcgccaa aaccggactc cgcccacttc 360
ctcgcccgcc ggtgcgaggg tgtggaatcc tccagacgct gggggagggg gagttgggag 420
cttaaaaact agtacccctt tgggaccact ttcagcagcg aactctcctg tacaccaggg 480
gtcagttcca cagacgcggg ccaggggtgg gtcattgcgg cgtgaacaat aatttgacta 540
gaagttgatt cgggtgtttc cggaaggggc cgagtcaatc cgccgagttg gggcacggaa 600
aacaaaaagg gaaggctact aagatttttc tggcgggggt tatcattggc gtaactgcag 660
ggaccacctc ccgggttgag ggggctggat ctccaggctg cggattaagc ccctcccgtc 720
ggcgttaatt tcaaactgcg cgacgtttct cacctgcctt cgccaaggca ggggccggga 780
ccctattcca agaggtagta actagcagga ctctagcctt ccgcaattca ttgagcgcat 840
ttacggaagt aacgtcgggt actgtctctg gccgcaaggg tgggaggagt acgcatttgg 900
cgtaaggtgg ggcgtagagc cttcccgcca ttggcggcgg atagggcgtt tacgcgacgg 960
cctgacgtag cggaagacgc gttagtgggg gggaaggttc tagaaaagcg gcggcagcgg 1020
ctctagcggc agtagcagca gcgccgggtc ccgtgcggag gtgctcctcg cagagttgtt 1080
tctcgagcag cggcagttct cactacagcg ccaggacgag tccggttcgt gttcgtccgc 1140
ggagatctct ctcatctcgc tcggctgcgg gaaatcgggc tgaagcgact gagtccgcga 1200
tggaggtaac gggtttgaaa tcaatgagtt attgaaaagg gcatggcgag gccgttggcg 1260
cctcagtgga agtcggccag ccgcctccgt gggagagagg caggaaatcg gaccaattca 1320
gtagcagtgg ggcttaaggt ttatgaacgg ggtcttgagc ggaggcctga gcgtacaaac 1380
agcttcccca ccctcagcct cccggcgcca tttcccttca ctgggggtgg gggatgggga 1440
gctttcacat ggcggacgct gccccgctgg ggtgaaagtg gggcgcggag gcgggaattc 1500
ttattccctt tctaaagcac gctgcttcgg gggccacggc gtctcc 1546

<210> 2
<211> 987

<212> ADN

<213> *Homo sapiens*

<400> 2

ccggaagggg ccgagtcaat ccgccgagtt ggggcacgga aaacaaaaag ggaaggctac 60

taagattttt ctggcggggg ttatcattgg cgtaactgca gggaccacct cccgggttga 120

gggggctgga tctccaggct gcggattaag cccctcccgt cggcgttaat ttcaaactgc 180

gcgacgtttc tcacctgcct tcgccaaggc aggggccggg accctattcc aagaggtagt 240

aactagcagg actctagcct tccgcaattc attgagcgca tttacggaag taacgtcggg 300

tactgtctct ggccgcaagg gtgggaggag tacgcatttg gcgtaaggtg gggcgtagag 360

ccttcccgcc attggcggcg gatagggcgt ttacgcgacg gcctgacgta gcggaagacg 420

cgttagtggg ggggaaggtt ctagaaaagc ggcggcagcg gctctagcgg cagtagcagc 480

agcgccgggt cccgtgcgga ggtgctcctc gcagagttgt ttctcgagca gcggcagttc 540

tcactacagc gccaggacga gtccggttcg tgttcgtccg cggagatctc tctcatctcg 600

ctcggctgcg ggaaatcggg ctgaagcgac tgagtccgcg atggaggtaa cgggtttgaa 660

atcaatgagt tattgaaaag ggcatggcga ggccgttggc gcctcagtgg aagtcggcca 720

gccgcctccg tgggagagag gcaggaaatc ggaccaattc agtagcagtg gggcttaagg 780

tttatgaacg gggtcttgag cggaggcctg agcgtacaaa cagcttcccc accctcagcc 840

tcccggcgcc atttcccttc actgggggtg ggggatgggg agctttcaca tggcggacgc 900

tgccccgctg gggtgaaagt ggggcgcgga ggcgggaatt cttattccct ttctaaagca 960

cgctgcttcg ggggccacgg cgtctcc 987