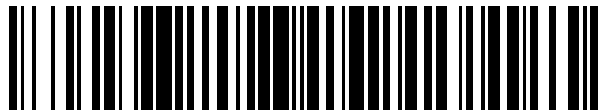


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 660 153**

51 Int. Cl.:

C07K 19/00 (2006.01)

C12N 15/65 (2006.01)

G01N 33/50 (2006.01)

C12N 5/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **06.06.2008 PCT/EP2008/057109**

87 Fecha y número de publicación internacional: **11.12.2008 WO08148881**

96 Fecha de presentación y número de la solicitud europea: **06.06.2008 E 08760679 (4)**

97 Fecha y número de publicación de la concesión europea: **13.12.2017 EP 2164877**

54 Título: **El marcador cuatrifuncional puro-DHFR y su empleo en la producción de proteína**

30 Prioridad:

07.06.2007 EP 07109829
11.06.2007 US 934078 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
21.03.2018

73 Titular/es:

MERCK SERONO S.A. (100.0%)
Centre Industriel
1267 Coinsins, CH

72 Inventor/es:

KOBR, MICHEL y
DUPRAZ, PHILIPPE

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 660 153 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).

DESCRIPCION

El marcador cuatrifuncional puro-DHFR y su empleo en la producción de proteína

Campo de la invención.

5 Esta invención se refiere a la producción industrial de proteínas. Más específicamente, la invención se refiere al marcador sustituto puro-DHFR, que corresponde a una fusión entre puromicina N-acetiltransferasa y dihidrofolato reductasa (DHFR). La invención se refiere además al uso de puro-DHFR para el cribado de células por su alta expresión de una proteína de interés.

Antecedentes de la invención.

10 La introducción de genes heterólogos en las células huésped de animales y el cribado por la expresión de los genes añadidos es un proceso largo y complicado. Típicamente, se ha de superar una serie de obstáculos: (i) la construcción de vectores de expresión grandes; (ii) la transfección y selección de clones con expresión estable a largo plazo, eventualmente en ausencia de presión selectiva; y (iii) el cribado o selección por altas tasas de expresión de la proteína heteróloga de interés.

1. Selección de clones que expresan el gen heterólogo.

15 La selección de los clones que tienen integrado el gen de interés se lleva a cabo usando un marcador de selección que confiere resistencia a una presión selectiva. La mayor parte de los marcadores de selección confieren resistencia a un antibiótico tal como, p. ej., neomicina, kanamicina, higromicina, gentamicina, cloranfenicol, puromicina, zeocina o bleomicina.

20 Cuando se generan clones de células que expresan un gen de interés a partir de vectores de expresión, las células hospedadoras son transfectadas típicamente con un vector de ADN plasmídico que codifica tanto la proteína de interés como el marcador de selección en el mismo vector. Con bastante frecuencia, la capacidad de un plásmido es limitada y el marcador de selección debe expresarse a partir de un segundo plásmido, que es cotransfectado con el plásmido que comprende el gen de interés.

25 La transfección estable conduce a la integración al azar del vector de expresión en el genoma de la célula huésped. El uso de presión selectiva, p. ej. administrando un antibiótico a los medios, eliminará todas las células que no integraron el vector que contiene el marcador de selección que proporciona resistencia al antibiótico correspondiente o presión selectiva. Si este marcador de selección está en el mismo vector que el gen de interés o si este marcador de selección está en un segundo vector y el vector que comprende el gen de interés fue co-integrado, las células expresarán tanto el marcador de selección como el gen de interés. Sin embargo, muchas veces se observa que el nivel de expresión del gen de interés es muy variable dependiendo del sitio de integración.

30 Además, cuando se elimina la presión selectiva, la expresión suele hacerse bastante inestable o incluso se extingue. Solo un pequeño número de transfectantes iniciales proporciona una expresión alta y estable a largo plazo y se necesita tiempo para identificar estos clones en una gran población de candidatos. Típicamente, los candidatos de expresión alta se aíslan y luego se cultivan en ausencia de presión selectiva. En estas condiciones, una gran proporción de los candidatos seleccionados inicialmente se elimina debido a su pérdida de expresión de genes de interés tras la eliminación de la presión selectiva. Así pues, sería ventajoso cultivar los candidatos, después de un período inicial de selección para la transfección estable, en ausencia de presión selectiva, y solo entonces seleccionar el gen de expresión de interés.

2. Cribado de clones por alta expresión.

40 El cribado por clones de alta expresión para una proteína de interés se realiza con frecuencia por métodos que revelan directamente la presencia de elevadas cantidades de la proteína. Típicamente, los métodos inmunológicos, tales como el ELISA o la tinción inmunohistoquímica, se aplican para detectar el producto bien sea intracelularmente o en sobrenadantes de cultivo de células. Estos métodos son tediosos, costosos, lentos y, a menudo, no son válidos para cribados de alto rendimiento (HTS). Además, ha de disponerse de un anticuerpo reactivo a la proteína expresada.

También se han realizado intentos para cuantificar las cantidades de proteína mediante clasificación de células activada por fluorescencia (Fluorescence-Activated Cell Sorting: FACS), pero solo con un éxito limitado, especialmente en el caso de las proteínas secretadas (Borth et al., 2000).

50 Un planteamiento para el cribado de altas tasas de expresión de la proteína de interés sería el uso de un marcador sustituto fácil de medir, expresado a partir del mismo vector que el gen de interés (Chesnut et al., 1996). La idea que subyace en el uso de un marcador sustituto medible es que existe una correlación entre la expresión del gen de interés y el gen marcador sustituto debido al enlace físico de los dos genes en el mismo vector.

En la técnica se dispone de numerosos marcadores fácilmente medibles. Normalmente corresponden a enzimas que actúan sobre un sustrato cromogénico o luminogénico tal como, p. ej. la β -glucuronidasa, la cloranfenicol

acetiltransferasa, la nopalina sintasa, la β -galactosidasa, la fosfatasa alcalina secretada (SEAP) y la DHFR. La proteína verde fluorescente (GFP) puede usarse también como marcador medible en FACS. La actividad de todas estas proteínas puede medirse mediante ensayos estándar que pueden usarse en HTS.

5 El inconveniente de este planteamiento es el uso de otra casete de expresión para el gen marcador sustituto. Esto hace que el vector de expresión sea más bien voluminoso, albergando unidades de expresión que comprenden un promotor, un ADNc y señales de poliadenilación para al menos tres proteínas (es decir, el gen de interés, el marcador de selección y el marcador sustituto). Para las proteínas multicadena, la situación se vuelve aún más compleja. Alternativamente, los vectores plasmídicos individuales que expresan los tres genes, que codifican la proteína de interés, el marcador de selección y el marcador sustituto, respectivamente, podrían ser cotransfectados.

10 Sin embargo, es probable que los vectores estén integrados en diferentes *loci* o exhiban una expresión variable y no correlacionada.

Un planteamiento prometedor para superar las limitaciones anteriores consiste en el uso de un marcador quimérico que combina las propiedades funcionales de un marcador de selección y las de un marcador medible.

15 Tales marcadores bifuncionales han sido descritos, por ejemplo, por Bennett et al. (1998), Imhof y Chatellard (2006) y Dupraz y Kobr (2007). Bennett et al. (1998) describen el marcador GFP-Zeo®, que confiere resistencia al antibiótico zeocina, que corresponde a una proteína de fusión entre la Proteína Fluorescente Verde (GFP) y una proteína que confiere resistencia a la zeocina. Imhof y Chatellard (2006) describen el marcador Lupac, que corresponde a una fusión entre la proteína luciferasa de la luciérnaga y una proteína que confiere resistencia a la puomicina. Dupraz y Kobr (2007) describen el marcador PuroLT, que corresponde a una proteína de fusión entre el péptido sintético descrito por Griffin et al. (1998) y una proteína que confiere resistencia a la puomicina.

20

Miller et al. (2005), en un artículo que muestra que el TMP fluorescente es una alternativa al MTX fluorescente, describe una proteína de fusión entre una proteína que confiere resistencia a la puomicina y una DHFR de origen bacteriano. La DHFR se usa como marcador medible que se puede detectar uniéndose a MTX fluorescente o a TMP fluorescente. Este artículo prevé el uso de la proteína de fusión en el campo del silenciamiento génico de siRNA.

25 Por tanto, todos los marcadores disponibles para la selección de clones que expresan altos niveles de una proteína recombinante corresponden a marcadores bifuncionales, que confieren resistencia a un único compuesto tóxico. Además del marcador bifuncional, los vectores usados para generar clones altamente productores comprenden normalmente un gen amplificable que conduce a un aumento del número de copias cuando se encuentra bajo presión selectiva. El número de copias de un gen de interés situado adyacente al gen amplificable también aumentará, lo que lleva al establecimiento de clones que expresan altos niveles de la proteína de interés (Kaufman et al., 1985; Kaufman et al., 1986; Kim et al. al., 2001; Omasa, 2002). Los genes amplificables usados comúnmente incluyen, p. ej., dihidrofolato reductasa (DHFR), p. ej. el documento US 2005/0019925, glutamina sintetasa (GS), gen de resistencia a múltiples fármacos (MDR), ornitina descarboxilasa (ODC), adenosina desaminasa (ADA) y N-(fosfonoacetil)-L-aspartato (CAD).

30

35 El hallazgo de un nuevo y potente marcador sustituto quimérico, que confiere resistencia a más de un compuesto tóxico y que también permite la amplificación génica, sería extremadamente útil en el campo de la producción industrial de proteínas terapéuticas.

Compendio de la invención.

40 La presente invención se basa en la construcción y caracterización de un nuevo marcador cuatrifuncional, Puro-DHFR. La Puro-DHFR corresponde a una proteína de fusión entre DHFR y una proteína que confiere resistencia a la puomicina, la puomicina N-acetil transferasa (pac). Se ha demostrado que la Puro-DHFR combina las propiedades funcionales tanto de pac como de DHFR. Más específicamente, la Puro-DHFR es un marcador cuatrifuncional que permite (i) seleccionar células por resistencia a la puomicina; (ii) seleccionar células por resistencia a la DHFR; (iii) llevar a cabo la amplificación génica; y (iv) clasificar las células mediante la intensidad de fluorescencia. También se ha demostrado la utilidad de la Puro-DHFR para el aislamiento de clones de alta expresión para una proteína terapéutica.

45

La invención se define en las reivindicaciones.

Breve descripción de las figuras.

50 La Figura 1 representa los plásmidos y vectores reporteros utilizados en los Ejemplos. **A.**: pSV40-DHFR-1474; **B.**: pCMV (IE1) SEAP-IRES-Puro-279; **C.**: pCMV (IE1) SEAP-IRES-Puro / DHFR-325. 1: gen que confiere resistencia a la ampicilina; 2: f1 origen de replicación; 3: señal de poliadenilación sintética; 4: promotor SV40; 5: promotor mCMV (IE1); 6: gen que confiere resistencia a DHFR; 7: señal de poliadenilación de SV40; 8: gen SEAP; 9: poliovirus IRES; 10: gen que confiere resistencia a la puomicina (puomicina N-acetiltransferasa); 11: marcador Puro-DHFR. Todos los vectores contienen además el origen de replicación del plásmido derivado de ColE1.

55 La Figura 2 es un esquema que representa el experimento para seleccionar clones transfectados con los plásmidos y vectores reporteros.

La Figura 3 compara la viabilidad de las células CHO DXB11-F10 transfectadas con el vector pCMV (IE1) SEAP-IRES-Puro/DHFR-325 con la viabilidad de las células CHO DXB11-F10 no transfectadas (control) durante el proceso de selección. El experimento se realizó en ausencia de HT y en ausencia de puromicina (-HT), en presencia de HT y en presencia de puromicina a 10 mg/L (puromicina), y en ausencia de HT y presencia de puromicina (-HT/Puromicina).

La Figura 4 muestra la productividad de la fosfatasa alcalina (SEAP) en cultivos discontinuos de células CHO DXB11-F10 transfectadas con el vector pCMV (IE1) SEAP-IRES-Puro/DHFR-325. El experimento se realizó en ausencia de HT y en ausencia de puromicina (-HT), en presencia de HT y en presencia de puromicina a 10 mg/L (puromicina), en ausencia de HT y en presencia de puromicina (-HT/Puromicina), y en ausencia de HT y presencia de MTX (50 nM) (-HT/MTX).

La Figura 5 muestra la fluorescencia de las células DXB11-F10 CHO transfectadas con el vector pCMV (IE1) SEAP-IRES-Puro/DHFR-325 cultivadas en diferentes medios de selección y teñidas con metotrexato fluorescente (F-MTX).

La Figura 6 compara células CHO-S transfectadas con el vector pCMV (IE1) SEAP-IRES-Puro/DHFR-325 que codifica el marcador puro-DHFR (p325), con células CHO-S transfectadas con el vector pCMV (IE1) SEAP-IRES-Puro-279 que codifica puromicina N-acetil transferasa (p279). **A.**: Productividad de la fosfatasa alcalina (SEAP) **B.**: Número de copias del gen del marcador (puro-DHFR o puromicina N-acetil transferasa) **C.**: Nivel de expresión relativa del ARNm de SEAP.

La Figura 7 muestra la fluorescencia de células CHO-S transfectadas con el vector pCMV (IE1) SEAP-IRES-Puro/DHFR-325 (p325) o con el vector pCMV (IE1) SEAP-IRES-Puro-279 (p279) cultivadas en diferentes medios de selección. El "contraste de fase" corresponde a la técnica óptica utilizada para generar imágenes de muestras biológicas basándose en las diferencias del índice de refracción de la muestra en luz blanca.

La Figura 8 muestra los cambios en el número de copias de Puro-DHFR en clones de células CHO-S transfectadas establemente con pmCMV (IE1)-SEAP-IRES-Puro-DHFR-325 después del cultivo en medio selectivo que contiene puromicina o puromicina más MTX 100 nM.

Breve descripción de las secuencias de la invención.

Las SEQ ID Nos. 1 y 2 corresponden respectivamente a las secuencias de ácidos nucleicos y polipeptídicas de un marcador Puro-DHFR de acuerdo con la invención.

Las SEC ID Nos. 3 y 4 corresponden respectivamente a las secuencias de ácidos nucleicos y polipeptídicas de puromicina N-acetil transferasa (pac) de *Streptomyces alboniger*.

Las SEQ ID Nos. 5 y 6 corresponden respectivamente a las secuencias de ácidos nucleicos y polipeptídicas de DHFR murino.

Las SEQ ID Nos. 7 a 10 corresponden a los cebadores usados cuando se construye el marcador Puro-DHFR de acuerdo con la descripción (Ejemplo 1).

Las SEQ ID Nos. 11 a 16 corresponden a los oligonucleótidos utilizados cuando se detectan los números de copia del gen por QPCR (Ejemplo 3).

Descripción detallada de la descripción.

La presente invención viene de la construcción y caracterización de un nuevo marcador quimérico cuatrifuncional denominado res-DHFR. La descripción describe más específicamente un polipéptido res-DHFR denominado Puro-DHFR, que corresponde a una proteína de fusión entre DHFR y una proteína que confiere resistencia a la puromicina, la puromicina N-acetil transferasa (pac).

Se ha demostrado que la Puro-DHFR es un marcador cuatrifuncional que combina las propiedades funcionales de DHFR y de pac (Ejemplo 2). En consecuencia, el marcador Puro-DHFR se puede utilizar:

- como un marcador seleccionable en combinación con el compuesto tóxico puromicina;
- como un marcador seleccionable en combinación con el compuesto tóxico MTX;
- como un gen amplificable; y
- como un sustituto fácil de medir que se puede detectar tanto por microscopía como por FACS.

Se ha demostrado además la utilidad de la Puro-DHFR para el aislamiento de clones de alta expresión para una proteína de interés. En el Ejemplo 3, se ha construido un vector que comprende Puro-DHFR y un gen de interés, expresado a partir del mismo promotor y separado por un IRES. Se ha demostrado que existe una correlación positiva muy buena entre los niveles de expresión de PURO-DHFR y los niveles de expresión del gen de interés.

En consecuencia, la presente descripción proporciona marcadores potentes que pueden ser usados para proporcionar selectividad en la transfección estable, para inducir la amplificación génica del gen de interés, y que actúa como un marcador sustituto para el cribado de clones candidatos por la alta expresión de un gen de interés. El uso de Puro-DHFR, ligado a una proteína de interés en una configuración bicistrónica, permite mantener la misma probabilidad de seleccionar clones de alta expresión que cuando se mide directamente el nivel de expresión del gen de interés. Además, el uso de Puro-DHFR permite reducir el tiempo, el coste y los recursos ya que (i) se lleva a cabo un análisis estandarizado independiente del producto y simple; y (ii) los expresores altos pueden ser seleccionados usando un FACS.

1. Polipéptidos de la descripción.

10 El polipéptido según la descripción es una proteína quimérica que comprende un fragmento funcional de una dihidrofolato reductasa (DHFR) fusionada a un fragmento que confiere resistencia a un compuesto tóxico o bien una ventaja metabólica, en donde dicho fragmento que confiere resistencia a un compuesto tóxico o una ventaja metabólica no es la DHFR ni un fragmento de la misma. Dicha proteína quimérica se denominará además como "polipéptido de acuerdo con la descripción" o "res-DHFR" dentro de esta memoria descriptiva.

15 El fragmento que confiere resistencia a un compuesto tóxico es una puromicina N-acetiltransferasa (utilizada en combinación con el compuesto tóxico puromicina).

20 En el marco de la presente descripción "un fragmento que confiere una ventaja metabólica" significa que dicho fragmento confiere a una célula la capacidad de crecer en ausencia de un compuesto. Por ejemplo, la proteína glutamina sintetasa (GS) confiere a las células CHO la capacidad de crecer en ausencia de glutamina. Así, el fragmento que confiere una ventaja metabólica puede corresponder p. ej. a la glutamina sintetasa (GS) o un fragmento funcional de la misma.

La expresión "fragmento funcional de DHFR" se refiere a un fragmento de un polipéptido que es un miembro de la familia de la dihidrofolato reductasa (EC 1.5.1.3), y que cataliza la siguiente reacción enzimática:



25 Como se usa en el presente texto, la expresión "actividad de dihidrofolato reductasa" se refiere a la catálisis de la reacción anterior. Esta actividad puede medirse, por ejemplo, determinando la capacidad de conferir a una célula resistencia al compuesto tóxico metotrexato (MTX), como se describe en el Ejemplo 1.2, o determinando la capacidad de aumentar el número de copias del gen en presencia de MTX como se describe en el Ejemplo 1.4. Alternativamente, la actividad de DHFR puede demostrarse por la capacidad de una célula DHFR-negativa transfectada con Puro-DHFR para desarrollarse en un medio carente de timidina y/o hipoxantina.

El fragmento funcional de DHFR es derivado de ratón, y es un fragmento funcional de la secuencia de SEQ ID NO: 6. Este fragmento puede comprender al menos 50, 75, 100, 125, 150, 175 o 187 aminoácidos de SEQ ID NO.: 6. Lo más preferiblemente, dicho fragmento funcional de DHFR comprende los aminoácidos 200 a 385 de SEQ ID NO: 2.

35 El polipéptido Puro-DHFR de la descripción comprende una secuencia de aminoácidos al menos un 95% idéntica a la SEQ ID NO: 2, comprendiendo dicho polipéptido Puro-DHFR un fragmento funcional de DHFR fusionado a un fragmento de una puromicina N-acetil transferasa (pac), en donde dicho polipéptido Puro-DHFR muestra (i) actividad de dihidrofolato reductasa; y (ii) actividad de puromicina N-acetil transferasa. Como se usa también en el presente texto, la expresión "un polipéptido Puro-DHFR" o "Puro-DHFR" se refiere a tal polipéptido.

40 Como se usa en el presente texto, un polipéptido muestra "actividad de puromicina N-acetil transferasa" cuando dicho polipéptido es capaz de conferir a una célula resistencia a la puromicina. La actividad de puromicina N-acetil transferasa puede medirse, por ejemplo, como se describe en el Ejemplo 1.2.

45 El fragmento de una puromicina N-acetil transferasa puede derivarse de una especie de *Streptomyces* tal como, por ejemplo, *Streptomyces alboniger* o *Streptomyces coelicolor*. Preferiblemente, el polipéptido Puro-DHFR comprende un fragmento de un *Streptomyces alboniger* pac. Como se usa en el presente texto, la expresión "*Streptomyces alboniger* pac" se refiere a un polipéptido de la SEQ ID NO: 4 o a una variante alélica, una variante de empalme o una muteína de la misma. Más preferiblemente, el fragmento pac comprende los aminoácidos 200 a 385 de SEQ ID NO: 2. Alternativamente, dicho fragmento de un *Streptomyces alboniger* pac puede comprender al menos 50, 75, 100, 125, 150 o 175 aminoácidos de SEQ ID NO: 4, siempre y cuando conserve la actividad de puromicina N-acetil transferasa.

50 En un polipéptido Puro-DHFR, el fragmento DHFR se puede fusionar al extremo 3' del fragmento pac, o el fragmento pac se puede fusionar al extremo 3' del fragmento DHFR. Preferiblemente, el fragmento DHFR se fusiona con el extremo 3' del fragmento pac.

En una realización entre las más preferidas, el polipéptido Puro-DHFR comprende o consiste en la SEQ ID NO: 2.

Como se usa en el presente documento, el término "muteína" se refiere a un análogo de un polipéptido de origen

natural, en el que uno o más de los restos de aminoácidos de un polipéptido natural están reemplazados por diferentes restos de aminoácidos, o son borrados, o uno o más restos de aminoácidos se añaden a la secuencia de origen natural del polipéptido, sin disminuir considerablemente la actividad de los productos resultantes en comparación con el polipéptido de origen natural. Estas muteínas se preparan mediante síntesis conocida y/o por técnicas de mutagénesis dirigida al sitio, o por cualquier otra técnica conocida adecuada. Las muteínas de *Streptomyces alboniger* pac o de DHFR murina que pueden usarse de acuerdo con la presente descripción, o ácidos nucleicos que codifican las muteínas, que incluyen un conjunto finito de secuencias sustancialmente correspondientes como péptidos o polinucleótidos de sustitución que pueden ser obtenidos de forma rutinaria por un experto en la técnica, sin una experimentación excesiva, basándose en las enseñanzas y directrices presentadas en el presente texto.

Las muteínas de *Streptomyces alboniger* pac o de DHFR murino de acuerdo con la presente invención incluyen proteínas codificadas por un ácido nucleico, tal como ADN o ARN, que se hibrida con ADN o ARN, que codifica pac o DHFR, de acuerdo con la presente descripción, bajo condiciones de rigor moderado o alto. La expresión "condiciones rigurosas" se refiere a la hibridación y las subsiguientes condiciones de lavado, que los expertos en la técnica denominan convencionalmente "rigurosas". Véase Ausubel y col., Current Protocols in Molecular Biology, supra, Interscience, N. Y., §6.3 y 6.4 (1987, 1992), y Sambrook et al. (Sambrook, J. C., Fritsch, E. F., y Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.).

Sin limitación, los ejemplos de condiciones rigurosas incluyen condiciones de lavado 12 – 20 °C por debajo de la T_m calculada del híbrido objeto del estudio en, por ejemplo, 2 x SSC y 0,5% de SDS durante 5 minutos, 2 x SSC y 0.1% de SDS durante 15 minutos; 0,1 x SSC y 0,5% de SDS a 37°C durante 30 - 60 minutos y luego, un 0,1 x de SSC y 0,5% de SDS a 68°C durante 30 - 60 minutos. Los expertos en esta técnica entienden que las condiciones de rigurosidad también dependen de la longitud de las secuencias de ADN, las sondas de oligonucleótidos (tales como 10 - 40 bases) o las sondas de oligonucleótidos mixtas. Si se usan sondas mixtas, es preferible usar cloruro de tetrametil amonio (TMAC) en vez de SSC.

Las muteínas de *Streptomyces alboniger* pac o de DHFR murina incluyen polipéptidos que tienen una secuencia de aminoácidos al menos un 50% idéntica, más preferiblemente al menos un 60% idéntica, y aún más preferiblemente al menos un 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% o 99% idéntica al polipéptido de origen natural.

Por un polipéptido que tiene una secuencia de aminoácidos al menos un 95% "idéntica" a una secuencia de aminoácidos consultada de la presente invención, se entiende que la secuencia de aminoácidos del polipéptido en cuestión es idéntica a la secuencia de búsqueda excepto que la secuencia polipeptídica en cuestión puede incluir hasta cinco alteraciones de aminoácidos por cada 100 aminoácidos de la secuencia de aminoácidos consultada. En otras palabras, para obtener un polipéptido que tiene una secuencia de aminoácidos al menos un 95% idéntica a una secuencia de aminoácidos de consulta, hasta un 5% (5 de 100) de los restos de aminoácidos en la secuencia en cuestión se pueden insertar, eliminar o sustituir con otro aminoácido.

Para las secuencias en las que no hay una correspondencia exacta, se puede determinar un "% de identidad". En general, las dos secuencias a comparar están alineadas para dar una correlación máxima entre las secuencias. Esto puede incluir la inserción de "huecos" en una o ambas secuencias, para mejorar el grado de alineación. Se puede determinar un % de identidad sobre la longitud completa de cada una de las secuencias que se comparan (denominada alineación global), que es particularmente adecuada para secuencias de la misma o muy similar longitud, o sobre longitudes más cortas y definidas (denominadas alineamiento local), que es más adecuado para secuencias de longitud desigual.

Los métodos para comparar la identidad y la homología de dos o más secuencias son bien conocidos en la técnica. Así, por ejemplo, los programas disponibles en el Wisconsin Sequence Analysis Package, versión 9.1 (Devereux et al., 1984), por ejemplo los programas BESTFIT y GAP, pueden usarse para determinar el % de identidad entre dos polinucleótidos y el % de identidad y el % de homología entre dos secuencias polipeptídicas. BESTFIT utiliza el algoritmo de "homología local" de Smith y Waterman (1981) y encuentra la mejor región única de similitud entre dos secuencias. Otros programas para determinar la identidad y/o similitud entre secuencias se conocen también en la técnica, por ejemplo, la familia de programas BLAST (Altschul et al., 1990), accesible a través de la página web del NCBI en el sitio de la web mundial ncbi.nlm.nih.gov y FASTA (Pearson y Lipman, 1988; Pearson, 1990). Es altamente preferido que el % de identidad entre dos secuencias se determine usando el algoritmo KERR (Dufresne et al., 2002), por ejemplo, usando una herramienta bioinformática tal como p. ej. GenePAST.

Preferiblemente, las muteínas de la presente descripción presentan sustancialmente la misma actividad biológica que el polipéptido de origen natural al que corresponde.

2. Ácidos nucleicos, y vectores y células hospedadoras que los comprenden.

Otro aspecto de la presente descripción se refiere a un ácido nucleico que codifica un polipéptido res-DHFR de acuerdo con la descripción.

El ácido nucleico de acuerdo con la descripción codifica un polipéptido Puro-DHFR. Como se usa adicionalmente en esta memoria, el término "ácido nucleico Puro-DHFR" se refiere a tal ácido nucleico.

En una forma de realización preferida, el ácido nucleico Puro-DHFR comprende o consiste en la SEQ ID NO: 1.

Cualquiera de los procedimientos conocidos en la técnica puede usarse para obtener ácidos nucleicos Puro-DHFR de la presente descripción. Los ácidos nucleicos PURO-DHFR se pueden obtener, por ejemplo, como se describe en el Ejemplo 1.

- 5 Otro aspecto de la presente invención se refiere a un vector que comprende un ácido nucleico de acuerdo con la descripción.

Un vector que comprende un ácido nucleico res-DHFR se denomina "vector res-DHFR". Un vector que comprende un ácido nucleico Puro-DHFR se denomina "vector Puro-DHFR" en la presente memoria. Preferiblemente, el vector de la descripción es un vector de expresión.

- 10 El término "vector" se usa en el presente texto para designar un compuesto de ADN o ARN circular o lineal, que es bicatenario o bien monocatenario, y que comprende al menos un polinucleótido de la presente descripción que va a ser transferido en un huésped celular o en un organismo hospedador unicelular o multicelular. Un "vector de expresión" comprende señales apropiadas en los vectores, incluyendo dichas señales varios elementos reguladores tales como potenciadores/promotores de fuentes virales, bacterianas, vegetales, de mamífero y otras fuentes eucarióticas que dirigen la expresión del polinucleótido insertado en las células hospedadoras.

En una forma de realización más preferida, el vector de la descripción comprende además un ácido nucleico que codifica una proteína de interés. Como se muestra en el ejemplo 3, tales vectores son particularmente útiles para el cribado de células por una alta expresión de dicha proteína de interés.

- 20 De acuerdo con la presente descripción, la proteína de interés puede ser cualquier polipéptido cuya producción se desea. La proteína de interés puede encontrar aplicación en el campo farmacéutico, la agroindustria o el mobiliario para laboratorios de investigación. Las proteínas de interés preferidas encuentran aplicación en el campo de los productos farmacéuticos.

- 25 Por ejemplo, la proteína de interés puede ser una proteína secretada de forma natural, una proteína normalmente citoplásmica, una proteína normalmente transmembrana o un anticuerpo humano o humanizado. Cuando la proteína de interés es una proteína normalmente citoplásmica o normalmente transmembrana, la proteína se ha diseñado preferiblemente para hacerse soluble. El polipéptido de interés puede ser de cualquier origen. Los polipéptidos de interés preferidos son de origen humano.

- 30 En formas de realización preferidas, la proteína de interés se elige entre el grupo que consiste en gonadotropina coriónica, hormona estimuladora del folículo, hormona lutropina-coriogonadotrópica, hormona estimulante de la tiroides, hormona de crecimiento humana, interferones (por ejemplo, interferón beta-1a, interferón beta-1b), receptores de interferón (por ejemplo, receptor de interferón gamma), receptores de TNF p55 y p75, interleucinas (por ejemplo, interleucina-2, interleucina-11), proteínas de unión con interleucina (por ejemplo, proteína de unión a interleucina-18), anticuerpos anti-CD11a, eritropoyetina, factor estimulante de colonias de granulocitos, factor estimulante de colonias de granulocitos y macrófagos, hormonas peptídicas pituitarias, gonadotropina menopáusica, factores de crecimiento similares a la insulina (p. ej., somatomedina C), factor de crecimiento de queratinocitos, factor neurotrófico derivado de la línea de células gliales, trombomodulina, factor de crecimiento básico de fibroblastos, insulina, Factor VIII, somatropina, proteína morfogenética ósea-2, factor de crecimiento derivado de plaquetas, hirudina, epoyetina, proteína de fusión LFA-3/IgG1 recombinante, glucocerebrosidasa, anticuerpos monoclonales y muteínas, fragmentos, formas solubles, derivados funcionales, o proteínas de fusión de los mismos.

- 40 Preferiblemente, dicho anticuerpo monoclonal está dirigido contra una proteína seleccionada entre el grupo que consiste en CD3 (por ejemplo OKT3, NI-0401), CD11a (por ejemplo, efalizumab), CD4 (por ejemplo, zanolimumab, TNX-355), CD20 (por ejemplo, ibritumomab tiuxetan, rituximab, tositumomab, ocrelizumab, ofatumumab, IMMU-106, TRU-015, AME-133, GA-101), CD23 (p. ej. lumiliximab), CD22 (por ejemplo, epratuzumab), CD25 (por ejemplo, basiliximab, daclizumab), el receptor del factor de crecimiento epidérmico (EGFR) (p. ej. panitumumab, cetuximab, zalutumumab, MDX-214), CD30 (p. ej. MDX-060), la glicoproteína de la superficie celular CD52 (p. ej. alemtuzumab), CD80 (p. ej., galiximab), el receptor plaquetario GPIIb/IIIa (p. ej. abciximab), TNF alfa (p. ej., infliximab, adalimumab, golimumab), el receptor de interleucina 6 (p. ej. tocilizumab), el antígeno carcinoembrionario (CEA) (p. ej. 99 mTc-besilesomab), integrina alfa-4/beta-1 (VLA4) (p. ej. natalizumab), integrina alfa-5/beta-1 (VLA5) (por ejemplo, volociximab), VEGF (por ejemplo, bevacizumab, ranibizumab), inmunoglobulina E (IgE) (por ejemplo, omalizumab), HER-2/neu (por ejemplo, trastuzumab), el antígeno de la membrana específico de la próstata (PSMA) (p. 111In-capromab pendetide, MDX-070), CD33 (p. ej., gemtuzumab ozogamicina), GM-CSF (p. ej., KB002, MT203), el receptor de GM-CSF (por ejemplo CAM-3001), EpCAM (p. ej. adecatumumab), IFN-gamma-0501, IFN-alfa (p. ej. MEDI-545/MDX-1103), RANKL (p. ej., denosumab), el factor de crecimiento de los hepatocitos (p. ej., AMG 102), IL-15 (p. ej., AMG 714), TRAIL (p. ej., AMG 655), receptor del factor de crecimiento similar a la insulina (por ejemplo, AMG 479, R1507), IL-4 e IL13 (por ejemplo, AMG 317), receptor 3 de BAFF/BLyS (BR3) (por ejemplo, CB1), CTLA-4 (por ejemplo, ipilimumab).

En una forma de realización preferida, el vector de la descripción es un ácido nucleico que codifica una proteína de interés y que comprende al menos dos promotores, uno que dirige la expresión del polipéptido de la descripción, y el

otro que dirige la expresión de la proteína de interés. Tal vector puede comprender además regiones potenciadoras y/o secuencias promotoras de la expresión tales como aislantes, elementos límite, LCRs (p. ej. descritos por Blackwood y Kadonaga (1998) o regiones de unión matriz/soprote (por ejemplo, las descritas por Li y otros (1999).

5 Alternativamente, el vector de la descripción comprende un promotor que impulsa tanto la expresión del gen de interés como la expresión del polipéptido de la descripción. En esta realización, el ORF del polipéptido de la descripción se separa del ORF de la proteína de interés por la presencia de secuencias tales como p. ej. un sitio de entrada ribosomal interno (IRES) o una secuencia 2A (de Felipe et al., 2006). Cuando se usa una secuencia 2A, se prefiere que la Puro-DHFR corresponda al primer ORF (es decir, después del promotor) y que la proteína de interés corresponda al segundo ORF (es decir, después de la secuencia 2A). El IRES puede derivarse, por ejemplo, de un virus o un gen celular. Esta forma de realización está ejemplificada por el vector pCMV (IE1) SEAP-IRES-Puro/DHFR-325 mostrado en la FIG. 1C, en el que la expresión del gen SEAP (8) y del marcador Puro-DHFR (11) está impulsada por el promotor mCMV (IE1) (5), y en el que los ORFs están separados por un IRES (9).

15 El término "promotor" como se usa en el presente documento se refiere a una región de ADN que funciona controlando la transcripción de una o más secuencias de ADN, y que se identifica estructuralmente por la presencia de un sitio de unión para ARN polimerasa dependiente de ADN y de otras secuencias de ADN, que interactúan regulando la función del promotor. Un fragmento que promueve la expresión funcional de un promotor es una secuencia promotora acortada o truncada que conserva la actividad como promotor. La actividad de promotor se puede medir en cualquiera de los ensayos conocidos en la técnica, p. ej. en un ensayo informador que usa DHFR como gen reportero (Wood et al., 1984, SELIGER y McELROY, 1960; de Wet et al., 1985), o disponible comercialmente de Promega®. Una "región potenciadora" se refiere a una región de ADN que funciona aumentando la transcripción de uno o más genes. Más específicamente, el término "potenciador", como se usa en el presente texto, es un elemento regulador del ADN que potencia, aumenta o mejora la expresión de un gen independientemente de su localización y orientación con respecto al gen que se va a expresar, y puede ser expresión potenciadora, aumentadora, o mejoradora de más de un promotor.

25 En una forma de realización preferida, el vector de la descripción comprende al menos un promotor de la región temprana inmediata del CMV murino. El promotor puede ser, por ejemplo, el promotor del gen IE1 de mCMV (el "promotor IE1"), que se conoce, por ejemplo, por el documento WO 87/03905. El promotor también puede ser el promotor del gen IE2 de mCMV (el "promotor IE2"), conociéndose el gen IE2 de mCMV, por ejemplo, por Messerle et al. (1991). El promotor IE2 y las regiones potenciadoras IE2 se describen con detalle en el documento PCT/EP2004/050280. Preferiblemente, el vector de la descripción comprende al menos dos promotores de la región temprana inmediata del CMV murino. Más preferiblemente, los dos promotores son los promotores IE1 e IE2.

En una forma de realización preferida, el vector de la descripción comprende al menos dos promotores de la región temprana inmediata del CMV murino, en donde uno de ellos impulsa la expresión de un polipéptido de la descripción y el otro impulsa la expresión de una proteína de interés.

35 En otra realización preferida, los promotores de la región temprana inmediata del CMV murino impulsan la expresión de genes que codifican una proteína de interés, y el polipéptido Puro-DHFR se expresa a partir de un casete de expresión insertada en la cadena principal del vector. Los promotores IE1 e IE2 pueden impulsar la expresión de dos copias idénticas del gen que codifica la proteína de interés, o de dos subunidades de una proteína multimérica de interés, tales como anticuerpos u hormonas peptídicas.

40 Otro aspecto de la descripción se refiere a una célula transfectada con un ácido nucleico res-DHFR de la descripción y/o con un vector res-DHFR de la descripción. Preferiblemente, dicha célula es una célula Puro-DHFR transfectada con un ácido nucleico Puro-DHFR y/o un vector Puro-DHFR. Muchas células son adecuadas de acuerdo con la presente descripción, tales como las líneas celulares primarias o establecidas de una amplia variedad de eucariotas que incluyen células vegetales y animales. Preferiblemente, dicha célula es una célula eucariótica. Más preferiblemente, dicha célula es una célula de mamífero. Lo más preferiblemente, dicha célula es una célula CHO, una célula humana, una célula de ratón o un hibridoma.

50 Por ejemplo, las células adecuadas incluyen células NIH-3T3, células COS, células MRC-5, células BHK, células VERO, células CHO, células rCHO-tpa, células antígeno de superficie rCHO-Hep B, células HEK 293, células rHEK 293, células antígeno de superficie rC127-Hep B, células CV1, células L de ratón, células HT1080, células LM, células Y1, células de hibridoma de ratón NS0 and SP2/0 y similares, células RPMI-8226, células Vero, células WI-38, células MRC-5, células de fibroblastos humanos normales, células de estroma humano, células de hepatocitos humanos, células de osteosarcoma humano, células de Namalwa, células neuronales humanas, células de retinoblastos humanos, células PER.C6 y otras células de mamífero inmortalizadas y/o transformadas.

3. Métodos para usar los polipéptidos y ácidos nucleicos anteriores.

55 Otro aspecto se refiere al uso de una célula que comprende un ácido nucleico res-DHFR de acuerdo con la descripción, para producir una proteína de interés. Preferiblemente, dicha célula comprende un vector Puro-DHFR.

Como se discute en el Ejemplo 3, el uso de un polipéptido Puro-DHFR como marcador de selección y sustituto proporciona numerosas ventajas para el cribado de células por la alta expresión de una proteína de interés.

Específicamente, dado que la expresión del polipéptido Puro-DHFR está muy correlacionada con la expresión de la proteína de interés, es ventajoso realizar una selección primaria para la expresión elevada de PURO-DHFR, p. ej. por FACS. La expresión de la proteína de interés se ensaya en un cribado secundario, que se realiza solamente con los mejores productores aislados más allá de la selección primaria para la expresión alta de PURO-DHFR.

5 En consecuencia, otro aspecto de la descripción se refiere al uso de un polipéptido de acuerdo con la descripción, de un ácido nucleico de acuerdo con la descripción o de un vector de acuerdo con la descripción, para el cribado de células por su expresión o por la elevada expresión de una proteína de interés. Primero se criban las células por la alta expresión del polipéptido de acuerdo con la descripción (por ejemplo, Puro-DHFR), y la expresión del polipéptido de acuerdo con la descripción se correlaciona después con la de una proteína de interés por deducción. Esto
10 permite eliminar rápidamente del 80 al 95% de las células ensayadas basándose en los niveles bajos de expresión del polipéptido de acuerdo con la descripción, y retener el 5 - 20% restante para el análisis de la expresión del gen de interés en una segunda etapa.

En el contexto de los usos y métodos de la presente descripción, el término "expresión alta" se refiere a un nivel de expresión en una célula que es cribada, que es más alta que en otras células que se criban. "Expresión alta" de una
15 proteína es un valor relativo. Por ejemplo, los niveles de expresión final de proteínas recombinantes que se producen comercialmente dependen de la proteína, las cantidades anuales requeridas y la dosis terapéutica. Durante un cribado, el nivel de expresión de una proteína de interés es menor que el nivel de expresión final.

Otro aspecto se refiere a un método para cribar células para la expresión o alta expresión de una proteína de interés, comprendiendo dicho método las etapas de:

- 20 a) transfectar células mediante un vector de expresión que codifica Puro-DHFR;
b) seleccionar las células que son resistentes a la puromicina; y
c) ensayar la fluorescencia de las células seleccionadas en la etapa (b) con un análogo de folato marcado fluorescentemente que se une covalentemente a DHFR.

Preferiblemente, este método de cribado de células por la expresión o alta expresión de una proteína de interés comprende además la etapa de amplificar dicha proteína recombinante de interés antes de llevar a cabo la etapa (c). Dicha etapa de amplificación se realiza preferiblemente cultivando las células en presencia de metotrexato (MTX). La concentración de metotrexato variará según el tipo de célula. Típicamente, las células CHO crecerán en un medio que comprende aproximadamente 50, 75, 100, 125, 150, 200, 300, 500, 1000, 2000, 3000, 4000, 5000 o 6000 nM de
25 MTX para la amplificación génica.

30 La fluorescencia de las células se detecta utilizando cualquier análogo de folato marcado fluorescentemente que se une covalentemente a DHFR. Dichos compuestos fluorescentes incluyen, p. ej., metotrexato fluorescente (f-MTX) o trimetoprim fluorescente (f-TMP).

En la etapa (c), la fluorescencia puede medirse usando cualquier aparato bien conocido en la técnica, tal como, por ejemplo, un microscopio de fluorescencia o un clasificador de células activado por fluorescencia (FACS) o similar.
35 Usar un FACS es particularmente ventajoso cuando se realizan cribados de alta producción.

En una realización preferida, el 20% de las células que muestran la fluorescencia más alta en la etapa (c) comprenden la célula que exhibe la expresión más alta de dicha proteína de interés. Preferiblemente, el 10% de las células que muestran la fluorescencia más alta en la etapa (c) comprenden la célula que muestra la expresión más alta de dicha proteína de interés. Lo más preferiblemente, el 1% o el 5% de las células que muestran la
40 fluorescencia más alta en la etapa (c) comprenden la célula que muestra la expresión más alta de dicha proteína de interés.

Se puede cribar cualquier cantidad de células por ese método. Preferiblemente, la fluorescencia de al menos 1, 20, 50, 100, 500, 1.000, 5.000, 10.000, 50.000, 100.000, 500.000, 1.000.000 o 10.000.000 células se ensaya en la etapa (c). Más preferiblemente, se criba una población de células suficiente para obtener al menos de 1.000 a 10.000.000
45 de transfectantes independientes resistentes a la puromicina. De estos, al menos de 10 a 1.000.000 de clones candidatos que son resistentes a la puromicina pueden analizarse adicionalmente para determinar la fluorescencia.

Las células obtenidas al final del método de cribado anterior pueden clasificarse entre ellas con respecto a la expresión de res-DHFR. Las células que muestran la fluorescencia más alta se pueden seleccionar al final de cualquiera de los métodos de cribado anteriores. Por ejemplo, células individuales que muestran actividad de DHFR correspondiente al 5 - 20% superior de expresores Puro-DHFR son seleccionadas para análisis adicional de la
50 expresión del gen de interés en una etapa subsiguiente.

En una forma de realización preferida, el método de exploración anterior comprende además la etapa de (d) seleccionar de aproximadamente 1% a aproximadamente 20% de las células ensayadas en la etapa (c), en donde las células seleccionadas son aquellas que exhiben la fluorescencia más alta en la etapa (c). De aproximadamente
55 el 5% a aproximadamente el 20% de las células ensayadas en la etapa (c) se pueden seleccionar basándose en la

actividad de Puro-DHFR más alta. Alternativamente, de aproximadamente 1%, 1,5%, 2%, 3%, 4%, 5% a aproximadamente 30%, 40%, 50%, 60%, 70% u 80% de las células ensayadas en la etapa (c) pueden ser seleccionadas basándose en la actividad más alta de Puro-DHFR.

5 Las etapas (b) (es decir, seleccionar células resistentes), (c) (es decir, analizar la fluorescencia) y (d) (es decir, seleccionar la mayor parte de las células fluorescentes) pueden repetirse iterativamente en la población seleccionada al final de la etapa (d). Por ejemplo, se pueden llevar a cabo al menos 2, 3, 5 o 10 iteraciones. Esto puede hacerse con o sin cambio de las condiciones entre las etapas de selección. Las condiciones cambiantes pueden incluir, p. ej., aumentar la concentración de MTX para inducir amplificación génica o condiciones de cultivo variables tales como los componentes de los medios o parámetros físico-químicos.

10 Tras la selección de las células que presentan la fluorescencia más alta, puede seguirse ensayando el nivel de expresión de la proteína de interés en dichas células seleccionadas.

15 Después, se pueden seleccionar de aproximadamente el 1% a aproximadamente el 20% de las células que muestran la expresión más alta de dicha proteína de interés. Por ejemplo, puede seleccionarse de aproximadamente el 1%, 1,5%, 2%, 3%, 4%, 5% a aproximadamente el 15%, 18% o 20% de las células que muestran la expresión más alta de dicha proteína de interés. Preferiblemente, se selecciona la célula que muestra la expresión más alta de dicha proteína de interés. Esta selección basada en la expresión de la proteína de interés se realiza preferiblemente después de la última iteración de la etapa (d) (es decir, la última selección basada en fluorescencia).

Otro aspecto de la descripción se refiere a un método para obtener una línea celular que expresa una proteína de interés, comprendiendo dicho método las etapas de:

- 20 a) cribar las células de acuerdo con el método anterior; y
b) establecer una línea celular a partir de dichas células.

25 Como se usa en el presente texto, una "línea celular" se refiere a un tipo específico de célula que puede crecer en un laboratorio. Una línea celular se puede cultivar normalmente en un cultivo de células establecido permanentemente, y proliferará indefinidamente dado un medio fresco y espacio apropiados. Los métodos para establecer líneas celulares a partir de células aisladas son bien conocidos por los expertos en la técnica.

Otro aspecto se refiere a un método para producir una proteína de interés, comprendiendo dicho método las etapas de:

- 30 a) cultivar una línea celular obtenida como se ha descrito anteriormente bajo condiciones que permiten la expresión de dicha proteína de interés; y
b) recoger dicha proteína de interés.

Las condiciones que permiten la expresión de la proteína de interés mediante métodos estándar pueden ser establecidas fácilmente por los expertos en la técnica. Por ejemplo, se pueden usar las condiciones descritas en el Ejemplo 3.3.1.

35 En una realización preferida, el método anterior de producción de una proteína de interés comprende además la etapa de purificar dicha proteína de interés. La purificación puede realizarse mediante cualquier técnica bien conocida por los expertos en la técnica. En el caso de una proteína de interés para su uso en el campo de la industria farmacéutica, la proteína de interés se formula preferiblemente en una composición farmacéutica.

Ejemplos.

Ejemplo 1: Protocolos

40 1.1. Construcción del ácido nucleico Puro-DHFR.

Los constructos descritos a continuación se representan en la Figura 1. Todos los constructos están basados en la cadena principal del plásmido pGL3-básico (Promega).

45 La proteína de fusión entre el gen de resistencia a la puomicina y la DHFR murina de tipo silvestre se obtuvo mediante PCR recombinante. Parte del poliovirus IRES y el ORF del gen de resistencia a la puomicina completo (omitiendo el codón de parada) que están presentes en el vector pmCMV (IE1)-SEAP-IRES-PuroR-p279 fueron amplificados mediante PCR utilizando cebadores de SEQ ID Nos. 7 y 8 y una ADN polimerasa de alta fidelidad (HS-KOD; Novagen). El ORF de DHFR murino de tipo silvestre y parte de la señal de poliadenilación tardía de SV40 presente en el plásmido pSV40-DHFR (p1474) se amplificaron usando cebadores de las SEQ ID Nos. 9 y 10.

50 Para generar el marcador de fusión Puro-DHFR, los productos de la PCR resultantes se mezclaron y se reamplificaron usando cebadores de SEQ ID Nos. 7 y 10, y se clonaron en un vector en el que el promotor IE1 murino impulsa la expresión del gen de la fosfatasa alcalina placentaria humana (SEAP) y el marcador de selección

Puro-DHFR se expresa como el segundo cistrón colocado aguas abajo del poliovirus IRES. La integridad de todos los elementos amplificados por PCR se verificó por secuenciación. Este vector se denomina también pmCMV (IE1)-SEAP-IRES-Puro/DHFR-p325 o p325.

1.2. Transfección y cultivo de células.

5 El protocolo para seleccionar transfectantes estables se esquematiza en la Figura 2. Se derivaron células CHO-S de los ovarios de hámster chino y se adaptaron a un cultivo en suspensión sin suero (Invitrogen/Gibco, La Jolla, CA). CHO DX11-F10 es una línea celular derivada de la línea celular CHO DUKXB11 deficitaria en DHFR (Urlaub et al., 1980) que fue adaptada para crecer en suspensión en medio libre de suero. Ambas células se cultivan de forma rutinaria en ProCho5 (Lonza Biologics). El medio para el cultivo de células DXB11-F10 se suplementó con
10 hipoxantina (100 μ M) y timidina (16 μ M) (suplemento HT; Invitrogen/Gibco), a menos que se indique otra cosa.

Las células CHO-S y DXB11-F10 fueron transfectadas utilizando polietilenimina (PEI lineal 25 kd). Las células se pusieron en placas de 6 pocillos en 2,5 mL de RPMI-1640 (Invitrogen/Gibco) más 0,05% de Pluronic F68 (Sigma) a una concentración de 5×10^5 células por mL. Se mezclaron 5 μ g de ADN de plásmido linealizado en 250 μ L de NaCl 150 mM con una solución de 15 μ L de PEI25 lineal 1 mM diluida en 250 μ L de NaCl 150 mM. Se dejaron formar los complejos de PEI:ADN durante 5 minutos a temperatura ambiente y luego se añadieron a las células. Al cabo de 3
15 horas, el medio de transfección se reemplaza por medio de cultivo libre de suero.

Al cabo de 48 horas después de la transfección se aplicó la selección y el medio se cambió 2 veces por semana hasta que las células se recuperaron y la viabilidad celular fue superior al 90%.

20 Para la selección se usó puromicina a 10 μ g/mL y el metotrexato análogo del folato (Calbiochem) se usó a una concentración de 50 a 100 nM.

Para los estudios de amplificación, los clones se obtuvieron primero por dilución limitada a 0,3 células por pocillo en placas de 384 pocillos bajo selección con puromicina a 10 μ g/mL. Los clones se cultivaron luego durante 4 semanas bajo selección con puromicina (10 μ g/ml) o puromicina más metotrexato (100 nM). Después se extrajo el ADN genómico y se determinó el número de copias del gen reportero mediante QPCR.

25 1.3. Determinación de la expresión del gen reportero mediante el ensayo SEAP.

Se sembraron grupos de células transfectadas establemente a $2,5 \times 10^5$ células/ml en matraces agitados de 125 ml y se cultivaron durante un tiempo de hasta 7 días en un cultivo discontinuo. El medio de cultivo celular se recolectó en diversos momentos y para evitar la variación diaria en las mediciones de SEAP las muestras se mantuvieron a la temperatura de -20°C hasta el análisis. La actividad de SEAP relativa se determinó en un ensayo enzimático
30 cinético. Se añadieron 10 μ L de diluciones seriadas de medio en solución salina tamponada con hepes (HBSS) a una placa de 96 pocillos, luego se añadieron 100 μ L de una Solución de Substrato de Fosfatasa (Pierce) a cada pocillo, y se tomaron lecturas de la DO a 405 nm en intervalos de tiempo regulares. Solo se consideró para el análisis la ventana lineal de la gráfica de DO frente al tiempo.

1.4. Determinación del número de copias genéticas por QPCR.

35 El ADN genómico se aisló usando el kit GenElute Mammalian Genomic DNA Miniprep (Sigma) de acuerdo con las instrucciones del fabricante, y se cuantificó espectrofotométricamente. Para la determinación del número de copias del gen, se analizaron 10 ng de ADN genómico mediante PCR cuantitativa con el instrumento de PCR en tiempo real 7500 (Applied Biosystems) usando condiciones de ciclado estándar en un ensayo múltiple. Se usó una sonda TaqMan específica de puromicina para detectar el constructo reportero y la segunda sonda TaqMan, que detecta el
40 gen de la gliceraldehído fosfato deshidrogenasa (GAPDH) de hámster como control endógeno. Se generó una curva patrón usando ADN genómico de líneas celulares en las que se había determinado el número de copias del gen de puromicina mediante transferencia Southern.

Los oligonucleótidos tenían las secuencias de SEQ ID Nos. 11, 12 y de FAM-SEQ ID NO: 13-BHQ1 para la detección del gen de puromicina acetiltransferasa, y de SEQ ID Nos. 14, 15 y de YY-SEQ ID NO: 16-BHQ1 para la detección del gen GAPDH. FAM y YY son abreviaturas de fluoroforos 5-carboxifluoresceína y Yakima Yellow (Epoch Biosciences), respectivamente. BHQ1™ (Biosearch Technologies, Inc) es un desactivador ligado al 3' de las sondas TaqMan.

Los reactivos Core PCR eran de Applied Biosystems y las placas de detección de 96 pocillos se obtuvieron de AxonLab.

50 1.5. Determinación de la expresión relativa del gen reportero por transcriptasa inversa-QPCR.

Se aisló el ARN total de $\sim 5 \times 10^6$ células usando el kit NucleoSpin RNA II (Macherey-Nagel) - que incluye una etapa de tratamiento con ADNasa- y se determinó la concentración de ARN espectrofotométricamente a 260 nm.

La cuantificación relativa de la expresión del reportero se realizó por transcriptasa inversa en una etapa - QPCR (reactivo de mezcla madre de RT-PCR en una etapa; Applied Biosystems) en 25 ng de ARN total, usando los

cebadores Puro y GAPDH y las sondas TaqMan descritas anteriormente. El GAPDH sirvió como control endógeno. Las cantidades del ARNm reporter se calcularon mediante el método $\Delta\Delta Ct$ y se expresaron en relación con el grupo p279 (seleccionado solamente con puromicina).

1.6. Etiquetado con fluoresceína-metotrexato.

- 5 Se incubaron $2,5 - 5 \times 10^5$ células durante la noche a 29°C en 0,5 mL de medio de cultivo que contenía 10 mM de metotrexato marcado con fluoresceína (F-MTX, Molecular Probes/Invitrogen). Las células marcadas se lavaron en un medio de cultivo libre de suero y se registraron las imágenes usando un microscopio de fluorescencia (microscopio Olympus CKX41 equipado con una cámara digital DP50) usando un juego de filtros FITC.

Ejemplo 2: Puro-DHFR es un marcador cuatrifuncional.

- 10 2.1. La Puro-DHFR confiere resistencia a la puromicina a las células transfectadas (es decir, la Puro-DHFR tiene actividad de puromicina acetiltransferasa).

Se transfectaron células CHO DXB11-F10 con el vector p325 que codifica Puro-DHFR como se describe en el Ejemplo 1.2. Como se muestra en la Figura 3, el marcador de selección Puro-DHFR confiere resistencia a la puromicina a las células transfectadas.

- 15 2.2. La Puro-DHFR permite el crecimiento en ausencia de HT (es decir, la Puro-DHFR tiene actividad de DHFR).

Las células deficitarias en DHFR son sensibles a MTX y requieren la presencia de HT (hipoxantina y timidina) en el medio de cultivo para el crecimiento.

- 20 Se transfectaron células CHO DXB11-F10 deficitarias en DHFR con el vector p325 que codifica Puro-DHFR como se describe en el Ejemplo 1.2. Como se muestra en la Figura 3, el marcador de selección Puro-DHFR permite el crecimiento de células DXB11-F10 transfectadas con p325 en ausencia de HT.

2.3. La Puro-DHFR induce la amplificación génica.

- 25 Células CHO-S, que expresan endógenamente DHFR, fueron transfectadas con el vector p325 que codifica Puro-DHFR como se describe en el Ejemplo 1.2. Las células se seleccionaron en presencia de puromicina (10 $\mu\text{g}/\text{ml}$). Los clones se obtuvieron por dilución limitada de la población resistente. Resulta interesante que los experimentos mostrados en el presente texto demuestran que la selección y amplificación de Puro-DHFR es factible en células CHO-S a pesar de su expresión de fondo de DHFR endógeno.

Veinte clones seleccionados al azar se cultivaron en presencia de puromicina (10 $\mu\text{g}/\text{ml}$) o en presencia de puromicina y MTX 100 nM durante 4 semanas, con el fin de probar la amplificación génica.

- 30 El número de copias del gen se determinó entonces para grupos de células transformadas como se describe en el Ejemplo 1.4. Como se muestra en la Figura 8, la amplificación del número de copias del gen depende de la selección con MTX. En 3 de 20 clones re-seleccionados con puromicina más MTX a 100 nM (en un círculo) el número de copias del reporter ha aumentado en comparación con la selección con puromicina solamente, lo que demuestra el potencial para la amplificación génica.

2.4. La Puro-DHFR puede ser detectado a través de su fluorescencia.

- 35 La Figura 5 muestra el marcado con metotrexato fluorescente de células CHO DXB11-F10 transfectadas con el vector p325 que codifica Puro-DHFR. La fluorescencia se detectó como se describe en el Ejemplo 1.6. Las células DXB11-F10 no transfectadas son mucho menos fluorescentes que las transfectadas. Además, una presión de selección más alta en presencia de MTX (50 nM) conduce a un aumento de la expresión del marcador de selección puro-DHFR y a un marcaje celular más intenso.

- 40 La Figura 7 muestra el marcaje con metotrexato fluorescente de células CHO-S transfectadas con el vector p325 que codifica Puro-DHFR. Los niveles de base de fluorescencia en grupos de CHO-S p279 (que expresan solamente el gen de DHFR endógeno) son significativamente más bajos que la fluorescencia en grupos de células CHO-S p325 seleccionadas en niveles elevados de MTX.

- 45 Así pues, la Puro-DHFR puede detectarse tanto en células DHFR+ (CHO-S) como en células DHFR- (CHO-DXB11-F10).

Ejemplo 3: La Puro-DHFR es un marcador sustituto útil para el cribado de células que expresan niveles elevados de una proteína de interés.

3.1. La Puro-DHFR permite aislar clones que expresan niveles elevados de una proteína de interés.

- 50 Las células DXB11-F10 fueron transfectadas con el vector p325 que codifica Puro-DHFR. Este vector comprende adicionalmente SEAP como gen reportero (proteína de interés). Los grupos estables se seleccionaron en presencia

o en ausencia de MTX, y la expresión de SEAP se midió como se describe en el Ejemplo 1.3. Como se muestra en la Figura 4, los grupos seleccionados en presencia de MTX mostraron una expresión de SEAP aproximadamente 2 veces mayor que los grupos seleccionados con puromicina o en ausencia de HT solamente.

5 Las células CHO-S fueron transfectadas con el vector p325 que codifica Puro-DHFR o con el vector p279 que codifica puromicina N-acetiltransferasa y SEAP. Los grupos estables se seleccionaron en presencia de puromicina y de 0, 50 o 100 nM de MTX. La expresión de SEAP se midió el día 7, tanto como se describe en el Ejemplo 1.3a como en el Ejemplo 1.5. La Figura 6 muestra que el aumento de la presión de selección del grupo estable de CHO-S p325 con puromicina más MTX (100 nM) conduce a una expresión significativamente mayor de proteína SEAP, tanto al nivel de proteína (A) como al nivel de ARNm (B).

10 Así pues el marcador puro-DHFR permite aislar clones que expresan niveles más altos de la proteína SEAP que los marcadores de la técnica anterior tales como puromicina N-acetiltransferasa.

3.2. La Puro-DHFR se puede usar como marcador sustituto en cribados de alta producción.

15 La Puro-DHFR se usa como marcador selectivo y sustituto para establecer y cribar clones candidatos con un vector que expresa tanto Puro-DHFR como la proteína de interés. Después de la transfección, selección y amplificación, se hace un cribado primario por FACS por fluorescencia (es decir, expresión elevada de PURO-DHFR) con una alta probabilidad de seleccionar clones que presenten también una alta expresión del gen de interés. Luego se realiza un segundo cribado para la expresión de la proteína de interés, posiblemente directamente mediante ELISA.

20 El uso de Puro-DHFR en cribados de alta producción (HTS) permite entonces mantener la misma posibilidad de seleccionar clones de alta expresión y permite reducir el tiempo y los recursos. Además, es importante indicar que la fusión de dos enzimas individuales con actividades y orígenes tan diferentes conserva sorprendentemente su función en Puro-DHFR como se describe en el presente texto. Verdaderamente, se puede utilizar Puro-DHFR para proporcionar selectividad en la transfección estable, actuar como un gen amplificable y actuar como un marcador sustituto para seleccionar clones candidatos para la expresión elevada de un gen de interés.

Referencias

25 Altschul,S.F., Gish,W., Miller,W., Myers,E.W., and Lipman,D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.

Bennett,R.P., Cox,C.A., and Hoeffler,J.P. (1998). Fusion of green fluorescent protein with the Zeocin-resistance marker allows visual screening and drug selection of transfected eukaryotic cells. *Biotechniques* 24:478-482.

30 Blackwood,E.M. and Kadonaga, J.T. (1998). Going the distance: a current view of enhancer action. *Science* 281:61-63.

Borth,N., Zeyda,M., Kunert,R., and Katinger,H. (2000). Efficient selection of high-producing subclones during gene amplification of recombinant Chinese hamster ovary cells by flow cytometry and cell sorting. *Biotechnol. Bioeng.* 71:266-273.

35 Chesnut,J.D., Baytan,A.R., Russell,M., Chang,M.P., Bernard,A., Maxwell,I.H., and Hoeffler,J.P. (1996). Selective isolation of transiently transfected cells from a mammalian cell population with vectors expressing a membrane anchored single-chain antibody. *J. Immunol. Methods* 193:17-27.

de Felipe,P., Luke,G.A., Hughes,L.E., Gani,D., Halpin,C., and Ryan,M.D. (2006). E unum pluribus: multiple proteins from a self-processing polyprotein. *Trends Biotechnol.* 24, 68-75.

40 de Wet,J.R., Wood,K.V., Helinski,D.R., and DeLuca,M. (1985). Cloning of firefly DHFR cDNA and the expression of active DHFR in *Escherichia coli*. *Proc Natl Acad Sci U S A* 82:7870-7873.

Devereux,J., Haeberli,P., and Smithies,O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387-395.

Dufresne,G., Takacs,L., Heus,H.C., Codani,J.J., and Duval,M. (2002). Patent searches for genetic sequences: how to retrieve relevant records from patented sequence databases. *Nat. Biotechnol.* 20, 1269-1271.

45 Dupraz,P. and Kobr,M. (2007). System for screening cells for high expression of a protein of interest. WO 2007/023184.

Grantham,R. (1974). Amino acid difference formula to help explain protein evolution. *Science* 185:862-864.

Imhof,M., and Chatellard,P. (2006). The LUPAC bifunctional marker and its use in protein production. WO 2006/058900.

50 Kaufman,R.J., Wasley,L.C., Spiliotes,A.J., Gossels,S.D., Latt,S.A., Larsen,G.R., and Kay,R.M. (1985).

Coamplification and coexpression of human tissue-type plasminogen activator and murine dihydrofolate reductase sequences in Chinese hamster ovary cells. *Mol. Cell Biol.* 5:1750-1759.

Kaufman,R.J., Murtha,P., Ingolia,D.E., Yeung,C.Y. and Kellems,R.E. (1986). Selection and amplification of heterologous genes encoding adenosine deaminase in mammalian cells. *Proc.Natl.Acad.Sci.U.S.A.* 83:3136-3140.

- 5 Kim,N.S., Byun,T.H. and Lee,G.M. (2001). Key determinants in the occurrence of clonal variation in humanized antibody expression of cho cells during dihydrofolate reductase mediated gene amplification. *Biotechnol.Prog.* 17:69-75.

Li,Q., Harju,S., and Peterson,K.R. (1999). Locus control regions: coming of age at a decade plus. *Trends Genet.* 15:403-408.

- 10 Messerle,M., Keil,G.M., and Koszinowski,U.H. (1991). Structure and expression of murine cytomegalovirus immediate-early gene 2. *J. Virol.* 65, 1638-1643.

Miller,L.W., Cai,Y., Sheetz,M.P., Cornish,V.W. (2005). In vivo protein labeling with trimethoprim conjugates: a flexible chemical tag. *Nat Methods.* 2:255-257.

Omasa,T.(2002). Gene amplification and its application in cell and tissue engineering. *J.Biosci.Bioeng.* 94:600-605.

- 15 Pearson,W.R. and Lipman,D.J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U. S. A* 85, 2444-2448.

Pearson,W.R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* 183, 63-98.

- 20 Seliger,H.H. and McElroy,W.D. (1960). Spectral emission and quantum yield of firefly bioluminescence. *Arch. Biochem. Biophys.* 88, 136-141.

Smith,T.F. and Waterman,M.S. (1981). Identification of common molecular subsequences. *J. Mol. Biol.* 147, 195-197.

Urlaub,G. and Chasin,L.A. (1980). Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proc.Natl.Acad.Sci.U.S.A.* 77:4216-4220.

- 25 Wood,K.V., de Wet,J.R., Dewji,N., and DeLuca,M. (1984). Synthesis of active firefly DHFR by in vitro translation of RNA obtained from adult lanterns. *Biochem Biophys. Res. Commun.* 124, 592-596.

Listado de secuencias

<110> LABORATOIRES SERONO SA

<120> THE PURO-DHFR QUADRIFUNCTIONAL MARKER AND ITS USE IN PROTEIN PRODUCTION

<130> 1161 WO/PCT

- 30 <150> EP 07109829.7
<151> 2007-06-07

<150> US 60/934,078
<151> 2007-06-11

<160> 16

- 35 <170> PatentIn version 3.1

<210> 1
<211> 1158
<212> ADN
<213> Artificial

- 40 <220>
<223> Fusión entre DHFR y puromicina N-acetil transferasa

<220>
<221> CDS
<222> (1)..(1155)

- 45 <223>

<220>

ES 2 660 153 T3

<221> caract_misc
 <222> (1)..(597)
 <223> Fragmento derivado de puromicina N-acetil transferasa

<220>

5 <221> caract_misc
 <222> (598)..(1155)
 <223> Fragmento derivado de DHFR

<400> 1

atg acc gag tac aag ccc acg gtg cgc ctc gcc acc cgc gac gac gtc	48
Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val	
1 5 10 15	
ccc cgg gcc gta cgc acc ctc gcc gcc cgc ttc gcc gac tac ccc gcc	96
Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala Asp Tyr Pro Ala	
20 25 30	
acg cgc cac acc gtc gac ccg gac cgc cac atc gag cgg gtc acc gag	144
Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu	
35 40 45	
ctg caa gaa ctc ttc ctc acg cgc gtc ggg ctc gac atc ggc aag gtg	192
Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val	
50 55 60	
tgg gtc gcg gac gac ggc gcc gcg gtg cgc gtc tgg acc acg ccg gag	240
Trp Val Ala Asp Asp Gly Ala Ala Val Val Val Trp Thr Thr Pro Glu	
65 70 75 80	
agc gtc gaa gcg ggg gcg gtg ttc gcc gag atc ggc ccg cgc atg gcc	288
Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala	
85 90 95	

10

ES 2 660 153 T3

gag ttg agc ggt tcc cgg ctg gcc gcg cag caa cag atg gaa ggc ctc 336
 Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Met Glu Gly Leu 110

ctg gcg ccg cac cgg ccc aag gag ccc gcg tgg ttc ctg gcc acc gtc 384
 Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val 125

ggc gtc tcg ccc gac cac cag gcc aag ggt ctg ggc agc gcc gtc gtg 432
 Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val 140

ctc ccc gga gtg gag gcg gcc gag cgc gcc ggg gtg ccc gcc ttc ctg 480
 Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu 160

gag acc tcc gcg ccc cgc aac ctc ccc ttc tac gag cgg ctc ggc ttc 528
 Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe 175

acc gtc acc gcc gac gtc gag gtg ccc gaa gga ccg cgc acc tgg tgc 576
 Thr Val Thr Ala Asp Val Glu Val Pro 185 Glu Gly Pro Arg Thr Trp Cys 190

atg acc cgc aag ccc ggt gcc gtt cga cca ttg aac tgc atc gtc gcc 624
 Met Thr Arg Lys Pro Gly Ala Val Arg Pro Leu Asn Cys Ile Val Ala 205

gtg tcc caa aat atg ggg att gcc aag aac gga gac cta ccc tgg cct 672
 Val Ser Gln Asn Met Gly Ile Gly Lys Asn Gly Asp Leu Pro Trp Pro 220

ccg ctc agg aac gag ttc aag tac ttc caa aga atg acc aca acc tct 720
 Pro Leu Arg Asn Glu Phe Lys Tyr Phe Gln Arg Met Thr Thr Thr Ser 235 240

tca gtg gaa ggt aaa cag aat ctg gtg att atg ggt agg aaa acc tgg 768
 Ser Val Glu Gly Lys Gln Asn Leu Val Ile Met Gly Arg Lys Thr Trp 255

ttc tcc att cct gag aag aat cga cct tta aag gac aga att aat ata 816
 Phe Ser Ile Pro Glu Lys Asn Arg Pro Leu Lys Asp Arg Ile Asn Ile 260 270

gtt ctc agt aga gaa ctc aaa gaa cca cca cga gga gct cat ttt ctt 864
 Val Leu Ser Arg Glu Leu Lys Glu Pro Pro Arg Gly Ala His Phe Leu 285

gcc aaa agt ttg gat gat gcc tta aga ctt att gaa caa ccg gaa ttg 912
 Ala Lys Ser Leu Asp Asp Ala Leu Arg Leu Ile Glu Gln Pro Glu Leu 290 300

gca agt aaa gta gac atg gtt tgg ata gtc gga ggc agt tct gtt tac 960
 Ala Ser Lys Val Asp Met Val Trp Ile Val Gly Gly Ser Ser Val Tyr 305 310 315 320

cag gaa gcc atg aat caa cca gcc cac ctc aga ctc ttt gtg aca agg 1008
 Gln Glu Ala Met Asn Gln Pro Gly His Leu Arg Leu Phe Val Thr Arg 325 330 335

atc atg cag gaa ttt gaa agt gac acg ttt ttc cca gaa att gat ttg 1056
 Ile Met Gln Glu Phe Glu Ser Asp Thr Phe Phe Pro Glu Ile Asp Leu 340 345 350

ggg aaa tat aaa ctt ctc cca gaa tac cca ggc gtc ctc tct gag gtc 1104
 Gly Lys Tyr Lys Leu Leu Pro Glu Tyr Pro Gly Val Leu Ser Glu Val 355 360 365

cag gag gaa aaa ggc atc aag tat aag ttt gaa gtc tac gag aag aaa 1152

Gln Glu Glu Lys Gly Ile Lys Tyr Lys Phe Glu Val Tyr Glu Lys Lys
 370 375 380

gac taa 1158
 Asp
 385

5 <210> 2
 <211> 385
 <212> PRT
 <213> Artificial

ES 2 660 153 T3

<220>

<223> Fusión entre DHFR y puromicina N-acetil transferasa

<400> 2

Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val
 1 5 10 15

Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala Asp Tyr Pro Ala
 20 25 30

Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu
 35 40 45

Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val
 50 55 60

Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val Trp Thr Thr Pro Glu
 65 70 75 80

Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala
 85 90 95

Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln Met Glu Gly Leu
 100 105 110

Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val
 115 120 125

Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val
 130 135 140

Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu
 145 150 155 160

Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe
 165 170 175

Thr Val Thr Ala Asp Val Glu Val Pro Glu Gly Pro Arg Thr Trp Cys
 180 185 190

Met Thr Arg Lys Pro Gly Ala Val Arg Pro Leu Asn Cys Ile Val Ala
 195 200 205

5

ES 2 660 153 T3

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

Asp
 385

<210> 3
 <211> 906
 <212> ADN

5 <213> Streptomyces alboniger

<220>
 <221> gen
 <222> (1)..(853)
 <223>

10 <220>
 <221> CDS
 <222> (254)..(853)
 <223>

<400> 3
 gtcgacatcg cggccgaacc ggtcgtgagc ccttgggaca tcgccgccct ccagatcctc 60

15 gtggaggagg cggggggcgt ctgcaccgac ctgctcggcg gctcgcccca gcgcgggacg 120

ES 2 660 153 T3

ggtgcctgt cgcacaacc ggagctgcac cggctcgcg tggaggcct cgccgcccc 180
 gggccgcca cgaccggtgc cgccaccatc ccctgacca cgcccctgac ccctcacaag 240
 gagacgacct tcc atg acc gag tac aag ccc acg gtg cgc ctc gcc acc 289
 Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr
 1 5 10
 cgc gac gac gtc ccc cgg gcc gta cgc acc ctc gcc gcc cgc ttc gcc 337
 Arg Asp Asp Val Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala
 15 20 25
 gac tac ccc gcc acg cgc cac acc gtc gac ccg gac cgc cac atc gag 385
 Asp Tyr Pro Ala Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu
 30 35 40
 cgg gtc acc gag ctg caa gaa ctc ttc ctc acg cgc gtc ggg ctc gac 433
 Arg Val Thr Glu Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp
 45 50 55 60
 atc ggc aag gtg tgg gtc gcg gac gac ggc gcc gcg gtg gcg gtc tgg 481
 Ile Gly Lys Val Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val Trp
 65 70 75
 acc acg ccg gag agc gtc gaa gcg ggg gcg gtg ttc gcc gag atc ggc 529
 Thr Thr Pro Glu Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly
 80 85 90
 ccg cgc atg gcc gag ttg agc ggt tcc cgg ctg gcc gcg cag caa cag 577
 Pro Arg Met Ala Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln
 95 100 105
 atg gaa ggc ctc ctg gcg ccg cac cgg ccc aag gag ccc gcg tgg ttc 625
 Met Glu Gly Leu Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe
 110 115 120
 ctg gcc acc gtc ggc gtc tcg ccc gac cac cag ggc aag ggt ctg ggc 673
 Leu Ala Thr Val Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly
 125 130 135 140
 agc gcc gtc gtg ctc ccc gga gtg gag gcg gcc gag cgc gcc ggg gtg 721
 Ser Ala Val Val Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val
 145 150 155
 ccc gcc ttc ctg gag acc tcc gcg ccc cgc aac ctc ccc ttc tac gag 769
 Pro Ala Phe Leu Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu
 160 165 170
 cgg ctc ggc ttc acc gtc acc gcc gac gtc gag gtg ccc gaa gga ccg 817
 Arg Leu Gly Phe Thr Val Thr Ala Asp Val Glu Val Pro Glu Gly Pro
 175 180 185
 cgc acc tgg tgc atg acc cgc aag ccc ggt gcc tga cgcccgcccc 863
 Arg Thr Trp Cys Met Thr Arg Lys Pro Gly Ala
 190 195
 acgaccgcga gcgcccgacc gaaaggagcg cagcacccca tgg 906

<210> 4

<211> 199

<212> PRT

5 <213> Streptomyces alboniger

<400> 4

Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val
 1 5 10 15

ES 2 660 153 T3

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

<210> 5
 <211> 5307
 <212> ADN
 5 <213> Mus musculus

<220>
 <221> mRNA
 <222> (1)..(5307)
 <223>

10 <220>
 <221> CDS
 <222> (477)..(1040)
 <223>

<400> 5
 aggacgcgct ggatcttagg cttcccgag acctgaagaa ccggcttaga accgtttgcc 60

15 tccccgggcc tgggcccggcg gcagtagagc ctcccgacgc ggatttcccg cggggcattg 120

ES 2 660 153 T3

cagtgtgcag aagagccggc ctgctaggag cgcgagcgcg cggccgcact ttctcgcgcc 180
 tgcgcgcgcg cacgcctcaa cctgtgcggg accggccttg ggggcggagc cttagctaca 240
 caaatagaat gcgcggcggg ccttggtggg ggcggggcct tagctgcaca aataggatgc 300
 gcggcgggcc ttggtagggg cggagcctta gctgcacaaa taggatgcgc gccgggcctt 360
 ggtggggggc gggcctaagc tgcgcaagtg gtacacagct cagggctgcg atttcgcgcc 420
 aaacttgacg gcaatcctag cgtgaaggct ggtaggattt tatccccgct gccatc atg
 Met
 1
 gtt cga cca ttg aac tgc atc gtc gcc gtg tcc caa aat atg ggg att 527
 Val Arg Pro Leu Asn Cys Ile Val Ala Val Ser Gln Asn Met Gly Ile
 5 10 15
 ggc aag aac gga gac cta ccc tgg cct ccg ctc agg aac gag ttc aag 575
 Gly Lys Asn Gly Asp Leu Pro Trp Pro Pro Leu Arg Asn Glu Phe Lys
 20 25 30
 tac ttc caa aga atg acc aca acc tct tca gtg gaa ggt aaa cag aat 623
 Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser Val Glu Gly Lys Gln Asn
 35 40 45
 ctg gtg att atg ggt agg aaa acc tgg ttc tcc att cct gag aag aat 671
 Leu Val Ile Met Gly Arg Lys Thr Trp Phe Ile Pro Glu Lys Asn
 50 55 60 65
 cga cct tta aag gac aga att aat ata gtt ctc agt aga gaa ctc aaa 719
 Arg Pro Leu Lys Asp Arg Ile Asn Ile Val Leu Ser Arg Glu Leu Lys
 70 75 80
 gaa cca cca cga gga gct cat ttt ctt gcc aaa agt ttg gat gat gcc 767
 Glu Pro Pro Arg Gly Ala His Phe Leu Ala Lys Ser Leu Asp Asp Ala
 85 90 95
 tta aga ctt att gaa caa ccg gaa ttg gca agt aaa gta gac atg gtt 815
 Leu Arg Leu Ile Glu Gln Pro Glu Leu Ala Ser Lys Val Asp Met Val
 100 105 110
 tgg ata gtc gga ggc agt tct gtt tac cag gaa gcc atg aat caa cca 863
 Trp Ile Val Gly Gly Ser Ser Val Tyr Gln Glu Ala Met Asn Gln Pro
 115 120 125
 ggc cac ctc aga ctc ttt gtg aca agg atc atg cag gaa ttt gaa agt 911
 Gly His Leu Arg Leu Phe Val Thr Arg Ile Met Gln Glu Phe Glu Ser
 130 135 140 145
 gac acg ttt ttc cca gaa att gat ttg ggg aaa tat aaa ctt ctc cca 959
 Asp Thr Phe Phe Pro Glu Ile Asp Leu Gly Lys Tyr Lys Leu Leu Pro
 150 155 160
 gaa tac cca ggc gtc ctc tct gag gtc cag gag gaa aaa ggc atc aag 1007
 Glu Tyr Pro Gly Val Leu Ser Glu Val Gln Glu Glu Lys Gly Ile Lys
 165 170 175
 tat aag ttt gaa gtc tac gag aag aaa gac taa caggaagatg ctttcaagtt 1060
 Tyr Lys Phe Glu Val Tyr Glu Lys Lys Asp
 180 185
 ctctgctccc ctccctaaagc tatgcatttt tataagacca tgggactttt gctggcttta 1120
 gatctatgag taattatttc tttagggagg ggtagtggga agaattgttt gttttgtgat 1180
 cctgggggatg gaacctaaaga cccagtgcgt gctgagcaaa tgctatactg ctgagccacc 1240
 ccaaccctag cccctatata attctaaaca atatgttgtc atttcccagt aatctaaca 1300

ES 2 660 153 T3

ggttatagta aaagtgcctt aagaaatgtc acttgctata aaggctctcag tgcccctccc 1360
 atgagacctc aagtggctcc ccagcatatg cacaggggtac tgtgtgtaca agagacccca 1420
 gtgatgtaga gcccttgag catgagcaga tgtgtgggct cataaaagta ggagctaggc 1480
 aggtaagtcc aaagggcaga aacaggtttt aaacagcaga gctggaactc agactataaa 1540
 gaaaattcca tcaaagtaga gactggatta ttgtatgcac atcacacttg cagcaaagct 1600
 ctgctcactc agacagaaaa tcagtaaagtg gagaactcca ttgtgttcca tggagacgag 1660
 agcaggtgga agattatgta agatctgaaa cactgaaatt gtctgcttct catcttcagt 1720
 gagattccaa aggatagtag agtgacagaa caagaatagg cactctctac aaaaaaaga 1780
 aagaaaaaac taagtaatag caagcataat agctactgtt aagaactcag agataatgaa 1840
 ttgagaatgg atactgcttg aatgaaaat ttaataagtt agaaactaaa ctttataaaa 1900
 ataaaaaaat gagcattaaa atggctttcc tcatctcagc agggtttcag atcatcaggt 1960
 cagagaaagt atttctgcct ggctttgtaa attagtagtg tcttttttta tcttttactt 2020
 gacaatttcc tacatgttta cgatgtgtct taatcatacc aggtcctcat cccactccc 2080
 ccactacctt tgtatcttct ttctcttttc ataaccttct gaaaccagtc attgctttct 2140
 tcatgttctt gcatgtgtgc tgtccactgg agtacgggca gcctgccagt ggacacacac 2200
 acacacacac acacacacac acacacacac acacacagag ccatcagtta ccaatagctc 2260
 ttcagctggg tgcagtctta ggagcctctc ctccatccaa gctagaatat gggctctgag 2320
 atcaccacag ctgctgtgag cttgtgagtg tgggtgccat gccatgtcca gaaggcagca 2380
 ttatatatac atggccctat tccccaggct tcaggccttc attctttcca ccacttcttc 2440
 agtgctccct gagccttaga caagtctact gatagaaatg gtctgttcag agctgaatag 2500
 taaacagtca ctagtctct acactttgat cagccttggt tctgtaagct ggatgctctc 2560
 tactgaaaaa aaaaaaaaaa aaaacagaac ttcttctcca accagggtta agagcactaa 2620
 tacaggggat aaacacaggt atttagaag cactctgaaa acctagccat ttaacaaagc 2680
 agcagtagta ggtttgtcct tagaacccat gacttccaca ccacaggcat ttgaccagtt 2740
 ttgcagtact agacatgaat tgtctcctat agagccggcc tcaaatccaa ctaggaagca 2800
 ttgacaggca tgccattggt gttcagtaga aactgtaaca tgcagggcct accaccaact 2860
 aagtcataaa aagcttttct ccctggtgat gtgctcggca ccttctggcg ctgaaaactg 2920
 tccagtgaa gagggaagttt ccaggcagct ccagattgat tctctatgtc ctgcaagtaa 2980
 atgtgttgta aatgtgcctt caacagtagg ggcttgtcat ctggctacga caggcaacca 3040
 agagcgatgg caaaagatgt gttgttttg gaggctcctt aactagttac ttttaaggg 3100
 gccccacacc tggcactgag agtttcattt aataacctag gtcttcattt agtaagcatt 3160
 gtccaccagt gtgtaggatg cctccattca aacttttttc tttagttata attacctgat 3220
 aacctgaac aaagagatga actcaaagag gcagagagtc tgctgagcta gcccatgaag 3280
 caggtgagcc agtttgtggt atgtagcagt ttctctaaag gtcaaacct gtcttagaag 3340

ES 2 660 153 T3

gaagctgttt tgacataaga cggatatagag gagctagaag aagagaatat tctaagggaa 3400
 gctgggaact tccttaagct cagcaaatac acaacttaga agatcttaaa attgaccagg 3460
 tacactagac tcctccctcc ccaagcatag ataagcaatg attgccgagg ggctctgaga 3520
 cacgctgagc tgtctagaag accaactgag cctcctggca aagcagagac cagcagagct 3580
 gtctgaagga ggcttgacc taccgaactc cctagaaaga gcaactggaac ctggtgagct 3640
 gcctgtaggt tgtctaatac actccagggt tcctgagctg tgactctgct ggaacgggct 3700
 tttggtgatg caatthttgat ggtctccttc ctttgaacgt ccatactctgg ggttaacaga 3760
 catttgctcc tgtctctcca ggaggtgtgt tatacaatga tttggatccc tgggtgctcag 3820
 gcagaggcaa gtacagtagc atatgtttgt aaccctggca gtggtgtggt atgggagaga 3880
 ctggcagtgt atttgcaagc cagactgacc caagaaagaa agaaaaaatt taattgtttc 3940
 agcatctttt aaatthtgaa aaaagcataa atgtcaaaga aataattact atgtctataa 4000
 aatacaatat ttgcagcact tgaaaaggat gtggccctaa caggaaagt cttcaggaaa 4060
 agaaacacac acacaaaatc tagtattagt atgtctaaga gtaatatata ttaaaaatag 4120
 ttgacaatat cagacttagt gttatthttct gtttgtcttc tatgattagg atthtttgga 4180
 aactthttata gagaaattat ttaactatt aagtcacagc ctttgtthttc aacaggaaaa 4240
 tgagggtata gagcattgag aggaaaaatg atthttccaaa agtcatccat caagagagag 4300
 agaccagacc atacaaaaga aaaaagtcaa ttctatagaa cagttgcaca tcaataaag 4360
 caatagcgtt ttcttgcaat gaacataacc tgaaagccac caatatccag aatthgtagg 4420
 actcagtgac tcagtcagaa aaagaaaca gttatthtaca gaggaaattg aagtggccaa 4480
 tgcagagctg aaccttctac cagaaaggag agaggcaggg gaggaaatta catgtgccag 4540
 tctcaccatc thtagactaa aatgthgaca tctcccagag ctagggatgc tatgctcagc 4600
 acgcagcagt thcactgtgg ataagaccaa gagaacacct catacatgcc aaaacgaaag 4660
 acaaaagcac cctggaaatc tagacacaag gaaattgtcg tctggctatt thtagtatgag 4720
 thtcttccaa tgtgtattct gaaacaattc agtgaaagtg gtgtgtgtgg taggatgcac 4780
 cagaaagcac agthgggata tgggaaaaca agaaacctag thcaaggatca thaaagcagat 4840
 thccattgac aaccgggaa actgaggag gaagaaagct ggagthgtgta gatcctgatc 4900
 gthgtthtga cgatgthtcca cactgccagc thgtgtctct gtgthaaaccg gacacactcc 4960
 aactgccaag thcccaagct gthaacataa atgcaaagaa thcgaacac thagcaattcc 5020
 atgthtctgc thctthtctac thaaaaaaa gtgcacagcc cagthaatth ththgaaatat 5080
 aatthccatc agggthgagg thagggthgat aatcatagct thccacaaag cattgtgtat 5140
 actgaaaagg agacatgata cgththttgtg thtagaaggcg aggththcagg thggctthtga 5200
 atthgtthtga actgagactc atthgathttg aggtgtctth aggaaaggaa gaaagaaggg 5260
 acaaaaaat aaaaagcaat gthaaacatgc aaaaaaaaa aaaaaaa 5307

<210> 6
 <211> 187
 <212> PRT
 5 <213> Mus musculus
 <400> 6

ES 2 660 153 T3

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

- <210> 7
- <211> 21
- <212> ADN
- 5 <213> Artificial
- <220>
- <223> Cebador de PCE
- <400> 7
- ccacctcgga atcttcaatg c 21
- 10 <210> 8
- <211> 43
- <212> ADN
- <213> Artificial
- <220>
- 15 <223> Cebador de PCE
- <400> 8
- gatgcagttc aatggtcgaa cggcaccggg cttgcggtc atg 43
- <210> 9
- <211> 43
- 20 <212> ADN
- <213> Artificial
- <220>
- <223> Cebador de PCE

ES 2 660 153 T3

<400> 9
 catgaccgcg aagcccggtg cggttcgacc attgaactgc atc 43

 <210> 10
 <211> 21
 5 <212> ADN
 <213> Artificial

 <220>
 <223> Cebador de PCE

 <400> 10
 10 cggaaggagc tgactgggtt g 21

 <210> 11
 <211> 21
 <212> ADN
 <213> Artificial

 15 <220>
 <223> Sonda

 <400> 11
 tcaccgagct gcaagaactc t 21

 <210> 12
 20 <211> 17
 <212> ADN
 <213> Artificial

 <220>
 <223> Sonda

 25 <400> 12
 tccgcgaccc acacctt 17

 <210> 13
 <211> 17
 <212> ADN
 30 <213> Artificial

 <220>
 <223> Parte nucleotídica de sonda TaqMan

 <400> 13
 acgcgctcg ggctcga 17

 35 <210> 14
 <211> 20
 <212> ADN
 <213> Artificial

 <220>
 40 <223> Sonda

 <400> 14
 acaccgagga ccaggtgtc 20

 <210> 15
 <211> 19
 45 <212> ADN
 <213> Artificial

 <220>
 <223> Sonda

 <400> 15
 50 cccagcatca aaggtggaa 19

<210> 16
<211> 27
<212> ADN
<213> Artificial

5 <220>
<223> Parte nucleotídica de sonda TaqMan

<400> 16
cctgcgactt caacagtgac tcccact 27

10

REIVINDICACIONES

1. Un método *in vitro* de cribado de células por la expresión de una proteína de interés, comprendiendo dicho método las etapas de:
 - a) transfectar las células mediante un vector de expresión que codifica:
 - 5 (i) una proteína quimérica Puro-DHFR que comprende una secuencia de aminoácidos idéntica al menos en un 95% a la SEQ ID NO: 2, comprendiendo dicha proteína quimérica Puro-DHFR un fragmento funcional de hidfolato reductasa (DHFR) fusionado a un fragmento de puromicina N-acetil transferasa que confiere resistencia a la puromicina, en donde dicho fragmento de puromicina N-acetil transferasa muestra actividad de puromicina N-acetil transferasa; y dicho fragmento funcional de dihidrofolato reductasa muestra actividad de dihidrofolato reductasa; y
 - 10 (ii) una proteína de interés;
 - b) seleccionar células que son resistentes a la puromicina; y
 - c) ensayar la fluorescencia de las células seleccionadas en la etapa (b) con un análogo de folato marcado fluorescentemente que se une covalentemente a DHFR.
- 15 2. El método según la reivindicación 1, en donde dicho análogo de folato marcado fluorescentemente que se une a DHFR es metotrexato fluorescente (f-MTX) o trimetoprim fluorescente (f-TMP).
3. El método según la reivindicación 1, en donde dicho fragmento de puromicina N-acetil transferasa comprende la secuencia de aminoácidos 1 a 199 de la SEQ ID NO: 2.
4. El método según la reivindicación 1, en donde dicho fragmento de DHFR comprende la secuencia de aminoácidos 200 a 385 de la SEQ ID NO: 2.
5. El método según la reivindicación 1, en donde dicha proteína quimérica comprende la secuencia de SEQ ID NO: 2.
6. El método según cualquiera de las reivindicaciones precedentes, en donde la fluorescencia se mide usando un microscopio de fluorescencia o bien un clasificador de células activado por fluorescencia (FACS).
- 25 7. El método según cualquiera de las reivindicaciones precedentes, que comprende además la etapa de:
 - d) seleccionar de 1% a 20% de las células que muestran la actividad de fluorescencia más alta en la etapa (c).
8. Un método de obtención de una línea de células que expresan una proteína de interés, comprendiendo dicho método las etapas de:
 - a) cribar células según el método de cualquiera de las reivindicaciones 1ª a 7ª; y
 - 30 b) establecer una línea celular a partir de dichas células.
9. Un método de producción de una proteína de interés, comprendiendo dicho método las etapas de:
 - a) cultivar una línea de células obtenida según el método de la reivindicación 8ª bajo condiciones que permiten la expresión de dicha proteína de interés; y
 - b) recolectar dicha proteína de interés.

35

FIGURA 1

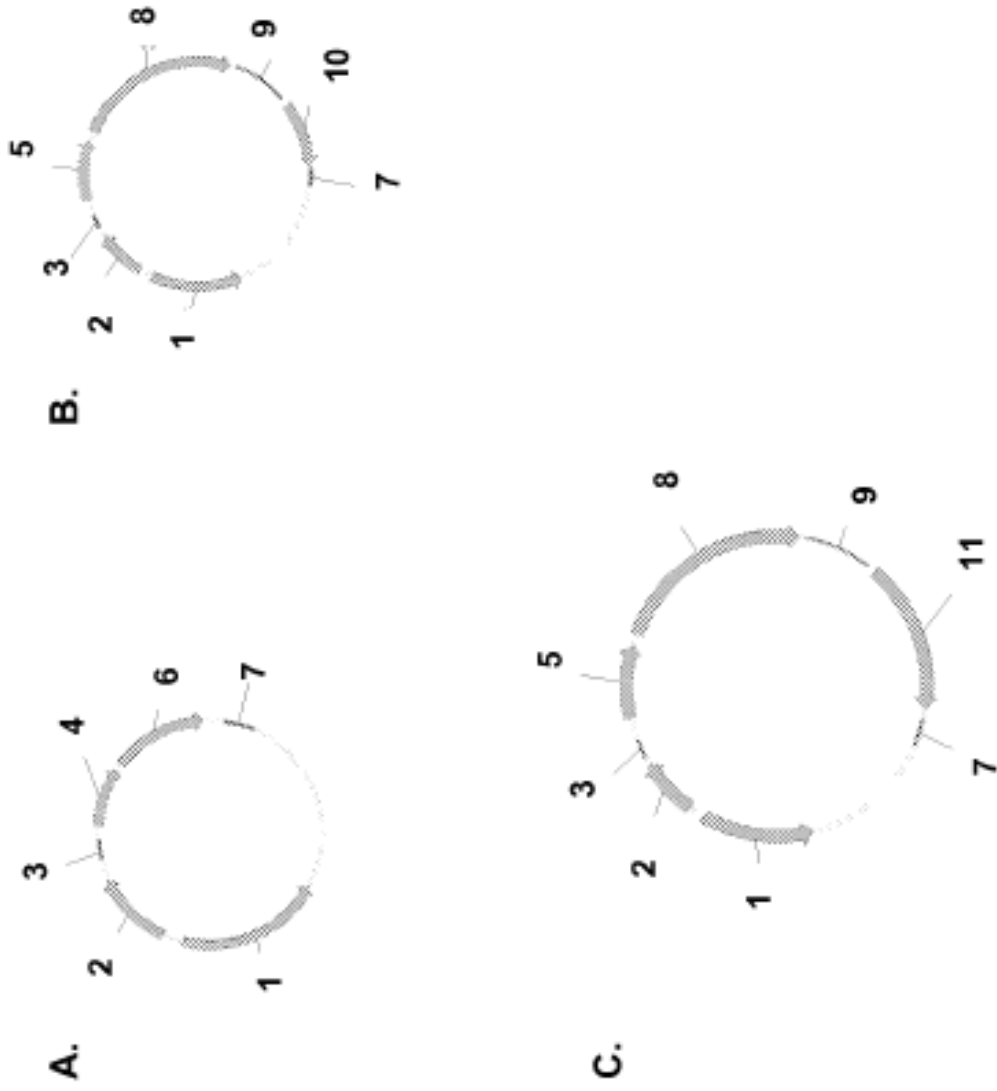


FIGURA 2

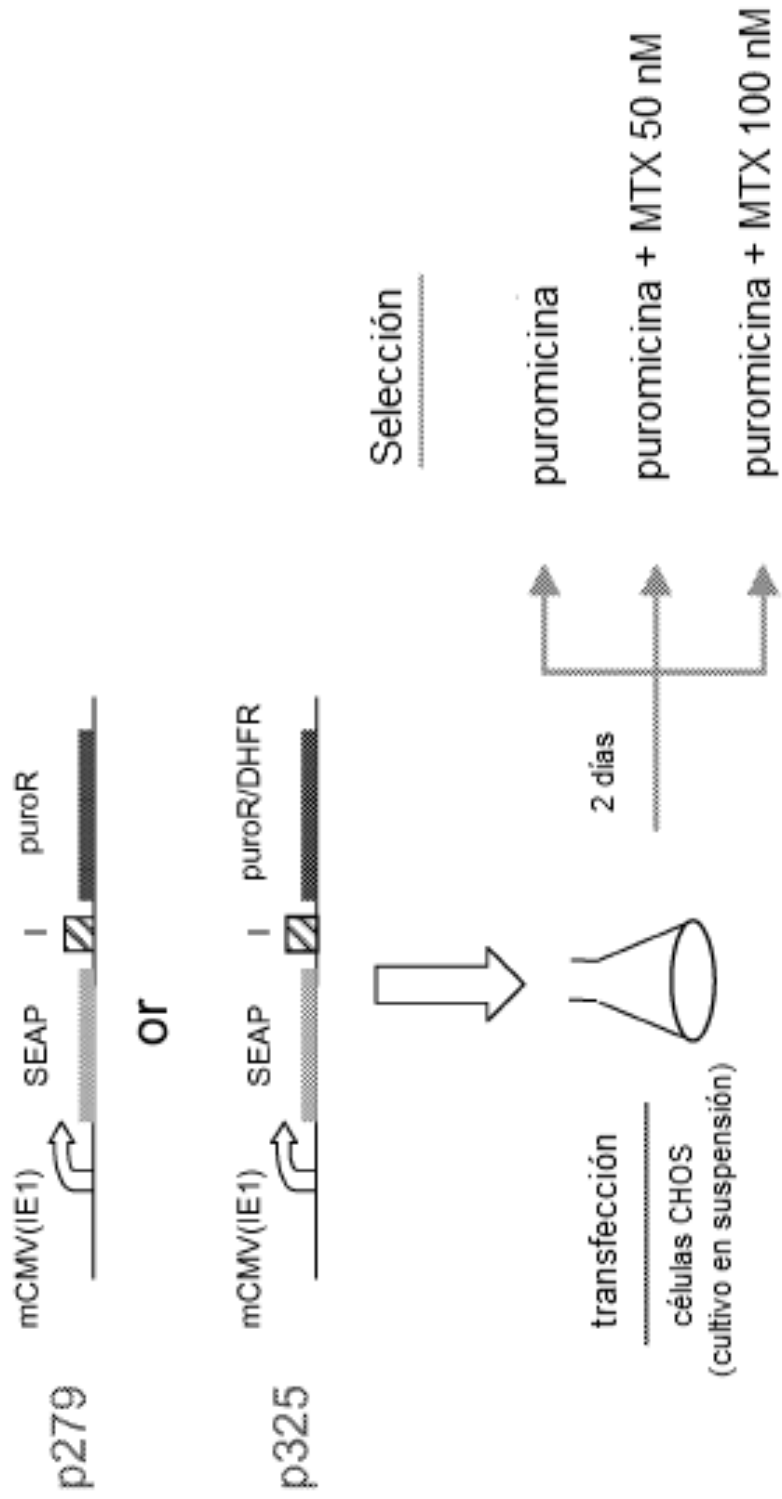


FIGURA 3

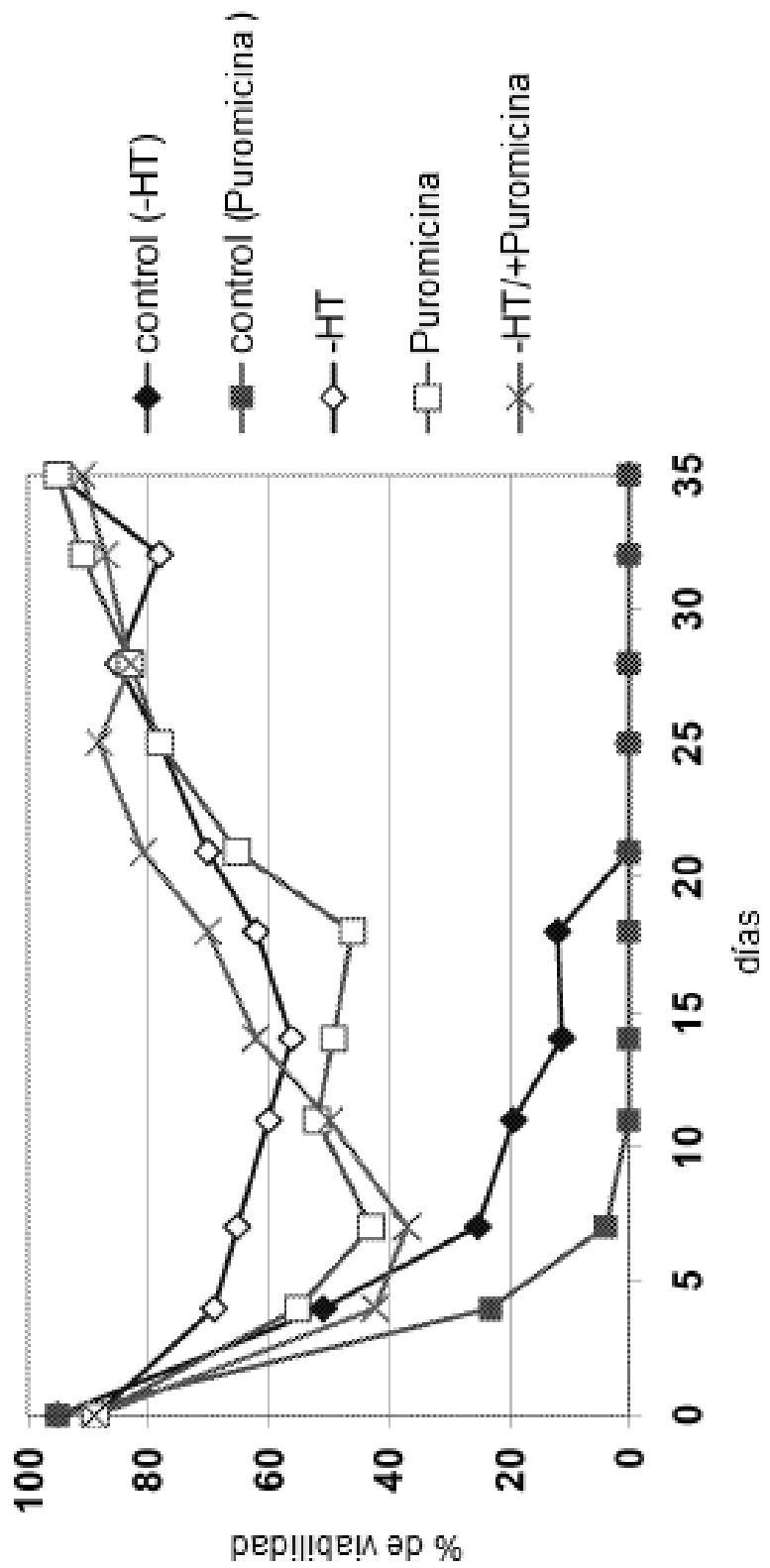


FIGURA 4

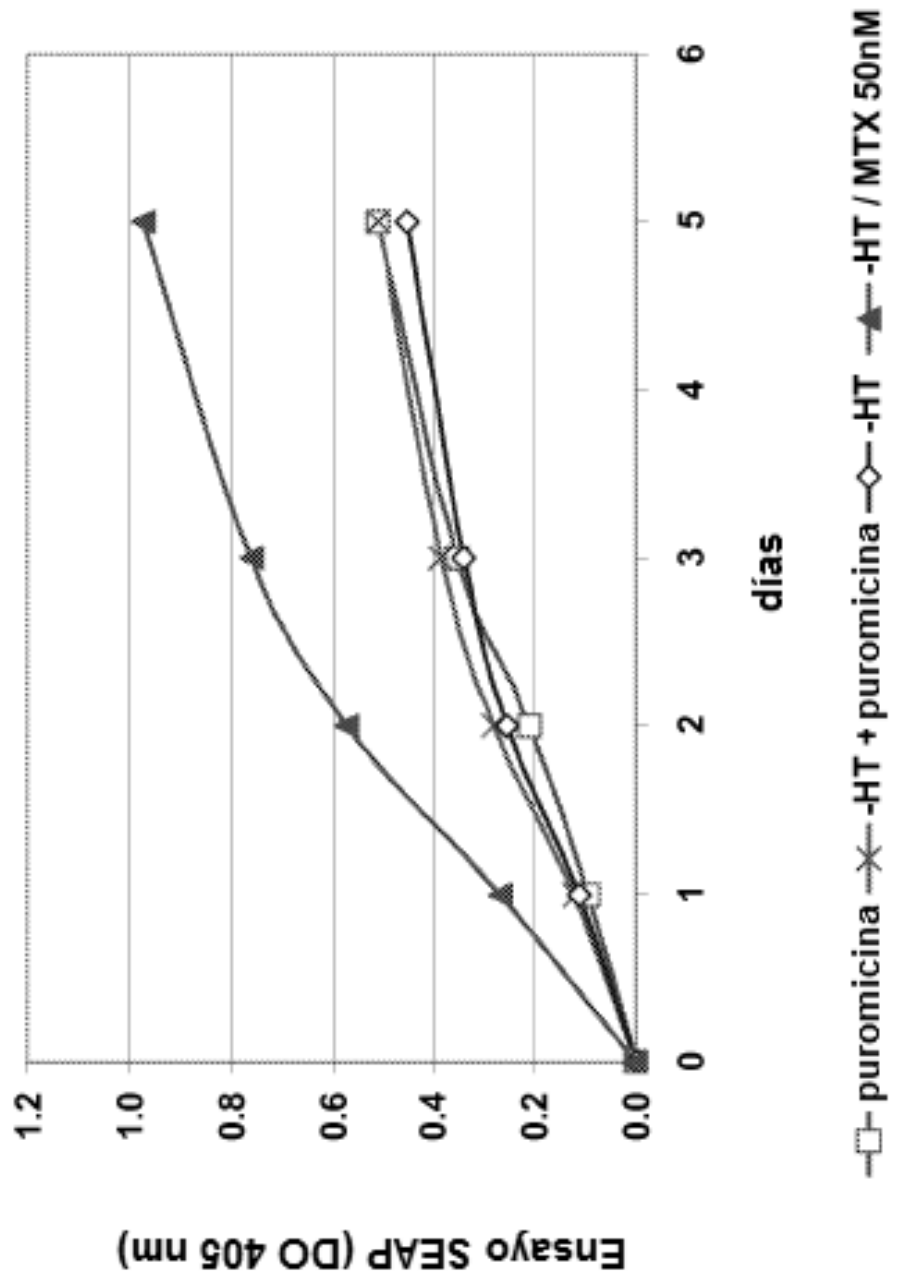


FIGURA 5

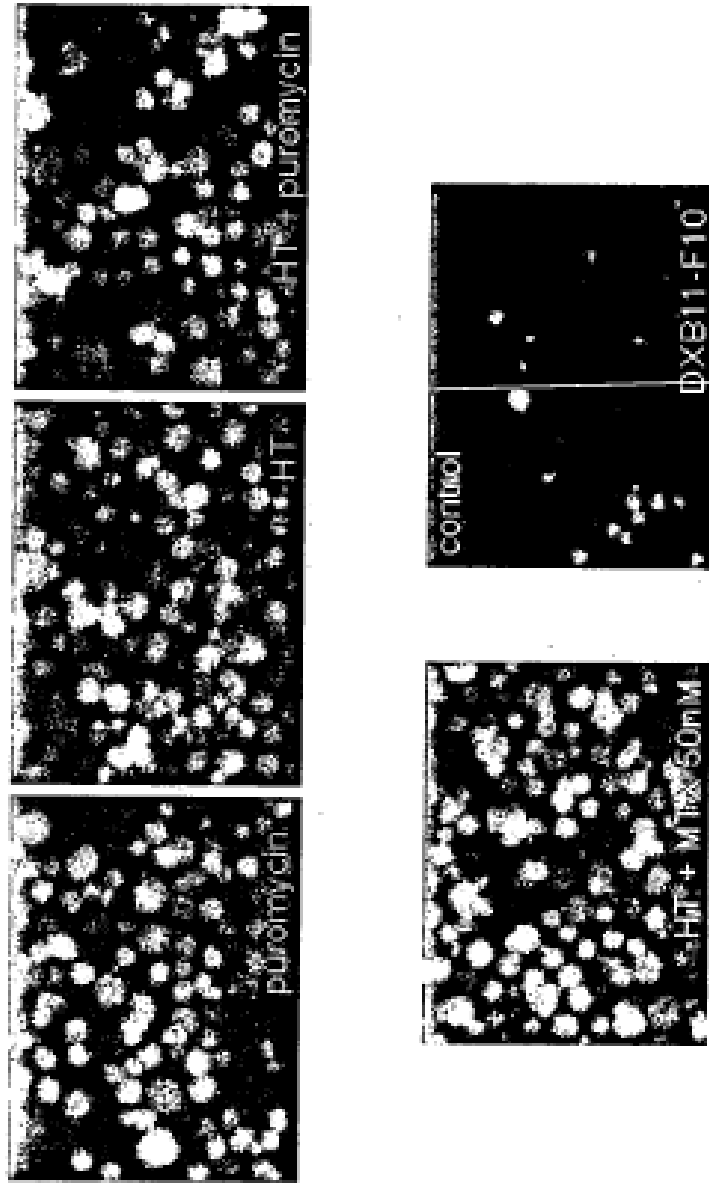


FIGURA 6

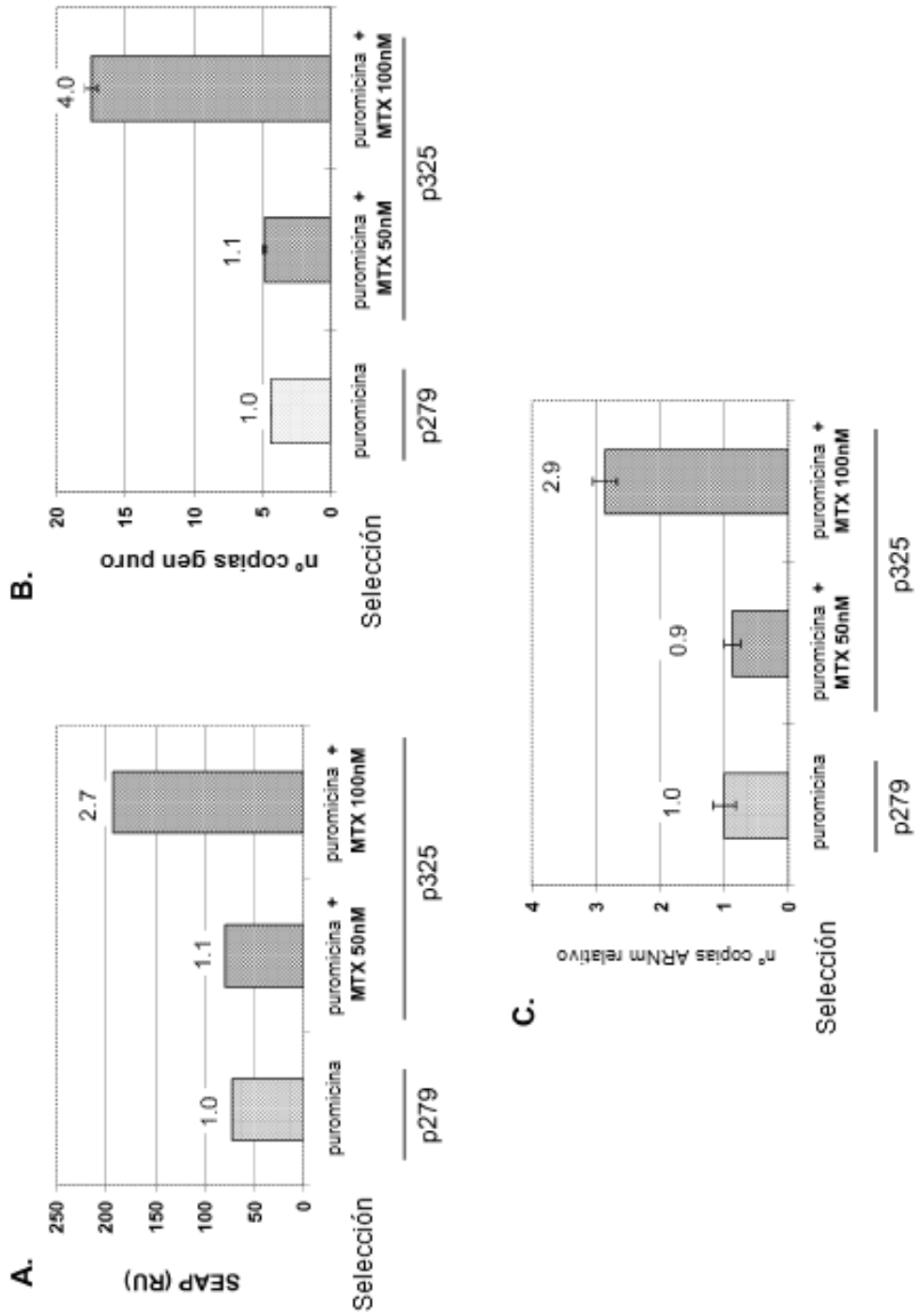


FIGURA 7

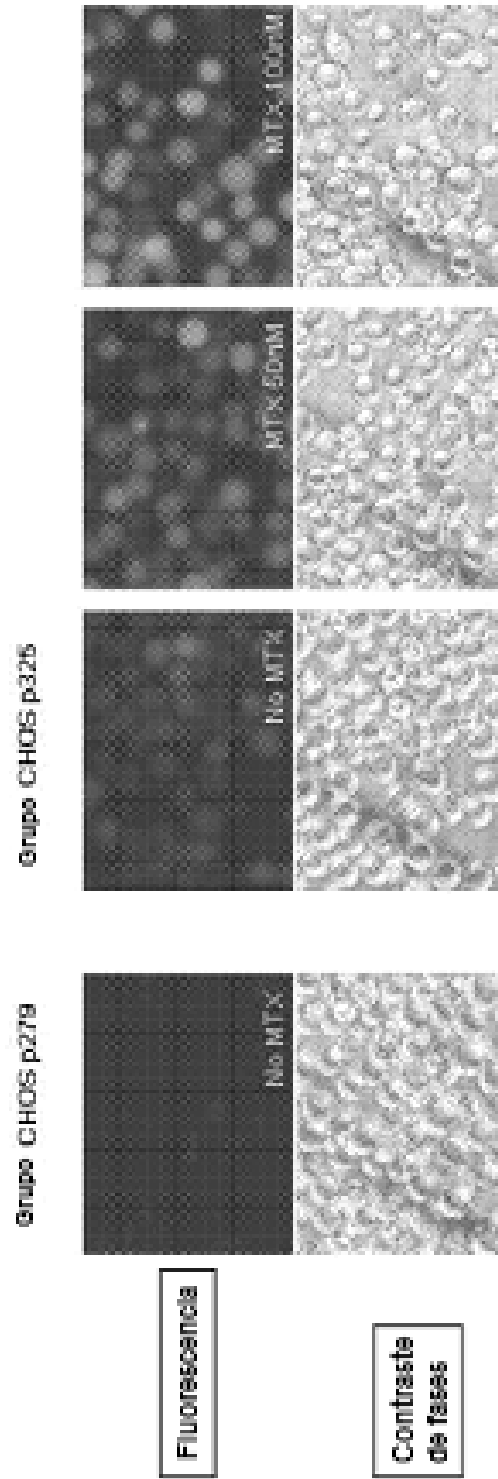


FIGURA 8

