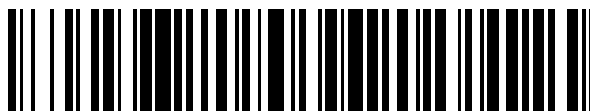


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 585 502**

51 Int. Cl.:

C12N 15/79

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.12.2008** **E 08865499 (1)**

97 Fecha y número de publicación de la concesión europea: **04.05.2016** **EP 2227546**

54 Título: **Vector de expresión en mamíferos**

30 Prioridad:

21.12.2007 EP 07150339

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

06.10.2016

73 Titular/es:

**NOVARTIS AG (100.0%)
Lichtstrasse, 35
4056 Basel , CH**

72 Inventor/es:

**KNOPF, HANS-PETER y
WILMS, BURKHARD**

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 585 502 T3

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

DESCRIPCIÓN

Vector de expresión en mamíferos

La presente invención es pertinente a un vector de expresión en mamíferos para expresar un polipéptido de interés así como a métodos para expresar un polipéptido de interés en una célula huésped de mamífero utilizando un vector respectivo y células huésped que comprenden dicho vector.

La capacidad de clonar y expresar péptidos y proteínas recombinantes en cantidades grandes se ha hecho crecientemente importante en años recientes. La capacidad de purificar altos niveles de proteínas es importante en el campo farmacéutico y biotecnológico humano, por ejemplo para producir agentes farmacéuticos proteínicos así como en el desarrollo de la investigación básica, por ejemplo, para cristalizar proteínas que permiten la determinación de la estructura tridimensional. Las proteínas que de otra manera son difíciles de obtener en cantidades pueden ser sobre expresadas en una célula huésped y subsecuentemente aisladas y purificadas.

La selección de un sistema de expresión para la producción de proteínas recombinantes depende de muchos factores, incluyendo características de crecimiento celular, niveles de expresión, expresión intracelular y extracelular, modificaciones postranslacionales y actividad biológica de la proteína de interés, así como dificultades regulatorias y consideraciones económicas en la producción de proteínas terapéuticas. Las ventajas claves de las células de mamíferos sobre otros sistemas de expresión tales como bacterias o levaduras son la capacidad de llevar a cabo un plegamiento apropiado de la proteína, glicosilación enlazada a N compleja y glicosilación enlazada a O autentica, así como un amplio espectro y otras modificaciones postranslacionales. Debido a las ventajas descritas, las células de mamíferos son actualmente el sistema de expresión de selección para producir proteínas terapéuticas complejas tales como anticuerpos monoclonales. La primera etapa en la generación de una línea celular recombinante es la construcción de un vector de expresión. El vector de expresión es el elemento clave para guiar la expresión del gen heterólogo en la célula huésped y para proveer marcadores de selección para generar la línea celular recombinante. Los elementos esenciales de los vectores de expresión en mamíferos usualmente incluyen un promotor constitutivo o inducible capaz de una actividad transcripcional robusta; el procesamiento y señales de traducción de ARN optimizados que usualmente incluyen una secuencia Kozak, un codón de terminación de traducción, escisión de ARNm y señales de poliadenilación, así como señales de división de ARNm; un terminador de transcripción; marcadores de selección para la preparación de líneas celulares estables y para la amplificación genética; un origen procariota de replicación y marcadores de selección para la propagación del vector en bacterias.

En años recientes el foco del desarrollo ha estado concentrándose en el diseño de vectores mejorados para la expresión genética en células de mamíferos. A pesar de la multitud de vectores disponibles sin embargo, aún es un reto la producción de polipéptidos/proteínas robustos en células de mamíferos. Varios factores pueden influir en la expresión recombinante en células de mamíferos, incluyendo la fuerza del promotor, el contexto de la 5' no traducida y la región de iniciación de la traducción, la eficiencia de la región no traducida 3', para poliadenilar y terminar la transcripción, el sitio de inserción del gen recombinante integrado aleatoriamente en el cromosoma huésped, y el número de copias integradas del gen que está siendo expresado. Los incrementos en el número de copias genéticas son alcanzados de la manera más común por amplificación genética utilizando líneas celulares deficientes en una enzima tal como dihidrofolato reductasa (DHFR) o glutamina sintetasa (GS) en conjunción con vectores de expresión que contienen genes que codifican estas enzimas y agentes tales como metrotexato (MTX), el cual inhibe la DHFR, y la metionina sulfoxamina (MSX) la cual inhibe la GS. Utilizando vectores de expresión que contienen el gen recombinante bajo control de un promotor fuerte y genes que codifican DHFR o GS, los transfectantes DHFR⁺ o GS⁺, respectivamente, son obtenidos primero y se logra entonces la amplificación del gen haciendo crecer los transfectantes en concentraciones progresivamente crecientes de MTX o MSX.

La WO2004/050879 enseña un vector de expresión que comprende un casete de expresión bicistrónico que comprende un promotor de ubiquitina de hámster/S27a para expresar el gen de interés y un gen informador tal como GFP. En el vector circular, el gen DHFR está localizado en 5' con respecto al casete de expresión que comprende el gen que codifica el polipéptido de interés y el gen informador de GFP.

La WO01/04306 enseña un vector de expresión que comprende un gen marcador seleccionable amplificable tal como DHFR y un gen informador, por ejemplo GFP. Enseña un sistema de expresión bicistrónico en donde la expresión del gen que codifica la proteína de interés está enlazado operativamente a la expresión del marcador seleccionable amplificable o al gen informador GFP.

Lin et al "Expression efficiency of the human thrombomodulin-encoding gene in various vector and host systems" in Gene, 147 (1994) 287 292 analiza el rendimiento de cinco vectores de expresión en mamíferos que expresan trombomodulina recombinante humana en diferentes células huésped.

La US 5,017,478 enseña plásmidos de expresión eucariota que comprenden un casete de gen seleccionable tal como DHFR y un casete genético que codifica el producto de interés, en donde el casete de gen seleccionable y el casete

de gen para expresar la proteína de interés están dispuestos adyacentes uno a otro en orientación transcripcional opuesta y divergente. El casete de gen seleccionable está localizado en 5' del casete genético para expresar la proteína de interés. Un casete genético marcador seleccionable adicional que codifica por ejemplo neo o higromicina puede estar localizado en 3' del casete genético que codifica el producto de interés. En comparación, inter alia un vector de expresión pPA206 fue diseñado en el cual el casete de gen seleccionable que codifica DHFR, el casete genético que codifica el producto de interés y el casete genético seleccionable que codifica neo están orientados en la misma dirección.

Es el objeto de la presente invención, proveer un vector de expresión mejorado así como un sistema de expresión para expresar un polipéptido de interés en una célula de mamífero.

10 De acuerdo con lo anterior, conforme a la reivindicación 1, se provee un ácido nucleico de vector linealizado adecuado para expresar al menos un polipéptido de interés en una célula de mamífero, que comprende

(a) al menos un casete de expresión (POI) adecuado para expresar un polipéptido de interés, en donde dicho casete de expresión comprende al menos un promotor o un elemento promotor/potenciador;

15 (b) un casete de expresión (MSM) que comprende un gen marcador seleccionable de mamífero, en donde dicho gen marcador seleccionable de mamífero es un gen resistente a antibióticos y en donde dicho casete de expresión comprende al menos un promotor o un elemento promotor/potenciador;

(c) un casete de expresión (MASM) que comprende un gen marcador seleccionable amplificable de mamífero que codifica una dihidrofolato reductasa (DHFR) enzimáticamente funcional, en donde dicho casete de expresión comprende al menos un promotor o un elemento promotor/potenciador; y

20 (d) un casete de expresión (PSM) que comprende un gen marcador seleccionable procariota,

en donde el casete de expresión para expresar el polipéptido de interés (POI) comprende un promotor y/o potenciador más fuerte que los casetes de expresión para expresar los marcadores seleccionables,

25 y en donde el casete de expresión (POI) está flanqueado en 5' por el casete de expresión (MASM), el casete de expresión (MSM) está localizado 3' desde el casete de expresión (POI) y en donde los casetes de expresión (MASM), (POI) y (MSM) están dispuestos en la misma orientación de 5' a 3',

y en donde dicho vector ha sido linealizado a través de un sitio de restricción de linealización que está localizado en la forma circular del vector entre los casetes de expresión (PSM) y (MASM) y en donde dicha forma circular del vector, el casete de expresión (PSM) está localizado entre los casetes de expresión (MSM) y (MASM), de tal manera que en el vector linealizado, el casete de expresión (POI) está flanqueado en 5' por el casete de expresión (MASM) y en 3' por el casete de expresión (MSM) y el casete de expresión (PSM) está localizado en 3' del casete de expresión (MSM),

30 y en donde el ácido nucleico del vector lineal comprende los siguientes elementos genéticos en la disposición indicada, en donde la dirección 5' a 3' es indicada por la →:

I. Promotor del casete de expresión (MASM) (→)

II. Gen que codifica el marcador seleccionable amplificable de mamífero del casete de expresión (MASM) (→)

35 III. Intrón del casete de expresión (MASM) (→)

IV. Sitio poliA del casete de expresión (MASM) (→)

V. Promotor del casete de expresión (POI) (→)

VI. Intrón del casete de expresión (POI) (→)

VII. Polinucleótido que codifica un polipéptido de interés, el cual es insertado en el casete de expresión (POI) (→)

40 VIII. Sitio PoliA del casete de expresión (POI) (→)

IX. Promotor del casete de expresión (POI') (→)

X. Intrón del casete de expresión (POI') (→)

XI. Polinucleótido que codifica un polipéptido adicional de interés, el cual está insertado en el casete de expresión (POI') (→)

XII. Sitio poliA del casete de expresión (POI') (→)

XIII. Promotor del casete de expresión (MSM) (→)

5 XIV. Gen que codifica el marcador seleccionable de mamíferos del casete de expresión (MSM) (→)

XV. Sitio poliA del casete de expresión (MSM) (→)

XVI. Casete de expresión PSM (→) o (←),

en donde el casete de expresión POI' tal como está indicado por los elementos IX a XII es opcional.

10 Un "ácido nucleico vector" de acuerdo con la presente invención es un polinucleótido que portan al menos un fragmento de ácido nucleico foráneo. Un ácido nucleico vector funciona como un "portador molecular" que suministra fragmentos de ácidos nucleicos respectivamente polinucleótidos a una célula huésped. Comprende al menos un casete de expresión que comprende secuencias reguladoras y codificadoras. Un casete de expresión permite la expresión apropiada de un polinucleótido incorporado. Los casetes de expresión decisivos para la presente invención serán descritos subsecuentemente en mayor detalle. Los polinucleótidos foráneos, por ejemplo que codifican el polipéptido de interés pueden ser insertados en los casetes de expresión del ácido nucleico vector con el fin de ser expresados. El ácido nucleico vector de acuerdo con la presente invención puede estar presente en forma circular o linealizada. Dicho sitio puede ser, por ejemplo, un sitio de clonación múltiple (MCS).

20 El casete de expresión (POI) define el casete de expresión para expresar un polipéptido de interés. Dicho casete de expresión (POI) comprende bien sea el polinucleótido que codifica el polipéptido de interés o comprende un sitio adecuado para insertar un polinucleótido respectivo que codifica el polipéptido de interés.

Un "polipéptido" se refiere a una molécula que comprende un polímero de aminoácidos enlazados entre sí mediante enlaces peptídicos. Los polipéptidos incluyen polipéptidos de cualquier longitud, incluyendo proteínas (por ejemplo, que tienen más de 50 aminoácidos) y péptidos (por ejemplo, que tienen 2-49 aminoácidos). Los polipéptidos incluyen proteínas y/o péptidos de cualquier actividad o bioactividad. Más adelante se delinean ejemplos adecuados.

25 El casete de expresión (MSM) define el casete de expresión que comprende un gen marcador seleccionable de mamífero el cual es un gen resistente a antibióticos. Genes marcadores seleccionables de mamíferos permiten la selección de células huésped de mamíferos que comprenden dichos genes y así las células huésped de mamíferos comprenden el vector.

30 El casete de expresión (MASM) define el casete de expresión que comprende un gen marcador seleccionable amplificable de mamífero que codifica una dihidrofolato reductasa enzimáticamente funcional (DHFR). Los genes marcadores seleccionables amplificables de mamíferos permiten la selección de células huésped de mamíferos que contienen el vector así como la amplificación de genes.

35 Los términos "5' " y "3' " son una convención utilizada para describir características de las secuencias de ácidos nucleicos relacionadas bien sea a la posición de elementos genéticos y/o la dirección de eventos (5' a 3'), tales como por ejemplo la transcripción por ARN polimerasa o traducción del ribosoma la cual avanza en la dirección 5' a 3'. Sinónimos son corriente arriba (5') y corriente abajo (3'). Convencionalmente, las secuencias de ADN, los mapas genéticos, las tarjetas de vector y las secuencias de ARN se dibujan con 5' a 3' de izquierda a derecha o la dirección 5' a 3' es indicada con flechas, en donde la cabeza de la flecha apunta en la dirección 3'. De acuerdo con lo anterior, 5' (corriente arriba) indica elementos genéticos posicionados hacia el lado izquierdo, y 3' (corriente abajo) indica elementos genéticos posicionados hacia la derecha, cuando se sigue esta convención.

40 La disposición y orientación de los casetes de expresión presentes en el vector de la invención es importante. De acuerdo con las enseñanzas de la presente invención, el casete de expresión (POI) está flanqueado en 5' por el casete de expresión (MASM). De acuerdo con lo anterior, el casete de expresión (MASM) está localizado 5' adyacente al casete de expresión (POI) y en cercana proximidad al mismo. Desde luego, las secuencias esqueleto del vector pueden separar los casetes de expresión (MASM) y (POI). Sin embargo, preferiblemente, ningún otro casete de expresión se localiza entre el casete de expresión (MASM) y el casete de expresión (POI). El casete de expresión (MSM) está localizado a 3' desde el casete de expresión (POI). Pueden insertarse casetes de expresión adicionales entre los casetes de expresión (POI) y (MSM) tales como, por ejemplo, un casete de expresión (POI') adicional para expresar un polipéptido adicional de interés (descrito en más detalle más adelante). Los casetes de expresión (MASM), (POI) y (MSM) están todos dispuestos en la misma orientación 5' a 3'. Los inventores encontraron, que esta configuración de ácidos nucleicos de vector particular permite la generación rápida de altos rendimientos de líneas

celulares.

El polinucleótido que codifica el polipéptido de interés puede ser incorporado en el casete de expresión (POI) utilizando métodos de clonación apropiados, por ejemplo utilizando enzimas de restricción con el fin de insertar el polinucleótido que codifica el polipéptido de interés en el casete de expresión (POI). Para este propósito, el casete de expresión (POI) puede comprender por ejemplo un sitio de clonación múltiple (MCS) el cual puede ser utilizado por ejemplo en todos los marcos de lectura. Más adelante se describen en detalle sitios de MCS adecuados. El casete de expresión (POI) puede comprender también un polinucleótido de reemplazo o una secuencia de ácido nucleico de relleno, la cual puede ser escindida y reemplazada por el polinucleótido que codifica el polipéptido de interés. La presente invención provee un ácido nucleico vector como se describe más arriba, que comprende un casete de expresión (POI) que comprende el polinucleótido que codifica el polipéptido de interés. El ácido nucleico vector de expresión final puede ser transfectado para expresión en la célula huésped.

Un polinucleótido es un polímero de nucleótidos que están enlazados usualmente desde una desoxirribosa o ribosa a otra. El término "polinucleótido" no comprende ninguna restricción de tamaño y también abarca polinucleótidos que comprenden modificaciones, en particular nucleótidos modificados.

En un ácido nucleico vector circular el casete de expresión (MSM) está dispuesto en 3' del casete de expresión (POI) y el casete de expresión (MASM) está dispuesto en 3' del casete de expresión (MSM). Una descripción alternativa de tal vector circular es un ácido nucleico vector circular para expresar al menos un polipéptido de interés en una célula de mamífero, que comprende

(a) al menos un casete de expresión (POI) para expresar un polipéptido de interés;

(b) un casete de expresión (MSM) que comprende un gen marcador seleccionable de mamíferos;

(c) un casete de expresión (MASM) que comprende un gen marcador seleccionable amplificable de mamíferos;

en donde el casete de expresión (MSM) está dispuesto en 3' del casete de expresión (POI) y el casete de expresión (MASM) está dispuesto en 3' del casete de expresión (MSM) y en donde los casetes de expresión (MASM), (POI) y (MSM) están dispuestos en la misma orientación 5' a 3'.

La linealización del ácido nucleico vector antes de la transfección mejora frecuentemente la eficiencia de una transfección estable. Esto también puesto que el punto de linealización puede ser controlado si el vector es linealizado antes de la transfección.

El vector de expresión circular comprende un sitio de restricción predefinido, el cual es utilizado para linealización del ácido nucleico vector antes de la transfección. La colocación inteligente de dicho sitio de restricción de linealización es importante, puesto que dicho sitio de restricción determina si el ácido nucleico vector es abierto/linealizado y así determina el orden/disposición de los casetes de expresión cuando el constructo es integrado en el genoma de la célula de mamífero.

De acuerdo con lo anterior, el ácido nucleico vector circular comprende un sitio de restricción de linealización para linealizar el vector, en donde dicho sitio de restricción de linealización está localizado entre los casetes de expresión (MSM) y (MASM) y entre los casetes de expresión (PSM) y (MASM) (véase más abajo). Dicho sitio de restricción de linealización es único y está presente solamente una vez en el ácido nucleico vector de expresión. Por ejemplo se puede utilizar un sitio de restricción de linealización que es reconocido por una enzima de restricción que tiene una frecuencia de corte baja con el fin de patronizar que el vector es solamente escindido en el sitio de restricción de linealización pero no (o solo raramente) por ejemplo dentro del casete de expresión o el esqueleto del vector. Esto por ejemplo puede promoverse proveyendo un sitio de restricción para una enzima de restricción que tiene una secuencia de reconocimiento de más de seis pares de bases o que reconoce secuencias que están subrepresentadas en ADN cromosómico. Un ejemplo adecuado es la enzima SmaI y el vector puede por lo tanto incorporar un sitio de reconocimiento de SmaI como sitio de restricción de linealización único. En caso de que dicho sitio de restricción de linealización esté presente más de una vez en la secuencia de ácido nucleico vector (incluyendo los polinucleótidos que codifican el polipéptido de interés), o en caso de que se utilice una enzima de restricción que corte varias veces en la secuencia del ácido nucleico vector, también está dentro del alcance de la presente invención por ejemplo, alterar/hacer mutar los sitios de restricción además del sitio de restricción de linealización que está localizado entre los casetes de expresión (MSM) y (MASM), con el fin de eliminar aquellos sitios de restricción adicionales y obtener un sitio de restricción de linealización único o al menos raro.

En caso de que se utilice el vector como el vector de expresión estándar previsto, por ejemplo como una herramienta para la expresión de varios diferentes polipéptidos, es ventajoso proveer un sitio de restricción de linealización que comprende múltiples sitios de reconocimiento para enzimas que tengan una frecuencia de corte baja. Las enzimas de restricción escogidas para la linealización deberían preferiblemente no cortar dentro de los casetes de expresión para

los marcadores seleccionables u otras secuencias de esqueleto de vector con el fin de asegurar que la enzima corta solamente una vez para la linealización apropiada del vector. Proveyendo un sitio de restricción de linealización que comprende sitios de reconocimiento múltiples para enzimas de restricción que tengan una frecuencia de corte baja, el usuario puede seleccionar una enzima de restricción adecuada para linealización a partir de las opciones provistas con el fin de evitar con seguridad la restricción dentro del polinucleótido que codifica el polipéptido de interés. Sin embargo, como se delineó más arriba los sitios de restricción adicional pueden ser mutados o puede llevarse a cabo una digestión de restricción parcial.

Al colocar el sitio de restricción de linealización entre el casete de expresión (MSM) y el casete de expresión (MASM) se obtiene el efecto de que el casete de expresión (POI) (y casetes de expresión adicionales para expresar los polipéptidos de interés -si están presentes) es flanqueado en 5' por el casete de expresión (MASM). El casete de expresión (MSM) está localizado en 3' del casete de expresión (POI) por linealización. Por lo tanto, los casetes de expresión (MSM) y (MASM) son separados por linealización del ácido nucleico vector circular. Puesto que un casete de expresión (PSM) para el marcador de selección bacteriana está presente (véase más abajo), el sitio de restricción de linealización es colocado entre los casetes de expresión (PSM) y (MASM). Esto tiene el efecto de que el gen marcador de selección bacteriana es 3' y así "por fuera" de las partes "de mamífero" del ácido nucleico vector linealizado. Esta disposición es favorable porque los genes bacterianos presumiblemente no son ventajosos para la expresión en mamíferos puesto que las secuencias bacterianas pueden llevar a una metilación incrementada u otros efectos de silenciamiento en las células de mamíferos.

Ejemplos no limitantes para genes marcadores seleccionables de mamíferos que pueden estar comprendidos en el casete de expresión (MSM) incluyen genes de resistencia a antibióticos por ejemplo que confieren resistencia a G418; higromicina (hig o hph, disponible comercialmente de Life Technologies, Inc. Gaithersboro, Md); neomicina (neo, disponible comercialmente de Life Technologies, Inc. Gaithersboro, Md.); zeocina (Sh Ble, disponible comercialmente de Pharmingen, San Diego Calif.); puromicina (pac, puromycin-N-acetil-transferasa, disponible de Clontech, Palo Alto Calif.), ouabaina (oua, disponible de Pharmingen) y blasticidina (disponible de Invitrogen). Los genes marcadores seleccionables de mamíferos respectivos son bien conocidos y permiten la selección de células huésped de mamíferos que comprenden dichos genes así de células huésped que comprenden el vector. El término "gen" tal como se utiliza aquí también se refiere a un polinucleótido natural o sintético que codifica una variante funcional del marcador seleccionable proveyendo la resistencia buscada. Por lo tanto, también versiones truncadas o mutadas de un gen tipo silvestre o polinucleótidos sintéticos están abarcados en tanto provean la resistencia pretendida. De acuerdo con una realización preferida, dicho casete de expresión (MSM) comprende un gen que codifica una neomicina fosfotransferasa enzimáticamente funcional (I o II). Esta realización trabaja bien en combinación con el uso de un gen que codifica una DHFR enzimáticamente funcional como un gen marcador seleccionable amplificable.

Genes marcadores de mamíferos seleccionables, amplificables permiten la selección de células huésped de mamíferos que contienen el vector así como la amplificación de genes. De acuerdo con la reivindicación 1, el gen marcador de mamíferos seleccionable amplificable es el gen de dihidrofolato reductasa (DHFR). Otros estudios actualmente en uso son entre otros el sistema glutamina sintetasa (gs) (Bebbington et al., 1992), y el sistema guiado por histidinol (Hartmann and Mulligan, 1988). Estos marcadores amplificables también son marcadores seleccionables y así pueden ser utilizados para seleccionar aquellas células que obtuvieron el vector. La DHFR provee buenos resultados. La selección usualmente se presenta en la ausencia del metabolito apropiado (hipoxantina y timidina en caso de DHFR, previniendo el crecimiento de células no transformadas. Con sistema amplificables tales como el sistema DHFR, la expresión de una proteína recombinante puede ser incrementada exponiendo las células a ciertos agentes que promueven la amplificación de genes tales como antifolatos (por ejemplo, metotrexato (MTX)) en caso del sistema DHFR.

De acuerdo con una realización, dicho casete de expresión (MSM) comprende un gen que codifica una neomicina fosfotransferasa enzimáticamente funcional y dicho casete de expresión (MASM) comprende un gen que codifica una dihidrofolato reductasa enzimáticamente funcional (DHFR).

El vector puede comprender al menos un casete de expresión adicional en diferentes (POI') para expresar cualquier polipéptido adicional de interés. Dicho casete de expresión (POI') adicional está localizado entre el casete de expresión (POI) y el casete de expresión (MSM). Dicho casete de expresión (POI') está dispuesto en la misma orientación 5' a 3' que los casetes de expresión (POI) y (MSM). De acuerdo con una realización, comprende el polinucleótido de acuerdo con el polipéptido adicional de interés.

De acuerdo con lo anterior, el ácido nucleico vector de acuerdo con la presente invención puede comprender más de un casete de expresión para expresar polipéptidos de interés. Por lo tanto, también es posible que varios casetes de expresión ((POI), (POI'), (POI'') etc.) para expresar diferentes polipéptidos de interés estén dispuestos en el ácido nucleico vector de expresión de acuerdo con la presente invención. Estos casetes de expresión están flanqueados en 5' por el casete de expresión (MASM) y en 3' por el casete de expresión (MSM). Por lo tanto, la presente invención también provee un ácido nucleico vector que comprende más de un casete de expresión que codifica por ejemplo subunidades de proteínas díméricas o multiméricas de orden más alto. Los casetes de expresión que codifican diferentes subunidades de una proteína multimérica, incorporados cada uno en un casete de expresión diferente

pueden ser colocados adyacentes uno a otro. Para proteínas multiméricas codificadas por al menos dos genes distintos (por ejemplo, en cadenas de inmunoglobulinas livianas y pesadas o fragmentos funcionales de las mismas tales como al menos las regiones variables de las cadenas de inmunoglobulinas livianas y pesadas), los polinucleótidos que codifican las subunidades deseadas del polipéptido de interés son insertados en los casetes de expresión (POI) y (POI'). Una realización respectiva que utiliza al menos dos casetes de expresión (POI) y (POI') para expresar polipéptidos de interés es particularmente ventajosa para la expresión de moléculas de inmunoglobulina tales como anticuerpos o fragmentos funcionales de los mismos. De acuerdo con lo anterior, se provee un ácido nucleico vector para expresar una molécula de inmunoglobulina que comprende en cada casete de expresión (POI) y (POI') un polinucleótido que codifica bien sea una cadena liviana o pesada de una molécula de inmunoglobulina o fragmentos de las mismas, en donde cada casete de expresión (POI) y (POI') comprende uno de dichos polinucleótidos. De acuerdo con lo anterior, el casete de expresión (POI) puede comprender bien sea el polinucleótido para expresar la cadena liviana, o el polinucleótido para expresar la cadena pesada de la molécula de inmunoglobulina.

De acuerdo con una realización preferida, el casete de expresión (POI) comprende el polinucleótido que codifica al menos parte de la cadena liviana de dicha molécula de inmunoglobulina o un fragmento funcional de la misma y el casete de expresión (POI') comprende un polinucleótido que codifica al menos parte de la cadena pesada de dicha molécula de inmunoglobulina o un fragmento funcional de la misma. El disponer el casete de expresión para la cadena liviana 5' con respecto al casete de expresión de la cadena pesada ha demostrado ser beneficioso con respecto a la tasa de expresión de las moléculas de inmunoglobulina. De acuerdo con una realización, el ácido nucleico vector de expresión está diseñado de tal manera que los casetes de expresión comprenden ya un polinucleótido que codifica al menos parte de las regiones constantes de una molécula de inmunoglobulina. Los fragmentos de polinucleótido que codifican las partes variables de las moléculas de inmunoglobulina pueden ser insertados entonces en el usuario/cliente en los casetes de expresión utilizando estrategias de clonación apropiadas con el fin de obtener el vector de expresión final.

Los casetes de expresión de mamíferos presentes en el vector de expresión de acuerdo con la presente invención están diseñados de tal manera que permiten la expresión de los polinucleótidos/genes incorporados en células de mamíferos. Para este propósito los casetes de expresión comprenden las secuencias reguladoras necesarias tal como una secuencia promotora y una de terminación de transcripción tal como un sitio poliA.

Los vectores usados para expresar polipéptidos de interés contienen usualmente elementos de control transcripcional adecuados para guiar la transcripción tales como por ejemplo promotores, potenciadores, señales de poliadenilación, señales de pausa o terminación de transcripción como elementos de un casete de expresión. Para una expresión apropiada de los polipéptidos, los elementos de control translacional adecuados se incluyen preferiblemente en el vector, tales como por ejemplo, regiones no traducidas en 5' que llevan a estructuras de tapa de 5' adecuadas para reclutar ribosomas y detener codones para terminar el proceso de traducción. El polinucleótido que sirve como gen marcador seleccionable así como el polinucleótido que codifica el polipéptido de interés son transcritos bajo el control de elementos de transcripción presente en promotores apropiados. Los transcritos resultantes de los genes marcadores seleccionables y los del polipéptido de interés acogen elementos de traducción funcionales que facilitan los niveles sustanciales de expresión de proteínas (esto es traducción) y terminación apropiada de la traducción. Una unidad de expresión funcional, capaz de guiar apropiadamente la expresión de un polinucleótido incorporado también se denomina como "casete de expresión" aquí. Preferiblemente, comprende una región UTR en 3'.

De acuerdo con lo anterior, se proveen ácidos nucleicos vector en donde los casetes de expresión comprenden al menos un elemento promotor y/o promotor/potenciador. Aunque las fronteras físicas entre estos dos elementos de control no están siempre claras, el término "promotor" se refiere usualmente a un sitio sobre la molécula de ácido nucleico al cual la ARN polimerasa y/o cualquier factor asociado se enlaza y en el cual se inicia la transcripción. Los potenciadores potencian la actividad del promotor, temporalmente así como espacialmente. Muchos promotores son activos transcripcionalmente en un amplio rango de tipos de células. Los promotores pueden ser divididos en dos clases, aquellos que funcionan constitutivamente y aquellos que son regulados por inducción o desrepresión. Los promotores usados para producción a alto nivel de proteínas en células de mamíferos deberían ser fuertes y preferiblemente activos en un amplio rango de tipos de células para permitir una evaluación cualitativa y cuantitativa del polipéptido recombinante. El promotor puede ser seleccionado del grupo consistente de un promotor SV40, un promotor CMV, un promotor EF1alfa, un promotor RSV, un promotor BROAD3, un promotor rosa 26 murínico, un promotor pCEFL y un promotor β -actina. Promotores constitutivos fuertes que guían la expresión en muchos tipos de células incluyen pero no se limitan a un promotor tardío principal de adenovirus, el promotor temprano inmediato de citomegalovirus humano, el promotor SV40 y del virus de sarcoma Rous, y el promotor murínico de 3-fosfoglicerato quinasa y EF1a. Se alcanzan buenos resultados con el vector de expresión de la presente invención cuando el promotor y/o potenciador se obtiene también de CMV y/o SV40.

Los casetes de expresión para expresar los polipéptidos de interés comprenden un promotor y/o potenciador más fuerte que los casetes de expresión para expresar los marcadores seleccionables. Esta disposición tiene el efecto de que se genera más transcritos del polipéptido de interés que para los marcadores de selección. Es ventajoso que la producción del polipéptido de interés que es secretado sea dominante sobre la producción de los marcadores de selección, puesto que la capacidad celular individual para producir proteínas heterólogas no es ilimitada y debería así

enfocarse al polipéptido de interés.

De acuerdo con una realización, los casetes de expresión (POI) y (POI') (si están presentes) los cuales son utilizados para expresar el polipéptido de interés comprenden un promotor/potenciador CMV. Se describen ejemplos específicos en detalle más adelante. Los casetes de expresión (MSM) y (MASM), que expresan el gen DHFR y preferiblemente el gen marcador de neomicina, comprenden un promotor SV40 o un promotor/potenciador SV40. El promotor CMV es conocido por ser uno de los promotores más fuertes disponibles para la expresión en mamíferos y lleva a una tasa de expresión muy buena. Se considera que da significativamente más transcritos que el promotor SV40.

Adicionalmente, los casetes de expresión pueden comprender un sitio de terminación de transcripción apropiado. Este, puesto que la transcripción continuada de un promotor corriente arriba a través de una segunda unidad de transcripción puede inhibir la función del promotor corriente abajo, es un fenómeno conocido como oclusión de promotor o interferencia transcripcional. Este evento ha sido descrito tanto en procariotas como en eucariotas. La colocación apropiada de señales de terminación transcripcionales entre dos unidades de transcripción puede evitar la oclusión del promotor. Los sitios de terminación de la transcripción están bien caracterizados y su incorporación en vectores de expresión ha demostrado tener múltiples efectos beneficiosos sobre la expresión genética.

La mayoría de los ARNm nacientes eucariotas poseen una cola poliA en su extremo 3' que se agrega durante un proceso complejo que involucra la escisión del transcrito primario y una reacción de poliadenilación acoplada. La cola de poliA es ventajosa para la estabilidad y transferabilidad del ARNm. Por lo tanto, los casetes de expresión del vector de acuerdo con la presente invención comprenden un sitio de poliadenilación. Hay varias señales de poliA eficientes que pueden ser utilizadas en vectores de expresión en mamíferos, incluyendo los derivados de hormona de crecimiento bovino (bgh), beta-globina de ratón, la unidad de transcripción temprana SV40 y el gen de timidina quinasa del virus de Herpes simplex. Sin embargo, también se conocen sitios de poliadenilación sintéticos (véase, por ejemplo, el vector de expresión pCI-neo de Promega el cual está basado en Levitt et al, 1989, Genes Dev. 3, (7): 1019-1025). El sitio de poliadenilación puede ser seleccionado a partir del grupo consistente del sitio SV40 poliA, tal como el SV40 tardío y el sitio poliA temprano (véase, por ejemplo, el plásmido pSV2-DHFR tal como se describe en Subramani et al, 1981, Mol. Cell. Biol. 854-864), un sitio poliA sintético (véase, por ejemplo, el vector de expresión pCI-neo de Promega que está basado en Levitt et al, 1989, Genes Dev. 3, (7): 1019-1025) y un sitio poliA bgh (hormona de crecimiento bovino).

Los casetes de expresión para expresar el polipéptido de interés comprenden un intrón. La mayoría de los genes de eucariotas altas contienen intrones que son retirados durante el procesamiento del ARN. Se encuentra, que los constructos genómicos son expresados más eficientemente en sistemas transgénicos que constructos idénticos que carezcan de intrones. Usualmente, los intrones son colocados en el extremo 5' del marco de lectura abierto. De acuerdo con lo anterior, un intrón comprende los casetes de expresión para expresar los polipéptidos de interés con el fin de incrementar la tasa de expresión. Dicho intrón puede estar localizado entre el elemento promotor y/o promotor/potenciador, en el extremo 5' del marco de lectura abierto del polipéptido que se va a expresar. Por lo tanto, se provee un ácido nucleico vector, en donde el casete de expresión (POI) comprende un intrón que está dispuesto entre el promotor y el codón de inicio del polinucleótido para expresar el polipéptido de interés. Varios intrones adecuados son conocidos en el estado del arte y que pueden ser utilizados en conjunción con la presente invención.

De acuerdo con una realización, el intrón usado en los casetes de expresión para expresar los polipéptidos de interés, es un intrón sintético tal como el intrón SIS o RK. El intrón RK es un intrón sintético fuerte que se coloca preferiblemente en el codón de inicio ATG del gen de interés. El intrón RK consiste del sitio de empalme del donante de intrón del promotor CMV y el sitio de empalme del receptor de la región variable de la cadena pesada de IgG de ratón (véase, por ejemplo Eaton et al., 1986, Biochemistry 25, 8343-8347, Neuberger et al., 1983, EMBO J 2(8), 1373-1378; puede ser obtenido a partir del vector pRK-5 (BD Pharmingen)).

Adicionalmente, se encontró sorprendentemente que la colocación de un intrón en el extremo 3' del marco de lectura abierto del gen DHFR tiene efectos ventajosos sobre la tasa de expresión/amplificación del constructo. El intrón usado en el casete de expresión DHFR lleva a una variante más pequeña no funcional del gen DHFR (Grillari et al., 2001, J. Biotechnol. 87, 59-65). Por lo tanto el nivel de expresión del gen DHFR disminuye. Esto lleva a una sensibilidad incrementada para MTX y a condiciones de selección más restrictivas. De acuerdo con lo anterior, se provee un ácido nucleico vector, en donde el casete de expresión (MASM) comprende un intrón que está localizado en 3' del gen marcado seleccionable amplificable. Puede obtenerse un intrón adecuado a partir del vector pSV2-DHFR (véase, por ejemplo más arriba).

Dicho vector comprende al menos un casete de expresión adicional (PSM) que comprende un gen marcador seleccionable procariota. Dicho casete de expresión (PSM) está localizado entre los casetes de expresión (MSM) y (MASM). Dicho marcador seleccionable puede proveer una resistencia a antibióticos tales como por ejemplo ampicilina, kanamicina, tetraciclina y/o cloranfenicol. Dicho casete de expresión (PSM) está dispuesto preferiblemente en la misma orientación 5' a 3' que los otros casetes de expresión (POI), (MSM) y (MASM).

El gen DHFR trabaja como un marcador y un gen de amplificación de genes. Puede presentarse selección cultivando

las células en ausencia de los metabolitos apropiados (hipoxantina y timidina) evitando por tanto el crecimiento de células no transformadas. Con el sistema DHFR, la expresión de los polipéptidos de interés puede ser incrementada exponiendo las células a antifolatos tales como el metotrexato por ejemplo (MTX) un fármaco que bloquea la actividad de la DHFR. Después de un cierto tiempo de exposición al metotrexato la mayoría de las células mueren, pero un pequeño número de células usualmente sobreviven para sobreproducir la DHFR. Con el tratamiento con metotrexato, con rampas hacia arriba en concentración, las células sobrevivientes pueden contener frecuentemente desde varios cientos hasta unos pocos miles de copias del vector integrado incrustado en cromosomas que son elongados frecuentemente. Células amplificadas con la mayor coincidencia producen más proteína recombinante que las células no amplificadas.

- 10 Varias enzimas DHFR adecuadas y genes concordantes son conocidos en el arte anterior y pueden ser utilizados en conjunción con la presente invención. La DHFR puede ser una DHFR tipo silvestre o una variante funcional o un derivado de la misma. El término una "variante" o "derivado" incluye enzimas DHFR que tienen uno o más intercambios en la secuencia de aminoácidos (por ejemplo, eliminaciones, sustituciones o adiciones) con respecto a la secuencia de aminoácidos de la enzima DHFR respectiva, proteínas de fusión que comprenden una enzima de DHFR o un fragmento funcional de la misma y enzimas de DHFR que han sido modificadas para proveer una estructura y/o función adicionales, así como fragmentos funcionales de las anteriores, los cuales tienen todavía al menos una función de una enzima de DHFR. El gen DHFR seleccionado preferiblemente del grupo consistente de DHFR silvestre, una variante DHFR que tiene una sensibilidad reducida al MTX en comparación con la DHFR tipo silvestre y una variante DHFR que tiene una sensibilidad a MTX potenciada en comparación con la DHFR tipo silvestre. De acuerdo con una realización, el gen de DHFR es el gen de DHFR tipo silvestre. De acuerdo con una realización diferente, el gen de DHFR codifica una variante menos funcional de DHFR. Esto lleva a una sensibilidad incrementada por MTX y a condiciones de selección más restrictivas. Esto puede lograrse por ejemplo, colocando un intrón en el extremo 3' del gen DHFR (véase más arriba). Una alternativa diferente podría basarse en el uso de un mutante/variante de DHFR que tiene una sensibilidad mayor hacia MTX que la DHFR tipo silvestre. Estas realizaciones son particularmente beneficiosas en el caso de que el vector sea usado en células huésped que son DHFR.

En el caso en que se desee utilizar el sistema DHFR en células huésped incorporando una copia del gen DHFR en su propio genoma (por ejemplo CHO DHFR⁺), se prefiere utilizar una mutante/variante del gen DHFR, que sea menos sensible hacia antifolatos tales como el MTX que la DHFR tipo silvestre y así hasta cierto grado antifolato resistente a MTX. Por ejemplo, debido a mutaciones en el gen la sensibilidad del gen de DHFR hacia MTX puede ser reducida significativamente de tal manera que dicha variante pueda ser utilizada a concentraciones de antifolato (MTX) más altas. Dicha variante de DHFR antifolato y en particular sensible a MTX puede poseer por ejemplo una afinidad de enlazamiento antifolato/MTX reducida. Pueden utilizarse variantes respectivas de DHFR para "sobretitular" la expresión endógena de DHFR en líneas celulares tipo silvestre. Variantes DHFR respectivas "resistentes" o menos sensibles son bien conocidas en el arte anterior.

- 35 El vector de expresión puede ser seleccionado del grupo consistente de

(a) un ácido nucleico vector lineal que comprende los siguientes elementos genéticos en la disposición indicada, en donde la dirección 5' a 3' es indicada por la →:

I. Promotor del casete de expresión (MASM) (→)

II. Gen que codifica el marcador seleccionable y amplificable de mamífero del casete de expresión MASM (→)

- 40 III. Intrón del casete de expresión (MASM) (→)

IV. Sitio poliA del casete de expresión (MASM) (→)

V. Promotor del casete de expresión (POI) (→)

VI. Intrón del casete de expresión (POI) (→)

VII. Polinucleótido que codifica un polipéptido de interés, el cual es insertado en el casete de expresión (POI) (→)

- 45 VIII. Sitio PoliA del casete de expresión (POI) (→)

IX. Promotor del casete de expresión (POI') (→)

X. Intrón del casete de expresión (POI') (→)

XI. Polinucleótido que codifica un polipéptido adicional de interés, el cual está insertado en el casete de expresión

(POI') (→)

XII. Sitio poliA del casete de expresión (POI') (→)

XIII. Promotor del casete de expresión (MSM) (→)

XIV. Gen que codifica el marcador seleccionable de mamíferos del casete de expresión (MSM) (→)

5 XV. Sitio poliA del casete de expresión (MSM) (→)

XVI. Casete de expresión PSM (→) o (←),

(b) un ácido nucleico vector como se muestra en Seq ID No. 1 o Seq ID No. 16.

La variante (a) en su forma circular se muestra en la figura 1 y en la Tabla 1 correspondiente.

10 El vector de acuerdo con la presente invención puede ser obtenido disponiendo los casetes de expresión en el orden y orientación apropiados como se describe en más detalle más arriba. La disposición de los casetes de expresión/elementos genéticos puede hacerse utilizando enzimas de restricción y estrategias de clonación adecuadas con el fin de ensamblar el vector de expresión.

15 De acuerdo con, lo anterior, se provee también un método para producir un ácido nucleico vector de acuerdo con al menos una de las reivindicaciones 1 a 6, en donde dicho método comprende disponer en un vector circular al menos los siguientes elementos genéticos

(a) al menos un casete de expresión (POI) para expresar un polipéptido de interés, en donde dicho casete de expresión comprende al menos un elemento promotor o promotor/potenciador;

20 (b) un casete de expresión (MSM) que comprende un gen marcador seleccionable de mamífero, en donde dicho gen marcador seleccionable de mamífero es un gen de resistencia a un antibiótico y en donde dicho casete de expresión comprende al menos un elemento promotor o promotor/potenciador;

(c) un casete de expresión (MASM) que comprende un gen marcador seleccionable amplificable de mamífero, en donde dicho casete de expresión comprende al menos un elemento promotor o promotor/potenciador;

(d) un casete de expresión (PSM) que comprende un gen marcador seleccionable procariota,

25 de tal forma que el casete de expresión (POI) está flanqueada en 5' por el casete de expresión (MASM), el casete de expresión (MSM) está localizado en 3' desde el casete de expresión (POI), el casete de expresión (MASM) está dispuesto en 3' desde el casete de expresión (MSM) y el casete de expresión (PSM) está localizado entre los casetes de expresión (MSM) y (MASM) y en donde los casetes de expresión (MASM), (POI) y (MSM) están dispuestos en la misma orientación 5' a 3' en donde dicho vector circular comprende un sitio de restricción de linealización único para
30 linealizar el vector el cual está localizado entre los casetes de expresión (PSM) y (MASM) y linealizar el vector circular a través de dicho sitio de restricción de linealización único para proveer un vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 6.

35 El vector de acuerdo con la presente invención puede ser utilizado para expresar el polipéptido de interés en muchas células huésped de mamíferos diferentes. El vector de expresión de la presente invención se integra usualmente en y se mantiene en el genoma. Hay dos formatos principales de células huésped, cultivos de células adherentes y cultivos en suspensión. Los cultivos de suspensión son los preferidos. Las líneas celulares más establecidas mantienen su carácter dependiente del anclaje a menos que se ejecuten esfuerzos especiales para adaptarlos al crecimiento en suspensión. Las formulaciones de medios comercialmente disponibles facilitan la transición. Básicamente cualquier célula huésped de mamífero puede ser utilizada en conjunción con la presente invención en tanto permita la expresión de un polipéptido. Células huésped de mamíferos adecuadas para los propósitos de la presente invención incluyen
40 pero no se limitan a células derivadas de ratones (por ejemplo COP, L, C127, Sp2/0, NS-0, NS-1, At20 o NIH3T3), ratas (PC12, PC12h, GH3, MtT), hámster (por ejemplo BHK, CHO y CHO defectuoso en el gen DHFR), monos (por ejemplo, COS1, COS3, COS7, CV1 y Vero) y humanos (por ejemplo Hela, HEK-293, PER-C6 derivador retina, células derivadas de fibroblastos diploides, células de mieloma y HepG2). Preferiblemente, la célula huésped es una célula CHO. La construcción del vector de acuerdo con la presente invención es particularmente adecuada para producir
45 polipéptidos en células de roedores tales como CHO y células CHO defectuosas en el gen DHFR.

También se proveen células huésped de mamíferos, que comprenden el vector de expresión linealizado de acuerdo con la presente invención el cual está integrado de manera estable en el genoma de la célula huésped de mamífero.

También se divulga una línea celular estable, que comprende un vector de expresión de acuerdo con la presente invención o un segmento de la misma en el genoma. El segmento comprenderá al menos los casetes de expresión decisivos para la presente invención. Puesto que el vector y sus características así como las células huésped adecuadas están descritas en detalle más arriba, nos referimos a la divulgación anterior. De acuerdo con lo anterior, también se provee un método para producir una célula huésped como se describió más arriba, en donde la célula huésped es transfectada de manera estable con el ácido nucleico vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 6.

Hay varios métodos apropiados conocidos en el arte anterior para introducir un vector de expresión en una célula huésped de mamífero. Los métodos respectivos incluyen pero no se limitan a transfección con fosfato de calcio, electroporación, lipofección, transferencia de genes mediada por procedimientos biolísticos y polímeros. Más arriba se describen las células huésped adecuadas.

Después de la introducción del ácido nucleico vector de expresión en las células huésped, los transformantes obtenidos son cultivados bajo condiciones selectivas adecuadas para ensayar la expresión del gen marcador seleccionable de mamífero incluido en el casete de expresión (MSM). Esto significa, que por ejemplo cuando el gen marcador seleccionable de mamífero es un gen con resistencia a antibióticos, los transformantes se cultivan en un medio que contiene el antibiótico correspondiente activo en células de mamífero y se seleccionan los transformantes que son viables bajo tales condiciones, permitiendo así la obtención de transformantes que expresan el gen marcador y ser así incorporados en el vector. Adicionalmente, puede llevarse a cabo una segunda etapa de selección cultivando los transformantes en un medio de selección adaptado para seleccionar el gen marcador seleccionable, amplificable, comprendido en el casete de expresión (MASM). De acuerdo con la invención se utiliza DHFR como gen marcador amplificable, seleccionable, y los transformantes pueden ser cultivados en un medio libre de nucleótidos o de purina en la presencia de un inhibidor de DHFR.

En caso que se utilice un promotor inducible en al menos un casete de expresión, se debería proveer una señal de inducción correspondiente con el fin de iniciar la expresión del polipéptido.

Con el fin de hacer uso del sistema de selección/amplificación de DHFR, dichas células huésped pueden ser cultivadas en la presencia de un inhibidor de DHFR. Inhibidores de DHFR adecuados son antifolatos tales como, por ejemplo, el MTX. La concentración de antifolato/MTX usada depende de la célula huésped y de la variante DHFR incorporada en el vector. El rango de concentración puede ser seleccionado para procedimientos de amplificación en etapas múltiples en células huésped de DHFR⁻ por ejemplo en valores alrededor de 5 nM-20 nM que varían hasta valores de 500 nM a 1000 nM o incluso mayores para etapas secundarias o posteriores de amplificación. Para células DHFR⁺ las concentraciones de inicio están usualmente en el rango de 100 nM a 750 nM, preferiblemente 500 nM en las primeras etapas y de 500 nM a 1000 nM y más para etapas de amplificación posteriores. Se describen más arriba variantes de DHFR adecuadas.

Con el vector de expresión de acuerdo con la presente invención, pueden expresarse/producirse varios polipéptidos diferentes de interés. El término polipéptido se refiere a una molécula que comprende un polímero de aminoácidos enlazados entre sí mediante enlaces peptídicos. Los polipéptidos incluyen polipéptidos de cualquier longitud, incluyendo proteínas (por ejemplo que tienen más de 50 aminoácidos) y péptidos (por ejemplo 2-49 aminoácidos). Los polipéptidos incluyen proteínas y/o péptidos de cualquier actividad o bioactividad, incluyendo por ejemplo polipéptidos bioactivos tales como proteínas o péptidos enzimáticos (por ejemplo, proteasas, quinasas, fosfatasa), proteínas o péptidos receptores, proteínas o péptidos transportadores, proteínas bactericidas y/o de enlazamiento a endotoxinas, proteínas o péptidos estructurales, polipéptidos inmunes, toxinas, antibióticos, hormonas, factores de crecimiento, vacunas o similares. Dicho polipéptido puede ser seleccionado del grupo consistente de hormonas peptídicas, interleucinas, activadores de plasminógeno de tejidos, citoquinas, inmunoglobulinas, en particular anticuerpos o fragmentos de anticuerpos o variantes de los mismos. Dicha inmunoglobulina puede ser de cualquier isotipo. Muy frecuentemente, se producen/necesitan moléculas de IgG (por ejemplo IgG1) como proteínas terapéuticas. Un fragmento de anticuerpos es cualquier fragmento de un anticuerpo que comprende al menos 20 aminoácidos de dicho anticuerpo entero, preferiblemente al menos 100 aminoácidos, el cual al menos tiene todavía una capacidad de enlazamiento al antígeno. El fragmento de anticuerpo puede comprender la región de enlazamiento del anticuerpo tal como un fragmento Fab, un fragmento F(ab)₂, multicuerpos que comprenden múltiples dominios de enlazamiento tales como diacuerpos, triacuerpos o tetracuerpos, anticuerpos o aficuerpos de dominio sencillo. Una variante de anticuerpos es un derivado de un anticuerpo o fragmento de anticuerpo que tiene la misma función de enlazamiento pero, por ejemplo, una secuencia de aminoácidos alterada. Dicho anticuerpo y/o fragmento de anticuerpo puede comprender una cadena liviana murínica, cadena liviana humana, cadena liviana humanizada, cadena pesada humana y/o cadena pesada murínica así como fragmentos activos o derivados de los mismos. Por lo tanto, puede ser por ejemplo murínico, humano, quimérico o humanizado.

La presente invención también provee métodos para producir un polipéptido de interés, comprendiendo dicho método cultivar al menos una célula huésped de acuerdo con la reivindicación 8 en un medio de cultivo celular bajo condiciones que permiten la expresión de dicho polipéptido de interés.

En una etapa siguiente, dicho polipéptido puede ser aislado/recolectado a partir del cultivo celular. El polipéptido expresado de interés puede ser obtenido destruyendo las células huésped. Los polipéptidos también pueden ser expresados, por ejemplo secretados en el medio de cultivo y pueden ser obtenidos a partir del mismo. También son posibles combinaciones de los métodos respectivos. Por lo tanto, pueden producirse productos, en particular polipéptidos y obtenerse/aislarse eficientemente con alto rendimiento. El polipéptido obtenido también puede ser sujeto de etapas posteriores de procesamiento tales como por ejemplo etapas de purificación y/o modificación con el fin de producir el producto de interés en la calidad deseada.

De acuerdo con una alternativa, dicho polipéptido de interés es secretado en el medio de cultivo celular y subsecuentemente aislado a partir del medio de cultivo celular. El polipéptido es preferiblemente una molécula de inmunoglobulina tal como un anticuerpo o un fragmento funcional del mismo. Con el fin de promover la secreción del polipéptido de interés puede utilizarse una secuencia líder. Preferiblemente, se utiliza la secuencia líder de una molécula de inmunoglobulina.

El vector de expresión de acuerdo con la presente invención así como células huésped y polipéptidos adecuados de interés se describen en detalle más arriba; nos referimos a la divulgación anterior.

El método para producir el polipéptido de interés puede comprender al menos una de las siguientes etapas:

- aislar el polipéptido de interés a partir de dicho medio de cultivo celular y/o a partir de dicha célula huésped; y/o
- procesar el polipéptido aislado de interés.

El polipéptido de interés producido de acuerdo con la invención puede ser recuperado, purificado adicionalmente, aislado y/o modificado por métodos conocidos en el arte. Por ejemplo, el producto puede ser recuperado a partir del medio nutritivo por procedimientos convencionales que incluyen, pero no se limitan a, centrifugación, filtración, ultrafiltración, extracción o precipitación. La purificación puede llevarse a cabo mediante una variedad de procedimientos conocidos en el arte incluyendo, pero no limitándose a, cromatografía (por ejemplo, intercambio de iones, afinidad, hidrófoba, cromatografía de exclusión por tamaño), procedimientos electroforéticos (por ejemplo, enfoque isoeléctrico preparativo), solubilidad diferencial (por ejemplo, precipitación con sulfato de amonio) o extracción.

De acuerdo con una realización que es ventajosa en particular para la producción de proteínas/péptidos farmacéuticos, la célula huésped es cultivada en suspensión bajo condiciones libres de suero. El polipéptido obtenido puede ser purificado posteriormente, por ejemplo, purificando el polipéptido presente en el cultivo celular sobrenadante utilizando métodos cromatográficos (por ejemplo, purificación por afinidad).

Los polipéptidos producidos de acuerdo con el método de la presente invención representan buenas propiedades de estabilidad. Los resultados también muestran que los polipéptidos son expresados en forma funcional y por lo tanto en la conformación correcta. De acuerdo con lo anterior, la invención también provee polipéptidos obtenidos por el método de producción de acuerdo con la presente invención usando el vector de expresión descrito en detalle más arriba. Como se delinea más arriba, los polipéptidos se obtienen con un buen rendimiento. El polipéptido es preferiblemente una molécula de inmunoglobulina tal como un anticuerpo o un fragmento funcional del mismo.

La figura 1 muestra un ácido nucleico vector circular.

La figura 2 muestra la versión linealizada del ácido nucleico vector de acuerdo con la figura 1 con el fin de demostrar la influencia de la posición del sitio de restricción de linealización.

Los números 1 a 17 mostrados en las figuras 1 y 2 indican los elementos/características genéticas del ácido nucleico vector y se describen en detalle en la subsiguiente Tabla 1. Si no se define otra cosa, las flechas blancas caracterizan elementos promotores y/o promotores/potenciadores; las cajas rayadas caracterizan elementos de intrón, las flechas negras simbolizan los polinucleótidos para expresar el polipéptido de interés; las cajas con cuadrículas caracterizan el sitio poliA; las flechas con cuadrículas los genes marcadores de mamíferos en el casete de expresión (MSM) y la flecha punteada el gen marcador seleccionable procariota; como puede verse, todos los elementos genéticos están dispuestos en la misma orientación 5' a 3' (indicada por la dirección de la flecha). Los ácidos nucleicos vectores pBW147, pBW154 y pBW160, que se describen en más detalle más adelante, se construyen respectivamente.

Tabla 1. Orientación y disposición de los elementos genéticos de acuerdo con un ácido nucleico vector como se muestra en las figuras 1 y 2

ES 2 585 502 T3

Numeración en figura 1 y 2	Elemento genético
1	Promotor del casete de expresión (POI) (diagramado por la flecha blanca). Es por ejemplo un promotor/potenciador de CMV.
2	Intrón del casete de expresión (POI) (diagramado por la caja rayada). Es por ejemplo un intrón RK como se describe más arriba.
3	Polinucleótido que codifica un polipéptido de interés, el cual es insertado en el casete de expresión (POI) (diagramado por la flecha negra). De acuerdo con la realización mostrada, es la cadena liviana de un anticuerpo monoclonal (mAB-LC).
4	Sitio poliA del casete de expresión (POI) (diagramado por la caja cuadrículada). Es por ejemplo un sitio poliA SV40.
5	Promotor del casete de expresión (POI') (diagramado por la flecha blanca). Es por ejemplo un promotor/potenciador de CMV.
6	Intrón del casete de expresión (POI') (diagramado por la caja rayada). Es por ejemplo un intrón RK, como se describió más arriba.
7	Polinucleótido que codifica un polipéptido adicional de interés, el cual es insertado en el casete de expresión (POI') (diagramado por la flecha negra). De acuerdo con la realización mostrada, es la cadena pesada de un anticuerpo monoclonal (mAB-HC).
8	Sitio poliA del casete de expresión (POI') (diagramado por la caja cuadrículada). Es por ejemplo un sitio poliA SV40.
9	Promotor del casete de expresión (MSM) (diagramado por la flecha blanca). Es por ejemplo un promotor/potenciador SV40.
10	Gen que codifica el marcador seleccionable de mamífero del casete de expresión (MSM) (diagramado por la flecha cuadrículada). Es por ejemplo el gen neo.
11	Sitio poliA del casete de expresión (MSM) (diagramado por la barra). Es por ejemplo un sitio poliA sintético como se describió más arriba.
12	Casete de expresión PSM (diagramado por la flecha rayada). Puede comprender por ejemplo un gen marcador seleccionable procariona que provee una resistencia contra ampicilina.
13	Sitio de restricción de linealización (diagramado por la barra). Dicho sitio es preferiblemente un sitio de restricción único.
14	Promotor del casete de expresión (MASM) (diagramado por la flecha blanca). Es por ejemplo un promotor SV40.
15	Gen que codifica el marcador seleccionable amplificable de mamífero del casete de expresión (MASM) (diagramado por la flecha cuadrículada). Es por ejemplo el gen DHFR.
16	Intrón en el casete de expresión (MASM) (diagramado por la barra). Este intrón está presente opcionalmente.
17	Sitio poliA del casete de expresión (MASM) (diagramado por la caja cuadrículada). Es por ejemplo un sitio poliA SV40.

Subsecuentemente, se dan ejemplos adecuados para los elementos de vector descritos, los cuales, sin embargo, no son limitantes.

5 Se utiliza un marcador DHFR seleccionable amplificable de mamífero. Un ejemplo adecuado de un polinucleótido de DHFR de ratón tipo silvestre se provee con Seq ID No. 5 el cual se utiliza preferiblemente con las células huésped de DHFR⁻. Una forma mutante adecuada de DHFR se provee con Seq ID No. 6. Una forma mutante respectiva se utiliza preferiblemente junto con las células DHFR⁺. La Seq ID No. 12 muestra un mutante de DHFR que incluye un intrón adecuado para incrementar adicionalmente la presión de selección (véase más arriba). Pueden utilizarse también variantes o fragmentos funcionales de los anteriores.

10 Se prefiere el uso de neo como gen marcador seleccionable de mamífero. Se provee una secuencia adecuada con Seq ID No. 7. Pueden usarse también variantes o fragmentos funcionales.

Como una secuencia promotora para guiar la expresión del polipéptido de interés se prefiere el uso de un promotor de CMV. Se provee una secuencia adecuada con Seq ID No. 8. También pueden utilizarse variantes o fragmentos funcionales de la misma.

15 Se prefiere como secuencia promotora para guiar la expresión de los genes marcadores seleccionables MSM y MASM el uso de un promotor SV40. Se provee una secuencia adecuada con Seq ID No. 9. También pueden usarse variantes o fragmentos funcionales de la misma.

Puede utilizarse como secuencia de poliadenilación para los polipéptidos de interés y/o el MASM un sitio poliA SV40. Una secuencia adecuada se muestra como Seq ID No. 10. También pueden utilizarse variantes o fragmentos funcionales de las mismas o una orientación invertida (sitio poliA de SV40 tardío o temprano).

20 Puede utilizarse una secuencia de intrón de un intrón RK para el casete de expresión (POI) que codifica el polipéptido de interés. Se provee una secuencia adecuada con Seq ID No. 11. Pueden utilizarse también variantes o fragmentos funcionales de la misma.

25 Como sitio de poliadenilación sintético que puede utilizarse, por ejemplo, en conjunción con el marcador seleccionable de mamífero (MSM), se muestra como Seq ID No. 13. Pueden usarse también variantes o fragmentos funcionales de la misma.

Un marcador seleccionable (PSM) bacteriano adecuado es por ejemplo el gen de beta-lactamasa que provee resistencia a la ampicilina. Se provee una secuencia adecuada con Seq ID No. 14. Pueden utilizarse también variantes o fragmentos funcionales de la misma.

30 Adicionalmente, el ácido nucleico vector puede comprender al menos un sitio de clonación múltiple (MCS) para insertar por ejemplo un polinucleótido que codifica un polipéptido de interés. Puede proveerse el MCS en 3' y 5' del polinucleótido que codifica el polipéptido de interés. Se proveen sitios MCS adecuados como Seq ID No. 4 (localizada preferiblemente en el sitio/región de 5') y Seq ID No. 15 (localizada preferiblemente en el sitio/región 3'). Estos sitios MCS pueden ser utilizados por ejemplo con el fin de introducir el polinucleótido que codifica el polipéptido de interés.

35 Ácidos nucleicos vectores preferidos particularmente se muestran como Seq ID No. 1 (que comprende el gen DHFR tipo silvestre, dicho vector es particularmente útil para el sistema DHFR⁻) y Seq ID No. 16 (que comprende un gen DHFR mutante, dicho vector es particularmente útil para el sistema DHFR⁺).

Ejemplos

La presente invención se describe ahora por medio de ejemplos no limitantes, los cuales sin embargo, constituyen realizaciones preferidas de la presente invención.

40 I. Métodos de cultivo celular y transfección

Subsecuentemente, se describen métodos apropiados para transfectar y cultivar las células huésped de acuerdo con la presente invención para expresar un polipéptido de interés, por medio de ejemplos.

Ejemplo 1: Cultivo celular

45 Se cultivan células CHO en un medio CHO adecuado tal como por ejemplo ExCell81134 (obtenida de SAFC Biosciences). Las células son replicadas 2-3 veces por semana en medio fresco y se mantiene en fase de crecimiento logarítmico a lo largo del estudio.

Ejemplo 2: Estrategia de transfección

Para transfección, se usan células CHO progenitoras en fase de cultivo exponencial con una viabilidad superior al 90%. Las transfecciones por lipofección se efectúan utilizando el reactivo DMRIE-C de acuerdo con las instrucciones del fabricante (Invitrogen). La cantidad de células es ajustada a 1×10^6 células en medio OptiMEM I (Invitrogen). Para la lipofección, se mezclan entre sí 2 μg o 4 μg del vector de expresión y 4 μl del reactivo DMRIE-C durante 28 minutos a temperatura ambiente y se agregan a las células durante 4 horas a 37°C . Las células son diluidas entonces a 2×10^5 células/ml de medio de cultivo e incubadas durante 2 días a 37°C y CO_2 al 5%.

Ejemplo 3. Selección de neomicina y amplificación de gen

El marcador de selección de neomicina localizado en el ácido nucleico vector de expresión permite la selección de resistencia G418. Para la selección de transfectantes, se cultivan células en la presencia de 0.8 mg/ml de G418 (Invitrogen) durante aproximadamente dos semanas. Dos semanas después de la transfección y selección de G418, surgen poblaciones reunidas que consisten predominantemente de células resistentes G418. Las células son cultivadas entonces en la ausencia de nucleótidos durante aproximadamente dos semanas. Se inicia entonces la amplificación genética mediante la adición de MTX 20 nM al medio de cultivo. Después de tres semanas de cultivo, se genera una reserva celular heterogénea amplificada. El marcador de selección/amplificación de DHFR (dihidrofolato reductasa) permite la amplificación del gen de DHFR así como del transgen agregando el análogo de ácido fólico metotrexato (MTX) al medio de cultivo, dando como resultado títulos incrementados para las reservas de transfección. Después de la recuperación de la crisis, las reservas experimentan una segunda y una tercera etapa de amplificación utilizando concentraciones más altas de MTX cada una durante aproximadamente dos semanas (MTX 100 nM y 500 nM). En cada etapa las células son congeladas después de la recuperación de la reserva.

Ejemplo 4. Establecimiento de líneas celulares clonales

Para obtener una línea celular clonal (esto es una línea celular derivada de una célula individual), la reserva de células transfectadas de manera estable puede ser diluida y sembrada en placas de 96 pozos con una densidad celular de 0.3-0.5 células por pozo (dilución limitada). Las células que forman una colonia distinguible son escaladas utilizando procedimientos estándar. Eventualmente, se evalúan clones individuales para la expresión de polipéptidos recombinantes, reteniendo los productores más altos después del cultivo y análisis. A partir de estos candidatos, se escogen líneas celulares con características de crecimiento y productividad apropiadas para la producción de la proteína recombinante. La productividad puede ser mejorada usualmente de manera adicional estableciendo/adaptando las condiciones del cultivo esto es agregando aditivos tales como peptonas.

II. Construcciones de vector

Son factibles varios ensamblajes de vector de acuerdo con las enseñanzas de la presente invención. Puesto que los elementos individuales del vector son conocidos en el arte anterior, pueden ensamblarse vectores adecuados por ejemplo secuenciando o amplificando y clonando de manera apropiada los elementos genéticos básicos y los casetes de expresión en la orientación deseada. Los métodos de clonación respectivos están en el estado del arte y también la secuencia de elementos genéticos descrita más arriba esta descrita en el arte anterior. Subsecuentemente, se describe a manera de ejemplo la generación de varios constructos de vector. Sin embargo, debe entenderse por los expertos en el arte que son adecuadas y fácilmente obtenibles otras varias realizaciones y maneras de obtener vectores respectivos.

Con el fin de facilitar el entendimiento de la disposición de los constructos de vector descritos aquí a manera de ejemplo y sus vectores precursores, las Tablas 1 y 2 proveen una visión de los elementos genéticos principales comprendidos en estos vectores, su orden y orientación. Desde luego, solamente se muestran los elementos principales, los vectores sin embargo, pueden comprender elementos adicionales genéticos o secuencias esqueleto. Cada columna en la tabla representa un constructo de vector. Desde la fila superior a la fila inferior los elementos genéticos de los casetes de expresión están listados en el orden de su disposición y su orientación sobre el vector. Puesto que los vectores descritos son circulares, el elemento mostrado en la última fila de cada columna es en efecto adyacente al elemento genético mostrado en la primera fila (desde luego, pueden estar presentes secuencias de esqueleto). La orientación de cada elemento genético está indicada por las flechas. La cabeza de la flecha apunta a la dirección 3' del elemento genético respectivo.

ES 2 585 502 T3

Tabla 2 - Orden de los elementos genéticos en los vectores precursores

pBW133 10416bp	pBW139 9213bp	pBW146 9247bp	pBW159 11122bp
CMVprom/pot →	CMVprom/pot →	CMVprom/pot, comprende un sitio de restricción <i>SpeI</i> (153) →	CMVprom/pot →
RK-intrón →	RK-intrón →	RK-intrón →	RK-intrón →
mAB-LC →	mAB-LC →	mAB-LC →	mAB-LC →
MCS I	MCS	MCS	MCS
poliA de SV40 →	poliA de SV40 →	poliA de SV40 →	poliA de SV40 →
Sitio de restricción <i>DraIII</i> (2365)	CMVprom/pot →	CMVprom/pot comprende un sitio de restricción <i>SpeI</i> (2519) →	CMVprom/pot →
DHFR* porta un sitio de restricción <i>Scal</i> (2870); orientación inversa ←	RK-intrón →	RK-intrón →	RK-intrón →
SV40prom/pot; orientación inversa ←	MCS	MCS2	mAB-HC →
Sitio de restricción <i>DraIII</i> (3561)	mAB-HC →	mAB-HC →	MCS
SV40prom/pot	MCS	MCS	Promotor T3
Neo →	poliA de SV40 →	Promotor T3	poliA de SV40 →
CMVprom/pot →	Región fago f1 que comprende un sitio de restricción <i>DraIII</i> (5456) →	poliA de SV40 →	Región fago f1 →
RK-intrón →	SV40prom/pot →	Región fago f1 →	Prom/pot SV40 →
MCS2	Neo →	Prom/potSV40 que comprende un origen mínimo de replicación de SV40 →	Neo →
mAB-HC →	SV40poliA →	Neo →	PoliA sintético
SV40poliA →	Amp que comprende un sitio de restricción <i>Scal</i> (7828) →	poliA sintético	Amp →
Amp que porta <i>Scal</i> (9031) →	Sitio de restricción <i>BglII</i> (9209), 5' al CMV prom/pot	Amp →	3 Sitios pA
		Sitio de restricción <i>BglII</i> (9243), adyacente en 5' del prom/pot de CMV	Intrón
			DHFR; orientación inversa ←

pBW133 10416bp	pBW139 9213bp	pBW146 9247bp	pBW159 11122bp
			promSV40; orientación inversa ←
			Sitio de restricción SwaI (11113)

Tabla 3 – Orden de los elementos genéticos en los vectores de expresión

pBW147 11053bp	pBW154 11109bp	pBW160 11122bp
prom/pot CMV, que comprende un sitio de restricción <i>SpeI</i> (153) →	CMVprom/pot →	CMVprom/pot →
RK-intrón →	RK-intrón →	RK-intrón →
mAB-LC →	mAB-LC →	mAB-LC →
PoliA de SV40 →	PoliA de SV40 →	PoliA de SV40 →
prom/pot de CMV, que comprende un sitio de restricción <i>SpeI</i> (2519) →	CMVprom/pot →	CMVprom/pot →
RK-intrón →	RK-intrón →	RK-intrón →
mAB-HC →	mAB-HC →	mAB-HC →
PoliA de SV40 →	PoliA de SV40 →	PoliA de SV40 →
SV40 prom/pot →	SV40 prom/pot →	SV40 prom/pot →
Neo →	Neo →	Neo →
poliA sintético	poliA sintético	poliA sintético
Amp →	Amp →	Amp →
Swa I sitio de restricción (9288)	Swa I sitio de restricción (9243)	Swa I sitio de restricción (9256)
SV40 prom/pot →	SV40 prom →	SV40 prom →
DHFR*→	DHFR →	DHFR →
Sitio Bgh pA →	Intrón	Intrón
	poliA de SV40	poliA de SV40

5 Las abreviaturas en las tablas anteriores 1 a 3 y en las figuras 1 y 2 tienen significados regulares como son evidentes para la persona de experiencia en el arte y como se describió más arriba, y tienen en particular los siguientes significados:

MCS = sitio de clonación múltiple

mAB-HC = cadena pesada de anticuerpo monoclonal

mAB-LC = cadena liviana de anticuerpo monoclonal

intrón = véase Grillari et al, 2001, 3. Biotechnol. 87, 59-65

prom/pot = promotor/potenciador

Ejemplo 5: Construcción del vector de expresión pBW147

5 En esta disposición la configuración en tándem de genes mAB y la DHFR* (variante mutante que tiene una sensibilidad más baja a MTX que la DHFR tipo silvestre) combinada con el sitio bgh pA, se prueba. El casete de DHFR* se coloca en 5' antes del casete de expresión (POI), que comprende la mAB-LC, de tal manera que todos los marcos de lectura abiertos son colocados en una dirección de lectura. El ensamblaje de pBW147 se muestra en la Tabla 3.

10 La pBW147 puede construirse a partir de pBW133 (por favor hágase referencia también a la Tabla 2). La construcción de pBW147 se describe aquí.

Construcción de pBW133

15 Las construcciones de vector descritas más abajo se basan en el vector de expresión pCI-neo disponible comercialmente (Promega Corporation, Estados Unidos). La secuencia de ADN completa está disponible públicamente (número de acceso GenBank/EMBL: U47120). Se introduce en pCIneo un sitio de clonación múltiple nuevo.

Las dos cadenas del sitio de clonación múltiples son sintetizadas de *ново*. El pCIneo es cortado con *NheI* y *XmaI*. El MCS antiguo es eliminado por electroforesis en gel. El sitio de clonación múltiple nuevo es sintetizado de manera que los nucleótidos 4 terminales del extremo 5' de la cadena antisentido y el extremo 3' terminal de la cadena superior de ADN no son sintetizados. Después de fusionar ambas cadenas se crean extremos compatibles para *NheI* y *XmaI*.

20 La secuencia del nuevo sitio de clonación múltiple es como sigue (véase secuencia Seq. ID No. 2):

<i>ApalSgrAI</i>				
<i>AgeI</i>	<i>PmeI</i>	<i>EcoRV</i>	<i>PshAI</i>	
<i>EcoO109I</i>	<i>BstEII</i>	<i>PmlI</i>	<i>BspEI</i>	<i>Ascl</i>

25 El plásmido resultante de la ligazón de pCIneo con MCS se denomina pCI-neo-2 para propósitos de descripción. El pCI-neo-2 es modificado adicionalmente por la introducción del intrón pRK de pRK5 (BD PharMingen). Por lo tanto, el pCI-neo-2 es digerido con *Apal*. Se crean extremos romos mediante tratamiento con T4 polimerasa. Luego el plásmido es dirigido con *NdeI*. Se digiere la pRK5 con *NdeI* y *NruI* (cortador de extremos romos). El intrón RK que contiene el fragmento es aislado y ligado con el esqueleto pCIneo2. El plásmido resultante es pCIneo2RK.

Para llevar ambos casetes de expresión a un vector, puede obtenerse el vector pCIneoDHFR*-RK. El pCIneoDHFR*-RK es obtenido como sigue:

30 El casete de expresión DHFR* es amplificado por PCR a partir del vector pCHI-LC (Simulect SP2/0 vector de expresión de cadena ligera). Los cebadores son BB35 (GGGCACTACGTGCCGCGGATTTAAATGCGGCCGCATATGGTGCAC - Seq. ID No. 3) y BB36 (GGGCACGTAGTGTATTAGGGGAGCAGAGAACTTGAA - Seq. ID No. 4).

35 El fragmento de PCR es clonado en pCIneoRK a través de digestión con restricción *DraIII* dando el vector pCIneoDHFR*-RK. pCIneoDHFR*-RK es abierta por digestión con *EcoO109I*. Con el fin de crear extremos romos, se lleva a cabo posteriormente tratamiento con la enzima Klenow.

40 El casete de expresión de pCIneo2RK es escindido digiriendo el plásmido con *BglII*, *NgoMIV* y *StuI*. Después de la ligación de los dos fragmentos se crea el vector de expresión pCHO2neoN "vacío". Se inserta en el pCHO2neoN el gen de anticuerpo de cadena liviana a través del sitio de restricción *MluI* y *Sall* creando por lo tanto un constructo de vector, al cual nos referimos como pBW108. El gen del anticuerpo es amplificado entonces utilizando cebadores que contienen los dos sitios de restricción.

La cadena pesada mAB es insertada en pBW108 a través de digestión en *PmeI* y *Ascl* del vector, obteniéndose por tanto un constructo de vector considerado como pBW111 para propósitos de descripción. La cadena pesada es amplificada por PCR con el extremo 5' generando extremos romos y el extremo 3' del gen conteniendo el sitio de

restricción *Ascl*.

En pBW111 la región 5' no traducida de la cadena liviana es intercambiada puesto que un codón adicional ATG está presente en el frente de la cadena de ADNc ligera. Esto se logra seccionando un fragmento *Bgl*II / *Mlu*I de pBW111 y reemplazándolo con el fragmento corregido. El nuevo plásmido se considera como pBW133. El pBW133 es el primer vector con todos los genes sobre un plásmido. La disposición de los genes es: LC-DHFR (dirección opuesta) – neo-HC (véase también Tabla 2). Este vector es uno de los materiales de partida que pueden ser utilizados para obtener vectores de acuerdo con las enseñanzas de la presente invención. Sin embargo, es claro que hay varias otras formas para obtener vectores respectivos.

Construcción de pBW139

- 10 El segundo constructo de vector que puede ser utilizado para obtener pBW147 tiene la configuración de pBW139. pBW139 puede crearse a partir de pBW115. Para la construcción de pBW115, el gen de cadena pesada es clonado en pCneoRK (véase más arriba). Sin embargo, el pCneoRK se digiere con *Mlu*I y *Nru*I (cortadores de extremo romo), mientras que el fragmento de PCR de cadena pesada es digerido con *Ascl* (3') (compatible con *Mlu*I) y es romo en 5'. El plásmido resultante se considera como pBW115.
- 15 EL pBW115 es digerido con *Scal* y *Bgl*III. Luego se hace un llenado con Klenow para clonar extremos romos. El casete de expresión de cadena liviana es escindido del pBW133 con *Scal* y *Dra*III. Con el fin de crear extremos romos, se hace un tratamiento con ADN-T4 polimerasa. El resultado de la ligación es un vector que tiene una configuración como pBW139 (véase Tabla 2).

Ensamblaje de pBW147

- 20 Para obtener pBW147, se digiere pBW133 con *Spe*I, *Xho*I. Las partes que contienen el fragmento del promotor CMV y la primera parte de la cadena pesada se aíslan y se ligan a pBW139, el cual también es cortado con *Spe*I, *Xho*I. En el vector resultante el casete de cadena pesada se encuentra sin un codón de ATG adicional de ruptura. Con el fin de recuperar la cadena liviana hacia el vector, el pBW139 se digiere con *Spe*I. El fragmento que contiene LC es insertado en pBW148 el cual es abierto con *Spe*I. El plásmido resultante tiene una configuración como pBW146 (véase Tabla 2).

- 25 En pBW146 se inserta el gen DHFR de pBW112 (vector de expresión de otro proyecto). Sin embargo, el gen DHFR también podría ser obtenido a partir de una fuente diferente, dependiendo de la clase de variante de DHFR deseada. El pBW146 es digerido con *Bgl*III. Después de esto el casete de DHFR es amplificado por PCR con cebadores que contienen sitios de restricción *Bgl*III y *Bam*HI. El fragmento de PCR es digerido con las dos enzimas e insertado en el sitio de ajuste *Bgl*III de pBW146. El plásmido resultante tiene una configuración como pBW147, en donde todos los casetes de expresión tienen la misma orientación. La estructura se muestra en la Tabla 3.

Este vector de expresión puede ser utilizado para obtener transfecciones estables. Con el fin de incrementar adicionalmente el rendimiento de la expresión, puede seleccionarse una variante de DHFR tipo silvestre muy "sensible" a MTX, en donde también las concentraciones de MTX deberían ser adaptadas adecuadamente.

35 Ejemplo 6. Construcción del vector de expresión pMW154

- Para este vector, se prueba la configuración en tándem de los genes mAB y el casete del gen DHFR de pSV2DHFR (versión tipo silvestre de DHFR con alta sensibilidad a MTX). El casete de expresión de DHFR de pSV2DHFR (ATCC#374146) es amplificado por PCR. El fragmento contenía el promotor y los sitios poliA. Como antes, los oligos tenían sitios de restricción *Bgl*III / *Bam*HI. El casete de expresión de DHFR es insertado en el sitio de restricción *Bgl*III de pBW146, dando como resultado un constructo de vector que tiene la misma estructura que pBW154. La estructura de pBW154 se muestra en la Tabla 3 y puede ser derivada de la figura 1 y figura 2 las cuales muestran un ejemplo general de un constructo de vector que tiene una estructura/configuración global respectiva de los elementos genéticos. La secuencia de pBW154 se provee como Seq. ID No 1. El polinucleótido de cadena liviana es marcado como n (en la solicitud de prioridad indicada con la referencia V), el polinucleótido de cadena pesada esta marcado como n (en la solicitud de prioridad indicado con la referencia Y) en Seq. ID No. 1. Las características de pBW154 se resumen en la Tabla 4. Desde luego, también pueden utilizarse otros elementos de vector, por ejemplo promotores diferentes, potenciadores diferentes, sitios poliA diferentes y otros elementos tales como oris. Además, es posible conmutar los casetes de expresión para la cadena liviana y la pesada de la molécula de inmunoglobulina. Sin embargo, la selección y la disposición mostradas de los elementos del vector son las preferidas. Como se delineó más arriba, también pueden utilizarse fragmentos funcionales de moléculas de inmunoglobulina. Por lo tanto, el "nnn" indicado solamente sirve para propósitos de ilustración y no indican ninguna restricción de tamaño puesto que pueden estar presentes secuencias de inmunoglobulina más pequeñas o más grandes en la posición correspondiente. Con el fin de facilitar la comparación con las figuras 1 y 2, las cuales muestran la construcción general de los vectores de acuerdo con la realización preferida de la invención, hemos indicado la numeración de los elementos correspondientes en las

figuras 1 y 2.

Tabla 4

Par de bases de inicio	Final	Característica	Numeración correspondiente en figura 1 y 2
1	743	Prom/pot de CMV	1
857	1000	RK-intrón	2
1054	1766	mAB-LC	3
1815	2036	poliA de SV40	4
2367	3109	Prom/pot de CMV	5
3223	3366	RK-intrón	6
3452	4863	mAB-HC	7
4931	5152	poliA de SV40	8
5766	6184	Prom de SV40	9
6229	7024	Neomicina fosfotransferasa	10
7087	7135	PoliA sintético	11
7546	8406	Gen con resistencia a antibiótico por beta lactamasa	12
9243		Sitio de linealización único	13
9422	9617	SV40prom	14
9640	10204	DHFR	15
9776	10426	Intrón (donante-receptor)	16
10909	11098	PoliA de SV40	17

5 Todos los elementos genéticos están dispuestos en la misma orientación 5' a 3'. Con este constructo de vector utilizando un DHFR tipo silvestre, la amplificación de genes como se describió más arriba trabaja muy eficientemente. El título en un experimento por los lotes estándar puede ser incrementado en 10-20 veces. Aquí, se observa un gran incremento en el título de expresión del anticuerpo por amplificación con MTX. Partiendo de un tratamiento G418, a lo largo del tratamiento sin nucleótidos hasta de varias concentraciones diferentes con MTX (MTX 20 y 100 nM), el título se incrementa constante y considerablemente. Sin embargo, utilizando concentraciones de MTX mucho más altas (por ejemplo, MTX 500 nM) no se obtienen usualmente ventajas adicionales con células CHO, aunque pueden utilizarse altas concentraciones respectivas. Los títulos de anticuerpos obtenidos para reservas a lo largo del proceso de selección/amplificación variaron desde 2 a más de 60 mg/L cuando se utilizan procedimientos de cultivo estándar. Los títulos de expresión pueden ser incrementados adicionalmente al establecer líneas celulares clonales a partir de las reservas utilizando medios personalizados para potenciar la expresión celular puesto que el título obtenible también depende del medio usado.

Ejemplo 7: Construcción del vector de expresión pBW160

Los experimentos también demuestran que la orientación del gen DHFR en el vector es decisiva. En pBW146 (véase más arriba) está presente un sitio de restricción *EcoRI*. Con el fin de tener el *EcoRI* como un cortador individual presente en el vector de expresión final, el sitio puede ser retirado digiriendo el pBW146 con *EcoRI*, Klenow-Fill y

relegación. Esto da como resultado un plásmido que tiene la configuración como pBW158 (no mostrado). El casete de DHFR puede ser integrado en pBW158, tal como se describe más arriba. Puesto que ambas orientaciones (orientación como se muestra en pBW159 y pBW160, véase Tabla 2 y 3) son generadas automáticamente, ambas pueden ser probadas en cuanto a los niveles de expresión. Nuestros resultados muestran un rendimiento superior de la configuración con todos los marcos de lectura abiertos orientados en una dirección de lectura 5' a 3'.

Los vectores que tienen una configuración tal como el vector pBW159 (véase Tabla 2), en donde la orientación de DHFR es en orden inverso de los genes mAB usualmente mostraron solo títulos de expresión muy bajos, incluso después de amplificación con MTX (y usualmente menores de 1 mg/L). Los vectores que tienen un diseño tal como pBW160 (véase Tabla 3), en donde la orientación de DHFR está en marco con los genes mAB, puede proveer títulos de anticuerpos más altos de más de 5 mg/L e incluso más de 10 o incluso más de 15 mg/L (obtenibles a partir de las reservas). De nuevo, al establecer una línea celular clonal y utilizando un medio de alto rendimiento el rendimiento en título puede incrementarse adicionalmente cuando se utilizan los constructos de vector de acuerdo con la presente invención.

Estos experimentos pueden demostrar la ventaja de la "configuración en marco" de los marcadores de selección y los genes de codificación mAB se utilizan de acuerdo con las enseñanzas de la presente invención. Este resultado soporta el hallazgo de que la orientación 5' a 3' de los elementos del vector es un factor importante para los vectores de expresión alta. Adicionalmente, la estabilidad de la expresión es muy favorable con los vectores de expresión de acuerdo con la presente invención.

La configuración del vector de acuerdo con la presente invención permite la generación en avance recto de reservas de células con altas productividades específicas de células. Los elementos claves son la orientación 5' a 3', las variantes DHFR seleccionadas y la colocación del marcador de selección DHFR sobre el vector así como la disposición de los genes de los anticuerpos y el segundo marcador de selección (neo) sobre el plásmido. El vector también puede ser utilizado para la producción de proteínas o péptidos que no son anticuerpos. Como se describió más arriba, con ligeras adaptaciones del casete de DHFR, este sistema también es utilizable para la amplificación genética en una línea celular CHO-K1 PD positiva a DHFR. Para la amplificación genética en células huésped DHFR⁺ se utiliza una versión mutada del gen DHFR (véase más arriba). El casete de expresión de DHFR completo de un vector que comprende una versión mutada del gen DHFR tal como pBW117 puede ser amplificada por PCR con cebadores que incorporan un sitio *Bam*HI. Este fragmento es clonado entonces en el sitio *Bg*/II del pBW158 dando como resultado el vector pBW165. Con vectores que tienen una configuración tal como pBW165 que comprenden un gen de DHFR mutado (y seguimientos con otros anticuerpos), pueden generarse líneas celulares de alto rendimiento en la célula huésped CHO-K1-PD, la cual es una línea celular DHFR⁺. Un ejemplo de una secuencia de vector respectiva se provee como Seq ID No. 16. Desde luego pueden utilizarse también diferentes elementos de vector que los mostrados, por ejemplo, promotores diferentes, potenciadores diferentes, sitios poliA diferentes y/o otros elementos tales como orís diferentes. Además es posible conmutar los casetes de expresión para la cadena liviana y la pesada de la molécula de inmunoglobulina. Sin embargo, la selección y disposición mostradas de los elementos de vector son las preferidas. Como se delineó más arriba, las moléculas de inmunoglobulina de longitud completa así como los fragmentos funcionales de las moléculas de inmunoglobulina pueden expresarse a partir del vector. En Seq ID No. 16 solamente el sitio está indicado como el sitio de inserción, en donde la secuencia respectiva de inmunoglobulina puede ser localizada/insertada en el vector de expresión final. Cualquier secuencia de inmunoglobulina puede estar presente en la posición correspondiente. Además, como se delineó más arriba, también es posible expresar diferentes polipéptidos de interés.

Ejemplo 8: Producción a escala pequeña de anticuerpos con células CHO transfectadas

Para prueba en cultivos en suspensión, las células se siembran a 1×10^5 células/ml en 50 ml de medio ExCell81134 (SAFC Biosciences) en un matraz de cultivo con tapa de filtro de fondo redondo de 250 ml. Las células son agitadas a 65 rpm en un incubador Kühner Shaker ISF-4-W a 37°C en un ambiente de CO₂ al 10% durante la duración del estudio. Se dan suministros individuales con soluciones de diseño particular de acuerdo con un esquema de suministro fijo que comienza en el 4° día de la expansión celular. En el día 13, se recolectan muestras de 1 ml y el título es medido utilizando HPLC y una columna A para proteína.

El cultivo celular sobrenadante resultante de los cultivos en los matraces con agitación de los mejores clones se purifica por cromatografía de afinidad con proteína A.

Ejemplo 9: Purificación de proteína A del anticuerpo expresado

Para la purificación de la proteína A se cargan aproximadamente 27 mL de sobrenadante de cultivo libre de células que contiene aproximadamente 32.4 mg de anticuerpo sobre una columna de afinidad MabSelect de 0.5x10 cm. Después de la carga, la columna se enjuaga y lava suficientemente. Luego el anticuerpo es eluido a pH 3-4. El eluido es analizado con HPLC estándar utilizando una columna para proteína A. Se obtuvieron aproximadamente 30.5 mg de anticuerpo.

III. Ejemplos para productividades y rendimientos específicos de las células

Los clones que son seleccionados después de la expansión clonal son probados en cuanto a su productividad.

Ejemplo 10: Expresión de un anticuerpo IgG1

- 5 Se expresó un anticuerpo IgG1. Los clones son cultivados en el medio comercialmente disponible ExCell81134 (SAFC Biosciences). Se agregan soluciones de suministro y aditivos convencionales tales como peptona. Pueden obtenerse ratas de alta productividad cuando se utiliza el vector de acuerdo con la presente invención:

Clon	Qp (pg/célula/día)
1	114
2	91
3	103
Qp = productividad específica de la célula	

Ejemplo 11: Expresión de un anticuerpo IgG1 y un anticuerpo IgG4

- 10 Un anticuerpo IgG1 y un anticuerpo IgG4 son expresados. Los clones se cultivan en un medio de cultivo apropiado. Se agregan soluciones de suministro y aditivos convencionales tales como peptona. Pueden obtenerse ratas de alta productividad cuando se utiliza el vector de acuerdo con la presente invención:

Clon	Anticuerpo IgG1 Qp (pg/célula/día)	Anticuerpo IgG4 Qp (pg/célula/día)
1	76	
2	96	
3		73
Qp = Productividad específica de célula.		

Ejemplo 12: Producción a gran escala de polipéptidos con células CHO transfectadas

- 15 La producción de polipéptidos a gran escala puede hacerse por ejemplo en biorreactores de onda, vidrio o acero inoxidable. Para este propósito las células son expandidas, partiendo usualmente de un vial congelado individual, por ejemplo un vial de un Master Cell Bank. Las células son descongeladas y expandidas a través de varias etapas. Los biorreactores de diferente escala son inoculados con cantidades apropiadas de células. La densidad celular puede incrementarse agregando soluciones de suministro y aditivos al biorreactor. Las células se mantienen a una alta viabilidad durante un tiempo prolongado. Se alcanzaron en la escala grande títulos de producto en el reactor que
- 20 varían desde unos pocos cientos de miligramos por litro hasta varios gramos por litro. La purificación puede hacerse por metodologías de cromatografía estándar, las cuales pueden incluir etapas de cromatografía de afinidad, intercambio de iones, interacción hidrófoba o exclusión por tamaño. El tamaño del biorreactor puede ser hasta de varios miles de litros de volumen en escala final (véanse también F. Wurm, Nature Biotechnology Vol. 22, 11, 2004, 1393-1398).

- 25 **Ejemplo 13:** Estrategia de clonación para introducir nuevos genes de anticuerpos en los vectores

Una estrategia -entre otras- para insertar nuevos polipéptidos de interés es como sigue (explicada a manera de ejemplo utilizando un vector que tiene una configuración tal como pBW154):

Clonación del gen de cadena liviana

El gen de cadena liviana puede ser amplificado por PCR con cebadores introduciendo un sitio en 5' *MluI* del codón

ATG y un sitio *Sa*/I en 3' del gen. El producto de PCR es introducido en pBW154 a través de estas dos enzimas de restricción. Esto lleva a un vector intermedio que consiste de solo la cadena liviana.

Clonación del gen de cadena pesada

5 El gen de cadena pesada puede ser amplificado por PCR con cebadores introduciendo un extremo romo para uso del sitio en 5' *Nru*I y del vector en el codón ATG con un sitio *Xba*I en 3' del gen. El producto de PCR es introducido en pBW154 a través de estas dos enzimas de restricción. Esto lleva a un vector intermedio con la antigua cadena liviana y la nueva cadena pesada.

Ensamblaje del vector final

10 El nuevo fragmento que contiene mAB-HC es escindido del vector HC a través de digestión con *Sa*/I. Este es insertado en el vector intermedio LC a través de *Sa*/I para dar como resultado el nuevo vector final LC-HC.

Listado de secuencias

<110> Novartis AG

<120> Compuestos orgánicos

15 <130> 52411-WO-PCT

<150> 07150339.5-2401

<151> 2007-12-21

20 <160> 16

<170> PatentIn versión 3.5

25 <210> 1

<211> 11109

<212> ADN

<213> Artificial

30 <220>

<223> secuencia de vector pBW154

<220>

<221> características_misceláneas

35 <222> (1)..(743)

<223> CMV prom//pot

<220>

<221> características_misceláneas

40 <222> (825)..(1024)

<223> RK-intrón

<220>

<221> características_misceláneas

45 <222> (857)..(1000)

<223> RK-región de intrón que es empalmada en la expresión

<220>

<221> características_misceláneas

50 <222> (1054)..(1767)

<223> mAB LC (indicada originalmente en Seq. Id. No: 1 con el símbolo "v")

<220>

<221> características_misceláneas

55 <222> (1815)..(2036)

<223> SV40poliA

<220>

<221> características_misceláneas

<222> (2367)..(3109)
 <223> CMV prom//pot

 <220>
 5 <221> características_misceláneas
 <222> (3191)..(3390)

 <223> RK-intrón

 <220>
 <221> características_misceláneas
 10 <222> (3223)..(3366)
 <223> RK-intrón que es empalmado en la expresión

 <220>
 <221> características_misceláneas
 15 <222> (3452)..(4864)
 <223> mAB HC (indicada originalmente en Seq. Id. No: 1 con el símbolo "y")

 <220>
 <221> características_misceláneas
 20 <222> (4931)..(5152)
 <223> SV40 poliA

 <220>
 <221> características_misceláneas
 25 <222> (5247)..(5702)
 <223> región f1

 <220>
 <221> características_misceláneas
 30 <222> (5766)..(6184)
 <223> SV40 prom

 <220>
 <221> características_misceláneas
 35 <222> (6229)..(7023)
 <223> Gen de neomicina fosfotransferasa

 <220>
 <221> características_misceláneas
 40 <222> (7087)..(7135)
 <223> PoliA sintético

 <220>
 <221> características_misceláneas
 45 <222> (7546)..(8406)
 <223> Gen de beta lactamasa resistente a antibiótico

 <220>
 <221> características_misceláneas
 50 <222> (9243)..(9243)
 <223> sitio de linealización único

 <220>
 <221> características_misceláneas
 55 <222> (9276)..(9623)
 <223> SV40prom

 <220>
 <221> características_misceláneas
 60 <222> (9521)..(9568)
 <223> SV40 origen mínimo de replicación

 <220>

<221> características_miscláneas
 <222> (9640)..(10203)
 <223> Gen DHFR

5 <220>
 <221> características_miscláneas
 <222> (9776)..(10426)
 <223> Intrón (donante-aceptor))

10 <220>
 <221> características_miscláneas
 <222> (10910)..(11097)
 <223> SV40poliA

15 <400> 1

tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta	60
ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc	120
aatatgaccg ccatgttggc attgattatt gactagttaa taatagtaat caattacggg	180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc	240
gcctggctga ccgccaacg acccccgccc attgacgtca ataatgacgt atgttcccat	300
agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatattac ggtaaactgc	360
ccacttggca gtacatcaag tgtatcatat gccaaagtac cccctattg acgtcaatga	420
cggtaaattg cccgcctggc attatgcccga gtacatgacc ttatgggact ttcctacttg	480
gcagtacatc tacgtattag tcatcgctat taccatggtg atgcgggtttt ggcagtacat	540
caatgggctg ggatagcggg ttgactcacg gggatttcca agtctccacc ccattgacgt	600
caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactc	660
cgccccattg acgcaaattg gcggtaggcg tgtacgggtg gaggtctata taagcagagc	720
tcgttttagtg aaccgtcaga tcgcctggag acgccatcca cgctgttttg acctccatag	780
aagacaccgg gaccgatcca gcctccgcgg ccgggaacgg tgcattggaa cgcggattcc	840
ccgtgccaag agtgacgtaa gtaccgccta tagagtctat agggccaccc ccttggttc	900
gttagaacgc ggctacaatt aatacataac cttatgtatc atacacatac gatttaggtg	960
acactataga ataacatcca ctttgccttt ctctccacag gtgtccactc ccagggtccaa	1020
ctgcacctcg gttctatcga aaacgcgtcc accnnnnnnn nnnnnnnnnn nnnnnnnnnn	1080
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1140
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1200
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1320
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1380
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1440
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1500
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1560
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1620
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1680
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1740
nnnnnnnnnn nnnnnnnnnn nnnnnnnngtc gaccggg gcg gccgcttccc tttagtgagg	1800

gttaatgctt	cgagcagaca	tgataagata	cattgatgag	tttggacaaa	ccacaactag	1860
aatgcagtga	aaaaaatgct	ttatttgtga	aatttgtgat	gctattgctt	tatttghtaac	1920
cattataagc	tgcaataaac	aagttaacaa	caacaattgc	attcatttta	tgtttcaggt	1980
tcagggggag	atgtgggagg	tttttttaaag	caagtaaaac	ctctacaaat	gtggtaaaat	2040
ccgataagga	tcgatccggg	ctggcgtaat	agcgaagagg	cccgcaccga	tcgcccttcc	2100
caacagttgc	gcagcctgaa	tggcgaatgg	acgcgccctg	tagcggcgca	ttaagcgcg	2160
cgggtgtggt	ggttacgcgc	agcgtgaccg	ctacacttgc	cagcgcccta	gcgcccgcctc	2220
ctttcgcttt	cttcccttcc	tttctcgcca	cgttcgccgg	ctttccccgt	caagctctaa	2280
atcgggggct	cccttttaggg	ttccgattta	gagctttacg	gcacctcgac	cgcaaaaaac	2340
ttgatttggg	tgatggttca	cgatcttcaa	tattggccat	tagccatatt	attcattggt	2400
tatatagcat	aaatcaatat	tggctattgg	ccattgcata	cgttgtatct	atatcataat	2460
atgtacattt	atattggctc	atgtccaata	tgaccgccat	gttggcattg	attattgact	2520
agttattaat	agtaatcaat	tacgggggtca	ttagttcata	gcccataatat	ggagttccgc	2580
gttacataac	ttacggtaaa	tggcccgcct	ggctgaccgc	ccaacgaccc	ccgcccattg	2640
acgtcaataa	tgacgtatgt	tcccatagta	acgccaatag	ggactttcca	ttgacgtcaa	2700
tgggtggagt	atttacggta	aactgcccac	ttggcagtac	atcaagtgtg	tcatatgccca	2760
agtaccccc	ctattgacgt	caatgacggg	aaatggcccc	cctggcatta	tgcccagtac	2820
atgaccttat	gggactttcc	tacttggcag	tacatctacg	tattagtcac	cgctattacc	2880
atgggtgatg	ggttttggca	gtacatcaat	gggcgtggat	agcggtttga	ctcacggggga	2940
tttccaagtc	tccaccccat	tgacgtcaat	gggagtttgt	tttggcacca	aaatcaacgg	3000
gactttccaa	aatgtcgtaa	caactccgcc	ccattgacgc	aaatgggagg	taggcgtgta	3060
cgggtgggagg	tctatataag	cagagctcgt	ttagtgaacc	gtcagatcgc	ctggagacgc	3120
catccacgct	gttttgacct	ccatagaaga	caccgggacc	gatccagcct	ccgcggccgg	3180
gaacggtgca	ttggaacgcg	gattccccgt	gccaaagagt	acgtaagtac	cgcttataga	3240
gtctataggc	ccacccccct	ggcttcgtta	gaacgcggct	acaattaata	cataacctta	3300
tgtatcatat	acatacgatt	taggtgacac	tatagaataa	catccacttt	gcctttctct	3360
ccacaggtgt	ccactcccag	gtccaactgc	acctcggttc	tatcgcgatt	gaattccccg	3420
gggatcctct	aggggtgaccg	tttgtgccac	cnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3660

nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	3720
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	3780
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	3840
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	3900
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	3960
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4020
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4080
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4140
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4200
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4260
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4320
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4380
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4440
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4500
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4560
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4620
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4680
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4740
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4800
nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	nnnnnnnnnnnn	4860
nnnnnggcgcg	tggtacctct	agagtcgacc	cgggcgggccg	cttcctttta	gtgaggggta	4920
atgcttcgag	cagacatgat	aagatacatt	gatgagtttg	gacaaaccac	aactagaatg	4980
cagtgaaaaa	aatgctttat	ttgtgaaatt	tgtgatgcta	ttgctttatt	tgtaaccatt	5040
ataagctgca	ataaacaagt	taacaacaac	aattgcattc	attttatgtt	tcaggttcag	5100
ggggagatgt	gggagggttt	ttaaagcaag	taaaacctct	acaaatgtgg	taaaatccga	5160
taaggatcga	tccgggctgg	cgtaatagcg	aagaggcccg	caccgatcgc	ccttcccaac	5220
agttgcgag	cctgaatggc	gaatggacgc	gccctgtagc	ggcgcattaa	gcgcggcggg	5280
tgtgggtggt	acgcgcagcg	tgaccgctac	acttgccagc	gccctagcgc	ccgctccttt	5340
cgctttcttc	ccttcctttc	tcgccacggt	cgccggcttt	ccccgtcaag	ctctaaatcg	5400
ggggctccct	ttaggggttc	gatttagagc	tttacggcac	ctcgaccgca	aaaaacttga	5460
tttgggtgat	ggttcacgta	gtgggccatc	gccctgatag	acggtttttc	gccctttgac	5520
gttgaggtcc	acgttcttta	atagtggact	cttgttccaa	actggaacaa	cactcaaccc	5580

tatctcgggc	tattcttttg	atttataagg	gattttgccc	atttcggcct	attggttaaa	5640
aatgagctg	atttaacaaa	tatttaacgc	gaattttaac	aaaatattaa	cgtttacaat	5700
ttcgctgat	gcggtatttt	ctccttacgc	atctgtgcgg	tatttcacac	cgcatacgcg	5760
gatctgcga	gcaccatggc	ctgaaataac	ctctgaaaga	ggaacttggc	taggtacctt	5820
ctgaggcggg	aagaaccagc	tgtggaatgt	gtgtcagtta	gggtgtggaa	agtccccagg	5880
ctccccagca	ggcagaagta	tgcaaagcat	gcattctcaat	tagtcagcaa	ccagggtgtg	5940
aaagtcccca	ggctccccag	caggcagaag	tatgcaaagc	atgcatctca	attagtcagc	6000
aaccatagtc	ccgccccctaa	ctccgcccct	ctccgcccct	actccgcccc	gttccgcccc	6060
ttctccgccc	catggctgac	taattttttt	tatttatgca	gaggccgagg	ccgcctcggc	6120
ctctgagcta	ttccagaagt	agtgaggagg	cttttttggg	ggcctaggct	tttgcaaaaa	6180
gcttgattct	tctgacacaa	cagtctcgaa	cttaaggcta	gagccaccat	gattgaacaa	6240
gatggattgc	acgcagggtc	tccggccgct	tgggtggaga	ggctattcgg	ctatgactgg	6300
gcacaacaga	caatcggtcg	ctctgatgcc	gccgtgttcc	ggctgtcagc	gcaggggccc	6360
ccggttcttt	ttgtcaagac	cgacctgtcc	gggtgccctga	atgaactgca	ggacgaggca	6420
gcgcggctat	cgtggctggc	cacgacgggc	gttccttgcg	cagctgtgct	cgacgttgtc	6480
actgaagcgg	gaagggactg	gctgctattg	ggcgaagtgc	cggggcagga	tctcctgtca	6540
tctcaccttg	ctcctgccga	gaaagtatcc	atcatggctg	atgcaatgcg	gcggctgcat	6600
acgcttgatc	cggctacctg	cccattcgac	caccaagcga	aacatcgcat	cgagcgagca	6660
cgtactcgga	tggaaagccg	tcttgtcgat	caggatgatc	tggacgaaga	gcattcaggg	6720
ctcgcgccag	ccgaactgtt	cgccaggctc	aaggcgcgca	tgcccagcgg	cgaggatctc	6780
gtcgtgaccc	atggcgatgc	ctgcttgccg	aatatcatgg	tggaaaatgg	ccgcttttct	6840
ggattcatcg	actgtggccg	gctgggtgtg	gcggaccgct	atcaggacat	agcgttggct	6900
acccgtgata	ttgctgaaga	gcttggcggc	gaatgggctg	accgcttcct	cgtgctttac	6960
ggatatcgcc	ctcccgattc	gcagcgcata	gccttctata	gccttcttga	cgagttcttc	7020
tgagcgggac	tctgggggtt	gaaatgaccg	accaagcgac	gccaacctg	ccattcacgat	7080
ggccgcaata	aaatatcttt	attttcatta	catctgtgtg	ttggtttttt	gtgtgaatcg	7140
atagcgataa	ggatccgcgt	atgggtgact	ctcagtacaa	tctgctctga	tgccgcatag	7200
ttaagccagc	cccgacaccc	gccaacaccc	gctgacgcgc	cctgacgggc	ttgtctgctc	7260
ccggcatccg	cttacagaca	agctgtgacc	gtctccggga	gctgcatgtg	tcagagggtt	7320
tcaccgtcat	caccgaaacg	cgcgagacga	aagggcctcg	tgatacgcct	atttttatag	7380
gttaatgtca	tgataataat	ggtttcttag	acgtcagggt	gcacttttcg	gggaaatgtg	7440

cgcggaaccc	ctatttgttt	atTTTTctaa	atācattcaa	ataTgtatcc	gctcatgaga	7500
caataaccct	gataaatgct	tcaataatat	tgaaaaagga	agagtatgag	tattcaacat	7560
ttccgtgtcg	cccttattcc	ctTTTTtgcg	gcattttgcc	ttcctgtttt	tgctcaccca	7620
gaaacgctgg	tgaaagtaaa	agatgctgaa	gatcagttgg	gtgcacgagt	gggttacatc	7680
gaactggatc	tcaacagcgg	taagatcctt	gagagttttc	gccccgaaga	acgttttcca	7740
atgatgagca	cttttaaaagt	tctgctatgt	ggcgcggtat	tatcccgtat	tgacgccggg	7800
caagagcaac	tcggtcgccg	catacactat	tctcagaatg	acttggttga	gtactcacca	7860
gtcacagaaa	agcatcttac	ggatggcatg	acagtaagag	aattatgcag	tgctgccata	7920
accatgagtg	ataacactgc	ggccaactta	cttctgacaa	cgatcggagg	accgaaggag	7980
ctaaccgctt	ttttgcacaa	catgggggat	catgtaactc	gccttgatcg	ttgggaaccg	8040
gagctgaatg	aagccatacc	aaacgacgag	cgtgacacca	cgatgcctgt	agcaatggca	8100
acaacgttgc	gcaaactatt	aactggcgaa	ctacttactc	tagcttcccg	gcaacaatta	8160
atagactgga	tggaggcgga	taaagttgca	ggaccacttc	tgcgctcggc	ccttccggct	8220
ggctggttta	ttgctgataa	atctggagcc	ggtgagcgtg	ggtctcgcgg	tatcattgca	8280
gcactggggc	cagatggtaa	gccctcccgt	atcgtagtta	tctacacgac	ggggagtcag	8340
gcaactatgg	atgaacgaaa	tagacagatc	gctgagatag	gtgcctcact	gattaagcat	8400
tggtaaactgt	cagaccaagt	ttactcatat	atactttaga	ttgatttaaa	acttcatttt	8460
taatttaaaa	ggatctaggt	gaagatcctt	tttgataatc	tcatgaccaa	aatcccttaa	8520
cgtgagtttt	cgttccactg	agcgtcagac	cccgtagaaa	agatcaaagg	atcttcttga	8580
gatccttttt	ttctgcgcgt	aatctgctgc	ttgcaaacaa	aaaaaccacc	gctaccagcg	8640
gtggtttggt	tgccggatca	agagctacca	actctttttc	cgaaggtaac	tggtttcagc	8700
agagcgcaga	taccaaatac	tgtccttcta	gtgtagccgt	agttaggcca	ccacttcaag	8760
aactctgtag	caccgcctac	atacctcgct	ctgctaatcc	tgttaccagt	ggctgctgcc	8820
agtggcgata	agtcgtgtct	taccgggttg	gactcaagac	gatagttacc	ggataaggcg	8880
cagcggtcgg	gctgaacggg	gggttcgtgc	acacagccca	gcttggagcg	aacgacctac	8940
accgaactga	gatacctaca	gcgtgagcta	tgagaaagcg	ccacgcttcc	cgaagggaga	9000
aaggcggaca	ggtatccggt	aagcggcagg	gtcggaacag	gagagcgcac	gagggagctt	9060
ccagggggaa	acgcctggta	tctttatagt	cctgtcgggt	ttcgccacct	ctgacttgag	9120
cgtcgatttt	tgtgatgctc	gtcagggggg	cggagcctat	ggaaaaacgc	cagcaacgcg	9180
gccttttttac	ggttcctggc	cttttgctgg	ccttttgctc	acatggctcg	acagatccat	9240
ttaaatttttc	accgtcatca	ccgaaacgcg	cgaggcagct	gtggaatgtg	tgtcagttag	9300
ggtgtggaaa	gtccccaggc	tccccagcag	gcagaagtat	gcaaagcatg	catctcaatt	9360

agtcagcaac	caggtgtgga	aagtccccag	gctccccagc	aggcagaagt	atgcaaagca	9420
tgcatctcaa	ttagtcagca	accatagtc	cgccccctaac	tccgccccatc	ccgccccctaa	9480
ctccgcccag	ttccgccc	tctccgcccc	atggctgact	aatttttttt	atttatgcag	9540
aggccgaggc	cgccctcgcc	tctgaactat	tccagaagta	gtgaggaggc	ttttttggag	9600
gcctaggctt	ttgcaaaaag	ctttatcccc	gctgccatca	tggttcgacc	attgaactgc	9660
atcgtcgcg	tgtcccaaga	tatggggatt	ggcaagaacg	gagacctacc	ctggcctccg	9720
ctcaggaacg	agttcaagta	cttccaaaga	atgaccacaa	cctcttcagt	ggaaggtaaa	9780
cagaatctgg	tgattatggg	taggaaaacc	tggttctcca	ttcctgagaa	gaatcgacct	9840
ttaaaggaca	gaattaatat	agttctcagt	agagaactca	aagaaccacc	acgaggagct	9900
cattttcttg	ccaaaagttt	ggatgatgcc	ttaagactta	ttgaacaacc	ggaattggca	9960
agtaaagtag	acatggtttg	gatagtcgga	ggcagttctg	tttaccagga	agccatgaat	10020
caaccaggcc	acctcagact	ctttgtgaca	aggatcatgc	aggaatttga	aagtgcacacg	10080
tttttcccag	aaattgattt	ggggaaatat	aaacttctcc	cagaataccc	aggcgtcctc	10140
tctgaggtcc	aggaggaaaa	aggcatcaag	tataagtttg	aagtctacga	gaagaaagac	10200
taacaggaag	atgctttcaa	gttctctgct	cccctcctaa	agctatgcat	ttttataaga	10260
ccatgggact	tttgctggct	ttagatcttt	gtgaaggaac	cttacttctg	tggtgtgaca	10320
taattggaca	aactacctac	agagatttaa	agctctaagg	taaatataaa	atttttaagt	10380
gtataatgtg	ttaaactact	gattctaatt	gtttgtgtat	tttagattcc	aacctatgga	10440
actgatgaat	gggagcagtg	gtggaatgcc	tttaatgagg	aaaacctgtt	ttgtcagaa	10500
gaaatgccat	ctagtgatga	tgaggctact	gctgactctc	aacattctac	tcctccaaaa	10560
aagaagagaa	aggtagaaga	ccccaggac	tttccttcag	aattgctaag	ttttttgagt	10620
catgctgtgt	ttagtaatag	aactcttgct	tgctttgcta	tttaccacc	aaaggaaaaa	10680
gctgcactgc	tatacaagaa	aattatggaa	aaatattctg	taacctttat	aagtaggcat	10740
aacagttata	atcataacat	actgtttttt	cttactccac	acaggcatag	agtgtctgct	10800
attaataact	atgctcaaaa	attgtgtacc	tttagctttt	taatttgtaa	aggggttaat	10860
aaggaatatt	tgatgtatag	tgccctgact	agagatcata	atcagccata	ccacatttgt	10920
agaggtttta	cttgctttta	aaaacctccc	acacctcccc	ctgaacctga	aacataaaat	10980
gaatgcaatt	gttggttgta	acttggttat	tgagcttat	aatgggttaca	aataaagcaa	11040
tagcatcaca	aatttcacaa	ataaagcatt	tttttactg	cattctagtt	gtgggttgaa	11100
ttcgatct						11109

<210> 2
 <211> 59
 <212> ADN
 <213> Artificial

5

<220>
 <223> sitio I MCS

ES 2 585 502 T3

<400> 2

ctagggccca ccggtgaccg tttaaacacg tgatatccgg acaattgtcg gcgcgccgg 59

5 <210> 3
<211> 46
<212> ADN
<213> Artificial

10 <220>
<223> Cebador

<400> 3

15 gggcactacg tgccgcggat ttaaatgcgg ccgcatatgg tgcact 46

<210> 4
<211> 38
<212> ADN
20 <213> Artificial

<220>
<223> Cebador

25 <400> 4

gggcacgtag tgttattag gggagcagag aacttgaa 38

<210> 5
30 <211> 564
<212> ADN
<213> Artificial

<220>
35 <223> DHFR tipo silvestre

<400> 5

atggttcgac cattgaactg catcgtcgcc gtgtcccaag atatggggat tggcaagaac	60
ggagacctac cctggcctcc gctcaggaac gagttcaagt acttccaaag aatgaccaca	120
acctcttcag tgggaaggtaa acagaatctg gtgattatgg gtaggaaaac ctggttctcc	180
attcctgaga agaatcgacc tttaaaggac agaattaata tagttctcag tagagaactc	240
aaagaaccac cacgaggagc tcattttctt gccaaaagtt tggatgatgc cttagactt	300
attgaacaac cggaattggc aagtaaagta gacatggttt ggatagtcgg aggcagttct	360
gtttaccagg aagccatgaa tcaaccaggc cacctcagac tctttgtgac aaggatcatg	420
caggaatttg aaagtgcac gtttttcca gaaattgatt tggggaaata taaacttctc	480
ccagaatacc caggcgtcct ctctgaggtc caggaggaaa aaggcatcaa gtataagttt	540
gaagtctacg agaagaaaga ctaa	564

40 <210> 6
<211> 564
<212> ADN
<213> Artificial

45 <220>

<223> DHFR mutante

<400> 6

atggttcgac cattgaactg catcgtcgcc gtgtcccaaa atatggggat tggcaagaac	60
ggagaccgac cctggcctcc gctcaggaac gagtccaagt acttccaaag aatgaccaca	120
acctcttcag tggaaggtaa acagaatctg gtgattatgg gtaggaaaac ctggttctcc	180
attcctgaga agaatcgacc tttaaaggac agaattaata tagttctcag tagagaactc	240
aaagaaccac cacgaggagc tcatttttctt gccaaaagtt tggatgatgc cttaaagactt	300
attgaacaac cggaattggc aagtaaagta gacatgggtt ggatagtcgg aggcagttct	360
gtttaccagg aagccatgaa tcaaccaggc cacctcagac tctttgtgac aaggatcatg	420
caggaatttg aaagtgcac gtttttccca gaaattgatt tggggaaata taaacttctc	480
ccagaatacc caggcgtcct ctctgaggtc caggaggaaa aaggcatcaa gtataagttt	540
gaagtctacg agaagaaaga ctaa	564

5 <210> 7
 <211> 795
 <212> ADN
 <213> Artificial

10 <220>
 <223> Gen de resistencia a neomicina

<400> 7

atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc	60
ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggtgtgca	120
gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg	180
caggacgagg cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg	240
ctcgacgttg tcaactgaagc gggaaggac tggtgctat tgggcgaagt gccggggcag	300
gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg	360
cggcggctgc atacgcttga tccggctacc tgccattcg accaccaagc gaaacatcgc	420
atcgagcgag cacgtactcg gatggaagcc ggtcttgctg atcaggatga tctggacgaa	480
gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg catgcccgc	540
ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat	600
ggccgctttt ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac	660
atagcgttgg ctaccctgga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc	720
ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt	780
gacgagttct tctga	795

15 <210> 8
 <211> 743
 <212> ADN
 <213> Artificial

<220>
<223> Promotor CMV

<400> 8

5

tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta	60
ttggccattg catacgttgt atctatatca taatatgtac atttatatattg gctcatgtcc	120
aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg	180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc	240
gcctggctga ccgcccacg acccccgcgc attgacgtca ataatgacgt atgttcccat	300
agtaacgcca atagggactt tccattgacg tcaatgggtg gagtattttac ggtaaactgc	360
ccacttgga gtacatcaag tgtatcatat gccaaagtac cccctattg acgtcaatga	420
cggtaaatgg ccgcctggc attatgccc gtacatgacc ttatgggact ttcctacttg	480
gcagtacatc tacgtattag tcatcgctat taccatgggtg atgcggtttt ggcagtacat	540
caatgggcgt ggatagcggc ttgactcacg gggatttcca agtctccacc ccattgacgt	600
caatgggagt ttgttttggc accaaaatca acgggacttt ccaaatgtc gtaacaactc	660
cgccccattg acgcaaattg gcggtaggcg tgtacgggtg gaggtctata taagcagagc	720
tcgttttagtg aaccgtcaga tcg	743

<210> 9
<211> 419
<212> ADN
<213> Artificial

10

<220>
<223> Promotor SV40

15 <400> 9

gcgcagcacc atggcctgaa ataacctctg aaagaggaac ttggttaggt accttctgag	60
gcggaaagaa ccagctgtgg aatgtgtgtc agttagggtg tggaaagtcc ccaggctccc	120
cagcaggcag aagtatgcaa agcatgcac tcaattagtc agcaaccagg tgtggaaagt	180
ccccaggctc ccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca	240
tagtcccgc cctaactccg cccatccgc cctaactcc gccagttcc gccattctc	300
cgccccatgg ctgactaatt ttttttattt atgcagaggc cgaggccgcc tcggcctctg	360
agctattcca gaagtagtga ggaggctttt ttggaggcct aggcttttgc aaaaagctt	419

<210> 10
<211> 222
<212> ADN

20

<213> Artificial

<220>
<223> Sitio SV40 PoliA
<400> 10

25

cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa 60
aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca 120
ataaacaagt taacaacaac aattgcattc attttatggt tcagggttcag ggggagatgt 180
gggagggtttt ttaaagcaag taaaacctct acaaatgtgg ta 222

<210> 11
<211> 200
<212> ADN
<213> Artificial

<220>
<223> intrón RK

<400> 11

ttggaacgcg gattccccgt gccaaagagt acgtaagtac cgcctataga gtctataggc 60
ccaccccctt ggcttcgcta gaacgcggct acaattaata cataacctta tgtatcatac 120
acatacgatt taggtgacac tatagaataa catccacttt gcctttctct ccacaggtgt 180
ccactcccag gtccaactgc 200

<210> 12
<211> 942
<212> ADN
<213> Artificial

<220>
<223> intrón que incluye DHFR mutante

<220>
<221> características_misceláneas
<222> (565)..(942)
<223> intrón

<400> 12

atggttcgac cattgaactg catcgtcgcc gtgtcccaa atatggggat tggcaagaac 60
ggagaccgac cctggcctcc gctcaggaac gagttcaagt acttccaaag aatgaccaca 120
acctcttcag tggaaggtaa acagaatctg gtgattatgg gtaggaaaac ctggttctcc 180
attcctgaga agaatcgacc tttaaaggac agaattaata tagttctcag tagagaactc 240

```

aaagaaccac cacgaggagc tcatttttctt gcāaaaagt tggātgatgc cttāagactt 300
attgaacaac cggaattggc aagtaaagta gacatggttt ggatagtcgg aggcagttct 360
gtttaccagg aagccatgaa tcaaccaggc cacctcagac tctttgtgac aaggatcatg 420
caggaatttg aaagtgacac gtttttccca gaaattgatt tggggāaata taaacttctc 480
ccagaatacc caggcgtcct ctctgaggtc caggaggāaa aaggcatcaa gtataagttt 540
gaagtctacg agaagāaaga ctaacaggāa gatgctttca agttctctgc tcccctccta 600
aagctatgca tttttataag accatggggg atgctcgatc ccctcgcgag ttggttcagc 660
tgctgcctga ggctggacga cctcgcgag ttctaccggc agtgcaaatc cgtcggcatc 720
caggāaacca gcagcggcta tccgcgcac catgccccg aactgcagga gtggggaggc 780
acgatggccg ctttggtccg gatctttgtg aaggaacctt acttctgtgg tgtgacataa 840
ttggacāaac tacctacaga gatttāaagc tctaaggtaa atataāaatt tttāagtgtā 900
taatgtgtta aactactgat tctaattgtt tgtgtatttt ag 942

```

<210> 13
 <211> 49
 <212> ADN
 <213> Artificial

5

<220>
 <223> Sitio PoliA sintético

10

<400> 13

aataaaatat ctttatttc attacatctg tgtgttggt tttgtgtg

49

<210> 14
 <211> 861
 <212> ADN
 <213> Artificial

15

<220>
 <223> Gen de beta lactamasa resistente a antibiótico

20

<400> 14

```

atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt ttgccttcct 60
gtttttgctc acccagāaac gctggtgāaa gtāāaagatg ctgaagatca gttgggtgca 120
cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc 180
gaagaacgtt ttccaatgat gagcactttt āaagtctctg tatgtggcgc ggtattatcc 240
cgtattgacg ccgggcaaga gcaactcggc cgccgcatac actattctca gaatgacttg 300
gttgagtact caccagtcac āgāāaagcat cttacggatg gcatgacagt āagagāatta 360
tgcagtgtct ccataaccat gagtgataac actcgggcca acttacttct gacaacgatc 420
ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgccct 480
gatcgttggt aaccggagct gaatgaagcc ataccaacg acgagcgtga caccacgatg 540

```

cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact tactctagct 600
 tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc acttctgcgc 660
 tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct 720
 cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac 780
 acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc 840
 tcactgatta agcattggta a 861

<210> 15
 <211> 39
 <212> ADN
 <213> Artificial

<220>
 <223> sitio II MCS

<400> 15

ctagcctcga gaattcacgc gtggtacctc tagagtcga 39

<210> 16
 <211> 9210
 <212> ADN
 <213> Artificial

<220>
 <223> Secuencia de vector pBW165

<220>
 <221> características_misceláneas
 <222> (1)..(743)
 <223> CMV prom//pot

<220>
 <221> características_misceláneas
 <222> (825)..(1024)
 <223> RK-intrón

<220>
 <221> características_misceláneas
 <222> (857)..(1000)
 <223> RK-región de intrón que es empalmada en la expresión

<220>
 <221> variación
 <222> (1053)..(1054)
 <223> inserción de mAB LC o fragmento funcional del mismo

<220>
 <221> características_misceláneas
 <222> (1101)..(1322)
 <223> SV40poliA

<220>
 <221> características_misceláneas
 <222> (1653)..(2395)

<223> CMV prom//pot

<220>
 <221> características_misceláneas

- <222> (2477)..(2676)
<223> RK-intrón
- <220>
5 <221> características_miscláneas
<222> (2509)..(2652)
<223> región de intrón RK que es empalmada en la expresión
- <220>
10 <221> variación
<222> (2750)..(2751)
<223> inserción de mAB HC o fragmento funcional del mismo
- <220>
15 <221> características_miscláneas
<222> (2817)..(3038)
<223> SV40poliA
- <220>
20 <221> características_miscláneas
<222> (3133)..(3588)
<223> región f1
- <220>
25 <221> características_miscláneas
<222> (3652)..(4070)
<223> SV40 prom
- <220>
30 <221> características_miscláneas
<222> (3968)..(4033)
<223> origen mínimo de replicación
- <220>
35 <221> características_miscláneas
<222> (4115)..(4909)
<223> Gen de neomicina fosfotransferasa
- <220>
40 <221> características_miscláneas
<222> (4973)..(5021)
<223> poliA sintético
- <220>
45 <221> características_miscláneas
<222> (5432)..(6292)
<223> resistencia a antibiótico de beta lactamasa
- <220>
50 <221> características_miscláneas
<222> (7170)..(7508)
<223> SV40 prom
- <220>
- <220>
55 <221> características_miscláneas
<222> (7585)..(8148)
<223> DHFR mutante
- <220>
60 <221> características_miscláneas
<222> (8149)..(8526)
<223> intrón DHFR mutante

<220>

<221> características_miscláneas

<222> (9010)..(9205)

<223> SV40 poliA

5

<400> 16

tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta	60
ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc	120
aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg	180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc	240
gcctggctga ccgcccacg acccccgccc attgacgtca ataatgacgt atgttcccat	300
agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaaactgc	360
ccacttggca gtacatcaag tgtatcatat gccaaagtacg cccctattg acgtcaatga	420
cggtaaatgg cccgcctggc attatgccca gtacatgacc ttatgggact ttcctacttg	480
gcagtacatc tacgtattag tcatcgctat taccatgggtg atgcggtttt ggcagtacat	540
caatgggagt ggatagcggg ttgactcacg gggatttcca agtctccacc ccattgacgt	600
caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactc	660
cgccccattg acgcaaattg gcggtaggcg tgtacgggtg gaggtctata taagcagagc	720
tcgttttagtg aaccgtcaga tcgcctggag acgccatcca cgctgttttg acctccatag	780
aagacaccgg gaccgatcca gcctccgcgg ccgggaacgg tgcatgggaa cgcggaattcc	840
ccgtgccaaag agtgacgtaa gtaccgccta tagagtctat agggccaccc ccttggttc	900
gttagaacgc ggctacaatt aatacataac cttatgtatc atacacatac gatttaggtg	960
acactataga ataacatcca ctttgccttt ctctccacag gtgtccactc ccaggtccaa	1020
ctgcacctcg gttctatcga aaacgcgtcc accgtcgacc cgggcggccg cttcccttta	1080
gtgaggggta atgcttcgag cagacatgat aagatacatt gatgagtttg gacaaaccac	1140
aactagaatg cagtgaaaaa aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt	1200
tgtaacattt ataagctgca ataaacaagt taacaacaac aattgcattc attttatgtt	1260
tcagggttcag ggggagatgt gggagggttt ttaaagcaag taaaacctct acaaatgtgg	1320
taaaatccga taaggatcga tccgggctgg cgtaatagcg aagaggcccg caccgatcgc	1380
ccttcccaac agttgcgcag cctgaatggc gaatggacgc gccctgtagc ggcgcattaa	1440
gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac acttgccagc gccctagcgc	1500
ccgctccttt cgctttcttc ccttcctttc tcgccacgtt cgccggcttt ccccgtaag	1560
ctctaaatcg ggggctccct ttaggggttc gatttagagc ttacggcac ctcgaccgca	1620

aaaaacttga	tttgggtgat	ggttcacgat	cttcaatatt	ggccattagc	catattattc	1680
attggttata	tagcataaat	caatattggc	tattggccat	tgcatagctt	gtatctatat	1740
cataatatgt	acatttatat	tggctcatgt	ccaatatgac	cgccatgttg	gcattgatta	1800
ttgactagtt	attaatagta	atcaattacg	gggtcattag	ttcatagccc	atatatggag	1860
ttccgcgtta	cataacttac	ggtaaatggc	ccgcctggct	gaccgcccac	cgacccccgc	1920
ccattgacgt	caataatgac	gtatgttccc	atagtaacgc	caatagggac	tttccattga	1980
cgtcaatggg	tggagtattt	acggtaaaact	gcccacttgg	cagtacatca	agtgtatcat	2040
atgccaahta	cgccccctat	tgacgtcaat	gacggtaaat	ggcccgcctg	gcattatgcc	2100
cagtacatga	ccttatggga	ctttcctact	tggcagtaca	tctacgtatt	agtcacgcgt	2160
attaccatgg	tgatgcggtt	ttggcagtac	atcaatgggc	gtggatagcg	gtttgactca	2220
cggggatttc	caagtctcca	ccccattgac	gtcaatggga	gtttgttttg	gcacccaaat	2280
caacgggact	ttccaaaatg	tcgtaacaac	tccgccccat	tgacgcaaht	gggcggtagg	2340
cgtgtacggt	gggagggtcta	tataagcaga	gctcgtttag	tgaaccgtca	gatcgcctgg	2400
agacgccatc	cacgctgttt	tgacctccat	agaagacacc	gggaccgatc	cagcctccgc	2460
ggccgggaac	ggtgcattgg	aacgcggatt	ccccgtgcca	agagtgcagt	aagtaccgcc	2520
tatagagtct	ataggcccac	ccccttggct	tcgttagaac	gcggctacaa	ttaatacata	2580
accttatgta	tcatacacat	acgatttagg	tgacactata	gaataacatc	cactttgcct	2640
ttctctccac	aggtgtccac	tcccaggctc	aactgcacct	cggttctatc	gcgattgaat	2700
taattccccg	gggatcctct	aggggtgaccg	tttaaaacac	cggtgccacc	ggcgcgtggt	2760
acctctagag	tcgacccggg	cggccgcttc	ccttttagtga	gggttaatgc	ttcgagcaga	2820
catgataaga	tacattgatg	agtttggaca	aaccacaact	agaatgcagt	gaaaaaaatg	2880
ctttattttgt	gaaattttgtg	atgctattgc	tttattttgta	accattataa	gctgcaataa	2940
acaagttaac	aacaacaatt	gcattcattt	tatgttttcag	gttcaggggg	agatgtggga	3000
ggtttttttaa	agcaagtaaa	acctctacaa	atgtggtaaa	atccgataag	gatcgatccg	3060
ggctggcgta	atagcgaaga	ggcccgacc	gatcgccctt	cccaacagtt	gcgcagcctg	3120
aatggcgaat	ggacgcgccc	tgtagcggcg	cattaagcgc	ggcgggtgtg	gtggttacgc	3180
gcagcgtgac	cgctacactt	gccagcgcct	tagcgcgcct	tcctttcgct	ttcttccctt	3240
cctttctcgc	cacgttcgcc	ggctttcccc	gtcaagctct	aaatcggggg	ctcccttttag	3300
ggttccgatt	tagagcttta	cggcacctcg	accgcaaaaa	acttgatttg	ggatgatggtt	3360
cacgtagtgg	gccatcgccc	tgatagacgg	tttttcgccc	tttgacgttg	gagtccacgt	3420
tctttaatag	tggactcttg	ttccaaaactg	gaacaacact	caaccctatc	tcggtctatt	3480

cttttgattt	ataagggatt	ttgccgattt	cggcctattg	gttaaaaaat	gagctgattt	3540
aacaaatatt	taacgcgaat	tttaacaaaa	tattaacggt	tacaatttcg	cctgatgcgg	3600
tattttctcc	ttacgcatct	gtgcgggtatt	tcacaccgca	tacgcggatc	tgcgcagcac	3660
catggcctga	aataacctct	gaaagaggaa	cttggttagg	taccttctga	ggcggaaaga	3720
accagctgtg	gaatgtgtgt	cagttagggg	gtggaaaagtc	cccaggctcc	ccagcaggca	3780
gaagtatgca	aagcatgcat	ctcaattagt	cagcaaccag	gtgtggaaag	tccccaggct	3840
ccccagcagg	cagaagtatg	caaagcatgc	atctcaatta	gtcagcaacc	atagtcccgc	3900
ccctaactcc	gcccattccc	cccctaactc	cgcccagttc	cgcccattct	ccgccccatg	3960
gctgactaat	tttttttatt	tatgcagagg	ccgaggccgc	ctcggcctct	gagctattcc	4020
agaagtagtg	aggaggcttt	tttggaggcc	taggcttttg	caaaaagctt	gattcttctg	4080
acacaacagt	ctcgaactta	aggctagagc	caccatgatt	gaacaagatg	gattgcacgc	4140
aggttctccg	gccgcttggg	tggagaggct	attcggtat	gactgggcac	aacagacaat	4200
cggctgctct	gatgccgcgg	tgttccggct	gtcagcgcag	gggcgcccgg	ttctttttgt	4260
caagaccgac	ctgtccggtg	ccctgaatga	actgcaggac	gaggcagcgc	ggctatcgtg	4320
gctggccacg	acgggcgttc	cttgccgcagc	tgtgtcgcac	gttgtcactg	aagcgggaag	4380
ggactggctg	ctattggggc	aagtgccggg	gcaggatctc	ctgtcatctc	accttgctcc	4440
tgccgagaaa	gtatccatca	tggctgatgc	aatgcggcgg	ctgcatacgc	ttgatccggc	4500
tacctgccca	ttcgaccacc	aagcgaaaca	tcgcatcgag	cgagcacgta	ctcggatgga	4560
agccggtctt	gtcgatcagg	atgatctgga	cgaagagcat	caggggctcg	cgccagccga	4620
actgttcgcc	aggctcaagg	cgcgcgatgc	cgacggcgag	gatctcgtcg	tgacctatgg	4680
cgatgcctgc	ttgccgaata	tcatggtgga	aaatggccgc	ttttctggat	tcatcgactg	4740
tggccggctg	ggtgtggcgg	accgctatca	ggacatagcg	ttggctaccc	gtgatattgc	4800
tgaagagctt	ggcggcgaat	gggctgaccg	cttcctcgtg	ctttacggta	tcgccgctcc	4860
cgattcgag	cgcacgcctt	tctatcgctt	tcttgacgag	ttcttctgag	cgggactctg	4920
gggttcgaaa	tgaccgacca	agcgacgccc	aacctgccat	cacgatggcc	gcaataaaat	4980
atctttattt	tcattacatc	tgtgtgttgg	ttttttgtgt	gaatcgatag	cgataaggat	5040
ccgcgtatgg	tgcactctca	gtacaatctg	ctctgatgcc	gcatagttaa	gccagccccg	5100
acacccgcca	acacccgctg	acgcgccctg	acgggcttgt	ctgctcccgg	catccgctta	5160
cagacaagct	gtgaccgtct	ccgggagctg	catgtgtcag	aggttttcac	cgtcatcacc	5220
gaaacgcgcg	agacgaaagg	gcctcgtgat	acgcctatct	ttataggtta	atgtcatgat	5280
aataatggtt	tcttagacgt	caggtggcac	ttttcgggga	aatgtgcgcg	gaacccttat	5340
ttgtttattt	ttctaaatac	attcaaatat	gtatccgctc	atgagacaat	aaccttgata	5400

aatgcttcaa	taatattgaa	aaaggaagag	tatgagtatt	caacatttcc	gtgtcgccct	5460
tattcccttt	tttgccgcat	tttgccttcc	tgtttttgct	caccagaaa	cgctggtgaa	5520
agtaaaagat	gctgaagatc	agttgggtgc	acgagtgggt	tacatcgaac	tggatctcaa	5580
cagcggtaag	atccttgaga	gttttcgccc	cgaagaacgt	ttccaatga	tgagcacttt	5640
taaagtctcg	ctatgtggcg	cggtattatc	ccgtattgac	gccgggcaag	agcaactcgg	5700
tcgccgcata	cactattctc	agaatgactt	ggttgagtac	tcaccagtca	cagaaaagca	5760
tcttacggat	ggcatgacag	taagagaatt	atgcagtgtc	gccataacca	tgagtgataa	5820
cactgcggcc	aacttacttc	tgacaacgat	cggaggaccg	aaggagctaa	ccgctttttt	5880
gcacaacatg	ggggatcatg	taactcgcct	tgatcgttgg	gaaccggagc	tgaatgaagc	5940
cataccaaac	gacgagcgtg	acaccacgat	gcctgtagca	atggcaacaa	cgttgcgcaa	6000
actattaact	ggcgaactac	ttactctagc	ttcccggcaa	caattaatag	actggatgga	6060
ggcggataaa	gttgcaggac	cacttctgcg	ctcggccctt	ccggctggct	ggtttattgc	6120
tgataaatct	ggagccggtg	agcgtgggtc	tcgcggtatc	attgcagcac	tggggccaga	6180
tggtaaagccc	tcccgtatcg	tagttatcta	cacgacgggg	agtcaggcaa	ctatggatga	6240
acgaaataga	cagatcgtcg	agataggtgc	ctcactgatt	aagcattggt	aactgtcaga	6300
ccaagtttac	tcatatatac	tttagattga	tttaaaactt	catttttaat	ttaaaaggat	6360
ctaggtgaag	atcctttttg	ataatctcat	gaccaaaatc	ccttaacgtg	agttttcggt	6420
ccactgagcg	tcagaccccg	tagaaaagat	caaaggatct	tcttgagatc	ctttttttct	6480
gcgcgtaatc	tgctgcttgc	aaacaaaaaa	accaccgcta	ccagcgggtg	tttgtttgcc	6540
ggatcaagag	ctaccaactc	tttttccgaa	ggtaactggc	ttcagcagag	cgcagatacc	6600
aaatactgtc	cttctagtgt	agccgtagtt	aggccaccac	ttcaagaact	ctgtagcacc	6660
gcctacatac	ctcgtctctg	taatcctgtt	accagtggct	gctgccagtg	gcgataagtc	6720
gtgtcttacc	gggttggact	caagacgata	gttaccggat	aaggcgcagc	ggtcgggctg	6780
aacggggggg	tcgtgcacac	agcccagctt	ggagcgaacg	acctacaccg	aactgagata	6840
cctacagcgt	gagctatgag	aaagcgccac	gcttcccga	gggagaaagg	cggacaggta	6900
tccggtaagc	ggcagggctg	gaacaggaga	gcgcacgagg	gagcttccag	ggggaaacgc	6960
ctgggtatctt	tatagtcctg	tcgggtttcg	ccacctctga	cttgagcgtc	gatttttgtg	7020
atgctcgtca	ggggggcgga	gcctatggaa	aaacgccagc	aacgcggcct	ttttacggtt	7080
cctggccctt	tgctggcctt	ttgctcacat	ggctcgacag	atccatttaa	attttcaccg	7140
tcatcaccga	aacgcgcgag	gcagctgtgg	aatgtgtgtc	agttaggggtg	tggaaagtcc	7200
ccaggctccc	cagcaggcag	aagtatgcaa	agcatgcata	tcaattagtc	agcaaccagg	7260

tgtggaaagt	ccccaggctc	cccagcaggc	agaagtatgc	aaagcatgca	tctcaattag	7320
tcagcaacca	tagtcccgcc	cctaactccg	cccatcccgc	ccctaactcc	gcccagttcc	7380
gcccattctc	cgccccatgg	ctgactaatt	ttttttatatt	atgcagaggc	cgaggccgcc	7440
tcggcctctg	agctattcca	gaagtagtga	ggaggctttt	ttggaggcct	aggcttttgc	7500
aaaaagctaa	ttcgagctcg	gtacccccaa	acttgacggc	aatcctagcg	tgaaggctgg	7560
taggatttta	tccccgctgc	catcatgggt	cgaccattga	actgcatcgt	cgccgtgtcc	7620
caaaatatgg	ggattggcaa	gaacggagac	cgaccctggc	ctccgctcag	gaacgagttc	7680
aagtactttc	aaagaatgac	cacaacctct	tcagtggaag	gtaaacagaa	tctggtgatt	7740
atgggtagga	aaacctgggt	ctccattcct	gagaagaatc	gacctttaaa	ggacagaatt	7800
aatatagttc	tcagtagaga	actcaaagaa	ccaccacgag	gagctcattt	tcttgccaaa	7860
agtttgatg	atgccttaag	acttattgaa	caaccggaat	tggcaagtaa	agtagacatg	7920
gtttggatag	tcggaggcag	ttctgtttac	caggaagcca	tgaatcaacc	agggcacctc	7980
agactctttg	tgacaaggat	catgcaggaa	tttgaaagtg	acacgttttt	cccagaaatt	8040
gatttgggga	aatataaact	tctcccagaa	taccagggcg	tcctctctga	ggtccaggag	8100
gaaaaaggca	tcaagtataa	gtttgaagtc	tacgagaaga	aagactaaca	ggaagatgct	8160
ttcaagttct	ctgctcccct	cctaaagcta	tgcattttta	taagaccatg	ggggatgctc	8220
gatcccctcg	cgagttgggt	cagctgctgc	ctgaggctgg	acgacctcgc	ggagttctac	8280
cggcagtgca	aatccgtcgg	catccaggaa	accagcagcg	gctatccgcg	catccatgcc	8340
cccgaactgc	aggagtgggg	aggcacgatg	gccgctttgg	tccggatctt	tgtgaaggaa	8400
ccttacttct	gtggtgtgac	ataattggac	aaactaccta	cagagattta	aagctctaag	8460
gtaaatataa	aatttttaag	tgtataatgt	gttaaaactac	tgattctaata	tgtttgtgta	8520
ttttagattc	caacctatgg	aactgatgaa	tgggagcagt	ggtggaatgc	ctttaatgag	8580
gaaaacctgt	tttgctcaga	agaaatgcca	tctagtgatg	atgaggctac	tgctgactct	8640
caacattcta	ctcctccaaa	aaagaagaga	aaggtagaag	accccaagga	ctttccttca	8700
gaattgctaa	gttttttgag	tcattgctgtg	tttagtaata	gaactcttgc	ttgctttgct	8760
atttacacca	caaaggaaaa	agctgcactg	ctatacaaga	aaattatgga	aaaatattct	8820
gtaaccttta	taagtaggca	taacagttat	aatcataaca	tactgttttt	tcttactcca	8880
cacaggcata	gagtgtctgc	tattaataac	tatgctcaaa	aattgtgtac	ctttagcttt	8940
ttaatttgta	aaggggttaa	taaggaatat	ttgatgtata	gtgccttgac	tagagatcat	9000
aatcagccat	accacatttg	tagaggtttt	acttgcttta	aaaaacctcc	cacacctccc	9060
cctgaacctg	aaacataaaa	tgaatgcaat	tgttggtgtt	aacttgttta	ttgcagctta	9120
taatggttac	aaataaagca	atagcatcac	aaatttcaca	aataaagcat	ttttttcact	9180

gcattctagt tgtggtttga attcggatct

9210

REIVINDICACIONES

1. Un ácido nucleico vector linealizado adecuado para expresar al menos un polipéptido de interés en una célula de mamífero que comprende
 - (a) al menos un casete de expresión (POI) adecuado para expresar un polipéptido de interés, en donde dicho casete de expresión comprende al menos un promotor o elemento promotor/potenciador;
 - (b) un casete de expresión (MSM) que comprende un gen marcador seleccionable de mamífero, en donde dicho gen marcador seleccionable de mamífero es un gen resistente a antibióticos y en donde dicho casete de expresión comprende al menos un promotor o elemento promotor/potenciador;
 - (c) un casete de expresión (MASM) que comprende un gen marcador seleccionable amplificable de mamífero que codifica una dihidrofolato reductasa (DHFR) enzimáticamente funcional, en donde dicho casete de expresión comprende al menos un promotor o elemento promotor/potenciador; y
 - (d) un casete de expresión (PSM) que comprende un gen marcador seleccionable procariota,
 en donde el casete de expresión para expresar el polipéptido de interés (POI) comprende un promotor y/o potenciador más fuerte que los casetes de expresión para expresar los marcadores seleccionables,
- 15 y en donde el casete de expresión (POI) está flanqueado en 5' por el casete de expresión (MASM), el casete de expresión (MSM) está localizado 3' desde el casete de expresión (POI) y en donde los casetes de expresión (MASM), (POI) y (MSM) está dispuestos en la misma orientación de 5' a 3',

y en donde dicho vector ha sido linealizado a través de un sitio de restricción de linealización que está localizado en la forma circular del vector entre los casetes de expresión (PSM) y (MASM) y en donde dicha forma circular del vector,
- 20 el casete de expresión (PSM) está localizado entre los casetes de expresión (MSM) y (MASM), de tal manera que en el vector linealizado, el casete de expresión (POI) está flanqueado en 5' por el casete de expresión (MASM) y en 3' por el casete de expresión (MSM) y el casete de expresión (PSM) está localizado en 3' del casete de expresión (MSM),

y en donde el ácido nucleico del vector lineal comprende los siguientes elementos genéticos en la disposición indicada, en donde la dirección 5' a 3' es indicada por la →:
- 25 I. Promotor del casete de expresión (MASM) (→)

II. Gen que codifica el marcador seleccionable y amplificable de mamífero del casete de expresión MASM (→)

III. Intrón del casete de expresión (MASM) (→)

IV. Sitio poliA del casete de expresión (MASM) (→)

V. Promotor del casete de expresión (POI) (→)
- 30 VI. Intrón del casete de expresión (POI) (→)

VII. Polinucleótido que codifica un polipéptido de interés, el cual es insertado en el casete de expresión (POI) (→)

VIII. Sitio PoliA del casete de expresión (POI) (→)

IX. Promotor del casete de expresión (POI') (→)

X. Intrón del casete de expresión (POI') (→)
- 35 XI. Polinucleótido que codifica un polipéptido adicional de interés, el cual está insertado en el casete de expresión (POI') (→)

XII. Sitio poliA del casete de expresión (POI') (→)

XIII. Promotor del casete de expresión (MSM) (→)

XIV. Gen que codifica el marcador seleccionable de mamíferos del casete de expresión (MSM) (→)

XV. Sitio poliA del casete de expresión (MSM) (→)

XVI. Casete de expresión PSM (→) o (←),

en donde el casete de expresión POI' tal como está indicado por los elementos IX a XII es opcional.

2. El ácido nucleico vector de acuerdo con la reivindicación 1, en donde los casetes de expresión (MSM) y (MASM) comprenden un promotor SV40 o un promotor/potenciador SV40 y el casete de expresión (POI) comprende un promotor/potenciador CMV.
 3. El ácido nucleico vector de acuerdo con la reivindicación 1 o 2, en donde dicho casete de expresión (MSM) comprenden un gen que codifica una neomicina fosfotransferasa enzimáticamente funcional y en donde dicho casete de expresión (MASM) comprende un gen que codifica una dihidrofolato reductasa enzimáticamente funcional (DHFR).
 4. El ácido nucleico vector de acuerdo con al menos una de las reivindicaciones 1 a 3, en donde el vector comprende un casete de expresión adicional (POI') para expresar un polipéptido adicional de interés.
 5. El ácido nucleico vector de acuerdo con la reivindicación 4, para expresar una molécula de inmunoglobulina que comprende en cada casete de expresión (POI) y (POI') un polinucleótido que codifica bien sea una cadena liviana o una pesada de una molécula de inmunoglobulina o fragmentos funcionales de la misma, en donde cada casete de expresión (POI) y (POI') comprende uno de dicho polinucleótidos.
 6. El ácido nucleico vector de acuerdo con al menos una de las reivindicaciones 1 a 5, en donde el vector de expresión es un ácido nucleico vector como se muestra en Seq ID No. 1 o Seq ID NO. 16.
 7. Un método para producir un ácido nucleico vector de acuerdo con al menos una de las reivindicaciones 1 a 6, en donde dicho método comprende disponer en un vector circular al menos los siguientes elementos genéticos
 - (a) al menos un casete de expresión (POI) adecuado para expresar un polipéptido de interés, en donde dicho casete de expresión comprende al menos un promotor o elemento promotor/potenciador;
 - (b) un casete de expresión (MSM) que comprende un gen marcador seleccionable de mamífero, en donde dicho gen marcador seleccionable de mamífero es un gen resistente a antibióticos y en donde dicho casete de expresión comprende al menos un promotor o elemento promotor/ potenciador;
 - (c) un casete de expresión (MASM) que comprende un gen marcador seleccionable amplificable de mamífero, en donde dicho casete de expresión comprende al menos un promotor o elemento promotor/potenciador; y
 - (d) un casete de expresión (PSM) que comprende un gen marcador seleccionable procariota,
- de tal manera que el casete de expresión (POI) está flanqueado en 5' por el casete de expresión (MASM), el casete de expresión (MSM) está localizado en 3' desde el casete de expresión (POI), el casete de expresión (MASM) está dispuesto en 3' del casete de expresión (MSM) y el casete de expresión (PSM) está localizado entre los casetes de expresión (MSM) y (MASM) y en donde los casetes de expresión (MASM), (POI) y (MSM) están dispuestos en la misma orientación 5' a 3' en donde dicho vector circular comprende de un sitio de restricción de linealización único para linealizar el vector que está localizado entre los casetes de expresión (PSM) y (MASM) y linealizar el vector circular a través de dicho sitio de restricción de linealización único para proveer un vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 6.
8. Una célula huésped de mamífero que comprende un ácido nucleico vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 6 el cual está integrado de manera estable en el genoma de la célula huésped de mamífero.
9. Un método para producir una célula huésped de mamífero, que comprende transfectar de manera estable la célula huésped con el ácido nucleico vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 6.
10. Un método para producir un polipéptido de interés, comprendiendo dicho método cultivar al menos una célula huésped de acuerdo con la reivindicación 8 en un medio de cultivo celular bajo condiciones que permiten la expresión de dicho polipéptido de interés.
11. El método de acuerdo con la reivindicación 10, en donde dicho polipéptido de interés es secretado hacia el medio de cultivo celular y el método comprende aislar el polipéptido de interés desde el medio de cultivo celular.
12. El método de acuerdo con la reivindicación 10, que comprende

- aislar el polipéptido de interés a partir de dicho medio de cultivo celular y/o a partir de dicha célula huésped; y/o
- procesar el polipéptido aislado de interés.

Fig. 1

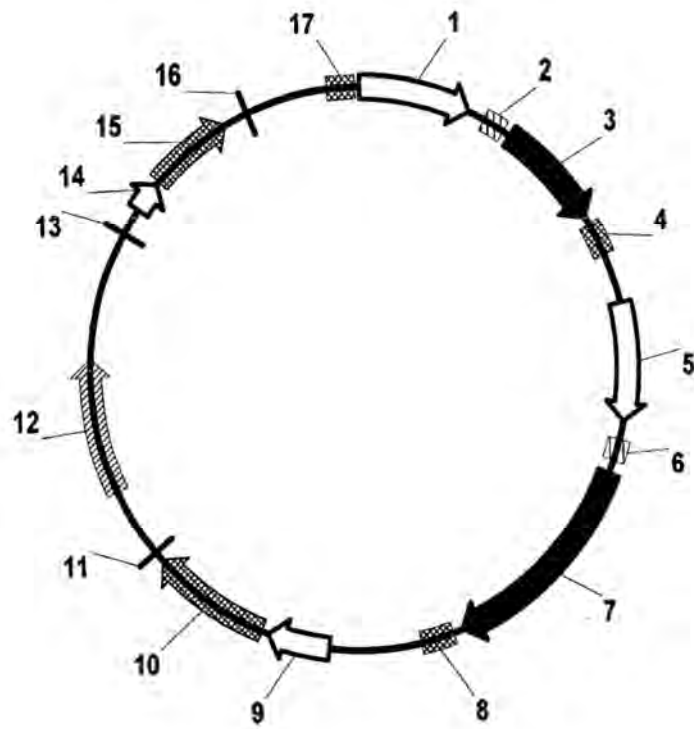


Fig. 2

