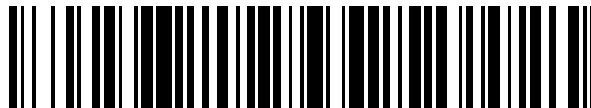


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 627 954**

51 Int. Cl.:

**C12N 9/00** (2006.01)

**C07K 14/00** (2006.01)

**C12N 15/85** (2006.01)

**C07K 16/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.06.2013 PCT/EP2013/062400**

87 Fecha y número de publicación internacional: **19.12.2013 WO13186371**

96 Fecha de presentación y número de la solicitud europea: **14.06.2013 E 13732112 (1)**

97 Fecha y número de publicación de la concesión europea: **12.04.2017 EP 2861741**

54 Título: **Sistema de expresión en CHO**

30 Prioridad:

**14.06.2012 EP 12305677**

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

**01.08.2017**

73 Titular/es:

**SANOFI (100.0%)  
54, rue La Boétie  
75008 Paris, FR**

72 Inventor/es:

**DEVAUD, CATHERINE;  
DUMAS, BRUNO y  
LOUNIS, NABIL**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

ES 2 627 954 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**

Sistema de expresión en CHO

**CAMPO DE LA INVENCION**

5 La presente invención está dentro del campo de la producción industrial de proteínas. Los inventores han diseñado y  
 10 construido un nuevo sistema de expresión que comprende un vector de expresión que codifica una glutamina  
 sintetasa de origen humano o de perro, y una línea de células CHO. Más específicamente, la invención se refiere a  
 una combinación de (i) un vector de ADN adecuado para la producción de una proteína recombinante, en la que  
 dicho vector comprende una secuencia que codifica una glutamina sintetasa, y (ii) una línea de células de ovario de  
 hámster chino (CHO), en la que dicha GS comprende una secuencia al menos el 94,5 % idéntica a la secuencia de  
 SEQ ID NO: 1 o a la secuencia de SEQ ID NO: 2.

**ANTECEDENTES DE LA INVENCION**

Cuando se producen proteínas recombinantes a escala industrial, deben aislarse clones que producen altas  
 cantidades de proteínas recombinantes.

15 La introducción de genes heterólogos en células huésped de animal y el cribado para la expresión de los genes  
 añadidos es un proceso largo y complicado. El proceso implica la transfección y la selección de clones con expresión  
 a largo plazo estable, y el cribado de altas velocidades de expresión de la proteína recombinante.

20 Cuando se generan clones que expresan una proteína recombinante a partir de vectores de expresión, normalmente  
 se transfectan células huésped con un vector de ADN que codifica tanto la proteína de interés como el marcador de  
 selección en el mismo vector. Un vector de expresión tal comprende así un marcador de selección que permite la  
 selección de clones en el que el vector de expresión está presente. Un marcador de selección tal puede también  
 conducir a una co-amplificación que tiene lugar, permitiendo así el aislamiento de clones altamente productores.

25 Se conocen en la técnica varios de tales marcadores de selección, que incluyen, por ejemplo, G418, higromicina,  
 puromicina, zeomicina, dihidrofolato reductasa (DHFR), glutamina sintetasa (GS) e hipoxantina-guanina  
 fosforribosiltransferasa (HPRT). En particular, la GS se usa ampliamente como marcador de selección en el campo  
 de la producción industrial de proteínas recombinantes en células eucariotas.

30 Más específicamente, el documento WO 87/04462 describe el uso de glutamina sintetasa (GS) como marcador de  
 selección. Los ejemplos enseñan un vector de expresión que comprende, como marcador de selección, la secuencia  
 que codifica una GS de origen de hámster chino. Se muestra además que un vector de expresión tal permite la  
 producción de una proteína recombinante tras la transfección del vector de expresión en células CHO, siendo la  
 proteína recombinante tPA.

Aunque el sistema de expresión en CHO anterior basado en el uso de GS como marcador de selección se describió  
 ya en los años 80, todavía hoy en día sigue siendo un estándar en la materia. En particular, no se ha publicado  
 mejora significativa al marcador de selección de GS original.

35 De hecho, la patente coreana KR10-0267720 desvela el uso de GS humana como marcador de selección. Sin  
 embargo, no se ha desvelado la secuencia exacta de la GS humana usada. Además, también se indica que el efecto  
 técnico (alto rendimiento) está solo asociado a tanto la GS humana como al promotor de SV40 específico que se usa  
 (es decir, un promotor de SV40 que carece de las posiciones 128 a 270).

40 Así, hay una necesidad en la materia de sistemas de expresión adicionales y/o mejorados que permitan el  
 aislamiento de un alto número de clones que expresan la proteína recombinante para la que se desea la producción,  
 presentando al menos algunos de estos clones altas velocidades de expresión de la proteína recombinante.

**SUMARIO DE LA INVENCION**

Los inventores han descubierto sorprendentemente que cuando se producen proteínas recombinantes en células  
 CHO, el uso de una GS de origen humano o de perro da mejores resultados que el uso de una GS de origen CHO  
 (véanse, por ejemplo, las Figuras 2 y 3).

45 En particular, se ha encontrado que el uso de una GS de origen humano es especialmente ventajoso, ya que permite  
 el aislamiento de más clones que expresan las proteínas recombinantes que cuando se usa una GS de origen CHO,  
 expresando algunos de ellos la proteína recombinante a niveles más altos que cuando se usa una GS de origen  
 CHO (véase, por ejemplo, la Figura 3).

50 Una realización de la invención proporciona una línea celular de ovario de hámster chino (CHO) que comprende un  
 vector de expresión de ácido desoxirribonucleico (ADN), y en la que el vector comprende una secuencia de  
 nucleótidos que codifica una glutamina sintetasa (GS) de mamífero heteróloga bajo el control de un promotor del  
 virus 40 vacuolante de simio (SV40), que incluye el potenciador de SV40 y al menos un casete de expresión para  
 expresar una proteína recombinante, en la que la GS comprende una secuencia de proteínas que es al menos el

94,5 % idéntica a la secuencia de SEQ ID NO: 1 o a la secuencia de SEQ ID NO: 2. En otra realización de la invención, la GS comprende una secuencia al menos el 94,5 % idéntica a la secuencia de SEQ ID NO: 1 y a la secuencia de SEQ ID NO: 2. En otra realización de la invención, la GS comprende una secuencia al menos el 97,5 % idéntica a la secuencia de SEQ ID NO: 1. En una realización particular de la invención, la GS es una GS humana y comprende una secuencia de SEQ ID NO: 1. En otra realización de la invención, la GS es una GS de perro y tiene una secuencia de SEQ ID NO: 2.

Los investigadores describen que la línea de células CHO comprende un vector de expresión de ácido desoxirribonucleico (ADN), y el vector comprende una secuencia de nucleótidos que codifica una glutamina sintetasa (GS) de mamífero heteróloga y al menos un casete de expresión para expresar una proteína recombinante, en la que los codones del triplete de dicha secuencia que codifica una GS han sido sesgados para la expresión en células CHO. Muestran que dicha secuencia que codifica una GS comprende una secuencia al menos el 80 % idéntica a la secuencia de SEQ ID NO: 8 o SEQ ID NO: 9. En particular, demuestran que la secuencia que codifica una GS comprende una secuencia de SEQ ID NO: 8 o SEQ ID NO: 9. En otra realización de la invención, la secuencia que codifica GS humana se pone bajo el control de un promotor del virus 40 vacuolante de simio (SV40) y la proteína recombinante es un anticuerpo monoclonal.

Los investigadores describen que la línea de células CHO comprende un vector de expresión de ácido desoxirribonucleico (ADN), y el vector comprende una secuencia de nucleótidos que codifica una glutamina sintetasa (GS) de mamífero heteróloga y al menos un casete de expresión para expresar una proteína recombinante en la que los codones del triplete de dicha secuencia que codifica una GS han sido sesgados para la expresión en células CHO. En otra realización, el vector comprende un primer casete de expresión adecuado para la clonación de una cadena ligera del anticuerpo, y un segundo casete de expresión adecuado para la clonación de una cadena pesada del anticuerpo. Los investigadores también muestran que el primer y segundo casetes de expresión comprenden cada uno un promotor de CMV y la línea de células CHO es capaz de ser cultivada en medio sin suero o medio libre de proteína libre de suero y derivado de animal.

En una realización de la invención, la línea de células CHO es la línea celular depositada con el N.º CCL-61 en la ATCC o se deriva de la línea celular depositada con el N.º CCL-61 en la ATCC. En otra realización de la invención, la línea de células CHO permite obtener clones que producen al menos 1 mg/l de proteína recombinante tras la transfección de dicho vector en la línea de células CHO depositada con el N.º CCL-61 en la ATCC.

En otra realización adicional de la invención, la línea de células CHO comprende un vector de expresión de ácido desoxirribonucleico (ADN), y el vector comprende una secuencia de nucleótidos de glutamina sintasa (GS) bajo el control de un promotor del virus (40) vacuolante de simio, que incluye el potenciador de SV40 (es decir, una secuencia de nucleótidos que codifica una glutamina sintetasa (GS)) y al menos un casete de expresión para expresar una proteína recombinante. Los investigadores describen que el vector no contiene un gen heterólogo para la expresión de una proteína recombinante. Muestran que el vector contiene al menos una secuencia que codifica una proteína recombinante. Demuestran que el vector contiene un gen heterólogo que codifica una proteína recombinante que es un anticuerpo monoclonal. Los investigadores también describen que el vector contiene un gen heterólogo que codifica una proteína recombinante que es una proteína inmunogénica para inducir una respuesta de anticuerpos. Muestran que el vector contiene un gen heterólogo que codifica una proteína recombinante que es una enzima para la terapia de reemplazo de enzima o para uso industrial.

Una realización de la invención proporciona un vector de expresión de ácido desoxirribonucleico (ADN), y el vector comprende una secuencia de nucleótidos que codifica una glutamina sintetasa (GS) bajo el control de un promotor del virus 40 vacuolante de simio (SV40), que incluye el potenciador de SV40 y un primer casete de expresión adecuado para la clonación de una proteína recombinante heteróloga bajo el control de un promotor de CMV. En una realización particular de la invención, la GS comprende una secuencia de proteínas que es al menos el 94,5 % idéntica a la secuencia de SEQ ID NO: 1 o a la secuencia de SEQ ID NO: 2.

El investigador también describe un vector de expresión de ácido desoxirribonucleico (ADN), y el vector comprende una secuencia de nucleótidos que codifica una glutamina sintetasa (GS) bajo el control de un promotor del virus 40 vacuolante de simio (SV40) y un primer casete de expresión adecuado para la clonación de una cadena ligera del anticuerpo bajo el control de un promotor de CMV, y un segundo casete de expresión adecuado para la clonación de una cadena pesada del anticuerpo bajo el control de un promotor de CMV. En particular, muestran que la GS comprende una secuencia de proteínas que es al menos el 94,5 % idéntica a la secuencia de SEQ ID NO: 1 o a la secuencia de SEQ ID NO: 2 o un fragmento de al menos 100 aminoácidos consecutivos de SEQ ID NO: 1 o SEQ ID NO: 2.

Los investigadores también describen un vector como se define en la Figura 1.

Una realización de la invención proporciona un método *in vitro* de producción de una proteína recombinante que comprende las etapas de proporcionar una línea de células CHO; cultivar la línea de células CHO obtenida en condiciones adecuadas para la producción de la proteína recombinante; y aislar y/o purificar dicha proteína recombinante. Otra realización proporciona otra etapa de formular la proteína recombinante en una composición farmacéutica.

Una realización de la invención se refiere a una combinación de:

- i) una línea celular eucariota (por ejemplo, una línea de células de ovario de hámster chino (CHO)); y
- ii) un vector de ADN adecuado para la producción de una proteína recombinante, en la que dicho vector comprende una secuencia que codifica una glutamina sintetasa (GS) de mamífero heteróloga bajo el control de un promotor del virus 40 vacuolante de simio (SV40), que incluye el potenciador de SV40 (por ejemplo, una GS que comprende una secuencia al menos el 94,5 % idéntica a la secuencia de SEQ ID NO: 1 o a la secuencia de SEQ ID NO: 2).

Otro aspecto de la presente invención se refiere a un kit que comprende la combinación anterior.

Todavía otro aspecto de la invención se refiere al vector de ADN como tal.

- Todavía otro aspecto de la invención se refiere a una línea de células CHO que comprende el vector de ADN.

En otro aspecto adicional, la invención se refiere a un método *in vitro* de producción de una proteína recombinante que comprende las etapas de:

- a) proporcionar un vector como se define aquí anteriormente;
- b) transfectar una línea celular con dicho vector;
- c) cultivar la línea celular transfectada obtenida en la etapa (b) en condiciones adecuadas para la producción de la proteína recombinante; y
- d) aislar y/o purificar dicha proteína recombinante.

Todavía otro aspecto de la invención se refiere al uso de una combinación tal, o de un vector tal, o de tal línea celular, para producir una proteína recombinante *in vitro*.

## **BREVE DESCRIPCIÓN DE LAS FIGURAS**

La Figura 1 muestra un esquema de los vectores usados en los ejemplos (pBH3695, pBH3700, pBH3694, pBH3699, pBH3698, pBH3697 y pBH3623).

- La Figura 2 muestra el número de pocillos ocupados y productividades logradas con una línea de células CHO transformada con los vectores pBH3695, pBH3700, pBH3694, pBH3699, pBH3698, pBH3697 y pBH3623. La Figura 3 muestra las productividades logradas por los clones obtenidos durante el experimento mostrado en la Figura 2, para los vectores pBH3695, pBH3700 y pBH3623. Cada barra representa un clon.

La Figura 4 muestra los resultados de un experimento llevado a cabo en la línea celular 9E4. Se muestran las productividades y el número de clones obtenidos. El número "0" en el eje horizontal indica que no se obtuvo ningún clon (que es el caso para los vectores pBH3700, pBH3694, pBH3697, pBH3698 y pBH3699).

- La Figura 5 muestra un alineamiento de secuencias entre la GS humana de SEQ ID NO: 1, la GS de perro de SEQ ID NO: 2 y la GS de CHO de SEQ ID NO: 3, que se preparó usando el programa de "alineamiento de múltiples secuencias" CLUSTAL 2.1". Los restos que son diferentes en GS humana y de perro, como en comparación con GS de CHO, se indican con una flecha negra. Estos restos se corresponden con los restos 12, 16, 18, 19, 33, 49, 80, 82, 91, 116, 191, 269, 282, 350, 355 y 356 de SEQ ID NO: 1 y de SEQ ID NO: 2 (las variaciones de aminoácido correspondientes a 12G, 16V, 18M, 19S, 33I, 49S, 80V, 82A, 91K, 116T, 191A, 269Y, 282Q, 350S, 355L y 356I, respectivamente). Los restos que son diferentes en GS humana, en comparación con GS de perro y CHO, se indican con una flecha gris. Estos restos se corresponden con restos en la posición 2, 68, 98, 107, 169, 213 de SEQ ID NO: 1 (las variaciones de aminoácido correspondientes a 2T, 68L, 98L, 107R, 169R y 213S, respectivamente).

- La Figura 6 muestra la concentración de anticuerpo obtenida después de la transfección transitoria de células CHO-S con vectores de control (Control 1 y Control 2), y con los vectores pBH3695 y pBH3772, respectivamente, que expresan los anticuerpos 13C3 y anti-CD38.

## **DESCRIPCIÓN DETALLADA DE LA INVENCION**

Los investigadores describen una combinación de:

- i) una línea celular eucariota; y
- ii) un vector de ADN (ácido desoxirribonucleico) adecuado para la producción de una proteína recombinante, en la que dicho vector comprende una secuencia que codifica una glutamina sintetasa (GS) de mamífero heteróloga.

5 La línea celular eucariota puede ser, por ejemplo, una línea celular de levadura (por ejemplo, una línea celular de *Saccharomyces cerevisiae* o *Yarrowia lipolytica*), una línea celular fúngica (por ejemplo, una línea celular de *Aspergillus niger*), una línea celular de insecto o una línea celular de mamífero (que incluye, pero no se limita a, líneas de células CHO, líneas celulares humanas tales como HEK293 o PERC.6, líneas celulares de ratón tales como NS0 y líneas celulares de mono). En una realización específica, la línea celular eucariota es una línea de células CHO. La GS codificada por el vector de ADN procede de una especie de mamífero heterólogo, y puede, por ejemplo, proceder de ser humano o perro.

En particular, muestran que dicha GS de mamífero heteróloga comprende o consiste en una secuencia:

- al menos el 94,5 % idéntica a al menos una de SEQ ID NO: 1 o SEQ ID NO: 2.

10 Una combinación tal constituye un sistema de expresión.

Más específicamente, un aspecto de la invención se refiere a una combinación de:

i) un vector de ADN adecuado para la producción de una proteína recombinante, en la que dicho vector comprende una secuencia que codifica una glutamina sintetasa (GS) bajo el control del promotor del virus 40 vacuolante de simio (SV40), que incluye el potenciador de SV40; y

15 ii) una línea de células de ovario de hámster chino (CHO);

en la que dicha GS comprende una secuencia:

- al menos el 94,5 % idéntica a la secuencia de SEQ ID NO: 1 o a la secuencia de SEQ ID NO: 2.

La combinación según la invención puede proporcionarse, por ejemplo, bajo la forma de un kit, por ejemplo con un vial que comprende el vector de ADN, y otro vial que comprende la línea celular.

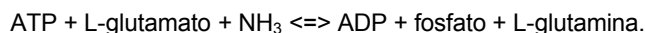
20 Cuando el sistema de expresión se usa para producir una proteína recombinante, el vector se introduce en la línea celular (puede ser, por ejemplo, establemente o transitoriamente transfectado en la línea celular).

Así, la presente invención engloba:

- una combinación en la que el vector está presente dentro de la línea celular, por una parte, y
- una combinación en la que el vector se aísla de la línea celular, por otra parte.

## 25 1. Vector según la invención

El vector de ADN para su uso en la combinación según la invención (denominado además "vector según la invención") es adecuado para la producción de una proteína recombinante, y comprende una secuencia que codifica una glutamina sintetasa (GS) bajo el control de un promotor del virus 40 vacuolante de simio (SV40), que incluye el potenciador de SV40. Como se usa en el presente documento, el término "glutamina sintetasa" o "GS" se refiere a un polipéptido capaz de catalizar la condensación de glutamato y amoníaco para formar glutamina, como se representa por la siguiente reacción bioquímica:



Un polipéptido tal se clasifica con el número de la Comisión de Enzimas (EC) 6.3.1.2. Polipéptidos capaces de catalizar la reacción anterior presentan "actividad de GS".

35 La GS que se usa en el marco de la presente invención (denominada además "GS según la invención") puede comprender o consistir en una secuencia al menos el 94,5 %, 95 %, 95,5 %, 96 %, 96,5 %, 97 %, 97,5 %, 98 %, 98,5 %, 99 %, 99,5 % o el 100 % idéntica a al menos una de SEQ ID NO: 1 o SEQ ID NO: 2. De hecho, se ha encontrado que una GS tal es ventajosa para su uso como un marcador de selección en células CHO (véase el Ejemplo 1). También puede comprender o consistir en un fragmento de al menos 100, 150, 200, 250, 300 o 350 aminoácidos consecutivos de SEQ ID NO: 1 o SEQ ID NO: 2, siempre que la proteína retenga su actividad de GS.

40 En una realización específica, la GS según la invención comprende o consiste en una secuencia al menos el 94,5 %, 95 %, 95,5 %, 96 %, 96,5 %, 97 %, 97,5 %, 98 %, 98,5 %, 99 %, 99,5 % o el 100 % idéntica tanto a la secuencia de SEQ ID NO: 1 como a la secuencia de SEQ ID NO: 2.

45 En una realización específica, la GS según la invención comprende o consiste en una secuencia al menos el 97,5 %, 98 %, 98,5 %, 99 %, 99,5 % o el 100 % idéntica a la secuencia de SEQ ID NO: 1. Una GS tal es particularmente ventajosa para su uso como un marcador de selección en células CHO (véase el Ejemplo 1), en particular en la línea de células CHO E94 (véase el Ejemplo 2).

En una realización específica, la GS según la invención es una GS humana, es decir, una GS de origen humano. Como se usa en el presente documento, el término "GS humana" se refiere a una secuencia que comprende o que

consiste en SEQ ID NO: 1, además de variantes de la misma que presentan actividad de GS. Tales variantes pueden corresponderse, por ejemplo, con variantes que se producen naturalmente en especies humanas (tales como variantes alélicas o variantes de corte y empalme). Alternativamente, tales variantes pueden corresponderse con variantes obtenidas por ingeniería genética. Lo más preferentemente, tales variantes solo se diferencian de la secuencia de SEQ ID NO: 1 por la presencia de como máximo 22, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 variaciones de aminoácido en comparación con SEQ ID NO: 1 (incluyendo dichas variaciones sustituciones, inserciones y deleciones).

En otra realización específica, la GS según la invención es una GS de perro, es decir, una GS de origen de perro. Como se usa en el presente documento, el término "GS de perro" se refiere a una secuencia que comprende o que consiste en SEQ ID NO: 2, además de variantes de la misma que presentan actividad de GS. Tales variantes pueden corresponderse, por ejemplo, con variantes que se producen naturalmente en especies de perro (tales como variantes alélicas o variantes de corte y empalme). Alternativamente, tales variantes pueden corresponderse con variantes obtenidas por ingeniería genética. Lo más preferentemente, tales variantes solo se diferencian de la secuencia de SEQ ID NO: 2 por la presencia de como máximo 22, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 variaciones de aminoácido en comparación con SEQ ID NO: 2 (incluyendo dichas variaciones sustituciones, inserciones y deleciones).

En una realización específica, la GS según la invención comprende al menos 1, 2, 3, 4, 5, 6, 10, 15, 16, 20 o 22 de los siguientes aminoácidos: 12G, 16V, 18M, 19S, 33I, 49S, 80V, 82A, 91 K, 116T, 191A, 269Y, 282Q, 350S, 355L, 356I, 2T, 68L, 98L, 107R, 169R y 213S, en la que el número indica la posición en SEQ ID NO: 1 y SEQ ID NO: 2, y la letra la naturaleza del aminoácido (usando el código genético de una letra). En una realización más específica, la GS según la invención comprende al menos 1, 2, 3, 4, 5 o 6 de los siguientes aminoácidos: 2T, 68L, 98L, 107R, 169R y 213S. En otra realización más específica, la GS según la invención comprende al menos 1, 2, 3, 4, 5, 6, 10, 15 o 16 de los siguientes aminoácidos: 12G, 16V, 18M, 19S, 33I, 49S, 80V, 82A, 91 K, 116T, 191A, 269Y, 282Q, 350S, 355L y 356I. Los aminoácidos anteriores parecen ser específicos para la GS humana y/o de perro, en comparación con la GS de CHO (véase la Figura 5).

Por un polipéptido que tiene una secuencia de aminoácidos al menos, por ejemplo, el 95 % "idéntica" a una secuencia de aminoácidos de búsqueda de la presente invención, se pretende que la secuencia de aminoácidos del polipéptido objeto sea idéntica a la secuencia de búsqueda, excepto que la secuencia de polipéptidos objeto puede incluir hasta cinco alteraciones de aminoácido por cada 100 aminoácidos de la secuencia de aminoácidos de búsqueda. En otras palabras, para obtener un polipéptido que tiene una secuencia de aminoácidos al menos el 95 % idéntica a una secuencia de aminoácidos de búsqueda, hasta el 5 % (5 de 100) de los restos de aminoácidos en la secuencia objeto pueden insertarse, deleccionarse, o sustituirse con otro aminoácido.

En el marco de la presente solicitud, el porcentaje de identidad se calcula usando un alineamiento global (es decir, se comparan las dos secuencias a lo largo de su longitud entera). Métodos de comparación de la identidad y homología de dos o más secuencias son muy conocidos en la técnica. Puede usarse el programa « needle », que usa el algoritmo de alineamiento global de Needleman-Wunsch (Needleman y Wunsch, 1970 J. Mol. Biol. 48:443-453) para encontrar el alineamiento óptimo (incluyendo huecos) de dos secuencias cuando se considera su longitud entera, por ejemplo, cuando se realiza un alineamiento global. Este programa needle está disponible, por ejemplo, en el sitio de la malla mundial de ebi.ac.uk. El porcentaje de identidad según la invención se calcula preferentemente usando el programa EMBOSS::needle (global) con un parámetro de "Hueco abierto" igual a 10,0, un parámetro de "Extensión de hueco" igual a 0,5, y una matriz Blosum62.

Variantes de una secuencia de referencia pueden comprender mutaciones tales como deleciones, inserciones y/o sustituciones en comparación con la secuencia de referencia. En caso de sustituciones, la sustitución se corresponde preferentemente con una sustitución conservativa como se indica en la tabla a continuación.

Sustituciones conservativas	Tipo de aminoácido
Ala, Val, Leu, Ile, Met, Pro, Phe, Trp	Aminoácidos con cadenas laterales hidrófobas alifáticas
Ser, Tyr, Asn, Gln, Cys	Aminoácidos con cadenas laterales no cargadas, pero polares
Asp, Glu	Aminoácidos con cadenas laterales ácidas
Lys, Arg, His	Aminoácidos con cadenas laterales básicas
Gly	Cadena lateral neutra

El vector de ADN según la invención comprende una secuencia que codifica una GS tal según la invención. La secuencia que codifica una GS tal según la invención puede ser la secuencia de nucleótidos que existe de forma natural. Alternativamente, los codones del triplete de la secuencia que codifica una GS tal pueden estar sesgados para la expresión en células CHO. Software y algoritmos para sesgar la secuencia con el fin de obtener una

expresión óptima se conocen en la técnica e incluyen, por ejemplo, el algoritmo descrito en Raab et al. (2010, Syst Synth Biol. 4:215-25). Este algoritmo no solo proporciona los mejores codones disponibles para la expresión, sino que también tiene en cuenta el contenido de GC y la ausencia de motivos de ADN no deseados.

5 Por ejemplo, la secuencia que codifica la GS según la invención puede comprender o consistir en una secuencia al menos el 80 %, 81 %, 82 %, 83 %, 84 %, el 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o el 100 % con respecto a la secuencia de SEQ ID NO: 8 (es decir, una secuencia que codifica la GS humana de SEQ ID NO: 1, que ha sido diseñada para la expresión óptima en células CHO) y/o con respecto a la secuencia de SEQ ID NO: 9 (es decir, una secuencia que codifica una GS de perro de SEQ ID NO: 2, que ha sido diseñada para la expresión óptima en células CHO).

10 En una realización específica, la secuencia que codifica la GS según la invención comprende o consiste en una secuencia de SEQ ID NO: 8 o SEQ ID NO: 9.

15 Por ejemplo, la secuencia que codifica la GS según la invención puede ponerse, por ejemplo, bajo el control de un promotor del virus 40 vacuolante de simio (SV40), por ejemplo, el promotor tardío o temprano de SV40. Un promotor temprano del SV40 se describe, por ejemplo, en Benoist y Chambon (1981, Nature. 290:304-10) y en Moreau et al. (1981, Nucleic Acids Res. 9:6047-68). En particular, dicho promotor de SV40 es un promotor de longitud completa, que incluye el potenciador de SV40. Dicho promotor de SV40 puede también tener un origen de replicación que contiene una repetición de 72 pb.

20 En una realización específica, dicho promotor de SV40 no es un promotor de SV40 en el que las posiciones 128 a 270 han sido eliminadas, es decir, dicho promotor de SV40 no es el promotor de SV40 descrito en la patente coreana N.º 10-0267720 y que transforma el transformante de *E. coli* depositado en Gene Bank, Institute of Bioengineering, KIST el 17 de diciembre de 1997 con el Número de depósito: KCTC 8860 P.

25 Vectores de ADN que son adecuados para la producción de proteínas recombinantes son conocidos para aquellos expertos en la materia. Tales vectores de ADN normalmente se corresponden con vectores de expresión que comprenden un origen de replicación y al menos un casete de expresión que permite la clonación y la expresión de la proteína recombinante para la que se desea la producción. Un casete de expresión normalmente comprende una región no traducida 5' (que comprende o que consiste en un promotor, y opcionalmente una secuencia de potenciador), uno o más sitios de restricción que permiten la clonación de una secuencia que codifica la proteína recombinante, una región no traducida 3' (por ejemplo, una señal de poliA), y opcionalmente uno o más intrones. La secuencia de promotor puede corresponderse con cualquier promotor fuerte muy conocido en la materia, tal como, por ejemplo, el promotor de CMV humano. El vector según la invención puede tener, por ejemplo, la estructura representada en la Figura 1, que se explica en más detalles en el Ejemplo 1, a condición de que la cadena pesada y la cadena ligera de 13C3 puedan sustituirse con otras dos secuencias codificantes (por ejemplo, secuencias que codifican la cadena pesada y la cadena ligera de otro anticuerpo).

35 La proteína recombinante puede corresponderse con cualquier proteína que sea de interés para aquellos expertos en la materia. Como se usa en el presente documento, el término "proteína" pretende englobar péptidos (es decir, cadenas de aminoácidos de menos de 50 aminoácidos), polipéptidos (es decir, cadenas de aminoácidos de al menos 50 aminoácidos), proteínas monoméricas (es decir, proteínas que consisten en una cadena de aminoácidos) y proteínas multiméricas (es decir, proteínas que consisten en dos o más cadenas de aminoácidos, tales como, por ejemplo, anticuerpos monoclonales).

40 El vector según la invención normalmente comprende un número de casetes de expresión que es idéntico al número de cadenas de aminoácidos diferentes que constituyen la proteína (por ejemplo, un casete de expresión en caso de una proteína monomérica o proteína homodimérica, dos en el caso de una proteína heterodimérica o de un anticuerpo monoclonal, etc.).

45 Alternativamente, el vector de ADN según la invención puede comprender solo un casete de expresión, incluso cuando se desee la producción de una proteína heterodimérica o de un anticuerpo monoclonal. En un caso tal, la(s) secuencia(s) que codifica(n) la(s) otra(s) cadena(s) de aminoácidos de la proteína está(n) presente(s) en un vector de expresión separado, que se co-transfecta con el vector según la invención en la línea de células CHO.

50 En una realización específica, el vector de ADN según la invención puede carecer de casete de expresión. En un caso tal, el (los) casete(s) de expresión adecuado(s) para la expresión de la proteína recombinante está(n) presente(s) en un vector separado, que se co-transfecta con el vector según la invención en la línea de células CHO.

55 En toda la presente memoria descriptiva, el término "proteína recombinante" se refiere a cualquier proteína recombinante para la que se desea la producción. Puede corresponderse, por ejemplo, con una proteína terapéutica y/o una profiláctica, es decir, una proteína prevista para su uso como un medicamento (incluyendo vacunas). En una realización específica, la proteína recombinante para la que se desea la producción no es una glutamina sintetasa (GS). En otra realización específica, la proteína recombinante para la que se desea la producción es un anticuerpo, por ejemplo un anticuerpo monoclonal. En otra realización específica adicional, la proteína recombinante para la que se desea la producción es una proteína antigénica. En otra realización específica adicional, la proteína recombinante para la que se desea la producción no es eritropoyetina (EPO).

El término "anticuerpo" se usa en el presente documento en el sentido más amplio y cubre específicamente anticuerpos monoclonales (incluyendo anticuerpos monoclonales de longitud completa) de cualquier isotipo tal como IgG, IgM, IgA, IgD e IgE, anticuerpos policlonales, anticuerpos multiespecíficos (incluyendo anticuerpos biespecíficos), fragmentos de anticuerpos (tales como, por ejemplo, fragmentos Fv, scFv, dsFv, Fab, Fab' o F(ab')<sub>2</sub>), y proteínas de fusión que comprenden un fragmento de anticuerpo. Un anticuerpo reactivo con un antígeno específico puede generarse por métodos recombinantes tales como selección de bibliotecas de anticuerpos recombinantes en fago o vectores similares, o inmunizando un animal con el antígeno o un ácido nucleico que codifica el antígeno.

Un "anticuerpo monoclonal", como se usa en el presente documento, es un anticuerpo obtenido a partir de una población de anticuerpos sustancialmente homogéneos, es decir, los anticuerpos que forman esta población son esencialmente idénticos, excepto por posibles mutaciones que existen de forma natural que podrían estar presentes en cantidades menores. Estos anticuerpos se dirigen contra un único epítipo (o un único grupo de epítopes en el caso de anticuerpos monoclonales multiespecíficos) y, por tanto, son altamente específicos.

Un anticuerpo monoclonal típico comprende dos cadenas pesadas idénticas y dos cadena ligeras idénticas que se unen por enlaces disulfuro. Cada cadena pesada y ligera contiene una región constante y una región variable. Cada región variable contiene tres segmentos llamados "regiones determinantes de la complementariedad" ("CDR") o "regiones hipervariables", que son principalmente responsables de la unión de un epítipo de un antígeno. Normalmente se denominan CDR1, CDR2 y CDR3, se numeran secuencialmente desde el extremo N (véase Kabat et al., Sequences of Proteins of Immunological Interest, 5ª edición, National Institute of Health, Bethesda, MD, 1991). Las porciones más altamente conservadas de las regiones variables se llaman las "regiones estructurales".

El anticuerpo monoclonal puede ser, por ejemplo, un anticuerpo murino, un anticuerpo quimérico, un anticuerpo humanizado, o un anticuerpo completamente humano.

Cuando la proteína recombinante para la que se desea la producción es un anticuerpo monoclonal, el vector según la invención puede comprender un primer casete de expresión adecuado para la clonación de la cadena ligera del anticuerpo, y un segundo casete de expresión adecuado para la clonación de la cadena pesada del anticuerpo.

En una realización específica, dicho primer y segundo casetes de expresión comprenden cada uno el promotor del citomegalovirus (CMV), por ejemplo un promotor de CMV de un CMV humano o murino. Más específicamente, dichos primer y segundo casetes de expresión pueden comprender:

- un promotor/potenciador inmediato/temprano del CMV (por ejemplo, aquél que tiene la secuencia descrita en Teschendorf et al., 2002, Anticancer Res. 22:3325-30); o
- una región de promotor/potenciador de IE2 de CMV de ratón (por ejemplo, aquella que tiene la secuencia descrita en Chatellard et al., 2007, Biotechnol Bioeng. 96:106-17); o
- un elemento regulador de hCMV-MIE (por ejemplo, aquél que tiene la secuencia descrita en el documento WO 89/01036).

El término "proteína antigénica" se usa en el presente documento en el sentido más amplio y cubre cualquier proteína capaz de generar una respuesta inmunitaria, tanto sola como en combinación con un adyuvante. Puede estar prevista para su uso tanto en una vacuna profiláctica como en una vacuna terapéutica. En una realización específica, la proteína antigénica es una proteína vacunal, es decir, una proteína prevista para su uso en una vacuna profiláctica.

En el marco de la presente invención, el vector de ADN podría tanto comprender al menos una secuencia que codifica la proteína recombinante de interés (por ejemplo, una secuencia que codifica una proteína monomérica, una secuencia que codifica una cadena de anticuerpo, o dos secuencias, que codifican una cadena ligera del anticuerpo y una cadena pesada del anticuerpo, respectivamente), o podría estar vacío (es decir, que carece de una secuencia tal que codifica la proteína recombinante de interés).

En un aspecto, la invención se refiere al vector según la invención de por sí. Un vector tal está preferentemente previsto para su uso en una línea de células CHO. Sin embargo, también puede usarse para expresar proteínas en otras líneas celulares eucariotas tales como líneas celulares de levadura, fúngicas, de insecto o de mamífero (por ejemplo, humanas, de ratón, mono, etc.).

## 2. Línea celular según la invención

La línea celular para su uso en la combinación según la invención (denominada además "línea celular según la invención") es una línea celular eucariota, por ejemplo una línea celular de mamífero tal como una línea de células CHO. Las líneas de células CHO se usan comúnmente para producción industrial de proteínas, y muchas líneas de células CHO son conocidas para aquellos expertos en la materia. Por ejemplo, tales líneas de células CHO incluyen, por ejemplo, la línea de células CHO-K1 (Número ATCC: CCL-61), la línea de células DP-12 de CHO (N.º ATCC



CRL-12444 y 12445) y la línea de células 1-15 de CHO (Número ATCC CRL-9606). Estas cepas están públicamente disponibles de la Colección Americana de Cultivos Tipo.

5 En una realización específica, la línea de células CHO según la invención es capaz de crecer en medio sin suero (por ejemplo, un medio químicamente definido) y/o en suspensión. Una línea celular tal puede obtenerse fácilmente por aquellos expertos en la materia adaptando la línea celular parental para crecer en medio sin suero y/o en suspensión (por ejemplo, mediante clonación de una sola célula, mediante adaptación progresiva y/o mediante un proceso de "privación y rescate").

La línea de células CHO según la invención puede tanto ser una línea celular deficiente en GS, como una línea celular que comprende un gen GS endógeno que codifica un polipéptido GS endógeno.

10 En una realización específica, la línea de células CHO es la línea celular depositada con el N.º CCL-61 en la ATCC. Como se usa en el presente documento, el término "línea celular depositada con el N.º CCL-61 en la ATCC" engloba el clon parental depositado en realidad en la ATCC, por una parte, y los clones derivados del mismo, por ejemplo mediante clonación de una sola célula, adaptación progresiva y/o o mediante un proceso de "privación y rescate", por otra parte. Más específicamente, la línea celular depositada con el N.º CCL-61 en la ATCC puede usarse para  
15 obtener clones capaces de crecer en medio sin suero y/o en suspensión.

En una realización específica, la combinación según la invención se caracteriza por que permite obtener clones que producen al menos 1, 2, 3, 4 o 5 mg/l de proteína recombinante tras la transfección del vector en la línea celular depositada con el N.º CCL-61 en la ATCC.

20 En otro aspecto, la invención se refiere a una línea de células CHO que comprende un vector según la invención. Preferentemente, dicha línea de células CHO se transfecta (establemente o se transfecta transitoriamente) con dicho vector. Lo más preferentemente, dicha línea de células CHO comprende dicho vector integrado en su genoma.

Más específicamente, la invención se refiere a la línea de células CHO que comprende un vector de ADN de expresión, y en la que dicho vector comprende una secuencia de nucleótidos que codifica una glutamina sintetasa (GS) de mamífero heteróloga y al menos un casete de expresión para expresar una proteína recombinante, en la  
25 que dicha GS comprende una secuencia de proteínas:

- a) al menos el 94,5 % idéntica a la secuencia de SEQ ID NO: 1 o a la secuencia de SEQ ID NO: 2; o
- b) que consiste en un fragmento de al menos 100 aminoácidos consecutivos de SEQ ID NO: 1 o SEQ ID NO: 2.

### 3. Kits, métodos y usos según la invención

30 Un aspecto de la invención se refiere a un kit que comprende o que consiste en una combinación según la invención. En un kit tal, el vector está preferentemente vacío, ya que esto permite la clonación de la proteína de interés para aquellos expertos en la materia. Además, el vector de ADN se aísla preferentemente de la línea celular en un kit tal. El kit puede comprender además medios adecuados para el cultivo de la línea celular, medios adecuados para la transfección del vector en la línea celular, y/o instrucciones para el uso del sistema de expresión.

35 Otro aspecto de la invención se refiere al uso de la combinación según la invención, o del vector según la invención, o de la línea celular según la invención, para producir una proteína recombinante *in vitro*.

Todavía otro aspecto de la invención se refiere a un método *in vitro* de producción de una proteína recombinante, comprendiendo o consistiendo dicho método en las siguientes etapas:

- a) proporcionar una combinación según la invención;
- b) transfectar dicha línea celular con dicho vector de ADN;
- 40 c) cultivar la línea celular transfectada obtenida en la etapa (b) en condiciones adecuadas para la producción de la proteína recombinante; y
- d) aislar y/o purificar dicha proteína recombinante.

Como es inmediatamente evidente para aquellos expertos en la materia, el aspecto anterior se refiere a una combinación según la invención en la que el vector de ADN se aísla de la línea celular en la etapa (a).

45 Todavía otro aspecto de la invención se refiere a un método *in vitro* de producción de una proteína recombinante, comprendiendo o consistiendo dicho método en las siguientes etapas:

- a) proporcionar una combinación según la invención;
- b) cultivar la línea celular transfectada en condiciones adecuadas para la producción de la proteína recombinante; y

c) aislar y/o purificar dicha proteína recombinante.

Como es inmediatamente evidente para aquellos expertos en la materia, el aspecto anterior se refiere a una combinación según la invención en la que la línea celular comprende el vector de ADN (por ejemplo, la línea celular ha sido previamente transfectada con el vector de ADN) en la etapa (a).

5 Otro aspecto más de la invención se refiere a un método *in vitro* de producción de una proteína recombinante, que comprende o que consiste en las siguientes etapas:

a) proporcionar un vector según la invención, en el que dicho vector comprende al menos una secuencia que codifica una proteína recombinante;

b) transfectar una línea celular según la invención con dicho vector;

10 c) cultivar la línea celular transfectada obtenida en la etapa (b) en condiciones adecuadas para la producción de la proteína recombinante; y

d) aislar y/o purificar dicha proteína recombinante.

Condiciones adecuadas para la producción de proteínas recombinantes son muy conocidas para aquellos expertos en la materia. Pueden usarse, por ejemplo, los protocolos descritos en los ejemplos.

15 En una realización específica, se añade un inhibidor de GS, tal como metionina sulfoximina (msx) o fosfotricina, cuando se cultiva la línea celular según la invención. En una realización más específica, se añaden concentraciones crecientes de un inhibidor de GS tal cuando se cultiva la línea celular. Esto permite seleccionar clones en los que se ha amplificado el gen GS derivado de vector (y así la secuencia que codifica la proteína recombinante).

20 Los métodos anteriores pueden comprender además la etapa de formular la proteína recombinante en una composición farmacéutica.

Todavía otro aspecto de la invención se refiere a un método de co-amplificación de una secuencia de ADN recombinante que codifica una proteína recombinante, que comprende o que consiste en las siguientes etapas:

a) proporcionar un vector según la invención, en el que dicho vector comprende una secuencia que codifica dicha proteína recombinante;

25 b) proporcionar una línea celular según la invención;

c) transfectar dicha línea celular con dicho vector; y

30 d) cultivar dicha línea celular transfectada en condiciones que permiten seleccionar transformantes que contienen un número amplificado de copias de una secuencia derivada de vector que codifica GS, en el que dichos transformantes también contienen un número amplificado de copias de la secuencia que codifica la secuencia de aminoácidos completa de la proteína recombinante.

La etapa (d) del método anterior puede comprender cultivar la línea celular transfectada en medios que contienen un inhibidor de GS y seleccionar células transformantes que son resistentes al nivel progresivamente elevado del inhibidor de GS. Los medios que contienen el inhibidor de GS pueden contener además metionina, por lo que pueden reducirse las concentraciones de inhibidor de GS en los medios.

35 La invención también se refiere a un método de uso de un vector de ADN como marcador de selección dominante en un proceso de co-transformación, en el que dicho método comprende o consiste en las siguientes etapas:

a) proporcionar un vector según la invención, en el que dicho vector comprende una secuencia que codifica una proteína recombinante;

b) proporcionar una línea celular según la invención;

40 c) transfectar dicha línea celular con dicho vector; y

d) seleccionar células transformantes que son resistentes a inhibidores de GS, por lo que se seleccionan las células transformantes, en el que una secuencia de ADN recombinante derivado de vector que codifica GS sirve de marcador de selección dominante y co-amplificable.

En una realización específica de los kits y métodos anteriores, la línea celular es una línea de células CHO.

45 En una realización específica, el uso de la combinación según la invención o del vector según la invención, o de la línea celular según la invención, permite (i) aumentar clones que expresan las proteínas recombinantes, y/o (ii) aumentar la producción de la proteína recombinante, más que cuando se usa una GS de origen CHO.

Se citan varios documentos en todo el texto de esta memoria descriptiva. Sin embargo, no se admite que cualquier documento citado en el presente documento sea de hecho estado de la técnica en relación con la presente invención.

- 5 La invención se describirá además con los siguientes dibujos y ejemplos, que son ilustrativos solo, y no pretenden limitar la presente invención. De hecho, la invención se define por las reivindicaciones, que deben interpretarse con la ayuda de la descripción y los dibujos.

#### **BREVE DESCRIPCIÓN de LAS SECUENCIAS**

SEQ ID NO: 1 muestra la secuencia de aminoácidos de una GS de origen humano.

SEQ ID NO: 2 muestra la secuencia de aminoácidos de una GS de origen de perro.

- 10 SEQ ID NO: 3 muestra la secuencia de aminoácidos de una GS de origen de hámster chino.

SEQ ID NO: 4 muestra la secuencia de aminoácidos de una GS de origen de levadura (*Saccharomyces cerevisiae*).

SEQ ID NO: 5 muestra la secuencia de aminoácidos de una GS que procede de sapo (*Xenopus laevis*).

SEQ ID NO: 6 muestra la secuencia de aminoácidos de una GS que procede de plantas (*Arabidopsis thaliana*).

SEQ ID NO: 7 muestra la secuencia de aminoácidos de una GS que procede de insectos (*Drosophila melanogaster*).

- 15 SEQ ID NO: 8 muestra una secuencia nucleotídica que codifica una GS de origen humano.

SEQ ID NO: 9 muestra una secuencia nucleotídica que codifica una GS de origen de perro.

SEQ ID NO: 10 muestra una secuencia nucleotídica que codifica una GS de origen de hámster chino.

SEQ ID NO: 11 muestra una secuencia nucleotídica que codifica una GS de origen de levadura (*Saccharomyces cerevisiae*).

- 20 SEQ ID NO: 12 muestra una secuencia nucleotídica que codifica una GS que procede de sapo (*Xenopus laevis*).

SEQ ID NO: 13 muestra una secuencia nucleotídica que codifica una GS que procede de plantas (*Arabidopsis thaliana*).

SEQ ID NO: 14 muestra una secuencia nucleotídica que codifica una GS que procede de insectos (*Drosophila melanogaster*).

#### **EJEMPLOS**

##### **Ejemplo 1: Identificación de una GS que da resultados mejorados**

Los inventores pretendieron desarrollar nuevos vectores para la expresión y producción de proteínas recombinantes en líneas de células de ovario de hámster chino (CHO). Se diseñó un conjunto de siete vectores como se describe a continuación.

- 30 Se usaron dos ADNc que codifican una versión humanizada del anticuerpo 13C3 (un ADNc que codifica la cadena pesada de 13C3 y otro ADNc que codifica la cadena ligera de 13C3, respectivamente, formando la combinación de dichas cadenas el anticuerpo 13C3 humanizado) como indicadores para evaluar la calidad del vector. El anticuerpo 13C3 murino es un anticuerpo que se une específicamente a la forma protofibrilar de la proteína de  $\beta$ -amiloide humana, como se describe en el documento WO 2009/065054. Como se usa además en el presente documento, el
- 35 término "13C3" se refiere a la versión humanizada del anticuerpo 13C3 murino.

Los siete vectores se representan esquemáticamente en la Figura 1. Estos ocho vectores comprenden todos:

- una secuencia que codifica una GS, puesta bajo el control del promotor de SV40 temprano;
- un primer casete de expresión, en el que la secuencia que codifica la cadena ligera del anticuerpo 13C3 se pone bajo el control del promotor de CMV;
- 40 - un segundo casete de expresión, en el que la secuencia que codifica la cadena pesada del anticuerpo 13C3 se pone bajo el control del promotor de CMV;
- un origen de replicación procariota;
- un origen de replicación eucariota; y

- un marcador de selección para su uso en células procariotas, concretamente una secuencia que codifica una proteína que confiere resistencia a ampicilina, puesto bajo el control de su promotor natural.

Más específicamente, la secuencia que codifica GS se pone bajo el control del promotor de SV40, que incluye el potenciador de SV40. Un promotor de SV40 temprano tal contiene los potenciadores de repeticiones en tándem de 72 pb de SV40 unidos a las repeticiones en tándem de 21 pb, y la secuencia de proteínas conductora temprana de SV40 que excluye cualquier secuencia codificante. El uso de esta región como promotor fuerte se describió por Benoist y Chambon (1981, Nature. 290:304-10) y en Moreau et al. (1981, Nucleic Acids Res. 9:6047-68). Se usa clásicamente como promotor para la expresión de marcadores de selección en células de mamífero. En los siete vectores pBH3694 a pBH3700, el sitio de restricción HindIII natural que se rompió, y los sitios de restricción únicos (Sall y XmaI) se añadieron en el extremo 5' y 3' de la región promotora, de tal forma que se permitiera un fácil intercambio de los diferentes ADNc de GS.

Los siete vectores se diferencian entre sí por la secuencia que codifica la GS. De hecho, se clonaron secuencias que codificaban GS que tenían diferentes orígenes en los vectores.

Más específicamente, se generaron siete ADNc que codificaban respectivamente una GS de hámster chino (*Cricetulus griseus*), humana (*Homo sapiens*), perro (*Canis lupus*), levadura (*Saccharomyces cerevisiae*), drosophila (*Drosophila melanogaster*), planta (*Arabidopsis thaliana*) y sapo (*Xenopus laevis*) usando las secuencias de aminoácidos que existen de forma natural que están disponibles en bases de datos públicas. A partir de estas secuencias, las proteínas se retro-tradujeron usando una matriz de los codones más frecuentes usados en CHO. A partir de aquí, los ADNc se modificaron para contener sitios de clonación apropiados y se optimizaron las secuencias nucleotídicas. Digno de mención, mientras que las secuencias nucleotídicas se optimizaron para la expresión de CHO, la secuencia de aminoácidos de proteínas codificadas sigue idéntica a la de las proteínas naturalmente codificadas.

Más específicamente, se seleccionaron las secuencias codificantes que existen de forma natural para las diferentes GS en diferentes genotecas de ADNc públicas. Por ejemplo, se usó el N.º de referencia de NCBI NM\_002065.5 para GS humana. Se usó el N.º de referencia de NCBI NM\_001002965.1 para GS de perro. Se usó el N.º de referencia de NCBI NM\_078568.2 para GS de drosophila. La secuencia que codifica GS de levadura se encontró en el sitio de la malla mundial disponible en yeastgenome punto org (N.º de referencia YPR035W). La secuencia de aminoácidos de GS de hámster chino se corresponde con aquella que se muestra en la secuencia de referencia de NCBI: XP\_003502909.1 (REFSEQ: acceso XM\_003502861.1). A partir de las secuencias de ADNc que se producen naturalmente, se sesgaron los codones del triplete de la secuencia que codifica una GS tal para la expresión en células CHO usando un software desarrollado por Wagner y colaboradores, que se basa en el algoritmo descrito en Raab et al. (2010, Syst Synth Biol. 4:215-25). Esta técnica no solo proporciona los mejores codones disponibles para la expresión, sino que también tiene en cuenta el contenido de GC y la ausencia de motivos de ADN no deseados.

Los ADNc obtenidos se clonaron en el esqueleto que lleva los casetes de expresión para el anticuerpo 13C3, dando así los vectores representados en la Figura 1.

El nombre de estos vectores, además del origen y la secuencia de la GS codificada, se muestran en la tabla a continuación.

Nombre	Origen de la GS	Secuencia de aminoácidos de la GS	Secuencia nucleotídica de la GS
pBH3695	Humano	SEQ ID NO: 1	SEQ ID NO: 8
pBH3700	Perro	SEQ ID NO: 2	SEQ ID NO: 9
pBH3623	CHO	SEQ ID NO: 3	SEQ ID NO: 10
pBH3694	Levadura	SEQ ID NO: 4	SEQ ID NO: 11
pBH3699	Sapo	SEQ ID NO: 5	SEQ ID NO: 12
pBH3698	Planta	SEQ ID NO: 6	SEQ ID NO: 13
pBH3697	Drosophila	SEQ ID NO: 7	SEQ ID NO: 14

Se nucleoporaron los vectores anteriores usando condiciones clásicas en una línea de células CHO. 24 horas después de la transfección, se sembraron aproximadamente 2000 células en 480 a 960 pocillos de placas de 96 pocillos, comprendiendo cada pocillo 200 µl de medio CD-CHO que contiene metionina sulfoximina (msx) a una concentración de 25 µM.

Aproximadamente 20 días después de la siembra, los medios de los pocillos se cambiaron a medio nuevo y selectivo (el mismo que se ha descrito anteriormente).

Cuatro días después, se contó el número de ocupados, es decir, se cuentan los números de pocillos que están conteniendo clones en crecimiento. Se probó cada sobrenadante de pocillos ocupados para su productividad de anticuerpos 13C3 usando una tecnología de fluorescencia resuelta en el tiempo homogénea (HTRF) desarrollada por Cisbio Bioassays (Bagnols/Seze, Francia).

- 5 Los resultados se muestran en las Figuras 2 y 3. Puede llegarse a la conclusión a partir de estas figuras que dos vectores, concretamente pBH3695 y PBH3700, dan mejores resultados que los otros vectores. Permiten obtener ambos más clones, y una mejor productividad.

10 Se muestran los porcentajes de identidad entre las secuencias de diferentes GS que se probaron en las tres tablas a continuación. Estos porcentajes de identidad se calcularon usando el programa EMBOSS Needle, usando los siguientes parámetros por defecto:

- Matriz: EBLOSUM62;
- Penalización\_hueco: 10,0; y
- Penalización\_extensión: 0,5.

Vector	Secuencia	Origen de GS	Porcentaje de identidad con la GS humana de SEQ ID NO: 1
pBH3695	SEQ ID NO: 1	Humano	100 %
pBH3700	SEQ ID NO: 2	Perro	97,3 %
pBH3623	SEQ ID NO: 3	CHO	94,1 %
pBH3694	SEQ ID NO: 4	Levadura	52,4 %
pBH3699	SEQ ID NO: 5	Sapo	85,8 %
pBH3698	SEQ ID NO: 6	Planta	50,3 %
pBH3697	SEQ ID NO: 7	Drosophila	62,8 %

Vector	Secuencia	Origen de GS	Porcentaje de identidad con la GS de perro de SEQ ID NO: 2
pBH3695	SEQ ID NO: 1	Humano	97,3 %
pBH3700	SEQ ID NO: 2	Perro	100 %
pBH3623	SEQ ID NO: 3	CHO	94,4 %

15

Vector	Secuencia	Origen de GS	Porcentaje de identidad con la GS de CHO de SEQ ID NO: 3
pBH3695	SEQ ID NO: 1	Humano	94,1 %
pBH3700	SEQ ID NO: 2	Perro	94,4 %
pBH3623	SEQ ID NO: 3	CHO	100 %

20 A partir de estas tablas, puede llegarse a la conclusión de que las dos secuencias que dan los mejores resultados, concretamente GS humana y de perro, se caracterizan por que su secuencia presenta al menos el 94,5 % de identidad con la secuencia de SEQ ID NO: 1 y/o 2. Esta característica no es cierta para las secuencias de las otras GS que se probaron, que condujeron a resultados menos óptimos.

**Ejemplo 2: Confirmación de las ventajosas propiedades del uso de vector que codifica una GS humana en una segunda línea de células CHO.**

Se ha repetido el experimento anterior con la línea de células CHO denominada "9E4", que es adecuada para la producción industrial de proteínas recombinantes.

25 La línea celular 9E4 se estableció a partir de un clon de la línea celular CHO-K1 mediante un proceso de clonación de una sola célula. La línea celular CHO-K1 se obtuvo por Puck en 1957 y se ha depositado en la ATCC con número

CCL-61. Parece que la línea celular 9E4 de CHO expresa una proteína GS endógena y funcional, ya que esta línea celular puede crecer en ausencia de glutamina. Así, debe usarse preferentemente metionina sulfoximina (msx) para la selección de clones transfectados.

5 El vector se introdujo en la línea celular 9E4 mediante nucleoporación. Se realizó un primer experimento usando los seis vectores construidos en el Ejemplo 1 (concretamente pBH3695, pBH3700, pBH3694, pBH3697, pBH3698 y pBH3699). Las condiciones de selección fueron idénticas a las condiciones descritas en el Ejemplo 1 (se añadió msx a una concentración de 25  $\mu$ M). Se midieron el número de pocillos ocupados y la concentración de anticuerpo 13C3 como se describe en el Ejemplo 1. Los resultados se muestran en la Figura 4.

10 En la línea celular 9E4, pBH3695 es el único vector capaz de generar clones que producen anticuerpos 13C3. Este plásmido es el único que lleva el ADNc que codifica GS humana, donde los codones del triplete se sesgaron para la expresión en células CHO. Así, el uso de un vector que comprende una secuencia que codifica una GS humana es particularmente ventajoso para producir proteínas recombinantes en líneas de células CHO.

**Ejemplo 3: Expresión transitoria de X14 en HEK 293 usando el vector pBH basado en GS humana y promotor de CMV humano.**

15 En este experimento, se ha usado un vector que contiene la GS humana de secuencia SEQ ID NO: 1 puesta bajo el control del promotor de SV40, y un único casete de expresión que contiene un ADNc que codifica el receptor de X14 humano o de ratón (también llamado la familia 14 del dominio de lectina de tipo C, miembro A (CLEC14A), y que tiene respectivamente el número de referencia de NCBI NP\_778230.1 y NP\_080085.3) bajo el control del promotor de CMV humano y un sitio de poliadenilación. El vector que contiene el ADNc que codifica el receptor de X14 humano en lo sucesivo vector pBH4590, y el vector que contiene el ADNc que codifica el receptor de X14 de ratón se denomina en lo sucesivo vector pBH4589.

20 Se introdujeron el vector pBH4590, vector pBH4589 o un vector de control (es decir, un vector plasmídico no relacionado) por transfección con Jet PEI en células HEK 293-FS como se describe por el fabricante Poly Plus transfection.

25 Las células se analizaron 24 h después de la transfección por inmunofluorescencia, citometría de flujo o inmunocitoquímica después del marcado apropiado para la detección de X14 humano o de ratón.

Detección por inmunofluorescencia

30 Para la detección por inmunofluorescencia, se centrifugaron las células transfectadas y se desecharon sus sobrenadantes. Se resuspendieron los sedimentos de células en tampón PBS que contenía 1 % de albúmina de suero bovino (PESO/VOLUMEN) y 0,1 % de Tween (V/V) (PBS T BSA) y se saturaron durante 10 minutos en este tampón. Las células se lavaron dos veces con el mismo tampón y se incubaron con suero 1 primario (es decir, suero obtenido justo antes de la inmunización del animal, llamados sueros pre-inmunosuero que representan el control negativo) o suero 2 (es decir, suero que es el inmunosuero no purificado obtenido después de la inmunización del animal) con dominio extracelular humano de X14 purificado como se describe más adelante a dilución 1/5000 en el tampón (PBS T BSA) durante 10 minutos a temperatura ambiente.

35 Después de lavar el anticuerpo primario no unido, se añade un anticuerpo anti-conejo de cabra secundario unido a un fluoróforo Alexa (Alexa 488nm Ref A11034 de Invitrogen). La inmunofluorescencia se realizó usando un microscopio de fluorescencia de Leica establecido para detectar Alexa 488 nm.

40 No puede observarse fluorescencia de Alexa 488 de ruido fondo tanto con el plásmido de control y el suero 1 como el plásmido de control y el suero 2, respectivamente. Esto indica la ausencia de fluorescencia no específica de ruido de fondo. Por el contrario, aparece una fuerte fluorescencia de Alexa 488 en la membrana plasmática de células transfectadas con vectores pBH4590 y pBH4589. Así, el uso de un vector que comprende una secuencia que codifica una GS humana es particularmente ventajoso para promover la expresión transitoria de proteínas unidas a membrana.

45 Detección por citometría de flujo.

Para la detección por citometría de flujo, se logró el marcado de X14 humano por dos incubaciones en serie con tres preparaciones de anticuerpo diferentes y un anticuerpo secundario anti-resto Fc de IgG de conejo. Se analizaron las dos líneas celulares transfectadas con tres diluciones diferentes del suero anti-X14 (1/5000, 1/1000, 1/500):

- 50
- Suero 1, obtener justo antes de la inmunización del animal, llamados sueros pre-inmunosuero que representan el control negativo,
  - Suero 2 que es el inmunosuero no purificado y obtener después de la inmunización del animal con dominio extracelular humano de X14 purificado,

- Suero 3 que se corresponde con la fracción de inmunoglobulina del Suero 2, dirigido contra lectina humana de X14.

Después de lavar el anticuerpo primario no unido, se ha añadido un anticuerpo anti-conejo de cabra secundario unido a un fluoróforo Alexa (Alexa 488nm, Número de catálogo A-11034, de Life Technologies). Las células transfectadas se analizaron usando citometría de flujo establecida para detectar Alexa 488 con tres diluciones diferentes de Suero 1, 2 y 3 respectivamente.

Merece la pena mencionar que todos los histogramas presentan un único pico de fluorescencia, que indica una población de células homogénea. La intensidad de fluorescencia media de este pico está entre  $10^2$  y  $10^3$  unidades de fluorescencia. Sin embargo, las células incubadas con suero 2, diluido a 1/500e, presentan una intensidad de fluorescencia media de aproximadamente 1000 unidades de fluorescencia. Esta intensidad de ruido de fondo mínima es adecuada para estudiar la detección en la membrana plasmática. El citómetro de flujo se calibró de tal forma que la fluorescencia observada con vector de control y suero 1 diluido a 1/5000e se tomara como la fluorescencia de referencia del ruido de fondo.

Entonces se analizaron las células transfectadas con el vector pBH4590. La intensidad de fluorescencia de la célula transfectada con X14 humano con el suero 1, para cada dilución, es similar a las señales de las células de control. Por el contrario, las señales de fluorescencia fueron marcadamente más intensas para el suero 2 y los anticuerpos policlonales purificados (Suero 3) que para el pre-inmunosuero (Suero 1). En realidad, la intensidad de fluorescencia media para el inmunosuero (Suero 2) y el anticuerpo purificado (Suero 3) en células transfectadas con X14 humano es  $10^4$  unidades de fluorescencia. Aumenta un factor de 10-20 a las tres concentraciones de antisueros específicos probados.

Estos resultados demuestran que en células HEK 293-FS transfectadas con el vector pBH4590, X14 humano se produce a un nivel claramente detectable como se observa por microscopía fluorescente y citometría de flujo. Además, X14 humano está siendo accesible a la detección extracelular, que indica que se expresa en la membrana plasmática.

#### 25 **Ejemplo 4: Expresión de receptores de X14 humano y murino en células CHO**

El objetivo de este experimento era probar si el mismo vector que se usó para la expresión transitoria de X14 humano o de ratón podría usarse para la expresión en células CHO como clones estables de tanto receptor de X14 de ratón como humano. Para hacer esto, se transfectó el vector pBH4590 o el vector pBH4589 que lleva ADNc de GS humana en la línea celular CHO-9E4, usando el protocolo desarrollado por el dispositivo de nucleoporación de Lonza/Amara. Se electroporaron dos millones de células CHO-9E4 con 10 µg de vector pBH4590 o vector pBH4589 descritos aquí anteriormente. Poco después del choque eléctrico, las células se diluyeron en 2 ml de medio fresco CD-CHO y 24 horas después las células se diluyeron otra vez en medio CD-CHO fresco que contenía metilsulfóxido (msx) 25 µM a una concentración de 10.000 células por ml. Se sembraron aproximadamente diez placas de 96 pocillos a 2000 células por pocillo.

35 Se cribaron semiclones de CHO, obtenidos después de la transfección de células CHO con el vector pBH4590 o el vector pBH4589, usando el marcador de selección GS. También se realizó una transfección con el control de referencia que expresa el anticuerpo 13C3. Dicho control se usó como control para el análisis por citómetro de flujo. Se obtuvieron treinta semiclones de X14 humano, cuarenta y un semiclones de X14 de ratón y veintinueve semiclones para el vector de control.

40 Después del pase de los semiclones en placas de 24 pocillos, que es al principio del proceso de amplificación, se ha realizado la detección de la presencia o ausencia de antígeno de X14 humano o de ratón usando las herramientas descritas previamente.

45 Se ha observado que, por ejemplo, el semiclón humano N.º 12 tiene intensidad de fluorescencia más baja que el semiclón murino N.º 30. En realidad, con semiclones murinos, se observó la presencia de fluorescencia de un solo pico que tiene una intensidad de fluorescencia máxima de  $10^3$  unidades de fluorescencia, mientras que en el caso de semiclón humano, el pico está extendido y la intensidad de fluorescencia media no supera 500 unidades de fluorescencia. Se realizó la amplificación de semiclones que se analizaron como positivos por citometría de flujo, permitiendo así generar finalmente 10 semiclones de líneas de CHO que expresan establemente lectina murina de X14 y 4 semiclones de líneas de CHO que expresan establemente lectina humana de X14.

50 Así, es particularmente ventajoso el uso de un vector que comprende una secuencia que codifica una GS humana para generar líneas de CHO que expresan establemente proteínas recombinantes.

#### **Ejemplo 5: Transfección transitoria en CHO-S de diferente plásmido de ADN que codifica diferentes dos anticuerpos diferentes**

55 Con el fin de estudiar la factibilidad de uso del vector de los presentes inventores que lleva los ADNc de GS humana para la transfección transitoria en CHO-S, se construyó un segundo vector, llamado, derivado de pBH3695, reemplazando los dos ADNc correspondientes a la cadena ligera y pesada del anticuerpo 13C3 por la cadena ligera

y pesada de los ADNc anti-CD38 usando tecnologías de clonación clásicas y cuatro sitios de restricción únicos diferentes. Por consiguiente, el vector pBH3772 comprende:

- una secuencia que codifica una GS, puesta bajo el control del promotor de SV40 temprano;
- 5 - un primer casete de expresión, en el que la secuencia que codifica la cadena ligera del anticuerpo anti-CD38 se pone bajo el control del promotor de CMV;
- un segundo casete de expresión, en el que la secuencia que codifica la cadena pesada del anticuerpo anti-CD38 se pone bajo el control del promotor de CMV;
- un origen de replicación procarionta;
- un origen de replicación eucariota; y
- 10 - un marcador de selección para su uso en células procariontas, concretamente una secuencia que codifica una proteína que confiere resistencia a ampicilina, puesta bajo el control de su promotor natural.

Entonces se transfectaron CHO-S usando el aparato Maxcyte® con (i) dos vectores de control clásicamente usados para la transfección transitoria y que contienen el ADNc de cadena ligera y pesada de dos anticuerpos diferentes (es decir, llamados Control 1 y Control 2), y (ii) los vectores pBH3695 y pBH3772 en las condiciones descritas por Maxcyte® Corporation. Se cultivaron células CHO-S en CD-CHO que contenía glutamina 8 mM en condiciones de cultivo de CHO-S clásicas (pase cada 2-3 días a  $0,3 \cdot 10^6$  células /ml). El día antes de la transfección, las células se separaron a  $1,2 \cdot 10^6$  células/ml y se transfectaron 24 horas después. Se hicieron desplazamiento de temperatura y alimentación de medio según el protocolo Maxcyte®. Se tomaron muestras de cultivo en el día 3, 6, 7, 8 y 9, y se midieron por SEC-HPLC usando anticuerpo purificado como curva de referencia estándar.

20 Los resultados de un experimento tal se muestran en la Figura 6.

Puede llegarse a la conclusión a partir de esta figura que los vectores pBH3695 y pBH3772 son capaces de producir anticuerpos al nivel que son equivalentes o mejores que los dos vectores de control clásicamente usados para la transfección transitoria.

25 En conclusión, los dos vectores pBH3695 y pBH3772 basados en glutamina sintasa humana son capaces de producir un nivel sorprendente de anticuerpos por encima de 100 mg/l.

El uso de un vector que comprende una secuencia que codifica una GS humana es así particularmente ventajoso para expresar establemente o transitoriamente la proteína unida a membrana o anticuerpos.

#### **Ejemplo 6: Expresión de eritropoyetina (EPO) humana en células CHO**

30 Con el fin de realizar la expresión de EPO humana, se digirió el vector pBH4590 con enzimas de restricción NheI y EcoRI y dos ADNc de EPO humana, es decir, ADNc1 o ADNc2 limitados con sitios NheI y EcoRI, se insertaron en dicho vector usando técnicas clásicas de biología molecular.

Esto permite obtener el vector pBH4614 que lleva el ADNc1 de EPO humana, y el vector pBH4615 que lleva el ADNc2 de EPO humana.

Los dos vectores se prepararon al nivel de maxi-preparación usando un kit desarrollado por Qiagen Corporation.

35 Se usaron vectores pBH4614 y pBH4615 para transfectar las tres líneas celulares CHO-S, CHO-9E4 y CHO 30D12 usando las técnicas de electroporación de Lonza, respectivamente. Para hacer esto, las células se separaron el día antes de la transfección para lograr una densidad celular de  $1 \cdot 10^6$  células/ml. Se centrifugaron dos millones de células y se suspendieron en 100  $\mu$ l de disolución V en presencia de 10  $\mu$ g de ADN, respectivamente, para cada ADN y línea celular. Las células se electroporaron usando el programa X05. Rápidamente después de la electroporación, las células se diluyeron en 2 ml de medio CD-CHO que contenía glutamina 6 mM (Life Technologies) y se incubaron durante 24 horas a 37 °C y 5 % de CO<sub>2</sub>. Después de esta incubación, las células se diluyen en 200 ml del mismo medio sin glutamina y en presencia de msx 25  $\mu$ M y se distribuyen en placas de 96 pocillos usando 200  $\mu$ l por pocillo. Después de 15 (CHO-S) a 25 días (CHO-9E4, CHO 30D12), se cambió medio fresco en pocillos que contenían células supervivientes. Cuatro a 5 días después, las células supervivientes se transfirieron a 1 ml de CD-CHO que contenía msx 25  $\mu$ M sin agitación, respectivamente, para cada pocillo. Para CHO-S, se amplificaron 24 semi-clones, mientras tanto se amplificaron 12 semi-clones para las otras dos líneas celulares, respectivamente para los dos ADNc de EPO. En general, se amplificaron 96 semi-clones, se cultivaron y se verificaron para su capacidad para producir eritropoyetina humana. Para hacer esto, después de 3-4 días de incubación, se diluyó 1 ml en 4 ml del medio CD-CHO que contenía msx 25  $\mu$ M y se puso en agitación a 37 °C y 5 % de CO<sub>2</sub>. Después de 3-4 días, los cultivos se diluyeron otra vez con 5 ml de medio fresco. Después de 3-4 días, se midió la densidad celular y células se diluyeron a  $3 \cdot 10^5$  células/ml y se cultivaron durante 3-4 días para un primer tiempo. Se midió la densidad celular y las células se sembraron a  $3 \cdot 10^5$  células/ml en medio CD-CHO que contenía msx 25  $\mu$ M y 30 % de Feed B (Life Technologies).



Las células se cultivaron en 10 ml del medio anterior (37 °C, 5 % de CO<sub>2</sub>) durante 10 días. Se midieron la densidad celular y viabilidad después de 8 y 10 días. Se tomaron igualmente muestras de 0,6 ml para evaluar la concentración de EPO humana. Se cribaron primero los sobrenadantes de cultivo (0,6 ml) usando la tecnología microfluídica de Caliper que evalúa la presencia de proteína al peso molecular aparente de EPO humana.

- 5 Se sometieron dieciséis sobrenadantes de los mejores clones, que tienen, por ejemplo, la señal de EPO más intensa, a un ELISA específico para la detección de EPO humana usando el kit desarrollado por R&D System® para el diagnóstico *in vitro* (Human Erythropoietin Quantikine® IVD ELISA Kit For In Vitro Diagnostic Use, referencia de catálogo DEP00). Se mostró que siete clones tenían interesantes productividades como se mide en el día 8 y día 10 (siguiente tabla).

Semi clones	Células viables /ml (x10 <sup>6</sup> ) Día8	Células viables /ml (x10 <sup>6</sup> ) Día 10	Concentración de EPO (g/l) Día 8	Concentración de EPO (g/l) Día 10
CHO-S ADNc1/ clon 17	3,9	0,4	0,9	1,5
CHO-S ADNc1/ clon 18	9,8	5,8	0,5	1,0
CHO-S ADNc1/ clon 21	2,4	2,6	0,5	0,9
CHO-S ADNc1/ clon 24	1,4	1,5	0,5	0,9
CHO-S ADNc2/ clon 30	4,6	1,5	0,4	0,6
CHO-30D12 ADNc2/ clon 90	8,4	0,8	1,1	2,0
CHO-30D12 ADNc2/ clon 92	19,4	20,3	0,4	0,7
CHO-30D12 ADNc2/ clon 93	23,1	26,7	0,3	0,7
CHO-30D12 ADNc2/ clon 94	18,8	17,8	1,3	1,4

- 10 Las productividades de los clones estuvieron oscilando de 0,3 a 2,0 g/l. El semiclón 90 fue el mejor clon productor con una productividad de 1,1 a 2,0 g/l en el día 8 y día 10, respectivamente.

- 15 La viabilidad de este clon en el día 10 fue de aproximadamente el 70 % con menos de un millón de células por ml (0,8 millones de células/ml), haciendo imposible calcular la productividad específica a medida que el número de células disminuía entre el día 8 y el día 10. No hace posible el cálculo de una actividad específica (tabla a continuación). Este fenómeno se observa para la mayoría de los semiclones, excepto para los semiclones 21, 24, 92 y 93. En ese caso, la productividad específica expresada en µg/10<sup>6</sup> células/día puede llegar hasta 441 y 854 (tabla a continuación).

Semiclones	Velocidad de crecimiento diaria	Productividad pg/célula en 2 días	Productividad específica (pg/célula/día)
CHO-S ADNc1/ clon 17	NA	NA	NA
CHO-S ADNc1/ clon 18	NA	NA	NA
CHO-S ADNc1/ clon 21	0,29	1542	441
CHO-S ADNc1/ clon 24	0,17	5125	854
CHO-S ADNc2/ clon 30	NA	NA	NA
CHO-30D12 ADNc2/ clon 90	NA	NA	NA
CHO-30D12 ADNc2/ clon 92	0,12	296	36
CHO-30D12 ADNc2/ clon 93	0,43	117	51
CHO-30D12 ADNc2/ clon 94	NA	117	NA

- 20 Estos resultados muestran así que, tanto en términos de volumen como productividad específica, el uso de un vector que comprende una secuencia que codifica una GS humana permite tener productividad por encima de 300 µg de proteína por millón de células.

**LISTADO DE SECUENCIAS**

<110> SANOFI

<120> SISTEMA DE EXPRESIÓN EN CHO

5

<130> FR2012-007

<160> 14

10

<170> PatentIn versión 3.3

<210> 1

<211> 373

<212> PRT

15

<213> Homo sapiens

<400> 1

ES 2 627 954 T3

Met Thr Thr Ser Ala Ser Ser His Leu Asn Lys Gly Ile Lys Gln Val  
 1 5 10 15

Tyr Met Ser Leu Pro Gln Gly Glu Lys Val Gln Ala Met Tyr Ile Trp  
 20 25 30

Ile Asp Gly Thr Gly Glu Gly Leu Arg Cys Lys Thr Arg Thr Leu Asp  
 35 40 45

Ser Glu Pro Lys Cys Val Glu Glu Leu Pro Glu Trp Asn Phe Asp Gly  
 50 55 60

Ser Ser Thr Leu Gln Ser Glu Gly Ser Asn Ser Asp Met Tyr Leu Val  
 65 70 75 80

Pro Ala Ala Met Phe Arg Asp Pro Phe Arg Lys Asp Pro Asn Lys Leu  
 85 90 95

Val Leu Cys Glu Val Phe Lys Tyr Asn Arg Arg Pro Ala Glu Thr Asn  
 100 105 110

Leu Arg His Thr Cys Lys Arg Ile Met Asp Met Val Ser Asn Gln His  
 115 120 125

Pro Trp Phe Gly Met Glu Gln Glu Tyr Thr Leu Met Gly Thr Asp Gly  
 130 135 140

His Pro Phe Gly Trp Pro Ser Asn Gly Phe Pro Gly Pro Gln Gly Pro  
 145 150 155 160

Tyr Tyr Cys Gly Val Gly Ala Asp Arg Ala Tyr Gly Arg Asp Ile Val  
 165 170 175

ES 2 627 954 T3

Glu Ala His Tyr Arg Ala Cys Leu Tyr Ala Gly Val Lys Ile Ala Gly  
 180 185 190

Thr Asn Ala Glu Val Met Pro Ala Gln Trp Glu Phe Gln Ile Gly Pro  
 195 200 205

Cys Glu Gly Ile Ser Met Gly Asp His Leu Trp Val Ala Arg Phe Ile  
 210 215 220

Leu His Arg Val Cys Glu Asp Phe Gly Val Ile Ala Thr Phe Asp Pro  
 225 230 235 240

Lys Pro Ile Pro Gly Asn Trp Asn Gly Ala Gly Cys His Thr Asn Phe  
 245 250 255

Ser Thr Lys Ala Met Arg Glu Glu Asn Gly Leu Lys Tyr Ile Glu Glu  
 260 265 270

Ala Ile Glu Lys Leu Ser Lys Arg His Gln Tyr His Ile Arg Ala Tyr  
 275 280 285

Asp Pro Lys Gly Gly Leu Asp Asn Ala Arg Arg Leu Thr Gly Phe His  
 290 295 300

Glu Thr Ser Asn Ile Asn Asp Phe Ser Ala Gly Val Ala Asn Arg Ser  
 305 310 315 320

Ala Ser Ile Arg Ile Pro Arg Thr Val Gly Gln Glu Lys Lys Gly Tyr  
 325 330 335

Phe Glu Asp Arg Arg Pro Ser Ala Asn Cys Asp Pro Phe Ser Val Thr  
 340 345 350

Glu Ala Leu Ile Arg Thr Cys Leu Leu Asn Glu Thr Gly Asp Glu Pro  
 355 360 365

Phe Gln Tyr Lys Asn  
 370

<210> 2

<211> 373

5 <212> PRT

<213> Canis lupus

<400> 2

10 Met Ala Thr Ser Ala Ser Ser His Leu Asn Lys Gly Ile Lys Gln Val  
 1 5 10 15

ES 2 627 954 T3

Tyr Met Ser Leu Pro Gln Gly Glu Lys Val Gln Ala Met Tyr Ile Trp  
 20 25 30  
 Ile Asp Gly Thr Gly Glu Gly Leu Arg Cys Lys Thr Arg Thr Leu Asp  
 35 40 45  
 Ser Glu Pro Lys Gly Val Glu Glu Leu Pro Glu Trp Asn Phe Asp Gly  
 50 55 60  
 Ser Ser Thr Phe Gln Ser Glu Gly Ser Asn Ser Asp Met Tyr Leu Val  
 65 70 75 80  
 Pro Ala Ala Met Phe Arg Asp Pro Phe Arg Lys Asp Pro Asn Lys Leu  
 85 90 95  
 Val Phe Cys Glu Val Phe Lys Tyr Asn Arg Lys Pro Ala Glu Thr Asn  
 100 105 110  
 Leu Arg His Thr Cys Lys Arg Ile Met Asp Met Val Ser Asn Gln His  
 115 120 125  
 Pro Trp Phe Gly Met Glu Gln Glu Tyr Thr Leu Met Gly Thr Asp Gly  
 130 135 140  
 His Pro Phe Gly Trp Pro Ser Asn Gly Phe Pro Gly Pro Gln Gly Pro  
 145 150 155 160  
 Tyr Tyr Cys Gly Val Gly Ala Asp Lys Ala Tyr Gly Arg Asp Ile Val  
 165 170 175  
 Glu Ala His Tyr Arg Ala Cys Leu Tyr Ala Gly Ile Lys Ile Ala Gly  
 180 185 190  
 Thr Asn Ala Glu Val Met Pro Ala Gln Trp Glu Phe Gln Ile Gly Pro  
 195 200 205  
 Cys Glu Gly Ile Asp Met Gly Asp His Leu Trp Val Ala Arg Phe Ile  
 210 215 220  
 Leu His Arg Val Cys Glu Asp Phe Gly Val Ile Ala Thr Phe Asp Pro  
 225 230 235 240  
 Lys Pro Ile Pro Gly Asn Trp Asn Gly Ala Gly Cys His Thr Asn Phe  
 245 250 255  
 Ser Thr Lys Ala Met Arg Glu Glu Asn Gly Leu Lys Tyr Ile Glu Glu  
 260 265 270

ES 2 627 954 T3

Ser Ile Glu Lys Leu Ser Lys Arg His Gln Tyr His Ile Arg Ala Tyr  
 275 280 285

Asp Pro Lys Gly Gly Leu Asp Asn Ala Arg Arg Leu Thr Gly Phe His  
 290 295 300

Glu Thr Ser Asn Ile Asn Asp Phe Ser Ala Gly Val Ala Asn Arg Gly  
 305 310 315 320

Ala Ser Ile Arg Ile Pro Arg Thr Val Gly Gln Glu Lys Lys Gly Tyr  
 325 330 335

Phe Glu Asp Arg Arg Pro Ser Ala Asn Cys Asp Pro Phe Ser Val Thr  
 340 345 350

Glu Ala Leu Ile Arg Thr Cys Leu Leu Asn Glu Thr Gly Asp Glu Pro  
 355 360 365

Phe Gln Tyr Lys Asn  
 370

<210> 3

<211> 373

5

<212> PRT

<213> Cricetulus griseus

<400> 3

Met Ala Thr Ser Ala Ser Ser His Leu Asn Lys Asn Ile Lys Gln Met  
 1 5 10 15

Tyr Leu Cys Leu Pro Gln Gly Glu Lys Val Gln Ala Met Tyr Ile Trp  
 20 25 30

Val Asp Gly Thr Gly Glu Gly Leu Arg Cys Lys Thr Arg Thr Leu Asp  
 35 40 45

Cys Glu Pro Lys Cys Val Glu Glu Leu Pro Glu Trp Asn Phe Asp Gly  
 50 55 60

Ser Ser Thr Phe Gln Ser Glu Gly Ser Asn Ser Asp Met Tyr Leu Ser  
 65 70 75 80

Pro Val Ala Met Phe Arg Asp Pro Phe Arg Arg Asp Pro Asn Lys Leu  
 85 90 95

Val Phe Cys Glu Val Phe Lys Tyr Asn Arg Lys Pro Ala Glu Thr Asn  
 100 105 110

ES 2 627 954 T3

Leu Arg His Ser Cys Lys Arg Ile Met Asp Met Val Ser Asn Gln His  
 115 120 125

Pro Trp Phe Gly Met Glu Gln Glu Tyr Thr Leu Met Gly Thr Asp Gly  
 130 135 140

His Pro Phe Gly Trp Pro Ser Asn Gly Phe Pro Gly Pro Gln Gly Pro  
 145 150 155 160

Tyr Tyr Cys Gly Val Gly Ala Asp Lys Ala Tyr Gly Arg Asp Ile Val  
 165 170 175

Glu Ala His Tyr Arg Ala Cys Leu Tyr Ala Gly Val Lys Ile Thr Gly  
 180 185 190

Thr Asn Ala Glu Val Met Pro Ala Gln Trp Glu Phe Gln Ile Gly Pro  
 195 200 205

Cys Glu Gly Ile Arg Met Gly Asp His Leu Trp Val Ala Arg Phe Ile  
 210 215 220

Leu His Arg Val Cys Glu Asp Phe Gly Val Ile Ala Thr Phe Asp Pro  
 225 230 235 240

Lys Pro Ile Pro Gly Asn Trp Asn Gly Ala Gly Cys His Thr Asn Phe  
 245 250 255

Ser Thr Lys Ala Met Arg Glu Glu Asn Gly Leu Lys His Ile Glu Glu  
 260 265 270

Ala Ile Glu Lys Leu Ser Lys Arg His Arg Tyr His Ile Arg Ala Tyr  
 275 280 285

Asp Pro Lys Gly Gly Leu Asp Asn Ala Arg Arg Leu Thr Gly Phe His  
 290 295 300

Glu Thr Ser Asn Ile Asn Asp Phe Ser Ala Gly Val Ala Asn Arg Ser  
 305 310 315 320

Ala Ser Ile Arg Ile Pro Arg Thr Val Gly Gln Glu Lys Lys Gly Tyr  
 325 330 335

Phe Glu Asp Arg Arg Pro Ser Ala Asn Cys Asp Pro Phe Ala Val Thr  
 340 345 350

Glu Ala Ile Val Arg Thr Cys Leu Leu Asn Glu Thr Gly Asp Glu Pro  
 355 360 365

Phe Gln Tyr Lys Asn  
 370

<210> 4

5 <211> 370

<212> PRT

ES 2 627 954 T3

<213> Saccharomyces cerevisiae

<400> 4

Met Ala Glu Ala Ser Ile Glu Lys Thr Gln Ile Leu Gln Lys Tyr Leu  
 1 5 10 15

Glu Leu Asp Gln Arg Gly Arg Ile Ile Ala Glu Tyr Val Trp Ile Asp  
 20 25 30

Gly Thr Gly Asn Leu Arg Ser Lys Gly Arg Thr Leu Lys Lys Arg Ile  
 35 40 45

Thr Ser Ile Asp Gln Leu Pro Glu Trp Asn Phe Asp Gly Ser Ser Thr  
 50 55 60

Asn Gln Ala Pro Gly His Asp Ser Asp Ile Tyr Leu Lys Pro Val Ala  
 65 70 75 80

Tyr Tyr Pro Asp Pro Phe Arg Arg Gly Asp Asn Ile Val Val Leu Ala  
 85 90 95

Ala Cys Tyr Asn Asn Asp Gly Thr Pro Asn Lys Phe Asn His Arg His  
 100 105 110

Glu Ala Ala Lys Leu Phe Ala Ala His Lys Asp Glu Glu Ile Trp Phe  
 115 120 125

Gly Leu Glu Gln Glu Tyr Thr Leu Phe Asp Met Tyr Asp Asp Val Tyr  
 130 135 140

Gly Trp Pro Lys Gly Gly Tyr Pro Ala Pro Gln Gly Pro Tyr Tyr Cys  
 145 150 155 160

Gly Val Gly Ala Gly Lys Val Tyr Ala Arg Asp Met Ile Glu Ala His  
 165 170 175

Tyr Arg Ala Cys Leu Tyr Ala Gly Leu Glu Ile Ser Gly Ile Asn Ala  
 180 185 190

Glu Val Met Pro Ser Gln Trp Glu Phe Gln Val Gly Pro Cys Thr Gly  
 195 200 205



ES 2 627 954 T3

Ile Asp Met Gly Asp Gln Leu Trp Met Ala Arg Tyr Phe Leu His Arg  
 210 215 220

Val Ala Glu Glu Phe Gly Ile Lys Ile Ser Phe His Pro Lys Pro Leu  
 225 230 235 240

Lys Gly Asp Trp Asn Gly Ala Gly Cys His Ala Asn Val Ser Thr Lys  
 245 250 255

Glu Met Arg Gln Pro Gly Gly Thr Lys Tyr Ile Glu Gln Ala Ile Glu  
 260 265 270

Lys Leu Ser Lys Arg His Ala Glu His Ile Lys Leu Tyr Gly Ser Asp  
 275 280 285

Asn Asp Met Arg Leu Thr Gly Arg His Glu Thr Ala Ser Met Thr Ala  
 290 295 300

Phe Ser Ser Gly Val Ala Asn Arg Gly Ser Ser Ile Arg Ile Pro Arg  
 305 310 315 320

Ser Val Ala Lys Glu Gly Tyr Gly Tyr Phe Glu Asp Arg Arg Pro Ala  
 325 330 335

Ser Asn Ile Asp Pro Tyr Leu Val Thr Gly Ile Met Cys Glu Thr Val  
 340 345 350

Cys Gly Ala Ile Asp Asn Ala Asp Met Thr Lys Glu Phe Glu Arg Glu  
 355 360 365

Ser Ser  
 370

<210> 5

<211> 373

5 <212> PRT

<213> Xenopus laevis

<400> 5

Met Ala Thr Ser Ala Ser Ala Gln Leu Ser Lys Ala Ile Lys Gln Met  
 1 5 10 15

Tyr Leu Glu Leu Pro Gln Gly Asp Lys Val Gln Ala Met Tyr Ile Trp  
 20 25 30

Ile Asp Gly Thr Gly Glu Gly Leu Arg Cys Lys Thr Arg Thr Leu Asp  
 35 40 45

10

ES 2 627 954 T3

Ser Glu Pro Lys Thr Ile Glu Asp Leu Pro Glu Trp Asn Phe Asp Gly  
50 55 60

Ser Ser Thr His Gln Ser Glu Gly Ser Asn Ser Asp Met Tyr Leu Ile  
65 70 75 80

Pro Val Ala Met Phe Arg Asp Pro Phe Arg Arg Asp Pro Asn Lys Leu  
85 90 95

Val Leu Cys Glu Val Leu Lys Tyr Asn Arg Lys Thr Ala Glu Thr Asn  
100 105 110

Leu Arg His Thr Cys Asn Gln Ile Met Asp Met Val Gly Asn Glu His  
115 120 125

Pro Trp Phe Gly Met Glu Gln Glu Tyr Thr Leu Leu Gly Met Asp Gly  
130 135 140

His Pro Phe Gly Trp Pro Ser Asn Gly Phe Pro Gly Pro Gln Gly Pro  
145 150 155 160

Tyr Tyr Cys Gly Val Gly Ala Asn Lys Ala Tyr Gly Arg Asp Ile Val  
165 170 175

Glu Ala His Tyr Arg Ala Cys Leu Tyr Ala Gly Val Lys Ile Ala Gly  
180 185 190

Thr Asn Ala Glu Val Met Pro Ala Gln Trp Glu Phe Gln Ile Gly Thr  
195 200 205

Cys Glu Gly Ile Asp Met Gly Asp His Leu Trp Ile Ala Arg Phe Ile  
210 215 220

Leu His Arg Val Cys Glu Asp Phe Gly Ile Ile Val Ser Phe Asp Pro  
225 230 235 240

Lys Pro Ile Thr Gly Asn Trp Asn Gly Ala Gly Cys His Thr Asn Phe  
245 250 255

Ser Thr Lys Ser Met Arg Glu Glu Gly Gly Leu Lys His Ile Glu Glu  
260 265 270

Ser Ile Glu Arg Leu Ser Lys Arg His Glu Tyr His Ile Arg Met Tyr  
275 280 285

Asp Pro Arg Gly Gly Lys Asp Asn Ala Arg Arg Leu Thr Gly Phe His

ES 2 627 954 T3

290 295 300

Glu Thr Ser Ser Ile His Glu Phe Ser Ala Gly Val Ala Asn Arg Gly  
 305 310 315 320

Ala Ser Ile Arg Ile Pro Arg Leu Val Gly Gln Glu Lys Lys Gly Tyr  
 325 330 335

Phe Glu Asp Arg Arg Pro Ser Ala Asn Cys Asp Pro Tyr Ala Val Thr  
 340 345 350

Glu Ala Val Ile Arg Thr Cys Leu Leu Asn Glu Thr Gly Asp Glu Pro  
 355 360 365

Leu Glu Tyr Lys Asn  
 370

<210> 6

<211> 354

5 <212> PRT

<213> Arabidopsis thaliana

<400> 6

Met Ser Leu Leu Ser Asp Leu Val Asn Leu Asn Leu Thr Asp Ala Thr  
 1 5 10 15

Gly Lys Ile Ile Ala Glu Tyr Ile Trp Ile Gly Gly Ser Gly Met Asp  
 20 25 30

Ile Arg Ser Lys Ala Arg Thr Leu Pro Gly Pro Val Thr Asp Pro Ser  
 35 40 45

Lys Leu Pro Lys Trp Asn Tyr Asp Gly Ser Ser Thr Gly Gln Ala Ala  
 50 55 60

Gly Glu Asp Ser Glu Val Ile Leu Tyr Pro Gln Ala Ile Phe Lys Asp  
 65 70 75 80

Pro Phe Arg Lys Gly Asn Asn Ile Leu Val Met Cys Asp Ala Tyr Thr  
 85 90 95

Pro Ala Gly Asp Pro Ile Pro Thr Asn Lys Arg His Asn Ala Ala Lys  
 100 105 110

Ile Phe Ser His Pro Asp Val Ala Lys Glu Glu Pro Trp Tyr Gly Ile  
 115 120 125

Glu Gln Glu Tyr Thr Leu Met Gln Lys Asp Val Asn Trp Pro Ile Gly

10

ES 2 627 954 T3

130		135		140											
Trp 145	Pro	Val	Gly	Gly	Tyr 150	Pro	Gly	Pro	Gln	Gly 155	Pro	Tyr	Tyr	Cys	Gly 160
Val	Gly	Ala	Asp	Lys 165	Ala	Ile	Gly	Arg	Asp 170	Ile	Val	Asp	Ala	His	Tyr 175
Lys	Ala	Cys	Leu 180	Tyr	Ala	Gly	Ile	Gly	Ile 185	Ser	Gly	Ile	Asn	Gly	Glu 190
Val	Met	Pro	Gly	Gln	Trp	Glu	Phe	Gln	Val	Gly	Pro	Val	Glu	Gly	Ile
		195					200					205			
Ser	Ser	Gly	Asp	Gln	Val	Trp	Val	Ala	Arg	Tyr	Leu	Leu	Glu	Arg	Ile
	210					215					220				
Thr	Glu	Ile	Ser	Gly	Val	Ile	Val	Ser	Phe	Asp	Pro	Lys	Pro	Val	Pro
225					230					235					240
Gly	Asp	Trp	Asn	Gly	Ala	Gly	Ala	His	Cys	Asn	Tyr	Ser	Thr	Lys	Thr
				245					250					255	
Met	Arg	Asn	Asp	Gly	Gly	Leu	Glu	Val	Ile	Lys	Lys	Ala	Ile	Gly	Lys
			260					265					270		
Leu	Gln	Leu	Lys	His	Lys	Glu	His	Ile	Ala	Ala	Tyr	Gly	Glu	Gly	Asn
		275					280					285			
Glu	Arg	Arg	Leu	Thr	Gly	Lys	His	Glu	Thr	Ala	Asp	Ile	Asn	Thr	Phe
	290					295					300				
Ser	Trp	Gly	Val	Ala	Asn	Arg	Gly	Ala	Ser	Val	Arg	Val	Gly	Arg	Asp
305					310					315					320
Thr	Glu	Lys	Glu	Gly	Lys	Gly	Tyr	Phe	Glu	Asp	Arg	Arg	Pro	Ala	Ser
				325					330					335	
Asn	Met	Asp	Pro	Tyr	Val	Val	Thr	Ser	Met	Ile	Ala	Glu	Thr	Thr	Ile
			340					345					350		
Leu	Gly														

<210> 7

<211> 369

5 <212> PRT

<213> Drosophila melanogaster

<400> 7

ES 2 627 954 T3

Met Ser Ala Arg Ile Leu Glu Asp Ser Pro Asn Ala Arg Ile Asn Lys  
 1 5 10 15

Thr Ile Leu Asp Arg Tyr Leu Ser Leu Pro Leu Gln Glu Asn Ile Val  
 20 25 30

Gln Ala Thr Tyr Val Trp Ile Asp Gly Thr Gly Glu Asp Leu Arg Cys  
 35 40 45

Lys Asp Arg Thr Leu Asp Phe Ile Pro Gln Ser Pro Lys Glu Leu Pro  
 50 55 60

Val Trp Asn Tyr Asp Gly Ser Ser Cys Tyr Gln Ala Glu Gly Ser Asn  
 65 70 75 80

Ser Asp Thr Tyr Leu Tyr Pro Val Ala Ile Tyr Lys Asp Pro Phe Arg  
 85 90 95

Arg Gly Asn Asn Ile Leu Val Met Cys Asp Thr Tyr Lys Phe Asp Gly  
 100 105 110

Thr Pro Thr Asp Thr Asn Lys Arg Lys Thr Cys Leu Glu Val Ala Asn  
 115 120 125

Lys Cys Ala Ala Glu Glu Pro Trp Phe Gly Ile Glu Gln Glu Tyr Thr  
 130 135 140

Phe Leu Asp Phe Asp Gly His Pro Leu Gly Trp Pro Lys Asn Gly Phe  
 145 150 155 160

Pro Gly Pro Gln Gly Pro Tyr Tyr Cys Gly Val Gly Ala Asn Lys Val  
 165 170 175

Tyr Ala Arg Asp Ile Val Asp Ala His Tyr Arg Ala Cys Leu Tyr Ala  
 180 185 190

Gly Ile Lys Val Ser Gly Thr Asn Ala Glu Val Met Pro Ala Gln Trp  
 195 200 205

Glu Phe Gln Val Gly Pro Cys Glu Gly Ile Ser Ile Gly Asp Asp Leu  
 210 215 220

Trp Met Ala Arg Phe Leu Leu His Arg Ile Ser Glu Glu Phe Gly Ile  
 225 230 235 240

ES 2 627 954 T3

Val Ser Thr Leu Asp Pro Lys Pro Met Pro Gly Asp Trp Asn Gly Ala  
 245 250 255

Gly Ala His Thr Asn Val Ser Thr Lys Ala Met Arg Glu Asp Gly Gly  
 260 265 270

Ile Arg Asp Ile Glu Lys Ala Val Ala Lys Leu Ser Lys Cys His Glu  
 275 280 285

Arg His Ile Arg Ala Tyr Asp Pro Lys Gln Gly Gln Asp Asn Ala Arg  
 290 295 300

Arg Leu Thr Gly Lys His Glu Thr Ser Ser Ile Asn Asp Phe Ser Ala  
 305 310 315 320

Gly Val Ala Asn Arg Gly Cys Ser Ile Arg Ile Pro Arg Gly Val Asn  
 325 330 335

Asp Asp Gly Lys Gly Tyr Phe Glu Asp Arg Arg Pro Ser Ser Asn Cys  
 340 345 350

Asp Pro Tyr Ser Val Val Glu Ala Ile Leu Arg Thr Ile Cys Leu Asp  
 355 360 365

Glu

<210> 8

<211> 1122

5

<212> ADN

<213> Homo sapiens

<400> 8

atgaccacct cgcctccag ccacctgaac aaggcatca aacaggtgta catgagcctg 60

ccccaggcgc agaaggtgca ggccatgtac atctggatcg acggcaccgg cgagggactg 120

cggtgcaaga ccagaaccct ggactccgag cctaagtgcg tggaagaact gcccagtggtg 180

aacttgacg gctcctccac cctgcagtcc gagggctcca actccgacat gtacctgggtg 240

cctgccgcca tgttccggga ccctttccgg aaggaccca acaagctggt gctgtgcgag 300

gtgttcaagt acaacagacg gcctgccgag acaaacctgc ggcatacctg caagcggatc 360

atggacatgg tgtccaacca gcacccttgg tttggcatgg aacaggagta caccctgatg 420

ggcaccgacg gccaccctt cggctggcct tctaacggct tccctggccc ccagggcccc 480

tactattgtg gcgtgggcgc cgaccgggcc tacggcagag atatcgtgga agcccactac 540

cgggcctgcc tgtacgccgg agtgaagatc gccggcacca acgccgaagt gatgcccgcc 600

cagtgggagt tccagatcgg cccttgcgag ggcatctcca tgggcatca cctgtgggtg 660

10

ES 2 627 954 T3

gcccggttca tctgacacag agtgtgagag gacttcggcg tgatcgccac cttcgacccc 720  
 aagcccatcc cgggcaactg gaacggcgct ggctgccaca ccaacttctc caccaaggcc 780  
 atgcgggaag agaacggcct gaagtacatc gaggaagcca tcgagaagct gtccaagcgg 840  
 caccagtacc acatcagagc ctacgaccct aagggcggcc tggacaacgc cagaaggctg 900  
 accggctttc acgagacatc caacatcaac gacttctctg cggcgctggc caacagatcc 960  
 gcctccatcc ggatccctag aaccgtgggc caggaaaaga agggctactt cgaggacaga 1020  
 cggccctccg ccaactgcga cccctttagc gtgaccgagc ccctgatccg gacctgcctg 1080  
 ctgaacgaga caggcgacga gcccttccag tacaagaact ga 1122

<210> 9

<211> 1122

5

<212> ADN

<213> Canis lupus

<400> 9

atggccacct cgcctccag ccacctgaac aagggcatca aacaggtgta catgagcctg 60  
 ccccaggcgc agaaggtgca ggccatgtac atctggatcg acggcaccgg cgagggactg 120  
 cggtgcaaga ccagaaccct ggactccgag cccaaggcgc tggagaact gcccgagtgg 180  
 aactcgacg gctcctccac cttccagtcc gagggctcca actccgacat gtacctgggtg 240  
 cctgcccaca tgttccggga ccctttccgg aaggacccca acaagctggg gttctgagag 300  
 gtgttcaagt acaaccggaa gcccgccgag acaaacctgc ggcatacctg caagcggatc 360  
 atggacatgg tgtccaacca gcacccttgg tttggcatgg aacaggagta caccctgatg 420  
 ggcaccgacg gccaccctt cggctggcct tctaaccgct tccttgccc ccagggcccc 480  
 tactattgtg gcgtgggcgc cgacaaggcc tacggcagag acatcgtgga agcccactac 540  
 cgggcctgcc tgtacgccgg catcaagatc gctggcacca acgccgaagt gatgccgcc 600  
 cagtgggagt tccagatcgg cccttgagag ggcatcgaca tgggcgatca cctgtgggtg 660  
 gcccggttca tctgacacag agtgtgagag gacttcggcg tgatcgccac cttcgacccc 720  
 aagcccatcc cgggcaactg gaacggcgct ggctgccaca ccaacttctc caccaaggcc 780  
 atgcgggaag agaacggcct gaagtacatc gaggaatcca tcgagaagct gtccaagcgg 840  
 caccagtacc acatccgggc ctacgacccc aagggcggcc tggataacgc cagacggctg 900  
 accggctttc acgagacatc caacatcaac gacttctctg cggcgctggc caacagaggc 960  
 gcctccatcc ggatcccccg gaccgtgggc caggaaaaga agggctactt cgaggacaga 1020  
 cggccctccg ccaactgcga cccctttagc gtgaccgagc ccctgatccg gacctgcctg 1080  
 ctgaacgaga caggcgacga gcccttccag tacaagaact ga 1122

10

ES 2 627 954 T3

<210> 10

<211> 1122

<212> ADN

<213> *Cricetulus griseus*

5

<400> 10

```

atggccacct cagcaagttc ccacttgaac aaaaacatca agcaaatgta cttgtgcctg      60
ccccaggggtg agaaagtcca agccatgtat atctggggtg atggtactgg agaaggactg      120
cgctgcacaaa cccgcaccct ggactgtgag cccaagtgtg tagaagagtt acctgagtgg      180
aatTTTgatg gctctagtac ctttcagtct gagggctcca acagtgacat gtatctcagc      240
cctgttgcca tgTTTcggga ccccttccgc agagatccca acaagctggg gttctgtgaa      300
gTTTTcaagt acaaccggaa gcctgcagag accaatttaa ggcactcgtg taaacggata      360
atggacatgg tgagcaacca gcacccttgg tttggaatgg aacaggagta tactctgatg      420
ggaacagatg ggcacccttt tggttggcct tccaatggct ttcctgggcc ccaaggtccg      480
tattactgtg gtgtgggccc agacaaagcc tatggcaggg atatcgtgga ggctcactac      540
cgcgcctgct tgtatgctgg ggtcaagatt acaggaacaa atgctgaggt catgcctgcc      600
cagtgggaat tccaaatagg accctgtgaa ggaatccgca tgggagatca tctctgggtg      660
gcccgtttca tcttgcacgc agtatgtgaa gactttgggg taatagcaac ctttgacccc      720
aagcccattc ctgggaactg gaatggtgca ggctgccata ccaactttag caccaaggcc      780
atgcgggagg agaatggtct gaagcacatc gagggaggcca tcgagaaact aagcaagcgg      840
caccgggtacc acattcgagc ctacgatccc aaggggggcc tggacaatgc ccgtcgtctg      900
actgggttcc acgaaacgtc caacatcaac gacttttctg ctggtgtcgc caatcgcagt      960
gccagcatcc gcattccccg gactgtcggc caggagaaga aaggttactt tgaagaccgc     1020
cgcccctctg ccaattgtga cccctttgca gtgacagaag ccatcgtccg cacatgcctt     1080
ctcaatgaga ctggcgacga gcccttccaa tacaaaaact aa                               1122

```

10

<210> 11

<211> 1113

<212> ADN

<213> *Saccharomyces cerevisiae*

15

<400> 11



ES 2 627 954 T3

atggccgagg cctccatcga aaagaccag atcctgcaga agtacctgga actggaccag 60  
 cggggcagaa tcattgccga gtacgtgtgg atcgacggca ccggcaacct gcggtctaag 120  
 ggccggacc tgaagaagcg gatcacctcc atcgaccagc tgcccagtg gaacttcgac 180  
 ggctcctcca ccaaccaggc ccctggccac gactccgaca tctacctgaa gccctggcc 240  
 tactaccccg accccttcag acggggcgac aacatcgtgg tgctggccgc ctgctacaac 300  
 aacgacggca ccccaacaa gttcaaccac cggcacgagg ccgccaagct gttcgccgcc 360  
 cacaaggacg aggaaatttg gttcggcctg gaacaggagt acaccctgtt cgatatgtac 420  
 gacgacgtgt acggctggcc caagggcggc tatcctgccc ctgagggccc ctactactgt 480  
 ggctggggcg ctggcaaggt gtacgccaga gacatgatcg aggccacta ccgggcctgc 540  
 ctgtacgccg gcctggaaat ctccggcatc aacgccgaag tgatgccctc ccagtgggag 600  
 ttccaggtgg gaccctgcac cggcatcgac atgggacgacc agctgtggat ggcccggtag 660  
 ttctgcacc gggtgccga ggaattcggc atcaagatca gcttccacc caagcccctg 720  
 aagggcgact ggaacggcgc tggatgccac gccaacgtgt ccaccaaga gatgcccag 780  
 cctggcggca ccaagtacat cgagcaggcc atcgagaagc tgtccaagcg gcacgccgag 840  
 cacatcaagc tgtacggctc cgacaacgac atgcccgtga ccggcagaca cgagacagcc 900  
 tccatgaccg ccttctccag cggcgtggcc aaccggggct cctccatccg gatccctaga 960  
 agcgtggcca aagagggcta cggctacttc gaggacagac ggctgcctc caacatcgac 1020  
 ccctacctgg tgacaggcat catgtgcgag acagtgtgcg gcgccatcga caacgccgac 1080  
 atgaccaaag agttcgagag agagtccctc tga 1113

<210> 12

5 <211> 1122

<212> ADN

<213> Xenopus laevis

<400> 12

ES 2 627 954 T3

```

atggccacct cgcctccgc ccagctgtcc aaggccatca agcagatgta cctggaactg      60
cctcagggcg acaaggtgca ggccatgtac atctggatcg acggcaccgg cgagggactg      120
cggtgcaaga ccagaaccct ggactccgag cccaagacca tcgaggacct gcccgagtgg      180
aacttcgacg gtcctccac ccaccagtcc gagggctcca actccgacat gtacctgatc      240
cccgtggcca tgttccggga ccctttccgg cgggaccca acaagctggg gctgtgagag      300
gtgctgaagt acaacagaaa gaccgccgag acaaacctgc ggcacacctg taaccagatc      360
atggacatgg tgggaaacga gcacccttgg tttggcatgg aacaggagta caccctgctg      420
ggcatggacg gccaccctt cggtggcct tccaacggct ttcctggccc ccagggcccc      480
tactattgcg gcgtggggcg caacaaggcc tacggcagag acatcgtgga agcccactac      540
cgggctgcc tgtacgccgg cgtgaagatc gccggcacca acgccgaagt gatgccgcc      600
cagtgggagt tccagatcgg cacatgagag ggcacgaca tgggagacca cctgtggatc      660
gcccggttca tcctgcacag agtgtgagag gacttcggca tcacgtgtc cttcgacccc      720
aagccatca ccggcaactg gaacggcgct ggctgccaca ccaacttctc caccaagtcc      780
atgcgggaag agggcgccct gaagcacatc gaggaatcca tcgagcggct gtccaagcgg      840
cacgagtacc acatcaggat gtacgacccc agaggcggca aggacaacgc cagacggctg      900
accggcttcc acgagacatc ctccatccac gagttctctg ctggcgtggc caacagaggc      960
gcctccatcc ggatccaag actggtggga caggaaaaga agggctactt cgaggacaga     1020
cggcctccg ccaactgca cccttacgct gtgaccgagg ccgtgatccg gacctgcctg     1080
ctgaacgaga caggcgacga gccctggag tacaagaact ga                          1122

```

<210> 13

5 <211> 1065

<212> ADN

<213> Arabidopsis thaliana

<400> 13

ES 2 627 954 T3

atgtccctgc tgtccgacct ggtgaacctg aacctgaccg acgccaccgg caagatcatt 60  
 gccgagtaca tctggatcgg cggctccggc atggacatcc ggtctaaggc cgggaccctg 120  
 cctggccctg tgaccgacct ctccaagctg cccaagtgga actacgacgg ctctccacc 180  
 ggcacggccg ctggcgagga ctccgaagtg atcctgtatc cacaggccat cttcaaggac 240  
 cctttccgga agggcaacaa catcctggtg atgtgcgacg cctacacccc tgccggcgac 300  
 cccatcccta ccaacaagcg gcacaacgcc gccaaagatct tctcccaccc cgacgtggcc 360  
 aaagaggaac cttggtacgg catcgagcag gagtacaccc tgatgcagaa agacgtgaac 420  
 tggcccatcg gctggcccgt gggcggctat cctggccctc agggacctta ctactgcggc 480  
 gtgggcgccc acaaggccat cggcagagac atcgtggacg cccactacaa ggcctgcctg 540  
 tacgccggca tcggcatctc tggcatcaac ggcaagtga tgcccggcca gtgggagttc 600  
 caggtgggac ccgtggaagg catctcctcc ggcgatcagg tgtgggtggc cagatacctg 660  
 ctggaacgga tcaccgagat ctccggcgtg atcgtgtcct tcgaccccaa gcccggtgcc 720  
 ggcgattgga atggcgctgg cggccactgc aactactcca ccaagacat gcggaacgac 780  
 ggcggcctgg aagtgatcaa gaaggctatc ggcaagctgc agctgaagca caaagagcat 840  
 atcgccgcct acggcgaggg caacgagcgg agactgaccg gcaagcacga gacagccgac 900  
 atcaacacct tcagctgggg cgtggccaac aggggcgctt ctgtgcgctg gggccgggat 960  
 accgagaaag agggcaaggg ctacttcgag gacagacggc ctgcctcaa catggacccc 1020  
 tacgtggtga catccatgat cgccgagaca accatcctgg gctga 1065

<210> 14

<211> 1110

5

<212> ADN

<213> Drosophila melanogaster

<400> 14

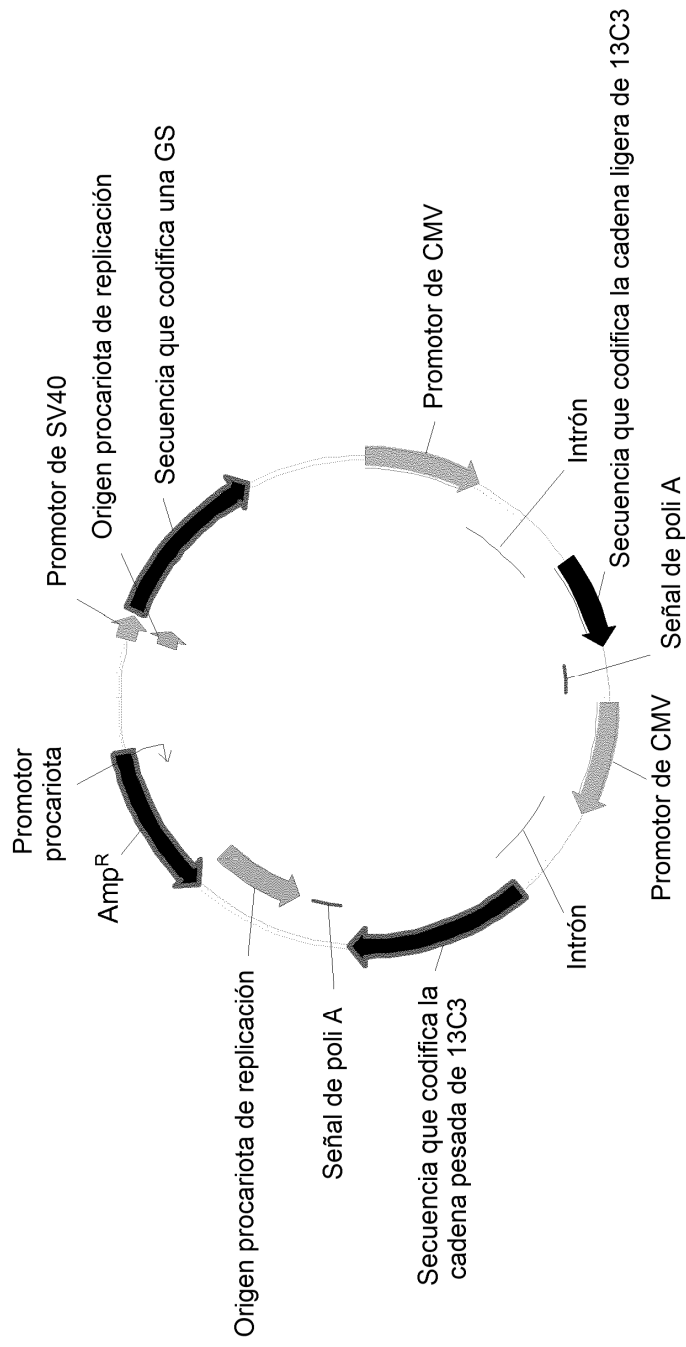
atgtccgccc ggtacctgga agattcccc aacgcccgga tcaacaagac catcctggac 60  
 agatacctga gcctgccact gcaggaaaac atcgtccagg ccacctacgt gtggatcgac 120  
 ggcacagggc aggcactgag gtgcaaggac cggaccctgg acttcatccc ccagtcccc 180

# ES 2 627 954 T3

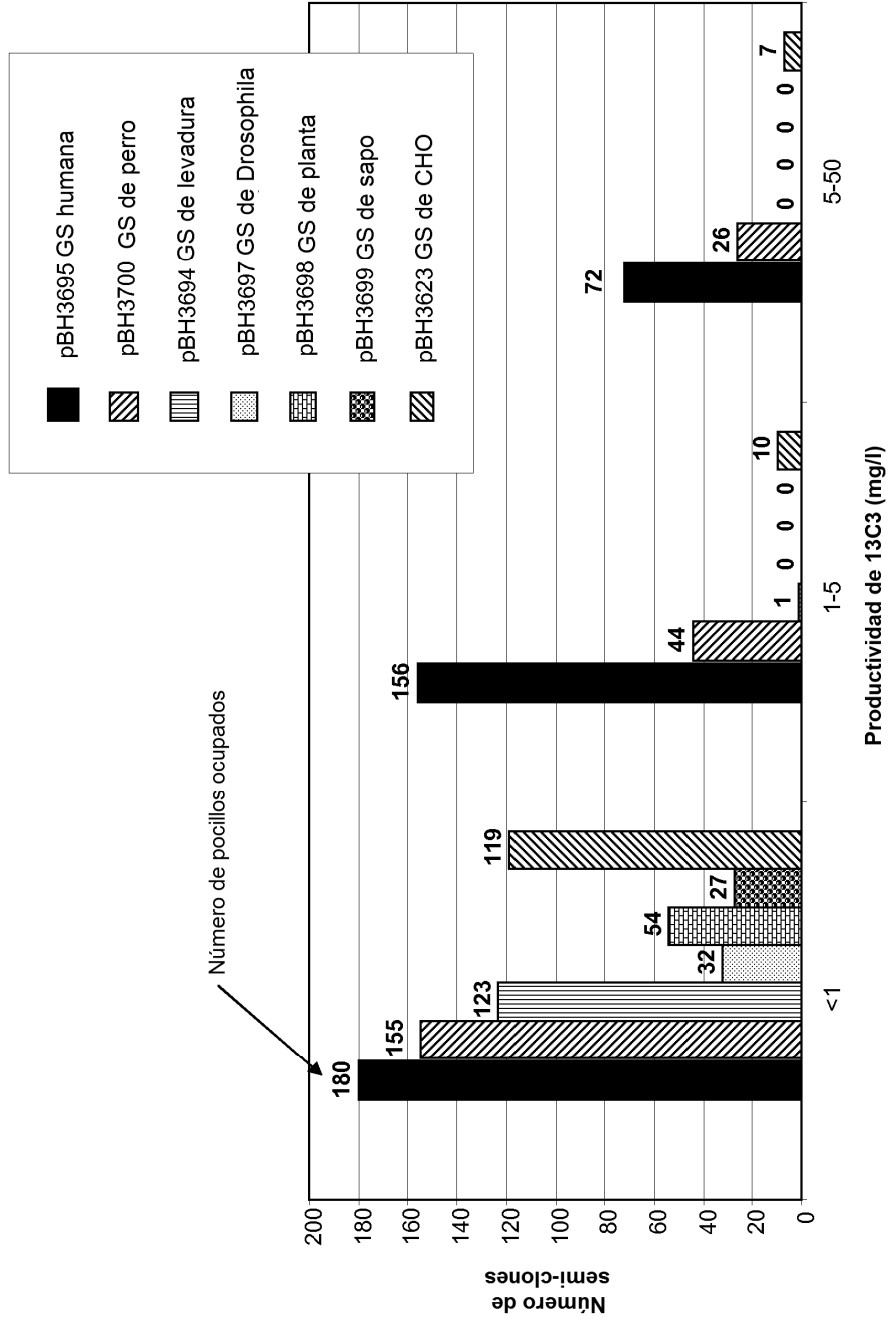
aaagaactgc ccgtgtggaa ctacgacggc tccagctgct accaggccga gggctccaac	240
tccgacacct acctgtacct cgtggccatc tacaaggacc ccttcagacg gggcaacaac	300
atcctggtga tgtgcgacac ctacaagttc gatggcacc ccaccgacac caacaagaga	360
aagacctgcc tggagtggc caacaagtgt gccgccgagg aaccttggtt tggcatcgag	420
caggagtaca ccttcctgga cttcgacggc caccctctgg gctggcccaa gaatggcttt	480
cctggacccc agggccccta ctattgcggc gtgggcgcca acaagtgta cgccagagac	540
atcgtggacg cccactaccg ggcctgcctg tacgccgaa tcaagtgtc cggcaccaac	600
gccgaagtga tgcccgccca gtgggagttc cagggtggac cttgagagg catctccatc	660
ggcgacgacc tgtggatggc ccggttcctg ctgcaccgga tctccgagga attcggcatc	720
gtgtccacc tggaccccaa gcccatgcc ggcgattgga atggcgctgg cgcccacacc	780
aacgtgtcca ccaaggccat gagagaggac ggcggcatcc gggacatcga gaaggcctg	840
gccaaagtgt ccaagtgcc cgagcggcac atccgggcct acgaccctaa gcagggccag	900
gacaacgcc gacggctgac cggcaagcac gagacatcct ccatcaacga cttctctgcc	960
ggcgtggcca accggggctg ctccatcaga atccctcggg gcgagaacga cgacggcaag	1020
ggctacttcg aggacagac gccctcctcc aactgcgacc cttactccgt ggtggaagcc	1080
atcctgcgga ccatctgcct ggacgagtga	1110

**REIVINDICACIONES**

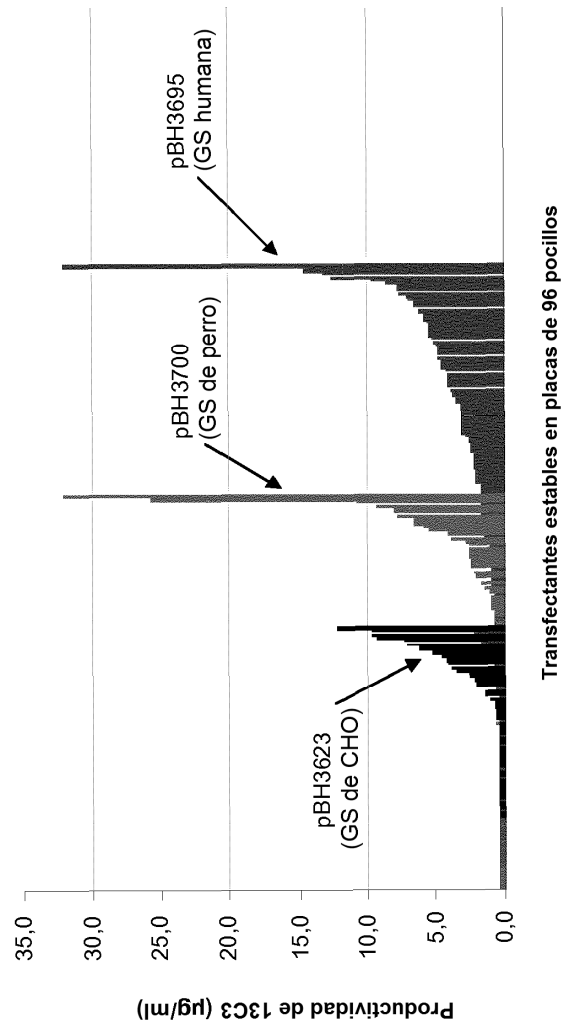
- 5 1. Una línea de células de ovario de hámster chino (CHO) que comprende un vector de expresión de ácido desoxirribonucleico (ADN), y en la que dicho vector comprende una secuencia de nucleótidos que codifica una glutamina sintetasa (GS) de mamífero heteróloga bajo el control de un promotor del virus 40 vacuolante de simio (SV40), que incluye el potenciador de SV40 y al menos un casete de expresión para expresar una proteína recombinante, en la que dicha GS comprende una secuencia al menos el 94,5 % idéntica a la secuencia de SEQ ID NO: 1 o a la secuencia de SEQ ID NO: 2.
2. La línea de células CHO según la reivindicación 1, en la que dicha GS es una GS humana.
- 10 3. La línea de células CHO según la reivindicación 2, en la que dicha GS comprende una secuencia de SEQ ID NO: 1.
4. La línea de células CHO según la reivindicación 1, en la que dicha GS es una GS de perro.
5. La línea de células CHO según la reivindicación 4, en la que dicha GS de perro comprende una secuencia de SEQ ID NO: 2.
- 15 6. La línea de células CHO según una cualquiera de las reivindicaciones 1 a 5, en la que dicha proteína recombinante es un anticuerpo monoclonal.
7. La línea de células CHO según la reivindicación 6, en la que dicho vector comprende un primer casete de expresión adecuado para la clonación de una cadena ligera del anticuerpo, y un segundo casete de expresión adecuado para la clonación de una cadena pesada del anticuerpo.
- 20 8. La línea de células CHO según una cualquiera de las reivindicaciones 1 a 7, en la que dicha línea de células CHO es la línea celular depositada con el N.º CCL-61 en la ATCC.
9. La línea de células CHO según una cualquiera de las reivindicaciones 1 a 8, caracterizada por que permite obtener clones que producen al menos 1 mg/l de proteína recombinante tras la transfección de dicho vector en la línea de células CHO depositada con el N.º CCL-61 en la ATCC.
- 25 10. Una combinación de:
  - i) una línea de células de ovario de hámster chino (CHO); y
  - ii) un vector de ADN (ácido desoxirribonucleico) adecuado para la producción de una proteína recombinante, en la que dicho vector comprende una secuencia que codifica una glutamina sintetasa (GS) de mamífero heteróloga bajo el control de un promotor del virus 40 vacuolante de simio (SV40) que incluye el potenciador de SV40, en la que dicha GS comprende una secuencia al menos el 94,5 % idéntica a la secuencia de SEQ ID NO: 1 o a la secuencia de SEQ ID NO: 2.
- 30 11. Un kit que comprende la combinación según la reivindicación 10.
12. Un método *in vitro* de producción de una proteína recombinante que comprende las etapas de:
  - a) proporcionar una línea de células CHO según una cualquiera de las reivindicaciones 1 a 9;
  - 35 b) cultivar dicha línea de células CHO obtenida en condiciones adecuadas para la producción de la proteína recombinante; y
  - c) aislar y/o purificar dicha proteína recombinante.
13. El método según la reivindicación 12, que comprende además la etapa de formular dicha proteína recombinante en una composición farmacéutica.
- 40 14. Uso de la línea de células CHO según una cualquiera de las reivindicaciones 1 a 9, o de la combinación según la reivindicación 10, para producir una proteína recombinante *in vitro*.



**Figura 1**

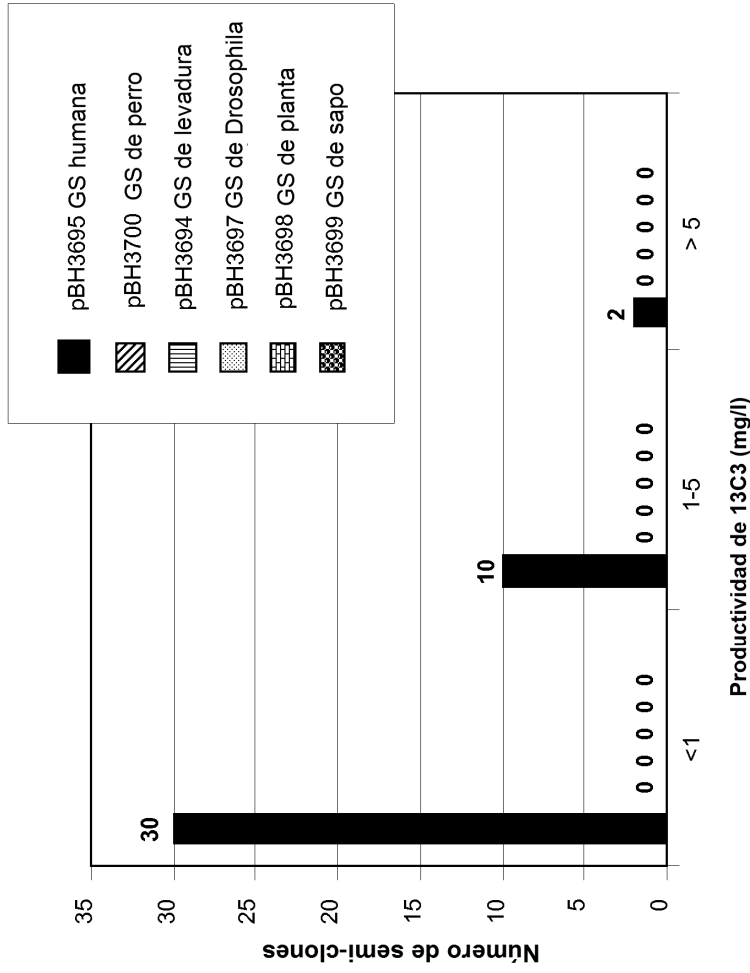


**Figura 2**



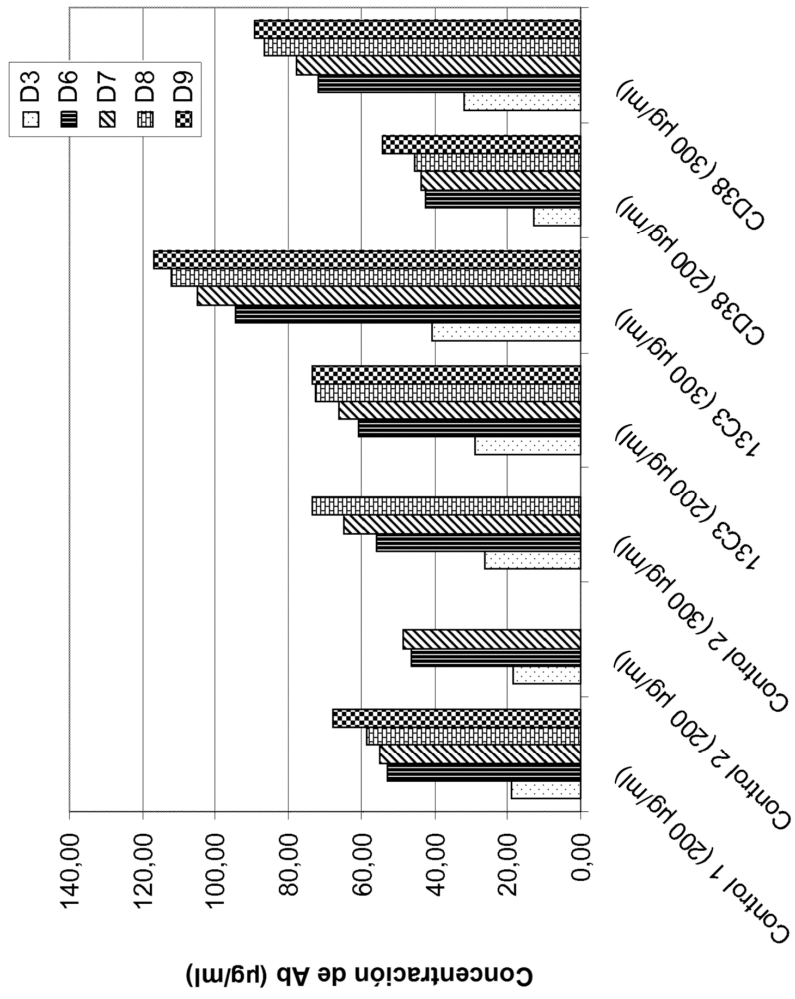
**Figura 3**





**Figure 4**





**Figura 6**