

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 468 317**

51 Int. Cl.:

C12N 15/90 (2006.01)

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **27.03.2009 E 09725747 (1)**

97 Fecha y número de publicación de la concesión europea: **05.03.2014 EP 2257631**

54 Título: **Métodos y materiales para la generación reproducible de líneas de células con alta producción de proteínas recombinantes**

30 Prioridad:

28.03.2008 US 40315 P
28.03.2008 EP 08006023

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

16.06.2014

73 Titular/es:

CELONIC AG (100.0%)
Eulerstrasse 55
4051 Basel, CH

72 Inventor/es:

HERRMANN, ANDREAS;
ABTS, HARRY y
GREULICH, BENEDIKT

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 468 317 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

Métodos y materiales para la generación reproducible de líneas de células con alta producción de proteínas recombinantes

1. Campo de la Invención

- 5 La invención pertenece al campo de la obtención de productos génicos recombinantes en células eucariotas. Particularmente, la invención se refiere a métodos y materiales para la generación rápida y reproducible de líneas de células productoras eucariotas de alto rendimiento adecuadas para producción en gran escala de productos génicos recombinantes.

2. Antecedentes de la Invención

- 10 La línea de células CHO es una de las células hospedadoras más ampliamente utilizadas para obtención de productos biofarmacéuticos. Clásicamente, la célula hospedadora se transfecta con un constructo de expresión que contiene un gen de interés (GOI) que codifica un producto génico recombinante, v.g. una proteína, y un marcador de selección tal como un antibiótico, v.g. un gen de resistencia a neomicina o metotrexato. En un pequeño porcentaje de las células hospedadoras el constructo de expresión se introduce aleatoriamente en el genoma por
15 recombinación no homóloga. Estas células transfectadas de manera estable se seleccionan bajo presión de selección del compuesto antibiótico. Para obtener una línea de células derivada de una sola célula cumplidora de las reglamentaciones, se realiza subsiguientemente un paso de subclonación por dilución limitante. Aunque se analizan varios centenares de clones derivados de células simples, el nivel de expresión era en muchos casos todavía débil, lo que requiere la optimización de los procedimientos de transfección o la amplificación de copias del constructo de
20 expresión integradas inicialmente, v.g., por aumento de la concentración de metotrexato u otro sistema de amplificación. El procedimiento completo requiere al menos 6 meses, a veces hasta un año.

- El problema principal con este enfoque es la naturaleza aleatoria del evento de integración. El DNA genómico circundante tiene un efecto de gran importancia sobre la actividad de traducción del constructo de expresión integrado ("efecto de posición"). Debido a la naturaleza estocástica de este proceso de integración no homóloga, la
25 integración no puede dirigirse a loci con una actividad de transcripción alta (puntos calientes de transcripción). Dado que la mayor parte del DNA de mamífero es no codificante e incluso la parte codificante es sólo transcripcionalmente activa en pequeña proporción, la mayoría de los eventos de integración conducirán a una expresión baja del transgén. En suma, la integración aleatoria de un constructo de expresión conducirá a una gran variación en los niveles de expresión del transgén, exhibiendo sólo una fracción de las células un nivel de expresión alto.

- 30 Con objeto de soslayar el laborioso proceso de cribado necesario y acortar el tiempo para la generación de una línea de células altamente productora, sería deseable un sistema para la inserción direccionada de cualquier GOI en un punto caliente de expresión del genoma de una célula hospedadora.

- A fin de introducir repetitivamente un GOI en este sitio, el sitio identificado puede suplementarse con secuencias que permitan la inserción por recombinasas específicas de secuencia tales como Cre o FLP. El uso de FLP para tales
35 modificaciones se describe en WO 92/15694 y con modificaciones para el intercambio de casetes de expresión de DNA repetitivas exentas de marcadores en US 2001/0032341. El uso de Cre para la modificación del genoma en células vegetales se describió en WO 91/09957 y con modificaciones adicionales en US 2006/0014264A1.

- 40 Ambas recombinasas son dependientes de sitios de reconocimiento específicos que tienen que introducirse primeramente en el genoma en una localización genómica apropiada. En el caso de animales transgénicos, esto se consigue por recombinación homóloga.

- Con objeto de utilizar estas recombinasas específicas para la generación de una línea de células de producción, el problema principal consiste en identificar un gen adecuado para inserción de sitios de reconocimiento por recombinación homóloga. Este gen precisa ser transcrito fuertemente, pero al mismo tiempo no debería ser esencial para la célula. El locus de inmunoglobulina en una célula de mieloma o hibridoma sería un ejemplo de un locus
45 apropiado de este tipo. La producción de anticuerpos por uso de una combinación de recombinación homóloga y recombinación específica de secuencia utilizando Cre se describe en WO 96/30498. Un enfoque similar se describe como sistema de expresión general de alto rendimiento en EP 11405908A1.

- Aunque aplicables, estos métodos están restringidos a genes celulares conocidos, no esenciales, y fuertemente expresados en la célula hospedadora respectiva. Adicionalmente, la recombinación homóloga espontánea es un
50 evento raro y su eficiencia depende del sitio respectivo. Para alcanzar una eficiencia alta tienen que utilizarse secuencias homólogas grandes, siendo sumamente eficaces las secuencias isogénicas. Por tanto, el uso de este enfoque en hospedadores de los que se dispone de información limitada de secuencia (como el hámster) limitará su aplicabilidad.

- 55 Con objeto de identificar loci genómicos aleatoriamente que permitan un alto nivel de expresión, se ha descrito la introducción en el sitio diana de un vector que contiene un gen informador y los sitios de reconocimiento requeridos para la recombinasa utilizada (WO 2004/029284A3. Si bien el gen informador permite la caracterización de la

capacidad de expresión del sitio de integración direccionado aleatoriamente, los sitios de reconocimiento permiten el intercambio del informador contra el GOI.

Puttini et al. (J. Biotechnol. 116 (2005), 145-151) dan a conocer un método para transgénesis direccionada mediada por rotura de la doble cadena en una línea de células. De este modo, se obtienen células hospedadoras que expresan un gen diana. El vector de direccionamiento comprende un solo sitio de reconocimiento de meganucleasas. Se informa de una eficiencia de direccionamiento baja de $2,5 \times 10^{-6}$ (10 clones/ 2×10^7 células con 20% de eficiencia de transfección). Así pues, este documento no proporciona un método fiable para obtener células fuertemente productoras.

WO 2004/029284 describe un método para la generación de una célula hospedadora productora que utiliza una recombinación específica del sitio catalizada por recombinasa FLP. Las líneas de células de expresión resultantes contienen todavía sitios de reconocimiento de recombinasa FLP, en algunos casos con la secuencia codificante después de introducción del gen de interés. Esto da como resultado una estabilidad reducida de la línea de células productoras. El documento sugiere también intrones del grupo II como sitios de reconocimiento de recombinación alternativos, pero sin describir evidencia experimental para ello. No se describen endonucleasas buscadoras codificadas por intrones del grupo I de sitios de reconocimiento de meganucleasas.

Sorrel & Kolb (Biotech. ADV 23 (2005), 431-469) describen una modificación específica de sitio de genomas de mamífero por recombinación homóloga utilizando vectores de direccionamiento que tienen un solo sitio de reconocimiento para la meganucleasa I-Sce-I. La integración de la diana/vector de reemplazamiento ocurre por recombinación homóloga direccionada y no por integración aleatoria.

Belfort & Roberts (Nucleic Acids Res. 25 (1997), 3379-3388) describen endonucleasas buscadoras que incluyen moléculas codificadas por intrones de grupo I y grupo II.

Con todos estos enfoques, la eficiencia de la reacción de intercambio depende de la eficiencia de la recombinasa utilizada. Adicionalmente, el locus seleccionado aleatorio podría tener una actividad de transcripción satisfactoria, pero podría no soportar una recombinación eficiente por la recombinasa utilizada. Con objeto de identificar un locus de integración apropiado, este enfoque requerirá un esfuerzo de cribado adicional. Un obstáculo potencial adicional en el uso de tales recombinasas es la presencia de sitios de pseudo-reconocimiento dentro del genoma de la célula hospedadora deseada. Estos sitios de pseudo-reconocimiento podían conducir a eventos de recombinación indeseables. Esto conducirá a una eficiencia reducida del sistema y a células genéticas potencialmente modificadas con fenotipo indeseable. Una vez más, tiene que realizarse un cribado para identificar las células más adecuadas. Aunque este enfoque parece funcionar, su eficiencia y aplicabilidad universal en la célula productora de alto rendimiento será baja.

Considerado en su conjunto, existe todavía necesidad de un sistema rápido, eficiente y universal para la generación de una línea de células de producción eucariotas de alto rendimiento con capacidades demostradas para la producción en gran escala de proteínas recombinantes.

La presente invención describe métodos y material a fin de satisfacer esta demanda.

3. Sumario de la Invención

La presente invención describe métodos y materiales relativos a una estrategia de ... para la generación de líneas de células eucariotas de alto rendimiento para obtención de productos biofarmacéuticos que implican la introducción específica de sitio de una casete de expresión que comprende un GOI en una línea de células hospedadora preformada y caracterizada (la célula ...). la célula comodín se ha seleccionado para propiedades necesarias para producción en gran escala en sistemas de cultivo dinámicos. La misma se genera por introducción aleatoria de una casete informadora direccionada en el genoma de una línea de células de partida y por caracterización de la capacidad de expresión del sitio de integración. Por intercambio del informador frente a un gen de interés (GOI), se obtiene una línea de células productora capaz de expresar eficientemente un producto génico deseado.

Como fuente para el GOI se utiliza un vector de intercambio, que no permite una expresión fuerte del GOI por sí mismo y preferiblemente no confiere una propiedad seleccionable a las células transfectadas.

Al contrario de los sistemas descritos, la estrategia de factor comodín no utiliza recombinasas a fin de mediar el intercambio de casete de expresión. El intercambio del primer gen informador y marcador de selección está mediado por recombinación homóloga inducida por rotura de la doble cadena (DSB). A fin de inducir la DSB en la posición genómica deseada, el locus direccionado contiene como marcador específico uno o dos sitios de reconocimiento para una enzima rara mediadora de DSB, v.g. una meganucleasa de corte raro o endonucleasa de ... (HE) tal como I-SceO. La HE se introduce en las células confiere la restricción de los sitios de reconocimiento en la posición definida. La reparación de rotura DSB es realizada a una frecuencia elevada por los mecanismos de reparación endógenos de la célula. El vector de intercambio transfectado paralelamente sirve como matriz de reparación debido a secuencias homólogas a las regiones flanqueantes de la casete intercambiada.

Una nueva célula hospedadora de este tipo facilita la clonación de un GOI en puntos calientes de expresión genómica predeterminados en el transcurso de un tiempo breve a fin de obtener línea de células reproducibles de producción altamente eficiente para productos génicos recombinantes, en particular para proteínas recombinantes.

El método incluye los pasos siguientes:

- 5
- Integración aleatoria de un vector diana que contiene un marcador seleccionable, un gen informador cuantificable y sitios de reconocimiento para mediación de la recombinación homóloga inducida por DSB.
 - Selección de una línea de células hospedadoras universal (célula "comodín") preferentemente con una sola copia integrada del vector diana en un locus transcripcionalmente activo, a juzgar por el gen informador.
- 10
- Intercambio del gen informador y el primer marcador seleccionable contra un gen de interés (GOI) y generación simultánea de al menos un marcador seleccionable adicional por recombinación homóloga inducida por DSB que implica un vector de intercambio con los elementos requeridos como matriz de reparación.

15 El proceso de intercambio en la célula hospedadora comodín universal puede realizarse para otros GOIs y suministrará cada vez una célula de producción con niveles de expresión predecibles altos, calidad del producto (es decir glicosilación) y propiedades de crecimiento que permita producción en gran escala del producto génico, v.g. de una proteína.

20 El método permite preferiblemente la producción de cualquier proteína recombinante. La proteína producida podría ser una enzima, en particular una proteasa, inhibidor de proteasas, hormona, citoquina, receptor con y sin dominios transmembranales o intracelulares (es decir fijado a la membrana o soluble), anticuerpos de longitud total o dominios de anticuerpos. En particular, la proteína producida podría asemejarse a una proteína de fusión que combine partes o dominios de las proteínas mencionadas.

El presente sistema permite por tanto la generación de células altamente productoras necesarias para la producción eficiente y económica de productos biofarmacéuticos.

El mismo ofrece varias ventajas:

- 25
- ahorro de tiempo: podría obtenerse una línea de células de producción estable en el transcurso de unas pocas semanas partiendo de una línea de células comodín prefabricada.
 - Debido a la integración en puntos calientes de expresión preidentificados, el nivel de expresión de la línea de células de producción puede predecirse de modo más fiable y podría llegar a ser reproducible para diferentes GOIs.
 - Alta eficiencia dado que únicamente se obtendrán células que expresen el GOI después de intercambio.
 - Ausencia total de integración adicional de secuencias vectoras eliminables durante la integración específica del sitio.
- 30
- Producción de material para prueba de experimentos de concepto podría realizarse en el transcurso de pocos meses. No habrá cambio alguno en la calidad del producto en fases posteriores del desarrollo debido a que estará disponible una población de células homogénea y genéticamente idénticas desde el mismo comienzo.

35 Así pues, un primer aspecto de la presente invención es un método para la generación de una célula hospedadora productora para la obtención de productos génicos recombinantes, particularmente polipéptidos, que comprende los pasos

(a) proporcionar un vector diana que comprende

- (i) un gen informador (RG1) enlazado operativamente a una primera secuencia de control de la expresión (P1) que comprende un promotor constitutivo (P1),
- 40 (ii) un primer gen marcador de selección (SM1) enlazado operativamente a una segunda secuencia de control de la expresión (P2),
- (iii) un segundo gen marcador de selección no funcional (SM2) sin enlace operativo a una secuencia de control de la expresión,

45 en donde la secuencia (i) está localizada en la posición 5', la secuencia (ii) está localizada entre (i) y (iii) y la secuencia (iii) está localizada en la posición 3',

(iv) un primer y un segundo sitio de reconocimiento para una enzima mediadora de rotura de la doble cadena, en donde el primer sitio de reconocimiento está localizado entre la primera secuencia de control de la expresión (P1) y el primer gen informador (RG1) y el segundo sitio de reconocimiento está localizado entre la secuencia (ii) y la secuencia (iii),

(b) introducir el vector diana en una célula hospedadora en condiciones tales que permiten la integración aleatoria del vector diana en el genoma de la célula hospedadora,

5 (c) determinar cuantitativamente el nivel de expresión del gen informador (RG1) y seleccionar una célula hospedadora que tenga integrado de manera estable el vector diana y que exhiba una actividad de transcripción del gen informador (RG1), que es al menos 10 veces mayor comparada con los niveles de expresión medios de la población de clones resultante,

(d) proporcionar un vector de intercambio que comprende:

(i) un gen de interés (GOI), y

10 (ii) un segundo gen marcador de selección inactivo (Δ SM2) enlazado operativamente a una tercera secuencia de control de la expresión (P3),

en donde el vector de intercambio comprende una primera secuencia homóloga 5' y una segunda secuencia homóloga 3', que permiten la recombinación con secuencias del vector diana,

15 (e) introducir el vector de intercambio en una célula hospedadora obtenida en el paso (c) en condiciones que permiten una rotura de la doble cadena en el primer y/o segundo sitio de reconocimiento, preferiblemente en el primer y segundo sitio de reconocimiento, como se define en (a) (iv) y una integración del vector de intercambio en el genoma de la célula hospedadora por recombinación homóloga mediada por rotura de la doble cadena,

con lo cual el segundo marcador de selección inactivo (Δ SM2) se activa por integración del vector de intercambio,

(f) seleccionar una célula productora que tiene integrado el vector de intercambio por recombinación homóloga con el vector diana integrado, en donde la célula productora expresa el GOI.

20 Un aspecto adicional es un método para generar una célula hospedadora universal (célula comodín) que permite la reproducción rápida de la línea de células de producción de rendimiento alto adecuadas para producción en gran escala de productos génicos recombinantes. Este método comprende:

(a) Proporcionar un vector diana que comprende:

25 (i) un gen informador (RG1) enlazado operativamente a una primera secuencia de control de la expresión (P1),

(ii) un primer gen marcador de selección (SM1) enlazado operativamente a una segunda secuencia de control de la expresión (P2),

(iii) un segundo gen marcador de selección no funcional (SM2) sin enlace operativo a una secuencia de control de la expresión,

30 En donde la secuencia (i) está localizada en la posición 5', la secuencia (ii) está localizada entre (i) y (iii) y la secuencia (iii) está localizada en la posición 3'.

35 (iv) un primer y un segundo sitio de reconocimiento para una enzima mediadora de rotura de la doble cadena, particularmente una meganucleasa, en donde el primer sitio de reconocimiento está localizado entre la primera secuencia de control de la expresión (P1) y el primer gen informador (RG1), y el segundo sitio de reconocimiento está localizado entre la secuencia (ii) y la secuencia (iii),

(b) Introducir el vector diana en una célula hospedadora en condiciones tales que permiten la integración aleatoria del vector diana en el genoma de la célula hospedadora, y

(c) Seleccionar una célula hospedadora que tiene integrado de manera estable el vector diana y que exhibe la actividad de transcripción del RG1,

40 Un aspecto adicional es un método para la obtención de productos génicos recombinantes, que comprende los pasos

(a) Producir una célula hospedadora comodín para introducción de un gen de interés (GOI) proporcionando un vector diana que comprende:

45 (i) un gen informador (RG1) enlazado operativamente a una primera secuencia de control de la expresión que comprende un promotor constitutivo (P1),

(ii) un primer gen marcador de selección (SM1) enlazado operativamente a una segunda secuencia de control de la expresión (P2),

(iii) un segundo gen marcador de selección no funcional (SM2) sin enlace operativo a una secuencia de control de la expresión,

en donde la secuencia (i) está localizada en la posición 5', la secuencia (ii) está localizada entre (i) y (iii) y la secuencia (iii) está localizada en la posición 3',

5 (iv) un primer y un segundo sitio de reconocimiento para una enzima mediadora de rotura de la doble cadena, particularmente una meganucleasa, en donde el primer sitio de reconocimiento está localizado entre la primera secuencia de control de la expresión (P1) y el primer gen informador (RG1), y el segundo sitio de reconocimiento 5 está localizado entre la secuencia (ii) y la secuencia (iii),

10 (b) Introducir el vector diana en una célula hospedadora en condiciones tales que permiten la integración aleatoria del vector diana en el genoma de la célula hospedadora, y

(c) Seleccionar una célula hospedadora comodín que tiene integrado de manera estable el vector diana y que exhibe la actividad de transcripción del gen informador (RG1) que es al menos 10 veces mayor comparada con los niveles de expresión medios de la población de clones resultante,

(d) Proporcionar un vector de intercambio que comprende:

15 (i) un gen de interés, y

(ii) un segundo gen marcador de selección inactivo (Δ SM2) enlazado operativamente a una tercera secuencia de control de la expresión (P3),

20 en donde el vector de intercambio comprende una primera secuencia homóloga 5' y una segunda secuencia homóloga 3', que permiten la recombinación con secuencias del vector diana, integrado en la célula hospedadora,

(e) Introducir el vector de intercambio en la célula hospedadora proporcionada en el paso (a) en condiciones que permiten una rotura de la doble cadena en el primer y/o segundo sitio de reconocimiento del vector diana y una integración del vector de intercambio en el genoma de la célula hospedadora por recombinación homóloga mediada por rotura de la doble cadena,

25 en donde el segundo marcador de selección inactivo (Δ SM2) está activado por una integración del vector de intercambio,

(f) Seleccionar una célula productora que tiene integrado el vector de intercambio por recombinación homóloga con el vector diana integrado, en donde la célula productora expresa el GOI.

30 El método adicional es un método para la generación de una célula hospedadora productora para la obtención de productos génicos recombinantes que comprende los pasos

(a) Introducir una secuencia de ácido nucleico que comprende un primer y un segundo sitio de reconocimiento para una enzima mediadora de rotura de la doble cadena por recombinación homóloga direccionada en un sitio del genoma de una célula hospedadora que soporta actividad de transcripción,

35 (b) Seleccionar una célula hospedadora comodín que tiene integrados de manera estable el primer y el segundo sitio de reconocimiento

(c) Proporcionar un vector de intercambio que comprende

(i) un gen de interés (GOI), y

(ii) una primera secuencia homóloga 5' y una segunda secuencia homóloga 3', que permiten la recombinación con secuencias en el sitio de integración del primer y/o segundo sitio de integración.

40 (d) Introducir el vector de intercambio en una célula hospedadora obtenida en el paso (b) en condiciones que permiten una rotura de la doble cadena en el primer y/o segundo sitio de reconocimiento definido en (a) y una integración del vector de intercambio en el genoma de la célula hospedadora por recombinación homóloga mediada por rotura de la doble cadena, y

45 (e) Seleccionar una célula productora que tiene integrado el vector de intercambio por recombinación homóloga, en donde la célula productora expresa el GOI.

Un aspecto adicional es un método para la producción de una célula hospedadora comodín para introducción de un gen de interés GOI, que comprende los pasos:

(a) Introducir una secuencia de ácido nucleico que comprende un primer y un segundo sitio de reconocimiento para una enzima mediadora de rotura de la doble cadena por recombinación homóloga direccionada en un sitio del genoma de una célula hospedadora que soporta actividad de transcripción,

5 (b) Seleccionar una célula hospedadora comodín que tiene integrados de manera estable el primer y el segundo sitio de reconocimiento.

Un aspecto adicional es un método para la obtención de productos génicos recombinantes:

(a) Proporcionar una célula hospedadora comodín que tiene integrada en su genoma una secuencia de ácido nucleico que comprende un primer y un segundo sitio de reconocimiento para una enzima mediadora de rotura de la doble cadena,

10 En donde la secuencia de ácido nucleico está integrada de manera estable en el genoma de la célula hospedadora en un sitio que soporta actividad de transcripción,

(b) Proporcionar un vector de intercambio que comprende

(i) un gen de interés (GOI), y

15 (ii) una primera secuencia homóloga 5' y una segunda secuencia homóloga 3', que permiten la recombinación con secuencias en el sitio de integración del primer y/o segundo sitio de integración.

(c) Introducir un vector de intercambio en una célula hospedadora proporcionada en el paso (a) en condiciones que permiten una rotura de la doble cadena en el primer y/o segundo sitio de reconocimiento definido en (a) y una integración del vector de intercambio en el genoma de la célula hospedadora por recombinación homóloga mediada por rotura de la doble cadena y

20 (d) Seleccionar una célula productora que tiene integrado el vector de intercambio por recombinación homóloga, en donde la célula productora expresa el GOI.

Un aspecto adicional es un vector diana que comprende

(a) Un gen informador (RG1) enlazado operativamente a una primera secuencia de control de la expresión (P1),

25 (b) Un primer gen marcador de selección (SM1) enlazado operativamente a una segunda secuencia de control de la expresión (P2),

(c) Un segundo gen marcador de selección no funcional (SM2) sin enlace operativo a una segunda secuencia de control de la expresión,

En donde la secuencia (i) está localizada en la posición 5', la secuencia (ii) está localizada entre (i) y (iii) y la secuencia (iii) está localizada en la posición 3',

30 Un primer y un segundo sitio de reconocimiento para una enzima mediadora de rotura de la doble cadena, particularmente una meganucleasa, en donde el primer sitio de reconocimiento está localizado entre la primera secuencia del control de la expresión (P1) y el primer gen informador (RG1) y el segundo sitio de reconocimiento está localizado entre la secuencia (ii) y la secuencia (iii).

Un aspecto adicional es un vector de intercambio que comprende

35 (i) gen de interés (GOI), y

(ii) una segundo gen marcador de selección inactivo (Δ SM2) enlazado operativamente a una tercera secuencia de control de la expresión (P3),

40 En donde el vector de intercambio comprende una primera secuencia homóloga 5' y una secuencia homóloga 3', que permite la recombinación con sitios predeterminados en el genoma en una célula hospedadora, v.g. con secuencias del vector diana.

Aspectos adicionales se refieren a células hospedadoras comodín universales para introducción y expresión eficiente de un gen de interés que pueden obtenerse por los métodos arriba descritos y células de producción de alto rendimiento para la producción en gran escala de producto génicos recombinantes producidas por utilización de tales células hospedadoras comodín utilizando los métodos descritos.

45 4. Descripción Detallada de la Invención

4.1 células Hospedadoras

La presente invención se refiere a la generación de una célula hospedadora productora o comodín. Preferiblemente, la célula hospedadora es una célula eucariota, v.g. una célula de levadura, una célula fúngica o una célula de vertebrado tal como una célula de insecto o una célula de mamífero. Más preferiblemente, la célula hospedadora es una célula de mamífero, v.g. una célula de roedor tal como una célula de ratón, rata o hámster o una célula de primate tal como una célula humana. En una realización especialmente preferida, la célula hospedadora es una célula de ovario de hámster chino (CHO). Célula preferidas adicionales son células NSO, células de hibridoma, v.g. de ratón o ratón/humanas, células HEK293 o células Per-C6. En una realización particular, la célula hospedadora es capaz de crecer a densidades de células altas en condiciones exentas de suero en cultivo de suspensión.

4.2 Sistemas Vectores

- 10 La presente invención implica el uso de diferentes vectores, a saber un vector diana para introducir un casete de gen informador en una célula hospedadora a fin de obtener una célula hospedadora comodín y un vector de intercambio o integración a fin de introducir un gen de interés (GOI) en una célula hospedadora comodín a fin de generar una célula productora. Opcionalmente se utiliza un tercer vector para la expresión de una enzima mediadora de rotura de la doble cadena de la célula hospedadora.
- 15 Los vectores se seleccionan a fin de ser compatibles con la célula hospedadora respectiva. Así, los vectores son particularmente adecuados para células eucariotas. Preferiblemente, los vectores son vectores no virales, v.g. vectores plasmídicos circulares o lineales.

...

Cinta No. 44 (cara B) y 45-46 (cara B)

20 4.2.1 Vector Diana

4.2.1.1 Vector Diana Para Integración Aleatoria

El vector diana de la presente invención comprende preferiblemente un gen informador (RG1) y un primer gen marcador de selección (SM1). Ambos genes forman preferiblemente un casete de expresión bicistónica, en donde el RG1 está enlazado operativamente a una primera secuencia de control de la expresión (P1) y el gen SM1 está enlazado operativamente a una segunda secuencia de control de la expresión (P2). El RG1 está localizado preferiblemente 5' respecto al gen SM1. Adicionalmente, el vector diana comprende un segundo gen marcador de selección no funcional SM2. Adicionalmente, el vector diana está de modo preferible sustancialmente exento de elementos de secuencia procariotas.

30 El RG1 puede ser cualquier gen informador, que expresa un producto génico detectable, v.g. una enzima o un producto génico luminiscente. Preferiblemente, el gen informador es un gen cuya expresión puede monitorizarse e identificarse. Ejemplos típicos de genes informadores son fosfatasa alcalina, v.g. una fosfatasa alcalina secretada (SEAP), luciferasa, β -galactosidasa o proteínas fluorescentes tales como GFP o variantes de las mismas.

El RG1 está enlazado operativamente a la primera secuencia de control de la expresión (P1), que comprende preferiblemente un promotor constitutivo capaz de impulsar la expresión del gen informador en la célula hasta ahora. En una realización, P1 puede comprender un promotor constitutivo fuerte, v.g. un promotor CMV (como el promotor inmediato/precoz con intensificador derivado de virus CMV humano o murino), que puede estar intensificado adicionalmente, v.g. por suplementación con el intrón A de CMV, con lo cual se aumentan la actividad de transcripción y la estabilidad del mRNA. En general, podría utilizarse cualquier secuencia de control que permita una expresión fuerte y constitutiva en una célula eucariota. Dichas secuencias podrían ser existentes naturalmente o combinaciones de elementos existentes naturalmente. Adicionalmente, pueden utilizarse también elementos promotores sintéticos. Ejemplos adicionales de primeras secuencias de control de expresión adecuadas (P1) son el promotor del factor 1 alfa de elongación humano (EF-1 alfa) (con y sin el primer intrón correspondiente), promotores derivados de la LTR del virus del sarcoma de Rous (RSV), o LTR de HIV2 o combinaciones o secuencias derivadas de ellos.

45 En una realización diferente, P1 puede comprender un promotor atenuado cuya fuerza está reducida comparada con la del promotor de tipo salvaje, v.g. un promotor CMV alternado. La fuerza del promotor puede reducirse, v.g. por modificación del efecto intensificador de la expresión de la secuencia del intrón A. Esto puede hacerse por incorporación de secuencias de nucleótidos heterólogos cortas (v.g. hasta 100 pb) en la región 3' del intrón. Estas inserciones no afectan negativamente a la recombinación homóloga con el vector de intercambio. Por selección de secuencias apropiadas en el vector de intercambio, la versión atenuada del promotor puede ser intercambiada por la versión plenamente activa cuando se inserta el vector de intercambio en el genoma de la célula hospedadora.

50 ... gen marcador de selección (SVM) puede ser cualquier gen marcador de selección que sea capaz de proporcionar un fenotipo seleccionable en la célula hospedadora. Por ejemplo, SM1 puede ser un gen resistente a antibióticos tales como un gen de resistencia a neomicina o higromicina. Adicionalmente, podrán utilizarse también marcadores seleccionables tales como genes que confieren resistencia a blasticidina, puromicina, ouabaina o el gen de glutamina-sintasa.

- El SM1 es un enlace operativo con una segunda secuencia de control de la expresión (P2), que comprende preferiblemente un promotor constitutivo, v.g. el promotor SV-40. En una realización particular, P2 puede comprender un promotor atenuado, es decir un promotor que es una versión atenuada del promotor de tipo salvaje. Alternativa o adicionalmente, el SM1 puede ser un gen marcador de selección atenuado con una actividad reducida comparado con el gen marcador de selección de tipo salvaje. Por utilización de un promotor atenuado y/o un gen marcador de selección, puede facilitarse la selección de una célula hospedadora que tiene integrado el vector diana en un sitio de integración con actividad de transcripción alta.
- 5
- En una realización alternativa, la expresión de SM1 se consigue por el uso de un elemento IRES situado entre el extremo 3' de la secuencia codificante RG1 y el extremo 5' de la secuencia codificante SM1.
- 10
- Adicionalmente, el vector diana contiene típicamente señales de poliadenilación en el extremo 3' de las secuencias codificantes para RG1 y SM1, v.g. las señales de poliadenilación de la hormona del crecimiento bovina o la señal de poliadenilación de SV-40.
- El vector diana comprende adicionalmente, un segundo gen marcador de selección no funcional (SM2) sin enlace operativo a una secuencia de control de la expresión. El SM2 está localizado fuera, preferiblemente en el lado 3' de la casete de expresión bicistrónica para RG1 y SM1. En una realización particular, la secuencia codificante para el SM2 carece del codón de inicio ATG y/o un codón de parada aguas arriba, preferiblemente con codones de parada aguas arriba en los tres marcos de lectura.
- 15
- El SM2 puede ser cualquier marcador de selección que es capaz de proporcionar un fenotipo seleccionable en la célula hospedadora. Preferiblemente, el SM2 es diferente del SM1. Por ejemplo, el SM2 puede ser un gen de resistencia a antibióticos tal como un gen de resistencia a neomicina o higromicina. Ejemplos adicionales son genes que confieren resistencia a blasticidina, puromicina, ouabaína o el gen de glutamina-sintasa.
- 20
- Con objeto de inducir roturas específicas de la doble cadena dentro del vector diana, el mismo comprende un primer y un segundo sitio de reconocimiento para una enzima mediadora de rotura de la doble cadena, particularmente una meganucleasa. El primer sitio de reconocimiento está localizado entre P1 y la secuencia codificante de RG1, y el segundo sitio de reconocimiento está localizado entre la secuencia codificante para SM1 y la secuencia no-funcional SM2. La parte del vector diana flanqueada por dos sitios de reconocimiento puede designarse como área de intercambio.
- 25
- En una realización particular, el segundo sitio de reconocimiento puede estar localizado entre el codón de parada de SM1 y la señal de poliadenilación siguiente (vector diana A). En una realización adicional, el segundo sitio de reconocimiento puede estar localizado entre la señal de poliadenilación SM1 y el comienzo de la secuencia codificante de SM2 (vector diana B).
- 30
- Los dos sitios de reconocimiento pueden tener la misma orientación o una orientación opuesta uno con respecto al otro. En una realización preferida, la parte del sitio de reconocimiento que corresponde al exón aguas abajo del sitio de tipo salvaje está localizada fuera del área de intercambio después del corte.
- 35
- La enzima mediadora de rotura de la doble cadena es preferiblemente una meganucleasa o nucleasa buscadora, que tiene un sitio de reconocimiento que se encuentra raras veces en el genoma de la célula hospedadora, por ejemplo una nucleasa buscadora codificada por un intrón del grupo I. Preferiblemente, la meganucleasa o nucleasa buscadora tiene un sitio de reconocimiento que no se encuentra naturalmente en el genoma de la célula hospedadora. El sitio de reconocimiento tiene típicamente una longitud de al menos 10 y preferiblemente de al menos 12 nucleótidos. Más preferiblemente, la nucleasa tiene un sitio de reconocimiento de 18 nucleótidos, perteneciendo por tanto a la familia de endonucleasas dodecapeptídicas. Ejemplos de tales enzimas se describen en US 2005/0032223. Preferiblemente, la meganucleasa se selecciona de I-SceI, I-SceII, I-SceIII, I-SceIV, I-CeuI, I-CreI, I-PpoI, I-TevI, I-TevII, I-TevIII, HO y Endo SceI. Más preferiblemente, la meganucleasa es I-SceI.
- 40
- Dos realizaciones del vector diana de la invención se muestran en las figuras 1 (vector diana A) y 2 (vector diana B). El gen informador está enlazado operativamente al promotor CMV intensificado (eCMV) con inclusión del intrón A. El primer gen marcador de selección (SM1) está enlazado operativamente a un promotor de SV40 (Psv40). El segundo gen marcador de selección (SM2) no está enlazado operativamente a un promotor y por consiguiente es no funcional.
- 45
- #### 4.2.1.2 Vector Diana para Integración No-Aleatoria
- 50
- En una realización diferente de la invención, el vector diana puede comprender adicionalmente una primera secuencia homóloga 5' y una segunda secuencia homóloga 3', que flanquean los elementos genéticos del vector diana como se ha descrito arriba y permiten recombinación homóloga no aleatoria en un locus génico predeterminado de la célula hospedadora, es decir un locus génico que se sabe soporta actividad de transcripción, v.g. un locus de inmunoglobulina. En esta realización, el gen informador puede estar enlazado operativamente a una secuencia de control de la expresión como se ha descrito arriba o, alternativamente, puede carecer de una secuencia de control de la expresión si las secuencias homólogas flanqueantes se seleccionan de modo que
- 55

permitan recombinación homóloga en una posición que sitúa el gen informador (RG1) en enlace operativo con una secuencia endógena de control de la expresión.

Alternativamente, puede utilizarse para esta realización un vector diana que carece de un gen informador.

4.2.2 El Vector de Intercambio

5 El vector de intercambio comprende preferiblemente un gen de interés (GOI) y una secuencia codificante inactiva, v.g. incompleta, para un segundo gen marcador de selección (Δ SM2). El gen de interés puede ser cualquier gen que codifique un producto génico deseado, particularmente una proteína recombinante, pero también un ácido nucleico recombinante, v.g. una molécula de RNA deseada. Preferentemente, el GOI está presente sin una secuencia funcional de control de la expresión. Más preferentemente, GOI contiene en su extremo 5' una secuencia parcial de control de la expresión, que es idéntica o sustancialmente idéntica a la parte 3' prima de la primera secuencia de control de la expresión (P1) en el vector diana. Adicionalmente, se prefiere que una señal de poliadenilación apropiada para el GOI esté localizada en el extremo 3' de la secuencia codificante como se describe para el vector diana.

15 La secuencia codificante inactiva para SM2 no permite la generación de un marcador de selección funcional. Para este propósito, típicamente la porción 3' de la secuencia codificante de SM2 puede estar ausente en el vector de intercambio. Sin embargo, la secuencia codificante inactiva para SM2 está enlazada operativamente a una secuencia funcional de control de la expresión, v.g. un promotor constitutivo tal como el promotor SV40.

20 En una realización adicional, el vector de intercambio comprende adicionalmente un tercer gen marcador de selección (SM3), que se cotranscribe con el GOI. Para este propósito, la secuencia codificante para SM3 enlazada operativamente a un elemento IRES puede estar situada entre el extremo 3' de la secuencia codificante para el GOI y el sitio de poliadenilación. Alternativamente, el orden del GOI y SM3 puede estar intercambiado. En una realización adicional, el GOI y SM3 se expresan desde casetes de expresión independientes con elementos de control de la expresión separados (promotores). El SM3 es diferente de SM2 y permite la selección de células con independencia de SM2. Así pues, una doble selección permite el enriquecimiento de células que expresan los dos marcadores de selección. Ambos marcadores de selección pueden seleccionarse, sin limitación, de genes que confieren resistencia a higromicina, neomicina, G418, blastidina, puromicina, ouabaína o el gen de glutamina-sintasa.

La longitud de la secuencia promotora parcial 5' del GOI y la longitud del SM2 incompleto se selecciona de una manera que es posible una recombinación homóloga eficiente entre los elementos homólogos respectivos en el vector diana.

30 Adicionalmente, el vector de intercambio comprende una primera secuencia homóloga 5' y una segunda secuencia homóloga 3', que permiten la recombinación con secuencias del vector diana. La longitud de las secuencias homólogas respectivas es con preferencia al menos aproximadamente 500 nucleótidos, con más preferencia al menos aproximadamente 700 nucleótidos. En una realización especialmente preferida, la secuencia homóloga 5' tiene una longitud de al menos aproximadamente 1000 nucleótidos y la secuencia homóloga 3' tiene una longitud de al menos aproximadamente 700 nucleótidos. La longitud máxima de las secuencias homólogas es con preferencia aproximadamente 2000 nucleótidos. El grado de identidad entre las secuencias homólogas en la diana y el vector de intercambio es preferiblemente al menos 80%, más preferiblemente al menos 90% referido a la secuencia total. De modo muy preferible, las secuencias homólogas son sustancialmente idénticas a fin de obtener una tasa de intercambio máxima. Adicionalmente, se prefiere que la secuencia homóloga 5' en el vector de intercambio no comprenda un promotor funcional. La secuencia homóloga 3' en el vector de intercambio no comprende preferiblemente una secuencia marcadora de selección funcional.

45 Adicionalmente, el vector de intercambio puede comprender un cuarto gen marcador de selección fuera de la región de secuencia definida por la primera secuencia homóloga 5' y la segunda secuencia homóloga 3'. El cuarto gen marcador de selección es un gen marcador de selección negativo, v.g. un gen suicida tal como HSV-timidina-quinasa, que permite selección contra cualquier integración aleatoria del vector de intercambio en el genoma de la célula hospedadora por el cual está integrado también el marcador de selección negativo.

50 Dos realizaciones del vector de intercambio se muestran en figura 3 (vector de intercambio A) y figura 4 (vector de intercambio B). El gen de interés (GOI) está enlazado operativamente al promotor eCMV parcial deletado en 5' (Δ CMV). El segundo marcador de selección (SM2) está enlazado operativamente al promotor SV40 (Psv40). Las secuencias homólogas con el vector diana se indican por líneas de puntos. El vector de intercambio B (Fig. 4) comprende adicionalmente un gen suicida (HCV-dk) fuera de las secuencias homólogas y un tercer marcador de selección co-transmitido con el gen GOI (SM3) enlazado operativamente a un elemento IRES.

55 En una realización alternativa, el vector de intercambio puede comprender un segundo gen marcador de selección activo (SM2) enlazado operativamente a una secuencia de control de la expresión en lugar del segundo gen marcador de selección inactivo.

4.3 Generación de Líneas de células comodín

4.3.1 Integración Aleatoria

A fin de generar una célula hospedadora comodín, puede introducirse un vector diana como se ha descrito arriba en la sección 4.2.1.1 en una célula hospedadora en condiciones que permitan la integración aleatoria del vector diana en el genoma de la célula hospedadora por recombinación no homóloga, es decir aleatoria.

- 5 En una realización particular, el vector diana se ha privado de elementos bacterianos innecesarios antes de la transfección. Esto eliminará un efecto de silenciamiento de genes mediado por elementos de DNA bacteriano.

Las condiciones para la transfección se seleccionan de tal modo que preferentemente se marca un solo sitio de integración. Adicionalmente, se ve favorecida la integración de una sola copia del vector diana.

- 10 La selección de células hospedadoras transfectadas de manera estable se facilita por el uso del SM1, que es típicamente un gen de resistencia a los antibióticos tal como un gen de resistencia a neomicina o higromicina. Así, el procedimiento de selección comprende preferiblemente un crecimiento de las células transfectadas en presencia de un antibiótico contra el cual SM1 media resistencia.

- 15 En un paso siguiente, se selecciona una célula hospedadora que tiene integrado el vector diana y que exhibe una expresión fuerte del gen informador (RG1). Preferiblemente, a este fin se genera una población de células clonal a oligo-clonal por una estrategia de dilución limitante. Preferiblemente, este paso se realiza en condiciones de cultivo exentas de suero utilizando células hospedadoras adaptadas para crecimiento en suspensión.

- 20 En una realización preferida de este procedimiento de selección, se siembran células transfectadas a una densidad que dé como resultado después de la selección una población de células que se deriva preferentemente de 1-5 células. Se expandirán poblaciones de células clonales u oligo-clonales y la capacidad de transcripción del locus genómico aleatorio marcado es analizada por el gen informador (RG1). Preferiblemente, se lleva a cabo un análisis cuantitativo de la expresión de RG1.

- 25 Basándose en los niveles expresados de RG1, se seleccionan clones altamente productores. Con objeto de obtener una población de células hospedadoras derivadas en una sola célula como línea de células comodín potenciales se lleva a cabo preferiblemente una segunda clonación a dilución limitante. En un paso ulterior preferido, las células se seleccionan de acuerdo con sus propiedades de crecimiento, exhibiendo particularmente las células preferidas una tasa de crecimiento rápido correspondiente a un tiempo de duplicación bajo, particularmente en condiciones de cultivo exentas de suero.

En una realización preferida, el procedimiento de selección puede implicar los pasos siguientes:

- 30 (i) seleccionar a partir de las células hospedadoras con vector diana integrado, aquéllas que exhiben expresión alta y estable del gen informador (RG1);
- (ii) seleccionar a partir de las células obtenidas en (i) aquéllas cuyo fenotipo permite el cultivo eficiente en gran escala en condiciones de cultivo industrial (es decir, sistemas de cultivo dinámicos como centrifugadoras, tanques agitados, reactores de onda, reactores de lecho fluidizado o reactores de lecho fijo); y
- 35 (iii) seleccionar a partir de las células obtenidas en (ii) aquéllas que permiten el intercambio eficiente de la casete informadora contra una casete de expresión para un gen de interés.

- 40 La selección de una línea de células "comodín" apropiada con capacidad de producción suficiente puede efectuarse por cribado respecto a la actividad de un informador secretado como la fosfatasa alcalina placentaria humana secretada (hSEAP) o una proteína informadora apropiada o cuantificable. Con objeto de obtener una población de células hospedadoras derivada de una sola célula como línea de células comodín potenciales, se realiza preferiblemente una segunda clonación a dilución limitante.

- 45 En una realización adicional, la selección de una célula hospedadora comodín de alta expresión puede realizarse también utilizando una proteína fluorescente semejante a la proteína fluorescente verde (GFP) o variantes, es decir, proteína fluorescente roja (RFP), proteína fluorescente azul (BFP), proteína fluorescente azul-verdosa (CFP), proteína fluorescente amarilla (YFP) o variantes adicionales. Por medio del informador expresado, las células con expresión alta pueden enriquecerse a partir de una agrupación de células transfectadas por clasificación de células activadas por fluorescencia (FACS) o después de marcación apropiada por enriquecimiento magnético (MACS o técnicas similares). Por varios ciclos de enriquecimiento subsiguientes, puede obtenerse una población con expresión alta para el informador respectivo. La célula comodín final se genera también preferiblemente por clonación a dilución limitante de esta población enriquecida de células policlonales.

- 50 Ejemplos preferidos adicionales de genes informadores son cloranfenicol-transferasa, β -galactosidasa, β -glucuronidasa, luciferasa, etc.

En una realización especialmente preferida, se selecciona una línea de células comodín "superrealizadoras" que tiene un nivel de expresión del gen informador que es preferiblemente al menos 10 veces, más preferiblemente al

menos 20 veces y todavía más preferiblemente al menos 25 veces mayor comparado con los niveles medios de expresión de la población de clones resultante.

5 En una realización preferida adicional, se analizan al menos 1000, v.g. al menos 2000 o al menos 3000 poblaciones de células primarias clonales u oligoclonales a fin de identificar líneas de células hospedadoras comodín de alta expresión, v.g. líneas de células hospedadoras comodín superrealizadoras.

10 Las líneas de células comodín preferidas tienen integrada una sola copia del vector diana en un sitio de transcripción altamente activo. Para la selección de células comodín apropiadas, el número de sitios de integración y el número de copias integradas de vectores diana deberá tenerse en cuenta preferentemente. El análisis del número de vectores diana integrados y el número de sitios de integración podría realizarse por PCR (preferentemente por PCR cuantitativa en tiempo real) y análisis de transferencia Southern o hibridación in situ.

15 En una realización adicional, la célula hospedadora comodín final podía albergar además del vector diana integrado estable una casete de expresión para la enzima mediadora de rotura de la doble cadena (DSB) utilizada durante la reacción de intercambio. Esta casete de expresión podría ser parte del vector diana o introducirse como un vector separado. Esta casete de expresión contiene preferentemente un promotor inducible para alcanzar una expresión controlada oportunamente de la enzima mediadora de DSB.

4.3.2 Integración no Aleatoria

20 En una realización diferente, los sitios de reconocimiento primero y segundo para la enzima mediadora de DSB pueden introducirse por recombinación homóloga direccionada en un sitio predeterminado del genoma de la célula hospedadora, que se sabe soporta actividad de transcripción, v.g. el locus inmunoglobulina. Para este propósito, se utiliza preferiblemente un vector diana para integración no aleatoria como se describe en la Sección 4.2.1.2. Puede llevarse a cabo un procedimiento de selección descrito en la Sección 4.3.1.

Sin embargo, debería indicarse que puede no ser necesario un procedimiento de selección en algunas variantes de esta realización, dado que los sitios de reconocimiento se introducen en un sitio predeterminado que se sabe soporta actividad de transcripción.

25 4.4. Intercambio de Casetes de Expresión Dentro de Células Comodín

30 Para un intercambio de la casete de expresión, se introduce un vector de intercambio como se ha descrito arriba en una célula hospedadora comodín, es decir una célula hospedadora comodín obtenida por integración aleatoria o no aleatoria del vector diana en el genoma de la célula hospedadora, en condiciones que permiten una rotura de la doble cadena en el primer y/o el segundo sitio de reconocimiento presente en el vector diana. De este modo, se consigue una integración del vector de intercambio en el genoma de la célula hospedadora por recombinación homóloga mediada por rotura de la doble cadena.

35 Para este propósito, la célula hospedadora comodín se co-transfecta preferiblemente con el vector de intercambio como se ha descrito arriba y un vector adicional que expresa la meganucleasa, particularmente una endonucleasa buscadora (HE) compatible con los sitios de reconocimiento presentes en el vector diana, v.g. I-Scel. Alternativamente, la enzima mediadora de rotura de la doble cadena puede expresarse en la célula hospedadora después de transfección con un mRNA apropiado codificante de la enzima deseada o por suministro directo de la enzima a la célula.

40 La enzima cortará el o los sitios de reconocimiento dentro del vector diana. Dependiendo del número de sitios de corte, esto abrirá simplemente el vector diana integrado en el genoma (un solo sitio de reconocimiento) o liberará la casete de expresión intercambiable del vector diana (dos sitios de reconocimiento flanqueantes de la casete de expresión). En ambos casos, la actividad de la HE conducirá a la generación de una rotura de la doble cadena (DSB) dentro del genoma de la célula comodín en las posiciones definidas dentro del vector diana.

45 Esta DSB inducirá recombinación homóloga (HR) con una frecuencia que es 100-1000 veces mayor que la observada para la HR endógena. Dentro de este proceso de reparación el vector de intercambio transfectado sirve como matriz de reparación, debido a sus tramos de secuencia homóloga flanqueantes de las casetes de expresión.

Como resultado, la casete de expresión completa con RG1 y SM1 se retirará del genoma de la célula comodín y será sustituida por la casete de expresión que comprende el GOI y en una realización particular adicionalmente el SM3, v.g. acoplada por un elemento IRES al GOI.

50 El SM3 permite la selección adicional de células que transcriben la casete de expresión con GOI y SM3 sin necesidad de analizar la expresión del GOI.

Un resultado adicional importante de la HR es la activación del SM2 carente de promotor en el vector diana. Debido a la HR con el vector de intercambio, el SM2 será suplementado con un promotor y facilita así la expresión de SM2 dentro de la célula comodín reparada con éxito.

Dado que el promotor está enlazado operativamente sólo a una versión truncada Δ SM2 en el vector de intercambio, el mismo no aporta resistencia a la célula. Después de recombinación homóloga, SM2 se activa y permite la selección eficiente de células que han reparado con éxito rotura de la doble cadena por introducción de la casete de expresión procedente del vector de intercambio.

- 5 Debido a esta selección simple (utilizando SM2) o doble (utilizando SM2 y SM3) puede generarse una población de células promotoras homogéneas con altos niveles de expresión para el GOI sin necesidad de cribado excesivo.

Dado que la modificación genómica de la célula comodín se limita al intercambio de la casete de expresión en el locus marcado, el fenotipo de la célula no se verá afectado.

- 10 Incluso en los casos en que se requiere una población de células derivada de una sola célula, esto puede lograrse en un tiempo breve, dado que una dilución limitante de la población obtenida después de la selección se realiza fácilmente con un grado de éxito alto.

Así pues, la presente invención proporciona una vía rápida y reproducible para generar células hospedadoras productoras para la obtención de productos génicos recombinantes deseados, v.g. proteínas o ácidos nucleicos recombinantes, en grandes cantidades.

- 15 Dos esquemas de una reacción de intercambio entre la diana y el vector de intercambio por recombinación homóloga se muestran en las figuras 5 y 6.

- 20 En una realización diferente, puede introducirse un vector de intercambio en el genoma de una célula hospedadora comodín, que se obtiene por integración no aleatoria de un primer y/o segundo sitio de reconocimiento de enzima mediadora de DSB en un locus predeterminado en su genoma, utilizando preferiblemente un vector diana como se describe en la Sección 4.2.1.2. En esta realización, el vector de intercambio comprende preferiblemente

(i) un gen de interés (GOI) y

(ii) una primera secuencia homóloga 5' y una segunda secuencia homóloga 3', que permiten recombinación con secuencias en el sitio de integración del primer y/o segundo sitio de integración.

Más preferiblemente, se utiliza un vector de intercambio como se describe en la Sección 4.2.2.

- 25 Adicionalmente, la presente invención se explicará con mayor detalle por las figuras y ejemplos que siguen.

Figuras

Fig. 1 Representación esquemática de elementos genéticos en una realización del vector diana (vector diana A)

- 30 El vector diana contiene una casete de intercambio con un gen informador (RG1) dirigido por el promotor CMV (P1) y un primer marcador de selección (SM1) dirigido por PSV40 (P2) para selección de integración estable del vector en el genoma de una célula hospedadora. La casete de intercambio está flanqueada por dos sitios de reconocimiento para una endonucleasa buscadora (I-SceI). El segundo marcador de selección (SM2) fuera de la casete de intercambio carece de promotor y por tanto no es funcional en el vector diana.

Fig. 2 Representación esquemática de elementos genéticos en una realización adicional del vector diana (vector diana B)

- 35 En la versión modificada del vector diana, el sitio 3'-I-SceI está situado delante de la señal poli-A (pA) del primer marcador de selección (SM1). En caso de religación (unión por los extremos en lugar de recombinación homóloga) del vector diana cortado I-SceI, el elemento pA sirve como aislador, impidiendo la transcripción de SM2. Adicionalmente, están situados codones de parada frente al marcador de selección carente de promotor 2 (SM2) a fin de terminar la traducción de transcritos no terminados por el elemento pA.

- 40 Fig. 3 Representación esquemática de elementos genéticos en una realización del vector de intercambio (vector de intercambio A)

- 45 El vector de intercambio contiene un CDS para un GOI y un CDS incompleto para el marcador de selección (Δ SM2). Únicamente el CDS para el GOI va seguido por una señal de poliadenilación apropiada. Mientras que el CDS incompleto para SM2 está dirigido por un promotor constitutivo funcional P3 (PSV40), el GOI contiene en su extremo 5' sólo un promotor parcial que es idéntico a la parte 3' del promotor CMV en el constructo diana.

Fig. 4 Representación esquemática de elementos genéticos en una realización adicional del vector de intercambio (vector de intercambio B)

En esta realización, el vector de intercambio contiene además del vector representado en Fig. 3, un tercer marcador de selección (SM3) cuya expresión está acoplada a la expresión del GOI por un elemento IRES. Esto permite el

aislamiento de células con casetes de expresión intercambiadas con éxito por una selección doble utilizando los marcadores de selección 2 y 3.

Para poder eliminar las células que tienen integrado aleatoriamente el vector de intercambio completo en su genoma, el vector de intercambio contiene en esta realización un gen HSV-TK como gen suicida.

- 5 Fig. 5 Esquema de la reacción de intercambio entre la diana y el vector de intercambio por recombinación homóloga (HR)

El vector diana en el locus genómico marcado con informador (RG1) y casete de selección (SM1) está flanqueado por sitios de reconocimiento de endonucleasas buscadoras (I-SceI). El promotor P1 está representado por el elemento "intensificador-PCMV-intrón". El promotor P2 está representado por el promotor SV40 (PSV40). El vector de intercambio alberga el GOI y un marcador de selección no funcional (Δ SM2). Se representan las regiones de homología utilizadas para reparación de la DSBs mediada por endonucleasas buscadoras por recombinación homóloga (HR). El promotor P3 en el vector de intercambio está representado por un promotor SV40 (PSV40).

- 10

La parte inferior muestra un locus genómico marcado después de intercambio con éxito de casetes de expresión.

Fig. 6 Esquema de una reacción de intercambio modificada utilizando un tercer marcador de selección (SM3)

- 15 La reacción de intercambio se realiza como se describe en Fig. 5 entre el vector diana en un locus genómico marcado y el vector de intercambio. En esta realización, el vector de intercambio contiene, además del GOI, un tercer marcador de selección (SM3) cuya expresión está acoplada al GOI por un elemento IRES.

Fig. 7 Medida de actividades hSEAP en paneles de clones de células primarias

- 20 Las poblaciones de células clonales u oligoclonales individuales (eje x) se seleccionan basándose en la actividad de SM1 para integración estable del vector diana. Las células se expandieron y cribaron respecto a la actividad de transcripción del locus de integración aleatorio marcado basado en la actividad del RG1 (hSEAP). La actividad del informador hSEAP se da en ng/ml en el eje y. A. Distribución de la actividad de hSEAP en 39 poblaciones de células individuales. B. Distribución de la actividad de hSEAP en 92 poblaciones de células individuales.

Fig. 8 Cribado primario de células hospedadoras universales potenciales

- 25 Se generan poblaciones de células como se describe en Fig. 7 por selección de células utilizando SM1 después de transfección con el vector diana. Los datos demuestran la identificación de una población de células altamente productoras deseadas denominadas "superadoras". La identificación está basada en la expresión de RG1 (hSEAP). La actividad del hSEAP informador se da en ng/ml en el eje y.

Fig. 9 Cribado secundario del superador 01A438

- 30 A partir de las células superadoras, caracterizadas por alta expresión del RG1 (hSEAP), se generan poblaciones de células clonales de alta expresión por clonación a dilución limitante (LD) de células simples en condiciones exentas de suero. Los gráficos A. y B. representan la diferente capacidad de expresión de poblaciones de células clonales individuales como se demuestra por la actividad del RG1 (hSEAP). La actividad del hSEAP informador se da en ng/ml en el eje y. El asterisco en B. denota un clon de células seleccionado como célula comodín (designado 07-022) para uso ulterior. La selección se basó en características de expresión y crecimiento.

- 35

Fig. 10

Intercambio mediado por rotura de la doble cadena: optimización de la cantidad de plásmido de meganucleasa

Intercambio de casetes de expresión en una célula comodín 07-022 preformada con vector diana integrado estable. La casete de expresión procedente de las células comodín marcadas se sometió a intercambio contra una casete GFP-IRES-SM3 con GFP como GOI modelo por medio de HR mediada por DSB entre el vector diana y el vector de intercambio. El intercambio fue realizado por co-transfección del vector de intercambio (ECV2) y un vector de expresión para la meganucleasa (Mnv). Con objeto de optimizar la reacción de intercambio se testó una cantidad constante del vector de intercambio con cantidades diferentes de plásmido de meganucleasa. El intercambio con éxito de las casetes de expresión por recombinación homóloga entre el vector diana y el vector de intercambio activará el SM2 (NeoR). Después de selección de 2×10^6 con G418 durante 11 días, se determinó el número de colonias resistentes a G418, así como la cantidad de colonias GFP-positivas. El gráfico representa el porcentaje de células GFP-positivas en los diferentes enfoques. Obsérvese que el enfoque 5 producía aproximadamente 6 veces más células clonales que expresaban GFP que el enfoque 3.

- 40
- 45

Fig. 11 Intercambio mediado por rotura de la doble cadena: optimización de la cantidad de plásmido meganucleasa

- 50 Células procedentes del enfoque descrito para Fig. 10 se seleccionan ulteriormente utilizando la actividad de SM3 (ZeoR) además de SM2. Después de la doble selección, se determinó por microscopía UV el número total de colonias doblemente resistentes, así como la cantidad de colonias GFP-positivas entre ellas. El enfoque 5 muestra el

porcentaje máximo (~ 94%) y con 261 clones el número absoluto máximo de células que expresaban GFP doblemente resistentes y brillantes.

Fig. 12 Intercambio mediado por rotura de la doble cadena: optimización de la cantidad de plásmido de meganucleasa

- 5 Ejemplo de células clonales obtenidas por intercambio de casetes de expresión en células comodín por recombinación homóloga inducida por rotura de la doble cadena (DSB) entre el vector diana y el vector de intercambio. La imagen representa una transparencia superpuesta (luz visible/UV) de una colonia GFP-positiva del enfoque 5 descrito en Fig. 10 y 11. La barra de escala indica 100 μ m.

Fig. 13 Colonia GFP-positiva derivada del clon 01C090 de células comodín

- 10 Ejemplo de un intercambio con éxito de casete por el método descrito utilizando una población de células comodín diferentes (01C090). La imagen representa células que expresaban GFP obtenidas después de intercambio de casete y doble selección basada en SM2 y SM3. La barra de escala indica 100 μ m.

Fig. 14 Colonias GFP-positivas después de reacción de intercambio y doble selección

- 15 Ejemplos adicionales de intercambio con éxito de casete por intercambio mediado por DSB de casetes de expresión en clones de células comodín establecidos. Se muestran en comparación células obtenidas después de intercambio y doble selección a partir del clon de células comodín 07-022 (A) y el clon 08-018 (B). La barra de escala indica 100 μ m.

Fig. 15 Confirmación de la reacción de intercambio basada en PCR

- 20 El DNA genómico de tres clones comodín diferentes (07-022, 08-018, 01C090) se analizó por PCR antes (pistas 1, 4, 6) y después de la reacción de intercambio (pistas 3, 5, 7). Para 07-022 se analizó también un control con vector de intercambio integrado aleatorio (pista 2). (A.) un producto de 368 pares de bases confirma la presencia de la casete de expresión hSEAP antes del intercambio (pistas 1, 4, 6). La banda que falta después del intercambio (pistas 3, 5, 7) indica la eliminación completa de esta casete de expresión. (B.) Un producto de 293 pares de bases confirma la presencia de la casete de expresión GFP después de la reacción de intercambio (pistas 3, 5, 7), pero también después de integración aleatoria (pista 2). (C.) un producto de 971 pares de bases se genera únicamente después de completarse SM2 e indica así el intercambio de casetes de expresión por HR (pistas 3, 5, 7). Plásmidos apropiados sirvieron como controles positivos (pista 8). Obsérvese que el control en (C) muestra una banda de 917 pares de bases debida al DNA plasmídico utilizado. Controles negativos contenían en lugar de DNA molde agua (pista 9) o DNA genómico de células con control parentales sin modificación alguna (pista 10). Los datos se recopilaron de diferentes geles de agarosa.
- 25
- 30

5. Ejemplos

5.1 Generación del Sistema Vector

5.1.1 Generación de un vector diana

- 35 El promotor CMV original del vector pcDNA 3.1(+) se separó por restricción con NheI y NruI. El promotor CMV extendido con el intrón A se separó del vector pMG con PaeI y después de hacer romos los extremos con XbaI. Este fragmento largo de 1,7 kb se ligó al fragmento pcDNA 3.1 cortado con NruI y NheI (compatible con XbaI). De este modo se generó el vector CV001.

- 40 Se introdujo el sitio I-SceI 5' en CV001 por una estrategia de adaptación. Oligonucleótidos complementarios con el sitio I-SceI TAGGGATAACAGGGTAAT y extremos colgantes HindIII se hibridaron y clonaron en el HindIII de CV001 dando como resultado CV001-MN.

- 45 El gen original de resistencia a neomicina (NeoR) del pcDNA 3.1(+) en CV001-MN se reemplazó por un gen de resistencia a higromicina (HygroR) como SM1. A este fin, el gen HygroR que incluía un sitio poliA se amplificó a partir de un plásmido de tal modo que el HygroR-CDS carecía de los 4 primeros nt (con inclusión del ATG de inicio) desde el extremo 5' y contiene en su extremo 3' un sitio de restricción PciI. Después de digestión con PciI, el producto PCR se ligó al CV001-MN no metilado restringido con BsaBI y PciI. Por clonación en el sitio BsaBI, se completó el HygroR-CDS por adición del ATGA ausente al extremo 5' del vector. En el vector CV001-MN-Hygro resultante se introdujo un gen NeoR sin promotor como SM2. A este fin, se amplificó por PCR un NeoR-CDS completo a partir de un DNA plasmídico apropiado. Se utilizó el cebador 5' para introducir un sitio I-SceI, y permite adicionalmente la clonación del NeoR-CDS por PciI. El cebador 3' introduce un sitio BstZ171 (TAC).

- 50 Después de restricción con PciI y BstZ171, el fragmento PCR se purificó en gel y se ligó al vector CV001-MN-Hygro cortado con las mismas enzimas de restricción.

Para finalizar el vector diana, el vector CV001-MN2-Hygro-Neo obtenido se suplementó con un gen secretado de fosfatasa alcalina (hSEAP) como RG1. El hSEAP-CDS que incluía la señal poli-A se separó de un vector apropiado por digestión con KpnI y XbaI. El fragmento se clonó en el sitio KpnI y XbaI del vector CV001-MN2-Hygro-Neo. El vector diana final se designó pCTV-SAP-Hyg-01.

5 5.1.2 Generación de un vector de intercambio

Basándose en el vector CV001, se generó el vector de intercambio B conforme a Fig. 4. Con objeto de generar una versión truncada no funcional de SM2, el NeoR-CDS de CV001 se separó por digestión con BstZ171 y BlnI. El NeoR-CDS se reemplazó por un fragmento NeoR de 631 pares de bases BlnI-NaeI. Este gen NeoR truncado carece de la secuencia codificante para los 24 aminoácidos de la parte terminal C de la enzima, que son cruciales para su función.

En el vector resultante CV001-ΔNeoR la región 5' del promotor CMV, que contenía el intensificador y el promotor de RNA-polimerasa II, se separó por doble digestión con BsmBI y MunI. Esto dejaba principalmente la secuencia de intrón A del elemento promotor del intrón A de CMV en CV001. Después de rellenar los extremos 3' rebajados con enzima Klenow, el vector se cerró por relegación de los extremos romos.

El vector resultante CV001-InA-ΔNeoR está listo para introducción de un GOI libre seleccionable y sirve como vector de intercambio general. Como ejemplo para un GOI se seleccionó la proteína verde fluorescente (GFP). En este caso particular, la expresión de GFP se acopló a la expresión de un tercer marcador de selección (SM3) por un elemento IRES. Como SM3 se ha seleccionado el gen de resistencia a Zeocina (ZeoR). La casete de expresión constituida por la secuencia para L-GFP seguida por un elemento IRES y el CDS para ZeoR se separó del vector pMono-Zeo-GFP (Invitrogen) como fragmento AgeI-Sall. Esta casete se ha ubicado por ligación de extremos romos en el CV001-InA-ΔNeoR cortado con EcoRV. El vector final se denominó pCEV-GFP-Zeo-01.

5.2 Generación de Líneas de Células Comodín Universales

Para la generación de líneas de células comodín potenciales, se utilizó una línea de células CHO adaptada para crecimiento en suspensión en condiciones exentas de suero y capaces de crecer hasta altas densidades de células (al menos 2.5×10^6 células/ml). Con objeto de transfectar las células en suspensión con el vector diana pCTV-SAP-Hyg-01 se utilizó la tecnología Nucleofector.

Antes de la nucleofección, el vector diana pCTV-SAP-Hyg-01 se digirió con SspI por dos razones:

- Eliminación de elementos bacterianos indeseables (origen de replicación, gen de resistencia a ampicilina) a fin de prevenir la silenciamiento del vector diana mediada por estos elementos.
- Aumento de la probabilidad de la integración intacta de elementos funcionales relevantes del vector diana en el genoma de las células hospedadoras.

La digestión con SspI produjo un fragmento de 1698 kb que contenía el origen de pUC y la mayor parte del gen AmpR. El elemento del vector diana de 7,41 kb restante relevante se purificó por electroforesis en gel de agarosa utilizando el kit de extracción de gel QIAquick de acuerdo con la recomendación del fabricante.

Para la nucleofección con el vector diana purificado en gel, se establecieron condiciones que permiten transfección con alta eficiencia de las células CHO adaptadas en suspensión. En general pudieron alcanzarse rutinariamente eficiencias de 70-90% de células transfectadas.

Se dejó que las células sometidas a nucleofección se recuperaran durante 24 horas y se sometieron luego a selección para células con vector diana integrado estable basándose en la expresión de Hygromicin-fosfotransferasa (HygroR) como SM1. Para la selección con Higromicina B (200 µg/ml) las células sometidas a nucleofección se transfirieron a placas de 96 pocillos a una densidad que permitía el aislamiento de poblaciones de células mono- u oligo-clonales (dilución limitante). Esto podría realizarse por densidades de células entre 1 y 60 células/pocillo, muy preferentemente a densidades entre 1 y 30 células/pocillo.

Cada una de estas poblaciones de células mono- u oligo-clonales representa sitios de integración aleatoria diferentes del vector diana. La actividad de transcripción de estos sitios de integración se determinó por la actividad del gen informador hSEAP en el sobrenadante utilizando un ensayo colorimétrico en microplacas.

A aproximadamente 70-80% de confluencia, las células seleccionadas se transfirieron a placas de 24 pocillos para expansión ulterior. Se utilizó el sobrenadante para ensayo de la actividad de hSEAP.

La fosfatasa alcalina endógena se desactivó por incubación a 65°C durante 30 min. La actividad de hSEAP se midió en tampón de reacción hSEAP (dietanolamina 2M, MgCl₂ 1 mM, L-homoarginina 20 mM, pH 9,8) a 37°C. La hidrólisis del sustrato fosfato de p-nitrofenol (120 mM) se midió por lectura a OD405. La actividad se cuantificó contra un estándar con fosfatasa alcalina placentaria.

La mayoría de los clones exhibían una actividad muy baja para el informador hSEAP. Sólo una fracción de los clones aislados con un vector diana integrado de manera estable exhibían una expresión fuerte del informador. La distribución típica de la variación en los niveles de expresión en estas poblaciones de células primarias clonales u oligoclonales se muestra por su expresión de hSEAP resumida en los ejemplos en Fig. 7A y B. A fin de identificar los clones raros de alta expresión denominados "superadores", se analizaron más de 3000 poblaciones de células primarias clonales/oligoclonales.

Para asegurar que se utilizará el clon de máxima expresión como célula comodín universal, se realizó una segunda dilución limitante con los clones que exhibían los niveles de hSEAP máximos a fin de identificar subclones con el nivel de expresión máximo. Se muestra un ejemplo para el clon 01A438 identificado como superador en una primera dilución limitante (Fig. 8). Células individuales del clon 01A438 se sembraron en una segunda dilución limitante que permitía la generación de una población de células derivada de una sola célula. Como se representa en Fig. 9A y B, los sub- clones resultantes exhibían en general una mayor expresión de hSEAP comparada con la mayoría de los clones iniciales del cribado primario. No obstante, fue posible identificar dentro de este grupo de clones de alta expresión aquéllos que tenían la expresión máxima para hSEAP. En esta etapa de la selección de clones para análisis ulterior, además del nivel de expresión del gen informador se tuvo en cuenta también la capacidad de crecimiento (tiempo de duplicación) de los clones. Basándose en este criterio adicional, se seleccionó el clon 07-022 (marcado con un asterisco en Fig. 9).

En suma, para obtener una línea de células hospedadoras con un nivel de expresión alto para hSEAP, se generaron varios miles de clones individuales por clonación a dilución limitante exenta de suero. Debido a la integración aleatoria en el genoma, sólo aproximadamente 50% de estos clones exhibían expresión medible de hSEAP. De estos clones que expresaban hSEAP sólo 0,3% de los clones se identificaron como clones de alta expresión, es decir "superadores" (expresión más de 25 veces mayor con respecto a la media de todos los clones).

5.3 Intercambio del gen Informador

Para intercambio del RG1 y el SM1 contra un GOI, las células comodín del subclón 07-022 se cotransfectaron con un vector de intercambio pEV-GFP-Zeo-01 de acuerdo con Fig. 4 y un plásmido de expresión para la endonucleasa buscadora de corte raro I-SceI. Como GOI se utilizó la secuencia codificante para una proteína verde fluorescente (GFP).

La región del intrón A remodelable delante de la región codificante de GFP, era lo bastante larga (1077 pares de bases) para permitir una recombinación homóloga eficiente, pero contiene actividad promotora atenuada.

El SM1 en el vector de intercambio es una versión truncada C-terminal del gen de resistencia a neomicina (NeoR). Debido a una delección de 163 pares de bases en la región codificante de NeoR, la neomicin-fosfotransferasa truncada resultante no confiere resistencia a G418. La región 5' del gen NeoR (719 pares de bases) es suficiente para permitir la recombinación homóloga con el NeoR completo en el vector diana. Esta energía permite la activación del gen NeoR sin promotor en el vector diana sin generación de células NeoR debido a integración aleatoria del vector de intercambio.

El sistema permitirá la selección de aquellas células que han reparado la DSBs inducida por la enzima buscadora por recombinación homóloga utilizando el vector de intercambio como matriz de reparación.

Como marcador de selección adicional se expresa un SM3 (gen de resistencia a Zeocina) a partir de la casete de expresión de GFP por medio de un elemento IRES entre la GFP y el gen de resistencia Zeo.

Este SM3 se utiliza para determinar la frecuencia de integración aleatoria del vector de intercambio y para comprobar si otros eventos de reparación podrían conducir a células resistentes a neomicina sin integración de la casete de expresión procedente del vector de intercambio, sólo por delección del área de intercambio en el vector diana.

A este fin, las células doblemente transfectadas se seleccionaron primeramente con G418, se registró el número de clones generados y se analizaron los clones respecto a la expresión de GFP por microscopía UV.

Como se representa en Fig. 10 después de selección durante 11 días con G418, se obtuvo una población heterogénea de células NeoR de las cuales el 45-62% exhibían una expresión brillante de GFP, mientras que el resto de las células NeoR eran GFP-negativas.

Dichas células podrían representar la reparación de la DSB por simple unión en los extremos o por recombinación homóloga en un solo lado. En el último caso, únicamente la DSB adyacente al gen NeoR es reparada por HR, mientras que la DSB adyacente al promotor CMV se repara por unión no homóloga en los extremos.

Por medio de una selección adicional con Zeocina, la fracción de células GFP-positivas se aumentó hasta 94% dependiendo de la cantidad de plásmido Mnv de expresión de I-SceI cotransfectado con una cantidad constante de plásmido de intercambio (Fig. 11). Cuando se utilizó la ratio optimizada entre el vector de expresión para la meganucleasa y el vector de intercambio (enfoque 5 en Fig. 11), eran detectables un total de 291 colonias que

expresaban GFP. Basándose en 2×10^6 células extendidas en placas después de transfección y seleccionadas con G418 y Zeocina, esto representa una frecuencia de $\sim 1 \times 10^{-4}$. Este es un resultado satisfactorio sorprendente comparado con frecuencias publicadas para recombinación homóloga inducida por rotura de la doble cadena, que están comprendidas en el intervalo de $1-2,5 \times 10^{-6}$.

- 5 En resumen, el uso de dos marcadores de selección permitió la selección de una población homogénea de células que exhibían una fluorescencia brillante de GFP (Fig. 12).

Además del clon 07-022, se utilizaron clones comodín adicionales para demostrar la intercambiabilidad del RG1 y SM1 contra la casete de expresión constituida por GFP-IRES-ZeoR (Fig. 13 y 14).

- 10 En general, el intercambio con utilización de clones comodín que expresaban hSEAP baja daba como resultado únicamente células no fluorescentes o débilmente fluorescentes después de doble selección con G418 y Zeocina.

En contraste, la utilización de clones comodín que expresaban hSEAP alta daba como resultado un intercambio eficiente de las casetes de expresión por el mecanismo de intercambio descrito mediado por meganucleasas y la generación de células doblemente resistentes que expresaban GFP brillante como se muestra para el clon 01-C090 en Fig. 13.

- 15 En Fig. 14 se muestra una comparación con el clon 07-022 para el clon comodín 08-018. Con el clon 08-018 fue posible obtener también después de intercambio de casetes de expresión mediado por meganucleasas una población homogénea de células que expresaban GFP. La intensidad de la expresión de GFP era comparable a las observadas en las células derivadas del clon 07-022.

- 20 Las células derivadas de las células comodín después de la reacción de intercambio basada en meganucleasas pueden expandirse en cultivo de suspensión en condiciones exentas de suero y utilizarse para propósitos de producción.

Con objeto de asegurar una línea de células de producción derivadas de una sola célula, las células intercambiadas podían generarse rápidamente con un enfoque de dilución limitante con poco esfuerzo.

5.4 Caracterización Molecular

- 25 Para monitorizar los eventos moleculares durante la reacción de intercambio, se utilizó un enfoque basado en PCR utilizando tres combinaciones de cebadores diferentes (PCR 1-3, véase Tabla 1).

Tabla 1 Combinaciones de cebadores para caracterización molecular

PCR No.	Fijación del cebador 5' – 3'	Diana	Fenotipo identificado	Tamaño del producto PCR [pb]
1	IntA-SEAP	Vector diana en la célula comodín hospedadora	SEAP+, HygroR GFP-	368
2	IntA- GFP	Vector de intercambio	ZeoR GFP (+)	293
3	SV40P- 3'Neo	SM2 activado (NeoR) después de recombinación homóloga	G418+ Zeo+ GFP+	971

- 30 Con PCR1 se detecta el vector diana integrado estable dentro del genoma de la célula hospedadora comodín. La PCR2 se utiliza para confirmar la presencia de la casete de expresión procedente del vector de intercambio en las células después de realización de la reacción de intercambio. Dado que el vector de intercambio puede incorporarse también aleatoriamente, se utiliza la PCR3 para confirmar la recombinación homóloga en el extremo 3' del área de intercambio dentro del constructo diana. El cebador 5' para esta PCR se fija al promotor SV40 que está presente en el constructo de intercambio, mientras que el cebador 3' se fija a la parte 3' del gen NeoR presente únicamente en el
- 35 constructo diana y no en el constructo de intercambio. Sólo después de activación del SM2 sin promotor (gen NeoR) por recombinación homóloga, la PCR3 puede generar un producto PCR de 971 pares de bases.

Por combinación de los resultados de los diferentes análisis PCR puede demostrarse la eliminación de la casete de expresión del constructo diana y la introducción de la casete de expresión del vector de intercambio.

Los resultados para un análisis de este tipo se resumen en Fig. 15 para tres líneas de células comodín universales diferentes para las cuales se realizó con éxito una reacción de intercambio. En la parte (A.) se confirmó la presencia del constructo diana por cebadores que se fijaban al intrón del elemento promotor CMV y la región 5' del CDS de hSEAP. El panel (B.) muestra la confirmación de la casete de expresión integrada estable del vector de intercambio B por el mismo cebador intrónico que en (A.) en combinación con un cebador que se fija en la región 5' del CDS para GFP. La amplificación con éxito del producto PCR correspondiente indica la introducción de la casete de expresión sea por HR o por integración aleatoria. El panel (C.) exhibe la confirmación por PCR de la presencia de un gen NeoR activado en el vector diana como resultado de HR con el vector de intercambio. Dado que los cebadores para esta PCR se fijan al promotor SV40 (cebador 5'), que está presente sólo en el vector de intercambio y a la parte 3' del NeoR-CDS (cebador 3'), presente sólo en el vector diana, la amplificación con éxito indica HR entre el vector de intercambio y el vector diana en la DSB mediada por 3'-I-SceI.

Como control positivo (pista 8) en el panel (A.) y (B.) se han utilizado el vector diana y el vector de intercambio 2, respectivamente. El control positivo en el panel (C.) consiste en un vector (CV001) con secuencias de fijación de cebadores idénticas pero con una distancia ligeramente diferente a otra. Por tanto, el control positivo da como resultado un producto PCR ligeramente más pequeño (917 pares de bases en lugar de 971 pb) que el esperado a partir del DNA genómico. Como control negativo, el DNA molde se ha reemplazado por agua (pista 9) o, para comprobar la amplificación inespecífica de CHO-DNA genómico, por la misma cantidad de DNA de la línea de células CHO parentales sin nucleofección del constructo diana (pista 10).

El clon 07-022 se caracterizó de modo muy intenso y además de las células de partida y sus derivados de células de intercambio con éxito, se analizó también un control para integración aleatoria del vector de intercambio B. El clon de células utilizado para este control exhibía después de nucleofección de sólo el vector de intercambio B, resistencia a Zeocina y una fluorescencia de GFP muy débil.

Para los clones 08-018 y 01C090, se analizaron las células iniciales y células de intercambio con éxito que sobrevivieron a la doble selección y exhiben una fluorescencia de GFP brillante.

Para todos los clones pudo demostrarse claramente la presencia del hSEAP-CDS claramente en la población inicial de células (panel A). Con respecto al clon 07-022, pudo confirmarse también la presencia de este CDS para el clon de células generado por integración aleatoria. Después de la reacción de intercambio mediada por meganucleasa, la hSEAP-CDS ya no era amplificable en la totalidad de los clones intercambiados testados, que exhibían NeoR/ZeoR y una fluorescencia brillante de GFP. Esto indica la eliminación de la casete de expresión inicial de los vectores diana integrados en todos los ejemplos analizados.

De acuerdo con dicho descubrimiento, podría demostrarse también por PCR la presencia de la casete de expresión del vector de intercambio B para la GFP-CDS sólo después de la realización de la reacción de intercambio (panel B, pistas 3, 5, 7). Las células de partida para los clones testados eran negativas en este ensayo PCR (panel B, pistas 1, 4, 6). Este ensayo PCR no es específico para introducción de la casete de expresión GFP por HR. El ejemplo para integración aleatoria del vector de intercambio B del clon 07-022 demostraba que también en este caso podría obtenerse una señal positiva con el cebador específico GFP-CDS.

La brillante intensidad de la expresión de GFP en las diferentes poblaciones de células derivadas después de la realización de la reacción de intercambio indica que la introducción del GFP-CDS había ocurrido en el punto caliente de expresión marcado dentro de las células utilizadas en las pistas 3, 5 y 7.

Finalmente, para confirmar que el fenotipo observado está basado en la eliminación de la casete de expresión en el vector diana y la introducción de la casete de expresión del vector de intercambio en la misma posición, se realizó un análisis PCR adicional (panel C.) Esta PCR dará únicamente un producto PCR esperado de 971 pares de bases cuando la DSB mediada por I-SceI 3' en el vector diana ha sido reparada por HR utilizando el vector de intercambio 2 como matriz de reparación. Como era de esperar, únicamente en las células doblemente resistentes, de GFP brillante pudo detectarse un producto PCR esperado (panel C., pistas 3, 5, 7) y no en las células resultantes de la integración aleatoria del vector de intercambio (panel C., pista 2).

LISTADO DE SECUENCIAS

<110> Celonic AG

<120> Métodos y materiales para la generación reproducible de líneas de células con alta producción de proteínas recombinantes

<130> 41754P EP

<140> 08006023.9

<141> 2008-06-12

<160> 1

<170> PatentIn version 3.3

<210> 1

<211> 18

<212> DNA

<213> Artificial

<220> Oligonucleótido

<223>

<400> 1
tagggataac agggtaat

18

REIVINDICACIONES

1. Método para la generación de una célula hospedadora productora para la obtención de productos génicos recombinantes que comprende los pasos

(a) proporcionar un vector diana que comprende

5 (i) un gen informador (RG1) enlazado operativamente a una primera secuencia de control de la expresión que comprende un promotor constitutivo (P1),

(ii) un primer gen marcador de selección (SM1) enlazado operativamente a una segunda secuencia de control de la expresión (P2),

10 (iii) un segundo gen marcador de selección no funcional (SM2) sin enlace operativo a una secuencia de control de la expresión,

en donde la secuencia (i) está localizada en la posición 5', la secuencia (ii) está localizada entre (i) y (iii) y la secuencia (iii) está localizada en la posición 3'.

15 (iv) un primer y un segundo sitio de reconocimiento para una enzima mediadora de rotura de la doble cadena, en donde el primer sitio de reconocimiento está localizado entre la primera secuencia de control de la expresión (P1) y el primer gen informador (RG1) y el segundo sitio de reconocimiento está localizado entre la secuencia (ii) y la secuencia (iii),

(b) introducir el vector diana en una célula hospedadora en condiciones tales que permiten la integración aleatoria del vector diana en el genoma de la célula hospedadora,

20 (c) determinar cuantitativamente el nivel de expresión del gen informador (RG1) y seleccionar una célula hospedadora que tenga integrado de manera estable el vector diana y que exhiba una actividad de transcripción del gen informador (RG1), que es al menos 10 veces mayor comparada con los niveles de expresión medios de la población de clones resultante,

(d) proporcionar un vector de intercambio que comprende:

(i) un gen de interés (GOI), y

25 (ii) un segundo gen marcador de selección inactivo (Δ SM2) enlazado operativamente a una tercera secuencia de control de la expresión (P3),

en donde el vector de intercambio comprende una primera secuencia homóloga 5' y una segunda secuencia homóloga 3', que permiten la recombinación con secuencias del vector diana,

30 (e) introducir el vector de intercambio en una célula hospedadora obtenida en el paso (c) en condiciones que permiten una rotura de la doble cadena en el primer y/o segundo sitio de reconocimiento, como se define en (a) (iv) y una integración del vector de intercambio en el genoma de la célula hospedadora por recombinación homóloga mediada por rotura de la doble cadena,

con lo cual el segundo gen marcador de selección inactivo (Δ SM2) se activa por integración del vector de intercambio por recombinación homóloga con el vector diana,

35 (f) seleccionar una célula productora que tiene integrado el vector de intercambio por recombinación homóloga con el vector diana integrado, en donde la célula productora expresa el GOI.

2. El método de la reivindicación 1, en donde la célula hospedadora es una célula eucariota, particularmente una célula de mamífero, v.g. una célula de roedor tal como una célula de ratón, rata o hámster, o una célula de primate tal como una célula humana, preferiblemente una célula CHO.

40 3. El método de la reivindicación 1 ó 2, en donde el gen informador (RG1) codifica una enzima o un producto génico luminiscente o fluorescente.

4. El método de una cualquiera de las reivindicaciones 1-3, en donde la primera secuencia de control de la expresión (P1) comprende como promotor constitutivo el promotor CMV, opcionalmente suplementado con el intrón A de CMV.

45 5. El método de una cualquiera de las reivindicaciones 1-4, en donde la enzima inductora de rotura de la doble cadena es una meganucleasa o una nucleasa buscadora que tiene un sitio de reconocimiento de al menos 10, preferiblemente al menos 12 y más preferiblemente 18 nucleótidos, y se selecciona v.g. de I-SceI, I-SceII, I-SceIII, I-SceIV, I-CeuI, I-CreI, I-PpoI, I-TevI, I-TevII, I-TevIII, HO y Endo Seel.

6. El método de una cualquiera de las reivindicaciones 1-5, en donde se selecciona una célula que tiene el vector diana integrado particularmente como una sola copia en un solo sitio del cromosoma de la célula hospedadora.
- 5 7. El método de una cualquiera de las reivindicaciones 1-6, en donde el paso de selección (c) está basado en una detección de la actividad del primer gen marcador de selección (SM1).
8. El método de una cualquiera de las reivindicaciones 1-7, en donde el vector de intercambio comprende el gen de interés (GOI) sin una secuencia funcional de control de la expresión, particularmente con una secuencia parcial de control de la expresión que es sustancialmente idéntica a la parte 3' de la primera secuencia de control de la expresión (P1) en el vector diana.
- 10 9. El método de una cualquiera de las reivindicaciones 1-8, en donde el vector de intercambio comprende el segundo gen marcador de selección incompleto (Δ SM2) enlazado operativamente a una tercera secuencia de control de la expresión (P3), que comprende particularmente un promotor constitutivo, v.g. el promotor SV40.
- 15 10. El método de una cualquiera de las reivindicaciones 1-9, en donde la primera secuencia homóloga 5' tiene una longitud de al menos aproximadamente 500 nucleótidos y/o la segunda secuencia homóloga 3' tiene una longitud de al menos aproximadamente 500 nucleótidos.
11. El método de una cualquiera de las reivindicaciones 1-10, en donde el vector de intercambio comprende adicionalmente un gen marcador de selección negativo, particularmente un gen suicida, v.g., el gen de timidina-quinasa HSV que está localizado fuera de la secuencia definida por la primera secuencia homóloga 5' y la segunda secuencia homóloga 3'.
- 20 12. El método de una cualquiera de las reivindicaciones 1-11, que comprende adicionalmente
- (i) introducir un vector que comprende una secuencia de ácido nucleico que codifica una enzima mediadora de rotura de la doble cadena o un mRNA que codifica una enzima mediadora de rotura de la doble cadena en la célula hospedadora antes o durante el paso (e) y que expresa la enzima a fin de permitir la realización del paso (e) o
- (ii) introducir una enzima mediadora de rotura de la doble cadena en la célula hospedadora antes o durante el paso (e) y proporcionar la enzima a fin de permitir la realización del paso (e).
- 25 13. El método de una cualquiera de las reivindicaciones 1-12, que comprende adicionalmente expresar el gen de interés (GOI) y obtener el producto génico expresado.
14. Un método para la obtención de productos génicos recombinantes, que comprende los pasos
- (a) producir una célula hospedadora comodín para la introducción de un gen de interés (GOI) proporcionando un vector diana que comprende
- 30 (i) un gen informador (RG1) enlazado operativamente a una primera secuencia de control de la expresión que comprende un promotor constitutivo (P1),
- (ii) un primer gen marcador de selección (SM1) enlazado operativamente a una segunda secuencia de control de la expresión (P2),
- 35 (iii) un segundo gen marcador de selección no funcional (SM2) sin enlace operativo a una secuencia de control de la expresión,
- en donde la secuencia (i) está localizada en la posición 5', la secuencia (ii) está localizada entre (i) y (iii), y la secuencia (iii) está localizada en la posición 3',
- (iv) un primer y un segundo sitio de reconocimiento para una enzima mediadora de rotura de la doble cadena, particularmente una meganucleasa, en donde el primer sitio de reconocimiento está localizado entre la primera secuencia de control de la expresión (P1) y el primer gen informador (RG1), y el segundo sitio de reconocimiento 5 está localizado entre la secuencia (ii) y la secuencia (iii),
- 40 (b) introducir el vector diana en una célula hospedadora en condiciones tales que permiten la integración aleatoria del vector diana en el genoma de la célula hospedadora, y
- 45 (c) seleccionar una célula hospedadora comodín que tiene integrado de manera estable el vector diana y que exhibe una actividad de transcripción del gen informador (RG1) que es al menos 10 veces mayor comparada con los niveles medios de expresión de la población de clones resultante,
- (d) proporcionar un vector de intercambio que comprende:
- (i) un gen de interés, y

(ii) un segundo gen marcador de selección inactivo (Δ SM2) enlazado operativamente a una tercera secuencia de control de la expresión (P3),

5 en donde el vector de intercambio comprende una primera secuencia homóloga 5' y una segunda secuencia homóloga 3', que permiten la recombinación con secuencias del vector diana integradas en la célula hospedadora,

(e) introducir el vector de intercambio en la célula hospedadora proporcionada en el paso (a) en condiciones que permiten una rotura de la doble cadena en el primer y/o el segundo sitio de reconocimiento del vector diana y una integración del vector de intercambio en el genoma de la célula hospedadora por recombinación homóloga mediada por rotura de la doble cadena,

10 con lo cual el segundo gen marcador de selección inactivo (Δ SM2) se activa por una integración del vector de intercambio,

(f) seleccionar una célula productora que tiene integrado el vector de intercambio por recombinación homóloga con el vector diana integrado, en donde la célula productora expresa el GOI.

Fig. 1

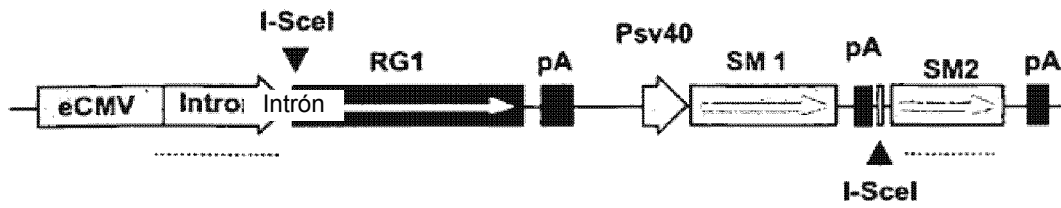


Fig. 2

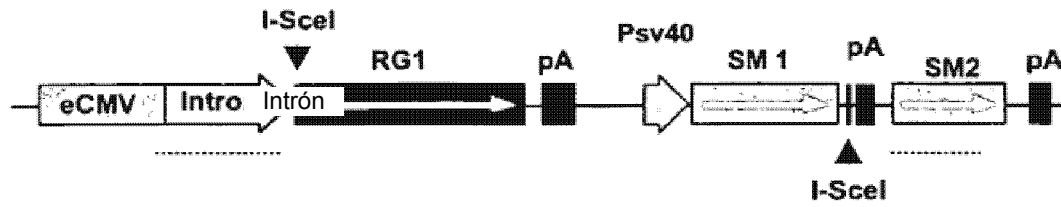


Fig. 3

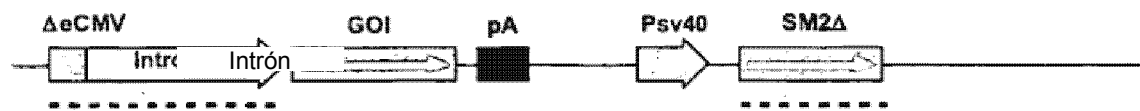


Fig. 4



Fig. 5

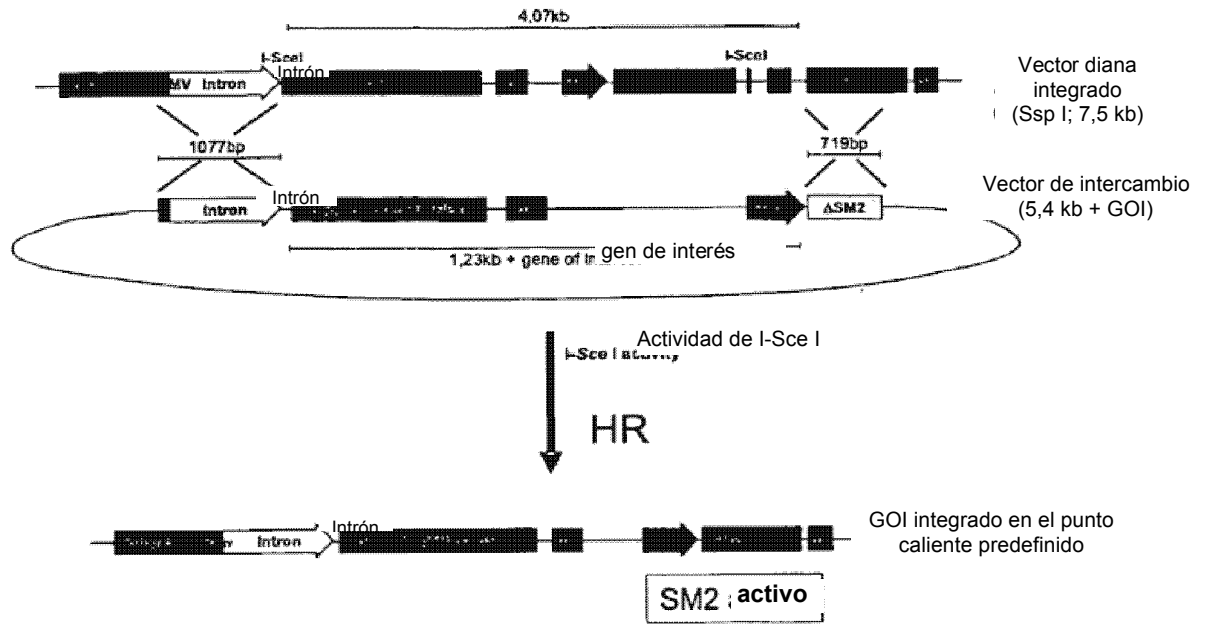


Fig. 6

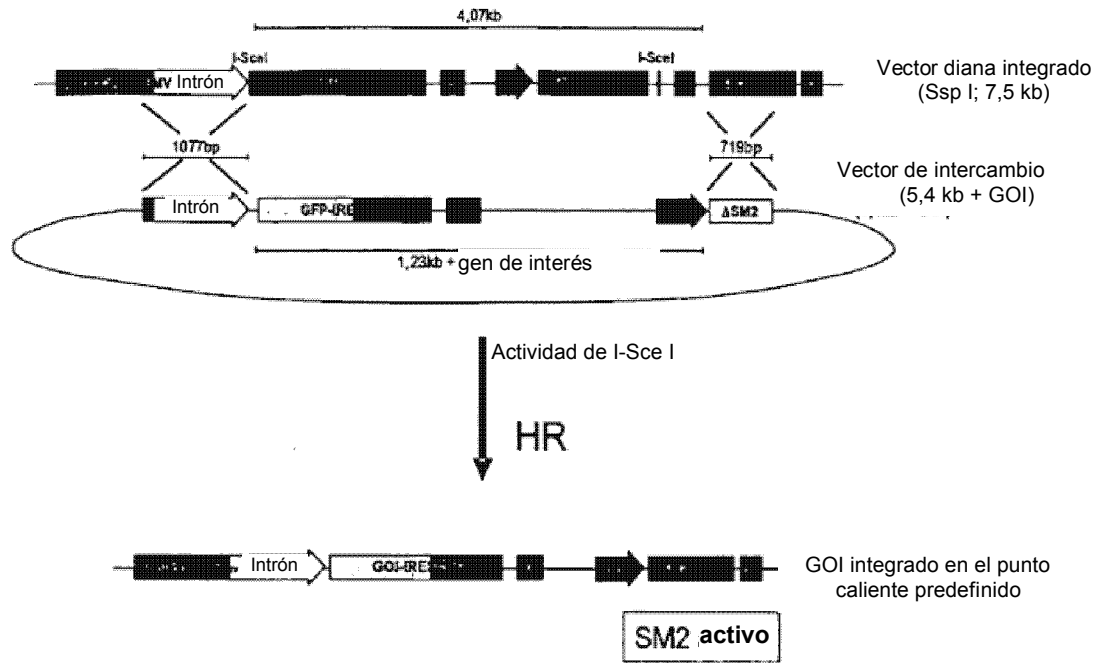


Fig. 7

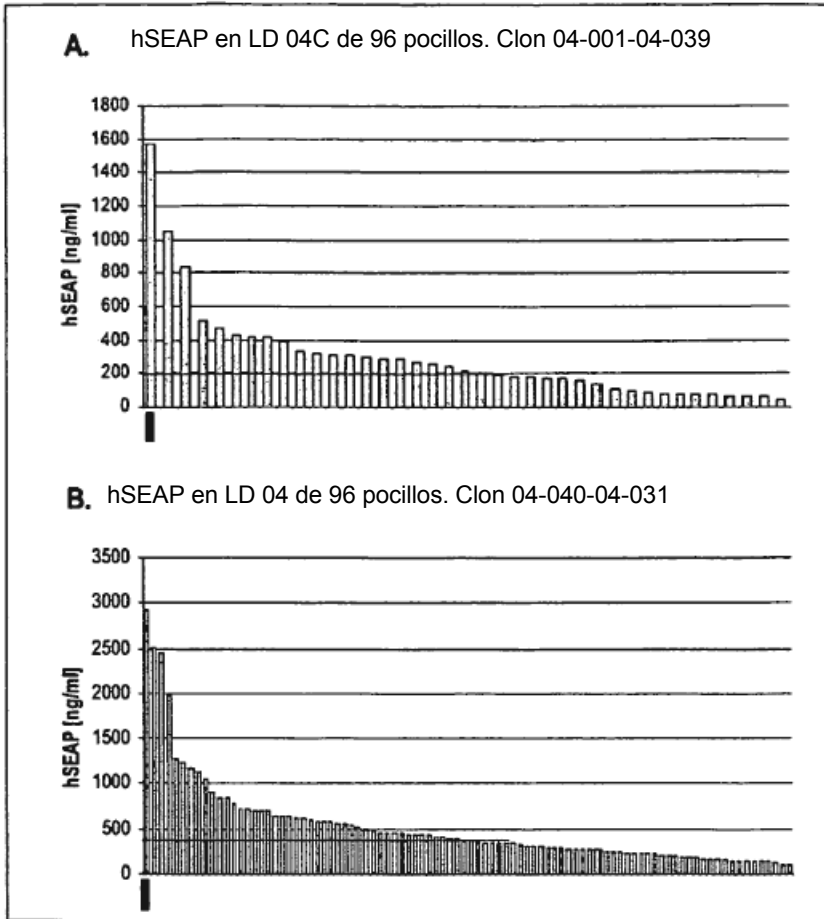


Fig. 8

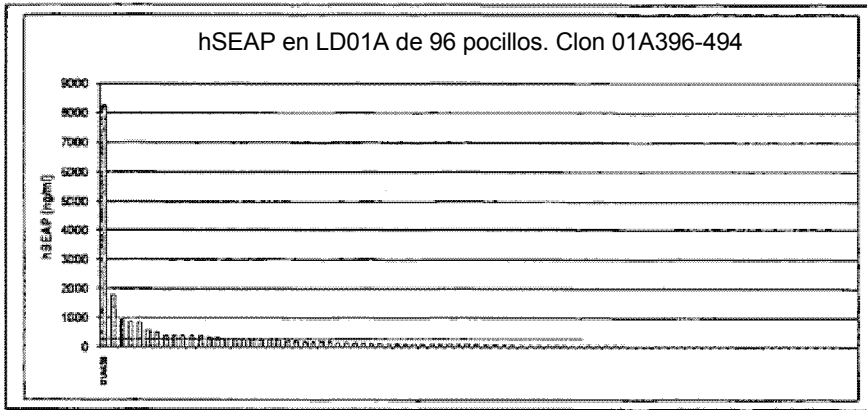


Fig. 9

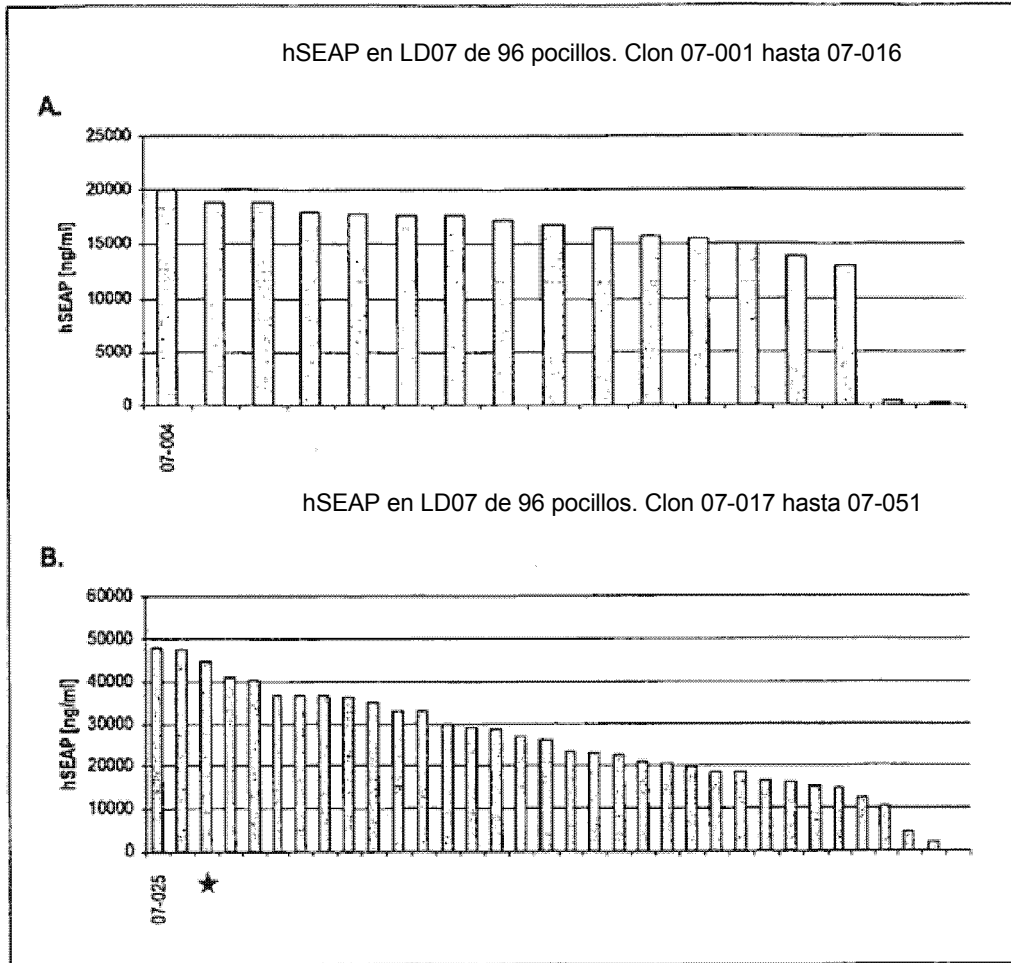


Fig. 10

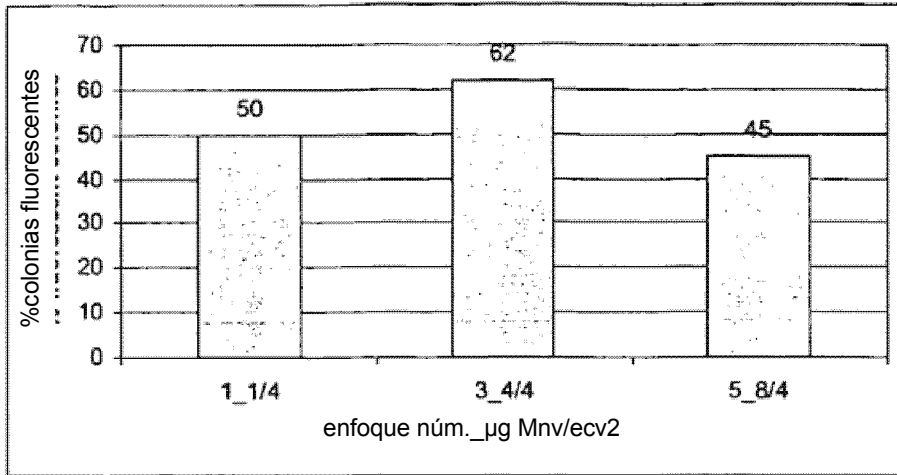


Fig.11

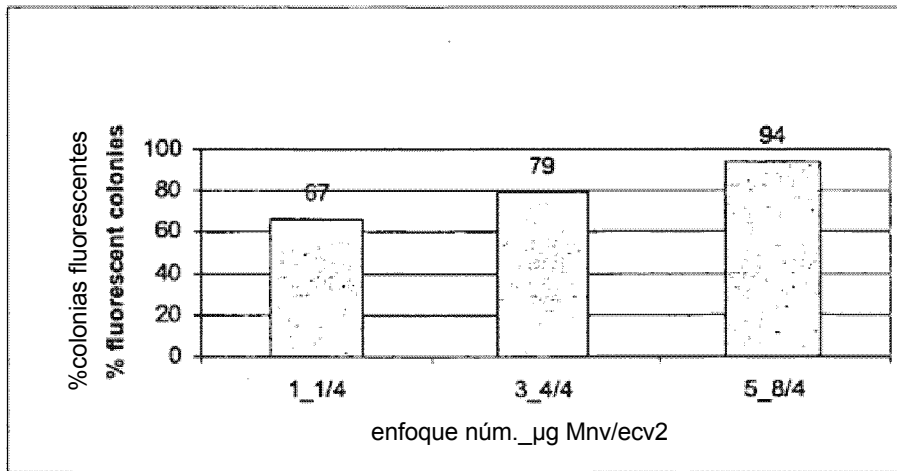


Fig. 12

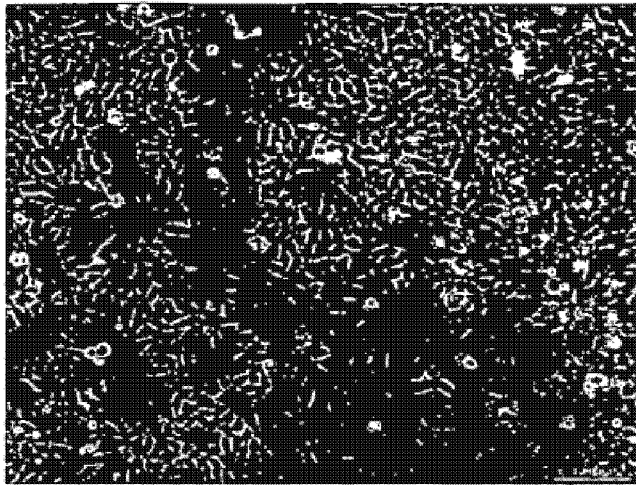


Fig. 13

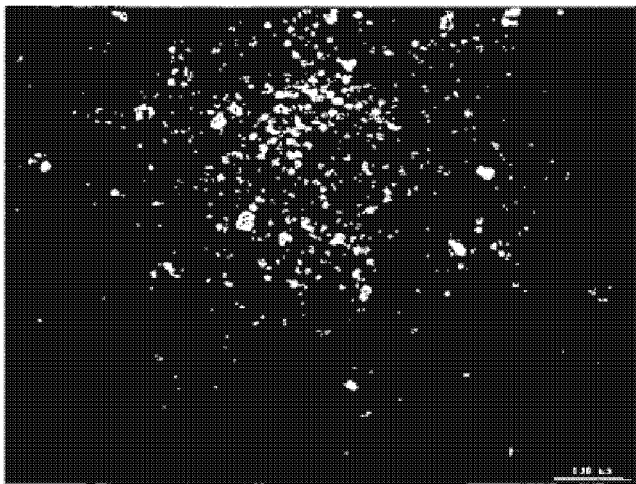


Fig. 14

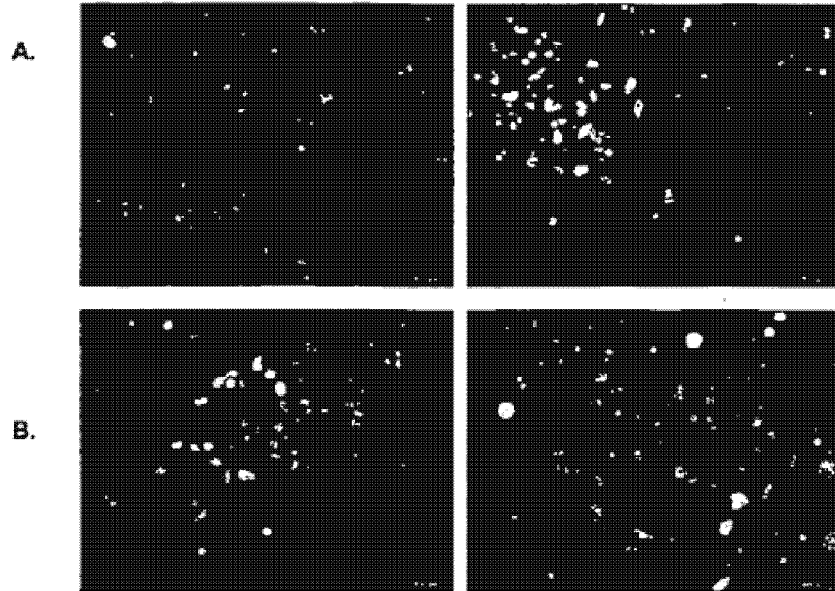


Fig. 15

