

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 811 251**

51 Int. Cl.:

C12N 15/09 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **18.09.2014 PCT/GB2014/052836**

87 Fecha y número de publicación internacional: **26.03.2015 WO15040401**

96 Fecha de presentación y número de la solicitud europea: **18.09.2014 E 14772197 (1)**

97 Fecha y número de publicación de la concesión europea: **08.07.2020 EP 3047022**

54 Título: **Producción de vectores de expresión y selección de células de alta capacidad de procesamiento**

30 Prioridad:
19.09.2013 GB 201316644

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.03.2021

73 Titular/es:
**KYMAB LIMITED (100.0%)
The Bennet Building (B930), Babraham Research
Campus
Cambridge CB22 3AT, GB**

72 Inventor/es:
**BRADLEY, ALLAN;
LEE, E-CHIANG;
LIANG, QI;
LIU, HUI;
WOOD, ANDREW y
WONG, VIVIAN**

74 Agente/Representante:
ELZABURU, S.L.P

ES 2 811 251 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

Producción de vectores de expresión y selección de células de alta capacidad de procesamiento

La presente invención y descripción se refiere, entre otras cuestiones, a la producción de vectores de expresión, así como a su aplicación a la producción de células hospedantes para la expresión de repertorios de proteínas y la selección de alta capacidad de procesamiento. La invención o descripción también se refiere a ácidos nucleicos, cebadores de PCR y mezclas útiles para la amplificación por PCR de secuencias de nucleótidos que codifican dominios variables de anticuerpos humanos o la recombinación homóloga con loci de Ig en células.

Antecedentes

La técnica reconoce el deseo de seleccionar repertorios de células o proteínas para identificar proteínas de interés ("proteins of interest", POI) o células que las expresan. Por ejemplo, la técnica comprende técnicas para seleccionar células B y repertorios de anticuerpos para identificar una o más células que expresan anticuerpos que muestran una característica deseada (generalmente, unión específica al antígeno). La selección no se limita a la selección de anticuerpos, sino que también puede aplicarse a la selección de colecciones de otros tipos de proteínas para una o más características deseadas.

Para expresar repertorios de proteínas, las correspondientes secuencias de nucleótidos pueden clonarse en los respectivos vectores de expresión e introducirse (por ejemplo, transfectarse) en células hospedantes que pueden expresar las proteínas. De modo habitual, se usa la clonación molecular para clonar secuencias que codifican proteínas a partir de un repertorio en células hospedantes. La clonación molecular se ha aplicado a técnicas de selección de células B (linfocitos), en las que las células B se expanden mediante cultivo, se clasifican en células individuales (por ejemplo, empleando un ensayo de placa hemolítica o un ensayo de focos fluorescentes), el ARNm de la cadena de anticuerpo (o región variable) procedente de las células clasificadas se somete a transcripción inversa y se amplifica usando RT-PCR, el ADN amplificado se somete a clonación molecular para introducir las secuencias que codifican cadenas de anticuerpo en vectores de ácidos nucleicos, y después los vectores se introducen en células hospedantes para la expresión transitoria (cuando el vector es episómico) o para la expresión estable (mediante integración aleatoria en el genoma de la célula hospedante). Véase, por ejemplo, Proc. Natl. Acad. Sci. U.S.A., 23 de julio de 1996, 23, 93(15):7843-8, "A novel strategy for generating monoclonal antibodies from single, isolated lymphocytes producing antibodies of defined specificities", J. S. Babcook *et al.* (conocida como la técnica SLAM), y doi: 10.1016/j.jala.2009.05.004, Journal of Laboratory Automation, octubre de 2009, vol. 14, n.º 5, 303-307, "High-Throughput Screening for High Affinity Antibodies", S. Tickle *et al.* (conocida como la técnica UGB SLAM).

La clonación molecular implica amplificar ADNc mediante PCR (producido por RT-PCR de ARNm de POI) con cebadores que portan sitios de clonación de enzimas de restricción. Esto produce productos de PCR en los que cada secuencia de nucleótidos de las POI está flanqueada por sitios de restricción. Los productos PCR después se digieren con las enzimas de restricción apropiadas y se subclonan en vectores de expresión vacíos que proporcionan un promotor y una señal de poliA mediante acoplamiento, así como portan un marcador de selección. Los productos acoplados después se transforman en *E. coli* y se seleccionan los clones mediante la presencia del marcador. Es necesario confirmar y seleccionar los clones con las inserciones de secuencia correctas mediante un cartografiado de restricción laborioso y la secuenciación de los clones individuales. Después, cada vector de expresión del clon correcto individual se purifica individualmente de *E. coli* y se transfecta en una célula hospedante (por ejemplo, una célula de mamífero CHO o HEK293) para la expresión transitoria o estable.

Por tanto, la clonación molecular es una técnica laboriosa que implica múltiples etapas y no está bien adaptada a la alta capacidad de procesamiento.

La expresión transitoria, tal como a partir de un módulo de expresión lineal, habitualmente es limitada, y se prefiere la expresión estable para una expresión a largo plazo. Para la expresión estable, generalmente la secuencia de interés que codifica la POI se inserta en el genoma de mamífero mediante el método de integración clásico de integración espontánea de ADN extraño (es decir, la integración aleatoria en el genoma). Esta estrategia a menudo conduce a una significativa variación transcripcional (y, por tanto, a una imprevisibilidad y expresión incoherente de la POI) como resultado de diferencias en el número de copias del transgén y el sitio de integración (véase, Henikoff S. (1992), "Position effect and related phenomena", Curr. Opin. Genet. Dev., 2(6):907-912; Martin D.I., Whitelaw E. (1996), "The variegating transgenes", Bioessays, 18(11):919-923; y Whitelaw E. *et al.* (2001), "Epigenetic effects on transgene expression", Methods Mol. Biol., 158:351-368). Además, los fragmentos de transgenes integrados de esta manera a menudo se insertan como concatámeros, que pueden provocar la inactivación del gen mediante el silenciamiento del gen inducido repetidamente (véase, Garrick D. *et al.* (1998), "Repeat-induced gene silencing in mammals", Nat. Genet., 18(1):56-59; y McBurney M.W. *et al.* (2002), "Evidence of repeat-induced gene silencing in cultured Mammalian cells: inactivation of tandem repeats of transfected genes", Exp. Cell. Res., 274(1):1-8). Estos problemas dificultan la producción de repertorios de POI y de poblaciones celulares para expresar dichos repertorios de POI.

- Existen muchas tecnologías para la generación de anticuerpos monoclonales (mAb) a partir de seres humanos o animales transgénicos que portan un repertorio de anticuerpos. En general, los mAb se obtienen a partir de la inmortalización de células B por fusión (tecnología del hibridoma) o transformación (transfección de virus o transformación de oncogenes). Sin embargo, estos métodos de inmortalización de células no son adecuados para una selección exhaustiva de grandes repertorios de anticuerpos, porque tienen un fuerte sesgo, son ineficaces y, generalmente, solo muestrean una proporción minúscula del repertorio disponible (generalmente, menos del 0,1% (células inmortalizadas/células de entrada) de un repertorio de células B obtenido de un ratón inmunizado, por ejemplo). Por tanto, el uso de métodos de selección de células B alternativos que no requieran de inmortalización resulta atractivo, pero en la actualidad, las técnicas se enfrentan a los problemas analizados anteriormente.
- 5
- 10 Duvall *et al.* ((2011), *Mabs*, 3:203-208) describen métodos usados en la producción de anticuerpos humanos derivados de células B sin estimular humanas.
- DeKosky *et al.* ((2013), *Nature Biotechnology*, 31:166-169) describen la secuenciación de alta capacidad de procesamiento de parejas de V_H:V_L procedentes de linfocitos B humanos individuales.
- 15 Kurosawa *et al.* ((2012), *BMC Biology*, 10:80) describen la producción de anticuerpos monoclonales específicos de antígeno a partir de una diversidad de animales.
- El documento WO2013/000982 (Vivalis) describe un método para seleccionar células.
- El documento US2007/0184546 (Hudson Control Group, Inc.) describe una célula de trabajo automática capaz de realizar experimentos de laboratorio genéticos de principio a fin moviendo muestras o microplacas entre los instrumentos para el análisis.
- 20 El documento WO2008/142028 (Boehringer Ingelheim RCV GMBH & CO) describe un método para producir una proteína recombinante a una escala de fabricación.
- El documento US7.884.054 (Amgen, Inc.) describe métodos para seleccionar anticuerpos multiméricos producidos por células de mamífero.
- 25 El documento US6.258.529 (OraVax, Inc.) describe métodos para aislar genes de región variable de inmunoglobulina y el uso de estos genes para la producción de anticuerpos quiméricos y de isotipo cambiado.
- El documento US6.919.183 (Regeneron Pharmaceuticals, Inc.) describe un método para identificar y aislar células que producen proteínas segregadas.
- El documento US2007/0141048 (Oleksiewicz) describe un método para conectar secuencias de interés.
- 30 Liao *et al.* ((2009), *J. Virol. Methods*, 158:171-179) describen el aislamiento de alta capacidad de procesamiento de genes de inmunoglobulina a partir de células B individuales y su expresión como anticuerpos monoclonales.
- Boria *et al.* ((2008), *BMC Immunology*, 9:50) describen conjuntos de cebadores para clonar el repertorio humano de regiones variables de receptores de células T.
- Hogrefe y Shopes ((1994), *Genome Res.*, 4:S109-S122) describen la construcción de bancos de presentación de fágmidos.
- 35 Smith *et al.* ((2009), *Nat. Protoc.*, 4:372-384) describen un protocolo para la producción de anticuerpos monoclonales humanos específicos de antígeno.
- Tiller *et al.* ((2008), *J. Immunol. Methods*, 329:112-124) describen la generación de anticuerpos monoclonales a partir de células B humanas individuales mediante RT-PCR de células individuales y clonación de vectores de expresión.
- 40 Franz *et al.* ((2011), *Blood*, 118:348-357) describen la caracterización *ex vivo* y el aislamiento de células B de memoria poco frecuentes con tetrámeros de antígenos.
- Beerli *et al.* ((2008), *PNAS*, 105:14336-14341) describen métodos para el aislamiento de anticuerpos monoclonales humanos mediante la presentación de células de mamífero.
- 45 Tiller ((2011), *New Biotechnology*, 28:453-457) es un análisis de tecnologías de células B individuales para el descubrimiento de anticuerpos.

Sumario de la invención

La invención soluciona la necesidad de técnicas de producción de vectores que sean adecuadas para la producción fiable de alta capacidad de procesamiento para repertorios y para la selección, en particular con un potencial para la automatización.

La invención es como se indica en las reivindicaciones. Según la invención, las proteínas de interés ("proteins of interest", POI) referidas en la presente son anticuerpos.

Una primera configuración de la descripción proporciona:

En un primer aspecto:

- 5 un método para producir células que codifican un repertorio de proteínas de interés (POI), comprendiendo dicho método:
- a) proporcionar una población de células que expresan un repertorio de POI;
 - b) clasificar la población de células para producir una población clasificada de células individuales, y cada célula comprende un ácido nucleico que codifica una respectiva POI;
 - 10 c) amplificar el ácido nucleico comprendido en la población de células individuales clasificadas para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican POI;
 - d) modificar los ácidos nucleicos que codifican POI amplificados clasificados procedentes de la etapa (c) para producir un repertorio clasificado de módulos de expresión, y cada módulo comprende una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI; y
 - 15 e) transferir los módulos de expresión de POI desde dicho repertorio de módulos a una población clasificada de células hospedantes, manteniendo al mismo tiempo la clasificación de los módulos de expresión de POI, y producir un repertorio clasificado de células hospedantes que expresan un repertorio clasificado de POI.

En una realización, la clasificación se mantiene en la etapa (e) mediante el suministro de repertorios en una pluralidad de recipientes cuyas localizaciones relativas y disposiciones globales están fijadas. Esto permite un procesamiento de alta capacidad de procesamiento y una automatización de los métodos de la invención y descripción, por ejemplo, para una selección de células rápida y eficaz para seleccionar una o más secuencias de interés de POI. Además, o como alternativa, la transferencia de los módulos de expresión en la etapa (e) se realiza mediante transferencia de lotes, y esto también permite un procesamiento de alta capacidad de procesamiento y una automatización del método de la invención.

25 En un segundo aspecto de la descripción:

un aparato automático para realizar el método de cualquier aspecto o configuración anterior, comprendiendo dicho aparato:

- a. un medio para contener una población de células individuales clasificadas en una pluralidad de recipientes, en el que cada célula individual está en un respectivo recipiente, y cada célula comprende un ácido nucleico que codifica una respectiva POI;
- 30 b. un medio para trasladar reactivos de PCR a los recipientes para amplificar el ácido nucleico comprendido en la población de células individuales clasificadas para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican POI;
- c. un medio para trasladar a los recipientes reactivos para modificar los ácidos nucleicos que codifican POI amplificados y clasificados para producir un repertorio clasificado de módulos de expresión, y cada módulo comprende una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI;
- 35 d. un medio para contener una población clasificada de células hospedantes en una pluralidad de recipientes;
- e. un medio para transferir los módulos de expresión POI desde dicho repertorio de módulos a una población clasificada de células hospedantes en los recipientes, manteniendo al mismo tiempo la clasificación de los módulos de expresión de POI; y
- 40 f. un medio para realizar la transfección de los módulos de expresión hacia el interior de las células hospedantes en los recipientes para producir un repertorio clasificado de células hospedantes que expresan un repertorio clasificado de POI.

45 Una segunda configuración de la presente descripción proporciona:

En un primer aspecto:

un módulo de expresión para la expresión de una POI en una célula hospedante, y el módulo es proporcionado por un ácido nucleico lineal que comprende un transposón, y el transposón comprende elementos de transposón 5'- y 3'-terminales con una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para la

expresión de la POI entre los elementos de transposón.

En un segundo aspecto:

5 una población clasificada de módulos de expresión que codifican un repertorio de POI que se corresponden con POI expresadas por una población de células, y cada módulo comprende una secuencia de nucleótidos que codifica un miembro del repertorio de POI y uno o más elementos reguladores para la expresión de POI, en la que cada uno de dichos módulos comprende la disposición (en la dirección 5' a 3'): elemento de transposón - [secuencia de nucleótidos de POI y uno o más elementos reguladores] - elemento de transposón, y los módulos de expresión para la expresión de POI de diferentes células están aislados entre sí en la población clasificada (por ejemplo, en diferentes pocillos de una placa).

10 En un tercer aspecto:

un método para preparar un transposón que comprende una secuencia de nucleótidos de interés ("nucleotide sequence of interest", NOI), comprendiendo dicho método:

15 a. proporcionar una primera secuencia de nucleótidos que comprende (en la dirección 5' a 3') A, B y C, en el que A es una primera secuencia de homología, B es una secuencia de nucleótidos que comprende la NOI, y C es una segunda secuencia de homología;

b. proporcionar una primera secuencia de nucleótidos de molde que comprende (en la dirección 5' a 3') W y X, en el que W es una secuencia de nucleótidos que comprende un primer elemento de transposón, y X es una tercera secuencia de homología; y

20 c. proporcionar una segunda secuencia de nucleótidos de molde que comprende (en la dirección 5' a 3') Y e Z, en el que Y es una cuarta secuencia de homología, y Z es una secuencia de nucleótidos que comprende un segundo elemento de transposón; y bien

25 d. (i) mezclar la primera secuencia de nucleótidos con el primer molde para hibridar entre sí el primer y el tercer brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la primera secuencia de nucleótidos usando el primer molde para producir una primera secuencia de nucleótidos extendida (primera ENS) que comprende (en la dirección 5' a 3') W, B y C; y (ii) mezclar la primera ENS con el segundo molde para hibridar entre sí el segundo y el cuarto brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la primera ENS para producir una segunda ENS que comprende (en la dirección 5' a 3') W, B y Z; o bien

30 (ii) mezclar la primera secuencia de nucleótidos con el segundo molde para hibridar entre sí el segundo y el cuarto brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la primera secuencia de nucleótidos usando el segundo molde para producir una tercera secuencia de nucleótidos extendida (tercera ENS) que comprende (en la dirección 5' a 3') A, B y Z; y (ii) mezclar la tercera ENS con el primer molde para hibridar entre sí el primer y el tercer brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la tercera ENS para producir una cuarta ENS que comprende (en la dirección 5' a 3') W, B y Z; o bien

35 (iii) mezclar la primera secuencia de nucleótidos con el primer y el segundo molde para hibridar entre sí el primer y el tercer brazo de homología y para hibridar entre sí el segundo y el cuarto brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la primera secuencia de nucleótidos usando el segundo molde para producir una quinta ENS que comprende (en la dirección 5' a 3') W, B y Z;

y

40 e. aislar una ENS que comprende (en la dirección 5' a 3') W, B y Z, produciendo con ello un transposón aislado que comprende una NOI flanqueada por elementos de transposón; y

f. opcionalmente, introducir el transposón aislado en una célula receptora, de modo que el transposón se integra en el genoma de la célula.

Una tercera configuración de la presente descripción proporciona:

un método para producir una célula hospedante para la expresión de una POI, comprendiendo dicho método:

45 a. proporcionar al menos un primer y un segundo módulo de expresión, en el que cada módulo de expresión comprende:

i. un primer elemento de integración y un segundo elemento de integración 3' de la secuencia de nucleótidos del primer elemento de integración; y

50 ii. entre los elementos de integración, una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI;

iii. en el que los elementos de integración son capaces de insertarse en un ácido nucleico mediante el reconocimiento de un motivo de secuencia de nucleótidos predeterminado del ácido nucleico usando una enzima de integración;

b. proporcionar una célula hospedante cuyo genoma comprende una pluralidad de dichos motivos; y

5 c. de modo simultáneo o secuencial, introducir el primer y el segundo módulo de expresión en la célula hospedante, en el que cada módulo se integra genómicamente en el genoma de la célula hospedante en dicho motivo para la expresión de las POI por la célula hospedante; y

d. opcionalmente, producir una expresión de las POI en la línea celular en una etapa que comprende cultivar la célula hospedante.

10 Una cuarta configuración de la presente descripción proporciona:

Una mezcla de ácidos nucleicos que comprende un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en la que el primer ácido nucleico es capaz de hibridarse con una secuencia de 5'UTR de una región V de un anticuerpo humano de un gen comprendido en un ácido nucleico diana, en la que el gen codifica una región V humana; y el segundo ácido nucleico es capaz de hibridarse con una segunda secuencia, en la que la segunda secuencia está comprendida en el ácido nucleico diana y se encuentra 3' con respecto a la secuencia de UTR, en la que el primer ácido nucleico aislado comprende una secuencia que es al menos 90% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-47.

Una mezcla de ácidos nucleicos que comprende un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en la que los ácidos nucleicos son diferentes y se seleccionan de ácidos nucleicos que comprenden una secuencia que es al menos 90% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-47.

Un método para amplificar un repertorio de secuencias de región variable humanas usando una o más de las secuencias.

Breve descripción de las figuras

25 La figura 1A es un esquema que ilustra un ejemplo no limitante de los métodos de la invención y descripción para producir un repertorio de sitios de unión de anticuerpos (POI) en células hospedantes (por ejemplo, células CHO o HEK293).

La figura 1B es un esquema alternativo que ilustra un ejemplo no limitante de los métodos de la invención y descripción para producir un repertorio de sitios de unión de anticuerpos (POI) en células hospedantes (por ejemplo, células CHO o HEK293).

30 Figura 2: Ejemplos de la estrategia de compuertas para diferentes poblaciones de células B, (a) un ejemplo de una gráfica de contorno citométrico de flujo que muestra la compartimentación de la población de CD38⁺ CD95⁺ células de memoria después de la exclusión de células B IgM e IgD, (b) cada célula de memoria individual positiva a CD19 (azul pacífico) y antígeno (ovoalbúmina-AlexaFluor-488) después se clasifica en pocillos separados en una placa de 96 pocillos.

35 Figura 3: Un diagrama de flujo que resume la producción de alta capacidad de procesamiento de anticuerpos a partir de células B individuales. Se realizó una síntesis de ADNc basada en células individuales usando cebadores específicos de región constante. Se usó una mezcla de cebadores específicos del gen V para la cadena pesada y la cadena ligera con un fragmento del promotor de citomegalovirus humano (hCMV) en el extremo 5' para la primera ronda de PCR para amplificar el fragmento del gen V. Se empleó un cebador directo genérico que se asoció al marcador de hCMV con un cebador anidado inverso para la región constante para la segunda ronda de PCR. Los productos amplificados después se unieron en puente con un módulo de Ig lineal con LTR de PB 5'-promotor de CMV y región constante-señal de poliA-LTR de PB 3'.

45 Figura 4: Producción de construcciones de expresión de cadena pesada y ligera mediante PCR en puente. (a) geles de agarosa teñidos con bromuro de etidio de las parejas de fragmentos de genes de V_H y V_L amplificados a partir de células B individuales. Cada carril contiene 10 µl de un producto de la segunda ronda de la PCR de V_H+V_L de 30 µl, (b) geles de agarosa teñidos con bromuro de etidio que muestran los resultados de la PCR en puente. Cada carril contiene 10 µl de un producto de puente de 30 µl para la expresión de Ig. Los controles negativos muestran el módulo de Ig de las cadenas pesada y ligera. Se muestran los productos de la PCR representativos para una plataforma de PCR de 96 pocillos de alta capacidad de procesamiento.

50 Figura 5: Análisis de las secuencias de anticuerpos de las células B individuales específicas de Ag clasificadas procedentes del proyecto 1. Las secuencias de anticuerpos expresadas por las células B individuales se dispusieron mediante la utilización de la familia de genes V de cadena pesada, y se agruparon para generar los árboles filogenéticos mostrados.

Figura 6: Un ejemplo de una familia agrupada que muestra la maduración por afinidad mediante hipermutación para la afinidad aparente y la potencia de neutralización. CNROR: no fue posible determinar la constante de disociación.

Figura 7: Nivel de expresión aumentado por el sistema de transposón PiggyBac. Se ensayó la expresión de anticuerpos de los productos de PCR en puente transfectados en células HEK293 usando una cotransfección con diferentes cantidades de PBasa: 0; 20; 100; 500 ng/pocillo. Se recogieron los sobrenadantes en cada pocillo en el día 0; el día 2; y el día 5. Los datos demuestran que el sistema de transposón aumentó el nivel de expresión en 2-4 veces en el día 5.

Figura 8: Ejemplo de cuantificación de la concentración de IgG en el sobrenadante de células HEK293 transfectadas con productos de la PCR en puente. La concentración se determinó mediante un ELISA de sándwich ocho días después de la transfección. La concentración es de aproximadamente 200 ng/ml a 3000 ng/ml, que es comparable a la concentración en el sobrenadante de técnicas de hibridoma de la técnica anterior.

Figuras 9A y B: Sensogramas de SPR representativos de anticuerpos que se unen a la ovoalbúmina. Aproximadamente dos terceras partes de los anticuerpos ensayados mostraron evidencias de unión al antígeno, con una gama diversa de afinidades de unión y cinéticas diferenciadas.

Figura 10: La afinidad aparente de anticuerpos contra dos antígenos diana diferentes. Se detectó una gama de ligantes (en blanco), así como de neutralizantes funcionales (en negro), detectándose la mayor afinidad en el intervalo picomolar. Este experimento valida la tecnología de clonación de células B individuales de los inventores como una herramienta poderosa en la identificación y extracción de anticuerpos de alta afinidad y funcionalmente competentes.

20 Descripción detallada de la invención

La figura 1 es un esquema que ilustra un ejemplo del método de la invención o descripción para producir un repertorio de sitios de unión de anticuerpos (POI) en células hospedantes (por ejemplo, células CHO o HEK293), que es útil en un proceso de selección de células B. El repertorio de células hospedantes producido al final del esquema en la figura 1 es útil, puesto que las células expresan los sitios de unión al anticuerpo (por ejemplo, parejas de V_H/V_L) que pueden seleccionarse frente a un antígeno predeterminado para identificar las células que expresan sitios de unión de interés. Por tanto, esto permite conectar el genotipo al fenotipo, permitiendo la identificación de una secuencia de nucleótidos que codifica un sitio de unión de interés. Esta secuencia de nucleótidos después puede usarse para construir líneas celulares para la producción estable de un anticuerpo deseado, por ejemplo, para producir productos farmacéuticos de anticuerpos o reactivos de diagnóstico o de investigación que se dirigen al antígeno predeterminado. En un ejemplo, la célula hospedante producida al final del esquema de la figura 1 es, en sí misma, una célula que expresa de modo estable el sitio de unión, y esto ofrece la posibilidad de generar una línea celular estable conveniente y eficaz para la fabricación a más largo plazo del sitio de unión. Tal como se analiza más a fondo a continuación, en un ejemplo puede realizarse la inserción mediada por sitio de módulos de expresión que codifican POI en células hospedantes, lo cual proporciona la ventaja añadida de una expresión de POI estable que evita la integración aleatoria y los concatámeros.

Aunque los ejemplos se proporcionan en términos de sitios de unión de anticuerpos, cadenas y dominios variables de anticuerpos, la invención no se limita a estas proteínas. En este sentido, la invención es aplicable a cualquier proteína de interés (POI). El término "proteína", en este contexto, también incluye polipéptidos y péptidos que son de interés, así como dominios o fragmentos de proteínas o polipéptidos.

Con respecto al ejemplo no limitante de la figura 1, en una primera etapa se proporciona una población de células B deseada, que puede ser de células germinales, células de memoria, plasmablastos, células plasmáticas o, en general, puede ser de células secretoras de anticuerpos ("antibody-secreting cells", ASC), o mezclas de dos o más de estos tipos de células B. Los expertos en la técnica están familiarizados con la selección de dichas poblaciones, por ejemplo, usando marcadores de la superficie celular opcionalmente con FACS. Estos marcadores pueden seleccionarse de CD19, IgM, IgD, CD30 y CD95, por ejemplo, y pueden incluir un panel de uno o más de estos (por ejemplo, 1, 2, 3, 4), o pueden incluir un panel de todos estos marcadores. Las células que portan estos marcadores pueden teñirse, por ejemplo, con cualquier fluoróforo de molécula pequeña que pueda detectarse mediante el sistema de clasificación de células, tales como Alexa-488, Alexa-647, azul pacífico, R-ficoeritrina, isotiocianato de fluoresceína, o alofocianina opcionalmente conjugada con un tinte de cianuro, por ejemplo, Cy7. La población de células B puede generarse en una etapa preliminar mediante el aislamiento a partir de uno o más animales (por ejemplo, ratones, ratas o seres humanos o animales no humanos), por ejemplo, un animal que ha sido inmunizado con un antígeno diana. Así, los métodos de la invención y descripción son útiles para seleccionar poblaciones de células B para identificar una o más secuencias de POI derivadas de una célula B, en el que la POI se une al antígeno diana con una característica deseada. Esta característica, por ejemplo, es la unión al antígeno diana y/o un antígeno ortólogo, homólogo o relacionado desde el punto de vista estructural (por ejemplo, procedente de una especie diferente) y/o la unión a un antígeno con una afinidad deseada y/o la competición con un anticuerpo conocido por la unión a un antígeno predeterminado (por ejemplo, el antígeno diana) y/o la unión a un epitopo predeterminado.

Opcionalmente, la población de células B de entrada deseada (por ejemplo, plasmablastos o células plasmáticas) puede seleccionarse realizando una selección de células específica de antígeno usando técnicas conocidas en la técnica (por ejemplo, véase Jin, A *et al.*, *Nature Medicine*, 2009, 15(9):1088-1093). El resultado es una población de células B que expresan sitios de unión de anticuerpos que se unen específicamente a un antígeno diana predeterminado. Los inventores han descubierto que resulta ventajoso realizar esta etapa para racionalizar el proceso de selección de células B global, consiguiendo de este modo una selección eficaz y con una alta capacidad de procesamiento aún mayor.

En general, las células GC (centro germinal) o células B de memoria específicas de antígeno pueden ser capturadas por un antígeno marcado porque expresan predominantemente anticuerpos transmembrana sobre su superficie celular. El antígeno puede marcarse de modo fluorescente (por ejemplo, con cualquier fluoróforo de molécula pequeña que pueda detectarse mediante el sistema de clasificación de células, tales como Alexa-488, Alexa-647, azul pacífico, R-ficoeritrina, isotiocianato de fluoresceína, o alofocianina opcionalmente conjugada con un tinte de cianuro, por ejemplo, Cy7). Si la población de células B ha sido previamente teñida para la selección inicial, resulta ventajoso usar un fluoróforo que tenga una longitud de onda de emisión diferente del fluoróforo empleado en esa selección inicial, para facilitar el proceso de clasificación, por ejemplo, utilizando FACS. En una alternativa, la célula puede teñirse simultáneamente para la presencia del marcador de selección celular y seleccionarse para la unión a VLP que portan antígenos/antígenos fluorescentes. Aunque no se pretenda limitación alguna por la teoría, se cree que, por otra parte, los plasmablastos o las células plasmáticas no serán capturados con tanta facilidad por el antígeno marcado, debido a su expresión dominante de anticuerpos secretores.

En una realización alternativa, el antígeno puede capturarse con partículas similares a virus ("virus-like particles", VLP) con un antígeno recombinante sobre la superficie. Las VLP pueden generarse a partir de células CHO, células KEK, fibroblastos embrionarios de ratón ("mouse embryonic fibroblasts", MEF) u otras líneas de células de mamífero con coexpresión del antígeno recombinante, la proteína gag de retrovirus, y MA-GFP (proteína de fusión del fragmento de matriz de gag p15-GFP). La expresión de la proteína gag permite que las VLP germinen de las células, y la MA-GFP marca las VLP para la detección de la fluorescencia. Ambas proteínas gag y MA-GFP están asociadas con la superficie interna de la membrana plasmática, y el antígeno recombinante está sobre la superficie de las VLP. Los antígenos sobre las VLP se presentan en la forma nativa, directamente expresados a partir de las células recombinantes sin ninguna etapa de purificación o modificación. La forma nativa de un antígeno debería proporcionar todos los epitopos naturales que ayudan en gran medida a la selección de anticuerpos neutralizantes. La alta densidad del antígeno sobre las VLP aumenta la proporción de señal/ruido para la detección de células que expresan anticuerpos específicos de antígeno sobre la superficie celular, y facilita en gran medida la etapa de clasificación. Las VLP recombinantes pueden generarse con la expresión de diferentes proteínas fluorescentes, tales como MA-CFP o MA-YFP. Usando la multiplexación de VLP con diferentes antígenos y diferentes proteínas de fluorescencia, pueden seleccionarse las células que expresan ligantes de alta afinidad, ligantes de reactividad cruzada o ligantes específicos de homólogo. Las células que expresan ligantes de alta afinidad pueden ser seleccionadas por células con una matriz de afinidad relativamente alta (matriz de afinidad = la proporción de actividad de unión a VLP con baja densidad de antígenos a VLP con alta densidad de antígenos). Las células que expresan ligantes de reactividad cruzada con ortólogos o antígenos diferentes (para el aislamiento de anticuerpos bispecíficos de 2 en 1) pueden ser seleccionadas por células que se unen a diferentes tipos de VLP al mismo tiempo. Las células que expresan ligantes específicos de homólogo también pueden ser seleccionadas por células que se unen solo al antígeno específico, pero no a su homólogo.

Por tanto, las células B se clasifican (por ejemplo, usando FACS) para proporcionar una población de células B individuales clasificadas. Generalmente, las células B se clasifican en pocillos de una placa convencional (por ejemplo, una placa de 96 pocillos o de 384 pocillos convencional) de modo que cada célula individual se encuentra en un respectivo pocillo y no se mezcla con otras células. Es posible que haya un número mínimo de pocillos (por ejemplo, menos del 5%, menos del 3%, menos del 2% o menos del 1%, o niveles no detectables) que no tengan células o que tengan más de una (por ejemplo, dos) células, y esto no altera la utilidad global del método de selección (y no se considera parte del repertorio deseado). Preferiblemente, cada pocillo de la placa contiene una única célula B. Opcionalmente, es posible cultivar células directamente antes y/o después de la etapa de clasificación de células, pero los inventores han descubierto que esto no es necesario, a diferencia de las técnicas en la técnica. Así, evitando esta etapa, los métodos de la invención y descripción se prestan aún más a la racionalización y la alta capacidad de procesamiento.

Después, en el ejemplo mostrado en la figura 1, se amplifica el ácido nucleico que codifica la POI. En el ejemplo, esto se realiza mediante la transcripción inversa del ARNm en las células de la población de células clasificada (es decir, el ARNm que codifica la POI se convierte en el correspondiente ADNc que codifica la POI) y este se amplifica usando PCR. Puede emplearse una RT-PCR convencional, tal como conocerán los expertos en la técnica (por ejemplo, véase Dixon AK *et al.*, *Trends Pharmacol. Sci.*, 2000, 21(2):65-70). Esto produce un repertorio de ADN que codifica un repertorio de regiones variables de anticuerpos. En este ejemplo, ambas secuencias de V_H y V_L son copiadas y amplificadas para cada célula clasificada. Los expertos en la técnica saben que cada célula expresa un solo tipo de sitio de unión de anticuerpo y, por tanto, solo se amplificará un único tipo de secuencia de V_H amplificada y un solo tipo de secuencia de V_L amplificada por cada célula clasificada individual (por tanto, cuando se emplean pocillos, solo se obtendrá un único tipo de V_H y V_L por pocillo, lo cual conserva el agrupamiento de

secuencias de V_H y V_L que forman sitios de unión).

Las secuencias que codifican POI también se modifican para producir un repertorio de módulos de expresión para expresar POI procedentes de células hospedantes en una etapa posterior del método. Los módulos de expresión contienen una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para la expresión (por ejemplo, un promotor y/o un potenciador y/o una poliA). En una realización, las etapas de amplificación y de modificación pueden realizarse simultáneamente usando PCR y los moldes y cebadores apropiados, tal como será evidente para los expertos en la técnica. En otra realización, la amplificación y la modificación se realizan en etapas distintas, por ejemplo, la amplificación y después la modificación. En el ejemplo de la figura 1, primero se realiza una RT-PCR y después las secuencias que codifican POI amplificadas se modifican para producir módulos de expresión que comprenden cada uno (en la dirección 5' a 3'): promotor-secuencia de nucleótidos de POI-poliA. En este ejemplo, también se añaden elementos de transposón flanqueantes. Los elementos de transposón pueden ser elementos de repetición terminal invertidos del transposón piggyBac (PB) para formar un ácido nucleico de transposón que comprende la estructura: [elemento de PB 5'] - [promotor] - [secuencia de nucleótidos de POI] - [poliA] - [elemento de PB 3']. En un ejemplo, cada módulo de expresión es suministrado por un ADN lineal que comprende o que consiste en el transposón.

Elemento de PB 5':

```
GATATCTATAACAAGAAAATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATT
ATCGTATGAGTTAAATCTTAAAAGTCACGTAAGATAATCATGCGTCATTTTGACTCACGCGGTCGTTA
TAGTTCAAATCAGTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGA
TGTCTAAATGCACAGCGACGGATTTCGCGCTATTTAGAAAGAGAGCAATATTTCAAGAATGCATGCGT
CAATTTTACGCAGACTATCTTTCTAGGGTTAA
```

Elemento de PB 3':

```
TTTGTTACTTTATAGAAGAAATTTTGAGTTTTTGTTTTTTTTAATAAATAAATAAACATAAATAAATTGT
TTGTTGAATTTATTATTAGTATGTAAGTGTAATATAATAAACTTAATATCTATTCAAATTAATAAATAA
ACCTCGATATACAGACCGATAAAACACATGCGTCAATTTTACGCATGATTATCTTTAACGTACGTCACAAT
ATGATTATCTTTCTAGGGTTAA
```

Tal como se muestra en el ejemplo de la figura 1, puede usarse una PCR en puente (véase, por ejemplo, Mehta R.K., Sihgh J., *Biotechniques*, 1999, 26(6):1082-1086) para construir el transposón añadiendo los elementos de transposón flanqueantes.

En esta etapa del método, se ha obtenido un repertorio de módulos de expresión que codifican un repertorio de secuencias de nucleótidos que codifican POI derivadas de la población de células de entrada. En este ejemplo, los procedimientos de amplificación y de modificación se realizan en las células individuales que están clasificadas en pocillos de una o más placas. Así, el resultado es un repertorio clasificado de módulos de expresión (en este caso por pocillo individual, y el pocillo contiene una pluralidad de un tipo de módulo de expresión de V_H y un tipo de módulo de expresión de V_L derivado de un tipo de sitio de unión de anticuerpo de una sola célula de entrada). Después, los módulos de expresión clasificados (en el presente ejemplo, contenidos en los respectivos transposones y, opcionalmente, como ADN lineales) se transfieren a células hospedantes para producir un repertorio clasificado de células hospedantes que expresan un repertorio clasificado de POI. Esto puede realizarse, en el presente ejemplo, transfiriendo los módulos de expresión desde los pocillos en una placa (o un conjunto de placas) a una o más placas distintas que tienen una pluralidad de células que contienen células hospedantes (por ejemplo, un tipo de células idéntico de la misma línea celular, por ejemplo, células CHO o HEK293 o de levadura). En esta realización, el repertorio de módulos clasificados puede transferirse de modo manual o con un robot usando una pipeta de múltiples canales (por ejemplo, una pipeta de 4, 8, 12, 16 (2 x 8), 64 (8 x 8), 96 (12 x 8), 384 o 1536 canales, tal como se conoce en la técnica) para extraer simultáneamente los módulos (en este caso, las parejas de módulos V_H/V_L clasificadas) desde los pocillos individuales en una primera placa y transferirlos a los correspondientes pocillos en la segunda placa, en la que estos pocillos contienen células hospedantes. Haciendo esto, se mantiene la localización relativa y la naturaleza clasificada de los módulos de expresión cuando se transfieren a las células hospedantes. Esto tiene la ventaja de mantener la conexión de pares de secuencias de módulos V_H/V_L sin mezclar (es decir, por pocillo en la segunda placa, y el pocillo contiene células hospedantes y secuencias de V_H y V_L derivadas solo de una única célula B de entrada). De modo ventajoso, si una parte de la muestra de módulo se deja sobre la primera placa, con ello se puede identificar con facilidad uno o más pocillos

5 como fuente de una secuencia que codifica una POI útil (y un módulo de expresión y transposón) remitiéndose a los positivos deseados descubiertos en la segunda placa después de la selección de la segunda placa para una característica de sitio de unión de anticuerpo deseada. Además, de modo ventajoso, las células hospedantes pueden proporcionarse en un medio que apoye, de modo deseable, el funcionamiento y el mantenimiento de las células hospedantes (generalmente diferente del entorno usado en las células de la primera placa para la amplificación y la modificación de la secuencia). En una realización alternativa, las células hospedantes se añaden a los pocillos de la placa que contiene el repertorio de módulos de expresión de modo que se conserva la disposición clasificada (pero esto entonces no proporciona las ventajas adicionales de la otra realización, en la que se conserva una placa maestra de módulos de expresión sin mezcla con las células hospedantes, útil, por ejemplo, para realizar una segunda transferencia a otra placa que tenga un tipo diferente de célula hospedante, tal como cuando se desea evaluar la actuación de la expresión de POI y la presentación/secreción por un tipo de célula hospedante diferente).

10 Un ejemplo de una pipeta de múltiples canales automática es el sistema de manipulación de líquidos automático de 96 canales Thermo Scientific Matrix Hydra II. V&P Scientific VP 177AD-1 y VP 179BJD son distribuidores de dispensación diseñados para el llenado rápido de placas de 96 y 384 pocillos, respectivamente, y cualquiera de los dos puede utilizarse en los métodos de la invención y descripción para transferir módulos a células hospedantes o para la manipulación de muestras en general.

15 Por tanto, el resultado de los métodos de la invención o descripción, en su aspecto más amplio, es la producción de un repertorio de células hospedantes que expresan un repertorio de POI. Tal como se analizó anteriormente, esto se puede usar después en una etapa posterior de selección (por ejemplo, empleando un ensayo de unión de células, ELISA, resonancia de plasmón de superficie u otro ensayo apropiado para la naturaleza concreta de la POI) para identificar una o más células hospedantes que expresan una POI con una característica deseada. Después se puede aislar la POI de la célula o del medio circundante y/o aislar (y, opcionalmente, replicar o amplificar) una secuencia de nucleótidos (por ejemplo, una secuencia de ADN) que codifica la POI deseada. Puede determinarse la secuencia de nucleótidos. Estas células aisladas pueden cultivarse para producir una línea celular que codifica una POI (lo cual es especialmente útil cuando el módulo se ha integrado de modo estable en el genoma de la célula, como es posible con la integración genómica dirigida a sitio, es decir, realizada usando un transposón). La secuencia de nucleótidos que codifica la POI puede insertarse en un vector de expresión diferente y/o mutarse (por ejemplo, madurarse por afinidad) o condensarse con otra secuencia de proteína.

20 La integración con transposones se realiza suministrando una correspondiente enzima transposasa a las células hospedantes (por ejemplo, mediante cotransfección de los módulos de expresión con un vector que comprende un gen transposasa inducible, o suministrando células hospedantes que portan dicho gen transposasa, por ejemplo, de modo inducible). En un ejemplo, cada módulo comprende elementos del transposón piggyBac, y el método usa una transposasa de piggyBac (por ejemplo, la transposasa de piggyBac de tipo salvaje o hiperactiva; véase, por ejemplo, Yusa, K *et al.*, PNAS USA, 2011, 10(4):1531-1536; y Yusa K *et al.*, "A hyperactive piggyBac transposase for mammalian applications", PNAS USA, 2010, 108(4):1531-1536).

PBasa de tipo salvaje:

MGCSLDDEHILSALLQSDDELVGEDSDSEISDHVSEDDVQSDTEEAFIDEVHEVQPTSSGSEILDEQNVIEQ
 PGSSLASNRILTLPQRTIRGKNKHCWSTSKSTRRSRVSALNIVRSQRGPTRMCRNIYDPLLCFKLFFTDEIISE
 IVKWTNAEISLKRRESMTGATFRDTNEDEIYAFFGILVMTAVRKDNHMSTDDLDFDRSLSMVYVSVMSRDRF
 DFLIRCLRMDKSI RPTLRENDVFTPVRKIWDLFIHQCIQNYTPGAHLTIDEQLLGFRGRCPFRMYIPNKPSK
 YGIKILMMCDSGTKYMINGMPYLGRGTQTNGVPLGEYVKELSKPVHGS CRNITCDNWFTSIPLAKNLLQEP
 YKLTIVGTVRSNKREIPEVLKNSR SRPVGTSMFCFDGPLTLVSYKPKPAKMVYLLSSCEDEDASINESTGKPQM
 VMYYNQTKGGVDTLDQMCSVMTCSRKTNRWPMALLYGMINIACINSFIIYSHNVSSKGEKVQSRKKFMRNL
 YMSLTSSFMRKRLEAPT LKRYLRDNISNILPNEVPGTSDDSTEPEVMKKRTYCTYCPSKIRRKANASCKKCKK
 VICREHNIDMCQSCF

PBasa hiperactiva:

MGSSLDDEHILSALLQSDDELVGEDSDSEVSDHVSEDDVQSDTEEAFIDEVHEVQPTSSGSEILDEQNVIEQ
 PGSSLASNRILTLPQRTIRGKNKHCWSTSKPTRRSRVSALNIVRSQRGPTRMCRNIYDPLLCFKLFFFTDEIISE
 IVKWTNAEISLKRRESMTSATFRDNEDEIYAFFGILVMTAVRKDNHMSTDDLDFRSLSMVYVSVMSRDRF
 DFLIRCLRMDKSIKIRPTLRENDVFTPVRKIWDLFIHQCIQNYTPGAHLTIDEQLLGFRGRCPPFRVYIPNKPSK
 YGIKILMMCDSGTKYMINGMPYLGRGTQTNGVPLGEYVVKELSKPVGHSRNRITCDNWFTSIPLAKNLLQEP
 YKLTIVGTVRSNKREIPEVLKNSRSPVGTSMFCFDGPLTLVSYKPKAKMVYLLSSCDEEDASINESTGKPQM
 VMYYNQTKGGVDTLDQMCSVMTCSRKTNRWPMALLYGMINIACINSFIIYSHNVSSKGEKVQSRKKFMRNL
 YMGLTSSFMRKRLEAPTALKRYLRDNISNILPKEVPGTSDDSTEETPMKKRTYCTYCPKIRRKASASCKKCKK
 VICREHNIDMCQSCF

Quando se emplea un transposón para la integración de módulos en la célula hospedante se obtienen beneficios adicionales. En primer lugar, los transposones pueden integrarse en varias copias en células hospedantes, proporcionando con ello múltiples copias de módulos de expresión para apoyar la expresión de alto nivel de POI. Esto puede estimularse aún más usando una enzima transposasa hiperactiva. Además, los transposones pueden integrarse preferentemente en sitios que son activos para la transcripción, con lo cual también se favorece una expresión eficiente y de alto nivel de la POI, tal como se demuestra mediante el análisis de cartografiado de los sitios de integración. Esto puede observarse, por ejemplo, con piggyBac (véase Wang W., *et al.*, "Chromosomal transposition of PiggyBac in mouse embryonic stem cells", 2008, PNAS USA, 105(27):9290-9295; Galvan D.L., *et al.*, "Genome-wide mapping of PiggyBac transposon integration in primary human T cells", J. Immunother., 2009, 32(8):837-844; y Yang W., *et al.*, "Development of a database system for mapping insertional mutations onto the mouse genome with large-scale experimental data", 2009, BMC Genomics, 10 (supl. 3):S7).

Así, la descripción proporciona los siguientes conceptos:

1. Un método para producir células que codifican un repertorio de proteínas de interés (POI), comprendiendo dicho método:

a) proporcionar una población de células que expresan un repertorio de POI;

b) clasificar la población de células para producir una población clasificada de células individuales, y cada célula comprende un ácido nucleico que codifica una respectiva POI;

c) amplificar el ácido nucleico comprendido en la población de células individuales clasificadas para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican POI;

d) modificar los ácidos nucleicos que codifican POI amplificados clasificados procedentes de la etapa (c) para producir un repertorio clasificado de módulos de expresión, y cada módulo comprende una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI; y

e) transferir los módulos de expresión de POI desde dicho repertorio de módulos a una población clasificada de células hospedantes, manteniendo al mismo tiempo la clasificación de los módulos de expresión de POI, y producir un repertorio clasificado de células hospedantes que expresan un repertorio clasificado de POI.

Las etapas (c) y (d) pueden realizarse por separado (en cualquier orden, por ejemplo, primero (c) y después (d)) o de modo simultáneo.

Opcionalmente, la etapa (c) comprende la amplificación por PCR de secuencias que codifican POI, por ejemplo, RT-PCR usando ARNm que codifican POI como molde (por ejemplo, uno, dos, tres o más cebadores que comprenden o consisten en una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-53, tal como se analiza más a fondo a continuación). Opcionalmente, el repertorio de ácidos nucleicos amplificados en la etapa (c) son ADN. En un ejemplo, cada POI es un dominio variable de anticuerpo, y la etapa (c) comprende la amplificación por PCR de secuencias que codifican POI usando uno o más cebadores 5' específicos de la región V y/o uno o más cebadores 3' de la región C (por ejemplo, CY, por ejemplo, cebador CY de ratón). Opcionalmente, la PCR comprende 5'- y/o 3'-RACE de secuencias de nucleótidos que codifican la POI. En un ejemplo, el 5'-RACE se realiza usando uno o más cebadores 5', y cada uno de ellos son homólogos con una 5' UTR o una secuencia de promotor de una región variable de anticuerpo. En un ejemplo, el 3'-RACE se realiza usando uno o más cebadores 5', y cada uno de ellos son homólogos con una región constante de anticuerpo, por ejemplo, una secuencia CH1 o Fc. En este caso, cada secuencia que codifica una POI amplificada codifica una cadena de anticuerpo que comprende un dominio variable y

una región constante de anticuerpo. En un ejemplo, el 3'-RACE usa una o más secuencias de la región constante humana como cebador; esto entonces produce secuencias que codifican regiones variables humanizadas, en las que cada región variable está condensada con una región constante humana (por ejemplo, una gamma CH1 o Fc (por ejemplo, gamma Fc) humana), proporcionando con ello una cadena de anticuerpo humano (POI) tras la posterior expresión. Esta humanización durante la etapa (c) es útil, puesto que las POI identificadas a partir de la selección posterior representan cadenas humanas que pueden usarse para producir productos terapéuticos de anticuerpos para un uso humano. En un ejemplo, el 5'-RACE usa un molde 5' que comprende un promotor de región variable para producir ácidos nucleicos amplificados que comprenden (en la dirección 5' a 3'): un promotor y una secuencia de nucleótidos que codifican una POI. Además, o como alternativa, se usa un 3'-RACE, en el que el RACE emplea un molde 3' que comprende una secuencia de poliA para producir ácidos nucleicos amplificados que comprenden (en la dirección 5' a 3'): una secuencia de nucleótidos que codifica una POI y una poliA. En este caso, la amplificación y la modificación para producir módulos de expresión pueden realizarse de modo simultáneo (es decir, las etapas (c) y (d) se realizan de modo simultáneo).

En un ejemplo, la modificación de la etapa (d) se realiza usando PCR, por ejemplo, PCR en puente. Por ejemplo, la etapa (d) se realiza después de la etapa (c), por ejemplo, después de la amplificación con RT-PCR o RACE. En este caso, la PCR en puente se realiza en una etapa que comprende hibridar un primer cebador con el extremo 5' de los productos de ácido nucleico de la etapa (c); e hibridar un segundo cebador al extremo 3' de los productos de ácido nucleico de la etapa (c) (o al extremo 3' del producto de ácido nucleico de la etapa de hibridación usando el primer cebador). Como alternativa, el segundo cebador puede usarse inicialmente (para hibridarse con el producto de la etapa (c)), y el producto de esto puede hibridarse con el primer cebador. Como alternativa, el primer y el segundo cebador y el producto de la etapa (c) pueden mezclarse y realizarse una PCR. En cualquier caso, el resultado es un producto extendido que comprende (en el orden 5' a 3'):

[una secuencia del primer cebador 5'] - [un promotor] - [una secuencia de nucleótidos que codifica una POI] - [una poliA] - [una secuencia del segundo cebador 3']

En una realización, el promotor y la poliA se combinan con la secuencia de nucleótidos que codifica la POI mediante la etapa (c), tal como se describió anteriormente (por ejemplo, usando 5'- y 3'-RACE con los cebadores apropiados). En otra realización, el primer cebador usado en la etapa (d) comprende una secuencia de promotor (por ejemplo, tal como una secuencia en el extremo 3' del cebador). El resultado de la etapa (d) después combina el promotor con la secuencia de nucleótidos que codifica la POI. Además, o como alternativa, el segundo cebador usado en la etapa (d) comprende una secuencia de poliA (por ejemplo, tal como una secuencia en el extremo 5' del cebador). El resultado de la etapa (d) después combina la poliA con la secuencia de nucleótidos que codifica la POI. Son posibles otras combinaciones, por ejemplo, el promotor se añade en la etapa (c) usando el cebador apropiado, y la poliA se añade en la etapa (d) usando el segundo cebador apropiado. En un ejemplo, la etapa (c) (por ejemplo, RACE) añade secuencias 5' y/o 3' a los ácidos nucleicos producto que pueden usarse para la hibridación con cebadores en la etapa (d), en la que esta usa una PCR en puente.

El resultado de las etapas (c) y (d) siempre es un repertorio de módulos de expresión para la expresión de un repertorio de POI. En un ejemplo, se omiten uno o más elementos reguladores requeridos o deseados para la expresión (o la expresión óptima) de cada módulo, pero estos son proporcionados por los genomas de la célula hospedante tras haber introducido los módulos en las células hospedantes.

En una realización, la etapa (d) añade secuencias de integración 5' y 3' que flanquean el promotor y la poliA, respectivamente. Por ejemplo, la secuencia 5' es un elemento de transposón 5' (por ejemplo, un elemento terminal de PB 5') y la secuencia 3' es un elemento de transposón 3' (por ejemplo, un elemento terminal de PB 3' que está en una orientación invertida con respecto al elemento 5'). Por ejemplo, la secuencia de integración 5' se proporciona en el extremo 5' terminal del primer cebador de PCR en puente y/o la secuencia de integración 3' se proporciona en el extremo 3' terminal del segundo cebador de PCR en puente. El resultado es un repertorio de módulos de expresión, y cada uno termina (5' y 3') en elementos de transposón, es decir, cada módulo se modifica para formar un transposón. Opcionalmente, cada módulo de expresión se produce como un ADN lineal que termina en el extremo 5' y 3' con una secuencia de integración (por ejemplo, un elemento de transposón, por ejemplo, que termina en elementos del transposón PB terminales invertidos). Pueden usarse secuencias de integración alternativas en lugar de elementos de transposón (véase a continuación); por ejemplo, las secuencias de integración pueden ser brazos de homología (por ejemplo, al menos 15, 20, 50 o 100 nucleótidos contiguos) para realizar la integración homóloga en los genomas de las células hospedantes en uno o más sitios diana específicos (que se hibridan con los brazos de homología). Como alternativa, las secuencias de integración pueden ser secuencias de recombinación específicas de sitio (por ejemplo, lox, rox o frt) para la integración específica de sitio en los genomas de las células hospedantes que portan los correspondientes sitios de recombinación específicos de sitio en uno o más sitios de integración deseados en el genoma (tras el suministro o la expresión de la respectiva integrasa (es decir, cre, dre o flp, respectivamente). En una realización se emplea RMCE para la inserción usando dos sitios de recombinación incompatibles (por ejemplo, loxP de tipo salvaje y un lox mutante, por ejemplo, lox2272 o lox511).

En una realización, se añaden uno o más elementos reguladores para la expresión de POI (por ejemplo, un promotor y/o un potenciador y/o una señal de poliA y/o una secuencia señal) mediante la etapa (c), por ejemplo, usando RACE. Además, o como alternativa, se añaden uno o más elementos reguladores para la expresión de POI

(por ejemplo, un promotor y/o un potenciador y/o una señal de poliA y/o una secuencia señal) mediante la etapa (d), por ejemplo, usando una PCR en puente.

Además, o como alternativa, en una realización, se añade una secuencia de integración 5' y/o 3' mediante la etapa (c), por ejemplo, usando RACE. Además, o como alternativa, se añade una secuencia de integración 5' y/o 3' mediante la etapa (d), por ejemplo, usando una PCR en puente.

En una realización, se añaden uno o más elementos reguladores para la expresión de POI (por ejemplo, un promotor y/o un potenciador y/o una señal de poliA y/o una secuencia señal) y una secuencia de integración 5' y/o 3' mediante la etapa (c), por ejemplo, usando RACE.

En una realización, se añaden uno o más elementos reguladores para la expresión de POI (por ejemplo, un promotor y/o un potenciador y/o una señal de poliA y/o una secuencia señal) y una secuencia de integración 5' y/o 3' mediante la etapa (d), por ejemplo, usando una PCR en puente.

2. El método del concepto 1, en el que, en la etapa (e), los módulos de expresión clasificados se transfieren en lotes a las células hospedantes clasificadas.

La transferencia de lotes según esta realización de la invención y descripción es mejor que los métodos de la técnica anterior que utilizan la clonación molecular para transferir secuencias de nucleótidos que codifican POI hacia células hospedantes. Tal como se explicó anteriormente, esto último requiere una secuenciación laboriosa y larga, un análisis y una subclonación de las secuencias que codifican POI individuales que se han modificado mediante la inclusión de sitios de restricción terminales (para permitir la introducción en células hospedantes en posteriores etapas). Generalmente, tras haber confirmado una PCRd correcta de la secuencia que codifica una POI con sitios de restricción, esta entonces se selecciona como una secuencia individual para proseguir con la introducción en células hospedantes, que después se cultivan para producir una población de célula que expresan la secuencia de POI elegida. Este proceso de clonación molecular, laborioso y de múltiples etapas, se realiza para cada variante de POI en un repertorio que se va a incluir en la posterior selección. Por consiguiente, los métodos de selección de la técnica anterior pueden tardar varias semanas (generalmente, del orden de 6-8 semanas) en obtener un repertorio útil de células de partida, tales como células productoras de anticuerpos. Por contraste, los métodos de la invención y descripción, que emplean una transferencia de lotes clasificada de los módulos de expresión completos para POI (es decir, que incluyen POI y elementos reguladores para la expresión), proporcionan una técnica mucho más rápida para producir un repertorio clasificado de POI para la selección. Esto hace que los presentes métodos puedan adaptarse a la automatización de alta capacidad de procesamiento de la selección. Por ejemplo, los presentes inventores, mediante una operación manual de los métodos, consiguieron la producción de un repertorio de células hospedantes clasificadas para la expresión de POI en tan solo 2 días empleando placas de 4 x 96 pocillos (aproximadamente 186 células B de entrada). La selección del repertorio de POI expresado puede realizarse a mano tan solo en aproximadamente 2 días. Evidentemente, la automatización acelera esto aún más (y, de modo ventajoso, minimiza la contaminación cruzada entre partes alícuotas clasificadas).

Para la transferencia de lotes de módulos de expresión, una pluralidad de módulos de expresión clasificados se mezclan en la misma operación (por ejemplo, una única etapa de aspiración y traslado del módulo, es decir, una única etapa de pipeteado) con las células hospedantes para su transferencia hacia el interior de las células (por ejemplo, mediante una transfección posterior o simultánea hacia el interior de las células). Esta operación, por ejemplo, es una única transferencia mediante pipeta (por ejemplo, usando una pipeta de múltiples canales, por ejemplo, usando una pipeta de 4, 8, 12, 16, 64, 96, 384 o 1536 canales en una única operación). En un ejemplo, al menos 4 partes alícuotas clasificadas de módulos de expresión se mezclan con células hospedantes clasificadas en una única operación, de modo que cada parte alícuota de módulos de expresión se mezcla con una respectiva parte alícuota de células (por ejemplo, en un respectivo recipiente, por ejemplo, en un pocillo de una placa o un tubo en una rejilla). En un ejemplo, al menos 4, 8, 12, 16, 24, 32, 40, 48, 56, 64, 96, 384 o 1535 partes alícuotas clasificadas de módulos de expresión se mezclan en la misma operación con células hospedantes clasificadas en una única operación, de modo que cada parte alícuota de módulos de expresión se mezcla con una respectiva parte alícuota de células (por ejemplo, en un respectivo recipiente, por ejemplo, en un pocillo de una placa o un tubo en una rejilla). En una realización, la operación es una operación manual (por ejemplo, mediante pipeteado usando una pipeta de múltiples canales). En otra realización, la operación es automática, por ejemplo, se realiza mediante un aparato de manipulación de líquidos automático (por ejemplo, un robot de manipulación de líquidos).

De modo ventajoso, los inventores han podido transferir en lotes los módulos de expresión procedentes de la etapa (d) a las células hospedantes clasificadas sin tener que purificar los módulos, pero consiguiendo, al mismo tiempo, un repertorio de células hospedantes útiles en la etapa (e). Esto proporciona velocidades y capacidades de procesamiento más veloces y hace que el proceso pueda adaptarse a una automatización más sencilla.

3. El método del concepto 1 o 2, en el que (i) el repertorio clasificado de los módulos de expresión producidos mediante la etapa (d) se proporciona en una pluralidad de recipientes cuyas localizaciones relativas entre sí están fijadas (por ejemplo, pocillos en una placa o tubos en una rejilla), en el que cada recipiente contiene un respectivo tipo de módulo de expresión, de modo que la localización relativa de los módulos de expresión entre sí está predeterminada; y (ii) los módulos de expresión se transfieren a las células hospedantes clasificadas en la etapa (e),

de modo que las localizaciones relativas de los módulos de expresión se mantienen.

En un ejemplo, cada recipiente contiene módulos que codifican un primer tipo de POI derivadas de una respectiva célula individual proporcionada en la etapa (a) y también módulos que codifican un segundo tipo de POI derivadas de dicha célula individual. Por ejemplo, cada recipiente contiene módulos que codifican secuencias de V_H y V_L derivadas de una célula individual. De esta forma, los dominios variables que forman los respectivos sitios de unión de las células de entrada (cuando estas codifican sitios de unión de anticuerpo, por ejemplo, plasmablastos o células plasmáticas) se mantienen juntos en el repertorio clasificado, pero no se mezclan con secuencias derivadas de otra célula. Por tanto, esta clasificación, de forma ventajosa, se mantiene en etapas posteriores del proceso y puede rastrearse en el resultado de la posterior selección.

4. El método de cualquier concepto anterior, en el que (i) el repertorio de los módulos de expresión producidos mediante la etapa (d) se clasifica proporcionando una pluralidad de recipientes (por ejemplo, pocillos en una placa o tubos en una rejilla), en el que cada recipiente comprende secuencias que codifican POI de una célula individual clasificada en la etapa (b); (ii) las células hospedantes clasificadas de la etapa (e) se proporcionan en una pluralidad de recipientes (por ejemplo, pocillos en una placa o tubos en una rejilla), y (iii) los módulos de expresión se transfieren a las células hospedantes clasificadas en la etapa (e), de modo que las células hospedantes en cada respectivo recipiente se mezclan solo con las secuencias que codifican POI derivadas de una célula individual clasificada en la etapa (b).

5. El método del concepto 3 o 4, en el que, en la etapa (d), los módulos de expresión clasificados se proporcionan en una pluralidad de recipientes cuyas localizaciones relativas entre sí están fijadas (por ejemplo, una pluralidad de recipientes con una disposición de X recipientes por Y recipientes, por ejemplo, una placa de 8 x 12 pocillos (placa de 96 pocillos) o una placa que presenta un múltiplo de una disposición de 8 x 12 recipientes (por ejemplo, una placa de 384 pocillos)); y, en la etapa (e), las células hospedantes clasificadas se proporcionan en una pluralidad de recipientes cuyas localizaciones relativas entre sí están fijadas y comprenden la misma disposición que los recipientes usados en la etapa (d) (por ejemplo, los repertorios de las etapas (d) y (e) se proporcionan ambos en placas de 96 pocillos o de 384 pocillos con las mismas o sustancialmente las mismas dimensiones), de modo que la clasificación se mantiene en la etapa (e).

6. El método de cualquier concepto anterior, en el que el repertorio clasificado de células hospedantes es capaz de expresar de modo estable el repertorio de POI.

Como alternativa, el repertorio clasificado de células hospedantes es capaz de realizar una expresión transitoria. La expresión estable (por ejemplo, como resultado de la integración genómica de los módulos en los genomas de las células hospedantes) resulta ventajosa para el suministro a largo plazo de células (y, por tanto, las POI expresadas) identificadas después de la selección (y también de células mientras esperan a la selección, si las células hospedantes producidas en la etapa (e) se conservan un tiempo antes de ser empleadas para la selección).

7. El método de cualquier concepto anterior, en el que el repertorio clasificado de células hospedantes que expresan POI producidas mediante la etapa (e) se proporciona en una pluralidad de recipientes (por ejemplo, tubos o pocillos), en el que cada recipiente contiene POI de una célula individual clasificada en la etapa (b).

Estos tubos pueden estar fijados en una rejilla o soporte; estos pocillos pueden fijarse mediante su suministro en una o más placas.

8. El método del concepto 7, en el que cada célula hospedante expresa una primera y una segunda POI, en el que las POI son diferentes, por ejemplo, subunidades de una proteína, por ejemplo, dominios variables de un anticuerpo o sitios de unión al receptor de células T.

9. El método de cualquier concepto anterior, en el que la etapa (b) comprende clasificar células individuales hacia sus respectivos recipientes (por ejemplo, sus respectivos pocillos en una o más placas), y realizar las etapas (c) y (d) en dichos recipientes, al mismo tiempo que se mantiene la clasificación.

10. El método de cualquier concepto anterior, que comprende además seleccionar el repertorio de POI clasificado para identificar una POI con una característica deseada (por ejemplo, la unión a un antígeno o a un anticuerpo; o una afinidad de unión por un antígeno o ligando cognado) y/o una secuencia de nucleótidos que codifica la POI identificada (por ejemplo, ADN o ARN, por ejemplo, ARNm o ADNc).

11. El método del concepto 10, que comprende además identificar, amplificar o sintetizar la secuencia de nucleótidos que codifica la POI identificada (por ejemplo, usando PCR o cultivando una célula hospedante seleccionada o una célula derivada de esta); y opcionalmente, producir una POI aislada usando dicha secuencia de nucleótidos identificada, amplificada o sintetizada.

12. El método de cualquier concepto anterior, que comprende además seleccionar el repertorio de POI clasificado para identificar una POI con una característica deseada (por ejemplo, la unión a un antígeno o a un anticuerpo; o una afinidad de unión por un antígeno o ligando cognado) y aislar una célula hospedante que expresa la POI identificada; y, opcionalmente, propagar la célula para producir una línea celular que expresa la POI.

13. El método de cualquier concepto anterior, en el que la etapa (e) comprende integrar genómicamente (por ejemplo, integrar cromosómicamente) módulos de expresión de POI en los respectivos genomas de las células hospedantes para expresar las respectivas POI. Como alternativa, uno o más de los módulos se proporcionan episómicamente en su respectiva célula hospedante para la expresión transitoria de la POI.
- 5 14. El método del concepto 13, en el que dicha integración genómica se realiza usando un motivo de secuencia de nucleótidos genómico predeterminado para la inserción de los módulos de expresión en el respectivo genoma celular.
- Por ejemplo, el motivo es una secuencia de nucleótidos usada por un transposón para la integración (por ejemplo, el motivo TTAA usado por PB); o una secuencia de nucleótidos que puede recombinarse con secuencias de integración 5' y 3' del módulo mediante recombinación homóloga; o un motivo usado para integrar un sitio de recombinación específico de sitio.
- 10 15. El método del concepto 14, en el que cada genoma celular comprende más de una copia del motivo de secuencia.
- 15 16. El método del concepto 13, 14 o 15, en el que dicha integración genómica se realiza mediante una integración mediada por transposones.
- Los elementos de transposón adecuados para su uso como secuencias de integración 5' y 3' de módulos son elementos de transposón de clase II (por ejemplo, elementos de repetición terminal invertidos del transposón piggyBac o elementos del transposón Mariner), o elementos del transposón Bella Durmiente ("Sleeping Beauty") o elementos similares a Tc1 (TLE).
- 20 17. El método de uno cualquiera de los conceptos 13 a 16, en el que la etapa (e) comprende múltiples inserciones de módulos de expresión en los respectivos genomas de células hospedantes.
- Cada módulo se proporciona opcionalmente como parte de un ácido nucleico lineal (por ejemplo, ADN lineal). Por ejemplo, cada módulo es un transposón lineal que comprende o que consiste en elementos de transposón 5'- y 3'-terminales (por ejemplo, elementos de repetición terminal invertidos de piggyBac) con secuencias de nucleótidos de POI y uno o más elementos reguladores para la expresión entre los elementos de transposón. En una realización, existe otra secuencia 5' con respecto al elemento de transposón 5' y/o 3' con respecto al elemento de transposón 3'; en otra realización, estos elementos están en los extremos 5' y 3' terminales del módulo, respectivamente.
- 25 18. El método de cualquier concepto anterior, en el que las células hospedantes son células de una línea de células de mamífero (por ejemplo, células CHO o HEK293) o células de levadura.
- 30 Por ejemplo, un método en el que cada célula hospedante es una célula de mamífero (por ejemplo, una célula humana o de un animal no humano, vegetal o de insecto o roedor o ratón o rata o conejo o pollo o camélido o pez), bacteriana o de levadura.
19. El método de cualquier concepto anterior, en el que, en la etapa (a), las células son células aisladas a partir de uno o más animales.
- 35 Opcionalmente, cada célula de la etapa (a) es una célula de mamífero (por ejemplo, una célula humana o de un animal no humano, vegetal o de insecto o roedor o ratón o rata o conejo o pollo o camélido o pez), bacteriana o de levadura.
- Opcionalmente, todas las células de la etapa (a) son células del mismo tipo de tejido o compartimento de uno o más organismos. Por ejemplo, todas son células de hígado, riñón, corazón, cerebro, sangre, linfocitos, de próstata, ovario o germinales de uno o más organismos, por ejemplo, un paciente humano o un animal no humano, o un roedor o ratón o rata o conejo o pollo o camélido o pez.
- 40 20. El método de cualquier concepto anterior, en el que, en la etapa (a), las células comprenden o consisten en células B, células del centro germinal, células B de memoria, células secretoras de anticuerpos, células plasmáticas o células de plasmablastos.
- 45 21. El método de cualquier concepto anterior, en el que cada POI es una cadena de inmunoglobulina (por ejemplo, un anticuerpo o un receptor de células T) o una de sus partes (por ejemplo, un dominio variable).
22. El método de cualquier concepto anterior, en el que cada POI comprende o consiste en un dominio variable de anticuerpo (por ejemplo, un dominio VH, VHH o VL o un dAb o un Nanobody™).
- 50 23. El método de cualquier concepto anterior, en el que cada célula de la etapa (a) expresa una primera y una segunda POI, en el que las POI son diferentes entre sí; en el que la etapa (b) comprende clasificar células individuales en sus respectivos recipientes (por ejemplo, sus respectivos pocillos en una o más placas) y realizar las etapas (c) y (d) en dichos recipientes, en el que, después de la etapa (c), cada recipiente comprende las primeras POI amplificadas mezcladas con POI amplificadas procedentes de la misma célula; y en el que la etapa (e)

- comprende mezclar los respectivos ácidos nucleicos que codifican la primera y la segunda POI procedentes de su respectivo recipiente con las células hospedantes; en el que la primera y la segunda POI procedentes de la misma célula en la etapa (a) se transfieren a la misma célula hospedante para la expresión de la primera y la segunda POI por la célula hospedante, produciendo con ello un repertorio de células hospedantes clasificadas que coexpresan, cada una, la respectiva primera y segunda POI.
24. El método del concepto 23, en el que la primera y la segunda POI procedentes de la misma célula son polipéptidos cognados que, juntos, forman una proteína funcional (por ejemplo, dominios V_H y V_L que forman un sitio de unión al antígeno).
25. El método del concepto 24, en el que la primera y la segunda POI comprenden o consisten en dominios V_H y V_L de anticuerpos, respectivamente, por ejemplo, la primera y la segunda POI son cadenas pesadas y ligeras de anticuerpos cognadas, respectivamente.
26. El método de cualquier concepto anterior, en el que la etapa (b) comprende unir las POI expresadas por células a un ligando cognado (por ejemplo, unir sitios de unión de anticuerpos expresados por células a un antígeno de interés); opcionalmente, en el que el ligando se une a una POI expresada en la superficie de una célula; y
- después clasificar y aislar las células que expresan las POI que se unen al ligando, produciendo con ello la población clasificada de células.
27. El método del concepto 26, en el que se emplea la clasificación de células FACS; opcionalmente, FACS de fluorescencia.
28. El método del concepto 26 o 27, en el que cada célula clasificada de la etapa (b) se proporciona en un respectivo recipiente (por ejemplo, un pocillo en una placa), de modo que cada uno de dichos recipientes (por ejemplo, pocillos) no comprende más de un tipo de célula.
29. El método del concepto 28, en el que la población de células clasificada se proporciona en pocillos en una o más placas, que comprenden en total menos del 5, 4, 3, 2, 1 o 0,5% de pocillos que contienen más de una célula y/o en total menos del 5, 4, 3, 2, 1 o 0,5% de pocillos que no contienen células.
30. El método de cualquier concepto anterior, en el que la etapa (c) se realiza usando PCR, por ejemplo, RT-PCR usando ARN que codifica una POI (por ejemplo, ARNm).
31. El método de cualquier concepto anterior, en el que la etapa (d) se realiza usando PCR, por ejemplo, PCR en puente.
32. El método de cualquier concepto anterior, en el que la etapa (d) comprende la modificación de ácidos nucleicos amplificados mediante la combinación con una secuencia predeterminada, de modo que dicha secuencia predeterminada flanquea en 5' y/o 3' a las secuencias de nucleótidos que codifican POI de los ácidos nucleicos; opcionalmente, en el que la modificación coloca un elemento regulador (por ejemplo, un promotor 5' y/o una poliA 3') y/o un elemento de transposón (por ejemplo, un elemento del transposón piggyBac) que flanquea en 5' y/o 3' a las secuencias de nucleótidos que codifican POI.
33. El método del concepto 32, en el que cada POI comprende un dominio variable de anticuerpo, y la etapa (c) o (d) combina secuencias de nucleótidos que codifican POI con una región constante de anticuerpo (opcionalmente, una región constante humana o de una especie que es diferente de la especie de la región C comprendida en las POI en las células de la etapa (a)) para producir una secuencia de nucleótidos que codifica una cadena de anticuerpo (opcionalmente, una cadena humanizada).
- Por ejemplo, las células de la etapa (a) codifican POI que comprenden regiones constantes de vertebrados no humanos (por ejemplo, roedor, por ejemplo, ratón o rata), y estas son reemplazadas por regiones constantes humanas mediante la etapa (c) o (d). Por ejemplo, las POI son cadenas de anticuerpo (por ejemplo, cadenas pesadas) que comprenden un dominio variable humano y una región constante de roedor (por ejemplo, ratón o rata) que se humaniza mediante el método de la invención. Esto resulta conveniente, puesto que proporciona un modo de alta capacidad de procesamiento para humanizar cadenas de anticuerpo y anticuerpos a escala, y permite la posterior selección, producción y desarrollo de células y vectores de expresión en el contexto de formatos finales de cadena/anticuerpo humano adecuados para un uso como fármaco terapéutico humano. Las técnicas anteriores no consiguen esto.
34. Un método según el concepto 33, que comprende además seleccionar el repertorio de POI clasificado para identificar una célula hospedante que expresa una cadena de anticuerpo con una característica deseada (por ejemplo, la unión específica a un antígeno o una afinidad de unión a un antígeno), identificar la secuencia de nucleótidos que codifica la cadena de anticuerpo de la célula hospedante, usar la secuencia de nucleótidos que codifica la cadena de anticuerpo para producir copias de la cadena de anticuerpo identificada, y formular las copias como una composición farmacéutica (opcionalmente, en combinación con uno o más fármacos adicionales, excipientes, diluyentes o vehículos) para la terapia médica humana; y, opcionalmente, administrar la composición a

un paciente humano para la terapia médica del paciente.

35. El método de cualquier concepto anterior, en el que la etapa (e) es automática; opcionalmente, en el que una o todas las etapas (b) a (c) también son automáticas.

5 La automatización puede incluir el control del proceso mediante un ordenador programado para realizar el método de cualquier aspecto, configuración, realización o ejemplo de la invención.

36. El método de cualquier concepto anterior, en el que (i) las etapas (b) a (e) inclusive se realizan en un equivalente de al menos 180 células (con la condición de que la etapa (a) se procesa en no más de 1 o 2 días); y/o (ii) el repertorio de las POI expresadas se selecciona para una característica deseada, y se identifican una o más correspondientes células hospedantes o secuencias de nucleótidos que codifican POI en un equivalente de no más de 4 días.

10

Los inventores han logrado esto usando aproximadamente 400 células B de entrada y seleccionando para anticuerpos específicos de antígeno, y (a) tarda 2 días y (b) tarda 3 días (todo se realizó manualmente). Evidentemente, sería aún más rápido si se emplea la automatización. Por tanto, la invención o descripción proporciona un ahorro de tiempo significativo frente a las técnicas de la técnica actual, que generalmente tardan 6-8 semanas para realizar dicha selección.

15

37. Un aparato automático para realizar el método de cualquier concepto anterior, comprendiendo dicho aparato:

a. un medio para contener una población de células individuales clasificadas en una pluralidad de recipientes (por ejemplo, pocillos en una o más placas, o tubos en una rejilla o soporte, tal como se describió anteriormente), en el que cada célula individual está en un respectivo recipiente, y cada célula comprende un ácido nucleico que codifica una respectiva POI;

20

b. un medio para trasladar reactivos de PCR a los recipientes para amplificar el ácido nucleico comprendido en la población de células individuales clasificadas para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican POI;

c. un medio para trasladar a los recipientes reactivos para modificar los ácidos nucleicos que codifican POI amplificados y clasificados para producir un repertorio clasificado de módulos de expresión, y cada módulo comprende una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI;

25

d. un medio para contener una población clasificada de células hospedantes en una pluralidad de recipientes;

e. un medio para transferir (opcionalmente, para transferir en lotes) los módulos de expresión POI desde dicho repertorio de módulos a una población clasificada de células hospedantes en los recipientes, manteniendo al mismo tiempo la clasificación de los módulos de expresión de POI; y

30

f. un medio para realizar la introducción (por ejemplo, la transfección) de los módulos de expresión hacia el interior de las células hospedantes en los recipientes para producir un repertorio clasificado de células hospedantes que expresan un repertorio clasificado de POI; y

g. opcionalmente, un ordenador programado para realizar el método de cualquier aspecto, configuración, realización o ejemplo de la invención o descripción.

35

38. El aparato del concepto 37, que comprende además un medio (por ejemplo, un medio para realizar FACS) para clasificar una población de células para producir la población clasificada de células individuales.

39. El aparato del concepto 37 o 38, que comprende además un medio para controlar el funcionamiento del aparato para la ejecución automática del método de uno cualquiera de los conceptos 1 a 36.

40

40. Un kit para realizar el método de uno cualquiera de los conceptos 1 a 36, comprendiendo dicho kit un aparato según el concepto 37, 38 o 39, junto con un ácido nucleico que comprende uno o más elementos de transposón para realizar el método del concepto 32.

Los elementos de transposón pueden ser transportados, por ejemplo, por un ADN lineal. En un ejemplo, los elementos son elementos de un transposón que media en la integración del ADN mediante un mecanismo de transposición de corte y empalme (por ejemplo, transposón de clase II). En un ejemplo, los elementos son elementos similares a PB o Mariner o elementos similares a Tc-1 (TLE).

45

41. Un módulo de expresión para la expresión de una POI en una célula hospedante, y el módulo es proporcionado por un ácido nucleico lineal (por ejemplo, ADN lineal) que comprende un transposón, y el transposón comprende elementos de transposón 5'- y 3'-terminales con una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para la expresión de la POI entre los elementos de transposón.

50

Estos módulos son útiles para integrar genómicamente las secuencias de POI expresables en las células hospedantes, por ejemplo, para producir una línea celular para proporcionar una fuente de POI y/o para su uso en el método de selección de la invención. Los elementos de transposón pueden ser cualquiera de estos elementos descritos en la presente.

- 5 En un ejemplo, el módulo comprende o consiste en un transposón que comprende elementos de transposón 5'- y 3'-terminales (por ejemplo, elementos de repetición terminal invertidos de piggyBac) con secuencias de nucleótidos que codifican POI y uno o más elementos reguladores para la expresión entre los elementos de transposón. En una realización, el módulo comprende otra secuencia 5' con respecto al elemento de transposón 5' y/o 3' con respecto al elemento de transposón 3'. En un ejemplo, el módulo comprende una secuencia de nucleótidos adicional que se
10 corresponde con una secuencia de nucleótidos del genoma de la célula hospedante, y dicha secuencia se encuentra 5' y/o 3' con respecto a la secuencia de nucleótidos que codifica una POI. Por ejemplo, la secuencia adicional se corresponde con una secuencia genómica de la célula hospedante que se transcribe activamente en el hospedante. Por tanto, la secuencia que codifica una POI se inserta en el hospedante en un entorno adecuado para la transcripción activa de la secuencia de POI.
- 15 En un ejemplo, una "población" (por ejemplo, una población de células o módulos) o un "repertorio", tal como se emplean en la presente, comprende al menos 10, 100, 1000, 10^4 , 10^5 o 10^6 miembros.

42. Una población de módulos de expresión según el concepto 41, en la que la población codifica un repertorio de POI.

43. Una población clasificada de módulos de expresión según el concepto 42.

- 20 44. Una población clasificada de módulos de expresión que codifican un repertorio de POI que se corresponden con POI expresadas por una población de células, y cada módulo comprende una secuencia de nucleótidos que codifica un miembro del repertorio de POI y uno o más elementos reguladores para la expresión de POI (cuando están en una célula hospedante), en la que cada uno de dichos módulos comprende la disposición (en la dirección 5' a 3'): elemento de transposón - [secuencia de nucleótidos de POI y uno o más elementos reguladores] - elemento de
25 transposón, y los módulos de expresión para la expresión de POI correspondientes a las POI de diferentes células están aislados entre sí en la población clasificada (por ejemplo, en diferentes pocillos de una placa, por ejemplo, una especie de módulo por pocillo en una o más placas).

En un ejemplo, cada módulo es capaz de expresar una POI de una célula individual (o derivada de esta), por ejemplo, una cadena pesada o ligera de anticuerpo, o uno de sus fragmentos, derivada de una única célula B.

- 30 En un ejemplo, se usan elementos de piggyBac.

45. La población del concepto 44, en la que cada módulo de expresión es proporcionado por un ADN lineal.

46. La población de módulos de uno cualquiera de los conceptos 42 a 45, en la que cada módulo se encuentra en una célula hospedante.

- 35 47. Una población clasificada de células hospedantes que comprende la población clasificada de módulos de expresión según uno cualquiera de los conceptos 43, 44 y 45 para la expresión de un repertorio clasificado de POI.

48. Un método para preparar un transposón que comprende una secuencia de nucleótidos de interés (NOI), comprendiendo dicho método:

- 40 a. proporcionar una primera secuencia de nucleótidos (por ejemplo, proporcionada por ADN o ARN) que comprende (en la dirección 5' a 3') A, B y C (y, opcionalmente, que consiste en la estructura 5'-A-B-C-3'), en el que A es una primera secuencia de homología, B es una secuencia de nucleótidos que comprende (o que consiste en) la NOI, y C es una segunda secuencia de homología;

- 45 b. proporcionar una primera secuencia de nucleótidos de molde que comprende (o que consiste en) (en la dirección 5' a 3') W y X, en el que W es una secuencia de nucleótidos que comprende (o que consiste en) un primer elemento de transposón (por ejemplo, un elemento de repetición terminal de piggyBac), y X es una tercera secuencia de homología; y

c. proporcionar una segunda secuencia de nucleótidos de molde que comprende (o que consiste en) (en la dirección 5' a 3') Y e Z, en el que Y es una cuarta secuencia de homología, y Z es una secuencia de nucleótidos que comprende (o que consiste en) un segundo elemento de transposón (por ejemplo, un elemento de repetición terminal de piggyBac); y bien

- 50 d. (i) mezclar la primera secuencia de nucleótidos con el primer molde para hibridar entre sí el primer y el tercer brazo de homología y realizar la amplificación y extensión del ácido nucleico (por ejemplo, usando una PCR) para extender la primera secuencia de nucleótidos usando el primer molde para producir una primera secuencia de nucleótidos extendida (primera ENS) que comprende (en la dirección 5' a 3') W, B y C; y (ii) mezclar la primera ENS

con el segundo molde para hibridar entre sí el segundo y el cuarto brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la primera ENS para producir una segunda ENS que comprende (o que consiste en) (en la dirección 5' a 3') W, B y Z; o bien

5 (ii) mezclar la primera secuencia de nucleótidos con el segundo molde para hibridar entre sí el segundo y el cuarto brazo de homología y realizar la amplificación y extensión del ácido nucleico (por ejemplo, usando una PCR) para extender la primera secuencia de nucleótidos usando el segundo molde para producir una tercera secuencia de nucleótidos extendida (tercera ENS) que comprende (en la dirección 5' a 3') A, B y Z; y (ii) mezclar la tercera ENS con el primer molde para hibridar entre sí el primer y el tercer brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la tercera ENS para producir una cuarta ENS que comprende (o que consiste en) (en la dirección 5' a 3') W, B y Z; o bien

10 (iii) mezclar la primera secuencia de nucleótidos con el primer y el segundo molde para hibridar entre sí el primer y el tercer brazo de homología y para hibridar entre sí el segundo y el cuarto brazo de homología y realizar la amplificación y extensión del ácido nucleico (por ejemplo, usando una PCR) para extender la primera secuencia de nucleótidos usando el segundo molde para producir una quinta ENS que comprende (o que consiste en) (en la dirección 5' a 3') W, B y Z; y

e. aislar una ENS que comprende (o que consiste en) (en la dirección 5' a 3') W, B y Z, produciendo con ello un transposón aislado que comprende una NOI flanqueada por elementos de transposón; y

f. opcionalmente, introducir el transposón aislado en una célula receptora, de modo que el transposón se integra en el genoma de la célula.

20 Opcionalmente, una, más o todas de las secuencias de homología comprenden una secuencia de nucleótidos de al menos 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más nucleótidos contiguos.

En un ejemplo, la NOI codifica una POI, un dominio de proteína o un fragmento de proteína o, en sí misma, es uno o más elementos reguladores. Por ejemplo, la NOI codifica una POI que es un ortólogo o un homólogo de una proteína en el genoma de la célula receptora o en un vertebrado humano o no humano.

25 En una realización, W y X están en los extremos 5' y 3' terminales de la secuencia del primer molde, respectivamente. Además, o como alternativa, Y e Z están en los extremos 5' y 3' terminales de la secuencia del segundo molde, respectivamente. Cuando W y X están en los extremos 5' y 3' terminales de la secuencia del primer molde, respectivamente, y Y e Z están en los extremos 5' y 3' terminales de la secuencia del segundo molde, respectivamente, el producto del método es un transposón lineal con elementos de transposón en sus extremos terminales que es muy adecuado para la integración genómica para modificar células hospedantes.

30 Opcionalmente, el primer molde consiste en 5'-W-X-3'. En un ejemplo, no existe una secuencia de nucleótidos intermedia entre W y X. En otra realización, existe otra secuencia de nucleótidos entre W y X, por ejemplo, un elemento regulador o exón u otra secuencia de nucleótidos deseada (por ejemplo, una secuencia que codifica una proteína) que se combinará inmediatamente cadena arriba de la NOI en el producto del método. Esto resulta útil, por ejemplo, para construir un módulo de expresión para combinar un promotor cadena arriba de una NOI (cuando la NOI codifica una POI) para la posterior expresión de la POI tras haberse insertado el transposón en el genoma de una célula hospedante.

35 Además, o como alternativa, opcionalmente, la secuencia de nucleótidos del segundo molde consiste en 5'-X-Y-3', o existe una secuencia de nucleótidos intermedia entre X e Y, por ejemplo, un elemento regulador o exón u otra secuencia de nucleótidos deseada (por ejemplo, una secuencia que codifica una proteína) que se combinará inmediatamente cadena abajo de la NOI en el producto del método. Esto resulta útil, por ejemplo, para construir un módulo de expresión para combinar una poliA cadena abajo de una NOI (cuando la NOI codifica una POI) para la posterior expresión de la POI tras haberse insertado el transposón en el genoma de una célula hospedante. En otro ejemplo, la secuencia intermedia codifica una proteína que se condensará con la POI tras la expresión para producir un producto de fusión. Por ejemplo, la POI comprende o consiste en un dominio variable de anticuerpo y la secuencia intermedia comprende o consiste en una secuencia que codifica una región constante de anticuerpo. Por ejemplo, la secuencia intermedia codifica un Fc de anticuerpo o un dominio CH1 o CL de anticuerpo. En un ejemplo, el Fc o región constante o proteína codificada por la secuencia intermedia es un Fc, una región constante o una proteína humanas. Esto resulta útil para humanizar la POI (por ejemplo, para producir una cadena de anticuerpo humanizada cuando la POI es un dominio variable, por ejemplo, un dominio variable humano).

49. El método del concepto 48, en el que existe una secuencia de nucleótidos intermedia entre W y X y/o una secuencia de nucleótidos intermedia entre Y e Z; opcionalmente, en el que la secuencia intermedia o cada una de las secuencias intermedias es un elemento regulador o una secuencia que codifica una proteína.

50. El método del concepto 49, en el que la NOI codifica un dominio de proteína (por ejemplo, un dominio variable de anticuerpo) y existe una secuencia de nucleótidos que codifica una región constante de anticuerpo (por ejemplo, un Fc de anticuerpo, por ejemplo, un Fc humano) entre Y e Z, por lo cual el producto de transposón codifica una proteína de fusión que comprende un dominio de proteína condensado con una región constante de anticuerpo (por

ejemplo, codifica una cadena de anticuerpo).

51. El método de uno cualquiera de los conceptos 48 a 50, en el que uno o más del primer y segundo brazo de homología se combina con la NOI mediante PCR (por ejemplo, 5'- y/o 3'-RACE) para formar dicha primera secuencia de nucleótidos antes de realizar dicha extensión.

5 52. Un método para preparar un repertorio de transposones, en el que los miembros del repertorio codifican diferentes POI (por ejemplo, diferentes dominios variables de anticuerpo), comprendiendo dicho método:

i. proporcionar una población de primeras secuencias de nucleótidos que comprende un repertorio de NOI; y

ii. para cada primera secuencia de nucleótidos, realizar el método de uno cualquiera de los conceptos 48 a 51, produciendo con ello un repertorio de transposones que codifican un repertorio de POI.

10 53. El método del concepto 52, que comprende clasificar las primeras secuencias de nucleótidos para proporcionar una población clasificada antes de realizar la etapa (ii), en el que se produce un repertorio clasificado de transposones que codifica un repertorio clasificado de POI.

54. El método del concepto 52, en el que los transposones de dicho repertorio de transposones se mezclan entre sí.

15 55. El método de uno cualquiera de los conceptos 52 a 54, en el que los transposones del repertorio se introducen en células receptoras, de modo que los transposones se integran en el genoma de las células, y cada transposón integrado comprende un módulo de expresión de POI flanqueado por elementos de transposón y el módulo comprende una NOI y uno o más elementos reguladores para la expresión de la POI en una célula hospedante.

20 56. El método del concepto 55 cuando depende del concepto 53, en el que la clasificación se mantiene cuando los transposones se introducen en las células, produciendo con ello un repertorio clasificado de células que expresan un repertorio clasificado de POI.

57. Un método para producir una célula hospedante para la expresión de una POI, comprendiendo dicho método:

a. proporcionar al menos un primer y un segundo módulo de expresión, en el que cada módulo de expresión comprende:

25 i. un primer elemento de integración y un segundo elemento de integración 3' de la secuencia de nucleótidos del primer elemento de integración; y

ii. entre los elementos de integración, una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI;

30 iii. en el que los elementos de integración son capaces de insertarse en un ácido nucleico mediante el reconocimiento de un motivo de secuencia de nucleótidos predeterminado del ácido nucleico usando una enzima de integración;

b. proporcionar una célula hospedante cuyo genoma comprende una pluralidad de dichos motivos; y

c. de modo simultáneo o secuencial, introducir el primer y el segundo módulo de expresión en la célula hospedante, en el que cada módulo se integra genómicamente en el genoma de la célula hospedante en dicho motivo para la expresión de las POI por la célula hospedante; y

35 d. opcionalmente, producir una expresión de las POI en la línea celular en una etapa que comprende cultivar la célula hospedante.

Este aspecto de la invención es útil para producir células hospedantes y líneas celulares para una expresión relativamente alta de una o más POI de interés. La integración genómica de los módulos de POI en múltiples sitios genómicos proporciona una expresión estable y también existe la posibilidad de dirigirse a regiones transcripcionalmente activas del genoma del hospedante. El uso de motivos de secuencia guía la inserción hacia sitios útiles, y esto es preferible a la integración aleatoria de secuencias, tal como se emplea en la técnica.

40 58. El método del concepto 57, en el que el primer y el segundo elemento de integración del primer módulo son idénticos al primer y segundo elemento de integración, respectivamente, del segundo módulo.

En un ejemplo, cada módulo comprende un primer y un segundo elemento de transposón, por ejemplo, elementos del mismo tipo de transposón (por ejemplo, transposón PB o de clase II). En un ejemplo, todos los elementos son sitios de recombinación específica de sitio, por ejemplo, sitios lox o sitios frt o una mezcla de estos. En otro ejemplo, todos los elementos son brazos de homología (secuencias de nucleótidos contiguas suficientes para la recombinación homóloga en la célula hospedante). En un ejemplo, los sitios de recombinación específica de sitio son iguales o son diferentes (por ejemplo, sitios mutuamente incompatibles (por ejemplo, loxP y lox511 o 2272) para realizar el RMCE (intercambio de módulos mediados por recombinasa, "recombinase-mediated cassette exchange")

50

para la inserción dirigida del módulo en el genoma.

59. El método del concepto 57 o 58, en el que el primer y el segundo elemento de integración de cada uno de dichos primer y segundo módulo están en una orientación mutuamente invertida (por ejemplo, elementos de transposón PB invertidos o sitios de recombinación específica de sitio invertidos).

5 60. El método de uno cualquiera de los conceptos 57 a 59, en el que uno o más de dichos motivos se introduce mediante ingeniería en el cromosoma de la célula hospedante antes de realizar la etapa (c), por ejemplo, se introducen mediante ingeniería parejas de sitios lox en uno o más cromosomas del hospedante, en el que las parejas se corresponden con parejas de lox en los módulos usados).

10 61. El método de uno cualquiera de los conceptos 57 a 59, en el que uno o más de dichos motivos es endógeno del genoma de la célula hospedante; opcionalmente, en el que cada uno de dichos motivos en donde se integra un módulo es un motivo endógeno.

Por ejemplo, los transposones reconocen motivos endógenos (por ejemplo, PB reconoce TTAA en los genomas).

15 62. El método de uno cualquiera de los conceptos 57 a 61, en el que al menos 3 módulos se integran genómicamente en el genoma de la célula hospedante, por ejemplo, en uno o más cromosomas del hospedante, lo cual es útil para la expresión estable.

63. El método de uno cualquiera de los conceptos 57 a 62, en el que los sitios de integración genómica del módulo son activos para la transcripción de las secuencias que codifican POI.

Esto puede lograrse usando transposones (por ejemplo, PB) en los módulos.

20 64. El método de uno cualquiera de los conceptos 57 a 63, en el que cada módulo se integra mediante recombinación homóloga entre los sitios de integración y el genoma del hospedante; mediante recombinación específica de sitio entre los sitios de integración y el genoma del hospedante; o mediante integración mediada por transposones.

25 65. El método del concepto 64, en el que la enzima se selecciona de una recombinasa o una transposasa (por ejemplo, una enzima que se corresponde con los elementos de integración de PBasa (por ejemplo, PBasa hiperactiva), flp o cre recombinasa).

En un ejemplo, la célula hospedante se ha modificado para que exprese dicha enzima o enzimas, por ejemplo, a partir de un gen genómicamente integrado (por ejemplo, un gen inducible). En otra realización, la enzima se expresa a partir de un vector episómico. En otro ejemplo, la enzima se introduce en la célula hospedante.

66. El método de uno cualquiera de los conceptos 57 a 64, en el que cada módulo es un transposón.

30 67. El método de uno cualquiera de los conceptos 57 a 66, que comprende proporcionar una población de células hospedantes y realizar el método de uno cualquiera de los conceptos 57 a 66 en una pluralidad de células hospedantes de dicha población; y, opcionalmente, aislar las células hospedantes producidas mediante la etapa (c) o (d).

35 68. El método de uno cualquiera de los conceptos 57 a 67, que comprende aislar una célula hospedante producida mediante la etapa (c) o (d) e identificar, amplificar o sintetizar la secuencia de nucleótidos que codifica la POI expresable por la célula; y, opcionalmente, producir una POI aislada usando dicha secuencia de nucleótidos identificada, amplificada o sintetizada, o uno de sus mutantes.

69. El método del concepto 68, que comprende formular la POI aislada en un fármaco para la medicina humana; y, opcionalmente, administrar el fármaco a un paciente humano.

40 70. Una población de células hospedantes que puede obtenerse mediante el método de uno cualquiera de los conceptos 57 a 69, y cada célula comprende una pluralidad de módulos de expresión integrados genómicamente para expresar POI, y cada célula hospedante comprende una pluralidad de motivos de secuencia de nucleótidos idénticos a través de su genoma adyacentes a un módulo de expresión integrado para la expresión de una POI a partir de cada módulo; comprendiendo cada módulo integrado:

45 a. una secuencia de un primer elemento de integración y una secuencia de un segundo elemento de integración 3' de la secuencia de nucleótidos del primer elemento de integración; y

b. entre las secuencias de los elementos de integración, una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI.

50 71. Las células hospedantes del concepto 70, en las que todas las POI expresadas por las células son la misma POI.

72. Las células hospedantes del concepto 70, en las que las células expresan una primera y una segunda POI (por ejemplo, dominios VH y VL de un único tipo de anticuerpo; o cadenas pesadas y ligeras de un único tipo de anticuerpo) que se asocian entre sí para formar una proteína funcional o un sitio de unión al ligando (por ejemplo, antígeno).

5 73. Un anticuerpo o un sitio de unión de anticuerpo de un anticuerpo para el tratamiento médico de un ser humano, en el que el anticuerpo o el sitio de unión se ha aislado a partir de una célula hospedante producida mediante el método de uno cualquiera de los conceptos 57 a 69, o se ha aislado a partir de una célula hospedante de una población según uno cualquiera de los conceptos 70 a 72.

10 74. Una mezcla de ácidos nucleicos que comprende un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en la que el primer ácido nucleico es capaz de hibridarse con una secuencia de 5'UTR de una región V de un anticuerpo humano (es decir, una secuencia de nucleótidos) de un gen comprendido en un ácido nucleico diana, en la que el gen codifica una región V humana; y el segundo ácido nucleico es capaz de hibridarse con una segunda secuencia, en la que la segunda secuencia está comprendida en el ácido nucleico diana y se encuentra 3' con respecto a la secuencia de UTR, en la que el primer ácido nucleico aislado comprende (o consiste en) una
15 secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica (o 100% idéntica) a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-47.

20 SEQ ID NO:1-47 comprenden secuencias específicas de región variable humanas, tal como se indica en la tabla 1 (más en concreto, secuencias de nucleótidos de 5'UTR de regiones variables humanas). "Específica de" significa que dichas secuencias pueden usarse como una secuencia de cebador 5' en una PCR convencional (por ejemplo, RT-PCR) de un ácido nucleico de región variable humana.

SEQ ID NO:1-17 comprenden secuencias específicas de región variable de cadena pesada humanas.

SEQ ID NO:18-26 comprenden secuencias específicas de región variable de cadena kappa humanas.

SEQ ID NO:27-47 comprenden secuencias específicas de región variable de cadena kappa humanas.

25 En una realización, la descripción proporciona un ácido nucleico (por ejemplo, un cebador de PCR o un vector para la recombinación homóloga) que comprende (o que consiste en) al menos 15 nucleótidos contiguos de una secuencia denominada X en la tabla 2, para hibridarse con la secuencia de 5'UTR de un segmento de gen denominado Y en la tabla 2, por ejemplo, para realizar una PCR para copiar el segmento de gen o para hibridar un vector de recombinación homóloga con la secuencia de 5'UTR para la modificación del segmento de gen. En un ejemplo, el ácido nucleico comprende (o consiste en) al menos 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 o toda la
30 secuencia X. En un ejemplo, los nucleótidos contiguos terminan con el nucleótido 3' de la secuencia X (es decir, se usan los nucleótidos contiguos se extienden 5' desde el extremo 3' de X). En una realización, la invención y descripción proporciona una mezcla de dos o más de los ácidos nucleicos, por ejemplo, para copiar por PCR dos o más secuencias de región variable (por ejemplo, usando ADN, ADNc o ARN procedente de las correspondientes células B). En un ejemplo, dos, más o todos los ácidos nucleicos en la mezcla copian segmentos del gen de V_H. En un ejemplo, dos, más o todos los ácidos nucleicos en la mezcla copia segmentos del gen de V_λ. Opcionalmente, el ácido nucleico o cada ácido nucleico comprende una secuencia de nucleótidos de promotor inmediatamente 5' de la secuencia de UTR (o la parte de 15 o más nucleótidos contiguos). Por ejemplo, la secuencia de promotor es una secuencia de promotor de CMV como sigue:

40 5'-CTTACTGGCTTATCGAAATTAATACGACTCAGATC-3' (SEQ ID NO:54)

En un ejemplo, la invención proporciona:

45 Un cebador de PCR o un vector de recombinación homóloga que comprende al menos 15 nucleótidos contiguos de una secuencia de UTR de un segmento de gen variable de anticuerpo humano para hibridarse con la secuencia de 5'UTR de un segmento de gen variable de anticuerpo denominado Y en la tabla 2, en el que la secuencia del cebador/vector se selecciona del grupo que consiste en las secuencias denominadas X en la tabla 2.

En un ejemplo de la mezcla de la invención, o del ácido nucleico, mezcla o cebador de la descripción, cada ácido nucleico o cebador se hibrida con su secuencia cognada a una temperatura de 45-70 °C (por ejemplo, a 50 °C, o 60 °C, o 68 °C) o de 60-75 °C, en una reacción de PCR. Los expertos en la técnica sabrán los tiempos y las temperaturas de los ciclos para realizar la reacción de PCR.

50 Cada ácido nucleico de la descripción y mezcla de la invención o descripción es útil para realizar la amplificación por PCR o la replicación de una secuencia de nucleótidos diana que codifica un dominio variable de anticuerpo humano o una proteína que comprende dicho dominio, por ejemplo, una PCR de una o más secuencias de nucleótidos que codifican una región variable humana aisladas a partir de una o más células (por ejemplo, células B). Por tanto, en una realización, cada ácido nucleico es un cebador de PCR.

55 Cada ácido nucleico de la descripción y mezcla de la invención o descripción es útil para realizar la recombinación

homóloga para modificar una secuencia de nucleótidos diana (por ejemplo, una secuencia formada por el genoma de una célula, por ejemplo, una célula de mamífero, por ejemplo, una célula ES o una célula CHO). Para la recombinación homóloga, tal como conocen los expertos en la técnica, se emplea un vector de ácido nucleico que comprende un brazo de homología 5', un brazo de homología 3' y, opcionalmente, una secuencia de nucleótidos de interés predeterminada entre ambos. La secuencia puede codificar, por ejemplo, una POI, un dominio de proteína o estar formada por un elemento regulador. Como alternativa, no existe una secuencia intermedia entre los brazos de homología (y, en este caso, el vector se emplea para delecionar una secuencia del genoma que se encuentra entre las regiones que se hibridan con los brazos de homología, tal como conocen los expertos en la técnica). En la presente realización, la descripción proporciona un vector recombinante homólogo, en el que el vector comprende un brazo 5' que comprende un brazo de homología 3' y, opcionalmente, una secuencia de nucleótidos entre ambos, en el que el brazo 5' comprende (o consiste en) una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-47 (o es 100% idéntica) y/o en el que el brazo 3' comprende (o consiste en) una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica (o 100% idéntica) a una secuencia seleccionada del grupo que consiste en SEQ ID NO:48 a 53. Esto permite la localización génica de segmentos de genes de V y/o C específicos de loci de Ig en un vertebrado usando la recombinación homóloga.

Por tanto, la descripción también proporciona un método para modificar un locus de Ig incluido en una célula de vertebrado, comprendiendo dicho método introducir el vector de la invención en la célula (por ejemplo, mediante transfección) y realizar una recombinación homóloga para modificar el locus de Ig; y, opcionalmente, expresar una cadena o dominio V de anticuerpo a partir del locus modificado. Opcionalmente, la secuencia del dominio V se identifica o se copia o se aísla de la célula y se usa para producir un anticuerpo o una composición farmacéutica que comprende dicho anticuerpo para un uso médico humano.

La descripción proporciona además un cebador de PCR que comprende (o que consiste en) una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-53 (o es 100% idéntica). Por ejemplo, el cebador está *in vitro*.

El término "aislado" excluye las secuencias que están presentes en el contenido cromosómico de un vertebrado o una célula de vertebrado.

El ácido nucleico, el cebador de PCR o la mezcla pueden proporcionarse *in vitro*, por ejemplo, mezclados con un reactivo o tampón de PCR. En un ejemplo, un ácido nucleico, un cebador o una mezcla de la descripción o una mezcla de la invención o descripción se proporciona en un recipiente, un vial, un tubo, una placa o una cubeta de PCR.

El primer ácido nucleico aislado comprende (o consiste en) una secuencia que es al menos 95, 96, 97, 98 o 99% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-47. Opcionalmente, el primer ácido nucleico aislado comprende (o consiste en) una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-47.

75. La mezcla del concepto 74, en la que el primer ácido nucleico aislado comprende una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-17.

Opcionalmente, el primer ácido nucleico aislado comprende una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-17.

76. La mezcla del concepto 74, en la que el primer ácido nucleico aislado comprende una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:18-26.

Opcionalmente, el primer ácido nucleico aislado comprende una secuencia seleccionada del grupo que consiste en SEQ ID NO:18-26.

77. La mezcla del concepto 74, en la que el primer ácido nucleico aislado comprende una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:27-47.

Opcionalmente, el primer ácido nucleico aislado comprende una secuencia seleccionada del grupo que consiste en SEQ ID NO:27-47.

78. La mezcla de uno cualquiera de los conceptos 74 a 77, en la que el segundo ácido nucleico aislado comprende una secuencia de región constante de anticuerpo, opcionalmente, una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:48 a 53.

Opcionalmente, el segundo ácido nucleico aislado comprende una secuencia seleccionada del grupo que consiste en SEQ ID NO:48 a 53. SEQ ID NO:48 -51 son secuencias procedentes de regiones constantes de ratón; SEQ ID NO:52 y 53 son secuencias procedentes de regiones constantes humanas (véase la tabla 1).

79. La mezcla del concepto 75, en la que el segundo ácido nucleico aislado comprende una secuencia de región constante de cadena pesada de anticuerpo; y, opcionalmente, comprende una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica (o 100% idéntica) a SEQ ID NO:48 o 49.
- 5 80. La mezcla del concepto 76, en la que el segundo ácido nucleico aislado comprende una secuencia de región constante de cadena kappa de anticuerpo; y, opcionalmente, comprende una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica (o 100% idéntica) a SEQ ID NO:50 o 51.
81. La mezcla del concepto 77, en la que el segundo ácido nucleico aislado comprende una secuencia de región constante de cadena lambda de anticuerpo; y, opcionalmente, comprende una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica (o 100% idéntica) a SEQ ID NO:52 o 53.
- 10 82. Una mezcla de ácidos nucleicos que comprende un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en la que los ácidos nucleicos son diferentes y se seleccionan de ácidos nucleicos que comprenden una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica (o 100% idéntica) a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-47.
- 15 En un ejemplo, los ácidos nucleicos son cebadores de PCR; en otra realización, comprenden vectores de recombinación homóloga para modificar un locus o loci de Ig.
83. La mezcla del concepto 82, que comprende una secuencia seleccionada del grupo que consiste en SEQ ID NO:18-26 y una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica (o 100% idéntica) a una secuencia seleccionada del grupo que consiste en SEQ ID NO:18-26 y/o seleccionada del grupo que consiste en SEQ ID NO:27-47.
- 20 84. La mezcla del concepto 82 o 83, en la que cada uno del primer y el segundo ácido nucleico aislado se selecciona de modo que sea al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntico (o 100% idéntico) a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-17.
- 25 85. La mezcla del concepto 84, que comprende al menos 3 ácidos nucleicos aislados diferentes, y cada uno es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntico (o 100% idéntico) a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-17.
86. La mezcla del concepto 82 o 83, en la que cada uno del primer y el segundo ácido nucleico aislado es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntico (o 100% idéntico) a una secuencia seleccionada del grupo que consiste en SEQ ID NO:18-26.
- 30 87. La mezcla del concepto 84, que comprende al menos 3 ácidos nucleicos aislados diferentes, y cada uno es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntico (o 100% idéntico) a una secuencia seleccionada del grupo que consiste en SEQ ID NO:18-26.
88. La mezcla del concepto 82 o 83, en la que cada uno del primer y el segundo ácido nucleico aislado es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntico (o 100% idéntico) a una secuencia seleccionada del grupo que consiste en SEQ ID NO:27-47.
- 35 89. La mezcla del concepto 84, que comprende al menos 3 ácidos nucleicos aislados diferentes, y cada uno es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntico (o 100% idéntico) a una secuencia seleccionada del grupo que consiste en SEQ ID NO:27-47.
90. La mezcla del concepto 84 o 85, en la que la mezcla comprende una secuencia de región constante de cadena pesada de anticuerpo; y, opcionalmente, comprende SEQ ID NO:48 y/o 49.
- 40 91. La mezcla del concepto 86 o 87, en la que la mezcla comprende una secuencia de región constante kappa de anticuerpo; y, opcionalmente, comprende SEQ ID NO:50 y/o 51.
92. La mezcla del concepto 87 o 88, en la que la mezcla comprende una secuencia de región constante lambda de anticuerpo; y, opcionalmente, comprende SEQ ID NO:52 y/o 53.
- 45 93. El método de uno cualquiera de los conceptos 1 a 36, en el que la etapa (c) se realiza mediante PCR usando una o más mezclas según uno cualquiera de los conceptos 74 a 92.
94. El método de uno cualquiera de los conceptos 48 a 56, en el que el método se realiza mediante PCR usando una o más mezclas según uno cualquiera de los conceptos 74 a 92.
95. Un kit que comprende una o más mezclas según uno cualquiera de los conceptos 74 a 92, y un aparato según uno cualquiera de los conceptos 36 a 39.
- 50 96. Un método para amplificar un repertorio de secuencias de región variable humanas, comprendiendo dicho método:

a. proporcionar una población de células que expresa un repertorio de regiones variables humanas, en el que las células comprenden secuencias de nucleótidos que codifican las regiones variables;

b. replicar una pluralidad de dichas secuencias de nucleótidos que codifican regiones variables usando PCR y moldes de PCR; y

5 c. aislar, secuenciar o identificar una o más de las secuencias de nucleótidos replicadas, o realizar las etapas (d) y (e) del método de uno cualquiera de los conceptos 1 a 36;

en el que uno o más moldes de la etapa (b) comprende una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idéntica (o 100% idéntica) a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-52.

10 97. El método del concepto 96, en el que la etapa (b) usa una o más mezclas según uno cualquiera de los conceptos 74 a 92 como molde de PCR.

98. El método del concepto 96 o 97, en el que las células en la etapa (a) son células individuales clasificadas (por ejemplo, clasificadas en pocillos de una o más placas).

15 99. El método del concepto 96 o 97, que comprende además producir una región variable humana (por ejemplo, como parte de una cadena de anticuerpo aislada o un anticuerpo aislado para la medicina humana) usando una secuencia replicada obtenida en la etapa (c) y, opcionalmente, producir una línea celular que expresa la región variable humana.

Las siguientes características opcionales son aplicables a cualquier configuración, aspecto, realización o ejemplo de la invención o descripción descritos en la presente.

20 Opcionalmente, la secuencia de nucleótidos que codifica una POI está unida operablemente a un promotor capaz de dirigir la expresión de la POI, en la que el promotor comprende un promotor eucariota que puede ser regulado por un activador o un inhibidor. En otra realización, el promotor eucariota está unido operablemente a un operador procariota, y la célula eucariota opcionalmente comprende además una proteína represora procariota.

25 Opcionalmente, cada módulo de expresión comprende una secuencia que codifica un marcador, tal como un marcador seleccionable, por ejemplo, un gen de resistencia a la higromicina o que codifica una proteína fluorescente verde (por ejemplo, la proteína fluorescente se selecciona de DsRed, GFP, eGFP, CFP, eCFP, e YFP).

Opcionalmente, uno o más o todos los módulos de expresión comprenden una primera y una segunda secuencia de nucleótidos que codifica una POI, por ejemplo, en tándem o como un módulo bicistrónico. En un ejemplo, las POI codificadas son diferentes (por ejemplo, VH y VL de un anticuerpo); en otro ejemplo, son idénticas. En un ejemplo, existen 1, 2, 3, 4, 5, 6 o más secuencias de nucleótidos que codifican una POI.

30 En un ejemplo, la célula hospedante o cada célula hospedante es una célula CHO (ovario de hámster chino) o una célula HEK293.

35 Por ejemplo, la proteína de interés puede ser un anticuerpo o sus fragmentos, un anticuerpo quimérico o sus fragmentos, un ScFv o sus fragmentos, una proteína marcada con Fc o sus fragmentos, un factor de crecimiento o sus fragmentos, una citoquina o sus fragmentos, o un dominio extracelular de un receptor de la superficie celular o sus fragmentos.

Construcciones de ácidos nucleicos

40 Los módulos de expresión recombinante (vectores) pueden comprender fragmentos de ADN sintéticos o derivados de ADNc que codifican una proteína de interés, unidos operablemente a un elemento regulador transcripcional y/o traduccional adecuado derivado de genes de mamífero, víricos o de insecto. Estos elementos reguladores incluyen promotores transcripcionales, potenciadores, secuencias que codifican sitios de unión ribosómicos del ARNm adecuados, y secuencias que controlan el fin de la transcripción y la traducción. Los módulos de expresión de mamífero también pueden comprender elementos no transcritos, tales como un origen de la replicación, otras secuencias no transcritas en los flancos 5' o 3', y secuencias no traducidas 5' o 3', tales como sitios donantes y aceptores de corte y empalme. También puede incorporarse un gen de marcador seleccionable para facilitar el reconocimiento de los transfectantes.

45 Las secuencias de control transcripcional y traduccional en los módulos de expresión útiles para transfectar células de vertebrado pueden ser proporcionadas por fuentes víricas. Por ejemplo, los promotores y potenciadores usados habitualmente se derivan de virus, tales como polioma, adenovirus 2, virus de simio 40 (SV40), y citomegalovirus humano (CMV). Los promotores genómicos, las secuencias de control y/o señal víricas pueden utilizarse para dirigir la expresión, con la condición de que dichas secuencias control sean compatibles con la célula hospedante elegida. También pueden usarse promotores celulares no víricos (por ejemplo, los promotores de beta-globina y de EF-1 alfa), dependiendo del tipo de célula en que se va a expresar la proteína recombinante.

Pueden emplearse secuencias de ADN derivadas del genoma vírico de SV40, por ejemplo, el origen, promotores

tempranos y tardíos, potenciadores, sitios de corte y empalme y de poliadenilación de SV40, para proporcionar otros elementos genéticos útiles para la expresión de la secuencia de ADN heteróloga. Los promotores tempranos y tardíos son particularmente útiles, porque ambos se obtienen con facilidad del virus SV40 como un fragmento que también comprende el origen de la replicación vírico del SV40 (Fiers *et al.*, Nature, 1978, 273:113). También pueden utilizarse fragmentos de SV40 más pequeños o más grandes. Generalmente, se incluye la secuencia de aproximadamente 250 pb que se extiende desde el sitio Hind III hacia el sitio BglI localizado en el origen de la replicación de SV40.

Se han descrito previamente vectores de expresión bicistronicos usados para la expresión de múltiples transcritos (Kim S. K. y Wold B. J., Cell, 1985, 42:129) y estos pueden usarse en combinación con una o más secuencias que codifican POI.

Células hospedantes y transfección

Opcionalmente, se usan células hospedantes eucariotas en los métodos de la invención y descripción, por ejemplo, son células hospedantes de mamífero que incluyen, por ejemplo, células CHO o células de ratón.

Las proteínas expresadas (POI) preferiblemente se segregarán hacia el medio de cultivo, dependiendo de la secuencia de ácido nucleico seleccionada, pero pueden ser retenidas dentro de la célula o depositadas en la membrana celular. Pueden emplearse diversos sistemas de cultivo de células de mamífero para expresar proteínas recombinantes. Los ejemplos de líneas de células hospedantes de mamífero adecuadas incluyen las líneas COS-7 de células de riñón de mono descritas por Gluzman (1981), Cell, 23:175, y otras líneas celulares capaces de expresar un vector apropiado que incluyen, por ejemplo, las líneas celulares CV-1/EBNA (ATCC CRL 10478), células L, C127, 3T3, CHO, HeLa y BHK. Otras líneas celulares desarrolladas para esquemas de amplificación o selección específicos también serán útiles con los métodos y las composiciones proporcionados en la presente. Una línea celular preferida es la línea celular CHO denominada K1. Para lograr el objetivo de un alto volumen de producción de proteínas recombinantes, opcionalmente la línea de células hospedantes se preadapta al medio del biorreactor en el caso apropiado.

En la técnica se conocen varios protocolos de transfección, y se analizan en Kaufman (1988), Meth. Enzymology, 185:537. El protocolo de transfección elegido dependerá del tipo de célula hospedante y de la naturaleza de la POI, y puede elegirse basándose en los experimentos habituales. Los requisitos básicos de cualquiera de estos protocolos son, en primer lugar, introducir un ADN que codifica la proteína de interés en una célula hospedante adecuada, y después identificar y aislar las células hospedantes que han incorporado el ADN heterólogo de una manera expresable y relativamente estable.

Otro método para introducir ADN heterólogo en una célula que se emplea habitualmente es la precipitación con fosfato de calcio, como describen, por ejemplo, Wigler *et al.* (Proc. Natl. Acad. Sci. USA, 77:3567, 1980). El ADN introducido en una célula hospedante mediante este método con frecuencia sufre una redistribución, lo cual hace que este procedimiento sea útil para la cotransfección de genes independientes.

La fusión inducida por polietileno de protoplastos bacterianos con células de mamífero (Schaffner *et al.* (1980), Proc. Natl. Acad. Sci. USA, 77:2163) es otro método útil para introducir ADN heterólogo. Los protocolos de fusión de protoplastos producen con frecuencia múltiples copias del ADN plasmídico integrado en el genoma de la célula hospedante de mamífero, y esta técnica requiere que el marcador de selección y amplificación estén en el mismo ácido nucleico que la POI.

También puede emplearse la electroporación para introducir ADN directamente en el citoplasma de una célula hospedante, por ejemplo, tal como describen Potter *et al.* (Proc. Natl. Acad. Sci. USA, 81:7161, 1988) o Shigekawa *et al.* (BioTechniques, 6:742, 1988). A diferencia de la fusión de protoplastos, la electroporación no requiere que el marcador de selección y la POI estén sobre el mismo ácido nucleico.

En fechas más recientes se han descrito varios reactivos útiles para introducir ADN heterólogo en una célula de mamífero. Estos incluyen el reactivo Lipofectin™ y el reactivo Lipofectamine™ (Gibco BRL, Gaithersburg, Md.). Ambos reactivos son reactivos disponibles en el mercado usados para formar complejos de lípido-ácido nucleico (o liposomas) que, cuando se aplican a células cultivadas, facilitan la captación del ácido nucleico hacia el interior de las células.

También resulta deseable un método para amplificar la POI para la expresión de la proteína recombinante, y este generalmente implica el uso de un marcador de selección (analizado en Kaufman, *supra*). La resistencia a fármacos citotóxicos es la característica que se usa con más frecuencia como marcador de selección, y esta puede ser el resultado de un rasgo dominante (por ejemplo, puede usarse independientemente del tipo de célula hospedante) o de un rasgo recesivo (por ejemplo, útil en tipos de células hospedantes concretas que son deficientes en la actividad que se esté seleccionando). Varios marcadores amplificables son adecuados para su uso en los vectores de expresión descritos en la presente (por ejemplo, según se describe en Maniatis, Molecular Biology: A Laboratory Manual, Cold Spring Harbor Laboratory, NY, 1989, pp. 16.9-16.14).

Los marcadores seleccionables útiles para la amplificación de genes en células de mamífero resistentes a fármacos

se muestran en la tabla de 1 de Kaufman, R. J., *supra*, e incluyen la resistencia a DHFR-MTX, P-glicoproteína y la resistencia a múltiples fármacos ("multiple drug resistance", MDR), diversos agentes citotóxicos lipófilos (por ejemplo, adriamicina, colchicina, vincristina), y adenosina desaminasa (ADA)-Xyl-A o adenosina y 2'-desoxicoformicina.

5 Otros marcadores seleccionables dominantes incluyen genes de resistencia a antibióticos que surgen de microbios, por ejemplo, resistencia a la neomicina, kanamicina o higromicina. Sin embargo, no se ha demostrado que estos marcadores de selección sean amplificables (Kaufman, R. J., *supra*). Existen varios sistemas de selección adecuados para hospedantes mamíferos (Maniatis, *supra*, pp. 16.9-16.15). También se han descrito protocolos de cotransfección que emplean dos marcadores seleccionables dominantes (Okayama y Berg, Mol. Cell. Biol., 5:1136, 1985).

También pueden incluirse elementos reguladores útiles, descritos previamente o conocidos en la técnica, en las construcciones de ácido nucleico usadas para transfectar células de mamífero. El protocolo de transfección elegido y los elementos seleccionados para su uso en él dependerán del tipo de célula hospedante usada. Los expertos en la técnica conocen numerosos protocolos y células hospedantes diferentes, y pueden seleccionar un sistema apropiado para la expresión de una proteína deseada, basándose en los requisitos del sistema de cultivo celular empleado.

Un aspecto proporciona una composición farmacéutica que comprende una POI aislada (por ejemplo, un anticuerpo, una cadena o un dominio variable) y un diluyente, un excipiente o un vehículo, opcionalmente en la que la composición está contenida en un recipiente IV (por ejemplo, una bolsa IV) o un recipiente conectado a una jeringa IV, y en la que la POI ha sido aislada a partir de una célula hospedante de la descripción o una población de células hospedantes.

Un aspecto proporciona el uso de la POI, según se describe en la presente, en la fabricación de un medicamento para el tratamiento y/o la profilaxis de una enfermedad o un trastorno en un paciente, por ejemplo, un ser humano.

Se entenderá que las realizaciones concretas descritas en la presente se muestran como ilustración y no como limitación de la invención. El uso de la palabra "un" o "una", cuando se emplea junto con el término "comprende" en las reivindicaciones y/o la memoria descriptiva puede significar "uno", pero también es coherente con el significado de "uno o más", "al menos uno", y "uno o más de uno". El uso del término "o" en las reivindicaciones se emplea para indicar "y/o", a menos que se indique explícitamente que se refiere solo a alternativas o que las alternativas son mutuamente excluyentes, aunque la memoria descriptiva apoya una definición que se refiere solo a alternativas e "y/o". A lo largo de esta solicitud, el término "aproximadamente" se emplea para indicar que un valor incluye la variación de error inherente para el dispositivo, y el método se emplea para determinar el valor, o la variación que existe entre los sujetos de estudio.

Ejemplo 1

La tecnología de clonación de células B de la presente invención incluye tres etapas principales: aislamiento de células B o ASC individuales específicas de antígeno a partir del bazo, nódulos linfáticos y médula ósea con los correspondientes marcadores celulares; rescate de la secuencia de anticuerpo de las células individuales y amplificación del módulo de expresión; expresión de los anticuerpos recombinantes en células de mamífero. El diagrama de flujo se muestra en la figura 1B. Los detalles de cada etapa se describen a continuación.

Ejemplo 1A: Aislamiento de células individuales específicas de antígeno

Las células específicas de antígeno incluyen las células de memoria/GC con anticuerpos unidos a la membrana y las células plasmáticas que segregan anticuerpos. Se usa un panel de marcadores de la superficie celular para definir y marcar células de memoria/GC de ratón (CD19; IgM; IgD; CD38; CD95) (figura 2). Las células específicas de antígeno se tiñeron con antígenos solubles marcados con fluorescencia (por ejemplo, cualquier fluoróforo de molécula pequeña que puede detectarse mediante el sistema de clasificación de células, tal como Alexa-488, Alexa-647, azul pacífico, R-ficoeritrina, isotiocianato de fluoresceína, o alofocianina opcionalmente conjugada con un tinte de cianuro, por ejemplo, Cy7) o antígenos de la superficie celular en partículas similares a virus ("virus-like particles", VLP). La figura 2 es un ejemplo de marcaje de células de memoria/GC específicas de antígeno en bazo de ratón inmunizado con OVA. Pueden clasificarse más de 10.000 células de IgG de memoria/GC específicas de OVA a partir de un bazo. La clasificación de células individuales se realizó usando un citómetro de flujo de inlujo BD equipado con una unidad de depósito de células automática. Las células clasificadas con FACS se depositaron en placas de PCR de 96 pocillos con tampón de lisis para la siguiente etapa.

En general, las células GC (centro germinal, "germinal centre") o células B de memoria específicas de antígeno pueden ser capturadas por un antígeno marcado porque expresan predominantemente anticuerpos transmembrana sobre su superficie celular. Por otra parte, se cree que los plasmablastos o células plasmáticas pueden ser capturadas con menor facilidad por el antígeno marcado, debido a su expresión dominante de anticuerpos secretores. Después se intentaron aislar ASPC usando un antígeno marcado de modo fluorescente y anti-CD138 para clasificar las células procedentes del resto de la población usando FACS (figura 3). Tal como se muestra en la

figura 2, la mayoría de las células de plasmablastos o plasmáticas específicas de antígeno aisladas, pero ninguno de los mismos tipos de células que quedaron, demostraron ser ASC específicas de antígeno en el ensayo ELISPOT. Esto demostró que el método de clasificación de células usando un antígeno marcado de modo fluorescente puede capturar, de modo eficaz, todas las ASC específicas de antígeno probablemente con anticuerpos transmembrana residuales o anticuerpos segregados de anclaje temporal sobre la superficie celular.

El antígeno marcado de modo fluorescente puede ser reemplazado por VLP con un antígeno recombinante sobre su superficie. Las VLP se generan a partir de células CHO, células KEK, MEF (fibroblastos embrionarios de ratón) u otras líneas de células de mamífero con coexpresión del antígeno recombinante, la proteína gag de retrovirus, y MA-GFP (proteína de fusión del fragmento de matriz de gag p15-GFP). La expresión de gag permite que las VLP germinen de las células, y la MA-GFP marca las VLP para la detección de la fluorescencia. Ambas proteínas gag y MA-GFP están asociadas con la superficie interna de la membrana plasmática, y el antígeno recombinante está sobre la superficie de las VLP. Los antígenos sobre las VLP se presentan en la forma nativa, directamente expresados a partir de las células recombinantes sin ninguna etapa de purificación o modificación. La forma nativa de un antígeno debería proporcionar todos los epitopos naturales que ayudan en gran medida a la selección de anticuerpos neutralizantes. La alta densidad del antígeno sobre las VLP aumenta la proporción de señal/ruido para la detección de células que expresan anticuerpos específicos de antígeno sobre la superficie celular, y facilita en gran medida la etapa de clasificación. Las VLP recombinantes pueden generarse con la expresión de diferentes proteínas fluorescentes, tales como MA-CFP o MA-YFP. Usando la multiplexación de VLP con diferentes antígenos y diferentes proteínas de fluorescencia, pueden seleccionarse las células que expresan ligantes de alta afinidad, ligantes de reactividad cruzada o ligantes específicos de homólogo. Las células que expresan ligantes de alta afinidad pueden ser seleccionadas por células con una matriz de afinidad relativamente alta (matriz de afinidad = la proporción de actividad de unión a VLP con baja densidad de antígenos a VLP con alta densidad de antígenos). Las células que expresan ligantes de reactividad cruzada con ortólogos o antígenos diferentes (para el aislamiento de anticuerpos biespecíficos de 2 en 1) pueden ser seleccionadas por células que se unen a diferentes tipos de VLP al mismo tiempo. Las células que expresan ligantes específicos de homólogo también pueden ser seleccionadas por células que se unen solo al antígeno específico, pero no a su homólogo.

Ejemplo 1B: Recuperación de alta capacidad de procesamiento de la secuencia de anticuerpo a partir de células individuales

Se desarrolló un método de alta capacidad de procesamiento rápido y eficaz para la generación de anticuerpos a partir de células B individuales sin ninguna etapa de clonación molecular. El método permite producir construcciones de expresión de Ig a partir de segmentos de genes variables amplificados de cadena pesada y cadena ligera a partir de una célula individual (figura 3). A través del procedimiento completo de la PCR, se amplificaron cadenas pesadas y cadenas ligeras a partir de una célula individual en el mismo pocillo.

Las secuencias de las regiones V de anticuerpo se recuperaron mediante RT-PCR y dos rondas de PCR mediante el siguiente procedimiento. Placas que portan células individuales clasificadas conservadas a -80 °C se descongelaron en hielo y se centrifugaron brevemente antes del uso. Las placas se incubaron en el termociclador a 65 °C durante 5 minutos e indefinidamente a 4 °C. Se añadieron 6 µl de mezcla de cebadores, Superscript III, dNTP, inhibidor de ARNasa y tampón a cada pocillo y se mezcló mediante pipeteado. Las placas se centrifugaron brevemente y se incubaron a 50 °C durante 60 minutos. Se usaron cebadores específicos de región constante para la cadena pesada y la cadena ligera para amplificar los segmentos de genes variables procedentes de una célula individual. Se emplearon cebadores inversos específicos de gen para amplificar las cadenas kappa, lambda y gamma, y las cadenas gamma fueron gamma RT1, kappa RT1 y lambda RT3.

Se realizó la primera ronda de PCR con cebadores específicos de genes de V directos con un fragmento del promotor de citomegalovirus humano (hCMV) en el extremo 5', y cebadores específicos de región constante inversos. El producto de la RT-PCR se utilizó como molde para la primera PCR. El producto de la PCR comprende la región variable de inmunoglobulina y parte de la región constante. Las condiciones de ciclación para la primera PCR incluyeron una etapa de desnaturalización inicial a 98 °C durante 30 minutos, seguida de 13 ciclos de rampa descendiente de temperatura ("touchdown") de 98 °C durante 10 minutos, de 72 °C a 60 °C durante 30 minutos, y 72 °C durante 30 minutos con una caída de 1 °C para cada posterior etapa de reasociación; 20 ciclos de 98 °C durante 10 minutos, 60 °C durante 30 minutos, y 72 °C durante 30 minutos, seguido de una extensión final a 72 °C durante 2 minutos y mantenimiento a 4 °C indefinidamente.

En la segunda ronda de PCR, se empleó un cebador directo genérico que se asoció al marcador de hCMV con un cebador anidado inverso para la región constante. Se usó 1 µl de los productos procedentes de la primera PCR como moldes para la segunda PCR anidada. Las condiciones de ciclación para la segunda incluyeron una etapa de desnaturalización inicial a 98 °C durante 30 minutos, seguida de 20 ciclos de 98 °C durante 10 minutos, 68 °C durante 30 minutos, y 72 °C durante 30 minutos; una extensión final a 72 °C durante 2 minutos y mantenimiento a 4 °C indefinidamente. Se ensayaron 1/3 del producto de la segunda ronda de PCR en un gel de agarosa al 1% para comprobar la tasa de recuperación de la secuencia de anticuerpo desde las células individuales para las etapas de RT y 2 rondas de PCR. Puesto que los cebadores para la cadena pesada y ligera se mezclan en los mismos pocillos, los productos de PCR contienen dos bandas representativas para la región VDJ de cadena pesada y la región VJ de cadena ligera. Los tamaños esperados de los productos de PCR fueron de aproximadamente 700 pb

para las cadenas ligeras kappa y lambda, y aproximadamente 500 pb para la cadena pesada gamma (figura 4a).

- 5 Para la expresión de anticuerpos en células de mamífero, los productos amplificados después se unieron en puente con un módulo de Ig lineal con LTR de PB 5'-promotor de CMV y región constante-señal de poliA-LTR de PB 3' (figura 3). El módulo de Ig contiene todos los elementos fundamentales para la expresión del anticuerpo, incluyendo el promotor de CMV, la región constante de cadena de inmunoglobulina y la señal de poli(A). Además, el módulo presenta regiones solapantes largas de CMV y homología de región constante en sus extremos. Se usaron 2 µl de los productos procedentes de la segunda ronda de como moldes para la PCR en puente. Las condiciones de ciclación para la PCR en puente fueron una etapa de desnaturalización inicial a 98 °C durante 30 minutos, seguida de 5 ciclos de 98 °C durante 10 minutos, 68 °C durante 30 minutos, y 72 °C durante 2 minutos; y 25 ciclos de 98 °C durante 10 minutos, 60 °C durante 30 minutos, y 72 °C durante 2 minutos; seguido de una extensión final a 72 °C durante 2 minutos y mantenimiento a 4 °C indefinidamente. Se ensayaron 1/3 del producto de la PCR en puente en un gel de agarosa al 1% para comprobar la tasa de recuperación de la PCR en puente. Los tamaños esperados de los productos de PCR fueron de aproximadamente 2600 pb para las cadenas ligeras kappa y lambda, y aproximadamente 3100 pb para la cadena pesada gamma (figura 4b).
- 10
- 15 La etapa de puente permite juntar todos los elementos de expresión y las LTR de PB para formar el transposón PB con genes de expresión de cadena pesada y cadena ligera. Independientemente del isotopo de los anticuerpos de ratón, pueden aplicarse IgG1, IgG2a, IgG2b, IgG3 de ratón o IgG1, IgG2, IgG3, IgG4 humana o cualquier variante de región constante en la etapa de puente para reformatear el Fc. El método aplicado en esta tecnología no requiere de ninguna etapa de purificación y puede ser automatizado en gran parte. La tasa de recuperación global para la tecnología de células B ("B-cell technology", BCT) a través de la clasificación de células y la PCR de células individuales es de aproximadamente 38-71% para diferentes poblaciones celulares (tabla 3).
- 20

Ejemplo 1C: Análisis de la secuencia mediante agrupaciones

- Los productos de la segunda ronda de PCR se enviaron para la secuenciación. Las secuencias de nucleótidos se determinaron usando un secuenciador de ADN Applied Biosystems 373. Las secuencias se analizaron mediante el programa Kymab seq-utils (Lee E.C. *et al.*, Nature Biotechnol., 2014, 32:356-363). El programa predice la secuencia de la línea germinal y la hipermutación de la secuencia de IG analizada. La región variable de inmunoglobulina comprende una región VDJ de una secuencia de nucleótidos de inmunoglobulina para genes pesados, y una región VJ de una secuencia de nucleótidos de inmunoglobulina para Igk y Igl. Una familia clonal en general se define mediante el uso de secuencias de V(D)J de cadena pesada y/o cadena ligera de inmunoglobulina relacionadas por 2 o más muestras. Las secuencias de V(D)J de cadena pesada de inmunoglobulina relacionadas pueden identificarse por medio de su utilización compartida de segmentos del gen de V(D)J codificadas en el genoma (figura 5).
- 25
- 30

- Dentro de una familia clonal existen, en general, subfamilias que varían basándose en mutaciones compartidas dentro de sus segmentos de V(D)J, que pueden surgir durante la recombinación de genes y la hipermutación somática de células B. Los clones con diferente utilización de segmentos de V(D)J habitualmente muestran diferentes características de unión. Además, los clones con la misma utilización de segmentos de V(D)J, pero diferentes mutaciones, también muestran diferentes características de unión. Las células B sufren una hipermutación somática, en la que se producen cambios aleatorios en las secuencias de nucleótidos de los genes de anticuerpos, y se seleccionan las células B cuyos anticuerpos tienen una mayor afinidad (figura 6). Si los clones de baja afinidad procedentes del mismo linaje tienen función de neutralización, la potencia habitualmente aumenta en clones con más mutaciones para adquirir una mayor afinidad.
- 35
- 40

Ejemplo 1D: Generación de anticuerpos monoclonales a partir de células individuales

- La etapa de PCR final amplifica el módulo de expresión lineal que codifica la cadena pesada y la cadena ligera. Los módulos amplificados para la cadena pesada y la cadena ligera y el vector de expresión de PB transposasa (PBasa) se cotransfectaron en una línea celular de mamífero sin purificación y clonación. Después se recogió el sobrenadante con la expresión transitoria o estable del anticuerpo en los correspondientes momentos. La transfección del vector de expresión convencional probablemente provoca la integración de concatámeros en el genoma, y el gen integrado se somete a su silenciamiento. La expresión mediada por el transposón PB proporciona una ventaja importante para un nivel de expresión alto y estable de los genes transfectados porque pueden transfectarse múltiples copias (10-100) de transposones PB e integrarse en el genoma dentro de regiones con transcripción activa, lo cual permite un alto nivel de expresión del anticuerpo (figura 7).
- 45
- 50

- Las transfecciones de los productos de la PCR en puente y los vectores de expresión de PBasa para la expresión transitoria se realizaron usando Lipofectamine 2000 siguiendo el protocolo del fabricante. Las transfecciones se realizaron en placas de pocillos profundos de 96 pocillos. Brevemente, se cultivaron células HEK293 en DMEM+FBS con contenido ultrabajo en IgG al 10% (Invitrogen) para evitar que la IgG bovina compita con la IgG humana segregada en la etapa de purificación de proteína A corriente abajo. Para las transfecciones de las placas de 96 pocillos, cada pocillo se sembró el día anterior con 5×10^5 células en 500 µl de medio y se dejó que creciesen hasta 1×10^6 células al día siguiente. Se incubaron 25 µl de los 30 µl de los productos de la PCR en puente en medio Optimem con 100 ng de vector de PBasa para un volumen final de 70 µl, y también se incubó Lipofectamine 2000 por separado con 70 µl de medio Optimem. Ambas incubaciones se realizaron durante 10 minutos. Después se
- 55

mezclaron la Lipofectamine 2000 y los productos de la PCR mediante un pipeteado suave y se incubaron durante 15 minutos antes de la adición a las células HEK293 y mezclar con suavidad. Los sobrenadantes del cultivo se recogieron en el día 8 después de la transfección para las posteriores selecciones. Se determina la concentración de IgG en los sobrenadantes que contienen anticuerpos de interés mediante un ELISA de IgG (figura 8). La concentración de los anticuerpos expresados es comparable a la que se obtiene normalmente con la tecnología del hibridoma, que es suficiente para la mayoría de las selecciones corriente abajo para la capacidad de unión del anticuerpo o los ensayos funcionales. Las tasas de éxito globales para la tecnología de células B a través de la clasificación de células para la identificación de IgG es de 36%-71% dependiendo de diferentes poblaciones de células (tabla 3).

10 Ejemplo 1E: Ensayo de selección de unión de anticuerpos usando un barrido LI-COR Odyssey NIR

Los anticuerpos expresados a partir de la transfección de células HEK293 de la tecnología de células B (BCT) se seleccionaron en primer lugar para su capacidad para unirse al antígeno de interés usando un barrido LI-COR Odyssey NIR, y después los clones positivos se seleccionaron para su afinidad aparente mediante resonancia de plasmón de superficie (ProteOn XPR36, BioRad) (véase a continuación).

15 Las células B que producen anticuerpos específicos de antígeno se identificaron mediante selección fluorescente. Cada pocillo de placas de fondo plano de 384 pocillos transparentes se sembró con 1×10^4 células CHO adherentes transfectadas de modo estable con un gen que codifica un antígeno transmembrana humano en 80 μ l de medio F12 que contiene FBS al 10% (en v/v) ($1,2 \times 10^5$ células/ml) usando un instrumento Multidrop. Las células se incubaron durante la noche a 37 °C en un incubador de CO₂. Al día siguiente, el medio de cultivo se retiró mediante aspiración y se añadieron 45 μ l de anticuerpo antirratón LI-COR IRDye 800CW a 500 ng/ml + 5 mM DRAQ5 (LI-COR) diluido 1:25.000 en tampón FACS (PBS + BSA al 1% + NaN₃ al 0,1%). Se añadieron 5 μ l de sobrenadante de BCT, 5 μ l de anticuerpo control (2 μ g/ml) en medio de cultivo de HEK293, o 5 μ l de anticuerpo control de IgG1 de ratón (Sigma, 2 μ g/ml) en medio de cultivo de HEK293 usando un manipulador de líquidos FluidX. Las placas se incubaron durante 1 hr a 4 °C y se aspiró el medio de cultivo. La reacción se detuvo y las células se fijaron mediante la adición de 25 μ l de paraformaldehído al 4% por pocillo y una incubación durante 15 minutos a temperatura ambiente. Las placas se lavaron dos veces con 100 μ l de PBS, y la disolución de lavado se retiró mediante secado sobre pañuelos de papel secante. Las placas se barrieron usando un instrumento Li-Cor Odyssey Classic. Las tasas de éxito globales para la BCT a través de la clasificación de células para la identificación específica de antígeno son del 25%-61% dependiendo de diferentes poblaciones de células (tabla 3).

30 Tabla 3. Tasa de recuperación de cada etapa de la tecnología de células B

Procedimientos de BCT	Tasa de recuperación			
	Células de memoria/GC de bazo	Células de memoria/GC de LN (RIMMS)	Plasmablato de bazo	Células plasmáticas de médula ósea
PCR de células individuales	38%	66%	71%	50%
Expresión y selección de IgG	36%	66%	71%	50%
Expresión y selección de ligantes específicos de Ag	25%	59%	61%	31%

Ejemplo 1F: Mediciones de afinidad mediante SPR usando el método de captura de anticuerpos

Los clones positivos que expresan anticuerpos específicos de antígenos después se seleccionaron para su afinidad aparente mediante resonancia de plasmón de superficie. Una IgG antirratón (GE Healthcare/Biacore) se acopló a GLM mediante acoplamiento de amina primaria. El chip de GLM (BioRad) se activó usando NHS/EDAC y la IgG antirratón se acopló a esta superficie activada, y después se bloqueó usando etanolamina 1 M. La inmovilización se realizó en HBS-EP (Teknova) o en HBS-N (GE Healthcare/Biacore) a temperatura ambiente o a 37 °C, respectivamente. La superficie de IgG antirratón en el chip de GLM se usó para capturar directamente los anticuerpos de interés. Para el análisis de la cinética se usaron 5 concentraciones de analito (256 nM, 64 nM, 16 nM, 4 nM y 1 nM). Para el análisis de los datos, los sensogramas de unión se referenciaron usando la "intermancha" interna de referencia exclusiva de ProteOn XPR36 que se referenció doblemente usando el sensograma de inyección de tampón. Por último, los datos se analizaron usando el modelo 1:1 inherente al software de análisis de ProteOn XPR36.

5 La figura 9 es un ejemplo de los datos de SPR de los anticuerpos procedentes de Kymouse® inmunizados con ovoalbúmina usando la tecnología de células B individuales de la invención. Aproximadamente dos terceras partes de los anticuerpos ensayados mostraron evidencias de unión al antígeno, con una gama diversa de afinidades de unión y cinéticas diferenciadas. La diversidad de características de unión revela que el procedimiento de clasificación de células es eficaz para capturar una calidad diversa de anticuerpos. A pesar de la escala de este experimento (dos animales), se aislaron muchos anticuerpos de alta afinidad en el intervalo de nM bajos a pM bajos, lo cual verifica la eficacia de la maduración de afinidad.

10 La figura 10 muestra la afinidad aparente de los anticuerpos contra dos antígenos diana Kymab diferentes. Se detectó una gama de ligantes (en blanco), así como de neutralizantes funcionales (en negro), detectándose la mayor afinidad en el intervalo picomolar. Esto valida a tecnología de clonación de células B individuales de los inventores como una herramienta poderosa en la identificación y extracción de anticuerpos de alta afinidad y funcionalmente competentes.

15 Tal como se emplean en esta memoria descriptiva y las reivindicaciones, los términos "comprende" (y cualquier forma de comprender, tal como "que comprende" y "comprendiendo"), "tiene" (y cualquier forma de tener, tal como "que tiene" y "teniendo"), "incluye" (y cualquier forma de incluir, tal como "que incluye" e "incluyendo") o "contiene" (y cualquier forma de contener, tal como "que contiene" y "conteniendo") son inclusivas o sin límites fijos, y no excluyen otras etapas o elementos del método no mencionados.

20 La expresión "o sus combinaciones", tal como se emplea en la presente, se refiere a todas las permutaciones y las combinaciones de los conceptos listados antes de la expresión. Por ejemplo, "A, B, C, o sus combinaciones" pretende incluir al menos uno de: A, B, C, AB, AC, BC, o ABC, y si el orden es importante en un contexto concreto, también BA, CA, CB, CBA, BCA, ACB, BAC, o CAB. Continuando con este ejemplo, se incluyen expresamente las combinaciones que contienen repeticiones de uno o más conceptos o términos, tales como BB, AAA, MB, BBC, AAABCCCC, CBBAAA, CABABB, etc. Los expertos en la técnica comprenderán que, en general, no existe límite en el número de conceptos o términos en cualquier combinación, a menos que sea evidente a partir del contexto.

25 Cualquier parte de esta descripción puede leerse en combinación con cualquier otra parte de la descripción, a menos que sea evidente a partir del contexto.

Todas las composiciones y/o métodos descritos y reivindicados en la presente pueden realizarse y ejecutarse sin experimentación indebida a la luz de la presente descripción.

30

ES 2 811 251 T3

Tabla 1: Secuencias de nucleótidos/ácidos nucleicos:

Oligos de V_H

SEQ ID NO: 1	TCTAGAGAAAACCCTGTGAGCACAGCTC
SEQ ID NO: 2	GAGAATCCCCTGAGAGCTCCGTTC
SEQ ID NO: 3	TCAGAAGCCCCCAGAGCACAACGC
SEQ ID NO: 4	TGGGAGAATCCCCTAGATCACAGCTC
SEQ ID NO: 5	ACAGAAGCCCCCAGAGCGCAGCAC
SEQ ID NO: 6	CCCACCATGGACACACTTTGCTCC
SEQ ID NO: 7	TGGACTIONAAGGCCTTTCCACTTGG
SEQ ID NO: 8	TGGACCTCCTGCACAAGAACATGAAACAC
SEQ ID NO: 9	GCAGTCACCAGAGCTCCAGACAATGTC
SEQ ID NO: 10	AAGAAGAAGCCCCTAGACCACAGCTCCAC
SEQ ID NO: 11	TGAGATTCCCAGGTGTTTCCATTTCAG
SEQ ID NO: 12	AGAGCCCCAGCCCCAGAATTCCCAGGAG
SEQ ID NO: 13	TTCAGTGATCAGGACTGAACACACA
SEQ ID NO: 14	CCCCAGCCTTGGGATTCCCAAGTGTTTTC
SEQ ID NO: 15	TGAGATTCCCACGTGTTTCCATTTCAG
SEQ ID NO: 16	ACTTGGTGATCAGCACGGAGCACCGA
SEQ ID NO: 17	CTGGGATTTTCAGGTGTTTTCATTTGG

Oligos de V_K

SEQ ID NO: 18	GGAGTCAGACCCAGTCAGGACACAGC
SEQ ID NO: 18	GGAGTCAGACCCACTCAGGACACAGC
SEQ ID NO: 20	GGAATCAGTCCCCTCAGGACACAGC
SEQ ID NO: 21	GGAGTCAGTCTCAGTCAGGACACAGC
SEQ ID NO: 22	ATCAGGACTCCTCAGTTCACCTTCTCAC
SEQ ID NO: 23	ATTAGGACTCCTCAGGTCACCTTCTCAC
SEQ ID NO: 24	GAGGAACTGCTCAGTTAGGACCCAGA
SEQ ID NO: 25	GCTACAACAGGCAGGCAGGGGCAGC
SEQ ID NO: 26	GACTACCACCTGCAGGTCAGGGCCAAG

ES 2 811 251 T3

Oligos de VL

SEQ ID NO: 27	atggcctgggtcctcctcctc
SEQ ID NO: 28	atggccggcttcctcctcctc
SEQ ID NO: 29	atgcctgggctctgctcctc

SEQ ID NO: 30	atgccctgggtcatgctcctc
SEQ ID NO: 31	atggcctgggctctgctgctc
SEQ ID NO: 32	atggcatggatccctctcttc
SEQ ID NO: 33	atggcctggaccctcctcctg
SEQ ID NO: 34	atggcctggaccctcctcctc
SEQ ID NO: 35	atggcctggaccctcctcctgg
SEQ ID NO: 36	atggcctggaccgttctcctc
SEQ ID NO: 37	atggcatggggccacactcctg
SEQ ID NO: 38	atggcctggatccctctactt
SEQ ID NO: 39	atggcctggatccctcctcctg
SEQ ID NO: 40	atggcctggaccgctcctctt
SEQ ID NO: 41	atggcctgggtctcctctac
SEQ ID NO: 42	atggcctggaccctcctcctc
SEQ ID NO: 43	atggcttgaccctcctcctc
SEQ ID NO: 44	atggcctggactcctcctcctc
SEQ ID NO: 45	atggcctggactcctcctttt
SEQ ID NO: 46	atggcctggatgatgcttctc
SEQ ID NO: 47	atggcctgggctcctctgctc

Oligos de la región C

	Oligo	Secuencia (5' a 3')
SEQ ID NO: 48	C _H 1	gctctgctgTAGCCCTTGACCAGGCATCC
SEQ ID NO: 49	C _H 2	CAGATCCAGGGGCCAGTGGATAGAC
SEQ ID NO: 50	C _K 1	gtttctgatcgaaCTAACACTCATTCTGTTGAAG
SEQ ID NO: 51	C _K 2	GACAATGGGTGAAGTTGATGTCTTGTGAG
SEQ ID NO: 52	C _L 1	cgacaaccactacctCTATGAACATTCTGTAGGGGC
SEQ ID NO: 53	C _L 2	CTTCTCCACGGTGCTCCCTTCATGC

Tabla 2:

Oligos de V_H

X	Y (es decir, la secuencia X usada para copiar o modificar el segmento de gen Y)
SEQ ID NO: 1	IGHV1-8 (p.e., IGHV1-8*01)
SEQ ID NO: 2	IGHV1-2 (p.e., IGHV1-2*04)
SEQ ID NO: 3	IGHV1-3*01 (p.e., IGHV1-3*01)
SEQ ID NO: 4	IGHV1-18 (p.e., IGHV1-18*01)
SEQ ID NO: 5	IGHV1-24 (p.e., IGHV1-24*01)

SEQ ID NO: 6	IGHV2-5 (p.e., IGHV2-5*10) y/o IGHV2-26 (p.e., IGHV2-26*01)
SEQ ID NO: 7	IGHV3-7 (p.e., IGHV3-7*01)
SEQ ID NO: 8	IGHV4-4 (p.e., IGHV4-4*02)
SEQ ID NO: 9	IGHV6-1 (p.e., IGHV6-1*01)
SEQ ID NO: 10	IGHV7-4-1 (p.e., IGHV7-4-1*01)
SEQ ID NO: 11	IGHV3-9 (p.e., IGHV3-9*01)
SEQ ID NO: 12	IGHV3-11 (p.e., IGHV3-11*01)
SEQ ID NO: 13	IGHV3-13 (p.e., IGHV3-13*01)
SEQ ID NO: 14	IGHV3-15 (p.e., IGHV3-15*01)
SEQ ID NO: 15	IGHV3-20 (p.e., IGHV3-20*01)
SEQ ID NO: 16	IGHV3-21 (p.e., IGHV3-21*01)
SEQ ID NO: 17	IGHV3-23 (p.e., IGHV3-23*01)

5 Oligos de V_K

X	Y (es decir, la secuencia X usada para copiar o modificar el segmento de gen Y)
SEQ ID NO: 18	Uno, más o todos de IGKV1-5, 1-12, 1-8, 1D-8, 1D-43, 1D-16, 1D-9
SEQ ID NO: 19	Uno, más o todos de IGKV1-6, 1-13, 1D-12, 1D-13
SEQ ID NO: 20	IGKV1-17 y/o 1D-17
SEQ ID NO: 21	Uno, más o todos de IGKV1-27, 1-33, 1D-39
SEQ ID NO: 22	Uno, más o todos de IGKV2-28, 2-30, 2D-40
SEQ ID NO: 23	IgKV2-24
SEQ ID NO: 24	Familia de IGKV3
SEQ ID NO: 25	IGKV4-1
SEQ ID NO: 26	IGKV5-2

Oligos de VL

X	Y (es decir, la secuencia X usada para copiar o modificar el segmento de gen Y)
SEQ ID NO: 27	IGLV1-40
SEQ ID NO: 28	IGLV1-47
SEQ ID NO: 29	IGLV10-54
SEQ ID NO: 30	IGLV2-23
SEQ ID NO: 31	IGLV3-1
SEQ ID NO: 32	IGLV3-10
SEQ ID NO: 33	IGLV3-12
SEQ ID NO: 34	IGLV3-19

SEQ ID NO: 35	IGLV3-21
SEQ ID NO: 36	IGLV3-22
SEQ ID NO: 37	IGLV3-25
SEQ ID NO: 38	IGLV3-16
SEQ ID NO: 39	IGLV3-9
SEQ ID NO: 40	IGLV4-3
SEQ ID NO: 41	IGLV3-2
SEQ ID NO: 42	IGLV5-45
SEQ ID NO: 43	IGLV7-43
SEQ ID NO: 44	IGLV9-49
SEQ ID NO: 45	IGLV1-40
SEQ ID NO: 46	IGLV1-47
SEQ ID NO: 47	IGLV10-54

5 Declaraciones de divulgación:

S1. Un método para producir células que codifican un repertorio de proteínas de interés (POI), comprendiendo dicho método:

- a) proporcionar una población de células que expresan un repertorio de POI;
- 10 b) clasificar la población de células para producir una población clasificada de células individuales, y cada célula comprende un ácido nucleico que codifica una respectiva POI;
- c) amplificar el ácido nucleico comprendido en la población de células individuales clasificadas para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican POI;
- 15 d) modificar los ácidos nucleicos que codifican POI amplificados clasificados procedentes de la etapa (c) para producir un repertorio clasificado de módulos de expresión, y cada módulo comprende una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI; y

- e) transferir los módulos de expresión de POI desde dicho repertorio de módulos a una población clasificada de células hospedantes, manteniendo al mismo tiempo la clasificación de los módulos de expresión de POI, y producir un repertorio clasificado de células hospedantes que expresan un repertorio clasificado de POI, en el que el repertorio clasificado de células hospedantes es capaz de expresar de modo estable el repertorio de POI;
- 5 en el que la POI son cadenas de inmunoglobulina, o sus partes, y en el que, en la etapa (a), las células son células aisladas a partir de uno o más animales, que comprenden células B, células del centro germinal, células B de memoria, células secretoras de anticuerpos, células plasmáticas o células de plasmablastos, y en el que la etapa (c) se realiza usando PCR.
- 10 S2. El método de la declaración S1, en el que, en la etapa (e), los módulos de expresión clasificados se transfieren en lotes a las células hospedantes clasificadas.
- 15 S3. El método de la declaración S1 o S2, en el que (i) el repertorio clasificado de los módulos de expresión producidos mediante la etapa (d) se proporciona en una pluralidad de recipientes cuyas localizaciones relativas entre sí están fijadas (por ejemplo, pocillos en una placa o tubos en una rejilla), en el que cada recipiente comprende un respectivo tipo de módulo de expresión, de modo que la localización relativa de los módulos de expresión entre sí está predeterminada; y (ii) los módulos de expresión se transfieren a las células hospedantes clasificadas en la etapa (e), de modo que las localizaciones relativas de los módulos de expresión se mantienen.
- 20 S4. El método de cualquier declaración anterior, en el que (i) el repertorio de los módulos de expresión producidos mediante la etapa (d) se clasifica proporcionando una pluralidad de recipientes (por ejemplo, pocillos en una placa o tubos en una rejilla), en el que cada recipiente comprende secuencias que codifican POI de una célula individual clasificada en la etapa (b); (ii) las células hospedantes clasificadas de la etapa (e) se proporcionan en una pluralidad de recipientes (por ejemplo, pocillos en una placa o tubos en una rejilla), y (iii) los módulos de expresión se transfieren a las células hospedantes clasificadas en la etapa (e), de modo que las células hospedantes en cada respectivo recipiente se mezclan solo con las secuencias que codifican POI derivadas de una célula individual clasificada en la etapa (b).
- 25 S5. El método de la declaración S3 o S4, en el que, en la etapa (d), los módulos de expresión clasificados se proporcionan en una pluralidad de recipientes cuyas localizaciones relativas entre sí están fijadas (por ejemplo, una pluralidad de recipientes con una disposición de X recipientes por Y recipientes, por ejemplo, una placa de 8 x 12 pocillos (placa de 96 pocillos) o una placa que presenta un múltiplo de una disposición de 8 x 12 recipientes (por ejemplo, una placa de 384 pocillos)); y, en la etapa (e), las células hospedantes clasificadas se proporcionan en una pluralidad de recipientes cuyas localizaciones relativas entre sí están fijadas y comprenden la misma disposición que los recipientes usados en la etapa (d) (por ejemplo, los repertorios de las etapas (d) y (e) se proporcionan ambos en placas de 96 pocillos o de 384 pocillos), de modo que la clasificación se mantiene en la etapa (e).
- 30 S6. El método de la declaración S5, en el que cada célula hospedante expresa una primera y una segunda POI, en el que las POI son diferentes.
- 35 S7. El método de cualquier declaración anterior, que comprende además seleccionar el repertorio de POI clasificado para identificar una POI con una característica deseada y/o una secuencia de nucleótidos que codifica la POI identificada; y, opcionalmente, comprende además identificar, amplificar o sintetizar la secuencia de nucleótidos que codifica la POI identificada; y, opcionalmente, producir una POI aislada usando dicha secuencia de nucleótidos identificada, amplificada o sintetizada.
- 40 S8. El método de cualquier declaración anterior, que comprende además seleccionar el repertorio de POI clasificado para identificar una POI con una característica deseada y aislar una célula hospedante que expresa la POI identificada; y, opcionalmente, propagar la célula para producir una línea celular que expresa la POI.
- 45 S9. El método de cualquier declaración anterior, en el que la etapa (e) comprende integrar genómicamente módulos de expresión de POI en los respectivos genomas de las células hospedantes para expresar las respectivas POI, en el que dicha integración genómica se realiza usando un motivo de secuencia de nucleótidos genómico predeterminado para la inserción de los módulos de expresión en el respectivo genoma celular.
- S10. El método de la declaración S9, en el que dicha integración genómica se realiza mediante una integración mediada por transposones.
- 50 S11. El método de la declaración S9 o S10, en el que la etapa (e) comprende múltiples inserciones de módulos de expresión en los respectivos genomas de células hospedantes.
- S12. El método de cualquier declaración anterior, en el que las células hospedantes son células de una línea de células de mamífero (por ejemplo, células CHO o HEK293) o células de levadura.
- S13. El método de cualquier declaración anterior, en el que cada POI comprende un dominio variable de anticuerpo (por ejemplo, un dominio V_H, V_{HH} o V_L).

- S14. El método de cualquier declaración anterior, en el que cada célula de la etapa (a) expresa una primera y una segunda POI, en el que las POI son diferentes;
- 5 en el que la etapa (b) comprende clasificar células individuales hacia sus respectivos recipientes (por ejemplo, sus respectivos pocillos en una o más placas), y realizar las etapas (c) y (d) en dichos recipientes, en el que, después de la etapa (c), cada recipiente comprende las primeras POI amplificadas mezcladas con las POI amplificadas procedentes de la misma célula; y
- en el que la etapa (e) comprende mezclar los respectivos ácidos nucleicos que codifican la primera y la segunda POI procedentes de un respectivo recipiente con células hospedantes;
- 10 en el que la primera y la segunda POI procedentes de la misma célula de la etapa (a) se transfieren a la misma célula hospedante para la expresión de la primera y la segunda POI por la célula hospedante, produciendo con ello un repertorio de células hospedantes clasificadas que coexpresan, cada una, la respectiva primera y segunda POI.
- S15. El método de la declaración S14, en el que la primera y la segunda POI comprenden o consisten en dominios VH y VL de anticuerpos, respectivamente, por ejemplo, la primera y la segunda POI son cadenas pesadas y ligeras de anticuerpos cognadas, respectivamente.
- 15 S16. El método de cualquier declaración anterior, en el que la etapa (b) comprende:
- unir las POI expresadas por células a un ligando cognado (por ejemplo, unir sitios de unión de anticuerpos expresados por células a un antígeno de interés); opcionalmente, en el que el ligando se une a una POI expresada en la superficie de una célula; y
- 20 después clasificar y aislar las células que expresan la POI que se unen al ligando, produciendo con ello la población clasificada de células;
- en el que se emplea la clasificación de células FACS; opcionalmente, FACS de fluorescencia.
- S17. El método de la declaración 16, en el que el ligando cognado es un antígeno marcado con fluorescencia; o una partícula similar a un virus (VLP) que comprende un antígeno expresado sobre la superficie y que comprende además una proteína marcada con fluorescencia.
- 25 S18. El método de cualquier declaración anterior, en el que la etapa (c) se realiza usando RT-PCR empleando un ARN que codifica POI.
- S19. El método de cualquier declaración anterior, en el que la etapa (d) se realiza usando PCR en puente.
- 30 S20. El método de cualquier declaración anterior, en el que la etapa (d) comprende la modificación de ácidos nucleicos amplificados mediante la combinación con una secuencia predeterminada, de modo que dicha secuencia predeterminada flanquea en 5' y/o 3' a las secuencias de nucleótidos que codifican POI de los ácidos nucleicos; opcionalmente, en el que la modificación coloca un elemento regulador (por ejemplo, un promotor 5' y/o una poliA 3') y/o un elemento de transposón (por ejemplo, un elemento del transposón piggyBac) que flanquea en 5' y/o 3' a las secuencias de nucleótidos que codifican POI.
- 35 S21. El método de la declaración S20, en el que cada POI comprende un dominio variable de anticuerpo, y la etapa (c) o (d) combina secuencias de nucleótidos que codifican POI con una región constante de anticuerpo (opcionalmente, una región constante humana) para producir una secuencia de nucleótidos que codifica una cadena de anticuerpo (opcionalmente, una cadena humanizada).
- 40 S22. Un método según la declaración S21, que comprende además seleccionar el repertorio de POI clasificado para identificar una célula hospedante que expresa una cadena de anticuerpo con una característica deseada (por ejemplo, la unión específica a un antígeno), identificar la secuencia de nucleótidos que codifica la cadena de anticuerpo de la célula hospedante, usar la secuencia de nucleótidos que codifica la cadena de anticuerpo para producir copias de la cadena de anticuerpo identificada, y formular las copias como una composición farmacéutica (opcionalmente, en combinación con uno o más fármacos adicionales, excipientes, diluyentes o vehículos) para la terapia médica humana; y, opcionalmente, administrar la composición a un paciente humano para la terapia médica del paciente.
- 45 S23. El método de cualquier declaración anterior, en el que la etapa (e) es automática; opcionalmente, en el que una o todas las etapas (b) a (c) también son automáticas.
- 50 S24. El método de cualquier declaración anterior, en el que la población de células en la etapa (a) se identifica mediante la presencia de uno o más marcadores de la superficie celular, tales como uno o más de: CD19, IgM, IgD, CD38, o CD95, por ejemplo, todos de CD19, IgM, IgD, CD38, y CD95.
- S25. Un aparato automático para realizar el método de cualquier declaración anterior, comprendiendo dicho aparato:

- a. un medio para contener una población de células individuales clasificadas en una pluralidad de recipientes, en el que cada célula individual está en un respectivo recipiente, y cada célula comprende un ácido nucleico que codifica una respectiva POI;
- 5 b. un medio para trasladar reactivos de PCR a los recipientes para amplificar el ácido nucleico comprendido en la población de células individuales clasificadas para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican POI;
- 10 c. un medio para trasladar a los recipientes reactivos para modificar los ácidos nucleicos que codifican POI amplificados y clasificados para producir un repertorio clasificado de módulos de expresión, y cada módulo comprende una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI;
- d. un medio para contener una población clasificada de células hospedantes en una pluralidad de recipientes;
- e. un medio para transferir (opcionalmente, para transferir en lotes) los módulos de expresión POI desde dicho repertorio de módulos a una población clasificada de células hospedantes en los recipientes, manteniendo al mismo tiempo la clasificación de los módulos de expresión de POI; y
- 15 f. un medio para realizar la transfección de los módulos de expresión hacia el interior de las células hospedantes en los recipientes para producir un repertorio clasificado de células hospedantes que expresan un repertorio clasificado de POI; y, opcionalmente,
- g. un medio para clasificar una población de células para producir la población clasificada de células individuales; y, opcionalmente,
- 20 h. un medio para controlar el funcionamiento del aparato para la ejecución automática del método de una cualquiera de las declaraciones 1 a 24.
- S26. Un kit para realizar el método de una cualquiera de las declaraciones S1 a S24, comprendiendo dicho kit un aparato según la declaración S25, junto con un ácido nucleico que comprende uno o más elementos de transposón para realizar el método de la declaración 20.
- 25 S27. Un módulo de expresión para la expresión de una POI en una célula hospedante, y el módulo es proporcionado por un ácido nucleico lineal que comprende un transposón, y el transposón comprende elementos de transposón 5'-y 3'-terminales con una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para la expresión de la POI entre los elementos de transposón, en el que la población codifica un repertorio de POI.
- S28. Una población clasificada de módulos de expresión según la declaración S27.
- 30 S29. Una población clasificada de módulos de expresión que codifican un repertorio de POI que se corresponden con POI expresadas por una población de células, y cada módulo comprende una secuencia de nucleótidos que codifica un miembro del repertorio de POI y uno o más elementos reguladores para la expresión de POI, en la que cada uno de dichos módulos comprende la disposición (en la dirección 5' a 3'): elemento de transposón - [secuencia de nucleótidos de POI y uno o más elementos reguladores] - elemento de transposón, y los módulos de expresión para la expresión de POI correspondientes a POI de diferentes células están aislados entre sí en la población clasificada (por ejemplo, en diferentes pocillos de una placa), en la que cada módulo de expresión es proporcionado por un ADN lineal.
- 35 S30. Una población clasificada de células hospedantes que comprende la población clasificada de módulos de expresión según la declaración S28 o la declaración S29 para la expresión de un repertorio clasificado de POI.
- 40 S31. Un método para preparar un transposón que comprende una secuencia de nucleótidos de interés (NOI), comprendiendo dicho método:
- a. proporcionar una primera secuencia de nucleótidos que comprende (en la dirección 5' a 3') A, B y C, en el que A es una primera secuencia de homología, B es una secuencia de nucleótidos que comprende la NOI, y C es una segunda secuencia de homología;
- 45 b. proporcionar una primera secuencia de nucleótidos de molde que comprende (en la dirección 5' a 3') W y X, en el que W es una secuencia de nucleótidos que comprende un primer elemento de transposón, y X es una tercera secuencia de homología; y
- c. proporcionar una segunda secuencia de nucleótidos de molde que comprende (en la dirección 5' a 3') Y e Z, en el que Y es una cuarta secuencia de homología, y Z es una secuencia de nucleótidos que comprende un segundo elemento de transposón;
- 50

y bien

- d. (i) mezclar la primera secuencia de nucleótidos con el primer molde para hibridar entre sí el primer y el tercer brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la primera secuencia de nucleótidos usando el primer molde para producir una primera secuencia de nucleótidos extendida (primera ENS) que comprende (en la dirección 5' a 3') W, B y C; y (ii) mezclar la primera ENS con el segundo molde para hibridar entre sí el segundo y el cuarto brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la primera ENS para producir una segunda ENS que comprende (en la dirección 5' a 3') W, B y Z; o bien
- (ii) mezclar la primera secuencia de nucleótidos con el segundo molde para hibridar entre sí el segundo y el cuarto brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la primera secuencia de nucleótidos usando el segundo molde para producir una tercera secuencia de nucleótidos extendida (tercera ENS) que comprende (en la dirección 5' a 3') A, B y Z; y (ii) mezclar la tercera ENS con el primer molde para hibridar entre sí el primer y el tercer brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la tercera ENS para producir una cuarta ENS que comprende (en la dirección 5' a 3') W, B y Z; o bien
- (iii) mezclar la primera secuencia de nucleótidos con el primer y el segundo molde para hibridar entre sí el primer y el tercer brazo de homología y para hibridar entre sí el segundo y el cuarto brazo de homología y realizar la amplificación y extensión del ácido nucleico para extender la primera secuencia de nucleótidos usando el segundo molde para producir una quinta ENS que comprende (en la dirección 5' a 3') W, B y Z;
- y
- e. aislar una ENS que comprende (en la dirección 5' a 3') W, B y Z, produciendo con ello un transposón aislado que comprende una NOI flanqueada por elementos de transposón; y
- f. opcionalmente, introducir el transposón aislado en una célula receptora, de modo que el transposón se integra en el genoma de la célula.
- S32. El método de la declaración S31, en el que existe una secuencia de nucleótidos intermedia entre W y X y/o una secuencia de nucleótidos intermedia entre Y e Z; opcionalmente, en el que la secuencia intermedia o cada una de las secuencias intermedias es un elemento regulador o una secuencia que codifica una proteína.
- S33. El método de la declaración S32, en el que la NOI codifica un dominio de proteína (por ejemplo, un dominio variable de anticuerpo) y existe una secuencia de nucleótidos que codifica una región constante de anticuerpo (por ejemplo, un Fc de anticuerpo, por ejemplo, un Fc humano) entre Y e Z, por lo cual el producto de transposón codifica una proteína de fusión que comprende un dominio de proteína condensado con una región constante de anticuerpo (por ejemplo, codifica una cadena de anticuerpo).
- S34. El método de una cualquiera de las declaraciones 30 a 33, en el que uno o más del primer y segundo brazo de homología se combina con la NOI mediante PCR (por ejemplo, 5'- y/o 3'-RACE) para formar dicha primera secuencia de nucleótidos antes de realizar dicha extensión.
- S35. Un método para preparar un repertorio de transposones, en el que los miembros del repertorio codifican diferentes POI (por ejemplo, diferentes dominios variables de anticuerpo), comprendiendo dicho método:
- i. proporcionar una población de primeras secuencias de nucleótidos que comprende un repertorio de NOI; y clasificar las primeras secuencias de nucleótidos para proporcionar una población clasificada antes de realizar la etapa (ii), en el que se produce un repertorio clasificado de transposones que codifica un repertorio clasificado de POI;
- ii. para cada primera secuencia de nucleótidos, realizar el método de una cualquiera de las declaraciones 31 a 34, produciendo con ello un repertorio de transposones que codifican un repertorio de POI.
- S36. El método de la declaración S36, en el que los transposones del repertorio se introducen en células receptoras, de modo que los transposones se integran en el genoma de las células, y cada transposón integrado comprende un módulo de expresión de POI flanqueado por elementos de transposón y el módulo comprende una NOI y uno o más elementos reguladores para la expresión de la POI en una célula hospedante.
- S37. Un método para producir una célula hospedante para la expresión de una POI, comprendiendo dicho método:
- a. proporcionar al menos un primer y un segundo módulo de expresión, en el que cada módulo de expresión comprende:
- i. un primer elemento de integración y un segundo elemento de integración 3' de la secuencia de nucleótidos del primer elemento de integración; y
- ii. entre los elementos de integración, una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI;
- iii. en el que los elementos de integración son capaces de insertarse en un ácido nucleico mediante el

reconocimiento de un motivo de secuencia de nucleótidos predeterminado del ácido nucleico usando una enzima de integración;

b. proporcionar una célula hospedante cuyo genoma comprende una pluralidad de dichos motivos; y

5 c. de modo simultáneo o secuencial, introducir el primer y el segundo módulo de expresión en la célula hospedante, en el que cada módulo se integra genómicamente en el genoma de la célula hospedante en dicho motivo para la expresión de las POI por la célula hospedante; y

d. opcionalmente, producir una expresión de las POI en la línea celular en una etapa que comprende cultivar la célula hospedante.

10 S38. El método de la declaración S37, en el que el primer y el segundo elemento de integración del primer módulo son idénticos al primer y segundo elemento de integración, respectivamente, del segundo módulo, opcionalmente en el que el primer y el segundo elemento de integración de cada uno de dichos primer y segundo módulo están en una orientación mutuamente invertida.

S39. El método de las declaraciones S37 o S38, en el que uno o más de dichos motivos se introduce mediante ingeniería en el cromosoma de la célula hospedante antes de realizar la etapa (c).

15 S40. El método de una cualquiera de las declaraciones S37 a S39, en el que uno o más de dichos motivos es endógeno del genoma de la célula hospedante; opcionalmente, en el que cada uno de dichos motivos en donde se integra un módulo es un motivo endógeno.

20 S41. El método de una cualquiera de las declaraciones S37 a S40, en el que cada módulo se integra mediante recombinación homóloga entre los sitios de integración y el genoma del hospedante; mediante recombinación específica de sitio entre los sitios de integración y el genoma del hospedante; o mediante integración mediada por transposones.

25 S42. El método de una cualquiera de las declaraciones S37 a S41, que comprende aislar una célula hospedante producida mediante la etapa (c) o (d) e identificar, amplificar o sintetizar la secuencia de nucleótidos que codifica la POI expresable por la célula; y, opcionalmente, producir una POI aislada usando dicha secuencia de nucleótidos identificada, amplificada o sintetizada, o uno de sus mutantes.

S43. El método de la declaración S42, que comprende formular la POI aislada en un fármaco para la medicina humana; y, opcionalmente, administrar el fármaco a un paciente humano.

30 S44. Una población de células hospedantes que puede obtenerse mediante el método de una cualquiera de las declaraciones 37 a 43, y cada célula comprende una pluralidad de módulos de expresión integrados genómicamente para expresar POI, y cada célula hospedante comprende una pluralidad de motivos de secuencia de nucleótidos idénticos a través de su genoma adyacentes a un módulo de expresión integrado para la expresión de una POI a partir de cada módulo; comprendiendo cada módulo integrado:

a. una secuencia de un primer elemento de integración y una secuencia de un segundo elemento de integración 3' de la secuencia de nucleótidos del primer elemento de integración; y

35 b. entre las secuencias de los elementos de integración, una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI.

40 S45. Un anticuerpo o un sitio de unión de anticuerpo de un anticuerpo para el tratamiento médico de un ser humano, en el que el anticuerpo o el sitio de unión se ha aislado a partir de una célula hospedante producida mediante el método de una cualquiera de las declaraciones S37 a S43, o se ha aislado a partir de una célula hospedante de una población según la declaración S44.

45 S46. Una mezcla de ácidos nucleicos que comprende un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en la que el primer ácido nucleico es capaz de hibridarse con una secuencia de 5'UTR de una región V de un anticuerpo humano de un gen comprendido en un ácido nucleico diana, en la que el gen codifica una región V humana; y el segundo ácido nucleico es capaz de hibridarse con una segunda secuencia, en la que la segunda secuencia está comprendida en el ácido nucleico diana y se encuentra 3' con respecto a la secuencia de UTR, en la que el primer ácido nucleico aislado comprende una secuencia que es al menos 90% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-47, y en la que el segundo ácido nucleico aislado comprende una secuencia de región constante de anticuerpo; opcionalmente, una secuencia que es al menos 90% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:48 a 53.

50 S47. Una mezcla de ácidos nucleicos que comprende un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en la que los ácidos nucleicos son diferentes y se seleccionan de ácidos nucleicos que comprenden una secuencia que es al menos 90% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-47.

S48. Un kit que comprende una o más mezclas según la declaración 46 o 47, y un aparato según la declaración

S25.

S49. Un método para amplificar un repertorio de secuencias de región variable humanas, comprendiendo dicho método:

5 a. proporcionar una población de células que expresa un repertorio de regiones variables humanas, en el que las células comprenden secuencias de nucleótidos que codifican las regiones variables, en el que las células en la etapa (a) son células individuales clasificadas (por ejemplo, clasificadas en pocillos de una o más placas);

b. replicar una pluralidad de dichas secuencias de nucleótidos que codifican regiones variables usando PCR y moldes de PCR; y

10 c. aislar, secuenciar o identificar una o más de las secuencias de nucleótidos replicadas, o realizar las etapas (d) y (e) del método de una cualquiera de las declaraciones 1 a 34;

en el que uno o más moldes de la etapa (b) comprende una secuencia que es al menos 90% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1- 52.

15 S50. Un cebador de PCR o un vector de recombinación homóloga que comprende al menos 15 nucleótidos contiguos de una secuencia de UTR de un segmento de gen variable de anticuerpo humano para hibridarse con la secuencia de 5'UTR de un segmento de gen variable de anticuerpo denominado Y en la tabla 2, en el que la secuencia del cebador/vector se selecciona del grupo que consiste en las secuencias denominadas X en la tabla 2.

REIVINDICACIONES

1.- Un método para producir células que codifican un repertorio de anticuerpos que comprende cadenas pesadas y cadenas ligeras de anticuerpo cognadas, comprendiendo dicho método:

5 a) proporcionar una población de células que expresan un repertorio de anticuerpos que comprende cadenas pesadas y cadenas ligeras de anticuerpo cognadas, cuyas células se aíslan de uno o más animales, y que comprende células B, células del centro germinal, células B de memoria, células secretoras de anticuerpos, células plasmáticas o células de plasmablastos;

10 b) clasificar la población de células mediante la unión de los anticuerpos a un antígeno de interés para producir una población clasificada de células individuales, y cada célula comprende un ácido nucleico que codifica una respectiva cadena pesada de anticuerpo y cadena ligera de anticuerpo cognadas, y cada célula clasificada se proporciona en un recipiente distinto;

15 c) amplificar el ácido nucleico comprendido en la población de células individuales clasificadas usando PCR para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican secuencias VH y VL cognadas, en el que las secuencias VH y VL de cada anticuerpo están comprendidas por el anticuerpo cognado procedente de la misma célula y se amplifican en el mismo recipiente;

d) modificar los ácidos nucleicos que codifican VH y VL cognadas amplificados y clasificados procedentes de la etapa (c) para producir un repertorio clasificado de módulos de expresión, y cada recipiente comprende:

un primer tipo de módulo de expresión de anticuerpo cognado que comprende:

20 (i) una secuencia de nucleótidos que codifica una secuencia de región constante flanqueada en 5' por una secuencia de VH cognada derivada de dicha célula individual para producir una secuencia de nucleótidos que codifica una cadena pesada de anticuerpo; y

(ii) uno o más elementos reguladores para expresar la cadena pesada de anticuerpo, y

un segundo tipo de módulo de expresión de anticuerpo cognado que comprende:

25 (iii) una secuencia de nucleótidos que codifica una secuencia de región constante flanqueada en 5' por una secuencia de VL cognada derivada de dicha célula individual para producir una secuencia de nucleótidos que codifica una cadena ligera de anticuerpo; y

(iv) uno o más elementos reguladores para expresar la cadena pesada de anticuerpo, y

en el que las secuencias de VH y VL de cada cadena pesada de anticuerpo y cadena ligera de anticuerpo cognadas se modifican en el mismo recipiente; y

30 e) transferir los módulos de expresión de anticuerpo cognado desde dicho repertorio de módulos a una población clasificada de células hospedantes, manteniendo al mismo tiempo la clasificación de los módulos de expresión de anticuerpo cognado, y producir un repertorio clasificado de células hospedantes que coexpresan un anticuerpo que comprende una cadena pesada de anticuerpo y una cadena ligera de anticuerpo cognadas, en el que el repertorio clasificado de células hospedantes es capaz de expresar de forma estable el repertorio de anticuerpos.

35 2.- El método de la reivindicación 1, en el que, en la etapa (e), los módulos de expresión clasificados se transfieren en lotes a las células hospedantes clasificadas.

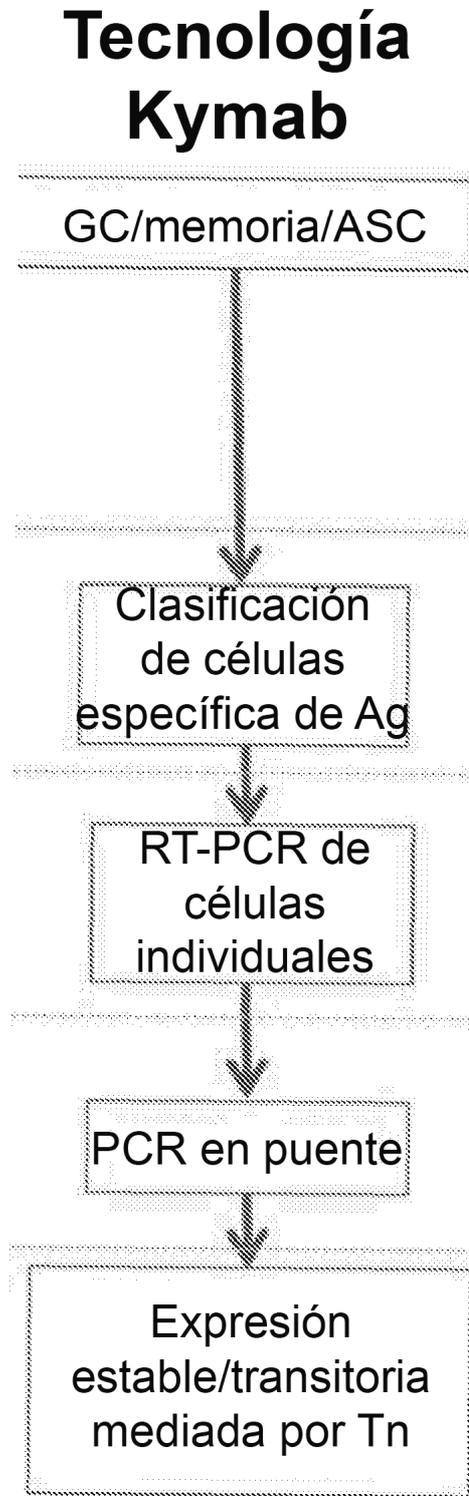
3.- El método de la reivindicación 2, en el que, en la etapa (e), los módulos de expresión clasificados se transfieren sin purificación.

40 4.- El método de una cualquiera de las reivindicaciones 1 a 3, en el que (i) el repertorio clasificado de los módulos de expresión producidos mediante la etapa (d) se proporciona en una pluralidad de recipientes cuyas localizaciones relativas entre sí están fijadas (por ejemplo, pocillos en una placa o tubos en una rejilla), en el que cada recipiente contiene un respectivo tipo de módulo de expresión, de modo que la localización relativa de los módulos de expresión entre sí está predeterminada; y (ii) los módulos de expresión se transfieren a las células hospedantes clasificadas en la etapa (e), de modo que las localizaciones relativas de los módulos de expresión se mantienen.

45 5.- El método de cualquier reivindicación anterior, en el que (i) el repertorio de los módulos de expresión producidos mediante la etapa (d) se clasifica proporcionando una pluralidad de recipientes (por ejemplo, pocillos en una placa o tubos en una rejilla), en el que cada recipiente comprende secuencias que codifican un sitio de unión de anticuerpo de una célula individual clasificada en la etapa (b); (ii) las células hospedantes clasificadas de la etapa (e) se proporcionan en una pluralidad de recipientes (por ejemplo, pocillos en una placa o tubos en una rejilla), y (iii) los
50 módulos de expresión se transfieren a las células hospedantes clasificadas en la etapa (e), de modo que las células hospedantes en cada respectivo recipiente se mezclan solo con las secuencias que codifican un sitio de unión de anticuerpo derivadas de una célula individual clasificada en la etapa (b).

- 6.- El método de cualquier reivindicación anterior, en el que, en la etapa (b), se emplea la clasificación de células FACS; opcionalmente, FACS de fluorescencia.
- 7.- El método de la cualquier reivindicación anterior, en el que el antígeno de interés es un antígeno marcado con fluorescencia; o una partícula similar a un virus (VLP) que comprende un antígeno expresado sobre la superficie y que comprende además una proteína marcada con fluorescencia
- 8.- El método de cualquier reivindicación anterior, en el que la etapa (d) comprende la modificación de ácidos nucleicos amplificados para colocar un elemento regulador (por ejemplo, un promotor 5' y/o una poliA 3') y/o un elemento de transposón (por ejemplo, un elemento del transposón piggyBac) que flanquea en 5' y/o 3' a las secuencias de nucleótidos que codifican una cadena de anticuerpo.
- 9.- El método de cualquier reivindicación anterior, en el que la etapa (e) comprende integrar genómicamente los módulos de expresión de anticuerpo cognado en los respectivos genomas de las células hospedantes para expresar los respectivos anticuerpos cognados, en el que dicha integración genómica se realiza usando un motivo de secuencia de nucleótidos genómico predeterminado para la inserción de los módulos de expresión en el respectivo genoma de una célula.
- 10.- El método de la reivindicación 9, en el que dicha integración genómica se realiza mediante integración mediada por transposones.
- 11.- El método de cualquier reivindicación anterior, en el que, en la etapa (d), la modificación se realiza usando una PCR en puente para producir un repertorio de primeros y segundos módulos de expresión que comprenden elementos de transposón 5'- y 3'-terminales con una respectiva secuencia de nucleótidos que codifica una cadena pesada de anticuerpo o una cadena ligera de anticuerpo cognadas, y uno o más elementos reguladores para la expresión de las cadenas de anticuerpo entre los elementos de transposón.
- 12.- El método de cualquier reivindicación anterior, que comprende además seleccionar el repertorio de anticuerpos clasificado para identificar una célula hospedante que expresa un anticuerpo con una característica deseada (por ejemplo, la unión específica a un antígeno o una afinidad de unión a un antígeno), identificar las secuencias de nucleótidos que codifican cadenas de anticuerpo de la célula hospedante, usar las secuencias de nucleótidos que codifican cadenas de anticuerpo para producir copias de las cadenas de anticuerpo identificadas, y formular las copias como una composición farmacéutica (opcionalmente, en combinación con uno o más fármacos adicionales, excipientes, diluyentes o vehículos) para la terapia médica humana.
- 13.- Una mezcla de ácidos nucleicos, en el que la mezcla de ácidos nucleicos comprende:
- (i) un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en la que el primer ácido nucleico es capaz de hibridarse con una secuencia de 5'UTR de una región V de un anticuerpo humano de un gen comprendido en un ácido nucleico diana, en la que el gen codifica una región V humana; y el segundo ácido nucleico es capaz de hibridarse con una segunda secuencia, en la que la segunda secuencia está comprendida en el ácido nucleico diana y se encuentra 3' con respecto a la secuencia de UTR, en la que el primer ácido nucleico aislado comprende una secuencia que es al menos 95% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-47, y en la que el segundo ácido aislado comprende una secuencia de región constante de anticuerpo; opcionalmente, una secuencia que es al menos 95% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:48 a 53, o
- (ii) un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en la que los ácidos nucleicos son diferentes y se seleccionan de ácidos nucleicos que comprenden una secuencia que es al menos 95% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-47.
- 14.- Un método para amplificar un repertorio de secuencias de región variable humanas, comprendiendo dicho método:
- a. proporcionar una población de células que expresa un repertorio de regiones variables humanas, en el que las células comprenden secuencias de nucleótidos que codifican las regiones variables, en el que las células en la etapa (a) son células individuales clasificadas (por ejemplo, clasificadas en pocillos de una o más placas);
- b. replicar una pluralidad de dichas secuencias de nucleótidos que codifican regiones variables usando PCR y moldes de PCR; y
- c. aislar, secuenciar o identificar una o más de las secuencias de nucleótidos replicadas, o realizar las etapas (d) y (e) del método de una cualquiera de las reivindicaciones 1 a 11;
- en el que uno o más moldes de la etapa (b) comprende una secuencia que es al menos 95% idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO:1-52.

FIGURA 1A



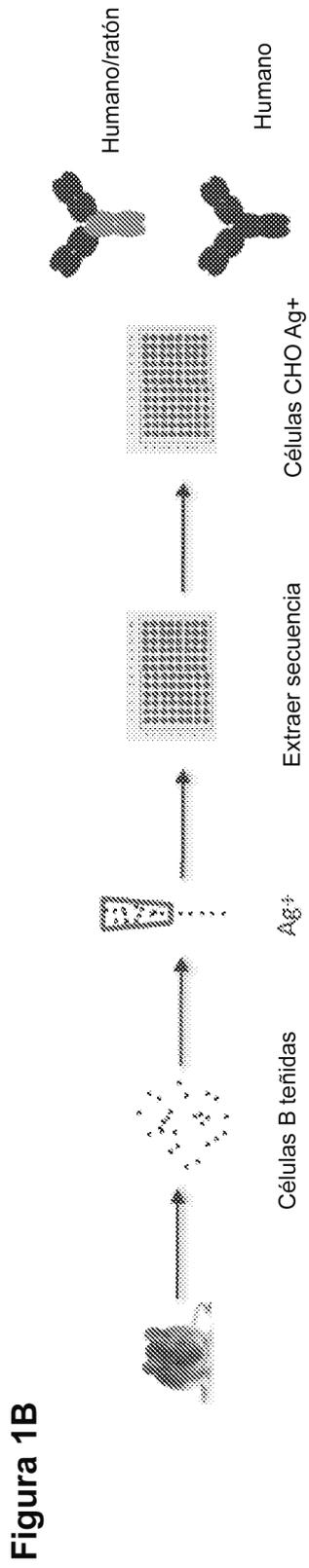


Figura 1B

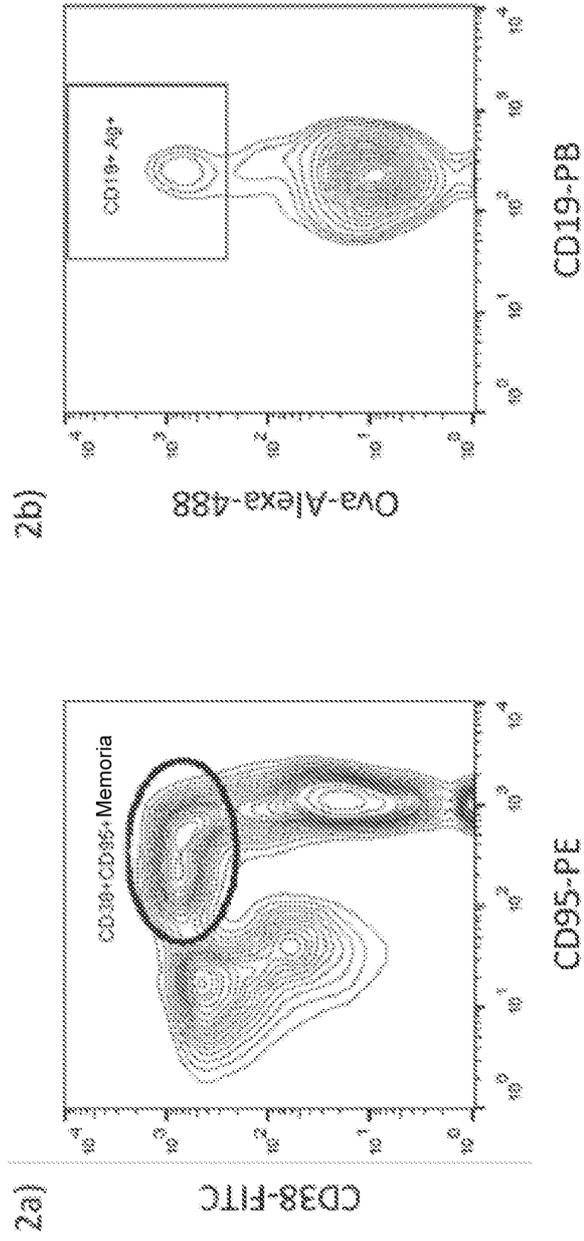


Figura 2

Figura 3

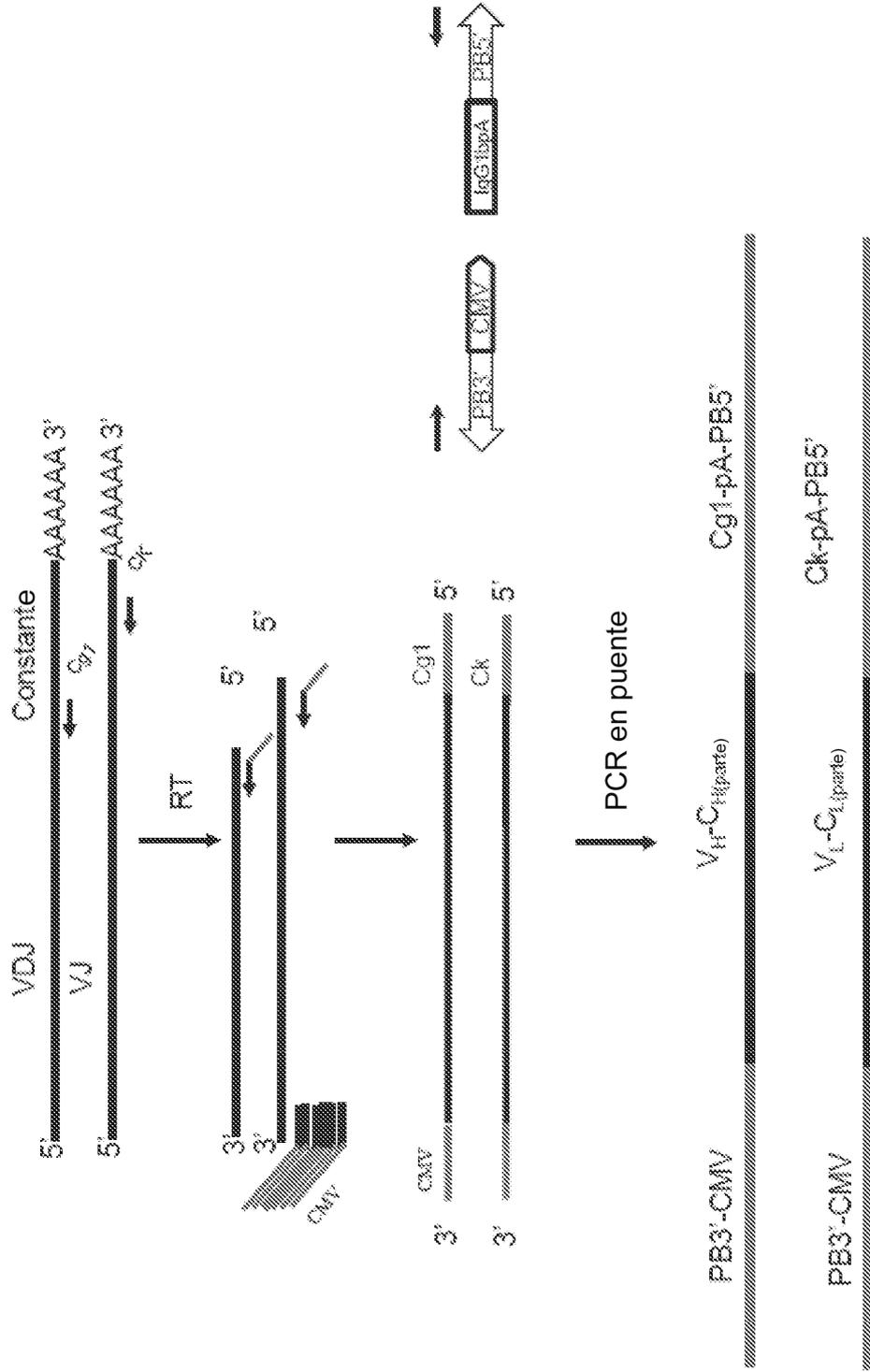


Figura 4

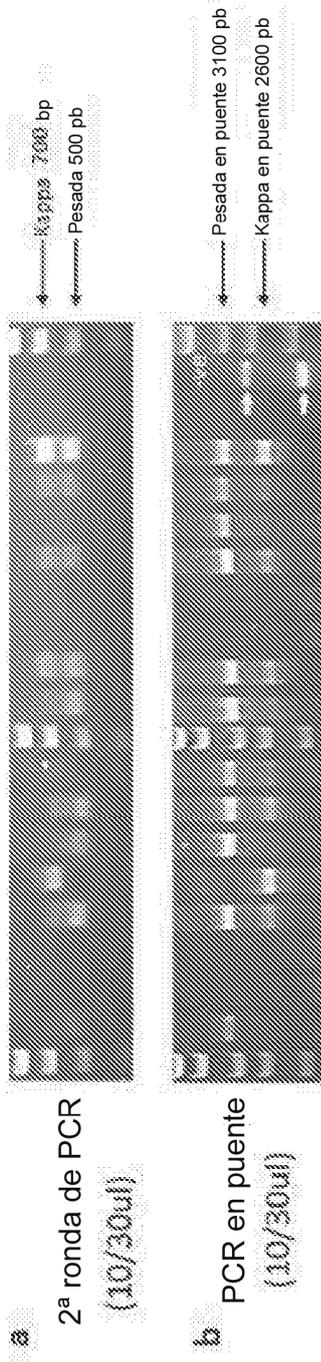


Figura 7

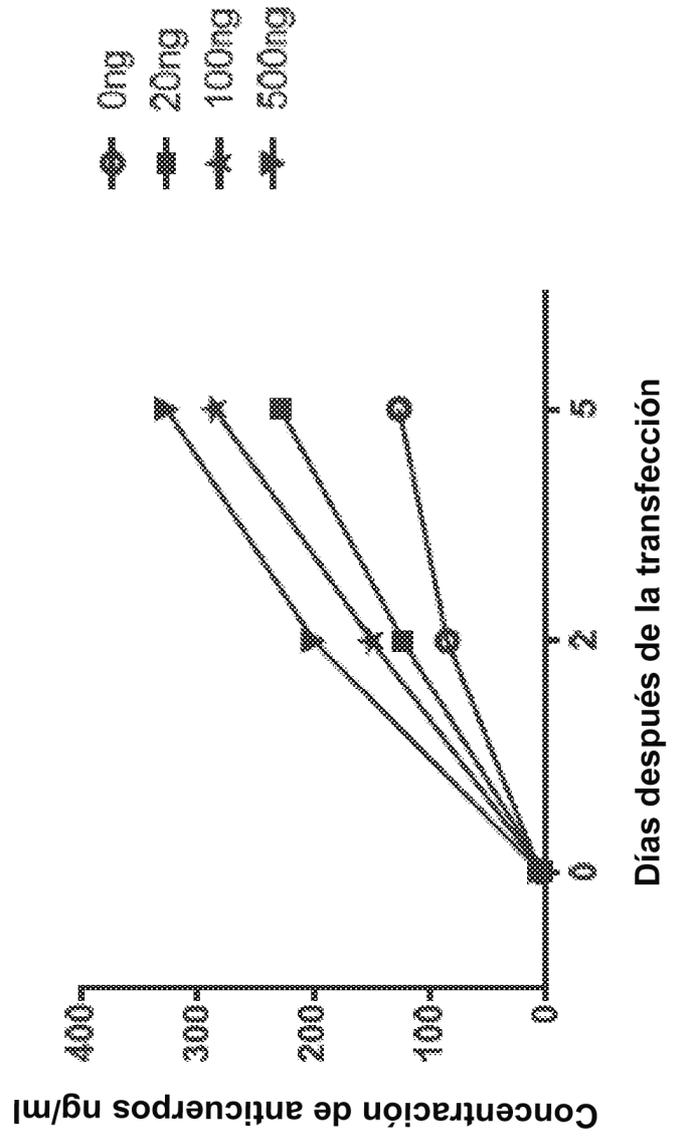
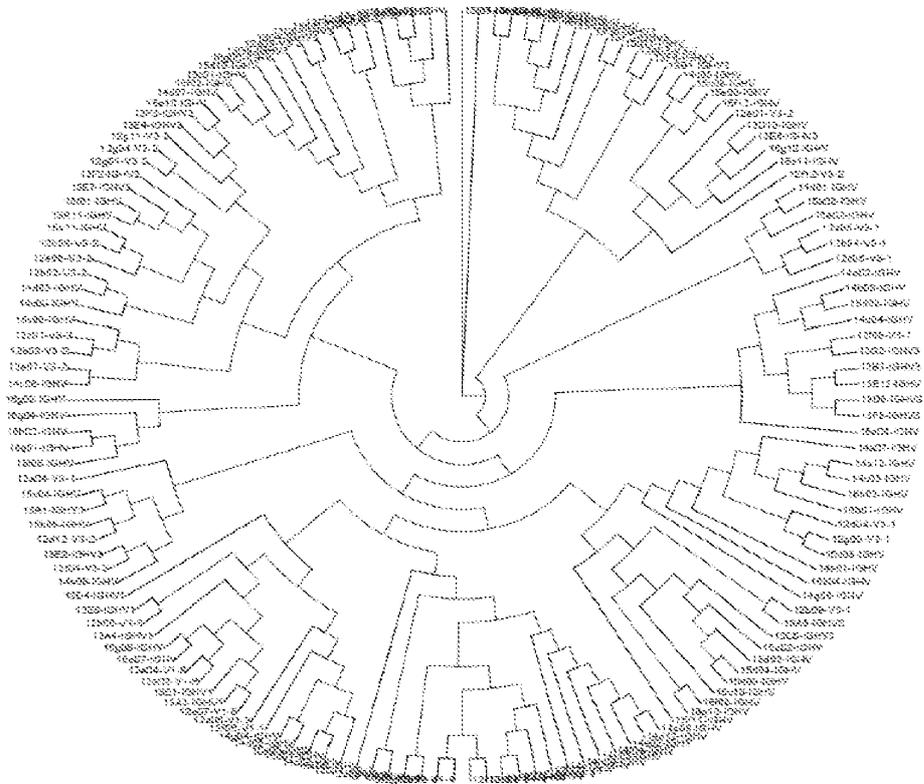


Figura 5

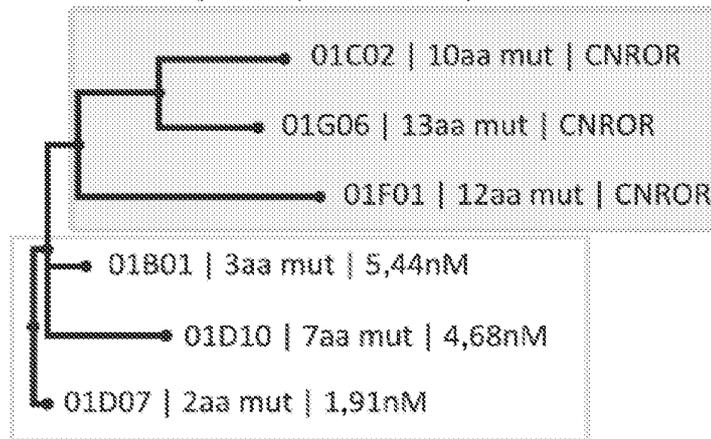


Según las secuencias de cadena pesada

Figura 6

Ejemplo de agrupación

VH1-18|D2-21|J4 & VK1-9|J4



■ Neutralizadores de alta potencia

□ Neutralizadores pero por debajo del 40% del umbral

Basado en un ensayo de neutralización de un solo punto

Figura 8

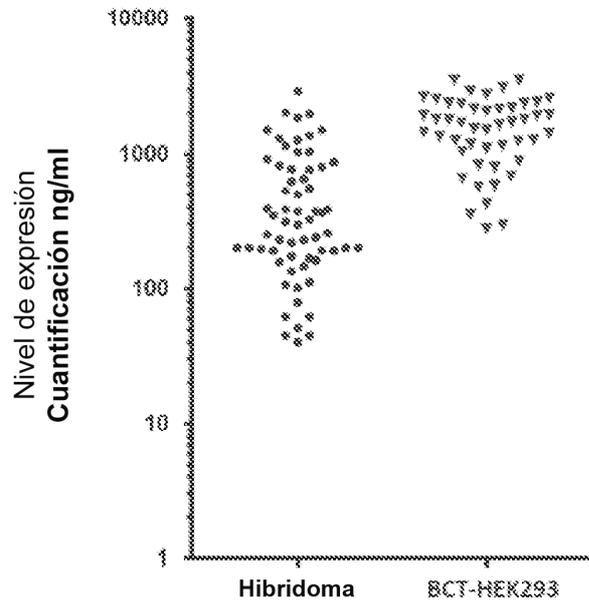
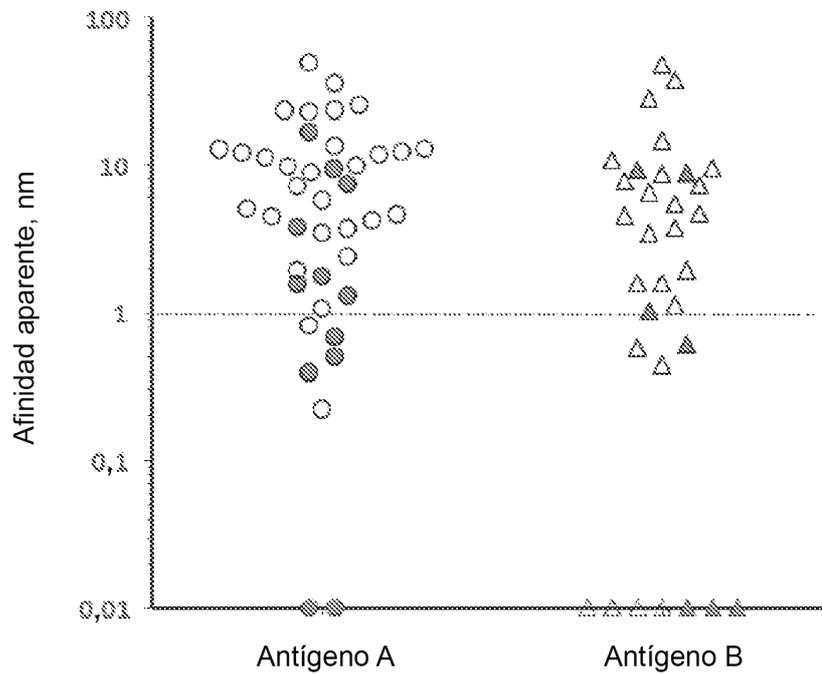


Figura 10



Círculos/triángulos oscuros: neutralizantes

Figura 9B

