

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 997 196**

51 Int. Cl.:

**C12N 15/10**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.09.2014** **E 20184286 (1)**

97 Fecha y número de publicación de la concesión europea: **11.09.2024** **EP 3766970**

54 Título: **Producción de vectores de expresión y cribado celular de alto rendimiento**

30 Prioridad:

**19.09.2013 GB 201316644**

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

**14.02.2025**

73 Titular/es:

**KYMAB LIMITED (100.00%)  
The Bennet Building (B930) Babraham Research  
Campus  
Cambridge Cambridgeshire CB22 3AT, GB**

72 Inventor/es:

**BRADLEY, ALLAN;  
LEE, E-CHIANG;  
LIANG, QI;  
LIU, HUI;  
WOOD, ANDREW y  
WONG, VIVIAN**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

ES 2 997 196 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 cribado celular de alto rendimiento

La presente invención y divulgación se refiere, *entre otras cosas*, a la producción de vectores de expresión, así como a la aplicación a la producción de células huésped para la expresión del repertorio de proteínas y el cribado de alto rendimiento. La divulgación también se refiere a ácidos nucleicos, cebadores de la PCR y mezclas útiles para la amplificación por PCR de secuencias de nucleótidos que codifican dominios variables de anticuerpos humanos o recombinación homóloga con loci de Ig en células.

## ANTECEDENTES

La técnica reconoce la conveniencia de cribar repertorios de células o proteínas para identificar proteínas de interés (POI, por sus siglas en inglés) o células que expresan éstas. Por ejemplo, la técnica comprende técnicas para cribar células B y repertorios de anticuerpos, con el fin de identificar una o más células que expresan anticuerpos que exhiben una característica deseada (típicamente unión específica a antígenos). El cribado no se limita al cribado de anticuerpos, sino que también puede ser aplicable a colecciones de cribado de otros tipos de proteínas para una o más características deseadas.

Con el fin de expresar repertorios de proteínas, las secuencias de nucleótidos correspondientes pueden ser clonadas en vectores de expresión respectivos e introducidas (p. ej., transfectadas) en células huésped que pueden expresar las proteínas. Comúnmente, la clonación molecular se utiliza para clonar secuencias codificantes de proteínas de un repertorio en células huésped. La clonación molecular se ha aplicado a técnicas de cribado de células B (linfocitos), en las que las células B se expanden mediante cultivo, se clasifican en células individuales (p. ej., utilizando un ensayo de placa hemolítica o un ensayo de focos fluorescentes), ARNm de cadena de anticuerpos (o región variable) de las células clasificadas se transcribe inversamente y se amplifica utilizando RT-PCR, el ADN amplificado se somete a clonación molecular para introducir secuencias codificantes de cadena de anticuerpos en vectores de ácido nucleico, y los vectores se introducen luego en células huésped para la expresión transitoria (en los casos en los que el vector es episomal) o para la expresión estable (por integración aleatoria en el genoma de la célula huésped). Véase, por ejemplo, Proc Natl Acad Sci USA. 23 de julio de 1996;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 de SLAM) y doi: 10.1016/j.jala.2009.05.004 Journal of Laboratory Automation octubre de 2009 vol. 14 n° 5 303-307, "High-Throughput Screening for High Affinity Antibodies", S Tickle *et al* (conocida como la técnica UCB SLAM).

La clonación molecular implica la amplificación por PCR del ADNc (producido por RT-PCR del ARNm de POIs) con los cebadores que llevan sitios de clonación de enzimas de restricción. Esto produce productos de PCR en los que cada una de las secuencias de nucleótidos de POI está flanqueada por sitios de restricción. Los productos de PCR se digieren luego con las enzimas de restricción apropiadas y se subclonan en vectores de expresión vacíos que proporcionan un promotor y señal poliA por ligamiento, así como llevan un marcador de selección. Los productos ligados se transforman después en *E. coli* y los clones se seleccionan en cuanto a la presencia del marcador. Es necesario confirmar y seleccionar clones con inserciones de secuencia correctas mediante un laborioso mapeo de restricción y secuenciación de clones individuales. Cada uno de los vectores de expresión de clon correcto se purifica individualmente de *E. coli* y se transfecta en una célula huésped (p. ej., una célula de mamífero CHO o HEK293) para una expresión transitoria o estable.

Por lo tanto, la clonación molecular es una técnica laboriosa que implica múltiples etapas y no es adecuada para un rendimiento elevado.

La expresión transitoria, tal como la de un casete de expresión lineal, es habitualmente limitada y la expresión estable es preferible para la expresión a largo plazo. Para una expresión estable, típicamente la secuencia de interés codificante de POI se inserta en el genoma de los mamíferos mediante el método clásico de integración espontánea de ADN extraño (es decir, integración aleatoria en el genoma). Este enfoque conduce a menudo a una variación transcripcional significativa (y, por lo tanto, impredecibilidad y expresión de POI inconsistente) como resultado de diferencias en el número de copias transgénicas y el sitio de integración (véase Henikoff S (1992), "Position effect and related phenomena", Curr Opin Genet DEV2(6): 907-912; Martin DI, Whitelaw E (1996), "The variegating transgenes", Bioessays 18(11): 919-923; y Whitelaw E et al. (2001), "Epigenetic effects on transgene expression", Métodos Mol Biol. 158: 351-368). Además, los fragmentos transgénicos integrados de esta manera se encuentran a menudo insertados como concatémeros, lo que puede resultar en la inactivación génica mediante el silenciamiento génico inducido por repetición (véase Garrick D et al (1998), "Repeat-induced gene silencing in mammals", Nat Genet 18(1): 56-59; y McBurney MW 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 POIs y poblaciones celulares para expresar repertorios de POI de este tipo.

Existen muchas tecnologías para la generación de anticuerpos monoclonales (mAbs) a partir de animales humanos o transgénicos portadores de un repertorio de anticuerpos. Generalmente, los mAbs se obtienen a partir de la

- 5 inmortalización de células B, ya sea por fusión (tecnología de hibridoma) o transformación (transfección de virus o transformación de oncogén). Estos métodos de inmortalización celular, sin embargo, son inadecuados para un cribado exhaustivo de repertorios de anticuerpos grandes, porque son muy sesgados, ineficientes y típicamente solo muestran una pequeña proporción del repertorio disponible (típicamente menos del 0,1 % (células inmortalizadas/células de entrada) de un repertorio de células B obtenido de un ratón inmunizado, por ejemplo). El uso de métodos alternativos de cribado de células B que no requieren inmortalización, por lo tanto, es atractivo, pero las técnicas actualmente enfrentan los problemas arriba discutidos.
- 10 Duvall et al., ((2011) MAb. 3:203-208) describe métodos utilizados en la producción de anticuerpos humanos derivados de células B vírgenes humanas.
- DeKosky et al., ((2013) Nature biotechnology. 31:166-169) describe una secuenciación de alto rendimiento de pares VH:VL de linfocitos B humanos individuales.
- 15 Kurosaw A et al., ((2012) BMC Biology. 10:80) describe la producción de anticuerpos monoclonales específicos para el antígeno de una diversidad de animales.
- El documento WO2013/000982 (Vivalis) describe un método para el cribado de células.
- 20 El documento US2007/0184546 (Hudson Control Group, Inc.) describe una célula de trabajo automatizada capaz de realizar experimentos de laboratorio genéticos de principio a fin, moviendo muestras o microplacas entre los instrumentos para su análisis.
- 25 El documento WO2008/142028 (Boehringer Ingelheim RCV GMBH & CO) describe un método para producir una proteína recombinante a escala de fabricación.
- El documento US 7.884.054 (Amgen, Inc.) describe métodos para el cribado de anticuerpos multiméricos producidos por células de mamíferos.
- 30 El documento US 6.258.529 (OraVax, Inc.) describe métodos para aislar genes de la región variable de inmunoglobulina y el uso de estos genes en la producción de anticuerpos quiméricos y de isotipo conmutados.
- El documento US 6.919.183 (Regeneron Pharmaceuticals, Inc.) describe un método para identificar y aislar células que producen proteínas secretadas.
- 35 El documento US 2007/0141048 (Oleksiewicz) describe un método para enlazar secuencias de interés.
- Liao et al., ((2009) J Virol Methods. 158:171-179) describe el aislamiento de alto rendimiento de genes de inmunoglobulina de células B humanas únicas y la expresión como anticuerpos monoclonales.
- 40 Boria et al., ((2008) BMC immunology. 9:50) describe conjuntos de cebadores para la clonación del repertorio humano de regiones variables receptoras de células T.
- 45 Hogrefe y Shopes ((1994) Genome Res. 4: S109-S122) describe la construcción de colecciones de visualización de fagémidos.
- Smith et al., ((2009) Nat protoc. 4:372-384) describe un protocolo para la producción de anticuerpos monoclonales humanos específicos para antígenos.
- 50 Tiller et al., ((2008) J Immunol Methods. 329:112-124) describe la generación de anticuerpos monoclonales a partir de células B humanas individuales por RT-PCR de una sola célula y clonación de vectores de expresión.
- Franz et al., ((2011) Blood. 118:348-357) describe la caracterización ex vivo y el aislamiento de células B de memoria raras con tetrámeros antígenos.
- 55 Beerli et al., ((2008) PNAS. 105:14336-14341) describe métodos para el aislamiento de anticuerpos monoclonales humanos mediante la exhibición de células de mamíferos.
- 60 Tiller ((2011) New Biotechnology 28:453-457) es una revisión de tecnologías de células B individuales para el descubrimiento de anticuerpos.

## **SUMARIO**

- 65 La invención aborda la necesidad de técnicas de producción de vectores que sean susceptibles a una producción fiable y de alto rendimiento, adecuada para repertorios y el cribado, especialmente con potencial para la automatización.

La invención se recoge en las reivindicaciones. De acuerdo con la invención, las proteínas de interés (POI) a las que se hace referencia en esta memoria son anticuerpos.

**5    Una primera configuración de la presente divulgación proporciona:-**

En un primer aspecto,

Un método de producir células que codifican un repertorio de proteínas de interés (POI), comprendiendo el método:-

- a) proporcionar una población de células que expresen un repertorio de POIs;
- b) clasificar la población de células para producir una población clasificada de células individuales, comprendiendo cada una de las células ácido nucleico que codifica una POI respectiva;
- c) amplificar el ácido nucleico compuesto por la población de células individuales clasificada para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican POIs;
- d) modificar los ácidos nucleicos que codifican POI amplificados clasificados de la etapa (c) para producir un repertorio clasificado de casetes de expresión, comprendiendo cada uno de los casetes una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI; y
- e) transferir casetes de expresión de POI de dicho repertorio de casetes a una población clasificada de células huésped al tiempo que se mantiene la clasificación de casetes de expresión de POI y producir un repertorio clasificado de células huésped que expresa un repertorio clasificado de POIs.

En una realización, la clasificación se mantiene en la etapa (e) mediante la provisión de repertorios en una pluralidad de recipientes, cuyas ubicaciones relativas y disposiciones generales son fijas. Esto permite el procesamiento de alto rendimiento y la automatización de los métodos de invención y divulgación, p. ej., para un cribado celular eficiente y rápido para seleccionar una o más secuencias de POI de interés. Adicional o alternativamente, la transferencia de los casetes de expresión en la etapa (e) se lleva a cabo mediante transferencia por lotes, y esto también permite el procesamiento de alto rendimiento y la automatización del método de la invención.

En un segundo aspecto de la divulgación,

un aparato automatizado para realizar el método de cualquier aspecto o configuración precedente, comprendiendo el aparato

- a. medios para mantener una población de células únicas clasificadas en una pluralidad de recipientes, en donde cada una de las células se encuentra en un recipiente respectivo, comprendiendo cada una de las células ácido nucleico que codifica una POI respectiva;
- b. medios para suministrar reactivos de PCR a los recipientes para amplificar el ácido nucleico compuesto por la población de células individuales clasificadas para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican POIs;
- c. medios para entregar a los recipientes reactivos para modificar ácidos nucleicos que codifican POI amplificados clasificados para producir un repertorio clasificado de casetes de expresión, comprendiendo cada uno de los casetes una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI;
- d. medios para mantener una población clasificada de células huésped en una pluralidad de recipientes;
- e. medios para transferir casetes de expresión de POI de dicho repertorio de casetes a la población clasificada de células huésped en los recipientes mientras se mantiene la clasificación de casetes de expresión de POI; y
- f. medios para llevar a cabo la transfección de casetes de expresión en células huésped en los recipientes para producir un repertorio clasificado de células huésped que expresa un repertorio clasificado de POIs.

**Una segunda configuración de la presente divulgación proporciona:-**

En un primer aspecto,

un casete de expresión para la expresión de una POI en una célula huésped, siendo proporcionado el casete por el ácido nucleico lineal que comprende un transposón, comprendiendo el transposón elementos de transposón 5'- y 3'-

terminales con una secuencia de nucleótidos que codifica POI y elemento(s) regulador(es) para la expresión de POI entre los elementos transposón.

En un segundo aspecto,

una población clasificada de casetes de expresión que codifica un repertorio de POIs correspondientes a POIs expresadas por una población de células, comprendiendo cada uno de los casetes una secuencia de nucleótidos que codifica un miembro del repertorio de POIs y uno o más elementos reguladores para la expresión de POI, en donde cada uno de dichos casetes comprende la disposición (en dirección 5' a 3'): elemento de transposón - [secuencia de nucleótidos POI y elemento(s) regulador(es)] - elemento de transposón, y casetes de expresión para la expresión de POIs de diferentes células están aislados entre sí en la población clasificada (p. ej., en diferentes pocillos de una placa).

En un tercer aspecto,

un método de hacer un transposón que comprende una secuencia de nucleótidos de interés (NOI, por sus siglas en inglés), comprendiendo el método

- a. proporcionar una primera secuencia de nucleótidos que comprende (en dirección 5' a 3') A, B y C, en donde A es una primera secuencia de homología, B es una secuencia de nucleótidos que comprende los NOI y C es una segunda secuencia de homología;
- b. proporcionar una primera secuencia de nucleótidos del molde que comprende (en dirección 5' a 3') W y X, en donde 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 del molde que comprende (en dirección 5' a 3') Y y Z, en donde Y es una cuarta secuencia de homología y Z es una secuencia de nucleótidos que comprende un segundo elemento de transposón; y cualquiera

4. d.

(i) mezclar la primera secuencia de nucleótidos con el primer molde para hibridar los brazos de homología primero y tercero juntos y llevar a cabo una amplificación y extensión de ácidos nucleicos para extender la primera secuencia de nucleótidos utilizando el primer molde para producir una primera secuencia de nucleótidos extendida (primera ENS, por sus siglas en inglés) que comprende (en dirección 5' a 3') W, B y C; y (ii) mezclar la primera ENS con el segundo molde para hibridar los brazos de homología segundo y cuarto juntos y llevar a cabo la amplificación y extensión del ácido nucleico para extender la primera ENS para producir una segunda ENS que comprende ((en dirección 5' a 3') W, B y Z; o

(ii) mezclar la primera secuencia de nucleótidos con el segundo molde para hibridar los brazos de homología segundo y cuarto juntos y llevar a cabo la amplificación y extensión del ácido nucleico para extender la primera secuencia de nucleótidos utilizando el segundo molde para producir una tercera secuencia de nucleótidos extendida (tercera ENS) que comprende (en dirección 5' a 3') A, B y Z; y (ii) mezclar la tercera ENS con el primer molde para hibridar los brazos de homología primero y tercero juntos y llevar a cabo la amplificación y extensión del ácido nucleico para extender la tercera ENS para producir una cuarta ENS que comprende ((en dirección 5' a 3') W, B y Z; o

(iii) mezclar la primera secuencia de nucleótidos con el primer y el segundo moldes para hibridar los brazos de homología primero y tercero juntos e hibridar los brazos de homología segundo y cuarto juntos y llevar a cabo amplificación y extensión de ácido nucleico para extender la primera secuencia de nucleótidos utilizando el segundo molde para producir una quinta ENS que comprende ((en dirección 5' a 3') W, B y Z; y

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

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

#### **Una tercera configuración de la presente divulgación proporciona:-**

Un método de producir una célula huésped para la expresión de una POI, comprendiendo el método

a. proporcionar al menos primero y segundo casetes de expresión, en donde cada uno de los casetes de expresión comprende

i. un primer elemento de integración y un segundo elemento de integración en posición 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 utilizando una enzima de integración;

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

c. introducir simultánea o secuencialmente los primero y segundo casetes de expresión en la célula huésped, en donde cada uno de los casetes está integrado genómicamente en el genoma de la célula huésped en dicho motivo para la expresión de POIs por la célula huésped; y

d. producir opcionalmente una línea celular de expresión de POIs en una etapa que comprende cultivar la célula huésped.

#### **Una cuarta configuración de la presente divulgación proporciona:-**

Una mezcla de ácido nucleico que comprende un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en donde el primer ácido nucleico es capaz de hibridarse a una secuencia 5'UTR de la región V de anticuerpos humanos de un gen compuesto por un ácido nucleico diana, en donde el gen codifica una región V humana; y el segundo ácido nucleico es capaz de hibridarse a una segunda secuencia, en donde la segunda secuencia está compuesta por el ácido nucleico diana y está en posición 3' a la secuencia UTR, en donde 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 NOs: 1-47.

Una mezcla de ácido nucleico que comprende un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en donde los ácidos nucleicos son diferentes y se seleccionan de los ácidos nucleicos que comprenden una secuencia que es al menos 90 % idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 1-47.

Un método de amplificación de un repertorio de secuencias de regiones variables humanas utilizando una o más de las secuencias.

#### **BREVE DESCRIPCIÓN DE LAS FIGURAS**

La **Figura 1A** es un esquema que ilustra un ejemplo no limitante del método de invención y divulgación para producir un repertorio de sitios de unión de anticuerpos (POIs) en células huésped (p. ej., células CHO o HEK293).

La **Figura 1B** es un esquema alternativo que ilustra un ejemplo no limitante del método de la invención y divulgación para producir un repertorio de sitios de unión de anticuerpos (POIs) en células huésped (p. ej., células CHO o HEK293).

**Figura 2:** Ejemplos de estrategia de selección para diferentes poblaciones de células B. (A) Un ejemplo de un gráfico de contorno citométrico de flujo que muestra la selección de la población de células de memoria CD38+ CD95+ después de la exclusión de células B IgM e IgD. (B) cada una de las células de memoria individuales positivas para CD19 (Pacific Blue) y antígeno (Ovalbumin-AlexaFluor-488) se clasifica luego en pocillos separados en una placa de 96 pocillos.

**Figura 3:** Un diagrama de flujo que resume la producción de anticuerpos de alto rendimiento a partir de células B individuales. La síntesis de ADNc basada en células individuales se realizó utilizando cebadores específicos para la región constante. Se utilizó una mezcla de cebadores específicos para el gen V para cadena pesada y cadena ligera con un fragmento promotor del citomegalovirus humano (hCMV) en el extremo 5' para la primera ronda de PCR para amplificar el fragmento del gen V. Se utilizó un cebador directo genérico que reasoció la etiqueta hCMV con un cebador anidado inverso para la región constante para la segunda ronda de PCR. Los productos amplificados se puentearon entonces con el casete de Ig lineal con promotor de LTR-CMV 5' PB y señal-3' LTR PB de la región-poliA constante.

**Figura 4:** Producción de construcciones de expresión de cadena pesada y ligera mediante PCR puente. A) Geles de agarosa teñidos con bromuro de etidio de los pares de fragmentos de genes V<sub>H</sub> y V<sub>L</sub> amplificados

a partir de células B individuales. Cada una de las pistas contiene 10  $\mu$ L de un producto de PCR de segunda ronda de 30  $\mu$ L  $V_H+V_L$ . B) Geles de agarosa teñidos con bromuro de etidio que muestran los resultados de la PCR puente. Cada una de las pistas contiene 10  $\mu$ L de un producto de puente de 30  $\mu$ L para la expresión de Ig. Los controles negativos mostraron el casete de Ig de cadenas pesada y ligera. Aquí se mostraron productos representativos de la PCR de la plataforma de PCR de 96 pocillos de alto rendimiento.

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

**Figura 6:** Un ejemplo de familia agrupada que mostró la maduración de afinidad *vía* hipermutación tanto para la afinidad aparente como para la potencia de neutralización. CNROR: incapaz de resolver la tasa de desactivación.

**Figura 7:** Nivel de expresión impulsado por el sistema de transposición PiggyBac. La expresión de anticuerpos de los productos de la PCR puente transfectados en células HEK293 se testó utilizando la co-transfección con diferentes cantidades de PBase: 0; 20; 100; 500 ng/pocillo. Los sobrenadantes en cada uno de los pocillos se recogieron el día 0; el día 2; y el día 5. Los datos mostraron que el sistema de transposón aumentó el nivel de expresión en 2-4 veces el día 5.

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

**Figuras 9A y B:** Sensorgramas de SPR representativos de anticuerpos que se unen a la ovoalbúmina. Alrededor de dos tercios de los anticuerpos testados mostraron evidencia de unión al antígeno, con una gama diversa de afinidades de unión y cinética.

**Figura 10:** La afinidad aparente de anticuerpos contra dos antígenos diana diferentes. Se detectó una gama de aglutinantes (círculos en blanco) así como neutralizadores funcionales (círculos en negro), con la mayor afinidad detectada en el intervalo picomolar. Este experimento valida la tecnología de clonación de células B individuales de los autores de la invención para ser una herramienta poderosa en la identificación y recuperación de anticuerpos de alta afinidad y funcionalmente competentes.

## DESCRIPCIÓN DETALLADA

La Figura 1 es un esquema que ilustra un ejemplo del método de la invención o divulgación para producir un repertorio de sitios de unión de anticuerpos (POIs) en células huésped (p. ej., células CHO o HEK293), que es útil en un procedimiento de cribado de células B. El repertorio de células huésped producido al final del esquema en la Figura 1 es útil, ya que las células expresan sitios de unión de anticuerpos (p. ej., pares  $V_H/V_L$ ) que pueden ser cribados contra un antígeno predeterminado para identificar qué células expresan sitios de unión de interés. Esto permite, por lo tanto, el enlace genotipo a fenotipo, permitiendo la identificación de la secuencia de nucleótidos que codifica un sitio de unión de interés. Esta secuencia de nucleótidos se puede utilizar entonces para construir líneas celulares para la producción estable del anticuerpo deseado, p. ej., para producir productos farmacéuticos de anticuerpos o reactivos de diagnóstico o de investigación que fijan como objetivo el antígeno predeterminado. En un ejemplo, la célula huésped producida al final del esquema de la Figura 1 es en sí misma una célula que expresa de manera estable el sitio de unión y esto ofrece la posibilidad de una generación de línea celular estable conveniente y eficiente para la fabricación a largo plazo del sitio de unión. Como se discute más adelante, en un ejemplo, la inserción mediada por el sitio de los casetes de expresión codificantes de POI se puede realizar en células huésped, lo que proporciona la ventaja añadida de la expresión de POI estable que evita la integración aleatoria y los concatémeros.

Si bien se proporcionan ejemplos en términos de sitios de unión de anticuerpos, dominios variables de anticuerpos y cadenas, la divulgación no se limita a proteínas de este tipo. En este sentido, la divulgación 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 fragmentos o dominios de proteínas o polipéptidos.

Volviendo al ejemplo no limitante de la Figura 1, en una primera etapa, se proporciona una población deseada de células B, que pueden ser células germinales, células de memoria, plasmablastos, células plasmáticas o generalmente pueden ser células secretoras de anticuerpos (ASCs, por sus siglas en inglés), o mezclas de dos o más de estos tipos de células B. La persona calificada está familiarizada con la selección de poblaciones de este tipo, por ejemplo, utilizando marcadores de la superficie celular opcionalmente con FACS. Marcadores de este tipo pueden seleccionarse de CD19, IgM, IgD, CD30 y CD95, por ejemplo, y pueden incluir un panel de uno o más de estos (p. ej. 1, 2, 3, 4), o pueden incluir un panel de todos estos marcadores. Las células que portan estos marcadores se pueden teñir, por ejemplo, con cualquier fluoróforo de moléculas pequeñas que pueda ser detectado por el sistema de clasificación

celular, tal como Alexa-488, Alexa-647, Pacific Blue, R-ficoeritrina, isotiocianato de fluoresceína o alofococianina, conjugado opcionalmente a un colorante de cianina, p. ej., Cy7. La población de células B puede generarse en una etapa preliminar por aislamiento de uno o más animales (p. ej., ratones, ratas o seres humanos o animales no humanos), por ejemplo, un animal que ha sido inmunizado con un antígeno diana. Por lo tanto, los métodos de la invención y la divulgación son útiles para cribar poblaciones de células B para identificar una o más secuencias de POI derivadas de una célula B, en donde la POI une el antígeno diana con una característica deseada. Una característica de este tipo es, p. ej., unirse al antígeno diana y/o a un antígeno homólogo u ortólogo estructuralmente relacionado (p. ej., uno de una especie diferente); y/o unirse a un antígeno con una afinidad deseada; y/o competir con un anticuerpo conocido para unirse a un antígeno predeterminado (p. ej., el antígeno diana); y/o unirse a un epítipo predeterminado.

Opcionalmente, la población de células B de entrada deseada (p. ej., plasmablastos o células plasmáticas) puede seleccionarse realizando la clasificación de células específicas para el antígeno utilizando técnicas conocidas en la técnica (p. ej., 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 encontrado que es ventajoso realizar esta etapa para agilizar el proceso global de cribado de células B, lo que permite un cribado de rendimiento eficiente y más alto.

Por lo general, las células GC (centro germinal) específicas para el antígeno o B de memoria pueden ser capturadas por el antígeno marcado, porque expresan predominantemente anticuerpos de la transmembrana en la superficie celular. El antígeno puede ser marcado de manera fluorescente (por ejemplo, cualquier fluoróforo de molécula pequeña que pueda ser detectado por el sistema de clasificación celular, tal como Alexa-488, Alexa-647, Pacific Blue, R-ficoeritrina, isotiocianato de fluoresceína o alofococianina conjugada opcionalmente a un colorante de cianina, p. ej., Cy7). Si la población de células B ha sido previamente teñida para la selección inicial, es ventajoso utilizar un fluoróforo que tenga una longitud de onda de emisión diferente al fluoróforo utilizado en esa selección inicial, con el fin de facilitar el proceso de clasificación, p. ej., utilizando FACS. En una alternativa, la célula puede teñirse simultáneamente para detectar la presencia del marcador de selección celular y cribarse para unirse a las VLPs portadoras de antígeno/antígeno fluorescentes. Sin limitarse a la teoría, se piensa que, por otro lado, plasmablastos o células plasmáticas no serían tan fácilmente capturados por el antígeno marcado, debido a su expresión dominante de anticuerpo secretor.

En una realización alternativa, el antígeno puede ser capturado con partículas similares a virus (VLPs, por sus siglas en inglés) con antígeno recombinante en la superficie. Las VLPs pueden generarse a partir de células CHO, células KEK, fibroblastos embrionarios de ratón (MEFs, por sus siglas en inglés) u otras líneas celulares de mamíferos con co-expresión del antígeno recombinante, la proteína gag del retrovirus y MA-GFP (proteína de fusión p15-GFP del fragmento de matriz gag). La expresión de la proteína gag permite que la VLP emerja de las células, y la MA-GFP marca las VLPs para la detección de fluorescencia. Tanto las proteínas gag como MA-GFP están asociadas con la superficie interna de la membrana plasmática, y el antígeno recombinante está en la superficie de la VLP. Los antígenos en las VLPs se presentan en forma nativa, expresados directamente a partir de células recombinantes sin ninguna etapa de purificación o modificación. La forma nativa de un antígeno debe proporcionar todos los epítopos naturales que ayudan en gran medida a la selección de anticuerpos neutralizantes. La alta densidad del antígeno en las VLPs aumenta la relación señal/ruido para la detección de células que expresan anticuerpos específicos para el antígeno en la superficie celular y facilita en gran medida la etapa de clasificación. Las VLPs recombinantes se pueden generar con la expresión de diferentes proteínas fluorescentes, tales como MA-CFP o MA-YFP. Utilizando la multiplexación de VLPs con diferentes antígenos y diferentes proteínas de fluorescencia, se pueden seleccionar células que expresan aglutinantes de alta afinidad, aglutinantes de reactividad cruzada o aglutinantes específicos de homólogos. Las células que expresan aglutinantes de alta afinidad se pueden seleccionar por células con matriz de afinidad relativa alta (matriz de afinidad = la relación de la actividad de VLPs de unión a antígeno de baja densidad a antígeno de alta densidad). Las células que expresan aglutinantes de reactividad cruzada a ortólogos o diferentes antígenos (para el aislamiento de anticuerpos biespecíficos 2 en 1) se pueden seleccionar por células que se unen a diferentes tipos de VLPs al mismo tiempo. Las células que expresan aglutinantes específicos del homólogo también se pueden seleccionar por la célula que solo se une al antígeno específico, pero no a su homólogo.

Las células B se clasifican por lo tanto (p. ej., utilizando FACS) para proporcionar una población de células B individuales y clasificadas. Típicamente, las células B se clasifican en pocillos de una placa estándar (p. ej., una placa estándar de 96 pocillos o 384 pocillos) de modo que cada una de las células se encuentra en un pocillo respectivo y no se mezcla con otra célula. Es posible que haya un número mínimo (p. ej., menos del 5 %, menos del 3 %, menos del 2 % o menos del 1 %, o niveles no detectables) de pocillos que no tengan ni más de una (p. ej., dos) células, y esto no obstaculiza la utilidad global del método de cribado (y no se considera parte del repertorio deseado). Preferiblemente, cada uno de los pocillos en la placa contiene una sola célula B. Opcionalmente, es posible cultivar células directamente antes y/o después de la etapa de clasificación de células, pero los inventores han encontrado que esto no es necesario a diferencia de las técnicas en la técnica. Por lo tanto, al evitar esta etapa, los métodos de la invención y la divulgación se prestan más por sí mismos a la racionalización y el alto rendimiento.

A continuación, en el ejemplo mostrado en la Figura 1, se amplifica el ácido nucleico codificante de POI. En el ejemplo, esto se realiza mediante la transcripción inversa del ARNm en células de la población de células clasificadas (es decir,



el ARNm codificante de POI se convierte en el correspondiente ADNc codificante de POI) y esto se amplifica utilizando la PCR. La RT-PCR estándar se puede utilizar como le resulte familiar a la persona calificada (p. ej., 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, las secuencias tanto V<sub>H</sub> como V<sub>L</sub> se copian y amplifican para cada una de las células clasificadas. La persona experta sabrá que cada una de las células expresa un solo tipo de sitio de unión de anticuerpos, por lo tanto, solo un solo tipo de secuencia V<sub>H</sub> amplificada y un solo tipo de secuencia V<sub>L</sub> serán amplificados por cada célula individual clasificada (por lo tanto, en los casos en los que se utilizan pocillos, solo se obtendrá un solo tipo V<sub>H</sub> y V<sub>L</sub> por cada pocillo, que por lo tanto conserva la agrupación de las secuencias V<sub>H</sub> y V<sub>L</sub> formando sitios de unión).

Las secuencias codificantes de POI también se modifican para producir un repertorio de casetes de expresión para expresar POIs de células huésped en una etapa posterior del método. Los casetes de expresión contienen una secuencia de nucleótidos que codifica POI y uno o más elementos reguladores para la expresión (p. ej., un promotor y/o un potenciador y/o un poliA). En una realización, las etapas de amplificación y modificación se pueden realizar simultáneamente utilizando PCR y moldes y cebadores apropiados como resultará evidente para la persona calificada. En otra realización, la amplificación y la modificación se llevan a cabo en etapas separadas, p. ej., amplificación y luego modificación. En el ejemplo de la Figura 1, primero se lleva a cabo una RT-PCR y luego se modifican las secuencias codificantes de POI amplificadas para producir casetes de expresión que comprenden cada uno (en dirección 5' a 3'): promotor-secuencia de nucleótidos de POI-poliA. En este ejemplo, también se agregan elementos de transposón flanqueantes. Los elementos de transposón pueden ser elementos de repetición terminales invertidos de transposón piggyBac (PB) para formar un ácido nucleico de transposón que comprende la estructura: [5' elemento PB] - [promotor] - [secuencia de nucleótidos POI] - [poliA] - [3' elemento PB]. En un ejemplo, cada uno de los casetes de expresión es proporcionado por ADN lineal que comprende o consiste en el transposón.

**elemento PB 5':**

```
GATATCTATAACAAGAAAATATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATT
ATCGTATGAGTTAAATCTTAAAAGTCACGTAAAAGATAATCATGCGTCATTTTGACTCACGCGGTCGTTA
TAGTTCAAATCAGTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGA
TGTCCTAAATGCACAGCGACGGATTTCGCGCTATTTAGAAAGAGAGCAATATTTCAAGAATGCATGCGT
CAATTTTACGCAGACTATCTTTCTAGGGTTAA
```

**elemento PB 3':**

```
TTTGTTACTTTATAGAAGAAATTTTGAGTTTTTGTTTTTTTTAATAAATAAATAAACATAAATAAATTGT
TTGTTGAATTTATTATTAGTATGTAAGTGTAATATAATAAACTTAATATCTATTCAAATTAATAAATAA
ACCTCGATATACAGACCGATAAAACATGCGTCAATTTTACGCATGATTATCTTTAACGTACGTACAAT
ATGATTATCTTTCTAGGGTTAA
```

Como se muestra en el ejemplo de la figura 1, la PCR puente (véase, p. ej., Mehta R.K., Sihgh J., Biotechniques, 1999, 26(6): 1082-1086) se puede utilizar para construir el transposón añadiendo los elementos de transposón flanqueantes.

En esta etapa del método, se ha obtenido un repertorio de casetes de expresión que codifican un repertorio de secuencias de nucleótidos codificantes de POI derivadas de la población de entrada de células. En este ejemplo, los procedimientos de amplificación y modificación se llevan a cabo en las células individuales que se clasifican en pocillos en una o más placas. Por lo tanto, el resultado es un repertorio clasificado de casetes de expresión (en este caso por pocillo individual, el pocillo contiene una pluralidad de un tipo de casete de expresión V<sub>H</sub> y un tipo de casete de expresión V<sub>L</sub> derivado de un tipo de sitio de unión de anticuerpos de una célula de entrada individual). A continuación, los casetes de expresión clasificados (en el presente ejemplo, contenidos en transposones respectivos y opcionalmente como ADNs lineales) se transfieren a las células huésped para producir un repertorio clasificado de células huésped que expresa un repertorio clasificado de POIs. Esto se puede realizar, en el presente ejemplo, transfiriendo casetes de expresión de pocillos en una placa (o un conjunto de placas) a una o más placas que tienen una pluralidad de células que contienen células huésped (p. ej., tipos idénticos de células de la misma línea celular, p. ej., células CHO o HEK293 o de levadura). En esta realización, el repertorio de casetes clasificados se puede transferir manualmente o con un robot utilizando una pipeta multicanal (p. ej., una pipeta de 4, 8, 12, 16 (2 x 8), 64 (8 x 8), 96 (12 x 8), 384 o 1536 canales como se conoce en la técnica) para recuperar simultáneamente casetes (en este caso, pares de casetes V<sub>H</sub>/V<sub>L</sub> clasificados) de pocillos individuales en una primera placa y transferirlos a pocillos correspondientes en una segunda placa, en donde esos pocillos contienen células huésped. Al hacer esto, la ubicación relativa y la naturaleza clasificada de los casetes de expresión se mantienen cuando se transfieren a células huésped. Esto tiene la ventaja de mantener el enlace de los pares de secuencias de casete V<sub>H</sub>/V<sub>L</sub> sin mezclar (es decir, por pocillo en la segunda placa, el pocillo contiene células huésped y secuencias V<sub>H</sub> y V<sub>L</sub> derivadas de una sola célula B de entrada solamente). Ventajosamente, también si se deja alguna muestra de casete en la primera placa, se puede

identificar fácilmente uno o más pocillos como una fuente de secuencia útil codificante de POI (y casete de expresión y transposón) por referencia a los positivos deseados encontrados en la segunda placa después del cribado de la segunda placa para una característica deseada del sitio de unión de anticuerpos. Además, ventajosamente se puede proporcionar a las células huésped en un medio que apoya deseablemente el funcionamiento y el mantenimiento de las células huésped (típicamente diferente del entorno utilizado en las células de la primera placa para la amplificación y modificación de la secuencia). En una realización alternativa, las células huésped se añaden a pocillos de la placa que contienen el repertorio de casetes de expresión para que se retenga la disposición clasificada (pero esto no proporciona las ventajas adicionales de la otra realización, en donde una placa maestra de casetes de expresión se conserva sin mezcla con células huésped - útil, por ejemplo, para llevar a cabo una segunda transferencia a otra placa que tenga un tipo diferente de célula huésped, como cuando uno quiere evaluar el rendimiento de la expresión de POI y la visualización/secreción por un tipo de célula huésped diferente).

Un ejemplo de una pipeta multicanal automatizada es el sistema automatizado de manipulación de líquidos Thermo Scientific Matrix Hydra II de 96 canales. V&P Scientific VP 177AD-1 y VP 179BJD son colectores de dosificación diseñados para el llenado rápido de placas de 96 y 384 pocillos, respectivamente, y cualquiera de ellos puede utilizarse en el método de la invención y la divulgación para transferir casetes a células huésped o para la manipulación general de muestras.

La salida del método de la invención o la divulgación, por lo tanto, en su aspecto más amplio, es la producción de un repertorio de células huésped que expresa un repertorio de POIs. Como se discutió anteriormente, esto puede utilizarse en una etapa posterior de cribado (p. ej., utilizando un ensayo de unión celular, ELISA, resonancia de plasmón de superficie u otro ensayo apropiado para la naturaleza particular de la POI) para identificar una o más células huésped que expresan una POI con una característica deseada. Se es entonces capaz de aislar POI de la célula o medio circundante y/o aislar (y opcionalmente replicar o amplificar) una secuencia de nucleótidos (p. ej., secuencia de ADN) codificando la POI deseada. Se puede determinar la secuencia de nucleótidos. Células aisladas de este tipo se pueden cultivar para producir una línea celular codificante de POI (lo cual es especialmente útil cuando el casete se ha integrado de forma estable en el genoma de la célula, como es posible con la integración genómica dirigida al sitio, p. ej., realizada utilizando un transposón). La secuencia de nucleótidos codificantes de POI se puede insertar en un vector de expresión diferente y/o mutar (p. ej., madurar por afinidad) o condensarse a otra secuencia de proteína.

La integración en el transposón se efectúa proporcionando una enzima transposasa correspondiente en las células huésped (p. ej., mediante la co-transfección de casetes de expresión con un vector que comprende un gen transposasa expresable o proporcionando células huésped que albergan un gen transposasa de este tipo, p. ej., de manera inducible). En un ejemplo, cada uno de los casetes comprende elementos de transposón piggyBac y el método utiliza una transposasa piggyBac (p. ej., transposasa piggyBac de tipo salvaje o hiperactiva; véase, p. ej., 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).

**Pbase WT:**

MGCSLDDEHILSALLQSDDELVGEDSDSEISDHVSEDDVQSDTEEFIDEVHEVQPTSSGSEILDEQNVIEQ  
PGSSLASNRILTLPQRTIRGKNKHCWSTSKSTRSRVSALNIVRSQRPTRMCRNIYDPLLCFKLFFTDIIESE  
IVKWTNAEISLKRRESMTGATFRDTNEDEIYAFFGILVMTAVRKDNHMSTDDLFDRLSMVYVSVMSRDRF  
DFLIRCLRMDDKSIRPTLRENDVFTPVRKIWDLFIHQCIQNYTPGAHLTIDEQLLGFRGRCPFRMYIPNKP SK  
YGIKILMMCDSGTKYMINGMPYLGRGTQNGVPLGEYVKELSKPVHGSCRNITCDNWFTSIPLAKNLLQEP  
YKLTIVGTVRSNKR EIPVLKNSR SRPVGTSMFCFDGPLTLVSYKPKPAKMVYLLSSCDEEDASINESTGKPKM  
VMYYNQTKGGVDTLDQMCSVMTCSRKTNRWPMALLYGMINIACINSFIIYSHNVSSKGEKVQSRKKFMRNL  
YMSLTSSFMRKRLEAPTLKRYLRDNISNILPNEVPGTSDDSTEPMKKRTYCTYCP SKIRRKANASCKKCKK  
VICREHNIDMCQSCF

**PBase hiperactiva:**

MGSSLDDEHILSALLQSDDELVGEDSDSEVSDHVEDDVQSDTEEFIDEVHEVQPTSSGSEILDEQNVIEQ  
 PGSSLASNRILTLPQRTIRGKNKHCWSTSKPTRRSRVSALNIVRSQRPTRMCRNIYDPLLCKLFFFTDEIIE  
 IVKWTNAEISLKRRESMTSATFRDTNEDEIYAFFGILVMTAVRKDNHMSTDDLFDRLSLMVVVSMSRDRF  
 DFLIRCLRMDDKSIRPTLRENDVFTPVRKIWDLFIHQIQNYTPGAHLTIDEQLLGFRGRCPFRVYIPNKPSK  
 YGIKILMMCDSGTKYMINGMPYLGRGTQTNGVPLGEYVVKELSKPVHGSCRNITCDNWFTSIPLAKNLLQEP  
 YKLTIVGTVRSNKREIPEVLKNSRSPVGTSMFCFDGPLTLVSYKPKPAKMVYLLSSCEDEDASINESTGKPQM  
 VMYYNQTKGGVDTLDQMCSVMTCSRKTNRWPMALLYGMINIACINSFIIYSHNVSSKGEKVQSRKKFMRNL  
 YMGLTSSFMKRLEAPTLKRYLRDNISNILPKEVPGTSDDSTEPMKKRTYCTYCPSKIRRKASASCKCKK  
 VICREHNIDMCQSCF

Quando se utiliza un transposón para la integración de casetes en la célula huésped, existen beneficios adicionales. En primer lugar, los transposones pueden integrarse en varias copias en las células huésped, proporcionando con ello casetes de expresión multicopia para soportar la expresión de POI de alto nivel. Esto se puede fomentar adicionalmente utilizando una enzima transposasa hiperactiva. Adicionalmente, los transposones pueden integrarse preferentemente en sitios que son activos para la transcripción, favoreciendo con ello la expresión de POI de alto nivel y eficiente, como se demuestra con el análisis de mapeo de sitios de integración. Esto se ve, 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).

Por lo tanto, la divulgación proporciona los siguientes conceptos:-

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

- a) proporcionar una población de células que expresen un repertorio de POIs;
- b) clasificar la población de células para producir una población clasificada de células individuales, comprendiendo cada una de las células ácido nucleico que codifica una POI respectiva;
- c) amplificar el ácido nucleico compuesto por la población de células individuales clasificada para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican POIs;
- d) modificar los ácidos nucleicos que codifican POI amplificados clasificados de la etapa (c) para producir un repertorio clasificado de casetes de expresión, comprendiendo cada una de las casetes una secuencia de nucleótidos que codifica una POI y uno o más elementos reguladores para expresar la POI; y
- e) transferir casetes de expresión de POI de dicho repertorio de casetes a una población clasificada de células huésped al tiempo que se mantiene la clasificación de casetes de expresión de POI y producir un repertorio clasificado de células huésped que expresa un repertorio clasificado de POIs.

Las etapas (c) y (d) pueden llevarse a cabo por separado (en cualquier orden, p. ej., (c) y (d)) o simultáneamente.

Opcionalmente, la etapa (c) comprende la amplificación por PCR de secuencias codificantes de POI, p. ej., RT-PCR utilizando ARNm codificantes de POI como molde (p. ej., uno, dos, tres o más cebadores que comprenden o consisten en una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 1-53, como se comenta más adelante). Opcionalmente, el repertorio de ácidos nucleicos amplificados en la etapa (c) son ADN. En un ejemplo, cada una de las POI es un dominio variable de anticuerpo y la etapa (c) comprende la amplificación por PCR de secuencias codificantes de POI utilizando uno o más cebadores 5' específicos para la región V y/o uno o más cebadores 3' de la región C (p. ej., CY, p. ej., el cebador CY de ratón). Opcionalmente, la PCR comprende secuencias de nucleótidos codificantes de POI 5'- y/o 3'-RACE. En un ejemplo, la 5'-RACE se lleva a cabo utilizando uno o más cebadores 5' cada uno homólogo a una UTR 5' o secuencia promotora de una región variable de anticuerpos. En un ejemplo, la 3'-RACE se lleva a cabo utilizando uno o más cebadores 5' cada uno homólogo a una región constante de anticuerpos, p. ej., una secuencia CH1 o Fc. En este caso, cada una de las secuencias codificantes de POI amplificada codifica una cadena de anticuerpos que comprende un dominio variable de anticuerpo y una región constante. En un ejemplo, la 3'-RACE utiliza una o más secuencias de regiones constantes humanas como un cebador; esto produce entonces secuencias que codifican regiones variables humanizadas en las que cada una de las regiones variables se condensa con una región constante humana (p. ej., una gamma CH1 o Fc humana (p. ej., gamma Fc)), proporcionando con ello

una cadena de anticuerpos humanos (POI) en la expresión posterior. Esta humanización durante la etapa (c) es útil, ya que las POIs identificadas a partir de un cribado posterior representan cadenas humanas que pueden utilizarse para producir agentes terapéuticos de anticuerpos para uso humano. En un ejemplo, la 5'-RACE utiliza un molde 5' que comprende un promotor de región variable para producir ácidos nucleicos amplificados que comprenden (en dirección 5' a 3'): un promotor y una secuencia de nucleótidos que codifican una POI. Adicional o alternativamente, se utiliza 3'-RACE, en donde la RACE utiliza una molde 3' que comprende una secuencia poliA para producir ácidos nucleicos amplificados que comprenden (en 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 casetes de expresión se pueden llevar a cabo simultáneamente (es decir, las etapas (c) y (d) se llevan a cabo simultáneamente).

En un ejemplo, la modificación de la etapa (d) se lleva a cabo utilizando la PCR, p. ej., la PCR puente. Por ejemplo, la etapa (d) se lleva a cabo después de la etapa (c), p. ej., después de la amplificación por RT-PCR o RACE. En este caso, la PCR puente se lleva a cabo en una etapa que comprende la hibridación de un primer cebador al extremo 5' de los productos de ácido nucleico de la etapa (c); y la hibridación de 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 utilizando el primer cebador). Alternativamente, el segundo cebador se puede utilizar inicialmente (para hibridar con el producto de la etapa (c)) y el producto de éste se puede hibridar con el primer cebador. Alternativamente, los primero y segundo cebadores y el producto de la etapa (c) pueden mezclarse juntos y se puede llevar a cabo una PCR. El resultado en cualquier caso es un producto extendido que comprende (en orden de 5' a 3'):

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

En una realización, el promotor y poliA se combinan con la secuencia de nucleótidos codificante de POI por la etapa (c), como se describe arriba (p. ej., utilizando 5'- y 3'-RACE con los cebadores apropiados). En otra realización, el primer cebador utilizado en la etapa (d) comprende una secuencia de promotor (p. ej., una secuencia en el extremo 3' del cebador). El resultado de la etapa (d) combina entonces el promotor con la secuencia de nucleótidos codificante de POI. Adicional o alternativamente, el segundo cebador utilizado en la etapa (d) comprende una secuencia poliA (p. ej., una secuencia en el extremo 5' del cebador). El resultado de la etapa (d) combina entonces poliA con la secuencia de nucleótidos codificante de POI. Son posibles otras combinaciones, p. ej., el promotor se añade en la etapa (c) utilizando el cebador apropiado y la poliA se añade en la etapa (d) utilizando el segundo cebador apropiado. En un ejemplo, la etapa (c) (p. ej., RACE) añade secuencias 5' y/o 3' en los ácidos nucleicos del producto que se pueden utilizar para la hibridación con cebadores en una etapa (d) en donde esta última utiliza la PCR puente.

El resultado de las etapas (c) y (d) es siempre un repertorio de casetes de expresión para la expresión de un repertorio de POIs. En un ejemplo, uno o más elementos reguladores requeridos o deseados para la expresión (o expresión óptima) se omiten de cada uno de los casetes, pero en su lugar son proporcionados por los genomas de las células huésped una vez que los casetes han sido introducidos en las células huésped.

En una realización, la etapa (d) añade secuencias de integración 5' y 3' flanqueando el promotor y poliA, respectivamente. Por ejemplo, la secuencia 5' es un elemento transposon 5' (p. ej., un elemento terminal 5' PB) y la secuencia 3' es un elemento transposon 3' (p. ej., un elemento terminal 3' PB que está en orientación invertida con respecto al elemento 5'). Por ejemplo, la secuencia de integración 5' se proporciona en el extremo 5' del primer cebador de la PCR puente y/o la secuencia de integración 3' se proporciona en el extremo 3' del segundo cebador de la PCR puente. El resultado es un repertorio de casetes de expresión, cada uno terminando (5' y 3') en elementos de transposón, es decir, cada uno de los casetes se modifica para formar un transposón. Opcionalmente, cada uno de los casetes de expresión se produce como un ADN lineal que termina en el extremo 5' y 3' por una secuencia de integración (p. ej., un elemento de transposón, p. ej., terminando en elementos de transposón PB terminales invertidos). Se pueden utilizar secuencias de integración alternativas en lugar de elementos de transposón (véase más adelante); por ejemplo, las secuencias de integración pueden ser brazos de homología (p. ej., al menos 15, 20, 50 o 100 nucleótidos contiguos) para llevar a cabo la integración homóloga en los genomas de las células huésped receptoras en uno o más sitios diana específicos (que hibridan con los brazos de homología). Alternativamente, las secuencias de integración pueden ser secuencias de recombinación específicas para el sitio (p. ej., lox, rox o frt) para la integración específica para el sitio en genomas de células huésped que transportan sitios de recombinación específicos del sitio correspondientes en uno o más sitios de integración deseados en el genoma (tras la provisión o expresión de la integrasa respectiva (es decir, cre, dre o flp, respectivamente). En una realización, RMCE se utiliza para insertar utilizando dos sitios de recombinación incompatibles (p. ej., loxP de tipo salvaje y un lox mutante, p. ej., lox2272 o lox511).

En una realización, uno o más elementos reguladores para la expresión de POI (p. ej., un promotor y/o potenciador y/o una poliA y/o una secuencia señal) se añaden mediante la etapa (c), p. ej., utilizando RACE. Adicional o alternativamente, uno o más elementos reguladores para la expresión de POI (p. ej., un promotor y/o un potenciador y/o una secuencia poliA y/o una secuencia señal) se añaden mediante la etapa (d), p. ej., utilizando una PCR puente.

Adicional o alternativamente, en una realización, una secuencia de integración de 5' y/o 3' se añade mediante la etapa (c), p. ej., utilizando RACE. Adicional o alternativamente, una secuencia de integración de 5' y/o 3' se añade mediante la etapa (d), p. ej., utilizando una PCR puente.

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

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

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

15 La transferencia por lotes de acuerdo con esta realización de la invención y la divulgación es superior a los métodos de la técnica anterior de utilizar la clonación molecular para transferir secuencias de nucleótidos codificantes de POI en células huésped. Como se ha explicado arriba, esta última requiere una secuenciación, análisis y subclonación laboriosa y lenta de secuencias codificantes de POI individuales que han sido modificadas por la inclusión de sitios de restricción terminales (para permitir la introducción en las células huésped en etapas posteriores). Típicamente, una vez que se ha confirmado una secuencia PCR codificante de POI correcta con sitios de restricción, ésta se selecciona entonces como una secuencia individual para avanzar para su introducción en las células huésped, estas últimas creciendo para producir una población de células que expresan la secuencia de POI elegida. Este proceso de clonación molecular laboriosa y multietapa se realiza para cada una de las variantes de POI en un repertorio que se ha de incluir en el cribado posterior. En consecuencia, los métodos de cribado de la técnica anterior pueden tardar varias semanas (típicamente del orden de 6-8 semanas) en funcionar para un repertorio útil de células de partida tales como las células productoras de anticuerpos. Los métodos de la invención y la divulgación, por el contrario, que utilizan la transferencia por lotes de casetes de expresión enteros para POIs (es decir, incluyendo POI y elementos reguladores para la expresión) proporcionan una técnica mucho más rápida para producir un repertorio clasificado de POIs para el cribado. Esto hace que el presente método sea susceptible de una automatización de alto rendimiento del cribado. Por ejemplo, los presentes inventores - realizando la operación manual del método - han sido capaces de realizar la producción de un repertorio de células huésped clasificadas para la expresión de POI en solo 2 días utilizando placas de 4 x 96 pocillos (aproximadamente 186 células B de entrada). La selección del repertorio de POI expresado se puede realizar manualmente en alrededor de 2 días solamente. Claramente, la automatización acelera aún más esto (y minimiza de manera ventajosa la contaminación cruzada entre partes alícuotas clasificadas).

Para la transferencia por lotes de casetes de expresión, una pluralidad de casetes de expresión clasificados se mezclan en la misma operación (p. ej., una sola etapa de aspiración y entrega de casetes, es decir, una sola etapa de pipeteo) con las células huésped para su transferencia a las células (p. ej., mediante transfección posterior o simultánea en las células). Esa operación es, por ejemplo, una sola transferencia mediante pipeta (p. ej., utilizando una pipeta multicanal, p. ej., utilizando una pipeta de 4, 8, 12, 16, 64, 96, 384 o 1536 canales en una sola operación). En un ejemplo, al menos 4 partes alícuotas clasificadas de casetes de expresión se mezclan con células huésped clasificadas en una sola operación, de modo que cada una de las partes alícuotas de casete de expresión se mezcla con una parte alícuota de célula respectiva (p. ej., en un recipiente respectivo, p. ej., en un pocillo de una placa o un tubo en un estante). En un ejemplo, al menos 4, 8, 12, 16, 24, 32, 40, 48, 56, 64, 96, 384 o 1535 partes alícuotas clasificadas de casetes de expresión se mezclan en la misma operación con células huésped clasificadas, de modo que cada una de las partes alícuotas de casete de expresión se mezcla con una parte alícuota de célula respectiva (p. ej., en un recipiente respectivo, p. ej., en un pocillo de una placa o un tubo en un estante). En una realización, la operación es una operación manual (p. ej., pipeteando utilizando una pipeta multicanal). En otra realización, la operación es automatizada, p. ej., realizada por un aparato automatizado de manipulación de líquidos (p. ej., un robot de manipulación de líquidos).

Ventajosamente, los inventores han encontrado que es posible transferir por lotes los casetes de expresión de la etapa (d) a las células huésped clasificadas sin la necesidad de purificar los casetes, pero aún así produciendo un repertorio útil de células huésped en la etapa (e). Esto proporciona mayores velocidades de procesamiento y rendimiento y hace que el procedimiento sea susceptible de una automatización más simple.

3. El método del concepto 1 o 2, en el que (i) el repertorio clasificado de casetes de expresión producidos por la etapa (d) se proporciona en una pluralidad de recipientes, cuyas ubicaciones relativas entre sí son fijas (p. ej., pocillos en una placa o tubos en un estante), en el que cada uno de los recipientes contiene un tipo respectivo de casete de expresión de tal manera que la ubicación relativa de los casetes de expresión relativa entre sí está predeterminada; y (ii) los casetes de expresión se transfieren a las células huésped clasificadas en la etapa (e) de tal manera que se mantienen las ubicaciones relativas de los casetes de expresión.

65 En un ejemplo, cada uno de los recipientes contiene casetes que codifican un primer tipo de POIs derivado de una célula individual respectiva proporcionada en la etapa (a) y también casetes que codifican un segundo tipo de POI

derivado de dicha célula individual. Por ejemplo, cada uno de los recipientes contiene casetes que codifican secuencias  $V_H$  y  $V_L$  derivadas de una célula individual. De esta manera, los dominios variables que componen los respectivos sitios de unión de las células de entrada (en los casos en los que estos codifican sitios de unión de anticuerpos, p. ej., plasmablastos o células plasmáticas) se mantienen juntos en el repertorio clasificado, pero no se mezclan con secuencias derivadas de otra célula. Esta clasificación se mantiene, por lo tanto, ventajosamente en etapas posteriores del procedimiento y es rastreable en el resultado del cribado posterior.

4. El método de cualquier concepto precedente, en el que (i) el repertorio de casetes de expresión producidos por la etapa (d) se clasifica proporcionando una pluralidad de recipientes (p. ej., pocillos en una placa o tubos en un estante), en el que cada uno de los recipientes comprende secuencias codificantes de POI de una célula individual clasificada en la etapa (b); (ii) las células huésped clasificadas de la etapa (e) se proporcionan en una pluralidad de recipientes (p. ej., pocillos en una placa o tubos en un estante) y (iii) los casetes de expresión se transfieren a las células huésped clasificadas en la etapa (e) de manera que las células huésped de cada uno de los recipientes respectivos se mezclen únicamente con secuencias codificantes de 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 casetes de expresión clasificados se proporcionan en una pluralidad de recipientes, cuyas ubicaciones relativas entre sí son fijas (p. ej., una pluralidad de recipientes con una disposición de X recipientes por Y recipientes, p. ej., una placa de  $8 \times 12$  pocillos (placa de 96 pocillos) o una placa que tiene un múltiplo de una disposición de  $8 \times 12$  recipientes (p. ej., una placa de 384 pocillos)); y en la etapa (e) las células huésped clasificadas se proporcionan en una pluralidad de recipientes, cuyas ubicaciones relativas entre sí son fijas y comprenden la misma disposición que los recipientes utilizados en la etapa (d) (p. ej., los repertorios de las etapas (d) y (e) se proporcionan ambos en placas de 96 pocillos o 384 pocillos de la misma o sustancialmente la misma dimensión) de modo que la clasificación se mantenga en la etapa (e).

6. El método de cualquier concepto precedente, en el que el repertorio clasificado de las células huésped son capaces de expresar de manera estable el repertorio de las POIs.

Alternativamente, el repertorio clasificado de células huésped es capaz de expresión transitoria. La expresión estable (p. ej., como resultado de la integración genómica de los casetes en los genomas de las células huésped) es ventajosa para el suministro a largo plazo de células - y, por lo tanto, las POIs expresadas - identificados después del cribado (y también para las células mientras esperan realizar el cribado si las células huésped producidas en la etapa (e) se almacenan durante un tiempo antes de ser utilizadas para el cribado).

7. El método de cualquier concepto precedente, en el que el repertorio clasificado de células huésped que expresan POI producidas mediante la etapa (e) se proporciona en una pluralidad de recipientes (p. ej., tubos o pocillos), en el que cada uno de los recipientes contiene POIs de una célula individual clasificada en la etapa (b).

Tubos de este tipo pueden fijarse en un estante o soporte; pocillos de este tipo pueden fijarse por disposición en una o más placas.

8. El método del concepto 7, en el que cada una de las células huésped expresa la primera y segunda POIs, en el que las POIs son diferentes, p. ej., subunidades de una proteína, p. ej., dominios variables de un sitio de unión de anticuerpos o receptor de células T.

9. El método de cualquier concepto precedente, en el que la etapa (b) comprende la clasificación de células individuales en recipientes respectivos (p. ej., pocillos respectivos en una o más placas) y la realización de las etapas (c) y (d) en dichos recipientes, al tiempo que se mantiene la clasificación.

10. El método de cualquier concepto precedente, que comprende, además, el cribado del repertorio de POI clasificado para identificar una POI con una característica deseada (p. ej., la unión a un antígeno o anticuerpo; o afinidad de unión para un ligando o antígeno cognado) y/o una secuencia de nucleótidos que codifica la POI identificada (p. ej., ADN o ARN, p. ej., 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 (p. ej., utilizando la PCR o mediante el cultivo de una célula huésped seleccionada o célula derivada de la misma); y, opcionalmente, producir POI aislada utilizando dicha secuencia de nucleótidos identificada, amplificada o sintetizada.

12. El método de cualquier concepto precedente, que comprende, además, el cribado del repertorio de POI clasificado para identificar una POI con una característica deseada (p. ej., la unión a un antígeno o anticuerpo; o afinidad de unión para un ligando o antígeno cognado) y el aislamiento de una célula huésped que expresa el POI identificado; y, opcionalmente, propagar la célula para producir una línea celular que expresa la POI.

13. El método de cualquier concepto precedente, en el que la etapa (e) comprende casetes de expresión de POI que se integran genómicamente (por ejemplo, se integran cromosómicamente) en los respectivos genomas de células

huésped para expresar las respectivas POIs. Alternativamente, uno o más de los casetes se proporcionan episomalmente en su respectiva célula huésped para la expresión transitoria de POI.

14. El método del concepto 13, en el que dicha integración genómica se lleva a cabo utilizando un motivo de secuencia de nucleótidos genómicos predeterminado para la inserción de los casetes de expresión en el genoma de la célula respectivo.

Por ejemplo, el motivo es una secuencia de nucleótidos utilizada por un transposón para la integración (p. ej., el motivo TTAA utilizado por PB); o una secuencia de nucleótidos que puede recombinarse con secuencias de integración de casete 5' y 3' por recombinación homóloga; o un motivo utilizado para integrar un sitio de recombinación específico para el sitio.

15. El método del concepto 14, en el que cada uno de los genomas celulares comprende más de una copia del motivo de secuencia.

16. El método del concepto 13, 14 o 15, en el que dicha integración genómica se lleva a cabo mediante la integración mediada por transposón.

Elementos de transposón adecuados para uso como secuencias de integración 5' y 3' de casetes son elementos de transposón de clase II (p. ej., elementos de repetición terminal invertida de transposón piggyBac o elementos de transposón Mariner), o elementos de transposón de la Bella Durmiente o elementos similares a TC1 (TLEs).

17. El método de uno cualquiera de los conceptos 13 a 16, en el que la etapa (e) comprende múltiples inserciones de casetes de expresión en los respectivos genomas de células huésped.

Cada uno de los casetes se proporciona opcionalmente como parte del ácido nucleico lineal (p. ej., ADN lineal). Por ejemplo, cada uno de los casetes es un transposón lineal que comprende o consiste en elementos de transposón 5'-y 3'-terminales (p. ej., elementos de repetición terminal invertidos piggyBac) con secuencia de nucleótidos POI y elemento(s) regulador(es) para la expresión entre los elementos de transposón. En una realización, hay una secuencia adicional 5' del elemento de transposón 5' y/o 3' del elemento de transposón 3'; en otra realización estos elementos están en los extremos 5' y 3' del casete, respectivamente.

18. El método de cualquier concepto precedente, en el que las células huésped son células de una línea celular de mamíferos (p. ej., células CHO o HEK293) o células de levadura.

Por ejemplo, el que cada una de las células huésped es una célula de mamífero (p. ej., célula animal humano o no humano, vegetal o de insecto o roedor o ratón, rata, conejo, pollo o camélido o de pescado), una célula bacteriana o de levadura.

19. El método de cualquier concepto precedente, en el que en la etapa (a) las células son células aisladas de uno o más animales.

Opcionalmente, cada una de las células de la etapa (a) es una célula de mamífero (p. ej., célula animal humano o no humano, vegetal o insecto o roedor o ratón, rata, conejo, pollo o camélido o pez), célula bacteriana o de levadura.

Opcionalmente, todas las células de la etapa (a) son células del mismo tipo de tejido o compartimiento de un(os) organismo(s). Por ejemplo, todos ellos son hígado, riñón, corazón, cerebro, sangre, linfocitos, próstata, ovarios o células germinales de uno o más organismos, p. ej., un paciente humano o un animal no humano, o un roedor o ratón o rata, conejo o pollo o camélido o pez.

20. El método de cualquier concepto precedente, en el que en la etapa (a) las células comprenden o consisten en células B, células centrales germinales, células B de memoria, células secretoras de anticuerpos, células plasmáticas o células plasmablasto.

21. El método de cualquier concepto precedente, en el que cada una de las POI es una cadena de inmunoglobulina (p. ej., un anticuerpo o receptor de células T) o parte de la misma (p. ej., un dominio variable).

22. El método de cualquier concepto precedente, en el que cada una de las POI comprende o consiste en un dominio variable de anticuerpo (p. ej., un dominio VH, VHH o VL o un dAb o Nanobody™).

23. El método de cualquier concepto precedente, en el que cada una de las células de la etapa (a) expresa la primera y segunda POIs, en la que las POIs son diferentes entre sí; en el que la etapa (b) comprende clasificar células individuales en recipientes respectivos (p. ej., pocillos respectivos en una o más placas) y llevar a cabo las etapas (c) y (d) en dichos recipientes, en donde después de la etapa (c) cada uno de los recipientes comprende las primeras POIs amplificadas mezcladas con POIs amplificadas de la misma célula; y en el que la etapa (e) comprende mezclar primer y segundo ácidos nucleicos de codificantes de POI respectivos de un recipiente respectivo con las células

huésped; en el que las primeras y segundas POIs de la misma célula de la etapa (a) se transfieren a la misma célula huésped para la expresión de primeras y segundas POIs por la célula huésped, produciendo con ello un repertorio de células huésped clasificadas, cada una de las cuales expresa conjuntamente las primeras y segundas POIs respectivas.

24. El método del concepto 23, en el que la primera y segunda POIs de la misma célula son polipéptidos cognados que juntos forman una proteína funcional (p. ej., dominios  $V_H$  y  $V_L$  que forman un sitio de unión del antígeno).

25. El método del concepto 24, en el que la primera y segunda POIs comprenden o consisten en dominios  $V_H$  y  $V_L$  de anticuerpo, respectivamente, p. ej., la primera y segunda POIs son cadenas pesadas y ligeras de anticuerpos cognados, respectivamente.

26. El método de cualquier concepto precedente, en el que la etapa (b) comprende las POIs de unión expresadas por células a un ligando cognado (p. ej., la unión a los sitios de unión de anticuerpos expresados por las células a un antígeno de interés); opcionalmente, en donde el ligando se une a la POI expresada en la superficie celular; y,

además, clasificar y aislar células que expresan POIs que se unen al ligando, produciendo con ello la población clasificada de células.

27. El método del concepto 26, en el que se utiliza la clasificación de células FACS; opcionalmente FACS de fluorescencia.

28. El método del concepto 26 o 27, en el que cada una de las células clasificada de la etapa (b) se proporciona en un recipiente respectivo (p. ej., un pocillo en una placa), de modo que cada uno de los recipientes de este tipo (p. ej., pocillo) comprende no más de un tipo de célula.

29. El método del concepto 28, en el que la población de células clasificadas se proporciona en pocillos en una o más placas que comprenden en total menos de 5, 4, 3, 2, 1 o 0,5 % pocillos que contienen más de una célula y/o en total menos de 5, 4, 3, 2, 1 o 0,5 % pocillos que no contienen célula alguna.

30. El método de cualquier concepto precedente, en el que la etapa (c) se realiza utilizando PCR, p. ej., RT-PCR utilizando ARN que codifica POI (p. ej., ARNm).

31. El método de cualquier concepto precedente, en el que la etapa (d) se realiza utilizando PCR, p. ej., PCR puente.

32. El método de cualquier concepto precedente, en el que la etapa (d) comprende la modificación de ácidos nucleicos amplificados mediante combinación con una secuencia predeterminada de modo que dicha secuencia predeterminada flanquee en 5' y/o 3' secuencias de nucleótidos codificantes de POI de los ácidos nucleicos; opcionalmente, en el que la modificación coloca un elemento regulador (p. ej., un promotor 5' y/o una poliA 3') y/o un elemento de transposón (p. ej., un elemento de transposón piggyBac) que flanquea en 5' y/o 3' secuencias de nucleótidos codificantes de POI.

33. El método del concepto 32, en el que cada una de las POI comprende un dominio variable de anticuerpos y la etapa (c) o (d) combina secuencias de nucleótidos codificantes de POI con una región constante de anticuerpo (opcionalmente una región constante humana o una de una especie que es diferente a la especie de región C comprendida por POIs en las células de la etapa (a)) para producir una secuencia de nucleótidos que codifica una cadena de anticuerpos (opcionalmente, una cadena humanizada).

Por ejemplo, las células de la etapa (a) codifican las POIs que comprenden regiones constantes de vertebrados no humanos (p. ej., roedores, p. ej., ratón o rata) y éstas son reemplazadas por regiones constantes humanas por la etapa (c) o (d). Por ejemplo, las POIs son cadenas de anticuerpos (p. ej., cadenas pesadas) que comprenden un dominio variable humano y una región constante de roedor (p. ej., ratón o rata) que se humaniza por el método de la invención. Esto es conveniente, ya que proporciona una manera de alto rendimiento de humanizar cadenas de anticuerpos y anticuerpos a escala y permite la selección posterior, producción y el desarrollo de vectores de células y de expresión en el contexto de formatos finales de anticuerpos/cadenas humanos adecuados para el uso de fármacos terapéuticos humanos. Las técnicas de la técnica anterior no hacen esto.

34. Un método de acuerdo con el concepto 33, que comprende, además, la selección del repertorio de POI clasificado para identificar una célula huésped que expresa una cadena de anticuerpos con una característica deseada (p. ej., unión específica al antígeno o afinidad de unión al antígeno), la identificación de la secuencia de nucleótidos codificante de la cadena de anticuerpos de la célula huésped, utilizando la secuencia de nucleótidos codificantes de la cadena de anticuerpos para producir copias de la cadena de anticuerpos identificada, y formular las copias como una composición farmacéutica (opcionalmente en combinación con uno o más fármacos, excipientes, diluyentes o soportes adicionales) 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 precedente, en el que la etapa (e) está automatizada; opcionalmente, en el que una o todas las etapas (b) a (c) también están automatizadas.

La automatización puede incluir el control del procedimiento por una computadora programada para llevar a cabo el método de cualquier aspecto, configuración, realización o ejemplo de la invención o la divulgación.

36. El método de cualquier concepto precedente, en el que (i) las etapas (b) a (e) inclusive se llevan a cabo en el equivalente de al menos 180 células (proporcionadas en la etapa (a) procesadas en no más de 1 o 2 días; y/o (ii) el repertorio de POIs expresadas se criba en busca de una característica deseada y una o más células huésped correspondientes o secuencias de nucleótidos codificantes de POI se identifican en un equivalente de no más de 4 días.

Los inventores han logrado esto utilizando aproximadamente 400 células B de entrada y el cribado de anticuerpos específicos para el antígeno, con (a) durando 2 días y (b) durando 3 días - todo hecho manualmente. Claramente, esto sería aún más rápido si se utiliza la automatización. Por lo tanto, la invención o la divulgación proporciona un ahorro significativo de tiempo frente a las técnicas del estado de la técnica que típicamente tardan 6-8 semanas en realizar un cribado de este tipo.

37. Un aparato automatizado para realizar el método de cualquier concepto precedente, comprendiendo el aparato

a. medios para mantener una población de células individuales clasificadas en una pluralidad de recipientes (p. ej., pocillos en una o más placas, o tubos en un estante o soporte como se describe arriba), en donde cada una de las células individuales está en un recipiente respectivo, comprendiendo cada una de las células ácido nucleico que codifica una POI respectiva;

b. medios para suministrar reactivos de PCR a los recipientes para amplificar el ácido nucleico compuesto por la población de células individuales clasificadas para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican POIs;

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

d. medios para mantener una población clasificada de células huésped en una pluralidad de recipientes;

e. medios para transferir (opcionalmente transferir por lotes) casetes de expresión de POI de dicho repertorio de casetes a la población clasificada de células huésped en los recipientes al tiempo que se mantiene la clasificación de casetes de expresión de POI; y

f. medios para llevar a cabo la introducción (p. ej., la transfección) de casetes de expresión en células huésped en los recipientes para producir un repertorio clasificado de células huésped que expresa un repertorio clasificado de POIs; y

g. opcionalmente una computadora programada para llevar a cabo el método de cualquier aspecto, configuración, realización o ejemplo de la invención o divulgación.

38. El aparato del concepto 37, que comprende, además, medios (p. ej., medios 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, medios para controlar el funcionamiento del aparato para el funcionamiento automatizado del método de uno cualquiera de los conceptos 1 a 36.

40. Un kit para llevar a cabo el método de uno cualquiera de los conceptos 1 a 36, comprendiendo el kit un aparato de acuerdo con el concepto 37, 38 o 39 junto con ácido nucleico que comprende el (los) elemento(s) de transposón para realizar el método del concepto 32.

Los elementos de transposón pueden ser transportados, p. ej., por 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 cortar y pegar (p. ej., transposón Clase II). En un ejemplo, los elementos son PB o elementos similares a Mariner o elementos similares a Tc-1 (TLEs).

41. Un casete de expresión para la expresión de una POI en una célula huésped, siendo el casete proporcionado por el ácido nucleico lineal (p. ej., ADN lineal) que comprende un transposón, comprendiendo el transposón elementos de transposón 5'- y 3'-terminales con una secuencia de nucleótidos codificante de POI y elemento(s) regulador(es) para la expresión de POI entre los elementos de transposón.

Casetes de este tipo son útiles para integrar genómicamente secuencias de POI expresables en células huésped, p. ej., para producir una línea celular para proporcionar una fuente de POI y/o para su uso en el método de cribado de la invención o divulgación. Los elementos de transposón pueden ser cualquiera de dichos elementos descritos en esta memoria.

En un ejemplo, el casete comprende o consiste en un transposón que comprende elementos de transposón 5'- y 3'- terminales (p. ej., elementos de repetición terminal invertida piggyBac) con una secuencia de nucleótidos codificante de POI y uno o más elementos reguladores para la expresión entre los elementos de transposón. En una realización, el casete comprende una secuencia adicional 5' del elemento de transposón 5' y/o 3' del elemento de transposón 3'. En un ejemplo, el casete comprende una secuencia de nucleótidos adicional correspondiente a una secuencia de nucleótidos del genoma de la célula huésped, estando dicha secuencia adicional en 5' y/o 3' de la secuencia de nucleótidos codificante de POI. Por ejemplo, la secuencia adicional corresponde a la secuencia genómica de células huésped que se transcribe activamente en el huésped. Por lo tanto, la secuencia codificante de POI se inserta en el huésped en un entorno adecuado para la transcripción activa de la secuencia POI.

En un ejemplo, una "población" (p. ej., una población de células o casetes) o "repertorio" como se utiliza en esta memoria comprende al menos 10, 100, 1000,  $10^4$ ,  $10^5$  o  $10^6$  miembros.

42. Una población de casetes de expresión de acuerdo con el concepto 41, en donde la población codifica un repertorio de POIs.

43. Una población clasificada de casetes de expresión de acuerdo con el concepto 42.

44. Una población clasificada de casetes de expresión que codifica un repertorio de POIs correspondientes a POIs expresadas por una población de células, comprendiendo cada una de las casetes una secuencia de nucleótidos que codifica un miembro del repertorio de POIs y uno o más elementos reguladores para la expresión de POI (cuando se encuentra en una célula huésped), en donde cada uno de los casetes comprende la disposición (en dirección 5' a 3'): elemento de transposón - [secuencia de nucleótidos POI y elemento(s) regulador(es)] - elemento de transposón y casetes de expresión para la expresión de POIs correspondientes a POIs de diferentes células están aislados entre sí en la población clasificada (p. ej., en diferentes pocillos de una placa, p. ej., una especie de casete por pocillo en una o más placas).

En un ejemplo, cada uno de los casetes es capaz de expresar una POI de (derivada de) una célula individual, p. ej., una cadena pesada o ligera de anticuerpo o fragmento de la misma derivado de una célula B individual.

En un ejemplo, se utilizan elementos piggyBac.

45. La población del concepto 44, en donde cada uno de los casetes de expresión es proporcionado por un ADN lineal.

46. La población de casetes de uno cualquiera de los conceptos 42 a 45, en donde cada uno de los casetes está en una célula huésped.

47. Una población clasificada de células huésped que comprende la población clasificada de casetes de expresión de acuerdo con uno cualquiera de los conceptos 43, 44 y 45 para la expresión de un repertorio clasificado de POIs.

48. Un método de hacer un transposón que comprende una secuencia de nucleótidos de interés (NOI), comprendiendo el método

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

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

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

4. d.

- 5 (i) mezclar la primera secuencia de nucleótidos con el primer molde para hibridar los brazos de homología primero y tercero juntos y llevar a cabo la amplificación y extensión de ácidos nucleicos (p. ej., utilizando la PCR) para extender la primera secuencia de nucleótidos utilizando el primer molde para producir una primera secuencia ampliada de nucleótidos (primera ENS) que comprende (en dirección 5' a 3') W, B y C; y (ii) mezclar la primera ENS con el segundo molde para hibridar los brazos de homología segundo y cuarto juntos y llevar a cabo la amplificación y extensión del ácido nucleico para extender la primera ENS para producir una segunda ENS que comprenda (o consista en) (en dirección 5' a 3') W, B y Z; o
- 10 (ii) mezclar la primera secuencia de nucleótidos con el segundo molde para hibridar los brazos de homología segundo y cuarto juntos y llevar a cabo la amplificación y extensión del ácido nucleico (p. ej., utilizando la PCR) para extender la primera secuencia de nucleótidos utilizando el segundo molde para producir una tercera secuencia extendida de nucleótidos (tercera ENS) que comprende (en dirección 5' a 3') A, B y Z; y (ii) mezclar la tercera ENS con el primer molde para hibridar los
- 15 brazos de homología primero y tercero juntos y llevar a cabo la amplificación y extensión del ácido nucleico para extender la tercera ENS para producir una cuarta ENS que comprende (o consiste en) (en dirección 5' a 3') W, B y Z; o
- 20 (iii) mezclar la primera secuencia de nucleótidos con el primer y el segundo moldes para hibridar los brazos de homología primero y tercero juntos e hibridar los brazos de homología segundo y cuarto juntos y llevar a cabo la amplificación y extensión de ácidos nucleicos (p. ej., utilizando la PCR) para extender la primera secuencia de nucleótidos utilizando el segundo molde para producir una quinta ENS que comprende (o consiste en) (en dirección 5' a 3') W, B y Z; y
- 25 e. aislar una ENS que comprende (o consiste en) (en dirección 5' a 3') W, B y Z, produciendo con ello un transposón aislado que comprende un NOI flanqueado por elementos de transposón; y
- f. opcionalmente, introducir el transposón aislado en una célula receptora de modo que el transposón se integre en el genoma de la célula.
- 30 Opcionalmente, una, más o todas 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.
- 35 En un ejemplo, el NOI codifica una POI, dominio de proteína o fragmento de proteína o es en sí mismo uno o más elementos reguladores. Por ejemplo, el NOI codifica una POI que es un ortólogo u homólogo de una proteína en el genoma de la célula receptora o en un vertebrado humano o no humano.
- 40 En una realización, W y X están en los extremos 5' y 3' de la primera secuencia del molde, respectivamente. Adicional o alternativamente, Y y Z están en los extremos 5' y 3' de la segunda secuencia del molde, respectivamente. Cuando W y X están en los extremos 5' y 3' de la primera secuencia del molde, respectivamente, e Y y Z están en los extremos 5' y 3' de la segunda secuencia del molde, respectivamente, el producto del método es un transposón lineal con elementos de transposón en sus extremos que es muy adecuado para la integración genómica para modificar las células huésped.
- 45 Opcionalmente, el primer molde consiste en 5'-W-X-3'. En un ejemplo, no hay una secuencia de nucleótidos intermedia entre W y X. En otra realización, hay secuencia de nucleótidos adicional entre W y X, p. ej., un elemento regulador o exón u otra secuencia de nucleótidos deseada (p. ej., secuencia codificante de proteínas) que se combinará inmediatamente aguas arriba del NOI en el producto del método. Esto es útil, por ejemplo, para construir un casete de expresión para combinar un promotor aguas arriba de un NOI (en que el NOI codifica una POI) para la expresión
- 50 posterior de la POI una vez que el transposón se ha insertado en un genoma de la célula huésped.
- 55 Adicional o alternativamente, opcionalmente la segunda secuencia de nucleótidos del molde consiste en 5'-X-Y-3' o hay una secuencia de nucleótidos intermedia entre X e Y, p. ej., un elemento regulador o exón u otra secuencia de nucleótidos deseada (p. ej., secuencia codificante de proteínas) que se combinará inmediatamente aguas abajo del NOI en el producto del método. Esto es útil, por ejemplo, para construir un casete de expresión para combinar una políA aguas abajo de un NOI (en que el NOI codifica una POI) para la expresión posterior de la POI una vez que el transposón se ha insertado en un genoma de la célula huésped. En otro ejemplo, la secuencia intermedia codifica una proteína que se fusionará con la POI tras la expresión para producir un producto de fusión. Por ejemplo, la POI comprende o consiste en un dominio variable de anticuerpos y la secuencia intermedia comprende o consiste en una
- 60 secuencia codificante de la región constante de anticuerpos. Por ejemplo, la secuencia intermedia codifica un anticuerpo Fc o un dominio CH1 o CL de un anticuerpo. En un ejemplo, el Fc o la región constante o proteína codificada por la secuencia intermedia es un Fc humano, región constante o proteína. Esto es útil para humanizar la POI (p. ej., para producir una cadena de anticuerpo humanizado cuando la POI es un dominio variable, p. ej., un dominio variable humano).
- 65

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

50. El método del concepto 49, en el que el NOI codifica un dominio de proteína (p. ej., un dominio variable de anticuerpo) y hay una secuencia de nucleótidos que codifica una región constante de anticuerpos (p. ej., un Fc de anticuerpos, p. ej., un Fc humano) entre Y y Z, por lo que el producto de transposón codifica una proteína de fusión que comprende un dominio de proteína condensado a una región constante de anticuerpos (p. ej., codificando una cadena de anticuerpos).

51. El método de una cualquiera de los conceptos 48 a 50, en el que uno o más de los brazos de homología primero y segundo se combinan con el NOI mediante la PCR (p. ej., 5'- y/o 3'-RACE) para formar dicha primera secuencia de nucleótidos antes de llevar a cabo dicha extensión.

52. Un método para hacer un repertorio de transposones, en el que los miembros del repertorio codifican diferentes POIs (p. ej., diferentes dominios variables de anticuerpos), comprendiendo el método

i. proporcionar una población de primeras secuencias de nucleótidos que comprenden un repertorio de NOIs; y

ii. para cada primera secuencia de nucleótidos, llevar a cabo el método de uno cualquiera de los conceptos 48 a 51, produciendo con ello un repertorio de transposones que codifica un repertorio de POIs.

53. El método del concepto 52, que comprende clasificar las primeras secuencias de nucleótidos para proporcionar una población clasificada antes de llevar a cabo la etapa (ii), en el que se produce un repertorio clasificado de transposones codificando un repertorio clasificado de POIs.

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

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

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 POIs.

57. Un método de producir una célula huésped para la expresión de una POI, comprendiendo el método

a. proporcionar al menos primero y segundo casetes de expresión, en donde cada uno de los casetes de expresión comprende

i. un primer elemento de integración y un segundo elemento de integración en posición 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 utilizando una enzima de integración;

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

c. introducir simultánea o secuencialmente los primero y segundo casetes de expresión en la célula huésped, en donde cada uno de los casetes está integrado genómicamente en el genoma de la célula huésped en dicho motivo para la expresión de POIs por la célula huésped; y

d. producir opcionalmente una línea celular de expresión de POIs en una etapa que comprende cultivar la célula huésped.

Este aspecto de la divulgación es útil para producir células huésped y líneas celulares para una expresión relativamente alta de uno o más POIs de interés. La integración genómica de casetes de POI en múltiples sitios genómicos proporciona una expresión estable y también existe la posibilidad de fijar como objetivo regiones transcripcionalmente activas del genoma huésped. El uso de motivos de secuencia guía la inserción a sitios útiles y esto es preferible a la integración aleatoria de secuencias como se utiliza en la técnica.

58. El método del concepto 57, en el que los elementos de integración primero y segundo del primer casete son idénticos a los elementos de integración primero y segundo, respectivamente, del segundo casete.

59. En un ejemplo, cada uno de los casetes comprende elementos de transposón primero y segundo, p. ej., elementos del mismo tipo de transposón (p. ej., PB o un transposón de Clase II). En un ejemplo, todos los elementos son sitios de recombinación específicos para el sitio, p. ej., sitios lox o sitios frt o una mezcla de estos. En otro ejemplo, todos los elementos son brazos de homología (secuencias de nucleótidos contiguos suficientes para la recombinación homóloga en la célula huésped). En un ejemplo, los sitios de recombinación específicos para el sitio son los mismos o son diferentes (p. ej., sitios mutuamente incompatibles (p. ej., loxP y lox511 o 2272) para llevar a cabo un RMCE (siglas inglesas de intercambio de casetes mediado por recombinasa) para la inserción dirigida del casete en el genoma.

59. El método del concepto 57 o 58, en el que los elementos de integración primero y segundo de cada uno de dichos casetes primero y segundo están en orientación mutuamente invertida (p. ej., elementos de transposón PB invertidos o sitios de recombinación invertidos específicos para el sitio).

60. El método de uno cualquiera de los conceptos 57 a 59, en el que uno o más de dichos motivos se diseñan en el cromosoma de la célula huésped antes de llevar a cabo la etapa (c), p. ej., los pares de sitios lox se diseñan en uno o más cromosomas huésped, en donde se utilizan pares correspondientes a pares lox en casetes).

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

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

62. El método de uno cualquiera de los conceptos 57 a 61, en el que al menos 3 casetes están genómicamente integrados en el genoma de la célula huésped, p. ej., en uno o más cromosomas huésped - que 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 de casetes están activos para la transcripción de las secuencias codificantes de POI.

Esto se puede lograr utilizando transposones (p. ej., PB) en los casetes.

64. El método de uno cualquiera de los conceptos 57 a 63, en el que cada uno de los casetes se integra por recombinación homóloga entre los sitios de integración y el genoma del huésped; recombinación específica para el sitio entre los sitios de integración y el genoma del huésped; o por integración mediada por transposón.

65. El método de los conceptos 64, en el que la enzima se selecciona de una recombinasa o una transposasa (p. ej., una enzima correspondiente a los elementos de integración PBase (p. ej., PBase hiperactiva), flp o cre recombinasa).

En un ejemplo, la célula huésped ha sido diseñada para expresar enzima(s) de este tipo, por ejemplo, a partir de un gen genómicamente integrado (p. ej., un gen inducible). En otra realización, la enzima se expresa a partir de un vector episomal. En otro ejemplo, la enzima se introduce en la célula huésped.

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

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

68. El método de uno cualquiera de los conceptos 57 a 67, que comprende aislar una célula huésped producida por 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 POI aislada utilizando dicha secuencia de nucleótidos identificada, amplificada o sintetizada o un mutante de la misma.

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

70. Una población de células huésped, obtenible por el método de uno cualquiera de los conceptos 57 a 69, cada una de las células comprende una pluralidad de casetes de expresión genómicamente integrados para expresar POIs, comprendiendo cada una de las células huésped una pluralidad de motivos de secuencia de nucleótidos idénticos a lo largo de su genoma adyacente a un casete de expresión integrado para la expresión de una POI de cada uno de los casetes de este tipo; comprendiendo cada uno de los casetes integrados

a. una primera secuencia de elementos de integración y una segunda secuencia de elementos de integración 3' de la primera secuencia de nucleótidos de elementos de integración; y

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

71. Las células huésped del concepto 70, en donde todas las POIs expresadas por las células son la misma POI.

10 72. Las células huésped del concepto 70, en donde las células expresan POIs primera y segunda (p. ej., dominios VH y VL de un solo tipo de anticuerpo; o cadenas pesadas y ligeras de un solo tipo de anticuerpo) que se asocian para formar un sitio de unión a proteína funcional o ligando (p. ej., antígeno).

15 73. Un sitio de unión de anticuerpos o antígenos de un anticuerpo para el tratamiento médico de un ser humano, en donde el anticuerpo o el sitio de unión ha sido aislado de una célula huésped producida por un método de uno cualquiera de los conceptos 57 a 69 o aislado de una célula huésped de una población de acuerdo con uno cualquiera de los conceptos 70 a 72.

20 74. Una mezcla de ácido nucleico que comprende un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en donde el primer ácido nucleico es capaz de hibridarse a una secuencia de anticuerpos humanos V de la región 5'UTR (es decir, una secuencia de nucleótidos) de un gen compuesto por un ácido nucleico diana, en donde el gen codifica una región V humana; y el segundo ácido nucleico es capaz de hibridarse a una segunda secuencia, en donde la segunda secuencia está compuesta por el ácido nucleico diana y es 3' a la secuencia UTR, en donde el primer ácido nucleico aislado comprende (o consiste en) una secuencia que es al menos 90, 91, 92, 93, 94, 95, 96, 25 97, 98 o 99 % idéntico (o 100 % idéntico) a una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 1-47.

30 SEQ ID NOs: 1-47 comprende las secuencias específicas para la región variable humana, como se indica en la Tabla 1 (más particularmente las secuencias específicas de nucleótidos 5'UTR de regiones variables humanas). Por "específico para" se entiende que secuencias de este tipo se pueden utilizar como una secuencia de cebador 5' en la PCR estándar (p. ej., RT-PCR) de ácido nucleico de la región variable humana.

SEQ ID NOs: 1-17 comprende secuencias específicas para la región variable de la cadena pesada humana.

35 SEQ ID NOs: 18-26 comprende secuencias específicas para la región variable de la cadena kappa humana.

SEQ ID NOs: 27-47 comprende secuencias específicas para la región variable de la cadena kappa humana.

40 En una realización, la divulgación proporciona un ácido nucleico (p. ej., un cebador de la PCR o un vector para recombinación homóloga) que comprende (o consiste en) al menos 15 nucleótidos contiguos de una Secuencia denominada X en la Tabla 2 para hibridarse a la secuencia 5'UTR de un segmento de gen denominado Y en la Tabla 2, p. ej., para realizar una PCR para copiar el segmento de gen o para hibridarse a un vector de recombinación homólogo para la secuencia 5'UTR para la modificación del segmento del 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 todos de la secuencia X. En un ejemplo, los nucleótidos contiguos terminan con el nucleótido 3' de la secuencia X (es decir, se utilizan los nucleótidos contiguos que se extienden 5' desde el extremo 3' de X). En una realización, la divulgación proporciona una mezcla 45 de dos o más de los ácidos nucleicos, p. ej., para la copia por PCR de dos o más secuencias de región variable (p. ej., utilizando ADN, ADNc o ARN de células B correspondientes). En un ejemplo, dos, más o todos los ácidos nucleicos en la mezcla copian segmentos del gen V<sub>H</sub>. En un ejemplo, dos, más o todos los ácidos nucleicos de la mezcla copian segmentos del gen V<sub>λ</sub>. En un ejemplo, dos, más o todos los ácidos nucleicos de la mezcla copian los segmentos del gen V<sub>κ</sub>. Opcionalmente, el ácido nucleico, o cada uno, comprende una secuencia de nucleótidos del promotor inmediatamente 5' de la secuencia UTR (o la parte de nucleótidos contigua 15 o más). Por ejemplo, la secuencia del promotor es una secuencia del promotor de CMV de la siguiente manera:

55 **5'-CTTACTGGCTTATCGAAATTAATACGACTCAGATC-3' (SEQ ID NO: 54)**

En un ejemplo, la divulgación proporciona:

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

65 En un ejemplo del ácido nucleico, mezcla o cebador de la divulgación, cada uno de los ácidos nucleicos o cebadores se hibrida con su secuencia cognada a una temperatura de 45-70 °C, (p. ej., a 50 °C, o 60 °C, o 68 °C) o 60-75 °C, en una reacción PCR. Una persona experta en la técnica será consciente de los tiempos de ciclo y las temperaturas para llevar a cabo la reacción PCR.

Cada uno de los ácidos nucleicos de la divulgación y de las mezclas de la divulgación es útil para realizar la amplificación o replicación de una secuencia de nucleótidos diana que codifica un dominio variable de anticuerpo humano o una proteína que comprende un dominio de este tipo, por ejemplo, una PCR de secuencia(s) de nucleótidos que codifican la región variable humana aisladas de una o más células (p. ej., células B). Por lo tanto, en una realización, cada uno de los ácidos nucleicos es un cebador de la PCR.

Cada uno de los ácidos nucleicos de la divulgación y de las mezclas de la divulgación es útil para realizar la recombinación homóloga para modificar una secuencia de nucleótidos diana (p. ej., una secuencia compuesta por el genoma de una célula, p. ej., una célula de mamífero, p. ej., una célula ES o célula CHO). Para la recombinación homóloga, como es conocido por la persona calificada, se utiliza un vector de ácido nucleico que comprende un brazo de homología 5', un brazo de homología 3' y, opcionalmente, una secuencia de nucleótidos predeterminada de interés entre ellos. La secuencia puede, por ejemplo, codificar una POI, un dominio de proteínas o puede estar compuesta por un elemento regulador. En una alternativa no hay una secuencia intermedia entre los brazos de homología (y en este caso el vector se utiliza para suprimir la secuencia del genoma que se encuentra entre regiones que se hibridan con los brazos de homología, como es conocido por la persona experta). En la presente realización, la divulgación proporciona un vector de recombinación homóloga, en donde el vector comprende un brazo 5' que comprende un brazo de homología 3' y, opcionalmente, una secuencia de nucleótidos entre ellos, en donde 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 NOs: 1-47 (o es 100 % idéntica) y/o en donde 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 es 100 % idéntica) a una secuencia seleccionada del grupo que consiste en los SEQ ID NOs: 48 a 53. Esto permite la fijación como objetivo génica de segmentos de genes V y/o C específicos de loci de Ig en un vertebrado utilizando recombinación homóloga.

La divulgación, por lo tanto, también proporciona un método para modificar un locus de Ig compuesto por una célula de vertebrado, comprendiendo el método introducir el vector de la divulgación en la célula (p. ej., por transfección) y llevar a cabo la recombinación homóloga para modificar el locus de Ig; y, opcionalmente, expresar un dominio V o cadena de anticuerpo a partir del locus modificado. Opcionalmente, la secuencia del dominio V se identifica o copia o aísla de la célula y se utiliza para producir un anticuerpo o una composición farmacéutica que comprende un anticuerpo de este tipo para uso médico humano.

La divulgación proporciona, además, un cebador de la PCR que 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 NOs: 1-53 (o es 100% idéntica). Por ejemplo, el cebador es *in vitro*.

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

El ácido nucleico, el cebador o mezcla de la PCR se pueden proporcionar *in vitro*, p. ej., mezclado con un tampón o reactivo de PCR. En un ejemplo, un ácido nucleico, cebador o mezcla de la divulgación se proporciona en un recipiente, un vial, un tubo, una placa o una cubeta de PCR.

Opcionalmente, el primer ácido nucleico aislado 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 formado por SEQ ID NOs: 1-47. Opcionalmente, el primer ácido nucleico aislado comprende (o consiste en) una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 1-47.

75. La mezcla del concepto 74, en donde 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 NOs: 1-17.

Opcionalmente, el primer ácido nucleico aislado comprende una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 1-17.

76. La mezcla del concepto 74, en donde 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 NOs: 18-26.

Opcionalmente, el primer ácido nucleico aislado comprende una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 18-26.

77. La mezcla del concepto 74, en donde 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 NOs: 27-47.

Opcionalmente, el primer ácido nucleico aislado comprende una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 27-47.

78. La mezcla de uno cualquiera de los conceptos 74 a 77, donde el segundo ácido nucleico aislado comprende una secuencia de región constante de anticuerpos; 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 NOs: 48 a 53.

Opcionalmente, el segundo ácido nucleico aislado comprende una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 48 a 53. SEQ ID NOs: 48 -51 son secuencias de regiones constantes de ratón; SEQ ID NOs: 52 y 53 son secuencias de regiones constantes humanas (véase la Tabla 1).

79. La mezcla del concepto 75, en donde el segundo ácido nucleico aislado comprende una secuencia de región constante de cadena pesada de anticuerpos; 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.

80. La mezcla del concepto 76, en donde el segundo ácido nucleico aislado comprende una secuencia de región constante de cadena kappa de anticuerpos; 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 donde el segundo ácido nucleico aislado comprende una secuencia de región constante de cadena lambda de anticuerpos; 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.

82. Una mezcla de ácidos nucleicos que comprende un primer ácido nucleico aislado y un segundo ácido nucleico aislado, en donde 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 NOs: 1-47.

En un ejemplo, los ácidos nucleicos son cebadores de la PCR; en otra realización comprenden vectores homólogos de recombinación 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 NOs: 18-26 y una secuencia que sea 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 NOs: 18-26 y/o seleccionada del grupo formado por SEQ ID NOs: 27-47.

84. La mezcla del concepto 82 o 83, en donde se selecciona cada uno de los ácidos nucleicos aislados primero y segundo, que es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99 % idénticos (o 100 % idénticos) a una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 1-17.

85. La mezcla del concepto 84, que comprende al menos 3 ácidos nucleicos aislados diferentes, siendo cada uno al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99 % idénticos (o 100 % idénticos) a una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 1-17.

86. La mezcla del concepto 82 o 83, en donde cada uno de los ácidos nucleicos aislados primero y segundo es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99 % idénticos (o 100 % idénticos) a una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 18-26.

87. La mezcla del concepto 84, que comprende al menos 3 ácidos nucleicos aislados diferentes, cada uno de los cuales es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99 % idénticos (o 100 % idénticos) a una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 18-26.

88. La mezcla del concepto 82 o 83, en donde cada uno de los ácidos nucleicos aislados primero y segundo es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99 % idénticos (o 100 % idénticos) a una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 27-47.

89. La mezcla del concepto 84, que comprende al menos 3 ácidos nucleicos aislados diferentes, siendo cada uno al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99 % idénticos (o 100 % idénticos) a una secuencia seleccionada del grupo que consiste en SEQ ID NOs: 27-47.

90. La mezcla del concepto 84 o 85, en donde la mezcla comprende una secuencia de región constante de cadena pesada de anticuerpos; y, opcionalmente, comprende SEQ ID NO: 48 y/o 49.

91. La mezcla del concepto 86 o 87, en donde la mezcla comprende una secuencia de región constante de cadena kappa de anticuerpo; y, opcionalmente, comprende SEQ ID NO: 50 y/o 51.



92. La mezcla del concepto 87 o 88, en donde la mezcla comprende una secuencia de región constante de cadena lambda de anticuerpos; y, opcionalmente, comprende SEQ ID NO: 52 y/o 53.

93. El método de uno cualquiera de los conceptos 1 a 36, en el que la etapa (c) se realiza mediante PCR utilizando una o más mezclas de acuerdo con 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 utilizando una o más mezclas de acuerdo con uno cualquiera de los conceptos 74 a 92.

95. Un kit que comprende una o más mezclas de acuerdo con uno cualquiera de los conceptos 74 a 92 y un aparato de acuerdo con uno cualquiera de los conceptos 36 a 39.

96. Un método de amplificar un repertorio de secuencias de regiones variables humanas, comprendiendo el método

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

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

c. aislar, secuenciar o identificar una o más de las secuencias de nucleótidos replicadas o llevar a cabo 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) 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 NOs: 1-52.

97. El método del concepto 96, en el que la etapa (b) utiliza una o más mezclas de acuerdo con 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 de la etapa (a) se clasifican en células individuales (p. ej., se clasifican en pocillos de una o más placas).

99. El método del concepto 96 o 97, que comprende, además, producir una región variable humana (p. ej., como parte de una cadena de anticuerpos aislados o un anticuerpo aislado para la medicina humana) utilizando 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 divulgación descrita en esta memoria.

Opcionalmente, la secuencia de nucleótidos codificantes de POI está enlazada operativamente a un promotor capaz de impulsar la expresión de la POI, en donde el promotor comprende un promotor eucariótico que es regulable por un activador o inhibidor. En otra realización, el promotor eucariótico está operativamente enlazado a un operador procariótico, y la célula eucariótica comprende opcionalmente, además, una proteína represora procariótica.

Opcionalmente, cada uno de los casetes de expresión comprende una secuencia que codifica un marcador tal como un marcador seleccionable, por ejemplo, un gen de resistencia a la higromicina o codifica una proteína fluorescente (p. ej., la proteína fluorescente se selecciona de DsRed, GFP, eGFP, CFP, ECFP e YFP).

Opcionalmente, uno o más o todos los casetes de expresión comprenden secuencias de nucleótidos primera y segunda codificantes de POI, p. ej., en tándem o como un casete bicistrónico. En un ejemplo, las POIs codificadas son diferentes (p. ej., VH y VL de un anticuerpo); en otro ejemplo, son diferentes. En un ejemplo, 1, 2, 3, 4, 5, 6 o más secuencias de nucleótidos codificantes de POI.

En un ejemplo, la o cada una de las células huésped es una célula CHO (ovario de hámster chino) o célula HEK293.

Por ejemplo, la proteína de interés puede ser un anticuerpo o fragmento del mismo, un anticuerpo quimérico o fragmento del mismo, un scFv o fragmento del mismo, una proteína marcada con Fc o fragmento de la misma, un factor de crecimiento o un fragmento del mismo, una citoquina o un fragmento del mismo, o un dominio extracelular de un receptor de la superficie celular o fragmento del mismo.

## Construcciones de Ácido Nucleico

Los casetes de expresión recombinante (vectores) pueden comprender fragmentos de ADN sintéticos o derivados de ADNc que codifican una proteína de interés, operativamente enlazados a un elemento regulador transcripcional y/o traslacional adecuado, derivado de genes de mamíferos, virales o de insectos. Elementos reguladores de este tipo incluyen promotores de transcripciones, potenciadores, secuencias que codifican sitios de unión ribosomal de ARNm adecuados y secuencias que controlan la terminación de la transcripción y la traducción. Casetes de expresión de

mamíferos también pueden incluir elementos no transcritos, tales como un origen de replicación, otras secuencias no transcritas flanqueantes 5' o 3', y secuencias no traducidas 5' o 3', tales como sitios de donante y aceptor de corte y empalme. También se puede incorporar un gen marcador seleccionable para facilitar el reconocimiento de transfectantes.

Secuencias de control de la transcripción y la traducción en casetes de expresión, útiles para la transfección de células de vertebrados, pueden ser proporcionadas por fuentes virales. Por ejemplo, promotores y potenciadores de uso común se derivan de virus tales como el poliovirus, el adenovirus 2, el virus de simio 40 (SV40) y el citomegalovirus (CMV) humano. Se pueden utilizar promotores genómicos virales, secuencias de control y/o señal para impulsar la expresión, con la condición de que secuencias de control de este tipo sean compatibles con la célula huésped elegida. También se pueden utilizar promotores celulares no virales (p. ej., los promotores beta-globina y EF-1 alfa), dependiendo del tipo de célula en el que se ha de expresar la proteína recombinante.

Las secuencias de ADN derivadas del genoma viral SV40, por ejemplo, el origen SV40, sitios de promotor temprano y tardío, potenciadores, de corte y empalme y de poliadenilación se pueden utilizar para proporcionar otros elementos genéticos útiles para la expresión de la secuencia heteróloga del ADN. Los promotores tempranos y tardíos son particularmente útiles porque ambos se obtienen fácilmente del virus SV40 como un fragmento que también comprende el origen viral SV40 de la replicación (Fiers et al., Nature, 1978, 273:113). También se pueden utilizar fragmentos de SV40 más pequeños o más grandes. Típicamente, se incluye la secuencia de aproximadamente 250 bp que se extiende desde el sitio de Hind III hacia el sitio de BglI situado en el origen SV40 de la replicación.

Los vectores de expresión bicistónica utilizados para la expresión de múltiples transcripciones han sido descritos anteriormente (Kim S. K. y Wold B. J., Cell, 1985, 42:129) y se pueden utilizar en combinación con una o más secuencias codificantes de POI.

### **Células Huésped y Transfección**

Opcionalmente, las células huésped eucariotas se utilizan en los métodos de la invención y divulgación, p. ej., son células huésped de mamíferos, incluyendo, por ejemplo, células CHO o células de ratón.

Las proteínas expresadas (POIs) serán secretadas preferiblemente en el medio de cultivo, dependiendo de la secuencia de ácido nucleico seleccionada, pero pueden ser retenidas en la célula o depositadas en la membrana celular. Se pueden emplear diversos sistemas de cultivo celular de mamíferos para expresar proteínas recombinantes. Ejemplos de líneas celulares huésped de mamíferos adecuadas incluyen las líneas COS-7 de células renales de mono, descritas por Gluzman (1981) Cell 23:175, y otras líneas celulares capaces de expresar un vector apropiado incluyendo, por ejemplo, CV-1/EBNA (ATCC CRL 10478), células L, C127, 3T3, CHO, líneas celulares HeLa y BHK. Otras líneas celulares desarrolladas para esquemas específicos de selección o amplificación también serán útiles con los métodos y las composiciones proporcionados en esta memoria. Una línea celular preferida es la línea de células CHO designada K1. Con el fin de lograr el objetivo de producción de alto volumen de proteínas recombinantes, la línea celular huésped se pre-adapta opcionalmente al medio biorreactor en el caso apropiado.

Varios protocolos de transfección son conocidos en la técnica, y son revisados en Kaufman (1988) Meth. Enzymology 185:537. El protocolo de transfección elegido dependerá del tipo de célula huésped y de la naturaleza de la POI, y se puede elegir basándose en la experimentación de rutina. Los requisitos básicos de cualquier protocolo de este tipo son primero introducir el ADN que codifica la proteína de interés en una célula huésped adecuada, y luego identificar y aislar las células huésped que han incorporado el ADN heterólogo de una manera relativamente estable y expresable.

Un método comúnmente utilizado para introducir ADN heterólogo en una célula es la precipitación de fosfato de calcio, por ejemplo, como se describe por Wigler et al. (Proc. Natl. Acad. Sci. USA 77:3567, 1980). El ADN introducido en una célula huésped mediante este método se somete frecuentemente a un reordenamiento, lo que hace que este procedimiento sea útil para la co-transfección de genes independientes.

Fusión inducida por polietileno de protoplastos bacterianos con células de mamíferos (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 con frecuencia producen múltiples copias del ADN del plásmido integrado en el genoma de la célula huésped de los mamíferos, y esta técnica requiere que el marcador de selección y amplificación esté en el mismo ácido nucleico que la POI.

La electroporación también se puede utilizar para introducir el ADN directamente en el citoplasma de una célula huésped, por ejemplo, como se describe por Potter et al (Proc. Natl. Acad. Sci. USA 1:7161, 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 en el mismo ácido nucleico.

Más recientemente, 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 comercialmente utilizados para formar complejos de lípido-ácido nucleico (o liposomas) que, cuando se aplican a células cultivadas, facilitan la absorción del ácido nucleico en las células.

Un método para amplificar la POI también es deseable para la expresión de la proteína recombinante, y típicamente implica el uso de un marcador de selección (revisado en Kaufman *supra*). La resistencia a los fármacos citotóxicos es la característica más frecuentemente utilizada como un marcador de selección, y puede ser el resultado de un rasgo dominante (p. ej., puede ser utilizado independientemente del tipo de célula huésped) o un rasgo recesivo (p. ej., útil en tipos de células huésped particulares que son deficientes en cualquier actividad que se está seleccionando). Varios marcadores amplificables son adecuados para su uso en los vectores de expresión descritos en esta memoria (p. ej., como se describe en Maniatis, *Molecular Biology: A Laboratory Manual*, Cold Spring Harbor Laboratory, NY, 1989; pp 16.9-16.14).

Marcadores seleccionables útiles para la amplificación génica en células de mamífero resistentes a los fármacos se muestran en la Tabla 1 de Kaufman, R. J., *supra*, e incluyen resistencia a DHFR-MTX, P-glicoproteína y diversos agentes citotóxicos lipófilos de resistencia múltiple a fármacos (MDR, por sus siglas en inglés) (p. ej., adriamicina, colchicina, vincristina) y adenosina desaminasa (ADA-Xyl-A o adenosina y 2'-desoxicoformicina).

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

Elementos reguladores útiles, descritos anteriormente o conocidos en la técnica, también se pueden incluir en las construcciones de ácido nucleico utilizadas para transfectar células de mamíferos. El protocolo de transfección elegido y los elementos seleccionados para su uso dependerán del tipo de célula huésped utilizada. Aquellos con experiencia en la técnica son conscientes de numerosos protocolos y células huésped diferentes, y pueden seleccionar un sistema apropiado para la expresión de una proteína deseada, basado en los requisitos del sistema de cultivo celular utilizado.

Un aspecto proporciona una composición farmacéutica que comprende una POI aislada (p. ej., anticuerpo, cadena o dominio variable) y un diluyente, excipiente o soporte, opcionalmente en donde la composición está contenida en un recipiente IV (p. ej., y bolsa IV) o un recipiente conectado a una jeringa IV y en donde la POI ha sido aislada de una célula huésped de la divulgación o población de células huésped.

Un aspecto proporciona el uso de la POI como se describe en esta memoria en la fabricación de un medicamento para el tratamiento y/o la profilaxis de una enfermedad o afección en un paciente, p. ej., un ser humano.

Se entenderá que realizaciones particulares descritas en esta memoria se muestran a modo de ilustración y no como limitaciones de la invención. El uso de la palabra "un" o "una", cuando se utiliza junto con la expresión "que comprende" en las reivindicaciones y/o la memoria descriptiva puede significar "uno"/"una", pero también es consistente con el significado de "uno/una o más, "al menos uno/una" y "uno/una o más de uno/una". El uso del término "o" en las reivindicaciones significa "y/o" a menos que se indique explícitamente que se refiere solo a alternativas o las alternativas se excluyen mutuamente, aunque la divulgación apoya una definición que se refiere solo a alternativas e "y/o". A lo largo de esta solicitud, el término "aproximadamente" se utiliza para indicar que un valor incluye la variación inherente de error para el dispositivo, empleándose el método 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 individuales o ASCs específicas para el antígeno del bazo, ganglios linfáticos y médula ósea con los marcadores celulares correspondientes; rescate de la secuencia de anticuerpos de células individuales y amplificación de casetes de expresión; expresión de anticuerpos recombinantes en células de mamíferos. El diagrama de flujo se muestra en la Figura 1B. Los detalles de cada una de las etapas se describen a continuación

### **Ejemplo 1A: Aislamiento de células individuales específicas para el antígeno**

Las células específicas para el antígeno incluyen las células de memoria/GC con anticuerpos unidos a membrana y las células plasmáticas secretoras de anticuerpos. Se utilizó un panel de marcadores de superficie celular para definir y marcar células de memoria de ratón/GC (CD19; IgM; IgD; CD38; CD95) (Figura 2). Las células específicas para el antígeno fueron teñidas utilizando antígenos solubles marcados por fluorescencia (por ejemplo, cualquier fluoróforo de moléculas pequeñas que pueda ser detectado por el sistema de clasificación celular, tal como Alexa-488, Alexa-647, Pacific Blue, R-ficoeritrina, isotiocianato de fluoresceína o alofococianina, conjugado opcionalmente a un colorante de cianina, p. ej., Cy7) o antígenos de la superficie celular en partículas similares a virus (VLPs). La Figura 2 es un ejemplo de marcaje de células de memoria/GC específicas para el antígeno en el bazo de ratón inmunizado con OVA. Más de 10.000 células IgG de memoria/GC específicas para OVA podrían clasificarse a partir de un bazo. La

clasificación de una célula individual se realizó utilizando un citómetro de flujo de inyección de BD equipado con una unidad automática de deposición de células. Las células clasificadas por FACS se depositaron en placas de PCR de 96 pocillos con tampón de lisis para la siguiente etapa.

Por lo general, las células GC específicas para el antígeno (centro germinal) o B de memoria pueden ser capturadas por el antígeno marcado, porque expresan de forma dominante anticuerpos transmembrana en la superficie celular. Por otro lado, se pensaba que plasmablastos o células plasmáticas eran menos fácilmente capturados por el antígeno marcado debido a su expresión dominante de anticuerpos secretores. A continuación, los inventores intentaron aislar ASCs utilizando antígeno marcado fluorescentemente y anti-CD138 para clasificar las células del resto de la población celular utilizando FACS (Figura 3). Como se muestra en la Figura 2, la mayoría de las células plasmáticas o plasmablastos específicos para el antígeno aislados, pero ninguna de esas células sobrantes del mismo tipo de células demostró que eran ASCs específicas para el antígeno en el ensayo ELISPOT. Esto demostró que el método de clasificación celular que utiliza el antígeno marcado por fluorescencia puede capturar eficientemente todas las ASCs específicas para el antígeno probablemente con anticuerpos transmembrana residuales o anclaje temporal del anticuerpo secretado en la superficie celular.

El antígeno marcado fluorescentemente puede ser reemplazado por VLPs con antígeno recombinante en su superficie. Las VLPs se generan a partir de células CHO, células KEK, MEFs (siglas inglesas de fibroblastos embrionarios de ratón) u otras líneas celulares de mamíferos con la co-expresión del antígeno recombinante, la proteína gag del retrovirus, y MA-GFP (proteína de fusión p15-GFP del fragmento de matriz gag). La expresión de gag permite la aparición de VLP de las células, y la MA-GFP marca las VLPs para la detección de fluorescencia. Tanto las proteínas gag como MA-GFP están asociadas con la superficie interna de la membrana plasmática, y el antígeno recombinante está en la superficie de la VLP. Los antígenos en las VLPs se presentan en forma nativa expresada directamente a partir de células recombinantes sin etapa de purificación o modificación alguna. La forma nativa de un antígeno debe proporcionar todos los epítomos naturales que ayudan en gran medida a la selección de anticuerpos neutralizantes. La alta densidad del antígeno en las VLPs aumenta la relación señal/ruido para la detección de células que expresan anticuerpos específicos para el antígeno en la superficie celular y facilita en gran medida la etapa de clasificación. Las VLPs recombinantes pueden generarse con la expresión de diferentes proteínas fluorescentes tales como MA-CFP o MA-YFP. Utilizando la multiplexación de VLPs con diferentes antígenos y diferentes proteínas de fluorescencia, se pueden seleccionar células que expresan aglutinantes de alta afinidad, aglutinantes de reactividad cruzada o aglutinantes específicos de homólogos. Las células que expresan aglutinantes de alta afinidad pueden ser seleccionadas por células con matriz de afinidad relativa alta (matriz de afinidad = la relación de la actividad de unión a antígeno de baja densidad frente a VLPs de antígeno de alta densidad). Las células que expresan aglutinantes de reactividad cruzada a ortólogos o diferentes antígenos (para el aislamiento de anticuerpos biespecíficos 2 en 1) se pueden seleccionar por células que se unen a diferentes tipos de VLPs al mismo tiempo. Las células que expresan aglutinantes específicos del homólogo también se pueden seleccionar por la célula que solo se une al antígeno específico, pero no a su homólogo.

#### **Ejemplo 1B: Recuperación de alto rendimiento de la secuencia de Anticuerpos de células individuales**

Se desarrolló un método rápido, eficiente y de alto rendimiento para la generación de anticuerpos a partir de células B individuales sin etapa de clonación molecular alguna. El método permitió a los inventores producir construcciones de expresión de Ig a partir de segmentos de genes variables amplificados de cadena pesada y cadena ligera de una célula individual (Figura 3). A través de todo el procedimiento de la PCR, las cadenas pesadas y las cadenas ligeras de una célula individual se amplificaron en el mismo pocillo.

Las secuencias de regiones de anticuerpo V fueron recuperadas por RT-PCR y dos rondas de PCRs por el siguiente procedimiento. Las placas clasificadas de una célula individual almacenadas a -80 °C se descongelaron en hielo y se centrifugaron brevemente antes de su uso. Las placas fueron incubadas en el termociclador a 65 °C durante 5 minutos e indefinidamente a 4 °C. 6 µL de mezcla de cebadores, Superscript III, dNTP, inhibidor de RNasa y tampón se añadieron a cada uno de los pocillos y se mezclaron por pipeteo. Las placas se centrifugaron brevemente y se incubaron a 50 °C durante 60 minutos. Se utilizaron cebadores específicos para la región constante para la cadena pesada y la cadena ligera para amplificar los segmentos de genes variables de una célula individual. Se utilizaron cebadores inversos específicos para genes para amplificar las cadenas kappa, lambda y gamma, gamma RT1, kappa RT1 y lambda RT3.

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

En la segunda ronda de PCR, se utilizó un cebador genérico directo que reasoció la etiqueta hCMV con un cebador anidado inverso para la región constante. 1 µL de los productos de la primera PCR se utilizaron como moldes para la segunda PCR anidada. Las condiciones de ciclación para la segunda PCR incluían una etapa inicial de desnaturalización a 98 °C durante 30 minutos, seguido 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 mantenido a 4 °C indefinidamente. 1/3 del producto de la PCR de segunda ronda se ejecutaron en gel de agarosa al 1 % para comprobar la tasa de recuperación de la secuencia de anticuerpos de células individuales para RT y 2 rondas de etapas de PCR. Dado que los cebadores para cadenas pesadas y ligeras se mezclaron en los mismos pocillos, los productos de la 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 son ~ 700 pb para las cadenas ligeras kappa y lambda, y ~ 500 pb para la cadena pesada gamma (Figura 4a).

Para la expresión de anticuerpos en las células de mamíferos, los productos amplificados se puentearon con un casete de Ig lineal con promotor LTR-CMV 5' PB y región constante-señal poliA-3' PB LTR (Figura 3). El casete de Ig contiene todos los elementos esenciales para la expresión del anticuerpo, incluyendo el promotor CMV, la región constante de la cadena de inmunoglobulina y la señal poli (A). Adicionalmente, el casete tiene largas regiones de CMV solapantes y homología de región constante en sus extremos. 2 µL de los productos de la segunda ronda de PCR se utilizaron como molde para la PCR puente. Las condiciones de ciclación para la PCR puente fueron una etapa inicial de desnaturalización a 98 °C durante 30 minutos, seguido 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 mantenido a 4 °C indefinidamente. 1/3 del producto de PCR puente se ejecutaron en gel de agarosa al 1 % para comprobar la tasa de recuperación de la PCR puente. Los tamaños esperados de los productos de PCR son ~ 2600 pb para las cadenas ligeras kappa y lambda, y ~ 3100 pb para la cadena pesada gamma (Figura 4b).

La etapa puente permite reunir todos los elementos de expresión y PB LTRs juntos para formar el transposón PB con genes de expresión de cadena pesada y cadena ligera. No importa qué isotipo tengan los anticuerpos IgG1, IgG2a, IgG2b, IgG3 de ratón o IgG1, IgG2, IgG3, IgG4 humana o cualquier variante de región constante se puede aplicar en la etapa puente para reformatear el Fc. El método aplicado en esta tecnología no requiere etapa de purificación alguna y puede ser ampliamente automatizado. La tasa de recuperación general para la tecnología de células B (BCT, por sus siglas en inglés) a través de la clasificación celular y la PCR de células individuales es de aproximadamente 38-71% para diferentes poblaciones celulares (Tabla 3).

#### **Ejemplo 1C: Análisis de secuencias por agrupaciones**

Los productos de PCR de segunda ronda se enviaron para la secuenciación. Las secuencias de nucleótidos se determinaron utilizando un secuenciador de ADN 373 de Applied Biosystems. Las secuencias se analizaron por 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 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 se define generalmente por el uso de secuencias relacionadas de cadena pesada de inmunoglobulina y/o cadena ligera V(D)J por 2 o más muestras. Las secuencias de la cadena pesada V(D)J de inmunoglobulina relacionadas pueden ser identificadas por su uso compartido de los segmentos del gen V(D)J codificados en el genoma (Figura 5).

Dentro de una familia clonal, generalmente hay subfamilias que varían según mutaciones compartidas dentro de sus segmentos V(D)J, que pueden surgir durante la recombinación del gen de células B y la hipermutación somática. Clones con diferente uso del segmento V(D)J exhiben habitualmente diferentes características de unión. Además, los clones con el mismo uso del segmento V(D)J pero diferentes mutaciones exhiben diferentes características de unión. Las células B se someten a hipermutación somática, en que se producen cambios aleatorios en las secuencias de nucleótidos de los genes de los anticuerpos, y se seleccionan células B cuyos anticuerpos tienen una afinidad más alta de las células B (Figura 6). Si los clones de baja afinidad del mismo linaje tienen función de neutralización, la potencia aumenta habitualmente en los clones con más mutación para adquirir mayor afinidad.

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

La etapa final de la PCR amplificó el casete de expresión lineal que codifica la cadena pesada y la cadena ligera. Los casetes amplificados para la cadena pesada y la cadena ligera, y el vector de expresión de PB transposasa (PBase) fueron co-transfectados en una línea celular de mamífero sin purificación ni clonación. El sobrenadante con expresión transitoria o estable de anticuerpo se recogió en los momentos correspondientes. La transfección del vector de expresión convencional probablemente provoca la integración del concatémico en el genoma y el gen integrado está sujeto a ser silenciado. La expresión mediada por transposón PB proporciona una ventaja importante para el nivel de expresión alto y estable de los genes transfectados porque múltiples copias (10-100) de transposones PB pueden ser transpuestas e integradas al genoma dentro de las regiones activas de transcripción, permitiendo la expresión de anticuerpos de alto nivel (Figura 7).

Las transfecciones de productos de la PCR puente y vectores de expresión de PBase para la expresión transitoria se realizaron utilizando Lipofectamine 2000 siguiendo el protocolo del fabricante. Las transfecciones se llevaron a cabo en placas de pocillos profundos de 96 pocillos. En resumen, células HEK293 fueron cultivadas en DMEM+ FBS de IgG ultrabajo al 10 % (Invitrogen) para evitar que IgG bovino compita con IgG humana secretada en la fase de purificación de la proteína A aguas abajo. Para las transfecciones en placas de 96 pocillos, cada uno de los pocillos se sembró el día anterior con  $5 \times 10^5$  células en 500  $\mu$ L de medio, y se dejó que creciera a razón de  $1 \times 10^6$  células al día siguiente. 25  $\mu$ L de los 30  $\mu$ L de los productos de PCR puente fueron incubados en medios Optimem con 100 ng de vector PBase para un volumen final de 70  $\mu$ L, y Lipofectamine 2000 también fue incubada por separado con 70  $\mu$ L de medios Optimem. Ambas incubaciones fueron durante 10 minutos. Lipofectamine 2000 y los productos de PCR se mezclaron mediante pipeteo suave y se incubaron durante 15 minutos antes de añadirlos a las células HEK293 y se mezclaron suavemente. Los sobrenadantes de cultivo se recogieron el día 8 después de la transfección para los siguientes cribados. La concentración de IgG en sobrenadantes que contienen anticuerpos de interés está determinada por IgG ELISA (Figura 8). La concentración de los anticuerpos expresados es equiparable a la que normalmente obtienen los inventores a partir de la tecnología del hibridoma, que es suficiente para la mayoría de los cribados aguas abajo para la capacidad de unión de anticuerpos o los ensayos funcionales. Las tasas de éxito generales para la tecnología de células B a través de la clasificación celular para la identificación de IgG es de 36 %-71 % dependiendo de las diferentes poblaciones celulares (Tabla 3).

#### **Ejemplo 1E: Análisis de cribado de unión de anticuerpos utilizando escaneo LI-COR Odyssey NIR**

Los anticuerpos expresados de la transfección de células HEK293 de la tecnología de células B (BCT) se cribaron primero en cuanto a su capacidad de unirse al antígeno de interés utilizando la exploración LI-COR Odyssey NIR, y luego los clones positivos fueron cribados en cuanto a su afinidad aparente por Resonancia de Plasmón de Superficie (ProteOn XPR36, BioRad) (véase más adelante).

Células B que producen anticuerpos específicos de antígeno se identificaron mediante detección fluorescente. Cada uno de los pocillos de placas transparentes de fondo plano de 384 pocillos se sembró con  $1 \times 10^4$  células CHO adherentes transfectadas de forma estable con un gen que codifica un antígeno de transmembrana humana en 80  $\mu$ L de medio F12 que contenía FBS al 10 % (v/v) ( $1,25 \times 10^5$  células/mL) utilizando un instrumento Multidrop. Las células se incubaron durante la noche a 37 °C en una incubadora de CO<sub>2</sub>. Al día siguiente se retiró el medio de cultivo por aspiración y se añadieron 45  $\mu$ L de anticuerpo LI-COR IRDye 800CW anti-Ratón a 500 ng/mL + DRAQ5 5 mM (LI-COR) diluido 1:25.000 en Tampón FACS (PBS+BSA al 1 % + Na<sub>3</sub> al 0,1 %). Se añadieron 5  $\mu$ L de sobrenadante BCT, 5  $\mu$ L de anticuerpo control (2  $\mu$ g/mL) en medio de cultivo HEK293 o 5  $\mu$ L de anticuerpo control IgG1 de ratón (Sigma, 2  $\mu$ g/mL) en medio de cultivo HEK293 utilizando un manipulador de líquidos FluidX. Se incubaron placas durante 1 h a 4 °C y se aspiraron medios de cultivo. La reacción se detuvo y las células se fijaron mediante la adición de 25  $\mu$ L de paraformaldehído al 4 % por pocillo e incubación durante 15 minutos a TA (temperatura ambiente). Las placas se lavaron dos veces con 100  $\mu$ L de PBS y la solución de lavado se eliminó secando en toallas de papel. Las placas fueron escaneadas utilizando un instrumento Li-Cor Odyssey Classic. Las tasas generales de éxito para la clasificación celular a través de BCT para la identificación específica para el Antígeno es del 25 %-61 % dependiendo de las diferentes poblaciones celulares (Tabla 3).

**Tabla 3** Tasa de recuperación de cada una de las etapas de la tecnología de células B.

Procedimientos de BCT	Tasa de recuperación			
	Células memoria bazo/GC	de del	Células memoria LN/GC (RIMMS)	Células plasmáticas de la médula ósea
PCR de células individuales	38%		66%	71%
IgG de expresión y cribado	36%		66%	71%
Aglutinantes específicos para Ag de expresión y cribado	25%		59%	61%

#### **Ejemplo 1F: Mediciones de afinidad por SPR utilizando el Método de Captura de Anticuerpos**

Clones positivos que expresaban anticuerpos específicos para el antígeno se examinaron para determinar su afinidad aparente por Resonancia de Plasmón de Superficie. IgG anti-ratón (GE Healthcare/Biacore) se acopló al GLM por acoplamiento de aminas primarias. El chip GLM (BioRad) fue activado utilizando NHS/EDAC y la IgG anti-ratón se acopló a esta superficie activada y luego se bloqueó utilizando etanolamina 1 M. La inmovilización se llevó a cabo en HBS-EP (Teknova) o HBS-N (GE Healthcare/Biacore) a temperatura ambiente o 37 °C, respectivamente. La superficie de IgG anti-ratón en el chip GLM se utilizó para capturar directamente anticuerpos de interés. Para el análisis cinético se utilizaron 5 concentraciones de analito (256 nM, 64 nM, 16 nM, 4 nM y 1 nM). Para el análisis de los datos, los sensorgramas de unión fueron referenciados utilizando la referencia interna "interspot" única al Proteon XPR36 que

son doble referenciados utilizando el sensorgrama de inyección de tampón. Finalmente, los datos fueron analizados utilizando el modelo 1:1 inherente al software de análisis ProteOn XPR36.

La Figura 9 es un ejemplo de los datos de SPR de los anticuerpos del Kymouse® inmunizado con ovalbúmina, utilizando la tecnología de células B individuales de la invención. Alrededor de dos tercios de los anticuerpos testados mostraron evidencia de unión al antígeno, con una gama diversa de afinidades de unión y cinética. La diversidad de características de unión revela que el procedimiento de clasificación celular 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 bajo-nM a bajo-pM verificando la eficiencia de la maduración de afinidad.

La Figura 10 muestra la afinidad aparente de los anticuerpos contra dos antígenos diana de Kymab diferentes. Se detectó una gama de aglutinantes (círculos en blanco) así como neutralizadores funcionales (círculos en negro), con la mayor afinidad detectada en el intervalo picomolar. Esto valida la tecnología de clonación de células B individuales de la invención para ser una herramienta poderosa en la identificación y recuperación de anticuerpos de alta afinidad y funcionalmente competentes.

Tal como se utiliza en esta memoria descriptiva y la o las reivindicaciones, las palabras "comprendiendo" (y cualquier forma de comprendiendo, tal como "comprenden" y "comprende"), "teniendo" (y cualquier forma de teniendo, tal como "tienen" y "tiene"), "incluyendo" (y cualquier forma de incluyendo, tal como "incluye" e "incluyen") o "conteniendo" (y cualquier forma de conteniendo, tal como "contiene" y "contienen") son inclusivas o indefinidas y no excluyen elementos adicionales ni etapas del método no enumerados.

La expresión "o combinaciones de los mismos", tal como se utiliza en esta memoria, se refiere a todas las permutaciones y combinaciones de los elementos enumerados que preceden a la expresión. Por ejemplo, "A, B, C o combinaciones de los mismos" pretende incluir al menos uno de: A, B, C, AB, AC, BC o ABC, y si el orden es importante en un contexto particular, también BA, CA, CB, CBA, BCA, ACB, BAC o CAB. Continuando con este ejemplo, se incluyen expresamente las combinaciones que contengan repeticiones de uno o más elementos o términos, tales como BB, AAA, MB, BBC, AAABCCCC, CBBAAA, CABABB, etcétera. El experto en la materia entenderá que típicamente no hay límite en el número de elementos o términos en cualquier combinación, a menos que resulte evidente de otro modo del contexto.

**Tabla 1: Secuencias de Nucleótidos/Ácidos Nucleicos:**

Oligos V <sub>H</sub>	
SEQ ID NO: 1	TCTAGAGAAAACCCTGTGAGCACAGCTC
SEQ ID NO: 2	GAGAATCCCCTGAGAGCTCCGTTT
SEQ ID NO: 3	TCAGAAGCCCCCAGAGCACAACGC
SEQ ID NO: 4	TGGGAGAATCCCCTAGATCACAGCTC
SEQ ID NO: 5	ACAGAAGCCCCCAGAGCGCAGCAC
SEQ ID NO: 6	CCCACCATGGACACACTTTGCTCC
SEQ ID NO: 7	TGGACTCCAAGGCCTTCCACTTGG
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	CCCCAGCCTTGGGATTCCCAAGTGTTC
SEQ ID NO: 15	TGAGATTCCCACGTGTTTCCATTTCAG
SEQ ID NO: 16	ACTTGGTGATCAGCACGGAGCACCGA
SEQ ID NO: 17	CTGGGATTTTCAGGTGTTTTCATTTGG
Oligos V <sub>K</sub>	
SEQ ID NO: 18	GGAGTCAGACCCAGTCAGGACACAGC
SEQ ID NO: 18	GGAGTCAGACCCACTCAGGACACAGC

Oligos V <sub>H</sub>	
SEQ ID NO: 20	GGAATCAGTCCCACTCAGGACACAGC
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

Oligos V <sub>L</sub>	
SEQ ID NO: 27	atggcctggctcctctcctc
SEQ ID NO: 28	atggccggctccctctcctc
SEQ ID NO: 29	atgccctgggctctgctcctc
SEQ ID NO: 30	atgccctgggtcatgctcctc
SEQ ID NO: 31	atggcctgggctctgctgctc
SEQ ID NO: 32	atggcatggatccctctcttc
SEQ ID NO: 33	atggcctggaccctctcctg
SEQ ID NO: 34	atggcctggaccctctcctc
SEQ ID NO: 35	atggcctggaccctctctgg
SEQ ID NO: 36	atggcctggaccgttctcctc
SEQ ID NO: 37	atggcatgggccacactcctg
SEQ ID NO: 38	atggcctggatccctctactt
SEQ ID NO: 39	atggcctggatccctctcctg
SEQ ID NO: 40	atggcctggaccgctctcctt
SEQ ID NO: 41	atggcctgggtctcctttac
SEQ ID NO: 42	atggcctggacccactcctc
SEQ ID NO: 43	atggcttgacccactcctc
SEQ ID NO: 44	atggcctggactcctctcctc
SEQ ID NO: 45	atggcctggactcctctctt
SEQ ID NO: 46	atggcctggatgatgcttctc
SEQ ID NO: 47	atggcctgggtcctctgctc
Oligos Región C	
	Oligo    Secuencia (5' a 3')
SEQ ID NO: 48	C <sub>H1</sub> gctcttgccgTAGCCCTTGACCAGGCATCC
SEQ ID NO: 49	C <sub>H2</sub> CAGATCCAGGGGCCAGTGATAGAC
SEQ ID NO: 50	C <sub>K1</sub> gttctgatcgaaCTAACTCATTCTGTTGAAG
SEQ ID NO: 51	C <sub>K2</sub> GACAATGGGTGAAGTTGATGTCTTGTGAG
SEQ ID NO: 52	C <sub>L1</sub> cgacaaccactacctCTATGAACATTCTGTAGGGGC
SEQ ID NO: 53	C <sub>L2</sub> CTTCTCCACGGTGCTCCCTTCATGC

**Tabla 2:**



<b>Oligos V<sub>H</sub></b>	
<b>X</b>	<b>Y (es decir, la secuencia X utilizada para copiar o modificar el segmento génico Y)</b>
SEQ ID NO: 1	IGHV1-8 (p. ej., IGHV1-8*01)
SEQ ID NO: 2	IGHV1-2 (p. ej., IGHV1-2*04)
SEQ ID NO: 3	IGHV1-3*01 (p. ej., IGHV1-3*01)
SEQ ID NO: 4	IGHV1-18 (p. ej., IGHV1-18*01)
SEQ ID NO: 5	IGHV1-24 (p. ej., IGHV1-24*01)
SEQ ID NO: 6	IGHV2-5 (p. ej., IGHV2-5*10) y/o IGHV2-26 (p. ej., IGHV2-26*01)
SEQ ID NO: 7	IGHV3-7 (p. ej., IGHV3-7*01)
SEQ ID NO: 8	IGHV4-4 (p. ej., IGHV4-4*02)
SEQ ID NO: 9	IGHV6-1 (p. ej., IGHV6-1*01)
SEQ ID NO: 10	IGHV7-4-1 (p. ej., IGHV7-4-1*01)
SEQ ID NO: 11	IGHV3-9 (p. ej., IGHV3-9*01)
SEQ ID NO: 12	IGHV3-11 (p. ej., IGHV3-11*01)
SEQ ID NO: 13	IGHV3-13 (p. ej., IGHV3-13*01)
SEQ ID NO: 14	IGHV3-15 (p. ej., IGHV3-15*01)
SEQ ID NO: 15	IGHV3-20 (p. ej., IGHV3-20*01)
SEQ ID NO: 16	IGHV3-21 (p. ej., IGHV3-21*01)
SEQ ID NO: 17	IGHV3-23 (p. ej., IGHV3-23*01)
<b>Oligos V<sub>K</sub></b>	
<b>X</b>	<b>Y (es decir, la secuencia X utilizada para copiar o modificar el segmento génico Y)</b>
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 IGKV3
SEQ ID NO: 25	IGKV4-1
SEQ ID NO: 26	IGKV5-2
<b>Oligos V<sub>L</sub></b>	
<b>X</b>	<b>Y (es decir, la secuencia X utilizada para copiar o modificar el segmento génico Y)</b>
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

Oligos VL	
X	Y (es decir, la secuencia X utilizada para copiar o modificar el segmento génico Y)
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

## REIVINDICACIONES

1. Un método de producir células que codifican un repertorio de anticuerpos que comprende cadenas pesadas y cadenas ligeras de anticuerpos cognados, comprendiendo el método:-

a) proporcionar una población de células que expresen un repertorio de anticuerpos que comprenden cadenas pesadas de anticuerpos y cadenas ligeras de anticuerpos cognados, cuyas células se aíslan de uno o más animales, y que comprenden células B, células centrales germinales, células B de memoria, células secretoras de anticuerpos, células plasmáticas o células plasmablastos;

b) clasificar la población de células uniendo los anticuerpos a un antígeno de interés para producir una población clasificada de células individuales, comprendiendo cada una de las células ácido nucleico que codifica una cadena pesada de anticuerpos y una cadena ligera de anticuerpos cognados respectivos, y cada una de las células clasificadas se proporciona en un recipiente separado;

c) amplificar el ácido nucleico compuesto por la población de células individuales clasificada utilizando PCR para producir un repertorio clasificado de ácidos nucleicos amplificados que codifican las secuencias de VH y VL cognadas, en donde las secuencias VH y VL de cada uno de los anticuerpos están compuestas por el anticuerpo cognado de la misma célula y se amplifican en el mismo recipiente;

d) modificar los ácidos nucleicos cognados amplificados clasificados que codifican VH y VL a partir de la etapa (c) para producir un repertorio clasificado de casetes de expresión, comprendiendo cada uno de los recipientes:

un primer tipo de casete de expresión de anticuerpos cognados que comprende:

(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 anticuerpos; y

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

un segundo tipo de casete de expresión de anticuerpos cognados. que comprende:

(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 anticuerpos; y

(iv) uno o más elementos reguladores para expresar la cadena ligera de anticuerpos, en donde las secuencias VH y VL de cada una de las cadenas pesadas de anticuerpos y de las cadenas ligeras del anticuerpo cognados se modifican en el mismo recipiente; y

e) transferir los casetes de expresión de anticuerpos cognados de dicho repertorio de casetes a una población clasificada de células huésped, manteniendo al mismo tiempo la clasificación de los casetes de expresión de anticuerpos cognados y produciendo un repertorio clasificado de células huésped que co-expresan cada una un anticuerpo que comprende una cadena pesada de anticuerpos y una cadena ligera de anticuerpos cognados, en donde el repertorio clasificado de células huésped es capaz de expresar el repertorio de anticuerpos.

2. El método de la reivindicación 1, en el que en la etapa (e) los casetes de expresión clasificados se transfieren por lotes a las células huésped clasificadas.

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

4. El método de una cualquiera de las reivindicaciones 1 a 3, en el que (i) el repertorio clasificado de casetes de expresión producidos por la etapa (d) se proporciona en una pluralidad de recipientes, cuyas ubicaciones relativas entre sí están fijadas (p. ej., pocillos en una placa o tubos en un estante), en el que cada uno de los recipientes comprende un tipo respectivo de casete de expresión de tal manera que la ubicación relativa de casetes de expresión relativa entre sí está predeterminada; y (ii) los casetes de expresión se transfieren a las células huésped clasificadas en la etapa (e) de tal manera que se mantienen las ubicaciones relativas de los casetes de expresión.

5. El método de cualquier reivindicación precedente, en el que (i) el repertorio de casetes de expresión producidos por la etapa (d) se clasifica proporcionando una pluralidad de recipientes (p. ej., pocillos en una placa o tubos en un estante), en el que cada uno de los recipientes comprende secuencias codificantes del sitio de unión de anticuerpos de una célula individual clasificada en la etapa (b); (ii) las células huésped clasificadas de la etapa (e) se proporcionan en una pluralidad de recipientes (p. ej., pocillos en una placa o tubos en un estante) y (iii) los casetes de expresión se transfieren a las células huésped clasificadas en la etapa (e) de manera que las células huésped en cada cada uno de los recipientes respectivos se mezclen únicamente con secuencias codificantes del sitio de unión de anticuerpos derivadas de una célula individual clasificada en la etapa (b).

6. El método de cualquier reivindicación precedente, en el que, en la etapa (b), se utiliza la clasificación de células FACS; opcionalmente FACS de fluorescencia.

7. El método de cualquier reivindicación precedente, en el que el antígeno de interés es un antígeno marcado de manera fluorescente; o una partícula similar al virus (VLP) que comprende un antígeno expresado en la superficie y que, además, comprende una proteína marcada de manera fluorescente.
8. El método de cualquier reivindicación precedente, en el que la etapa (d) comprende la modificación de los ácidos nucleicos amplificados para colocar un elemento regulador (p. ej., un promotor 5' y/o una poliA 3') y/o un elemento de transposón (p. ej., un elemento de transposón piggyBac) flanqueando 5' y/o 3' las secuencias de nucleótidos codificantes de la cadena de anticuerpos.
9. El método de cualquier reivindicación precedente, en el que la etapa (e) comprende la integración genómica de los casetes de expresión de anticuerpos cognados en los respectivos genomas de células huésped para expresar los respectivos anticuerpos cognados, en donde dicha integración genómica se lleva a cabo utilizando un motivo de secuencia de nucleótidos genómicos predeterminado para la inserción de los casetes de expresión en el respectivo genoma de la célula.
10. El método de la reivindicación 9, en el que dicha integración genómica se lleva a cabo por integración mediada por transposón.
11. El método de cualquier reivindicación precedente, en el que en la etapa (d) la modificación se lleva a cabo utilizando una PCR puente para producir un repertorio de casetes de expresión primero y segundo que comprenden elementos de transposón 5'- y 3'- terminales con una respectiva cadena pesada de anticuerpo cognado o secuencia de nucleótidos codificante de la cadena ligera de anticuerpo y elemento(s) regulador(es) para la expresión de la cadena de anticuerpos entre los elementos de transposón.
12. El método de cualquier reivindicación precedente, que comprende, además, cribar el repertorio de anticuerpos clasificados para identificar una célula huésped que expresa un anticuerpo con una característica deseada (p. ej., unión específica al antígeno o afinidad de unión al antígeno), identificar las secuencias de nucleótidos codificantes de la cadena de anticuerpos de la célula huésped, utilizar las secuencias de nucleótidos codificantes de la cadena de anticuerpos para producir copias de las cadenas de anticuerpos identificadas, y formular las copias como una composición farmacéutica (opcionalmente en combinación con uno o más fármacos, excipientes, diluyentes o soportes adicionales) para la terapia médica humana.
13. El método de la reivindicación 12, en el que la composición farmacéutica que comprende un excipiente o soporte diluyente está contenida en un recipiente IV (p. ej., una bolsa IV) o un recipiente conectado a una jeringa IV.

FIGURA 1A

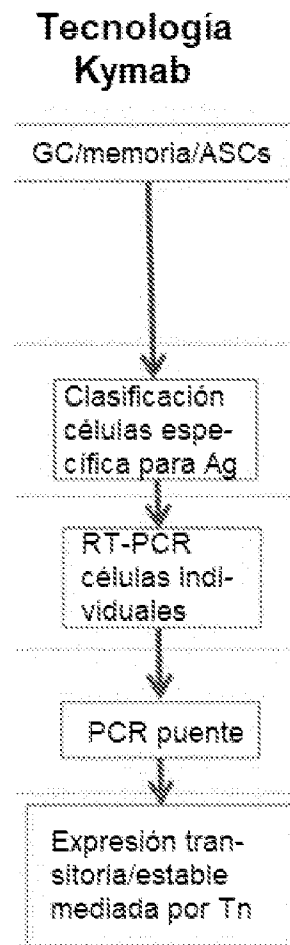


Figura 1B

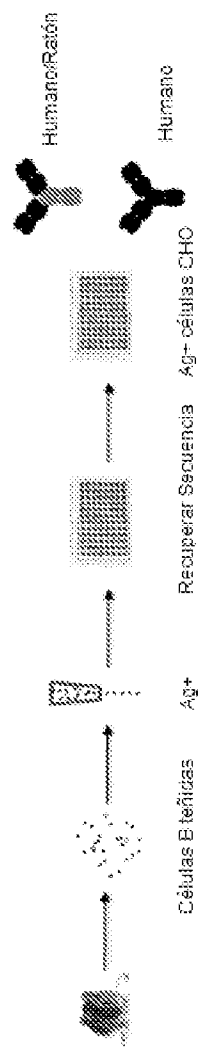


Figura 2

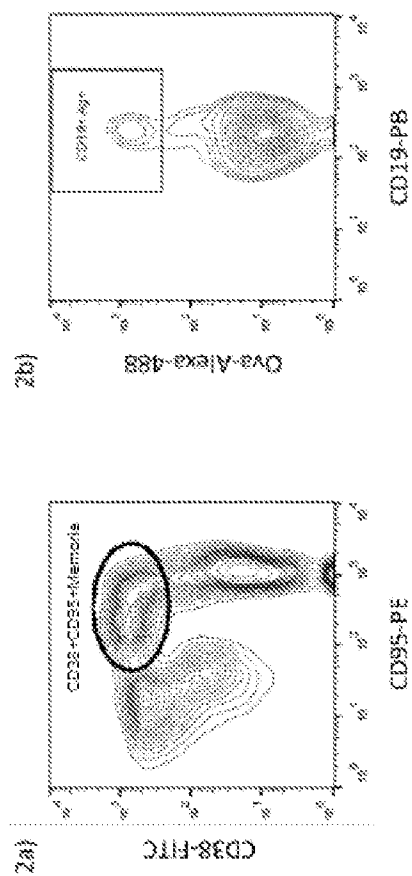


Figura 3

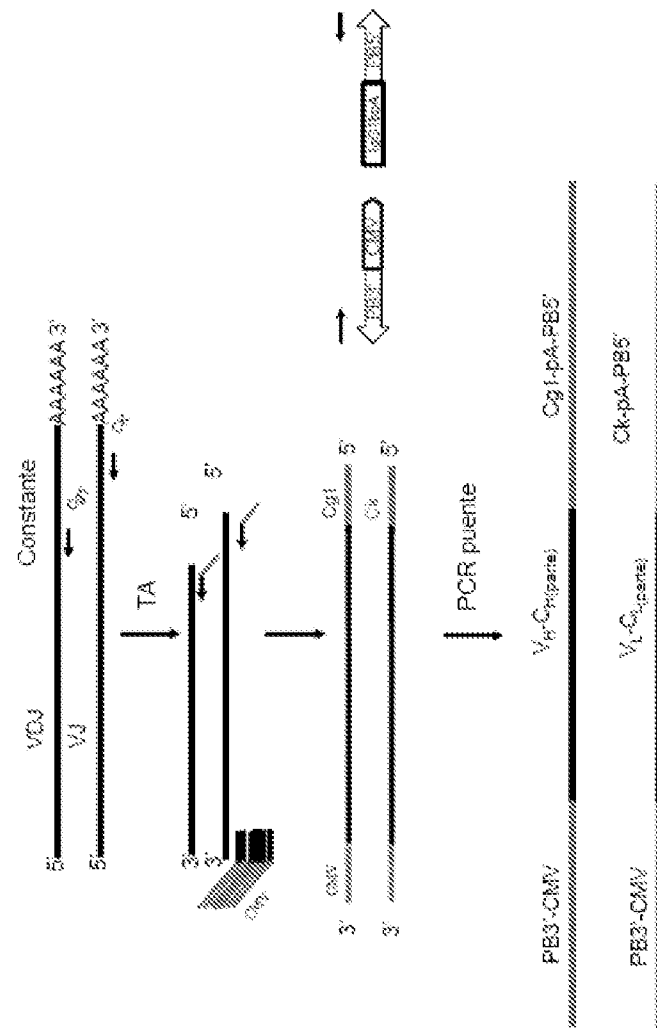


Figura 4

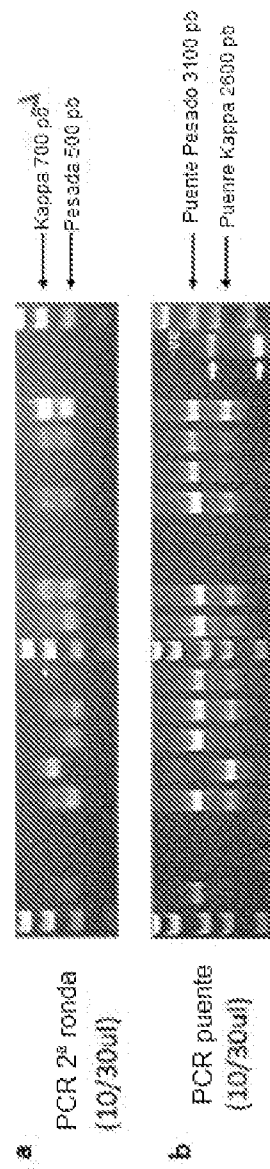
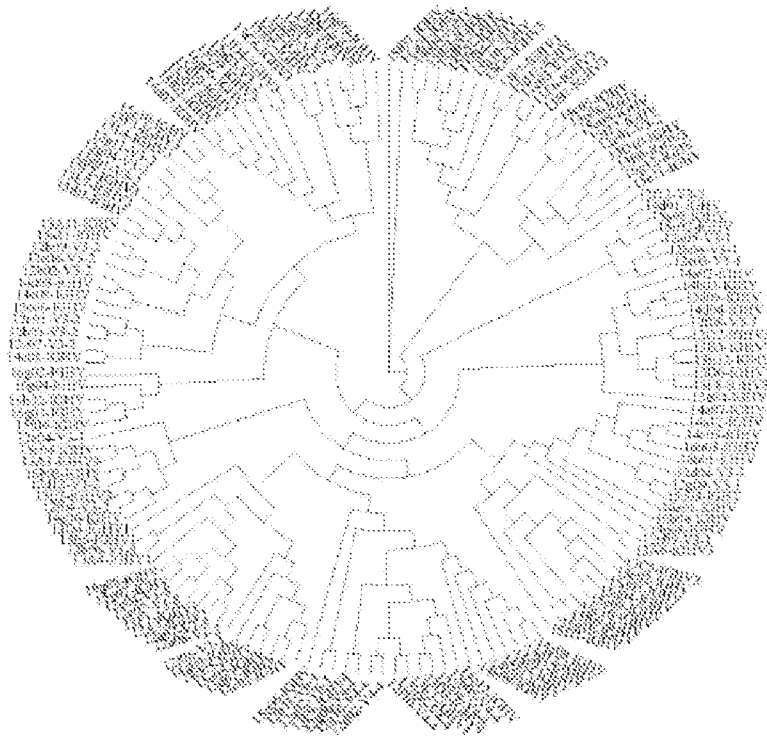




Figura 5



De acuerdo con secuencias de Cadena Pesada

Figura 6

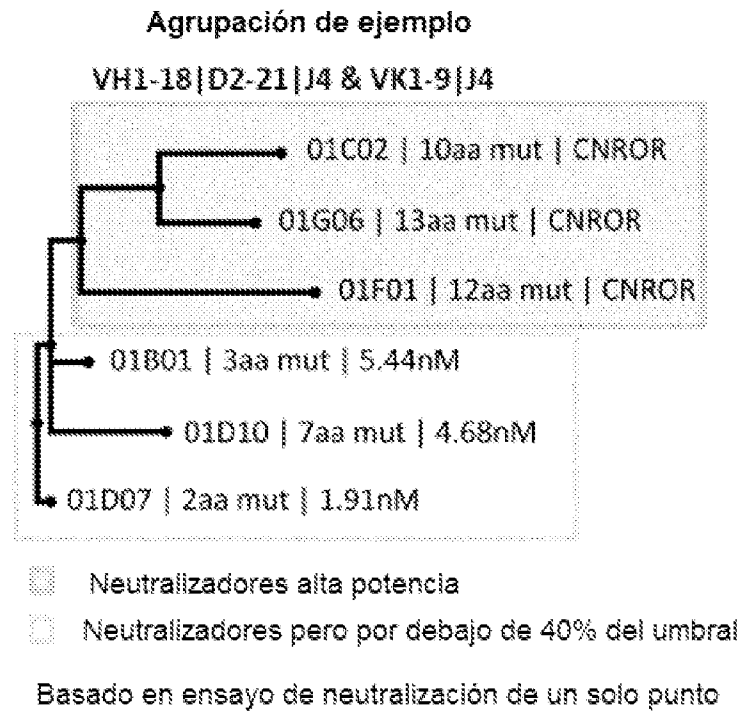


Figura 7

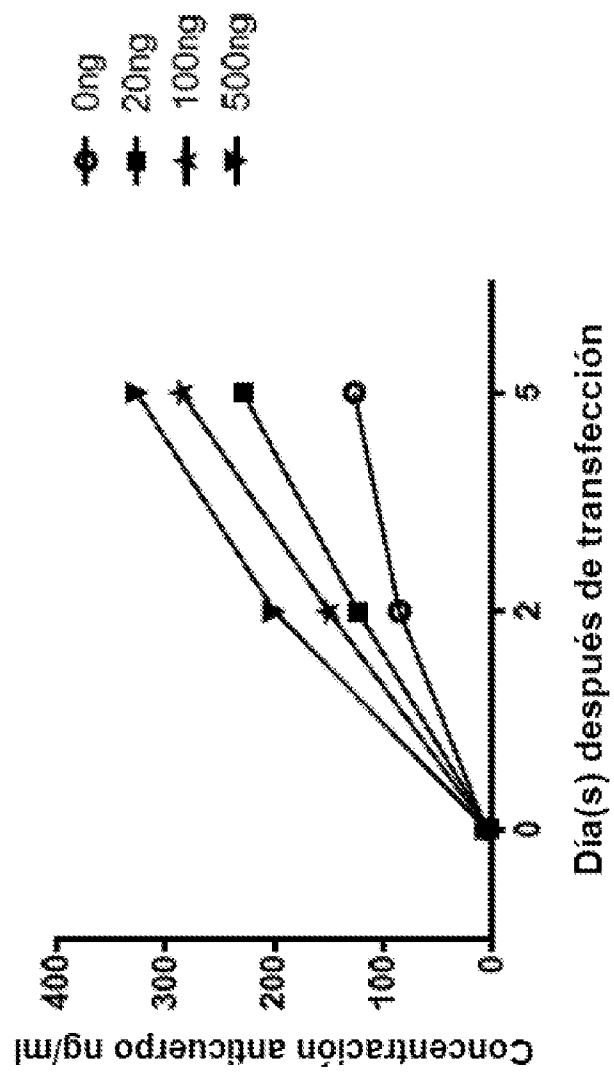


Figura 8

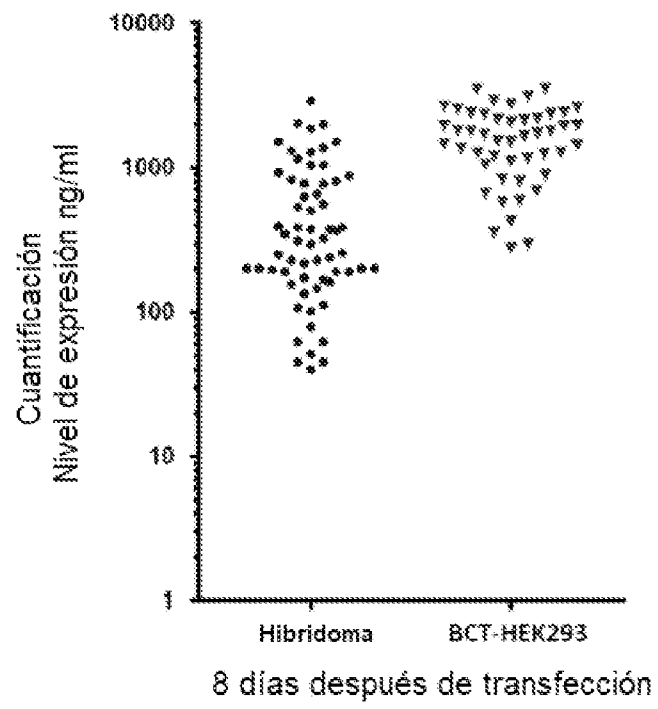


Figura 9a

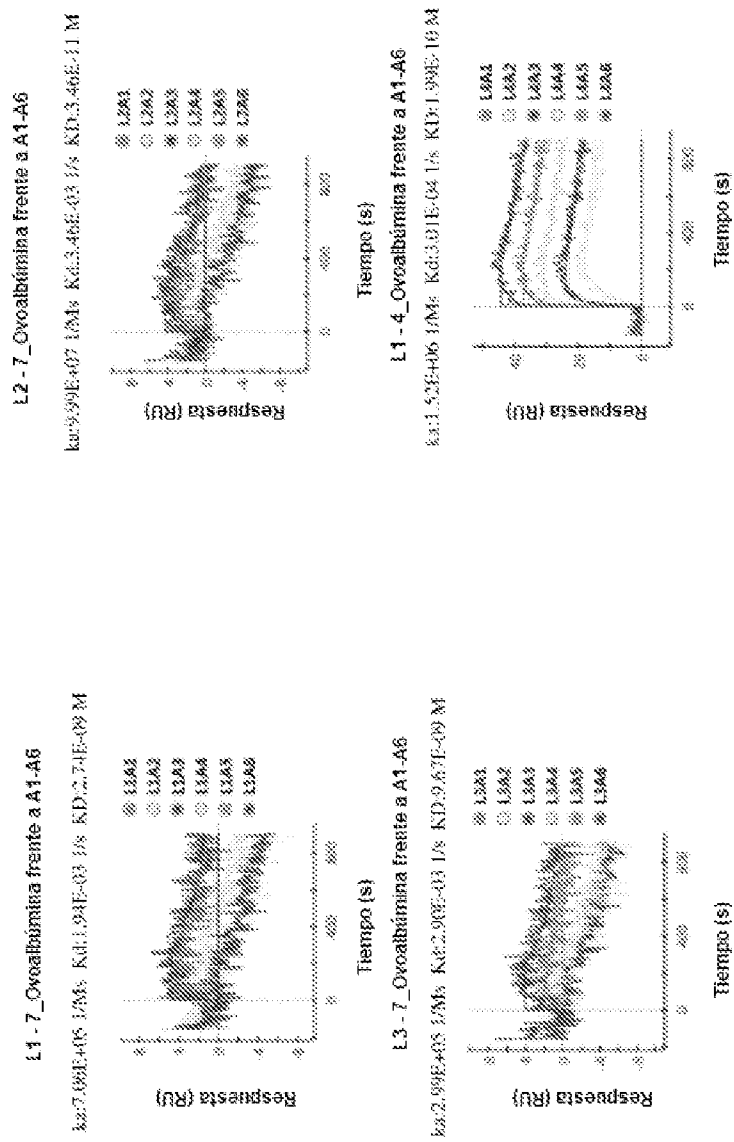


Figura 9a continúa

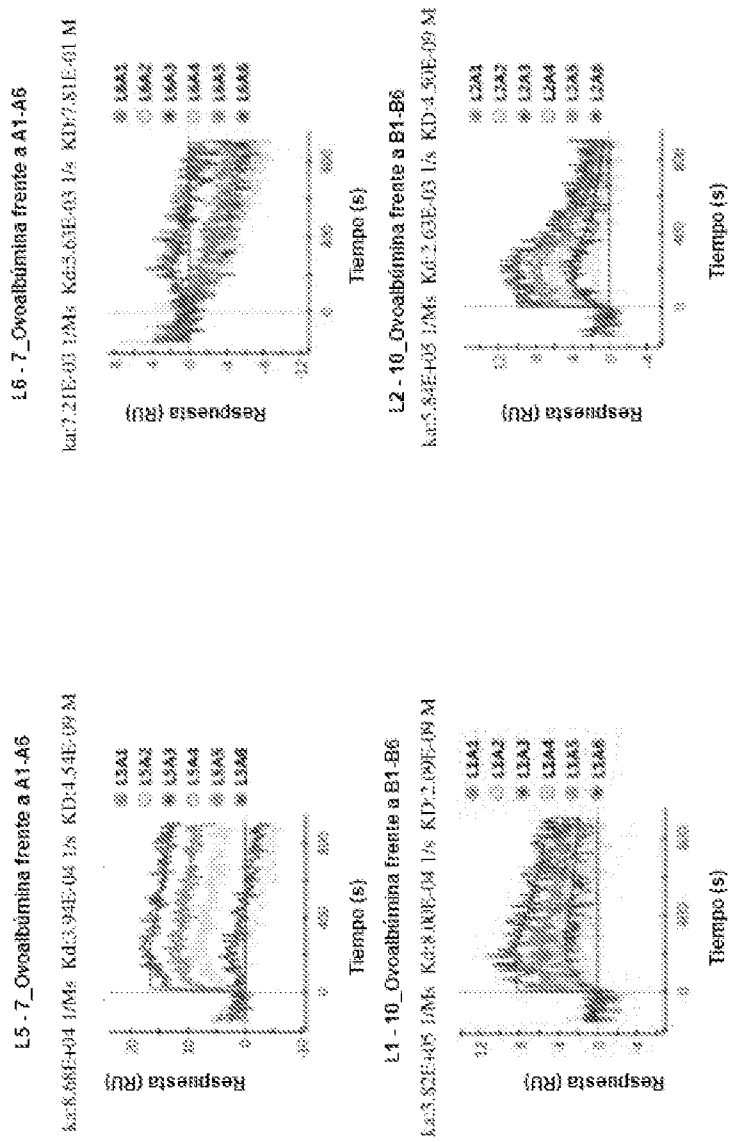


Figura 9a continúa

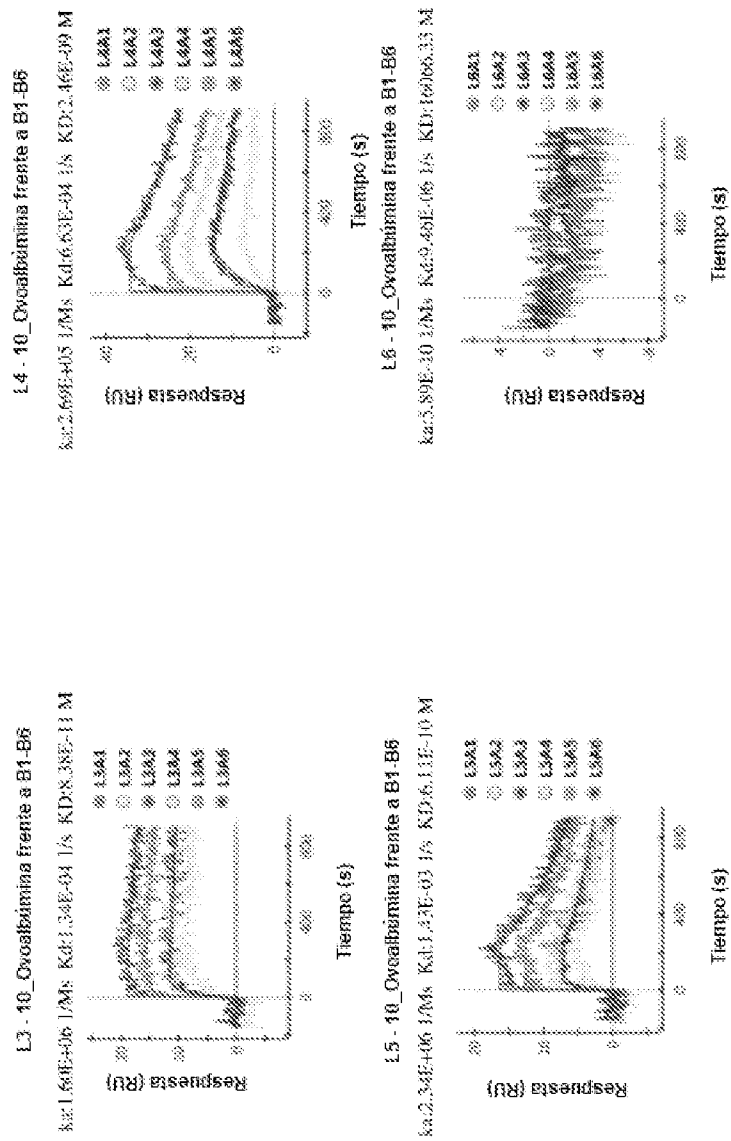


Figura 9a continúa

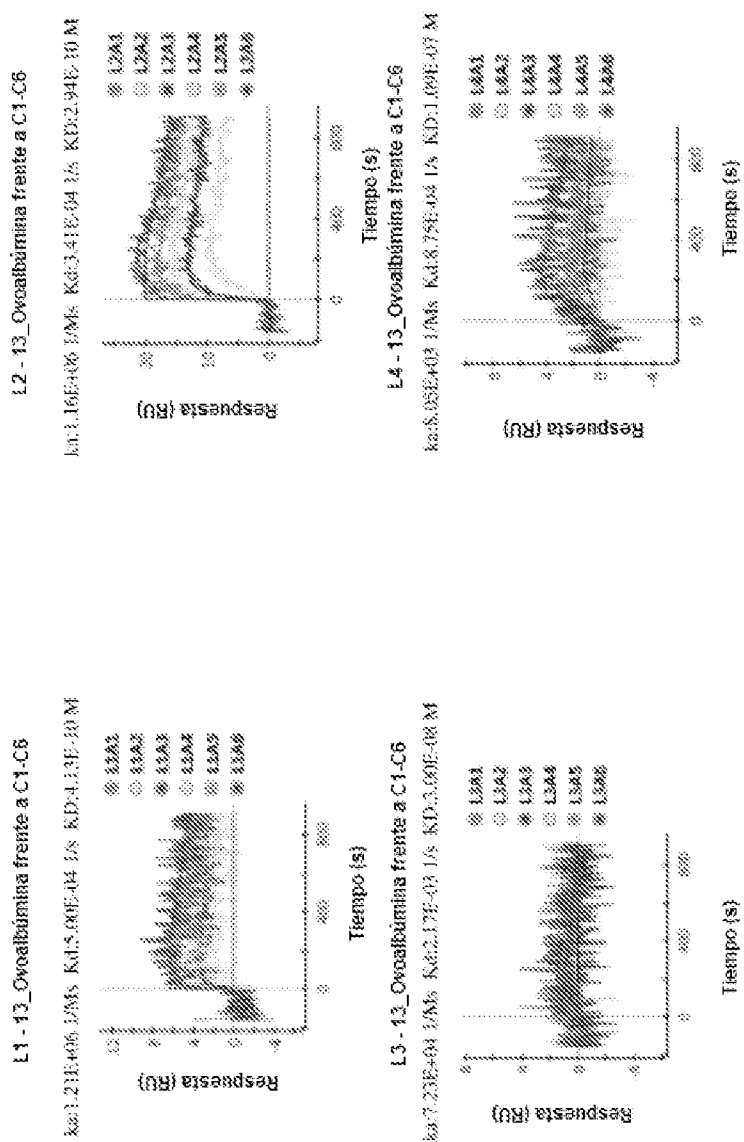




Figura 9a continúa

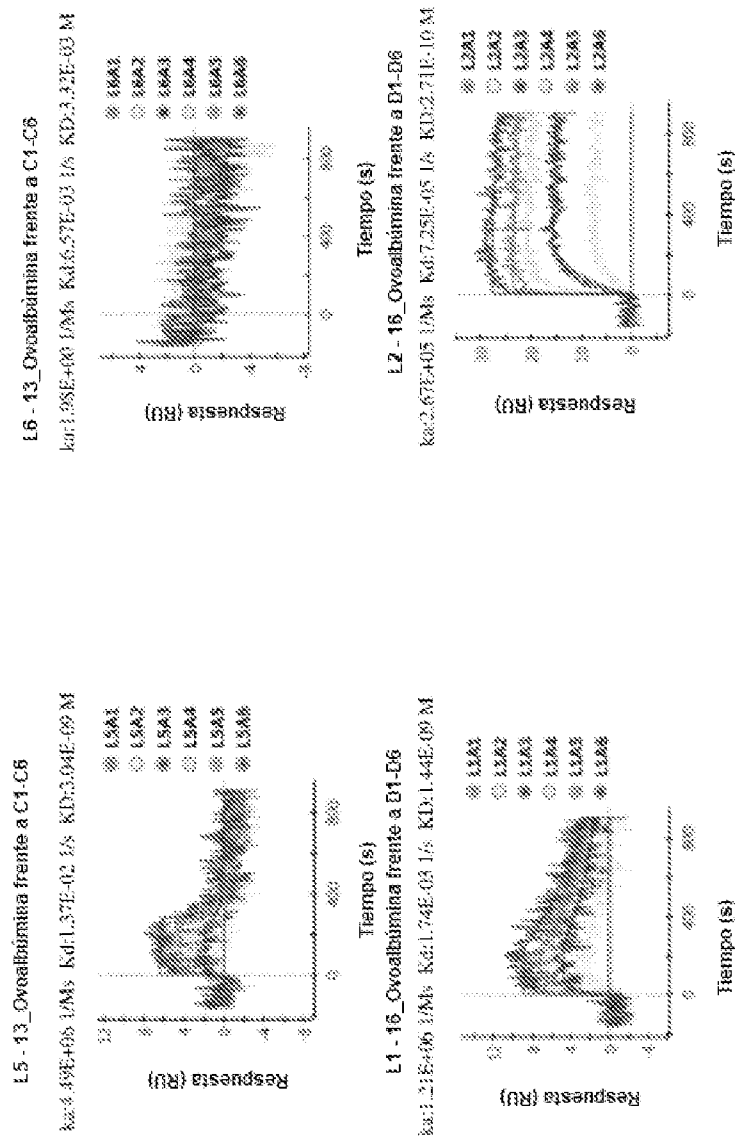


Figura 9a continúa

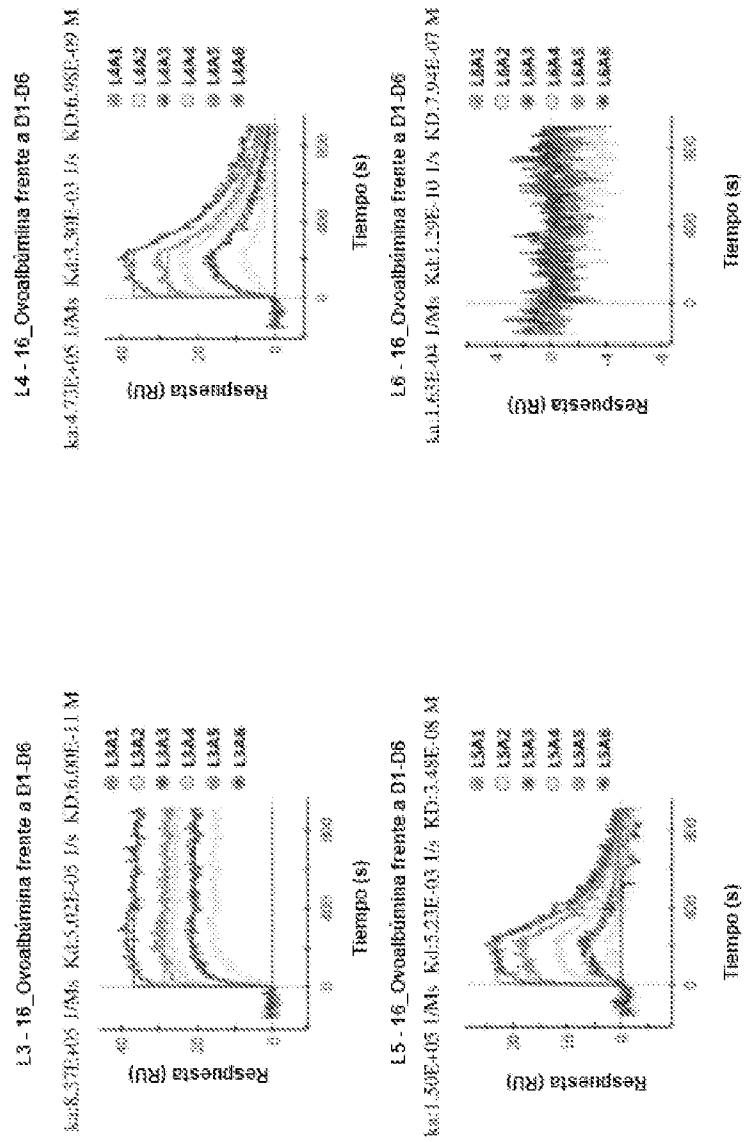


Figura 9a continúa

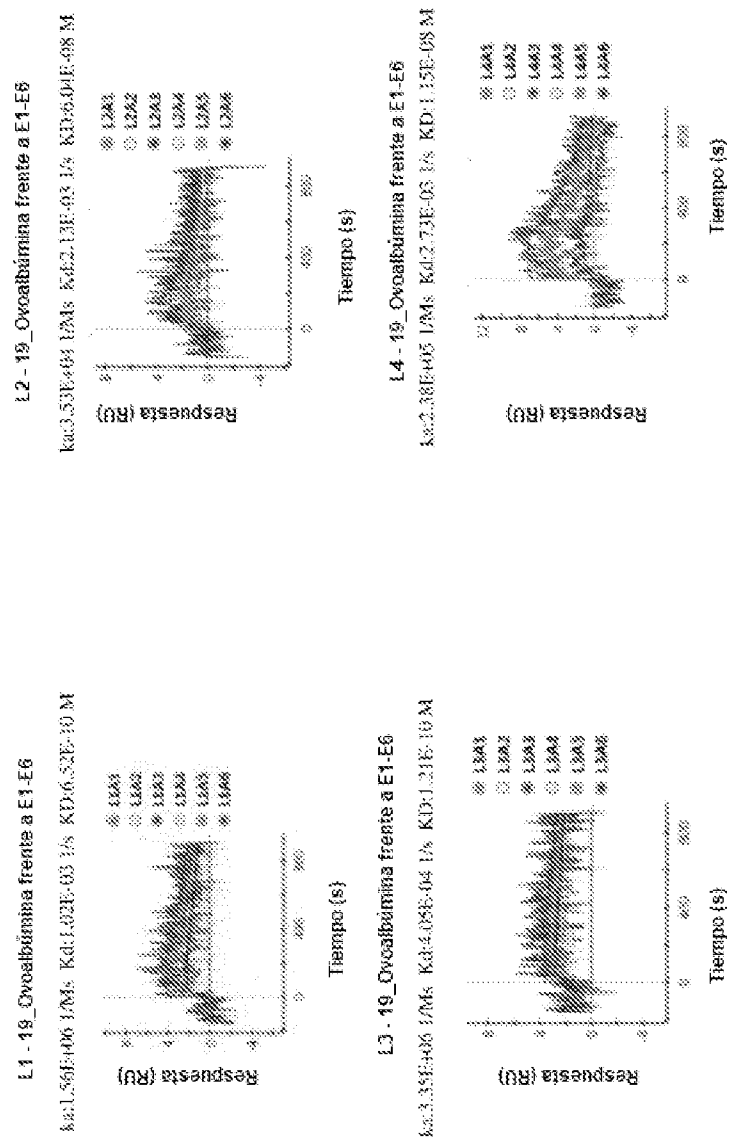


Figura 9a continúa

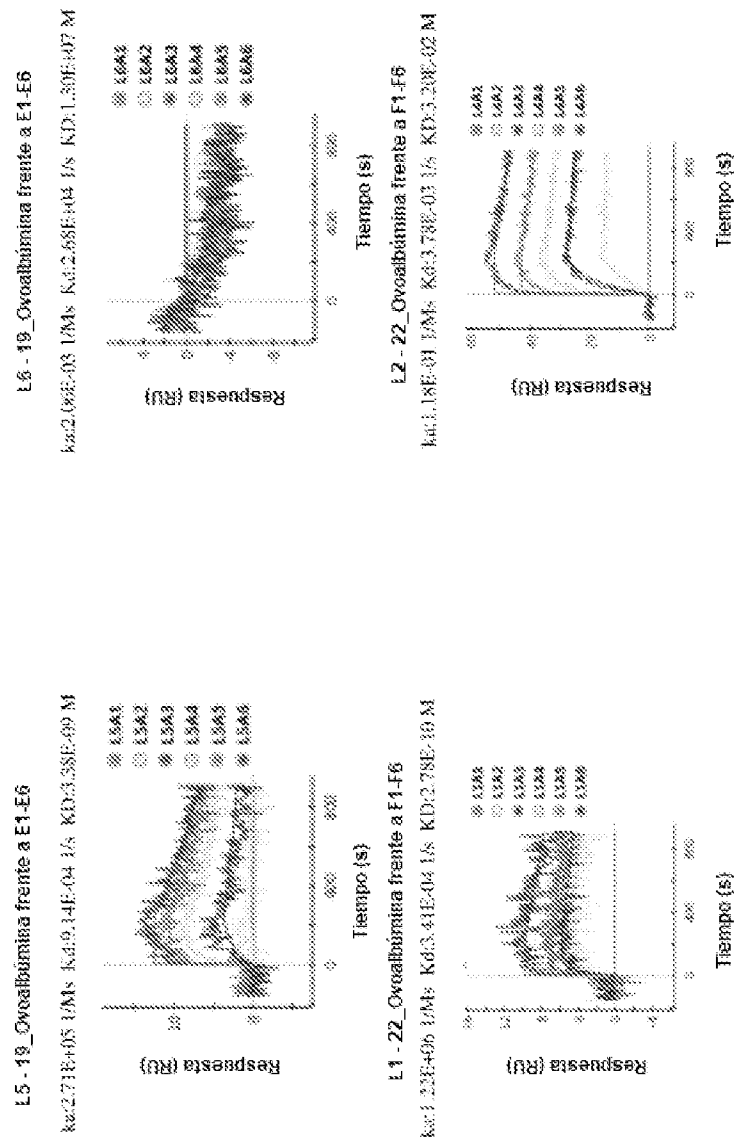


Figura 9a continúa

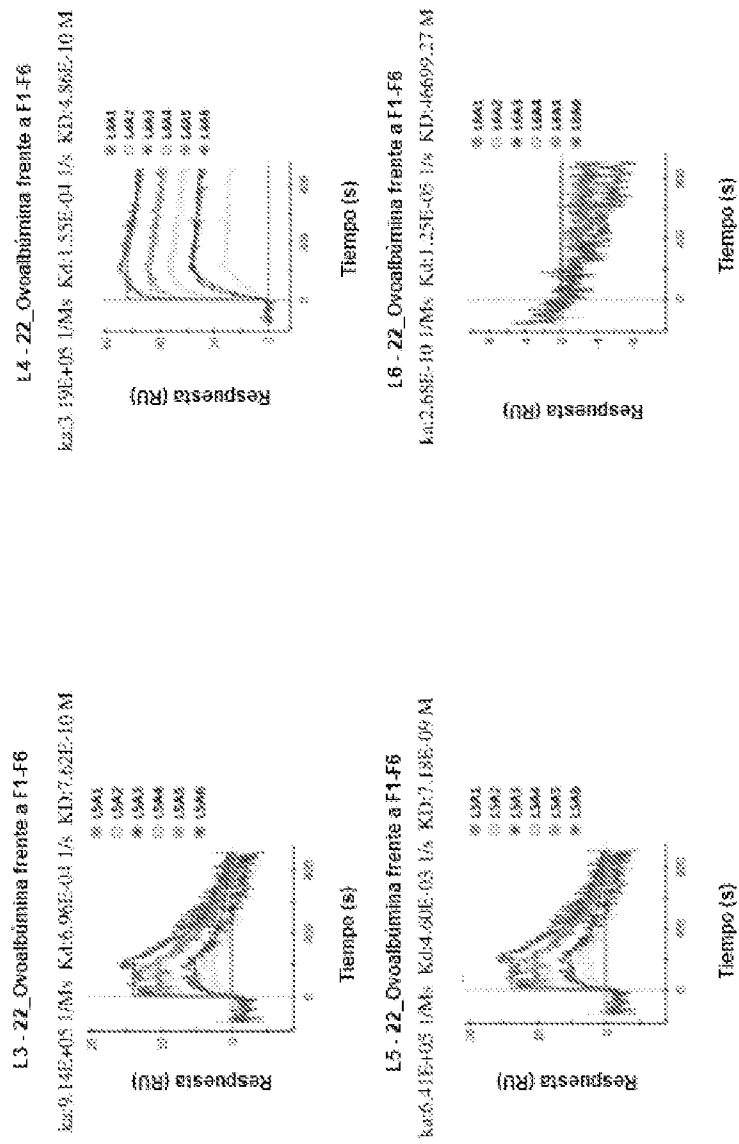


Figura 9a continúa

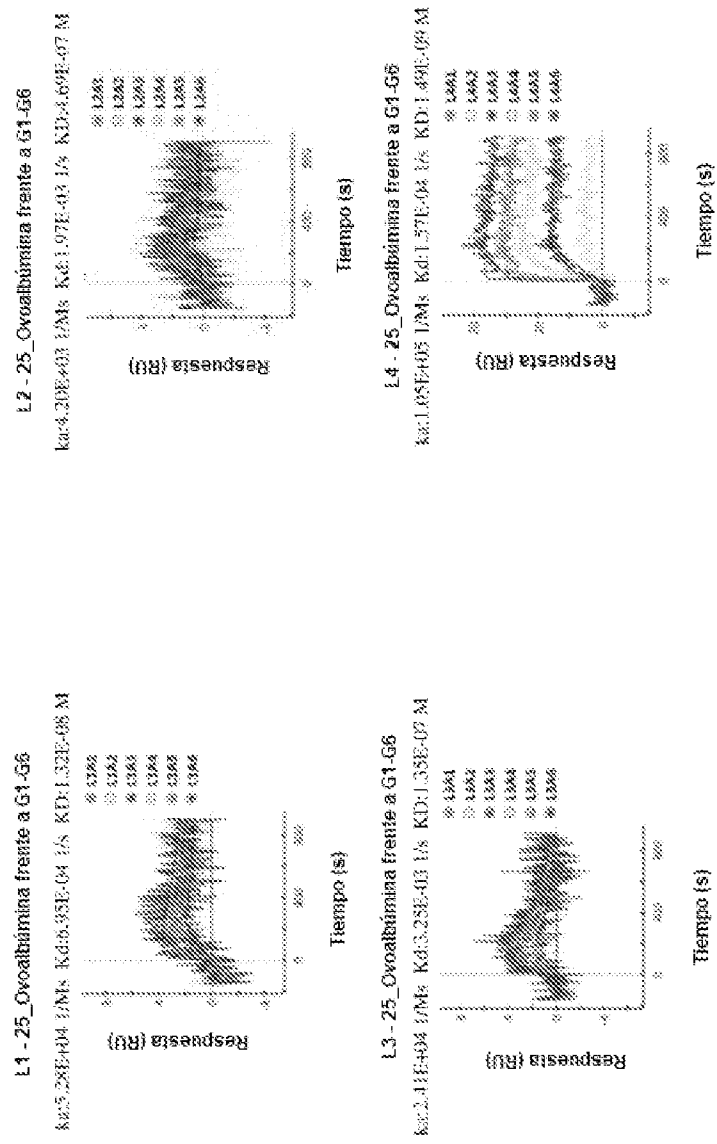


Figura 9a continúa

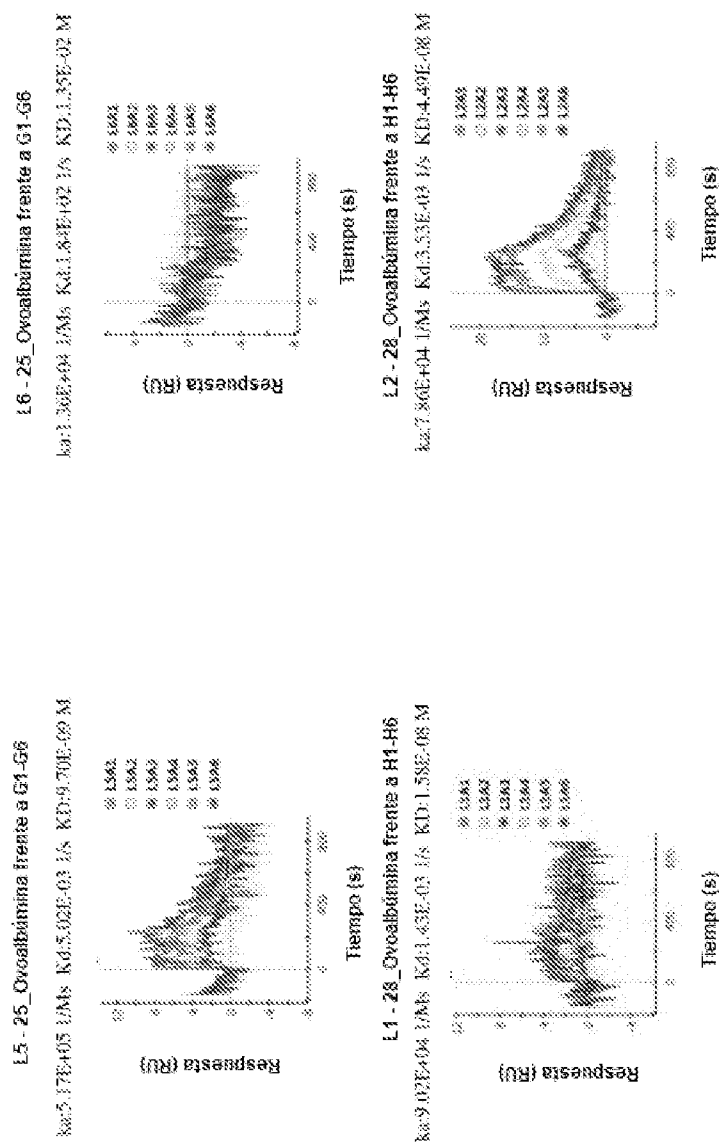


Figura 9a continúa

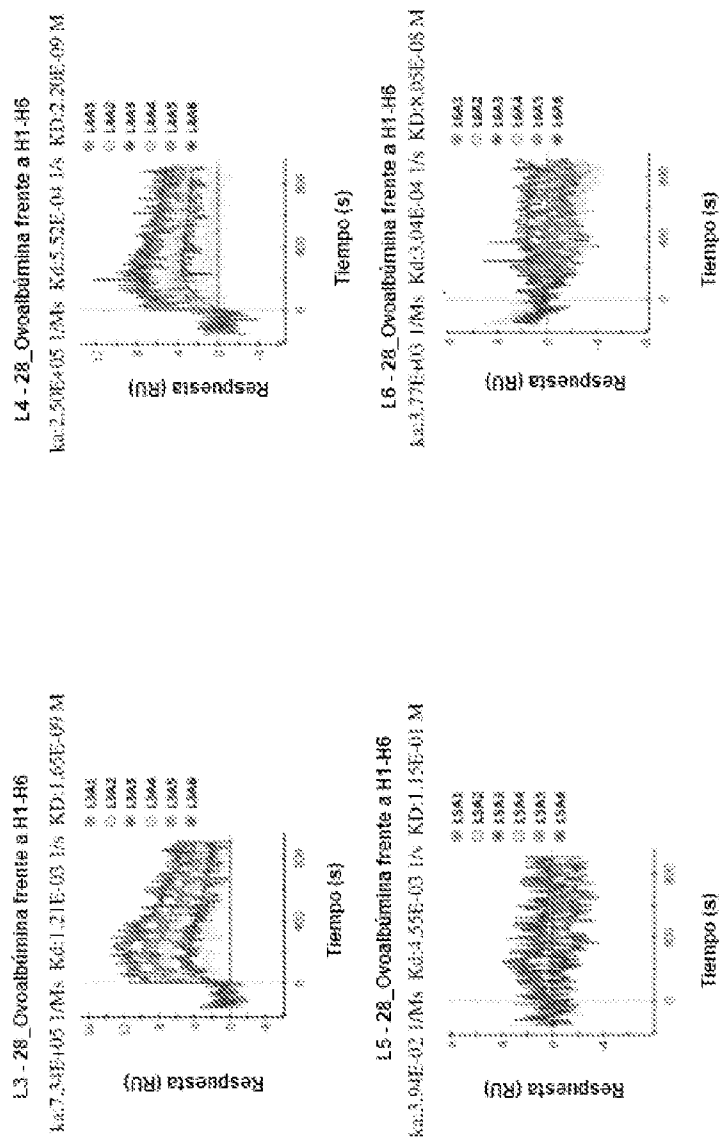




Figura 9a continúa

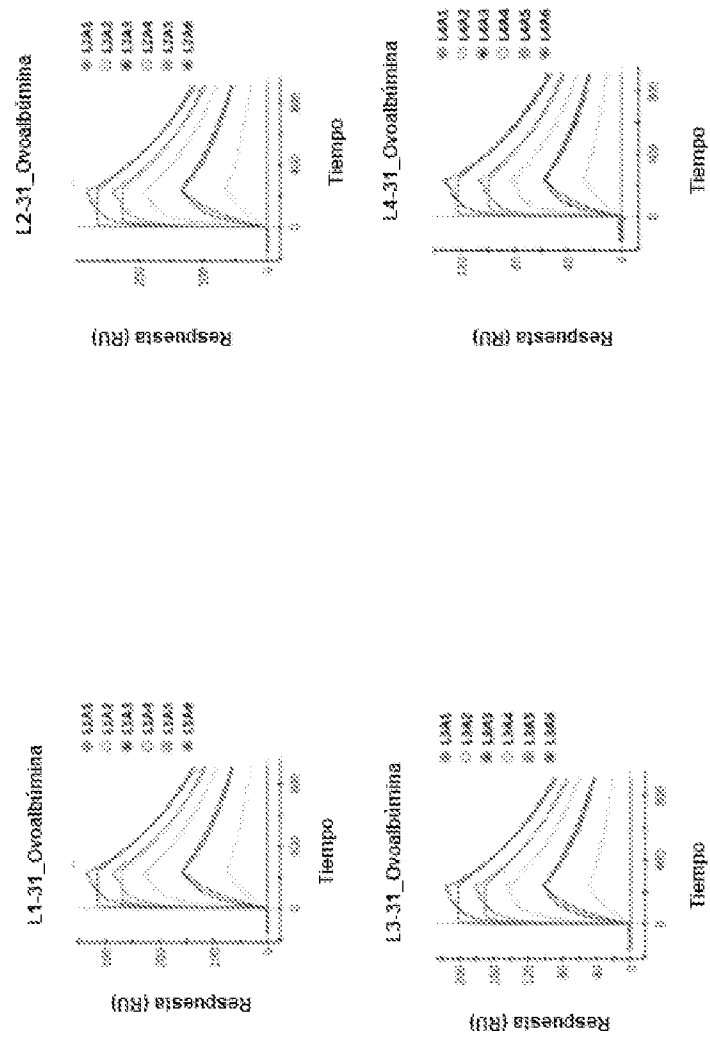


Figura 9a continúa



Figura 9B

1.1-16\_Ovoalbumina frente a D1-B6

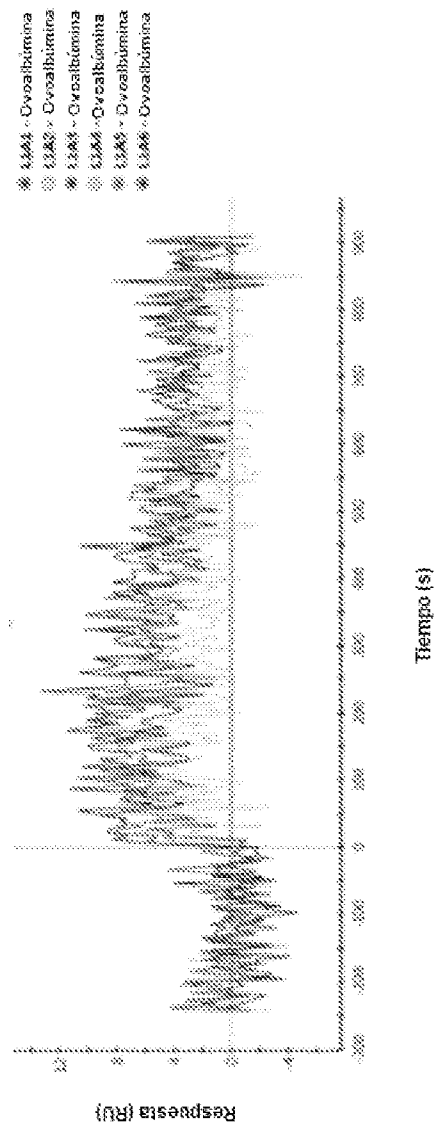


Figura 9B continúa

12-16\_Ovalbúmina frente a D1-D6

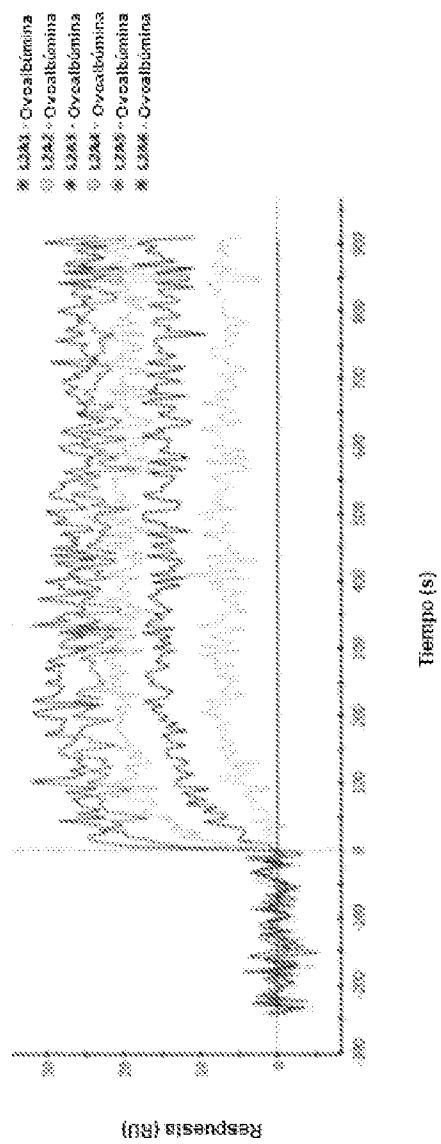


Figura 9B continúa

L3-16\_Ovalbumina frente a D1-D6

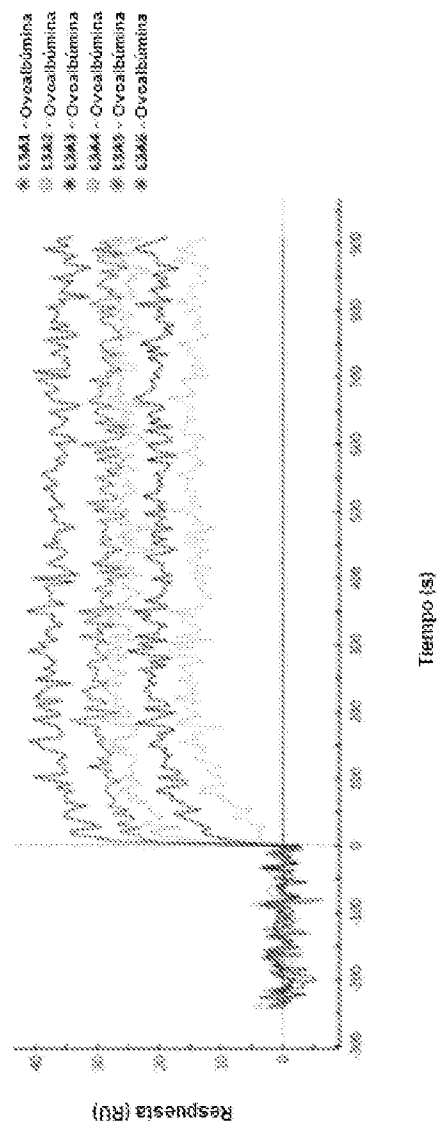


Figura 9B continúa

L4-16\_Ovalbúmina frente a D1-D6

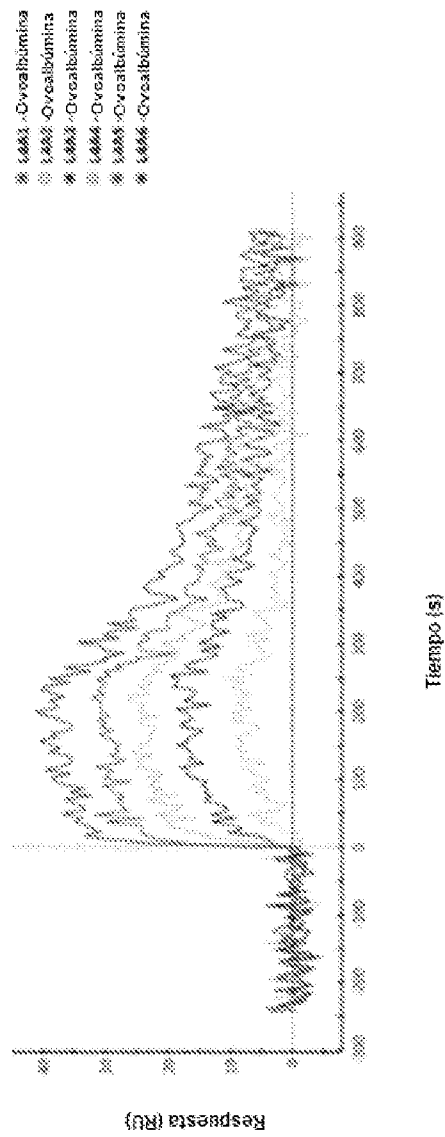


Figura 9B continúa

15-16\_Oxalbumina frente a D1-E6

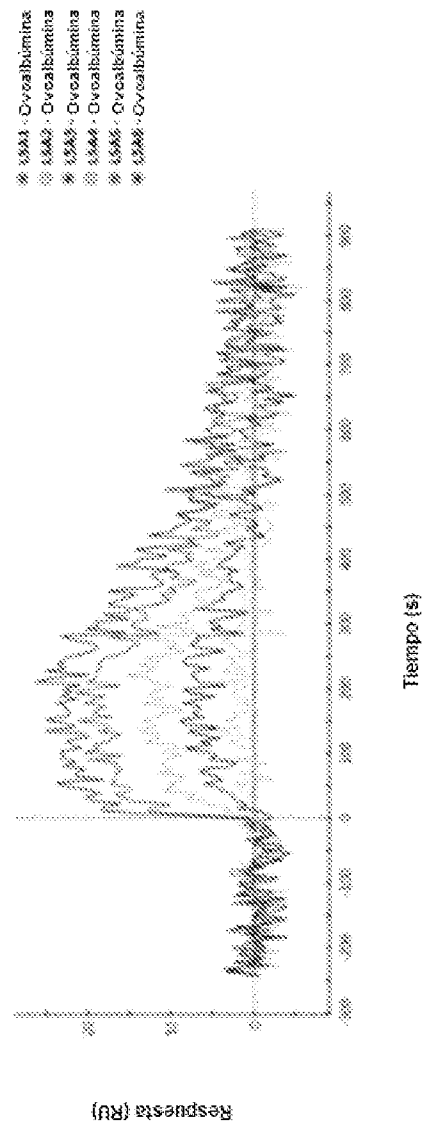


Figura 98 continúa

L6-16\_Ovoalbumina frente a D1-D6

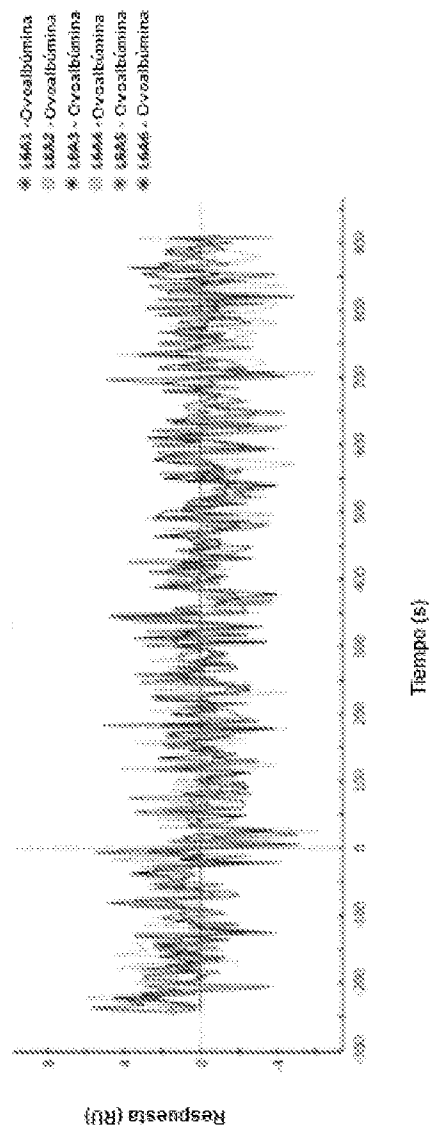




Figura 10

