

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 549 440**

51 Int. Cl.:

C07K 16/30 (2006.01)

C07K 16/00 (2006.01)

C07K 16/18 (2006.01)

C12N 15/63 (2006.01)

C12N 15/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.05.1999 E 09151233 (5)**

97 Fecha y número de publicación de la concesión europea: **22.07.2015 EP 2067788**

54 Título: **Bibliotecas de fragmentos de Fab y métodos para su uso**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.10.2015

73 Titular/es:

**DYAX CORP. (100.0%)
55 Network Drive
Burlington, MA 01803, US**

72 Inventor/es:

**HOOGENBOOM, HENDRICUS RENERUS
JACOBUS MATTHEUS**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 549 440 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

Bibliotecas de fragmentos de Fab y métodos para su uso

- 5 La presente invención se refiere en general a métodos para construir bibliotecas de presentación de fagos de fragmentos Fab humanos, tal como se define en las reivindicaciones.

10 La presentación en un fago filamentosos, en combinación con la selección, forma una herramienta poderosa para la identificación de fármacos basados en péptidos o proteínas (Winter *et al.*, 1994; Clackson *et al.*, 1994). De estos, los anticuerpos son de especial interés, debido a su capacidad para reconocer una variedad de dianas con alta especificidad y afinidad. En particular, el uso de anticuerpos humanos parciales o completos, que no provocan o provocan una respuesta inmunitaria mínima cuando se administran a los pacientes, está produciendo una lista creciente de fármacos basados en proteínas aprobados por la FDA (Holliger *et al.* 1998). La tecnología de presentación de fagos permite la generación de grandes repertorios de anticuerpos humanos (Marks *et al.*, 1991, Hoogenboom *et al.*, 1992; Griffiths *et al.*, 1993; Vaughan *et al.*, 1996), y los procedimientos de bioselección permiten la selección de anticuerpos individuales con una especificidad deseada.

20 Las claves del éxito de la tecnología fueron dos observaciones críticas: (i) la expresión de fragmentos funcionales de anticuerpos mediante la secreción en el periplasma de *E. coli* (Better *et al.*, 1988; Skerra *et al.*, 1988) y (ii) el acceso rápido a grupos génicos de la región variable por medio de la reacción en cadena de la polimerasa (Larrick *et al.*, 1989; Ward *et al.*, 1989; Marks *et al.*, 1991). Para la construcción de bibliotecas de anticuerpos, se amplifican genes V a partir del ADNc de las células B y genes de la cadena pesada y ligera se combinan aleatoriamente y se clonan para codificar una biblioteca combinatoria de fragmentos Fv monocatenarios (scFv, *single chain Fv*) o Fab (Marks *et al.*, 1991; Clackson *et al.* 1991; Persson *et al.*, 1991; Orum *et al.*, 1993). El repertorio de anticuerpos naturales primarios (no seleccionados) dentro de las células B contiene una gran matriz de anticuerpos que reconocen una variedad de antígenos; esta matriz puede clonarse como un repertorio 'virgen' de genes reorganizados, recogiendo los genes V del ARNm de la IgM de las células B de donantes humanos no inmunizados, aislados de linfocitos de sangre periférica (Marks *et al.*, 1991), médula ósea o amígdalas (Vaughan *et al.*, 1996) o de fuentes animales similares (Gram *et al.*, 1992). Este procedimiento proporciona el acceso a anticuerpos para los que no se ha encontrado aún un antígeno, aunque la frecuencia de estos anticuerpos genuinos de 'línea germinal' tendrá una fuerte dependencia de la fuente de células B (Klein *et al.*, 1997). Una sola biblioteca 'virgen', si es suficientemente grande y diversa, puede, en efecto, usarse para generar anticuerpos contra un gran panel de antígenos, incluyendo autoantígenos y antígenos no inmunogénicos y relativamente tóxicos (Griffiths *et al.*, 1993; Marks *et al.*, 1991). En un enfoque distinto, los anticuerpos pueden construirse artificialmente, mediante un ensamblaje *in vitro* de los segmentos génicos V y los segmentos D/J, produciendo anticuerpos 'sintéticos' (Hoogenboom *et al.*, 1992). Una desventaja principal de estos procedimientos es que a partir de las bibliotecas iniciales 'vírgenes' y 'sintéticas', únicamente se aislaron anticuerpos de afinidad moderada (Marks *et al.*, 1991; Nissim *et al.*, 1994). A lo largo de los últimos años se han desarrollado técnicas más eficientes para construir bibliotecas más grandes de fragmentos de anticuerpos, usando métodos de recombinación *in vivo* sofisticados (Griffiths *et al.*, 1993) o procedimientos de clonación de fuerza bruta (Vaughan *et al.*, 1996; Sheets *et al.*, 1998). Dichas grandes bibliotecas han producido un mayor número de anticuerpos humanos por antígenos ensayados, con una afinidad promedio mucho más alta (por encima de sub-nanomolar). Sin embargo, las restricciones técnicas en el tamaño de las bibliotecas que pueden obtenerse o manejarse en la selección, la pérdida de diversidad de las bibliotecas tras la amplificación de la biblioteca y el análisis relativamente largo de la ruta aguas abajo de los anticuerpos seleccionados, es decir, análisis de afinidad a gran escala, ha limitado la difusión de estas bibliotecas como herramientas genéricas para la generación de anticuerpos.

50 La mayor parte de las bibliotecas producidas hasta la fecha usan el formato monocatenario para la presentación de fagos (Vaughan *et al.*, 1996; Sheets *et al.* 1998). Un informe describió el uso de una biblioteca de Fab humanos vírgenes en fagos (sin permitir la exploración inmediata de los fragmentos Fab solubles seleccionados) (Griffiths *et al.*, 1994). Los scFv tienen tendencia a formar dímeros y multímeros de orden más alto de un modo dependiente del clon y relativamente impredecible (Weidner, *et al.* 1992; Holliger, *et al.* 1993; Marks *et al.*, 1993). En consecuencia, el ensayo de afinidad usado, (tal como análisis por BIAcore) a menudo necesita la purificación de los fragmentos de anticuerpos seleccionados. Por ejemplo, la clasificación de las velocidades de disociación, usando BIAcore no es fácilmente posible con fragmentos scFv no purificados; la fracción monomérica de los clones scFv seleccionados necesitan purificarse primero mediante cromatografía de afinidad y filtración en gel (Sheets *et al.*, 1998; Schier *et al.*, 1996).

60 Tal como postularon y observaron Griffiths y colaboradores (Griffiths *et al.*, 1994), el tamaño de la biblioteca de anticuerpos determina la posibilidad de la selección de anticuerpos de alta afinidad con respecto al antígeno. La comparación del primer repertorio de scFv vírgenes que contenía $2,9 \times 10^7$ clones (Marks *et al.*, 1991), con un repertorio de scFv recientemente construido de aproximadamente 10^{10} clones (Vaughan *et al.*, 1996; Sheets *et al.* 1998), confirma este postulado: el aumento del tamaño de la biblioteca en 500 veces dio como resultado afinidades aproximadamente 100 veces más altas. Este aumento se produce disminuyendo las velocidades de disociación de 10^{-1} - 10^{-2} s⁻¹, para los fragmentos seleccionados de la biblioteca de menor tamaño, a 10^{-3} - 10^{-4} s⁻¹ para los seleccionados de la biblioteca de mayor tamaño.

Huse *et al.* Science 246:1275-1281 (1989) divulgan la generación de una biblioteca combinatoria grande del repertorio de inmunoglobulinas en fago Lambda.

5 Es un objeto de la invención crear una biblioteca de Fab que sea una fuente valiosa de anticuerpos para muchas dianas distintas y que desempeñará un papel vital en el descubrimiento de dianas y la validación en el área de la genómica funcional.

La invención proporciona un método para fabricar una biblioteca de Fab, comprendiendo la biblioteca una pluralidad de vectores en la que cada vector de la pluralidad de vectores comprende:

- 10
- una primera región de clonación y una segunda región de clonación, en las que
 - cada región de clonación comprende al menos un sitio de escisión de enzimas de restricción único para el vector,
 - estando cada región de clonación flanqueada en el extremo 5' por un sitio de unión a ribosomas y una secuencia señal,
 - 15 - un polinucleótido que codifica una región de anclaje, localizada en el extremo 3' de la segunda región de clonación,
 - un polinucleótido que codifica una región constante de la cadena pesada o una porción de la misma localizada entre el sitio de restricción de enzimas único para el vector de la segunda región de clonación y la región de anclaje,
 - 20 - un polinucleótido que codifica un marcador,

comprendiendo el método:

25 introducir en la primera región de clonación un miembro de una primera pluralidad de polinucleótidos variables, codificando dicha pluralidad de polinucleótidos variables una primera pluralidad de polipéptidos, en los que el miembro de la primera pluralidad de polinucleótidos codifica un polipéptido que comprende una región variable de la cadena ligera del anticuerpo posiblemente seguida por una región constante de la cadena ligera del anticuerpo, introduciendo por separado en la segunda región de clonación de cada vector un miembro de una

30 segunda pluralidad de polinucleótidos variables, codificando dicha pluralidad de polinucleótidos variables una segunda pluralidad de polipéptidos, en la que el miembro de la segunda pluralidad de polinucleótidos codifica un polipéptido que comprende una región variable de cadena pesada de un anticuerpo, e introducir cada uno de los vectores en una célula hospedadora para proporcionar una pluralidad de partículas de la cápside comprendiendo cada una un miembro de la primera pluralidad de polinucleótidos variables y un

35 miembro de la segunda pluralidad de polinucleótidos, para fabricar de este modo la biblioteca de anticuerpos o fragmentos de los mismos.

Se entenderá que la expresión " sitio de escisión de enzimas de restricción único para el vector " se refiere a la presencia de uno de dicho un sitios de restricción en la secuencia del vector, sin tener en cuenta la posible

40 presencia de dicho un sitio en el primer y/o segundo de los polinucleótidos mencionados anteriormente, que codifican una región variable completa del anticuerpo o una parte de una región variable de un anticuerpo, posiblemente seguida por una región constante completa de un anticuerpo o una parte de una región constante de un anticuerpo. Dicho primer y segundo polinucleótido puede comprender sitios de restricción idénticos al sitio "único". Esto significa que dicho sitio de restricción fue "único" antes de que se clonaran tanto la primera como la

45 segunda secuencia de polinucleótidos en el vector.

El primer y segundo polinucleótido variable se clona preferentemente en la región de clonación en una orientación predeterminada. Por lo tanto, en el caso de que la región de clonación comprenda un solo único sitio de restricción, este sitio es preferentemente de tal tipo que se generan extremos de restricción no idénticos, tales como, por

50 ejemplo, los generados por la enzima de restricción *SfiI*. Sin embargo, la región de clonación puede comprender dos o más sitios de restricción únicos, de modo que los polinucleótidos variables pueden clonarse convenientemente como un fragmento de restricción que tiene los extremos correspondientes.

Preferentemente, en el vector usado en el método de acuerdo con la invención, la primera y la segunda región de clonación, ambos sitios de unión a ribosomas, las secuencias señal y la secuencia de anclaje son parte de una sola

55 secuencia polienlazadora. Ambas regiones de clonación pueden, por lo tanto, ser parte de un solo casete, que comprende la primera región de clonación, flanqueada en el extremo 5' por un sitio de unión a ribosomas y una secuencia señal, situada adyacente a la segunda región de clonación, también flanqueada en el extremo 5' por su sitio de unión a ribosomas correspondiente y una secuencia señal, y flanqueada en el extremo 3' por la secuencia de anclaje.

60

Preferentemente, la primera pluralidad de polinucleótidos variable es de polinucleótidos V_L , y la segunda pluralidad de polinucleótidos variables es de polinucleótidos V_H . Más preferentemente, los polinucleótidos V_L son polinucleótidos V_K , polinucleótidos $V_K C_K$, polinucleótidos V_λ , polinucleótidos $V_\lambda C_\lambda$, una mezcla de polinucleótidos V_K

65 y V_λ o una mezcla de polinucleótidos $V_K C_K$ y $V_\lambda C_\lambda$.

En los polinucleótidos usados en el método de acuerdo con la invención, el vector comprende adicionalmente un marcador para la purificación o detección de un anticuerpo, comprendiendo dicho marcador para la purificación del anticuerpo preferentemente una cola de polihistidina; el marcador para la detección del anticuerpo es preferentemente un marcador derivado de *c-myc*.

5 En otra realización de los polinucleótidos usados en el método de acuerdo con la invención, el vector comprende adicionalmente un codón de terminación ámbar localizado entre el segundo polinucleótido variable y la proteína de anclaje.

10 En otra realización más de los polinucleótidos usados en el método de acuerdo con la invención, el vector comprende además un dominio C_{H1} localizado entre el segundo polinucleótido variable y la proteína de anclaje, siendo el dominio C_{H1} preferentemente un dominio C_{H1} gamma-1 humano.

15 La "proteína de anclaje" se define como una proteína o una parte de la misma que puede acomodarse al menos parcialmente en el recubrimiento externo de una partícula generada por un organismo que expresa la biblioteca, tal como una partícula fágica o vírica, o en el recubrimiento externo de un organismo propio, en el caso de que el propio organismo exprese la biblioteca. El recubrimiento externo se define en el presente documento como la estructura de una célula, partícula vírica o fágica que define la superficie externa de la misma. En el caso de un fago o fagémido que expresa la biblioteca, la proteína de anclaje puede ser una proteína de recubrimiento, tal como, el producto del
20 gen III. Sin embargo, pueden usarse otros sistemas conocidos por la persona experta, para obtener una biblioteca de acuerdo con la presente invención. Por lo tanto, por ejemplo, pueden contemplarse proteínas transmembrana, o el dominio transmembrana de las mismas, para su uso como la proteína de anclaje en sistemas de expresión eucariota. En la invención, la proteína de anclaje puede fusionarse a una región variable de anticuerpo o a una parte de la misma, dando como resultado la presentación de dicha región variable en el entorno externo del organismo,
25 estando la región anclada a su recubrimiento externo. En una realización preferida de los polinucleótidos usados de acuerdo con la invención, la proteína de anclaje es una proteína de recubrimiento menor III de un fago filamentoso f_{ϕ} . En una realización de la invención, los polinucleótidos usados en el método de acuerdo con la invención, y por lo tanto la biblioteca de Fab, codifican al menos 10^9 Fab distintos. En otra realización de la invención, la biblioteca de Fab codifica al menos 10^{10} Fab distintos. Aún en otra realización de la invención, la biblioteca de Fab codifica al
30 menos $3,7 \times 10^{10}$ Fab distintos. Aún en otra realización de la invención, la biblioteca de Fab codifica de 10^9 a $3,7 \times 10^{10}$ Fab distintos.

Además, la divulgación proporciona una biblioteca de Fab, que comprende

- 35 - una pluralidad de vectores tal como se define anteriormente,
- la segunda región de clonación en cada vector que forma un polinucleótido de fusión que codifica una pluralidad de proteínas de fusión,
- una pluralidad de partículas de la cápside, en la que la pluralidad de vectores que contiene la primera y la segunda pluralidad de polinucleótidos variables está empaquetada en las partículas de la cápside,
40 en las que
- al menos algunas de las partículas de la cápside presentan la proteína de fusión codificada por el vector empaquetado en la cápside sobre la superficie de la cápside.

45 Además la invención se refiere a un método para producir una pluralidad de polinucleótidos que codifican una biblioteca de Fab, que adicionalmente comprende las etapas de:

- amplificar una primera pluralidad de polinucleótidos variables con un primer conjunto de cebadores,
- amplificar una segunda pluralidad de polinucleótidos variables con un segundo conjunto de cebadores,
50 - en el que cada conjunto de cebadores comprende oligonucleótidos diseñados para que sean homólogos con respecto al extremo 5' y 3' de los polinucleótidos variables que codifican las regiones variables del anticuerpo o partes de las mismas, de modo que estos puedan usarse para amplificar grupos de polinucleótidos variables a partir de fuentes naturales o sintéticas de genes conservando al mismo tiempo todo o parte del sitio de combinación antigénica del anticuerpo.

55 En una realización, el método para construir la biblioteca de Fab comprende las etapas de: amplificar una pluralidad de grupos génicos variables con un conjunto de cebadores, en los que los cebadores comprenden oligonucleótidos diseñados para que sean homólogos con respecto al extremo 5' y 3' de los polinucleótidos variables que codifican las regiones variables del anticuerpo o partes de las mismas, de modo que estos puedan usarse para amplificar grupos de polinucleótidos variables a partir de fuentes naturales o sintéticas de genes conservando al mismo tiempo
60 todos o parte de los sitios de combinación antigénica del anticuerpo; clonar los grupos de genes variables amplificados en un vector con un procedimiento de dos etapas para obtener una biblioteca de Fab; en la que el vector comprende un fago o un vector fagémido que acomodará la expresión de los polinucleótidos variables de anticuerpos clonados como fragmentos Fab de anticuerpos, en los que una de las dos cadenas de anticuerpos se fusiona a una de las proteínas de recubrimiento de fagos (por ejemplo, producto génico III).

65 En una realización, se diseñaron cebadores BACK para que tuvieran a lo sumo tres mutaciones en un total de

veintiuno a veintitrés nucleótidos cuando se comparaban con la región del segmento génico de la línea germinal humana al que tendrían que unirse, pero con al menos 3 restos homólogos hacia el sitio 3' del oligonucleótido. Este conjunto de oligonucleótidos reconocerá aproximadamente el 90 % de los segmentos génicos de la línea germinal humana y como tal proporciona acceso a la mayor parte de la diversidad presente de las células B en fuentes no inmunizadas. En otra realización, los cebadores de la cadena pesada deberían terminar con 'GG' para asegurar la unión estable a altas temperaturas de hibridación (al menos 55 °C). De un modo similar, los cebadores V λ BACK y la mayor parte de los cebadores V λ BACK se diseñarán para terminar preferentemente en 'CC'. En una realización alternativa, los cebadores consisten en la secuencias de la Figura 2.

10 La divulgación también se refiere a un vector tal como se define anteriormente, que comprende una de la primera y una de la segunda pluralidad de los polinucleótidos variables clonados en la primera y la segunda región de clonación respectivamente.

15 La presente divulgación se refiere además a células hospedadoras que contienen las bibliotecas de Fab obtenidas mediante el método de la invención o los polinucleótidos que codifican las bibliotecas de Fab.

20 En un aspecto el método de la invención implica la unión de un miembro de un par de unión específico deseado, tal como una molécula de anticuerpo, a una proteína de recubrimiento fágica. Enriqueciendo después el miembro del par de unión específico, tal como mediante técnicas de afinidad, por ejemplo, también se enriquece y puede aislarse después el ADN que codifica el miembro del par de unión específico. El ADN obtenido de este modo puede clonarse y expresarse después en otros sistemas, produciendo posiblemente grandes cantidades del miembro del par de unión específico deseado, o puede someterse a secuenciación y clonación adicional y a manipulaciones genéticas antes de la expresión.

25 Típicamente la diana miembro del par de unión específico, por ejemplo, un antígeno o un hapteno cuando el miembro del par de unión específico es un anticuerpo, es conocida y los métodos del presente documento proporcionan un medio para crear y/o identificar un miembro de par de unión específico que se una específicamente a la diana de interés. Por lo tanto, cuando la proteína es un anticuerpo, la presente divulgación proporciona un nuevo medio para producir anticuerpos, particularmente anticuerpos monoclonales, con especificidad para dianas predeterminadas, evitando de este modo el proceso laborioso, largo y a menudo impredecible de la tecnología de anticuerpos monoclonales convencional.

Breve descripción de los dibujos

35 Figura 1. Vector fagémido pCES1 para la presentación de fragmentos Fab de anticuerpos. La representación esquemática (A) y la región polienlazadora (B) de pCES1. La región polienlazadora comprende dos secuencias señal ("S"; secuencia líder de p λ B y del gen III), el dominio C κ , el sitio de unión a ribosomas (rbs), el dominio CH1, el marcador de hexahistidina (H6) y una secuencia derivada de *c-myc*. Los genes de dominio variable pueden clonarse como fragmentos *Apa*LI - *Xho*I o *Apa*LI - *Asc* (para VL o VLCL respectivamente) y fragmentos *Sfi*I / *Pst*I - *Bst*EII o *Sfi*I - *Not*I (para VH o VHCH1 respectivamente). El codón de terminación ámbar (*) entre los genes del anticuerpo y el gen III del bacteriófago permite la producción de fragmentos Fab solubles en una cepa no supresora de *E. coli*. La expresión del operón bicistrónico está bajo el control del promotor LacZ (*pLacZ*).

45 Figura 2. Esta figura describe oligonucleótidos usados en una realización para la amplificación por PCR de las regiones V de la cadena pesada y ligera humana.

El término "activo" se refiere a aquellas formas del polipéptido que conservan las actividades biológica y/o inmunológica de cualquier polipéptido de origen natural.

50 La expresión células "activadas" tal como se usa en esta solicitud son aquellas que están implicadas en el tránsito de la membrana extracelular o intracelular incluyendo la exportación de moléculas neurosecretoras o enzimáticas como una parte de un proceso normal o de enfermedad.

55 El término "anticuerpo" significa una inmunoglobulina natural o producida parcial o completamente de un modo sintético. El término también engloba cualquier proteína o polipéptido que tenga un dominio de unión que sea homólogo a un dominio de unión a una inmunoglobulina. Estas proteínas pueden proceder de fuentes naturales o pueden producirse parcial o completamente de un modo sintético. Son ejemplos de anticuerpos los isotipos de inmunoglobulinas y los fragmentos Fab, scFv, Fv, dAab, VHH, Fd.

60 La expresión "dímero polipeptídico de anticuerpo" significa una asociación de dos componentes de cadena polipeptídica de un anticuerpo, capaz de unirse a un antígeno. Por lo tanto, puede haber un brazo de un anticuerpo que consista en una cadena pesada y una cadena ligera, puede haber un fragmento Fab que consista en los dominios de anticuerpo V_L, V_H, C_L y C_{H1}, o un fragmento Fv que consista en un dominio V_L y un dominio V_H. El término "cápside" significa un paquete de presentación genética replicable, con o sin la información genética. Las cápsides presentan un miembro de un par de unión específico en su superficie. El paquete puede ser una población de bacteriófagos que expone un dominio de unión antigénica, por ejemplo, un Fab, en la superficie de algunas o

todas las cápsidas dentro de la población. Este tipo de paquete se ha denominado anticuerpo fágico (pAb).

La expresión "dominio C_{H1}" significa la primera región constante de la cadena pesada de un anticuerpo o parte de la misma o se extiende a los aminoácidos de las regiones bisagras de modo que permitan el emparejamiento del fragmento (VH)CH1 expresado con la cadena ligera del anticuerpo y la posible formación de puentes disulfuro. Este

5 puede ser el dominio CH1 de un anticuerpo humano de isotipo gamma-1.

Un "componente parte de un sitio de unión antigénica de un anticuerpo" puede ser o corresponder a un componente de cadena polipeptídica, por ejemplo, un dominio V_H o un V_L. Sin embargo, puede haber una CDR o una secuencia V_L más la CDR de un V_H, una secuencia V_H más la CDR de un V_L, una V_H más la secuencia de V_L que carece

10 únicamente de una CDR y así sucesivamente. La condición es que la primera y la segunda parte del componente de un sitio de unión antigénica de un anticuerpo deben formar en combinación (conjuntamente) un sitio de unión antigénica. Por lo tanto, si la segunda parte del componente de un sitio de unión antigénica de un anticuerpo no humano específico para un antígeno de interés es una CDR, entonces la primera parte del componente de un sitio

15 de unión antigénica de un anticuerpo humano comprende el resto de una región V_H y V_L necesario para formar un sitio de unión antigénica (con o sin dominios constantes de anticuerpo asociados (en un formato Fab) o con o sin una secuencia peptídica enlazadora (en un formato Fv)). La segunda parte del componente de un sitio de unión antigénica de un anticuerpo no humano puede comprender un dominio V_L más parte de un dominio V_H, siendo esa parte una o más CDR, por ejemplo, quizá la CDR3. En dicho caso, la primera parte del componente de un sitio de unión antigénica de un anticuerpo humano comprendería el resto de una secuencia V_H que, en combinación con la

20 segunda parte del componente, forma un sitio de unión antigénica. Por supuesto, se mantiene la situación inversa y la persona experta en la materia podrá contemplar otras combinaciones de la primera y la segunda parte del componente que forman conjuntamente un sitio de unión antigénica. La expresión "condicionalmente defectuoso" significa un gen que no expresa un polipéptido particular en un conjunto de condiciones, pero lo expresa en otro conjunto de condiciones. Un ejemplo, es un gen que contiene una mutación ámbar en hospedadores no supresores o supresores respectivamente. Como alternativa, un gen puede expresar una proteína o un polipéptido que es defectuoso en un conjunto de condiciones pero no en otro conjunto. Un ejemplo es un gen con una mutación sensible a la temperatura.

El término "derivado" se refiere a polipéptidos químicamente modificados por técnicas tales como la ubiquitinación, el marcaje (por ejemplo, con radionúclidos o diversas enzimas), la pegilación (derivatización con polietilenglicol) y la inserción o sustitución mediante síntesis química de aminoácidos tales como ornitina, que no se dan normalmente en las proteínas humanas. El término "dominio" significa una parte de una proteína o un polipéptido que está plegada sobre sí misma e independientemente de otras partes de la misma proteína o polipéptido e independientemente de un miembro de unión complementario.

El término "eluyente" significa una solución usada para romper la unión entre dos moléculas. La unión puede ser mediante uno o más enlaces covalentes o no covalentes. Las dos moléculas pueden ser miembros de una sbp (*pareja de unión específica*). La expresión "fragmento modulador de la expresión", FME, significa una serie de nucleótidos que modula la expresión de una ORF unida operativamente u otro FME.

Tal como se usa en el presente documento, se dice que una secuencia "modula la expresión de una secuencia unida operativamente" cuando la expresión de la secuencia está alterada por la presencia del FME. Los FME incluyen, pero sin limitación, promotores y secuencias moduladoras de promotores (elementos inducibles). Una clase de FME son los fragmentos que inducen la expresión de una ORF unida operativamente en respuesta a un factor regulador específico o evento fisiológico.

El término "Fab" se refiere a fragmentos de anticuerpo que incluyen fragmentos que comprenden dos porciones N-terminales del polipéptido de cadena pesada unidas por al menos un puente disulfuro en la región bisagra y dos polipéptidos de cadena ligera completos, en los que cada cadena ligera forma un complejo con una porción N-terminal de una cadena pesada. Fab también incluye fragmentos Fab que comprenden toda o una porción grande de un polipéptido de cadena ligera (por ejemplo, V_LC_L) formando complejo con la porción N-terminal de un polipéptido de cadena pesada (por ejemplo, V_HC_{H1}).

La expresión "biblioteca de Fab" se refiere a una colección de secuencias polinucleotídicas de Fab dentro de los clones; o a una colección genéticamente diversa de polipéptidos de Fab presentados en un rgdp (paquete de expresión genética replicable, del inglés *replicable genetic display package*) que pueden seleccionarse o explorarse para proporcionar un polipéptido Fab individual o una población mixta de polipéptidos Fab. La expresión "unidad plegada" significa una combinación específica de estructura de hélice alfa y/o cadena-beta y/o giro-beta. Los dominios y las unidades plegadas contienen estructuras que asocian aminoácidos que no son adyacentes en la estructura primaria.

La expresión "población genéticamente diversa" significa anticuerpos o componentes polipeptídicos de los mismos, esta se refiere no solamente a la diversidad que puede existir en la población natural de células u organismos, sino también a la diversidad que puede crearse por mutación artificial *in vitro* o *in vivo*. La mutación *in vitro* puede, por ejemplo, implicar mutagénesis aleatoria usando oligonucleótidos que tienen mutaciones aleatorias de la secuencia deseada a variar. La mutagénesis *in vivo* puede, por ejemplo, usar cepas mutantes de microorganismos

hospedadores para alojar el ADN (véase el Ejemplo 38 del documento WO 92/01047).

La expresión "población única" puede usarse para indicar una pluralidad de, por ejemplo, cadenas polipeptídicas, que no son genéticamente diversas, es decir, todas son iguales. Una población restringida es una que es diversa pero lo es menos que el repertorio completo de un animal. La diversidad puede haberse reducido mediante la selección anterior, por ejemplo, usando la especificidad de unión antigénica.

La expresión "fago auxiliar" significa un fago que se usa para infectar células que contienen un genoma fágico defectuoso y que funciona para complementar el efecto. El genoma fágico defectuoso puede ser un fagémido o un fago con eliminación de alguna función que codifica secuencias génicas. Son ejemplos de fagos auxiliares son el gen III de M13K07, M13K07 N° 3; y fagos que exponen o codifican una molécula de unión fusionada a una proteína de la cápside.

El término "homólogos" significa polipéptidos que tienen el mismo resto o restos conservados en una posición correspondiente en su estructura primaria, secundaria o terciaria. El término también se extiende a dos o más secuencias nucleotídicas que codifican los polipéptidos homólogos. Son ejemplos de péptidos homólogos los isotipos de inmunoglobulina. La expresión "célula hospedadora" se refiere a una célula procariota o eucariota en la cual pueden introducirse, expresarse y/o propagarse los vectores del método de la invención. Las células hospedadoras de procariotas típicas incluyen diversas cepas de típicas de *E. coli*. Las células hospedadoras eucariotas típicas son levaduras u hongos filamentosos o células de mamíferos, tales como, las células de ovario de hámster chino, los fibroblastos murinos NIH 3t3 o las células 193 de riñón embrionario.

La expresión "superfamilia de inmunoglobulinas" significa una familia de polipéptidos, cuyos miembros tienen al menos un dominio con una estructura relacionada con la del dominio variable o constante de las moléculas de inmunoglobulina. El dominio contiene dos láminas B y habitualmente un enlace disulfuro conservado (véase A. F. Williams y A. N. Barclay 1988 Ann. Rev Immunol. 6 381-405). Son ejemplos de miembros de una superfamilia de inmunoglobulina CD4, el receptor del factor de crecimiento derivado de plaquetas (PDGFR), la molécula de adhesión intercelular. (ICAM). Excepto cuando el contexto determine claramente lo contrario, las referencias a inmunoglobulinas y a homólogos de inmunoglobulinas en esta solicitud incluyen miembros de la superfamilia de las inmunoglobulinas y homólogos de las mismas.

El término "infección" se refiere a la introducción de ácidos nucleicos en una célula hospedadora adecuada mediante el uso de un virus o vector viral.

La expresión "fragmento intermedio" significa un ácido nucleico de entre 5 y 1000 bases de longitud y preferentemente entre 10 y 40 pb de longitud.

El término "aislado" tal como se usa en el presente documento, se refiere a un ácido nucleico o polipéptido separado no solamente de otros ácidos nucleicos o polipéptidos que están presentes en la fuente natural del ácido nucleico o el polipéptido, sino también de polipéptidos y preferentemente se refiere a un ácido nucleico o polipéptido encontrado en presencia de (si es que hay alguno) únicamente un disolvente, tampón, ion u otro componente normalmente presente en una solución del mismo. Los términos "aislado" y "purificado" no abarcan ácidos nucleicos o polipéptidos presentes en su fuente natural.

La expresión "cepa mutante" significa una célula hospedadora que tiene un defecto genético que produce que el ADN replicado dentro de ésta mute con respecto a su ADN parental. Son ejemplos de cepas mutantes NR9046mutD5 y NR9046 mut T1 (Véase el Ejemplo 38 del documento WO 92/01047).

La expresión "polipéptido de origen natural" se refiere a polipéptidos producidos por células que no se han modificado genéticamente y específicamente contempla diversos polipéptidos que surgen a partir de modificaciones post-traduccionales del polipéptido que incluyen, pero no se limitan a, acetilación, carboxilación, glucosilación, fosforilación, lipidación y acilación.

La expresión "secuencia de nucleótidos" se refiere a un heteropolímero de nucleótidos o a la secuencia de estos nucleótidos. La expresión "ácido nucleico" y el término "polinucleótido" también se usan indistintamente en el presente documento para referirse a un heteropolímero de nucleótidos. Generalmente, los segmentos de ácidos nucleicos proporcionados por esta invención pueden ensamblarse a partir de fragmentos del genoma y enlazadores oligonucleotídicos cortos, o a partir de una serie de oligonucleótidos, o a partir de nucleótidos individuales, para proporcionar un ácido nucleico sintético que es capaz de expresarse en una unidad transcripcional recombinante que comprende elementos reguladores derivados de un operón microbiano o viral, o un gen eucariota.

Las expresiones "fragmento de oligonucleótido" o "fragmento de polinucleótido" y los términos "porción" o "segmento" son un tramo de restos de nucleótidos del polipéptido que es lo suficientemente largo para usar la reacción en cadena de la polimerasa (PCR) o diversos procedimientos de hibridación para identificar o amplificar partes idénticas o relacionadas de moléculas de ARNm o ADN.

El término "oligonucleótidos" o la expresión "sondas de ácidos nucleicos" se preparan basándose en las secuencias

- polinucleotídicas proporcionadas en la presente divulgación. Los oligonucleótidos comprenden porciones de dicha secuencia de polinucleótidos que tienen al menos aproximadamente 15 nucleótidos y habitualmente al menos 20 nucleótidos. Las sondas de ácidos nucleicos comprenden porciones de dicha secuencia polinucleotídica que tienen menos de aproximadamente 6 kb de nucleótidos, habitualmente menos de aproximadamente 1 kb. Después del ensayo apropiado para eliminar falsos positivos, estas sondas pueden utilizarse, por ejemplo, para determinar si las moléculas de ARNm específicas están presentes en una célula o tejido o para aislar secuencias de ácidos nucleicos similares de ADN cromosómico tal como describen Walsh *et al.* (Walsh, P.S. *et al.*, 1992, PCR Methods Appl 1:241-250).
- 5
- 10 La expresión "fase de lectura abierta", ORF (del inglés "*Open Reading Frame*", significa una serie de tripletes de nucleótidos que codifican aminoácidos sin ningún codón de terminación y es una secuencia traducible en proteína.
- La expresión "vector fágico" significa un vector derivado por modificación de un genoma fágico, que contiene un origen de replicación para un bacteriófago, pero no uno para un plásmido.
- 15
- La expresión "vector fagémido" significa un vector derivado por modificación de un genoma plasmídico, que contiene un origen de replicación para un bacteriófago así como el origen de replicación de un plásmido. El término "fenotipo" se refiere a una propiedad física (por ejemplo, pigmento o forma celular) y/o a una propiedad metabólica de una célula que puede medirse o aprovecharse de algún modo y que efectúa el gen indicador.
- 20
- La expresión "región polienlazadora" significa un polinucleótido que contiene al menos dos sitios de enzimas de restricción que son únicos en el vector que contiene la región polienlazadora, es decir, estos sitios de restricción se usan fácilmente como sitios de restricción en el vector.
- 25
- Un "fragmento", "porción" o "segmento" polipeptídico es un tramo de restos de aminoácidos de al menos aproximadamente 5 aminoácidos, a menudo de al menos 7 aminoácidos, típicamente al menos de 9 a 13 aminoácidos, y en diversas realizaciones, al menos aproximadamente 17 o más aminoácidos. Para que sea activo, cualquier péptido debe tener la suficiente longitud para mostrar actividad biológica y/o inmunológica.
- 30
- El término "sondas" incluye ácidos nucleicos mono o bicatenarios, de origen natural o sintetizados de manera recombinante o química. Estos pueden marcarse mediante traducción por cortes, reacción de relleno de Klenow, PCR u otros métodos bien conocidos en la técnica. Las sondas de la presente divulgación, su preparación y/o su marcaje se explican en Sambrook, J. *et al.*, 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; o Ausubel, F.M. *et al.*, 1989, Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York NY.
- 35
- El término "purificado" tal como se usa en el presente documento, indica que el ácido nucleico o el polipéptido indicado está presente en ausencia sustancial de otras macromoléculas biológicas, por ejemplo, polinucleótidos, proteínas y similares. En una realización, el polinucleótido o polipéptido se purifica de tal modo que este constituye al menos el 95 % en peso, más preferentemente al menos el 99,8 % en peso, de las macromoléculas biológicas indicadas presentes (aunque puede haber agua, tampones y otras moléculas pequeñas, especialmente moléculas que tienen un peso molecular menor de 1000 daltons).
- 40
- Cuando el término "recombinante" se usa en el presente documento para referirse a un polipéptido o a una proteína, significa que un polipéptido o una proteína proceden de sistemas de expresión recombinantes (por ejemplo, microbiano o de mamífero). 'Microbiano' se refiere a polipéptidos o proteínas recombinantes producidos en sistemas de expresión bacterianos o fúngicos (por ejemplo, levaduras). Como un producto, 'microbiano recombinante' define a un polipéptido o a una proteína que carece esencialmente de sustancias nativas endógenas, y no viene acompañado de glucosilación nativa asociada. Los polipéptidos o proteínas expresados en la mayor parte de los cultivos bacterianos, por ejemplo, *E. coli*, carecerán de modificaciones de glucosilación: los polipéptidos o proteínas expresados en levaduras tendrán un patrón de glucosilación en general distinto del patrón expresado en células de mamífero.
- 45
- 50
- La expresión "vehículo o vector de expresión recombinante" se refiere a un plásmido o a un fago o a un virus o a un vector, para expresar un polipéptido a partir de una secuencia de ADN (ARN). Un vehículo de expresión puede comprender una unidad transcripcional que comprenda un ensamblaje de (1) uno o más elementos genéticos que tengan un papel regulador en la expresión génica, por ejemplo, promotores o potenciadores, (2) una secuencia estructural o codificante que se transcribe en ARNm y se traduce a proteína y (3) secuencias de inicio y terminación de la transcripción apropiadas. Las unidades estructurales destinadas al uso en sistemas de expresión en levaduras o eucariotas incluyen, preferentemente, una secuencia líder que permite la secreción extracelular de la proteína traducida por una célula hospedadora. Como alternativa, cuando se expresa la proteína recombinante sin una secuencia líder o transportadora, esta puede incluir un resto de metionina N-terminal. Este resto puede escindirse o no posteriormente de la proteína o del polipéptido recombinante expresado para proporcionar un producto final.
- 55
- 60
- 65 La expresión "sistema de expresión recombinante" significa células hospedadoras que han integrado de un modo estable una unidad transcripcional recombinante en el ADN cromosómico o que portan la unidad transcripcional

- recombinante extracromosómicamente. Los sistemas de expresión recombinante, como se definen en el presente documento, expresarán polipéptidos o proteínas heterólogos tras la inducción de los elementos reguladores ligados al segmento de ADN o al gen sintético a expresar. Esta expresión también significa células hospedadoras que han integrado de un modo estable uno o más elementos genéticos recombinantes que tienen un papel regulador en la expresión génica, por ejemplo, promotores o potenciadores. Los sistemas de expresión recombinante tal como se define en el presente documento expresarán polipéptidos o proteínas endógenos para la célula tras la inducción de los elementos reguladores ligados al segmento de ADN endógeno o al gen a expresar. Las células pueden ser procariontas o eucariotas.
- 10 La expresión "variante recombinante" se refiere a cualquier polipéptido que difiera de los polipéptidos de origen natural debido a inserciones, deleciones y sustituciones de aminoácidos creadas usando técnicas de ADN recombinante. Una guía para determinar qué restos de aminoácidos pueden reemplazarse, añadirse o eliminarse, sin anular las actividades de interés, tales como el tránsito celular, pueden encontrarse comparando la secuencia del polipéptido particular con la de los péptidos homólogos y minimizando el número de cambios en la secuencia de aminoácidos realizados en las regiones de alta homología. Preferentemente, las "sustituciones" de aminoácidos son el resultado de reemplazar un aminoácido por otro aminoácido que tiene propiedades estructurales y/o químicas similares, es decir, reemplazos de aminoácidos conservativos. Las sustituciones de aminoácidos pueden realizarse basándose en su similitud en la polaridad, carga, solubilidad, hidrofobicidad, hidrofiliidad y/o en la naturaleza anfipática de los restos implicados. Por ejemplo, los aminoácidos no polares (hidrófobos) incluyen alanina, leucina, isoleucina, valina, prolina, fenilalanina, triptófano y metionina; los aminoácidos polares neutros incluyen glicina, serina, treonina, cisteína, tirosina, asparagina y glutamina; los aminoácidos cargados positivamente (básicos) incluyen arginina, lisina e histidina; y los aminoácidos cargados negativamente (ácidos) incluyen ácido aspártico y ácido glutámico.
- 25 Las "inserciones" o "deleciones" están típicamente en el intervalo de aproximadamente 1 a 5 aminoácidos. La variación permitida puede determinarse experimentalmente mediante la realización sistemática de inserciones, deleciones o sustituciones de aminoácidos en una molécula polipeptídica usando técnicas de ADN recombinante y ensayando las variantes recombinantes resultantes con respecto a su actividad.
- 30 Como alternativa, cuando se desea una alteración de la función, pueden modificarse por ingeniería inserciones, deleciones, alteraciones no conservativas para producir polipéptidos alterados. Dichas alteraciones pueden, por ejemplo, alterar una o más de las funciones biológicas o características bioquímicas de los polipéptidos. Por ejemplo, dichas alteraciones pueden cambiar características del polipéptido tales como las afinidades de unión a ligando, las afinidades intercatenarias o la tasa de degradación/renovación. Además, dichas alteraciones pueden seleccionarse de modo que generen polipéptidos que se ajusten mejor a la expresión, mejora y similar en las células hospedadoras elegidas para su expresión. Por ejemplo, pueden eliminarse o sustituirse restos de cisteína por otros restos de aminoácidos con el fin de eliminar los puentes disulfuro. Como alternativa, pueden sintetizarse o seleccionarse variantes recombinantes que codifican estos mismos polipéptidos o similares haciendo uso de la "redundancia" del código genético. Diversas sustituciones de codones, tales como los cambios silenciosos que producen diversos sitios de restricción, que pueden producir diversos sitios de restricción, pueden introducirse para optimizar la clonación en un plásmido o vector viral o la expresión en un sistema procarionta o eucariota particular. Las mutaciones en la secuencia polinucleotídica pueden reflejarse en el polipéptido o en los dominios de otros péptidos añadidos al polipéptido para modificar las propiedades de cualquier parte del polipéptido, para cambiar las características tales como las afinidades de unión a ligando, las afinidades intercatenarias o la tasa de degradación/renovación.
- 45 La expresión "repertorio de genes de inmunoglobulina reorganizados artificialmente" significa una colección de nucleótidos, por ejemplo, secuencias de ADN procedentes completa o parcialmente de una fuente distinta de la de las secuencias de inmunoglobulina reorganizadas de un animal. Esto puede incluir, por ejemplo, secuencias de ADN que codifican dominios VH mediante la combinación de segmentos V desorganizados con segmentos D y J y secuencias de ADN que codifican los dominios VL mediante la combinación de segmentos V y J. Parte o todas las secuencias de ADN pueden obtenerse por síntesis de oligonucleótidos.
- 50 La expresión "repertorio de genes de inmunoglobulina reorganizados" significa una colección de oligonucleótidos de origen natural, por ejemplo, secuencias de ADN que codifican genes de inmunoglobulina expresados en un animal. Las secuencias se generan mediante la reorganización *in vivo* de, por ejemplo, segmentos V, J y D para las cadenas H y, por ejemplo, los segmentos V y J para las cadenas L. Como alternativa, las secuencias pueden generarse a partir de una línea celular inmunizada *in vitro* y en las que la reorganización en respuesta a la inmunización se da intracelularmente. El vocablo "repertorio" se usa para indicar la diversidad genética. La expresión "paquete de presentación genética replicable" (Rgdp, del inglés *replicable genetic display package*) significa una partícula biológica que tiene información genética que proporciona a la partícula la capacidad de replicarse. La partícula puede presentar en su superficie al menos parte de un polipéptido. El polipéptido puede codificarse mediante la información genética nativa en la partícula y/o colocarse artificialmente en la partícula o en un ancestro de esta. El polipéptido presentado puede ser cualquier miembro de un par de unión específico, por ejemplo, los dominios de la cadena pesada o ligera basados en una molécula de inmunoglobulina, una enzima o un receptor, etc. La partícula puede ser un virus, por ejemplo, un bacteriófago tal como fd o M13.

- La expresión "gen indicador" se refiere a un ácido nucleico que codifica una proteína o un polipéptido que produce un cambio fenotípico en la célula hospedadora que puede medirse y/o usarse para separar células hospedadoras. Por ejemplo, el gen indicador puede codificar una proteína o un polipéptido que tiene propiedades fluorescentes, por ejemplo, β -galactosidasa, proteína auto-fluorescente GFP, etc; o el gen indicador puede codificar un marcador de selección, por ejemplo, resistencia a antibióticos; o un epítipo que se exprese en la superficie de la célula hospedadora.
- La expresión "sitio de unión a ribosomas" significa un polirribonucleótido que permite que un ribosoma seleccione el codón de inicio apropiado durante el inicio de la traducción. En algunos procariotas, este polinucleótido se denomina secuencia de Shine-Dalgarno y los pares de bases de la secuencia de Shine-Dalgarno con el ARN de 16 S del ribosoma.
- La expresión proteína o polipéptido "secretado" se refiere a una proteína o a un polipéptido que se transporta a través o a lo largo de una membrana, incluyendo el transporte como resultado de secuencias señal en su secuencia de aminoácidos cuando se expresa en una célula hospedadora adecuada. Las proteínas o los polipéptidos "secretados" incluyen, sin limitación, proteínas o polipéptidos secretados completamente (por ejemplo, proteínas secretadas) o parcialmente (por ejemplo, receptores) a partir de la célula en la que se expresan. Las proteínas o los polipéptidos "secretados" también incluyen, sin limitación, proteínas o polipéptidos que se transportan a través de la membrana del retículo endoplasmático.
- La expresión "secuencia señal" significa una secuencia de aminoácidos que se encuentra en el extremo amino-terminal de un polipéptido y dirige el transporte del polipéptido a través o a lo largo de una membrana. Las secuencias señal incluyen polipéptidos amino-terminales que tienen 13-36 restos de longitud y que tienen un núcleo hidrófobo de 7 a 13 restos flanqueado por varios restos hidrófilos que habitualmente incluyen uno o más restos básicos cerca del extremo N-terminal.
- El término "riguroso" se usa para referirse a condiciones que se entienden comúnmente en la técnica como rigurosas. Un conjunto de condiciones ejemplares incluyen una temperatura de 60-70 °C, (preferentemente de aproximadamente 65 °C) y una concentración de sales de 0,70 M a 0,80 M (preferentemente de aproximadamente 0,75 M). Otras condiciones ejemplares incluyen, condiciones de hibridación que (1) emplean una fuerza iónica baja y una temperatura de lavado alta, por ejemplo, NaCl 0,015 M/citrato de sodio 0,0015 M/SDS al 0,1 % a 50 °C ; (2) emplean durante la hibridación un agente desnaturizante tal como la formamida, por ejemplo, formamida al 50 % (vol/vol) con albúmina de suero bovino al 0,1 % /Ficoll al 0,1 %/polivinilpirrolidona al 0,1 %/ tampón de fosfato de sodio 50 mM a pH 6,5 con NaCl 750 mM, citrato de sodio 75 mM a 42 °C ; o (3) emplean formamida al 50 %, SSC 5 x (NaCl 0,75 M, pirofosfato de sodio 0,075 M, solución de Denhardt 5 x, ADN de esperma de salmón sometido a ultrasonido (50 g/ml), SDS al 0,1 % y sulfato de dextrano al 10 % a 42 °C , con lavados a 42 °C en SSC 0,2 x y SDS al 0,1 %.
- En casos que conciernen a la hibridación de desoxirribonucleótidos, las condiciones ejemplares de hibridación rigurosa adicionales incluyen lavados en SSC 6x/pirofosfato de sodio al 0,05 % a 37 °C (para oligos de 14 bases), 48 °C (para oligos de 17 bases), 55 °C (para oligos de 20) y 60 °C (para oligos de 23 bases).
- Tal como se usa en el presente documento, "sustancialmente equivalente" puede referirse a secuencias tanto de nucleótidos como de aminoácidos, por ejemplo, una secuencia mutante, que varía de una secuencia de referencia en una o más sustituciones, deleciones o adiciones, cuyo efecto neto no produce ninguna disimilitud funcional adversa entre las secuencias de referencia y las secuencias sujeto. Típicamente, dicha secuencia sustancialmente equivalente varía de una de aquellas enumeradas en el presente documento en no más de aproximadamente el 20 % (es decir, el número de sustituciones, adiciones y/o deleciones de restos individuales en una secuencia sustancialmente equivalente, en comparación con la referencia correspondiente, dividido entre el número total de restos en la secuencia sustancialmente equivalente es aproximadamente de 0,2 o menor). Se dice que dicha secuencia tiene el 80 % de identidad de secuencia con respecto a la secuencia enumerada. En una realización, una secuencia sustancialmente equivalente, por ejemplo, mutante, varía de una secuencia enumerada en no más del 10 % (90 % de identidad de secuencia); en una variación de esta realización en no más del 5 % (95 % de identidad de secuencia); y en una variación adicional de esta realización, en no más del 2 % (98 % de identidad de secuencia). Las secuencias de aminoácidos sustancialmente equivalentes, por ejemplo, mutantes, de acuerdo con la presente invención generalmente tienen al menos un 95 % de identidad de secuencia con respecto a una secuencia de aminoácidos enumerada, mientras que la secuencia de nucleótidos sustancialmente equivalente de la divulgación puede tener unos porcentajes de identidades de secuencia menores, teniendo en cuenta, por ejemplo, la redundancia o degeneración del código genético. Para los fines de la presente divulgación, las secuencias tienen una actividad biológica sustancialmente equivalente y unas características de expresión sustancialmente equivalentes. Para los fines de determinación de la equivalencia, no debería prestarse atención al truncamiento de la secuencia madura (por ejemplo, por medio de una mutación que crea un codón de terminación espurio).
- Las secuencias de ácidos nucleicos que codifican dichas secuencias sustancialmente equivalentes, por ejemplo, los porcentajes de identidad de secuencia enumerados, pueden aislarse de un modo rutinario e identificarse por medio

de procedimientos de hibridación convencionales bien conocidos por los expertos en la materia.

La expresión "codón de terminación traduccional suprimible" significa un codón que permite la traducción de una secuencia de nucleótidos aguas abajo del codón en un conjunto de condiciones, pero en otro conjunto de condiciones la traducción termina en el codón. Son ejemplos de codones traduccionales suprimibles los codones ámbar, ocre y ópalo.

El término "marcador" significa una extensión del fragmento Fab del anticuerpo, por ejemplo expresado en el extremo carboxi-terminal de la cadena pesada, que comprende al menos un aminoácido, pero más típicamente de cinco a quince aminoácidos, y que puede reconocer específicamente un anticuerpo u otro ligando de unión o matriz de unión para la secuencia. Los marcadores pueden combinarse en el mismo Fab. Son ejemplos un tramo de cinco restos de histidina que puede reconocer anticuerpos específicos e iones metálicos inmovilizados definidos y un tramo de los siguientes 12 aminoácidos (EQKLISEEDLN) reconocido por el anticuerpo 9E10 (Marks *et al.*, 1991).

El término "diana" significa cualquier molécula que sea antigénica, por ejemplo, que pueda reconocer con especificidad razonablemente un anticuerpo de la biblioteca de Fab.

La expresión "elemento diana" se refiere a una secuencia de ácido nucleico que altera la expresión del gen diana. Los elementos diana incluyen, pero no se limitan a, promotores y secuencias moduladoras promotoras (elementos inducibles). Una clase de elementos diana son los fragmentos que inducen la expresión en respuesta a un factor regulador específico o evento fisiológico.

El término "transfección" se refiere a la incorporación de un vector de expresión en una célula hospedadora adecuada, se expresen de hecho o no cualquiera de las secuencias codificantes.

El término "transformación" significa la introducción de ADN en una célula hospedadora adecuada de modo que el ADN sea replicable, bien como un elemento extracromosómico o mediante integración cromosómica.

La expresión "conjunto universal" se refiere a un conjunto de ácidos nucleicos, más preferentemente a un conjunto de oligonucleótidos, que representa todas las combinaciones de secuencias posibles para una longitud de nucleótidos dada por ejemplo, los 4096 insertos de oligonucleótidos de seis nucleótidos de longitud. En una realización preferente, la expresión conjunto universal se refiere al conjunto de todos los oligonucleótidos posibles de una longitud dada, en los que una o más de las posiciones en el oligonucleótido se mantienen constantes (es decir, el mismo nucleótido en esta posición está presente en todos los miembros del conjunto).

Tal como se usa en el presente documento, un "fragmento modulador de la captación", FMC, significa una serie de nucleótidos que media la captación del ADN unido en una célula. Los FMC pueden identificarse fácilmente usando FMC conocidos como una secuencia o motivo diana con los sistemas informáticos descritos más adelante.

La presencia y actividad de un FMC puede confirmarse uniendo el FMC sospechoso con una secuencia marcadora. La molécula de ácido nucleico resultante se incuba después con un hospedador apropiado en condiciones apropiadas y se determina la captación de la secuencia marcadora. Tal como se describe anteriormente, un FMC aumentará la frecuencia de captación de una secuencia marcadora ligada.

El término "vector" se refiere a un plásmido o fago o virus o vector, para la expresión de un polipéptido a partir de una secuencia de ADN (ARN). El vector puede comprender una unidad transcripcional que comprende un ensamblaje (1) un elemento o elementos genéticos que tienen un papel regulador en la expresión génica, por ejemplo, promotores o potenciadores, (2) a una secuencia estructural o codificante que se transcribe en ARNm y se traduce en proteína y (3) secuencias apropiadas de inicio y terminación de la traducción. Las unidades estructurales destinadas para el uso en sistemas de expresión en levaduras o eucariotas pueden incluir una secuencia líder que permite la secreción extracelular de la proteína traducida por una célula hospedadora.

La expresión "polinucleótidos V_L " significa polinucleótidos que codifican los dominios que contienen las CDR de algunos o todos los genes de la cadena ligera de las familias de V_{K-} y/o $V_{\lambda-}$

La expresión "polinucleótidos V_H " significa polinucleótidos que codifican los dominios que contienen las CDR de algunos o todos los genes de la cadena pesada de la familia de genes de cadena pesada.

Cada uno de los términos y expresiones anteriores tiene por objeto incluir todo lo que se describe para cada uno de ellos, a menos que el contexto indique claramente lo contrario.

Las construcciones recombinantes de la presente divulgación comprenden un vector, tal como un plásmido o un vector viral, en el que puede insertarse uno o más ácidos nucleicos de interés. El vector puede comprender adicionalmente secuencias reguladoras, incluyendo por ejemplo, un promotor, unido operativamente a uno o más ácidos nucleicos de interés. Un gran número de vectores y promotores adecuados son conocidos para los expertos en la materia y están disponibles en el comercio para generar las construcciones recombinantes del método de la presente invención. Los siguientes vectores se proporcionan a modo de ejemplo. Bacterianos: pBs, phagescript,

PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eucariotas: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

5 Pueden usarse métodos muy conocidos para los expertos en la materia para construir vectores que contienen un polinucleótido usado en el método de la invención y señales de control transcripcional/traduccionales apropiadas. Estos métodos incluyen técnicas de ADN recombinante *in vitro*, técnicas de síntesis y técnicas de recombinación/genética *in vivo*. Véanse, por ejemplo, las técnicas descritas en Maniatis *et al.*, Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y. (1989) y Ausubel *et al.* Current Protocols in Molecular
10 Biology, Greene Publishing Associates y Wiley Interscience, N.Y. (1989).

Pueden seleccionarse regiones promotoras de cualquier gen deseado usando vectores CAT (cloranfenicol transferasa) u otros vectores con marcadores de selección. Son dos vectores apropiados pKK232-8 y pCM7. Los promotores bacterianos particulares nombrados incluyen *lacI*, *lacZ*, T3, T7, *gpt*, *lambda P* y *trc*. Los promotores eucariotas incluyen el promotor temprano inmediato del CMV, el de timidina quinasa del VHS y el temprano y tardío de SV40, los LTR de retrovirus y la metalotioneína-I de ratón.

Generalmente, los vectores de expresión recombinante incluirán orígenes de replicación y marcadores de selección que permitan la transformación de la célula hospedadora, por ejemplo, el gen de resistencia a ampicilina de *E. coli* y el gen TRP1 de *S. cerevisiae*, y un promotor derivado de un gen de alta expresión para dirigir la transcripción de una secuencia estructural aguas abajo. Dichos promotores pueden proceder de operones que codifican enzimas glucolíticas tales como la 3-fosfoglicerato quinasa (PGK), el factor *a*, la fosfatasa ácida o proteínas de choque térmico, entre otras. El polinucleótido usado en el método de la invención se ensambla en una fase apropiada con secuencias de inicio y de terminación, y preferentemente con una secuencia líder capaz de dirigir la secreción de la proteína traducida al interior del espacio periplasmático o al medio extracelular. Opcionalmente, el polinucleótido usado en el método de la invención puede codificar una proteína de fusión que incluye un péptido de identificación N-terminal que otorga características deseadas, por ejemplo, estabilización o purificación simplificada del producto recombinante expresado.

30 Los vectores de expresión útiles para bacterias se construyen insertando un polinucleótido usado en el método de la invención junto con señales adecuadas de inicio y de terminación de la traducción, opcionalmente en fase de lectura operativa con un promotor funcional. El vector comprenderá uno o más marcadores fenotípicos de selección y un origen de replicación para asegurar el mantenimiento del vector y, si es deseable proporcionar amplificación dentro del hospedador. Los hospedadores procariotas adecuados para la transformación incluyen *E. coli*, *Bacillus subtilis*,
35 *Salmonella typhimurium* y diversas especies dentro del género *Pseudomonas*, *Streptomyces* y *Staphylococcus*, aunque también pueden emplearse otros como una cuestión de elección.

Como un ejemplo representativo pero no limitante, los vectores de expresión útiles para bacterias pueden comprender un marcador de selección y un origen de replicación bacteriano derivado de plásmidos disponibles en el comercio que comprenden elementos genéticos del vector de clonación bien conocido pBR322 (ATCC 37017). Dichos vectores comerciales incluyen, por ejemplo, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Suecia) y GEM 1 (Promega Biotec, Madison, WI, EE.UU.). Estas secciones de "estructura" de pBR322 se combinan con un promotor apropiado y la secuencia estructural a expresar.

45 Adicionalmente, la presente divulgación proporciona células hospedadoras que contienen los vectores de la presente divulgación, en las que se ha introducido el ácido nucleico en la célula hospedadora usando métodos de transformación, transfección o infección conocidos. La célula hospedadora puede ser una célula hospedadora eucariota superior, tal como una célula de mamífero, una célula hospedadora eucariota inferior, tal como una célula de levadura o una célula hospedadora puede ser una célula procariota, tal como una célula bacteriana. La introducción de la construcción recombinante en la célula hospedadora puede efectuarse, por ejemplo, mediante transfección con fosfato de calcio, transfección mediada por DEAE, dextrano o electroporación (Davis, L. *et al.*, *Basic Methods in Molecular Biology* (1986)). Puede usarse cualquier sistema de hospedador/vector para identificar uno o más de los elementos diana de la presente divulgación. Estos incluyen, pero no se limitan a, hospedadores eucariotas tales como las células HeLa, células Cv-1, células COS y células Sf9, así como hospedadores
50 procariotas tales como *E. coli* y *B. subtilis*. Las células más preferentes son aquellas que no expresan normalmente el polipéptido o proteína indicadora particular o que expresan el polipéptido o proteína indicadora particular a un nivel natural bajo.

El hospedador también puede ser una levadura u otro hongo. En las levaduras, pueden usarse una serie de vectores que contienen promotores constitutivos o inducibles. Para una revisión, véase, Current Protocols in Molecular Biology, Vol. 2, Ed. Ausubel *et al.*, Greene Publish. Assoc. & Wiley Interscience, Capítulo. 13 (1988); Grant *et al.*, Expression and Secretion Vectors for Yeast, en Methods in Enzymology, Ed. Wu & Grossman, Acad. Press, N.Y. 153: 516-544 (1987); Glover, DNA Cloning, Vol. II, IRL Press, Wash., D.C., capítulo 3 (1986); Bitter, Heterologous Gene Expression in Yeast, en Methods in Enzymology, Berger & Kimmel editores, Acad. Press, N.Y. 152: 673-684 (1987); y The Molecular Biology of the Yeast Saccharomyces, Strathern *et al.* editores, Cold Spring Harbor Press, Vols. I y II (1982).

El hospedador también puede ser una célula procariota tal como *E. coli*, otras *Enterobacteriaceae* tales como *Serratia marescans*, bacilos, diversas pseudomonas u otros procariotas que pueden transformarse, transfectarse, infectarse, etc. (es decir, existe un método para la introducción de ácidos nucleico en la célula hospedadora).

5 La presente divulgación proporciona adicionalmente células hospedadoras modificadas genéticamente para que contengan los polinucleótidos usados en el método de la invención. Por ejemplo, dichas células hospedadoras pueden contener ácidos nucleicos de la divulgación introducidos en la célula hospedadora usando métodos de transformación, transfección o infección conocidos. La presente divulgación aún proporciona adicionalmente células
10 hospedadoras modificadas genéticamente para que expresen los polinucleótidos de la divulgación, en los que dichos polinucleótidos están en asociación operativa con una secuencia reguladora heteróloga para la célula hospedadora que conduce la expresión de los polinucleótidos en la célula.

15 La célula hospedadora puede ser una célula hospedadora eucariota superior, tal como una célula de mamífero, una célula hospedadora eucariota inferior, tal como una célula de levadura, o la célula hospedadora puede ser una célula procariota, tal como una célula bacteriana. La introducción de la construcción recombinante en la célula hospedadora puede efectuarse por transfección de fosfato de calcio, transfección mediada por DEAE, dextrano o electroporación (Davis, L. *et al.*, *Basic Methods in Molecular Biology* (1986)). Las células hospedadoras que contienen uno de los polinucleótidos de la divulgación, pueden usarse de modos convencionales para producir el
20 producto génico codificado por el fragmento aislado (en el caso de una ORF) o pueden usarse para producir una proteína heteróloga bajo el control del FME.

Puede usarse cualquier sistema de hospedador/vector para expresar una o más de las ORF de la presente divulgación. Estos incluyen, pero no se limitan a, hospedadores eucariotas tales como células HeLa, células Cv-1, células COS y células Sf9 así como hospedadores procariotas tales como *E. coli* y *B. subtilis*. Las células más
25 preferentes son aquellas que no expresan normalmente el polipéptido o proteína particular o que expresan el polipéptido o proteína particular a un nivel natural bajo. Las proteínas maduras pueden expresarse en células de mamíferos, levaduras, bacterias u otras células bajo el control de promotores apropiados. Los sistemas de traducción libres de células también pueden emplearse para producir dichas proteínas usando ARN derivados de las construcciones de ADN de la presente divulgación. Los vectores de clonación y expresión apropiados para su uso con hospedadores procariotas y eucariotas se describen por Sambrook, *et al.*, in *Molecular Cloning: A Laboratory Manual*, segunda edición, Cold Spring Harbor, Nueva York (1989).

35 También pueden emplearse diversos sistemas de cultivos celulares en mamíferos para expresar proteínas recombinantes. Los ejemplos de sistemas de expresión en mamíferos incluyen las líneas COS-7 de fibroblastos de riñón de mono, descritas por Gluzman, *Cell* 23: 175 (1981) y otras líneas celulares capaces de expresar un vector compatible, por ejemplo, las líneas celulares C127, 3T3, CHO, HeLa y BHK. Los vectores de expresión en mamíferos comprenderán un origen de replicación, un promotor adecuado y también cualquiera de los sitios de unión a ribosomas, sitios de poliadenilación, sitios de escisión donantes y aceptores, secuencias de terminación
40 transcripcional y secuencias no transcritas flanqueantes de 5' necesarios. Las secuencias de ADN derivadas del genoma viral de SV40, por ejemplo, el origen SV40, el promotor temprano, el potenciador, los sitios de escisión y poliadenilación pueden usarse para proporcionar los elementos genéticos no transcritos necesarios. Los polipéptidos y proteínas recombinantes producidos en cultivos bacterianos se aíslan habitualmente mediante la extracción inicial a partir de los sedimentos celulares, seguida de uno o más de etapas de desalinización, cromatografía acuosa de intercambio iónico o exclusión molecular. También pueden usarse etapas de repliegue de las proteínas, siempre que sea necesario, para completar la configuración de la proteína madura. Finalmente, puede emplearse la cromatografía líquida de alto rendimiento (HPLC) para las etapas de purificación final. Las células microbianas empleadas en la expresión de proteínas pueden romperse por cualquier método conveniente, incluyendo ciclos de congelación, sonicación, rotura mecánica o el uso de agentes de lisado celular.

50 Diversos tipos de células pueden actuar como células hospedadoras adecuadas para la expresión de la proteína. Las células hospedadoras de mamíferos incluyen, por ejemplo, células COS de mono, células de ovario de hámster chino (CHO), células 293 de riñón humano, células humanas epidérmicas A431, células humanas Colo205, células 3T3, células CV-1, otras células transformadas de primates, células normales diploides, cepas celulares derivadas del cultivo de tejidos primarios *in vitro*, explantes primarios, células HeLa, células L de ratón, células BHK, HL-60, U937, HaK o Jurkat.

60 Como alternativa, puede ser posible producir la proteína en eucariotas inferiores tal como una levadura o en procariotas tales como bacterias. Las cepas de levadura posiblemente adecuadas incluyen *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* o cualquier cepa de levadura capaz de expresar proteínas heterólogas. La cepas de levaduras posiblemente adecuadas incluyen *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium* o cualquier cepa bacteriana capaz de expresar proteínas heterólogas. Si la proteína se produce en levaduras o bacterias, puede ser necesario modificar la proteína producida en las mismas, por ejemplo mediante la fosforilación o glucosilación en los sitios apropiados, con el fin de obtener la proteína funcional. Dichas uniones covalentes pueden llevarse a cabo usando métodos químicos o enzimáticos conocidos.

65

Las células y tejidos pueden modificarse por ingeniería genética para que expresen un gen endógeno que comprenda los polinucleótidos de la divulgación bajo el control de elementos reguladores inducibles, en cuyo caso las secuencias reguladoras del gen endógeno pueden reemplazarse por recombinación homóloga. Tal como se describe en el presente documento, el direccionamiento génico puede usarse para reemplazar una región génica reguladora existente por una secuencia reguladora aislada de un gen distinto o una secuencia reguladora nueva sintetizada mediante métodos de ingeniería genética. Dichas secuencias reguladoras pueden estar comprendidas de promotores, potenciadores, regiones de plegamiento-unión, elementos reguladores negativos, sitios de inicio de la transcripción, sitios de unión a proteínas reguladoras o combinaciones de dichas secuencias. Como alternativa, las secuencias que afectan a la estructura o la estabilidad del ARN o de la proteína pueden reemplazarse, eliminarse, añadirse o modificarse de otro modo mediante el direccionamiento, incluyendo las señales de poliadenilación, elementos de estabilidad del ARNm, sitios de escisión, secuencias líder para potenciar o modificar las propiedades de transporte o secreción de la proteína u otras secuencias que alteran o mejoran la función o estabilidad de las moléculas de proteína o de ARN.

El acontecimiento de direccionamiento puede ser una inserción simple de la secuencia reguladora, colocando el gen bajo el control de una nueva secuencia reguladora, por ejemplo, insertando un nuevo promotor o potenciador o ambos aguas arriba de un gen. Como alternativa, el evento de direccionamiento puede ser una delección simple de un elemento regulador, tal como la delección de un elemento regulador negativo específico de tejido. Como alternativa, el evento de direccionamiento puede reemplazar un elemento existente; por ejemplo, un potenciador específico de tejido puede reemplazarse por un potenciador que tenga una especificidad más amplia o para distintos tipos celulares que los elementos de origen natural. En el presente documento, las secuencias de origen natural se eliminan y se añaden nuevas secuencias. En todos los casos, la identificación del acontecimiento de direccionamiento puede facilitarse mediante el uso de uno o más genes marcadores de selección que son contiguos al ADN de direccionamiento, permitiendo la selección de células en las que se integra el ADN exógeno en el genoma de célula hospedadora. La identificación del acontecimiento de direccionamiento también puede facilitarse mediante el uso de uno o más genes marcadores que muestran las propiedades de la regulación negativa, tales como los marcadores de selección que se unen negativamente al ADN exógeno, pero que se configuran de tal modo que el marcador de selección flanquea las secuencias de direccionamiento y de tal modo que un acontecimiento de recombinación homóloga correcto con las secuencias del genoma de la célula hospedadora no da como resultado la integración estable del marcador de selección de manera negativa. Los marcadores útiles para este fin incluyen el gen de la timidina quinasa (TK) del virus del herpes simple o el gen bacteriano de la xantina-guanina fosforribosil-transferasa (gpt).

Las técnicas de direccionamiento génico o activación génica que pueden usarse de acuerdo con este aspecto de la invención se describen más particularmente en la Patente de Estados Unidos N° 5.272.071 de Chappel; Patente de Estados Unidos N° 5.578.461 a Sherwin *et al.*; documento WO 93/09222 por Selden *et al.*; y el documento WO 91/06667 por Skoultchi *et al.*

En general, las técnicas para preparar anticuerpos policlonales y monoclonales así como hibridomas capaces de producir el anticuerpo deseado son muy conocidas en la técnica (Campbell, A.M., *Monoclonal Antibodies Technology: Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Publishers, Amsterdam, Países Bajos (1984); St. Groth *et al.*, *J. Immunol.* 35: 1-21 (1990); Kohler y Milstein, *Nature* 256: 495-497 (1975)), la técnica del trioma, la técnica del hibridoma de células B humanas (Kozbor *et al.*, *Immunology Today* 4:72 (1983); Cole *et al.*, en *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985), pág. 77-96).

Los métodos para la inmunización son muy conocidos en la técnica. Dichos métodos incluyen la inyección subcutánea o intraperitoneal del polipéptido. Un experto en la materia reconocerá que la cantidad de proteína codificada por el gen indicador usada para la inmunización variará en función del animal al que se le inmuniza, de la antigenicidad del péptido y del sitio de inyección.

El polipéptido o la proteína de la divulgación que se usa como un inmunógeno puede modificarse o administrarse en un adyuvante con el fin de aumentar la antigenicidad del polipéptido o la proteína: Los métodos para aumentar la antigenicidad de un polipéptido o proteína son bien conocidos en la técnica e incluyen, pero no se limitan a, acoplamiento del antígeno con una proteína heteróloga (tal como una globulina o β -galactosidasa) o a través de la inclusión de un adyuvante durante la inmunización.

Para los anticuerpos monoclonales, se extraen células del bazo de los animales inmunizados, se fusionan con células de mieloma, tales como las células de mieloma SP2/0-Ag14 y se les permite que se conviertan en células de hibridoma productoras de un anticuerpo monoclonal.

Puede usarse uno cualquiera de los diversos métodos bien conocidos en la técnica para identificar la célula de hibridoma que produce un anticuerpo con las características deseadas. Estos incluyen la exploración de hibridomas con un ensayo ELISA, análisis de transferencia western o un radioinmunoensayo (Lutz *et al.*, *Exp. Cell Research* 175: 109-124 (1988)).

Los hibridomas que secretan los anticuerpos deseados se clonan y la clase y la subclase se determinan usando procedimientos conocidos en la técnica (Campbell, A.M., *Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Editores, Ámsterdam, Países Bajos (1984)).

Para los anticuerpos policlonales, se aíslan los antisueros que contienen anticuerpos del animal inmunizado y se exploran con respecto a la presencia de anticuerpos con la especificidad deseada usando uno de los procedimientos anteriormente descritos.

5 Las células hospedadoras se transfectan o preferentemente se infectan o se transforman con los vectores descritos anteriormente, y se cultivan en medios con nutrientes apropiados para la selección de los transductantes o transformantes que contienen el vector.

10 Las células hospedadoras que expresan el polipéptido o proteína del producto de la divulgación pueden identificarse mediante al menos cuatro enfoques generales; (a) hibridación de ADN-ADN o ADN-ARN; (b) la presencia o ausencia de funciones génicas; (c) la evaluación del nivel de la transcripción tal como se mide por la expresión de los transcritos de ARNm en la célula hospedadora; y (d) la detección del producto génico tal como se mide mediante inmunoensayo o mediante su actividad biológica.

15 En el primer enfoque, la presencia del polipéptido o de la proteína de la divulgación insertado(a) en el vector puede detectarse mediante hibridación del ADN-ADN o ADN-ARN usando sondas que comprenden secuencias de nucleótidos que son homólogas al polipéptido o a la proteína de la divulgación, respectivamente, o a porciones o a sus derivados.

20 En el segundo enfoque, el sistema vector/hospedador de expresión recombinante puede identificarse y seleccionarse basándose en la presencia o ausencia de determinadas funciones génicas "marcadoras" (por ejemplo, actividad timidina quinasa, resistencia a antibióticos, resistencia a metotrexato, fenotipo de transformación, formación de cuerpos de oclusión en baculovirus, etc.). Por ejemplo, si el polipéptido o proteína de la divulgación se inserta dentro de una secuencia génica marcadora del vector, las células recombinantes que contienen el polipéptido de la proteína de la divulgación pueden identificarse por la ausencia de la función génica del marcador. Como alternativa, puede colocarse un gen marcador en tándem con el polipéptido o proteína de la divulgación bajo el control del mismo o distinto promotor usado para controlar la expresión del polipéptido o proteína de la divulgación. La expresión del marcador en respuesta a la inducción o la selección indica la expresión del polipéptido o proteína de la divulgación.

25 En el tercer enfoque, la actividad transcripcional del polipéptido o proteína de la divulgación puede evaluarse mediante ensayos de hibridación. Por ejemplo, puede aislarse el ARN e hibridar mediante transferencia de Northern usando una sonda homóloga al polipéptido o proteína de la divulgación o porciones particulares de la misma. Como alternativa, los ácidos nucleicos totales de la célula hospedadora pueden extraerse y ensayarse con respecto a la hibridación con dichas sondas.

30 En el cuarto enfoque, la expresión de un producto del polipéptido o de la proteína de la divulgación puede evaluarse inmunológicamente, por ejemplo mediante transferencias de Western, inmunoensayos tales como radioinmunoprecipitación, inmunoensayos ligados a enzimas y similares.

35 Puede unirse un polinucleótido de acuerdo con la divulgación a cualquiera de una variedad de otras secuencias de nucleótidos mediante técnicas de ADN recombinantes bien establecidas (véase Sambrook J *et al.* (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Las secuencias de nucleótidos útiles para la unión con los polipéptidos incluyen una variedad de vectores, por ejemplo, plásmidos, cósmidos, derivados del fago lambda, fagémidos y similares, que son muy conocidos en la técnica. Por consiguiente, también se proporciona un vector que incluye un polinucleótido de la divulgación y una célula hospedadora que contiene el polinucleótido. En general, el vector contiene un origen de replicación funcional en al menos un organismo, sitios de restricción por endonucleasas convenientes y un marcador de selección para la célula hospedadora. Los vectores de acuerdo con la divulgación incluyen vectores de expresión, vectores de replicación, vectores de generación de sondas y vectores de secuenciación. Una célula hospedadora de acuerdo con la divulgación puede ser una célula procariota o eucariota y puede ser un organismo unicelular o parte de un organismo multicelular.

40 La presente divulgación proporciona adicionalmente construcciones recombinantes que comprenden un ácido nucleico de la divulgación o un fragmento del mismo. Las construcciones recombinantes de la presente divulgación comprenden un vector, tal como un plásmido o vector viral, en el que se ha insertado un ácido nucleico de la divulgación o un fragmento del mismo, en una orientación directa o inversa. En el caso de un vector que comprende una de las ORF de la presente divulgación, el vector puede comprender además secuencias reguladoras, que incluyen por ejemplo, un promotor, unido operativamente a la ORF. Para los vectores que comprenden los FME y los FMC de la presente divulgación, el vector puede comprender además una secuencia marcadora o una ORF heteróloga unida operativamente al FME o UMF. Un gran número de vectores y promotores adecuados son conocidos para los expertos en la materia y están disponibles en el comercio para generar las construcciones recombinantes de la presente divulgación. Los siguientes vectores proporcionan a modo de ejemplo. Bacterianos: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eucariotas: pWLeo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

El polinucleótido aislado de la divulgación puede estar unido operativamente a una secuencia de control de la expresión tal como los vectores de expresión pMT2 o pED divulgados en Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), con el fin de producir la proteína o polipéptido recombinantemente. Muchas secuencias de control de la expresión adecuadas son conocidas en la técnica. Los métodos generales de expresión de proteínas recombinantes también son conocidos y se ejemplifican en R. Kaufman, Methods in Enzymology 185, 537-566 (1990). Tal como se define en el presente documento, "unido operativamente" significa que el polinucleótido aislado y una secuencia de control de la expresión están situados dentro de un vector o célula de tal modo que la proteína o el polipéptido se expresan por una célula hospedadora que se ha transformado (transfectado) con el polinucleótido/secuencia de control de la expresión ligado.

Las regiones promotoras pueden seleccionarse de cualquier gen deseado usando vectores CAT (cloranfenicol transferasa) u otros vectores con marcadores de selección. Dos vectores apropiados son pKK232-8 y pCM7. Los promotores bacterianos particulares nombrados incluyen, *lacZ*, T3, T7, *gpt*, λ P_R y *trc*. Los promotores eucariotas incluyen el intermedio temprano de CMV, el de timidina quinasa de VSH, y el temprano y tardío de SV40, los LTR de retrovirus y la metalotioneína-I de ratón. La selección del vector apropiado y el promotor está muy dentro del nivel de una experiencia habitual en la técnica. Generalmente, los vectores de expresión recombinantes incluirán orígenes de replicación y marcadores de selección que permiten la transformación de la célula hospedadora, por ejemplo, el gen de resistencia a ampicilina de *E. coli* y el gen TRP1 de *S. cerevisiae* y un promotor derivado de un gen altamente expresado para dirigir la transcripción de una secuencia estructural aguas abajo. Dichos promotores pueden derivarse de operones que codifican enzimas glucolíticas tales como 3-fosfoglicerato quinasa (PGK), factor a, fosfatasa ácida o proteínas del choque térmico, entre otras. La secuencia estructural heteróloga se ensambla en una fase adecuada con las secuencias de inicio y de terminación de la traducción, y preferentemente, una secuencia líder capaz de dirigir la secreción de la proteína o glucopéptido traducido al espacio periplasmático o el medio extracelular. Opcionalmente, la secuencia de heteróloga puede codificar una proteína de fusión que incluye un péptido de identificación N-terminal que aporta las características deseadas, por ejemplo, estabilización o purificación simplificada del producto recombinante expresado. Los vectores de expresión útiles para el uso bacteriano se construyen insertando una secuencia estructural de ADN que codifica una proteína polipeptídica deseada junto con señales de inicio y de terminación de la traducción adecuadas en una fase de lectura operativa con un promotor funcional. El vector comprenderá uno o más marcadores fenotípicos de selección y un origen de replicación para asegurarse del mantenimiento del vector para, si es deseable, proporcionar la amplificación dentro del hospedador. Los hospedadores eucariotas adecuados para la transformación incluyen *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* y diversas especies dentro del género *Pseudomonas*, *Streptomyces* y *Staphylococcus*, aunque también pueden emplearse otros como una materia de elección.

Como un ejemplo representativo pero no limitante, los vectores de expresión útiles para el uso en bacterias pueden comprender un marcador de selección y un origen de replicación bacteriano procedente de plásmidos disponibles en el comercio que comprenden elementos genéticos del vector de clonación bien conocido pBR322 (ATCC 37017). Dichos vectores comerciales incluyen, por ejemplo, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Suecia) y GEM 1 (Promega Biotec, Madison, WI, EE.UU.). Estas secciones de la 'estructura' de pBR322 se combinan con un promotor apropiado y la secuencia estructural a expresar. Después de la transformación en una cepa de hospedador adecuada y del crecimiento de la cepa hospedadora hasta una densidad celular apropiada, el promotor seleccionado se induce o se suprime por medios apropiados (por ejemplo, desplazamiento de la temperatura o inducción química) y las células se cultivan durante un periodo adicional. Las células se recogen típicamente mediante centrifugación, se rompen mediante un método físico o químico, y el extracto sin procesar resultante se retiene para la purificación posterior.

Los polinucleótidos de la divulgación se dirigen adicionalmente a secuencias que codifican variantes de polipéptidos o proteínas de la divulgación. Estas variantes de la secuencia de aminoácidos pueden prepararse mediante métodos conocidos en la técnica introduciendo cambios apropiados en los nucleótidos en un polinucleótido nativo o de la variante. Existen dos variables en la construcción de variantes de la secuencia de aminoácidos: la localización de la mutación y la naturaleza de la mutación. Las variantes de la secuencia de aminoácidos de los ácidos nucleicos se construyen preferentemente mutando el polinucleótido para producir una secuencia de aminoácidos que no se da en la naturaleza. Estas alteraciones de los aminoácidos pueden producirse en sitios que difieren de los ácidos nucleicos de distintas especies (posiciones variables) o en regiones altamente conservadas (regiones constantes). Los sitios en dichas localizaciones típicamente se modificarán en serie, por ejemplo, sustituyendo primero las elecciones conservativas (por ejemplo, un aminoácido hidrófobo por un aminoácido hidrófobo distinto) y después con elecciones más distantes (por ejemplo, un aminoácido hidrófobo por un aminoácido cargado), y después pueden hacerse deleciones e inserciones en el sitio diana. Las deleciones en la secuencia de aminoácidos generalmente varían de aproximadamente 1 a 30 restos, preferentemente de aproximadamente 1 a 10 restos y típicamente son contiguas. Las inserciones de aminoácidos incluyen fusiones amino-y/o carboxilo terminales que varían en su longitud de uno a cien o más restos, así como inserciones intrasecuencia de uno solo o múltiples restos de aminoácidos. Las inserciones intrasecuencia pueden variar generalmente de aproximadamente 1 a 10 restos de aminoácidos, preferentemente de 1 a 5 restos. Los ejemplos de inserciones terminales incluyen las secuencias señal heterólogas necesarias para la secreción o para el direccionamiento intracelular en distintas celulares hospedadoras.

Las deleciones en la secuencia de aminoácidos varían de aproximadamente 1 a 30 restos, preferentemente de aproximadamente 1 a 10 restos y típicamente son contiguas. Las inserciones de aminoácidos incluyen fusiones amino- y/o carboxilo-terminales que varían en su longitud de uno a cien o más restos, así como inserciones intrasecuencia de uno o múltiples restos de aminoácidos. Las inserciones intrasecuencia pueden variar generalmente de aproximadamente 1 a 10 restos de aminoácidos, preferentemente de 1 a 5 restos. Los ejemplos de inserciones terminales incluyen las secuencias de inserción heterólogas necesarias para la secreción o el direccionamiento intracelular para la secreción en distintas células hospedadoras.

También puede usarse la PCR para crear variantes de la secuencia de aminoácidos de los polinucleótidos de la divulgación. Cuando se usan cantidades pequeñas de ADN molde como material inicial, el/los cebador(es) que difiere(n) ligeramente de la secuencia de la región correspondiente en el ADN molde pueden generar la variante de aminoácidos deseada. En la amplificación por PCR da como resultado una población de productos de fragmentos de ADN que difieren del molde de polinucleótido que codifica el polipéptido o proteína en la posición especificada por el cebador. Los fragmentos de ADN de los productos reemplazan la región correspondiente en el plásmido y esto proporciona la variante de aminoácidos deseada.

En un método preferente, los polinucleótidos que codifican los polinucleótidos de la divulgación cambian mediante mutagénesis de sitio dirigido. Este método usa secuencias de oligonucleótidos que codifican la secuencia de polinucleótidos de la variante de aminoácidos deseada, así como un nucleótido adyacente suficiente en ambos lados del aminoácido cambiado para formar un duplo estable en cualquiera de las caras del sitio que se cambia. En general, las técnicas de mutagénesis de sitio dirigido son bien conocidas para los expertos en la materia y esta técnica se ejemplifica mediante publicaciones tales como, Edelman *et al.*, DNA 2:183 (1983). Un método versátil y eficaz para producir cambios específicos de sitio en una secuencia de polinucleótidos se publicó por Zoller y Smith, Nucleic Acids Res. 10:6487-6500 (1982). También puede usarse la PCR para crear variantes en la secuencia de aminoácidos de los nuevos ácidos nucleicos. Cuando se usan pequeñas cantidades de ADN molde como material inicial, el/los cebador(es) que difieren ligeramente de la secuencia de la región correspondiente en el ADN molde pueden generar la variante de aminoácidos deseada. La amplificación por PCR da como resultado una población de productos de fragmentos de ADN que difieren del polinucleótido molde que codifica el polipéptido en la posición especificada por el cebador. El producto de los fragmentos de ADN reemplaza la región correspondiente en el plásmido y esto proporciona la variante de aminoácido deseada.

Una técnica adicional para generar variantes de aminoácidos es la técnica de mutagénesis en casete descrita en Wells *et al.*, Gene 34:315 (1985); y otras técnicas de mutagénesis bien conocidas en la materia, tales como, por ejemplo, las técnicas de Sambrook *et al.*, anteriormente mencionado, y Current Protocols in Molecular Biology, Ausubel *et al.*

La presente divulgación también proporciona formas tanto de longitud completa como maduras de los polipéptidos o proteínas de la divulgación. La forma de longitud completa de dichos polipéptidos o proteínas puede identificarse mediante la traducción de la secuencia de nucleótidos de cada uno de los polinucleótidos de la divulgación. La forma madura de dicho polipéptido o proteína puede obtenerse mediante la expresión del polinucleótido de longitud completa en una célula de mamífero adecuada u otra célula hospedadora. La secuencia de la forma madura del polipéptido o proteína también puede ser determinable a partir de la secuencia de aminoácidos de la forma de longitud completa.

Cuando la proteína o el péptido de la presente divulgación está unida a la membrana (por ejemplo, es un receptor), la presente divulgación también proporciona formas solubles de dicha proteína o polipéptido. En dichas formas, parte o todos los dominios intracelulares y transmembrana de la proteína y del polipéptido se eliminan de tal modo que la proteína o polipéptido se secreta completamente a partir de la célula en la que se expresa. Los dominios intracelulares y transmembrana de las proteínas o polipéptidos de la divulgación pueden identificarse de acuerdo con técnicas conocidas para la determinación de dichos dominios a partir de la información de la secuencia.

La divulgación también se refiere a métodos para producir un polipéptido o proteína de la divulgación que comprenden el crecimiento de un cultivo de las células de la divulgación en un medio de cultivo adecuado y la purificación de la proteína o polipéptido de la divulgación a partir del cultivo. Por ejemplo, los métodos de la invención incluyen un proceso para producir un polipéptido o proteína de la divulgación en el que una célula hospedadora que contiene un vector de expresión adecuado que incluye un polinucleótido o proteína de la divulgación se cultiva en condiciones que permiten la expresión del polipéptido o proteína codificada. El polipéptido o proteína puede recuperarse a partir del cultivo, convenientemente a partir del medio de cultivo y purificarse adicionalmente.

La presente divulgación además proporciona polipéptidos o proteínas aislados codificados por los polinucleótidos de la presente divulgación o por variantes degeneradas de los polipéptidos de la presente divulgación. Por "variante degenerada" se entienden los polinucleótidos que difieren de un fragmento de ácido nucleico de la presente divulgación (por ejemplo, una ORF en la secuencia de nucleótidos pero, debido a la degeneración del código genético, codifica una secuencia polipeptídica idéntica. Los polinucleótidos preferentes de la presente divulgación

son las ORF que codifican proteínas o polinucleótidos. Pueden utilizarse una variedad de metodologías conocidas en la técnica para obtener uno o cualquiera de los polinucleótidos o proteínas aislados de la presente divulgación. En el nivel más simple, la secuencia de aminoácidos puede sintetizarse usando sintetizadores peptídicos disponibles en el comercio. Esto es particularmente útil para producir pequeños péptidos y fragmentos de péptidos mayores. Los fragmentos son útiles, por ejemplo, para la generación de anticuerpos contra el polipéptido nativo. En un método alternativo, el polipéptido o proteína se purifica a partir de células bacterianas que producen naturalmente el polipéptido o la proteína. Un experto en la materia puede seguir fácilmente métodos conocidos para aislar polipéptidos y proteínas con el fin de obtener uno de los polipéptidos o proteínas aislados de la presente divulgación. Estos incluyen, pero no se limitan a, inmunocromatografía, HPLC, cromatografía de exclusión molecular, cromatografía de intercambio iónico y cromatografía de inmutofinidad. Véase, por ejemplo, Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, *et al.*, en *Molecular Cloning: A Laboratory Manual*; Ausubel *et al.*, *Current Protocols in Molecular Biology*. Los polipéptidos y proteínas de la presente divulgación pueden purificarse como alternativa a partir de células que se han alterado para que expresen el polipéptido o proteína deseados. Tal como se usa en el presente documento, se dice que una célula está alterada para que exprese un polipéptido o proteína deseado cuando se hace que la célula, a través de manipulación genética, produzca un polipéptido o proteína que normalmente no produce o que la célula normalmente produce a un nivel bajo. Un aspecto de la materia puede adaptar fácilmente procedimientos para introducir o expresar bien secuencias recombinantes o sintéticas en células eucariotas o procariontas con el fin de generar una célula que produzca uno de los polipéptidos o proteínas de la presente divulgación.

La proteína o polipéptido de la divulgación puede expresarse también como un producto de animales transgénicos no humanos, por ejemplo, como un componente de la leche de vacas, cabras, cerdos u ovejas transgénicos que están caracterizados por que las células somáticas o germinales contienen un polinucleótido que codifica la proteína o el polipéptido de la divulgación.

Las proteínas o polipéptidos de la divulgación proporcionados en el presente documento también incluyen proteínas o polipéptidos caracterizados por secuencias de aminoácidos similares a las de las proteínas o polipéptidos purificados de la divulgación pero en las que las modificaciones se han proporcionado naturalmente o se han modificado deliberadamente. Por ejemplo, los expertos en la materia pueden realizar modificaciones de las secuencias del péptido o del ADN usando técnicas conocidas. Las modificaciones de interés en las secuencias proteicas pueden incluir la alteración, sustitución, reemplazo, inserción o delección de un resto de aminoácido seleccionado en la secuencia codificante. Por ejemplo, uno o más de los restos de cisteína pueden eliminarse o reemplazarse por otro resto de aminoácido para alterar la conformación de la molécula. Las técnicas para dicha alteración, sustitución, reemplazo, inserción o delección son bien conocidas para los expertos en la materia (véase, por ejemplo, patente de los EE.UU. N° 4.518.584). Preferentemente, dicha alteración, sustitución, reemplazo, inserción o delección retiene la actividad deseada de la proteína o polipéptido.

Otros fragmentos y derivados de los polipéptidos o proteínas de la divulgación de los que podría esperarse que retuvieran la actividad proteica completamente o en parte pueden ser, por lo tanto, útiles para la exploración u otras metodologías inmunológicas que pueden realizarse fácilmente por aquellos expertos en la materia dadas las divulgaciones del presente documento.

La proteína o polipéptido de la divulgación también puede producirse uniendo operativamente un polinucleótido de la divulgación a secuencias de control adecuadas en uno o más vectores de expresión en insectos y empleando un sistema de expresión en insectos. Los materiales y métodos para los sistemas de expresión en células de baculovirus/insectos están disponibles en el comercio en forma de un kit, por ejemplo, de Invitrogen, San Diego, Calif., EE.UU. (el kit MaxBat.RTM.), y dichos métodos son muy conocidos en la técnica, tal como se describe en Summers y Smith, Texas Agricultural Experiment Station Boletín N°. 1555 (1987). Tal como se usa en el presente documento, se "transforma" una célula de insecto capaz de expresar un polinucleótido de la presente divulgación.

La proteína o polipéptido de la divulgación puede prepararse cultivando células hospedadoras transformadas en condiciones de cultivo adecuadas para expresar la proteína o polipéptido recombinante. La proteína o polipéptido expresado resultante puede purificarse después a partir de dicho cultivo (es decir, a partir del medio de cultivo o de los extractos celulares) usando procesos de purificación conocidos, tales como, filtración en gel y cromatografía de intercambio iónico. La purificación de la proteína o polipéptido también puede incluir una columna de afinidad que contiene agentes que unirán la proteína o el polipéptido; una o más de las etapas de columna a lo largo de resinas de afinidad tales como concanavalina A-agarosa, heparina Toyopearl.RTM o azul de Cibacrom 3GA Sefarosa.RTM.; una o más etapas que implican cromatografía de integración hidrófoba usando resinas tales como éter de fenilo, éter de butilo o éter de propilo; o cromatografía de inmutofinidad.

Como alternativa, la proteína o polipéptido de la divulgación también puede expresarse en una forma que facilitará la purificación. Por ejemplo, puede expresarse como una proteína de fusión, tal como la de la proteína de unión a maltosa (MBP), glutation-S-transferasa (GST) o tiorredoxina (TRX). Los kits para la expresión y purificación de dichas proteínas de fusión están disponibles en el comercio de England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) e In Vitrogen, respectivamente. La proteína o polipéptido también puede estar marcada con un epítipo y purificarse posteriormente usando un anticuerpo específico dirigido a dicho epítipo. Uno de dichos epítipos ("Flag") está disponible en el comercio de Kodak (New Haven, Conn.).

Finalmente, pueden emplearse una o más etapas de cromatografía líquida de alto rendimiento en fase inversa (RP-HPLC) que emplean medias hidrófobas de RP-HPLC, por ejemplo, gel de sílice que tiene grupos metilo colgantes u otros grupos alifáticos, para purificar adicionalmente la proteína o el polipéptido. También pueden emplearse algunas o todas las etapas de purificación anteriores, en diversas combinaciones, para proporcionar una proteína o polipéptido recombinante sustancialmente homogéneo. La proteína o el polipéptido purificado de este modo están sustancialmente libre de otras proteínas de mamíferos y se define de acuerdo con la presente divulgación como una "proteína aislada".

Como una elección de formato de anticuerpo, se prefirió el formato Fab sobre el formato scFv, ya que el formato Fab permite realizar ensayos de exploración de afinidad de flujo continuo alto para preparaciones de anticuerpo sin procesar. Muchos de los scFv forman, en efecto, especies de un peso molecular más alto que incluyen dímeros (Weidner, *et al.*, (1992) *J. Biol. Chem.* **267**, 10281-10288; Holliger, *et al.*, (1993) *Proc. Natl. Acad. Sci. EE.UU.* **90**, 6444-6448) y trímeros (Korttet *et al.* (1997) *Protein Eng.* **10**, 423-433), lo que complica tanto la selección como la caracterización. Se ha elegido el formato de presentación de Fab en el que un dominio variable de un gen de una cadena pesada o una ligera se une a una proteína de recubrimiento fágico, y en algunas realizaciones, también porta un marcador para la detección y la purificación. La otra cadena se expresa como un fragmento distinto secretado en el periplasma, donde puede emparejarse con el gen que está en una proteína de fusión con la proteína de recubrimiento fágico (Hoogenboom, *et al.*, (1991) *Nucleic Acids Res.* **19**, 4133-4137). En algunas realizaciones, la proteína de recubrimiento fágico es una proteína de recubrimiento pIII. En otras realizaciones, el dominio variable de un gen de la cadena pesada se fusiona a la proteína de recubrimiento fágico y el gen de la cadena ligera se expresa como un fragmento distinto.

La elección del formato Fab se basó en la noción de que el aspecto monomérico de los Fab permite la exploración rápida de un gran número de clones con respecto a su cinética de unión (velocidad de disociación) con fracciones de proteínas sin procesar. Esto reduce drásticamente el tiempo del análisis post-selección cuando se compara con el necesario para los anticuerpos de Fv monocatenarios (scFv) a partir de bibliotecas de fagémidos (Vaughan, *et al.*, (1996) *Nat. Biotechnol.* **14**, 309-314; Sheets, *et al.*, (1998) *Proc. Natl. Acad. Sci. EE.UU.* **95**, 6157-6162) o fragmentos Fab de otras fagotecas (Griffiths, *et al.*, (1993) *EMBOJ.* **12**, 725-734).

La biblioteca de Fab obtenible mediante el método de la invención produjo, por término medio, 14 Fab distintos contra los 6 antígenos que se ensayaron. Estos incluyeron el toxoide tetánico, el hapteno de fenil-oxazolona, el antígeno MUC1 asociado a cáncer de mama y tres hormonas glucoproteicas altamente relacionadas: la gonadotropina coriónica humana ("hCG"), la hormona luteinizante humana ("hLH") y la hormona estimulante del folículo humana ("hFSH"). Para las hormonas glucoproteicas, la biblioteca de Fab obtenible por el método de la invención produjo un panel bien de anticuerpos específicos de hormonas o de anticuerpos de reactividad cruzada. Por lo tanto, sin usar protocolos de selección sofisticados, se recuperaron los Fab específicos así como los que tenían reactividad cruzada contra estas glucocorticoides altamente homólogas, demostrando que la biblioteca es una fuente rica de especificidades de anticuerpo. Las afinidades de los anticuerpos anti-glucocorticoides varió entre 2,7 y 38 nM. Finalmente, el formato de Fab permitió, en efecto, la exploración rápida y la clasificación fiable de clones individuales basándose en la velocidad de disociación usando fracciones sin procesar. Además, las especificidades de los anticuerpos obtenidos mediante selecciones en las gonadotropinas son únicas: debido al alto grado de homología entre hLH y hCG ha sido muy difícil aislar anticuerpos monoclonales específicos para hCG con la tecnología del hibridoma, mientras que hay muy pocos anticuerpos específicos de hLH (Moyle, *y col.*, (1990) *J. Biol. Chem.* **265**, 8511-8518; Cole, (1997) *Clin. Chem.* **43**, 2233-2243). Usando un procedimiento de selección directa, sin tomar ninguna precaución para impedir la selección de Fab con reactividad cruzada, se han aislado fácilmente fragmentos con todas las especificidades posibles: Los Fab específicos para cualquiera de las tres hormonas hCG, hLH y hFSH, y los Fab de reactividad cruzada que reconocen las cadenas- α o epítomos comunes en la cadena β compartida por hCG y hLH. Estas selecciones demostraron que pueden recuperarse anticuerpos dirigidos contra distintos epítomos dentro de moléculas antigénicas únicas a partir de la biblioteca. La biblioteca de Fab obtenible por el método de la invención permite el control de las selecciones con preparaciones de fagos policlonales y la exploración a gran escala de las velocidades de disociación de los anticuerpos con fragmentos de Fab no purificados. En su conjunto, se recuperaron anticuerpos con unas velocidades de disociación en el orden de 10^{-2} a 10^{-4} s $^{-1}$ y afinidades de hasta 2,7 nM. La cinética de estas fagotecas es del mismo orden de magnitud que la de los anticuerpos asociados con una respuesta inmunitaria secundaria.

Una indicación de que los anticuerpos de la biblioteca de Fab se comportan de una manera similar o mejor que los anticuerpos de una biblioteca de scFv con respecto a su afinidad procede de una comparación de las selecciones de dos bibliotecas de anticuerpos distintas en los mismos dos antígenos en condiciones idénticas. Los anticuerpos para MUC1 seleccionados a partir de una biblioteca de scFv vírgenes de gran tamaño (Henderikx *et al.*, (1998) *Cancer Res.* **58**, 4324-4332) pueden tener velocidades de disociación más rápidas que los Fab equivalentes aislados a partir de la biblioteca descrita en este estudio. Además, estos muestran un uso muy distinto de los genes V y tienen una especificidad fina distinta. De un modo similar, cuando se comparan las velocidades de disociación de los anticuerpos fágicos contra el marcador de pancarcinoma Glucoproteína Epitelial 2, uno de los Fab seleccionados a partir de la presente biblioteca parece tener una velocidad de disociación 10 veces más lenta que el mejor scFv (Vaughan *et al.*, (1996) *Nat. Biotechnol.* **14**, 309-314).

Las afinidades de los fragmentos de anticuerpo seleccionados son, sin embargo, muy dependientes del antígeno usado para la selección. Sheets y colaboradores notificaron una afinidad que variaba entre 26 y 71 nM para los fragmentos scFv específicos seleccionados para el anti-neurotoxina de tipo A de *Clostridia botulinum*, mientras que para los anticuerpos para el dominio extracelular del Erb-2 humano, se encontraron K_d entre 0,22 y 4,03 nM (Sheets *et al.*, (1998) *Proc. Natl. Acad. Sci. EE.UU.* **95**, 6157-6162). Las afinidades de los Fab específicos de gonadotropinas seleccionados a partir de esta biblioteca variaron entre 2,7 y 38 nM, lo que es comparable a los valores de la proteína de unión de la biblioteca virgen producida por Vaughan *et al.* y Sheets *et al.* y se aproxima a los valores de los mejores anticuerpos de su tipo. El tamaño de la biblioteca de Fab obtenible mediante el método de la invención no solamente es importante por su afinidad, sino que también determina la tasa de éxito de la selección de anticuerpos contra un gran conjunto de antígenos distintos. A este respecto, la biblioteca de Fab obtenible por el método de la invención funciona muy bien: se seleccionaron por encima de 24 anticuerpos para el hapteno pH_{0x} y un promedio de 13 anticuerpos contra los otros antígenos.

En el conjunto limitado de 14 clones de Fab que se secuenciaron, se identificaron anticuerpos con genes de la región variable de la totalidad de las grandes familias génicas V, incluyendo $V_{H1/3/4}$, $V_{K1/3}$ y $V_{\lambda 1/2}$, pero también se recuperaron los segmentos usados con menor frecuencia de la familia V_{H6} , $V_{K2/7}$ y $V_{\lambda 7}$. Lo más probable, es que el uso de un conjunto extenso de cebadores de genes de la región variable, diseñados basándose en la información de secuencia más reciente de las regiones V de la línea germinal y/o las PCR separadas, combinadas con la clonación parcialmente separada, aseguró el acceso a una muestra altamente diversa del repertorio de genes V humanos.

Se prepara una biblioteca a partir de polinucleótidos que son capaces de codificar el miembro del par de unión específico deseado. Existe una variedad de técnicas para la preparación de la biblioteca, que puede prepararse, por ejemplo, bien a partir de ADN genómico o ADNc. Véase, por ejemplo, Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual*, 2ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989. Las células pueden servir como la fuente de los polinucleótidos que codifican los miembros de pares específicos de unión de interés. Pueden emplearse procedimientos enriquecimiento y medios para amplificar las regiones que contienen el/los gene(s). Por ejemplo, cuando el miembro del par de unión específico es un anticuerpo, puede prepararse ARN y/o ADN genómico, por ejemplo, a partir de células de bazo obtenidas a partir de un animal no inmunizado, a partir de un animal inmunizado con la(s) diana(s) de interés, a partir de células de hibridoma o a partir de células linfoblastoides. La biblioteca de anticuerpos obtenidos a partir de los animales no inmunizados contiene una representación no sesgada del repertorio completo de anticuerpos, mientras que la biblioteca de anticuerpos obtenidos a partir de los animales inmunizados contiene una población sesgada de anticuerpos dirigidos contra epítopos de la(s) diana(s). Las células de bazo o las células inmunitarias de otros tejidos o del sistema circulatorio pueden obtenerse a partir de una variedad de especies animales, tales como seres humanos, ratón, rata, equinos, bovinos, aves, etc.

La amplificación del ARN mensajero (ARNm) aislado de células de interés, tales como de células de bazo o de hibridoma, puede realizarse de acuerdo con protocolos esquematizados, por ejemplo, en la patente de EE.UU. N° 4.683.202, Orlandi, *et al.* *Proc. Natl. Acad. Sci. EE.UU.* **86**:3833-3837 (1989). Sastry *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* **86**:5728-5732 (1989) y Huse *et al.* *Science* **246**:1275-1281 (1989), Abelson, J. y Simon, M. (eds), *Methods in Enzymology, combinatorial chemistry*, Vol. 267, San Diego: Academic Press (1996), Kay, B.K., Winter, J., McCafferty, J. (eds), *Phage Display of peptides and Proteins, a Laboratory Manual*, San Diego: Academic Press (1996). Los cebadores oligonucleotídicos útiles en los protocolos de amplificación pueden ser únicos o degenerar o incorporar inosina en posiciones degeneradas. Por lo tanto, para las inmunoglobulinas multicaténarias, deberían usarse los cebadores generalmente para la amplificación de secuencias que codifican las regiones variables tanto de cadena pesada como de cadena ligera. Pueden incorporarse secuencias de reconocimiento de endonucleasas de restricción en los cebadores para permitir la clonación del fragmento amplificado en un vector en una fase de lectura abierta predeterminada para su expresión.

Las bibliotecas de expresión que contienen el ADNc amplificado se preparan típicamente en un vector tal como un bacteriófago o un fagémido. Las características del bacteriófago o fagémido adecuado dependen de la realización específica empleada, y generalmente serán aquellos que permiten convenientemente la inserción de los polinucleótidos recombinantes en células hospedadoras mediante el empaquetamiento *in vitro* o la transformación.

Después, se infectan las células hospedadoras con el fago o fagémido y el fago auxiliar y se cultivan en condiciones que permitan la expresión y el ensamblaje de las partículas fágicas. En una realización, se eligen las células hospedadoras apropiadas para los bacteriófagos o fagémidos del método de la invención son diversas cepas de *E. coli*, los ejemplos específicos de las cuales dependen de los diversos vectores adecuados. Por supuesto, también puede usarse un fago o fagémido que tenga hospedadores bacterianos distintos a *E. coli*.

Para enriquecer y aislar las partículas fágicas o fagos que contienen las secuencias clonadas de la biblioteca que codifican un miembro de un par de unión específico deseado y, por lo tanto, aislar finalmente las secuencias de ácidos nucleicos en sí mismas, se purifican por afinidad partículas fágicas o fagos recogidos de las células hospedadoras. Se usa una diana o un compañero de unión para el miembro del par de unión específico en la purificación por afinidad. Por ejemplo, cuando el miembro del par de unión específico deseado es un anticuerpo, que

se une específicamente una diana particular, se usa la diana para recuperar partículas fágicas o fagos que tienen el anticuerpo deseado en su superficie externa. La diana se adsorbe típicamente a un sustrato insoluble, tal como una partícula o una esfera o una placa. Las partículas fágicas o fagos obtenidos de este modo pueden amplificarse después infectando células hospedadoras (con fagos auxiliares con respecto a las partículas fágicas que contienen los fagémidos). Pueden emplearse rondas de enriquecimiento y amplificación por afinidad adicionales hasta que se alcance el nivel deseado de enriquecimiento o las partículas fágicas o el fago deseado no se enriquezcan más con respecto a las partículas fágicas o el fago de fondo.

Las partículas anticuerpo-fago o fágicas enriquecidas también se exploran con técnicas de detección adicionales tales como levantamiento de placas (o colonias) de expresión (véase, por ejemplo, Young y Davis, *Science*, **222**: 778-782 (1983)) de modo que se usa el mismo compañero de unión u otro como una sonda. La exploración puede emplear ensayos adicionales (para una actividad catalítica, por ejemplo) que se usan para detectar, *in situ*, placas que expresan miembros de pares de unión específicos que tienen las características deseadas. Las partículas fágicas o fagos obtenidos a partir del protocolo de exploración se infectan en células, se propagan y se aísla el ADN de la partícula fágica o fago y se secuencian, y/o se reclonan en un vector destinado a la expresión génica en procariontes o eucariotes para obtener cantidades más grandes del miembro del par de unión específico seleccionado.

En otra realización, el miembro del par de unión específico codificado por la biblioteca (o múltiples cadenas que codifican dicho miembro del par de unión específico) se transporta a un compartimento extra-citoplasmático de la célula hospedadora, habitualmente el espacio periplasmático, para facilitar el procesamiento y/o el ensamblaje apropiado. Cuando se emplea el transporte extra-citoplasmático del miembro del par de unión específico deseado, las secuencias que codifican el miembro del par de unión específico se clonan adyacentes a señales transcripcionales y traduccionales apropiadas y a líderes del péptido señal que dirigirán las cadenas maduras al periplasma. Tal como se hizo anteriormente, al menos una de las cadenas se clona como una proteína de fusión con una proteína de recubrimiento fágico de modo que la proteína de recubrimiento fágico no interfiera sustancialmente con la capacidad del miembro del par de unión de interés para la unión a una diana que se usa en el protocolo de enriquecimiento por afinidad.

Un ejemplo preferente de esta realización es la colocación de un miembro de un par específico de unión de la región N-terminal de la proteína de recubrimiento menor pIII del bacteriófago fd. Antes de la incorporación en el fago, pIII reside en la membrana interna de la célula hospedadora con su extremo N-terminal protruyendo en el periplasma. En esta configuración el polipéptido de un miembro de un par específico de unión en el extremo N-terminal de pIII está disponible para la unión a otras cadenas polipeptídicas que constituyen el miembro del par de unión específico de interés. Este complejo se incorpora después en la partícula fágica o el fago maduro cuando sale de la célula y el extremo C-terminal se incrusta en el recubrimiento de la partícula fágica o fago.

En esta realización la síntesis y amplificación de polinucleótidos es tal como se describe anteriormente, y después se clona en o cerca de una secuencia de un vector que codifica una proteína de recubrimiento, en la que el vector es, o deriva de, un fago filamentoso, tal como f1, fd, Pf1, M13, etc. En una realización preferente, el fago filamentoso es fd-tet. El vector fágico se elige para que contenga un sitio de clonación localizado en la región 5' de un gen que codifica una proteína de recubrimiento fágico, tal como, por ejemplo, la proteína de recubrimiento de pIII. Un vector apropiado (por ejemplo, fd-tet B1 que se describe más adelante) permite la clonación orientada de secuencias extrañas de modo que se expresen en o cerca del extremo N-terminal de la proteína de recubrimiento madura.

Se construye una biblioteca clonando los polinucleótidos (por ejemplo, la región V_H) a partir de las células donantes en un sitio de clonación de un gen de una proteína de recubrimiento (por ejemplo, gen III, "gIII"). Las secuencias clonadas de, por ejemplo, los dominios V_H se expresan finalmente como polipéptidos o proteínas fusionadas al extremo N-terminal de la proteína de recubrimiento madura en la superficie externa accesible de las partículas fágicas o el fago ensamblados.

Cuando la proteína deseada es una proteína multicaténaria, dicho anticuerpo o fragmento de unión del mismo, el polinucleótido que codifica la(s) cadena(s) no clonada(s) en una proteína de recubrimiento fágico puede clonarse directamente en un sitio apropiado (tal como se describe anteriormente) del vector que contiene la biblioteca de la primera cadena-proteína de recubrimiento; o preferentemente, la(s) cadena(s) posterior(es) pueden clonarse como una biblioteca separada en un vector plasmídico distinto, amplificarse, y posteriormente, instalarse los fragmentos en el vector de la biblioteca de la primera cadena-proteína de recubrimiento. Por ejemplo, cuando la primera cadena es una cadena pesada de un anticuerpo o un fragmento de la misma, el destino final de la secuencia de ADNc del V_L de la cadena ligera está en un vector que ya contiene una secuencia de V_H en un gen de una proteína de recubrimiento, recombinándose aleatoriamente las secuencias V_H y V_L en un solo vector.

La segunda cadena o posterior de la proteína multicaténaria deseada, tal como V_L, se clona de tal modo que se expresa con una secuencia líder de un péptido señal que dirigirá su secreción al periplasma de la célula hospedadora. Por ejemplo, se ha mostrado que varias secuencias líder dirigen la secreción de secuencias de anticuerpos en *E. coli*, tales como OmpA (Hsiung, *et al.*, *Biotechnology* 4:991-995 (1986)), pelB (Better, *et al.*, *Science* 240:1041-1043 (1988)), phoA (Skerra y Pluckthun, *Science* 240:1038-1043 (1988)), beta-lactamasa (Zemel-

Dreazen y Zamir, *Gene* 27:315-322 (1984). y aquellas descritas en Abelson, J. y Simon, M. (eds), *Methods in Enzymology, combinatorial chemistry*, Vol. 267, San Diego: Academic Press (1996), y Kay, B.K., Winter, J., McCafferty, J. (eds), *Phage Display of peptides and Proteins, a Laboratory Manual*, San Diego: Academic Press (1996).

5 Generalmente, la estrategia de clonación exitosa que utiliza una proteína de recubrimiento fágico, tal como pIII del fago filamentoso fd, proporcionará; (1) la expresión de una cadena proteica (o una primera cadena polipeptídica cuando la proteína deseada es multicaténaria, por ejemplo, la cadena V_H) fusionada al extremo N-terminal de una proteína de recubrimiento de tamaño completo (o casi de tamaño completo) (por ejemplo, pIII) y transporte a la
10 membrana interna del hospedador donde el dominio hidrófobo de la región C-terminal de la proteína de recubrimiento ancla la proteína de fusión en la membrana, conteniendo el extremo N-terminal la cadena que protruye en el espacio periplasmático y que está disponible para la interacción con una segunda cadena o posterior (por ejemplo, V_L para formar un fragmento Fab) que, por lo tanto, está unido a la proteína de recubrimiento; y (2) la expresión adecuada de una segunda cadena polipeptídica o posterior si está presente (por ejemplo, V_L) y el
15 transporte de esta cadena al compartimento soluble del periplasma.

En una realización para el enriquecimiento por afinidad de los clones deseados, se incuban aproximadamente de 10³ a 10⁴ equivalentes de biblioteca (un equivalente de biblioteca es cada uno de los equivalentes recombinantes -- 10⁴ de una biblioteca de 10⁹ miembros es 10⁹ x 10⁴ = 10¹³ partículas fágicas o fagos) con una diana para la cual se
20 busca el miembro del par de unión específico deseado (por ejemplo, anticuerpo). La diana está en una de varias formas apropiadas para los esquemas enriquecimiento por afinidad. En un ejemplo la diana está inmovilizada sobre una superficie o una partícula, opcionalmente anclada mediante una fijación de la longitud suficiente (de 3 a 12 carbonos, por ejemplo) por ejemplo para mantener la diana lo suficientemente lejos de la superficie para permitir la interacción libre con el sitio de combinación del anticuerpo. La biblioteca de partículas fágicas o fagos que portan los
25 anticuerpos se criba después sobre la diana inmovilizada generalmente de acuerdo con procedimientos bien conocidos en la técnica, por ejemplo aquellos descritos en Abelson, J. y Simon, M. (eds), *Methods in Enzymology, combinatorial chemistry*, vol. 267, San Diego: Academic Press (1996), Kay, B.K., Winter, J., McCafferty, J. (eds), *Phage Display of peptides and Proteins, a Laboratory Manual*, San Diego: Academic Press (1996).

30 Un segundo ejemplo de presentación de diana es una diana unida a un ligando reconocible (de nuevo opcionalmente con una fijación de alguna longitud). Un ejemplo específico de dicho ligando es la biotina. La diana, modificada de este modo, se incuba con la biblioteca partículas fágicas o fagos y se da la unión con ambos reactantes en solución. Los complejos resultantes se unen después con estreptavidina (o avidina) a través del resto de biotina. La estreptavidina puede inmovilizarse sobre una superficie tal como una placa de plástico o sobre
35 partículas, en cuyo caso los complejos se retienen físicamente; o la estreptavidina puede estar marcada, con un fluoróforo, por ejemplo, al marcador del fago/anticuerpo activo para la detección y/o aislamiento mediante procedimientos de separación, por ejemplo, en un separador celular de fluorescencia activada.

En una realización, las partículas fágicas o fagos que portan anticuerpos sin la especificidad deseada se retiran por
40 diversos medios, por ejemplo, por lavado. El grado y la rigurosidad del lavado necesario se determinarán para cada miembro de par de unión específico de interés. Puede ejercerse un determinado grado de control sobre las características de unión de los anticuerpos recuperados ajustando las condiciones de la incubación de unión y los lavados posteriores. La temperatura, el pH, la fuerza iónica, la concentración de los cationes divalentes y el volumen y la duración del lavado se seleccionarán para anticuerpos dentro intervalos particulares de afinidad para el hapteno.
45 La selección basada en la velocidad de disociación lenta que, es habitualmente predictiva de una alta afinidad, es la ruta más práctica. Esta puede hacerse bien mediante la incubación continua en presencia de una cantidad saturante del hapteno libre, o aumentando el volumen, número y longitud de los lavados. En cada caso, se impide la re-unión del anticuerpo-fago disociado, y con tiempo creciente, se recupera el anticuerpo-fago con una afinidad cada vez más alta.

50 Los anticuerpos con determinadas actividades catalíticas pueden enriquecerse en grupos de anticuerpos con alta afinidad para reactantes (sustratos e intermedios) pero baja afinidad para los productos. Una doble exploración para enriquecer en anticuerpos con estas características puede ser útil para encontrar anticuerpos para catalizar determinadas reacciones. Además, también pueden seleccionarse anticuerpos catalíticos capaces de determinadas
55 reacciones de escisión. Una categoría de dichas reacciones es la escisión de un grupo terminal específico de una molécula. Por ejemplo, puede seleccionarse un anticuerpo catalítico para que escinda un aminoácido específico de un extremo de un péptido mediante la inmovilización del péptido y el cribado de la biblioteca de anticuerpos en condiciones de las que se espera que promuevan la unión pero no la escisión (por ejemplo, baja temperatura, fuerza iónica particular, pH, concentraciones de cationes, etc., dependiendo de la naturaleza del grupo terminal y de la
60 reacción de escisión) y seguido por un lavado. Esto permite que los anticuerpos que reconocen el grupo terminal se unan y se inmovilicen y a partir de este grupo estos serán capaces de la escisión. Para encontrar aquellos que son capaces de la escisión, las condiciones se desplazan a aquellas que son favorables para la escisión. Esta etapa liberará los anticuerpos-fago capaces de escindirse en sí mismos, liberándose del péptido inmovilizados.

65 Un modo alternativo de llevar a cabo esto es hacer un cribado de anticuerpos que se unen al grupo terminal específico mediante la unión del grupo terminal a un enlace distinto del que se va a escindir (un enlace no peptídico,

por ejemplo). Para el cribado posterior (del fago positivo a partir de la primera exploración) del grupo terminal unido por medio del enlace apropiado en condiciones de escisión, la fracción de hibridación se enriquecerá para aquellos con la actividad catalítica deseada.

5 Para eluir la partícula de anticuerpo-fago activa o el fago de la diana inmovilizada, después de la lavar con la rigurosidad apropiada, la partícula fágica o fago unido (activo) puede recuperarse eluyendo mediante un desplazamiento del pH. Por ejemplo, puede usarse pH 2 o pH 11 que después se neutraliza y el fago eluido se amplifica mediante la infección o transformación de las células hospedadoras. Las células crecen después como colonias resistentes a tetraciclina. Las colonias se raspan y el fago extruido se purifica mediante procedimientos
10 convencionales tal como se hizo anteriormente. Estos fagos se usan después en otra ronda de enriquecimiento por afinidad (cribado) y este ciclo se repite hasta que se alcanza el nivel de enriquecimiento deseado o hasta que el fago diana no se enriquece más con respecto a las partículas fágicas o fago de fondo. Para aislar clones individuales, las partículas fágicas o fago de la ronda final de cribado y elución se infectan en células o su ADN se transforman en células y crece sobre agar (habitualmente agar L) y antibióticos (habitualmente tet) para formar colonias individuales bien separadas, cada una de las cuales es un clon porta vectores tanto con secuencias V_H como V_L . Puede aislarse el ADN monocatenario de las partículas fágicas o fagos extruidos a partir de cada colonia y secuenciarse el ADN que codifica para los fragmentos V_H y V_L . La forma replicativa del ADN del fago (bicatenario) puede aislarse mediante medios convencionales y el ADN de los sitios de clonación (secuencias V_H y V_L) reclonarse en un vector diseñado para la expresión de producto génico en procariotas o eucariotas para obtener cantidades mayores de los anticuerpos particulares seleccionados en el proceso de exploración.

15 Los fagos identificados por tener un anticuerpo reconocido por el ligando diana se propagan tal como sea apropiado para el vector fágico particular usado. Para fd-tet, esto se hace en un cultivo líquido de un medio rico (caldo L, por ejemplo) con selección de antibiótico (Tet). Se recogen los fagos y se prepara el ADN y se secuencia mediante métodos convencionales para determinar la secuencia de ADN y de aminoácidos del anticuerpo particular.

20 El ADN puede reclonarse en un vector de expresión eucariota o procariota adecuado y transfectarse en un hospedador apropiado para la producción de grandes cantidades de proteína. El anticuerpo se purifica a partir del sistema de expresión usando procedimientos convencionales. La afinidad de unión del anticuerpo se confirma mediante inmunoensayos bien conocidos con el antígeno diana o actividad catalítica tal como se describe en Harlow y Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor, N.Y. (1988), Abelson, J. y Simon, M. (eds), *Methods in Enzymology, combinatorial chemistry*, vol. 267, San Diego: Academic Press (1996), Kay, B.K., Winter, J., McCafferty, J. (eds), *Phage Display of peptides and Proteins, a Laboratory Manual*, San Diego: Academic Press (1996).

25 En otra realización, se purifican por afinidad las partículas fágicas o fagos que muestran miembro del par de unión específico deseado del siguiente modo: aproximadamente de 10^3 - 10^4 equivalentes de bibliotecas de partículas fágicas o fagos reaccionan durante toda la noche con 1 picogramo de anticuerpo purificado a 4 °C. La mezcla se criba mediante un procedimiento del siguiente modo. Se recubre una placa petri de poliestireno con 1 ml de solución de estreptavidina (1 mg/ml en NaHCO_3 0,1 M, pH 8,6, NaN_3 al 0,02 %) y se incuba durante toda la noche a 4 °C. Al siguiente día se retira la solución de estreptavidina. Se rellena la placa con 10 ml de solución de bloqueo (30 mg/ml de BSA, 3 microgramos/ml de estreptavidina en NaHCO_3 0,1 M, pH 9,2, NaN_3 al 0,02 %) y se incuba durante 2 horas a temperatura ambiente. Se añaden dos microgramos de IgG de cabra anti-ratón biotinilada (BRL) a la biblioteca de anticuerpos con sometida a reacción con el anticuerpo y se incuba durante 2 horas a 4 °C. Inmediatamente antes del cribado, se retira la solución de bloqueo de la placa recubierta de estreptavidina y se lava la placa 3 veces con TBS/Tween 20 al 0,05 %. Después, se añade la biblioteca que se somete a reacción con el anticuerpo a la placa y se incuba durante 30 minutos a temperatura ambiente. También pueden usarse esferas de agarosa recubiertas de estreptavidina (BRL) para esta purificación por afinidad. Se retira la solución de la biblioteca y se lava la placa diez veces con TBS/Tween 20 al 0,05 % a lo largo del periodo de 60 minutos. Se retira el fago unido mediante la adición de un tampón de elución (1 mg/ml de BSA, HCl 0,1 N, pH ajustado a 2,2 con glicina) a la placa petri y se incuba durante 10 minutos para disociar los complejos inmunitarios. Se retira el eluato, se neutraliza con Tris 2 M (pH sin ajustar) y se usa para infectar células bacterianas que contienen F' en la fase log. Estas células se siembran después en placas de agar LB que contienen tetraciclina (20 .mu.g/ml) y crecen durante toda la noche a 37 °C. Las partículas fágicas o los fagos se aíslan a partir de estas placas tal como se describe y el proceso de purificación por afinidad se repite durante dos o tres rondas. Después en la última ronda de purificación, se usa una porción del eluato para infectar células y se siembra a baja densidad en placas de LB tetraciclina. Las colonias individuales se transfirieren a tubos de cultivo que contienen 2 ml de LB tetraciclina y crecen hasta la saturación. El ADN del fago o fagémido se aísla usando un método diseñado para la estación de trabajo Beckman Biomek (Mardis y Roe., *Biotechniques*, 7:840-850 (1989)) que emplea placas de microtitulación de 96 pocillos. Se secuencia El ADN monocatenario mediante el método didesoxi usando una Secuencia (U.S. Biochemicals) y un cebador oligonucleotídico de secuenciación (5'-CGATCTAAAGTTTTGTCGTCT-3' SEC ID N°: 2) que es complementario a la secuencia localizada a 40 nucleótidos 3' del segundo sitio de BstXI en fdTetB1.

30 Se consideran una serie de variables para abordar la construcción de una nueva biblioteca fágica de un tamaño muy grande; (i) el diseño de los cebadores se optimizó para la amplificación de conjuntos génicos variables para mantener la máxima diversidad; (ii) se desarrolló un método de clonación en dos etapas muy eficaz para obtener una biblioteca virgen de gran tamaño (iii) se seleccionó un formato de anticuerpo y un vector de clonación

compatibles, que debían permitir el análisis rápido aguas abajo de los clones seleccionados.

Con el fin de alcanzar el acceso al mayor número de segmentos génicos de la región V de cadena la pesada y la ligera humanas como fuera posible, se desarrolló un nuevo conjunto de cebadores oligonucleotídicos (Tabla I), cuyo diseño se basó en la información de secuencia más reciente proporcionada por la base V.

Los cebadores se diseñaron para que fuesen de la secuencia, o de varias secuencias, consenso que pudiesen tener al menos un 70 % de homología con los extremos 5' o 3' respectivos, basándose en la región codificante de los segmentos génicos de la línea germinal humana de la familia génica específica V que tendrían que amplificar. Los cebadores amplificarían al menos un segmento génico V usando las condiciones de PCR descritas más adelante, y en una realización están anexos a los sitios de restricción apropiadamente colocados para la clonación en el vector para la expresión de Fab.

Los cebadores deberían permitir la amplificación eficaz de todos los segmentos génicos V comúnmente usados. Además, para obtener las bibliotecas de Fab de gran tamaño de acuerdo con la invención (por encima de 10^{10} en diversidad), se usa un procedimiento de clonación en dos etapas: los genes variables de la cadena pesada y de la cadena ligera se clonan primero por separado como productos de PCR digeridos, y después se combinan mediante clonación de fragmentos de restricción para formar una biblioteca de gran tamaño de fragmentos Fab. Este procedimiento de clonación debería ser una ruta más eficaz para la construcción de bibliotecas que la clonación directa relativamente ineficaz de productos de PCR digeridos, mientras que se impide la inestabilidad del ADN asociada a menudo con los sistemas de recombinación *in vivo* (Griffiths, *et al.*, (1994) *EMBO J.* 13, 3245-3260).

Se construye un nuevo vector fagémido, pCES1 (Fig. 1), que permite la clonación en etapas de los fragmentos de anticuerpo en formato Fab. En este sistema de vector, los genes variables de región de cadena pesada se clonan como fragmentos génicos V_H ; el vector suministra todos los Fab con un dominio C_{H1} - gamma-1 humano. El $V_H C_{H1}$ formado mediante la inserción de fragmentos génicos V_H en el vector se fusiona (en el vector) con dos marcadores para la purificación y detección (una cola de histidina para la cromatografía de afinidad de metales inmovilizados (Hochuli, *et al.*, (1988) *BioTechnology* 6,1321-1325) y un marcador derivado de *c-myc* (Munro, *et al.*, H. R. (1986) *Cell* 46,291-300)), seguido de un codón de terminación ámbar (Hoogenboom, *et al.*, (1991) *Nucleic Acids Res.* 19, 4133-4137) y la proteína de recubrimiento menor III del fago filamentososo f_d . La cadena ligera del anticuerpo se clona como un fragmento $V_L C_L$ completo, para la secreción dirigida y se ensambla con $V_H C_{H1}$ en la partícula fágica.

En una realización, el vector comprende un casete de expresión con un casete de expresión bicistrónico o doble para permitir la expresión ligada (para el bicistrónico) o independiente (para el doble cistrónico) de la cadena ligera y pesada del anticuerpo o sus fusiones, de tal modo que el casete de expresión consiste en los siguientes elementos: (1) un promotor ajustado para expresión no inducible e inducible (por ejemplo *lacZ*); (2) un sitio de unión a ribosomas y una secuencia similar que predice las regiones de clonación de la cadena ligera y pesada; (3) posiblemente, pero no necesariamente, una región que sigue a la región de clonación de la cadena pesada o ligera que codifica una secuencia marcadora tal como un tramo de 5-6 histidinas o una secuencia reconocida por un anticuerpo y un codón ámbar; (4) una proteína de recubrimiento fágico codificada como una fusión del extremo 3' bien de la cadena pesada o ligera.

Esta nueva biblioteca fágica será una fuente valiosa de anticuerpos esencialmente para cualquier diana. Los anticuerpos pueden usarse como reactivos de investigación o como un punto inicial para el desarrollo de anticuerpos terapéuticos o productos agrícolas. Así como la lista de genomas secuenciados y productos génicos relacionados con enfermedades que se expanden rápidamente, existirá una necesidad creciente de un método *in vitro* y eventualmente automatizado para el aislamiento de anticuerpos. Los anticuerpos han sido y serán sondas ideales para investigar la naturaleza, localización y purificación de nuevos productos génicos, está previsto que esta biblioteca desempeñe un papel importante en la validación de dianas y el descubrimiento de dianas en el área de la genómica funcional.

Las variantes de proteínas expresadas sobre la superficie del bacteriófago se seleccionan basándose en su afinidad con respecto al ligando (antígeno) usando cromatografía, cribado o absorción a las células. La elución a partir de matrices de afinidad se ha alcanzado mediante la elución específica usando el ligando (antígeno o un compuesto relacionado) o elución específica usando, por ejemplo, trietilamina 100 mM. Los procedimientos de lavado retiran los fagos unidos no específicamente. El fago se une y se eluye de la matriz de acuerdo con la afinidad o la naturaleza de la interacción de unión. Los fagos específicamente unidos se usan después para infectar células *E. coli* macho que expresan el pilus F, permitiendo la recuperación del fago contiene el ADN que codifica proteínas con las características de unión deseada.

La selección puede realizarse sin basarse únicamente en la especificidad, sino también basándose en la afinidad. La separación es fácilmente obtenible mediante cromatografía de afinidad entre el fago que expresa un anticuerpo con una constante de disociación de 10^{-8} M y con una constante de disociación de 10^{-5} M. Clackson, T. *et al.* (1991). *Nature* 352: 624-628. El aislamiento de este último anticuerpo a partir de un repertorio inmunitario demuestra que los anticuerpos con afinidades características de la respuesta inmunitaria primaria pueden aislarse usando la tecnología fágica.

Los anticuerpos dirigidos contra antígenos de superficie celular también pueden aislarse por la adsorción selectiva de un fago sobre la superficie de las células. De un modo similar, puede ser posible la incorporación de una selección negativa a las células para eliminar reactividades cruzadas no deseadas con los marcadores de superficie celular. Así como estas son más bien difíciles y aún son métodos poco entendidos, los métodos basados en la selección de antígenos purificados deberían usarse siempre que fuera posible.

Cualquier selección de aglutinantes dentro de una población tenderá automáticamente a seleccionar las variantes de alta afinidad a expensas de la más baja, enriqueciendo la población de alta afinidad. Esto se ha usado para ejercer un buen efecto en el reciente aislamiento de anticuerpos humanos de alta afinidad a partir de un repertorio virgen. Marks, J.D. *et al.* (1991) *J. Mol. Biol.* **222**, 581-597. Para la selección óptima, la concentración de antígeno debería ser menor que la constante de afinidad. Esto debería tenerse en cuenta cuando se aísla un anticuerpo con características predefinidas. En las revisiones de este manual se ofrecen detalles adicionales de diversos métodos de selección.

Con dichos paneles grandes de anticuerpos aislados, es útil disponer de métodos para determinar fácilmente los parámetros cinéticos de cada interacción antígeno-anticuerpo individual. Se ha mostrado que es factible determinar rápidamente y con exactitud la velocidad de disociación de los anticuerpos no purificados en fracciones periplasmáticas preparadas de cultivos a pequeña escala usando resonancia de plasmón superficial. Usando este método, una serie de Fab específicos del toxoide tetánico mostraron una disociación monofásica, la cual se esperaba para una unión de un fragmento Fab verdaderamente monomérico a una baja densidad de superficie antigénica. Usando este ensayo de la exploración de la velocidad de disociación, se determinó que las velocidades de disociación para los mejores los Fab específicos del toxoide tetánico y de MUC1 estaban en el orden de 10^{-2} a 10^{-4} s⁻¹.

Se ensayó la integridad de los Fab seleccionados obtenidos de las fracciones periplasmáticas usando transferencias de Western. Cuando se incubaron en un tampón de muestra no reductor, se detectaron dos productos con el anticuerpo 9E10, que reconoce el marcador *myc* en el extremo del dominio CH1, el producto principal es la molécula de Fab intacta, en la cual un puente disulfuro intermolecular une covalentemente los fragmentos de la cadena pesada y la ligera; el producto de bajo peso molecular procede más probablemente de las cadenas pesadas unidas por un puente no disulfuro. El análisis con sueros anti-cadena ligera revela un patrón similar y muestra que los clones usan un porcentaje prácticamente igual de cadenas kappa y lambda (encontrado en seis y siete clones respectivamente de un total de 13 ensayados). Tras la reducción de los Fab purificados, funcionales para la unión antígeno-anticuerpo, se observaron las mismas cantidades de la cadena ligera y la pesada, mientras que en condiciones no reductoras, el producto principal estaba representado por la molécula de Fab unida mediante disulfuro, con la misma cantidad de productos V_HCH₁ y V_LCL visibles unidos no covalentemente. La producción dio lugar a Fab específicos de hormonas seleccionadas que variaban entre 160 µg y 1,43 mg de Fab por litro de cultivo, que estaba en el mismo intervalo que se encontró para los Fab no seleccionados.

Se secuenció por completo un panel de 14 Fab específicos de antígeno (3 anticuerpos anti-MUC1; 11 anticuerpos anti-gonadotropina). Los genes de cadena pesada procedían de las cuatro familias de V_H (V_{H1}, V_{H3}, V_{H4} y V_{H6}) más grandes; los genes V_L pertenecían a una de las cuatro familias de V_K o a una de las cuatro familias de V_λ. La promiscuidad de las cadenas se observó para el clon SC#4G específico de la cadena α, los clones específicos de α/β-LH, LH#2H y LH#3G y el clon específico de β-FSH, FS#8B, que todos usaron un segmento génico de cadena ligera V_{K2} altamente homólogo, se combinaron con distintos fragmentos de cadena pesada. Los 3 anticuerpos anti-MUC1 usaron genes de cadena pesada y ligera procedentes de 2 familias V_H y V₆ distintas; el clon MUC#9 usó una V_H con un cruce de 2 segmentos.

La presente invención se ilustra adicionalmente en los siguientes ejemplos:

Como una fuente de tejidos linfoides, se usaron linfocitos de sangre periférica de 4 donantes sanos y parte de un bazo libre de tumores extraído de un paciente con carcinoma gástrico. Se aislaron los linfocitos B 2 I de sangre en un gradiente Ficoll-Paque. Para el aislamiento del ARN, el sedimento celular se disolvió inmediatamente en 50 ml de tiocianato de guanidinio 8 M/2-mercaptoetanol 0,1 M (Chirgwin, et al., (1979) *Biochemistry* 18, 5294-5299). El ADN cromosómico se sometió a cizalladura hasta su terminación pasándolo a través de una jeringa estrecha (de calibre 1,2/0,5 mm) y se retiraron los residuos insolubles mediante centrifugación a baja velocidad (15 min, 2.934 x g a temperatura ambiente). El ARN se sedimentó por centrifugación a través de un gradiente de bloqueo con CsCl (12 ml de sobrenadante en una capa de 3,5 ml de CsCl 5,7 M / EDTA 0,1 M; en total 4 tubos) durante 20 h a 125.000 x g a 20 °C en un rotor SW41 (Beckman). El rendimiento de ARN total fue de aprox. 600 µg. El ARN se conservó a -20 °C en etanol. A partir del bazo, se usaron 2 g de tejido para la homogenización con un politrón en 20 ml de tiocianato de guanidinio 8 M/2-mercaptoetanol 0,1 M. El volumen total aumentó hasta 80 ml con tampón de tiocianato de guanidinio y después del paso a través de una jeringuilla estrecha para la cizalla y la eliminación de residuos, se sedimentó el ARN tal como se describe anteriormente, excepto que fue durante 15 h a 85.000 x g a 20 °C en un rotor SW28.1 (12 ml de sobrenadante en 3,5 ml de CsCl 5,7 M/ EDTA 0,1 M en 5 tubos SW28.1). A partir de 2 g de tejido, se extrajeron 3 mg de ARN total.

Se preparó ADNc cebado aleatoriamente con 250 µg de ARN de PBL, mientras que se usaron 300 µg de ARN de bazo como molde en una reacción por separado. Se desnaturalizó el ARN con calor durante 5 min a 65 °C en presencia de 20 µg de cebador aleatorio (Promega), posteriormente se añadieron tampón y DTT de acuerdo con las instrucciones del proveedor (Gibco-BRL), así como dNTP 250 µM (Pharmacia), 800 U de ARNsin (40 U/µl; Promega) y 2.000 U de MMLV-RT (200 U/µl; Gibco-BRL) en un volumen total de 500 µl. Después de 2 h a 42 °C, se detuvo la incubación mediante una extracción en fenol / cloroformo; se precipitó el ADNc y se disolvió en 85 µl de agua.

Los oligonucleótidos usados para la amplificación por PCR de las regiones V de la cadena pesada y ligera humanas se describen en la Figura 2. Las regiones variables de cadena pesada derivadas de IgM se obtuvieron mediante una PCR primaria con un cebador de la región constante de IgM. Todas las PCR primarias se llevaron a cabo con cebadores BACK distintos y cebadores FOR combinados, para mantener la diversidad máxima. Los productos de PCR se reamplifican con una combinación de cebadores JHFOR, que hibridaban con el extremo 3' de V_H y cebadores VHBACK marcados con Sfi, que hibridaban con el extremo 5' y se clonaron posteriormente como fragmentos V_H. Los genes V de cadena ligera de las familias kappa y lambda se obtuvieron mediante PCR con un conjunto de cebadores CκFOR- o CλFOR- que hibridaban con el extremo 3' del dominio constante y cebadores BACK, que cebaban el extremo 5' de las regiones V. Los segmentos de ADN se reamplificaron con cebadores marcados con sitios de restricción y se clonaron como fragmentos V_κC_κ- y V_λC_λ-.

Se realizó la PCR en un volumen de 50 µl usando la polimerasa AmpliTaq (Cetus) y 500 pM de cada cebador durante 28 ciclos (1 min a 94 °C, 1 min a 55 °C y 2 min a 72 °C), se generaron 9 amplificaciones de VH derivados de IgM por separado con 2 µl de ADNc cebado aleatoriamente (equivalente a 6 µg de ARN de PBL o a 7 µg de ARN de bazo) como un molde para cada reacción. Para las familias de cadena ligera, se obtuvieron 6 productos de V_κC_κ distintos y productos de 11 V_λC_λ (se combinaron cebadores C_{λ2}- y C_{λ7}- en cada reacción). Se purificaron todos los productos a partir de un gel de agarosa con el kit de extracción QIAex-II (Qiagen). Como una entrada para la reamplificación para introducir sitios de restricción, se usaron 100-200 ng del fragmento de ADN purificado como un molde en un volumen de reacción de 100 µl. La gran cantidad de entrada, que asegura el mantenimiento de la variabilidad, se ensayó mediante análisis de 4 λl de la mezcla de PCR "sin amplificar" en un gel de agarosa.

Para la construcción de los repertorios de cadena pesada primaria y de dos cadenas ligeras primarias, los productos de PCR, con sitios de restricción anexos, se purificaron en gel antes de la digestión y se combinaron las distintas familias V_H-, V_κ- y V_λ- en tres grupos. Los fragmentos V_κC_κ- y V_λC_λ- se digirieron con *Apa*LI y *As*cl, y se clonaron en el vector fagémido pCES1. Se digirieron 1,5 µg en total de los fragmentos V_H con *S*fiI y *B*stEII y se ligaron en 100-200 µl de mezcla de reacción con 9 U de ADN ligasa de T₄ a temperatura ambiente para 4 µg del vector purificado en gel pUC119-CES1 (similar al vector pCES1 pero con el gen pIII delecionado). La mezcla de ligamiento desalinizada para los conjuntos de la cadena pesada o ligera se usó para la electroporación de la cepa TG1 de *E. coli*, para crear las bibliotecas monocatenarias.

La biblioteca de Fab se obtuvo clonando los fragmentos V_H, digeridos a partir de ADN plasmídico preparado a partir de los repertorios de cadena pesada, en la colección de plásmidos que contenía los repertorios de cadena ligera. El ADN plasmídico aislado a partir de al menos 3 x 10⁹ bacterias de la biblioteca V_H se digirió con *S*fiI y *B*stEII para la clonación en el vector que ya contenía bibliotecas de la cadena ligera λ y κ. Para retener los clones con un sitio interno *B*stEII en el V_λ- (este sitio es relativamente frecuente en algunos segmentos V de la línea germinal λ (Persic, *et al.*, (1997) *Gene* **187**,9-18) y también en el dominio constante de una de las familias λ), la clonación de V_HC_{H1} en el vector que contenía repertorio de cadena ligera λ también se llevó a cabo usando sitios de clonación de *S*fiI y *Not*I, para crear una biblioteca V_λ- menos sesgada por la restricción.

El rescate de partículas de fagémido con el fago auxiliar M13-KO7 se realizó de acuerdo con (Marks, *et al.*, (1991) *J. Mol. Biol.* **222**,581-597) a una escala de 10 l, usando números representativos de bacterias a partir de la biblioteca para la inoculación, para asegurar la presencia de al menos 10 bacterias de cada clon en el inoculo inicial. Para las selecciones, se usaron 10¹³ ufc (unidades formadoras de colonias) con antígenos inmovilizados en inmunotubos (inmunotubos Maxisorp, Nunc) (Marks, *et al.*, (1991) *J. Mol. Biol.* **222**, 581-597) o con antígenos solubles biotinilados (Hawkins, *et al.*, (1992b) *J. Mol. Biol.* **226**, 889-896). La cantidad de antígenos del toxoide tetánico y el hapteno fenil-oxazolona inmovilizados (conjugado a BSA en una proporción de 17 a 1) se redujo 10 veces durante las rondas de selección posteriores, comenzando a 100 µg/ml en la ronda 1. Se usó la captura con antígeno biotinilado en solución para un péptido 100-mero que codificaba cinco copias de la repetición en tándem de MUC1 (Henderikx, *et al.*, (1998) *Cancer Res.* **58**, 4324-4332) o con Gonadotropina Coriónica humana (hCG), Hormona Luteneizante humana (hLH), Hormona Estimulante del Folículo humana (hFSH) y su derivado quimérico (hFSH-CTP, que contiene el péptido carboxi-terminal de la subunidad β de la hCG fusionado con la subunidad β de hFSH). Se biotinilaron los antígenos a una proporción de diez a veinte moléculas de NHS-Biotina (Pierce) por antígeno molecular de acuerdo con las recomendaciones del proveedor. A menos que se indique lo contrario, se usaron los antígenos para la selección a concentraciones de 100 nM, 30 nM y 10 nM durante 1, 2 y 3 rondas respectivamente. Para hFSH-CTP se usaron 50, 15 y 10 nM respectivamente; para el péptido MUC1, se usaron 500, 100, 20 y 5 nM.

Se produjo Fab soluble a partir de clones individuales tal como se describe anteriormente (Marks, *et al.*, (1991) *J. Mol. Biol.* **222**,581-597). Se ensayaron los sobrenadantes de cultivo en un ELISA con un antígeno recubierto

- directamente o un antígeno indirectamente capturado biotinilado por medio de BSA-estreptavidina biotinilada inmovilizada. El toxoide tetánico y phOx-BSA se recubrieron a 10 µg/ml en NaHCO₃ 0,1 M pH 9,6 durante 16 h a 4°C. Para el recubrimiento de hCG y hFSH-CTP se usó una concentración de 4 µg/ml en NaHCO₃ 50 mM pH 9,6. Para la captura de los antígenos biotinilados, la BSA biotinilada se recubrió a 2 µg/ml en PBS durante 1 h a 37 °C.
- 5 Después de 3 lavados con PBS - Tween 20 al 0,1 % (v/v) (PBST), se incubaron las placas durante 1 h con estreptavidina (10 µg/ml en PBS / gelatina al 0,5 %) (Henderikx, *et al.*, (1998) *Cancer Res.* 58, 4324-4332). Después de lavar tal como se hizo anteriormente, se añadió un antígeno biotinilado durante un tiempo de incubación de una noche a 4 °C a una concentración de 0,5 µg/ml para el péptido MUC-1, 3 µg/ml para hLH, y 0,6 µg/ml para hFSH (se ensayó la unión a hCG con el antígeno usado directamente para el recubrimiento). Se bloquearon las placas durante
- 10 30 min a temperatura ambiente con 2 % de leche en polvo semidesnatada (p/v) (Marvel) en PBS. El sobrenadante de cultivo se diluyó 1 o 5 veces en Marvel al 2 % (p/v) / PBS y se incubó durante 2 h; el Fab unido se detectó con un anticuerpo anti-*myc* 9E10 (5 µg/ml) que reconocía el marcador peptídico *myc* en el extremo carboxi-terminal de la cadena pesada de Fd y el conjugado HRP anti-ratón de conejo (DAKO) (Marks, *et al.*, (1991) *J. Mol. Biol.* 222, 581-597). Después de la última incubación, se realizó una tinción con tetrametilbencidina (TMB) y H₂O₂ como sustrato y se detuvo añadiendo medio volumen de H₂SO₄ 2 N; se midió la densidad óptica a 450 nm. Los clones que producían una señal positiva en el ELISA (por encima de 2x el fondo), se analizaron mediante identificación de huella de *Bst*NI de los productos de PCR obtenidos mediante la amplificación con los cebadores oligonucleotídicos M13 inverso y genIII-directo (Marks, *et al.*, (1991) *J. Mol. Biol.* 222, 581-597).
- 20 La inducción a gran escala de los fragmentos solubles de Fab a partir de clones individuales se realizó a una escala de 50 ml en TY 2x que contenía 100 µg/ml de ampicilina y glucosa al 2 %. Después del crecimiento a 37 °C a una DO₆₀₀ de 0,9, se sedimentaron las células (10 min a 2.934 x g) y se resuspendieron en TY 2x con ampicilina e IPTG 1 mM. Se recogieron las bacterias después de 3,5 h que crecieron a 30 °C mediante centrifugación (como se hizo anteriormente); se prepararon las fracciones periplasmáticas resuspendiendo el sedimento celular en 1 ml de PBS enfriado en hielo. Después de someterlos a rotación sobre sí mismos de 2 a 16 h a 4 °C, se retiraron los esferoplastos mediante dos etapas de centrifugación: después de la centrifugación durante 10 min a 3.400 x g, se aclaró el sobrenadante mediante una etapa de centrifugación adicional durante 10 min a 13.000 x g en una centrífuga eppendorf. La fracción periplásmica obtenida se usó directamente para la determinación de las especificidades finas mediante resonancia de plasmones de superficie para estudios de transferencia de western.
- 25 30 Para la secuenciación, se preparó el plásmido de ADN a partir de cultivos de 50 ml que crecieron a 30 °C en medio LB, que contenía 100 µg/ml de ampicilina y glucosa al 2 %, usando el midi-kit de QIAGEN (Qiagen).
- 35 La secuenciación se realizó con el kit de termociclado (Amersham) con cebadores marcados con CY5 CH1 FOR (5'-GTC CTT GAC CAG GCA GCC CAG GGC-3' - SEC ID N°: 3) y M13REV (5'-CAG GAA ACA GCT ATG AC-3" - SEC ID N°: 4); las muestras se procesaron en un ALF-Express (Pharmacia). Se alinearon las secuencias de los genes V con la base V (Tomlinson *et al.*, BASE V, Centro para Modificación de Proteínas MRC, 1997, <http://www.mrc-cpe.cam.ac.uk/imt-doc/public/INTRO.html>) o la del centro Sanger (Búsqueda de Línea Germinal del Centro Sanger, 1997, <http://www.sanger.ac.uk/Data/Search/gq-search.html>).
- 40 45 Se usó una preparación de hCG purificada de orina y la hLH recombinante purificada mediante inmunoafinidad, la hFSH y hFSH-CTP producidas en células CHO (Matzuk, *et al.*, (1989) *J. Cell. Biol.* 109,1429-1438; Muyan, *et al.*, (1996) *Mol. Endocrinol.* 10, 1678-1687) para los estudios de transferencia de western tal como se describió (Moyle, *et al.*, (1990) *J. Biol. Chem.* 265, 8511-8518). Se cargaron entre 0,5 y 1 µg de cada hormona por línea; se diluyeron las proteínas en un tampón de muestra no reductor y se hirvieron durante 5 min o se aplicaron directamente en el gel sin tratamiento con calor; se transfirieron las proteínas a la membrana de transferencia mediante electrotransferencia. Se incubaron las inmunotransferencias posteriormente durante 16 h a temperatura ambiente con una fracción periplasmática diluida 10 veces en PBS/Marvel al 4 %. El Fab unido se detectó con anticuerpo anti-*myc* 9E10 (5 µg/ml) y conjugado de fosfatasa alcalina anti-ratón diluido 4.000 veces (Promega), usando los sustratos 5-bromo-1-cloro-3-indolil fosfato (BCIP) y azul de nitro-tetrazolio (NBT) (Boehringer Mannheim) para su visualización.
- 50 55 La especificidad de los Fab se caracterizó adicionalmente mediante resonancia de plasmón superficial (BIAcore 2000, Biacore). La hLh, hFSH recombinantes y la hCG de la orina se inmovilizaron sobre las celdas de flujo de una microplaca CM usando el kit NHS/EDC (Pharmacia), produciendo una superficie de 1906 UR para hLH, 1529 UR para hFSH y 1375 UR para hCG. Se diluyeron las fracciones periplasmáticas tres veces en Solución Salina de Hepes Tamponada (HBS; Hepes 10 mM, EDTA 3,4 mM, NaCl 150 mM, tensioactivo P20 al 0,05 % (v/v), pH 7,4) y se analizaron usando una velocidad de flujo de 10 µl/min.
- 60 65 Los Fab se obtuvieron replegando las proteínas bacterianas totales de un cultivo de 50 ml (de Haard, *et al.*, (1998) *Protein Eng.*, 11:1267-1276). Brevemente, las células sedimentadas de un cultivo bacteriano de 50 ml inducido se resuspendieron en 8 ml de urea 8 M (en PBS). Después de someter a ultrasonidos, la mezcla se sometió a rotación sobre sí misma durante 30 min y el material insoluble se retiró por centrifugación durante 30 min a 13.000 x g. Se dializó el sobrenadante contra PBS con cuatro cambios de tampón. Se retiraron las proteínas insolubles por centrifugación y la fracción de flujo continuo, obtenida mediante filtración a través de una membrana de 0,2 µm, se cargó inmediatamente en una columna de hCG (volumen del lecho de 0,3 ml). Se preparó el material de la columna

acoplado 8,4 mg de proteína a un gramo de Tresil sefarosa de acuerdo con las instrucciones del proveedor (Pierce). La columna (1 ml de material de columna) se lavó con 10 volúmenes de Tris 100 mM, NaCl 500 mM pH 7,5, posteriormente con 10 volúmenes de Tris 100 mM / NaCl 500 mM pH 9,5 y con 2 volúmenes de NaCl al 0,9 %, el Fab unido se eluyó con dos volúmenes de TEA 0,1 M y se neutralizaron inmediatamente con 0,5 volúmenes de Tris 1 M pH 7,5. La fracción Fab se dializó contra PBS usando un filtro de diálisis de centrifugación Microcon 30 (Amicon). Finalmente, se llevó a cabo un análisis de filtración en gel en una columna Superdex 75HR (Pharmacia). Se determinó el rendimiento midiendo la densidad óptica a 280 nm (usando un coeficiente de extinción molar de 13 para los Fab).

- 10 Las cinéticas de unión se analizaron mediante resonancia de plasmones de superficie en tres superficies de hCG distintas (303 UR, 615 UR y 767 UR inmovilizadas, con 4955 UR de BSA en una celda de flujo separada como un control negativo). El Fab presente en los extractos periplasmáticos sin tratar se cuantificó en una superficie de alta densidad de anticuerpo policlonal de Fab anti-humano purificado (Pierce) tal como se describe (Kazemier, *et al.*, (1996) *J. Immunol. Methods* **194**, 201-209). Los controles de Fab anti-hCG se purificaron mediante cromatografía de afinidad en columnas de hCG tal como se describe anteriormente y se usaron para calibrar el sistema.

La biblioteca de Fab se construyó en dos etapas. En la primera etapa, se amplificaron los conjuntos génicos de la región variable a partir de aprox. 4×10^6 células B a partir de los PBL de cuatro donantes sanos y, como una fuente de anticuerpos de IgM posiblemente más fuertemente mutados, a partir de un segmento de un bazo (libre de tumores) retirado de un paciente con carcinoma gástrico, que contenía aproximadamente $1,5 \times 10^8$ células B (Roit, *et al.*, (1985) *Immunology*, Gower Medical Publishing, Ltd., Londres). Únicamente los segmentos V_H derivados de IgM se amplificaron usando una amplificación con un cebador oligonucleotídico localizado en el primer dominio constante de este isotipo. Se clonaron estos productos en un vector fagémido pCES1 para V_L y en pUC119-CES1 para V_H (la clonación fue más eficaz en el vector de tamaño más pequeño, en el que se eliminó el gen III). Las bibliotecas de V_H , V_K y V_λ derivado de los PBL y del bazo se clonaron por separado para mantener la diversidad, para producir bibliotecas monocatenarias en un tamaño típico para las bibliotecas producidas mediante clonación de fragmentos de PCR (Marks, *et al.*, (1991) *J. Mol. Biol.* **222**, 581-597): $1,75 \times 10^8$ clones individuales para la cadena pesada, $9,4 \times 10^7$ clones para V_K y $5,2 \times 10^7$ clones para V_λ . En la segunda etapa, los fragmentos de la cadena pesada se digirieron a partir del ADN plasmídico aislado a partir del repertorio primario de V_H y se clonaron en el vector que contenía los repertorios de la cadena ligera (nuevamente por separado para los repertorios de los PBL y del bazo). Se combinaron las bibliotecas usando este procedimiento de clonación eficaz, para crear un repertorio de Fab vírgenes con $3,7 \times 10^{10}$ clones individuales ($4,3 \times 10^{10}$ clones recombinantes, 86 % de los cuales tienen un inserto de Fab de longitud completa), alojando el 70 % de los clones una cadena ligera kappa, el 30 % una cadena lambda. Los 20 clones con un Fab de longitud completa ensayados tuvieron una puntuación positiva en el análisis de transferencia de puntos con el anticuerpo 9E10 para indicar un nivel de expresión de Fab soluble de al menos 0,2 mg/l.

Se evaluó la biblioteca mediante la selección con distintos antígenos. Primero, se analizaron los resultados de tres antígenos modélicos, la proteína del toxoide tetánico, el hapteno 2-fenoxiloxazol-5-ona (phOx) (Griffiths, *et al.*, (1984) *Nature* 312,271-275 y el péptido MUC1. Tres rondas de bioselección en el toxoide tetánico produjeron un conjunto diverso de Fab positivos en el ELISA, en una serie de 47 Fab de unión al toxoide tetánico, al menos 21 fueron distintos con respecto a la identificación de huella *Bst*I/II. De manera similar, se recuperó un panel extenso de Fab específicos de phOx después de tres rondas de cribado: se identificaron al menos 24 clones distintos en una serie de 50 clones positivos en el ELISA. La solución de captura con el péptido MUC1 biotinilado dio como resultado la selección de 14 fragmentos de anticuerpos distintos de 37 clones positivos en el ELISA seleccionados después de 3 rondas.

Como un panel de ensayo más riguroso de antígenos para ensayar el rendimiento de la biblioteca, se eligió derivar anticuerpos para tres glucoproteínas estructuralmente relacionadas: Gonadotropina Coriónica humana (hCG), Hormona Luteinizante humana (hLH) y Hormona Estimulante del Folículo (hFSH) (revisado en (Cole, (1997) *Clin. Chem.* **43**, 2233-2243)). Estas hormonas son heterodímeros que comparten una cadena α idéntica con 92 restos de aminoácidos, pero tienen subunidades β de distinta composición y longitud. La cadena β de la hCG contiene 145 restos de aminoácidos, y la de la hLH únicamente 121 restos, mostrando esta última el 85 % de homología con respecto a la hCG- β . La cadena β de hFSH es únicamente de 111 aminoácidos y comparte el 36 % de los restos con hCG. Los anticuerpos que detectan específicamente hCG se han usado exhaustivamente en las pruebas de embarazo (Cole, (1997) *Clin. Chem.* **43**, 2233-2243) y para el diagnóstico de cáncer (Masure, *et al.*, (1981) *J. Clin. Endocrinol. Metab.* **53**, 1014-1020; Papapetrou, *et al.*, (1980) *Cancer* **45**, 2583-2592). Un gran conjunto de anticuerpos para estas dianas extendería el número limitado de anticuerpos específicos de hormonas (especialmente contra hLH), obtenidos usando la tecnología de hibridoma (Cole, (1997) *Clin. Chem.* **43**, 2233-2243). El origen humano de los anticuerpos podría ser beneficioso cuando se usan estos para la imagen o la terapia del cáncer testicular y el cáncer de vejiga (Masure, *et al.*, (1981) *J. Clin. Endocrinol. Metab.* **53**, 1014-1020; Papapetrou, *et al.*, (1980) *Cancer* **45**, 2583-2592).

Por lo tanto, se realizaron las selecciones en hCG urinaria biotinilada, hLh, hFSH y hFSH-CTP recombinantes (esta última es una molécula quimérica que contiene el péptido carboxi-terminal de hCG- β fusionado con la cadena β de FSH (Fares, *et al.*, (1992) *Proc. Natl. Acad. Sci. EE.UU.* **89**, 4304-4308)). Se descubrió el grado más alto de

enriquecimiento con respecto al aumento en el número de partículas fágicas eluidas en la ronda 3 frente a la ronda 1 para hCG (10.000 veces), seguido de hFSH-CTP (1.000 veces), hFSH (300 veces) y hLH (150 veces). Los fagos policlonales de poblaciones seleccionadas se ensayaron con respecto a su unión usando chips sensores que contenían hormonas inmovilizadas (Schier, *et al.*, (1996) *Hum. Antibodies Hybridomas* 7, 97-105). El fago policlonal seleccionado con hCG mostró la unión después de las rondas dos y tres de selección para las tres proteínas, es decir, hCG, hLH y hFSH, siendo la señal visible más fuerte para hCG. Un análisis similar de las poblaciones fágicas policlonales seleccionadas para las tres rondas en hFSH mostró una dominancia de la unión específica de hFSH, mientras que se produjeron las selecciones en hFSH-CTP tanto para hFSH como para hCG. Las selecciones en hLH produjeron anticuerpos reactivos con hFSH y hCG. Por lo tanto, esta exploración de fagos policlonales proporciona un ensayo rápido para comprobar la calidad total de los clones del repertorio seleccionado y también puede usarse para guiar la elección de las condiciones para la siguiente ronda de selección (Schier, *et al.*, (1996) *Hum. Antibodies Hybridomas* 7, 97-105).

Los ELISA de los anticuerpos fágicos monoclonales revelaron que las tres rondas de selección con hCG, en efecto, dieron como resultado el aislamiento de un alto porcentaje (74 %) de los clones positivos para la gonadotropina. El 27 % de estos clones tuvieron reactividad cruzada para hLH; ninguno de ellos fue reactivo contra estreptavidina. El análisis de la huella BsfNI de los clones positivos para ELISA reveló un alto grado de diversidad (8 patrones distintos). A partir de un clon de hCG específico representativo (CG#4F codificado) y un clon con reactividad cruzada para hLH (CG#5C), se ensayó la especificidad en un BIAcore usando fragmentos de Fab solubles sin purificar. El clon CG#4F produjo una alta respuesta en hCG, sin unión visible bien para hLH o hFSH-CTP. Por el contrario, el clon CG#5C se unió a hCG y hLH, pero no a hFSH-CTP. Las transferencias de Western, con las distintas hormonas en forma no reducida, mostraron el reconocimiento específico de la subunidad β de la hCG por el clon CG#4F, mientras que el clon con reactividad cruzada CG#5C reaccionó con la subunidad β tanto de hCG como de hLH.

La selección con la hormona hLH dio como resultado el aislamiento de los clones específicos de hLH y con reactividad cruzada para hCG. El examen de clones individuales a partir de la ronda de selección tres en ELISA reveló una gran fracción de clones específicos de hLH (69 %), y un grupo menor de clones con reactividad cruzada (16 %); no se seleccionaron clones reactivos con estreptavidina. Dentro del grupo de clones específicos, podría discriminarse una matriz grande de distintas especies (>21) con un análisis de la huella; sin embargo, todas las especies con reactividad cruzada tenían un patrón único. La especificidad única para hLH se confirmó para los clones representativos LH#2H y LH#3G, mostrados en la resonancia de plasmones de superficie; y en la transferencia de western. LH#3G únicamente reconoce el heterodímero α/β de hLH. Dos clones representativos de un anticuerpo pan-reactivo en ELISA, codificados LH#1C y LH#3F, reaccionaron en un BIAcore con hFSH-CTP, hCG y hLH y en transferencia de western con las cadenas α de las tres hormonas.

Cuando hFSH se usó como antígeno durante la selección, se aislaron 6 anticuerpos distintos a partir de la biblioteca, con un tipo, representado por el clon FS#8B, dominando la población seleccionada. Este Fab únicamente reconoció a hFSH en el BIAcore y, tal como demostró el análisis por transferencia de Western, en particular su subunidad β . Además, la especificidad de un clon de unión a una cadena α , SC#2B, se confirmó en BIAcore y transferencia de western.

Tras la selección con FSH-CTP, se identificaron 7 Fab específicos para la cadena α distintos mediante análisis de la huella, a partir de lo cual los clones codificados SC#2B, SC#2F, SC#2G y SC#4G se examinaron con más detalle. Los análisis de inmunotransferencia con el Fab recombinante como un anticuerpo de detección confirmaron la especificidad de la cadena α .

Se determinaron las afinidades y las velocidades de disociación de los Fab reactivos con hCG purificados por afinidad LH#1C, SC#2B, LH#3F y CG#5C. Las velocidades de disociación de la mayor parte de los Fab fueron del orden de 10^{-2} y 10^{-3} s⁻¹. Los valores de las velocidades de disociación obtenidas usando fracciones periplasmáticas sin tratar tenían una buena concordancia con los valores descubiertos para los Fab purificados, validando la utilidad de la exploración de la velocidad de disociación con los fragmentos de Fab sin purificar. Las afinidades, 23 nM u 38 nM para el anticuerpo específico de la subunidad α LH#1 C y el anticuerpo con reactividad cruzada para la subunidad β - de hCG/hLH CG#5C respectivamente, fueron comparables con la afinidad de los anticuerpos seleccionados a partir de una biblioteca de anticuerpos fágicos inmunitarios murinos (H.d.H., B. Kazemier, *et al.*, no publicado); la afinidad superior, 2,7 nM para el Fab específico de la cadena α SC#2B, se aproximó a los valores de los mejores anticuerpos monoclonales anti-hCG (H.d.H., B. Kazemier, *et al.*, no publicado).

El objetivo de este procedimiento era seleccionar y enriquecer en anticuerpos fágicos a un antígeno con el que se recubre la superficie de los inmunotubos. El antígeno se usó para recubrir el inmunotubo (por ejemplo, un inmunotubo Nunc) y se incubó con la fagoteca. Se lavaron los fagos no unidos y se eluyeron los fagos unidos, por lo tanto, la fagoteca se enriqueció en anticuerpos fágicos que se unían específicamente al antígeno.

El objetivo de este procedimiento era biotinylar proteínas o péptidos. A un pH neutro o superior, los grupos amino primarios reaccionaron con la NHS-SS-Biotina y se liberó N-hidroxisulfosuccinimida. Los grupos NH₂- N-terminales así como las lisinas (K) de la proteína reaccionaron con la NHS-S-S-Biotina, en este intervalo de pH.

La NHS-SS-Biotina es un análogo de biotina único con un brazo espaciador extendido de aproximadamente 24,3 Å de longitud, el brazo espaciador de la NHS-LC-Biotina es de 22,4 Å. Estos análogos de cadena larga reducen los

impedimentos estéricos en la unión de moléculas biotiniladas a avidina o estreptavidina.

La presencia del enlazador S-S en la NHS-S-S-Biotina permitió la ruptura de la unión usando agentes reductores (DTT, DTE, β -mercaptoetanol). La NHS-LC-Biotina se usó cuando se necesitó que la proteína/péptido biotinilado no fuera sensible a agentes reductores.

El objetivo de este procedimiento fue seleccionar anticuerpos fágicos contra un antígeno biotinilado. La selección se hizo en solución, y pudo usarse para seleccionar anticuerpos fágicos contra antígenos que con propensos a la desnaturalización cuando recubren superficies sólidas.

Primero se incubó el antígeno biotinilado con la biblioteca de anticuerpos fágicos. Después la adición de Dynabeads (Dyna) recubiertas con estreptavidina, la biotina del complejo antígeno-anticuerpo se unió a la estreptavidina. Este complejo Dynabead-antígeno-anticuerpo se extrajo con un imán (por ejemplo, un imán Dynal) y por lo tanto debería contener los anticuerpos específicos.

El objetivo de este procedimiento fue seleccionar aquellos anticuerpos de una biblioteca que se unían a los antígenos presentes en la membrana celular, usando células adherentes en crecimiento o células en suspensión. El método pudo usarse para la selección de anticuerpos contra dianas expresadas en líneas celulares (tumores).

Incubando células completas, orgánulos o fracciones de membranas con un repertorio de anticuerpos fágicos de alta variedad, tales como las bibliotecas de Fab obtenidas de acuerdo con la invención (concentradas mediante precipitación con PEG), únicamente (o preferentemente) anticuerpos relevantes, para una de las moléculas expuestas en la superficie de la(s) membrana(s) celular(es), se retendrá mientras que se separan los anticuerpos fágicos no unidos se separan de los anticuerpos unidos a las células, orgánulos o fracciones de membrana (mediante métodos bien conocidos en la técnica para la separación de células, orgánulos o fracciones de membrana a partir de moléculas en solución). La población de fagos retenida se enriqueció para aquellos clones que sean específicos para las moléculas relacionadas con células. En principio los siguientes factores influirán positivamente en el enriquecimiento de clones individuales: la afinidad, abundancia antigénica y la baja toxicidad de la construcción de anticuerpos para un hospedador TG1.

El objetivo de este procedimiento fue preparar fragmentos de anticuerpos solubles a partir del periplasma de *E. coli*. En el periplasma existen: una actividad proteasa menor, menos proteínas contaminantes que las del citoplasma o los sobrenadantes y el anticuerpo está más concentrado. Por lo tanto, las preparaciones periplásmicas son más estables y más puras que los sobrenadantes de cultivo.

Como una consecuencia de la inducción de un fagémido contenido en los cultivos bacterianos en un medio bajo en glucosa con IPTG, se produjeron fragmentos de anticuerpos solubles y se dirigieron al periplasma donde se concentraron dentro de las 4 horas. El cultivo durante toda la noche en estas circunstancias haría que la membrana bacteriana tuviera fugas y se encontrarán anticuerpos en el sobrenadante. Para la preparación de fracciones periplásmicas, la pared celular bacteriana se lisó primero mediante choque osmótico frío (TES enfriado en hielo) y después se diluyó rápidamente en una solución fría de fuerza osmótica baja (TES/H₂O). El EDTA hizo que la membrana externa fuera más permeable y el frío inhibió la actividad proteasa. Posteriormente, se centrifugaron las células bacterianas y entonces el sobrenadante contuvo las proteínas periplásmicas.

Los anticuerpos de la fracción periplásmica podían usarse como un "extracto sin tratar" o los anticuerpos podían purificarse mediante medios convencionales bien conocidos en la técnica, por ejemplo, los enumerados en la Sección 5.6 y 5.7.

El objetivo de este procedimiento era purificar anticuerpos marcados con un marcador His6 de fracciones periplásmicas de los Fab producidas tal como se describe en el ejemplo 6.18.

Cromatografía de afinidad con metales inmovilizados (CAMI) para la purificación de proteínas recombinantes marcadas con 6xHis en condiciones nativas: Se capturaron proteínas recombinantes marcadas con histidina en una resina que contenía metales quelados a través de la coordinación de los átomos de N libres de las histidinas con el metal (mayoritariamente Ni²⁺ o Co²⁺). Después de retirar con lavado las proteínas contaminantes y otros constituyentes celulares, la proteína marcada con his se eluyó específicamente de la resina con imidazol que compite por la unión de restos de histidina con el ion metálico.

LISTADO DE SECUENCIAS

<110> Dyax Corp.

<120> Nuevas bibliotecas de fragmentos Fab y método para su uso

<130> P24015EP01

<140>

<141> 18-05-1999

<160> 69

5 <170> PatentIn Ver. 2.1

<210> 1
 <211> 24
 <212> ADN
 10 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: oligonucleótido

15 <400> 1
 tggaagaggc acgttctttt cttt 24

<210> 2
 <211> 24
 20 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: oligonucleótido

25 <400> 2
 acactctccc ctgttgaagc tctt 24

<210> 3
 <211> 23
 30 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: oligonucleótido

35 <400> 3
 tgaacattct gtaggggcca ctg 23

<210> 4
 <211> 23
 40 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: oligonucleótido

45 <400> 4
 agagcattct gcaggggcca ctg 23

50 <210> 5
 <211> 23
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido

60 <400> 5
 cagrtgcagc tgggtcartc tgg 23

<210> 6
 <211> 23
 <212> ADN
 <213> Secuencia artificial

65 <220>

<223> Descripción de la secuencia artificial: oligonucleótido

5 <400> 6
saggtccagc tggtrcagtc tgg 23

<210> 7
<211> 23
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

15 <400> 7
cagrtcacct tgaaggagtc tgg 23

<210> 8
<211> 23
<212> ADN
<213> Secuencia artificial

20 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

25 <400> 8
saggtgcagc tggggagtc tgg 23

30 <210> 9
<211> 23
<212> ADN
<213> Secuencia artificial

35 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

40 <400> 9
gaggtgcagc tggggagwc ygg 23

<210> 10
<211> 23
<212> ADN
<213> Secuencia artificial

45 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

50 <400> 10
caggtgcagc tacagcagtc ggg 23

<210> 11
<211> 23
<212> ADN
<213> Secuencia artificial

55 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

60 <400> 11
cagstgcagc tgcaggagtc sgg 23

<210> 12
<211> 23
<212> ADN
<213> Secuencia artificial

65 <220>

<223> Descripción de la secuencia artificial: oligonucleótido

5 <400> 12
gargtgcagc tgggcagtc tgg 23

<210> 13
<211> 23
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

15 <400> 13
caggtacagc tgcagcagtc agg 23

<210> 14
<211> 23
<212> ADN
<213> Secuencia artificial

20 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

25 <400> 14
gacatccagw tgaccagtc tcc 23

30 <210> 15
<211> 23
<212> ADN
<213> Secuencia artificial

35 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

<400> 15
gatgtgtga tgactcagtc tcc 23

40 <210> 16
<211> 23
<212> ADN
<213> Secuencia artificial

45 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

<400> 16
gaaattgtw tgacrcagtc tcc 23

50 <210> 17
<211> 23
<212> ADN
<213> Secuencia artificial

55 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

60 <400> 17
gatattgtga tgaccacac tcc 23

<210> 18
<211> 23
<212> ADN
<213> Secuencia artificial

65 <220>

<223> Descripción de la secuencia artificial: oligonucleótido
 <400> 18
 5 gaaacgacac tcacgcagtc tcc 23
 <210> 19
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 10
 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido
 <400> 19
 15 gaaattgtgc tgactcagtc tcc 23
 <210> 20
 <211> 23
 <212> ADN
 20 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido
 <400> 20
 25 cagtctgtgc tgactcagcc acc 23
 <210> 21
 <211> 23
 <212> ADN
 30 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido
 <400> 21
 35
 <210> 22
 <211> 23
 <212> ADN
 40 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido
 <400> 22
 45 cagtctgtcg tgacgcagcc gcc 23
 <210> 23
 <211> 21
 <212> ADN
 50 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido
 <400> 23
 55 cartctgccc tgactcagcc t 21
 <210> 24
 <211> 23
 <212> ADN
 60 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido
 65

<400> 24
 tcctatgwgcc tgactcagcc acc 23

5

<210> 25
 <211> 23
 <212> ADN
 <213> Secuencia artificial

10

<220>
 <223> Descripción de la secuencia artificial: oligonucleótido

<400> 25
 tcttctgagc tgactcagga ccc 23

15

<210> 26
 <211> 23
 <212> ADN
 <213> Secuencia artificial

20

<220>
 <223> Descripción de la secuencia artificial: oligonucleótido

<400> 26
 cacgttatac tgactcaacc gcc 23

25

<210> 27
 <211> 23
 <212> ADN
 <213> Secuencia artificial

30

<220>
 <223> Descripción de la secuencia artificial: oligonucleótido

<400> 27
 caggctgtgc tgactcagcc gtc 23

35

<210> 28
 <211> 23
 <212> ADN
 <213> Secuencia artificial

40

<220>
 <223> Descripción de la secuencia artificial: oligonucleótido

45

<400> 28
 aattttatgc tgactcagcc cca 23

50

<210> 29
 <211> 23
 <212> ADN
 <213> Secuencia artificial

55

<220>
 <223> Descripción de la secuencia artificial: oligonucleótido

<400> 29
 cagrctgtgg tgacycagga gcc 23

60

<210> 30
 <211> 23
 <212> ADN
 <213> Secuencia artificial

65

<220>
 <223> Descripción de la secuencia artificial: oligonucleótido

ES 2 549 440 T3

	<400> 30 cwgctgtgc tgactcagcc mcc	23
5	<210> 31 <211> 51 <212> ADN <213> Secuencia artificial	
10	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
15	<400> 31 accgcctcca cgggcgcg cttattaaca ctctcccctg ttgaagctct t	51
20	<210> 32 <211> 50 <212> ADN <213> Secuencia artificial	
25	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
30	<400> 32 accgcctcca cgggcgcg cttattatga acattctgta ggggccactg	50
35	<210> 33 <211> 50 <212> ADN <213> Secuencia artificial	
40	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
45	<400> 33 accgcctcca cgggcgcg cttattaaga gcattctgca ggggccactg	50
50	<210> 34 <211> 56 <212> ADN <213> Secuencia artificial	
55	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
60	<400> 34 gtcctcgcaa ctgcgccca gccggccatg gccagrtgc agctgggca rctgg	56
65	<210> 35 <211> 56 <212> ADN <213> Secuencia artificial	
70	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
75	<400> 35 gtcctcgcaa ctgcgccca gccggccatg gccsaggtcc agctggtrca gtctgg	56
80	<210> 36 <211> 56 <212> ADN <213> Secuencia artificial	
85	<220> <223> Descripción de la secuencia artificial: oligonucleótido	

ES 2 549 440 T3

	<400> 36 gtcctcgcaa ctgcgGCCCA gccggccatg gcccagrtca cctgaagga gtctgg	56
5	<210> 37 <211> 56 <212> ADN <213> Secuencia artificial	
10	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
15	<400> 37 gtcctcgcaa ctgcgGCCCA gccggccatg gccsaggtgc agctggtgga gtctgg	56
20	<210> 38 <211> 56 <212> ADN <213> Secuencia artificial	
25	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
30	<400> 38 gtcctcgcaa ctgcgGCCCA gccggccatg gccgaggtgc agctggtgga gwcyyg	56
35	<210> 39 <211> 56 <212> ADN <213> Secuencia artificial	
40	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
45	<400> 39 gtcctcgcaa ctgcgGCCCA gccggccatg gccaggtgc agctacagca gtgggg	56
50	<210> 40 <211> 56 <212> ADN <213> Secuencia artificial	
55	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
60	<400> 40 gtcctcgcaa ctgcgGCCCA gccggccatg gccagstgc agctgcagga gtcsgg	56
65	<210> 41 <211> 56 <212> ADN <213> Secuencia artificial	
70	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
75	<400> 41 gtcctcgcaa ctgcgGCCCA gccggccatg gccgargtgc agctggtgca gtctgg	56
80	<210> 42 <211> 56 <212> ADN <213> Secuencia artificial	
85	<220> <223> Descripción de la secuencia artificial: oligonucleótido	

ES 2 549 440 T3

<400> 42
gtcctcgcaa ctgcggccca gccggccatg gcccaggtac agctgcagca gtcagg 56

5 <210> 43
<211> 24
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

<400> 43
tgaggagacg gtgaccaggg tgcc 24

15 <210> 44
<211> 24
<212> ADN
<213> Secuencia artificial

20 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

<400> 44
tgaagagacg gtgaccattg tccc 24

25 <210> 45
<211> 24
<212> ADN
<213> Secuencia artificial

30 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

<400> 45
tgaggagacg gtgaccaggg ttcc 24

35 <210> 46
<211> 24
<212> ADN
<213> Secuencia artificial

40 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

<400> 46
tgaggagacg gtgaccgtgg tccc 24

45 <210> 47
<211> 44
<212> ADN
<213> Secuencia artificial

50 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

<400> 47
accgcctcca ccagtgact tgacatccag wtgaccaggt ctcc 44

55 <210> 48
<211> 44
<212> ADN
<213> Secuencia artificial

60 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

65 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

ES 2 549 440 T3

	<400> 48 accgcctcca ccagtgact tgatgttg atgactcagt ctcc	44
5	<210> 49 <211> 44 <212> ADN <213> Secuencia artificial	
10	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
15	<400> 49 accgcctcca ccagtgact tgaatttg wtgacrcagt ctcc	44
20	<210> 50 <211> 44 <212> ADN <213> Secuencia artificial	
25	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
30	<400> 50 accgcctcca ccagtgact tgatatttg atgaccaca ctcc	44
35	<210> 51 <211> 44 <212> ADN <213> Secuencia artificial	
40	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
45	<400> 51 accgcctcca ccagtgact tgaacgaca ctcacgagt ctcc	44
50	<210> 52 <211> 44 <212> ADN <213> Secuencia artificial	
55	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
60	<400> 52 accgcctcca ccagtgact tgaatttg ctgactcagt ctcc	44
65	<210> 53 <211> 41 <212> ADN <213> Secuencia artificial	
70	<220> <223> Descripción de la secuencia artificial: oligonucleótido	
75	<400> 53 accgcctcca ccagtgaca gtctgtgctg actcagccac c	41
80	<210> 54 <211> 41 <212> ADN <213> Secuencia artificial	
85	<220> <223> Descripción de la secuencia artificial: oligonucleótido	

ES 2 549 440 T3

<400> 54
 accgcctcca ccagtgacaca gtctgtgytg acgcagccgc c 41

5 <210> 55
 <211> 41
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido

<400> 55
 accgcctcca ccagtgacaca gtctgtctgtg acgcagccgc c 41

15 <210> 56
 <211> 39
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido

<400> 56
 accgcctcca ccagtgacaca rtctgccttg actcagcct 39

25 <210> 57
 <211> 44
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido

<400> 57
 accgcctcca ccagtgacact ttctatgwg ctgactcagc cacc 44

35 <210> 58
 <211> 44
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido

<400> 58
 accgcctcca ccagtgacact ttcttctgag ctgactcagg accc 44

45 <210> 59
 <211> 41
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido

<400> 59
 accgcctcca ccagtgacaca cggtatactg actcaaccgc c 41

55 <210> 60
 <211> 41
 <212> ADN
 <213> Secuencia artificial

60 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido

65 <220>
 <223> Descripción de la secuencia artificial: oligonucleótido

ES 2 549 440 T3

<400> 60
accgcctcca ccagtcgaca ggctgtgctg actcagccgt c 41

5 <210> 61
<211> 44
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

<400> 61
accgcctcca ccagtcgact taatttatg ctgactcagc ccca 44

15 <210> 62
<211> 41
<212> ADN
<213> Secuencia artificial

20 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

<400> 62
accgcctcca ccagtcgaca grctgtggg acycaggagc c 41

25 <210> 63
<211> 41
<212> ADN
<213> Secuencia artificial

30 <220>
<223> Descripción de la secuencia artificial: oligonucleótido

<400> 63
accgcctcca ccagtcgacw gcctgtgctg actcagccmc c 41

35 <210> 64
<211> 89
<212> ADN
<213> Secuencia artificial

40 <220>
<223> Descripción de la secuencia artificial: fragmento de pCES1 que comprende una secuencia señal, sitios de restricción, la parte 5' del gen Ck humano

45 <400> 64

ttatttcgcaa ttccttttagt tgttcctttc tattctcaca gtgcacaggt ccaactgcag 60
gtcgacctcg agatcaaacg tggaaactgtg 90

50 <210> 65
<211> 163
<212> ADN
<213> Secuencia artificial

55 <220>
<223> Descripción de la secuencia artificial: fragmento de pCES1

<400> 65

ggagagtgtt aataaggcgc gccaatctta tttcaaggag acagtcataa tgaaatacct 60
attgcctacg gcagccgctg gattgttatt actcgcggcc cagccggcca tggcccaggt 120
gcagctgcag gagagcgggg tcaccgtctc aagcgcctcc acc 163

60

ES 2 549 440 T3

<210> 66
 <211> 96
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Descripción de la secuencia artificial: fragmento de pCES1 que comprende la parte 3' del gen CH1 humano, el sitio de restricción, el marcador de hexahistidina, el marcador cMyc, el codón de terminación y una parte 5' del gen III

10

<400> 66

aaatcttgtg cggccgcaca tcatcatcat catcacgggg ccgcagaaca aaaactcatc 60
tcagaagagg atctgaatgg ggccgcatag actggt 96

15

<210> 67
 <211> 30
 <212> PRT
 <213> Secuencia artificial

20

<220>
 <223> Descripción de la secuencia artificial: fragmento de pCES1

<400> 67

Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser His Ser Ala Gln
1 5 10 15
Val Gln Leu Gln Val Asp Leu Glu Ile Lys Arg Gly Thr Val
20 25 30

25

<210> 68
 <211> 38
 <212> PRT
 <213> Secuencia artificial

30

<220>
 <223> Descripción de la secuencia artificial: fragmento codificado por pCES1 que comprende la secuencia señal y la parte amino terminal del producto del gen CH1 humano

35

<400> 68

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15
Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Glu Ser Gly Val Thr
20 25 30
Val Ser Ser Ala Ser Thr
35

40

<210> 69
 <211> 29
 <212> PRT
 <213> Secuencia artificial

45

<220>
 <223> Descripción de la secuencia artificial: fragmento codificado por pCES1 que comprende la parte carboxi terminal del producto del gen CH1 humano, el marcador de hexahistidina y el marcador de cMyc

<400> 69

ES 2 549 440 T3

Lys Ser Cys Ala Ala Ala His His His His His His Gly Ala Ala Glu
1 5 10 15
Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Gly Ala Ala
20 25

REIVINDICACIONES

1. Método de producción de una biblioteca de Fab, comprendiendo la biblioteca una pluralidad de vectores en la que cada vector de la pluralidad de vectores comprende:

- 5
- una primera región de clonación y una segunda región de clonación, en la que
 - cada región de clonación comprende al menos un sitio de escisión de enzimas de restricción único para el vector,
 - estando cada región de clonación flanqueada en el extremo 5' por un sitio de unión a ribosomas y una
 - 10 secuencia señal,
 - un polinucleótido que codifica una región de anclaje, localizada en el extremo 3' de la segunda región de clonación,
 - un polinucleótido que codifica una región constante de cadena pesada o una porción de la misma localizada entre el sitio de escisión de enzimas de restricción único para el vector de la segunda región de clonación y la
 - 15 región de anclaje,
 - un polinucleótido que codifica un marcador,

comprendiendo el método:

- 20 introducir en la primera región de clonación un miembro de una primera pluralidad de polinucleótidos variables, codificando dicha pluralidad de polinucleótidos variables una primera pluralidad de polipéptidos, en la que el miembro de la primera pluralidad de polinucleótidos codifica un polipéptido que comprende una región variable de cadena ligera de un anticuerpo posiblemente seguida por una región constante de cadena ligera de un anticuerpo,
- 25 introducir por separado en la segunda región de clonación de cada vector un miembro de una segunda pluralidad de polinucleótidos variables, codificando dicha pluralidad de polinucleótidos variables una segunda pluralidad de polipéptidos, en la que el miembro de la segunda pluralidad de polinucleótidos codifica un polipéptido que comprende una región variable de cadena pesada de un anticuerpo,
- 30 e
- introducir cada uno de los vectores en una célula hospedadora para proporcionar una pluralidad de partículas de la cápside comprendiendo cada una de ellas un miembro de la primera pluralidad de polinucleótidos variables y un miembro de la segunda pluralidad de polinucleótidos, para constituir de este modo la biblioteca de anticuerpos o fragmentos de los mismos.

35 2. Método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la pluralidad de polinucleótidos codifica una biblioteca de Fab de al menos 10^9 Fab distintos.

3. Método de acuerdo con la reivindicación 1, en el que la pluralidad de polinucleótidos codifica una biblioteca de Fab de al menos 10^{10} Fab distintos.

40 4. Método de acuerdo con la reivindicación 1, en el que la pluralidad de polinucleótidos codifica una biblioteca de Fab de al menos $3,7 \times 10^{10}$ Fab distintos.

45 5. Método de producción de acuerdo con cualquiera de las reivindicaciones 1-4, que además comprende las etapas de:

 amplificar una primera pluralidad de polinucleótidos con un primer conjunto de cebadores, y

 amplificar una segunda pluralidad de polinucleótidos con un segundo conjunto de cebadores,

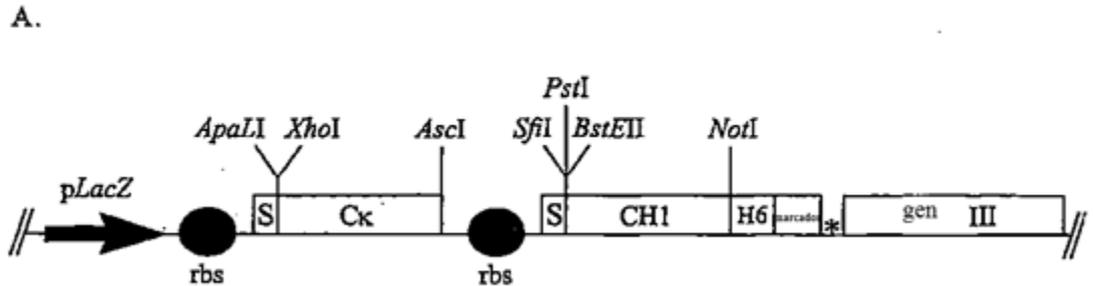
50 en el que cada conjunto de cebadores comprende oligonucleótidos diseñados para que sean homólogos al extremo 5' y 3' de los polinucleótidos que codifican las regiones variables del anticuerpo o partes del mismo, de modo que estos puedan usarse para amplificar los grupos de polinucleótidos variables a partir de fuentes naturales o sintéticas de genes conservando al mismo tiempo todo o parte del sitio de combinación antigénica del anticuerpo.

55 6. Método de acuerdo con cualquiera de las reivindicaciones 1-5, en el que la región de anclaje comprende una proteína de recubrimiento.

7. Método de acuerdo con la reivindicación 6, en el que la proteína de recubrimiento es un producto del gen III.

60 8. Método de acuerdo con la reivindicación 7, en el que la proteína de recubrimiento comprende la proteína de recubrimiento menor III de un fago filamentoso f_d .

Figura 1. Representación esquemática (A) y región polienlazadora (B) del vector fagémido pCES1.



B.

Secuencia Señal

```

--- TTA TTC GCA ATP CCT TTA GTT GTT CCT TTC TAT TCT CAC AGT GCA CAG GTC CAA CTG CAG GTC GAC CTC GAG
    L  F  A  I  P  L  V  V  P  F  Y  S  H  S  A  Q  V  Q  L  Q  V  D  L  E
    
```

ApaLI +1

XhoI

ATC AAA CGT GGA ACT GTG --- GGA GAG TGT TAA TAA GAC GCG CCA ATT CTA TTT CAA GGA GAC AGT CAT A

I K R G T V G E C * *

Gen Ck humano Terminación

Secuencia Señal

```

ATG AAA TAC CTA TTG CCT ACG GCA GGC GCT GGA TTG TTA TTA CTC GCG GCC CAG CCG GCC ATG GCC CAG GTG
M  K  Y  L  L  P  T  A  A  A  G  L  L  L  L  A  A  Q  P  A  M  A  Q  V
    
```

SfiI +1

PstI

BstEII

NotI

```

CAG CTG CAG GAG AGC GCG GTC ACC GTC TCA AGC GGC TCC ACC --- AAA TCT TGT GCG GCC GCA CAT CAT CAT CAT
Q  L  Q  B  S  G  V  T  V  S  S  A  S  T  K  S  C  A  A  A  H  H  H  H
    
```

Gen CH1 (y1) humano Marcador de

```

CAT CAC GCG GCC GCA GAA CAA AAA CTC ATC TCA GAA GAG GAT CTG AAT GGG GCC GCA TAG ACT GTT ---
H  H  G  A  A  E  Q  K  L  I  S  R  E  D  L  N  G  A  A  *  T  Y
Hexahistidina Marcador c-Myc Ámbar Gen III
    
```

Figura 2: Cebadores oligonucleotídicos usados para la construcción de la biblioteca

A. Amplificaciones primarias		B. Amplificaciones secundarias	
Región constante de la cadena pesada de IgM			
HuIgMFOR	5'-TGG AAG AGG CAC GTT CTT TTC TTT-3'		
x región constante de la cadena ligera			
HuCxFOR	5'-ACA CTC TCC CCT GTT GAA GCT CTT-3'		
λ región constante de la cadena ligera			
HuCi2-FOR	5'-TGA ACA TTC TGT AGG GGC CAC TG-3'		
HuCi7-FOR	5'-AGA GCA TTC TGC AGG GGC CAC TG-3'		
VH back			
HuVH1B/7A-BACK	5'-CAG RTG CAG CTG GTG CAR TCT GG-3'		HuVH1B/7A-BACK-SFI
HuVH1C-BACK	5'-SAG GTC CAG CTG GTR CAG TCT GG-3'		HuVH1C-BACK-SFI
HuVH2B-BACK	5'-CAG RTC ACC TTG AAG GAG TCT GG-3'		HuVH2B-BACK-SFI
HuVH3B-BACK	5'-SAG GTG CAG CTG GTG GAG TCT GG-3'		HuVH3B-BACK-SFI
HuVH3C-BACK	5'-GAG GTG CAG CTG GTG GAG WCY GG-3'		HuVH3C-BACK-SFI
HuVH4B-BACK	5'-CAG GTG CAG CTA CAG CAG TGG GG-3'		HuVH4B-BACK-SFI
HuVH4C-BACK	5'-CAG STG CAG CTG CAG GAG TCS GG-3'		HuVH4C-BACK-SFI
HuVH5B-BACK	5'-GAR GTG CAG CTG GTG CAG TCT GG-3'		HuVH5B-BACK-SFI
HuVH6A-BACK	5'-CAG GTA CAG CTG CAG CAG TCA GG-3'		HuVH6A-BACK-SFI
Vκ back			
HuVk1B-BACK	5'-GAC ATC CAG WTG ACC CAG TCT CC-3'		HuVk1B-BACK-APA
HuVk2-BACK	5'-GAT GTT GTG ATG ACT CAG TCT CC-3'		HuVk2-BACK-APA
HuVk3B-BACK	5'-GAA ATT GTG WTG ACR CAG TCT CC-3'		HuVk3B-BACK-APA
HuVk4B-BACK	5'-GAT ATT GTG ATG ACC CAC ACT CC-3'		HuVk4B-BACK-APA
HuVk5-BACK	5'-GAA ACG ACA CTC ACG CAG TCT CC-3'		HuVk5-BACK-APA
HuVk6-BACK	5'-GAA ATT GTG CTG ACT CAG TCT CC-3'		HuVk6-BACK-APA
Vλ back			
VH directo			
			HuJH1/2-FOR
			HuJH9-FOR
			HuJH4/5-FOR
			HuJH6-FOR

Figura 2: Cebadores oligonucleotídicos usados para la construcción de la biblioteca

HuV1A-BACK	5'-CAG TCT GTG CTG ACT CAG CCA CC-3'	HuV1A-BACK-APA
HuV1B-BACK	5'-CAG TCT GTG YTG ACG CAG CCG CC-3'	HuV1B-BACK-APA
HuV1C-BACK	5'-CAG TCT GTG CTG ACG CAG CCG CC-3'	HuV1C-BACK-APA
HuV12-BACK	5'-CAR TCT GGC CTG ACT CAG CCT-3'	HuV12-BACK-APA
HuV13A-BACK	5'-TCC TAT GWG CTG ACT CAG CCA CC-3'	HuV13A-BACK-APA
HuV13B-BACK	5'-TCT TCT GAG CTG ACT CAG GAC CC-3'	HuV13B-BACK-APA
HuV14-BACK	5'-CAC GTT ATA CTG ACT CAA CCG CC-3'	HuV14-BACK-APA
HuV15-BACK	5'-CAG GCT GTG CTG ACT CAG CCG TC-3'	HuV15-BACK-APA
HuV16-BACK	5'-AAT TTT ATG CTG ACT CAG CCC CA-3'	HuV16-BACK-APA
HuV17B-BACK	5'-CAG RCT GTG GTG ACY CAG GAG CC-3'	HuV17B-BACK-APA
HuV19-BACK	5'-CWG CCT GTG CTG ACT CAG CCM CC-3'	HuV19-BACK-APA

Figura 2: Cebadores oligonucleotídicos usados para la construcción de la biblioteca

5'-ACC GCC TCC ACC GGG CGC GCC TTA TTA ACA CTC TCC CCT GTT GAA GCT CTT-3'

5'-ACC GCC TCC ACC GGG CGC GCC TTA TTA TGA ACA TTC TGT AGG GGC CAC TG-3'

5'-ACC GCC TCC ACC GGG CGC GCC TTA TTA AGA GCA TTC TGC AGG GGC CAC TG-3'

5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG RTG CAG CTG GTG CAR TCT GG-3'

5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG RTG CAG CTG GTR CAG TCT GG-3'

5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG RTC ACC TTG AAG GAG TCT GG-3'

5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG RTG CAG CTG GTG GAG TCT GG-3'

5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG RTG CAG CTG GTG GAG WCY GG-3'

5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG RTG CAG CTA CAG CAG TSS GG-3'

5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG STG CAG CTG CAG GAG TGS GG-3'

5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC GAR GTG CAG CTG GTG CAG TCT GG-3'

5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG GTA CAG CTG CAG TCA GG-3'

5'-TGA GCA GAC GGT GAC CAG GGT GCC-3'

5'-TGA AGA GAC GGT GAC CAT TGT CCC-3'

5'-TGA GGA GAC GGT GAC CAG GGT TCC-3'

5'-TGA GGA GAC GGT GAC CGT GGT CCC-3'

5'-ACC GCC TCC ACC AGT GCA CTT GAC ATC CAG WTG ACC CAG TCT CC-3'

5'-ACC GCC TCC ACC AGT GCA CTT GAT GTG ATG ACT CAG TCT CC-3'

5'-ACC GCC TCC ACC AGT GCA CTT GM ATT GTG WTG ACR CAG TCT CC-3'

5'-ACC GCC TCC ACC AGT GCA CTT GAT ATT GTG ATG ACC CAC ACT CC-3'

5'-ACC GCC TCC ACC AGT GCA CTT GM ADG ADA CTC AGG CAG TCT CC-3'

5'-ACC GCC TCC ACC AGT GCA CTT GM ATT GTG CTG ACT CAG TCT CC-3'

Figura 2: Cebadores oligonucleotídicos usados para la construcción de la biblioteca

5'-ACC GCC TCC ACC AGT GCA CAG TCT GTG CTG ACT CAG CCA CC-3'
 5'-ACC GCC TCC ACC AGT GCA CAG TCT GTG YTG ACG CAG CCG CC-3'
 5'-ACC GCC TCC ACC AGT GCA CAG TCT GTC GTG ACG CAG CCG CC-3'
 5'-ACC GCC TCC ACC AGT GCA CAR TCT GCC CTG ACT CAG CCT-3'
 5'-ACC GCC TCC ACC AGT GCA CTT TCC TAT GWG CTG ACT CAG CCA CC-3'
 5'-ACC GCC TCC ACC AGT GCA CTT TCT TCT GAG CTG ACT CAG GAC CC-3'
 5'-ACC GCC TCC ACC AGT GCA CAC GTT ATA CTG ACT CAA CCG CC-3'
 5'-ACC GCC TCC ACC AGT GCA CAG GCT GTG CTG ACT CAG CCG TC-3'
 5'-ACC GCC TCC ACC AGT GCA CTT AAT TTT ATG CTG ACT CAG CCC CA-3'
 5'-ACC GCC TCC ACC AGT GCA CAG RCT GTG GTG ACG CAG GAG CC-3'
 5'-ACC GCC TCC ACC AGT GCA CWG CCT GTG CTG ACT CAG CCM CC-3'