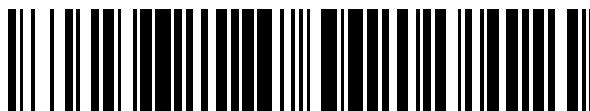


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 388 932**

51 Int. Cl.:  
**C07K 16/32** (2006.01)  
**C07K 16/28** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **06849911 .0**  
96 Fecha de presentación: **01.12.2006**  
97 Número de publicación de la solicitud: **1957540**  
97 Fecha de publicación de la solicitud: **20.08.2008**

54 Título: **Polipéptidos de unión y usos de los mismos**

30 Prioridad:  
**02.12.2005 US 742185 P**  
**22.06.2006 US 805553 P**

45 Fecha de publicación de la mención BOPI:  
**19.10.2012**

45 Fecha de la publicación del folleto de la patente:  
**19.10.2012**

73 Titular/es:  
**GENENTECH, INC.**  
**1 DNA WAY**  
**SOUTH SAN FRANCISCO, CA 94080-4990, US**

72 Inventor/es:  
**SIDHU, Sachdev S.;**  
**BIRTALAN, Sara C. y**  
**FELLOUSE, Frederic A.**

74 Agente/Representante:  
**Ponti Sales, Adelaida**

ES 2 388 932 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

Polipéptidos de unión y usos de los mismos.

## REFERENCIAS CRUZADAS CON SOLICITUDES RELACIONADAS

5 [0001] La presente solicitud es una solicitud no provisional que reivindica la prioridad de la U.S. Serie No. 60/742,185 presentada el 2 de diciembre y la U.S. Serie No. 60/805,553 presentada el 22 de junio del 2006.

## CAMPO DE LA INVENCION

10 [0002] La presente invención se refiere en general a variantes de CDR diversificadas utilizando repertorios de aminoácidos altamente limitados y bibliotecas que comprenden una pluralidad de dichas secuencias. La presente invención también se refiere a polipéptidos de fusión que comprenden estas variantes de CDR. La presente invención se refiere a métodos y composiciones útiles para identificar polipéptidos de unión nuevos que se pueden utilizar terapéuticamente o como reactivos.

## ANTECEDENTES

15 [0003] La tecnología de expresión en fagos ha proporcionado una herramienta potente para la generación y selección de nuevas proteínas que se unen a un ligando, tal como un antígeno. Utilizando las técnicas de expresión en fagos se permite la generación de amplias bibliotecas de variantes de proteínas que se pueden clasificar rápidamente por las secuencias que se unen a un antígeno diana con afinidad elevada. Los ácidos nucleicos que codifican polipéptidos variantes se fusionan a una secuencia de ácido nucleico que codifica una proteína de cubierta viral, tal como la proteína del gen III o la proteína del gen VIII. Se han desarrollado sistemas de expresión en fagos monovalentes en los que la secuencia de ácidos nucleicos que codifica la proteína o el polipéptido se fusiona a una secuencia de ácido nucleico que codifica una parte de la proteína del gen III. (Bass, S., Proteins, 8:309 (1990); Lowman and Wells, Methods: A Companion to Methods in Enzymology, 3:205 (1991)). En un sistema de expresión en fagos monovalentes, la fusión génica se expresa a niveles bajos y también se expresan proteínas del gen III de tipo salvaje, de manera que se mantiene la infectividad de las partículas. En muchas patentes se han descrito métodos de generación de bibliotecas de péptidos y cribado de estas bibliotecas (por ejemplo, patente de Estados Unidos No. 5,723,286, patente de Estados Unidos No. 5,432, 018, patente de Estados Unidos No. 5,580,717, patente de Estados Unidos No. 5,427,908 y patente de Estados Unidos No. 5,498,530).

20 [0004] La demostración de la expresión de péptidos en la superficie de fagos filamentosos y la expresión de fragmentos de anticuerpo funcionales en el periplasma era importante en el desarrollo de bibliotecas de expresión en fagos de anticuerpos. (Smith et al., Science (1985), 228:1315; Skerra and Pluckthun, Science (1988), 240:1038). Se han preparado bibliotecas de anticuerpos o polipéptidos de unión a antígeno de varias maneras mediante la alteración de un único gen mediante la inserción de secuencias de ADN aleatorias o mediante la clonación de una familia de genes relacionados. Se han descrito métodos para expresar anticuerpos o fragmentos de unión a antígeno utilizando la expresión de fagos en la patente de Estados Unidos Nos. 5,750,373, 5,733,743, 5,837,242, 5,969,108, 6,172,197, 5,580,717, y 5,658,727. A continuación, la biblioteca se criba por la expresión de anticuerpos o proteínas de unión a antígeno con las características deseadas.

25 [0005] La tecnología de expresión de fagos presenta varias ventajas sobre los métodos convencionales de hibridoma y recombinantes para la preparación de anticuerpos con las características deseadas. Esta tecnología permite el desarrollo de amplias bibliotecas de anticuerpos con diversas secuencias en menos tiempo y sin la utilización de animales. La preparación de hibridomas o la preparación de anticuerpos humanizados pueden requerir fácilmente varios meses de preparación. Además, dado que no se requiere inmunización, se pueden generar bibliotecas de anticuerpos en fagos para antígenos que son tóxicos o presentan una antigenicidad baja (Hogenboom, Immunotechniques (1988), 4:1-20). Las bibliotecas de expresión en fagos también se pueden utilizar para generar e identificar anticuerpos humanos nuevos.

30 [0006] Los anticuerpos se han convertido en muy útiles como agentes terapéuticos para una amplia variedad de afecciones. Por ejemplo, los anticuerpos humanizados para Her-2, un antígeno tumoral, son útiles en el diagnóstico y el tratamiento del cáncer. Otros anticuerpos, tales como un anticuerpo anti-IFN- $\gamma$ , son útiles en el tratamiento de afecciones inflamatorias, tales como la enfermedad de Crohn. Se han utilizando las bibliotecas de expresión en fagos para generar anticuerpos humanos de humanos inmunizados y no inmunizados, secuencias de la línea germinal, o repertorios de Ig de células B intactas (Barbas & Burton, Trends Biotech (1996), 14:230; Griffiths et al., EMBO J. (1994), 13:3245; Vaughan et al., Nat. Biotech. (1996), 14:309; Winter EP 0368 684 B1). Se han generado bibliotecas de unión a antígeno intactas o no inmunes utilizando un conjunto de tejidos linfoidales. Algunas de estas bibliotecas están disponibles comercialmente, tales como las desarrolladas por Cambridge Antibody Technology and Morphosys (Vaughan et al., Nature Biotech 14:309 (1996); Knappik et al., J. Mol. Biol. 296:57 (1999)). Sin embargo, muchas de estas bibliotecas presentan una diversidad limitada.

35 [0007] La capacidad de identificar y aislar anticuerpos de afinidad elevada de una biblioteca de expresión en fagos es importante para aislar anticuerpos humanos nuevos para uso terapéutico. Se cree, tradicionalmente,

que el aislamiento de anticuerpos de afinidad elevada de una biblioteca depende, por lo menos en parte, del tamaño de la biblioteca, la eficacia de producción en células bacterianas y la diversidad de la biblioteca. Véase, por ejemplo, Knappik et al., J. Mol. Biol. (1999), 296:57. El tamaño de la biblioteca disminuye por la ineficiencia en la producción debido a un plegamiento incorrecto del anticuerpo o proteína de unión a antígeno y la presencia de codones de parada. La expresión en células bacterianas se puede inhibir si el anticuerpo o el dominio de unión a antígeno no están correctamente plegados. La expresión se puede mejorar mediante la mutación de residuos en giros en la superficie de la interzona variable/constante o en residuos de CDR seleccionados. (Deng et al., J. Biol. Chem. (1994), 269:9533, Ulrich et al., PNAS (1995), 92:11907-11911; Forsberg et al., J. Biol. Chem. (1997), 272 :12430). La secuencia de la región armazón es un factor en la proporción para el plegamiento correcto cuando se producen bibliotecas de fagos de anticuerpos en células bacterianas.

**[0008]** La generación de una biblioteca diversa de anticuerpos o proteínas de unión a antígeno también es importante para el aislamiento de anticuerpos de afinidad elevada. Se han generado bibliotecas con diversificación en CDR limitadas utilizando una variedad de estrategias. Véase, por ejemplo, Tomlinson, Nature Biotech. (2000), 18:989-994. Las regiones CDR3 son de interés en parte porque se ha encontrado que a menudo participan en la unión a antígeno. Las regiones CD3 en la cadena pesada varían ampliamente de tamaño, secuencia y conformación estructural.

**[0009]** Otros también han generado diversidad mediante la aleatorización de regiones de CDR de las cadenas pesada y ligera variable utilizando los 20 aminoácidos en cada posición. Se creyó que utilizando los 20 aminoácidos se daría lugar a una amplia diversidad de secuencias de anticuerpos variantes y aumentaría la posibilidad de identificar nuevos anticuerpos (Barbas, PNAS 91:3809 (1994); Yelton, DE, J. Immunology, 155:1994 (1995); Jackson, J.R., J. Immunology, 154:3310 (1995) y Hawkins, RE, J. Mol. Biology, 226:889 (1992)).

**[0010]** Han habido también intentos de crear diversidad mediante la limitación del grupo de sustituciones de aminoácidos en algunas CDR para reflejar la distribución de aminoácidos en anticuerpos naturales. Véase, Garrard & Henner, Gene (1993), 128: 103; Knappik et al., J. Mol. Biol. (1999), 296:57. Sin embargo, estos intentos han tenido un éxito variable y no se han aplicado de una manera sistemática y cuantitativa. Crear la diversidad en las regiones CDR a la vez que minimizar el número de cambios de aminoácidos ha sido un desafío. Además, en algunos casos, una vez se ha generado una primera biblioteca según un grupo de criterios, puede ser deseable aumentar adicionalmente la diversidad de la primera biblioteca. Sin embargo, esto requiere que la primera biblioteca tenga una diversidad suficiente y aún permanezca suficientemente pequeña de tamaño, de manera que se pueda introducir una diversidad adicional sin superar sustancialmente las limitaciones prácticas, tales como el rendimiento, etc.

**[0011]** Algunos grupos han descrito análisis teóricos y experimentales del número mínimo del repertorio de aminoácidos que se necesita para generar proteínas. Sin embargo, estos análisis se han limitado en general en alcance y naturaleza y existe un escepticismo sustancial y cuestiones con respecto a la viabilidad de la generación de polipéptidos que presentan funciones complejas utilizando un grupo limitado de tipos de aminoácidos. Véanse, por ejemplo, Riddle et al., Nat. Struct. Biol. (1997), 4(10): 805-809; Shang et al., Proc. Natl. Acad. Sci. USA (1994), 91:8373-8377; Heinz et al., Proc. Natl. Acad. Sci. USA (1992), 89:3751-3755; Regan & Degrade, Science (1988), 241:976-978; Kamteker et al., Science (1993), 262:1680-1685; Wang & Wang, Nat. Struct. Biol. (1999), 6(11):1033-1038; Xiong et al., Proc. Natl. Acad. Sci. USA (1995), 92:6349-6353; Heinz et al., Proc. Natl. Acad. Sci. USA (1992), 89:3751-3755; Cannata et al., Bioinformatics (2002), 18(8):1102-1108; Davidson et al., Nat. Struct. Biol. (1995), 2(10):856-863; Murphy et al., Prot. Eng. (2000), 13(3):149-152; Brown & Sauer, Proc. Natl. Acad. Sci. USA (1999), 96:1983-1988; Akanuma et al., Proc. Natl. Acad. Sci. (2002), 99(21):13549-13553; Chan, Nat. Struct. Biol. (1999), 6(11):994-996.

**[0012]** De este modo, aún existe la necesidad de mejorar métodos de generación de bibliotecas que comprenden polipéptidos funcionales que presentan un grado suficiente de diversidad de secuencia, aunque son suficientemente susceptibles de manipulaciones adicionales dirigidas a una mayor diversificación, expresión de rendimiento elevado, etc. La invención aquí descrita cumple con esto.

#### DESCRIPCIÓN DE LA INVENCIÓN

**[0013]** La presente invención proporciona métodos de selección de un polipéptido que se une a HER2, tal como se establece en las reivindicaciones.

**[0014]** La presente invención proporciona métodos simplificados y flexibles de generación de polipéptidos que comprenden variantes de CDR que comprenden secuencias con una diversidad limitada, aunque mantienen la capacidad de unirse a un antígeno diana. A diferencia de los métodos convencionales que se basan en la proposición de que se puede generar una diversidad adecuada de enlazadores diana sólo si un CDR particular o todos los CDR se diversifican, y a diferencia de las ideas convencionales de que una diversidad adecuada depende del rango más amplio de sustituciones de aminoácidos (en general mediante la sustitución de todos o la mayoría de los 20 aminoácidos), la presente invención proporciona métodos capaces de generar enlazadores diana de alta calidad que no son necesariamente dependientes de la diversificación de un CDR particular o un

número particular de CDR de un polipéptido de referencia o anticuerpo de origen. La presente invención se basa, por lo menos en parte, en el hallazgo sorprendente e inesperado de que pueden generarse bibliotecas altamente diversas de calidad elevada que comprenden polipéptidos funcionales capaces de unirse a antígenos diana mediante la diversificación de un número mínimo de posiciones de aminoácidos con un número altamente limitado de residuos de aminoácidos. Los métodos de la invención son rápidos, convenientes y flexibles, se basan en la utilización de grupos de codones limitados que codifican un número bajo de aminoácidos. La diversidad de secuencia limitada, y de este modo, el tamaño generalmente menor de las poblaciones (por ejemplo bibliotecas) de polipéptidos generados por los métodos de la invención permite una mayor diversificación de estas poblaciones, cuando sea necesario o deseable. Ésta es una ventaja generalmente no proporcionada por métodos convencionales. Los polipéptidos enlazadores candidatos generados por la invención poseen características de unión a diana de alta calidad y presentan características estructurales que proporciona un rendimiento elevado de producción en cultivo celular. La presente invención proporciona métodos para generar estos polipéptidos enlazadores, tal como se define en las reivindicaciones.

**[0015]** Se describen aquí polipéptidos de fusión que comprenden CDR diversificadas y una secuencia de polipéptido heteróloga (en ciertas realizaciones, la de por lo menos una parte de un polipéptidos viral), como polipéptidos individuales y como un miembro de una pluralidad de polipéptidos individuales únicos que son enlazadores candidatos a dianas de interés. Las composiciones (tales como bibliotecas) que comprenden dichos polipéptidos son útiles en una variedad de aplicaciones, por ejemplo, como grupos de polipéptidos inmunoglobulina candidatos (por ejemplo, anticuerpos y fragmentos de anticuerpo) que se unen a dianas de interés. Dichos polipéptidos también se pueden generar utilizando estructuras que no son inmunoglobulina (por ejemplo, proteínas, tales como hormona del crecimiento humano, etc.). La presente descripción presenta varios aspectos, incluyendo polinucleótidos y polipéptidos generados según los métodos de la invención y sistemas, kits, artículos de fabricación para realizar métodos de la invención, y/o utilizando los polipéptidos /polinucleótidos y/o composiciones aquí descritos.

**[0016]** Se describen aquí un método de generación de un polipéptido que comprende por lo menos uno, dos, tres, cuatro, cinco o todas las CDR variantes seleccionadas del grupo que consiste en H1, H2, H3, L1, L2 y L3, en el que dicho polipéptido es capaz de unirse a un antígeno diana de interés, comprendiendo dicho método identificar por lo menos una (o cualquier número hasta todas) posición de aminoácido accesible al disolvente y altamente diversa en una CDR de referencia correspondiente a la CDR variante; y (ii) variar el aminoácido en la posición accesible al disolvente y altamente diversa mediante la generación de copias variantes de la CDR utilizando un grupo de codones limitado (la definición de "grupo de codones limitados" se proporciona a continuación).

**[0017]** Varios aspectos y realizaciones de los métodos de la invención son útiles para generar y/o utilizar un grupo que comprende una pluralidad de polipéptidos para seleccionar e identificar enlazadores candidatos para HER2, tal como se define en las reivindicaciones. Se describe aquí un método de generación de una composición que comprende una pluralidad de polipéptidos, cada polipéptido comprendiendo por lo menos una, dos, tres, cuatro, cinco o todas las variantes de CDR seleccionadas del grupo que consiste en H1, H2, H3, L1, L2 y L3, en el que dicho polipéptido es capaz de unirse a un antígeno diana de interés, comprendiendo dicho método identificar por lo menos una posición (o cualquier número hasta todas) de aminoácido accesible a disolvente y altamente diversa en una CDR de referencia correspondiente a la CDR variante; y (ii) variar el aminoácido en la posición accesible a disolvente y altamente diversa mediante la generación de copias de variantes de la CDR utilizando un grupo de codones limitados; en el que se genera una pluralidad de polipéptidos mediante la amplificación de un polinucleótido plantilla con un grupo de oligonucleótidos que comprende una degeneración altamente limitada en la secuencia que codifica un aminoácido variante, en el que dicha degeneración limitada refleja el número limitado de secuencias de codones del grupo de codones limitados.

**[0018]** Se describe aquí un método que comprende: construir un vector de expresión que comprende una secuencia de polinucleótidos que codifica dominios variables de una cadena ligera, una cadena pesada, o tanto una cadena ligera como una cadena pesada de un anticuerpo origen que comprende por lo menos una, dos, tres, cuatro, cinco o todas las CDR seleccionadas del grupo que consiste en CDR L1, L2, L3, H1, H2 y H3; y mutar por lo menos una, dos, tres, cuatro, cinco o todas las CDR del anticuerpo de origen en por lo menos una (o cualquier número hasta todas) posición de aminoácido accesible a disolvente y altamente diversa utilizando un grupo de codones limitado.

**[0019]** Se describen aquí un método que comprende: construir una biblioteca de fagos o partículas fagémidas que muestran una pluralidad de polipéptidos de la invención; poner en contacto la biblioteca de partículas con un antígeno diana bajo condiciones adecuadas para unir las partículas al antígeno diana; y separar las partículas que se unen de las que no se unen al antígeno diana.

**[0020]** Una posición de aminoácido accesible a disolvente y/o altamente diversa puede ser cualquiera que cumpla los criterios aquí descritos, en particular cualquier combinación de las posición aquí descritas, por ejemplo, cualquier combinación de las posiciones descritas para los polipéptidos de la invención (aquí descritos con mayor detalle). Las variantes de aminoácidos adecuados pueden ser cualquiera que cumplan los criterios aquí descritos, por ejemplo variantes de aminoácidos en polipéptidos de la invención descritos con mayor detalle

a continuación.

**[0021]** El diseño de la diversidad en CDR puede implicar diseñar la diversidad en la longitud y/o en la secuencia de la CDR. Por ejemplo, CDH3 se puede diversificar en la longitud para que tenga, por ejemplo, 7 a 21 aminoácidos de largo y/o en su secuencia, por ejemplo, variando las posiciones altamente diversas y/o accesibles a disolvente con aminoácidos codificados por un grupo de codones limitado. En algunos casos, una parte de CDRH3 presenta una longitud que varía desde 5 a 21, 7 a 20, 9 a 15 u 11 a 13 aminoácidos, y presenta una variante de aminoácido en una o más posiciones codificadas por un grupo de codones limitado que codifica un número limitado de aminoácidos, tales como grupos de codones que codifican no más de 19, 15, 10, 8, 6, 4 ó 2 aminoácidos. En algunos casos, el extremo C terminal presenta una secuencia de aminoácidos AM, AMDY, o DY.

**[0022]** Los polipéptidos aquí descritos pueden estar en una variedad de formar siempre que se mantenga la función de unión a la diana de los polipéptidos.

**[0023]** Por ejemplo, un polipéptido puede ser un polipéptido de fusión (es decir, una fusión de dos o más secuencias de polipéptidos heterólogos). Los polipéptidos con CDR diversificadas se pueden preparar como polipéptidos de fusión a por lo menos una parte de una proteína de cubierta viral, por ejemplo, para utilizar en la expresión de fagos. Las proteínas de cubierta viral que se pueden utilizar para la expresión de los polipéptidos de la invención comprenden la proteína pIII, la proteína pVIII de cubierta principal, Soc (fago T4), Hoc (fago T4), gpD

(fago lambda), pVI, o fragmentos variantes de los mismos. Por ejemplo, el polipéptido de fusión se fusiona a por lo menos una parte de una proteína de cubierta viral, tal como una proteína de cubierta viral seleccionada del grupo que consiste en pIII, pVIII, Soc, Hoc, gpD, pVI, y variantes o fragmentos de las mismas.

**[0024]** Cuando el polipéptido con CDR diversificadas es uno o más dominios variables de anticuerpo, los dominios variables de anticuerpo se pueden expresar en la superficie del virus en una variedad de formatos, que incluyen ScFv, Fab, ScFv2, F(ab')<sub>2</sub> y F(ab)<sub>2</sub>. Para la expresión de los polipéptidos de manera bivalente, la proteína de fusión en ciertas realizaciones incluye un dominio de dimerización. El dominio de dimerización puede comprender una secuencia de dimerización y/o una secuencia que comprende uno o más residuos cisteína. El dominio de dimerización puede estar unido, directa o indirectamente, al extremo c-terminal de un dominio variable o constante de cadena pesada (por ejemplo, CH1). La estructura del dominio de dimerización puede variar dependiendo de si el dominio variable de anticuerpo es producido como un componente proteína de fusión con el componente proteína de la cubierta viral (por ejemplo, un codón de parada amber después del dominio de dimerización) o si el dominio variable de anticuerpo es producido predominantemente sin el componente proteína de cubierta viral (por ejemplo, con un codón de parada amber después del dominio de dimerización). Cuando el dominio variable de anticuerpo es producido predominantemente como una proteína de fusión con el componente proteína de cubierta viral, uno o más enlaces disulfuro y/o una secuencia única de dimerización se proporciona una expresión bivalente. Para los dominios variables de anticuerpo predominantemente producidos sin estar fusionados a un componente de proteína de cubierta viral (por ejemplo, con un codón de parada amber), el dominio de dimerización puede comprender un residuo de cisteína y una secuencia de dimerización.

**[0025]** Además, opcionalmente, un polipéptido de fusión puede comprender una etiqueta que puede ser útil en la purificación, detección, y/o cribado tal como FLAG, poli-his, gD tag, c-myc, proteína de fluorescencia o B-galactosidasa. En una realización, un polipéptido de fusión comprende un dominio variable o constante de cadena ligera fusionado a un polipéptido etiqueta.

**[0026]** Un polipéptido, tal como un dominio variable de anticuerpo, se puede obtener de una única fuente o de una molécula plantilla. La fuente o molécula plantilla se pueden seleccionar o diseñar por características, tales como un buen rendimiento y estabilidad cuando se producen en un cultivo de células procariontas o eucariotas y/o para acomodar las regiones CDRH3 de longitudes variables. La secuencia de la molécula plantilla se puede alterar para mejorar el plegamiento y/o expresar el dominio variable cuando se presenta como una proteína de fusión con un componente proteína de cubierta de fago. Por ejemplo, un anticuerpo origen puede comprender la secuencia de aminoácidos de los dominios variables de anticuerpo humanizado 4D5 (dominio variable de cadena ligera (Figura 1; SEQ ID NO: 1)); (dominio variable de cadena pesada (Figura 1; SEQ ID NO: 2)). Por ejemplo, en un dominio variable de anticuerpo de una cadena pesada o ligera, los residuos de la región de armazón se pueden modificar o alterar a partir de la fuente o la molécula plantilla para mejorar el plegamiento, el rendimiento, la expresión o la afinidad del dominio variable de anticuerpo. En algunas realizaciones, los residuos de armazón se seleccionan para modificarse a partir del origen o la molécula plantilla cuando el aminoácido en la posición del armazón de la molécula de origen es diferente del aminoácido o aminoácidos hallados normalmente en esa posición en anticuerpos naturales o en un subgrupo de secuencia consenso. Los aminoácidos en esas posiciones se pueden cambiar a los aminoácidos hallados más normalmente en los anticuerpos naturales o en un subgrupo de secuencia consenso en esa posición. El residuo de armazón 71 de la cadena pesada puede ser R, V o A. El residuo de armazón 93 de la cadena pesada puede ser S o A.

**[0027]** El residuo de armazón 94 puede ser R, K o T o codificado por MRT. El residuo de armazón 93 es A y el residuo de armazón 94 es R. El residuo de armazón 49 en la cadena pesada puede ser alanina o glicina. Los

residuos de armazón en la cadena ligera también pueden cambiar. Por ejemplo, el aminoácido en la posición 66 puede ser arginina o glicina. Las regiones de armazón para las secuencias de cadena ligera y cadena pesada del anticuerpo humanizado 4D5-8 de tipo natural se muestran en la figura 6 (SEQ ID NOs: 6-9 y 10-13, respectivamente). Las regiones de armazón para versiones variantes de las secuencias de cadena ligera y cadena pesada del anticuerpo humanizado 4D5-8, en las que la cadena ligera se modifica en la posición 66 y la cadena pesada se modifica en las posiciones 71, 73 y 78 se muestran en la figura 7 (SEQ ID NOs:14-17 y 18-21, respectivamente).

**[0028]** Los métodos de la invención son capaces de generar una amplia variedad de polipéptidos que comprenden un grupo diverso de secuencias de CDR. En una realización, se forman una o más bibliotecas utilizando los métodos de la invención tal como se describe aquí y se criban las bibliotecas para la unión a antígeno HER-2 diana.

**[0029]** Se describen aquí dominios variables de cadena pesada de inmunoglobulina aleatorizados para proporcionar diversidad.

**[0030]** Tal como se define en las reivindicaciones, los métodos de la invención comprenden generar una pluralidad de polipéptidos, en los que los polipéptidos comprenden un dominio variable de cadena pesada de inmunoglobulina, en el que:

(i) CDRH1 comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ID NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;

(ii) CDRH2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y

(iii) CDRH3 comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:24), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 50% Y, 25% S, y 25% G; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 50% Y, 25% S, y 25% G, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.

**[0031]** Los polipéptidos pueden comprender un dominio variable de cadena pesada de inmunoglobulina,

(i) CDRH1 comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ID NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;

(ii) CDRH2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y

(iii) CDRH3 comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:26), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 25% Y, 50% S, y 25% R; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 25% Y, 50% S, y 25% R, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.

**[0032]** Los polipéptidos pueden comprender un dominio variable de cadena pesada de inmunoglobulina, en el que:

(i) CDRH1 comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ID NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;

(ii) CDRH2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona

entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y

(iii) CDRH3 comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:27), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 38% Y, 25% S, 25% G, y 12% R; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 38% Y, 25% S, 25% G, y 12% R, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.

10 **[0033]** Los polipéptidos pueden comprender un dominio variable de cadena pesada de inmunoglobulina, en el que:

(i) CDRH1 comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ID NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;

(ii) CDRH2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y

(iii) CDRH3 comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:28), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 20% Y, 26% S, 26% G, 13% R, 1% A, 1% D, 1% E, 1%F, 1%H, 1%I, 1%K, 1% L, 1% M, 1%N, 1%P, 1%Q, 1%T, 1%V, y 1% W; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 20% Y, 26% S, 26% G, 13% R, 1% A, 1% D, 1% E, 1% F, 1% H, 1% I, 1% K, 1% L, 1% M, 1% N, 1% P, 1% Q, 1% T, 1% V, y 1% W, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.

**[0034]** Tal como se define en las reivindicaciones, los métodos de la invención comprenden generar una pluralidad de polipéptidos. Dichos polipéptidos pueden comprender además un dominio variable de cadena ligera, en el que:

(i) CDRL3 comprende una secuencia de aminoácidos Q-Q-X1-X2-X3-X4-P-X5-T (SEQ ID NO:25); en la que X1 está en la posición 91 y se selecciona entre Y, H y S; X2 se selecciona entre Y y S; X3 se selecciona entre Y, S y T; X4 se selecciona entre Y, S y T; y X5 se selecciona entre S, P y Y. En una realización, un CDRL1 comprende una secuencia de aminoácidos RASQDVNTAVA (SEQ ID NO:29). En una realización, un CDRL2 comprende una secuencia de aminoácidos SASSLYS (SEQ ID NO:30).

**[0035]** En otra realización, un polipéptido comprende un dominio variable de cadena pesada de inmunoglobulina, en el que (i) CDRH1 comprende la secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ID NO:22); en la que X1 está en la posición 28 según el sistema de numeración Kabat y se selecciona entre S y Y; X2 se selecciona entre S y Y; X3 se selecciona entre S y Y; X4 se selecciona entre S y Y; y X5 se selecciona entre S y Y; (ii)CDRH2 comprende una secuencia de aminoácidos de X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23); en la que X1 está en la posición de aminoácido 50 según el sistema de numeración Kabat y se selecciona entre S y Y; X2 se selecciona entre S y Y; X3 se selecciona entre S y Y; X4 se selecciona entre S y Y; X5 se selecciona entre S y Y; y X6 se selecciona entre S y Y; y (iii) CDRH3 comprende la secuencia de aminoácidos X1-X2-X3-X4-X5-X6-X7-D-Y (SEQ ID NO:582), en la que X1 está en la posición de aminoácido 95 según el sistema de numeración Kabat y se selecciona entre Y y R; X2 se selecciona entre Y, S y R; X3 se selecciona entre S, G, Y y H; X4 se selecciona entre S, G, Y y R; X5 se selecciona entre G y A; X6 se selecciona entre F, M, L, y A; y X7 se selecciona entre F, M, y L o está ausente.

**[0036]** Un CDRH 1 puede comprender por lo menos una secuencia de aminoácidos seleccionada entre cualquiera de SEQ ID NOs:189-198. CDRH2 comprende por lo menos una secuencia de aminoácidos seleccionada entre SEQ ID NOs:295-304. CDRH3 comprende por lo menos una secuencia de aminoácidos seleccionada entre SEQ ID NOs:401-410.

**[0037]** Los polipéptidos pueden comprender un dominio variable de cadena pesada de inmunoglobulina, en el que (i) CDRH1 comprende la secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ID NO:22); en la que X1 está en la posición 28 según el sistema de numeración Kabat y se selecciona entre S y Y; X2 se selecciona entre S y Y; X3 se selecciona entre S y Y; X4 se selecciona entre S y Y; y X5 se selecciona entre S y Y;(ii)CDRH2 comprende la secuencia de aminoácidos de X1-I-X2-P-X3-X4-G-X5-TX6-Y-A-D-S-V-K-G (SEQ ID NO:22); en la que X está en la posición de aminoácido 50 según el sistema de numeración Kabat y se selecciona entre S e Y; X2 se selecciona entre S e Y; X3 se selecciona entre S e Y; X4 se selecciona entre S e Y; y X5 se

- selecciona entre S e Y; (iii)CDRH3 comprende una secuencia de aminoácidos X1-X2-X3-X4-X5-X6-X7-X8-D-Y (SEQ ID NO:583), en la que X1 está en la posición de aminoácido 95 según el sistema de numeración Kabat y se selecciona entre Y, S y G; X2 se selecciona entre Y, S, G, R, A, y M; X3 se selecciona entre G, Y, S y R; X4 se selecciona entre G, Y y F; X5 se selecciona entre Y, S, N, y G; X6 se selecciona entre Y, R, H y W; X7 se selecciona entre G y A; y X8 se selecciona entre F, M, L e I.
- 5
- [0038]** Un CDRH1 puede comprender por lo menos una secuencia de aminoácidos seleccionada entre cualquiera de SEQ ID NOs:199-216. CDRH2 comprende por lo menos una secuencia de aminoácidos seleccionada entre SEQ ID NOs:305-322. CDRH3 comprende por lo menos una secuencia de aminoácidos seleccionada entre SEQ ID NOs:411-428.
- 10
- [0039]** Los polipéptidos pueden comprender un dominio variable de cadena pesada de inmunoglobulina, en el que
- (i) CDRH1 comprende la secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ID NO:22); en la que X1 está en la posición 28 según el sistema de numeración Kabat y se selecciona entre S e Y; X2 se selecciona entre S e Y; X3 se selecciona entre S e Y; X4 se selecciona entre S e Y; y X5 se selecciona entre S e Y;
- 15
- (ii) CDRH2 comprende una secuencia de aminoácidos de X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23); en la que X1 está en la posición de aminoácido 50 según el sistema de numeración Kabat y se selecciona entre S e Y; X2 se selecciona entre S e Y; X3 se selecciona entre S e Y; X4 se selecciona entre S e Y; y X5 se selecciona entre S e Y; y
- 20
- (iii) CDRH3 comprende una secuencia de aminoácidos de X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:584), en la que X1 está en la posición de aminoácido 95 y se selecciona entre Y, S, R, G y E; X2 se selecciona entre Y, S, R y G; X3 se selecciona entre S, Y, G y W; X4 se selecciona entre S, Y, G y Q; X5 se selecciona entre G, Y y S; X6 se selecciona entre G, Y, S, R y V; X7 se selecciona entre S, Y, G y R; X8 se selecciona entre Y, S, G, R, P y V; X9 se selecciona entre G, A, Y, S y R; X10 se selecciona entre M, F, G, Y, S y R; X11 se selecciona entre A, Y, S, G y R o no está presente; X12 se selecciona entre I, M, F, L, A, G, S, Y, R, y T o no está presente; X13 se selecciona entre F, M, L, G, A, Y, T, y S o no está presente; X14 se selecciona entre L, M, F, I, G, Y, A, y T o no está presente; X15 se selecciona entre M, L, Y, G y R o no está presente; X16 se selecciona entre Y y G o no está presente; X17 se selecciona entre R, M, y G o no está presente; X18 se selecciona entre P y A o no está presente; y X18 es L o no está presente.
- 25
- [0040]** Los polipéptidos pueden comprender un dominio variable de cadena pesada de inmunoglobulina, en el que:
- 30
- (i) CDRH1 comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H, en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;
- 35
- (ii) CDR2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G, en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y
- 40
- (iii) CDRH3 comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19, en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X17 se seleccionan entre S y uno de A, C, F, G, I, L, N, P, R, T, W, o Y, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre F, L, I, y M.
- 45
- [0041]** Un CDRH1 puede comprender por lo menos una secuencia de aminoácidos seleccionada entre cualquiera de SEQ ID NOs:816-842 o por lo menos una secuencia de aminoácidos de CDRH1 seleccionada entre cualquiera de las secuencias en la figura 21A. CDRH2 comprende por lo menos una secuencia de aminoácidos seleccionada entre SEQ ID NOs:843-869 o por lo menos una secuencia de aminoácidos de CDRH2 seleccionada entre cualquiera de las secuencias en la figura 21A. CDRH3 comprende por lo menos una secuencia de aminoácidos seleccionada entre SEQ ID NOs:870-896 o por lo menos una secuencia de aminoácidos de CDRH3 seleccionada entre cualquiera de las secuencias en la Figura 21A.
- 50
- [0042]** Un polipéptido se puede unir a HER-2 humano. En algunos casos, el polipéptido comprende un anticuerpo. El anticuerpo puede comprender un dominio variable de cadena pesada que comprende
- i) un CDRH1 que comprende una secuencia de aminoácidos GFSIYSSYIH (SEQ ID NO:821);
- ii) un CDRH2 que comprende una secuencia de aminoácidos SIYPYSGYTSYADSVKG (SEQ ID NO:848); y



iii) un CDRH3 que comprende una secuencia de aminoácidos WWSSAFDY (SEQ ID NO: 875).

**[0043]** En otro caso, el anticuerpo comprende un dominio variable de cadena pesada que comprende

i) un CDRH1 que comprende una secuencia de aminoácidos GFYISSSIH (SEQ ID NO:230);

ii) un CDRH2 que comprende una secuencia de aminoácidos YIYPSSGYTSYADSVKG (SEQ ID NO:336); y

5 iii) un CDRH3 que comprende una secuencia de aminoácidos GYYYSYSGYALDY (SEQ ID NO:442).

**[0044]** En otro caso, el anticuerpo comprende además un dominio variable de cadena ligera que comprende una secuencia de CDRL3, en la que CDRL3 comprende una secuencia de aminoácidos de Q-Q-X1-X2-X3-X4-P-X5-T (SEQ ID NO:25), en la que X1 está en la posición 91 según el sistema de numeración Kabat y se selecciona entre S y Y; X2 se selecciona entre S, Y y F; X3 se selecciona entre Y, S y F; X4 se selecciona entre Y y S; X5 se selecciona entre S y Y.

10

**[0045]** En un caso, CDRL3 comprende una secuencia de aminoácidos seleccionada entre SEQ ID NO LS: 83-188 y 789-815 o cualquiera las secuencias de CDRL3 en las figuras 11 ó 21.

**[0046]** En ciertos casos, se proporciona un polipéptido que comprende por lo menos dos dominios variables de anticuerpos que comprenden: (a) un dominio variable de anticuerpo de cadena pesada que comprende cualquiera de los polipéptidos de cadena pesada indicadas anteriormente, y (b) un dominio variable de anticuerpo de cadena ligera que comprende cualquiera de los polipéptidos de cadena ligera indicados anteriormente.

15

**[0047]** En ciertos casos, se proporciona un anticuerpo que comprende un polipéptido que comprende un dominio variable de cadena pesada de inmunoglobulina según cualquiera de los polipéptidos de cadena pesada indicados anteriormente, y un polipéptido que comprende un dominio variable de cadena ligera de inmunoglobulina según cualquiera de los polipéptidos de cadena ligera indicados anteriormente.

20

**[0048]** En ciertos casos, los polipéptidos y anticuerpos indicados anteriormente pueden comprender un dominio de dimerización unido a la región C-terminal de un dominio variable de anticuerpo de cadena pesada. En ciertos casos, el dominio de dimerización comprende un dominio de cremallera de leucina o una secuencia que comprende por lo menos un residuo de cisteína. En dichos casos, el dominio de dimerización comprende una región bisagra de un anticuerpo y una cremallera de leucina. En ciertos otros aspectos, el dominio de dimerización es una única cisteína.

25

**[0049]** En un caso, se proporciona un polipéptido de fusión que comprende cualquiera de los polipéptidos indicados anteriormente, en el que un dominio variable de anticuerpo que comprende el polipéptido indicado anteriormente está fusionado a por lo menos una parte de una proteína de cubierta viral. En un caso, la proteína de cubierta viral se selecciona del grupo que consiste en proteína pIII, proteína pVIII de cubierta principal, Soc, Hoc, gpD, pv1, y variantes de las mismas. En un caso, el polipéptido de fusión comprende además un dominio de dimerización entre en dominio variable y la proteína de cubierta viral. En dicho caso, el dominio variable es un dominio variable de cadena pesada. En otro caso, el polipéptido de fusión comprende además un dominio variable fusionado a un péptido etiqueta. En dicho caso, el dominio variable es un dominio variable de cadena ligera. En otro de dichos casos, el péptido etiqueta se selecciona del grupo que consiste en gD, c-myc, poly-his, una proteína fluorescente, y  $\beta$ -galactosidasa.

30

35

**[0050]** En un caso, uno o más de los polipéptidos descritos anteriormente comprenden además regiones almacén FR1, FR2, FR3, y/o FR4 para un dominio variable de anticuerpo correspondiente a la variante CDRH1, CDRH2, CDRH3, y/o CDRL3, donde las regiones de almacén se obtienen de una única plantilla de anticuerpo. En algunos de dichos casos, cada una de las regiones de almacén comprende una secuencia de aminoácidos correspondiente a las secuencias de aminoácidos de la región de almacén del anticuerpo 4D5 (SEQ ID NOS: 6-9 y 10-13) o una variante del anticuerpo 4D5 (SEQ ID NOS: 14-17 y 18-21).

40

**[0051]** Se describe una biblioteca que comprende una pluralidad de uno o más de los polipéptidos descritos anteriormente, en la que la biblioteca tiene por lo menos  $1 \times 10^4$  secuencias distintas de dominio variable de anticuerpo.

45

**[0052]** Se describe aquí un método de generación de una composición que comprende una pluralidad de polipéptidos, que comprende:

(a) generar una pluralidad de polipéptidos que comprenden:

50

(i) CDRH1 que comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ID NO: ), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;

(ii) CDRH2 que comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO: ), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y

- 5 (iii) CDRH3 que comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO: ), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 50% Y, 25% S, y 25% G; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 50% Y, 25% S, y 25% G, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.

**[0054]** En un caso, el método comprende además:

(b) generar una pluralidad de polipéptidos que comprenden:

- 15 (i) CDRL1 que comprende una primera secuencia hipervariable consenso o variante de la misma que comprende una sustitución en una o más de las posiciones en comparación con una secuencia hipervariable consenso correspondiente;

(ii) CDRL2 que comprende una segunda secuencia hipervariable consenso o variante de la misma que comprende la sustitución en una o más posiciones en comparación con una secuencia hipervariable consenso correspondiente; y

- 20 (iii) CDRL3 que comprende una secuencia de aminoácidos: Q-Q-X1-X2-X3-X4-P-X5-T (SEQ ID NO: ), en la que X1 está en la posición 91 según el sistema de numeración Kabat, y en la que X1 se selecciona entre Y y S, en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S.

En un caso, la pluralidad de polipéptidos es codificada por una pluralidad de polinucleótidos.

- 25 **[0055]** Se describe aquí un método de generación de una composición que comprende una pluralidad de polipéptidos, que comprende:

(a) generar una pluralidad de polipéptidos que comprenden:

- 30 (i) CDRH1 que comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ID NO: ), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;

- 35 (ii) CDRH2 que comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO: ), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y

- 40 (iii) CDRH3 que comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO: ), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 25% Y, 50% S, y 25% R; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 25% Y, 50% S, y 25% R; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.

**[0056]** En un caso, el método comprende además:

(b) generar una pluralidad de polipéptidos que comprende:

- 45 (i) CDRL1 que comprende una primera secuencia hipervariable consenso o variante de la misma que comprende sustituciones en una o más posiciones en comparación con una secuencia hipervariable consenso correspondiente;

(ii) CDRL2 que comprende una segunda secuencia hipervariable consenso o variante de la misma que comprende sustituciones en una o más posiciones en comparación con una secuencia hipervariable consenso correspondiente; y

- 50 (iii) CDRL3 que comprende una secuencia de aminoácidos: Q-Q-X1-X2-X3-X4-P-X5-T (SEQ ID NO: ), en la que X1 está en la posición 91 según el sistema de numeración Kabat, y en la que X1 se selecciona entre Y y S, en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S;

S; y en la que X5 se selecciona entre Y y S.

**[0057]** En un caso, la pluralidad de polipéptidos son codificados por una pluralidad de polinucleótidos.

**[0058]** Se describe aquí un método de generación de una composición que comprende una pluralidad de polipéptidos que comprenden:

5 (a) generar una pluralidad de polipéptidos que comprenden:

(i)CDRH1 que comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQIDNO: ), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;

10 (ii) CDRH2 que comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO: ), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y

15 (iii) CDRH3 que comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO: ), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 38% Y, 25% S, 25% G, y 12% R; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 38% Y, 25% S, 25% G, y 12% R, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.

20

**[0059]** En un caso, el método comprende además:

(b) generar una pluralidad de polipéptidos que comprenden:

25 (i) CDRL1 que comprende una primera secuencia hipervariable consenso o variante de la misma que comprende sustituciones en una o más posiciones en comparación con una secuencia hipervariable consenso correspondiente;

(ii) CDRL2 que comprende una segunda secuencia hipervariable consenso o variante de la misma que comprende sustituciones en una o más posiciones en comparación con una secuencia hipervariable consenso correspondiente; y

30 (iii) CDRL3 que comprende una secuencia de aminoácidos: Q-Q-X1-X2-X3-X4-P-X5-T (SEQ ID NO: ), en la que X1 está en la posición 91 según el sistema de numeración Kabat, y en la que X1 se selecciona entre Y y S, en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S.

30

**[0060]** En un caso, la pluralidad de polipéptidos son codificados por una pluralidad de polinucleótidos.

**[0061]** Se describe aquí un método de generación de una composición que comprende una pluralidad de polipéptidos que comprenden:

35

(a) generar una pluralidad de polipéptidos que comprenden:

40 (i) CDRH1 que comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H(SEQIDNO: ), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;

40

(ii) CDRH2 que comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO: ), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que I1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y

45 (iii) CDRH3 que comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO: ), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 20% Y, 26% S, 26% G, 13% R, 1% A, 1% D, 1% E, 1% F, 1% H, 1% I, 1% K, 1% L, 1% M, 1% N, 1% P, 1% Q, 1% T, 1% V, y 1% W; en la que los aminoácidos en cada una de las posiciones X7-X17 se selecciona de un grupo de aminoácidos en una proporción molar de 20% Y, 26% S, 26% G, 13% R, 1% A, 1% D, 1% E, 1% F, 1% H, 1% I, 1% K, 1% L, 1% M, 1% N, 1% P, 1% Q, 1% T, 1% V, y 1% W, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.

50

**[0062]** En un caso, el método comprende además:

(b) generar una pluralidad de polipéptidos que comprenden:

(i) CDRL1 que comprende una primera secuencia hipervariable consenso o variante de la misma que comprende una sustitución en una o más posiciones en comparación con una secuencia hipervariable consenso correspondiente;

(ii) CDRL2 que comprende una segunda secuencia hipervariable consenso o variante de la misma que comprende una sustitución en una o más de las posiciones en comparación con una secuencia hipervariable consenso correspondiente; y

(iii) CDRL3 que comprende una secuencia de aminoácidos: Q-Q-X1-X2-X3-X4-P-X5-T (SEQ ID NO: ), en la que X1 está en la posición 91 según el sistema de numeración Kabat, y en la que X1 se selecciona entre Y y S, en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S.

**[0063]** En un caso, la primera secuencia hipervariable consenso comprende una secuencia CDRL1 consenso Kabat. En dicho caso, la primera secuencia hipervariable consenso es R-A-S-Q-D-V-N-T-A-V-A (SEQ ID NO: 29). En un caso, la segunda secuencia hipervariable consenso comprende una secuencia CDRL2 consenso Kabat. En dicho aspecto, la segunda secuencia hipervariable consenso es S-A-S-S-L-Y-S (SEQ ID NO: 30). En un caso, la pluralidad de polipéptidos son codificados por una pluralidad de polinucleótidos.

**[0064]** Se describen aquí un método de generación de una composición que comprende una pluralidad de polipéptidos que comprende:

(a) generar una pluralidad de polipéptidos que comprenden:

(i) CDRH1 que comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-XS-I-H (SEQ ID NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;

(ii) CDRH2 que comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y

(iii) CDRH3 comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:24), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que X1 se selecciona entre Y,S,G,R y E; X2 se selecciona entre Y,S,G,R,M y A; X3 se selecciona entre G,Y, S,R,W, y H, X4 se selecciona entre Y,S,G,R, F y Q; X5 se selecciona entre G, Y, N, A, y S; X6 se selecciona entre F, M, L, A, R, G, H, W, V, Y y S; X7 se selecciona entre M, L, G, A, , R, F, Y y S o no está presente; X8 se selecciona entre M, L, F, I, ,R, G, P, V, Y y S o no está presente; X9 se selecciona entre G, Y, R, y S o no está presente; X10 se selecciona entre M, F, G, Y, R, y S o no está presente; X11 se selecciona entre A, G, Y, R, y S o no está presente; X12 se selecciona entre I, M, L, F, A, G, R, T, Y y S o no está presente; X13 se selecciona entre F, M, L, G, A, T, Y y S o no está presente; X14 se selecciona entre L, F, M, I, G, A, T, e Y o no está presente; X15 se selecciona entre M, Y G, L, y R o no está presente; X16 se selecciona entre Y y G o no está presente; X17 se selecciona entre R, M, y G o no está presente; X18 se selecciona entre P y A o no está presente; y X19 es L o no está presente.

**[0065]** En un caso, CDRH1 comprende una secuencia de aminoácidos seleccionada entre SEQ ID NOs: 189-294. En un caso, CDRH2 comprende una secuencia de aminoácidos seleccionada entre SEQ ID NOs: 295-400. En un caso, CDRH3 comprende una secuencia de aminoácidos seleccionada entre SEQ ID NOs: 401-506.

**[0066]** En un caso, el método comprende además:

(b) generar una pluralidad de polipéptidos que comprenden:

(i) CDRL1 que comprende una primera secuencia hipervariable consenso o variante de la misma que comprende una sustitución en una o más posiciones en comparación con una secuencia hipervariable consenso correspondiente;

(ii) CDRL2 que comprende una segunda secuencia hipervariable consenso o variante de la misma que comprende una sustitución en una o más posiciones en comparación con una secuencia hipervariable consenso correspondiente; y

(iii) CDRL3 que comprende una secuencia de aminoácidos: Q-Q-X1-X2-X3-X4-P-X5-T (SEQ ID NO:25), en la que X1 está en la posición 91 según el sistema de numeración Kabat, y en la que X1 se selecciona entre Y y S,

en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S.

**[0067]** En un caso, la pluralidad de polipéptidos son codificados por una pluralidad de polinucleótidos.

5 **[0068]** Se describe aquí un método de generación de una composición que comprende una pluralidad de polipéptidos que comprende:

(a) generar una pluralidad de polipéptidos que comprende:

10 (i) CDRH1 que comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQIDNO: ), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;

(ii) CDRH2 que comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO: ), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S;

15 (iii) CDRH3 que comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19 (SEQ ID NO: ), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X17 se seleccionan entre S uno de A, C, F, G, I, L, N, P, R, T, W, o Y, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre F, L, I, y M.

20 **[0069]** En un caso, el método comprende además:

(b) generar una pluralidad de polipéptidos que comprenden:

(i) CDRL1 que comprende una primera secuencia hipervariable consenso o variante de la misma que comprende una sustitución en una o más posiciones en comparación con una secuencia hipervariable consenso correspondiente;

25 (ii) CDRL2 que comprende una segunda secuencia hipervariable consenso o variante de la misma que comprende una sustitución en una o más posiciones en comparación con una secuencia hipervariable consenso correspondiente; y

30 (iii) CDRL3 que comprende una secuencia de aminoácidos: Q-Q-X1-X2-X3-X4-P-X5-T (SEQ ID NO: \_), en la que X1 está en la posición 91 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S.

**[0070]** En un caso, se proporciona un método de generación de una o más de las secuencias CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, y CDRL3 descritas anteriormente, que comprende:

35 (a) construir un vector de expresión que comprende una secuencia de polinucleótidos que codifica un dominio variable de cadena ligera, un dominio variable de cadena pesada, o ambos de un anticuerpo origen que comprende por lo menos una, dos, tres, cuatro, cinco o todas las CDR del anticuerpo origen seleccionadas del grupo que consiste en CDRL1, CDRL2, CDRL3, CDRH1, CDRH2, y CDRH3; y

(b) mutar por lo menos una, dos, tres, cuatro, cinco o todas las CDR del anticuerpo origen para generar una o más de las regiones hipervariables descritas anteriormente.

40 **[0071]** Se describe aquí un método de selección de un polipéptido que se une a un antígeno diana, que comprende:

(a) generar una composición con una pluralidad de uno o más de los polipéptidos descritos anteriormente;

(b) seleccionar uno o más polipéptidos de la composición que se unen a un antígeno diana;

45 (c) aislar el uno o más polipéptidos que se unen al antígeno diana de los polipéptidos que no se unen al antígeno diana; y

(d) identificar el uno o más polipéptidos que se unen al antígeno diana que presentan una afinidad deseada por el antígeno diana.

**[0072]** Se describe aquí un método de selección de un dominio variable de unión a antígeno que se une a un antígeno diana de una biblioteca de dominios variables de anticuerpo, que comprende:

- (a) poner en contacto una o más de las bibliotecas descritas anteriormente con un antígeno diana;
- (b) separar uno o más polipéptidos que se unen específicamente al antígeno diana de los polipéptidos que no se unen específicamente al antígeno diana, recuperar dichos uno o más polipéptidos que se unen específicamente al antígeno diana e incubar dichos uno o más polipéptidos que se unen específicamente al antígeno diana en una serie de diluciones que comprenden cantidades decrecientes del antígeno diana en una concentración de aproximadamente 0,1 nM a aproximadamente 1000 nM; y
- (c) seleccionar dichos uno o más polipéptidos que se unen específicamente al antígeno diana y que se pueden unir a la concentración más baja del antígeno diana o que presentan una afinidad de aproximadamente 0,1 nM a aproximadamente 200 nM.
- 10 **[0073]** El antígeno diana es HER2. En un aspecto, la concentración del antígeno diana es de aproximadamente 100 a aproximadamente 250 nM. En un aspecto, la concentración de antígeno diana es de aproximadamente 25 a aproximadamente 100 nM. En algunos casos, se criban una o más de las bibliotecas, clones o polipéptidos contra un panel de antígenos que incluyen el antígeno diana. En algunos casos, se seleccionan aquellos clones o polipéptidos que se unen específicamente al antígeno diana y que no reaccionan de manera cruzada de forma sustancial con ninguno de los otros antígenos en el panel. El panel de antígenos puede incluir por lo menos tres y hasta 100 antígenos diferentes. En algunos casos, el panel de antígenos incluye de 3 a 100, 3 a 50, 3 a 25, o 3 a 10 antígenos diferentes.
- 15 **[0074]** En una realización, se proporciona un método de selección de un polipéptido que se une a un antígeno diana de una biblioteca de polipéptidos, que comprende
- 20 (a) aislar uno o más polipéptidos que se une específicamente al antígeno diana mediante el contacto de una biblioteca que comprende una pluralidad de cualquiera de los polipéptidos descritos anteriormente con un antígeno diana inmovilizado bajo condiciones adecuadas para la unión;
- (b) separar dicho uno o más polipéptidos que se unen específicamente al antígeno diana de los polipéptidos que no se unen específicamente al antígeno diana, y recuperar dichos uno o más polipéptidos que se unen específicamente al antígeno diana para obtener una subpoblación enriquecida en uno o más polipéptidos que se unen específicamente al antígeno diana; y
- 25 (c) opcionalmente, repetir las etapas (a)-(b) por lo menos dos veces, utilizando cada repetición la subpoblación enriquecida en dichos uno o más polipéptidos que se unen específicamente al antígeno diana obtenida de la ronda previa de selección tal como se define en las reivindicaciones.
- 30 **[0075]** Se describe aquí un método que comprende además:
- (d) incubar la subpoblación con una concentración de antígeno diana marcado en el intervalo de aproximadamente 0,1 nM a aproximadamente 1000 nM para formar una mezcla bajo condiciones adecuadas para la unión;
- (e) poner en contacto la mezcla con un agente inmovilizado que se une al marcador en el antígeno diana;
- 35 (f) detectar dicho uno o más polipéptidos que se unen específicamente al antígeno diana marcado y recuperar dicho uno o más polipéptidos que se unen específicamente al antígeno diana marcado del antígeno diana marcado; y
- (g) opcionalmente, repetir las etapas (d)-(f) por lo menos dos veces, utilizando cada repetición la subpoblación enriquecida en dichos uno o más polipéptidos que se unen específicamente al antígeno diana obtenida de la ronda previa de selección y utilizando una concentración inferior de antígeno diana marcado que la ronda previa de selección.
- 40 **[0076]** En un aspecto, el método comprende además añadir un exceso de antígeno diana marcado a la mezcla e incubar la mezcla durante un periodo de tiempo suficiente para recuperar uno o más polipéptidos que se unen específicamente al antígeno diana con afinidad baja. En algunos casos, en cualquiera de los métodos aquí descritos, se criban una o más de las bibliotecas, clones o polipéptidos contra un panel de antígenos que incluyen el antígeno diana. En algunos casos, se seleccionan aquellos clones o polipéptidos que se unen específicamente al antígeno diana y que no reaccionan de manera cruzada de forma sustancial con ninguno de los otros antígenos en el panel. El panel de antígenos puede incluir por lo menos tres y hasta 100 antígenos diferentes. En algunos casos, el panel de antígenos incluye de 3 a 100, 3 a 50, 3 a 25, o 3 a 10 antígenos diferentes.
- 45 **[0077]** Se describen aquí un método de aislamiento de uno o más polipéptidos que se unen específicamente a un antígeno diana con afinidad elevada, que comprende:
- 50 (a) poner en contacto una biblioteca que comprende una pluralidad de los polipéptidos descritos anteriormente con un antígeno diana a una concentración de por lo menos aproximadamente 0,1 nM a aproximadamente 1000

nM para aislar uno o más polipéptidos que se unen específicamente al antígeno diana;

(b) recuperar dichos uno o más polipéptidos que se unen específicamente al antígeno diana del antígeno diana para obtener una subpoblación enriquecida en dichos uno o más polipéptidos que se unen específicamente al antígeno diana; y

- 5 (c) opcionalmente, repetir las etapas (a) y (b) por lo menos dos veces, utilizando cada repetición la subpoblación obtenida de la ronda previa de selección y utilizando una concentración disminuida de antígeno diana de la utilizada en la ronda previa para aislar uno o más polipéptidos que se unen específicamente al antígeno diana a la concentración más baja del antígeno diana.

10 **[0078]** Se describe aquí un ensayo para seleccionar uno o más polipéptidos que se unen a un antígeno diana de una biblioteca que comprende una pluralidad de cualquiera de los polipéptidos descritos anteriormente, que comprende:

15 (a) poner en contacto la biblioteca con una concentración de antígeno diana marcado a un intervalo de concentración de aproximadamente 0,1 nM a aproximadamente 1000 nM bajo condiciones adecuadas para la formación de uno o más complejos entre el antígeno diana marcado y uno o más polipéptidos que se unen específicamente al antígeno diana;

(b) aislar dichos uno o más complejos y separar dichos uno o más polipéptidos que se unen específicamente al antígeno diana del antígeno diana marcado para obtener una subpoblación enriquecida en dichos uno o más polipéptidos que se unen específicamente al antígeno diana; y

20 (c) opcionalmente, repetir las etapas (a) y (b), por lo menos dos veces, utilizando cada vez la subpoblación obtenida de la ronda previa de selección y utilizando una concentración inferior de antígeno diana que la utilizada en la ronda previa.

25 **[0079]** En un aspecto, el ensayo comprende además añadir un exceso de antígeno diana no marcado al uno o más complejos. En un aspecto, las etapas (a) y (b) se repiten dos veces, en las que la concentración de antígeno diana en la primera ronda de selección es aproximadamente de 100 nM a aproximadamente 250 nM, en las que la concentración de antígeno diana en la segunda ronda de selección es de aproximadamente 25 nM a aproximadamente 100 nM, y en las que la concentración de antígeno diana en la tercera ronda de selección es de aproximadamente 0,1 nM a aproximadamente 25 nM.

**[0080]** Se describe aquí un método de cribado de una biblioteca que comprende una pluralidad de cualquiera de los polipéptidos descritos anteriormente que comprende:

30 (a) incubar una primera muestra de la biblioteca con un antígeno diana bajo condiciones adecuadas para la unión de los polipéptidos al antígeno diana;

(b) incubar una segunda muestra de la biblioteca en ausencia de un antígeno diana;

(c) poner en contacto cada uno de la primera muestra y la segunda muestra con antígeno diana inmovilizado bajo condiciones adecuadas para la unión del polipéptido al antígeno diana inmovilizado;

35 (d) detectar el polipéptido unido a antígeno diana inmovilizado para cada muestra; y

(e) determinar la afinidad del polipéptido para el antígeno diana mediante el cálculo de la proporción de las cantidades de polipéptido unido de la primera muestra sobre la cantidad de polipéptido unido de la segunda muestra.

40 **[0081]** El antígeno diana es HER-2. En un aspecto, la concentración del antígeno diana es de aproximadamente 100 a aproximadamente 250 nM. En un aspecto, la concentración de antígeno diana es de aproximadamente 25 a aproximadamente 100 nM. En algunos casos, se criban una o más de las bibliotecas, clones o polipéptidos contra un panel de antígenos que incluyen el antígeno diana. En algunos casos, se seleccionan aquellos clones o polipéptidos que se unen específicamente al antígeno diana y que no reaccionan de manera cruzada de forma sustancial con ninguno de los otros antígenos en el panel. El panel de antígenos puede incluir por lo menos tres y hasta 100 antígenos diferentes. En algunos casos, el panel de antígenos incluye de 3 a 100, 3 a 50, 3 a 25, o 3 a 10 antígenos diferentes.

45 **[0082]** Los métodos aquí descritos también proporcionan el aislamiento de un anticuerpo anti-HER-2. Un anticuerpo anti-HER-2 puede comprender un dominio variable de cadena pesada, en el que (i) CDRH1 comprende la secuencia de aminoácidos G-FX1-I-X2-X3-X4-XS-I-H (SEQ ID NO:22); en la que X1 está en la posición 28 según el sistema de numeración Kabat y se selecciona entre S e Y; X2 se selecciona entre S e Y; X3 se selecciona entre S e Y; X4 se selecciona entre S e Y; y X5 se selecciona entre S e Y; (ii) CDRH2 comprende una secuencia de aminoácidos de X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23); en la que X1 está en la posición de aminoácido 50 según el sistema de numeración Kabat y se selecciona entre S e Y; X2 se selecciona entre S e Y; X3 se selecciona entre S e Y; X4 se selecciona entre S e Y; X5 se selecciona entre S e

Y; y X6 se selecciona entre S e Y; y (iii) CDRH3 comprende una secuencia de aminoácidos X1-X2-X3-X4-X5-X6-X7-D-Y (SEQ ID NO:582), en la que X1 está en la posición de aminoácido 95 según el sistema de numeración Kabat y se selecciona entre Y y R; X2 se selecciona entre Y, S y R; X3 se selecciona entre S, G, Y y H; X4 se selecciona entre S, G, Y y R; X5 se selecciona entre G y A; X6 se selecciona entre F, M, L, y A; y X7 se selecciona entre F, M, y L o está ausente.

**[0083]** En otro caso un anticuerpo anti-HER-2 comprende un dominio variable de cadena pesada de inmunoglobulina, en el que (i) CDRH1 comprende la secuencia de aminoácidos G-F-XI-I-X2-X3-X4-X5-1-H (SEQ ID NO:22); en la que X1 está en la posición 28 según el sistema de numeración Kabat y se selecciona entre S e Y; X2 se selecciona entre S e Y; X3 se selecciona entre S e Y; X4 se selecciona entre S e Y; y X5 se selecciona entre S e Y; (ii) CDRH2 comprende una secuencia de aminoácidos de X1-IX2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23); en la que X1 está en la posición de aminoácido 50 según el sistema de numeración Kabat y se selecciona entre S e Y; X2 se selecciona entre S e Y; X3 se selecciona entre S e Y; X4 se selecciona entre S e Y; y X5 se selecciona entre S e Y; y (iii) CDRH3 comprende una secuencia de aminoácidos X1-X2-X3-X4-X5-X6-X7-X8-D-Y (SEQ ID NO:583), en la que X1 está en la posición de aminoácido 95 según el sistema de numeración Kabat y se selecciona entre Y, S y G; X2 se selecciona entre Y, S, G, R, A, y M; X3 se selecciona entre G, Y, S y R; X4 se selecciona entre G, Y y F; X5 se selecciona entre Y, S, N, y G; X6 se selecciona entre Y, R, H y W; X7 se selecciona entre G y A; y X8 se selecciona entre F, M, L y I.

**[0084]** En otro caso un anticuerpo anti-HER-2 comprende un dominio variable de cadena pesada de inmunoglobulina, en el que (i) CDRH1 comprende la secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ID NO:22); en la que X1 está en la posición 28 según el sistema de numeración Kabat y se selecciona entre S e Y; X2 se selecciona entre S e Y; X3 se selecciona entre S e Y; X4 se selecciona entre S e Y; y X5 se selecciona entre S e Y; (ii) CDRH2 comprende una secuencia de aminoácidos de X1-IX2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23); en la que X1 está en la posición de aminoácido 50 según el sistema de numeración Kabat y se selecciona entre S e Y; X2 se selecciona entre S e Y; X3 se selecciona entre S e Y; X4 se selecciona entre S e Y; y X5 se selecciona entre S e Y; y (iii) CDRH3 comprende una secuencia de aminoácidos de X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:584), en la que X1 está en la posición de aminoácido 95 y se selecciona entre Y, S, R, G y E; X2 se selecciona entre Y, S, R y G; X3 se selecciona entre S, Y, G y W; X4 se selecciona entre S, Y, G y Q; X5 se selecciona entre G, Y y S; X6 se selecciona entre G, Y, S, R y V; X7 se selecciona entre S, Y, G y R; X8 se selecciona entre Y, S, G, P, P y V; X9 se selecciona entre G, A, Y, S y R; X10 se selecciona entre M, F, G, Y, S y R; X11 se selecciona entre A, Y, S, G y R o no está presente; X12 se selecciona entre I, M, F, L, A, G, S, Y, R, y T o no está presente; X13 se selecciona entre F, M, L, G, A, Y, T, y S o no está presente; X14 se selecciona entre L, M, F, I, G, Y, A, y T o no está presente; X15 se selecciona entre M, L, Y, G y R o no está presente; X16 se selecciona entre Y y G o no está presente; X17 se selecciona entre R, M, t G o no está presente; X18 se selecciona entre P t A o no está presente; t X18 es L o no está presente.

**[0085]** En otro caso, un anticuerpo anti-HER-2 comprende un dominio variable de cadena pesada de inmunoglobulina, en el que: (i) CDRH1 comprende una secuencia de aminoácidos G-F-X1-1-X2-X3-X4-X5-1-H (SEQ ID NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S; (ii) CDRH2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-YA-D-S-V-K-G (SEQ ID NO:23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y (iii) CDRH3 comprende la secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:24), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 50% Y, 25% S, y 25% G; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 50% Y, 25% S, y 25% G, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.

**[0086]** En otro caso, un anticuerpo anti-HER-2 comprende un dominio variable de cadena pesada de inmunoglobulina, en el que: (i) CDRH1 comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ED NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S; (ii) CDRH2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-YA-D-S-V-K-G (SEQ ID NO:23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y (iii) CDRH3 comprende la secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:26), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de



aminoácidos en una proporción molar de 25% Y, 50% S, y 25% R; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 25% Y, 50% S, y 25% R, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.

5 **[0087]** En otro caso un anticuerpo anti-HER-2 comprende un dominio variable de cadena pesada de inmunoglobulina, en el que: (i) CDRH1 comprende una secuencia de aminoácidos G-F-XI-I-X2-X3-X4-XS-I-H (SEQ ID NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S; (ii)  
10 CDRH2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y (iii)  
15 CDRH3 comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:27), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se selecciona de un grupo de aminoácidos en una proporción molar de 38% Y, 25% S, 25% G, y 12% R; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 38% Y, 25% S, 25% G, y 12% R, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.  
20

**[0088]** En otro caso un anticuerpo anti-HER-2 comprende un dominio variable de cadena pesada de inmunoglobulina, en el que: (i) CDRH1 comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-XS-I-H (SEQ ID NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S; (ii)  
25 CDRH2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-KG (SEQ ID NO:23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y (iii)  
30 CDRH3 comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:28), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 20% Y, 26% S, 26% G, 13% R, 1% A, 1% D, 1% E, 1% F, 1% H, 1% I, 1% K, 1% L, 1% M, 1% N, 1% P, 1% Q, 1% T, 1% V, y 1% W; en la que los aminoácidos en cada una de las  
35 posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 20% Y, 26% S, 26% G, 13% R, 1% A, 1% D, 1% E, 1% F, 1% H, 1% I, 1% K, 1% L, 1% M, 1% N, 1% P, 1% Q, 1% T, 1% V, y 1% W, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F.

**[0089]** En otro caso un anticuerpo anti-HER-2 comprende un dominio variable de cadena pesada de inmunoglobulina, en el que: (i) CDRH1 comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ DD NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S; (ii)  
40 CDRH2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y (iii)  
45 CDRH3 que comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-x18-X19 (SEQ ID NO: ), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X19 se seleccionan entre S y uno de A, C, F, G, I, L, N, P, R, T, W, o Y, o no están presentes; en la que X18 se selecciona entre G y A o no está presente; y en la que X19 se selecciona entre F, L, I, y M o no está presente.  
50

**[0090]** En otro caso un anticuerpo anti-HER-2 comprende un dominio variable de cadena pesada de inmunoglobulina, en el que: (i) CDRH1 comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-XS-I-H (SEQ ID NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S; (ii)  
55 CDRH2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO:23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona Y y S; y en la que X6 se selecciona entre Y y S; y (iii) CDRH3  
60 comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19 (SEQ ID NO: ), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X19 se seleccionan entre S y uno de Y, W, R, o F, o no

están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre F, L, I, y M.

**[0091]** En algunos casos, el anticuerpo anti-HER-2 puede comprender un CDRH1 que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOS:189 a 294 tal como se muestra en la figura 11. El anticuerpo anti-HER-2 puede comprender también un CDRH2 que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOS:295 a 400 tal como se muestra en la figura 11. El anticuerpo anti-HER-2 también puede comprender un CDRH3 que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOS:401-506 tal como se muestra en la figura 11. En algunas realizaciones, el anticuerpo anti-HER-2 puede comprender un CDRH que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOS:816-842 tal como se muestra en la figura 21 A. El anticuerpo anti-HER-2 también puede comprender un CDRH2 que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOS:843-869 tal como se muestra en la figura 21A. El anticuerpo anti-HER-2 también puede comprender un CDRH3 que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOS:870-896 tal como se muestra en la figura 21A. En algunos casos, el anticuerpo anti-HER-2 puede comprender un CDRH1 que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOS:924-950 tal como se muestra en la figura 24A. El anticuerpo anti-HER-2 puede comprender también un CDRH2 que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOS:951-977 tal como se muestra en la figura 24A. El anticuerpo anti-HER-2 puede comprender también un CDRH3 que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOS:978-1004 tal como se muestra en la figura 24A.

**[0092]** En un caso, un anticuerpo que se une específicamente a HER2 humano comprende las secuencias CDRH1, CDRH2, CDRH3, y CDRL3 que corresponde con las secuencias CDRH1, CDRH2, CDRH3, y CDRL3 establecidas en la figura 11 para cualquiera de los Fabs 1-106. En un caso, un anticuerpo que se une específicamente a HER2 humano comprende las secuencias CDR1, CDRH2, CDRH3, y CDRL3 correspondiente a las secuencias CDRH1, CDRH2, CDRH3, y CDRL3 establecidas en la figura 21A para cualquiera de los clones B1-B28. En otro caso, un anticuerpo que se une específicamente a HER2 humano comprende las secuencias CDRH21, CDRH2, CDRH3, y CDRL3 correspondientes a las secuencias CDRH1, CDRH2, CDRH3, y CDRL3 establecidas en la figura 24A para cualquiera de los clones G29-G61.

**[0093]** En algunos casos, un anticuerpo anti-HER-2 comprende un CDRH1 que comprende una secuencia de aminoácidos GFSIX2X3SYIH (SEQ ED NO:588), en la que X2 y X3 son Y o S. Un anticuerpo anti-HER-2 puede comprender también un CDRH2 que comprende una secuencia de aminoácidos SIYPX3SGYTSYADSKVG (SEQ ID NO:589), en la que X3 es Y o S. Un anticuerpo anti-HER-2 puede comprender también un dominio variable de cadena ligera que comprende un CDRL3 que comprende una secuencia de aminoácidos QQSYYX4PST (SEQ ID NO:587), en la que X4 es Y o S.

**[0094]** En algunos casos, un anticuerpo anti-HER-2 comprende una secuencia de aminoácidos GFX1ISYSSIH (SEQ ID NO: 590), en la que X1 es Y o S. Un anticuerpo anti-HER-2 puede comprender además un CDRH2 que comprende una secuencia de aminoácidos SIYPX3YGXSTX6YADSKVG (SEQ ID NO:591), en la que X3, X5 y X6 son Y o S.

**[0095]** En otro caso un anticuerpo anti-HER-2 comprende un CDRH1 que tiene una secuencia de aminoácidos GFXIISSSSIH (SEQ ID NO:593), en la que X1 es Y o S. Un anticuerpo anti-HER-2 puede comprender también un CDRH2 que tiene una secuencia de aminoácidos X1IX2PSSGYTX6YADSKVG (SEQ ID NO:594), en la que X1, X2 y X6 son Y o S. Un anticuerpo anti-HER-2 puede comprender también un CDRH3 que tiene una secuencia de aminoácidos XIX2X3X4YYSYYX10GX12X13X14DY (SEQ ID NO:592), en la que X1 se selecciona entre Y, S and R; X2 se selecciona entre Y y S; X3 se selecciona entre G, Y y S; X4 se selecciona entre G, Y y S; X4 se selecciona entre Y, S, R y G; X10 se selecciona entre Y, S y G; X12 se selecciona entre Y, S, G y R; X13 se selecciona entre G y A y X14 se selecciona entre I, F, M y L.

**[0096]** En algunos casos, el anticuerpo anti-HER-2 puede comprender opcionalmente además un dominio variable de cadena ligera que comprende una secuencia CDRL3, en la que CDRL3 comprende una secuencia de aminoácidos de Q-Q-X1X2-X3-X4-P-X5-T (SEQ ID NO:25), en la que X1 está en la posición 91 según el sistema de numeración Kabat y se selecciona entre S e Y; X2 se selecciona entre S, Y y F; X3 se selecciona entre Y, S y F; X4 se selecciona entre Y y S; X5 se selecciona entre S e Y. El dominio variable de cadena ligera puede comprender también un CDRL3 que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOS:83 a 188 tal como se muestra en la figura 11, SEQ ID NOS:789-815 en la figura 21A, y SEQ ID NOS: 897-923 en la figura 24A. El anticuerpo también puede comprender además un CDRL1 que comprende una secuencia de aminoácidos RASQDVNTAVA (SEQ ID NO:29). El anticuerpo puede comprender además un CDRL2 que comprende una secuencia de aminoácidos SASSLYS (SEQ ID NO:30).

**[0097]** En un caso, el polipéptido es un anticuerpo que se une específicamente a HER2. En dicho aspecto, el anticuerpo comprende regiones de armazón del anticuerpo 4D5. En dicho aspecto, el anticuerpo comprende las regiones de armazón de una variante del anticuerpo 4D5. En dicho aspecto, el anticuerpo es un anticuerpo monoclonal. En dicho aspecto, el anticuerpo es un anticuerpo biespecífico.

**[0098]** En un caso, se proporciona un polinucleótido aislado que codifica cualquiera de los anticuerpos descritos anteriormente que se unen específicamente a HER2. En un caso, se proporciona un vector que comprende un polinucleótido aislado que codifica cualquiera de los anticuerpos descritos anteriormente que se unen específicamente a HER2. En un caso, se proporciona una célula huésped transformada con un vector que comprende un polinucleótido aislado que codifica cualquiera de los anticuerpos descritos anteriormente que se unen específicamente a HER2. En un caso, se proporciona un proceso de producción de un anticuerpo, que comprende cultivar una célula huésped transformada con un vector que comprende un polinucleótido aislado que codifica cualquiera de los anticuerpos descritos anteriormente que se unen específicamente a HER2, de manera que se expresa el polinucleótido. En una forma, el proceso comprende además recuperar el anticuerpo del cultivo de células huésped. En un aspecto, el proceso comprende además recuperar el anticuerpo del medio de cultivo de células huésped.

**[0099]** En un caso, se proporciona un método de utilización de uno o más de los anticuerpos descritos anteriormente que se unen específicamente a HER2 para tratar un trastorno relacionado con HER2, que comprende la etapa de administrar dicho uno o más anticuerpos al mamífero. En otro caso, el tratamiento comprende además la etapa de administrar un segundo agente terapéutico de manera simultánea o secuencial con el anticuerpo. En dicho caso, el segundo agente terapéutico se selecciona entre un agente anti-angiogénico, un agente anti-neoplásico, un agente quimioterapéutico, y un agente citotóxico.

**[0100]** En un caso, se proporciona un método de tratamiento de un mamífero que padece o presenta el riesgo de desarrollar un trastorno relacionado con HER2, que comprende la etapa de tratamiento de un mamífero con uno o más Fabs de uno o más de los anticuerpos descritos anteriormente que se unen específicamente a HER2. En un caso, uno o más de los polipéptidos descritos anteriormente se unen específicamente a HER2.

**[0101]** En un caso, un polipéptido de la memoria comprende por lo menos uno, o ambos, de dominios variables de anticuerpo de cadena pesada y cadena ligera, donde el dominio variable de anticuerpo comprende una, dos o tres variantes de CDR tal como se describe aquí (por ejemplo, tal como se ha descrito anteriormente).

**[0102]** En algunos casos un polipéptido de la invención (en particular, aquellos que comprenden un dominio variable de anticuerpo) comprende además una secuencia de armazón de anticuerpo, por ejemplo, FR1, FR2, FR3 y/o FR4 para un dominio variable de anticuerpo correspondiente a la variante de CDR, las secuencias de FR obtenidas de una única plantilla de anticuerpo. En un caso, las secuencias FR se obtienen de un anticuerpo humano. En un caso, las secuencias FR se obtienen de una secuencia consenso humana (por ejemplo, secuencia consenso del subgrupo III). En un caso, las secuencias de armazón comprenden una secuencia consenso modificada tal como se describe aquí (por ejemplo, que comprende las modificaciones en la posición 49, 71, 93 y/o 94 en la cadena pesada, y/o en la posición 66 en la cadena ligera). En un caso, las regiones armazón presentan las secuencias de las regiones armazón de la cadena ligera y la cadena pesada del anticuerpo 4D5-8 humanizado de tipo natural (mostrado en la figura 16 (SEQ ID NOS: 6-9 y 10-13, respectivamente)). En un caso, las regiones armazón presentan las secuencias de las regiones armazón de una versión variante de la cadena ligera y la cadena pesada del anticuerpo 4D5-8 humanizado, donde la cadena ligera está modificada en la posición 66 y la cadena pesada está modificada en las posiciones 71, 73, y 78 (mostrado en la figura 17 (SEQ ID NOS: 14-17 y 18-21)).

**[0103]** En algunos casos, un polipéptido de la invención comprende un dominio variable de anticuerpo de cadena ligera y cadena pesada, en el que el dominio variable de cadena ligera comprende por lo menos 1, 2 ó 3 CDR variantes seleccionadas del grupo que consiste en CDR L1, L2 y L3, y el dominio variable de cadena pesada comprende por lo menos 1, 2 o 3 variantes de CDR seleccionadas del grupo que consiste en CDR H1, H2 y H3.

**[0104]** En algunos casos, un polipéptido de la invención es un ScFv. En algunas realizaciones, es un fragmento Fab. En algunas realizaciones, es un  $F(ab)_2$  o  $F(ab')_2$ . Por consiguiente, en algunos casos, un polipéptido de la invención comprende además un dominio de dimerización. En algunos casos, el dominio de dimerización se localiza entre un dominio variable de cadena pesada o cadena ligera de anticuerpo y, por lo menos una parte de una proteína de cubierta viral. El dominio de dimerización puede comprender una secuencia de dimerización y/o una secuencia que comprende uno o más residuos de cisteína. El dominio de dimerización puede estar unido, directa o indirectamente, al extremo C-terminal de un dominio variable o constante de cadena pesada. La estructura del dominio de dimerización puede variar dependiendo de si el dominio variable de anticuerpo se produce como un componente de proteína de fusión con el componente proteína de cubierta viral (sin un codón de parada amber después del dominio de dimerización) o si el dominio variable de anticuerpo se produce predominantemente sin el componente proteína de cubierta viral (por ejemplo, con un codón de parada amber después del dominio de dimerización). Cuando el dominio variable de anticuerpo se produce predominantemente como una proteína de fusión con un componente proteína de cubierta viral, uno o más enlaces disulfuro y/o una única secuencia de dimerización proporcionan una expresión bivalente. Para dominios variables de anticuerpo producidos predominantemente sin fusionarse a un componente proteína de cubierta viral (por ejemplo, con parada amber), es preferible, aunque no necesario, tener un dominio de dimerización que comprende tanto un residuo de cisteína como una secuencia de dimerización. En algunos casos, las cadenas pesadas del  $F(ab)_2$  se dimerizan en un dominio de dimerización que no incluye la región bisagra. El dominio de dimerización puede comprender una secuencia de cremalleras de leucina (por ejemplo, una secuencia GCN4, tal como

GRMKQLEDKVEELLSKNY14LENEVARLKKLVGERG (SEQ ID NO: 3)).

**[0105]** En algunos casos, un polipéptido comprende además un dominio constante de cadena ligera fusionado a un dominio variable de cadena ligera, que, en algunos casos, comprende por lo menos una, dos o tres variantes de CDR. En algunos casos de polipéptidos, el polipéptido comprende un dominio constante de cadena pesada fusionado a un dominio variable de cadena pesada, que en algunos casos, comprende por lo menos una, dos o tres variantes de CDR.

**[0106]** En algunos casos, puede ser preferible mutar el residuo de armazón de manera que es una variante con respecto a un polipéptido de referencia o anticuerpo origen. Por ejemplo, el residuo 71 del armazón de la cadena pesada puede ser el aminoácido R, V o A. En otro ejemplo, el residuo 93 del armazón de la cadena pesada puede ser el aminoácido S o A. En otro ejemplo, el residuo 94 del armazón de la cadena pesada puede ser el aminoácido R, K o T o codificado por MRT. En otro ejemplo, el residuo de armazón 49 de la cadena pesada puede ser el aminoácido A o G. Los residuos de armazón en la cadena ligera también se pueden mutar. Por ejemplo, el residuo 66 de armazón en la cadena ligera puede ser el aminoácido R o G.

**[0107]** Tal como se describe aquí, una variante de CDR se refiere a una CDR con una varianza de secuencia en comparación con la correspondiente CDR de un único polipéptido de referencia/anticuerpo origen. Por consiguiente, una variante de CDR se refiere a una CDR con una varianza de secuencia en comparación con la correspondiente CDR de un único polipéptido de referencia/anticuerpo de origen. Por consiguiente, las CDR de un único polipéptido de la invención, en ciertos casos, pueden corresponder al grupo de CDR de un único polipéptido de referencia o anticuerpo origen. Los polipéptidos pueden comprender cualquiera de las variantes de CDR o combinaciones. Por ejemplo, un polipéptido puede comprender una variante de CDRH1 y una variante de CDRH2. Un polipéptido puede comprender una variante de CDRH1, una variante de CDRH2 y una variante de CDRH3. En otro ejemplo, un polipéptido puede comprender una variante de CDRH1, una variante de CDRH2, una variante de CDRH3 y una variante de CDRL3. En otro ejemplo, un polipéptido comprende una variante de CDRL1, una variante de CDRL2 y una variante de CDRL3. Cualquier polipéptido puede comprender adicionalmente una variante de CDRL3. Cualquier polipéptido puede comprender también una variante de CDRH3.

**[0108]** En un caso, un polipéptido comprende una o más secuencias de variantes de CDR representadas en las figuras 7, 19 y 22. En un caso, un polipéptido de la invención comprende una o más secuencias de variantes de CDR representadas en la figura 11A. En un caso, un polipéptido de la invención comprende una o más secuencias de variantes de CDR representadas en la figura 15. En otro caso, un polipéptido comprende una o más secuencias de variantes de representadas en las figuras 21A-21B.

**[0109]** Los polipéptidos pueden estar en un complejo con otro. Por ejemplo, la presente invención proporciona un complejo de polipéptidos que comprende dos polipéptidos, en el que cada polipéptido es un polipéptido, y en el que uno de dichos polipéptidos comprende por lo menos una, dos o todas las variantes de CDR H1, H2 y H3, y el otro polipéptido comprende una variante de CDR de cadena ligera (por ejemplo, CDR L3). Un complejo de polipéptidos puede comprender un primer y segundo polipéptido en el que el primer polipéptido comprende por lo menos una, dos o tres variantes de CDR de cadena ligera, y el segundo polipéptido comprende por lo menos una, dos o tres variantes de CDR de cadena pesada. Se describen aquí complejos de polipéptidos que comprenden las mismas secuencias variantes de CDR. La formación de complejos puede estar mediada por cualquier técnica adecuada, incluyendo mediante la dimerización/multimerización en un dominio de dimerización/multimerización, tal como aquellos aquí descritos o interacciones covalentes (tales como a través de un enlace disulfuro) (que, en algunos contextos, es parte de un dominio de dimerización, por ejemplo, un dominio de dimerización puede contener una secuencia de cremalleras de leucinas y una cisteína).

**[0110]** Se describen aquí composiciones que comprenden polipéptidos y/o polinucleótidos de la invención. Por ejemplo, se describe aquí una composición que comprende una pluralidad de cualquiera de los polipéptidos. Dicha pluralidad puede comprender polipéptidos codificados por una pluralidad de polinucleótidos generados utilizando un grupo de oligonucleótidos que comprenden la degeneración en la secuencia que codifica una variante de aminoácido, donde dicha degeneración es la de las múltiples secuencias de codones del grupo limitado de codones que codifica la variante de aminoácido. Una composición que comprende un polinucleótido o polipéptido o biblioteca puede estar en forma de un kit o un artículo de fabricación (opcionalmente empaquetado con instrucciones, tampones, etc.).

**[0111]** Se describe aquí un polinucleótido que codifica un polipéptido tal como se describe aquí. También se describe un vector que comprende una secuencia que codifica un polipéptido. El vector puede ser, por ejemplo, un vector de expresión replicable (por ejemplo, el vector de expresión replicable puede ser el fago M13, f1, fd, Pf3 o un derivado del mismo, o un fago lambda, tal como lambda, 21, phi80, phi81, 82, 424, 434, etc., o un derivado del mismo). El vector puede comprender una región promotora unida a la secuencia que codifica un polipéptido de la invención. El promotor puede ser cualquier adecuado para la expresión del polipéptido, por ejemplo, el sistema de promotores lac Z, el promotor de fosfatasa alcalina pho A (Ap), el promotor de bacteriófago 1<sub>PL</sub> (un promotor sensible a la temperatura), el promotor tac, el promotor de triptófano, y el promotor del bacteriófago T7. Se describe aquí un vector que comprende un promotor seleccionado del grupo que consiste

en los sistemas promotores anteriores.

**[0112]** Los polipéptidos se pueden expresar en cualquier forma adecuada según las necesidades y deseo del técnico. Por ejemplo, un polipéptido de la invención se puede expresar en una superficie viral, por ejemplo, un fago o una partícula viral fagémida. Se describen aquí partículas virales que comprenden un polipéptido y/o polinucleótido que codifica un polipéptido.

**[0113]** Se describe aquí una población que comprende una pluralidad de polipéptido o polinucleótido donde cada tipo de polipéptido o polinucleótido es tal como se describe aquí.

**[0114]** En algunos casos, los polipéptidos y/o polinucleótidos se proporcionan como una biblioteca, por ejemplo, una biblioteca que comprende una pluralidad de por lo menos aproximadamente  $1 \times 10^4$ ,  $1 \times 10^5$ ,  $1 \times 10^6$ ,  $1 \times 10^7$ ,  $1 \times 10^8$  secuencias distintas de polipéptidos y/o polinucleótidos de la invención. También se describe aquí una biblioteca que comprende una pluralidad de los virus o las partículas virales de la invención, expresando cada virus o partícula viral un polipéptido de la invención. Una biblioteca puede comprender virus o partículas virales que expresan cualquier cantidad de polipéptidos (secuencias) distintos, por ejemplo, por lo menos aproximadamente  $1 \times 10^4$ ,  $1 \times 10^5$ ,  $1 \times 10^6$ ,  $1 \times 10^7$ ,  $1 \times 10^8$  polipéptidos distintos.

**[0115]** Se describen aquí células huésped que comprenden un polinucleótido o vector que comprende una secuencia que codifica un polipéptido.

**[0116]** Se describen aquí métodos para seleccionar enlazadores de afinidad elevada a antígenos diana específicos. El antígeno diana específico es HER2.

**[0117]** Los métodos de la invención proporcionan poblaciones de polipéptidos (por ejemplo, bibliotecas de polipéptidos (por ejemplo, dominios variables de anticuerpo) con una o más regiones CDR diversificadas tal como se definen en las reivindicaciones. Estas bibliotecas se clasifican (seleccionan) y/o criban para identificar enlazadores de afinidad elevada para un antígeno diana. En un aspecto, los enlazadores de polipéptidos de la biblioteca se seleccionan para la unión a HER2 y para afinidad. Los enlazadores de polipéptido seleccionados utilizando una o más de estas estrategias de selección, se pueden cribar a continuación por la afinidad y/o la especificidad (unión sólo a antígeno diana y no a antígenos no dianas).

**[0118]** En un aspecto, un método de la invención se refiere a generar una pluralidad de polipéptidos con una o más regiones CDR diversificadas, clasificar la pluralidad de polipéptidos para enlazadores para HER2 mediante el contacto de la pluralidad de polipéptidos con HER2 bajo condiciones adecuadas para la unión; separar los enlazadores a HER2 de aquellos que no se unen; aislar los enlazadores; e identificar los enlazadores de afinidad elevada (o cualquier enlazador que presenta una afinidad de unión deseada). La afinidad de los enlazadores que se unen a Her 2 se puede determinar utilizando una variedad de técnicas conocidas en el sector, por ejemplo, ELISA de competición tal como se describen aquí. Opcionalmente, los polipéptidos se pueden fusionar a un polipéptido tag, tal como gD, poly his o FLAG, que se pueden utilizar para clasificar enlazadores en combinación con la clasificación para HER2.

**[0119]** Otra realización proporciona un método de aislar o seleccionar un dominio variable de anticuerpo que se une a HER2 de una biblioteca de dominios variables de anticuerpo, comprendiendo dicho método: a) poner en contacto una población que comprende una pluralidad de polipéptidos de la invención con un antígeno diana inmovilizado bajo condiciones adecuadas para la unión para aislar los enlazadores de polipéptido de antígeno diana; b) separar los enlazadores de polipéptido de los no enlazadores y eluir los enlazadores del antígeno diana; c) opcionalmente, repetir las etapas a-b menos una vez (en algunas realizaciones, por lo menos dos veces).

**[0120]** En algunos casos, un método puede comprender también: d) incubar los enlazadores de polipéptido con una concentración de antígeno diana marcado en el intervalo de 0,1 nM a 1000 nM bajo condiciones adecuadas para la unión para formar una mezcla; e) poner en contacto la mezcla con un agente inmovilizado que se une al marcador en el antígeno diana; f) eluir los enlazadores de polipéptido del antígeno diana marcado; g) opcionalmente, repetir las etapas d) a f) por lo menos una vez (en algunos casos, por lo menos dos veces), utilizando una concentración sucesivamente inferior de antígeno diana marcado cada vez. Opcionalmente, el método puede comprender añadir un exceso de antígeno diana no marcado a la mezcla e incubar durante un periodo de tiempo suficiente para eluir enlazadores de afinidad baja del antígeno diana marcado.

**[0121]** Se describe aquí un método de aislar o seleccionar enlazadores de afinidad elevada (o enlazadores que tienen una afinidad de unión deseada) a un antígeno diana. En una realización, dicho método comprende: a) poner en contacto una población que comprende una pluralidad de polipéptidos con un antígeno diana, donde el antígeno se dispone a una concentración en el intervalo de aproximadamente 0,1 nM a 1000 nM para aislar enlazadores de polipéptido al antígeno diana; b) separar los enlazadores de polipéptido del antígeno diana; c) opcionalmente, repetir las etapas a-b por lo menos una vez (en algunas realizaciones, por lo menos dos veces), cada vez con una concentración sucesivamente inferior de antígeno diana para aislar enlazadores de polipéptido que se unen a la concentración más baja de antígeno diana; d) seleccionar el enlazador de polipéptido que se une a la concentración inferior del antígeno diana para afinidad elevada (o cualquier afinidad deseada) mediante

la incubación de los enlazadores de polipéptido con varias diluciones diferentes del antígeno diana y determinar la IC50 del enlazador de polipéptido; y e) identificar un enlazador de polipéptido que presenta una afinidad deseada para el antígeno diana. Dicha afinidad puede ser, por ejemplo, aproximadamente 0,1 nM a 200 nM, 0,5 nM a 150 nM, 1 nM a 100 nM, y/o 25 nM a 75 nM.

5 **[0122]** Se describe aquí un ensayo para aislar o seleccionar enlazadores de polipéptido que comprende (a) poner en contacto una población que comprende una pluralidad de polipéptidos con un antígeno diana marcado, donde el antígeno diana marcado se dispone a una concentración en un intervalo de 0,1 nM a 1000 nM, bajo condiciones adecuadas para la unión para formar un complejo de un enlazador de polipéptido y el antígeno diana marcado; b) aislar los complejos y separar el enlazador de polipéptido del antígeno diana marcado; c) opcionalmente, repetir las etapas a-b por lo menos una vez, cada vez utilizando una concentración inferior de antígeno diana. Opcionalmente, el método puede comprender también poner en contacto el complejo de enlazador de polipéptido y antígeno diana con un exceso de antígeno diana no marcado. En un caso, las etapas del método se repiten dos veces y la concentración de diana en una primera ronda de selección se encuentra en el intervalo de aproximadamente 100 nM a 250 nM, y, en una segunda ronda de selección (si se realiza) se encuentra en el intervalo de aproximadamente 25 nM a 100 nM, y en la tercera ronda de selección (si se realiza) se encuentra en el intervalo de aproximadamente 0,1 nM a 25 nM.

10 **[0123]** Se describe aquí un método de cribado de una población que comprende una pluralidad de polipéptidos de la invención, comprendiendo dicho método: a) incubar una primera muestra de la población de polipéptidos con un antígeno diana bajo condiciones adecuadas para la unión de los polipéptidos al antígeno diana; b) someter una segunda muestra de la población de polipéptidos a una incubación similar pero en ausencia del antígeno diana; (c) poner en contacto cada una de la primera y segunda muestra con antígeno diana inmovilizado bajo condiciones adecuadas para la unión de los polipéptidos al antígeno diana inmovilizado; d) detectar la cantidad de polipéptidos unidos a antígeno diana inmovilizado para cada muestra; e) determinar la afinidad de un polipéptido particular para el antígeno diana mediante el cálculo de la proporción de la cantidad del polipéptido particular que está unido en la primera muestra sobre la cantidad del polipéptido particular que esta unido en la segunda muestra.

**[0124]** Las bibliotecas generadas tal como se describe aquí también se pueden cribar para la unión a una diana específica y por la falta de unión a antígenos no diana.

20 **[0125]** Se describen aquí un método de cribado de un polipéptido, tal como un dominio variable de anticuerpo, que se une a un antígeno diana específico de una biblioteca de dominios variables de anticuerpo, comprendiendo dicho método: a) generar una población que comprende una pluralidad de polipéptidos de la invención; b) poner en contacto la población de polipéptidos con un antígeno diana bajo condiciones adecuadas para la unión; c) separar un polipéptido enlazador en la biblioteca de polipéptidos no enlazadores; d) identificar un polipéptido enlazador específico de antígeno diana determinando si el polipéptido enlazador se une a un antígeno no diana; y e) aislar un polipéptido enlazador específico de antígeno diana. En algunos casos, la etapa (e) comprende eluir el polipéptido enlazador del antígeno diana y amplificar un vector de expresión replicable que codifica dicho polipéptido enlazador. En algunos casos se criban una o más de las bibliotecas, clones o polipéptidos contra un panel de antígenos que incluyen el antígeno diana. En algunos casos, se seleccionan aquellos clones o polipéptidos que se unen específicamente al antígeno diana y que no reaccionan de manera cruzada de forma sustancial con ninguno de los otros antígenos en el panel. El panel de antígenos puede incluir por lo menos tres y hasta 100 antígenos diferentes. En algunos casos, el panel de antígenos incluye de 3 a 100, 3 a 50, 3 a 25, o 3 a 10 antígenos diferentes.

30 **[0126]** Las combinaciones de cualquiera de los métodos de clasificación/selección descritos anteriormente se pueden combinar con los métodos de cribado. En un caso, los polipéptidos enlazadores se seleccionan en primer lugar para la unión a un antígeno diana inmovilizado. Los polipéptidos enlazadores se seleccionan en primer lugar por la unión a un antígeno diana inmovilizado. Los polipéptidos enlazadores que se unen al antígeno diana inmovilizado se pueden cribar a continuación para la unión al antígeno diana y por la falta de unión a antígenos no diana. Los polipéptidos enlazadores que se unen específicamente al antígeno diana se pueden amplificar según sea necesario. Estos polipéptidos enlazadores se pueden seleccionar por una afinidad más elevada mediante el contacto con una concentración de un antígeno diana marcado para formar un complejo, en el que el intervalo de concentración del antígeno diana marcado es de aproximadamente 0,1 nM a 1000 nM, y los complejos se aíslan mediante el contacto con un agente que se une al marcador en el antígeno diana. Un polipéptido enlazador se puede eluir a continuación del antígeno diana marcado y opcionalmente, se repiten las rondas de selección y, cada vez se utiliza una concentración inferior de antígeno diana marcado. Los polipéptidos enlazadores que se pueden aislar utilizando este método de selección se pueden cribar a continuación por una afinidad elevada utilizando, por ejemplo, el ensayo ELISA en fase solución tal como se describe, por ejemplo, en los ejemplos 2 y 4 u otros métodos convencionales conocidos en la técnica. Las poblaciones de polipéptidos de la invención utilizados en métodos de la invención se pueden proporcionar en cualquier forma adecuada para las etapas de selección/cribado. Por ejemplo, los polipéptidos pueden estar en forma soluble libre, unidos a una matriz, o presentes en la superficie de una partícula viral, tal como fago o partícula fagémida. En algunos métodos, la pluralidad de polipéptidos están codificados por una pluralidad de vectores replicables en forma de una biblioteca. En los métodos de selección/cribado aquí

descritos, los vectores que codifican un polipéptidos enlazador se pueden amplificar adicionalmente para proporcionar cantidades suficientes del polipéptido para utilizar en repeticiones de las etapas de selección/cribado (que, tal como se indica anteriormente, son opcionales en los métodos de la invención).

5 **[0127]** Se describe aquí un método de selección de un polipéptido que se une a un antígeno diana que comprende:

a) generar una composición que comprende una pluralidad de polipéptidos de la invención tal como se describen aquí;

b) seleccionar un polipéptido enlazador que se une a un antígeno diana de la composición;

c) aislar el polipéptido enlazador de los no enlazadores;

10 d) identificar enlazadores de la afinidad deseada de los polipéptidos enlazadores aislados.

**[0128]** Se describe aquí un método de selección de un dominio variable de unión a antígeno que se une a un antígeno diana de una biblioteca de dominios variables de anticuerpo que comprende:

a) poner en contacto la biblioteca de dominios variable de anticuerpo de la invención (tal como se describen aquí) con un antígeno diana;

15 b) separar los enlazadores de los no enlazadores y eluir los enlazadores del antígeno diana e incubar los enlazadores con una solución con cantidades decrecientes del antígeno diana en una concentración desde aproximadamente 0,1 nM a 1000 nM;

c) seleccionar los enlazadores que se pueden unir a la concentración más baja del antígeno diana y que tienen una afinidad de aproximadamente 0,1 nM a 200 nM.

20 **[0129]** En algunas realizaciones, la concentración de antígeno diana es de aproximadamente 100 a 250 nM, o de aproximadamente 25 a 100 nM.

**[0130]** Se describe aquí un método de selección de un polipéptido que se une a un antígeno diana de una biblioteca de polipéptidos que comprende:

25 a) aislar polipéptidos enlazadores a un antígeno diana mediante el contacto de una biblioteca que comprende una pluralidad de polipéptidos (tal como se describen aquí) con un antígeno diana inmovilizado bajo condiciones adecuadas para la unión;

b) separar los polipéptidos enlazadores en la biblioteca de los no enlazadores y eluir los enlazadores del antígeno diana para obtener una subpoblación enriquecida de los enlazadores; y

30 c) opcionalmente, repetir las etapas a-b por lo menos una vez (en algunas realizaciones por lo menos dos veces), utilizando cada repetición la subpoblación de enlazadores obtenidos de la ronda previa de selección.

**[0131]** En algunas realizaciones, los métodos de la invención comprenden además las etapas de:

d) incubar la subpoblación de polipéptidos enlazadores con una concentración de antígeno diana marcado en el intervalo de 0,1 nM a 1000 nM bajo condiciones adecuadas para la unión para formar una mezcla;

e) poner en contacto la mezcla con un agente inmovilizado que se une al marcador en el antígeno diana;

35 f) detectar los polipéptidos enlazadores unidos a los antígenos diana marcados y eluir los polipéptidos enlazadores del antígeno diana marcado;

g) opcionalmente, repetir las etapas d) a f) por lo menos una vez (en algunas realizaciones, por lo menos dos veces), utilizando cada repetición la subpoblación de enlazadores obtenida de la ronda previa de selección y utilizando una concentración inferior de antígeno diana marcado que la ronda previa.

40 **[0132]** En algunos casos, estos métodos comprenden además añadir un exceso de antígeno diana no marcado a la mezcla e incubar durante un periodo de tiempo suficiente para eluir los enlazadores de afinidad baja del antígeno diana marcado.

**[0133]** Se describe aquí un método de aislar enlazadores de afinidad elevada a un antígeno diana que comprende:

45 a) poner en contacto una biblioteca que comprende una pluralidad de polipéptidos (tal como se describen aquí) con un antígeno diana en una concentración de por lo menos aproximadamente 0,1 nM a 1000 nM para aislar polipéptidos enlazadores al antígeno diana;

b) separar los polipéptidos enlazadores del antígeno diana para obtener una subpoblación enriquecida en los polipéptidos enlazadores; y

5 c) opcionalmente, repetir las etapas a) y b) por lo menos una vez (en algunas realizaciones, por lo menos dos veces), utilizando cada repetición la subpoblación de enlazadores obtenida de la ronda previa de selección y utilizando una concentración menor de antígeno diana que la ronda previa para aislar polipéptidos enlazadores que se unen a la concentración más baja de antígeno diana.

[0134] Se describe aquí un ensayo para seleccionar polipéptidos enlazadores de una biblioteca que comprende una pluralidad de polipéptidos de la invención (tal como se describen aquí) que comprende:

10 a) poner en contacto la biblioteca con una concentración de antígeno diana marcado en un intervalo de concentración de 0,1 nM a 1000 nM, bajo condiciones adecuadas para la unión para formar un complejo de un polipéptido enlazador y el antígeno diana marcado;

b) aislar los complejos y separar los polipéptidos enlazadores del antígeno diana marcado para obtener una subpoblación enriquecida en los enlazadores;

15 c) opcionalmente, repetir las etapas a-b por lo menos una vez (en algunas realizaciones, por lo menos dos veces), utilizando cada vez la subpoblación de enlazadores obtenida de la ronda previa de selección y utilizando una concentración inferior de antígeno diana que la ronda previa.

20 [0135] En algunos casos, el método comprende además añadir un exceso de antígeno diana no marcado al complejo de enlazador de polipéptido y antígeno diana. En algunos casos, las etapas indicadas anteriormente se repiten por lo menos una vez (en algunas realizaciones, por lo menos dos veces) y la concentración de diana en la primera ronda de selección es de aproximadamente 100 nM a 250 nM, y en la segunda ronda de selección de aproximadamente 25 nM a 100 nM, y en la tercera ronda de selección de aproximadamente 0,1 nM a 25 nM.

[0136] Se describe aquí un método de cribado de una biblioteca que comprende una pluralidad de polipéptidos, comprendiendo dicho método:

25 a) incubar una primera muestra de la biblioteca con una concentración de un antígeno diana bajo condiciones adecuadas para la unión de los polipéptidos al antígeno diana;

b) incubar una segunda muestra de la biblioteca sin un antígeno diana;

c) poner en contacto cada una de la primera y segunda muestra con antígeno diana inmovilizado bajo condiciones adecuadas para la unión del polipéptido al antígeno diana inmovilizado;

d) detectar el polipéptido unido a antígeno diana inmovilizado para cada muestra;

30 e) determinar la afinidad del polipéptido por el antígeno diana mediante el cálculo de la proporción de las cantidades de polipéptido unido de la primera muestra sobre la cantidad de polipéptido unido de la segunda muestra.

35 [0137] Se contemplan usos de diagnóstico y terapéuticos para polipéptidos enlazadores aquí descritos. Una aplicación de diagnóstico es un método para determinar la presencia de una proteína de interés que comprende exponer una muestra sospechosa de contener la proteína a un polipéptido enlazador y determinar la unión del polipéptido enlazador a la muestra. Para este uso, un kit comprende el polipéptido enlazador e instrucciones para utilizar el polipéptido enlazador para detectar la proteína.

40 [0138] Se describen aquí: ácido nucleico aislado que codifica el polipéptido enlazador; un vector que comprende el ácido nucleico, opcionalmente, unido operativamente a secuencias de control reconocidas por una célula huésped transformada con el vector, una célula huésped transformada con el vector, un proceso para producir el polipéptido enlazador que comprende cultivar esta célula huésped, de manera que se expresa el ácido nucleico y, opcionalmente, recuperar el polipéptido enlazador del cultivo de células huésped (por ejemplo, del medio de cultivo de células huésped).

45 [0139] Por claridad, en la presente descripción, a menos que se indique específicamente o contextualmente lo contrario, todas las numeraciones de los aminoácidos son según Kabat et al. (véase más información en las Definiciones" a continuación).

## BREVE DESCRIPCIÓN DE LAS FIGURAS

[0140]

50 La figura 1 representa las secuencias de un dominio variable de cadena ligera y pesada de 4D5 (SEQ ID NOs:1 & 2, respectivamente).



- 5 La figura 2 muestra una estructura modelada en tres dimensiones de 4D5 humanizado que muestra los residuos de CDR que forman parches contiguos. Los parches contiguos están formados por los residuos de aminoácidos 28, 29,30,31 y 32 en CDRL1; residuos de aminoácidos 50 y 53 de CDRL2; residuos de aminoácidos 91,92, 93, 94 y 96 de CDRL3; residuos de aminoácidos 28, 30, 31, 32,33 en CDRH1; y residuos de aminoácidos 50,52,53,54,56, y 58 en CDRH2.
- 10 La figura 3 muestra la frecuencia de aminoácidos (identificados por el código de una letra) en secuencias de CDR de cadena ligera de anticuerpo humano a partir de la base de datos Kabat. La frecuencia de cada aminoácido en una posición de aminoácido particular se muestra partiendo con el aminoácido más frecuente en esa posición a la izquierda y continuando a la derecha hacia el aminoácido menos frecuente. El número debajo del aminoácido representa el número de secuencias naturales en la base de datos Kabat que presentan ese aminoácido en esa posición.
- 15 La figura 4 muestra la frecuencia de aminoácidos (identificados por el código de una letra) en secuencias de CDR de cadena pesada de anticuerpo humano a partir de la base de datos Kabat. La frecuencia de cada aminoácido en una posición de aminoácido particular se muestra partiendo con el aminoácido más frecuente en esa posición a la izquierda y continuando a la derecha hacia el aminoácido menos frecuente. El número debajo del aminoácido representa el número de secuencias naturales en la base de datos Kabat que presentan ese aminoácido en esa posición. También se muestran las posiciones de aminoácidos 71, 93 y 94 de armazón.
- 20 La figura 5 ilustra esquemáticamente un vector bicistrónico que permite la expresión de transcritos separados para la expresión de F(ab)<sub>2</sub>. Un promotor adecuado dirige la expresión del primer y segundo cistron. El primer cistron codifica una secuencia señal de secreción (malE o stII), un dominio variable y constante de cadena ligera y una etiqueta gD. El segundo cistron codifica una señal de secreción, una secuencia que codifica un dominio variable de cadena pesada y un dominio constante 1 (CH1) de dominio de dimerización de cisteína y por lo menos una parte de la proteína de cubierta viral.
- 25 La figura 6 representa secuencias de la región de armazón de cadenas ligera y pesada de huMAb4D5-8. Los números en superíndice/negrita indican las posiciones de los aminoácidos según Kabat.
- La figura 7 representa secuencias modificadas/variantes de la región de armazón de cadenas ligera y pesada de huMAb4D5-8. Los números en superíndice/negrita indican las posiciones de los aminoácidos según Kabat.
- La figura 8 ilustra el esquema de aleatorización para cada posición de CDR diversificada en las bibliotecas YSGR-A, YSGR-B, YSGRC, e YSGR-D, tal como se describe en el ejemplo 1.
- 30 Las figuras 9A-9D muestran oligonucleótidos mutagénicos en la construcción de las bibliotecas YSGR-A, YSGR-B, YSGR-C, e YSGR-D, tal como se describen en el ejemplo 3. Las degeneraciones de ADN equimolar se representan en los grupos de codones (W = A/T, K = G/T, M = A/C, N = A/C/G/T, R = A/G, S = G/C, Y = T/C). Los grupos de codones se representan en el código IUB. La notación "XXX" en los oligonucleótidos H3-A6-H3-A17 representa codones que codifican Tyr/Ser/Gly en una proporción molar de 50/25/25, respectivamente. La notación "XXX" en los oligonucleótidos H3-B6-H3-B17 representa codones que codifican Tyr/Ser/Arg en una proporción molar 25/50/25, respectivamente. La notación "XXX" en los oligonucleótidos H3-C6-H3-C17 representa codones que codifican Tyr/Ser/Gly/Arg en una proporción molar 38/25/25/12, respectivamente. La notación "XXX" en los oligonucleótidos H3-D6 a H3-D17 representa codones que codifican Tyr/Ser/Gly/Arg/Ala/Asp/Glu/Phe/His/Ile/Lys/Leu/Met/Asn/Pro/Gln/Thr/Val/Tr en una proporción molar 20/26/26/13/1/1/1/1/1/1/1/1/1/1/1/1/1/1/1/1/1, respectivamente.
- 35 40 La figura 10 muestra las proporciones de enriquecimiento para la biblioteca YSGR-A-D después de 5 rondas de selección contra DR5 humana o HER-2 humano, tal como se describe en el ejemplo 2. Los números se muestran como X/Y, representando X el número de clones únicos y representando Y el número de clones que se unen específicamente a DR5 humano o HER-2 humano. Los clones específicos se identifican como aquellos que muestran una unión a DR5 humano o a HER-2 humano por lo menos diez veces más (en base a la señal de ELISA leída a 450 nm) que la unión a albúmina de suero bovino (BSA).
- 45 La figura 11A muestra las secuencias de CDRH1, CDRH2, CDRH3 y CDRL3 para 10<sup>6</sup> clones que se unen a HER-2 humano. La figura 11B muestra los resultados de ensayos ELISA para cada uno de los clones indicados en la figura 11A. Los números en negrita indican una unión fuerte (señal de 2 a 10).
- 50 Las figuras 12 muestra las secuencias de aminoácidos para CDRL3, CDRH1, CDRH2, y CDRH3 de enlazadores específicos para HER-2 humano con regiones CDRH3 cortas (por ejemplo, 6-7 residuos) aisladas de la biblioteca YSGR-A-D, tal como se describe en el Ejemplo 2. Las secuencias consenso se muestran para CDRL3, CDRH1, y CDRH2. (Los números de clones corresponden a los mostrados en la figura 11.)
- 55 La figura 13 muestra las secuencias de aminoácidos para CDRH1, CDRH2, y CDRH3 de enlazadores específicos para HER-2 humano con una CDR3 que tiene 8 aminoácidos aisladas de la biblioteca YSGR-A-D tal como se describe en el ejemplo 2. Las secuencias consenso se muestran para CDRH1, CDRH2 y CDRH3. (Números de clones 1-19 corresponden a números de clones en la figura 11 de la siguiente manera: 17, 97, 18,

19, 98, 99, 100, 20, 21, 22, 23, 24, 101, 102, 25, 103, 26, 27 y 28, respectivamente.)

5 La figura 14 muestra las secuencias de aminoácidos para CDRH1, CDR2, y CDRH3 de los enlazadores específicos a HER-2 humano con regiones CDRH3 de longitud media (por ejemplo, aproximadamente 12-14 aminoácidos) aisladas de la biblioteca YSGR-A-D, tal como se describe en el Ejemplo 2. Las secuencias consenso se muestran para CDRH2, CDRH2, y CDRH3. La secuencia consenso se determinó para CDRH3 mediante el desplazamiento de las secuencias CDRH3 sobre dos aminoácidos, de manera que la secuencia CDRH3 empieza en la posición 97 en lugar de la posición 95.

La figura 15 muestra las secuencias CDRL3, CDRH1, CDRH2, y CDRH3 de enlazadores a DR5 humano y la IC50 para algunos de los enlazadores para DR5 humano.

10 La figura 16 muestra las secuencias de aminoácidos para CDRL3, CDRH1, CDRH2, y CDRH3 de los enlazadores específicos para DR5 humano aisladas de la biblioteca YSGR-A-D, tal como se describen en el ejemplo 2. Se muestran las IC50 de los clones para la unión a DR5 humano. También se identifican los clones que reaccionan de forma cruzada con DR5 murino. Las secuencias consenso se muestran para CDRL3, CDRH1, CDRH2, y CDRH3. (Números de clones 1-11 corresponden a números de clones 10, 11, 12, 8, 7, 13, 5, 9, 6, 15 y 14 de la figura 15.)

La figura 17 muestra las curvas de unión para enlazadores específicos para DR5 humano aislado de la biblioteca YSGR-A-D. Algunos de los enlazadores específicos también se unen a DR-5 humano.

20 La figura 18 muestra un modelo tridimensional que representa donde el ligando Apo-2L, el anticuerpo YSD1 y el anticuerpo BFD1 se unen al receptor DR5. La región de unión de los anticuerpos se solapa entre sí. El sitio de unión de estos anticuerpos es distinto de la mayoría de residuos del sitio de unión del ligando Apo-2L.

Las figuras 19A y 19B ilustran el esquema de aleatorización para cada posición de CDR diversificada en las bibliotecas H3 binarias (SAH3, SCH3, SFH3, SGH3, SIH3, SLH3, SNH3, SPH3, SRH3, STH3, SWH3, y SYH3), tal como se describe en el ejemplo 4. Las posiciones de los aminoácidos indicados se numeran según Kabat.

25 Las figuras 20A-20L muestran oligonucleótidos mutagénicos utilizados en la construcción de las bibliotecas H3 binarias (SAH3 (Figura 20A), SCH3 (Figura 20B), SFH3 (Figura 20C), SGH3 (Figura 20D), SIH3 (Figura 20E), SLH3 (Figura 20F), SNH3 (Figura 20G), SPH3 (Figura 20H), SRH3 (Figura 20I), STH3 (Figura 20J), SWH3 (Figura 20K), y SYH3 (Figura 20L)), tal como se describe en el ejemplo 4 (SEQ ID NOS:618-788). Las degeneraciones de ADN equimolar se representan en los grupos de codones (W = A/T, K = G/T, M = A/C, N = A/C/G/T, R = A/G, S = G/C, Y = T/C). Los grupos de codones se representan en el código IUB.

30 La figura 21A muestra las secuencias de aminoácidos para CDRL3, CDRH1, CDRH2, y CDRH3 de los enlazadores específicos para HER2 aisladas de las bibliotecas H3 binarias agrupadas (SXH3), tal como se describe en el ejemplo 5 (SEQ ID NOS:789-896).

La figura 21B muestra los resultados de ensayos ELISA para cada uno de los clones indicados en la figura 21 A. El sombreado oscuro indica la unión fuerte (señal de 2 a 10).

35 La figura 22 ilustra el esquema de aleatorización para cada posición de CDR diversificada en las bibliotecas de superficie binaria (SY, SW, SR, y SF), tal como se describe en el ejemplo 6. Las posiciones de los aminoácidos indicados se numeran según Kabat.

40 La figura 23 muestra oligonucleótidos mutagénicos utilizados en la construcción de ciertas bibliotecas de superficie binaria (SW, SR, y SF), tal como se describe en el ejemplo 6 (SEQ ID NOS:1005-1013). Las degeneraciones de ADN equimolar se representan en los grupos de codones (W = A/T, K = G/T, M = A/C, N = A/C/G/T, R = A/G, S = G/C, Y = T/C). Los grupos de codones se representan en el código IUB.

45 La figura 24A muestra las secuencias de aminoácidos para CDRL3, CDRH1, CDRH2, y CDRH3 se los enlazadores específicos para HER2 aisladas las bibliotecas de superficie binaria agrupadas (SX-superficie), tal como se describe en el ejemplo 8 (SEQ ID NOS: 897-1004). La figura 24B muestra los resultados de los ensayos ELISA para cada uno de los clones indicados en la figura 24A. El sombreado oscuro indica una unión fuerte (señal de 2 a 10) y el sombreado claro indica una unión débil (señal de 0,25 a 2).

La figura 25 representa gráficamente la especificidad de Fabs que contienen combinaciones de aminoácidos binarios diferentes (Ser:Tyr, Ser:Trp, Ser:Arg, o Ser:Phe) obtenida aquí a partir de la biblioteca SXH3 binaria o la biblioteca de superficie SX binaria.

50 Las figuras 26A y B representan análisis de unión por resonancia de plasmón de superficie de proteínas Fab solubles de los clones de unión a HER2 (clon B11, clon G54, y clon YSGR-A-42) a HER2 inmovilizado. El clon B11 presentaba una  $k_a$  de  $1,9 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ , una  $k_d$  de  $1,7 \times 10^{-3} \text{ s}^{-1}$ , y una  $K_D$  de 890 pM. El clon G54 presentaba una  $k_a$  de  $2,0 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ , una  $k_d$  de  $2,2 \times 10^{-3} \text{ s}^{-1}$ , y una  $K_D$  de 11 nM. El clon YSGR-A-42 presentaba una  $k_a$  de  $2,7 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ , una  $k_d$  de  $1,5 \times 10^{-3} \text{ s}^{-1}$ , y una  $K_D$  de 570 pM.

La figura 27 muestra los resultados de análisis citométricos de flujo de unión de Fabs anti-HER2 aislados cada uno de las bibliotecas YSGR (clon A-42), SX-superficie (clones G37 y G54), y SXH3 (clon B 11) a células NR6 o H2NR6-4D5, tal como se describen en el ejemplo 7.

5 La figura 28 muestra las secuencias para CDRH1, CDRH2, CDRH3, y CDRL3 para cada una de las IgG que se unen a HER2 B11, G37, G54, YSGR-A-42, YSGR-A-27, B27, G43, e YSGR-D-104. La figura 28 también muestra los valores de IC50 para la versión Fab de cada clon.

La figura 29 muestra los resultados de ensayos de unión competitiva descritos en el ejemplo 7 para determinar la capacidad de cada una de las IgG específicas de HER2 indicadas de competir por la unión a HER2 con Omnitarg, Herceptin, y cada una de las otras IgG.

## 10 MODOS PARA LLEVAR A CABO LA INVENCION

**[0141]** La presente invención se refiere a métodos nuevos, no convencionales, ampliamente simplificados y flexibles para diversificar secuencias de CDR, tal como se establece en las reivindicaciones.

15 **[0142]** También se describen aquí bibliotecas que comprenden una multiplicidad, generalmente una amplia multiplicidad de CDR diversificadas (incluyendo secuencias de dominio variable de anticuerpo). Dichas bibliotecas proporcionan bibliotecas combinatorias útiles para, por ejemplo, seleccionar y/o cribar clones de anticuerpos sintéticos con actividades deseables, tales como la afinidad y avidez de unión. Estas bibliotecas son útiles para identificar secuencias polipeptídicas de inmunoglobulina que son capaces de interactuar con cualquiera de una amplia variedad de antígenos dianas. Por ejemplo, las bibliotecas que comprenden polipéptidos de inmunoglobulina diversificados de la invención expresados como expresiones en fagos son particularmente útiles y proporcionan sistemas de alto rendimiento, eficientes y automatizables de selección y/o cribado de moléculas de unión a antígeno de interés. Los métodos de la invención se diseñan para proporcionar enlazadores de afinidad elevada a HER2 con mínimos cambios en una molécula origen o molde y proporcionar buenos rendimientos de producción cuando se producen el anticuerpo o fragmentos de unión a antígeno en cultivo celular.

25 **[0143]** Los métodos de la invención proporcionan numerosas ventajas adicionales. Por ejemplo, se pueden generar secuencias variantes de CDR relativamente simples utilizando grupos de codones que codifican un número limitado de aminoácidos (contrapuesto a la estrategia convencional de utilizar grupos de codones que codifican el número máximo de aminoácidos), manteniendo la diversidad suficiente de secuencias únicas de unión a diana. La naturaleza simplificada (y generalmente de tamaño relativamente más pequeño) de las poblaciones de secuencia generadas según la invención permite una diversificación adicional una vez la población o subpoblación de la misma, se ha identificado por poseer las características deseadas.

35 **[0144]** La naturaleza simplificada de las secuencias de enlazadores a antígeno diana obtenidas mediante los métodos de la invención deja un espacio significativamente mayor para modificaciones de secuencia posteriores individualizadas para conseguir los resultados deseados. Por ejemplo, dichas modificaciones de secuencia se realizan de manera rutinaria en maduración de afinidad, humanización, etc. Al basar la diversificación en grupos de codones limitados que codifican sólo un número limitado de aminoácidos, es posible reconocer diferentes epítopos utilizando diferentes grupos de codones limitados, proporcionando así al técnico un mayor control de la estrategia de diversificación en comparación con la aleatorización basada en un número máximo de aminoácidos. Una ventaja añadida de utilizar grupos de codones limitados es que los aminoácidos no deseados se pueden eliminar del proceso, por ejemplo, metionina o codones de parada, mejorando así la calidad y productividad global de una biblioteca. Además, en algunos casos, puede ser deseable limitar la diversidad conformacional de enlazadores potenciales. Los métodos de la invención proporcionan la flexibilidad para conseguir este objetivo. Por ejemplo, la presencia de ciertos aminoácidos, tales como tirosina, en una secuencia da lugar a menos conformaciones rotacionales.

## 45 DEFINICIONES

**[0145]** Los aminoácidos se representan aquí como un código de una letra o como el código de tres o ambos.

**[0146]** El término "purificación de afinidad" significa la purificación de una molécula basada en una atracción específica o unión de la molécula a una pareja química o de unión para formar una combinación o complejo que permite que la molécula se separe de las impurezas manteniéndose unida o atraída al grupo de pareja.

50 **[0147]** El término "anticuerpo" se usa en el sentido más amplio y cubre específicamente anticuerpos monoclonales individuales (incluyendo anticuerpos agonistas y antagonistas), composiciones de anticuerpos con especificidad poliepitópica, anticuerpos madurados por afinidad, anticuerpos humanizados, anticuerpos quiméricos, así como fragmentos de unión a antígeno (por ejemplo, Fab, F(ab')<sub>2</sub>, scFv y Fv), siempre que muestren la actividad biológica deseada. En una realización, el término "anticuerpo" también incluye anticuerpos humanos. Tal como se utiliza aquí, "dominio variable de anticuerpo" se refiere a las partes de las cadenas ligera y pesada de las moléculas de anticuerpo que incluyen las secuencias de aminoácidos de las Regiones Determinantes de Complementariedad (CDRs; es decir, CDR1, CDR2, y CDR3), y Regiones Armazón (FRs). V<sub>H</sub>

se refiere al dominio variable de la cadena pesada.  $V_L$  se refiere a dominio variable de la cadena ligera. Según las composiciones y los métodos utilizados en la invención, las posiciones de los aminoácidos asignados a CDR y FR se pueden definir según Kabat (Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md., 1987 y 1991)). La numeración de los aminoácidos de los anticuerpos o fragmentos de unión a antígeno también es según Kabat.

**[0148]** Tal como se utiliza aquí, el término "Regiones Determinantes de Complementariedad (CDRs; es decir, CDR1, CDR2, y CDR3) se refiere a los residuos de aminoácidos de un dominio variable de anticuerpo, la presencia de los cuales es necesaria para la unión a antígeno. Cada dominio variable presenta habitualmente tres regiones CDR identificadas como CDR1, CDR2 y CDR3. Cada región determinante de complementariedad puede comprender residuos de aminoácidos de una "región determinante de complementariedad" tal como se define por (es decir, aproximadamente los residuos 24-34 (L1), 50-56 (L2) y 89-97 (L3) en el dominio variable de cadena ligera y 31-35 (H1), 50-65 (H2) y 95-102 (H3) en el dominio variable de cadena pesada; Kabat et al., Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD. (1991)) y/o aquellos residuos de un "bucle hipervariable" (es decir, aproximadamente los residuos 26-32 (L1), 50-52 (L2) y 91-96 (L3) en el dominio variable de cadena ligera y 26-32 (H1), 53-55 (H2) y 96-101 (H3) en el dominio variable de cadena pesada; Chothia and Lesk J. Mol. Biol. 196:901-917 (1987)). En algunos casos, una región determinante de complementariedad puede incluir aminoácidos de una región CDR definida según Kabat y un bucle hipervariable. Por ejemplo, la CDRH1 de la cadena pesada del anticuerpo 4D5 incluye los aminoácidos 26 a 35. La secuencia consenso para CDRL1 (según la definición de Kabat) en el anticuerpo 4D5 es R-A-S-Q-D-V-N-T-A-V-A (SEQ ID NO:29). La secuencia consenso para CDRL2 (según la definición de Kabat) en el anticuerpo 4D5 es S-A-S-S-L-Y-S (SEQ ID NO:30).

**[0149]** "Regiones armazón" (de aquí en adelante "FR") son aquellos residuos de dominio variable diferentes de los residuos CDR. Cada dominio variable tiene habitualmente cuatro FR identificadas como FR1, FR2, FR3 y FR4. Si las CDR se definen según Kabat, los residuos de FR de cadena ligera se localizan aproximadamente en los residuos 1-23 (LCFR1), 35-49 (LCFR2), 57-88 (LCFR3), y 98-107 (LCFR4) y los residuos de FR de cadena pesada se localizan aproximadamente en los residuos 1-30 (HCFR1), 36-49 (HCFR2), 66-94 (HCFR3), y 103-113 (HCFR4) en los residuos de cadena pesada. Si las CDR comprenden residuos de aminoácidos de bucles hipervariables, los residuos de FR de cadena ligera se localizan aproximadamente en los residuos 1-25 (LCFR1), 33-49 (LCFR2), 53-90 (LCFR3), y 97-107 (LCFR4) en la cadena ligera y los residuos de FR de cadena pesada se localizan aproximadamente en los residuos 1-25 (HCFR1), 33-52 (HCFR2), 56-95 (HCFR3), y 102-113 (HCFR4) en los residuos de cadena pesada. En algunos casos, cuando la CDR comprende aminoácidos de una CDR tal como se define por Kabat y de un bucle hipervariable, los residuos de FR se pueden ajustar correspondientemente. Por ejemplo, cuando CDRH1 incluye los aminoácidos H26-H35, los residuos de FR1 de cadena pesada están en las posiciones 1-25 y los residuos de FR2 están en las posiciones 36-49.

**[0150]** Tal como se utiliza aquí, "grupo de codones" se refiere a un grupo de diferentes secuencias de tres nucleótidos utilizadas para codificar variantes de aminoácidos deseadas. Se puede sintetizar un grupo de oligonucleótidos, por ejemplo, mediante síntesis en fase sólida, incluyendo secuencias que representan todas las posibles combinaciones de tripletes de nucleótidos proporcionadas por el grupo de codones y que codificarán el grupo deseado de aminoácidos. Una forma estándar de designación de codones es la del código IUB, que es conocido en la técnica y se describe aquí. Un grupo de codones se representa normalmente por 3 letras mayúsculas en cursiva, por ejemplo, *NNK*, *NNS*, *XYZ*, *DVK* y similares. La síntesis de oligonucleótidos con una "degeneración" de nucleótidos seleccionada en ciertas posiciones es conocida en la técnica, por ejemplo la estrategia TRIM (Knappek et al.; J. Mol. Biol. (1999), 296:57-86); Garrard & Henner, Gene (1993), 128:103). Dichos grupos de oligonucleótidos que tiene ciertos grupos de codones se pueden sintetizar utilizando sintetizadores comerciales de ácidos nucleicos (disponible en, por ejemplo, Applied Biosystems, Foster City, CA), o se pueden obtener comercialmente (por ejemplo, de Life Technologies, Rockville, MD). Por lo tanto, un grupo de oligonucleótidos sintetizado que tiene un grupo de codones particular incluirá habitualmente una pluralidad de oligonucleótidos con diferentes secuencias, las diferencias establecidas por el grupo de codones en la secuencia global. Los oligonucleótidos, tal como se utiliza en la presente invención, presentan secuencias que permiten la hibridación a una plantilla de ácido nucleico del dominio variable y también pueden, pero no necesariamente, incluir sitios de enzimas de restricción útiles para, por ejemplo, objetivos de clonación.

**[0151]** El término "grupo de codones limitado", y variaciones del mismo, tal como se utiliza aquí, se refiere a un grupo de codones que codifica un número mucho más limitado de aminoácidos que los grupos de codones utilizados normalmente en los métodos de la técnica de generación de diversidad de secuencia. En un aspecto de la invención, los grupos de codones limitados para la diversificación de secuencias codifican de 2 a 10, de 2 a 8, de 2 a 6, de 2 a 4 o sólo 2 aminoácidos. En algunas realizaciones, un grupo de codones limitado utilizado para la diversificación de secuencias codifican por lo menos 2 excepto 10 o menos, 8 o menos, 6 o menos, 4 o menos aminoácidos. En un ejemplo típico, se utiliza un codón tetranominal. Entre los ejemplos de grupos de codones tetranominales se incluyen RMC, RMG, RRC, RSA, MKC, YMT, RST, KMT, SRC, MRT y WMT, tal como se conoce en la técnica. En otro ejemplo típico, se utiliza un codón binomial. Entre los ejemplos de grupos de codones binomiales se incluyen TMT, KAT, YAC, WAC, TWC, TYT, YTC, WTC, KTT, YCT, MCG, SCG, MGC, SGT, GRT, GKT y GYT. La determinación de codones limitados adecuados y la identificación de aminoácidos específicos codificados por un codón limitado particular son conocidos en la técnica y serían evidentes para un

experto en la materia. La determinación de grupos de aminoácidos adecuados a utilizar para la diversificación de una secuencia de CDR puede ser empírica y/o guiada por criterios conocidos en la técnica (por ejemplo, la inclusión de una combinación de tipos de aminoácidos hidrofóbicos e hidrofílicos, etc.).

5 **[0152]** Un fragmento "Fv" es un fragmento de anticuerpo que contiene un sitio de reconocimiento del antígeno y unión al antígeno completo. Esta región consiste en un dímero de un dominio variable de la cadena pesada y otro de la cadena ligera en estrecha asociación, que puede ser de naturaleza covalente, por ejemplo en scFv. Es en esta configuración en la que las tres CDR de cada dominio variable interaccionan para definir un sitio de unión al antígeno en la superficie del dímero  $V_H$ - $V_L$ . Colectivamente, las seis CDR o un subgrupo de las mismas confieren al anticuerpo especificidad de unión al antígeno. Sin embargo, incluso un único dominio variable (o la mitad de un Fv que comprende sólo tres CDR específicas para un antígeno) tiene la capacidad de reconocer y unirse al antígeno, aunque normalmente a una afinidad más baja que la del sitio de unión completo.

10 **[0153]** El fragmento "Fab" contiene un dominio variable y constante de la cadena ligera y un dominio variable y el primer dominio constante (CH1) de la cadena pesada. Los fragmentos de anticuerpo  $F(ab')_2$  comprenden una pareja de fragmentos fab que generalmente está unidos covalentemente próximos a sus extremos carboxilo por las cisteínas bisagra entre ellos. También se conocen otros acoplamientos químicos de fragmentos de anticuerpo.

15 **[0154]** Los fragmentos de anticuerpo "Fv de cadena sencilla" o "scFv" comprenden los dominios  $V_H$  y  $V_L$  del anticuerpo, donde estos dominios se presentan en una única cadena polipeptídica. Generalmente, el polipéptido Fv además comprende un enlazador polipeptídico entre los dominios  $V_H$  y  $V_L$  que permite al scFv formar la estructura deseada para unirse al antígeno. Para una revisión de los scFv, véase Plückthun en *The Pharmacology of Monoclonal Antibodies*, vol. 113, Rosenberg y Moore eds., Springer-Verlag, Nueva York, pág. 269-315 (1994).

20 **[0155]** El término "diabodies" se refiere a fragmentos pequeños de anticuerpo con dos sitios de unión a antígeno, cuyos fragmentos comprenden un dominio variable de cadena pesada ( $V_H$ ) conectado a un dominio variable de cadena ligera ( $V_L$ ) en la misma cadena polipeptídica ( $V_H$  y  $V_L$ ). Mediante la utilización de un enlazador que es demasiado corto para permitir el emparejamiento entre los dos dominios en la misma cadena, los dominios son forzados a emparejarse con los dominios complementarios de otra cadena y crean dos sitios de unión a antígeno. Los diabodies se describen de manera más completa en, por ejemplo, EP 404,097; WO 93/11161; y Hollinger et al., *Proc. Natl. Acad. Sci. USA*, 90:6444-6448 (1993).

25 **[0156]** La expresión "anticuerpos lineales" se refiere a los anticuerpos descritos en Zapata et al., *Protein Eng.*, 8(10): 1057-1062 (1995). Brevemente, estos anticuerpos comprenden una pareja de segmentos Fd en tándem ( $V_H$ -CH1- $V_H$ -CH1) que, junto con los polipéptidos de la cadena ligera complementaria, forman una pareja de regiones de unión a antígeno. Los anticuerpos lineales pueden ser biospecíficos o monoespecíficos.

30 **[0157]** El término "anticuerpo monoclonal" tal como se utiliza aquí, se refiere a un anticuerpo obtenido a partir de una población de anticuerpos sustancialmente homogénea, es decir, los anticuerpos individuales que componen la población son idénticos excepto por posibles mutaciones naturales que puedan presentarse en cantidades menores. Los anticuerpos monoclonales son muy específicos estando dirigidos contra un único sitio antigénico. Además, a diferencia de las preparaciones de anticuerpos convencionales (policlonales) que normalmente incluyen diferentes anticuerpos dirigidos contra diferentes determinantes (epítomos), cada anticuerpo monoclonal se dirige contra un único determinante en el antígeno. El modificador "monoclonal" indica el carácter del anticuerpo como obtenido a partir de una población sustancialmente homogénea de anticuerpos y no debe interpretarse como que la producción del anticuerpo requiere ningún procedimiento en particular. Por ejemplo, los anticuerpos monoclonales utilizados según la presente invención pueden obtenerse mediante el procedimiento de obtención de hibridomas descrito por primera vez por Kohler et al., *Nature*, 256:495 (1975) o pueden obtenerse mediante procedimientos de ADN recombinante (véase, por ejemplo, la patente de EE. UU. N.º 4.816.567). Los "anticuerpos monoclonales" también pueden aislarse a partir de bibliotecas de anticuerpos en fagos usando las técnicas descritas en Clackson et al.; *Nature*, 352:624-628 (1991) y Marks et al., *J. Mol. Biol.*, 222:581-597 (1991), por ejemplo.

35 **[0158]** Los anticuerpos monoclonales de este documento incluyen específicamente anticuerpos (inmunoglobulinas) "quiméricos" en los que una porción de la cadena pesada y/o ligera es idéntica, u homóloga, a las secuencias correspondientes en anticuerpos derivados de una especie en particular o pertenecen a una clase o subclase de anticuerpos en particular, mientras que el resto de las cadenas es idéntico, u homólogo, a secuencias correspondientes en anticuerpos derivados de otras especies o pertenecientes a otra clase o subclase de anticuerpos, así como fragmentos de estos anticuerpos, siempre que muestren la actividad biológica deseada (patente de EE.UU. N.º 4.816.567 y Morrison et al., *Proc. Natl. Acad. Sci. USA* 81:6851-6855 (1984)).

40 **[0159]** Las formas "humanizadas" de anticuerpos no humanos (por ejemplo, murinos) son anticuerpos quiméricos que contienen una secuencia mínima derivada de una inmunoglobulina no humana. En la mayoría de los casos, los anticuerpos humanizados son inmunoglobulinas humanas (anticuerpo receptor) en las que los restos de una región hipervariable del receptor se sustituyen por restos de una región hipervariable de una especie no humana

(anticuerpo donador), como ratón, rata, conejo o primate no humano que tienen la especificidad, afinidad y capacidad deseadas. En algunos casos, los restos de la región armazón (FR) de la inmunoglobulina humana se sustituyen por los restos no humanos correspondientes. Adicionalmente, los anticuerpos humanizados pueden comprender restos que no se encuentran en el anticuerpo receptor ni en el anticuerpo donante. Estas modificaciones se hacen para refinar adicionalmente el funcionamiento del anticuerpo. En general, el anticuerpo humanizado comprenderá sustancialmente todos de al menos uno, y normalmente dos, dominios variables, en los que todas o sustancialmente todas los bucles hipervariables corresponden a las de una inmunoglobulina humana y todas, o sustancialmente todas, las regiones FR son las de una secuencia de inmunoglobulina humana. Opcionalmente, el anticuerpo humanizado también comprenderá al menos una porción de una región constante de la inmunoglobulina (Fc), que normalmente es el de una inmunoglobulina humana. Para detalles adicionales, véanse Jones et al., Nature 321:522-525 (1986); Riechmann et al., Nature 332:323-329 (1988) y Presta, Curr. Op. Struct. Biol. 2:593-596 (1992).

**[0160]** Un "anticuerpo dependiente de especie" es aquel que tiene una afinidad de unión más fuerte para un antígeno de una primera especie de mamífero que la que tiene para un homólogo de ese antígeno de una segunda especie de mamífero. Normalmente, el anticuerpo dependiente de especie "se une específicamente" a un antígeno humano (es decir, tiene un valor de afinidad de unión (Kd) no superior a aproximadamente  $1 \times 10^{-7}$  M, por ejemplo no más de aproximadamente  $1 \times 10^{-8}$  M y como ejemplo adicional no más de aproximadamente  $1 \times 10^{-9}$  M), pero tiene una afinidad de unión por un homólogo del antígeno de una segunda especie de mamífero no humano que es por lo menos aproximadamente 50 veces, o por lo menos aproximadamente 500 veces, o por lo menos aproximadamente 1000 veces, más débil que su afinidad de unión por el antígeno humano. El anticuerpo dependiente de especie puede ser cualquiera de los varios tipos de anticuerpos definidos anteriormente, pero preferiblemente es un anticuerpo humanizado o humano.

**[0161]** Tal como se utiliza aquí, "anticuerpo mutante" o "variante de anticuerpo" se refiere a una variante de secuencia de aminoácidos del anticuerpo dependiente de especie, en el que uno o más de los residuos de aminoácidos del anticuerpo dependiente de especie han sido modificados. Dichos mutantes tienen necesariamente menos de un 100% de identidad o similitud en la secuencia con el anticuerpo dependiente de especie. En una realización, el anticuerpo mutante tendrá la secuencia de aminoácidos que tiene por lo menos un 75% de identidad o similitud en la secuencia de aminoácidos con la secuencia de aminoácidos del dominio variable de cadena pesada o cadena ligera del anticuerpo dependiente de especie, por ejemplo por lo menos el 80%, por ejemplo por lo menos el 85%, por ejemplo por lo menos el 90%, y por ejemplo por lo menos el 95%. La identidad o similitud con respecto a esta secuencia se define aquí como el porcentaje de residuos de aminoácidos en la secuencia candidata que son idénticos (es decir, el mismo residuo) o similares (es decir, residuos de aminoácidos del mismo grupo basado en propiedades comunes de la cadena lateral, ver a continuación) con los residuos del anticuerpo dependiente de especie, después de la alineación de las secuencias y la introducción de espacios, si es necesario, para conseguir el máximo porcentaje de identidad de secuencia. Ninguna de las extensiones, deleciones o inserciones N-terminal, C-terminal o internos en la secuencia del anticuerpo fuera del dominio variable se interpretará que afectan a la identidad o similitud de la secuencia.

**[0162]** Un anticuerpo "aislado" es aquel que se ha identificado y separado y/o recuperado a partir de un componente de su ambiente natural. Los componentes contaminantes de su ambiente natural son materiales que podrían interferir con los usos diagnósticos o terapéuticos del anticuerpo y pueden incluir enzimas, hormonas y otros solutos proteicos o no proteicos. En ciertas realizaciones, el anticuerpo se purificará (1) hasta más del 95% en peso de anticuerpo según se determina mediante el procedimiento de Lowry y, más preferiblemente, más del 99% en peso, (2) hasta un nivel suficiente para obtener al menos 15 restos de la secuencia de aminoácidos del extremo N-terminal o internos usando un secuenciador de taza giratoria o (3) a homogeneidad mediante PAGE-SDS en condiciones reductoras o no reductoras usando tinción con azul de Coomassie o, preferiblemente, con plata. El anticuerpo aislado incluye el anticuerpo *in situ* dentro de células recombinantes, ya que al menos un componente del medio natural del anticuerpo no estará presente. Sin embargo, generalmente el anticuerpo aislado se preparará mediante al menos una etapa de purificación.

**[0163]** El término "antagonista" se utiliza en el sentido más amplio, e incluye cualquier molécula que aumenta, estimula o activa parcial o completamente una o más actividades biológicas de moléculas diana aquí descritas (por ejemplo, DR5 o HER-2) *in vitro*, *in situ* o *in vivo*. Ejemplos de dichas actividades biológicas de DR5 incluyen la unión de Apo2L/TRAIL a DR5, la inducción de apoptosis, así como las descritas en la literatura. Ejemplos de dichas actividades biológicas de HER-2 incluyen la unión de ligandos, tales como heregulinas, fosforilación de tirosina de HER-2, inducción de la proliferación, así como la apoptosis, y también las descritas en la literatura. Un antagonista puede actuar de una manera directa o indirecta. Por ejemplo, el antagonista puede actuar para aumentar, estimular o activar, parcial o totalmente, una o más actividades biológicas de un ligando de una molécula diana *in vitro*, *in situ* o *in vivo* como resultado de su unión directa a la molécula diana. El antagonista puede actuar también indirectamente para aumentar, estimular o activar, parcial o totalmente, una o más actividades biológicas de molécula diana *in vitro*, *in situ* o *in vivo* como resultado de, por ejemplo, el bloqueo o inhibición de otra molécula efectora. La molécula antagonista puede comprender una actividad antagonista "dual" donde la molécula es capaz de aumentar, estimular o activar, parcial o totalmente, una actividad biológica de la molécula diana.

**[0164]** El término "agonista" se utiliza en el sentido más amplio, e incluye cualquier molécula que aumenta, estimula o activa parcial o completamente una o más actividades biológicas de moléculas diana aquí descritas (por ejemplo, DR5 o HER-2) in vitro, in situ o in vivo. Ejemplos de dichas actividades biológicas de DR5 incluyen la unión de Apo2L y la apoptosis, así como las descritas en la literatura. Ejemplos de dichas actividades biológicas de HER-2 incluyen la unión de ligandos, tales como heregulinas, fosforilación de tirosina de HER-2, inducción de la proliferación, así como la apoptosis, y también las descritas en la literatura. Un agonista puede actuar de una manera directa o indirecta. Por ejemplo, el agonista puede actuar para aumentar, estimular o activar, parcial o totalmente, una o más actividades biológicas de un ligando de una molécula diana in vitro, in situ o in vivo como resultado de su unión directa a la molécula diana. El agonista puede actuar también indirectamente para aumentar, estimular o activar, parcial o totalmente, una o más actividades biológicas de molécula diana in vitro, in situ o in vivo como resultado de, por ejemplo, la estimulación de otra molécula efectora que entonces provoca la activación de la molécula diana o la transducción de señal. Se contempla que un agonista puede actuar como molécula potenciadora que actúa indirectamente para aumentar o incrementar la activación o la actividad de la molécula diana. Por ejemplo, un agonista puede aumentar la actividad de Apo-2L endógeno en un mamífero. Esto se puede realizar, por ejemplo, mediante la precomplejación de DR5 o mediante la estabilización de complejos del ligando respectivo con el receptor de DR5.

**[0165]** "Célula", "línea celular", y "cultivo celular" se usan aquí indistintamente y dichas designaciones incluyen toda la progenie de una célula o línea celular. De este modo, por ejemplo, términos, tales como "transformantes" o "células transformadas" incluyen las células primarias en cuestión y los cultivos derivados de las mismas sin considerar el número de transferencias. También se entiende que toda la progenie puede no ser precisamente idéntica en el contenido de ADN, debido a mutaciones deliberadas o inadvertidas. Se incluye la progenie mutante que presenta la misma función o actividad biológica cribada en la célula originalmente transformada. Cuando se pretendan diferentes designaciones, quedará claro a partir del texto.

**[0166]** El término "secuencias de control" cuando se refiere a la expresión significa secuencias de ADN necesarias para la expresión de una secuencia codificante unida operativamente en un organismo huésped particular. Las secuencias de control que son adecuadas para procariontes incluyen, por ejemplo, un promotor, opcionalmente una secuencia operadora, un sitio de unión a ribosoma y, posiblemente, otras secuencias aunque poco entendidas. Se sabe que las células eucariotas utilizan promotores, señales de poliadenilación y potenciadores.

**[0167]** El término "proteína de cubierta" significa una proteína, por lo menos una parte de la cual, que está presente en la superficie de la partícula viral. Desde una perspectiva funcional, una proteína de cubierta es cualquier proteína que se asocia con una partícula viral durante el ensamblaje viral en una célula huésped, y permanece asociada con el virus ensamblado hasta que infecta otra célula huésped. La proteína de cubierta puede ser la proteína de cubierta mayor o principal o puede ser la proteína de cubierta menor. Una proteína de cubierta "mayor" es generalmente una proteína de cubierta que está presente en la cubierta viral en por lo menos aproximadamente 5, por lo menos aproximadamente 7, por lo menos aproximadamente 10 copias de la proteína o más. Una proteína de cubierta mayor puede estar presente en decenas, centenas o incluso miles de copias por virión. Un ejemplo de una proteína de cubierta mayor es la proteína p8 de fago filamentoso.

**[0168]** El "límite de detección" para una entidad química en un ensayo particular es la concentración mínima de esa entidad que se puede detectar por encima del nivel de fondo para ese ensayo. Por ejemplo, en el ELISA de fagos, el "límite de detección" para un fago particular que expresa un fragmento de unión a antígeno particular es la concentración de fago a la que el fago particular produce un señal de ELISA por encima de la producida por un fago de control que no expresa el fragmento de unión a antígeno.

**[0169]** Una "proteína de fusión" y un "polipéptido de fusión" se refieren a un polipéptido que tiene dos partes unidas covalentemente, donde cada una de las partes es un polipéptido que tiene una propiedad diferente. La propiedad puede ser una propiedad biológica, tal como una actividad in vitro o in vivo. La propiedad puede ser también una simple propiedad química o física, tal como la unión a un antígeno diana, catálisis de una reacción, etc. Las dos partes pueden estar unidas directamente por un enlace peptídico sencillo o a través de un enlazador peptídico que contiene uno o más residuos de aminoácidos. Generalmente, las dos partes y el enlazador estarán entre sí en el marco de lectura. En ciertas realizaciones, las dos partes del polipéptido se obtienen de polipéptidos heterólogos o diferentes.

**[0170]** "ADN heterólogo" es cualquier ADN que se introduce en una célula huésped. El ADN puede derivar de una variedad de orígenes que incluyen ADN genómico, ADNc, ADN sintético y fusiones o combinaciones de éstos. El ADN puede incluir ADN de la misma célula o tipo de células que el huésped o célula receptora o ADN de un tipo de célula diferente, por ejemplo, de un mamífero o planta. El ADN puede incluir, opcionalmente genes marcadores o de selección, por ejemplo, genes de resistencia a antibióticos, genes de resistencia a la temperatura, etc.

**[0171]** Tal como se utiliza aquí, "posición altamente diversa" se refiere a una posición de un aminoácido localizado en las regiones variables de las cadenas ligera y pesada que tienen un número de aminoácidos diferentes representados en la posición cuando se comparan las secuencias de aminoácidos de anticuerpos o

fragmentos de unión a antígeno conocidos y/o naturales. Las posiciones altamente diversas son habitualmente en las regiones de CDR. En un aspecto, la capacidad de determinar las posiciones altamente diversas en anticuerpos conocidos y/o naturales es facilitada por los datos proporcionados por Kabat, Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md., 1987 and 1991). La base de datos de internet localizada en <http://www.bioinf.org.uk.abs.structures.html> proporciona una colección extensa y la alineación de secuencias de cadena ligera (<http://www.bioinf.org.uk.abs.lc.align>) y cadena pesada (<http://www.bioinf.org.uk.abs.hc.align>) y facilita la determinación de posiciones altamente diversas en estas secuencias. Según la presente invención, una posición de aminoácido es altamente diversa si tiene de aproximadamente 2 a aproximadamente 11, de aproximadamente 4 a aproximadamente 9, y/o de aproximadamente 5 a aproximadamente 7 posibles variaciones diferentes de residuos de aminoácidos en esa posición. En algunas realizaciones, una posición de aminoácido es altamente diversa si tiene por lo menos aproximadamente 2, por lo menos aproximadamente 4, por lo menos aproximadamente 6, y/o por lo menos aproximadamente 8 posibles variaciones diferentes de residuos de aminoácidos en esa posición

**[0172]** Tal como se utiliza aquí, "biblioteca" se refiere a una pluralidad de secuencias de anticuerpo o fragmento de anticuerpo (por ejemplo, los polipéptidos de la invención), o los ácidos nucleicos que codifican estas secuencias, siendo las secuencias diferentes en la combinación de aminoácidos variantes que se introducen en estas secuencias según los métodos de la invención.

**[0173]** "Unión" o "ligamiento" es el proceso de formación de enlaces fosfodiéster ente dos fragmentos de ácido nucleico. Para el ligamiento de los dos fragmentos, los extremos de los fragmentos deben ser compatibles entre sí. En algunos casos, los extremos serán directamente compatibles después de la digestión con endonucleasa. Sin embargo, puede ser necesario primero convertir los extremos alternados producidos habitualmente después de la digestión con nucleasas en extremos romos para hacerlos compatibles para el ligamiento. Para hacer que los extremos sean romos, el ADN se trata en un tampón adecuado durante por lo menos 15 minutos a 15°C con aproximadamente 10 unidades del fragmento Klenow de la ADN polimerasa I o T4 ADN polimerasa en presencia de los cuatro desoxirribonucleótidos trifosfatos. A continuación, el ADN se purifica mediante extracción con fenolcloroformo y precipitación con etanol o mediante purificación con sílice. Los fragmentos de ADN que se ligan se ponen en solución en cantidades aproximadamente equimolares. La solución contendrá también ATP, tampón de ligasa y una ligasa, tal como T4 ADN ligasa en aproximadamente 10 unidades por 0,5 mg de ADN. Si el ADN a ligar en un vector, el vector se linealiza primero mediante digestión con la endonucleasa o endonucleasas de restricción apropiadas. El fragmento linealizado se trata a continuación con fosfatasa alcalina bacteriana o fosfatasa intestinal de ternera para evitar el auto-ligamiento durante la etapa de ligamiento. Otros métodos de ligamiento son conocidos en la técnica.

**[0174]** Una "mutación" es una delección, inserción o sustitución de un nucleótido o nucleótidos en relación a una secuencia de nucleótidos de referencia, tal como una secuencia de tipo salvaje.

**[0175]** Tal como se utiliza aquí, anticuerpos "naturales" se refiere a anticuerpos identificados a partir de un origen no sintético, por ejemplo, de una célula B específica de antígeno diferenciada obtenida ex vivo, o su correspondiente línea celular de hibridoma correspondiente, o de anticuerpos obtenidos del suero de un animal. Estos anticuerpos pueden incluir anticuerpos generados en cualquier tipo de respuesta inmune, ya sea natural o inducidos de otro modo. Los anticuerpos naturales incluyen las secuencias de aminoácidos y las secuencias de nucleótidos que constituyen o codifican estos anticuerpos, por ejemplo, identificadas en la base de datos Kabat. Tal como se utiliza aquí, los anticuerpos naturales son diferentes de los "anticuerpos sintéticos", refiriéndose los anticuerpos sintéticos a secuencias de anticuerpo que han cambiado a partir de la secuencia de origen o plantilla, por ejemplo, mediante la sustitución, delección o adición de un aminoácido, o más de un aminoácido, en una cierta posición con un aminoácido diferente, proporcionando el aminoácido diferente una secuencia de anticuerpo diferente de la secuencia de anticuerpo origen.

**[0176]** "Unido operativamente" cuando se refiere a ácidos nucleicos significa que los ácidos nucleicos están situados en una relación funcional con otra secuencia de ácidos nucleicos. Por ejemplo, el ADN para una presecuencia o secuencia líder secretora está unido operativamente a ADN para un polipéptido si se expresa como una preproteína que participa en la secreción del polipéptido; un promotor o potenciador está unido operativamente a una secuencia codificante si afecta a la transcripción de la secuencia; o un sitio de unión a ribosoma está unido operativamente a una secuencia codificante si está situado para facilitar la traducción. Generalmente, "unido operativamente" significa que las secuencias de ADN que se unen están contiguas y, en el caso de una secuencia líder secretora, contiguas y en fase de lectura. Sin embargo, los potenciadores no tienen que estar contiguos. La unión se realiza mediante la unión en los sitios de restricción convenientes. Si dichos sitios no existen, se utilizan los adaptadores o enlazadores de oligonucleótidos sintéticos según la práctica convencional.

**[0177]** "Expresión de fagos" es una técnica por la cual las variantes de polipéptidos se expresan como proteínas de fusión a por lo menos una parte de proteína de cubierta en la superficie del fago, por ejemplo, fago filamentoso, partículas. Una utilidad de la expresión en fagos se encuentra en el hecho de que amplias bibliotecas de variantes de proteínas aleatorizadas se pueden clasificar rápida y eficazmente en aquellas secuencias que se unen a un antígeno diana con afinidad elevada. La expresión de bibliotecas de péptidos y



proteínas en fagos se ha utilizado para cribar millones de polipéptidos para aquellos con propiedades de unión específicas. Los métodos de expresión en fagos polivalentes se han utilizado para expresar pequeños péptidos aleatorios y pequeñas proteínas a través de fusiones al gen II o el gen VIII de fago filamentoso. Wells and Lowman, *Curr. Opin. Struct. Biol.*, 3:355-362 (1992), y las referencias citadas. En la expresión en fagos monovalentes, se fusiona una biblioteca de proteínas o péptidos a un gen III o una parte del mismo y se expresa a niveles bajos en presencia de la proteína del gen III de tipo salvaje, de manera que las partículas de fagos expresan una copia o ninguna de las proteínas de fusión. Los efectos de avidéz se reducen en relación al fago polivalente, de manera que la clasificación es en base a la afinidad intrínseca al ligando y se utilizan vectores fagémidos, que simplifican las manipulaciones del ADN. Lowman and Wells, *Methods: A companion to Methods in Enzymology*, 3:205-0216 (1991).

**[0178]** Un "fagémido" es un vector plásmido que tienen un origen de replicación bacteriano, por ejemplo, ColE1, y una copia de una región intergénica de un bacteriófago. El fagémido se puede utilizar en cualquier bacteriófago conocido, incluyendo bacteriófago filamentoso y bacteriófago lambda. El plásmido contendrá también en general un marcador seleccionable para resistencia a antibiótico. Los segmentos de ADN clonado en estos vectores se pueden propagar como plásmidos. Cuando las células que albergan estos vectores están dispuestas con todos los genes necesarios para la producción de partículas de fagos, el modo de replicación del plásmido cambia a la replicación en círculo enrollado para generar copias de una cadena del ADN plasmídico y partículas de fagos de empaquetamiento. El fagémido puede formar partículas de fago infecciosas o no infecciosas. Este término incluye fagémidos que contienen un gen de proteína de cubierta de fago o fragmento del mismo unido a un gen de polipéptido heterólogo como una fusión de genes, de manera que el polipéptido heterólogo se expresa en la superficie de la partícula de fago.

**[0179]** El término "vector fago" significa una forma replicativa de doble cadena de un bacteriófago que contiene un gen heterólogo y capaz de la replicación. El vector fago tiene un origen de replicación de fago que permiten la replicación del fago y la formación de la partícula de fago. En ciertas realizaciones, el fago es un bacteriófago filamentoso, tal como un fago M13, fl, fd, Pf3 o un derivado del mismo, o un fago lambda, tal como lambda, 21, phi80, phi81, 82, 424, 434, etc., o un derivado del mismo.

**[0180]** "Oligonucleótidos" son polidesoxinucleótidos de cadena corta, de cadena sencilla o doble que se sintetizan químicamente mediante métodos conocidos (tales como la química de fosfotriéster, fosfito o fosforamídita, utilizando técnicas de fase sólida, tales como las descritas en EP266,032 publicada el 4 de mayo de 1988, o a través de los intermedios de desoxinucleósido H-fosfonato tal como se describe por Froeshler et al., *Nucl. Acids, Res.*, 14:5399-5407 (1986)). Métodos adicionales incluyen la reacción en cadena de la polimerasa definida a continuación y otros métodos de autocebadores y síntesis de oligonucleótidos en soportes sólidos. Todos estos métodos se describen en Engels et al., *Agnew. Chem. Int. Ed. Engl.*, 28:716-734 (1989). Estos métodos se utilizan si la secuencia completa de ácidos nucleicos del gen es conocida, o la secuencia de ácido nucleico complementaria a la cadena codificante está disponible. Alternativamente, si la secuencias de aminoácidos diana es conocida, se pueden deducir potenciales secuencias de ácido nucleico utilizando residuos codificantes conocidos y preferidos para cada residuo de aminoácido. Los oligonucleótidos se pueden purificar en geles de poliacrilamida o columnas de separación por tamaño molecular o mediante precipitación.

**[0181]** El ADN se "purifica" cuando el ADN se separa de impurezas que no son ácidos nucleicos. Las impurezas puede ser polares, no polares, iónicas, etc.

**[0182]** Un "anticuerpo origen", tal como se utiliza aquí, se refiere a un anticuerpo o fragmento de unión a antígeno cuya secuencia de unión a antígeno sirve como secuencia plantilla sobre la cual se realiza la diversificación según el criterio aquí descrito. En ciertas realizaciones, una secuencia de unión a antígeno incluye en general una región variable de anticuerpo, y por lo menos una CDR que incluye regiones armazón.

**[0183]** Tal como se utiliza aquí, "posición accesible al disolvente" se refiere a una posición de un residuo de aminoácido en las regiones variables de las cadenas pesada y ligera de un anticuerpo origen o fragmento de unión a antígeno que está determinada, en base a la estructura, ensamblaje de estructuras y/o estructura modelada del anticuerpo o fragmento de unión a antígeno, como disponible potencialmente para el acceso al disolvente y/o el contacto con una molécula, tal como un antígeno específico de anticuerpo. Estas posiciones se hallan habitualmente en las CDR y en el exterior de la proteína. Las posiciones accesibles al disolvente de un anticuerpo o fragmento de unión a antígeno, tal como se define aquí, se pueden determinar utilizando cualquier de un conjunto de algoritmos conocidos en la técnica. En ciertas realizaciones, las posiciones accesibles a los disolventes se determinan utilizando coordenadas a partir de un modelo tridimensional de un anticuerpo (o parte del mismo, por ejemplo, un dominio variable de anticuerpo, o segmento o segmentos de CDR), utilizando un programa informático, tal como el programa Insight II (Accelrys, San Diego, CA). Las posiciones accesibles al disolvente también se pueden determinar utilizando algoritmos conocidos en la técnica (por ejemplo, Lee and Richards, *J. Mol. Biol.* 55, 379 (1971) y Connolly, *J. Appl. Cryst.* 16, 548 (1983)). La determinación de las posiciones accesibles al disolvente se puede realizar utilizando software adecuado para el modelaje de las proteínas y la información estructural tridimensional obtenida a partir de un anticuerpo (o parte del mismo). El software que se puede utilizar para estos objetivos incluye el software SYBYL Biopolymer Module (Tripos Associates). En general, en ciertas realizaciones, cuando un algoritmo (programa) requiere un parámetro de

tamaño de entrada del usuario, el "tamaño" de una sonda que se utiliza en el cálculo se fija en aproximadamente 1,4 Angstrom o más pequeño de radio. Además, la determinación de las regiones accesibles a disolvente y los métodos de áreas utilizando software para ordenadores personales se ha descrito en Pacios ((1994)"ARVOMOL/CONTOUR:molecular surface areas and volumes on Personal Computers." Comput. Chem. 18(4): 377-386; and (1995). "Variations of Surface Areas and Volumes in Distinct Molecular Surfaces of Biomolecules." J. Mol. Model. 1: 46-53.)

**[0184]** Un "elemento regulador de la transcripción" contendrá uno o más de los siguientes componentes: un elemento potenciador, un promotor, una secuencia operativa, un gen represor, y una secuencia de terminación de la transcripción. Estos componentes son conocidos en la técnica. Patente de Estados Unidos No. 5,667,780.

**[0185]** Un "transformante" es una célula que ha captado y mantenido ADN evidenciado por la expresión de un fenotipo asociado con el ADN (por ejemplo, resistencia a antibiótico conferido por una proteína codificada por el ADN).

**[0186]** "Transformación" significa un proceso mediante el cual una célula capta ADN se convierte en un "transformante". La captación de ADN puede ser permanente o transitoria.

**[0187]** Un "anticuerpo humano" es aquel que posee una secuencia de aminoácidos que corresponde con la de un anticuerpo producido por un humano y/o se ha fabricado utilizando cualquiera de las técnicas para fabricar anticuerpos humanos tal como se describe aquí. Esta definición de un anticuerpo humano excluye específicamente un anticuerpo humanizado que comprende residuos de unión a antígeno no humano.

**[0188]** Un anticuerpo "madurado para afinidad" es aquel con una o más alteraciones en una o más de CDR del mismo que dan lugar a una mejora en la afinidad del anticuerpo para el antígeno, en comparación con un anticuerpo parental que no posee esta alteración o alteraciones. En ciertas realizaciones, los anticuerpos madurados para afinidad tendrán afinidades nanomolares o incluso picomolares para el antígeno diana. Los anticuerpos madurados para afinidad se producen mediante procedimientos conocidos en la técnica. Marks et al. Bio/Technology 10:779-783 (1992) describe la maduración para afinidad por mezcla de los dominios de VH y VL. La mutagénesis aleatoria de los residuos de CDR y/o armazón se describe por: Barbas et al. Proc Nat. Acad. Sci, USA 91:3809-3813 (1994); Schier et al. Gene 169:147-155 (1995); Yelton et al. J. Immunol. 155:1994-2004 (1995); Jackson et al., J. Immunol. 154(7):3310-9 (1995); y Hawkins et al, J. Mol. Biol. 226:889-896 (1992).

**[0189]** "Kd" o el "valor Kd" es la constante de disociación para la interacción de una molécula con otra. En una realización, el valor Kd se mide mediante un ensayo de unión a proteína radiomarcada (RIA). En una realización, se puede realizar un RIA para DR5 o HER-2 con la versión Fab de un anticuerpo anti-DR5 o HER-2 y una molécula de DR5 o HER-2 respectivamente, tal como se describe por el siguiente ensayo que mide la afinidad de unión en solución de Fabs para DR5 o HER-2 mediante el equilibrio de un Fab con una concentración mínima de DR5 o HER-2 marcados con <sup>125</sup>I en presencia de una serie de valoración de molécula de DR5 o HER-2 no marcada respectivamente, a continuación la captura de DR5 o HER-2 unidos respectivamente con una placa cubierta de anticuerpo anti-Fab (Chen, et al., (1999) J. Mol Biol 293:865-881). Para establecer las condiciones para el ensayo, las placas de microtitulación (Dynex) se cubren durante la noche con 5 mg/ml de un anticuerpo anti-Fab de captura (Cappel Labs) en carbonato de sodio 50 mM (pH 9,6), y posteriormente se bloquearon con albúmina de suero bovino al 2% (p/v) en PBS durante dos a cinco horas a temperatura ambiente (aproximadamente 23°C). En una placa no adsorbente (Nunc #269620), se mezclan 100 pM o 26 pM de [<sup>125</sup>I] DR5 o HER-2 con diluciones en serie de un Fab de interés, por ejemplo, Fab-12 (Presta et al., (1997) Cancer Res. 57:4593-4599). A continuación, el Fab de interés se incuba durante la noche; sin embargo, la incubación puede continuar durante 65 horas para asegurar que se alcanza el equilibrio. A continuación, las mezclas se transfieren a la placa de captura para la incubación a temperatura ambiente durante una hora. A continuación, la solución se extrae y la placa se lava ocho veces con Tween-20 al 0,1% en PBS. Cuando las placas se han secado, se añade 150 ml/pocillo de centellador (MicroScint-20; Packard), y se cuentan las placas en un contador Topcount gamma (Packard) durante diez minutos. Se eligen las concentraciones de cada Fab que producen menos o igual al 20% de unión máxima para utilizar en ensayos de unión competitiva.

**[0190]** Según otra realización, la Kd o valor Kd se puede medir utilizando ensayos de resonancia de plasmón de superficie utilizando un instrumento BIAcore™-2000 o BIAcore™-3000 (BIAcore, Inc., Piscataway, NJ). En una realización, el valor Kd de anticuerpos anti-molécula de DR5 o HER-2 para DR5 o HER-2 respectivamente, se determinar utilizando el análisis BIAcore™ según el siguiente protocolo. Brevemente, se activan chips de biosensores de dextrano carboximetilado (CM5, BIAcore Inc.) con clorhidrato de N-etil-N'-(3-dimetilaminopropil)-carbodiimida (EDC) y N-hdroxisuccinimida (NHS) según las instrucciones del proveedor. La molécula de DR5 o HER-2 humana se diluye con acetato de sodio 10 mM, pH 4,8, hasta 5 mg/ml (~0,2mM) antes de la inyección a una velocidad de flujo de 5 ml/minuto para conseguir aproximadamente 10 unidades de respuesta (RU) de proteína acoplada. Después de la inyección de DR5 o HER-2 humano, se inyecta etanolamina 1 M a grupos de bloque no reaccionados. Para las mediciones cinéticas, se inyectan diluciones en serie de dos veces de Fab (0,78 nM a 500 nM) en PBS con Tween 20 (PBST) al 0,05% a 25°C a una velocidad de flujo de aproximadamente 25 ml/min. Las velocidades de asociación (kon) y disociación (koff) se calculan utilizando un modelo simple de unión uno a uno de Langmuir (BIAcore Evaluation Software versión 3.2) mediante el ajuste

simultáneo de los sensogramas de asociación y disociación. La constante de disociación de equilibrio (Kd) se calculó como la relación koff/kon. Véase, por ejemplo, Chen, Y., et al., (1999) J. Mol Biol 293:865-881.

**[0191]** Una "variante" o "mutante" de un polipéptido de partida o de referencia (por ejemplo, un anticuerpo origen o su dominio o dominios variables/CDR), tal como una proteína (polipéptido) de fusión o un polipéptido heterólogo (heterólogo a un fago) es un polipéptido que 1) tiene una secuencia de aminoácidos diferente de la del polipéptido de partida o de referencia y 2) derivaba del polipéptido de partida o de referencia a través de mutagénesis natural o artificial (por el hombre). Dichas variantes incluyen, por ejemplo, deleciones y/o inserciones y/o sustituciones de residuos en la secuencia de aminoácidos del polipéptido de interés. Por ejemplo, un polipéptido de fusión de la invención generado utilizando un oligonucleótido que comprende un grupo de codones limitado que codifica una secuencia con una variante de aminoácido (con respecto al aminoácido hallado en la correspondiente posición en un anticuerpo origen/fragmento de unión a antígeno) sería una variante de polipéptido con respecto a un anticuerpo origen y/o fragmento de unión a antígeno y/o CDR. De este modo, una variante de CDR se refiere a una CDR que comprende una variante de secuencia con respecto a una secuencia de polipéptido de partida o de referencia (tal como la de un anticuerpo origen y/o fragmento de unión a antígeno y/o CDR). Una variante de aminoácido, en este contexto, se refiere a un aminoácido diferente del aminoácido en la posición correspondiente en una secuencia de polipéptido de partida o de referencia (tal como la de un anticuerpo origen y/o fragmento de unión a antígeno y/o CDR). Se puede realizar cualquier combinación de deleción, inserción y sustitución para llegar al constructo final variante o mutante, siempre que el constructo final posea las características funcionales deseadas. En algunos de los ejemplos aquí descritos, las secuencias enlazadoras contienen mutaciones puntuales, tales como deleciones o adiciones. Los cambios de aminoácidos también pueden alterar los procesos post-traduccionales del polipéptido, tales como el cambio del número o posición de los sitios de glicosilación. Los métodos para la generación de las variantes de las secuencias de aminoácidos de los polipéptidos se describen en la patente de Estados Unidos No. 5,534,615.

**[0192]** Una secuencia "de tipo natural" o "de referencia" o la secuencia de una proteína/polipéptido de "topo natural" o "de referencia", tal como una proteína de cubierta, o una CDR o dominio variable de un anticuerpo origen, pueden ser la secuencia de referencia a partir de la cual derivan las variantes de polipéptido a través de la introducción de mutaciones. En general, la secuencia "de tipo natural" para una proteína determinada es la secuencia que es más habitual en la naturaleza. De manera similar, una secuencia de gen "de tipo natural" es la secuencia para ese gen que se halla de forma más habitual en la naturaleza. Las mutaciones se pueden introducir en un gen "de tipo salvaje" (y de este modo la proteína que lo codifica) a través de procesos naturales o a través de medios inducidos por el hombre. Los productos de dichos procesos son formas "variantes" o "mutantes" de la proteína o gen "de tipo natural" original.

**[0193]** Una "pluralidad" de una sustancia, tal como un polipéptido o polinucleótido de la invención, tal como se utiliza aquí, se refiere en general a una colección de dos o más tipos o clases de sustancia. Existen dos o más tipos o clases de una sustancia si dos o más de las sustancias difieren entre sí con respecto a una característica particular, tal como la variante de aminoácido hallado en una posición de aminoácido particular. Por ejemplo, existe una pluralidad de polipéptidos de la invención si existen dos o más polipéptidos de la invención que son sustancialmente iguales, o son idénticos en la secuencia a excepción de la secuencia de una CDR variante o a excepción de la variante de aminoácidos en una posición de aminoácido particular accesible a disolvente y altamente diversa. En otro ejemplo, existe una pluralidad de polinucleótidos de la invención si existen dos o más polinucleótidos de la invención que son sustancialmente iguales o son idénticos en la secuencia a excepción de la secuencia que codifica una variante de CDR o a excepción de la secuencia que codifica una variante de aminoácido para una posición de aminoácido particular accesible a disolvente y altamente diversa

**[0194]** Se describen aquí métodos para generar y aislar nuevos polipéptidos de unión a antígeno diana, tales como anticuerpos o fragmentos de unión a antígeno, que pueden tener una afinidad elevada para un antígeno seleccionado. Se prepara una pluralidad de diferentes polipéptidos enlazadores mediante la mutación (diversificación) de una o más posiciones de aminoácido seleccionadas en un dominio variable de cadena ligera y/o dominio variable de cadena pesada de anticuerpo origen con grupos de codones limitados para generar una biblioteca de polipéptidos enlazadores con variantes de aminoácidos en por lo menos una secuencia de CDR, donde el número de tipos de variantes de aminoácidos se mantienen en un mínimo (es decir, 19 o menos, 15 o menos, 10 o menos, 8 o menos, 6 o menos, 4 o menos, o sólo 2, pero en general por lo menos 2). Las posiciones de aminoácido incluyen aquellas que son accesibles al disolvente, por ejemplo, determinadas analizando la estructura de un anticuerpo origen y/o que son altamente diversas entre los polipéptidos inmunoglobulina conocidos y/o naturales. Una ventaja adicional proporcionada por la naturaleza limitada de la diversificación es que las posiciones de aminoácido adicionales diferentes de las que son altamente diversas y/o accesibles a disolvente también se pueden diversificar según la necesidad o el deseo del técnico; se describen aquí ejemplos de estas realizaciones.

**[0195]** Las posiciones de aminoácidos que son accesibles a disolvente y altamente diversas son aquellas en las regiones de CDR de los dominios variables de anticuerpo seleccionadas del grupo que consiste en CDRL1, CDRL2, CDRL3, CDRH1, CDRH2, CDRH3, y mezclas de las mismas. Las posiciones de aminoácido se mutan cada una utilizando un grupo de codones limitado que codifican un número limitado de aminoácidos, siendo la elección de los aminoácidos generalmente independiente de los aminoácidos habituales en cada posición. En

algunos casos, cuando va a mutarse una posición accesible a disolvente y altamente diversa en una región CDR, se selecciona un grupo de codones que codifica de 2 a 10, de 2 a 8, de 2 a 6, de 2 a 4, y/o sólo 2 aminoácidos. En algunos casos, cuando va a mutarse una posición accesible a disolvente y altamente diversa en una región CDR, se selecciona un grupo de codones que codifica de 2 a 19, 2 a 15, 2 a 10, de 3 a 9, de 4 a 8, y/o de 5 a 7 aminoácidos. En algunos casos, un grupo de codones codifica por lo menos 2, excepto 19 o menos, 15 o menos, 10 o menos, 8 o menos, 6 o menos, 4 o menos aminoácidos. Las secuencias de CDR también se pueden diversificar variando la longitud. Por ejemplo, para CDRH3, se pueden generar variantes de las regiones CDRH3 que tienen diferentes longitudes y/o están aleatorizadas en posiciones seleccionadas utilizando grupos de codones limitados.

10 **[0196]** La diversidad de la biblioteca de los polipéptidos que comprende variantes de CDR se diseña utilizando grupos de codones que codifican sólo un número limitado de aminoácidos, de manera que se introduce una cantidad mínima, pero suficiente, de diversidad de secuencia en una CDR. El número de posiciones mutadas en la CDR se minimiza y se diseñan las variantes de aminoácido en cada posición para incluir un número limitado de aminoácidos, independiente de los aminoácidos que se estima que son habituales en esa posición en CDR conocidas y/o naturales. En ciertos casos, se utiliza un único anticuerpo, incluyendo por lo menos una CDR, como anticuerpo origen. Es sorprendente que se pueda generar una biblioteca de dominios variables de anticuerpo que tienen diversidad en las secuencias y el tamaño utilizando un único anticuerpo original como plantilla y dirigiendo la diversidad a posiciones particulares utilizando un número no limitado normalmente de sustituciones de aminoácidos.

#### 20 **Diseño de la diversidad de dominios variables de anticuerpo**

**[0197]** Tal como se describen aquí, se generan bibliotecas de alta calidad de dominios variables de anticuerpo. Las bibliotecas presentan una diversidad limitada de diferentes secuencias de secuencias CDR, por ejemplo, la diversidad de los dominios variables de anticuerpo. Las bibliotecas incluyen dominios variables de anticuerpo de unión de afinidad elevada para HER-2 humano. La diversidad en la biblioteca se diseña mediante la selección de posiciones de aminoácidos que son accesibles a disolvente y altamente diversas en un único anticuerpo origen y la mutación de aquellas posiciones en por lo menos una CDR utilizando grupos de codones limitados. El grupo limitado de codones puede codificar, en ciertas realizaciones, menos de 19, 15, 10, 8, 6, ó 4 aminoácidos, o codifica sólo 2 aminoácidos.

30 **[0198]** Un anticuerpo origen es el anticuerpo humanizado 4D5, pero los métodos para la diversificación se pueden aplicar a otros anticuerpos origen cuya secuencia es conocida. Un anticuerpo origen puede ser un anticuerpo natural, un anticuerpo sintético, un anticuerpo recombinante, un anticuerpo humanizado, un anticuerpo derivado de línea germinal, un anticuerpo quimérico, un anticuerpo madurado por afinidad o fragmento de unión a antígeno de los mismos. Los anticuerpos se pueden obtener de una variedad de especies de mamíferos, incluyendo humanos, ratones y ratas. En algunos casos, un anticuerpo origen es un anticuerpo que se obtiene después de una o más rondas de cribado inicial por afinidad, pero antes de una etapa o etapas de maduración de afinidad. Un anticuerpo origen se puede seleccionar o modificar para proporcionar un rendimiento y estabilidad elevados cuando se produce en un cultivo celular.

40 **[0199]** El anticuerpo 4D5 es un anticuerpo humanizado específico para un antígeno asociado a cáncer conocido como HER-2 (erbB2). El anticuerpo incluye dominios variables que tienen regiones armazón de consenso; se invirtieron algunas posiciones a la secuencia de ratón durante el proceso de incremento de la afinidad del anticuerpo humanizado. La secuencia y la estructura cristalina del anticuerpo humanizado 4D5 se ha descrito en U.S. 6,054,297, Carter et al, PNAS 89:4285 (1992), la estructura cristalina se muestra en J Mol. Biol. 229:969 (1993) y en línea en <http://www.ncbi.nlm.nih.gov/Structure/mmdb/mmdbsrv.cgi?form=6&db=t&Dopt=s&uid=990>, <http://www.ncbi.nlm.nih.gov/Structure/mmdb/mmdbsrv.cgi?form=6&db=t&Dopt=s&uid=991>, y <http://www.ncbi.nlm.nih.gov/Structure/mmdb/mmdbsrv.cei?form=6&db=t&Dopt=s&uid=992>.

50 **[0200]** Un criterio para generar diversidad en dominio variables de anticuerpo es mutar residuos en posiciones que son accesibles al disolvente (definido anteriormente). Estas posiciones se hallan normalmente en las CDR y habitualmente están en el exterior de la proteína. En ciertas realizaciones, las posiciones accesibles a disolvente se determinan utilizando coordenadas a partir de un modelo tridimensional de un anticuerpo, utilizando un programa informático, tal como el programa InsightII (Accelrys, San Diego, CA). Las posiciones accesibles a disolvente también se pueden determinar utilizando algoritmos conocidos en la técnica (por ejemplo, Lee y Richards, J. Mol. Biol. 55, 379 (1971) y Connolly, J. Appl. Cryst. 16, 548 (1983)). La determinación de las posiciones accesibles a disolvente se puede realizar utilizando software adecuado para el modelaje de proteínas y la información estructural tridimensional obtenida de un anticuerpo. El software que se puede utilizar para estos objetivos incluye el software SYBYL Biopolymer Module (Tripos Associates). En general, en ciertas realizaciones, cuando un algoritmo (programa) requiere un parámetro de tamaño de entrada del usuario, el "tamaño" de una sonda que se utiliza se fija en aproximadamente 1,4 Angstrom o menor de radio. Además, la determinación de regiones accesibles a disolvente y los métodos de área utilizando software para ordenadores personales se ha descrito en Pacios ((1994) "ARVOMOL/CONTOUR: molecular surface areas and volumes on Personal Computers", Comput. Chem. 18(4): 377-386; y "Variations of Surface Areas and Volumes in Distinct Molecular Surfaces of Biomolecules." J. Mol. Model. (1995), 1: 46-53).

**[0201]** En algunos casos, la selección de residuos accesibles a disolvente se refina adicionalmente mediante la elección de los residuos accesibles a disolvente que colectivamente forman un parche mínimo contiguo, por ejemplo cuando el polipéptido de referencia o el anticuerpo origen están en su estructura plegada en tres dimensiones. Por ejemplo, tal como se muestra en la figura 2, se forma un parche compacto (mínimo) contiguo por los residuos seleccionados para CDRH1/H2/H3/L1/L2/L3 de 4D5 humanizado. Un parche compacto (mínimo) contiguo puede comprender sólo un subgrupo (por ejemplo, 2-5 CDR) del rango completo de CDR, por ejemplo, CDRH1/H2/H3/L3. Los residuos accesibles a disolvente que no contribuyen a la formación de dicho parche puede opcionalmente excluirse de la diversificación. El refinamiento de la selección mediante este criterio permite al técnico minimizar, según se desee, el número de residuos a diversificar. Por ejemplo, el residuo 28 en H1 se puede excluir opcionalmente en la diversificación, ya que está en el extremo del parche. Sin embargo, este criterio de selección también se puede utilizar, cuando se desee, para elegir residuos a diversificar que pueden no necesariamente estimarse como accesibles a disolvente. Por ejemplo, se puede seleccionar para la diversificación un residuo que no se estima como accesible a disolvente, pero forma un parche contiguo en la estructura plegada tridimensional con otros residuos que se estiman accesibles a disolvente. Un ejemplo de esto es CDRL1-29. La selección de dichos residuos sería evidente para un experto en la materia, y si es apropiado se puede determinar empíricamente y según las necesidades y deseos del técnico.

**[0202]** Las posiciones accesibles al disolvente identificadas a partir de la estructura cristalina del anticuerpo humanizado 4D5 para cada CDR son las siguientes (posición de residuo según Kabat):

CDRL1: 28,30,31,32

CDRL2: 50, 53

CDRL3: 91, 92, 93, 94, 96

CDRH1: 28, 30, 31, 32, 33

CDRH2: 50, 52, 52A, 53, 54, 55, 56, 57, 58.

Además, en algunos casos, el residuo 29 de CDRL1 se puede seleccionar también en base a su inclusión en un parche contiguo que comprende otros residuos accesibles al disolvente. Todas o un subgrupo de las posiciones accesibles al disolvente tal como se han establecido anteriormente pueden diversificarse en métodos y composiciones de la invención. Por ejemplo, en algunas realizaciones, sólo las posiciones 50, 52, 52a, 53-56, y 58 está aleatorizadas en CDRH2.

**[0203]** Otro criterio para seleccionar posiciones a mutar es aquellas posiciones que muestran variabilidad en la secuencia de aminoácidos cuando se comparan las secuencias de anticuerpos conocidos y/o naturales. Una posición altamente diversa se refiere a una posición de un aminoácido localizado en las regiones variables de las cadenas ligera o pesada que tienen un número de diferentes aminoácidos representados en la posición cuando se comparan las secuencias de aminoácidos de anticuerpos/fragmentos de unión a antígeno conocidos y/o naturales. Las posiciones altamente diversas pueden estar en las regiones de CDR. Las posiciones de CDRH3 son todas consideradas altamente diversas. En ciertas realizaciones, los residuos de aminoácidos son altamente diversos si tienen desde aproximadamente 2 a aproximadamente 19 (aunque los valores pueden variar tal como se describe aquí) diferentes variaciones posibles de residuos de aminoácidos en esa posición.

**[0204]** En un aspecto, la identificación de las posiciones altamente diversas en anticuerpos conocidos y/o naturales es facilitada por los datos proporcionados por Kabat, Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md., 1987 and 1991). La base de datos de internet localizada en <http://www.bioinf.org.uk.abs.structures.html> proporciona una colección extensa y la alineación de secuencias de cadena ligera (<http://www.bioinf.org.uk.abs.lc.align>) y cadena pesada (<http://www.bioinf.org.uk.abs.hc.align>) y facilita la determinación de posiciones altamente diversas en estas secuencias. La diversidad en las posiciones accesibles a disolvente de anticuerpo humanizado 4D5 en cadenas ligera y pesada conocidas y/o naturales se muestra en las figuras 3 y 4.

**[0205]** Tal como se describen aquí, se mutan los residuos altamente diversos y accesibles al disolvente en por lo menos una, dos, tres, cuatro, cinco o todas las CDR seleccionadas del grupo que consiste en CDRL1, CDRL2, CDRL3, CDRH1, CDRH2, CDRH3, y mezclas de las mismas (es decir, aleatorizadas utilizando grupos de codones limitados tal como se describe aquí). Por ejemplo, se puede generar una población de polipéptidos mediante la diversificación de por lo menos un residuo accesible a disolvente y/o altamente diverso en CDRL3 y CDRH3 utilizando codones limitados. Los métodos aquí descritos proporcionan un gran número de nuevas secuencias de anticuerpo formadas mediante la sustitución de por lo menos una posición accesible a disolvente y altamente diversa de por lo menos una CDR del dominio variable del anticuerpo origen con aminoácidos variantes codificados por un codón limitado. Por ejemplo, una variante de CDR o dominio variable de anticuerpo puede comprender una variante de aminoácido en una o más de las posiciones de aminoácidos 28, 30, 31, 32, 33, y/o 34 de CDRH1; y/o en una o más de las posiciones de aminoácido 50, 52, 52a, 53, 54, 55, 56 y/o 58 de CDRH2; y/o en una o más de las posiciones de aminoácido 95-100, 100a, 100b, y 100c de CDRH3; y/o en una o más de las posiciones de aminoácido 28, 29, 30 y/o 31 de CDRL1; y/o en una o más de las posiciones de

aminoácido 50 y/o 53 en CDRL2; y/o en una o más de las posiciones de aminoácido 91, 92, 93, 94, 95 y/o 96 en CDRL3. En otro ejemplo, una variante de CDR o dominio variable de anticuerpo puede comprender una variante de aminoácido en una o más de las posiciones de aminoácido 28, 30, 31, 32, y/o 33 de CDRH1; y/o en una o más de las posiciones de aminoácido 50, 52, 53, 54, 56 y/o 58 de CDRH2; y/o en una o más de las posiciones de aminoácido 95-100, 100a, 100b, y 100c de CDRH3; y/o en una o más de las posiciones de aminoácido 28, 29, 30 y/o 31 de CDRL1; y/o en una o más de las posiciones de aminoácido 50 y/o 53 en CDRL2; y/o en una o más de las posiciones de aminoácido 91, 92, 93, 94, y/o 96 en CDRL3. Los aminoácidos variantes en estas posiciones son codificados por grupos de codones limitados tal como se describen aquí.

**[0206]** Tal como se ha descrito anteriormente, las variantes de aminoácidos son codificados por los grupos limitados de codones. Un grupo de codones es un grupo de secuencias de tres nucleótidos diferentes que se pueden utilizar para formar un grupo de oligonucleótidos utilizados para codificar el grupo deseado de aminoácidos. Se puede sintetizar un grupo de oligonucleótidos, por ejemplo, mediante síntesis en fase sólida, que contienen secuencias que representan todas las posibles combinaciones de tripletes de nucleótidos proporcionados por el grupo de codones y que codificarán el grupo deseado de aminoácidos. La síntesis de oligonucleótidos con la “degeneración” de nucleótidos seleccionada en ciertas posiciones es conocida en la técnica. Dichos grupos de nucleótidos que tienen ciertos grupos de codones se pueden sintetizar utilizando sintetizadores comerciales de ácidos nucleicos (disponible en, por ejemplo, Applied Biosystems, Foster City, CA), o se pueden obtener comercialmente (por ejemplo, de Life Technologies, Rockville, MD). Por lo tanto, un grupo de oligonucleótidos sintetizados que tienen un grupo de codones particulares incluirá habitualmente una pluralidad de oligonucleótidos con diferentes secuencias, las diferencias establecidas por el grupo de codones en la secuencia global. Los oligonucleótidos, tal como se utilizan en la invención, presentan secuencias que permiten la hibridación a una plantilla de ácido nucleico de dominio variable y también pueden incluir sitios para enzimas de restricción para fines de clonación.

**[0207]** En un caso, se determina el repertorio limitado de aminoácidos que pretende ocupar una o más de las posiciones accesibles al disolvente y altamente diversas en CDR del anticuerpo humanizado 4D5 (en base al deseo del técnico, que se puede basar en cualquier criterio, incluyendo aminoácidos específicos deseados para posiciones particulares, aminoácidos específicos que se desea que estén ausentes de una posición particular, tamaño de la biblioteca deseada, característica de enlazadores de antígeno buscados, etc.).

**[0208]** Las CDR3 (CDRH3) de cadena pesada en anticuerpos conocidos presentan diversas secuencias, conformaciones estructurales y longitudes. Las CDRH3 se hallan a menudo en el centro del bolsillo de unión a antígeno y a menudo participan en el contacto con el antígeno. El diseño de CDRH3 se puede desarrollar de este modo de manera separada de la de otras CDR porque puede ser difícil predecir la conformación estructural de CDRH3 y la diversidad de aminoácidos en esta región es especialmente diversa en anticuerpos conocidos. Según la presente invención, se diseña CDH3 para generar diversidad en posiciones específicas en CDRH3, por ejemplo, las posiciones 95, 96, 97, 98, 99, 100, 100a, 100b, y 100c (por ejemplo, según el sistema de numeración Kabat en el anticuerpo 4D5). En algunas realizaciones, la diversidad también se genera variando la longitud de CDRH3 utilizando grupos de codones limitados. La diversidad de longitud puede ser de cualquier rango determinado empíricamente por ser adecuado para generar una población de polipéptidos que contienen proporciones sustanciales de proteínas de unión a antígeno. Por ejemplo, los polipéptidos que comprenden una variante de CDRH3 se pueden generar con la secuencia X1-X2-X3-X4-X5-(X6)<sub>n</sub>-X7-X8-X9-D-Y (SEQ ID NO:4), en la que X1-X9 son aminoácidos codificados por los grupos de codones limitados, y n es de varias longitudes, por ejemplo, n=1-11, 5-11, o 7-11. Otros ejemplos de posibles valores de n son 1,2, 3, 4, 5, 6, 7, 8, 9, 10, y 11. Ejemplos ilustrativos de oligonucleótidos que se pueden utilizar para proporcionar una variedad en la longitud de la secuencia de CDRH3 incluyen los mostrados en la figura 9A-9D, la figura 20A-L y la figura 23.

**[0209]** Se contempla que la diversidad de secuencia de bibliotecas creadas mediante la introducción de variantes de aminoácidos en una CDR particular, por ejemplo, CDRH3, se puede incrementar mediante la combinación de la variante de CDR con otras CDR que comprenden variaciones en otras regiones del anticuerpo, específicamente en otras CDR de las secuencias variables de cadena ligera o cadena pesada. Se contempla que las secuencias de ácido nucleico que codifican los miembros de este grupo se pueden diversificar adicionalmente mediante la introducción de otras variante de aminoácidos en las CDR de las secuencias de cadena ligera o pesada, a través de grupos de codones. De este modo, por ejemplo, en una realización, las secuencias CDRH3 de los polipéptidos de fusión que se unen a antígeno diana se pueden combinar con las secuencias CDRL3, CDRH1, o CDRH2 diversificadas, o cualquier combinación de CDR diversificadas.

**[0210]** Cabe indicar que en algunos casos los residuos de armazón pueden variar en relación a la secuencia de un anticuerpo origen o fragmento de unión a antígeno, por ejemplo, para reflejar una secuencia consenso o para mejorar la estabilidad o la expresión. Por ejemplo, se pueden variar los residuos de armazón 49, 93, 94 ó 71 en la cadena pesada. El residuo 93 del armazón de cadena pesada puede ser serina o alanina (que es el aminoácido de la secuencia consenso humana en esa posición). El residuo 94 de armazón de cadena pesada se puede cambiar para reflejar la secuencia consenso de armazón de treonina a arginina o lisina. Otro ejemplo del residuo de armazón que se puede alterar es el residuo 71 del armazón de cadena pesada, que es R en aproximadamente 1970 polipéptidos, V en aproximadamente 627 polipéptidos y A en aproximadamente 527 polipéptidos, tal como se halla en la base de datos Kabat. El residuo 49 del armazón de cadena pesada puede

ser alanina o glicina. Además, opcionalmente, se pueden eliminar los 3 aminoácidos N-terminales del dominio variable de cadena pesada. En la cadena ligera, opcionalmente, la arginina en la posición de aminoácido 66 se puede cambiar a glicina. En un caso, el residuo 93 de armazón de cadena pesada es alanina y el residuo 94 del armazón de cadena pesada es arginina.

5 [0211] Se describen aquí constructos de vectores para generar polipéptidos de fusión que se unen con afinidad significativa a ligandos potenciales. Estos constructos comprenden un dominio dimerizable que cuando está presente en un polipéptido de fusión proporciona una mayor tendencia para que las cadenas pesadas se dimericen para formar dímero de fragmentos/porciones de anticuerpo Fab o Fab'. Estos dominios de dimerización pueden incluir, por ejemplo, una secuencia bisagra de cadena pesada (por ejemplo, una secuencia que comprende TCPPCPAPELLG (SEQ ID NO: 5) que puede estar presente en el polipéptido de fusión). Los dominios de dimerización en los polipéptidos de fago de fusión llevan dos grupos de polipéptidos de fusión (LC/HC-proteína de fago/fragmento (tal como pIII)) juntos, permitiendo así la formación de enlaces adecuados (tales como puentes disulfuro entre cadenas pesadas) entre los dos grupos de polipéptidos de fusión. Los constructos de vectores que contienen dichos dominios de dimerización se pueden utilizar para conseguir la expresión divalente de dominios variables de anticuerpo, por ejemplo, las proteínas de fusión diversificadas aquí descritas, en fagos. En ciertos casos, la afinidad intrínseca de cada fragmento de anticuerpo monomérico (polipéptido de fusión) no se ve significativamente alterada por la fusión del dominio de dimerización. En ciertos casos, la dimerización da lugar a una expresión en fagos divalentes que proporciona una mayor avidéz de la unión a fagos con un descenso significativo en la velocidad off, que se puede determinar mediante métodos conocidos en la técnica y tal como se describen aquí. Los vectores que contienen dominios de dimerización de la invención pueden incluir o no también un codón de parada amber después del dominio de dimerización.

[0212] La dimerización se puede variar para conseguir diferentes características de expresión. Los dominios de dimerización pueden comprender una secuencia que comprende un residuo de cisteína, una región bisagra de un anticuerpo de longitud completa, una secuencia de dimerización, tal como secuencia de cremallera de leucinas o secuencia de cremallera de GCN4 o mezclas de los mismos. Las secuencias de dimerización son conocidas en la técnica e incluyen, por ejemplo, la secuencia de cremallera de GCN4 (GRMKQLEDKVEELLSKNYHLENEVARLKKLVGERG) (SEQ ID NO: 3). El dominio de dimerización se localiza, en ciertos casos, en el extremo C-terminal de la secuencia del dominio variable o constante de cadena pesada y/o entre la secuencia del dominio variable o constante de cadena pesada y cualquier secuencia de componente de proteína de cubierta viral. También puede estar presente un codón de parada amber en o después del extremo C-terminal del dominio de dimerización. En un caso, en el que está presente un codón de parada amber, el dominio de dimerización codifica por lo menos una cisteína y una secuencia dimerizante, tal como la cremallera de leucina. En otro caso, en el que el codón de parada amber no esté presente, el dominio de dimerización puede comprender un único residuo de cisteína.

35 [0213] Los polipéptidos aquí descritos también se pueden fusionar a otros tipos de polipéptidos a efectos de proporcionar la expresión de las variantes de polipéptidos o para proporcionar la purificación, cribado o clasificación y la detección del polipéptido. Para la realización que implica la expresión en fagos, los polipéptidos se fusionan a toda o parte de una proteína de cubierta viral. Ejemplos de proteína de cubierta viral incluyen proteína PIII, proteína de cubierta principal, pVIII, Soc, Hoc, gpD, pVI y variantes de las mismas. Además, las variantes de los polipéptidos generados según los métodos de la invención pueden fusionarse opcionalmente a un polipéptido marcador o etiqueta, tal como FLAG, polihistidina, gD, c-myc, B-galactosidasa y similares.

#### **Métodos de generación de bibliotecas de dominios variables aleatorizados**

[0214] Los métodos de sustitución de un aminoácido de elección en un ácido nucleico plantilla están bien establecidos en la técnica, algunos de los cuales se describen aquí. Por ejemplo, las bibliotecas se pueden crear mediante el reconocimiento de posiciones accesibles a disolvente y/o altamente diversas en por lo menos una región de CDR para la sustitución de aminoácidos con variantes de aminoácidos utilizando el método Kunkel. Véase, por ejemplo, Kunkel et al., *Methods Enzymol.* (1987), 154:367-382. La generación de secuencias aleatorizadas también se describe a continuación en los ejemplos.

50 [0215] La secuencia de oligonucleótidos incluye uno o más de los grupos de codones limitados para diferentes longitudes de CDRH3 o para las posiciones accesibles a disolvente y altamente diversas en una CDR. Un grupo de codones es un grupo de secuencias de tres nucleótidos diferentes utilizadas para codificar las variantes de aminoácidos deseadas. Los grupos de codones se pueden representar utilizando símbolos para designar nucleótidos particulares o mezclas equimolares de nucleótidos tal como se muestra a continuación según el código IUB. Habitualmente, se representa un grupo de codones por tres letras mayúsculas, por ejemplo, KMT, TMT y similares.

#### **CÓDIGOS IUB**

[0216]

G Guanina

A Adenina

T Timina

C Citosina

R (A o G)

5 Y (C o T)

M (A o C)

K(G o T)

S (C o G)

W (A o T)

10 H (A o C o T)

B (C o G o T)

V (A o C o G)

D (A o G o T)

N(A o C o G o T)

15 **[0217]** Por ejemplo, en el grupo de codones TMT, T es el nucleótido timina; y M puede ser A o C. Este grupo de codones puede presentar múltiples codones y puede codificar sólo un número limitado de aminoácidos, concretamente tirosina y serina.

20 **[0218]** Los grupos de oligonucleótidos o sondas se pueden sintetizar utilizando métodos estándar. Se puede sintetizar un conjunto de oligonucleótidos, por ejemplo, mediante síntesis en fase sólida, que contienen secuencias que representan todas las posibles combinaciones de tripletes de nucleótidos proporcionadas por el grupo de codones limitados y que codificarán el grupo deseado limitado de aminoácidos. La síntesis de oligonucleótidos con "degeneración" de nucleótidos seleccionados en ciertas posiciones es conocida en la técnica. Dichos grupos de oligonucleótidos que tienen grupos de codones se pueden sintetizar utilizando sintetizadores de ácidos nucleicos comerciales (disponibles en, por ejemplo, Applied Biosystems, Foster City, CA), o se pueden obtener comercialmente (por ejemplo, de Life Technologies, Rockville, MD). Por lo tanto, un grupo de oligonucleótidos sintetizados que tienen un grupo de codones particulares incluirá habitualmente una pluralidad de oligonucleótidos con secuencias diferentes, las diferencias establecidas por el grupo de codones en la secuencia general. Los oligonucleótidos, tal como se utilizan en la presente invención, presentan secuencias que permiten la hibridación a un molde de ácido nucleico de CDR (por ejemplo, contenido en un dominio variable) y también pueden incluir sitios para enzimas de restricción para fines de clonación.

30 **[0219]** En un método, las secuencias de ácido nucleico que codifican las variantes de aminoácidos se pueden crear mediante mutagénesis mediada por oligonucleótido de una secuencia de ácido nucleico que codifica un polipéptido origen o de molde, tal como el dominio variable de anticuerpo de 4D5. Esta técnica es bien conocida en la técnica tal como se describe en Zoller *et al.* Nucleic Acids Res. 10:6487-6504 (1987). Brevemente, las secuencias de ácidos nucleicos que codifican aminoácidos variantes se crean mediante hibridación de un conjunto de oligonucleótidos que codifican los grupos de codones deseados limitados a un ADN molde, donde el molde es la forma en cadena sencilla del plásmido que contiene una secuencia de la plantilla de ácido nucleico de la región variable. Después de la hibridación, se utiliza ADN polimerasa para sintetizar una segunda hebra complementaria entera del molde que incorporará de este modo el cebador de oligonucleótidos y contendrá los grupos de codones limitado proporcionados por el grupo de oligonucleótidos. Los ácidos nucleicos que codifican otras moléculas de origen o de plantilla se conocen o se pueden determinar fácilmente.

35 **[0220]** En general, se utilizan oligonucleótidos de por lo menos 25 nucleótidos de longitud. Un oligonucleótido óptimo tendrá por lo menos de 12 a 15 nucleótidos que son completamente complementarios a la plantilla en cualquier lado de los nucleótidos que codifican la mutación o mutaciones. Esto asegura que el oligonucleótido se hibridará correctamente a la molécula plantilla de ADN de cadena sencilla. Los oligonucleótidos se sintetizan fácilmente utilizando técnicas conocidas en el sector, tales como las descritas por Crea *et al.*, Proc. Natl. Acad. Sci. USA, 75:5765 (1978).

40 **[0221]** El molde de ADN es generado por aquellos vectores que derivan de los vectores del bacteriófago M13 (los vectores M13mp18 y M13mp19 disponibles comercialmente son adecuados), o aquellos vectores que contienen un origen de replicación del fago de cadena única tal como se describe por Viera *et al.*, Meth. Enzymol., 153:3 (1987). De este modo, el ADN a mutar se puede insertar en uno de estos vectores con el fin de generar un molde



de cadena única. La producción del molde de cadena única se describe en las secciones 4.21-4.41 de Sambrook *et al.*, anterior

**[0222]** Para alterar la secuencia de ADN nativo, el oligonucleótido se hibrida al molde de cadena única en condiciones de hibridación adecuadas. A continuación, se añade una enzima polimerizante de ADN, normalmente ADN polimerasa de T7 o el fragmento Klenow de ADN polimerasa I, para sintetizar la cadena complementaria del molde utilizando el oligonucleótido como cebador para la síntesis. De este modo, se forma una molécula de cadena doble heteróloga, de modo que una cadena de ADN codifica la forma mutada del gen 1 y la otra cadena (el molde original) codifica la secuencia nativa inalterada del gen 1. Esta molécula de cadena doble heteróloga se transforma a continuación en una célula huésped adecuada, normalmente una procarionte, tal como *E. coli* JM101. Después del crecimiento de las células, se emplazan sobre placas de agarosa y se criban utilizando el cebador de oligonucleótidos radiomarcado con fosfato-32 para identificar las colonias bacterianas que contienen el ADN mutado.

**[0223]** El método descrito inmediatamente antes se puede modificar de manera que se crea una molécula de doble cadena homogénea donde ambas cadenas del plásmido contienen la mutación o mutaciones. Las modificaciones son las siguientes: el oligonucleótido de cadena sencilla se aparea a la plantilla de cadena sencilla tal como se ha descrito anteriormente. Se combina una mezcla de tres desoxirribonucleótidos, desoxirriboadenosina (dATP), desoxirriboguanosina (dGTP), y desoxirribotimidina (dTTP), con una tiodesoxirribocitosina modificada denominada dCTP-(aS) (que se puede obtener de Amersham). Esta mezcla se añade al complejo plantilla-oligonucleótido. Tras la adición de ADN polimerasa a la mezcla, se genera una cadena de ADN idéntica a la plantilla a excepción de las bases mutadas. Además, esta nueva cadena de ADN contendrá dCTP-(aS) en lugar de dCTP, que sirve para protegerlo de la digestión con endonucleasas de restricción. Después de modificar la cadena plantilla de la doble cadena heterogénea con una enzima de restricción apropiada, la cadena plantilla se puede digerir con ExoIII nucleasa u otra nucleasa apropiada después de la región que contiene el sitio o sitios a mutagenizar. A continuación, se detiene la reacción para dejar una molécula que sólo tiene parcialmente una cadena sencilla. A continuación, se forma una doble cadena homóloga de ADN de doble cadena completa utilizando la ADN polimerasa en presencia de los cuatro desoxirribonucleótidos trifosfatos, ATP y ADN ligasa. Esta molécula de doble cadena homóloga se puede transformar a continuación en una célula huésped adecuada.

**[0224]** Tal como se ha indicado previamente, la secuencia del grupo de oligonucleótidos es de suficiente longitud para hibridarse al ácido nucleico plantilla y también puede contener, pero no necesariamente, sitios de restricción. La plantilla de ADN se puede generar por aquellos vectores que derivan de los vectores del bacteriófago M13 o vectores que contienen un origen de replicación de fago de cadena sencilla tal como se ha descrito por Viera *et al.* ((1987) *Meth. Enzymol.*, 153:3). De este modo, el ADN a mutar debe insertarse en uno de estos vectores a efectos de generar plantillas de cadena sencilla. La producción de la plantilla de cadena sencilla se describe en las secciones 4.21-4.41 de Sambrook *et al.*, supra.

**[0225]** Según otro método, se puede generar una biblioteca mediante la disposición de grupos de oligonucleótidos en dirección 5' y en dirección 3', teniendo cada grupo una pluralidad de oligonucleótidos con diferentes secuencias, las diferentes secuencias establecidas por los grupos de codones proporcionados en la secuencia de los oligonucleótidos. Los grupos de oligonucleótidos en dirección 5' y en dirección 3', junto con una secuencia de ácido nucleico plantilla del dominio variable, se pueden utilizar en una reacción en cadena de la polimerasa para generar una "biblioteca" de productos PCR. Los productos PCR se pueden referir como "cassettes de ácidos nucleicos", ya que se pueden fusionar con otras secuencias de ácidos nucleicos relacionados o no relacionados, por ejemplo, componentes de proteína de cubierta viral y dominios de dimerización, utilizando técnicas de biología molecular establecidas.

**[0226]** La secuencia de los cebadores de PCR incluye uno o más de los grupos de codones designados para las posiciones accesibles al disolvente y altamente diversas en una región CDR. Tal como se ha descrito anteriormente, un grupo de codones es un grupo de secuencias diferentes de tres nucleótidos utilizados para codificar variantes de aminoácidos deseadas.

**[0227]** Se pueden utilizar grupos de oligonucleótidos en una reacción en cadena de la polimerasa utilizando una secuencia de plantilla de ácido nucleico de región variable como la plantilla para crear cassettes de ácido nucleico. La secuencia de plantilla de ácido nucleico de la región variable puede ser cualquier parte de las cadenas ligera o pesada de inmunoglobulina que contiene las secuencias de ácido nucleico diana (es decir, secuencias de ácido nucleico que codifican aminoácidos dirigidos para la sustitución). La secuencia de plantilla de ácido nucleico de la región variable es una parte de una molécula de ADN de doble cadena que tiene una primera cadena de ácido nucleico y una segunda cadena de ácido nucleico complementaria. La secuencia de la plantilla de ácido nucleico de la región variable contiene por lo menos una parte de un dominio variable y tiene por lo menos una CDR. En algunos casos, la secuencia de la plantilla de ácido nucleico de la región variable contiene más de una CDR. Se puede dirigir una parte en dirección 5' y una parte en dirección 3' de la secuencia de la plantilla de ácido nucleico de la región variable para la hibridación con miembros de un grupo de oligonucleótidos en dirección 5' y un grupo de oligonucleótidos en dirección 3'.

5 [0228] Se puede hibridar un primer oligonucleótido del grupo de cebadores en dirección 5' a la primera cadena de ácido nucleico y se puede hibridar un segundo oligonucleótido del grupo de cebadores en dirección 3' a la segunda cadena de ácido nucleico. Los cebadores de oligonucleótidos pueden incluir uno o más grupos de codones y se pueden diseñar para hibridarse a una parte de la secuencia de la plantilla de ácido nucleico de la región variable. La utilización de estos oligonucleótidos puede introducir dos o más grupos de codones en el producto de PCR (es decir, el cassette de ácidos nucleicos) después de la PCR. El cebador de oligonucleótidos que se hibrida a regiones de la secuencia de ácido nucleico que codifica el dominio variable de anticuerpo incluye partes que codifican residuos de CDR que están dirigidas a la sustitución de aminoácidos.

10 [0229] Los grupos de oligonucleótidos en dirección 5' y en dirección 3' también se pueden sintetizar para incluir sitios de restricción en la secuencia de oligonucleótidos. Estos sitios de restricción pueden facilitar la inserción de los cassetes de ácido nucleico [es decir, productos de reacción de PCR] en un vector de expresión que tiene secuencias de anticuerpos adicionales. En ciertas realizaciones, los sitios de restricción se diseñan para facilitar la clonación de los cassetes de ácido nucleico sin introducir secuencias de ácido nucleico extrañas o eliminar las secuencias de ácidos nucleicos de CDR o de armazón originales.

15 [0230] Los cassetes de ácidos nucleicos se pueden clonar en cualquier vector adecuado para la expresión de una parte o la secuencia completa de cadena ligera o pesada que contiene las sustituciones de aminoácidos reconocidas generadas. Según los métodos detallados en la invención, el cassette de ácidos nucleicos se clona en un vector que permite la producción de una parte o la secuencia completa de la cadena ligera o pesada fusionada a toda o una parte de una proteína de cubierta viral (es decir, creando una proteína de fusión) y se expresa en la superficie de una partícula o célula. Aunque están disponibles varios tipos de vectores y se pueden utilizar para realizar la presente invención, son convenientes vectores fagémidos, ya que se pueden construir con una facilidad relativa y se pueden amplificar fácilmente. Los vectores fagémidos contienen en general una variedad de componentes que incluyen promotores, secuencias señal, genes de selección fenotípicos, sitios de origen de replicación, y otros componentes necesarios conocidos para los expertos en la materia.

25 [0231] En otra realización, en que se expresa una combinación particular de variantes de aminoácidos, el cassette de ácidos nucleicos contiene una secuencia que es capaz de codificar todos o una parte de los dominios variables de cadena pesada o ligera, y es capaz de codificar combinaciones de variantes de aminoácidos. Para la producción de anticuerpos que contienen estas variantes de aminoácidos o combinaciones de variantes de aminoácidos, como en una biblioteca, los cassetes de ácidos nucleicos se pueden insertar en un vector de expresión que contiene una secuencia de anticuerpo adicional, por ejemplo, todas o partes de los dominios variables o constantes de las regiones variables de cadena ligera y pesada. Estas secuencias de anticuerpo adicionales también se pueden fusionar a otras secuencias de ácido nucleico, tales como secuencias que codifican componentes de proteína de cubierta viral y por lo tanto permiten la producción de una proteína de fusión.

### 35 **Vectores**

[0232] Se describe aquí un vector de expresión replicable que comprende una secuencia de ácido nucleico que codifica una fusión de genes, donde la fusión de genes codifica una proteína de fusión que comprende un polipéptido que contiene CDR (tal como un dominio variable de anticuerpo), o un dominio variable de anticuerpo y un dominio constante, fusionado a toda o una parte de una proteína de cubierta viral. También se incluye una biblioteca de diversos vectores de expresión replicables que comprenden una pluralidad de fusiones de genes que codifican una pluralidad de diferentes proteínas de fusión que incluyen una pluralidad de los polipéptidos de fusión generados con diversas secuencias descritas anteriormente. Los vectores pueden incluir una variedad de componentes y se pueden construir para permitir el movimiento de dominio variable de anticuerpo entre diferentes vectores y/o para proporcionar la expresión de las proteínas de fusión en diferentes formatos.

45 [0233] Ejemplos de vectores incluyen vectores de fagos y vectores fagémidos (que se ilustran aquí, y se describen con más detalle anteriormente). Un vector de fago tiene en general un origen de replicación de fago que permite la replicación del fago y la formación de partículas de fago. El fago es en general un bacteriófago filamentoso, tal como un fago M13, f1, fd, Pf3 o un derivado del mismo, o un fago lambda, tal como lambda, 21, phi80, phi81, 82, 424, 434, etc., o un derivado del mismo.

50 [0234] Ejemplos de proteínas de cubierta viral incluyen la proteína de infectividad PIII (algunas veces designadas como p3), proteína mayor de cubierta PVIII, Soc (T4), Hoc (T4), gpD (de bacteriófago lambda), proteína menor de cubierta de bacteriófago 6 (pVI) (fago filamentoso; J Immunol Methods. 1999 Dec 10;231(1-2):39-51), variantes de la proteína mayor de cubierta del bacteriófago M13 (P8) (Protein Sci 2000 Apr;9(4):647-54). La proteína de fusión se puede expresar en la superficie de un fago y los sistemas de fagos adecuados incluyen el fago auxiliar M13KO7, M13R408, M13-VCS, y Phi X 174, el sistema de fagos pJuFo (J Virol. 2001 Aug;75(15): 7107-13.v), hiperfago (Nat Biotechnol. 2001 Jan;19(1):75-8). En ciertas realizaciones, el fago auxiliar es M13KO7, y la proteína de cubierta es la proteína de cubierta del gen III del fago M13. En ciertas realizaciones, el huésped es E. coli, y cepas deficientes de proteasa de E. coli. Los vectores, tales como el vector fth1 (Nucleic Acids Res. 2001 May 15;29(10):E50-0) pueden ser útiles para la expresión de la proteína de fusión.

**[0235]** El vector de expresión también puede tener una secuencia señal secretora fusionada al ADN que codifica un polipéptido de fusión que contiene CDR (por ejemplo, cada subunidad de un anticuerpo o fragmento del mismo). Esta secuencia se localiza normalmente inmediatamente 5' del gen que codifica la proteína de fusión, y de este modo, se transcribirá en el extremo amino de la proteína de fusión. Sin embargo, en ciertos casos, la secuencia señal se ha demostrado que se localiza en las posiciones diferentes a 5' con respecto al gen que codifica la proteína a secretar. Esta secuencia reconoce la proteína a la que se une a través de una membrana interna de la célula bacteriana. El ADN que codifica la secuencia señal se puede obtener como un fragmento de endonucleasa de restricción de cualquier gen que codifica una proteína que tiene una secuencia señal. Las secuencias señal procariontas adecuadas se pueden obtener de genes que codifican, por ejemplo, LamB o OmpF (Wong et al., *Gene*, 68:1931 (1983), MalE, PhoA y otros genes. En una realización, una secuencia señal procarionta para realizar esta invención es la secuencia señal de la enterotoxina II estable al calor (STII) de *E. coli* descrita por Chang et al., *Gene* 55:189 (1987), y/o malE.

**[0236]** Tal como se ha indicado anteriormente, un vector también incluye habitualmente un promotor para dirigir la expresión del polipéptido de fusión. Los promotores utilizados más habitualmente en vectores procariontas incluyen el sistema de promotores lac Z, el promotor de fosfatasa alcalina pho A (Ap), el promotor 1PL de bacteriófago (un promotor sensible a la temperatura), el promotor tac (un promotor híbrido trp-lac que es regulado por el represor lac), el promotor de triptófano, y el promotor T7 de bacteriófago. Para descripciones generales de promotores, Para descripciones generales de promotores, véase la sección 17 de Sambrook et al. supra. Aunque estos son los promotores más utilizados habitualmente, se pueden utilizar también otros promotores adecuados.

**[0237]** El vector también puede incluir otras secuencias de ácido nucleico, por ejemplo,. Secuencias que codifican etiquetas gD, epítopos c-Myc, etiquetas de polihistidina, proteínas fluorescentes (por ejemplo, GFP), o proteína beta-galactosidasa que pueden ser útiles para la detección o purificación de la proteína de fusión expresada en la superficie del fago o la célula. El ácido nucleico que codifica, por ejemplo, una etiqueta gD también proporciona una selección positiva o negativa de células o virus que expresan la proteína de fusión. En algunas realizaciones, la etiqueta gD se fusiona a un dominio variable de anticuerpo que no está fusionado al componente de proteína de cubierta viral. Las secuencias de ácido nucleico que codifican, por ejemplo, una etiqueta de polihistidina, son útiles para la identificación de proteínas de fusión que incluyen dominios variables de anticuerpo que se unen a un antígeno específico utilizando inmunohistoquímica. Las etiquetas útiles para detección de unión a antígeno se pueden fusionar a un dominio variable de anticuerpo no fusionado a un componente de proteína de cubierta viral o un dominio variable de anticuerpo fusionado a un componente de proteína de cubierta viral.

**[0238]** Otro componente de los vectores utilizados para realizar esta invención es los genes de selección fenotípicos. Los genes de selección fenotípicos típicos son aquellos que codifican proteínas que confieren resistencia a antibiótico en la célula huésped. A modo de ilustración, se utilizan fácilmente para este objetivo el gen de resistencia a ampicilina (amp<sup>r</sup>), y el gen de resistencia a tetraciclina (tetr).

**[0239]** El vector también puede incluir secuencias de ácido nucleico que contienen sitios de restricción únicos y codones de parada suprimibles. Los sitios de restricción únicos son útiles para mover los dominios variables de anticuerpo entre diferentes vectores y sistemas de expresión, especialmente útiles para la producción de anticuerpo de longitud completa o fragmentos de unión a antígeno en cultivos celulares. Los codones de parada suprimibles son útiles para controlar el nivel de expresión de la proteína de fusión y para facilitar la purificación de fragmentos de anticuerpo solubles. Por ejemplo, un codón de parada amber se puede leer como Gln en un huésped supE para permitir la expresión en el fago, mientras que en un huésped no supE se lee como un codón de parada para producir fragmentos de anticuerpo solubles sin fusión a proteínas de cubierta del fago. Estas secuencias sintéticas se pueden fusionar a uno o más dominios variables de anticuerpo en el vector.

**[0240]** A veces es beneficioso utilizar sistemas de vectores que permiten que el ácido nucleico codifique una secuencia de anticuerpo de interés, por ejemplo una CDR que tiene variantes de aminoácidos, para extraerse fácilmente del sistema de vectores y colocarse en otro sistema de vectores. Por ejemplo, se pueden diseñar sitios de restricción apropiados en un sistema de vectores para facilitar la extracción de la secuencia de ácido nucleico que codifica un anticuerpo o dominio variable de anticuerpo que tiene variantes de aminoácidos. Las secuencias de restricción se eligen normalmente para que sean únicas en los vectores para facilitar la excisión y ligamiento eficiente en nuevos vectores. Los anticuerpos o dominios variables de anticuerpo se pueden expresar entonces a partir de vectores sin secuencias de fusión externas, tales como proteínas de cubierta viral u otras secuencias etiqueta.

**[0241]** Entre el ácido nucleico que codifica el dominio variable o constante de anticuerpo (gen 1) y el componente de proteína de cubierta viral (gen 2), se puede insertar el gen que codifica un codón de terminación o de parada, tal como codones de terminación que incluyen UAG (amber), UAA (ocher) y UGA (opel). (*Microbiology*, Davis et al., Harper & Row, New York, 1980, pp. 237, 245-47 y 374). El codón de terminación o parada expresado en una célula huésped de tipo natural da lugar a la síntesis del producto proteico del gen 1 sin la proteína del gen 2 unida. Sin embargo, el crecimiento en una célula huésped supresora da lugar a la síntesis de cantidades detectables de proteína fusionada. Dichas células huésped supresoras son conocidas y se han descrito, tales

como cepa supresora de E. coli (Bullock et al., BioTechniques 5:376-379 (1987)). Se puede utilizar cualquier método aceptable para colocar dicho codón de terminación en el ARNm que codifica el polipéptido de fusión.

5 **[0242]** El codón suprimible se puede insertar entre el primer gen que codifica un dominio variable o constante del anticuerpo y un segundo gen que codifica por lo menos una parte de una proteína de cubierta de fago. Alternativamente, el codón de terminación suprimible se puede insertar adyacente al sitio de fusión mediante la sustitución del último triplete de aminoácidos en el dominio variable de anticuerpo o el primer aminoácido en la proteína de cubierta del fago. El codón de terminación suprimible se puede localizar en o después del extremo C-terminal de un dominio de dimerización. Cuando el plásmido que contiene el codón suprimible se desarrolla en una célula huésped supresora, da lugar a la producción detectable de un polipéptido de fusión que contiene el polipéptido y la proteína de cubierta. Cuando el plásmido se desarrolla en una célula huésped no supresora, el dominio variable de anticuerpo se sintetiza sustancialmente sin fusión a la proteína de cubierta de fago debido a la terminación en el triplete suprimible insertado UAG, UAA, o UGA. En la célula no supresora, el dominio variable de anticuerpo se sintetiza y se secreta de la célula huésped debido a la ausencia de la proteína de cubierta de fago fusionada que de otro modo está anclada a la membrana huésped.

15 **[0243]** En algunos casos, la CDR que se diversifica (aleatoriza) puede tener un codón de parada modificado en la secuencia de plantilla (referido en este documento como "plantilla de parada"). Esta característica proporciona la detección y selección de secuencias diversificadas de manera satisfactoria en base a la reparación satisfactoria del codón o codones de parada en la secuencia plantilla debido a la incorporación del oligonucleótido u oligonucleótidos que comprende la secuencia o secuencias de variantes de aminoácidos de interés. Esta característica se ilustra adicionalmente en los siguientes ejemplos.

20 **[0244]** Los dominios variables o constantes de anticuerpo de cadena ligera y/o pesada también se pueden fusionar a una secuencia de péptido adicional, proporcionando ésta última la interacción de uno o polipéptidos de fusión en la superficie de la partícula viral o célula. Estas secuencias de péptidos se refieren aquí como "dominios de dimerización". Los dominios de dimerización pueden comprender por lo menos una o más de una secuencia de dimerización, o por lo menos una secuencia que comprende un residuo de cisteína o ambos. Las secuencias de dimerización adecuadas incluyen aquellas proteínas que tienen helices alfa antipáticas en que los residuos hidrofóbicos están espaciados regularmente y permiten la formación de un dímero mediante la interacción de los residuos hidrofóbicos de cada proteína; dichas proteínas y partes de proteínas incluyen, por ejemplo, regiones de cremallera de leucinas. Los dominios de dimerización pueden comprender también uno o más residuos de cisteína (por ejemplo, proporcionados por la inclusión de una secuencia bisagra de anticuerpo en el dominio de dimerización). Los residuos de cisteína pueden proporcionar la dimerización mediante la formación de uno o más enlaces disulfuro. En una realización, en la que está presente un codón de parada después del dominio de dimerización, el dominio de dimerización comprende por lo menos un residuo de cisteína. En algunas realizaciones, los dominios de dimerización se localizan entre el dominio variable o constante de anticuerpo y el componente de proteína de cubierta viral.

25 **[0245]** En algunos casos, el vector codifica un único polipéptido anticuerpo-fago en forma de una única cadena, por ejemplo, las regiones variables de cadena pesada y ligera fusionadas a una proteína de cubierta. En estos casos, el vector se considera que es "monocistrónico", expresando un transcrito bajo el control de un cierto promotor. Por ejemplo, un vector puede utilizar un promotor (tal como la fosfatasa alcalina (AP) o promotor Tac) para dirigir la expresión de una secuencia monocistrónica que codifica los dominios VL y VH con un péptido enlazador entre los dominios VL y VH. Esta secuencia cistrónica puede estar conectada en el extremo 5' a una secuencia señal (tal como una secuencia señal malE o enterotoxina II estable al calor (STII) de E. coli) y en su extremo 3' a toda o una parte de una proteína de cubierta viral (tal como la proteína del bacteriófago pIII). El polipéptido de fusión codificado por un vector de esta realización se refiere aquí como "ScFv-pIII". En algunas realizaciones, un vector puede comprender también una secuencia que codifica un dominio de dimerización (tal como una cremallera de leucina) en su extremo 3', entre la segunda secuencia de dominio variable (por ejemplo, VH) y la secuencia de la proteína de cubierta viral. Los polipéptidos de fusión que comprenden el dominio de dimerización son capaces de dimerizar para formar un complejo de dos polipéptidos scFv (referido aquí como "(ScFv) 2-pIII").

30 **[0246]** En otros casos, las regiones variables de las cadenas pesada y ligera se pueden expresar como polipéptidos separados, siendo el vector "bicistrónico", permitiendo la expresión de transcritos separados. En estos vectores, se utiliza un promotor adecuado, tal como el promotor Ptac o PhoA, para dirigir la expresión de un mensaje bicistrónico. Un primer cistrón que codifica, por ejemplo, un dominio variable y constante de cadena ligera, puede estar conectado en el extremo 5' a una secuencia señal, tal como la secuencia señal malE o enterotoxina II estable al calor (STII) de E. coli, y en extremo 3' a una secuencia de ácido nucleico que codifica una secuencia etiqueta, tal como la etiqueta gD. Un segundo cistrón que codifica, por ejemplo, un dominio variable y dominio constante CH1 de cadena pesada, está conectado en su extremo 5' a una secuencia señal, tal como la secuencia señal malE o enterotoxina II estable al calor (STII) de E. coli, y en el extremo 3' a toda o parte de una proteína de cubierta viral.

35 **[0247]** En una realización de un vector que proporciona un mensaje bicistrónico y para la expresión de F(ab')<sub>2</sub>-pIII, un promotor adecuado, tal como un promotor Ptac o PhoA (AP), dirige la expresión de un primer cistrón que

codifica un dominio variable y constante de cadena ligera unido operativamente en el extremo 5' a una secuencia señal, tal como la secuencia señal malE o enterotoxina II estable al calor (STII) de E. coli, y en el extremo 3' a una secuencia de ácido nucleico que codifica una secuencia etiqueta, tal como la etiqueta gD. El segundo cistrón codifica, por ejemplo, un dominio variable y constante de cadena pesada unido operativamente en 5' a una secuencia señal, tal como la secuencia señal malE o enterotoxina II estable al calor (STII) de E. coli, y en el extremo 3' tiene un dominio de dimerización que comprende una secuencia bisagra de IgG y una secuencia de cremallera de leucinas seguido de por lo menos una parte de proteína de cubierta viral.

### **Expresión de polipéptidos de fusión**

[0248] Los polipéptidos de fusión de un polipéptido que contiene CDR (por ejemplo, un dominio variable de anticuerpo) se pueden expresar en la superficie de una célula, virus o partícula fagémida en una variedad de formatos. Estos formatos incluyen fragmento Fv de cadena única (scFv), fragmento F(ab) y formas multivalentes de estos fragmentos. Por ejemplo, las formas multivalentes incluyen un dímero de ScFv, Fab, o F(ab'), referidos aquí como (ScFv)<sub>2</sub>, F(ab)<sub>2</sub> y F(ab')<sub>2</sub>, respectivamente. Las formas multivalentes de expresión son ventajosas en algunos contextos en parte porque tienen más de un sitio de unión a antígeno que en general da lugar a la identificación de clones de afinidad inferior y también permite una clasificación más eficaz de clones raros durante el proceso de selección.

[0249] Los métodos para expresar los polipéptidos de fusión que comprenden fragmentos de anticuerpos en la superficie del bacteriófago, son conocidos en la técnica, por ejemplo, tal como se describe en la publicación de patente número WO 92/01047 y aquí. Otras publicaciones de patente WO 92/20791; WO 93/06213; WO 93/11236 y WO 93/19172, describen métodos relacionados. Otras publicaciones han mostrado la identificación de anticuerpos con repertorios de genes V redistribuidos artificialmente contra una variedad de antígenos expresados en la superficie del fago (por ejemplo, H. R. Hoogenboom & G. Winter, J. Mol. Biol. 227 381-388 (1992); y tal como se describe en WO 93/06213 y WO 93/11236).

[0250] Cuando se construye un vector para la expresión en un formato scFv, incluye secuencias de ácido nucleico que codifican un dominio variable de cadena ligera de anticuerpo y un dominio variable de cadena pesada de anticuerpo. Habitualmente, la secuencia de ácido nucleico que codifica un dominio variable de cadena pesada de anticuerpo se fusiona a un componente de proteína de cubierta viral. Uno o ambos dominios variables del anticuerpo pueden tener variantes de aminoácidos en por lo menos una región de CDR. La secuencia de ácido nucleico que codifica la cadena ligera variable de anticuerpo está conectada al dominio variable de cadena pesada de anticuerpo por una secuencia de ácidos nucleicos que codifica un péptido enlazador. El péptido enlazador contiene habitualmente aproximadamente de 5 a 15 aminoácidos. Opcionalmente, otras secuencias que codifican, por ejemplo, etiquetas útiles para la purificación o detección se pueden fusionar al extremo 3' de la secuencia de ácido que codifica el dominio variable de cadena ligera de anticuerpo o el dominio variable de cadena pesada de anticuerpo o ambos.

[0251] Cuando se construye un vector para la expresión de F(ab), se incluye secuencias de ácidos nucleicos que codifican dominios variables de anticuerpo y dominios constantes de anticuerpo. Un ácido nucleico que codifica un dominio variable de cadena ligera se fusiona a una secuencia de ácido nucleico que codifica un dominio constante de cadena ligera. Una secuencia de ácido nucleico que codifica un dominio variable de cadena pesada de anticuerpo se fusiona a una secuencia de ácido nucleico que codifica un dominio constante CH1 de cadena pesada. Habitualmente, la secuencia de ácido nucleico que codifica los dominios variable y constante de cadena pesada se fusionan a una secuencia de ácido nucleico que codifica toda o una parte de una proteína de cubierta viral. Uno o ambos dominios variables de cadena ligera o pesada de anticuerpo pueden tener variantes de anticuerpo en por lo menos una CDR. En algunas realizaciones, los dominios variables y constantes de cadena pesada se expresan como una fusión con por lo menos una parte de una proteína de cubierta viral, y los dominios variable y constante de cadena ligera se expresan por separado de la proteína de fusión de cubierta viral de cadena pesada. Las cadenas pesada y ligera se asocian entre sí, lo cual puede ser mediante enlaces covalentes o no covalentes. Opcionalmente, se pueden fusionar otras secuencias que codifican, por ejemplo, etiquetas polipeptídicas útiles para la purificación o detección en el extremo 3' de la secuencia de ácido nucleico que codifica el dominio constante de la cadena ligera del anticuerpo o el dominio constante de la cadena pesada de anticuerpo o ambos.

[0252] En algunas realizaciones, se utiliza un grupo bivalente, por ejemplo, un dímero F(ab)<sub>2</sub> o F(ab')<sub>2</sub>, para expresar fragmentos de anticuerpo con las variantes de sustituciones de aminoácidos en la superficie de una partícula. Se ha observado que los dímeros F(ab')<sub>2</sub> presentan en general la misma afinidad que los dímeros F(ab) en un ensayo de unión a antígeno en fase solución pero la velocidad off para F(ab')<sub>2</sub> se reduce debido a una avidéz mayor. Por lo tanto, el formato bivalente (por ejemplo, F(ab')<sub>2</sub>) es un formato particularmente útil ya que permite la identificación de clones de afinidad inferior y también permite una clasificación más eficaz de clones raros durante el proceso de selección.

### **Introducción de vectores en células huésped**

[0253] Los vectores construidos tal como se ha descrito se introducen en una célula huésped para la

amplificación y/o expresión. Los vectores se pueden introducir en células huésped utilizando métodos de transformación estándar que incluyen electroporación, precipitación con fosfato de calcio, y similares. Si el vector es una partícula infecciosa, tal como un virus, el propio vector proporciona la entrada en la célula huésped. La transfección de células huésped que contienen un vector de expresión replicable que codifica la fusión de genes y la producción de partículas de fagos según los procedimientos estándar proporciona partículas de fagos en las que la proteína de fusión se expresa en la superficie de la partícula de fago.

**[0254]** Los vectores de expresión replicables se introducen en las células huésped utilizando una variedad de métodos. En una realización, los vectores se pueden introducir en las células utilizando electroporación tal como se describe en WO/00106717. Las células se desarrollan en cultivo en medio de cultivo estándar, opcionalmente durante aproximadamente 6-48 horas (o hasta DO600 = 0,6 – 0,8) a aproximadamente 37°C, y a continuación, el caldo se centrifuga y se extrae el sobrenadante (por ejemplo, se decanta). En algunos casos, la purificación inicial incluye la resuspensión del residuo celular en una solución tampón (por ejemplo, HEPES 1,0 mM, pH 7,4) seguido de la recentrifugación y la extracción del sobrenadante. El residuo celular resultante se resuspende en glicerol diluido (por ejemplo, 5-20% v/v) y de nuevo se recentrifuga para formar un residuo celular y se extrae el sobrenadante. La concentración celular final se obtiene mediante la resuspensión del residuo celular en agua o glicerol diluido hasta la concentración deseada.

**[0255]** En ciertos casos, la célula receptora es la cepa de E. Coli competente de electroporación de la presente invención, que es la cepa SS320 de E. Coli (Sidhu et al., Methods Enzymol. (2000), 328:333-363). La cepa SS320 se preparó mediante el apareamiento de células MC1061 con células XL1-BLUE bajo condiciones suficientes para transferir el episoma de fertilidad (plásmido F') o XL1-BLUE en las células MC1061. La cepa SS320 se ha depositado con la American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, Virginia USA, el 18 de junio de 1998 y se le ha asignado el Depósito de Acceso No. 98795. Cualquier episoma F' que permita la replicación de fagos en la cepa se puede utilizar en la invención. Los episomas adecuados están disponibles de cepas depositadas con la ATCC o están disponibles comercialmente (CJ236, CSH18, DHF', JM101, JM103, JM105, JM107, JM109, JM110), KS1000, XL1-BLUE, 71-18 y otros).

**[0256]** La utilización de concentraciones de ADN superiores durante la electroporación (aproximadamente 10 veces) incrementa la eficacia de transformación e incrementa la cantidad de ADN transformado en las células huésped. La utilización de concentración celulares elevadas también incrementa la eficacia (aproximadamente 10 veces). La mayor cantidad de ADN transferido produce bibliotecas más grandes que tienen una mayor diversidad y que representan un mayor número de miembros únicos de una biblioteca combinatoria. Las células transformadas se seleccionan en general mediante el crecimiento en un medio que contiene antibióticos.

#### **Selección (clasificación) y cribado de enlazadores para dianas de elección**

**[0257]** La utilización de la expresión de fagos para identificar enlazadores de antígeno diana, con sus diversas permutaciones y variaciones en la metodología, es bien conocida en la técnica. Una estrategia implica la construcción de una familia de variantes de vectores replicables que contienen un elemento regulador de la transcripción unido operativamente a una fusión de genes que codifica un polipéptido de fusión, la transformación de células huésped adecuadas, el cultivo de las células transformadas para formar partículas de fago que expresan el polipéptido de fusión en la superficie de la partícula del fago, seguido de un proceso que comprende la selección o clasificación mediante la puesta en contacto de las partículas de fago recombinantes con un antígeno diana, de manera que por lo menos una parte de la población de las partículas se unen a la diana con el objetivo de incrementar y enriquecer los subgrupos de partículas que se unen a partir de las partículas en relación con las partículas que no se unen en el proceso de selección. El grupo seleccionado se puede amplificar mediante la infección de células huésped, tales como células XL1-Blue frescas, para otra ronda de clasificación en la misma diana con la misma o diferente rigidez. El grupo resultante de variantes se criban a continuación contra los antígenos diana para identificar nuevas proteínas de unión con afinidad elevada. Estas nuevas proteínas de unión con afinidad elevada pueden ser útiles como agentes terapéuticos como antagonistas o agonistas y/o como reactivos de diagnóstico e investigación.

**[0258]** Los polipéptidos de fusión, tales como dominios variables de anticuerpo que comprenden variantes de aminoácidos, se pueden expresar en la superficie de un fago, partícula fagémida o una célula y, a continuación, se pueden seleccionar y/o cribar por la capacidad de los miembros del grupo de polipéptidos de fusión de unirse a un antígeno diana que es habitualmente un antígeno de interés. Los procesos de selección para los enlazadores a diana también pueden incluir la clasificación en una proteína genérica que tiene afinidad por dominios variables de anticuerpo, tales como proteína L o un anticuerpo específico etiqueta que se une a un anticuerpo o fragmentos de anticuerpo expresados en el fago, que se pueden utilizar para enriquecer los miembros de bibliotecas que expresan fragmentos de anticuerpo correctamente plegados (polipéptidos de fusión).

**[0259]** Las proteínas diana, tales como receptor DR5 y HER-2, se pueden aislar de fuentes naturales o prepararse mediante métodos recombinantes por procedimientos conocidos en la técnica. Las secuencias de DR5 humano y murino se proporcionan en la tabla 1. La secuencia y preparación de ECD de HER-2 se ha descrito en Franklin MC. Carey KD. Vajdos FF. Leahy DJ. de Vos AM. Sliwkowski MX., Insights into ErbB

signaling from the structure of the ErbB2-pertuzumab complex, Cancer Cell. 5 (4):317-28, 2004. La secuencia de los aminoácidos 23-646 del dominio extracelular de HER-2 se proporciona en el Protein DataBank Record 1878 (2004). Los antígenos diana pueden incluir un número de moléculas de interés terapéutico.

5 **[0260]** Se puede utilizar una variedad de estrategias de selección (clasificación) por afinidad. Un ejemplo es un método de soporte sólido o clasificación de placas o clasificación de dianas inmovilizadas. Otro ejemplo es un método de unión en solución.

10 **[0261]** Para el método de soporte sólido, la proteína diana se puede unir a una matriz sólida o semisólida adecuada. Dichas matrices son conocidas en la técnica, tales como perlas de agarosa, perlas de acrilamida, perlas de vidrio, celulosa, varios copolímeros acrílicos, geles de hidroxialquil metacrilato, copolímeros de poliacrílico y polimetacrílico, nylon, portadores neutros e iónicos, y similares. La unión de la proteína diana a la matriz se puede llevar a cabo mediante métodos descritos, por ejemplo, en *Methods in Enzymology*, 44 (1976), o mediante otros medios conocidos en la técnica.

15 **[0262]** Después de la unión del antígeno diana a la matriz, la diana inmovilizada se pone en contacto con la biblioteca que expresa los polipéptidos de fusión bajo condiciones adecuadas para la unión de por lo menos un subgrupo de la población de partículas de fago con antígeno diana inmovilizado. Normalmente, las condiciones, incluyendo el pH, la fuerza iónica, la temperatura y similares, imitarán las condiciones fisiológicas. Las partículas unidas ("enlazadores") a la diana inmovilizada se separan de las partículas que no se unen a la diana mediante lavado. Las condiciones de lavado se pueden ajustar para dar lugar a la eliminación de todos, a excepción de los enlazadores de afinidad elevada. Los enlazadores se pueden disociar de la diana inmovilizada mediante una serie de métodos. Estos métodos incluyen la disociación competitiva utilizando el ligando de tipo natural (por ejemplo, exceso de antígeno diana), alterando el pH y/o la fuerza iónica, y métodos conocidos en la técnica. La selección de enlazadores implica habitualmente la elución de una matriz de afinidad con un material de elución adecuado, tal como ácido como HCl 0,1 M o ligando. La elución con concentraciones crecientes de ligando podría eluir moléculas de unión expresadas de afinidad creciente.

25 **[0263]** Los enlazadores se pueden aislar y a continuación se pueden reamplificar en células huésped adecuadas mediante la infección de las células con las partículas virales que son enlazadores (y el fago auxiliar si es necesario, por ejemplo, cuando la partícula viral es una partícula fagémida) y las células huésped se cultivan bajo condiciones adecuadas para la amplificación de las partículas que expresan el polipéptido de fusión deseado. A continuación, las partículas de fagos se recogen y se repite el proceso de selección una o más veces hasta que se enriquecen los enlazadores del antígeno diana. Se puede utilizar cualquier cantidad de rondas de selección o clasificación. Uno de los procedimientos de selección o clasificación puede implicar el aislamiento de enlazadores que se unen a una proteína de afinidad genérica, tal como proteína L o un anticuerpo para una etiqueta polipeptídica presente en un polipéptido expresado, tal como un anticuerpo para la proteína gD o la etiqueta de polihistidina. Se puede incluir contraselección en una o más rondas de selección o clasificación para aislar enlazadores que también muestran una unión no deseada a uno o más que no son antígenos diana.

35 **[0264]** Se describe aquí un método que implica la selección contra bibliotecas de la invención utilizando un método de selección nuevo que se denomina "método de unión en solución". Éste permite la clasificación en fase solución con una eficiencia muy mejorada sobre métodos de clasificación en solución convencionales. El método de unión en solución se puede utilizar para hallar enlazadores originales de una biblioteca aleatoria o hallar enlazadores mejorados de una biblioteca que se designó para mejorar la afinidad de un clon o grupo de clones particulares de unión. El método comprende poner en contacto una pluralidad de polipéptidos, tales como los expresados en partículas de fagos o fagémidos (biblioteca), con un antígeno diana marcado o fusionado con una molécula etiqueta. La etiqueta podría ser biotina u otros grupos para los que están disponibles enlazadores específicos. El rigor de la fase en solución se puede variar mediante la utilización de concentraciones decrecientes de antígeno diana marcado en la primera fase de unión en solución. Para incrementar adicionalmente el rigor, la primera fase de unión en solución puede estar seguida de una segunda fase en solución que tiene una concentración elevada de antígeno diana no marcado después de la unión inicial con la diana marcada en la primera fase en solución. Normalmente, se utiliza de 100 a 1000 veces de diana no marcada sobre diana marcada en la segunda fase (si se incluye). El tiempo de incubación de la primera fase en solución puede variar de unos minutos a una a dos horas o más para alcanzar el equilibrio. Utilizando un tiempo más corto para la unión en esta primera fase se puede predisponer o seleccionar enlazadores que presentan una velocidad on rápida. El tiempo y la temperatura de incubación en la segunda fase pueden variar para incrementar el rigor. Esto proporciona una predisposición de la selección por enlazadores que presentan una velocidad baja de desprendimiento de la diana (velocidad off). Después de poner en contacto la pluralidad de polipéptidos (expresados en las partículas de fago/fagémido) con un antígeno diana, las partículas de fago o fagémido que están unidas a dianas marcadas se separan del fago que no se unen. La mezcla de partículas-diana de la fase en solución de unión se aísla al ponerla en contacto con el grupo diana marcado y permitiendo su unión a una molécula que se une al grupo diana marcado durante un periodo de tiempo corto (por ejemplo, 2-5 minutos). La concentración inicial del antígeno diana marcado puede variar de aproximadamente 0,1 nM a aproximadamente 1000 nM. Las partículas unidas se eluyen y se pueden propagar para la siguiente ronda de clasificación. En ciertos casos, se realizan múltiples rondas de clasificación utilizando una concentración inferior de antígeno diana marcado con cada ronda de clasificación.

- 5 [0265] Por ejemplo, una clasificación o selección inicial utilizando aproximadamente 100 a 250 nM de antígeno diana marcado debería ser suficiente para capturar un amplio rango de afinidad, aunque este factor se puede determinar empíricamente y/o adecuarse al deseo del experto. En la segunda ronda de selección, se puede utilizar de aproximadamente 25 a 100 nM de antígeno diana marcado. En la tercera ronda de selección, se puede utilizar de aproximadamente 0,1 a 0,25 nM de antígeno diana marcado. Por ejemplo, para mejorar la afinidad de un enlazador 100 nM, puede ser deseable empezar con 20 nM y a continuación progresar hasta 5 y 1 nM de diana marcada, a continuación, seguido por concentraciones incluso más bajas, tales como aproximadamente 0,1 nM de antígeno diana marcado.
- 10 [0266] La clasificación en solución convencional implica la utilización de perlas del tipo de perlas cubiertas de estreptavidina, que es muy engorrosa de utilizar y a menudo da lugar a una eficiencia muy baja de recuperación de enlazador de fago. La clasificación en solución convencional con perlas necesita mucho más tiempo que 2-5 minutos y es menos viable para adaptarse a la automatización de alto rendimiento que el método descrito anteriormente.
- 15 [0267] Tal como se describe aquí, se pueden utilizar ventajosamente combinaciones de métodos de clasificación en soporte sólido y solución para aislar enlazadores que tienen características deseadas. Después de la selección/clasificación de antígeno diana durante varias rondas, se realiza en general el cribado de clones individuales del grupo seleccionado para identificar enlazadores específicos con las propiedades/características deseadas. En algunos casos, el proceso de cribado se lleva a cabo mediante sistemas automatizados para permitir el cribado de alto rendimiento de candidatos de la biblioteca.
- 20 [0268] A continuación se describen dos métodos principales de cribado. Sin embargo, en los métodos de la invención también se pueden utilizar otros métodos conocidos en la técnica. El primer método de cribado comprende un ensayo de ELISA en fago con antígeno diana inmovilizado, lo cual proporciona la identificación de un clon de unión específica de un clon que no se une. La especificidad se puede determinar mediante el ensayo simultáneo del clon en el pocillo cubierto de diana y BSA u otros pocillos cubiertos de proteínas no diana. Este ensayo es automático para el cribado de alto rendimiento.
- 25 [0269] Se describe aquí un método de selección de un dominio variable de anticuerpo que se une a un antígeno diana específico de una biblioteca de dominios variables de anticuerpo mediante la generación de una biblioteca de vectores de expresión replicable que comprende una serie de polipéptidos; poner en contacto la biblioteca con un antígeno diana y por lo menos un antígeno no diana bajo condiciones adecuadas para la unión; separar los polipéptidos enlazadores en la biblioteca de los no enlazadores; identificar los enlazadores que se unen al antígeno diana y no se une al antígeno no diana; eluir los enlazadores del antígeno diana; y amplificar los vectores de expresión replicables que comprenden el polipéptido enlazador que se une a un antígeno específico.
- 30 [0270] El segundo ensayo de cribado es un ensayo de cribado por afinidad que proporciona el cribado de clones que presentan una afinidad elevada de clones que presentan una afinidad baja en un cribado de rendimiento elevado. En el ensayo, cada clon se analiza con y sin incubarse primero con el antígeno diana de cierta concentración durante un periodo de tiempo (por ejemplo, 30-60 minutos) antes de la aplicación a pocillos cubiertos de diana brevemente (por ejemplo, 5-15 minutos). A continuación, se mide el fago unido mediante un método habitual de ELISA en fagos, por ejemplo, utilizando los conjugados anti-M13 HRP. La proporción de la señal de unión de los dos pocillos, un pocillo habiéndose preincubado con diana y el otro pocillo no preincubado con antígeno diana es una indicación de la afinidad. La selección de la concentración de diana para una primera incubación depende del intervalo de afinidad de interés. Por ejemplo, si se desean enlazadores con una afinidad superior a 10 nM, a menudo se usa 100 nM de diana en la primera incubación. Una vez se hallan los enlazadores de una ronda particular de clasificación (selección), estos clones se pueden cribar con un ensayo de cribado de afinidad para identificar enlazadores con una afinidad más elevada.
- 35 [0271] Las combinaciones de cualquier método de clasificación/selección descritos anteriormente se pueden combinar con los métodos de cribado. Por ejemplo, en una realización, se selecciona primero los polipéptidos enlazadores para la unión a antígeno diana inmovilizado. Los polipéptidos enlazadores que se unen al antígeno diana inmovilizado se pueden entonces amplificar y cribar por la unión al antígeno diana y por la falta de unión a antígenos no dianas. Se amplifican los polipéptidos enlazadores que se unen específicamente al antígeno diana. Estos polipéptidos enlazadores se pueden seleccionar a continuación por una mayor afinidad por contacto con una concentración de un antígeno diana marcado para formar un complejo, donde la concentración de antígeno diana marcado varía desde aproximadamente 0,1 nM a aproximadamente 1000 nM, los complejos se aíslan mediante el contacto con un agente que se une al marcador en el antígeno diana. A continuación, los polipéptidos enlazadores se eluyen del antígeno diana marcado y opcionalmente, las rondas de selección se repiten, se utiliza cada vez una concentración menor de antígeno diana marcado. Los polipéptidos enlazadores de afinidad elevada aislados utilizando este método de selección se pueden cribar a continuación por una afinidad elevada utilizando una variedad de métodos conocidos en la técnica, algunos de los cuales aquí descritos.
- 45 [0272] Estos métodos pueden proporcionar el hallazgo de clones con una afinidad elevada sin tener que realizar ensayos de afinidad por competición largos y complejos en un gran número de clones. El aspecto intenso de
- 50
- 55
- 60



realizar ensayos complejos de muchos clones a menudo es un obstáculo significativo para encontrar los mejores clones de una selección. Este método es especialmente útil en esfuerzos de mejora de la afinidad donde se pueden recuperar múltiples enlazadores con afinidad similar del proceso de selección. Clones diferentes pueden tener una eficiencia muy diferente de expresión en partículas de fago o fagémidos. Los clones más altamente expresados tienen mejores oportunidades de recuperarse. Es decir, la selección se puede predisponer por el nivel de expresión de las variantes. El método de clasificación por unión en solución de la invención puede mejorar el proceso de selección para hallar enlazadores con afinidad elevada. Este método es un ensayo de cribado de afinidad que proporciona una ventaja significativa en el cribado para los mejores enlazadores de manera rápida y fácil.

5  
10 **[0273]** Los anticuerpos o fragmentos de unión a antígeno se pueden seleccionar adicionalmente por la actividad funcional, por ejemplo, la actividad antagonista o agonista. Por ejemplo, se pueden seleccionar anticuerpos anti-HER2 por la capacidad de inhibir la fosforilación de tirosina de HER-2, la proliferación de células cancerosas o de inducir la apoptosis de células cancerosas. Los ensayos para identificar y medir estas actividades se describen por ejemplo en WO98/17797.

15 **[0274]** Además, los anticuerpos anti-DR5 se pueden seleccionar por la capacidad de inducir la apoptosis de células cancerosas y/o inhibir la función de células inflamatorias. En otras realizaciones, los anticuerpos anti-DR5 se seleccionan por la capacidad de competir con Apo-2L para la unión a DR5. En otras realizaciones, los anticuerpos anti-DR5 humano se seleccionan por la unión a DR5 murino y/o cinomolgo. Los ensayos para determinar la actividad biológica se pueden realizar utilizando métodos conocidos en la técnica, tales como la fragmentación de ADN (véase, por ejemplo, Marsters et al., Curr. Biology, 6: 1669 (1996)), inactivación de caspasa, unión a DR5 (véase, por ejemplo, WO98/51793, publicada el 19 de noviembre de 1998). La apoptosis se puede medir mediante la identificación de la condensación del citoplasma, la pérdida de las microvellosidades de la membrana plasmática, segmentación del núcleo, degradación del ADN cromosómico o pérdida de función mitocondrial. Esta actividad se puede determinar y medir, por ejemplo, mediante ensayos de viabilidad celular (tal como ensayos de azul Alamar o ensayos MTT), análisis FACS, activación de caspasa, fragmentación de ADN (véase, por ejemplo, Nicoletti et al., J. Immunol. Methods, -139:271-279 (1991), y poli-ADP ribosa polimerasa, "PARP", ensayos de separación conocidos en la técnica.

20  
25  
30 **[0275]** Tal como se describe aquí, un ensayo para la apoptosis implica producir diluciones en serie de dos veces de patrón de control y anticuerpo anti-DR5 en placas de cultivo de tejido de 96 pocillos. Se analiza para comparación el ligando Apo-2 (aminoácidos 114-281, descrito en PCT US00/17579). Se siembran células de carcinoma de colon humano Colo-205 (20000 células-pocillo) (ATCC) en placas de 96 pocillos. Las placas se incuban a 37°C durante 24 horas. Se añade AlmazBlue (Trek Diagnostic Systems, Inc.) a los pocillos durante las últimas 3 horas del tiempo de incubación de 24 horas. La fluorescencia se lee utilizando un fluorómetro de 96 pocillos con excitación a 530 nm y emisión de 590 nm. Los resultados se expresan en unidades relativas de fluorescencia (RFU).

35  
40 **[0276]** Después de identificar los enlazadores mediante la unión al antígeno diana, y/o ensayos funcionales, se puede extraer el ácido nucleico. El ADN extraído se puede utilizar directamente a continuación para transformar células huésped de E. coli o alternativamente, se pueden amplificar las secuencias codificantes, por ejemplo utilizando PCR con cebadores adecuados, y se secuencia mediante cualquier método de secuenciación habitual. El ADN de dominio variable de los enlazadores se puede digerir con enzimas de restricción y a continuación se inserta en un vector para la expresión de las proteínas.

45 **[0277]** Las poblaciones que comprenden polipéptidos que tienen CDR con una diversidad de secuencia limitada se pueden utilizar para aislar enlazadores contra una serie de dianas, incluyendo las indicadas en las figuras 11, 15, 21A, y 24A. Estos enlazadores pueden comprender una o más variantes de CDR que comprenden secuencias diversas generadas utilizando codones limitados. En algunos casos, una variante de CDRH3 que comprende diversidad de secuencia generada por la sustitución de aminoácidos por grupos de codones limitados y/o inserciones de aminoácidos resultantes de variar las longitudes de CDRH3. Los oligonucleótidos ilustrativos útiles para la generación de polipéptidos de fusión de la invención incluyen los indicados en la figura 9A-D, la figura 20A-L, y la figura 23. Se pueden combinar una o más variantes de CDR. En algunos casos, sólo se diversifica CDRH3. En otros casos, son variantes dos o más CDR de cadena pesada, incluyendo CDRH3. En otros casos, son variantes una o más CDR de cadena pesada excluyendo CDRH3. En algunos casos, son variantes por lo menos una CDR de cadena pesada y por lo menos una CDR de cadena ligera. En algunos casos, son variantes por lo menos una, dos, tres, cuatro, cinco o todas las CDR H1, H2, H3, L1, L2 y L3.

50  
55 **[0278]** En algunos casos, puede ser beneficioso combinar una o más CDR de cadena ligera diversificada con nuevos enlazadores aislados de una población de polipéptidos que comprenden una o más CDR de cadena pesada diversificada. Este proceso se puede referir como un proceso de 2 etapas. Un ejemplo de un proceso de 2 etapas comprende primero determinar enlazadores (generalmente enlazadores de afinidad inferior) en una o más bibliotecas generadas aleatorizando una o más CDR, donde las CDR aleatorizadas en cada biblioteca son diferentes o, cuando la misma CDR está aleatorizada, se aleatorizan para generar diferentes secuencias. Los enlazadores de una biblioteca de cadena pesada se pueden aleatorizar a continuación con una diversidad de CDR en CDR de cadena ligera mediante, por ejemplo, una técnica de mutagénesis, tal como la de Kunkel, o

mediante clonación (copiar y pegar (por ejemplo, mediante la unión de secuencias de CDR diferentes juntas)), de la nueva biblioteca de cadenas ligeras en los enlazadores de cadena pesada existentes que tienen sólo una cadena ligera fijada. El grupo se puede clasificar a continuación contra una o más dianas para identificar enlazadores que poseen una afinidad incrementada. Por ejemplo, los enlazadores (por ejemplo, enlazadores de baja afinidad) obtenidos de la clasificación de H1/H2/H3 se pueden fusionar con una biblioteca de una diversidad de L1/L2/L3 para sustituir su L1/L2/L3 original fijada, donde las nuevas bibliotecas se clasifican a continuación adicionalmente contra una diana de interés para obtener otro grupo de enlazadores (por ejemplo, enlazadores de afinidad elevada). Se pueden identificar nuevas secuencias de anticuerpos que muestran una mayor afinidad de unión a cualquiera de un conjunto de antígenos diana.

5  
10  
15 [0279] En algunos casos, las bibliotecas que comprenden polipéptidos se someten a un conjunto de rondas de clasificación, donde cada ronda de clasificación comprende poner en contacto los enlazadores obtenidos de la ronda anterior con un antígeno diana distinto del antígeno o antígenos diana de la ronda o rondas anteriores. Preferiblemente, pero no necesariamente, los antígenos diana son homólogas en la secuencia, por ejemplo miembros de una familia de polipéptidos relacionados, pero diferentes, tales como, pero sin limitación, citoquinas (por ejemplo, subtipos alfa interferón).

### **Generación de bibliotecas que comprenden polipéptidos que contienen variantes de CDR**

20 [0280] Las bibliotecas de variantes de polipéptidos de CDR se pueden generar mediante la mutación de las posiciones accesibles al disolvente y/o altamente diversas en por lo menos una CDR de un dominio variable de anticuerpo. Algunas o todas las CDR se pueden mutar. En algunos casos, puede ser preferible generar diversas bibliotecas de anticuerpos mediante la mutación de posiciones en CDRH1, CDRH2 y CDRH3 para formar una única biblioteca o mediante la mutación de posiciones en CDR3 y CDRH3 para formar una única biblioteca o mediante la mutación de posiciones en CDRL3 y CDRH1, CDRH2 y CDRH3 para formar una única biblioteca.

25 [0281] Se puede generar una biblioteca de dominios variables de anticuerpo, por ejemplo, con mutaciones en las posiciones accesibles a disolvente y/o altamente diversas de CDRH1, CDRH2 y CDRH3. Se puede generar otra biblioteca con mutaciones en CDRL1, CDRL2 y CDRL3. Estas bibliotecas también se pueden utilizar conjuntamente entre sí para generar enlazadores de afinidades deseadas. Por ejemplo, después de una o más rondas de selección de bibliotecas de cadena pesada para la unión a un antígeno diana, se puede sustituir una biblioteca de cadenas ligeras en la población de enlazadores de cadena pesada para rondas adicionales de selección para incrementar la afinidad de los enlazadores.

30 [0282] En un caso, se crea una biblioteca mediante la sustitución de aminoácidos originales con un grupo limitado de variantes de aminoácidos en la región CDRH1, CDRH2, y/o CDRH3 de la región variable de la secuencia de cadena pesada y/o la región CDRL3 de la región variable de la secuencia de cadena ligera.

[0283] Esta biblioteca puede contener una serie de secuencias de anticuerpo, en las que la diversidad de secuencia está principalmente en la región CDRH3 de la secuencia de cadena pesada.

35 [0284] En un aspecto, la biblioteca se crea en el contexto de la secuencia de anticuerpo humanizado 4D5, o la secuencia de los aminoácidos del armazón de la secuencia de anticuerpo humanizado 4D5. En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52-54, 56, y 58 de CDRH2, los residuos 95, 96, 97, 98, 99, 100, 100a, 100b, y 100c de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 8 para a biblioteca "YSGR-A". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52-54, 56, y 58 de CDRH2, los residuos 95, 96, 97, 98, 99, 100, 100a, 100b, y 100c de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 8 para la biblioteca "YSGR-B". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52-54, 56, y 58 de CDRH2, los residuos 95, 96, 97, 98, 99, 100, 100a, 100b, y 100c de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 8 para la biblioteca "YSGR-C". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52-54, 56, y 58 de CDRH2, los residuos 95, 96, 97, 98, 99, 100, 100a, 100b, y 100c de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 8 para la biblioteca "YSGR-D". Las posiciones de aminoácidos de 100b y 100c pueden tener un identificador alfabético diferente dependiendo de la longitud de CDRH3, pero estas posiciones son las dos últimas posiciones de CDRH3 antes de la posición 101. Ejemplos de secuencias adecuadas de oligonucleótido incluyen, pero sin limitación, las indicadas en las figuras 20A-20L, y se pueden determinar por un experto en la materia según los criterios aquí descritos.

55 [0285] En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19A para la biblioteca "SAH3". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de

CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19A para la biblioteca "SCH3". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19A para la biblioteca "SFH3". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19A para la biblioteca "SGH3". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19A para la biblioteca "SIH3". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19A para la biblioteca "SLH3". A medida que varía la longitud de CDRH3, las últimas dos posiciones antes de la posición 101 puede tener un identificador alfabético diferente dependiendo de la longitud de CDRH3, pero estas posiciones son las dos últimas posiciones de CDRH3 antes de la posición 101. Ejemplos de secuencias adecuadas de oligonucleótido incluyen, pero sin limitación, las indicadas en las figuras 20A-20L, y se pueden determinar por un experto en la materia según los criterios aquí descritos.

**[0286]** En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52,53,54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19B para la biblioteca "SNH3". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19B para la biblioteca "SPH3". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19B para la biblioteca "SRH3". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19B para la biblioteca "STH3". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19B para la biblioteca "SWH3". En ciertas realizaciones, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 19B para la biblioteca "SYH3". A medida que varía la longitud de CDRH3, las últimas dos posiciones antes de la posición 101 puede tener un identificador alfabético diferente dependiendo de la longitud de CDRH3, pero estas posiciones son las dos últimas posiciones de CDRH3 antes de la posición 101. Ejemplos de secuencias adecuadas de oligonucleótido incluyen, pero sin limitación, las indicadas en las figuras 20A-20L, y se pueden determinar por un experto en la materia según los criterios aquí descritos.

**[0287]** Se describe aquí una biblioteca creada mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, residuos 50, 52, 53, 54, 56, y 58 de CDRH2, residuos 95-100m de CDRH3, y residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestra en la figura 22 para la biblioteca "SY". En ciertos casos, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestran en la figura 22 para la biblioteca "SW". En ciertos casos, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestran en la figura 22 para la biblioteca "SR". En ciertos casos, la biblioteca se crea mediante la sustitución de por lo menos los residuos 28 y 30-33 de CDRH1, los residuos 50, 52, 53, 54, 56, y 58 de CDRH2, los residuos 95-100m de CDRH3, y los residuos 91-94 y 96 de CDRL3 con los aminoácidos establecidos tal como se muestran en la figura 22 para la biblioteca "SF". A medida que varía la longitud de CDRH3, las últimas dos posiciones antes de la posición 101 pueden tener un identificador alfabético diferente dependiendo de la longitud de la CDRH3, pero estas posiciones son las dos últimas posiciones de CDRH3 antes de la posición 101. Ejemplos de secuencias adecuadas de oligonucleótidos incluyen, pero sin limitación, las indicadas en la figura 23, y se pueden determinar por un experto en la materia según los criterios aquí descritos.

**[0288]** En ciertos casos, se crea una biblioteca agrupando otras bibliotecas. En una realización, la biblioteca "SXH3" tal como se utiliza aquí comprende las bibliotecas SAH3, SCH3, SFH3, SGH3, SIH3, SLH3, SNH3, SPH3, SRH3, STH3, SWH3, y SYH3. En otro caso, la biblioteca "SX-superficie" comprende las bibliotecas "SY", "SW", "SR", y "SF".

**[0289]** En otro caso, se utilizan diferentes diseños de CDRH3 para aislar enlazadores de afinidad elevada y para

aislar enlazadores para una variedad de epítomos. Para la diversidad en CDRCH3, se pueden construir múltiples bibliotecas por separado con diferentes longitudes de H3 y a continuación combinarse para seleccionar enlazadores a antígenos diana. El rango de longitudes de CDRH3 generadas en esta biblioteca puede ser de 10-21, 11-21, 12-21, 13-21, 14-21, 15-21, 16-21, 17-21, 18-21, 19-21, 20-21, aminoácidos, aunque también se pueden generar longitudes diferentes a partir de éstas. La diversidad también se puede generar en CDRH1 y CDRH2, tal como se ha indicado anteriormente. En una realización de una biblioteca, la diversidad en H1 y H2 se genera utilizando los oligonucleótidos ilustrados en las figuras 9A-D, 20A-L, y la figura 23. También se pueden utilizar otros oligonucleótidos con secuencias variables. Los oligonucleótidos se pueden utilizar individualmente o agruparse en cualquiera de una variedad de combinaciones dependiendo de las necesidades prácticas y los deseos del experto. En algunas realizaciones, las posiciones aleatorizadas en CDR de cadena pesada incluyen las indicadas en las figuras 8, 9, 11, 15, 19, 21, 22, 23, y 24.

[0290] Se pueden agrupar múltiples bibliotecas y clasificarse utilizando métodos de selección en soporte sólido y clasificación en solución tal como se describen aquí. Se pueden utilizar múltiples estrategias de clasificación. Por ejemplo, una variación implica la clasificación en diana unida a un sólido, seguida de la clasificación de una etiqueta que puede estar presente en el polipéptido de fusión (por ejemplo, anti-etiqueta gD) y seguida de otra clasificación de diana unida a sólido. Alternativamente, las bibliotecas se pueden clasificar primero en la diana unida a una superficie sólida, a continuación se clasifican los enlazadores eluidos utilizando la unión en fase de solución con concentraciones decrecientes de antígeno diana. Utilizando combinaciones de diferentes métodos de clasificación se proporciona una minimización de la selección de sólo secuencias altamente expresadas y proporciona una selección de un número de diferentes clones de afinidad elevada.

[0291] De los enlazadores aislados de las bibliotecas agrupadas tal como se han descrito anteriormente, se ha descubierto que, en algunos casos, se puede mejorar adicionalmente la afinidad mediante la disposición de una diversidad limitada en la cadena ligera. La diversidad de cadena ligera puede ser, pero no necesariamente, generada mediante la diversificación de las posiciones de aminoácidos 91-96 en CDRL3 o un subgrupo de las mismas. En un ejemplo, las posiciones aleatorizadas son aquellas indicadas en las figuras 8, 9, 11, 15, 19, 21, 22, 23, y 24.

[0292] Los enlazadores de afinidad elevada aislados de las bibliotecas se producen fácilmente en el cultivo celular bacteriano y eucariota en rendimiento elevado. Los vectores se pueden diseñar para extraer fácilmente secuencias, tales como secuencias de etiqueta gD, componente de proteína de cubierta viral y/o para añadir en las secuencias de región constante para proporcionar la producción de anticuerpos de longitud completa o fragmentos de unión a antígeno en un rendimiento elevado. Según los métodos de la invención se puede diversificar cualquier combinación de grupos de codones y CDR.

### **Vectores, células huésped y métodos recombinantes**

[0293] Para la producción recombinante de un polipéptido anticuerpo, el ácido nucleico que lo codifica se aísla e inserta en un vector replicable para la clonación posterior (amplificación del ADN) o para la expresión. El ADN que codifica el anticuerpo se aísla fácilmente y se secuencia utilizando procedimientos convencionales (por ejemplo, utilizando sondas de oligonucleótidos que son capaces de unirse específicamente a genes que codifican las cadenas pesada y ligera del anticuerpo). Existen muchos vectores. La elección del vector depende en parte de la célula huésped a utilizar. En general, las células huésped son de origen procarionota o eucariota (en general mamífero)

#### **Generación de anticuerpos utilizando células huésped procarionotas:**

##### **Construcción de vectores**

[0294] Las secuencias de polinucleótidos que codifican los componentes polipeptídicos del anticuerpo se pueden obtener utilizando técnicas de recombinación estándar. Las secuencias de polinucleótidos deseadas se pueden aislar y secuenciar de células productoras de anticuerpos, tales como células de hibridoma. Alternativamente, los polinucleótidos se pueden sintetizar utilizando un sintetizador de nucleótidos o técnicas PCR. Una vez obtenidos, las secuencias que codifican los polipéptidos se insertan en un vector recombinante capaz de replicar y expresar polinucleótidos heterólogos en huéspedes procarionotas. Existen muchos vectores disponibles y conocidos en la técnica que se pueden utilizar para el objetivo de la presente invención. La selección del vector apropiado dependerá principalmente del tamaño de los ácidos nucleicos a insertar en el vector y de la célula huésped particular a transformar con el vector. Cada vector contiene varios componentes dependiendo de su función (amplificación o expresión de polinucleótidos heterólogo, o ambos) y de su compatibilidad con la célula huésped particular en la que reside. Los componentes del vector incluyen en general, pero sin limitación, un origen de replicación, un gen marcador de selección, un promotor, un sitio de unión a ribosoma (RBS), una secuencia señal, el inserto de ácido nucleico heterólogo y una secuencia de terminación de la transcripción.

[0295] En general, los vectores de plásmidos que contienen replicón y secuencias de control que derivan de especies compatibles con la célula huésped se utilizan en relación con estos huéspedes. El vector contiene normalmente un sitio de replicación, así como secuenciar marcadoras que son capaces de proporcionar una

selección fenotípica en células transformadas. Por ejemplo, E. Coli se transforma normalmente utilizando pBR322, un plásmido derivado de una especie de E. Coli. pBR322 contiene genes que codifican resistencia a ampicilina (Amp) y tetraciclina (Tet) y, de este modo, proporciona medios fáciles para identificar células transformadas. pBR322, sus derivados, o u otros plásmidos microbianos o bacteriófagos también pueden  
5 contener, o ser modificados para contener, promotores que se pueden utilizar por el organismo microbiano para la expresión de proteínas endógenas. Ejemplos de derivados de pBR322 utilizados para la expresión de anticuerpos concretos se describen en detalle en Carter et al., Patente de estados Unidos No. 5,648,237.

[0296] Además, los vectores de fagos que contienen replicón y secuencias de control que son compatibles con el microorganismo huésped se pueden utilizar como vectores transformantes en relación con estos huéspedes. Por  
10 ejemplo, se puede utilizar un bacteriófago como  $\lambda$ GEM.TM.-11 en la fabricación de un vector recombinante que se puede utilizar para transformar células huéspedes susceptibles, tales como E. coli LE392.

[0297] El vector de expresión puede comprender dos o más parejas promotor-cistrón, que codifican cada uno de los componentes polipeptídicos. Un promotor es una secuencia reguladora no traducida localizada en dirección 5' con respecto al cistrón que modula su expresión. Los promotores procariontes se clasifican normalmente en dos  
15 clases, inducible y constitutivo. El promotor inducible es un promotor que inicia mayores niveles de transcripción del cistrón bajo su control en respuesta a los cambios en la condición del cultivo, por ejemplo, la presencia o ausencia de un nutriente o un cambio en la temperatura.

[0298] Se conoce una gran cantidad de promotores reconocidos por una serie de células huésped potenciales. El promotor seleccionado se puede unir operativamente a ADN de cistrón que codifica la cadena ligera o pesada  
20 mediante la extracción del promotor del ADN de origen a través de la digestión con enzima de restricción y la inserción de la secuencia del promotor aislada en el vector de la presente invención. Se pueden utilizar tanto la secuencia del promotor nativo como de promotores heterólogos para dirigir la amplificación y/o expresión de los genes diana. En algunas realizaciones, se utilizan promotores heterólogos, ya que permiten en general una mayor transcripción y rendimientos más elevados del gen diana expresado en comparación con el promotor de polipéptido diana nativo.  
25

[0299] Los promotores adecuados para su uso con huéspedes procariontes incluyen el promotor PhoA, sistemas de promotores de la  $\beta$ -galactamasa y lactosa, un sistema de promotores de triptófano (trp) y promotores híbridos, tales como el promotor tac tac o el promotor trc. Sin embargo, también son adecuados otros promotores que son  
30 funcionales en las bacterias (tales como otros promotores bacterianos o de fagos conocidos). Se han publicado sus secuencias de nucleótidos, permitiendo así a un técnico unirlos a cistrones que codifican las cadenas ligera y pesada diana (Siebenlist et al. (1980) Cell 20: 269) utilizando enlazadores o adaptadores para suministrar cualquier sitio de restricción requerido.

[0300] En un caso, cada cistrón en el vector recombinante comprende un componente de la secuencia señal de secreción que dirige la translocación de los polipéptidos expresados a través de la membrana. En general, la  
35 secuencia señal puede ser un componente del vector, o puede ser una parte del ADN polipeptídico diana que se inserta en el vector. La secuencia señal seleccionada para el objetivo de la presente invención debería ser aquella que es reconocida y procesada (es decir, dividida por una peptidasa señal) por la célula huésped. Para células huésped procariontes que no reconocen y procesan las secuencias señal nativas a los polipéptidos heterólogos, la secuencia señal es sustituida por una secuencia señal procarionte seleccionada, por ejemplo, del  
40 grupo que consiste en secuencias líderes de fosfatasa alcalina, penicilinas, lpp, o enterotoxina II (STII) estable al calor, LamB, PhoE, PelB, OmpA y MBP. En un caso, las secuencias señal utilizadas en ambos cistrones del sistema de expresión son secuencias señal STII o variantes de la misma.

[0301] En otro caso, la producción de las inmunoglobulinas según la presente invención puede tener lugar en el citoplasma de la célula huésped, y, por tanto, no requiere la presencia de secuencias señal de secreción en cada  
45 cistrón. En este aspecto, se expresan las cadenas ligera y pesada de las inmunoglobulinas, se pliegan y se ensamblan para formar inmunoglobulinas funcionales en el citoplasma. Ciertas cepas huésped (por ejemplo, las cepas trx-B de E. coli) proporcionan condiciones del citoplasma que son favorables para la formación de enlaces disulfuro, permitiendo así el pliegue y ensamblaje correctos de subunidades de proteínas expresadas. Proba y Pluckthun Gene, 159:203 (1995).

[0302] Se describe aquí un sistema de expresión en el que la proporción cuantitativa de los componentes del polipéptido expresado se puede modular con el fin de maximizar el rendimiento de los anticuerpos secretados y  
50 correctamente ensamblados. Dicha modulación se lleva a cabo por lo menos en parte mediante la modulación simultánea de fuerzas traduccionales para los componentes del polipéptido.

[0303] Una técnica para modular la fuerza traduccional se describe en Simmons et al., Patente de Estados Unidos. No. 5,840,523. Utiliza variantes de la región de inicio de la traducción (TIR) en un cistrón. Para una TIR  
55 determinada, se puede crear una serie de variantes de la secuencia de aminoácidos o ácidos nucleicos con un intervalo de fuerzas traduccionales, proporcionando así un medio conveniente por el cual ajustar este factor al nivel de expresión deseado de la cadena específica. Las variantes de TIR se pueden generar mediante técnicas de mutagénesis convencionales que dan lugar a cambios de codones que pueden alterar la secuencia de

aminoácidos, aunque se prefieren cambios silenciosos en la secuencia de nucleótidos. Las alteraciones en la TIR pueden incluir, por ejemplo, alteraciones en el número o espaciado de las secuencias Shine-Dalgarno, junto con alteraciones en la secuencia señal. Un método para generar secuencias señal mutantes es la generación de un "banco de codones" en el inicio de una secuencia codificante que no cambia la secuencia de aminoácidos de la secuencia señal (es decir, los cambios son silenciosos). Esto se puede conseguir mediante el cambio de la

5 tercera posición de nucleótidos de cada codón; adicionalmente, algunos aminoácidos, tales como leucina, serina y arginina, tienen múltiples primeras y segundas posiciones que pueden añadir complejidad en la fabricación del banco. Este método de mutagénesis se describe en detalle en Yansura et al. (1992) METHODS: A Companion to Methods in Enzymol. 4:151-158.

10 **[0304]** En ciertas realizaciones, se genera un conjunto de vectores con un intervalo de fuerzas TIR para cada cistrón en el mismo. Este conjunto limitado proporciona una comparación de los niveles de expresión de cada cadena, así como el rendimiento de los productos de anticuerpo deseados bajo varias combinaciones de fuerza de TIR. Las fuerzas de TIR se pueden determinar mediante la cuantificación del nivel de expresión de un gen informador tal como se describe en detalle en Simmons et al. Patente de Estados Unidos No. 5, 840,523. En

15 base a la comparación de la fuerza traduccional, las TIR individuales deseadas se seleccionan para combinarse en las construcciones de vectores de expresión de la presente invención.

**[0305]** Entre las células huésped procariotas adecuadas para expresar anticuerpos se incluyen arqueobacterias y eubacterias, tales como organismos Gram-negativo o Gram-positivo. Ejemplos de bacterias útiles incluyen *Escherichia* (por ejemplo, *E. coli*), *Bacilli* (por ejemplo, *B. subtilis*), Enterobacteria, especies de *Pseudomonas* (por ejemplo, *P. aeruginosa*), *Salmonella typhimurium*, *Serratia marcescans*, *Klebsiella*, *Proteus*, *Shigella*, *Rhizobia*, *Vitreoscilla*, o *Paracoccus*. En una realización, se utilizan células gram negativas. En una realización, se utilizan células *E. coli* como huéspedes para la invención. Ejemplos de cepas de *E. coli* incluyen la cepa W3110 (Bachmann, Cellular and Molecular Biology, vol. 2 (Washington, D.C.: American Society for Microbiology, 1987), pp. 1190-1219; depósito ATCC No. 27,325) y derivados de la misma, incluyendo la cepa 33D3 que tiene el genotipo W3110  $\Delta$ fhuA ( $\Delta$ tonA) ptr3 lac lq lacL8  $\Delta$ ompT( $\Delta$ nmpc-fepE) degP41 kanR (patente de Estados Unidos No. 5,639,635). También son adecuadas otras cepas y derivados de las mismas, tales como *E. coli* 294 (ATCC 31,446), *E. coli* B, *E. coli* 1776 (ATCC 31,537) y *E. coli* RV308 (ATCC 31,608). Estos ejemplos son más ilustrativos que limitantes. Los métodos para construir derivados de cualquiera de las bacterias mencionadas anteriormente que tiene genotipos definidos son conocidos en la técnica y se describen en, por ejemplo, Bass et al., Proteins 8:309-314 (1990). Generalmente es necesario seleccionar la bacteria apropiada tomando en consideración la replicabilidad del replicón en las células de una bacteria. Por ejemplo, se pueden utilizar de manera adecuada especies *E. coli*, *Serratia*, o *Salmonella* como huésped cuando se utilizan plásmidos conocidos, tales como pBR322, pBR325, pACYC177, o pKN410 para suministrar el replicón. Habitualmente, la célula huésped debe secretar mínimas cantidades de enzimas proteolíticas y de manera deseable se pueden incorporar inhibidores de proteasa adicionales en el cultivo celular.

#### Producción de anticuerpos

**[0306]** Las células huésped se transforman con los vectores de expresión descritos anteriormente y se cultivan en medios de nutrientes convencionales modificados según sea apropiado para inducir promotores, seleccionar transformantes, o amplificar los genes que codifican las secuencias deseadas.

40 **[0307]** Transformación significa introducir ADN en el huésped procariota de tal forma que el ADN es replicable, bien como un elemento extracromosómico o mediante integración cromosómica. Dependiendo de la célula huésped usada, la transformación se realiza usando técnicas estándares apropiadas para tales células. El tratamiento con calcio empleando cloruro cálcico se usa generalmente para células bacterianas que contienen barreras sustanciales de pared celular. Otro método para la transformación utiliza polietilenglicol/DMSO. Otra

45 técnica utilizada es la electroporación.

**[0308]** Las células procariotas utilizadas para producir los polipéptidos de la invención se desarrollan en medios conocidos en la técnica y son adecuados para el cultivo de las células huésped seleccionadas. Ejemplos de medios adecuados incluyen caldo luria (LB) más complementos nutricionales necesarios. En algunas realizaciones, el medio también contiene un agente de selección, elegido en base a la construcción del vector de expresión, para permitir de manera selectiva el crecimiento de células procariotas que contienen el vector de expresión. Por ejemplo, se añade ampicilina al medio para el crecimiento de células que expresan el gen resistente a ampicilina.

**[0309]** También se pueden incluir en las concentraciones apropiadas cualquier suplemento necesario además de fuentes de carbono, nitrógeno y fosfato inorgánico introducidos solos o como una mezcla con otro suplemento o medio, tal como una fuente de nitrógeno compleja. Opcionalmente, el medio de cultivo puede contener uno o más agentes reductores seleccionados del grupo que consiste en glutatión, cisteína, cistamina, tioglicolato, ditioeritritol y ditiotreitrol.

**[0310]** Las células huésped procariotas se cultivan a temperaturas adecuadas. Para el crecimiento de *E. coli*, por ejemplo, se pueden utilizar intervalos de temperatura de aproximadamente 20°C a aproximadamente 39°C, de

aproximadamente 25°C a aproximadamente 37°C, y/o de aproximadamente 30°C. El pH del medio puede ser cualquier pH que varíe desde aproximadamente 5 a aproximadamente 9, dependiendo principalmente del organismo huésped. Para *E. coli*, el pH puede ser de aproximadamente 6,8 a aproximadamente 7,4, y puede ser aproximadamente 7,0.

5 **[0311]** Si se utiliza un promotor inducible en el vector de expresión, se induce la expresión de proteínas bajo condiciones adecuadas para la activación del promotor. En un aspecto de la presente invención, los promotores PhoA se utilizan para controlar la transcripción de los polipéptidos. Por consiguiente, las células huésped transformadas se cultivan en un medio limitante de fosfato para la inducción. El medio limitante de fosfato puede ser el medio C.R.A.P (véase, por ejemplo, Simmons et al., *J. Immunol. Methods* (2002), 263:133-147). Se puede utilizar un conjunto de otros inductores, según la construcción de vector utilizada, tal como se conoce en la técnica.

10 **[0312]** En un caso, los polipéptidos expresados se secretan en el periplasma y se recuperan del periplasma de las células huésped. La recuperación de proteínas implica habitualmente la ruptura del microorganismo, generalmente mediante medios, tales como choque osmótico, sonicación o lisis. Una vez se han roto las células, se pueden eliminar la debris celular o las células completas mediante centrifugación o filtración. Las proteínas se pueden purificar adicionalmente, por ejemplo, mediante cromatografía por afinidad de resina. Alternativamente, las proteínas se pueden transportar en el medio de cultivo y aislarse en el mismo. Las células se pueden extraer del cultivo y el sobrenadante de cultivo se puede filtrar y concentrar para una purificación posterior de las proteínas producidas. Los polipéptidos expresados se pueden aislar posteriormente e identificarse utilizando métodos conocidos habitualmente, tales como electroforesis en gel de poliacrilamida (PAGE) y ensayo de transferencia Western.

15 **[0313]** En un caso, la producción de anticuerpos se realiza en grandes cantidades mediante un proceso de fermentación. Existen varios procedimientos a gran escala de fermentación con alimentación por lotes para la producción de proteínas recombinantes. Las fermentaciones a gran escala tienen al menos 1000 litros de capacidad; en ciertas realizaciones, los fermentadores a gran escala tienen aproximadamente 1.000 a 100.000 litros de capacidad. Estos fermentadores utilizan impulsores agitadores para distribuir oxígeno y nutrientes, especialmente glucosa (la fuente de carbono/energía preferida). La fermentación a escala pequeña se refiere en general a la fermentación en un fermentador que no tiene más de 100 litros de capacidad volumétrica, y puede variar de aproximadamente 1 litro a aproximadamente 100 litros.

20 **[0314]** En un proceso de fermentación, la inducción de la expresión de proteínas se inicia habitualmente después de que las células hayan crecido bajo condiciones adecuadas hasta una densidad deseada, por ejemplo, DO550 de aproximadamente 180-220, en cuya fase las células se encuentran en una fase estacionaria inicial. Se pueden utilizar un conjunto de inductores, según la construcción de vector utilizada, tal como se conoce en la técnica y se ha descrito anteriormente. Las células se pueden desarrollar durante periodos de tiempo más cortos antes de la inducción. Las células se inducen habitualmente durante aproximadamente 12-50 horas, aunque se pueden utilizar tiempos de inducción más largos o más cortos.

25 **[0315]** Para mejorar el rendimiento de producción y la calidad de los polipéptidos, se pueden modificar varias condiciones de fermentación. Por ejemplo, para mejorar el ensamblaje correcto y el pliegue de los polipéptidos anticuerpo secretados, se pueden utilizar vectores adicionales que sobreexpresan proteínas chaperonas, tales como proteínas Dsb (DsbA, DsbB, DsbC, DsbD y/o DsbG) o FkpA (una peptidilproilil cis,trans-isomerasa con actividad de chaperona) para co-transformar las células huésped procariotas. Se ha demostrado que las proteínas chaperonas facilitan el correcto pliegue y la solubilidad de proteínas heterólogas producidas en células huésped bacterianas. Chen et al. (1999) *J Bio Chem.* 274:19601-19605; Georgiou et al., Patente de Estados Unidos No. 6,083,715; Georgiou et al., Patente de Estados Unidos No. 6,027,888; Bothmann y Pluckthun (2000) *J. Biol. Chem.* 275: 17100-17105; Ramm y Pluckthun (2000) *J. Biol. Chem.* 275:17106-17113; Arie et al. (2001) *Mol. Microbiol.* 39:199-210.

30 **[0316]** Para minimizar la proteólisis de proteínas heterólogas expresadas (especialmente aquellas que son proteolíticamente sensibles), se pueden utilizar ciertas cepas huésped deficientes en enzimas proteolíticas para la presente invención. Por ejemplo, las cepas de células huésped se pueden modificar para realizar una mutación o mutaciones genéticas en los genes que codifican proteasas bacterianas conocidas, tales como Proteasa III, OmpT, DegP, Tsp, Proteasa I, Proteasa Mi, Proteasa V, Proteasa VI y combinaciones de las mismas. Algunas cepas de *E. coli* deficientes en proteasa están disponibles y se describen en, por ejemplo, Joly et al. (1998), supra; Georgiou et al., Patente de Estados Unidos No. 5,264,365; Georgiou et al., Patente de Estados Unidos No. 5,508,192; Hara et al., *Microbial Drug Resistance*, 2:63-72 (1996).

35 **[0317]** En un caso, las cepas de *E. coli* deficientes en enzimas proteolíticas y transformadas con plásmidos que sobreexpresan una o más proteínas chaperonas se utilizan como células huésped en el sistema de expresión.

#### *Purificación de anticuerpos*

**[0318]** En un caso, la proteína anticuerpo producida aquí se purifica adicionalmente para obtener preparaciones

que son sustancialmente homogéneas para ensayos y usos posteriores. Se pueden utilizar métodos de purificación de proteínas estándar conocidos en la técnica. Los siguientes procedimientos son ejemplos de procedimientos de purificación adecuados: fraccionamiento en columnas de inmunoafinidad o intercambio iónico, precipitación con etanol, HPLC de fase inversa, cromatografía en sílice o en una resina de intercambio catiónico, tal como DEAE, "chromatofocusing", SDS-PAGE, precipitación con sulfato de amonio, y filtración en gel utilizando, por ejemplo, Sephadex G-75.

[0319] En un caso, se utiliza la proteína A inmovilizada en una fase sólida para la purificación por inmunoafinidad de los productos anticuerpos de la presente invención. La proteína A es una proteína de la pared celular de 41 kD de *Staphylococcus aureus* que se une con una gran afinidad a la región Fc de los anticuerpos. Lindmark et al (1983) J. Immunol. Meth. 62:1-13. En ciertos casos, la fase sólida a la que se inmoviliza la proteína A es una columna que comprende una superficie de vidrio o sílice. En ciertos casos, la fase sólida a la que se inmoviliza la proteína A es una columna de vidrio de poro controlado o una columna de ácido silícico. En algunas aplicaciones, la columna se ha recubierto con un reactivo, tal como glicerol, en un intento por prevenir la adherencia no específica de contaminantes.

[0320] Como primera etapa de purificación, la preparación derivada del cultivo celular tal como se ha descrito anteriormente, se aplica a la proteína A inmovilizada en la fase sólida para permitir la unión específica del anticuerpo de interés a la proteína A. A continuación, la fase sólida se lava para eliminar contaminantes no unidos específicamente a la fase sólida. Finalmente el anticuerpo de interés se recupera de la fase sólida mediante elución.

Generación de anticuerpos utilizando células huésped eucariotas:

[0321] Entre los componentes del vector se incluyen generalmente, pero no se limitan a, uno o más de los siguientes: una secuencia señal, un origen de replicación, uno o más genes marcadores, un elemento potenciador, un promotor, y una secuencia de terminación de la transcripción.

*(i) Componente secuencia señal*

[0322] Un vector para usar en una célula huésped eucariota también puede contener una secuencia señal u otro polipéptido que tiene un sitio de división específica en el extremo N-terminal de la proteína o polipéptido maduros de interés. En ciertos casos, la secuencia señal heteróloga seleccionada preferiblemente es la que es reconocida y procesada (es decir, dividida por una peptidasa señal) por la célula huésped. En la expresión de células de mamíferos, se disponen las secuencias señal de mamíferos, así como las secuencias líderes secretoras víricas, por ejemplo, la señal gD de herpes simplex.

[0323] El ADN para dicha región de precursor está ligada en el marco de lectura a ADN que codifica el anticuerpo.

*(ii) Origen de replicación*

[0324] Generalmente, no es necesario un componente origen de replicación para vectores de expresión de mamíferos. Por ejemplo, el origen SV40 se puede utilizar habitualmente sólo porque contiene el promotor temprano.

*(iii) Componente de gen de selección*

[0325] Los vectores de expresión y clonación pueden contener un gen de selección, también denominado un marcador seleccionable. Los genes de selección habituales codifican proteínas que (a) confieren resistencia a antibióticos u otras toxinas, por ejemplo, ampicilina, neomicina, metotrexato o tetraciclina, (b) complementan deficiencias auxotróficas, cuando sea relevante, o (c) suministran nutrientes críticos no disponibles del medio complejo.

[0326] Un ejemplo de un esquema de selección utiliza un fármaco para detener el crecimiento de una célula huésped. Las células que se transforman de forma satisfactoria con un gen heterólogo producen una proteína que confiere resistencia al fármaco y sobreviven de esta manera al régimen de selección. Algunos ejemplos de dicha selección dominante utilizan los fármacos neomicina, ácido micofenólico e higromicina.

[0327] Otro ejemplo de marcadores seleccionables adecuados para células de mamíferos son los que permiten la identificación de células competentes para captar el ácido nucleico del anticuerpo, tal como DHFR, timidina quinasa, metalotioneina-I y -II (por ejemplo, genes de metalotioneina de primate), adenosina desaminasa, ornitina descarboxilasa, etc.

[0328] Por ejemplo, las células transformadas con el gen de selección de DHFR se identificaron por primera vez mediante el cultivo de todos los transformantes en un medio de cultivo que contiene metotrexato (Mtx), un antagonista competitivo de DHFR. Una célula huésped apropiada cuando se utiliza la DHFR de tipo salvaje es la línea celular de ovario de hámster chino (CHO) deficiente en la actividad de DHFR (por ejemplo, ATCC CRL-



9096).

**[0329]** Alternativamente, se pueden seleccionar células huésped (particularmente huéspedes de tipo salvaje que contienen DHFR endógena) transformadas o cotransformadas con secuencias de ADN que codifican un anticuerpo, proteína DHFR de tipo salvaje, y otro marcador seleccionable, tal como aminoglicósido 3'-fosfotransferasa (APH), mediante el crecimiento celular en medio que contiene un agente de selección para el marcador seleccionable, tal como un antibiótico aminoglicosídico, por ejemplo, kanamicina, neomicina, o G418. Véase la Patente de Estados Unidos No. 4.965.199.

*(iv) Componente promotor*

**[0330]** Los vectores de expresión y clonación contienen habitualmente un promotor que es reconocido por el organismo huésped y está unido operativamente al ácido nucleico del polipéptido anticuerpo. Las secuencias de los promotores son conocidas para eucariotas. Prácticamente, todos los genes eucariotas tienen una región rica en AT localizada aproximadamente 25 a 30 bases en dirección 5' desde el sitio en el que se inicia la transcripción. Otra secuencia encontrada de 70 a 80 bases en dirección 5' desde el inicio de la transcripción de muchos genes es una región CNCAAT (SEQ ID NO:585) en la que N puede ser cualquier nucleótido. En el extremo 3' de la mayoría de genes eucariotas es una secuencia AATAAA (SEQ ID NO:586) que puede ser la señal para la adición de la cola poli A al extremo 3' de la secuencia codificante. Todas estas secuencias se insertan de manera adecuada en vectores de expresión eucariotas.

**[0331]** La transcripción de polipéptidos anticuerpos de vectores en células huésped de mamíferos se controla, por ejemplo, mediante promotores obtenidos de los genomas de virus, tales como el virus del polio, virus de la viruela aviar, adenovirus (tal como Adenovirus 2), virus de papiloma bovino, virus de sarcoma aviar, citomegalovirus, un retrovirus, virus de la hepatitis B y Virus del Simio 40 (SV40), de promotores de mamífero heterólogos, por ejemplo, el promotor de actina o un promotor de inmunoglobulina, de promotores de choque térmico, siempre y cuando dichos promotores sean compatibles con los sistemas de células huésped.

**[0332]** Los promotores tempranos y tardíos del virus SV40 se obtienen convenientemente como un fragmento de restricción de SV40 que también contiene el origen de replicación vírico SV40. El promotor temprano inmediato del citomegalovirus humano se obtiene convenientemente como un fragmento de restricción HindIII E. En la Patente de Estados Unidos No. 4.419.446 se describe un sistema para expresar ADN en huéspedes de mamíferos que utiliza el virus de papiloma bovino como vector. En la Patente de Estados Unidos No. 4.601.978 se describe una modificación de este sistema. Véase también Reyes et al., *Nature*, 297: 598-601 (1982) que describen la expresión de ADNc de  $\beta$ -interferón humano en células de ratón bajo el control de un promotor de timidina quinasa del virus de herpes simplex. Alternativamente, la repetición terminal larga del virus de sarcoma de Rous se puede utilizar como promotor.

*(v) Componente elemento potenciador*

**[0333]** La transcripción de un ADN que codifica el polipéptido anticuerpo por eucariotas superiores se incrementa frecuentemente mediante la inserción de una secuencia potenciadora en el vector. Actualmente, se conocen muchas secuencias potenciadoras de genes de mamíferos (globina, elastasa, albúmina,  $\alpha$ -fetoproteína e insulina). Habitualmente, sin embargo, se utilizará un potenciador de un virus de células eucariotas. Entre los ejemplos se incluyen el potenciador de SV40 en la cara tardía del origen de replicación (pb 100-270), el potenciador del promotor temprano de citomegalovirus, el potenciador de polio en la cara tardía del origen de replicación, y potenciadores de adenovirus. Véase también Yaniv. *Nature* 297:17-18 (1982) en elementos potenciadores para la activación de promotores eucariotas. El potenciador se puede empalmar en el vector en la posición 5' ó 3' con respecto a la secuencia codificante del polipéptido anticuerpo. En ciertas realizaciones, el potenciador se localiza en un sitio 5' desde el promotor.

*(vi) Componente de la terminación de la transcripción*

**[0334]** Los vectores de expresión utilizados en las células huésped eucariotas también contendrán habitualmente las secuencias necesarias para la terminación de la transcripción y para estabilizar el ARNm. Dichas secuencias están disponibles habitualmente de las regiones no traducidas 5', y alguna vez desde 3', de ADNs o ADNcs eucariotas o víricos. Estas regiones contienen segmentos de nucleótidos transcritos como fragmentos poliadenilados en la parte no traducida del ARNm que codifica un anticuerpo. Un componente de terminación de la transcripción útil es la región de poliadenilación de la hormona de crecimiento bovino. Véase WO 94/11026 y el vector de expresión descrito en la misma.

*(vii) Selección y transformación de células huésped*

**[0335]** Entre las células huésped adecuadas para la clonación o expresión del ADN en los vectores de la presente invención se incluyen células eucariotas superiores, descritas en la presente invención, incluyendo células huésped de vertebrados. La propagación de células de vertebrados en un cultivo (cultivo de tejidos) se ha convertido en un procedimiento rutinario. Entre los ejemplos de líneas celulares de huéspedes mamíferos útiles están la línea CV1 de riñón de mono transformada por SV40 (COS-7, ATCC CRL 1651); línea de riñón de

embrión humano (células 293 ó 293 subclonadas para el crecimiento en cultivo de suspensión, Graham et al., *J. Gen Virol.*, 36:59 (1977)); células de riñón de hámster bebé (BHK, ATCC CCL 10; células de ovario de hámster chino/-DHFR (CHO. Urlaub et al., *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980)); células de sertoli de ratón (TM4, Mather, *Biol. Reprod.*, 23:243-251 (1980)); células de riñón de mono (CV1 ATCC CCL 70); células de riñón de mono verde africano (VERO-76, ATCC CRL-1587); células de carcinoma cervical humano (HELA, ATCC CCL2); células de riñón canino (MDCK, ATCC CCL34); células de hígado de rata búfalo (BRL 3A, ATCC CRL 1442); células de pulmón humano (W138, ATCC CCL 75); células de hígado humano (Hep G2, HB 8065); tumor mamario de ratón (MMT 060562, ATCC CCL51); células TRI (Mather et al., *Annals N.Y. Acad. Sci.* 383: 44-68 (1982)); células MRC 5; células FS4 y una línea de hepatoma humano (Hep G2).

- 10 **[0336]** Las células huésped se transforman con los vectores de expresión o clonación descritos anteriormente para la producción del anticuerpo y se cultivan en un medio con nutrientes habituales modificado según sea apropiado para inducir promotores, seleccionar transformantes, o amplificar los genes que codifican las secuencias deseadas.

(viii) *Cultivo de células huésped*

- 15 **[0337]** Las células huésped utilizadas para producir un anticuerpo se pueden cultivar en una serie de medios. Los medios comercialmente disponibles, tales como F10 de Ham (Sigma), Medio Mínimo Esencial (MEM, Sigma), RPMI-1640 (Sigma), y Medio Eagle modificado por Dulbecco (DMEM, Sigma) son adecuados para el cultivo de células huésped. Además, cualquiera de los medios descritos en Ham et al., *Meth. Enz.* 58: 44 (1979), Barnes et al., *Anal. Biochem.* 102: 255 (1980), Patentes de Estados Unidos Nos. 4.767.704; 4.657.866; 20 4.927.762; 4.560.655 ó 5.122.469; WO 90/03430; WO 87/00195 o la Patente de Estados Unidos Re. 30.985, se pueden utilizar como medios de cultivo para las células huésped. Cualquiera de estos medios se puede suplementar según sea necesario con hormonas y/u otros factores de crecimiento (tales como insulina, transferrina, o factor de crecimiento epidérmico), sales (tales como cloruro sódico, calcio, magnesio y fosfato), 25 soluciones tampón (tales como HEPES), nucleótidos (tales como adenosina y timidina), antibióticos (tales como el fármaco GENTAMYCIN<sup>TM</sup>), elementos traza (definidos como compuestos inorgánicos presentes habitualmente en concentraciones finales a nivel micromolar), y glucosa o una fuente de energía equivalente. Cualquier otro suplemento necesario se puede también incluir en concentraciones apropiadas que serían conocidas por los expertos en la materia. Las condiciones de cultivo, tales como la temperatura, pH y similares, son las utilizadas previamente con la célula huésped seleccionada para la expresión y serán obvias para un técnico habitual.

30 (ix) *Purificación de anticuerpos*

- [0338]** Cuando se utilizan técnicas recombinantes, el anticuerpo se puede producir intracelularmente, o se secreta directamente en el medio. Si el anticuerpo se produce intracelularmente, como primera etapa, la debris particulada, ya sea células huésped o fragmentos lisados, se eliminan, por ejemplo, mediante centrifugación o ultrafiltración. Cuando el anticuerpo se secreta al medio, los sobrenadantes de dichos sistemas de expresión se 35 concentran generalmente en primer lugar utilizando un filtro de concentración de proteínas disponible comercialmente, por ejemplo, una unidad de filtración Amicon o Millipore Pellicon. Se puede incluir un inhibidor de proteasa tal como PMSF en cualquiera de las etapas anteriores para inhibir la proteólisis y se pueden incluir antibióticos para evitar el crecimiento de contaminantes accidentales.

- [0339]** La composición de anticuerpos preparada a partir de células se puede purificar utilizando, por ejemplo, 40 cromatografía de hidroxilapatito, electroforesis en gel, diálisis, y cromatografía de afinidad. La idoneidad de la proteína A como ligando de afinidad depende de la especie y el isótopo de cualquier dominio Fc de inmunoglobulina que esté presente en el anticuerpo. La proteína A se puede utilizar para purificar anticuerpos que están basados en cadenas pesadas humanas  $\gamma 1$ ,  $\gamma 2$  o  $\gamma 4$  (Lindmark et al., *J. Immunol Meth* 62:1-13 (1983)). La proteína G se recomienda para todos los isótopos de ratón y para  $\gamma 3$  humanas (Guss et al., *EMBO J.* 5:1567- 45 1575 (1986)). La matriz a la que se une el ligando de afinidad es frecuentemente agarosa, pero se disponen de otras matrices. Las matrices mecánicamente estables, tales como el vidrio de poro controlado o poli(estirenodivinil)benceno permite mayores velocidades de flujo y tiempos de procesado más cortos que los que se pueden conseguir con agarosa. Cuando el anticuerpo comprende un dominio C<sub>H</sub>3, la resina Bakerbond ABX<sup>TM</sup> (J.T. Baker Philipsburg, NJ) es útil para la purificación. Otras técnicas para la purificación de proteínas, 50 tales como el fraccionamiento en una columna de intercambio iónico, precipitación con etanol, HPLC de fase inversa, cromatografía en sílice, cromatografía en heparina-Sefarosa<sup>TM</sup>, cromatografía en una resina de intercambio de anión o catión (tal como una columna de ácido poliaspártico), "chromatofocusing", SDS-PAGE y precipitación con sulfato amónico también están disponibles dependiendo del anticuerpo a recuperar.

- [0340]** Tras cualquier etapa o etapas de purificación preliminar, la mezcla que comprende el anticuerpo de interés 55 y los contaminantes se puede someter a una cromatografía de interacción hidrofóbica a pH bajo utilizando un tampón de elución a un pH entre aproximadamente 2,5 y 4,5. En ciertas realaciones, la cromatografía de interacción hidrofóbica a pH bajo se realiza a concentraciones bajas de sales (por ejemplo, sal de aproximadamente 0-0,25 M).

Ensayos de actividad

**[0341]** Los anticuerpos se pueden caracterizar por sus propiedades físicas/químicas y funciones biológicas mediante varios ensayos conocidos en la técnica.

5 **[0342]** Las inmunoglobulinas purificadas se pueden caracterizar adicionalmente mediante una serie de ensayos que incluyen, pero sin limitación, secuenciación N-terminal, análisis de aminoácidos, cromatografía líquida a presión elevada (HPLC), exclusión por tamaño no desnaturalizante, espectrometría de masas, cromatografía de intercambio iónico y digestión con papaína.

10 **[0343]** En ciertos casos, las inmunoglobulinas producidas aquí se analizan por su actividad biológica. En algunos casos, las inmunoglobulinas de la presente invención se analizan por su actividad de unión a antígeno. Los ensayos de unión a antígeno que son conocidos en la técnica y que se pueden utilizar aquí incluyen sin limitación cualquier ensayo de unión directa o competitiva utilizando técnicas, tales como transferencias western, radioinmunoensayos, ELISA (ensayo inmunoabsorbente unido a enzima), inmunoensayos de "sándwich", ensayos de inmunoprecipitación, inmunoensayos fluorescentes, e inmunoensayos de proteína A.

15 **[0344]** Los anticuerpos o fragmentos de unión a antígeno se pueden seleccionar adicionalmente por la actividad funcional, por ejemplo, actividad antagonista o agonista. Por ejemplo, los anticuerpos anti-HER-2 se pueden seleccionar por la capacidad de inhibir la fosforilación de tirosina de HER-2, inhibir la proliferación de células cancerosas o inducir la apoptosis de células cancerosas. Los ensayos para identificar y medir estas actividades se describen en por ejemplo WO98/17797.

20 **[0345]** Tal como se describe aquí, se pueden seleccionar anticuerpos anti-DR5 por la capacidad de inducir apoptosis de células cancerosas y/o inhibir la función de células inflamatorias. En otros casos, los anticuerpos anti-DR5 se seleccionan por la capacidad de competir con Apo-2L por la unión a DR5. En otras realizaciones, los anticuerpos anti-DR5 humano se selecciona por la unión a DR5 murino y/o cinomolgo. Los ensayos para determinar la actividad biológica se pueden realizar utilizando métodos conocidos en la técnica, tales como la fragmentación de ADN (véase, por ejemplo, Marsters et al., Curr. Biology, 6: 1669 (1996)), inactivación de caspasa, unión a DR5 (véase, por ejemplo, WO98/51793, publicada el 19 de noviembre de 1998). La apoptosis se puede medir identificando la condensación del citoplasma, la pérdida de microvellosidades de la membrana plasmática, la segmentación del núcleo, degradación del ADN cromosómico o pérdida de función mitocondrial. Esta actividad se puede determinar y medir, por ejemplo, mediante ensayos de viabilidad celular (tales como ensayos de azul alamar o ensayos MTT), análisis FACS, activación de caspasa, fragmentación de ADN (véase, por ejemplo, Nicoletti et al., J. Immunol. Methods, 139:271-279 (1991), y poli ADP ribosa polimerasa, "PARP", ensayos de separación conocidos en la técnica.

35 **[0346]** En un caso, un ensayo para la apoptosis implica producir diluciones en serie de dos veces de patrón de control y anticuerpo anti-DR5 en placas de cultivo de tejido de 96 pocillos. Se analiza para comparación el ligando Apo-2 (aminoácidos 114-281, descrito en PCT US00/17579). Se siembran células de carcinoma de colon humano Colo-205 (20000 células-pocillo) (ATCC) en placas de 96 pocillos. Las placas se incuban a 37°C durante 24 horas. Se añade AlamazBlue (Trek Diagnostic Systems, Inc.) a los pocillos durante las últimas 3 horas del tiempo de incubación de 24 horas. La fluorescencia se lee utilizando un fluorómetro de 96 pocillos con excitación a 530 nm y emisión de 590 nm. Los resultados se expresan en unidades relativas de fluorescencia (RFU).

40 **[0347]** Se describe aquí un anticuerpo alterado que posee ciertas funciones efectoras, pero no todas, lo cual lo hace un candidato deseado para muchas aplicaciones en las que la vida media del anticuerpo in vivo es importante, aunque ciertas funciones efectoras (tales como el complemento y ADCC) sean innecesarias o perjudiciales. En ciertos casos, las actividades de Fc de la inmunoglobulina producida se miden para asegurar que se mantienen sólo las propiedades deseadas. Los ensayos de citotoxicidad in vitro y/o in vivo se pueden realizar para confirmar la reducción /agotamiento de las actividades de CDC y/o ADCC. Por ejemplo, los ensayos de unión a receptor Fc (FcR) se pueden realizar para asegurar que el anticuerpo carece de la unión a FcγR (por tanto, probablemente carece de la actividad de ADCC), pero mantiene la capacidad de unión a FcRn. Las células primarias para mediar la ADCC, células NK, sólo expresan FcγRIII, mientras que los monocitos expresan FcγRI, FcγRII y FcγRIII. La expresión de FcR en células hematopoyéticas se resume en la Tabla 3, página 464 de Ravetch y Kinet, Annu. Rev. Immunol 9:457-92 (1991). Un ejemplo de un ensayo in vitro para valorar la actividad de ADCC de una molécula de interés se describe en la patente de Estados Unidos No. 5,500,362 ó 5,821,337. Las células efectoras útiles para dichos ensayos incluyen células mononucleares de sangre periférica (PBMC) y células asesinas naturales (NK). Alternativamente, o adicionalmente, la actividad ADCC de la molécula de interés se puede valorar in vivo, por ejemplo, en un modelo animal, tal como el descrito en Clynes et al. PNAS (USA) 95:652-656 (1998). Los ensayos de unión a C1q también se pueden llevar a cabo para confirmar que el anticuerpo es incapaz de unirse a C1q y por tanto carece de la actividad CDC. Para valorar la activación de complemento, se puede realizar un ensayo CDC, por ejemplo tal como se describe en Gazzano-Santoro et al., J. Immunol. Methods 202:163 (1996). La unión a FcRn y las determinación de la depuración/vida media in vivo también se pueden llevar a cabo utilizando métodos conocidos en la técnica, por ejemplo, los descritos en la sección de ejemplo.

Anticuerpos humanizados

5 [0348] Se describen aquí anticuerpos humanizados. Se conocen en la técnica varios métodos para humanizar anticuerpos no humanos. Por ejemplo, un anticuerpo humanizado puede tener uno o más residuos de aminoácidos introducidos en el mismo a partir de una fuente que no es humana. Estos residuos de aminoácidos no humanos se denominan frecuentemente como residuos "importados", que habitualmente se obtienen de un dominio variable "importado". La humanización se puede realizar esencialmente siguiendo el método de Winter y colaboradores (Jones et al., *Nature*, 321: 522-525 (1986); Riechmann et al., *Nature*, 332: 323-327 (1988); Verhoeyen et al., *Science*, 239: 1534-1536 (1988)), mediante la sustitución de secuencias de la región hipervariable por las secuencias correspondientes de un anticuerpo humano. Por consiguiente, dichos anticuerpos "humanizados" son anticuerpos quiméricos (Patente de Estados Unidos No. 4.816.567) en los que sustancialmente menos de un dominio variable humano intacto se ha sustituido por la secuencia correspondiente de una especie no humana. En la práctica, los anticuerpos humanizados son habitualmente anticuerpos humanos en los que algunos residuos de la región hipervariable y posiblemente algunos residuos de FR se sustituyen por residuos de sitios análogos en anticuerpos de roedores.

15 [0349] La elección de dominios variables humanos, tanto ligeros como pesados, para utilizar en la fabricación de los anticuerpos humanizados es muy importante para reducir la antigenicidad. Según el método denominado "mejor-ajuste", la secuencia del dominio variable de un anticuerpo de roedor se criba frente a toda la biblioteca de secuencias de dominios variables humanos conocidos. A continuación, la secuencia humana que está más próxima a la del roedor se acepta como armazón ("framework") humano para el anticuerpo humanizado (Sims et al., *J. Immunol.*, 151: 2296 (1993); Chothia et al., *J. Mol. Biol.*, 196: 901 (1987)). Otro método utiliza un armazón particular derivado de la secuencia consenso de todos los anticuerpos humanos de un subgrupo particular de cadenas ligeras o pesadas. El mismo armazón se puede utilizar para varios anticuerpos humanizados diferentes (Carter et al., *Proc. Natl Acad. Sci. USA*, 89: 4285 (1992); Presta et al., *J. Immunol.*, 151: 2623 (1993)).

25 [0350] Es también importante que los anticuerpos se humanicen manteniendo una afinidad elevada por el antígeno y otras propiedades biológicas favorables. Para conseguir este objetivo, según un método preferido, los anticuerpos humanizados se preparan mediante un proceso de análisis de las secuencias parentales y varios productos conceptuales humanizados utilizando modelos tridimensionales de las secuencias parentales y humanizadas. Los modelos tridimensionales de inmunoglobulinas están disponibles normalmente y son familiares para los expertos en la materia. Existen programas informáticos que muestran y visualizan probables estructuras de conformaciones tridimensionales de secuencias de inmunoglobulinas candidatas seleccionadas. La observación de estas visualizaciones permite el análisis de la probable función de los residuos en el funcionamiento de la secuencia de inmunoglobulina candidata, es decir, el análisis de residuos que influyen en la capacidad de la inmunoglobulina candidata para unirse a su antígeno. De esta manera, los residuos de FR se pueden seleccionar y combinar a partir del receptor y secuencias importadas, de manera que se consigue la característica del anticuerpo deseado, tal como una mayor afinidad para el antígeno o antígenos diana. En general, los residuos de la región hipervariable están directamente y más sustancialmente implicados en la influencia sobre la unión al antígeno.

Variantes de anticuerpos

40 [0351] Se describen aquí fragmentos de anticuerpo que comprenden modificaciones en la interfase de los polipéptidos Fc que comprende la región Fc, donde las modificaciones facilitan y/o inducen la heterodimerización. Estas modificaciones comprenden la introducción de una protuberancia en un primer polipéptido Fc y una cavidad en un segundo polipéptido Fc, donde la protuberancia es posicionable en la cavidad para inducir la formación de complejos entre el primer y segundo polipéptido Fc. Los métodos de generación de anticuerpos con estas modificaciones son conocidos en la técnica, por ejemplo, tal como se describe en la Patente de Estados Unidos No. 5,731,168.

50 [0352] También se describen aquí la modificación o modificaciones en la secuencia de aminoácidos de los anticuerpos descritos aquí. Por ejemplo, puede ser deseable mejorar la afinidad de unión y/o otras propiedades biológicas del anticuerpo. Las variantes de la secuencia de aminoácidos del anticuerpo se preparan mediante la introducción de cambios apropiados de nucleótidos en el ácido nucleico del anticuerpo, o mediante síntesis peptídica. Dichas modificaciones incluyen, por ejemplo, deleciones y/o inserciones y/o sustituciones de residuos en las secuencias de aminoácidos del anticuerpo. Cualquier combinación de deleción, inserción y sustitución consigue llegar a la construcción final, siempre que la construcción final posea las características deseadas. Las alteraciones de aminoácidos se pueden introducir en la secuencia de aminoácidos del anticuerpo en cuestión en el momento en que se produce la secuencia.

55 [0353] Para incrementar la vida media de los anticuerpos o polipéptido que contienen las secuencias de aminoácidos de la presente invención, se puede unir un epítipo de unión a receptor de rescate al anticuerpo (especialmente un fragmento de anticuerpo), tal como se describe, por ejemplo, en la patente de Estados Unidos 5,739,277. Por ejemplo, se puede unir una molécula de ácido nucleico que codifica el epítipo de unión a receptor de rescate en el marco a un ácido nucleico que codifica una secuencia de polipéptido de la invención, de manera que la proteína de fusión expresada por la molécula de ácido nucleico modificada comprende el epítipo de unión

60

a receptor de rescate y una secuencia de polipéptido de la presente invención. Tal como se utiliza aquí, el término "epítipo de unión a receptor de rescate" se refiere a un epítipo de la región Fc de una molécula IgG (por ejemplo, IgG1, IgG2, IgG3, o IgG4) que es responsable del incremento de la vida media en suero in vivo de la molécula de IgG (por ejemplo, Ghetie, V et al., (2000) Ann. Rev. Immunol. 18:739-766, Tabla 1). Los anticuerpos con sustituciones en una región Fc de los mismos y vidas medias en suero incrementadas también se describen en WO00/42072 (Presta, L.), WO 02/060919; Shields, R.L., et al., (2001) JBC 276(9):6591-6604; Hinton, P.R., (2004) JBC 279(8):6213-6216). En otro caso, la vida media en suero también se puede incrementar, por ejemplo, mediante la unión a otras secuencias de polipéptido. Por ejemplo, los anticuerpos u otras secuencias de polipéptido se pueden unir a albúmina de suero o una parte de albúmina de suero que se une al receptor FcRn o un péptido de unión a albúmina de suero, de manera que la albúmina de suero se une al anticuerpo o polipéptido, por ejemplo, las secuencias de polipéptido descritas en WO01/45746. En un caso, el péptido de albúmina de suero a unir comprende la secuencia de aminoácidos de DICLPRWGCLW (SEQ ID NO:608). En otro caso, la vida media de un Fab según la presente invención se incrementa mediante estos métodos. Véase también, Dennis, M.S., et al., (2002) JBC 277(38):35035-35043 para secuencias de péptido de unión a albúmina de suero.

[0354] Un procedimiento útil para la identificación de ciertos residuos o regiones del anticuerpo que son posiciones preferidas para la mutagénesis se denomina "mutagénesis por rastreo de alanina", tal como se describe por Cunningham y Wells Science, 244: 1081-1085 (1989). Aquí, se identifican un residuo o grupo de residuos diana (por ejemplo, residuos cargados, tales como arg, asp, his, lys y glu) y se sustituyen por un aminoácido neutro o cargado negativamente (por ejemplo, alanina o polialanina) para influir en la interacción de los aminoácidos con el antígeno. Aquellas posiciones de aminoácidos que demuestran una sensibilidad funcional a las sustituciones se refinan a continuación mediante la introducción de variantes adicionales u otras en los sitios de sustitución o para los mismos. De este modo, mientras que el sitio para la introducción de una variación de secuencia de aminoácidos está predeterminado, la naturaleza de la mutación *per se* no necesita estar predeterminada. Por ejemplo, para analizar la acción de una mutación en un sitio determinado, se realiza la mutagénesis de rastreo de alanina o aleatorio en el codón o región diana y se criban las inmunoglobulinas expresadas por la actividad deseada.

[0355] Las inserciones en la secuencia de aminoácidos incluyen fusiones amino y/o carboxi terminales que varían de longitud desde un residuo a polipéptidos que contienen cien o más residuos, así como inserciones intrasecuencias de residuos de aminoácidos individuales o múltiples. Entre los ejemplos de inserciones terminales se incluyen un anticuerpo con un residuo metionilo N terminal o el anticuerpo fusionado a un polipéptido citotóxico. Otras variantes insercionales de la molécula de anticuerpo incluyen la fusión al N o C terminal del anticuerpo a una enzima (por ejemplo, ADEPT) o un polipéptido que aumenta la vida media del anticuerpo en el suero.

[0356] Otro tipo de variante es una variante por sustitución de aminoácidos. Estas variantes tienen por lo menos un residuo de aminoácido en la molécula de anticuerpo sustituida por un residuo diferente. Los sitios de mayor interés para la mutagénesis sustitucional incluyen las regiones hipervariables, pero también se contemplan alteraciones en la FR. Las sustituciones conservativas se muestran en la Tabla 2 bajo el encabezamiento de "sustituciones preferidas". Si dichas sustituciones dan lugar a un cambio en la actividad biológica, entonces se pueden introducir cambios más sustanciales, denominados "ejemplos de sustituciones", en la Tabla siguiente o tal como se describe posteriormente en referencia a clases de aminoácidos, y cribar los productos.

Tabla 2

Residuo original	Ejemplos de sustituciones	Substituciones preferidas
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Asp; Lys; Arg	Gln
Asp (U)	Glu; Asn	Glu
Cys(C)	Ser; Ala	Ser
Gln(Q)	Asn; Glu	Asn
Glu(E)	Asp; Gln	Asp
Gly(G)	Ala	Ala
His(H)	Asn; Gln; Lys; Arg	Arg

Ile(I)	Leu; Val; Met; Ala; Phe; norleucina	Leu
Leu(L)	norleucina; Ile; Val; Met; Ala; Phe	Ile
Lys(K)	Arg; Gln; Asn	Arg
Met(M)	Leu; Phe; Ile	Leu
Phe(F)	Trp; Leu; Val; Ile; Ala; Tyr	Tyr
Pro(P)	Ala	Ala
Ser(S)	Thr	Thr
Thr(T)	Val; Ser	Ser
Trp(W)	Tyr; Phe	Tyr
Tyr(Y)	Trp; Phe; Thr; Ser	Phe
Val(V)	Ile; Leu; Met; Phe; Ala; norleucina	Leu

5 **[0357]** Las modificaciones sustanciales en las propiedades biológicas del anticuerpo se realizan mediante la selección de sustituciones que difieren significativamente en su efecto de mantener (a) la estructura del esqueleto del polipéptido en el área de la sustitución, por ejemplo, como una conformación de hélice o lámina, (b) la carga o hidrofobicidad de la molécula en el sitio diana, o (c) el volumen de la cadena lateral. Los aminoácidos se pueden agrupar según las similitudes en las propiedades de sus cadenas laterales (en A. L. Lehninger, en Biochemistry, segunda ed., pág. 73-75, Worth Publishers, Nueva York (1975)):

- (1) no polares : Ala (A), Val (V), Leu (L), Ile (I), Pro (P), Phe (F), Trp (W), Met (M)
- (2) polares no cargados: Gly (G), Ser (S), Thr (T), Cys (C), Tyr (Y), Asn (N), Gln (Q)
- (3) ácidos: Asp (D), Glu (E)
- 10 (4) básicos : Lys (K), Arg (R), His (H).

**[0358]** Alternativamente, los residuos naturales se pueden dividir en grupos en base a las propiedades comunes de la cadena lateral:

- (1) hidrofóbico: norleucina, met, ala, val, leu, ile;
- (2) hidrofílico neutro: cys, ser, thr; Asn, Gln;
- 15 (3) ácido: asp, glu;
- (4) básico: his, lys, arg;
- (5) residuos que influyen en la orientación de la cadena: gly, pro; y
- (6) aromáticos: trp, tyr, phe.

20 **[0359]** Las sustituciones no conservativas comprenderán el intercambio de un miembro de una de estas clases por otra clase. Dichos residuos sustituidos también se pueden introducir en los sitios de sustitución conservativa o, en los sitios restantes (no conservativos).

25 **[0360]** Un tipo de variante por sustitución implica la sustitución de uno o más residuos de la región hipervariable de un anticuerpo parental (por ejemplo, un anticuerpo humanizado o humano). En general, la variante o variantes resultantes seleccionadas para el desarrollo posterior tendrán propiedades biológicas mejoradas en relación al anticuerpo parental del cual se generan. Una manera conveniente para generar dichas variantes por sustitución implica la maduración por afinidad utilizando la expresión en fagos. Brevemente, se mutan varios sitios de la región hipervariable (por ejemplo, 6-7 sitios) para generar todas las posibles sustituciones amino en cada sitio. Los anticuerpos generados de esta manera se expresan a partir de partículas de fagos filamentosos como fusiones al producto del gen III de M13 empaquetados en cada partícula. A continuación, las variantes expresadas en el fago se criban por su actividad biológica (por ejemplo, afinidad de unión) tal como se describe en la presente invención. Con el fin de identificar los sitios candidatos de la región hipervariable para la modificación, se puede aplicar la mutagénesis por rastreo de alanina para identificar residuos de la región hipervariable que contribuyen significativamente a la unión a antígeno. Alternativamente, o adicionalmente, puede ser beneficioso analizar una estructura del cristal del complejo antígeno-anticuerpo para identificar los puntos de contacto entre el anticuerpo y el antígeno. Dichos residuos de contacto y residuos próximos son

35

candidatos para la sustitución según las técnicas elaboradas en la presente invención. Una vez se generan dichas variantes, el panel de variantes se somete a cribado tal como se describe en la presente invención y se pueden seleccionar anticuerpos con propiedades superiores en uno o más ensayos pertinentes para un desarrollo posterior.

5 [0361] Las moléculas de ácido nucleico que codifican las variantes en la secuencia de aminoácidos del anticuerpo se preparan mediante una serie de procedimientos conocidos en la técnica. Estos procedimientos incluyen, pero no se limitan a, el aislamiento de una fuente natural (en el caso de variantes en las secuencias de aminoácidos naturales) o la preparación por mutagénesis mediada por oligonucleótidos (o dirigida de sitio), la mutagénesis de PCR y la mutagénesis de cassette de una variante preparada anteriormente o una versión no  
10 variante del anticuerpo.

[0362] Puede ser deseable para introducir una o más modificaciones de aminoácidos en una región Fc de los polipéptidos inmunoglobulina de la presente invención, generando de este modo una variante de la región Fc. La variante de la región Fc puede comprender una secuencia de la región Fc humana (por ejemplo, una región Fc de IgG1, IgG2, IgG3 o IgG4 humana) que comprende una modificación de aminoácido (por ejemplo, una sustitución) en una o más posiciones de aminoácidos que incluyen la de una cisteína bisagra.

[0363] Según esta descripción y los conocimientos de la técnica, se contempla que un anticuerpo puede comprender una o más alteraciones en comparación con el anticuerpo homólogo de tipo salvaje, por ejemplo, en la región Fc. Estos anticuerpos, sin embargo, mantendrían sustancialmente las mismas características requeridas para la utilidad terapéutica en comparación con su homólogo de tipo salvaje. Por ejemplo, se cree que se pueden realizar ciertas alteraciones en la región Fc que daría lugar a una unión C1q alterada (es decir, mejorada o disminuida) y/o Citotoxicidad Dependiente de Complemento (CDC), por ejemplo, tal como se describe en WO99/51642. Véase también Duncan & Winter Nature 322:738-40 (1988); Patente de Estados Unidos No. 5,648,260; Patente de Estados Unidos No. 5,624,821; y WO94/29351 con respecto a otros ejemplos de variantes de la región Fc.

[0364] Habiendo descrito en general la invención, la misma se entenderá más fácilmente por referencia a los siguientes ejemplos, los cuales se proporcionan a modo de ilustración y no pretenden ser limitantes.

#### **Ejemplo 1. Construcción de bibliotecas de Fab expresados en fagos con residuos de CDR enriquecidos en Tyr, Ser, Gly, y Arg.**

30 [0365] Se construyeron bibliotecas de Fab expresados en fagos utilizando un vector fagémido, Fab-C, que dio lugar a la expresión de grupos Fab bivalentes dimerizados por una cisteína libre insertada entre la cadena pesada de Fab y el dominio C-terminal de la proteína de recubrimiento menor del gen-3 (P3C). Este vector se construyó tal como se describe en la publicación de solicitud de patente de Estados Unidos No. US20050119455 y en Lee et al., J. Immunol. Meth. 284: 119-132 (2004). El vector (ilustrado esquemáticamente en la figura 5) comprende dominios variables del anticuerpo 4D5 humanizado bajo el control del promotor Ptac inducible por  
35 IPTG. El anticuerpo 4D5 humanizado presenta mayoritariamente regiones de armazón de secuencia consenso humanas en las cadenas pesada y ligera y regiones CDR de un anticuerpo monoclonal de ratón específico para HER-2. Los métodos de producción del anticuerpo anti-HER2 y la identidad de las secuencias de los dominios variables se proporcionan en la patente de Estados Unidos Nos. 5,821,337 and 6,054,297.

40 [0366] Se construyeron cuatro bibliotecas: YSGR-A, YSGR-B, YSGR-C, y YSGR-D. Las bibliotecas se construyeron con residuos aleatorizados en las tres CDR de cadena pesada y CDR3 de cadena ligera. Cada biblioteca se aleatorizó en las posiciones 91-94 y 96 de CDRL3, las posiciones 28 y 30-33 de CDRH1, las posiciones 50, 52-54, 56, y 58 de CDRH2, y las posiciones 95-100, 100a, 100b, y 100c de CDRH3. El tipo y la proporción de los aminoácidos permitidos en cada una de las posiciones aleatorizadas se describe en la figura 8. Además, la longitud de CDRH3 se varió utilizando oligonucleótidos que sustituían los siete codones de tipo natural de las posiciones 95 a 100a por seis a diecisiete codones. De este modo, en ciertos casos, el codón correspondiente a la posición 100a de la cadena pesada no estaba presente (por ejemplo, cuando la mutagénesis se realizaba con los oligonucleótidos mutagénicos H3-A6 (SEQ ID NO:35), H3-B6 (SEQ ID NO:47), H3-C6 (SEQ ID NO:59) o H3-D6 (SEQ ID NO:71), tal como se describe a continuación. Véase la figura 9A-D. El tipo y la proporción de los aminoácidos permitidos en estas posiciones fueron los mismos que los descritos en la  
50 figura 8 para las posiciones 95-100a de CDRH3.

[0367] Se construyeron bibliotecas utilizando el método de Kunkel (Kunkel, T.A., Roberts, J.D. & Zakour, R.A., Methods Enzymol. (1987), 154, 367-382) con métodos descritos previamente (Sidhu, S.S., Lowman, H.B., Cunningham, B. C. & Wells, J.A., Methods Enzymol. (2000), 328, 333-363). Se utilizó una única versión de "plantilla de detención" del vector de expresión de Fab Fab-C para generar las cuatro bibliotecas tal como se describe en el ejemplo 1.

[0368] Se utilizaron oligonucleótidos mutagénicos con codones degenerados en las posiciones a diversificar para simultáneamente (a) introducir diversidad de CDR y (b) reparar los codones de detención. Las secuencias de

estos oligonucleótidos mutagénicos se muestran en las figuras 9A-9D. Para todas las bibliotecas, se introdujo diversidad en CDRH1, CDRH2, y CDRH3 con oligonucleótidos H1, H2, y L3, respectivamente (SEQ ID NOs: ). Para la biblioteca YSGR-A, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-A6, H3-A7, H3-A8, H3-A9, H3-A10, H3-A11, H3-A12, H3-A13, H3-A14, H3-A15, H3-A16, y H3-A17 (SEQ ID NOs: 35-46). Para la biblioteca YSGR-B, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-B6, H3-B7, H3-B8, H3-B9, H3-B10, H3-B11, H3-B12, H3-B13, H3-B14, H3-B15, H3-B16, y H3-B17 (SEQ ID NOs: 47-58). Para la biblioteca YSGR-C, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-C6, H3-C7, H3-C8, H3-C9, H3-C10, H3-C11, H3-C12, H3-C13, H3-C14, H3-C15, H3-C16, y H3-C17 (SEQ ID NOs: 59-70). Para la biblioteca YSGR-D, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-D6, H3-D7, H3-D8, H3-D9, H3-D10, H3-D11, H3-D12, H3-D13, H3-D14, H3-D15, H3-D16, y H3-D17 (SEQ ID NOs:71-82). Cada uno de los oligonucleótidos mutagénicos H3-A6 a H3-A17 (SEQ ID NOs: 35-46), H3-B6 a H3-B17 (SEQ ID NOs:47-58), H3-C6 a H3-C17 (SEQ ID NOs:59-70) y H3-D6 a H3-D17 (SEQ ID NOs: 71-82) codificaba una alanina en la posición 93 de la cadena pesada. Los oligonucleótidos mutagénicos para todas las CDR a aleatorizar se incorporaron simultáneamente en una única reacción de mutagénesis, de manera que la incorporación simultánea de todos los oligonucleótidos mutagénicos dio lugar a la introducción de la diversidad diseñada en cada posición y simultáneamente la reparación de todos los codones de detención TAA. De este modo, se generó un marco de lectura abierto que codificaba un miembro de la biblioteca de Fab fusionado a un puente de cisteínas homodimerizante y P3C. Después de la mutagénesis, las cuatro bibliotecas se combinaron para crear una única biblioteca denominada YSGR-A-D.

**[0369]** Las reacciones de mutagénesis se sometieron a electroporación en *E. coli* SS320 (Sidhu et al., supra). Las células transformadas se desarrollaron durante la noche en presencia del fago auxiliar M13-K07 (New England Biolabs, Beverly, MA) para producir partículas de fagos que encapsulaban el ADN fagémido y expresaban fragmentos de Fab en sus superficies. La biblioteca combinada contenía más de  $3 \times 10^{10}$  miembros únicos.

## **Ejemplo 2. Selección de anticuerpos específicos de la biblioteca intacta YSGR-A-D.**

**[0370]** Se ciclaron fagos de la biblioteca YSGR-A-D (descrita en el Ejemplo 1 anterior) mediante rondas de selección de unión para enriquecer la unión de clones a DR5 o HER-2 humanas. Las selecciones de unión se realizaron utilizando procedimientos anteriormente descritos (Sidhu et al., supra).

**[0371]** En la Tabla 1 se muestra una secuencia de DR5 humana. Se utilizó en la selección de unión un dominio extracelular de DR5 tal y como se muestra en la Tabla 1. Asimismo, se prepara un dominio extracelular de una secuencia de HER-2 humana tal y como se describe en Franklin MC. Carey KD. Vajdos FF. Leahy DJ. de Vos AM. Sliwkowski MX. Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. *Cancer Cell*. 5(4):317-28, 2004. Se proporciona una secuencia de ECD de HER-2 humana (aminoácidos 23-646) del Protein DataBank Record 1S78 (2004).

**[0372]** Se cubrieron inmunoplasmas Maxisorp de 96 pocillos NUNC durante toda la noche a 4°C con proteína diana 5 µg/mL (DR5 humana o HER-2 humana) y se bloquearon durante 2 horas con una solución de PBT (solución salina tamponada con fosfato que contiene adicionalmente BSA al 0,2% y Tween 20 (Sigma) al 0,05%). Después de crecer durante toda la noche a 37°C, se concentraron los fagos mediante precipitación con PEG/NaCl y se resuspendieron en PBT, tal y como se ha descrito anteriormente (Sidhu et al., supra). Se añadieron soluciones de fagos (aproximadamente  $10^{12}$  fagos/mL) a las inmunoplasmas recubiertas. Después de una incubación de dos horas para permitir la unión de los fagos, se lavaron las placas diez veces con PBT. Se eluyeron los fagos de unión con HCl 0,1 M durante diez minutos y se neutralizó el eluyente con base Tris 1,0 M. Se amplificaron los fagos eluidos en *E. coli* XL1-blue y se utilizaron para rondas adicionales de selección.

**[0373]** Se sometieron las bibliotecas a cinco rondas de selección contra cada proteína diana. Se desarrollaron clones individuales de cada ronda de selección en un formato de 96 pocillos en 500 µL de caldo 2YT complementado con carbenicilina y M13-K07. Se utilizaron los sobrenadantes de cultivo directamente en ELISA de fagos (Sidhu et al., supra) para detectar Fabs expresados en fagos que se unían a placas cubiertas con proteína diana pero no a placas cubiertas con BSA. Se definieron los enlazadores específicos como aquellos clones de fagos que exhibieron una señal ELISA por lo menos 10 veces mayor en las placas cubiertas con diana en comparación con placas cubiertas con BSA. Se cribaron los clones individuales después de 4 y 5 rondas de selección para unirse a DR5 humana o HER-2 humana. Se sometieron los enlazadores específicos a análisis de secuencia. También se analizaron enlazadores específicos utilizando ELISA de afinidad spot, ELISA de especificidad, ELISA de especificidad, y afinidad por HER2 utilizando métodos tal y como se describen en la presente invención. (Ver Figura 11B.) Tal y como se muestra en la Figura 10, la biblioteca YSGR-A-D produjo enlazadores específicos contra ambas proteínas diana.

**[0374]** De los 240 clones identificados que se unían específicamente a HER-2 humana, 106 de ellos tuvieron secuencias de CDR únicas (ver Figura 11). Las secuencias únicas se encontraban en 3 categorías: 1) secuencias de CDRH3 de 6-7 residuos; 2) secuencias de CDRH3 de ocho residuos; y 3) secuencias de CDR de longitud media con tirosina, serina y glicina en la secuencia.



**[0375]** Los dominios variables de cadena pesada de anti-HER-2 muestran una preferencia por secuencias de CDRH3 cortas (por ejemplo 6-7 aminoácidos en posiciones correspondientes de 95 a 100a) no incluidas en la biblioteca de oligo. (Ver Figura 12). La CDRH3 muestra un residuo de tirosina conservado en el extremo N terminal en la posición 95. La otra posición conservada se encuentra en la posición 99 que es predominantemente una glicina. Se muestran secuencias consenso para CDRL3: QQSYYX4PST (SEC ID N°:587); CDRH1:GFSIX2X3SYIH (SEC ID N°:588); y CDRH2:SIYPX3SGYTSYADSKVG (SEC ID N°:589), en las que la X representa una posición de aminoácido para la que no se identificó un residuo de consenso y en las que las posiciones X en cada CDR son Y o S.

**[0376]** Los dominios variables de cadena pesada que tienen CDRH3 con ocho aminoácidos han conservado glicinas en los extremos terminales de N y C de CDRH3 (en las posiciones 95 y 100a). También se conserva la posición 98 con una tirosina. La posición 99 de una CDRH3 de ocho aminoácidos muestra una preferencia por un aminoácido pequeño como G, S, A, o T. Esta posición va seguida de un aminoácido grande en la posición 100, como R, H, Y, y W. Se muestran las secuencias consenso para CDRH1: GFX1ISYSSIH (SEC ID N°:590); y CDRH2:SIYPX3YGX5TX6YADSKVG(SEC ID N°:591), en las que la X representa una posición de aminoácido para la cual no se identificó un residuo consenso y es Y o S.

**[0377]** El análisis de los dominios variables de cadena pesada con regiones CDRH3 de longitud media (por ejemplo aproximadamente de 12 a 14 aminoácidos) proporciona una secuencia consenso de CDRH3 X1X2X3X4YYSYYX10GX12X13X14DY (SEC ID N°:592), en la que la X representa una posición de aminoácido para la cual no se identificó un residuo consenso, en la que X1 se selecciona entre Y, S y R; X2 se selecciona entre Y y S; X3 se selecciona entre G, Y y S; X4 se selecciona entre Y, S, R y G; X10 se selecciona entre Y, S y G; X12 se selecciona entre Y, S, G y R; X13 se selecciona entre G y A; y X14 se selecciona entre I, F, F y L. (Ver Figura 14). Como la CDRH3 forma un bucle, se desarrolló el consenso para CDRH3 mediante el desplazamiento de la secuencia para algunos clones alrededor de dos aminoácidos de modo que la posición 95 de la secuencia se alinearía con la posición 97 de la secuencia de referencia, que en este caso fue la secuencia del clon 52. También se muestran las secuencias consenso para CDRH1: GFX1ISSSSIH (SEC ID N°:593); y CDRH2:X1X2PSSGYTX6YADSKVG (SEC ID N°:594), en las que X representa una posición de aminoácido para la cual no se identificó un residuo consenso y es Y o S.

**[0378]** Tal y como se describe en la Figura 11B, varios de los clones se unieron a HER2 con entre de 0,1 a 10 nm de IC<sub>50</sub>. Para la mayoría de ellos, estos enlazadores de alta afinidad tuvieron poca o ninguna reactividad cruzada con otros antígenos, tales como VEGF, DR5, insulina, neutravidina, hormona del crecimiento humana o IGF-1.

**[0379]** Tal y como se muestra en la Figura 10, se identificaron 144 clones que expresaban Fabs que eran enlazadores específicos para DR5 humana. El análisis de secuencia identificó 18 secuencias únicas de aminoácidos de los 144 clones mostrados en la Figura 15.

**[0380]** Se determinó la IC<sub>50</sub> de cada uno de estos enlazadores mediante ELISA de fago competitivo, tal y como se indica a continuación. Se cubrieron inmunoplasmas Maxisorp de 96 pocillos NUNC durante toda la noche a 4°C con hDR5-ECD (5 µg/ml) y se bloquearon con BSA. Se propagaron los Fabs que se expresaron en fagos en *E. Coli* XL1-blue con la adición de fago auxiliar M13-KO7. Después del crecimiento durante toda la noche a 37°C en medios 2YT, se concentraron los fagos mediante precipitación con PEG/NaCl y se resuspendieron en solución salina tamponada con fosfato (PBS), BSA al 0,5% (p/v), Tween 20 0,1 % (v/v) (tampón PBT). Los fagos se diluyeron en serie en tampón PBT y se midió la unión para determinar una concentración de fago que produce ~50% de la señal en la saturación. Se preincubó una concentración fija de subsaturación de fago durante 2 h con diluciones en serie de hDR5-ECD y a continuación se transfirieron a placas de ensayo cubiertas con hDR5-ECD. Después de 15 min de incubación, se lavaron las placas con PBS, Tween 20 al 0,05% y se incubaron 30 min con conjugado de peroxidasa de rábano/anticuerpo anti-M13 (dilución 1:5000 en tampón PBT). Se lavaron las placas, se revelaron con sustrato TMB, se inhibieron con H<sub>3</sub>PO<sub>4</sub> 1,0 M, y se leyeron espectrofotométricamente a 450 nm. Se determinaron las afinidades de unión de los ligandos anti-hDR5 como valores de IC<sub>50</sub> definidos como la concentración de hDR5-ECD que bloqueaba el 50% de las uniones de fago a los hDR5-ECD inmovilizados.

**[0381]** Se analizaron los clones por la unión a los dominios extracelulares de DRS humana (SEC ID N°:595). Los enlazadores con menor IC<sub>50</sub> tienen predominantemente serina en CDRH1 y arginina en CDRH3 en las posiciones 96, y 99. El análisis de las regiones CDRH3 de los dominios variables de la cadena pesada proporciona una secuencia consenso de CDRH3 YRX3YRYGX8X9X10GSYX14X15DY (SEC ID N°:596), en la que X3 se selecciona entre Y, S, R, P y G; X8 se selecciona entre R, Y y S; X9 se selecciona entre G y Y; X10 se selecciona entre S, Y y R, X14 se selecciona entre G y A; y X15 se selecciona entre L y F, donde X representa una posición de aminoácido para la cual no se identificó un residuo de consenso. (Ver Figura 16). También se muestran las secuencias consenso para CDRL3:QQX1X2X3SPST (SEC ID N°:597), en las que X1, X2 y X3 son Y o S; CDRH1:GFX1X2SSSIH (SEC ID N°:598); y CDRH2:X1ISPX3X4GYTX6YADSKVG (SEC ID N°:599), donde X representa una posición de aminoácido para la cual no se identificó un residuo consenso y es Y o S. Pueden utilizarse las secuencias consenso, entre otras cosas, para formar nuevas bibliotecas de dominios variables de anticuerpo. En CDRH3, los aminoácidos en las posiciones 97, 100b, 100c, 100 h, y 100i pueden contribuir a una mayor afinidad.

[0382] También se analizaron los clones según la unión a DR5 de murino utilizando el ELISA de fago competitivo descrito anteriormente. Se aislaron diversos clones de la biblioteca YSGR A-D que unía tanto DR5 humano como DR5 de murino, aunque la unión de DR5 de murino fue de mucha menor afinidad. (Ver Figura 17) Esta biblioteca proporcionó el aislamiento de anticuerpos que pueden unirse a DR5 de murino y de humano indicando que los enlazadores identificados eran únicos en comparación con un CDRH3 aleatorio total (los veinte aminoácidos) y una biblioteca YS CDRH3. Cambiando la diversidad de aminoácidos en cada posición se permitió que se pudieran proporcionar anticuerpos que se unen a diferentes epítopos y tengan funciones biológicas únicas. Los anticuerpos anti-DR5 que se unen a CDRs de murino y de humano pueden unirse a diferentes epítopos de los anticuerpos anti-DR5 de bibliotecas desarrolladas previamente.

### 10 Ejemplo 3 (ejemplo comparativo) Análisis de enlazadores a DR5

[0383] El sitio de unión para el ligando Apo2L para DR5 humana se ha localizado previamente y se ha determinado la estructura cristalina para el sitio de unión (véase Hymowitz et al, Molecular Cell 4:564 (1999); WO01/19861). La estructura cristalina y los modelos se pueden utilizar para localizar la unión de los anticuerpos anti-DR5.

15 [0384] Estudios previos han identificado anticuerpos que se unen a DR5 humana. Estos anticuerpos se designan como BDF1 y YSD1. El anticuerpo BDF1 se aisló de una biblioteca en la que se aleatorizaron CDR con los 20 aminoácidos y tenía las secuencias CDR: 1) secuencia CDRH1 de IGKSGIH (SEQ ID NO:600); 2) secuencia CDR2 de VAVIYPHDGNTAYA (SEQ ID NO:601); y 3) secuencia CDRH3 de RLALVRMWM (SEQ ID NO:602). El anticuerpo YSD1 se aisló de una biblioteca en la que las posiciones de CDR se variaban con tirosina y serina y tiene una secuencia CDRH3 de YSSYYSSYYSSSSYSY (SEQ ID NO:603). El sitio de unión de estos anticuerpos en DR5 humana se localiza en el extremo N terminal de la molécula y presenta un solapamiento pequeño con el del ligando Apo2L, que se halla predominantemente en el extremo C terminal (aminoácidos del bucle en los 50, por ejemplo, aminoácidos 50-65 y aminoácidos del bucle en los 90, por ejemplo, aminoácidos 91 a 104 de DR5). En la figura 18 se muestra un modelo que muestra la unión de las regiones CDRH3 de cada uno de estos anticuerpos. Las CDRH3 de BDF1 y YSD1 se solapan y forman un punto clave para la unión a CD5. La unión de CDRH3 de YSD1 está mediada por tirosinas y la unión de BDF1 está mediada por la secuencia LAL. La unión de YSD1 a DR5 implica residuos de leucina, glutamina, alanina, fenilalanina, y arginina de DR5.

### 30 Ejemplo 4. Construcción de bibliotecas de Fab expresados en fagos con residuos de CDR1, H2, y L3 enriquecidos en Tyr y Ser y residuos de CDRH3 enriquecidos en Ser y Ala, Cys, Phe, Gly, Ile, Leu, Asn, Pro, Arg, Thr, Trp, o Tyr.

[0385] Se construyeron bibliotecas de Fab expresados en fagos utilizando un vector fagémido, Fab-C, que dio lugar a la expresión de grupos Fab bivalentes dimerizados por una cisteína libre insertada entre la cadena pesada de Fab y el dominio C-terminal de la proteína de recubrimiento menor del gen-3 (P3C), tal como se ha descrito previamente en el ejemplo 1.

35 [0386] Se construyeron doce bibliotecas: SAH3, SCH3, SFH3, SGH3, SLH3, SNH3, SPH3, SRH3, STH3, SWH3, y SYH3. Las bibliotecas se construyeron con residuos aleatorizados en las tres CDR de cadena pesada y CDR3 de cadena ligera. Cada biblioteca se aleatorizó en las posiciones 91-94 y 96 de CDRL3, las posiciones 28 y 30-33 de CDRH1, las posiciones 50, 52-54, 56, y 58 de CDRH2, y las posiciones 95-100, 100a a 100m de CDRH3. El tipo y la proporción de los aminoácidos permitidos en cada una de las posiciones aleatorizadas se describe en las figuras 19A-19B. Además, la longitud de CDRH3 se varió utilizando oligonucleótidos que sustituyeron los seis codones de tipo natural entre las posiciones 95 y 100 por 4 a 17 codones. El tipo y proporción de los aminoácidos permitidos en estas posiciones fueron los mismos que los descritos en las figuras 19A-19B para las posiciones 95-100 de CDRH3.

45 [0387] Se construyeron bibliotecas utilizando el método de Kunkel (Kunkel et al., Methods Enzymol. (1987) 154: 367-382) con métodos descritos previamente (Sidhu et al., Methods Enzymol. (2000) 328: 333-363). Se utilizó una única versión de "plantilla de detención" del vector de expresión de Fab Fab-C para generar las cuatro bibliotecas tal como se describe en el ejemplo 1.

50 [0388] Se utilizaron oligonucleótidos mutagénicos con codones degenerados en las posiciones a diversificar para simultáneamente (a) introducir diversidad de CDR y (b) reparar los codones de detención. Las secuencias de estos oligonucleótidos mutagénicos se muestran en las figuras 20A-20L. Para todas las bibliotecas, se introdujo diversidad en CDRH1, CDRH2, y CDRL3 con oligonucleótidos H1, H2, y L3, respectivamente (SEQ ID NOs: ).

[0389] Para la biblioteca SAH3, se introdujo diversidad en CDRH3 con una mezcla equimolar de oligonucleótidos H3-SA4, H3-SA5, H3-SA6, H3-SA7, H3-SA8, H3-SA9, H3-SA10, H3-SA11, H3-SA12, H3-SA13, H3-SA14, H3-SA15, H3-SA16, y H3-SA17 (SEQ ID NOs: 621-634).

55 [0390] Para la biblioteca SCH3, se introdujo diversidad en CDRH3 con una mezcla equimolar de oligonucleótidos H3-SC4, H3-SC5, H3-SC6, H3-SC7, H3-SC8, H3-SC9, H3-SC10, H3-SC11, H3-SC12, H3-SC13, H3-SC14, H3-SC15, H3-SC16, y H3-SC17 (SEQ ID NOs: 635-648).

[0391] Para la biblioteca SFH3, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SF4, H3-SFS, H3-SF6, H3-SF7, H3-SF8, H3-SF9, H3-SF10, H3-SF11, H3-SF12, H3-SF13, H3-SF14, H3-SF15, H3-SF16, y H3-SF17 (SEQ ID NOs: 649-662).

5 [0392] Para la biblioteca SGH3, se introdujo diversidad en CDRH3 con una mezcla equimolar de oligonucleótidos H3-SG4, H3-SGS, H3-SG6, H3-SG7, H3-SG8, H3-SG9, H3-SG10, H3-SG11, H3-SG12, H3-SG13, H3-SG14, H3-SG15, H3-SG16, y H3-SG17 (SEQ ID NOs: 663-676).

[0393] Para la biblioteca SIH3, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SI4, H3-SI5, H3-SI6, H3-SI7, H3-SI8, H3-SI9, H3-SI10, H3-SI11, H3-SI12, H3-SI13, H3-SI14, H3-SI15, H3-SI16, y H3-SI17 (SEQ ID NOs: 677-690).

10 [0394] Para la biblioteca SLH3, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SLA, H3-SL5, H3-SL6, H3-SL7, H3-SL8, H3-SL9, H3-SL10, H3-SL11, H3-SL12, H3-SL13, H3-SL14, H3-SL15, H3-SL16, y H3-SL17 (SEQ ID NOs: 691-704).

15 [0395] Para la biblioteca SNH3, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SN4, H3-SN5, H3-SN6, H3-SN7, H3-SN8, H3-SN9, H3-SN10, H3-SN11, H3-SN12, H3-SN13, H3-SN14, H3-SN15, H3-SN16, y H3-SN17 (SEQ ID NOs: 705-718).

[0396] Para la biblioteca SPH3, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SP4, H3-SP5, H3-SP6, H3-SP7, H3-SP8, H3-SP9, H3-SP10, H3-SP11, H3-SP12, H3-SP13, H3-SP14, H3-SP15, H3-SP16, y H3-SP17 (SEQ ID NOs: 719-732).

20 [0397] Para la biblioteca SRH3, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SR4, H3-SR5, H3-SR6, H3-SR7, H3-SR8, H3-SR9, H3-SR10, H3-SR11, H3-SR12, H3-SR13, H3-SR14, H3-SR15, H3-SR16, y H3-SR17 (SEQ ID NOs: 733-746).

[0398] Para la biblioteca STH3, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-ST4, H3-ST5, H3-ST6, H3-ST7, H3-ST8, H3-ST9, H3-ST10, H3-ST11, H3-ST12, H3-ST13, H3-ST14, H3-ST15, H3-ST16, y H3-ST17 (SEQ ID NOs: 747-760).

25 [0399] Para la biblioteca SWH3, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SW4, H3-SW5, H3-SW6, H3-SW7, H3-SW8, H3-SW9, H3-SW10, H3-SW11, H3-SW12, H3-SW13, H3-SW14, H3-SW15, H3-SW16, y H3-SW17 (SEQ ID NOs: 761-774).

30 [0400] Para la biblioteca SYH3, se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SY4, H3-SY5, H3-SY6, H3-SY7, H3-SY8, H3-SY9, H3-SY10, H3-SY11, H3-SY12, H3-SY13, H3-SY14, H3-SY15, H3-SY16, y H3-SY17 (SEQ ID NOs: 775-788).

35 [0401] Los oligonucleótidos mutagénicos para todas las CDR a aleatorizar se incorporaron en una única reacción de mutagénesis, de manera que la incorporación simultánea de todos los oligonucleótidos mutagénicos dio lugar a la introducción de la diversidad diseñada en cada posición y la reparación de todos los codones de detención TAA. De este modo, se generó un marco de lectura abierto que codificaba un miembro de la biblioteca de Fab fusionado a un puente de cisteínas homodimerizante y P3C. Después de la mutagénesis, las doce bibliotecas se combinaron para crear una única biblioteca denominada SXH3.

40 [0402] Las reacciones de mutagénesis se sometieron a electroporación en *E. coli* SS320 (Sidhu et al., supra). Las células transformadas se desarrollaron durante la noche en presencia del fago auxiliar M13-K07 (New England Biolabs, Beverly, MA) para producir partículas de fagos que encapsulaban el ADN fagémido y expresaban fragmentos de Fab en sus superficies. La biblioteca combinada contenía más de  $3 \times 10^{10}$  miembros únicos.

### Ejemplo 5. Selección de anticuerpos específicos de la biblioteca intacta SXH3.

45 [0403] Se ciclaron fagos de la biblioteca SXH3 (descrita en el Ejemplo 4 anterior) mediante rondas de selección de unión para enriquecer la unión de clones a HER-2 humana. Las selecciones de unión se realizaron utilizando procedimientos anteriormente descritos (Sidhu et al., supra).

50 [0404] Se cubrieron inmunoplasmas Maxisorp de 96 pocillos NUNC durante toda la noche a 4°C con proteína diana 5 µg/mL (HER-2) y se bloquearon durante 2 horas con una solución de PBT (Sigma). Después de crecer durante toda la noche a 37°C, se concentraron los fagos mediante precipitación con PEG/NaCl y se resuspendieron en PBT, tal y como se ha descrito anteriormente (Sidhu et al., supra). Se añadieron soluciones de fagos (aproximadamente  $10^{12}$  fagos/mL) a las inmunoplasmas recubiertas. Después de una incubación de dos horas para permitir la unión de los fagos, se lavaron las placas diez veces con PBT. Se eluyeron los fagos de unión con HCl 0,1 M durante diez minutos y se neutralizó el eluyente con base Tris 1,0 M. Se amplificaron los fagos eluidos en *E. coli* XL1-blue y se utilizaron para rondas adicionales de selección.

[0405] Se sometieron las bibliotecas a seis rondas de selección contra cada proteína diana. Se desarrollaron

clones individuales de cada ronda de selección en un formato de 96 pocillos en 500  $\mu$ L de caldo 2YT complementado con carbenicilina y M13-K07. Se utilizaron los sobrenadantes de cultivo directamente en ELISA de fagos (Sidhu et al., supra) para detectar Fabs expresados en fagos que se unían a placas cubiertas con proteína diana pero no a placas cubiertas con BSA. Se definieron los enlazadores específicos como aquellos clones de fagos que exhibieron una señal ELISA por lo menos 10 veces mayor en las placas cubiertas con diana en comparación con placas cubiertas con BSA. Se cribaron los clones individuales después de 4, 5 y 6 rondas de selección para unirse a HER-2 humana. Se sometieron los enlazadores específicos a análisis de secuencia. Tal y como se muestra en la Figura 21, la biblioteca SXH3 produjo enlazadores específicos para la proteína diana.

**[0406]** De los 72 clones identificados que se unían específicamente a HER2, 27 de ellos presentaban secuencias de CDR únicas (véase la figura 21A). Las secuencias únicas se encontraban en tres categorías: (1) secuencias CDR con posiciones aleatorizadas limitadas al binomio Tyr/Ser (clones nos. B1-5 y B28); (b) secuencias CDR con posiciones aleatorizadas limitadas al binomio Trp/Ser (clones nos. B6-24); (c) secuencias de CDR con posiciones aleatorizadas limitadas al binomio Phe/Ser (clones nos. B25-27). Estos clones también fueron altamente específicos para HER2 y no mostraban una reactividad cruzada con otras cinco proteínas de control, VEGF humano, DR5 humana, insulina humana, neutravidina, IGF-1 humana o HGH (véase la figura 21B). La concentración inhibitoria para cada clon se muestra en la figura 21B.

**[0407]** Se utilizó un ELISA de fagos para analizar la capacidad de todos los clones de reaccionar de manera cruzada con un panel de seis antígenos diferentes del antígeno diana. Los fagos se produjeron en un formato de 96 pocillos tal como se ha descrito y los sobrenadantes de los fagos se diluyeron tres veces en el tampón PBT. El sobrenadante de fagos diluido se transfirió a placas recubiertas con VEGF humano, HER2, DR5 humana, insulina humana, neutravidina, IGF-1 humana, HGH o BSA y se incubó durante una hora con agitación suave a temperatura ambiente. Las placas se lavaron con PBS que incluía Tween 20 al 0,05% y se incubaron durante 30 minutos con conjugado de peroxidasa de rábano/anticuerpo anti-M13 (diluido 1:5000 en tampón PT) (Farmacia). Las placas se lavaron, se revelaron con sustrato tetrametilbencidina (TMB) (Kirkegaard and Perry Laboratories) y se inhibieron con  $H_3PO_4$  1,0 M. Se determinó la absorbancia espectrofotométricamente a 450 nm. Se definió la reactividad cruzada débil como una señal entre 0,2 y 3,0 y una reactividad cruzada fuerte se definió como una señal de aproximadamente 2,0. Los resultados para los clones de unión a HER2 se muestran en la figura 21B. Tal como se muestra en la figura 25, de los clones de SXH3 aislados, los clones S:R exhibieron la mayor unión no específica promedio (DO 0,5-0,6 a 450 nm mediante ensayo ELISA), mientras que los clones S:W, S:Y, y S:F exhibieron cada uno niveles bajos similares de unión no específica promedio (DO 0-0,1 OD a 450 nm mediante ensayo ELISA).

**[0408]** Se utilizó un ELISA de fagos competitivo para estimar las afinidades de unión de Fabs que expresaban fagos de unión a HER2. Los fagos se produjeron en un formato de 96 pocillos tal como se ha descrito, y los sobrenadantes de fagos se diluyeron en serie en tampón PBT, a continuación se incubaron en placas cubiertas con HER2 durante 15 minutos. Las placas se lavaron con PBS que incluía Tween 20 al 0,05% y se incubaron durante 30 minutos con conjugado de peroxidasa de rábano/anticuerpo anti-M13 (diluido 1:5000 en tampón PT) (Farmacia). Las placas se lavaron, se revelaron con sustrato tetrametilbencidina (TMB) (Kirkegaard and Perry Laboratories) y se inhibieron con  $H_3PO_4$  1,0 M. Se midió la absorbancia espectrofotométricamente a 450 nm para determinar la concentración de fago que produce aproximadamente la 50% de la señal en la saturación. Se diluyeron dos veces una concentración fija de subsaturación de fago en tampón PBT o tampón PBT que contenía diluciones en serie de dos veces de proteína HER2 de HER2 250 nM a HER2 0,12 nM. Las mezclas se incubaron durante una hora con agitación suave a temperatura ambiente, se transfirieron a placas cubiertas con HER2 y las placas se incubaron durante 15 minutos. Las placas se lavaron y trataron exactamente como antes. Las afinidades de unión se estimaron como los valores  $IC_{50}$  (definido como la concentración de antígeno que bloqueaba el 50% de la unión de fago al antígeno inmovilizado). Los resultados se muestran en la figura 21B.

#### **Ejemplo 6 (ejemplo comparativo). Construcción de bibliotecas de Fab expresados en fagos con residuos de CDR enriquecidos en Ser y Phe, Arg, Trp o Tyr.**

**[0409]** Se construyeron bibliotecas de Fab expresados en fagos utilizando un vector fagémido, Fab-C, que dio lugar a la expresión de grupos Fab bivalentes dimerizados por una cisteína libre insertada entre la cadena pesada de Fab y el dominio C-terminal de la proteína de recubrimiento menor del gen-3 (P3C), tal como se ha descrito previamente en el ejemplo 1.

**[0410]** Se construyeron cuatro bibliotecas: SFH3, SRH3, SWH3, y SYH3. Las bibliotecas se construyeron con residuos aleatorizados en las tres CDR de cadena pesada y CDR3 de cadena ligera. Cada biblioteca se aleatorizó en las posiciones 91-94 y 96 de CDRL3, las posiciones 28 y 30-33 de CDRH1, las posiciones 50, 52-54, 56, y 58 de CDRH2, y las posiciones 95-100, 100a a 100m de CDRH3. El tipo y la proporción de los aminoácidos permitidos en cada una de las posiciones aleatorizadas se describen en la figura 22. Además, la longitud de CDRH3 se varió utilizando oligonucleótidos que sustituían los seis codones de tipo natural entre las posiciones 95 y 100 con 4 a 17 codones. El tipo y la proporción de los aminoácidos permitidos en estas posiciones fueron los mismos que los descritos en la figura 22 para las posiciones 95-100 de CDRH3.

**[0411]** Se construyeron bibliotecas utilizando el método de Kunkel (Kunkel, T.A., Roberts, J.D. & Zakour, R.A.,

Methods Enzymol. (1987), 154, 367-382) con métodos descritos previamente (Sidhu, S.S., Lowman, H.B., Cunningham, B. C. & Wells, J.A., Methods Enzymol. (2000), 328, 333-363). Se utilizó una única versión de "plantilla de detención" del vector de expresión de Fab Fab-C para generar las cuatro bibliotecas tal como se describe en el ejemplo 1.

5 **[0412]** Se utilizaron oligonucleótidos mutagénicos con codones degenerados en las posiciones a diversificar para simultáneamente (a) introducir diversidad de CDR y (b) reparar los codones de detención. Las secuencias de estos oligonucleótidos mutagénicos se muestran en las figuras 20 y 23. Para la biblioteca SF-superficie, se introdujo diversidad en CDR-L3, CDR-H1 y CDRH2 con los oligonucleótidos L3-SF, H1-SF y H2-SF respectivamente (SEQ ID NOs:) (Figura 23) y se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SF4, H3-SF5, H3-SF6, H3-SF7, H3-SF8, H3-SF9, H3-SF10, H3-SF11, H3-SF12, H3-SF13, H3-SF14, H3-SF15, H3-SF16, y H3-SF17 (SEQ ID NOs:649-662) (Figura 20 C).

10 **[0413]** Para la biblioteca SR-superficie, se introdujo diversidad en CDR-L3, CDR-H1 y CDR-H2 con los oligonucleótidos L3-SR, H1-SR y H2-SR respectivamente (SEQ ID NOs: ) (Figura 23) y se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SR4, H3-SR5, H3-SR6, H3-SR7, H3-SR8, H3-SR9, H3-SR10, H3-SR11, H3-SR12, H3-SR13, H3-SR14, H3-SR15, H3-SR16, y H3-SR17 (SEQ ID NOs: 733-746) (Figura 20I).

15 **[0414]** Para la biblioteca SW-superficie, se introdujo diversidad en CDR-L3, CDR-H1 y CDR-H2 con los oligonucleótidos L3-SW, H1-SW y H2-SW respectivamente (SEQ ID NOs: 747-760) (Figura 23) y se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SW4, H3-SW5, H3-SW6, H3-SW7, H3-SW8, H3-SW9, H3-SW10, H3-SW11, H3-SW12, H3-SW13, H3-SW14, H3-SW15, H3-SW16, y H3-SW17 (SEQ ID NOs: 761-774) (Figura 20K).

20 **[0415]** Para la biblioteca SY-superficie, se introdujo diversidad en CDR-L3, CDR-H1 y CDR-H2 con los oligonucleótidos L3, H1 y H2 respectivamente (SEQ ID NOs: ) (Figura 20A) y se introdujo diversidad en CDR-H3 con una mezcla equimolar de oligonucleótidos H3-SY4, H3-SY5, H3-SY6, H3-SY7, H3-SY8, H3-SY9, H3-SY10, H3-SY11, H3-SY12, H3-SY13, H3-SY14, H3-SY15, H3-SY16, y H3-SY17 (SEQ ID NOs: 775-788) (Figura 20L).

25 **[0416]** Los oligonucleótidos mutagénicos para todas las CDR a aleatorizar se incorporaron en una única reacción de mutagénesis, de manera que la incorporación simultánea de todos los oligonucleótidos mutagénicos dio lugar a la introducción de la diversidad diseñada en cada posición y la reparación de todos los codones de detención TAA. De este modo, se generó un marco de lectura abierto que codificaba un miembro de la biblioteca de Fab fusionado a un puente de cisteínas homodimerizante y P3C. Después de la mutagénesis, las cuatro bibliotecas se combinaron para crear una única biblioteca denominada SX-superficie.

30 **[0417]** Las reacciones de mutagénesis se sometieron a electroporación en *E. coli* SS320 (Sidhu et al., supra). Las células transformadas se desarrollaron durante la noche en presencia del fago auxiliar M13-K07 (New England Biolabs, Beverly, MA) para producir partículas de fagos que encapsulaban el ADN fagémido y expresaban fragmentos de Fab en sus superficies. La biblioteca combinada contenía más de  $3 \times 10^{10}$  miembros únicos.

#### **Ejemplo 7 (ejemplo comparativo) Selección de anticuerpos específicos de la biblioteca intacta SX-superficie.**

35 **[0418]** Se ciclaron fagos de la biblioteca SX-superficie (descrita en el Ejemplo 6 anterior) mediante rondas de selección de unión para enriquecer la unión de clones a HER-2 humana. Las selecciones de unión se realizaron utilizando procedimientos anteriormente descritos (Sidhu et al., supra).

40 **[0419]** Se cubrieron inmunoplasmas Maxisorp de 96 pocillos NUNC durante toda la noche a 4°C con proteína diana 5 µg/mL (HER-2 humana) y se bloquearon durante 2 horas con una solución de PBT (Sigma). Después de crecer durante toda la noche a 37°C, se concentraron los fagos mediante precipitación con PEG/NaCl y se resuspendieron en PBT, tal y como se ha descrito anteriormente (Sidhu et al., supra). Se añadieron soluciones de fagos (aproximadamente  $10^{12}$  fagos/mL) a las inmunoplasmas recubiertas. Después de una incubación de dos horas para permitir la unión de los fagos, se lavaron las placas diez veces con PBT. Se eluyeron los fagos de unión con HCl 0,1 M durante diez minutos y se neutralizó el eluyente con base Tris 1,0 M. Se amplificaron los fagos eluidos en *E. coli* XL1-blue y se utilizaron para rondas adicionales de selección.

45 **[0420]** Se sometieron las bibliotecas a seis rondas de selección contra cada proteína diana. Se desarrollaron clones individuales de cada ronda de selección en un formato de 96 pocillos en 500 µL de caldo 2YT complementado con carbenicilina y M13-K07. Se utilizaron los sobrenadantes de cultivo directamente en ELISA de fagos (Sidhu et al., supra) para detectar Fabs expresados en fagos que se unían a placas cubiertas con proteína diana pero no a placas cubiertas con BSA. Se definieron los enlazadores específicos como aquellos clones de fagos que exhibieron una señal ELISA por lo menos 10 veces mayor en las placas cubiertas con diana en comparación con placas cubiertas con BSA. Se cribaron los clones individuales después de 4, 5 y 6 rondas de selección para unirse a HER-2 humana. Se sometieron los enlazadores específicos a análisis de secuencia. Tal y como se muestra en la Figura 24, la biblioteca SX-superficie produjo enlazadores específicos contra la proteína

diana.

**[0421]** De los 81 clones identificados que se unían específicamente a HER2, 27 de ellos presentaban secuencias de CDR únicas (G49-61); (b) secuencias CDR con posiciones aleatorizadas limitadas al binomio Trp/Ser (clones nos. G29-48). Los clones de Tyr/Ser fueron altamente específicos para HER2 y no mostraban una reactividad cruzada con otras cinco proteínas de control, VEGF humano, DR5 humana, insulina humana, neutravidina, IGF-1 humana o HGH (véase la figura 24B). Sin embargo, parte de los clones de Trp/Ser reaccionaban de forma cruzada (véase la figura 24B). La concentración inhibidora para cada clon se muestra en la figura 24B.

**[0422]** Se utilizó un ELISA de fagos para analizar la capacidad de todos los clones de reaccionar de manera cruzada con un panel de seis antígenos diferentes del antígeno diana. Los fagos se produjeron en un formato de 96 pocillos tal como se ha descrito y los sobrenadantes de los fagos se diluyeron tres veces en tampón PBT. El sobrenadante de fagos diluido se transfirió a placas recubiertas con VEGF humano, HER2, DR5 humana, insulina humana, neutravidina, IGF-1 humana, HGH o BSA y se incubaron durante una hora con agitación suave a temperatura ambiente. Las placas se lavaron con PBS que incluía Tween 20 al 0,05% y se incubaron durante 30 minutos con conjugado de peroxidasa de rábano/anticuerpo anti-M13 (diluido 1:5000 en tampón PT) (Farmacia). Las placas se lavaron, se revelaron con sustrato tetrametilbencidina (TMB) (Kirkegaard and Perry Laboratories) y se inhibieron con  $\text{H}_3\text{PO}_4$  1,0 M. Se determinó la absorbancia espectrofotométricamente a 450 nm. Se definió la reactividad cruzada débil como una señal entre 0,2 y 3,0 y una reactividad cruzada fuerte se definió como una señal por encima de 2,0. Los resultados para los clones de SX-superficie se muestran en la figura 24B. Tal como se muestra en la figura 27, de los clones de SX-superficie aislados, los clones S:R y S:W exhibieron la mayor unión no específica promedio (DO 0,5-0,6 y aproximadamente 4,0, respectivamente, a 450 nm mediante ensayo ELISA), mientras que los clones S:Y y S:F exhibieron cada uno niveles bajos similares de unión no específica promedio (DO 0-0,1 OD a 450 nm mediante ensayo ELISA).

**[0423]** Se utilizó un ELISA de fagos competitivo para estimar las afinidades de unión de Fabs que expresaban fagos de unión a HER2. Los fagos se produjeron en un formato de 96 pocillos tal como se ha descrito, y los sobrenadantes de fagos se diluyeron en serie en tampón PBT, a continuación se incubaron en placas cubiertas con HER2 durante 15 minutos. Las placas se lavaron con PBS que incluía Tween 20 al 0,05% y se incubaron durante 30 minutos con conjugado de peroxidasa de rábano/anticuerpo anti-M13 (diluido 1:5000 en tampón PT) (Farmacia). Las placas se lavaron, se revelaron con sustrato tetrametilbencidina (TMB) (Kirkegaard and Perry Laboratories) y se inhibieron con  $\text{H}_3\text{PO}_4$  1,0 M. Se midió la absorbancia espectrofotométricamente a 450 nm para determinar la concentración de fago que produce menos del 50% de la señal en la saturación. Se diluyeron dos veces una concentración fija de subsaturación de fago en tampón PBT o tampón PBT que contenía diluciones en serie de dos veces de proteína HER2 de HER2 250 nM a HER2 0,12 nM. Las mezclas se incubaron durante una hora con agitación suave a temperatura ambiente, se transfirieron a placas cubiertas con HER2 y las placas se incubaron durante 15 minutos. Las placas se lavaron y trataron exactamente como antes. Las afinidades de unión se estimaron como los valores  $\text{IC}_{50}$  (definido como la concentración de antígeno que bloqueaba el 50% de la unión de fago al antígeno inmovilizado). Los resultados se muestran en la figura 24B.

**[0424]** En base a este análisis, el análisis de clones de unión a HER2 de la biblioteca SXH3 (ejemplo 6) y la biblioteca YSGRA-D (ejemplo 1), las proteínas Fab solubles de tres clones (clones nos. 42 (YSGR-A) y B11 (SXH3) y G54 (SX-superficie)) se purificaron y sometieron a análisis de resonancia de plasmón superficial de unión a HER2 humana. Se obtuvieron datos BIAcore® según Chen et al., J. Mol. Biol. (1999), 293(4): 865-81. Brevemente, se calcularon las afinidades de unión de las Fab purificadas para Her2 humana a partir de las constantes de velocidad de asociación y disociación mediadas utilizando un sistema de resonancia de plasmón BIAcore®-A100 (BIACORE, Inc., Piscataway, N.J.). HER2 se acopló covalentemente a dos concentraciones diferentes utilizando clorhidrato de N-etil-N'-(3-dimetilaminopropil)-carbodiimida (EDC) e N-hidroxisuccinimida (NHS) según las instrucciones del suministrador (BIAcore, Inc., Piscataway, N.J.). Se intercambió el tampón de HER2 por acetato de sodio 10 mM, pH 5,0 y se diluyó hasta aproximadamente 2,5 ó 5,0  $\mu\text{g}/\text{ml}$ . Se inyectaron alícuotas de HER2 a una velocidad de flujo de 5  $\mu\text{L}/\text{min}$  hasta conseguir aproximadamente 50-170 unidades de respuesta (RU) de la proteína acoplada. Se inyectó una solución de etanolamina 1 M como agente bloqueante. Para las mediciones cinéticas, se inyectaron diluciones en serie de dos veces de cada Fab en HBT a 25 °C a una velocidad de flujo de 10 mL/minuto sobre cada celda de flujo. Los valores de  $k_{\text{on}}$  y  $k_{\text{off}}$  se determinaron a partir de las curvas de unión utilizando el paquete informático BIAevaluation (BIACORE, Inc., Piscataway, N.J.) utilizando un ajuste global de dos puntos y combinando los datos de ambas celdas de flujo. La constante de disociación en equilibrio,  $K_D$ , se calculó como  $K_{\text{off}}/k_{\text{on}}$ . Los datos de BIAcore® se resumen en la figura 26A y B. El clon B11 presentaba una  $k_a$  de  $1,9 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ , una  $k_d$  de  $1,7 \times 10^{-3} \text{ s}^{-1}$ , y una  $K_D$  de 890 pM.  $R_{\text{max}1}$  para los experimentos del B11 fue 19 RU, y  $R_{\text{max}2}$  para los experimentos del clon B11 fue 29 RU. (Figura 26A) El clon G54 presentaba una  $k_a$  de  $2,0 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ , una  $k_d$  de  $2,2 \times 10^{-3} \text{ s}^{-1}$ , y una  $K_D$  de 11 nM.  $R_{\text{max}1}$  para los experimentos del clon G54 fue 21 RU y  $R_{\text{max}2}$  para los experimentos del clon G54 fue 34 RU. (Figura 26A) El clon YSGR-A-42 presentaba una  $k_a$  de  $2,7 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ , una  $k_d$  de  $1,5 \times 10^{-3} \text{ s}^{-1}$ , y una  $K_D$  de 570 pM.  $R_{\text{max}1}$  para los experimentos del clon 42 fue 25 RU, y  $R_{\text{max}2}$  para los experimentos del clon 42 fue 38 RU. (Figura 26B) El clon que contenía triptófano (B11) presentaba una  $k_{\text{on}}$  más rápida y la correspondiente  $K_D$  más pequeña que el clon que contenía tirosina (G54).

**[0425]** Para estudiar la unión de anticuerpos anti-HER2 a HER 2 expresada en células de mamífero, se estudió

mediante citometría de flujo la unión de proteína Fab purificada de los clones 42 (YSGR-A), B11 (SXH3), G54 (SX-superficie), y G37 (SX-superficie) a células de fibroblasto NR6 que sobreexpresan HER2 (NR6-HER2). Se incubaron un millón de células NR6-HER2 con 10 µg/ml de Fab durante 1 hora, seguido de la incubación con un anticuerpo IgG anti-humano murino conjugado a Alexa488 durante 1 hora. Como control negativo, se estudió la unión de Fab a células NR6 no expresantes. Como control positivo, se utilizó Fab de 4D5. Tal como se muestra en la figura 27, los clones 42, B11, G54, y G37 se unen específicamente a Her2 en células NR6.

**[0426]** Se utilizó un ELISA competitivo para analizar la competición de unión con Herceptina y Omnitarg y entre varios clones de HER2 en formato IgG (véase la figura 28 para las secuencias CDR de los clones pertinentes). Se diluyó en serie la proteína HER2 biotinilada desde 200 nM a 0,39 nM en tampón PBT, a continuación se incubó en placas cubiertas de proteínas IgG purificadas durante 15 minutos. Las placas se lavaron con PBS que incluía Tween 20 al 0,05% y se incubaron durante 30 minutos con conjugado de peroxidasa de rábano/anticuerpo anti-M13 (diluido 1:5000 en tampón PT) (Pharmacia). Las placas se lavaron, se revelaron con sustrato tetrametilbencidina (TMB) (Kirkegaard and Perry Laboratories) y se inhibieron con H<sub>3</sub>PO<sub>4</sub> 1,0 M. Se midió la absorbancia espectrofotométricamente a 450 nm para determinar la concentración de HER2 biotinilada que produce aproximadamente el 50% de la señal en la saturación. Se diluyeron dos veces una concentración fija de subsaturación de HER2 biotinilada en tampón PBT o tampón PBT que contenía proteínas IgG 100 nM purificadas. Las mezclas se incubaron durante una hora con agitación suave a temperatura ambiente, se transfirieron a placas cubiertas con proteínas IgG y las placas se incubaron durante 15 minutos. Las placas se lavaron y trataron exactamente como antes. Tal como se muestra en la figura 29, ninguna de las IgG que se unían a HER2 bloqueó la unión de Her2 biotinilada a Omnitarg o Herceptina. Las IgG bloquearon la unión ente sí en dos grupos. Un grupo formado de los clones B11, G37, G54, y YSGR-A-42 competían por el mismo epítipo y bloqueaban la unión a HER2 biotinilada que se había incubado previamente con cualquiera de estos clones. Un segundo grupo formado de los clones YSGR-A-27, B27, G43, y YSGR-D-104 competían por el mismo epítipo en HER2 y bloqueaban la unión a HER2 biotinilada. Los clones del grupo uno eran todos enlazadores de afinidad más elevada que los clones del grupo dos.

Tabla 1

Polipéptido DR5-ECD humano

**MSALLILALVGAAVADYKDDDDKLSALITQQDLAPQQRVAPQQKRSSPSEG  
LCPPGHHIS  
EDGRDCISCKYGDYSTHWNDLLFCLRCTRCDSGEVELSPCTTTRNTVCQC  
EEGTFREED  
SPEMCRKCRGTGCPRGMVKVGDCTPWSDIECVHKESGTHKHSGEAPAVEETVT  
SSPGTPASP  
CSLS (SEQ ID NO:595)**

Polipéptido DR5 humano

**meqrqnapa asgarkrhgp gpreargarp glrvpktivl vvaavlllvs aesalitqqd lapqgraapq  
qkrsspsegl cppghised grdcisckyg qdysthwndll fclrctrcd sgevelspct ttntvcqce  
egtfreedsp emcrkcertgc prgmvkvgcd tpwsdiecvh kesgiiigvt vaavvlivav fvckslwkk  
vlpylkgics ggggdpervd rssqrpgaed nvlneivsil qptqvpeqem evqepaeptg vnmlspgese  
hlepaeaeer sqrrllvpa negdptetr qcfdfdadlv pfdswepmlr klglmdneik vakaeaaghr  
dtlytmlikw vnktrdasv htlldaletl gerlakqkie dhllssgkfm  
ylegnadsal s (SEQ ID NO:604)**

ECD DR5 murino

**GLQRPEESPSRGPCLAGQYLSEGNCKPCREGIDYTSHSNHSLDSCILCTVCKE  
DKVVETR  
CNITNTNVCRCCKPGTFEDKDSPEICQSCSNCTDGEEELTSCTPRENRKCVSKT  
AWASWHK  
(SEQ ID NO:605)**

Secuencia de polipéptido Apo-2L

1 MetAlaMetMetGluValGlnGlyGlyProSerLeuGlyGlnThrCysValLeuIleValIle  
PheThrValLeuLeuGlnSerLeuCys

5 31 ValAlaValThrTyrValTyrPheThrAsnGluLeuLysGlnMetGlnAspLysTyrSerLy  
sSerGlyIleAlaCysPheLeuLysGlu

61 AspAspSerTyrTrpAspProAsnAspGluGluSerMetAsnSerProCysTrpGlnValLy  
sTrpGlnLeuArgGlnLeuValArgLys

10 91 MetIleLeuArgThrSerGluGluThrIleSerThrValGlnGluLysGlnGlnAsnIleSerPr  
oLeuValArgGluArgGlyProGln

121 ArgValAlaAlaHisIleThrGlyThrArgGlyArgSerAsnThrLeuSerSerProAsnSerL  
ysAsnGluLysAlaLeuGlyArgLys

15 151 IleAsnSerTrpGluSerSerArgSerGlyHisSerPheLeuSerAsnLeuHisLeuArgAsn  
GlyGluLeuValIleHisGluLysGly

181 PheTyrTyrIleTyrSerGlnThrTyrPheArgPheGlnGluGluIleLysGluAsnThrLysA  
snAspLysGlnMetValGlnTyrIle

20 211 TyrLysTyrThrSerTyrProAspProIleLeuLeuMetLysSerAlaArgAsnSerCysTrp  
SerLysAspAlaGluTyrGlyLeuTyr

241 SerIleTyrGlnGlyGlyIlePheGluLeuLysGluAsnAspArgIlePheValSerValThrA  
snGluHisLeuIleAspMetAspHis

25 VRERGPQRVA AHITGTRGRS NTLSSPNSKN EKALGRKINS WESSRSGHSF  
LSNLHLRNGE LVIHEKGFYY IYSQTYFRFQ EEKENTKND KQMVQYTYKY  
TSYPDPILLM KSARNSCWSK DAEYGLYSIY QGGIFELKEN DRIFVSVTNE  
HLIDMDHEAS FFGAFLVG (SEQ ID NO:607)



## REIVINDICACIONES

1. Método de selección de un polipéptido que se une a un antígeno diana específico, comprendiendo dicho método:

5 (A) generar una composición que comprende una pluralidad de polipéptidos, en la que dichos polipéptidos comprenden un dominio variable de cadena pesada de inmunoglobulina, en el que:

(i) CDRH1 comprende una secuencia de aminoácidos G-F-X1-I-X2-X3-X4-X5-I-H (SEQ ID NO:22), en la que G está en la posición 26 y X1 está en la posición 28 según el sistema de numeración Kabat; en la que X1 se selecciona entre S e Y; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; y en la que X5 se selecciona entre Y y S;

10 (ii) CDRH2 comprende una secuencia de aminoácidos: X1-I-X2-P-X3-X4-G-X5-T-X6-Y-A-D-S-V-K-G (SEQ ID NO: 23), en la que X1 está en la posición 50 según el sistema de numeración Kabat; en la que X1 se selecciona entre Y y S; en la que X2 se selecciona entre Y y S; en la que X3 se selecciona entre Y y S; en la que X4 se selecciona entre Y y S; en la que X5 se selecciona entre Y y S; y en la que X6 se selecciona entre Y y S; y

15 (iii) CDRH3 comprende:

20 (a) la secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:24), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 50% Y, 25% S, y 25% G; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 50% Y, 25% S, y 25% G, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F;

25 (b) la secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:26), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 25% Y, 50% S, y 25% R; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 25% Y, 50% S, y 25% R, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F;

30 (c) la secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:27), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 38% Y, 25% S, 25% G, y 12% R; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 38% Y, 25% S, 25% G, y 12% R, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F; o

35 (d) la secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-D-Y (SEQ ID NO:28), en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X6 se seleccionan de un grupo de aminoácidos en una proporción molar de 20% Y, 26% S, 26% G, 13% R, 1% A, 1% D, 1% E, 1% F, 1% H, 1% I, 1% K, 1% L, 1% M, 1% N, 1% P, 1% Q, 1% T, 1% V, y 1% W; en la que los aminoácidos en cada una de las posiciones X7-X17 se seleccionan de un grupo de aminoácidos en una proporción molar de 20% Y, 26% S, 26% G, 13% R, 1% A, 1% D, 1% E, 1% F, 1% H, 1% I, 1% K, 1% L, 1% M, 1% N, 1% P, 1% Q, 1% T, 1% V, y 1% W, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre I, M, L, y F; o

40 (e) CDRH3 que comprende una secuencia de aminoácidos: X1-X2-X3-X4-X5-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19, en la que X1 está en la posición 95 según el sistema de numeración Kabat, y en la que los aminoácidos en cada una de las posiciones X1-X17 se seleccionan entre S y uno de A, C, F, G, I, L, N, P, R, T, W, o Y, o no están presentes; en la que X18 se selecciona entre G y A; y en la que X19 se selecciona entre F, L, I, y M.

45 (B) seleccionar uno o más polipéptidos de la composición que se unen a un antígeno diana;

50 (C) aislar dicho uno o más polipéptidos que se unen al antígeno diana de los polipéptidos que no se unen al antígeno diana; y

55

(D) identificar dicho uno o más polipéptidos que se unen al antígeno diana que tienen una afinidad deseada para el antígeno diana;

en el que dicho antígeno diana es HER2.

2. Método, según la reivindicación 1, en el que los polipéptidos son anticuerpos.

5 3. Método, según la reivindicación 1 o la reivindicación 2, en el que los polipéptidos comprenden además un dominio variable de cadena ligera en el que

CDRL3 comprende una secuencia de aminoácidos Q-Q-X1-X2-X3-X4-P-X5-T (SEQ ID NO:25); en la que X1 está en la posición 91 y se selecciona entre Y, H y S; X2 se selecciona entre Y y S; X3 se selecciona entre Y, S y T; X4 se selecciona entre Y, S y T; y X5 se selecciona entre S, P e Y.

10 4. Método, según cualquiera de las reivindicaciones anteriores, cuyo método comprende proporcionar una biblioteca que comprende la pluralidad de polipéptidos, en el que la biblioteca comprende por lo menos  $1 \times 10^4$  polipéptidos distintos.

5. Método, según la reivindicación 4, en el que la biblioteca comprende por lo menos  $1 \times 10^5$ ,  $1 \times 10^6$ ,  $1 \times 10^7$ , ó  $1 \times 10^8$  polipéptidos distintos.

15 6. Método, según cualquiera de las reivindicaciones anteriores, cuyo método comprende construir una biblioteca de partículas de fago o fagémido que expresan la pluralidad de polipéptidos; poner en contacto la biblioteca de partículas con el antígeno diana; y separar las partículas que se unen de las que no se unen al antígeno diana.

7. Método, según cualquiera de las reivindicaciones anteriores, cuyo método comprende:

20 (a) aislar uno o más polipéptidos que se unen específicamente al antígeno diana al poner en contacto una biblioteca que comprende la pluralidad de polipéptidos con un antígeno diana inmovilizado bajo condiciones adecuadas para la unión;

25 (b) separar dichos uno o más polipéptidos que se unen específicamente al antígeno diana de los polipéptidos que no se unen específicamente al antígeno diana, y recuperar dichos uno o más polipéptidos que se unen específicamente al antígeno diana para obtener una subpoblación enriquecida en dichos uno o más polipéptidos que se unen específicamente al antígeno diana; y

(c) opcionalmente, repetir las etapas (a)-(b) por lo menos dos veces, utilizando cada repetición la subpoblación enriquecida en dichos uno o más polipéptidos que se unen específicamente al antígeno diana obtenida de la ronda previa de selección.

## FIG. 1

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys  
Arg Ala Ser Gln Asp Val Asn Thr Ala Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys  
Leu Leu Ile Tyr Ser Ala Ser Ser Leu Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Arg Ser Gly  
Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln His  
Tyr Thr Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr

(SEQ ID NO:1)

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys  
Ala Ala Ser Gly Phe Asn Ile Lys Asp Thr Tyr Ile His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu  
Glu Trp Val Ala Arg Ile Tyr Pro Thr Asn Gly Tyr Thr Arg Tyr Ala Asp Ser Val Lys Gly Arg Phe  
Thr Ile Ser Ala Asp Thr Ser Lys Asn Thr Ala Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp  
Thr Ala Val Tyr Tyr Cys Ser Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln  
Gly Thr Leu Val Thr Val Ser Ser

(SEQ ID NO:2)

Figura 2

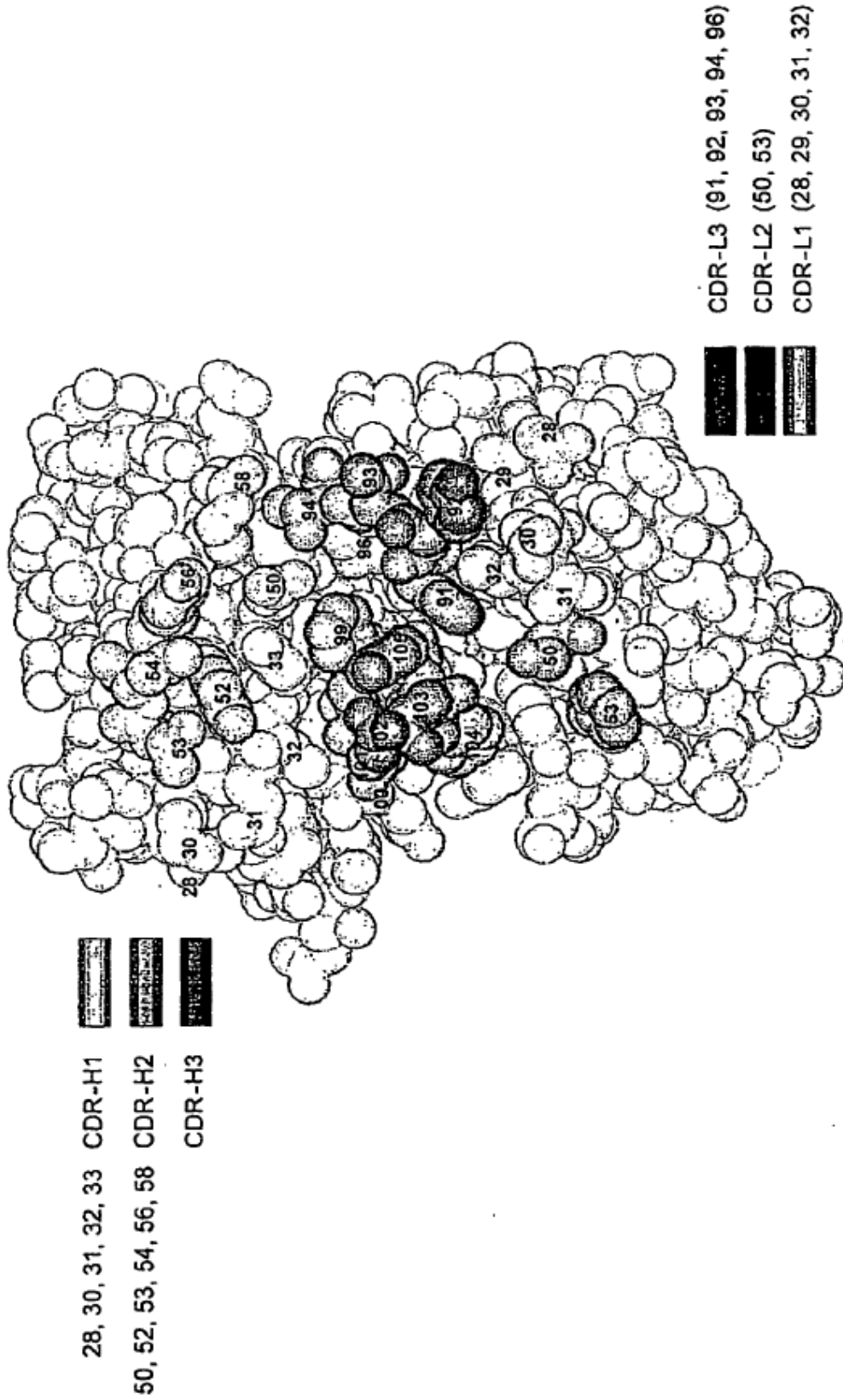


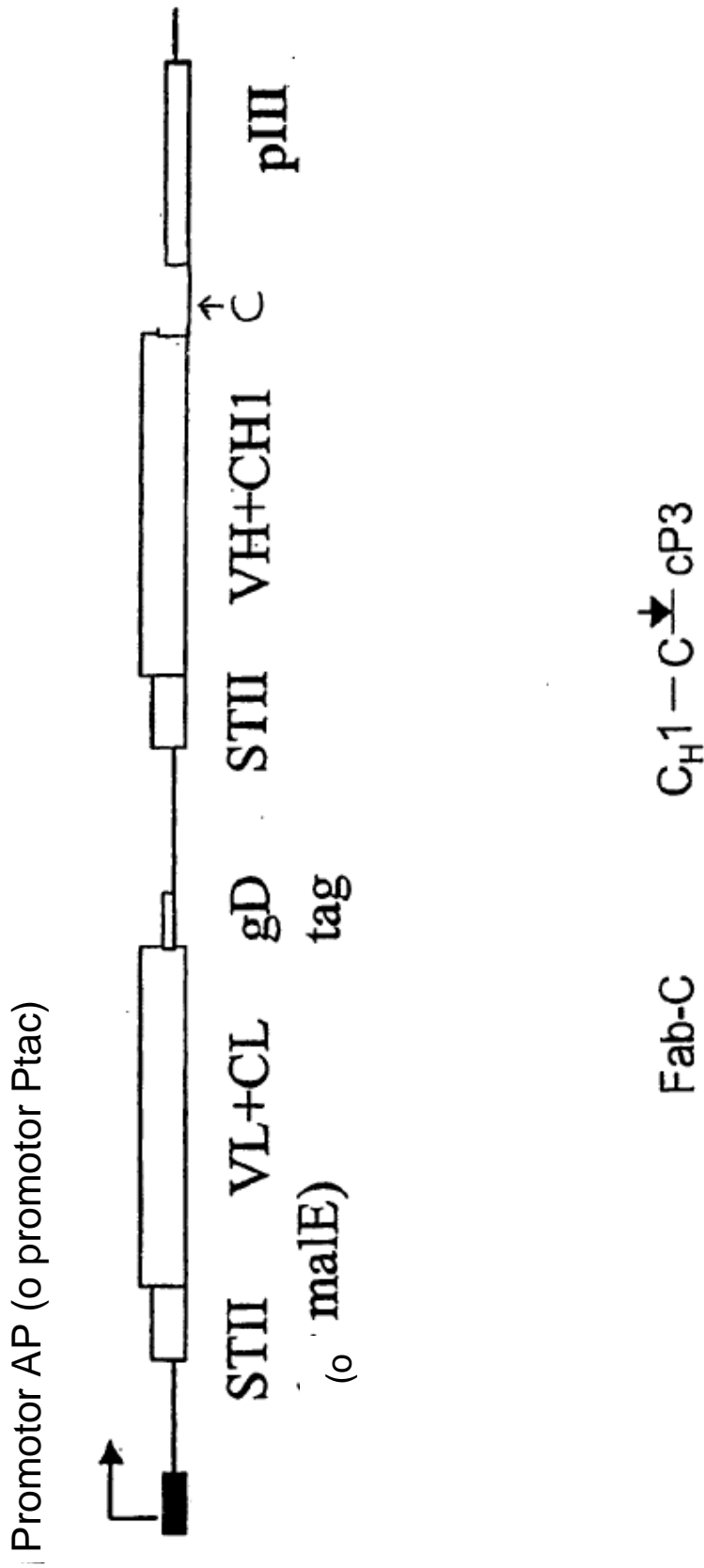
Figura 3

Frecuencia LC

28	S	511	N	262	V	D	G	I	T	L	X	
29	I	612	S	272	V	G	N	X	39	16	35	
30	S	849	N	176	K	G	R	Y	T	D	A	X
31	S	676	N	496	T	R	I	D	29	28	17	45
32	Y	1055	N	128	W	F	S	D	25	18	X	53
					97	77	61	40	25	69		
50	G	386	A	341	D	W	K	L	E	S	X	82
53	S	545	N	438	T	K	I	R	39	30		
					407	41	23	23	X	58		
91	Y	849	S	196	R	A	G	H	X			
92	Y	362	G	356	N	S	D	L	148	H	I	X
93	S	738	N	346	Q	T	H	G	64	43	38	91
94	S	386	T	365	W	Y	L	F	D	R	X	
96	L	264	Y	205	W	F	I	R	47	35	112	
					176	140	117	115	A	P	Y	I
									46	43	33	24
									P	X	N	X
									46	121	18	40



Figura 5



Secuencia de armazón de cadena ligera de huMAb4D5-8

- LC-FR1      <sup>1</sup>Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys<sup>23</sup> (SEQ ID NO:6)
- LC-FR2      <sup>35</sup>Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr<sup>49</sup> (SEQ ID NO:7)
- LC-FR3      <sup>57</sup>Gly Val Pro Ser Arg Phe Ser Gly Ser Arg Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys<sup>88</sup> (SEQ ID NO:8)
- LC-FR4      <sup>98</sup>Phe Gly Gln Gly Thr Lys Val Glu Ile Lys<sup>107</sup> (SEQ ID NO:9)

Secuencia de armazón de cadena pesada de huMAb4D5-8

- HC-FR1      <sup>1</sup>Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser<sup>25</sup> (SEQ ID NO:10)
- HC-FR2      <sup>36</sup>Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val<sup>48</sup> (SEQ ID NO:11)
- HC-FR3      <sup>66</sup>Arg Phe Thr Ile Ser Ala Asp Thr Ser Lys Asn Thr Ala Tyr Leu Gln Met Asn<sup>83</sup> Ser<sup>83a</sup> Leu<sup>83b</sup> Arg<sup>83c</sup> Ala Glu Asp Thr Ala Val Tyr Tyr Cys<sup>92</sup> (SEQ ID NO:12)
- HC-FR4      <sup>103</sup>Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser<sup>113</sup> (SEQ ID NO: 13)

FIG. 6



- Secuencia de armazón de cadena ligera de huMAb4D5-8 modificada en posición 66 (subrayado)
- LC-FR1      <sup>1</sup>Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys<sup>23</sup> (SEQ ID NO:14)
- LC-FR2      <sup>35</sup>Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr<sup>49</sup> (SEQ ID NO:15)
- LC-FR3      <sup>57</sup>Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys<sup>88</sup> (SEQ ID NO:16)
- LC-FR4      <sup>98</sup>Phe Gly Gln Gly Thr Lys Val Glu Ile Lys<sup>107</sup> (SEQ ID NO:17)
- Secuencia de armazón de cadena pesada de huMAb4D5-8 modificada en las posiciones 71, 73 y 78 (subrayado)
- HC-FR1      <sup>1</sup>Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser<sup>25</sup> (SEQ ID NO:18)
- HC-FR2      <sup>36</sup>Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val<sup>48</sup> (SEQ ID NO:19)
- HC-FR3      <sup>66</sup>Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn<sup>83</sup> Ser<sup>83a</sup> Leu<sup>83b</sup> Arg<sup>83c</sup> Ala Glu Asp Thr Ala Val Tyr Tyr Cys<sup>92</sup> (SEQ ID NO:20)
- HC-FR4      <sup>103</sup>Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser<sup>113</sup> (SEQ ID NO:21)

FIG. 7

CDR	Posición	biblioteca		biblioteca		biblioteca		biblioteca	
		YSGR-A	relación molar aminoácido	YSGR-A	relación molar aminoácido	YSGR-C	relación molar aminoácido	YSGR-D	relación molar aminoácido
L3	91	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	92	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	93	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	94	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	96	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
H1	28	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	30	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	31	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	32	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	33	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
H2	50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	52	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	53	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	54	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	56	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
	58	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50	Y, S	50, 50
H3	95								
	96								
	97	Y, S, G	50, 25, 25	Y, S, R	25, 50, 25	Y, S, G, R	38, 25, 25, 12	Y, S, G, R, A, D, E, F, H, I, K, L, M, N, P, Q, T, V, W	20, 26, 26, 13, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
	98								
	99								
	100								
	100a	G, A	50, 50	G, A	50, 50	G, A	50, 50	G, A	50, 50
100b	I, M, L, F	25, 25, 25, 25	I, M, L, F	25, 25, 25, 25	I, M, L, F	25, 25, 25, 25	I, M, L, F	25, 25, 25, 25	
100c									

FIG. 8

Figura 9A: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas YSGR-A, YSGR-B, YSGR-C e YSGR-D. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = T/C, R = A/G y N =A/T/G/C). La notación "XXX" representa codones que codifican Tyr/Ser/Gly en una proporción molar de 50/25/25, respectivamente.

**H1**

GCA GCT TCT GGC TTC TMT ATT TMT TMT TMT TMT ATA CAC TGG GTG CGT (SEQ ID NO:32)

**H2**

CTG GAA TGG GTT GCA TMT ATT TMT CCA TMT TMT GGT TMT ACT TMT TAT GCC GAT AGC GTC (SEQ ID NO:33)

**L3**

ACT TAT TAC TGT CAG CAA TMT TMT TMT TMT CCA TMT ACG TTC GGA CAG GGT ACC (SEQ ID NO:34)

**H3-A6**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:35)

**H3-A7**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:36)

**H3-A8**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:37)

**H3-A9**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:38)

**H3-A10**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:39)

**H3-A11**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:40)

**H3-A12**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:41)

**H3-A13**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:42)

**H3-A14**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:43)

**H3-A15**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:44)

**H3-A16**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:45)

**H3-A17**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:46)

Figura 9B: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas YSGR-B. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = T/C, R = A/G y N =A/T/G/C). La notación "XXX" representa codones que codifican Tyr/Ser/Arg en una proporción molar de 25/50/25, respectivamente.

**H3-B6**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT  
CAA GGA (SEQ ID NO:47)

**H3-B7**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG  
GGT CAA GGA (SEQ ID NO:48)

**H3-B8**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC  
TGG GGT CAA GGA (SEQ ID NO:49)

**H3-B9**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC  
TAC TGG GGT CAA GGA (SEQ ID NO:50)

**H3-B10**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK  
GAC TAC TGG GGT CAA GGA (SEQ ID NO:51)

**H3-B11**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST  
WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:52)

**H3-B12**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:53)

**H3-B13**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:54)

**H3-B14**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:55)

**H3-B15**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:56)

**H3-B16**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:57)

**H3-B17**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:58)

Figura 9C: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas YSGR-C. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = T/C, R = A/G y N =A/T/G/C). La notación "XXX" representa codones que codifican Tyr/Ser/Gly/Arg en una proporción molar de 38/25/25/12, respectivamente.

**H3-C6**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT  
CAA GGA (SEQ ID NO:59)

**H3-C7**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG  
GGT CAA GGA (SEQ ID NO:60)

**H3-C8**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC  
TGG GGT CAA GGA (SEQ ID NO:61)

**H3-C9**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC  
TAC TGG GGT CAA GGA (SEQ ID NO:62)

**H3-C10**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK  
GAC TAC TGG GGT CAA GGA (SEQ ID NO:63)

**H3-C11**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST  
WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:64)

**H3-C12**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:65)

**H3-C13**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:66)

**H3-C14**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:67)

**H3-C15**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:68)

**H3-C16**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:69)

**H3-C17**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:70)

Figura 9D: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas YSGR-D. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = T/C, R = A/G y N =A/T/G/C). La notación "XXX" representa codones que codifican Tyr/Ser/Gly/Arg/Ala/Asp/Glu/Phe/His/Ile/Lys/Leu/Met/Asn/Pro/Gln/Thr/Val/Trp en una proporción molar de 20/26/26/13/1/1/1/1/1/1/1/1/1/1/1/1/1/1/1/1/1, respectivamente.

5

**H3-D6**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT  
CAA GGA (SEQ ID NO:71)

**H3-D7**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC TGG  
GGT CAA GGA (SEQ ID NO:72)

**H3-D8**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC TAC  
TGG GGT CAA GGA (SEQ ID NO:73)

**H3-D9**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK GAC  
TAC TGG GGT CAA GGA (SEQ ID NO:74)

**H3-D10**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST WTK  
GAC TAC TGG GGT CAA GGA (SEQ ID NO:75)

**H3-D11**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX GST  
WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:76)

**H3-D12**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:77)

**H3-D13**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:78)

**H3-D14**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:79)

**H3-D15**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:80)

**H3-D16**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:81)

**H3-D17**

GTC TAT TAT TGT GCT CGC XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX  
XXX XXX XXX XXX XXX GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO:82)

Figura 10. Número de enlazadores específicos observados mediante ELISA de fagos. (clones únicos/enlazadores específicos)

	Biblioteca
	YSGR-A-D
5 Enlazadores HER-2 humanos	106/240
Enlazadores DR5 humanos	18/144







Figura 12: Fabs alfa-HER2 contienen secuencias cortas de CDRH3

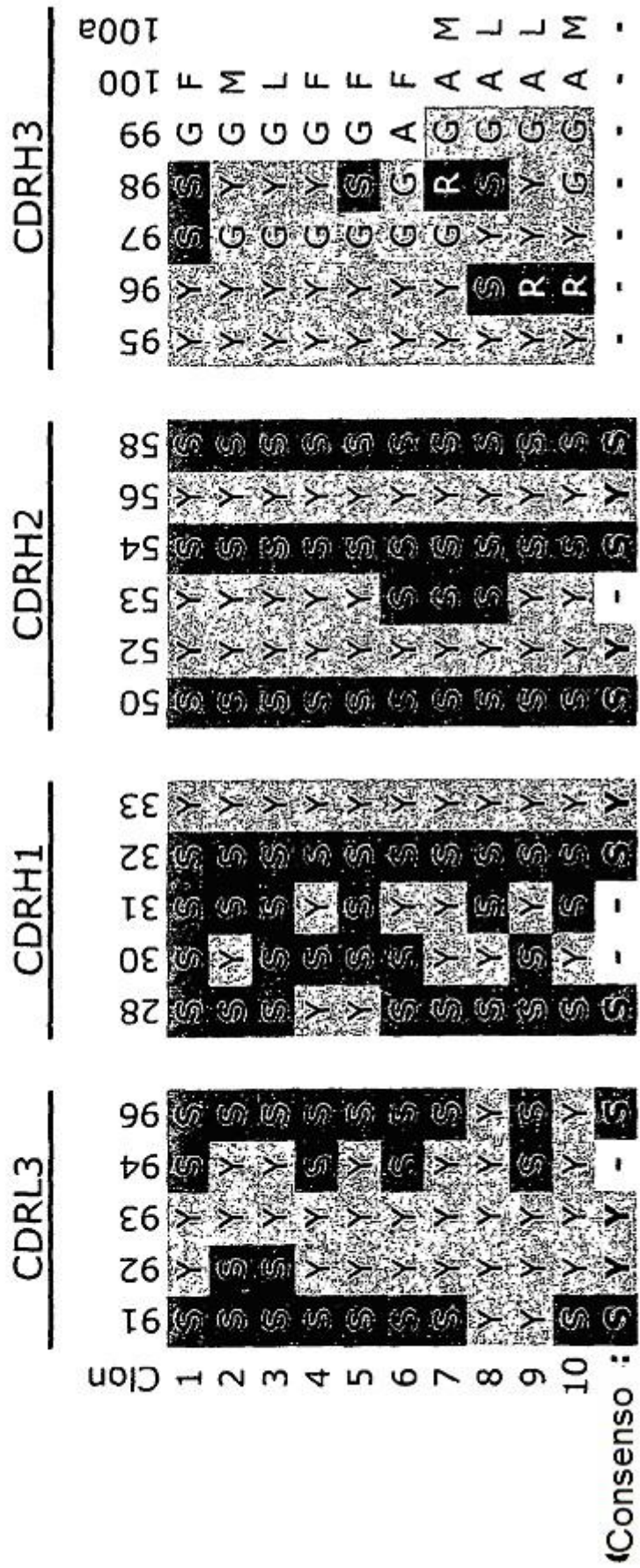
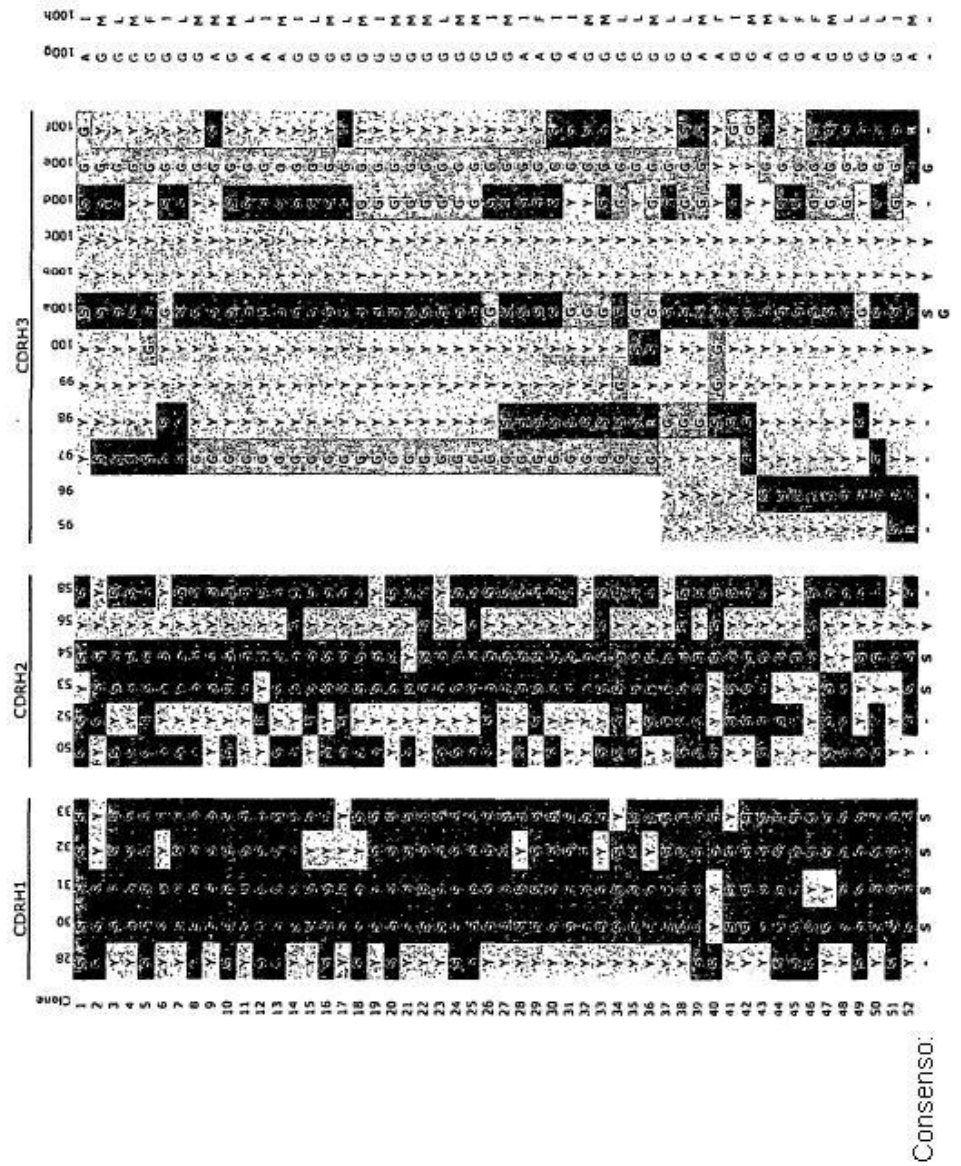




Figura 14: YSG de longitud media anti-HER2



clones de unión a DR5	CDRH1			CDRH2			CDRH3			reacción de forma cruzada con DR5 murino
	SEQ ID NO	CDRH1	CDRH2	CDRH3	SEQ ID NO	CDRH2	CDRH3	SEQ ID NO	IC50 nM	
1	97	Q Q Q Y S S Y P S T 507	S I S P Y Y G G Y T S Y A D S V K G 541	R Y S S S Y S Y S G L	66	66	66	66	ND	
2	98	Q Q Y S S Y P S T 508	S I P Y S G S T S Y A D S V K G 542	R Y S S S Y S Y S G L	67	67	67	67	ND	
3	99	Q Q Y S S Y P S T 509	S I P Y S G S T S Y A D S V K G 543	R Y S S S Y S Y S G L	68	68	68	68	ND	
4	99	Q Q Y S S Y P S T 510	S I P Y S G S T S Y A D S V K G 544	R Y S S S Y S Y S G L	69	69	69	69	ND	
5	99	Q Q Y T P P T 511	S I P Y S G Y T Y A D S V K G 545	R R S Y R Y G Y R G Y G F	70	70	70	70	160	
6	99	Q Q Y S S P S T 512	S I P Y S G Y T Y A D S V K G 546	R R S Y R Y G Y R G Y G F	71	71	71	71	400	
7	99	Q Q Y S S P S T 513	S I P Y S G Y T Y A D S V K G 547	R R S Y R Y G Y R G Y G F	72	72	72	72	130	
8	99	Q Q Y S S P S T 514	S I P Y S G Y T Y A D S V K G 548	R R S Y R Y G Y R G Y G F	73	73	73	73	25	
9	99	Q Q Y S S P S T 515	S I P Y S G Y T Y A D S V K G 549	R R Y Y G Y G Y G Y G F	74	74	74	74	1300	
10	99	Q Q Y T P P T 516	S I P Y S G Y T Y A D S V K G 550	R R Y Y G Y G Y G Y G F	75	75	75	75	250	
11	99	Q Q Y S S P S T 517	S I P Y S G Y T Y A D S V K G 551	R R Y R Y G Y S G Y G F	76	76	76	76	800	
12	99	Q Q Y S S P Y T 518	Y I P Y S G Y T S Y A D S V K G 552	R R Y R Y G Y S G Y G F	77	77	77	77	250	
13	99	Q Q Y S S P S T 519	S I P S G Y T S Y A D S V K G 553	R R Y R Y G Y S G Y G F	78	78	78	78	800	
14	99	Q Q Y S S P S T 520	S I P S G Y T S Y A D S V K G 554	R R Y R Y G Y S G Y G F	79	79	79	79	250	
15	99	Q Q S Y S P S T 521	S I P S G Y T S Y A D S V K G 555	R R Y R Y G Y S R G S Y A F	80	80	80	80	800	
16	99	Q Q Y S S P S T 522	Y I S P Y Y G S T S Y A D S V K G 556	R R Y R Y G S Y R G S Y A F	81	81	81	81	3000	
17	99	Q Q Y S S P S T 523	Y I S P S Y G S T Y A D S V K G 557	R R R Y H Y G R Y S G A Y G L	82	82	82	82	20	
										130
										ND

FIG. 15

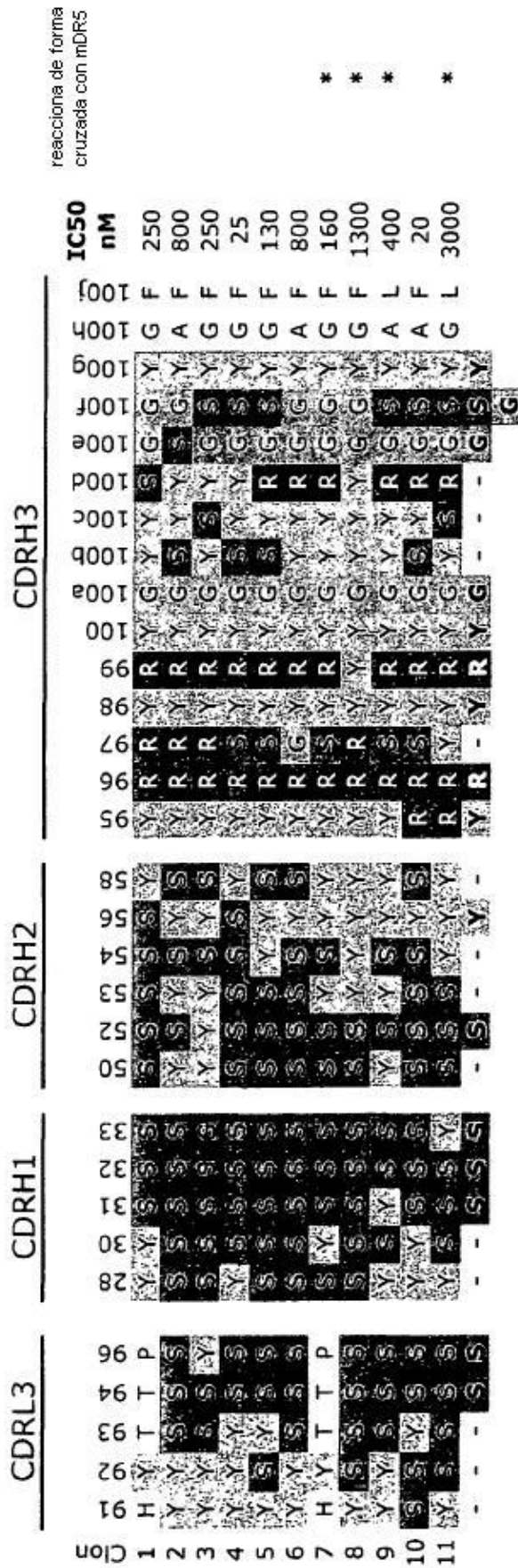
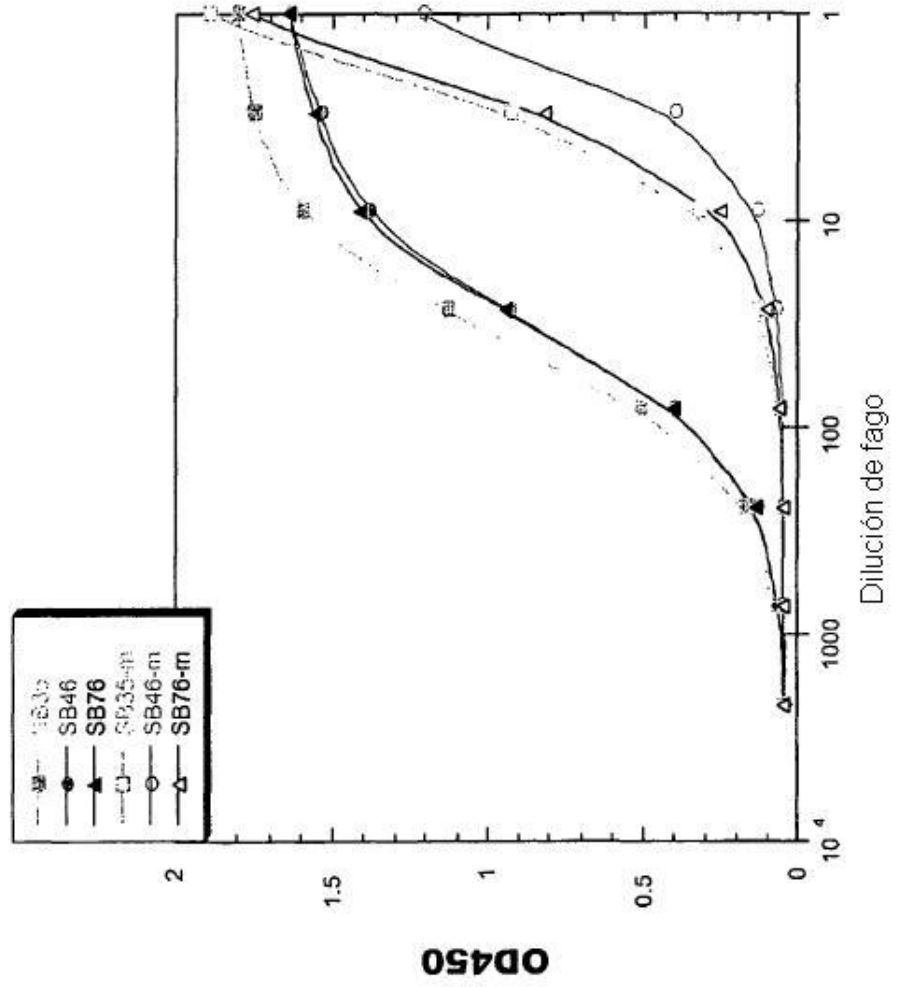
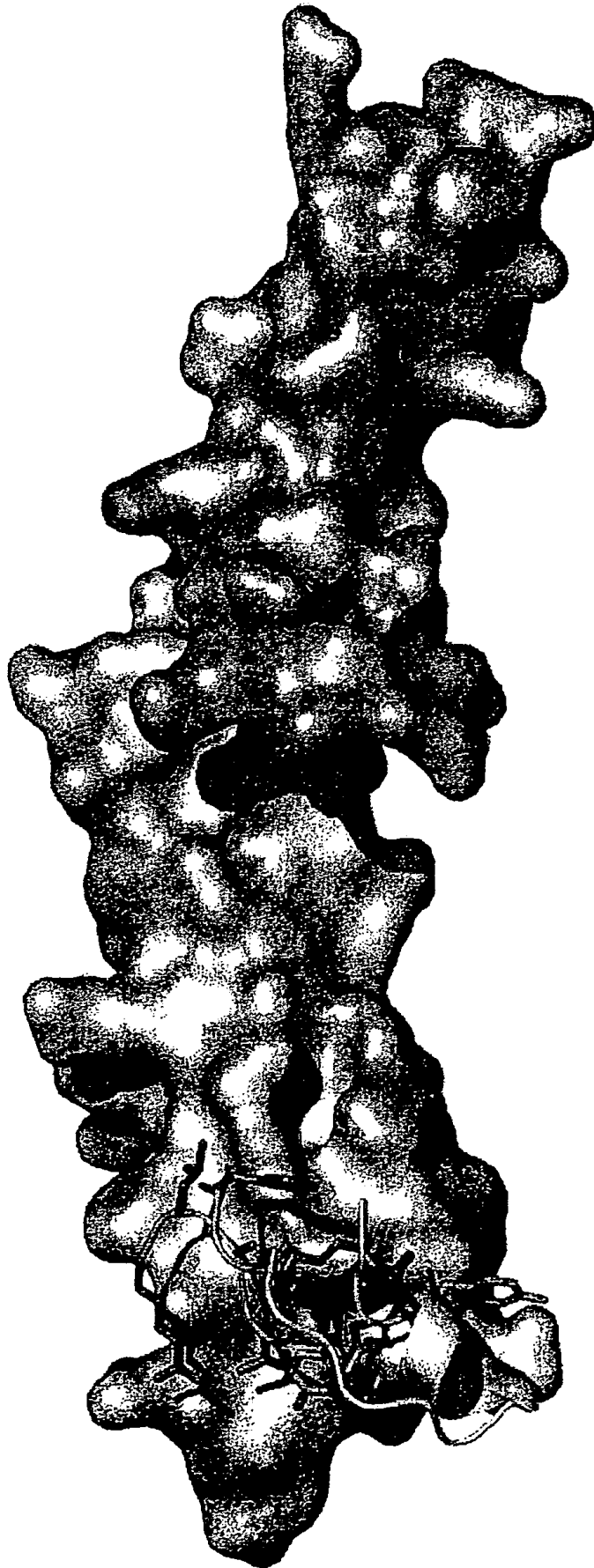


FIG. 16

Figura 17: Unión de DR5 a DR5 humano y murino





**FIG. 18**



Figura 19A

biblioteca H3 binaria

CDR	Posición	SAH3		SCH3		SFH3		SGH3		SIH3		SLH3	
		tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido
L3	91	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	92	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	93	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	94	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	96	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
H1	28	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	30	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	31	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	32	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	33	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	52	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	53	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
H2	54	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	56	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	58	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	58	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
H3	95												
	96												
	97												
	98	S,A	50,50	S,C	50,50	S,F	50,50	S,G	50,50	S,I	50,50	S,L	50,50
	99												
	100												
	100x	G,A	50,50	G,A	50,50	G,A	50,50	G,A	50,50	G,A	50,50	G,A	50,50
	100x	I,M,L,F	25,25,25,25	I,M,L,F	25,25,25,25	I,M,L,F	25,25,25,25	I,M,L,F	25,25,25,25	I,M,L,F	25,25,25,25	I,M,L,F	25,25,25,25

Figura 19B

CDR	posición	SNH3		SPH3		SRH3		STH3		SWH3		SYH3	
		tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido
L3	91	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	92	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	93	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	96	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
H1	28	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	30	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	31	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	32	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	33	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
H2	52	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	53	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	54	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	56	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	58	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
	50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50	S,Y	50,50
H3	95												
	96												
	97	S,N	50,50	S,P	50,50	S,R	50,50	S,T	50,50	S,W	50,50	S,Y	50,50
	98												
	99												
	100												
	100x	G,A	50,50	G,A	50,50	G,A	50,50	G,A	50,50	G,A	50,50	G,A	50,50
	100x	I,M,L,F	25,25,25,25	I,M,L,F	25,25,25,25	I,M,L,F	25,25,25,25	I,M,L,F	25,25,25,25	I,M,L,F	25,25,25,25	I,M,L,F	25,25,25,25

Figura 20A: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SAH3, SCH3, SFH3, SGH3, SIH3, SLH3, SNH3, SPH3, SRH3, STH3, SWH3 y SYH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

**H1**  
GCA GCT TCT GGC TTC TMT ATT TMT TMT TMT TMT ATA CAC TGG GTG CGT (SEQ ID NO: 618)

**H2**  
CTG GAA TGG GTT GCA TMT ATT TMT CCA TMT TMT GGT TMT ACT TMT TAT GCC GA'T AGC  
GTC (SEQ ID NO: 619)

**L3**  
ACT TAT TAC TGT CAG CAA TMT TMT TMT TMT CCA TMT ACG TTC GGA CAG GGT ACC (SEQ  
ID NO: 620)

**H3-SA4**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID  
NO: 621)

**H3-SA5**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT GST WTK GAC TAC TGG GGT CAA GGA  
(SEQ ID NO: 622)

**H3-SA6**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT GST WTK GAC TAC TGG GGT CAA  
GGA (SEQ ID NO: 623)

**H3-SA7**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT KCT GST WTK GAC TAC TGG GGT CAA  
GGA (SEQ ID NO: 624)

**H3-SA8**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT KCT KCT GST WTK GAC TAC TGG GGT  
CAA GGA (SEQ ID NO: 625)

**H3-SA9**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT KCT KCT KCT GST WTK GAC TAC TGG  
GGT CAA GGA (SEQ ID NO: 626)

**H3-sA10**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT GST WTK GAC TAC  
TGG GGT CAA GGA (SEQ ID NO: 627)

**H3-SA11**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT GST WTK GAC  
TAC TGG GGT CAA GGA (SEQ ID NO: 628)

**H3-SA12**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT GST WTK  
GAC TAC TGG GGT CAA GGA (SEQ ID NO: 629)

**H3-SA13**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT GST  
WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 630)

**H3-SA14**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT  
GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 631)

**H3-SA15**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT  
KCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 632)

**H3-SA16**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT  
KCT KCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 633)

**H3-SA17**  
GTC TAT TAT TGT GCT CGC KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT KCT  
KCT KCT KCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 634)

## ES 2 388 932 T3

Figura 20B: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SCH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

### H3-SC4

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 635)

### H3-SC5

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 636)

### H3-SC6

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 637)

### H3-SC7

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 638)

### H3-SC8

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 639)

### H3-SC9

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 640)

### H3-SC10

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 641)

### H3-SC11

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 642)

### H3-SC12

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 643)

### H3-SC13

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 644)

### H3-SC14

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 645)

### H3-SC15

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 646)

### H3-SC16

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 647)

### H3-SC17

GTC TAT TAT TGT GCT CGC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC TSC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 648)

Figura 20C: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SFH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

**H3-SF4**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 649)

**H3-SF5**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 650)

**H3-SF6**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 651)

**H3-SF7**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 652)

**H3-SF8**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 653)

**H3-SF9**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 654)

**H3-SF10**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 655)

**H3-SF11**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 656)

**H3-SF12**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 657)

**H3-SF13**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 658)

**H3-SF14**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 659)

**H3-SF15**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 660)

**H3-SF16**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 661)

**H3-SF17**

GTC TAT TAT TGT GCT CGC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC TYC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 662)

Figura 20D: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SGH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

**H3-SG4**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 663)

**H3-SG5**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 664)

**H3-SG6**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 665)

**H3-SG7**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 666)

**H3-SG8**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 667)

**H3-SG9**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 668)

**H3-SG10**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 669)

**H3-SG11**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 670)

**H3-SG12**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 671)

**H3-SG13**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 672)

**H3-SG14**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 673)

**H3-SG15**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 674)

**H3-SG16**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 675)

**H3-SG17**

GTC TAT TAT TGT GCT CGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC RGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 676)

Figura 20E: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SIH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

**H3-SI4**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 677)

**H3-SI5**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 678)

**H3-SI6**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 679)

**H3-SI7**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 680)

**H3-SI8**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 681)

**H3-SI9**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 682)

**H3-SI10**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 683)

**H3-SI11**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 684)

**H3-SI12**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 685)

**H3-SI13**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 686)

**H3-SI14**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 687)

**H3-SI15**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 688)

**H3-SI16**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 689)

**H3-SI17**

GTC TAT TAT TGT GCT CGC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC AKC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 690)

Figura 20F: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SLH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

**H3-SL4**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 691)

**H3-SL5**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 692)

**H3-SL6**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 693)

**H3-SL7**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 694)

**H3-SL8**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 695)

**H3-SL9**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 696)

**H3-SL10**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 697)

**H3-SL11**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 698)

**H3-SL12**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 699)

**H3-SL13**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 700)

**H3-SL14**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 701)

**H3-SL15**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 702)

**H3-SL16**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 703)

**H3-SL17**

GTC TAT TAT TGT GCT CGC TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA TYA GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 704)



Figura 20G: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SNH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

**H3-SN4**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 705)

**H3-SN5**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 706)

**H3-SN6**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 707)

**H3-SN7**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 708)

**H3-SN8**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 709)

**H3-SN9**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 710)

**H3-SN10**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 711)

**H3-SN11**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 712)

**H3-SN12**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 713)

**H3-SN13**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 714)

**H3-SN14**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 715)

**H3-SN15**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 716)

**H3-SN16**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 717)

**H3-SN17**

GTC TAT TAT TGT GCT CGC ARC ARC ARC ARC ARC ARC ARC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 718)

Figura 20H: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SPH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

**H3-SP4**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 719)

**H3-SP5**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 720)

**H3-SP6**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 721)

**H3-SP7**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 722)

**H3-SP8**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 723)

**H3-SP9**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 724)

**H3-SP10**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 725)

**H3-SP11**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 726)

**H3-SP12**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 727)

**H3-SP13**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 728)

**H3-SP14**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 729)

**H3-SP15**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 730)

**H3-SP16**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 731)

**H3-SP17**

GTC TAT TAT TGT GCT CGC YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT YCT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 732)

Figura 20l: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SRH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

**H3-SR4**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 733)

**H3-SR5**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 734)

**H3-SR6**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 735)

**H3-SR7**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 736)

**H3-SR8**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 737)

**H3-SR9**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 738)

**H3-SR10**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 739)

**H3-SR11**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 740)

**H3-SR12**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 741)

**H3-SR13**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 742)

**H3-SR14**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 743)

**H3-SR15**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 744)

**H3-SR16**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 745)

**H3-SR17**

GTC TAT TAT TGT GCT CGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC MGC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 746)

Figura 20J: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas STH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

**H3-ST4**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 747)

**H3-ST5**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 748)

**H3-ST6**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 749)

**H3-ST7**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 750)

**H3-ST8**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 751)

**H3-ST9**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 752)

**H3-ST10**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 753)

**H3-ST11**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 754)

**H3-ST12**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 755)

**H3-ST13**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 756)

**H3-ST14**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 757)

**H3-ST15**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 758)

**H3-ST16**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 759)

**H3-ST17**

GTC TAT TAT TGT GCT CGC ASC ASC ASC ASC ASC ASC ASC GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 760)

Figura 20L: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SYH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

**H3-SY4**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 775)

**H3-SY5**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 776)

**H3-SY6**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 777)

**H3-SY7**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 778)

**H3-SY8**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 779)

**H3-SY9**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 780)

**H3-SY10**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 781)

**H3-SY11**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 782)

**H3-SY12**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 783)

**H3-SY13**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 784)

**H3-SY14**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 785)

**H3-SY15**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 786)

**H3-SY16**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 787)

**H3-SY17**

GTC TAT TAT TGT GCT CGC TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT TMT GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 788)

Figura 20K: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SWH3. Las degeneraciones del ADN equimolar se representan en el código IUB (W = A/T, K = G/T, M = A/C, S = G/C, Y = C/T, R = A/G).

**H3-SW4**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 761)

**H3-SW5**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 762)

**H3-SW6**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 763)

**H3-SW7**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 764)

**H3-SW8**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 765)

**H3-SW9**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 766)

**H3-SW10**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 767)

**H3-SW11**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 768)

**H3-SW12**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 769)

**H3-SW13**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 770)

**H3-SW14**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 771)

**H3-SW15**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 772)

**H3-SW16**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 773)

**H3-SW17**

GTC TAT TAT TGT GCT CGC TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG TSG GST WTK GAC TAC TGG GGT CAA GGA (SEQ ID NO: 774)

Figura 21A

Clones SXH3 de unión a HER2

Clon	CDRH3	CDRH2	CDRH1	CDRH3
81	Y Y S Y A F	S I Y P Y S G Y T S Y A D S V K G	G F Y I S S S Y I H	Q Q S Y Y S P S T 789
82	Y Y S S G F	S I Y P Y S G Y T S Y A D S V K G	G F Y I S S S Y I H	Q Q S Y Y S P S T 790
83	Y Y S S G F	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q S Y Y S P S T 791
84	Y Y S S G F	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q S Y Y S P S T 792
85	Y Y S Y A L	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q S Y Y S P S T 793
86	W W S S A F	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q Y Y Y S P S T 794
87	W W S S A M	S I Y P S S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q S Y Y Y P S T 795
876	W W S S A F	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q S Y Y Y P S T 796
877	W W S W G F	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q S Y Y Y P S T 797
878	W W S S A F	S I Y P S S G Y T S Y A D S V K G	G F S I Y S Y I H	Q Q S Y Y S P S T 797
879	W W S S G F	S I Y P S S G Y T S Y A D S V K G	G F S I Y S Y I H	Q Q S Y Y S P S T 798
880	W W S S A F	S I Y P Y S G Y T S Y A D S V K G	G F S I Y S S Y I H	Q Q S Y Y Y P S T 799
881	W W S W A F	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q Y Y Y S P S T 800
882	W W S S A F	S I Y P S S G Y T S Y A D S V K G	G F S I Y S Y I H	Q Q S Y Y Y P S T 801
883	W W S S A F	S I Y P S S G Y T S Y A D S V K G	G F S I Y S Y I H	Q Q Y Y Y P Y T 802
884	W W S S A L	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q S Y Y S P S T 803
885	W W S S A F	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q Y Y Y P S T 804
886	W W S S G F	S I Y P Y S G Y T S Y A D S V K G	G F Y I Y S S Y I H	Q Q S Y Y Y P S T 805
887	W W S S A F	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q Y Y Y S P S T 806
888	W W S S A L	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q S Y Y S P S T 807
889	W S S S G F	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q Y Y Y S P S T 808
890	W W S S A M	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q S Y Y Y P S T 809
891	W W S S A L	S I Y P Y S G Y T S Y A D S V K G	G F Y I S S S Y I H	Q Q S Y Y S P S T 810
892	W W S S A L	S I Y P Y S G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q S Y Y S P S T 811
893	F S F S S S F A M	S I Y P Y Y G Y T Y A D S V K G	G F S I S S S Y I H	Q Q Y S Y Y P Y T 812
894	F S F S S S F A I	S I Y P Y Y G Y T Y A D S V K G	G F S I S S S Y I H	Q Q S S Y P Y T 813
895	F F F F S S F F G F	S I Y P Y Y G Y T S Y A D S V K G	G F S I S S S Y I H	Q Q S Y Y P Y T 814
896	Y Y S S Y S Y S F G L	Y I S P S S G Y T S Y A D S V K G	G F S I S S S S I H	Q Q S Y S S P Y T 815

Figura 21B

ID	señal ELISA a 450 nm										Fab-fago IC50 (nM)	
	humano VEGF	humano HER2	DR5	Insulin	humano dln	Neutravi humano	IGF-1	suero bovino	albumina de	IC50 (nM)	error +/-	
B1	0.03	0.01	0.01	0.08	0.01	0.02	0.02	0.00	6.28	0.38		
B2	0.01	0.01	0.00	0.05	0.01	0.02	0.02	-0.01	4.40	0.42		
B3	0.01	0.01	-0.01	0.05	0.00	0.02	0.02	-0.01	5.77	0.47		
B4	0.05	0.01	0.01	0.08	0.00	0.01	0.01	-0.02	14.64	1.04		
B5	0.00	0.00	0.00	0.08	-0.01	0.00	0.00	-0.02	9.26	0.80		
B6	0.05	0.01	0.00	0.05	0.02	0.02	0.02	-0.02	0.82	0.09		
B7	0.05	0.01	0.00	0.05	0.05	0.01	0.01	-0.03	2.54	0.48		
B8	0.10	0.01	0.01	0.07	0.04	0.07	0.07	-0.03	2.10	0.30		
B9	0.04	0.00	0.00	0.02	0.04	0.02	0.02	-0.02	0.72	0.10		
B10	0.03	0.00	0.00	0.04	0.02	0.03	0.03	-0.02	3.76	0.50		
B11	0.07	-0.02	0.12	0.12	0.08	0.05	0.05	-0.02	0.85	0.03		
B12	0.09	0.00	0.06	0.06	0.09	0.10	0.10	0.00	6.66	0.42		
B13	0.01	0.01	0.12	0.12	0.03	0.04	0.04	-0.01	1.10	0.07		
B14	0.13	0.02	0.09	0.07	0.07	0.03	0.03	0.01	1.36	0.32		
B15	0.03	0.01	0.05	0.05	0.02	0.02	0.02	0.01	1.75	0.24		
B16	0.03	0.01	0.06	0.06	0.02	0.01	0.01	0.01	1.77	0.10		
B17	0.03	0.01	0.04	0.04	0.04	0.02	0.02	0.01	5.76	0.32		
B18	0.01	0.01	0.04	0.04	0.02	0.03	0.03	0.01	0.93	0.37		
B20	0.04	0.06	0.03	0.03	0.01	0.05	-0.03	0.00	1.56	0.39		
B21	0.01	0.01	0.08	0.02	0.02	0.03	0.03	0.00	1.64	0.04		
B22	0.00	0.03	0.03	0.02	0.03	0.02	0.02	0.03	2.72	0.29		
B23	0.03	0.06	0.03	0.03	0.03	0.03	0.01	-0.01	1.46	0.07		
B24	0.03	0.04	0.01	0.01	0.00	-0.02	-0.02	-0.02	2.03	0.21		
B25	0.04	0.03	0.03	0.08	0.03	0.03	0.01	0.02	70.40	6.80		
B26	0.01	0.02	0.04	0.04	0.02	-0.01	0.00	0.00	>100	N/D		
B27	0.03	0.00	0.06	0.06	0.01	-0.02	0.01	0.01	71.70	6.50		
B28	0.00	-0.01	0.07	0.07	0.03	-0.01	0.01	0.01	N/D	N/D		



Figura 22

biblioteca de superficie binaria		SY		SW		SR		SF	
CDR	Posición	tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido	tipo aminoácido	relación molar aminoácido
L3	91	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	92	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	93	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	94	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	96	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
H1	28	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	30	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	31	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	32	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	33	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
H2	50	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	52	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	53	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	54	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	56	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
H3	58	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	95								
	96								
	97	S,Y	50, 50	S,W	50, 50	S,R	50, 50	S,F	50, 50
	98								
99									
100									
100x	G, A	50, 50	G, A	50, 50	G, A	50, 50	G, A	50, 50	
100x	I, M, L, F	25, 25, 25, 25	I, M, L, F	25, 25, 25, 25	I, M, L, F	25, 25, 25, 25	I, M, L, F	25, 25, 25, 25	



Figura 23: Oligonucleótidos mutagénicos utilizados en la construcción de bibliotecas SF, SR, SW. Las degeneraciones del ADN equimolar se representan en el código IUB (M = A/C, S = G/C, Y = C/T).

**H1-SF**

GCA GCT TCT GGC TTC TYC ATT TYC TYC TYC TYC ATA CAC TGG GTG CGT (SEQ ID NO: 1005)

**H2-SF**

CTG GAA TGG GTT GCA TYC ATT TYC CCA TYC TYC GGT TYC ACT TYC TAT GCC GAT AGC GTC (SEQ ID NO: 1006)

**L3-SF**

ACT TAT TAC TGT CAG CAA TYC TYC TYC TYC CCA TYC ACG TTC GGA CAG GGT ACC (SEQ ID NO: 1007)

**H1-SR**

GCA GCT TCT GGC TTC MGC ATT MGC MGC MGC MGC ATA CAC TGG GTG CGT (SEQ ID NO: 1008)

**H2-SR**

CTG GAA TGG GTT GCA MGC ATT MGC CCA MGC MGC GGT MGC ACT MGC TAT GCC GAT AGC GTC (SEQ ID NO: 1009)

**L3-SR**

ACT TAT TAC TGT CAG CAA MGC MGC MGC MGC CCA MGC ACG TTC GGA CAG GGT ACC (SEQ ID NO: 1010)

**H1-SW**

GCA GCT TCT GGC TTC TSG ATT TSG TSG TSG TSG ATA CAC TGG GTG CGT (SEQ ID NO: 1011)

**H2-SW**

CTG GAA TGG GTT GCA TSG ATT TSG CCA TSG TSG GGT TSG ACT TSG TAT GCC GAT AGC GTC (SEQ ID NO: 1012)

**L3-SW**

ACT TAT TAC TGT CAG CAA TSG TSG TSG TSG CCA TSG ACG TTC GGA CAG GGT ACC (SEQ ID NO: 1013)

Figura 24A

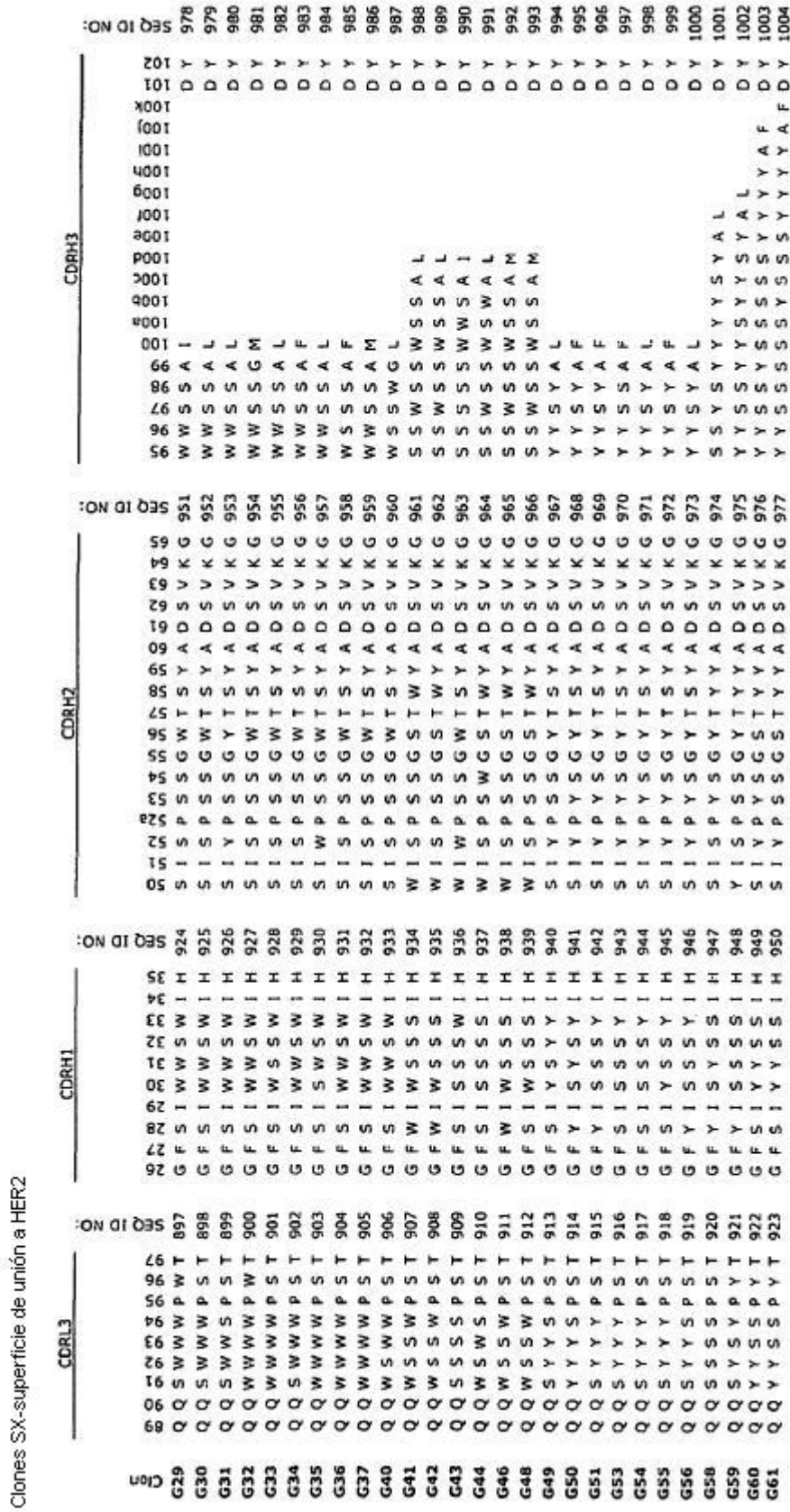


Figura 24B

		señal ELISA a 450 nm										Fab-fago IC50		
		albúmina de												
		Humano	Humano	Humano	Humano	Humano	Humano	Humano	Humano	Humano	Humano	Humano	IC50 (nM)	error +/-
		HER2	DR5	Insulin	IGF-1	IGF-1	IGF-1	IGF-1	IGF-1	IGF-1	IGF-1	IGF-1		
G29	0.33	0.06	0.26	0.19	0.17	0.29	0.02	0.32	0.02	1.96	0.22			
G30	0.13	0.04	0.19	0.17	0.29	0.01	0.26	0.01	1.26	0.21				
G31	0.09	0.07	0.22	0.12	0.26	-0.01	0.26	-0.01	1.66	0.07				
G32	0.35	0.13	0.48	0.33	0.52	0.08	0.52	0.08	1.12	0.14				
G33	0.09	0.11	0.10	0.06	0.15	0.00	0.15	0.00	1.70	0.10				
G34	0.16	0.13	0.18	0.18	0.35	0.01	0.35	0.01	1.06	0.27				
G35	0.31	0.12	0.37	0.29	0.25	0.03	0.25	0.03	0.65	0.14				
G36	0.38	0.14	0.29	0.26	0.30	0.01	0.30	0.01	1.20	0.19				
G37	0.23	0.21	0.45	0.32	0.32	0.03	0.32	0.03	0.31	0.09				
G40	0.31	0.10	0.33	0.45	0.65	0.03	0.65	0.03	1.80	0.30				
G41	0.04	0.04	0.08	0.08	0.07	0.01	0.07	0.01	51.00	6.00				
G42	0.02	0.17	0.14	0.06	0.07	0.03	0.07	0.03	43.00	7.00				
G43	0.13	0.06	0.14	0.29	0.37	0.03	0.37	0.03	22.00	3.00				
G44	0.04	0.06	0.11	0.13	0.29	-0.02	0.29	-0.02	29.00	3.00				
G46	0.02	0.03	0.07	0.07	0.10	0.00	0.10	0.00	73.00	11.00				
G48	0.09	0.04	0.04	0.04	0.01	0.01	0.01	0.01	67.00	11.00				
G49	0.02	0.04	0.04	0.00	0.05	0.01	0.05	0.01	3.13	0.37				
G50	0.12	0.11	0.05	0.10	0.10	0.02	0.10	0.02	7.94	0.42				
G51	0.05	0.10	0.03	0.04	0.08	-0.01	0.08	-0.01	3.28	0.24				
G53	0.00	0.02	0.02	0.08	0.08	-0.03	0.08	-0.03	9.75	1.76				
G54	0.05	0.06	0.05	0.03	0.05	-0.02	0.05	-0.02	2.44	0.73				
G55	0.05	0.06	0.04	0.06	0.05	-0.03	0.05	-0.03	6.52	0.73				
G56	0.01	0.08	0.11	0.05	0.07	-0.01	0.07	-0.01	7.13	0.88				
G58	0.05	0.08	0.08	0.06	0.36	0.00	0.36	0.00	15.10	0.82				
G59	0.02	0.05	0.02	0.05	0.04	0.00	0.04	0.00	6.80	0.99				
G60	0.07	0.08	0.07	0.03	0.09	-0.01	0.09	-0.01	79.80	17.20				
G61	0.09	0.09	0.07	0.12	0.09	-0.02	0.09	-0.02	145.00	42.00				

Figura 25

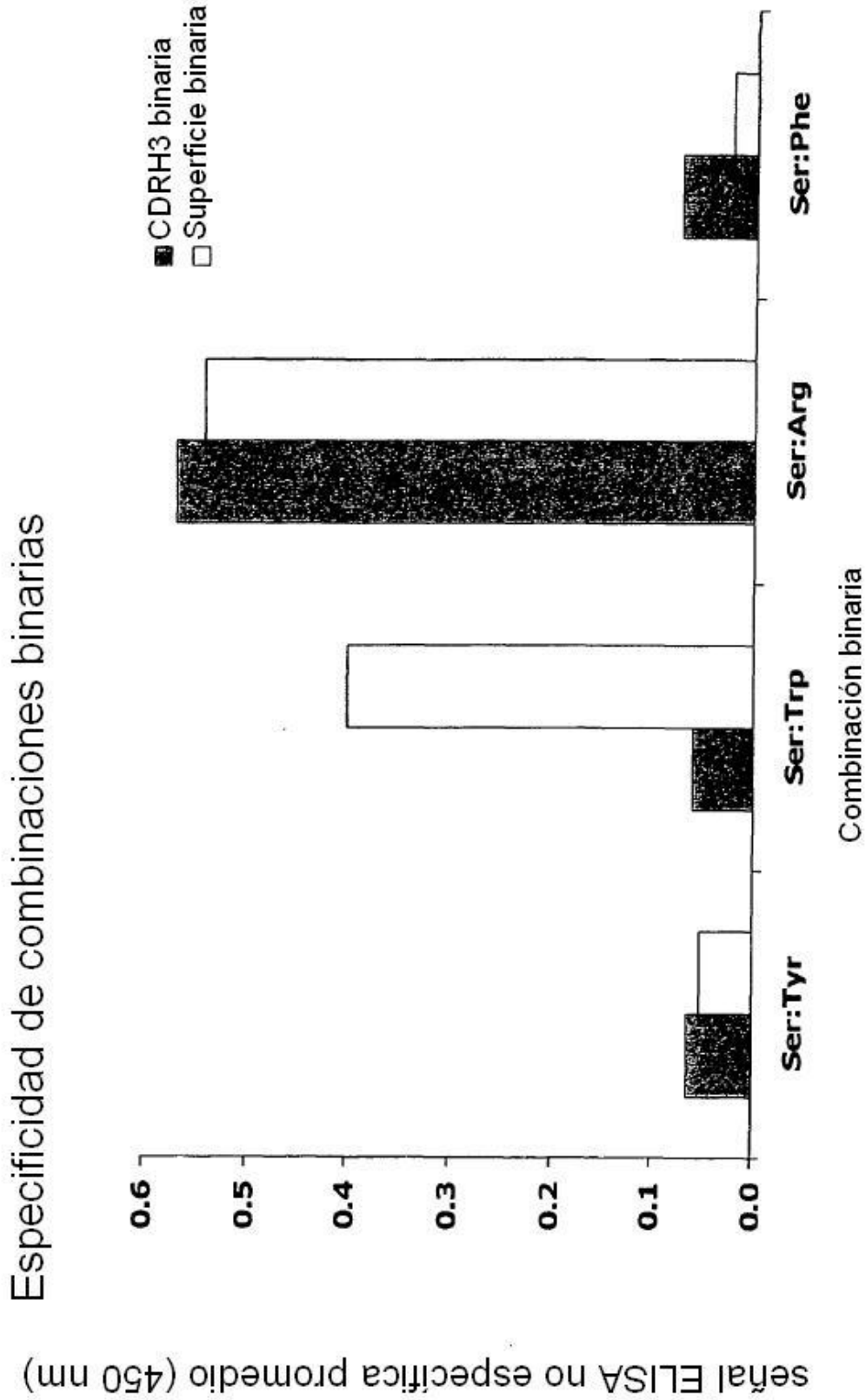
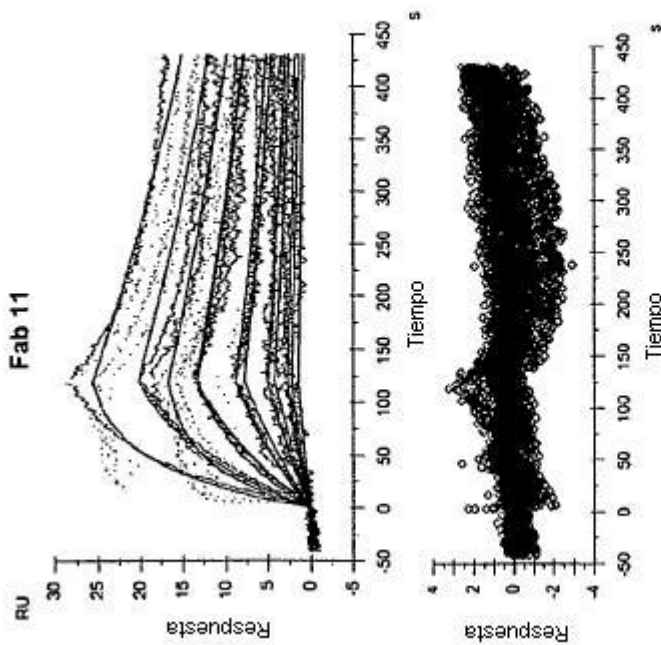


Figura 26A

**SXH3-B11**



$$k_a = 1.9 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$$

$$k_d = 1.7 \times 10^{-3} \text{ s}^{-1}$$

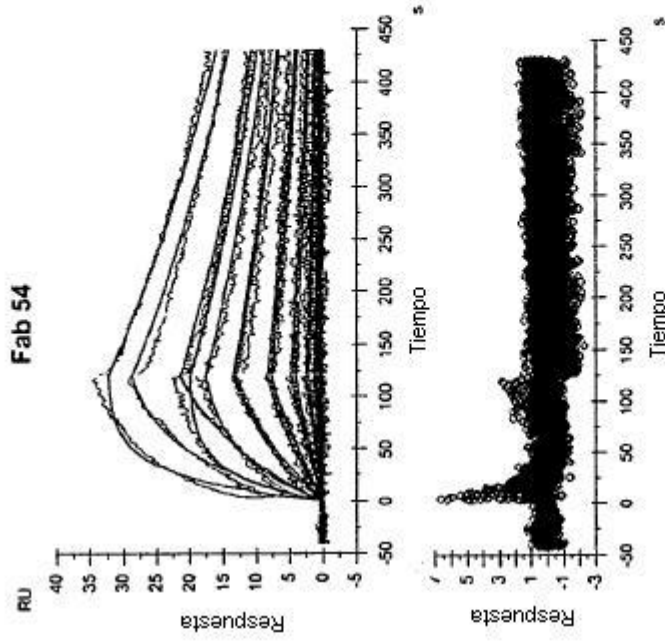
$$K_D = 890 \text{ pM}$$

$$R_{\text{max},1} = 19 \text{ RU}$$

$$R_{\text{max},2} = 29 \text{ RU}$$

\*Concentraciones de analito más elevadas eliminadas debido a la regeneración incompleta

**SX- superficie -G54**



$$k_a = 2.0 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$$

$$k_d = 2.2 \times 10^{-3} \text{ s}^{-1}$$

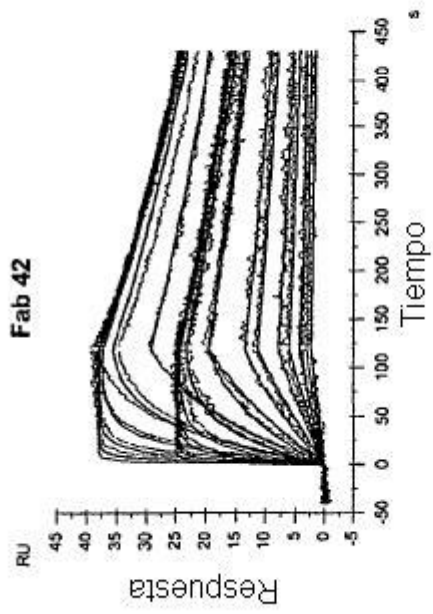
$$K_D = 11 \text{ nM}$$

$$R_{\text{max},1} = 21 \text{ RU}$$

$$R_{\text{max},2} = 34 \text{ RU}$$

Figura 26B

# YSGR-A-42



$$k_a = 2.7 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$$

$$k_d = 1.5 \times 10^{-3} \text{ s}^{-1}$$

$$K_D = 570 \text{ pM}$$

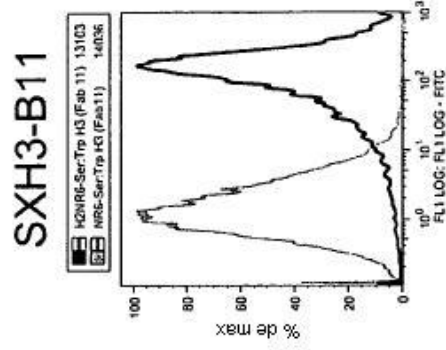
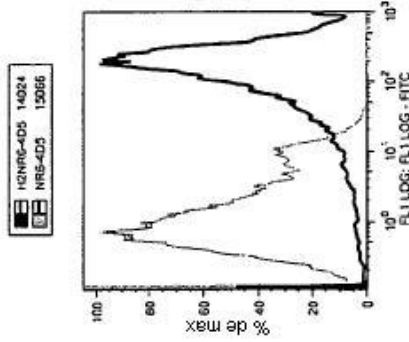
$$R_{\text{max},1} = 25 \text{ RU}$$

$$R_{\text{max},2} = 38 \text{ RU}$$



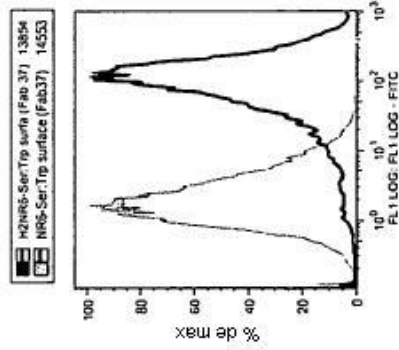
FACs con Fabs anti-HER2 purificados

control positivo Fab  
4D5

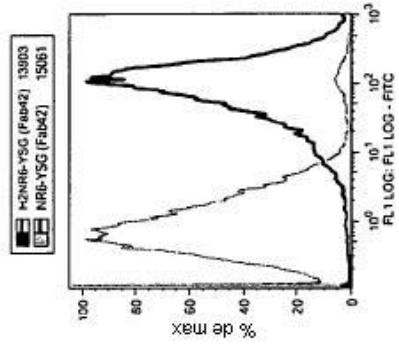


**FIG. 27**

SX-superficie-G37



YSGR-A-42



SX-superficie-G54

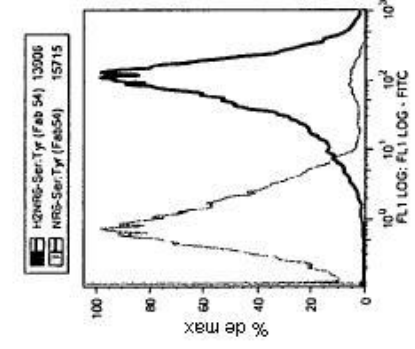


Figura 28

IgC de unión a HER2

número clon	amplificador de unión Fab-fago										IC50 (nM)	error +/-	
	CDRL3	CDRH1	CDRH2	CDRH3	CDRL1	CDRL2	CDRL3	CDRH4	CDRH5	CDRH6			
B11	89 00 91 92 93 94 95 96 97	26 27 28 29 30 31 32 33 34 35	50 51 52 53a 53b 54 55 56 57 58 59	60 61 62 63 64 65	95 96 97 98 99	100	1000	1000	1000	1000	1000	1000	1000
G37	O O S Y Y P S T	G F S I W S W I H	S I S P Y S G Y T S Y A D S V K G	60 61 62 63 64 65	95 96 97 98 99	100	1000	1000	1000	1000	1000	1000	1000
G34	O O S W W P S T	G F S I W S W I H	S I S P Y S G Y T S Y A D S V K G	60 61 62 63 64 65	95 96 97 98 99	100	1000	1000	1000	1000	1000	1000	1000
Y5GR-A-42	O O S Y Y P S T	G F S I W S W I H	S I S P Y S G Y T S Y A D S V K G	60 61 62 63 64 65	95 96 97 98 99	100	1000	1000	1000	1000	1000	1000	1000
Y5GR-A-27	O O S Y Y P V T	G F S I W S W I H	S I S P Y S G Y T S Y A D S V K G	60 61 62 63 64 65	95 96 97 98 99	100	1000	1000	1000	1000	1000	1000	1000
B37	O O S Y Y P V T	G F S I W S W I H	S I S P Y S G Y T S Y A D S V K G	60 61 62 63 64 65	95 96 97 98 99	100	1000	1000	1000	1000	1000	1000	1000
G43	O O S Y Y P S T	G F S I W S W I H	S I S P Y S G Y T S Y A D S V K G	60 61 62 63 64 65	95 96 97 98 99	100	1000	1000	1000	1000	1000	1000	1000
Y5GR-D-104	O O S S S S P S T	G F S I W S W I H	S I S P Y S G Y T S Y A D S V K G	60 61 62 63 64 65	95 96 97 98 99	100	1000	1000	1000	1000	1000	1000	1000

Figura 29

señal ELISA a 450 nm

**Coated IgG**

	<b>SXH3-B11</b>	<b>SX<sup>-superficie</sup> G37</b>	<b>YSGR-A-42</b>	<b>SX<sup>-superficie</sup> G54</b>	<b>YSGR-A-27</b>	<b>SXH3-B27</b>	<b>SX<sup>-superficie</sup> G43</b>	<b>YSGR-D-104</b>	<b>Omnitarg</b>	<b>Herceptin</b>
<b>SXH3-B11</b>	0.10	0.06	0.07	0.06	1.54	0.24	1.22	0.96	1.60	1.55
<b>SX<sup>-superficie</sup> G37</b>	0.11	0.06	0.08	0.06	1.73	0.49	1.58	1.22	1.72	1.68
<b>YSGR-A-42</b>	0.09	0.06	0.06	0.05	1.87	1.18	1.57	1.40	1.53	1.45
<b>SX<sup>-superficie</sup> G54</b>	0.32	0.16	0.30	0.13	1.55	0.27	1.37	1.23	1.48	1.15
<b>YSGR-A-27</b>	2.10	1.17	1.90	2.02	0.06	0.08	0.05	0.16	1.28	1.13
<b>SXH3-B27</b>	1.97	1.12	1.95	1.70	0.11	0.06	0.05	0.16	1.33	1.16
<b>SX<sup>-superficie</sup> G43</b>	2.02	1.20	1.93	1.76	0.12	0.07	0.04	0.20	1.30	1.14
<b>YSGR-D-104</b>	2.25	1.32	2.28	1.83	0.10	0.07	0.06	0.14	1.78	1.02

control positivo para bloqueo de unión

en negra

competición de unión

IgG incubada a 100 nM