



①9



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

①1 Número de publicación: **2 360 479**

⑤1 Int. Cl.:
C40B 40/02 (2006.01)
C12N 15/10 (2006.01)

⑫

TRADUCCIÓN DE PATENTE EUROPEA

T3

⑨6 Número de solicitud europea: **01998098 .6**
⑨6 Fecha de presentación : **18.12.2001**
⑨7 Número de publicación de la solicitud: **1360288**
⑨7 Fecha de publicación de la solicitud: **12.11.2003**

⑤4 Título: **Bibliotecas focalizadas de paquetes genéticos.**

③0 Prioridad: **18.12.2000 US 256380 P**

④5 Fecha de publicación de la mención BOPI:
06.06.2011

④5 Fecha de la publicación del folleto de la patente:
06.06.2011

⑦3 Titular/es: **DYAX Corp.**
300 Technology Square
Cambridge, Massachusetts 02139, US

⑦2 Inventor/es: **Ladner, Robert, Charles**

⑦4 Agente: **Curell Aguilá, Marcelino**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Bibliotecas focalizadas de paquetes genéticos.

- 5 Esta solicitud reivindica los derechos según la § 120 de la USC 35 de la solicitud provisional 60/256.380, presentada el 18 de diciembre de 2001.

La presente invención se refiere a bibliotecas focalizadas de paquetes genéticos que presenta, presenta y expresa o comprende un miembro de una familia diversa de péptidos, polipéptidos o proteínas y en conjunto presentan, presentan y expresan o comprenden por lo menos una parte de una diversidad concentrada de la familia. La diversidad concentrada de las bibliotecas de la presente invención comprende tanto la diversidad de secuencias como la diversidad de longitud. En una forma de realización preferida, la diversidad concentrada de las bibliotecas de la presente invención tiende hacia la diversidad natural de la familia seleccionada. En una forma de realización más preferida, las bibliotecas tienden hacia la diversidad natural de los anticuerpos humanos y se caracterizan por la variegación en sus regiones determinantes de complementariedad ("RDC") de la cadena pesada y la cadena ligera.

Antecedentes de la invención

Actualmente, es práctica corriente en la técnica preparar bibliotecas de paquetes genéticos que presentan individualmente, presentan y expresan o comprenden un miembro de una familia diversa de péptidos, polipéptidos o proteínas y en conjunto presentan, presentan y expresan o comprenden por lo menos una parte de la diversidad de aminoácidos de la familia. En muchas bibliotecas corrientes, los péptidos, polipéptidos o proteínas están relacionados con anticuerpos (por ejemplo, Fv monocatenario (scFv), Fv, Fab, anticuerpos o minicuerpos completos (es decir, dímeros constituidos por V_H unido a V_L)). A menudo, comprenden una o más de las RDC y regiones marco de las cadenas pesada y ligera de los anticuerpos humanos.

Las bibliotecas de péptidos, polipéptidos o proteínas se han producido de muchas maneras en la técnica anterior. Véase por ejemplo, Knappik *et al.*, *J. Mol. Biol.*, 296, págs. 57-86, (2000), que se incorpora a la presente memoria como referencia. Un procedimiento consiste en capturar la diversidad de donantes naturales, ya sean naturales o inmunizados. Otra manera consiste en generar bibliotecas con diversidad sintética. Un tercer procedimiento consiste en una combinación de los dos primeros. Por lo general, la diversidad producida por estos procedimientos se limita a la diversidad de secuencias, es decir, cada miembro de la biblioteca se diferencia de los otros miembros de la familia porque presentan diferentes aminoácidos o variegación en una posición dada en la cadena peptídica, polipeptídica o proteica. Naturalmente diversos péptidos, polipéptidos o proteínas, sin embargo, no se limitan a la diversidad solamente en sus secuencias de aminoácidos. Por ejemplo, los anticuerpos humanos no se limitan a la diversidad de secuencias en sus aminoácidos, también son variados en las longitudes de sus cadenas de aminoácidos.

Para los anticuerpos, la diversidad de longitud se produce, por ejemplo, durante las reorganizaciones de la región variable. Véase, por ejemplo, Corbett *et al.*, *J. Mol. Biol.*, 270, págs. 587-97, (1997). La unión de los genes V con los genes J, produce la inclusión de un segmento D reconocible en RDC3 en aproximadamente la mitad de las secuencias de anticuerpos con cadena pesada, creando de este modo regiones que codifican longitudes variables de aminoácidos. Lo siguiente puede producirse también durante la articulación de los segmentos génicos del anticuerpo: (i) el final del gen V puede presentar suprimidas o cambiadas de cero a varias bases; (ii) el final del segmento D puede presentar de cero a muchas bases eliminadas o cambiadas; (iii) numerosas bases al azar pueden insertarse entre V y D o entre D y J; y (iv) el extremo 5' de J puede evitarse para eliminar o cambiar varias bases. Estas reorganizaciones dan como resultado antibióticos variados tanto en la secuencia de aminoácidos como en longitud.

Las bibliotecas que contienen solamente diversidad de secuencia de aminoácidos están, por tanto, en desventaja porque no reflejan la diversidad natural de péptidos, polipéptidos o proteínas que la biblioteca pretende simular. Además, la diversidad en longitud puede ser importante para el funcionamiento en última instancia de la proteína, péptido o polipéptido. Por ejemplo, con respecto a una biblioteca que comprende regiones de anticuerpo, muchos de los péptidos, polipéptidos, proteínas presentados, presentados y expresados o compuestos de paquetes genéticos de la biblioteca no pueden doblarse de manera apropiada o su fijación a un antígeno puede resultar desventajosa, si la diversidad tanto en la secuencia como en la longitud no están representadas en la biblioteca.

Otro inconveniente de las bibliotecas de la técnica anterior de paquetes genéticos que presentan, presentan y expresan, o comprenden péptidos, polipéptidos y proteínas es que no están focalizados en los miembros que están basados en la diversidad natural y por tanto en los miembros que es más probable que sean funcionales. Más bien, las bibliotecas de la técnica anterior, por lo general, intentan incluir tanta diversidad o variegación en cada resto de aminoácido como sea posible. Esto hace que la construcción de la biblioteca emplee mucho tiempo y sea menos eficiente de lo posible. El número de miembros que se producen tratando de capturar la diversidad completa también hace la identificación más problemática que lo que sería necesario. Esto es particularmente cierto dado que muchos miembros de la biblioteca no serán funcionales.

Sumario de la invención

Un objeto de la presente invención son las bibliotecas de vectores o paquetes genéticos focalizados que codifican miembros de una familia diversa de péptidos, polipéptidos o proteínas en el que las bibliotecas codifican poblaciones que son variadas tanto en longitud como en secuencias. La diversa longitud que comprende componentes que contienen motivos que probablemente son para plegar y funcionar en el contexto del péptido, polipéptido o proteína original.

Otro objeto de la presente invención son las bibliotecas focalizadas de paquetes genéticos que presentan, presentan y expresan o comprenden un miembro de una familia de péptidos, polipéptidos o proteínas y en conjunto presentan, presentan y expresan o comprenden por lo menos una parte de la diversidad focalizada de la familia. Estas bibliotecas son variadas no solamente en sus secuencias de aminoácidos, sino también en sus longitudes. Y, su diversidad está focalizada para que simulen más estrechamente o consideren la diversidad natural de la familia específica que representa la biblioteca.

Otro objeto de la presente invención son las poblaciones variadas, pero focalizadas, secuencias de ADN que codifican péptidos, polipéptidos o proteínas adecuados para presentar o presentar y expresar utilizando paquetes genéticos (tal como fagos o fagómidos) u otros regímenes que permiten la selección de los componentes de la fijación específica de una biblioteca.

Un objeto adicional de la presente invención son las bibliotecas focalizadas que comprenden las RDC de anticuerpos humanos que presentan una diversidad en sus secuencias de aminoácidos y en sus longitudes (ejemplos de dichas bibliotecas incluyen bibliotecas de Fv monocatenario (scFv), Fv, Fab, anticuerpos completos o minicuerpos (es decir, dímeros que constan de V_H unida a V_L)). Dichas regiones pueden ser de cadena pesada o ligera o ambas o pueden incluir una o más de las RDC de estas cadenas. Más preferentemente, la diversidad de la variación se produce en todas las RDC de cadena pesada y cadena ligera.

La presente invención se refiere a:

Una biblioteca de ADN focalizada que comprende secuencias de ADN variegadas que codifican:

1) por lo menos una RDC1 de cadena pesada seleccionada de entre el grupo constituido por:

(a) $\langle 1 \rangle_1 Y_2 \langle 1 \rangle_3 M_4 \langle 1 \rangle_5$, en el que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y;

(b) $(S/T)_1 (S/G/X)_2 (S/G/X)_3 Y_4 Y_5 W_6 (S/G/X)_7$.

en la que (S/T) es una mezcla 1:1 de restos S y T, (S/G/X) es una mezcla de 0,2025 S, 0,2025 G y 0,035 de cada uno de los restos de aminoácidos A, D, E, F, H, I, K, L, M, N, P, Q, R, T, V, W e Y;

(c) $V_1 S_2 G_3 G_4 S_5 I_6 S_7 \langle 1 \rangle_8 \langle 1 \rangle_9 \langle 1 \rangle_{10} Y_{11} Y_{12} W_{13} \langle 1 \rangle_{14}$, en los que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y;

(d) mezclas de vectores o de paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores;

siendo los miembros de este grupo las únicas RDC1 de cadena pesada presentes en la biblioteca;

(2) por lo menos una de entre una RDC2 de cadena pesada seleccionada de entre el grupo constituido por:

(a) $\langle 2 \rangle I \langle 2 \rangle \langle 3 \rangle SGG \langle 1 \rangle T \langle 1 \rangle YADSVKG$, en la que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; $\langle 2 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos Y, R, W, V, G y S; y $\langle 3 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos P, S y G o una mezcla equimolar de P y S;

(b) $\langle 1 \rangle I \langle 4 \rangle \langle 1 \rangle \langle 1 \rangle \langle G \rangle \langle 5 \rangle \langle 1 \rangle \langle 1 \rangle YADSVKG$, en la que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; $\langle 4 \rangle$ es una mezcla equimolar de los restos de D, I, N, S, W, Y; y $\langle 5 \rangle$ es una mezcla equimolar de los restos de S, G, D y N;

(c) $\langle 1 \rangle I \langle 4 \rangle \langle 1 \rangle \langle 1 \rangle \langle G \rangle \langle 5 \rangle \langle 1 \rangle \langle 1 \rangle YNPSLKG$, en la que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y $\langle 4 \rangle$ y $\langle 5 \rangle$ son como se definieron anteriormente;

(d) $\langle 1 \rangle I \langle 8 \rangle S \langle 1 \rangle \langle 1 \rangle \langle 1 \rangle GGY \langle 1 \rangle YAASVKG$, en la que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y $\langle 8 \rangle$ es 0,27 R y 0,027 de cada uno

de ADEFGHIKLMNPQSTVWY; y

- (e) mezclas de vectores o de paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores;

siendo los miembros de este grupo las únicas RDC2 de cadena pesada presentes en la biblioteca; y

(3) por lo menos una de las RDC3 de cadena pesada seleccionada de entre el grupo constituido por:

- (a) YYCA21111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de K y R;

- (b) YYCA211111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de K y R;

- (c) YYCA2111111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de K y R;

- (d) YYCAR111S2S3111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de S y G; y 3 es una mezcla equimolar de Y y W;

- (e) YYCA2111CSG11CY1YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de K y R;

- (f) YYCA211S1TIFG1111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de K y R;

- (g) YYCAR111YY2S334411YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; 2 es una mezcla equimolar de D y G; y 3 es una mezcla equimolar de S y G;

- (h) YYCAR1111YC2231CY111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; 2 es una mezcla equimolar de S y G; y 3 es una mezcla equimolar de T, D y G; y

- (i) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores;

siendo los miembros de este grupo las únicas RDC3 de cadena pesada presentes en la biblioteca.

Preferentemente, las RDC1 de cadena pesada (a), (b) y (c) están presentes en la biblioteca en la relación 0,80:0,17:0,02.

Preferentemente, una mezcla de las RDC2 (a)/(b) (equimolar) de cadena pesada (c) y (d) están presentes en la biblioteca en una relación de 0,54:0,43:0,03.

Preferentemente, 1 en una o todas las RDC3 de cadena pesada (a) hasta (h) es 0,095 de cada una de G e Y y 0,048 de cada una de A, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V y W.

En una forma de realización preferida, las RDC3 de cadena pesada (a) hasta (h) están presentes en la biblioteca en las proporciones siguientes:

- (a) 0,10
- (b) 0,14
- (c) 0,25
- (d) 0,13
- (e) 0,13
- (f) 0,11
- (g) 0,04 y
- (h) 0,10.

Resulta asimismo preferido que las RDC3 de cadena pesada (a) hasta (h) estén presentes en la biblioteca en las siguientes proporciones:

- (a) 0,02
- (b) 0,14

- (c) 0,25
- (d) 0,14
- (e) 0,14
- (f) 0,12
- (g) 0,08 y
- (h) 0,11.

En una forma de realización preferida, la biblioteca de ADN focalizada comprende además, secuencias de ADN variegadas que codifican por lo menos uno de entre:

(1) por lo menos una RDC1 de cadena pesada kappa seleccionada de entre el grupo constituido por:

(a) RASQ<1>V<2><2><3>LA

(b) RASQ<1>V<2><2><2><3>LA;

en la que <1> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVWY; <2> es 0,2 S y 0,044 de cada ADEFGHIKLMNPQRTVWY; y <3> es 0,2 Y y 0,044 cada ADEFGHIKLMNPQRTVW e Y; y

(c) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores;

siendo los miembros de este grupo las únicas RDC1 de cadena ligera kappa presentes en la biblioteca;

(2) una RDC2 de cadena ligera kappa con la secuencia:

<1>AS<2>R<4><1>

en la que <1> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVWY; <2> es 0,2 S y 0,044 de cada ADEFGHIKLMNPQRTVWY; y <4> es 0,2 A y 0,044 cada ADEFGHIKLMNPQRTVWY;

siendo los miembros de este grupo las únicas RDC1 de cadena ligera kappa presentes en la biblioteca; o

(3) por lo menos una RDC3 de cadena ligera kappa seleccionada de entre el grupo constituido por:

(a) QQ<3><1><1><1>P<1>T,

en la que <1> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVWY; <3> es 0,2 Y y 0,044 de cada ADEFGHIKLMNPQRTVW;

(b) QQ33111P, en la que 1 y 3 son como se definieron en (1) anteriormente;

(c) QQ3211PP1T, en la que 1 y 3 son como se definieron en (1) anteriormente y 2 es 0,2 S y 0,044 cada ADEFGHIKLMNPQRTVWY; y

(d) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores;

siendo los miembros de este grupo las únicas RDC3 de cadena ligera kappa presentes en la biblioteca.

En esta forma de realización resulta preferido que la biblioteca de ADN focalizada comprenda secuencias de ADN variegadas que codifican por lo menos un miembro de por lo menos dos de los grupos (1), (2) y (3) de la RDC de cadena ligera kappa.

En esta forma de realización resulta además preferido que la biblioteca focalizada comprenda las secuencias de ADN variegadas que codifican por lo menos un miembro de la totalidad de los grupos de RDC de cadena ligera kappa (1), (2) y (3).

Preferentemente, las RDC1 (a) y (b) de cadena ligera kappa están presentes en la biblioteca en una proporción de 0,68:0,32.

Resulta asimismo preferido que las RDC3 (a), (b) y (c) de cadena ligera kappa están presentes en la biblioteca en una proporción de 0,65:0,1:0,25.

En una forma de realización más preferida la biblioteca de ADN focalizada comprende además secuencias de ADN variegadas que codifican por lo menos uno de entre:

(1) por lo menos una RDC1 de cadena ligera lambda seleccionada de entre el grupo constituido por:

(a) TG<1>SS<2>VG<1><3><2><3>VS en la que <1> es 0,27 T, 0,27 G y 0,027 cada ADEFHIKLMNPQRSVWY, <2> es 0,27 D, 0,27 N y 0,27 cada AEFHIKLMNPQRSTVWY y <3> es 0,36 Y y 0,036 cada ADEFGHIKLMNPQRSTVW;

(b) G<2><4>L<4><4><3><4><4>en la que <2> es tal como se definió en (1) anteriormente y <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVWY; y

(c) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores; siendo los miembros de este grupo las únicas RDC1 de cadena ligera lambda solamente presentes en la biblioteca;

(2) una RDC2 de cadena ligera lambda que tiene la secuencia:

<4><4><4><2>RPS

en la que <2> es 0,27 D, 0,27 N y 0,027 cada AEFHIKLMNPQRSTVWY y <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVW;

siendo los miembros de este grupo las únicas RDC2 de cadena ligera lambda presentes en la biblioteca;

(3) por lo menos una RDC3 de cadena ligera lambda seleccionada de entre el grupo constituido por:

(a) <4><5><4><2><4>S<4><4><4>V, en la que <2> es 0,27 D, 0,27 N y 0,027 cada AEFHIKLMNPQRSTVWY; <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVW; y <5> es 0,36 S y 0,0355 cada ADEFGHIKLMNPQRSTVWY;

(b) <5>SY<1><5>S<5><1><4>V en la que <1> es una mezcla equimolar de ADEFGHIKLMNPQRSTVWY; y <4> y <5> son como se definieron en (1) anteriormente; y

(c) mezclas de vectores o de paquetes genéticos caracterizados por algunas de las secuencias de ADN anteriores;

siendo los miembros de este grupo las únicas RDC3 de cadena ligera lambda presentes en la biblioteca.

En esta forma de realización es preferible que la biblioteca de ADN focalizada, comprenda secuencias de ADN variegadas que codifiquen por lo menos un miembro de por lo menos dos grupos (1), (2) y (3) de RDC de cadena ligera lambda.

Preferentemente, la biblioteca de ADN focalizada de esta forma de realización comprende secuencias de ADN variegadas que codifican por lo menos un miembro de cada uno de los grupos (1), (2) y (3) de RDC de cadena ligera lambda.

Resulta más preferido en esta forma de realización que las RDC1 (a) y (b) de cadena ligera lambda estén presentes en la biblioteca en una proporción de 0,67:0,33.

Resulta además preferido, en esta forma de realización que las RDC3 (a) y (b) de cadena ligera lambda estén presentes en la biblioteca en una mezcla equimolar.

Descripción detallada de las formas de realización preferidas

Los anticuerpos ("Ab") concentran su diversidad en las regiones que están implicadas en la determinación de la afinidad y especificidad del Ab para dianas específicas. Estas regiones pueden ser de secuencia o longitud variada. Generalmente, son variadas en ambos aspectos. Sin embargo, dentro de las familias de anticuerpos humanos, las diversidades, tanto de secuencia como de longitud, no son verdaderamente aleatorias. Más bien, algunos restos de aminoácidos se prefieren en determinadas posiciones de las RDC y se prefieren algunas longitudes de RDC. Estas diversidades preferidas explican la diversidad natural de la familia de anticuerpos.

Según la presente invención, y como se describe con mayor detalle a continuación, son preparadas y utilizadas las bibliotecas de vectores y paquetes genéticos que reflejan más detenidamente la diversidad natural, tanto en secuencia como en longitud, de familias de anticuerpos, o las porciones de los mismos.

Diversidad de secuencias y longitudes de la cadena pesada de anticuerpos humanos(a) Estructura

5 El gen de la línea germinal (GLC) 3-23 de la cadena pesada ("HC") (conocido también como VP-47) explica aproximadamente 12% de todos los Ab humanos y resulta preferido como marco en la forma de realización preferida de la invención. Debe apreciarse, sin embargo, que otros marcos bien conocidos, tales como 4-34, 3-30, 3-30.3 y 4-30.1 pueden utilizarse también sin apartarse de los principios de las diversidades focalizadas de la presente invención.

10 Además, JH4 (YFDYWGGQGLTVTUS) se produce más a menudo que JH3 en los anticuerpos naturales. Por consiguiente, es preferible para las bibliotecas focalizadas de la presente invención. Sin embargo, podría utilizarse también JH3 (AFDIWGGQGTMTVSS).

15 (b) Diversidad de la longitud focalizada: RDC1, 2 y 3

(i) RDC1

20 Para RDC1, las GLG proporcionan las RDC1 solamente de las longitudes 5, 6 y 7. Las mutaciones durante la maduración del gen con dominio V, sin embargo, pueden conducir a las RDC con longitudes de 2 como mínimo y de 16 como máximo. No obstante, predomina la longitud 5. Por consiguiente, en la forma de realización preferida de la presente invención, la RDC1 de HC preferida es de 5 aminoácidos, con las RDC1 menos preferidas que tienen longitudes de 7 y 14. En las bibliotecas más preferidas de la presente invención, las tres longitudes se utilizan en proporciones similares a las encontradas en los anticuerpos naturales.

25 (ii) RDC2

30 Las GLG proporcionan RDC2 solamente de las longitudes 15-19 pero las mutaciones durante la maduración pueden dar como resultados las RDC2 de longitudes entre 16 y 28 aminoácidos. Las longitudes 16 y 17 predominan en los genes Ab maduros. Por consiguiente, la longitud 17 es la longitud preferida para RDC2 de HC de la presente invención. Las RDC2 de HC menos preferidas de la presente invención tienen longitudes entre 16 y 19. En las bibliotecas focalizadas más preferidas de la presente invención, las tres longitudes están incluidas en proporciones similares a las encontradas en las familias de anticuerpos naturales.

35 (iii) RDC3

40 Las RDC3 de HC varían en longitud. Aproximadamente la mitad de las HC humanas constan de los componentes siguientes: V::nz::D::ny::JHn en la que V es un gen V, nz es una serie de bases (media 12) que son esencialmente aleatorias, D es un segmento D, con frecuencia con dirección en ambos extremos, ny es una serie de bases (media 6) que son esencialmente aleatorias y JH es uno de los seis segmentos JH con frecuencia con edición pesada en el extremo 5'. Los segmentos D parecen proporcionar segmentos espaciadores que permiten el plegamiento de la IgG. La mayor diversidad está en las uniones de V con D y de D con JH.

45 En las bibliotecas preferidas de la presente invención se utilizan ambos tipos de RDC3 HC. En las RDC3 de HC que no tienen segmento D identificable, la estructura es V::nz::JHn en la que JH se edita normalmente en el extremo 5'. En las RDC3 de HC que tienen un segmento D identificable la estructura es V::nz::D::ny::JHn.

(c) Diversidad de secuencias focalizadas: RDC1, 2 y 3

50 (i) RDC1

55 En la RDC1, con longitud de 5 aminoácidos, el examen de un modelo 3D de un Ab humanizado se demostró que los grupos laterales de los restos 1, 3 y 5 se dirigieron hacia la bolsa de combinación. Por lo tanto, en las bibliotecas focalizadas de la presente invención, cada una de estas posiciones puede seleccionarse de entre cualquiera de los restos de aminoácidos naturales, excepto cisteína ("C"). La cisteína puede formar enlaces disulfuro que son un componente importante del plegamiento canónico de Ig. Teniendo grupos tiol libres podría interferir, por tanto, con el plegamiento apropiado de la HC y podría conducir a problemas en la producción o manipulación de los Ab seleccionados. Por lo tanto, en las bibliotecas focalizadas de la presente invención se excluye la cisteína de las posiciones 1, 3 y 5 de las RDC1 de 5 aminoácidos preferidos. Los otros 19 restos de aminoácidos naturales pueden utilizarse en las posiciones 1, 3 y 5. Preferentemente, cada uno está presente en relaciones equimolares en las bibliotecas variegadas de la presente invención.

65 El modelado en 3D sugiere además que los grupos laterales del resto 2 en la RDC1 de 5 aminoácidos se dirijan lejos de la bolsa de combinación. Aunque esta posición presenta sustancial diversidad tanto en GLG como en los genes maduros, en las bibliotecas focalizadas de la presente invención este residuo es preferentemente Tyr (Y) porque

produce en los genes de los anticuerpos maduros 681/820. Sin embargo, alguno de los demás restos de aminoácidos naturales, excepto Cys (C), podrían utilizarse también en esta posición.

En la posición 4, existe también alguna diversidad en GLG y en los genes de anticuerpos maduros. Sin embargo, casi todos los genes maduros tienen restos de aminoácidos hidrófobos sin carga: A, G, L, P, F, M, W, I, V en esta posición. La inspección de un modelo en 3D demuestra además que el grupo lateral del resto 4 está dispuesto en las entrañas de la HC. Por lo tanto, en la forma de realización preferida de la presente invención que utiliza el marco 3-23 el resto 4 es preferentemente mET porque es probable que se fije muy bien en el marco de 3-23. Con otras estructuras, se utiliza una consideración de la fijación similar para asignar el resto 4.

Por lo tanto, la RDC1 con HC más preferida de la presente invención consta de la secuencia de aminoácidos de $\langle 1 \rangle Y \langle 1 \rangle M \langle 1 \rangle$ donde $\langle 1 \rangle$ puede ser alguno de los restos de aminoácidos: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y (no C), preferentemente presente en cada posición en una cantidad equimolar. Esta diversidad se presenta en el contexto de una estructura 3-23:JH4 en la Tabla 1. Presenta una diversidad de 6859 veces.

Las dos RDC1 con HC menos preferidas de la presente invención presentan una longitud 7 y una longitud 14. Para la longitud 7, una variegación preferida es $(S/T)_1(S/G/\langle 1 \rangle)_2(S/G/\langle 1 \rangle)_3Y_4Y_5W_6(S/G/\langle 1 \rangle)_7$; indicando S/T la mezcla equimolar de los codones Ser y Thr; $(S/G/\langle 1 \rangle)$ indica una mezcla de 0,2025 S, 0,2025 G y 0,035 para cada A, D, E, F, H, I, K, L, M, N, P, Q, R, T, V, W, Y. En este diseño predominan Ser y Gly en las posiciones 2, 3 y 7, como ocurre en los genes maduros con HC. Para la longitud 14, una variegación preferida es VSGGSIS $\langle 1 \rangle \langle 1 \rangle \langle 1 \rangle YYW \langle 1 \rangle$, en el que $\langle 1 \rangle$ es una mezcla equimolar de los 19 restos de aminoácidos naturales, excepto Cys (C).

El ADN que codifica estos RDC1 con HC preferidos se sintetiza preferentemente utilizando bloques de construcción de trinucleótidos de modo que cada resto de aminoácido está presente en cantidades esencialmente equimolares u otras descritas. Los codones preferidos para los restos de aminoácidos de $\langle 1 \rangle$ son gct, gat, gag, ttt, ggt, cat, att, aag, ctt, atg, aat, cct, cag, cgt, tgt, act, gtt, tgg y tat. Desde luego, podrían utilizarse también otros codones para el resto de aminoácido seleccionado.

La diversidad de oligonucleótidos (ON) se sintetiza preferentemente desde *BspEI* hasta *BstXI* (como se muestra en la Tabla 1) y puede, por consiguiente, incorporarse ya sea por síntesis de RCP utilizando ON solapantes o introducido por ligadura de fragmentos de corte *BspEI/BstXI*. La Tabla 2 presenta los oligonucleótidos que incorporan las variegaciones especificadas de longitudes 5 HC RDC1 preferidas de la presente invención. La RCP que utiliza ON-R1V1vg, ON-R1top y ON-R1bot proporciona un producto de ADNds de 73 pares de bases, se escinde con las bases 11 y 13 de los de los extremos de los recortes *BspEI* y *BstXI* y proporciona extremos cohesivos que pueden ligarse igualmente al vector de corte que tiene el dominio 3-23 mostrado en la Tabla 1. La sustitución de ON-R1V1vg por ON-R1V2vg o ON-R1V3vg (véase Tabla 2) permite la síntesis de los dos modelos de diversidad alternativos, la RDC1 con HC de 7 restos de longitud y 14 restos de longitud.

Las bibliotecas más preferidas de la presente invención comprenden 3 diversidades de longitud de RDC1 con HC preferidas. Más preferentemente las 3 longitudes se incorporarían en aproximadamente las relaciones que se observan en los anticuerpos seleccionados con referencia a la longitud de las RDC. Por ejemplo, una muestra de 1095 genes para HC tienen las tres longitudes presentes en la relación: L=5:L=7:L=14::820:175:23::0,80:0,17:0,02. Esta es la relación preferida según la presente invención.

(ii) RDC2

La diversidad en RDC2 de HC se diseñó con las mismas consideraciones que para RDC1 de HC: secuencias GLG, secuencias maduras y estructura 3D. Una longitud preferida para RDC2 es 17, como se muestra en la Tabla 1. Para esta RDC2 de longitud 17 preferida, la variegación preferida según la invención es: $\langle 2 \rangle I \langle 2 \rangle \langle 3 \rangle SGG \langle 1 \rangle T \langle 1 \rangle YADSVKG$, en el que $\langle 2 \rangle$ indica cualquier resto de aminoácidos seleccionado del grupo Y, R, W, V, G y S (mezcla equimolar), $\langle 3 \rangle$ es P, S y G o P y S solamente (mezcla equimolar) y $\langle 1 \rangle$ es cualquier resto de aminoácido natural excepto C (mezcla equimolar).

El ON-R2V1vg presentado en la Tabla 3 incorpora este modelo de diversidad. Se sintetiza preferentemente de modo que los fragmentos de dsADN que contienen la secuencia *BstXI* y *XbaI* pueden generarse por RCP. RCP con ON-R2V1vg, ONR2top y ON-R2bot proporciona un producto de dsADN de 122 pares de bases. La escisión con *BstXI* y *XbaI* elimina aproximadamente 10 bases de cada extremo y produce extremos cohesivos que pueden ligarse para cortar igualmente el vector que contiene el gen 3-23 mostrado en la Tabla 1.

En una forma de realización alternativa para una longitud de 17 RDC2 de HC, puede utilizarse la variegación siguiente: $\langle 1 \rangle I \langle 4 \rangle \langle 1 \rangle \langle 1 \rangle G \langle 5 \rangle \langle 1 \rangle \langle 1 \rangle \langle 1 \rangle YADSVKG$, en la que $\langle 1 \rangle$ es como se describió anteriormente para la alternativa más preferida de RDC2 de HC; $\langle 4 \rangle$ indica una mezcla equimolar de DINSWY, y $\langle 5 \rangle$ indica una mezcla equimolar de SGD. Este modelo de diversidad está incorporado en ON-R2V2vg mostrado en la Tabla 3. En las bibliotecas de la presente invención las dos formas de realización se utilizan preferentemente en mezclas equimolares.

Otras RDC2 con HC preferidas presentan longitudes 16 y 19. Longitud 16: <1>I<4><1><1>G<5><1><1>YNPSLKG; Longitud 19: <1>I<8>S<1><1><1>GGYY<1>YAASVKG, en el que <1> es una mezcla equimolar de todos los restos de aminoácidos naturales excepto C; <4> es una mezcla equimolar de DINSWY; <5> es una mezcla equimolar de SGDNI; y <8> es 0,27 R y 0,027 de cada uno de los restos ADEFGHIKLMNPQSTVWY. La Tabla 3 presenta ON-R2V3vg que incorpora una variegación de RDC2 preferida de longitud 16 y ON-R2V4vg que incorpora una variegación de RDC2 preferida de longitud 19. Para preparar estas variegaciones ON-R2V3vg puede ampliarse por RCP con ON-R2top y ON-R2bo3 y ON-R2V4vg puede ampliarse por RCP por ON-R2top y ON-R2-bo4. Véase la Tabla 3. En la forma de realización más preferida de la presente invención, se utilizan las tres longitudes de RDC2 con HC. Preferentemente, están presentes en una proporción 17:16:19::579:464:31::0,54:0,43:0,03.

(iii) CDR3

Las bibliotecas preferidas de la presente invención comprenden varios componentes de RDC3 con HC. Algunos de éstos presentarán solamente diversidad de secuencia. Otros tendrán diversidad de secuencia con segmentos D incrustados para ampliar la longitud, mientras que incorporan además secuencias conocidas para permitir que se plieguen las Ig. Los componentes de RDC3 con HC de las bibliotecas preferidas de la presente invención y sus diversidades se representan en la Tabla 4: Componentes 1-8.

Esta serie de componentes se seleccionó después de estudiar las secuencias de 1383 secuencias con HC humanas. Los componentes propuestos significan que cumplen los objetivos siguientes:

- 1) aproximadamente la misma distribución de longitudes observada en los genes para Ab natural;
- 2) gran nivel de diversidad de secuencias en los lugares que tienen gran diversidad en los genes para Ab naturales;
- 3) incorporación de secuencias constantes observadas con frecuencia en los genes para Ab naturales.

El componente 1 representa todos los genes que tiene longitudes 0 a 8 (contando desde el motivo YYCAR al final de FR3 para el motivo WG cerca del comienzo de la región J, es decir, FR4). El componente 2 corresponde a todos los genes que tienen longitudes 9 ó 10. El componente 3 corresponde a los genes que tienen longitudes 11 ó 12 más la mitad de los genes que tienen longitud 13. El componente 4 corresponde a los que tienen longitud 14 más la mitad de los que tienen longitud 13. El componente 5 corresponde a los genes que tienen longitud 15 y la mitad de los que tienen longitud 16. El componente 6 corresponde a los genes de longitud 17 más la mitad de los de longitud 16. El componente 7 corresponde a los que tienen longitud 18. El componente 8 corresponde a los que tienen longitud 19 y mayores. Véase la Tabla 4.

Para cada resto de RDC3 con HC que presenta la diversidad <1>, las relaciones equimolares no se utilizan preferentemente. En su lugar, se utilizan las siguientes relaciones 0,095 [G e Y] y 0,048 [A, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V y W]. Por tanto, existe una doble dosis de G e Y con los demás restos que están en relaciones equimolares. Para las demás diversidades, por ejemplo, KR o SG, los restos están presentes en mezclas equimolares.

En las bibliotecas preferidas de la presente invención, los ocho componentes están presentes en las siguientes fracciones: 1 (0,10), 2 (0,14), 3 (0,25), 4 (0,13), 5 (0,13), 6 (0,11), 7 (0,04) y 8 (0,10). Véase la Tabla 4.

En la forma de realización más preferida de la presente invención, las cantidades de los ocho componentes se ajustan porque el primer componente no es lo bastante complejo para justificar incluirlo como 10% de la biblioteca. Por ejemplo, si la biblioteca final fuera la que presenta 1×10^9 miembros, entonces 1×10^8 secuencias procederían del componente 1, pero presenta solamente $2,6 \times 10^5$ secuencias RDC3 de modo que cada una produciría ~ 385 contextos de RDC1/2. Por consiguiente, las cantidades más preferidas de los ocho componentes son 1 (0,02), 2 (0,14), 3 (0,25), 4 (0,14), 5 (0,14), 6 (0,12), 7 (0,08), 8 (0,11). Según la forma de realización más preferida el componente 1 se produce en contextos ~ 77 RDC1/2 y el otro, los RDC3 mayores se producen con más frecuencia.

La Tabla 5 presenta ADNvg que incorpora cada uno de los ocho componentes de RDC3 de HC mostrados en la Tabla 4. En la Tabla 5, los oligonucleótidos (ON) Ctp25, CtpmA, CBpmB y CBot25 permiten la ampliación por RCP de cada uno de los ON variegados (vgADN): C1t08, C2t10, C3t12, C4t14, C5t15, C6t17, C7t18 y C8t19. Tras la ampliación, el dsADN puede escindirse con *AflI* y *BstEII* (o *KpnI*) y ligarse al vector asimismo escindido que contiene el resto del dominio 3-23. Preferentemente, este vector ya contiene diversidad en uno o ambos de RDC1 y RDC2 tal como se da a conocer en la presente memoria. Todavía más preferentemente contiene diversidad tanto en las regiones RDC1 como en RDC2. Desde luego, debe apreciarse que varias diversidades puedan incorporarse en el vector en cualquier orden.

Preferentemente, el vector receptor contiene originalmente un cebador en lugar de RDC1, RDC2 y RDC3 de modo que será una secuencia no original la que se produzca a continuación en la biblioteca resultante. La Tabla 6 presenta una versión del segmento génico V3-23 con cada RDC sustituida por un segmento corto que contiene tanto

codones de terminación como secuencias de restricción que permitirán la escisión específica de cualquier vector que no haya eliminado el cebador. El cebador puede ser corto y contener una secuencia de la enzima de restricción que no ocurre en la biblioteca acabada, permitiendo la eliminación de vectores que no son escindidos tanto por *Afl* como por *BstEII* (o *KpnI*) y se vuelven a ligar. Alternativamente, el cebador podría tener una longitud de 200 a 400 bases de modo que el vector no escindido o escindido una vez puede separarse fácilmente del vector escindido dos veces.

Cadena ligera de anticuerpo humano: diversidad de secuencia y longitud

(i) Cadena kappa

(a) Estructura

En la forma de realización preferida de la presente invención, la cadena ligera kappa se construye en una estructura A27 con una región JK1. Estas son las regiones más comunes de V y J en los genes naturales. Pueden utilizarse otras estructuras, tales como O12, L2 y A11, y otras regiones J, tal como JK4, sin embargo, sin apartarse del alcance de la presente invención.

(b) RDC1

En las cadenas kappa humanas naturales, se observaron las RDC1 con longitudes de 11, 12, 13, 16 y 17 siendo predominante con longitud 11 y estando bien representada la longitud 12. Por lo tanto, en las formas de realización preferidas de la presente invención RDC1 con LC de longitud 11 y 12 se utilizan en una mezcla similar (a la observada en los anticuerpos naturales), siendo la longitud 11 la más preferida. La longitud 11 tiene la secuencia siguiente: RASQ<1>V<2><2><3>LA y la longitud 12 tiene la secuencia siguiente: RASQ<1>V<2><2><2><3>LA, en la que <1> es una mezcla equimolar de todos los restos de aminoácidos naturales excepto C, <2> es 0,2 S y 0,044 de cada uno de ADEFGHIKLMNPQRTVWY, y <3> es 0,2 Y y 0,044 cada uno de A, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W e Y. En la forma de realización más preferida de la presente invención, se utilizan ambas longitudes de RDC1. Preferentemente, están presentes en una proporción de 11:12::154:73::0,68:0,32.

(c) RDC2

En kappa natural, RDC2 presenta solo la longitud 7. Esta longitud se utiliza en las formas de realización preferidas de la presente invención. Tiene la secuencia <1>AS<2>R<4><1>, en la que <1> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRTVWY; <2> es 0,2 S y 0,004 de cada ADEFGHIKLMNPQRTVWY; y <4> es 0,2 A y 0,044 de cada DEFGHIKLMNPQRSTUWY.

(d) RDC3

En kappa natural, RDC3 presenta longitudes de 1, 4, 6, 7, 8, 9, 10, 11, 12, 13 y 19. Aunque algunas de estas longitudes y mezclas de ellas pueden utilizarse en la presente invención, resultan preferidas las longitudes 8, 9 y 10, siendo más preferida la longitud 9. Para la longitud 9 preferida la secuencia es QQ<3><1><1><1>P<1>T, en la que <1> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRTVWY y <3> es 0,2 Y y 0,044 cada uno de ADEFGHIKLMNPQRSVW. La longitud 8 es preferentemente QQ33111P y la longitud 10 es preferentemente QQ3211PP1T, en la que 1 y 3 son como se definieron para la longitud 9 y 2 es S (0,2) y 0,044 cada uno de ADEFGHIKLMNPQRTVWY. Siendo una mezcla de las 3 longitudes la más preferida (relaciones como en los anticuerpos naturales), es decir, 8:9:10::28:166:63::0,1:0,65:0,25.

La Tabla 7 representa un gen de la cadena kappa de la presente invención, que incluye un activador PlacZ, una secuencia que une al ribosoma y una secuencia señal (señal M13 III). La secuencia de ADN codifica la secuencia de aminoácidos GLG, pero no comprende la secuencia de GLG ADN. Las secuencias de restricción se diseñan para que estén comprendidas dentro de cada región de la estructura de modo que la diversidad pueda clonarse en las RDC. *XmaI* y *EspI* están en FR1, *SexAI* está en FR2, *RsrII* está en FR3 y *KpnI* (o *Acc65I*) están en FR4. Se proporcionan secuencias adicionales en la cadena kappa constante para facilitar la construcción del gen.

La Tabla 7 presenta asimismo un esquema adecuado de variegación para kappa. En RDC1, está representada la longitud 11 más preferida. Sin embargo, aún más preferentemente se utilizan ambas longitudes 11 y 12. La longitud 12 en RDC1 puede construirse introduciendo el codón 51 como <2> (es decir una mezcla con tendencia a Ser). RDC2 de kappa es siempre de 7 codones. La Tabla 7 representa un esquema de variegación preferida para RDC2. La Tabla 7 muestra un esquema de variegación para los RDC3 más preferidos (longitud 9). Pueden utilizarse variegaciones similares para las RDC de longitud 8 y 10. En la forma de realización preferida de la presente invención, estas tres longitudes (8, 9 y 10) se incluyen en las bibliotecas de la presente invención en las proporciones naturales, tal como se describió anteriormente.

La Tabla 9 presenta unas series de oligonucleótidos y cebadores con diversidad que pueden utilizarse para construir las diversidades de la cadena kappa representadas en la Tabla 7.

(ii) Cadena lambda(a) Marco

- 5 La cadena lambda se construye preferentemente en una estructura 2a2 con una región L2J. Éstas son las regiones V y J más frecuentes en los genes naturales. Otros marcos, tales como 31, 4b, 1a y 6a, y otras regiones J, tales como L1J, L3J y L7J, sin embargo, pueden utilizarse sin apartarse del alcance de la presente invención.

(b) RDC1

- 10 En las cadenas lambda naturales humanas, las RDC1 con longitud 14 predominante, también se producen las longitudes 11, 12 y 13. Aunque alguna de éstas puede utilizarse en la presente invención, resultan preferidas las longitudes 11 y 14. Para la longitud 11 la secuencia es: TG<2><4>L<4><4><3><4> y para la longitud 14 la secuencia es: TG<1>SS<2>VG<1><3><2><3>VS, en la que <1> es 0,27 T, 0,27 G y cada ADEFHIKLMNPQRSVWY es 0,027; <2> es 0,27 D, 0,27 N y cada AEFHGIKLMNPQRSTVWY es 0,027, <3> es 0,36 Y y cada ADEFGHIKLMNPQRSTVW es 0,0355; y <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVWY. Aún más preferentemente, se utilizan mezclas (similares a las que se producen en los anticuerpos naturales) preferentemente, la proporción es 11:14::23:46::0,33:0,67 de las tres longitudes.

(c) RDC2

- 20 En las cadenas lambda naturales humanas, las RDC2 con longitud 7 son con mucho las más frecuentes. Esta longitud resulta preferida en la presente invención. La secuencia de esta RDC2 de longitud 7 es <4><4><4><2>RPS, en el que <2> es 0,27 D, 0,27 N y cada uno de AEFHGIKLMNPQRSTVWY es 0,027 y <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVW.

(d) RDC3

- 30 En las cadenas lambda naturales humanas, predominan las RDC3 de longitud 10 y 11, aunque la longitud 9 también es frecuente. Cualquiera de estas tres longitudes pueden utilizarse en la invención. Resulta preferida la longitud 11 y las mezclas de 10 y 11 son más preferidas. La secuencia de la longitud 11 es <4><5><4><2><4>S<4><4><4>V, en la que <2> y <4> son como se define para la RDC1 lambda y <5> es 0,36 S y cada uno de ADEFGHIKLMNPQRSTVWY es 0,0355. La secuencia de la longitud 10 es <5>SY<1><5>S<5><1><4>V, en la que <1> es una mezcla equimolar de ADEFGHIKLMNPQRSTVWY; y <4> y <5> son como se definen para la longitud 11. Las mezclas preferidas de la presente invención comprenden una mezcla equimolar de longitud 10 y longitud 11. La Tabla 8 presenta una diversidad de la cadena ligera lambda focalizada preferida según la presente invención.

- La Tabla 9 presenta una serie de oligonucleótidos y cebadores con diversidad que pueden utilizarse para construir las diversidades de la cadena lambda esbozadas en la Tabla 7.

Procedimiento de construcción del paquete genético

- Las variedades de cadena pesada y de las cadenas ligeras kappa y lambda se construyen mejor en vectores separados. En primer lugar se diseña un gen sintético para incorporar cada uno de los dominios sintéticos variables. Las cadenas ligeras se unen por las secuencias de restricción para *Apa*LI (colocadas en el mero extremo de la secuencia señal) y *As*cl (colocada después del codón de terminación). La cadena pesada está unida por *Sfi*II (colocada dentro de la secuencia señal Pe1B) y *Not*I (colocada en el enlazador entre CH1 y la proteína de anclaje). Otras secuencias señales además de Pe1B pueden ser necesarias también, por ejemplo, una secuencia señal M13 pIII.

- 50 Los genes iniciales se preparan con secuencias "cebadoras" en lugar de las RDC deseadas. Un "cebador" es una secuencia que ha de cortarse y sustituirse por ADN diverso pero que no permite la expresión de un gen para anticuerpo funcional. Por ejemplo, el cebador puede contener varios codones de terminación y secuencias de restricción que no se producirán en el vector de la biblioteca acabado correcto. Por ejemplo, en la Tabla 10, el cebador para RDC1 de kappa A27 contiene una secuencia *Stu*I. El vgADN para RDC1 se introduce como casete procedente de *Es*pl, *X*maI o *A*flIII para *Sex*AI o *K*asI. Después de la ligadura, el ADN se escinde con *Stu*I; no debería haber ninguna secuencia *Stu*I en los vectores deseados.

- 60 Las secuencias del gen con cadena pesada con cebadores están esbozadas en la Tabla 6. Las secuencias del gen con cadena ligera kappa con cebadores está representada en la Tabla 10. La secuencia del gen con cadena ligera lambda con cebadores está expuesta en la Tabla 11.

- En otra forma de realización de la presente invención las variedades de cadena pesada y de las cadenas ligeras kappa o lambda se construyen en un solo vector o en paquetes genéticos (por ejemplo, para la presentación o presentación y expresión) con secuencias de restricción apropiadas que permiten la clonación de estas cadenas. Los procedimientos para construir dichos vectores son bien conocidos y son muy utilizados en la técnica.

Preferentemente, una cadena pesada y una biblioteca de la cadena ligera kappa y una cadena pesada y la biblioteca de la cadena ligera lambda se prepararían por separado. Las dos bibliotecas aún más preferentemente, se mezclarán a continuación en cantidades equimolares para alcanzar la máxima diversidad.

5 Más preferentemente, la presentación se realiza en la superficie de un derivado del fago M13. El vector más preferido contiene todos los genes de M13, un gen con resistencia a antibióticos y la casete de presentación. El vector preferido se proporciona con secuencias de restricción que permiten la introducción y escisión de los miembros de diversas familias de genes, como casetes. El vector preferido es estable frente a la reorganización en las condiciones de crecimiento utilizadas para ampliar el fago.

10 En otra forma de realización de la presente invención, la variedad capturada por los procedimientos de la presente invención, puede presentarse y/o expresarse en un vector fagómico (por ejemplo, pCES1) que presenta y/expresa el péptido, polipéptido o proteína. Dichos vectores pueden utilizarse también para almacenar la variedad de presentación y/o expresión posteriores utilizando otros vectores o fagos.

15 En otra forma de realización de la presente invención, la variedad capturada por los procedimientos de la presente invención puede presentarse y/o expresarse en un vector de levadura.

```

20 21 22
A M A
ctgtctgaac cc atg gcc
Scab..... NcoI....

FR1 (VP47/V3-23) -----
23 24 25 26 27 28 29 30
E V Q L L E S G
gaa|gtt|caa|ttg|tta|gag|tct|ggt|
MfeI

```

```

-----FR1-----
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
G G L V Q P G G S L R L S C A
|ggc|ggg|ctt|ggt|cag|cct|ggg|ggt|tct|tta|cgt|ctt|tct|tgc|gct|

Puntos de variación          <1>          <1> <1> <1>
-----FR1----->|....CDR1.....|-----FR2-----
46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
A S G F T F S - Y - M - W V R
|gct|tcc|gga|ttc|act|ttc|tct| - |tac| - |atg| - |tgg|ggt|cgc|

BspEI                               BsiWI                               BstXI.

Puntos de variación -><2>          <2> <3>
-----FR2----->|...CDR2.....
61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
Q A P G G K G L E W V S - I - -
|caa|gct|cct|ggt|aaa|ggt|ttg|gag|tgg|ggt|tct| - |atc| - | - |

....BstXI

```

```

<1>          <1>  diversidad de 25992 veces en RDC2
.....CDR2.....|-----FR3-----
76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
S  G  G  -  T  -  Y  A  D  S  V  K  G  R  F
|tct|ggg|ggc| - |act| - |tat|gct|gac|tcc|gtt|aaa|ggg|cgc|ttc|

-----FR3-----
91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
T  I  S  R  D  N  S  K  N  T  L  Y  L  Q  M
|act|atc|tct|aga|gac|aac|tct|aat|aat|act|ctc|tac|ttg|cag|atg|
XbaI

-----FR3----->|
106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
N  S  L  R  A  E  D  T  A  V  Y  Y  C  A  K
|aac|agc|tta|agg|gct|gag|gac|acc|gct|gtc|tac|tgc|gcc|aaa|
AflII

.....CDR3.....|Reemplazadas por varios componentes!
121 122 123 124 125 126 127
D  Y  E  G  T  G  Y
|gac|tat|gaa|ggg|act|ggg|tat|

|-----FR4 --- (JH4) -----
Y  F  D  Y  W  G  Q  G  T  L  V  T  V  S  S
|tat|ttc|gat|tat|tgg|ggg|caa|ggg|acc|ctg|gtc|acc|gtc|tct|agt|...
KpnI      BstEII

```

<1> = Codones para ADEFGHIKLMNPQRSTVWY (mezcla equimolar)
 <2> = Codones para YRWVGS (mezcla equimolar)
 <3> = Codones para PS o PS y G (mezcla equimolar)

Tabla 2: Oligonucleótidos utilizados para RDC1 abigarrada de HC humana

CDR1 - 5 restos	
(ON-R1V1vg):	5'-ct tcc gga ttc act ttc tct <1> tac <1> atg <1> tgg gtt cgc caa gct cct gg-3'
<1> = Codones de ADEFGHIKLMNPQRSTVWY 1:1	
(ON-R1top):	5'-cctactgtct tcc gga ttc act ttc tct-3'
(ON-R1bot) [RC]:	5'-tgg gtt cgc caa gct cct ggttgctcactc-3'
CDR1 - 7 restos	
(ON-R1V2vg):	5'-ct tcc gga ttc act ttc tct <6> <7> tac tac tgg <7> tgg gtt cgc caa gct cct gg-3'
<6> = Codones para ST, 1:1	
<7> = 0,2025 (Codones para SG) + 0,035 (Codones para ADEFGHIKLMNPQRTVWY)	
CDR1 - 14 restos	
(ON-R1V3vg):	5'-ct tcc gga ttc act ttc tct atc agc ggt ggt tct atc tcc <1> <1> <1> tac tac tgg <1> tgg gtt cgc caa gct cct gg-3'
<1> = Codones para ADEFGHIKLMNPQRSTVWY 1:1	

Tabla 3: Oligonucleótidos utilizados para RDC2 abigarrada de HC humana

CDR2 - 17 restos	
(ON-R2V1vg):	5'-ggg ttg gag tgg gtt tct <2> atc <2> <3> tct ggt ggc <1> act <1> tat gct - gac tcc gtt aaa gg-3'
(ON-R2top):	5'-ct tgg gtt cgc caa gct cct ggt aaa ggt ttg gag tgg gtt tct-3'
(ON-R2bot) [RC]:	5'-tat gct gac tcc gtt aaa ggt cgc ttc act atc tct aga ttcctgtcac-3'
<1> = Codones para A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W	Y (mezcla equimolar)
<2> = Codones para Y, R, W, V, G	Y S (mezcla equimolar)
<3> = Codones para P	Y S (mezcla equimolar) o P, S Y G (mezcla equimolar)
(ON-R2V2vg):	5'-ggg ttg gag tgg gtt tct <1> atc <4> <1> <1> ggt <5> <1> <1> tat gct - gac tcc gtt aaa gg-3'
<4> = Codones para DINSWY	(mezcla equimolar)
<5> = Codones para SGDN,	(mezcla equimolar)

CDR2 - 16 restos

(ON-R2V3vg): 5'-ggg|ttg|gag|tgg|ggt|tct|<1>|atc|<4>|<1>|<1>|ggt|<5>|<1>|<1>|tat|aac|cct|tcc|ctt|aag|gg-3'

(ON-R2bo3) [RC]: 5'-tat|aac|cct|tcc|ctt|aag|ggt|cgc|ttc|act|atc|tct|aga|ttcctgtcac-3'

CDR2 - 19 restos

(ON-R2V4vg): 5'-ggg|ttg|gag|tgg|ggt|tct|<1>|atc|<8>|agt|<1>|<1>|<1>|ggg|ggt|act|act|<1>|tat|gcc|gct|tcc|gtt|aag|gg-3'

(ON-R2bo4) [RC]: 5'-tat|gcc|gct|tcc|gtt|aag|ggt|cgc|ttc|act|atc|tct|aga|ttcctgtcac-3'

28

<1>, <2>, <3>, <4> y <5> son tal como se han definido anteriormente

<8> es 0,27 R y ,027 cada uno de ADEFGHIKLMNPQSTVWY

Tabla 4: Componentes preferidos de RDC3 con HC

<u>Componente</u>	<u>Longitud</u>	<u>Complejidad</u>	<u>Fracción de genoteca</u>	<u>Fracción ajustada preferida</u>
1 YYCA21111YFDYWG. (1=cualquier resto de aminoácidos,excepto C; 2 = K y R)	8	2,6 x 10 ⁵	,10	,02
2 YYCA211111YFDYWG. (1=cualquier resto de aminoácidos,excepto C; 2 = K y R)	10	9,4 x 10 ⁷	,14	,14
3 YYCA21111111YFDYWG. (1=cualquier resto de aminoácidos,excepto C; 2 = K y R)	12	3,4 x 10 ¹⁰	,25	,25
4 YYCAR111S2S3111YFDYWG. (1=cualquier resto de aminoácidos,excepto C; 2 = S y G 3 = Y y W)	14	1,9 x 10 ⁸	,13	,14
5 YYCA21111CSG11CY1YFDYWG. (1=cualquier resto de aminoácidos,excepto C; 2 = K y R)	15	9,4 x 10 ⁷	,13	,14
6 YYCA211S1TIFG11111YFDYWG. (1=cualquier resto de aminoácidos,excepto C; 2 = K y R)	17	1,7 x 10 ¹⁰	,11	,12
7 YYCAR111YY2S33YY111YFDYWG. (1=cualquier resto de aminoácidos,excepto C; 2 = D o G; 3 = S y G)	18	3,8 x 10 ⁸	,04	,08
8 YYCAR1111YC2231CY111YFDYWG. (1=cualquier resto de aminoácidos,excepto C; 2 = S y G; 3 = T, D y G)	19	2,0 x 10 ¹¹	,10	,11

(C2t10) :

$\langle 1 \rangle = 0,095$ Y + 0,095 G + 0,048 cada ADEFGHIKLMNPQRSTVW, sin C; $\langle 2 \rangle = K$ Y R (mezcla equimolar)

(C3t12) :

$\langle 1 \rangle = 0,095 \text{ Y} + 0,095 \text{ G} + 0,048$ cada ADEEHKILMNPQRSTVW sin C; $\langle 2 \rangle = \text{K Y R (mezcla equimolar)}$

Componente 4

(C4t140):

5'-cc|gct|gtc|tac|tac|tgc|gcc|cgt|<1>|<1>|<1>|tct|<2>|tct|<3>|<1>|<1>|tac|ttc|gat|-
tac|tgg|ggc|caa|gg-3'

<1> = 0,095 Y + 0,095 G + 0,048 cada ADEFHIKLMNPQRSTVW, sin C; <2> = S Y G (mezcla
equimolar); <3> = Y Y W (mezcla equimolar)

Componente 5

(C5t15):

5'-cc|gct|gtc|tac|tac|tgc|gcc|<2>|<1>|<1>|tgc|tct|ggt|<1>|<1>|tgc|tat|<1>|tac|-
ttc|gat|tac|tgg|ggc|caa|gg-3'

<1> = 0,095 Y + 0,095 G + 0,048 cada ADEFHIKLMNPQRSTVW, sin C; <2> = K Y R (mezcla
equimolar)

Componente 6

(C6t17):

5'-cc|gct|gtc|tac|tac|tgc|gcc|<2>|<1>|<1>|tct|<1>|act|atc|ttc|ggt|<1>|<1>|<1>|<1>|-
<1>|tac|ttc|gat|tac|tgg|ggc|caa|gg-3'

<1> = 0,095 Y + 0,095 G + 0,048 cada ADEFHIKLMNPQRSTVW, sin C; <2> = K Y R (mezcla
equimolar)

Componente 7

(C7t18):

5'-cc|gct|gtc|tac|tac|tgc|gcc|cgt|<1>|<1>|tat|tac|<2>|tct|<3>|<3>|tac|tat|-
<1>|<1>|<1>|tac|ttc|gat|tac|tgg|ggc|caa|gg-3'

<1> = 0,095 Y + 0,095 G + 0,048 cada ADEFHIKLMNPQRSTVW, sin C; <2> = D Y G (mezcla
equimolar); <3> = S and G (mezcla equimolar)

Componente 8

(c8t19):

5'-cc|gct|gtc|tac|tac|tgc|gcc|cgt|<1>|<1>|<1>|<1>|<1>|tat|tgc|<2>|<2>|<3>|<1>|tgc|tat|-
 <1>|<1>|<1>|tac|ttc|gat|tac|tgg|ggc|caa|gg-3'

<1> = 0,095 Y + 0,095 G + 0,048 cada ADEFHIKLMNPQRSTVW, sin C; <2> = S Y G (mezcla
 equimolar); <3> = TDG (mezcla equimolar);

Tabla 6: 3-23::JH4 Cebadores en lugar de las RDC

FR1 (DP47/V3-23) -----										
20	21	22	23	24	25	26	27	28	29 30	
A	M	A	E	V	Q	L	L	E	S G	
ctgtctgaac	cc	atg gcc	gaa gtt caa ttg tta gag tct ggt							
Scab.....	NcoI....	MfeI								
-----FR1-----										
31	32	33	34	35	36	37	38	39	40 41 42 43 44 45	
G	G	L	V	Q	P	G	G	S	L R L S C A	
ggc ggt ctt ggt cag cct ggt ggt tct tta cgt ctt tct tgc gct										
-----FR1-----> ...CDR1 cebador ----FR2-----										
46	47	48	49	50	51	52	53	54	55 56 57 58 59 60	
A	S	G	F	T	F	S	S	Y	A W V R	
gct tcc gga ttc act ttc tct tcg tac gct tag taa tgg gtt cgc										
BspEI					BsiWI					BstXI

```

-----FR2----->|...CDR2 cebador.
61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
Q A P G G K G L E W V S I P R I
|caa|gct|cct|ggt|aaa|ggt|ttg|gag|tgg|gtt|tct|taa|cct|agg|tag|
...BstXI
AvrII..
.....CDR2 cebador .....|---FR3---
91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
T I S R D N S K N T L Y L Q M
|act|atc|tct|aga|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|
XbaI

---FR3-----..> CDR3 cebador ----->|
106 107 108 109 110
N S L R A
|aac|agc|tta|agg|gct|tag|taa|agg|cct|taa
AflII
StuI...

```

```

|----- FR4 ---(JH4)-----|
Y  F  D  Y  W  G  Q  G  T  L  V  T  V  S  S
|tat|ttc|gat|tat|tgg|ggt|caa|ggt|acc|ctg|gtc|acc|gtc|tct|agt|...
                                KpnI    BstEII

```

Tabla 7: Gen para la cadena ligera kappa humana A27:JH1

```

gaggacc attgggcccc ctccgagact ctcgagcgca
Scab..... Eco0109I      XhoI..
      ApaI.

acgcaattaa tgtgagttag ctactcatt aggacccca ggctttacac tttatgcttc
      ..-35..      Plac      ..-10.

cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga

aacagctatg accatgatta

cgccaagctt tggagccttt tttttggaga ttttcaac
      PflMI.....
      Hind III

```

secuencia señal M13 III(AA sec) ----->
 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
 M K K L L L F A I P L V V P F Y
 gtg aag aag ctc cta ttt gct atc ccg ctt gtc gtt ccg ttt tac

--Señal -->FR1----->
 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
 S H S A Q S V L T Q S P G T L
 |agc|cat|agt|gca|caa|tcc|gtc|ctt|act|caa|tct|cct|ggc|act|ctt|
 ApaI...

----- FR1 ----->| CDR1----->
 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
 S L S S P G E R A T L S C R A S
 |tcg|cta|agc|ccg|ggg|gaa|cgt|gct|acc|tta|agt|tgc|cgt|gct|tcc|
 EspI..... AflII...
 XmaI.....

<1> ADEFGHIKLMNPQRSTVWY 1:1

$\langle 3 \rangle$ Y(0, 2) ADEFGHIKLMNPQRSTVW (0, 044 cada uno)

----- CDR1 ----->|----- ER2 ----->

46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

|cag| - |gtt| - | - | - | - |cct|gct|tgg|tat|caa|cag|aaa|cct|

SexAI...

Para CDR2:

- <1> ADEFGHIKLMNPQRSTVWY 1:1
- <2> S(0,2) ADEFGHIKLMNPQRSTVWY (0,044 cada uno)
- <4> A(0,2) DEFGHIKLMNPQRSTVWY (0,044 cada uno)

RDC2 instalada como casete (SexA1 o Kasi) a (BamHI o RsrII).

----- FR2 -----> ----- CDR2 ----->														
61	62	63	64	65	66	67	68	69	70	<1>	<2>	<4>		
G	Q	A	P	R	L	L	I	Y	-	A	S	-	R	-

```
CDR2-->|--- FR3 ----->
<1>
76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
- G I P D R F S G S G S G T D
| - |ggg|atc|cgg|gac|cgt|ttc|tct|ggc|tct|ggt|tca|ggt|act|gac|
    BamHI...
        RsrII.....

----- FR3 ----->
91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
F T L T I S R L E P E D F A V
|ttt|acc|ctt|act|att|tct|aga|ttg|gaa|cct|gaa|gac|ttc|gct|gct|
    XbaI...

Para CDR3 (longitud 9):
<1> ADEFGHIKLMNPQRSTVWY 1:1
<3> Y(0,2) ADEFGHIKLMNPQRTVW (0,044 cada uno)
```


Para RDC3 (longitud 8): QQ33111P

1 y 3 como se definió para la longitud 9

Para RDC3 (longitud 10): QQ3211PP1T

1 y 3 como se definió para la longitud 9

2 S(0,2) y 0,044 cada uno de ADEFGHIKLMNPQRTVWY

RDC3 instalada como XbaI para la casete (StyI o BsiWI)

```

----->|-----CDR3----->|-----FR4----->
          <3> <1> <1> <1> <1> <1>
106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
    Y  Y  C  Q  Q  -  -  -  -  -  P  -  T  F  G  Q
|tat|tat|tgc|caa|cag| - | - | - | - |cct| - |act|ttc|ggt|caa|
      BstXI.....

```

```

-----FR4----->|<----- Ckappa -----
121 122 123 124 125 126 127      128 129 130 131 132 133 134
G T K V E I K      R T V A A P S
|ggt|acc|aag|ggt|gaa|atc|aag|  |cgt|acg|ggt|gcc|gct|cct|agt|
StyI....      BsiWI..

135 136 137 138 139 140 141 142 143 144 145 146 147 148 149
V F I F P P S D E Q L K S G T
|gtg|ttt|atc|ttt|cct|cct|tct|gac|gaa|caa|ttg|aag|tca|ggt|act|
MfeI...

150 151 152 153 154 155 156 157 158 159 160 161 162 163 164
A S V V C L L N N F Y P R E A
|gct|tct|gtc|gta|tgt|ttg|ctc|aac|aat|ttc|tac|cct|cgt|gaa|gct|
BssSI...

```

165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
 K V Q W K V D N A L Q S G N S
 |aaa|gtt|cag|tgg|aaa|gtc|gat|aac|gcg|ttg|cag|tcg|ggt|aac|agt|
 MluI....
 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194
 Q E S V T E Q D S K D S T Y S
 |caa|gaa|tcc|gtc|act|gaa|cag|gat|agt|aag|gac|tct|acc|tac|tct|
 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209
 L S S T L T L S K A D Y E K H
 |ttg|tcc|tct|act|ctt|act|tta|tca|aag|gct|gat|tat|gag|aag|cat|
 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224
 K V Y A C E F V T H Q G L S S P
 |aag|gtc|tat|Gct|TGC|gaa|gtt|acc|cac|cag|ggt|ctg|agc|tcc|cct|
 SacI....

225 226 227 228 229 230 231 232 233 234
V T K S F N R G E C .
|gtt|acc|aaa|agt|ttc|aac|cgt|ggt|gaa|tgc|taa|tag ggcgcgcc
DsaI.... AscI....
BssHII

acgcatctcttaa gcggccgc aacaggaggag
NotI....

Tabla 8: Gen para la cadena lambda humana 2a2:JH2

gaggaccatt gggcccc ttactccgtgac
Scab..... EcoO109I
ApaI..

-----FR1----->															
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
S	A	Q	S	A	L	T	Q	P	A	S	V	S	G	P	G
agt gca caa tcc gct ctc act cag cct gct agc gtt tcc ggg tca cct ggt															
ApaLI...					NheI...					BstEII...					
SexAI.....															

Para CDR1 (longitud 14):
<1> = 0,27 T, 0,27 G, 0,027 cada uno de ADEFHIKLMNPQRSVWY, sin C
<2> = 0,27 D, 0,27 N, 0,027 cada uno de AEEFGHIKLMNPQRSTVWY, sin C
<3> = 0,36 Y, 0,0355 cada uno de ADEFFGHIKLMNPQRSTVW, sin C

```

-----FR1-----> |-----CDR1-----
16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Q S I T I S C T G - S S - V G
|caa|agt|atc|act|att|tct|tgt|aca|ggt| - |tct|tct| - |gtt|ggc|
BsrGI..

<1> <3> <2> <3> V S = vg Esquema #1, Longitud = 14
-----CDR1-----> |-----FR2-----
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
- - - - - V S W Y Q Q H P G K A
| - | - | - | - |gtt|tct|tgg|tat|caa|caa|cac|ccg|ggc|aag|gcg|
XmaI.... KasI.....
AvaI....
```

Un segundo esquema de Vg para RCD1 da segmentos de longitud 11:

T₂G<2><4>I<4><4><4><3><4><4> donde

<4> =mezcla equimolar de cada uno de ADEFGHIKLMNPQRSTVWY, sin C

<3> =como se definió anteriormente para la RDC1 alternativa

Para CDR2:

<2> y <4> son el mismo abigarramiento que para RDC1

<4> <4> <4> <2> R P S

---FR2-----> |-----CDR2----->|-----FR3---

46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

P K L M I Y - - - R P S G V

|ccg|aag|ttg|atg|atc|tac| - | - | - |cgt|cct|tct|ggt|ggt|

KasI....

-----FR3-----

61 62 63 64 65 66 67 68 69 70 71 72 73 74 75

S N R F S S G S K S G N T A S L

|agc|aat|cgt|ttc|tcc|gga|tct|aaa|tcc|ggt|aat|acc|gca|agc|tta|

BspEI..

HindIII.

BsaBI..... (romo)

-----FR3----->|
76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
T I S G L Q A E D E A D Y Y C
|act|atc|tct|tgt|ctg|cag|gct|gaa|gac|gag|gct|gac|tac|tat|tgt|
PstI...

CDR3 (longitud 11):

<2> y <4> son el mismo abigarramiento que para RDC1
<5> = 0,36 S, 0,0355 cada uno de ADEFGHIKLMNPQRTVWY sin C

CDR3 (longitud 10):<5> SY <1> <5> S <5> <1> <4> V

<1>mezcla equimolar de cada uno de ADEFGHIKLMNPQRTVWY, sin C

<4> Y <5> son como se definieron anteriormente para la longitud 11

<4> <5> <4> <2> <4> S <4> <4> <4> <4> V

121 122 123 124 125 126 127 128 129 130 131 132 133 134 135

T L F P P S S E E L Q A N K A

|act|ctc|ttc|cct|cct|agt|tct|gaa|gag|ctt|caa|gct|aac|aag|gct|

SapI.....

136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
T L V C L I S D F Y P G A V T
|act|ctt|gtt|tgc|ttg|atc|agt|gac|ttt|tat|cct|ggt|gct|gtt|act|

BclI....

151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
V A W K A D S S P V K A G V E
|gtc|gct|tgg|aaa|gcc|gat|tct|tct|cct|gtt|aaa|gct|ggt|gtt|gag|

BsmBI....

166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
T T T P S K Q S N N K Y A A S
|acg|acc|act|cct|tct|aaa|caa|tct|aac|aat|aag|tac|gct|gcg|agc|

BsmBI....

SacI....

181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
S Y L S L T P E Q W K S H K S
|tct|tat|ctt|tct|ctc|acc|cct|gaa|caa|tgg|aag|tct|cat|aaa|tcc|

SacI....

```

196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
Y   S   C   Q   V   T   H   E   G   S   T   V   E   K   T
|tat|tcc|tgt|caa|ggt|act|cat|gaa|ggt|tct|acc|ggt|gaa|aag|act|
                                     BspHI...

211 212 213 214 215 216 217 218 219
V   A   P   T   E   C   S   .   .
|gtt|gcc|cct|act|gag|tgt|tct|tag|tga|ggcgcgcc
                                     AscI....
                                     BssHII

aacgatgttc aag gcggccgc aacaggaggag
                                     NotI.... Scab.....

```

Tabla 9: Oligonucleótidos para el abigarramiento de la cadena ligera kappa y lambda

(Ctop25): 5'-gctctggtcaac|tta|agg|gct|gag|g-3'
(CtprmA): 5'-gctctggtcaac|tta|agg|gct|gag|gac|acc|gct|gtc|tac|tac|tgc|gcc-3'
AflII...

(CBprmB) [RC]: 5'-|tac|ttc|gat|tac|ttg|ggc|caa|ggg|acc|ctg|gtc|acc|tcgctccacc-3'
BstEII...

(CBot25) [RC]: 5'-|ggg|acc|ctg|gtc|acc|tcgctccacc-3'

Cadenas kappa: CDR1 ("1"), CDR2 ("2"), CDR3 ("3")

CDR1

(Ka1Top610): 5'-gggtctcagttg|cta|agc|ccg|ggg|gaa|cgt|gct|acc|tta|agt|tgc|cgt|gct|tcc|cag-3'
(Ka1STp615): 5'-gggtctcagttg|cta|agc|ccg|ggg|g-3'
(Ka1Bot620) [RC]: 5'-ctt|gct|tgg|tat|caa|cag|aaa|cct|ggg|cag|gcg|ccaagtcgtgtc-3'
(Ka1SB625) [RC]: 5'-cct|ggg|cag|gcg|ccaagtcgtgtc-3'

(Ka1vg600): 5'-gct|acc|tta|agt|tgc|cgt|gct|tcc|cag-
 |<1>|gtt|<2>|<2>|<3>|ctt|gct|tgg|tat|caa|cag|aaa|cc-3'
 (Ka1vg600-12): 5'-gct|acc|tta|agt|tgc|cgt|gct|tcc|cag-
 |<1>|gtt|<2>|<2>|<2>|<3>|ctt|gct|tgg|tat|caa|cag|aaa|cc-3'

CDR2

(Ka2Tshort657): 5'-cacgagtccta|cct|ggg|cag|gc-3'
 (Ka2Tlong655): 5'-cacgagtccta|cct|ggg|cag|gcg|cgt|tta|ctt|att|tat-3'
 (Ka2Bshort660): [RC]: 5'-|gac|cgt|ttc|tct|ggg|tctcacc-3'
 (Ka2vg650): 5'-cag|gcg|ccg|cgt|tta|ctt|att|tat|<1>|gct|tct|<2>|-
 |cgc|<4>|<1>|ggg|atc|ccg|gac|cgt|ttc|tct|ggg|tctcacc-3'

CDR3

(Ka3Tlon672): 5'-gacgagtccttct|aga|ttg|gaa|cct|gaa|gac|ttc|gct|gtt|tat|tat|tgc|caa|c-3'
 (Ka3BotL682) [RC]: 5'-act|ttc|ggg|caa|ggg|acc|aag|gtt|gaa|atc|aag|cgt|acg|tcacaggtgag-3'
 (Ka3Bsho694) [RC]: 5'-gaa|atc|aag|cgt|acg|tcacaggtgag-3'

```

(Ka3vg670): 5'-gac|ttc|gct|gtt|-
|tat|tat|tgc|caa|cag|<3>|<1>|<1>|<1>|cct|<1>|act|ttc|ggt|caa|-
|ggt|acc|aag|gtt|g-3'
(Ka3vg670-8): 5'-gac|ttc|gct|gtt|-
|tat|tat|tgc|caa|cag|<3>|<3>|<1>|<1>|cct|ttc|ggt|caa|-
|ggt|acc|aag|gtt|g-3'
(Ka3vg670-10): 5'-gac|ttc|gct|gtt|tat|-
|tat|tgc|caa|cag|<3>|<2>|<1>|cct|cct|<1>|act|ttc|ggt|caa|-
|ggt|acc|aag|gtt|g-3'

```

cadenas lambda: CDR1 ("1"), CDR2 ("2"), CDR3 ("3")

CDR1

```

(Im1TPri75): 5'-gacgagtcctgg|tca|cct|ggt|-3'
(Im1t1o715): 5'-gacgagtcctgg|tca|cct|ggt|caa|agt|atc|act|att|tct|tgt|aca|ggt-3'
(Im1blo724) [rc]: 5'-gtt|tct|tgg|tat|caa|caa|cac|cgc|ggc|aag|gcg|agatcttcacaggtgag-3'
(Im1bsh737) [rc]: 5'-gc|aag|gcg|agatcttcacaggtgag-3'
(Im1vg710b): 5'-gt|atc|act|att|tct|tgt|aca|ggt|<2>|<4>|ctc|<4>|<4>|<4>|-
|<3>|<4>|<4>|tgg|tat|caa|caa|cac|cc-3'

```

(Im1vg710): 5'-gt|atc|act|att|tct|tgt|aca|ggg|<1>|tct|tct|<2>|gtt|ggc|
|<1>|<3>|<2>|<3>|gtt|tct|tgg|tat|caa|cac|cc-3'

CDR2

(Im2TSh757): 5'-gagcagaggac|ccg|ggc|aag|gc-3'
(Im2TLo753): 5'-gagcagaggac|ccg|ggc|aag|gcg|ccg|aag|ttg|atg|atc|tac|-3'
(Im2BLo762) [RC]: 5'-cgt|cct|tct|ggg|gtc|agc|aat|cgt|ttc|tcc|gga|tcacaggtgag-3'
(Im2BSh765) [RC]: 5'-cgt|ttc|tcc|gga|tcacaggtgag-3'
(Im2vg750): 5'-g|ccg|aag|ttg|atg|atc|tac|
<4>|<4>|<4>|<2>|cgt|cct|tct|ggg|gtc|agc|aat|c-3'

CDR3

(Im3TSh822): 5'-ctg|cag|gct|gaa|gac|gag|gct|gac-3'
(Im3TLo819): 5'-ctg|cag|gct|gaa|gac|gag|gct|tac|tat|tgt|-3'
(Im3BLo825) [RC]: 5'-gtc|ttc|ggc|ggg|ggg|acc|aaa|ctt|act|gtc|ctc|ggt|caa|cct|aag|g-
acacaggtgag-3'
(Im3BSh832) [RC]: 5'-c|ggg|caa|cct|aag|gacacaggtgag-3'

```

(Im3vg817): 5'-gac|gag|gct|gac|tac|tat|tgt|-
             |<4>|<5>|<4>|<2>|<4>|tct|<4>|<4>|<4>|<4>|-
             Gtc|ttc|ggc|ggt|ggt|acc|aaa|ctt|ac-3'

(Im3vg817-10): 5'- gac|gag|gct|gac|tac|tat|tgt|-
                |<5>|agc|tat|<1>|<5>|tct|<5>|<1>|<4>|gtc|ttc|ggc|ggt|ggt|-
                |acc|aaa|ctt|ac-3'

```


Tabla 10: Gen A27:JH1 para la cadena ligera kappa con cebadores en lugar de las RDC

Cada cebador contiene por lo menos un codón de terminación y una secuencia de restricción que será única en el vector con diversidad

gaggacc attgggcccc ctccgagact ctcgagcgca

Scab.....EcoO109I

ApaI.

XhoI..

acgcaattaa tgtgagtttag ctcaactcatt aggcacccca ggctttacac tttatgcttc

...-35..

Plac

..-10.

cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatgac

catgatta cgccaagctt tggagccttt tttttggaga ttttcaac

PflMI.....

Hind3.

secuencia señal M13 III (AA seq) ----->

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
M	K	K	L	L	F	A	I	P	L	V	V	P	F	Y

gtg aag aag ctc cta ttt gct atc ccg ctt gtc gtt ccg ttt tac

--Señal --> FR1----->

16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
S	H	S	A	Q	S	V	L	T	Q	S	P	G	T	L

lagc|cat|agt|gca|caa|tcc|gtc|ctt|act|caa|tct|cct|ggc|act|ctt|

ApaI...

----- FR1 ----->|-----Cebador->

31	32	33	34	35	36	37	38	39	40	41	42	43
S	L	S	P	G	E	R	A	T	L	S		

|tcg|cta|agc|ccg|ggg|gaa|cgt|gct|acc|tta|agt|tag|taa|gct|ccc|

EspI.....

AflII...

XmaI.....

```

- Cebador para CDR1--> FR2 ----- FR2 ----->|-----Cebador para CDR2
      59 60 61 62 63 64 65 66
      K P G Q A P R
|agg|cct|ctt|tga|tct|g|aaa|cct|ggg|cag|gcg|ccg|cgt|taa|tga|aagcgctaataaggccaacagtg
StuI...      SexAI...      KasI....      AfeI..      MscI..

Cebador-->|--- FR3 ----->
76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
T G I P D R F S G S G S G T D
|act|ggg|atc|ccg|gac|cgt|ttc|tct|ggc|tct|ggg|tca|ggg|act|gac|
      BamHI...
      RsrII.....
```

```

----- FR3 ----->-----Cebador para CDR3----->
91 92 93 94 95 96 97
F  T  L  T  I  S  R  |  |
|ttt|acc|ctt|act|att|tct|aga|taa|tga|  gttaac tag acc tacgta acc tag
      XbaI...      HpaI..      SnaBI.

-----Cebador CDR3----->|-----FR4----->
118 119 120
      F  G  Q
      |ttc|ggg|caa|

-----FR4----->|<----- Ckappa ----->
121 122 123 124 125 126 127      128 129 130 131 132 133 134
G  T  K  V  E  I  K      R  T  V  A  A  P  S
|ggg|acc|aag|ggt|gaa|atc|aag|  |cgt|acg|ggt|gcc|ggt|cct|agt|
      StyI.....      BsiWI..

```

135 136 137 138 139 140 141 142 143 144 145 146 147 148 149
V F I F P P S D E Q L K S G T
|gtg|ttt|atc|ttt|cct|cct|tct|gac|gaa|caa|ttg|aag|tca|ggt|act|
MfeI...

acgcattctctaa gcggccgc aacaggaggag

NotI....

EagI..

Tabla 11: Gen 2a2:JH2 para la cadena ligera lambda humana con cebadores en lugar de las RDC

gaggaccatt gggcccc ttactccgtgac
Scab..... EcoO109I
ApaI..

-----FR1----->															
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
S	A	Q	S	A	L	T	Q	P	A	S	V	S	G	P	G
agt gca caa tcc gct ctc act cag cct gct agc gtt tcc ggg tca cct ggt															
ApaI...					NheI...					BstEII...					
SexAI....															
-----FR1-----> ----- cebador para CDR1-----															
16	17	18	19	20	21	22	23								
Q	S	I	T	I	S	C	T								
caa agt atc act att tct tgt aca tct tag tga ctc															
BsrGI..															

```
----- cebador ----->-----FR2----->
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
R  S  |  |  P  |              H  P  G  K  A
aga tct taa tga ccg tag      cac|ccg|ggc|aag|gcg|
BglII                      XmaI....   KsI.....
                              AvaI.....

--|----- cebador para CDR2 ----->
P
|ccg|taa|tga|atc|tcg|tac|g      ct|ggg|ggt|
KsI....      BsiWI...

-----FR3-----
61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
S  N  R  F  S  G  S  K  S  G  N  T  A  S  L
|agc|aat|cgt|ttc|tcc|gga|tct|aaa|tcc|ggg|aat|acc|gca|agc|tta|
BspEI..      HindIII.
              BsaBI..... (romo)
```

```

-----FR3----->|--- cebador para CDR3----->|
76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
   T  I  S  G  L  Q
|act|atc|tct|ggg|ctg|cag|ggt|ctg|tag|ttc|caattg|ctt|tag|tga|ccc
      PstI...           MfeI..

----- cebador ----->|---FR4-----
103 104 105
   G  G  G
   |ggc|ggg|ggg|
      KpnI...

-----FR4----->
106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
   T  K  L  T  V  L  G  Q  P  K  A  A  P  S  V
|acc|aaa|ctt|act|gtc|ctc|ggg|caa|cct|aag|gct|gct|cct|tcc|gtt|
      KpnI...           HincII..
                        Bsu36I...

```


121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
 T L F P P S S E E L Q A N K A
 |act|ctc|ttc|cct|cct|agt|tct|gaa|gag|ctt|caa|gct|aac|aag|gct|

SapI.....

136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
 T L V C L I S D F Y P G A V T
 |act|ctt|ggt|tgc|ttg|atc|agt|gac|ttt|tat|cct|ggg|gct|ggt|act|
 BclI.....

REIVINDICACIONES

1. Biblioteca de ADN focalizada que comprende secuencias de ADN variegadas que codifican:

5 (1) por lo menos una RDC1 de cadena pesada seleccionada de entre el grupo constituido por:

(a) $\langle 1 \rangle_1 Y_2 \langle 1 \rangle_3 M_4 \langle 1 \rangle_5$, en la que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y;

10 (b) $(S/T)_1 (S/G/X)_2 (S/G/X)_3 Y_4 Y_5 W_6 (S/G/X)_7$.

en la que (S/T) es una mezcla 1:1 de los restos S y T, (S/G/X) es una mezcla de 0,2025 S, 0,2025 G y 0,035 de cada uno de los restos de aminoácidos A, D, E, F, H, I, K, L, M, N, P, Q, R, T, V, W e Y;

15 (c) $V1S2G3G4S5I6S7 \langle 1 \rangle_8 \langle 1 \rangle_9 \langle 1 \rangle_{10} Y_{11} Y_{12} W_{13} \langle 1 \rangle_{14}$, en la que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y

(d) mezclas de vectores o de paquetes genéticos caracterizados porque presentan cualquiera de las secuencias de ADN anteriores;

20 siendo los miembros de este grupo las únicas RDC1 de cadena pesada presentes en la biblioteca;

(2) por lo menos una de entre una RDC2 de cadena pesada seleccionada de entre el grupo constituido por:

25 (a) $\langle 2 \rangle I \langle 2 \rangle \langle 3 \rangle SGG \langle 1 \rangle T \langle 1 \rangle YADSVKG$, en la que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; $\langle 2 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos Y, R, W, V, G y S; y $\langle 3 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos P, S y G o una mezcla equimolar de P y S;

30 (b) $\langle 1 \rangle I \langle 4 \rangle \langle 1 \rangle \langle 1 \rangle G \langle 5 \rangle \langle 1 \rangle \langle 1 \rangle YADSVKG$, en la que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; $\langle 4 \rangle$ es una mezcla equimolar de los restos D, I, N, S, W, Y; y $\langle 5 \rangle$ es una mezcla equimolar de los restos S, G, D y N;

35 (c) $\langle 1 \rangle I \langle 4 \rangle \langle 1 \rangle \langle 1 \rangle G \langle 5 \rangle \langle 1 \rangle \langle 1 \rangle YNPSLKG$, en la que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y $\langle 4 \rangle$ y $\langle 5 \rangle$ son tal como se definieron anteriormente;

40 (d) $\langle 1 \rangle I \langle 8 \rangle S \langle 1 \rangle \langle 1 \rangle \langle 1 \rangle GGY \langle 1 \rangle YAASVKG$, en la que $\langle 1 \rangle$ es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y $\langle 8 \rangle$ es 0,27 R y 0,027 de cada uno de ADEFGHIKLMNPQSTVWY; y

(e) mezclas de vectores o de paquetes genéticos caracterizados porque presentan cualquiera de las secuencias de ADN anteriores;

45 siendo los miembros de este grupo las únicas RDC2 de cadena pesada presentes en la biblioteca; y

(3) por lo menos una de una RDC3 de cadena pesada seleccionada de entre el grupo constituido por:

50 (a) YYCA21111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de K y R;

(b) YYCA211111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de K y R;

55 (c) YYCA21111111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de K y R;

60 (d) YYCAR111S2S3111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de S y G; y 3 es una mezcla equimolar de Y y W;

(e) YYCA2111CSG11CY1YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de K y R;

65 (f) YYCA211S1TIFG11111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; y 2 es una mezcla equimolar de K y R;

(g) YYCAR111YY2S3344111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; 2 es una mezcla equimolar de D y G; y 3 es una mezcla equimolar de S y G;

(h) YYCAR1111YC2231CY111YFDYWG, en la que 1 es una mezcla equimolar de cada uno de los restos de aminoácidos A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; 2 es una mezcla equimolar de S y G; y 3 es una mezcla equimolar de T, D y G; y

(i) mezclas de vectores o paquetes genéticos caracterizados porque presentan cualquiera de las secuencias de ADN anteriores;

siendo los miembros de este grupo las únicas RDC3 de cadena pesada presentes en la biblioteca.

2. Biblioteca focalizada según la reivindicación 1, en la que las RDC1 de cadena pesada (a), (b) y (c) están presentes en la biblioteca en la relación 0,80:0,17:0,02.

3. Biblioteca focalizada según la reivindicación 1, en la que una mezcla de las RDC2 de cadena pesada (a)/(b) (equimolar), (c) y (d) están presentes en la biblioteca en una relación de 0,54:0,43:0,03.

4. Biblioteca focalizada según la reivindicación 1, en la que 1 en una o la totalidad de las RDC3 de cadena pesada (a) hasta (h) es 0,095 de cada una de G e Y y 0,048 de cada una de A, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V y W.

5. Biblioteca focalizada según la reivindicación 1, en la que las RDC3 de cadena pesada (a) hasta (h) están presentes en la biblioteca en las proporciones siguientes:

- (a) 0,10
- (b) 0,14
- (c) 0,25
- (d) 0,13
- (e) 0,13
- (f) 0,11
- (g) 0,04 y
- (h) 0,10.

6. Biblioteca focalizada según la reivindicación 1, en la que las RDC3 de cadena pesada (a) hasta (h) están presentes en la biblioteca en las proporciones siguientes:

- (a) 0,02
- (b) 0,14
- (c) 0,25
- (d) 0,14
- (e) 0,14
- (f) 0,12
- (g) 0,08 y
- (h) 0,11.

7. Biblioteca de ADN focalizada según cualquiera de las reivindicaciones 1 a 6, que comprende además las secuencias de ADN variegadas que codifican por lo menos uno de entre:

(1) por lo menos una de una RDC1 de cadena ligera kappa seleccionada de entre el grupo constituido por:

(a) RASQ<1>V<2><2><3>LA

(b) RASQ<1>V<2><2><2><3>LA;

en la que <1> es una mezcla equimolar de los restos de aminoácidos ADEFGHIKLMNPQRSTVWY; <2> es 0,2 S y 0,044 de cada uno de ADEFGHIKLMNPQRTVWY; y <3> es 0,2 Y y 0,044 cada uno de ADEFGHIKLMNPQRTVW e Y; y

(c) mezclas de los vectores o paquetes genéticos caracterizados porque presentan cualquiera de las secuencias de ADN anteriores;

siendo los miembros de este grupo las únicas RDC1 de cadena ligera kappa presentes en la biblioteca;

(2) una RDC2 de cadena ligera kappa que presenta la secuencia:

<1>AS<2>R<4><1>

en la que <1> es una mezcla equimolar de los restos de aminoácidos ADEFGHIKLMNPQRSTVWY; <2> es 0,2 S y 0,044 de cada uno de ADEFGHIKLMNPQRTVWY; y <4> es 0,2 A y 0,044 cada uno de DEFGHIKLMNPQRSTVWY;

siendo los miembros de este grupo las únicas RDC1 de cadena ligera kappa presentes en la biblioteca; o

(3) por lo menos una RDC3 de cadena ligera kappa seleccionada de entre el grupo constituido por:

(a) QQ<3><1><1><1>P<1>T,

en la que <1> es una mezcla equimolar de los restos de aminoácidos ADEFGHIKLMNPQRSTVWY; <3> es 0,2 Y y 0,044 de cada uno de ADEFGHIKLMNPQRTVW;

(b) QQ33111P, en la que 1 y 3 son tal como se definieron en (1) anteriormente;

(c) QQ3211PP1T, en la que 1 y 3 son tal como se definieron en (1) anteriormente y 2 es 0,2 S y 0,044 de cada uno de ADEFGHIKLMNPQRTVWY; y

(d) mezclas de los vectores o paquetes genéticos caracterizados porque presentan cualquiera de las secuencias de ADN anteriores;

siendo los miembros de este grupo las únicas RDC3 de cadena ligera kappa presentes en la biblioteca.

8. Biblioteca de ADN focalizada según la reivindicación 7, que comprende las secuencias de ADN variegadas que codifican por lo menos un miembro de por lo menos dos de los grupos (1), (2) y (3) de la RDC de cadena ligera kappa.

9. Biblioteca de ADN focalizada según la reivindicación 7, que comprende las secuencias de ADN variegadas que codifican por lo menos un miembro de todos los grupos (1), (2) y (3) de la RDC de cadena ligera kappa.

10. Biblioteca focalizada según cualquiera de las reivindicaciones 7 a 9, en la que las RDC1 (a) y (b) de cadena ligera kappa están presentes en la biblioteca en una proporción de 0,68:0,32.

11. Biblioteca focalizada según cualquiera de las reivindicaciones 7 a 9, en la que las RDC3 (a), (b) y (c) de cadena ligera kappa están presentes en la biblioteca en una proporción de 0,65:0,1:0,25.

12. Biblioteca de ADN focalizada según cualquiera de las reivindicaciones 1 a 11, que comprende además las secuencias de ADN variegadas que codifican por lo menos uno de entre:

(1) por lo menos una RDC1 de cadena ligera lambda seleccionada de entre el grupo constituido por:

(a) TG<1>SS<2>VG<1><3><2><3>VS, en la que <1> es 0,27 T, 0,27 G y 0,027 cada ADEFHIKLMNPQRSVWY, <2> es 0,27 D, 0,27 N y 0,027 cada AEFGHIKLMPQRSTVWY y <3> es 0,36 Y y 0,036 cada ADEFGHIKLMNPQRSTVW;

(b) G<2><4>L<4><4><4><3><4><4>, en la que <2> es tal como se definió en (1) anteriormente y <4> es una mezcla equimolar de los restos de aminoácidos ADEFGHIKLMNPQRSTVWY; y

(c) mezclas de los vectores o paquetes genéticos caracterizados porque presentan cualquiera de las secuencias de ADN anteriores;

siendo los miembros de este grupo las únicas RDC1 de cadena ligera lambda presentes en la biblioteca;

(2) una RDC2 de cadena ligera lambda que presenta la secuencia:

<4><4><4><2>RPS,

en la que <2> es 0,27 D, 0,27 N y 0,027 cada AEFGHIKLMPQRSTVWY y <4> es una mezcla equimolar de los restos de aminoácidos ADEFGHIKLMNPQRSTVW;

siendo los miembros de este grupo las únicas RDC2 de cadena ligera lambda presentes en la biblioteca;

(3) por lo menos una de una RDC3 de cadena ligera lambda seleccionada de entre el grupo constituido por:

(a) <4><5><4><2><4>S<4><4><4>V, en la que <2> es 0,27 D, 0,27 N y 0,027 cada AEF₂GH₂IK₂LM₂P₂QR₂ST₂V₂W₂Y; <4> es una mezcla equimolar de los restos de aminoácidos ADEFGHIKLMNPQRSTVW; y <5> es 0,36 S y 0,0355 cada ADEFGHIKLMNPQRSTVWY;

5 (b) <5>SY<1><5>S<5><1><4>V, en la que <1> es una mezcla equimolar de ADEFGHIKLMNPQRSTVWY; y <4> y <5> son tal como se definieron en (1) anteriormente; y

(c) mezclas de los vectores o de paquetes genéticos caracterizados porque presentan cualquiera de las secuencias de ADN anteriores;

10 siendo los miembros de este grupo las únicas RDC3 de cadena ligera lambda presentes en la biblioteca.

13. Biblioteca de ADN focalizada según la reivindicación 12, que comprende las secuencias de ADN variegadas que codifican por lo menos un miembro de entre por lo menos dos de los grupos (1), (2) y (3) de RDC de cadena ligera lambda.

14. Biblioteca de ADN focalizada según la reivindicación 12, que comprende las secuencias de ADN variegadas que codifican por lo menos un miembro de cada uno de los grupos (1), (2) y (3) de RDC de cadena ligera lambda.

20 15. Biblioteca focalizada según cualquiera de las reivindicaciones 12 a 14, en la que las RDC1 (a) y (b) de cadena ligera lambda están presentes en la biblioteca en una proporción de 0,67:0,33.

16. Biblioteca focalizada según cualquiera de las reivindicaciones 12 a 14, en la que las RDC3 (a) y (b) de cadena ligera lambda están presentes en la biblioteca en una mezcla equimolar.

25