(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 430 857**

(51) Int. CI.:

C12N 15/10 (2006.01)
C12N 15/13 (2006.01)
C07K 16/00 (2006.01)
C12P 21/00 (2006.01)
C12N 15/09 (2006.01)
C40B 40/02 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **18.12.2001 E 10182768 (1)**

(97) Fecha y número de publicación de la concesión europea: **24.07.2013 EP 2316940**

(54) Título: **Bibliotecas focalizadas de paquetes genéticos**

(30) Prioridad:

18.12.2000 US 256380 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
22.11.2013

(73) Titular/es:

DYAX CORP. (100.0%)
300 Technology Square
Cambridge, MA 02139, US

(72) Inventor/es:

LADNER, ROBERT CHARLES

(74) Agente/Representante:

LEHMANN NOVO, María Isabel

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.

La presente invención se refiere a bibliotecas focalizadas de paquetes genéticos que presentan, presentan y expresan, o comprenden cada una 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 focalizada de la familia, como se define en las reivindicaciones. La diversidad focalizada 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 focalizada 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.

La presente descripción describe además vectores y paquetes genéticos (por ejemplo, células, esporas o virus) para presentar, o presentar y expresar una familia diversa focalizada de péptidos, polipéptidos o proteínas. En una realización preferda, los paquetes genéticos son fagos filamentosos o fagómidos o levaduras. Nuevamente, la diversidad focalizada de la familia comprende diversidad en secuencia y diversidad en longitud.

La presente descripción describe además métodos de cribado de las bibliotecas focalizadas de la invención, y a los péptidos, polipéptidos o proteínas identificados mediante tal cribado.

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 porción 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 completos o minicuerpos (es decir, dímeros constituidos por V_H unida 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). Un método consiste en capturar la diversidad de donantes naturales, ya sea sin tratamiento previo o inmunizados. Otra manera consiste en generar bibliotecas con diversidad sintética. Un tercer método consiste en una combinación de los dos primeros. Típicamente, la diversidad producida por estos métodos se limita a la diversidad de secuencias, es decir, cada miembro de la biblioteca se diferencia de los otros miembros de la familia porque presenta diferentes aminoácidos o variegación en una posición dada en la cadena peptídica, polipeptídica o proteica. Sin embargo, péptidos, polipéptidos o proteínas naturalmente diversos 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 diversos en las longitudes de sus cadenas de aminoácidos.

Para los anticuerpos, la diversidad en 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 a los genes J, por ejemplo, produce la inclusión de un segmento D reconocible en RDC3 en aproximadamente la mitad de las secuencias de anticuerpos de cadena pesada, creando de este modo regiones que codifican longitudes variables de aminoácidos. Lo siguiente puede producirse también durante la unión de los segmentos génicos del anticuerpo: (i) el extremo del gen V puede presentar de cero a varias bases suprimidas o cambiadas; (ii) el extremo 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 editarse 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 unión 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 gran número de miembros que se producen tratando de capturar la diversidad completa

también hace la identificación más problemática de 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 ADN focalizadas que comprenden secuencias de ADN variegadas de las RDC de anticuerpos humanos que presentan una diversidad tanto en sus secuencias de aminoácidos como en su longitud (ejemplos de dichas bibliotecas incluyen bibliotecas de Fv monocatenario (scFv), Fv, Fab, anticuerpos completos o minicuerpos (es decir, dímeros que consisten en V_H unida a V_L)). Dichas regiones pueden ser de las cadenas pesada o ligera o de ambas, y pueden incluir una o más de las RDC de estas cadenas. Más preferentemente, la diversidad o variegación se produce en todas las RDC de la cadena pesada y cadena ligera.

Entre las realizaciones preferidas de esta invención, están las siguientes:

1. Una biblioteca de ADN focalizada como se define en el párrafo 4 más abajo, que se caracteriza por comprender además secuencias de ADN variegadas que codifican una RDC1 de cadena pesada seleccionada del grupo que consiste en:

(1) <1>$_1$Y$_2$<1>$_3$M$_4$<1>$_5$, en el 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) (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;

(3) V$_1$S$_2$G$_3$G$_4$S$_5$I$_6$S$_7$<1>$_8$<1>$_9$<1>$_{10}$Y$_{11}$Y$_{12}$W$_{13}$<1>$_{14}$, en los 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

(4) mezclas de vectores o de paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores, preferiblemente en la relación: HC RDC1s (1):(2):(3)::0,80:0,17:0,02.

2. Una biblioteca de ADN focalizada como se define en el párrafo 4 más abajo, que se caracteriza por comprender además secuencias de ADN variegadas que codifican una RDC2 de cadena pesada seleccionada del grupo que consiste en:

(1) <2>I<2><3>SGG<1>T<1>YADSVKG, 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 cada uno de los restos de aminoácidos Y, R, W, V, G y S; y <3> es una mezcla equimolar de cada uno de los restos de aminoácidos P, S y G o una mezcla equimolar de P y S;

(2) <1>I<4><1><1><G><5><1><1><1>YADSVKG, 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; <4> es una mezcla equimolar de los restos de D, I, N, S, W, Y; y <5> es una mezcla equimolar de los restos de S, G, D y N;

(3) <1>I<4><1><1><1>G<5><1><1><1>YNPSLKG, 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 <4> y <5> son como se definieron anteriormente;

(4) <1>I<8>S<1><1><1><1>GGYY<1>YAASVKG, 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 <8> es 0,27 R y 0,027 de cada uno de ADEFGHIKLMNPQSTVWY; y

(5) mezclas de vectores o de paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores, preferiblemente en la relación: HC RDC2s: (1)/(2) (equimolar): (3):(4)::0,54:0,43:0,03.

3. Una biblioteca de ADN focalizada como se define en el párrafo 4 más abajo, que se caracteriza por comprender además secuencias de ADN variegadas que codifican una RDC3 de cadena pesada seleccionada del grupo que consiste en:

(1) 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;

(2) 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;

(3) YYCA211111111YFDYWG, 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;

(4) 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;

(5) 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;

(6) 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;

(7) 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;

(8) 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

(9) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores; preferiblemente las HC RDC3s (1) a (8) están en las siguientes proporciones en la mezcla:

(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, y más preferiblemente las HC RDC3s (1) a (8) están en las siguientes proporciones en la mezcla:

(1) 0,02

(2) 0,14

(3) 0,25

(4) 0,14

(5) 0,14

(6) 0,12

(7) 0,08 y

(8) 0,11.

Preferentemente, 1 en una o todas las HC RDC3s (1) a (8) 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.

4. Una biblioteca de ADN focalizada que codifica vectores o paquetes genéticos que presenta, presenta y expresa, o comprende un miembro de una familia diversa de péptidos, polipéptidos o proteínas relacionados con anticuerpos humanos, y colectivamente presenta, presenta y expresa, o comprende por lo menos una parte de la diversidad de la familia de anticuerpos, caracterizándose los vectores o paquetes genéticos por secuencias de ADN variegadas que codifican una RDC1 de cadena ligera kappa seleccionada del grupo que consiste en:

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

(2) 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 uno de ADEFGHIKLMNPQRTVWY; y <3> es 0,2 Y y 0,044 cada uno de ADEFGHIKLMNPQRSTVW; y

(3) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores, preferiblemente en las relaciones RDC1s (1) : (2) : :0,68:0,32; una RDC2 de cadena ligera kappa que tiene 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 uno de ADEFGHIKLMNPQRTVWY; y <4> es 0,2 A y 0,044 de cada uno de DEFGHIKLMNPQRSTVWY; y que comprende además una RDC3 de cadena ligera kappa seleccionada de los grupos que consisten en:

(1) 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 uno de ADEFGHIKLMNPQRTVW;

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

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

(4) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores, preferiblemente en la relación RDC3s (1) : (2) : (3) : :0,65:0,1:0,25.

5. Una biblioteca de ADN focalizada como se define en el párrafo 4 más abajo, que se caracteriza por comprender además secuencias de ADN variegadas que codifican una RDC1 de cadena pesada seleccionada del grupo que consiste en:

(1) 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 uno de ADEFHIKLMNPQRSVWY, <2> es 0,27 D, 0,27 N y 0,027 cada uno de AEFGHIKLMNPQRSTVWY y <3> es 0,36 Y y 0,036 cada uno de ADEFGHIKLMNPQRSTVW;

(2) G<2><4>L<4><4><4><3><4><4>

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

(3) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores, preferiblemente en la relación RDC1s (1) : (2) : :0,67:0,33; y/o

secuencias de ADN variegadas que codifican 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 uno de AEFGHIKLMPQRSTVWY, y <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVW; y/o secuencias de ADN variegadas que codifican una RDC3 de cadena ligera lambda seleccionada del grupo que consiste en:

(1) <4><5><4><2><4>S<4><4><4><4>V,

en la que <2> es 0,27 D, 0,27 N y 0,027 cada uno de AEFGHIKLMPQRSTVWY; <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVW; y <5> es 0,36 S y 0,0355 cada uno de ADEFGHIKLMNPQRTVWY;

(2) <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

(3) mezclas de vectores o de paquetes genéticos caracterizados por algunas de las secuencias de ADN anteriores, preferiblemente en la relación RDC3s (1):(2)::1:1.

12. Una población de secuencias de ADN variegadas como se describen en los párrafos 1-5 anteriores.

13. Una población de vectores que comprenden las secuencias de ADN variegadas como se describen en los párrafos 1-5 anteriores.

DESCRIPCIÓN DETALLADA DE LAS FORMAS DE REALIZACIÓN PREFERIDAS

Los anticuerpos (“Ab”) concentran su diversidad en aquellas regiones que están implicadas en la determinación de la afinidad y especificidad del Ab para dianas específicas. Estas regiones pueden ser diversas en secuencia o longitud. Generalmente, son diversas en ambos aspectos. Sin embargo, dentro de las familias de anticuerpos humanos, las diversidades, tanto en secuencia como en 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, se preparan y utilizan 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 longitud de la cadena pesada de anticuerpos humanos

(a) Estructura

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.

Además, JH4 (YFDYWGQGTLVTUSS) 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 (AFDIWGQGTMVTVSS).

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

(i) RDC1

Para RDC1, los 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 RDC1 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, teniendo las RDC1 menos preferidas 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.

(ii) RDC2

Los GLG proporcionan RDC2 solamente de las longitudes 15-19 pero las mutaciones durante la maduración pueden dar como resultado las RDC2 de longitudes entre 16 y 28 aminoácidos. Las longitudes 16 y 17 predominan en los genes de 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.

(iii) RDC3

Las RDC3 de HC varían en longitud. Aproximadamente la mitad de las HC humanas consisten en los componentes siguientes: V::nz::D::ny::JHn en el 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 edición pesada 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.

En las bibliotecas preferidas de la presente invención se utilizan ambos tipos de RDC3 de 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 secuencia focalizada: RDC1, 2 y 3

(i) RDC1

En la RDC1, con longitud de 5 aminoácidos, el examen de un modelo 3D de un Ab humanizado mostró que los grupos laterales de los restos 1, 3 y 5 se dirigieron hacia el bolsillo 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. La posesión de 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.

El modelado en 3D sugiere además que los grupos laterales del resto 2 en la RDC1 de 5 aminoácidos se dirijan lejos del bolsillo de combinación. Aunque esta posición presenta sustancial diversidad, tanto en GLG como en genes maduros, en las bibliotecas focalizadas de la presente invención este resto es preferentemente Tyr (Y) porque aparece en los genes de los anticuerpos maduros 681/820. Sin embargo, cualquiera de los demás restos de aminoácidos naturales, excepto Cys (C), podría utilizarse también en esta posición.

Para la posición 4, existe también cierta 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 muestra 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 ajuste muy bien en el marco de 3-23. Con otras estructuras, se utiliza una consideración del ajuste similar para asignar el resto 4.

Por lo tanto, la RDC1 de HC más preferida de la presente invención consiste en la secuencia de aminoácidos de <1>Y<1>M<1>, en la que <1> 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 muestra en el contexto de una estructura 3-23:JH4 en la Tabla 1. Presenta una diversidad de 6859 veces.

Las dos RDC1 de HC menos preferidas de la presente invención presentan una longitud 7 y una longitud 14. Para la longitud 7, una variegación preferida es $(S/T)_1(S/G/<1>)_2\ (S/G/<1>)_3Y_4Y_5W_6(S/G/<1>)_7$; en la que (S/T) indica una mezcla equimolar de los codones Ser y Thr; (S/G/<1>) indica una mezcla de 0,2025 S, 0,2025 G y 0,035 para cada uno de 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 de HC. Para la longitud 14, una variegación preferida es VSGGSIS<1><1><1>YYW<1>, en la que <1> es una mezcla equimolar de los 19 restos de aminoácidos naturales, excepto Cys (C).

El ADN que codifica estas RDC1 de HC preferidas 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 <1> 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.

El oligonucleótido de diversidad (ON) se sintetiza preferentemente a partir de *Bsp*EI hasta *Bst*XI (como se muestra en la Tabla 1), y puede, por consiguiente, incorporarse ya sea por síntesis de RCP utilizando ON solapantes o se puede introducir por ligadura de fragmentos de corte *Bsp*EI/*Bst*XI. La Tabla 2 presenta los oligonucleótidos que incorporan las variegaciones especificadas de RDC1 de HC de longitud 5 preferidas de la presente invención. La RCP que utiliza ON-R1V1vg, ON-R1top y ON-R1bot proporciona un producto de dsADN de 73 pares de bases, se escinde con las bases 11 y 13 de los de los extremos de los recortes *Bsp*EI y *Bst*XI y proporciona extremos cohesivos que pueden ligarse al vector cortado similarmente 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 de 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 de HC preferidas. Más preferentemente, las 3 longitudes se incorporarían en aproximadamente las relaciones en las que se observan en los anticuerpos seleccionados sin referencia a la longitud de las RDC. Por ejemplo, una muestra de 1095 genes de 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: <2>I<2><3>SGG<1>T<1>YADSVKG, en la que <2> indica cualquier resto de aminoácidos seleccionado del grupo Y, R, W, V, G y S (mezcla equimolar), <3> es P, S y G o P y S solamente (mezcla equimolar), y <1> 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 el sitio *Bst*XI y *Xba*I pueden generarse por RCP. RCP con ON-R2V1vg, ON-R2top y ONR2bot proporciona un producto de dsADN de 122 pares de bases. La escisión con *Bst*XI y *Xba*I elimina aproximadamente 10 bases de cada extremo y produce extremos cohesivos que pueden ligarse al vector cortado similarmente que contiene el gen 3-23 mostrado en la Tabla 1.

En una forma de realización alternativa para una RDC2 de HC de longitud 17, puede utilizarse la variegación siguiente: <1>I<4><1><1>G<5><1><1><1>YADSVKG, en la que <1> es como se describió anteriormente para la alternativa más preferida de RDC2 de HC; <4> indica una mezcla equimolar de DINSWY, y <5> indica una mezcla equimolar de SGDN. 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 de 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 las 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 SGDN; 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 con 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 de 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 de 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 por permitir que se plieguen las Ig. Los componentes de RDC3 de HC de las bibliotecas preferidas de la presente invención y sus diversidades se representan en la Tabla 4: Componentes 1-8.

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

1) aproximadamente la misma distribución de longitudes como la observada en los genes de Ab naturales;

2) gran nivel de diversidad de secuencias en los lugares que tienen gran diversidad en los genes de Ab naturales; y

3) incorporación de secuencias constantes observadas con frecuencia en los genes de Ab naturales.

El componente 1 representa todos los genes que tiene longitudes 0 a 8 (contando desde el motivo YYCAR en el extremo de FR3 hasta el motivo dipeptídico 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 de 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, estando los demás restos 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 a tener 1×10^9 miembros, entonces 1×10^8 secuencias procederían del componente 1, pero presenta solamente $2{,}6 \times 10^5$ secuencias de 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 ~77 contextos RDC1/2 y el otro, las RDC3 más largas, se producen con más frecuencia.

La Tabla 5 presenta vgADN que incorpora cada uno de los ocho componentes de RDC3 de HC mostrados en la Tabla 4. En la Tabla 5, los oligonucleótidos (ON) Ctop25, CtprmA, CBprmB 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 *Afl*II y *Bst*EII (o *Kpn*I) y ligarse al vector escindido de forma similar 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 las 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 parada 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 un sitio de enzima de restricción que no aparecerá en la biblioteca acabada, permitiendo la eliminación de vectores que no son escindidos tanto por *Afl*II como por *Bst*EII (o *Kpn*I) 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 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 RDC1 con longitudes de 11, 12, 13, 16 y 17, siendo predominante la longitud 11 y estando la longitud 12 bien representada. Por lo tanto, en la presente invención, RDC1 de 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, S, T, V, W. En la forma de realización más preferida de la presente invención, se utilizan ambas longitudes de RDC1. Preferentemente, están presentes en una relación 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 la presente invención. Tiene 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 0,004 de cada uno de ADEFGHIKLMNPQRTVWY; y <4> es 0,2 A y 0,044 de cada uno de 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 ADEFGHIKLMNPQRSTVWY 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, un sitio de unión 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 ADN de GLG. Los sitios 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. *Xma*I y *Esp*I están en FR1, *Sex*AI está en FR2, *Rsr*II está en FR3 y *Kpn*I (o Acc65I) están en FR4. Se proporcionan sitios 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 la RDC3 más preferida (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 relaciones naturales, tal como se describió anteriormente.

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

(ii) Cadena lambda

(a) Marco

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

En las cadenas lambda naturales humanas, predominan las RDC1 con longitud 14, también aparecen las longitudes 11, 12 y 13 . Aunque cualquiera 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><4><3><4><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 0,027 cada uno de ADEFHIKLMNPQRSVWY; <2> es 0,27 D, 0,27 N y 0,027cada uno de AEFGHIKLMPQRSTVWY, <3> es 0,36 Y y 0,0355 cada uno de ADEFGHIKLMNPQRSTVW; 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 relación es 11:14::23:46::0,33:0,67 de las tres longitudes.

(c) RDC2

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

(d) RDC3

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 puede 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><4>V, en la que <2> y <4> son como se definen para la RDC1 lambda, y <5> es 0,36 S y 0,0355 cada uno de ADFFGHIKLMNPQRTVWY. 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 de diversidad que pueden utilizarse para construir las diversidades de la cadena lambda representadas en la Tabla 7.

Método de construcción del paquete genético

Las diversidades 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 los sitios de restricción para *Apa*LI (colocada en el mero extremo de la secuencia señal) y *Asc*I (colocada después del codón de parada). La cadena pesada está unida por *Sfi*I (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.

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 de anticuerpo funcional. Por ejemplo, el cebador puede contener varios codones de parada y sitios de restricción que no aparecerán en el vector de la biblioteca acabado correcto. Por ejemplo, en la Tabla 10, el cebador para RDC1 de kappa A27 contiene un sitio *Stu*I. El vgADN para RDC1 se introduce como un casete procedente de *Esp*I, *Xma*I o *Afl*II a *Sex*AI o *Kas*I. Después de la ligadura, el ADN se escinde con *Stu*I; no debería haber ningún sitio *Stu*I en los vectores deseados.

Las secuencias del gen de cadena pesada con cebadores están representadas en la Tabla 6. Las secuencias del gen de cadena ligera kappa con cebadores están representadas en la Tabla 10. La secuencia del gen de cadena ligera lambda con cebadores está representada en la Tabla 11.

En otra forma de realización de la presente invención, las diversidades 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 sitios de restricción apropiados 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.

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 de resistencia a antibióticos, y el casete de presentación. El vector preferido se proporciona con sitios 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.

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

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

Se describen particularmente los siguientes aspectos:

1. Una biblioteca focalizada de vectores o paquetes genéticos que presenta, presenta y expresa, o comprende un miembro de una familia diversa de péptidos, polipéptidos o proteínas relacionados con anticuerpos humanos, y colectivamente presenta, presenta y expresa, o comprende por lo menos una parte de la diversidad de la familia de anticuerpos, caracterizándose los vectores o paquetes genéticos por secuencias de ADN variegadas que codifican una RDC1 de cadena pesada seleccionada del grupo que consiste en:

(1) <1>$_1$Y$_2$<1>$_3$M$_4$<1>$_5$, en el 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) (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;

(3) V$_1$S$_2$G$_3$G$_4$S$_5$I$_6$S$_7$<1>$_8$<1>$_9$<1>$_{10}$Y$_{11}$Y$_{12}$W$_{13}$<1>$_{14}$, en los 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

(4) mezclas de vectores o de paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores.

2. La biblioteca focalizada según el aspecto 1, en la que las RDC1 de HC (1), (2) y (3) están presentes en la biblioteca en una relación 0,80:0,17:0,02.

3. Una biblioteca focalizada de vectores o paquetes genéticos que presenta, presenta y expresa, o comprende un miembro de una familia diversa de péptidos, polipéptidos o proteínas relacionados con anticuerpos humanos, y colectivamente presenta, presenta y expresa, o comprende por lo menos una parte de la diversidad de la familia de anticuerpos, caracterizándose los vectores o paquetes genéticos por secuencias de ADN variegadas que codifican una RDC2 de cadena pesada seleccionada del grupo que consiste en:

(1) <2>I<2><3>SGG<1>T<1>YADSVKG, 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 cada uno de los restos de aminoácidos Y, R, W, V, G y S; y <3> es una mezcla equimolar de cada uno de los restos de aminoácidos P, S y G o una mezcla equimolar de P y S;

(2) <1>I<4><1><1><G><5><1><1><1>YADSVKG, 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; <4> es una mezcla equimolar de los restos de D, I, N, S, W, Y; y <5> es una mezcla equimolar de los restos de S, G, D y N;

(3) <1>I<4><1><1>G<5><1><1>YNPSLKG, 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 <4> y <5> son como se definieron anteriormente;

(4) <1>I<8>S<1><1><1>GGYY<1>YAASVKG, 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 <8> es 0,27 R y 0,027 de cada uno de ADEFGHIKLMNPQSTVWY; y

(5) mezclas de vectores o de paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores.

4. La biblioteca focalizada según el aspecto 3, en la que las RDC2 de HC (1)/(2) (equimolar), (3) y (4) están presentes en la biblioteca en una relación 0,54:0,43:0,03.

5. Una biblioteca focalizada de vectores o paquetes genéticos que presenta, presenta y expresa, o comprende un miembro de una familia diversa de péptidos, polipéptidos o proteínas relacionados con anticuerpos humanos,

y colectivamente presenta, presenta y expresa, o comprende por lo menos una parte de la diversidad de la familia de anticuerpos, caracterizándose los vectores o paquetes genéticos por secuencias de ADN variegadas que codifican una RDC3 de cadena pesada seleccionada del grupo que consiste en:

(1) 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;

(2) 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;

(3) YYCA211111111YFDYWG, 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;

(4) 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;

(5) 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;

(6) 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;

(7) 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;

(8) 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

(9) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores.

6. La biblioteca focalizada según el aspecto 5, en la que 1 en una o todas las RDC3s de HC (1) a (8) 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.

7. La biblioteca focalizada según los aspectos 5 ó 6, en la que las RDC3 de HC (1) a (8) están presentes en la biblioteca en las proporciones siguientes:

(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.

8. La biblioteca focalizada según los aspectos 5 ó 6, en la que las RDC3 de HC (1) a (8) están presentes en la biblioteca en las proporciones siguientes:

(1) 0,02

(2) 0,14

(3) 0,25

(4) 0,14

(5) 0,14

(6) 0,12

(7) 0,08 y

(8) 0,11.

9. Una biblioteca focalizada de vectores o paquetes genéticos que presenta, presenta y expresa, o comprende un miembro de una familia diversa de péptidos, polipéptidos o proteínas relacionados con anticuerpos humanos, y colectivamente presenta, presenta y expresa, o comprende por lo menos una parte de la diversidad de la familia de anticuerpos, caracterizándose los vectores o paquetes genéticos por secuencias de ADN variegadas que codifican una RDC1 de cadena ligera kappa seleccionada del grupo que consiste en:

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

(2) 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 uno de ADEFGHIKLMNPQRTVWY; y <3> es 0,2 Y y 0,044 cada uno de ADEFGHIKLMNPQRSTVW; y (3) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores.

10. La biblioteca focalizada del aspecto 9, en la que las RDC1 (1) y (2) están presentes en la biblioteca en una relación de 0,68:0,32.

11. Una biblioteca focalizada de vectores o paquetes genéticos que presenta, presenta y expresa, o comprende un miembro de una familia diversa de péptidos, polipéptidos o proteínas relacionados con anticuerpos humanos, y colectivamente presenta, presenta y expresa, o comprende por lo menos una parte de la diversidad de la familia de anticuerpos, caracterizándose los vectores o paquetes genéticos por secuencias de ADN variegadas que codifican una RDC2 de cadena ligera kappa que tiene 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 uno de ADEFGHIKLMNPQRTVWY; y <4> es 0,2 A y 0,044 cada uno de DEFGHIKLMNPQRSTVWY.

12. Una biblioteca focalizada de vectores o paquetes genéticos que presenta, presenta y expresa, o comprende un miembro de una familia diversa de péptidos, polipéptidos o proteínas relacionados con anticuerpos humanos, y colectivamente presenta, presenta y expresa, o comprende por lo menos una parte de la diversidad de la familia de anticuerpos, caracterizándose los vectores o paquetes genéticos por secuencias de ADN variegadas que codifican una RDC3 de cadena ligera kappa seleccionada del grupo que consiste en:

(1) 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 uno de ADEFGHIKLMNPQRTVW;

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

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

(4) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores.

13. La biblioteca focalizada del aspecto 12, en la que las RDC3 (1), (2) y (3) están presentes en la biblioteca en una relación de 0,65:0,1:0,25.

14. Una biblioteca focalizada de vectores o paquetes genéticos que presenta, presenta y expresa, o comprende un miembro de una familia diversa de péptidos, polipéptidos o proteínas relacionados con anticuerpos humanos, y colectivamente presenta, presenta y expresa, o comprende por lo menos una parte de la diversidad de la familia de anticuerpos, caracterizándose los vectores o paquetes genéticos por secuencias de ADN variegadas que codifican una RDC1 de cadena ligera lambda seleccionada del grupo que consiste en:

(1) 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 uno de ADEFHIKLMNPQRSVWY, <2> es 0,27 D, 0,27 N y 0,027 cada uno de AEFGHIKLMNPQRSTVWY y <3> es 0,36 Y y 0,036 cada uno de ADEFGHIKLMNPQRSTVW;

(2) G<2><4>L<4><4><4><3><4><4>

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

(3) mezclas de vectores o paquetes genéticos caracterizados por alguna de las secuencias de ADN anteriores.

15. La biblioteca focalizada del aspecto 14, en la que las RDC1 (1) y (2) están presentes en la biblioteca en una relación de 0,67:0,33.

16. Una biblioteca focalizada de vectores o paquetes genéticos que presenta, presenta y expresa, o comprende un miembro de una familia diversa de péptidos, polipéptidos o proteínas relacionados con anticuerpos humanos, y colectivamente presenta, presenta y expresa, o comprende por lo menos una parte de la diversidad de la familia de anticuerpos, caracterizándose los vectores o paquetes genéticos por secuencias de ADN variegadas que codifican 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 uno de AEFGHIKLMPQRSTVWY, y <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVW.

17. Una biblioteca focalizada de vectores o paquetes genéticos que presenta, presenta y expresa, o comprende un miembro de una familia diversa de péptidos, polipéptidos o proteínas relacionados con anticuerpos humanos, y colectivamente presenta, presenta y expresa, o comprende por lo menos una parte de la diversidad de la familia de anticuerpos, caracterizándose los vectores o paquetes genéticos por secuencias de ADN variegadas que codifican una RDC3 de cadena ligera lambda seleccionada del grupo que consiste en:

(1) <4><5><4><2><4>S<4><4><4><4>V, en la que <2> es 0,27 D, 0,27 N y 0,027 cada uno de AEFGHIKLMPQRSTVWY; <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVW; y <5> es 0,36 S y 0,0355 cada uno de ADEFGHIKLMNPQRTVWY;

(2) <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

(3) mezclas de vectores o de paquetes genéticos caracterizados por algunas de las secuencias de ADN anteriores.

18. La biblioteca focalizada según el aspecto 17, en la que las RDC3 (1) y (2) están presentes en la biblioteca en una mezcla equimolar.

19. La biblioteca focalizada según el aspecto 1 ó 2, que comprende además secuencias de ADN variegadas que codifican una RDC de cadena pesada seleccionada del grupo que consiste en:

(1) una o más de las RDC2 de cadena pesada definidas en el aspecto 3 ó 4;

(2) una o más de las RDC3 de cadena pesada definidas en los aspectos 5, 6, 7, ó 8; y

(3) mezclas de vectores o paquetes genéticos caracterizadas por (1) y (2).

20. La biblioteca focalizada según el aspecto 3, que comprende además secuencias de ADN variegadas que codifican una o más de las RDC3 de cadena pesada seleccionadas del grupo definido en los aspectos 5, 6, 7, ó 8.

21. La biblioteca focalizada según el aspecto 19 ó 20, que comprende además secuencias de ADN variegadas que codifican una RDC de cadena ligera seleccionada del grupo que consiste en

(1) una o más de las RDC1 de cadena ligera kappa definidas en el aspecto 9 ó 10;

(2) la RDC2 de cadena ligera kappa definida en el aspecto 11;

(3) una o más de las RDC3 de cadena ligera kappa definidas en el aspecto 12 ó 13;

(4) una o más de las RDC1 de cadena ligera kappa definidas en el aspecto 14 ó 15;

(5) la RDC2 de cadena ligera lambda definida en el aspecto 16;

(6) una o más de las RDC3 de cadena ligera lambda definidas en el aspecto 17 ó 18; y

(7) mezclas de vectores o paquetes genéticos caracterizadas por uno o más de (1) a (6).

22. Una población de secuencias de ADN variegadas que codifica una RDC1 de cadena pesada seleccionada del grupo que consiste en:

(1) $<1>_1Y_2<1>_3M_4<1>_5$, en el 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) $(S/T)_1(S/G/X)_2(S/G/X)_3Y_4Y_5W_6(S/G/X)_7$, 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;

(3) $V_1S_2G_3G_4S_5I_6S_7$<1>$_8$<1>$_9$<1>$_{10}Y_{11}Y_{12}W_{13}$<1>$_{14}$, en los 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 (4) mezclas de secuencias de ADN variegadas caracterizadas por alguna de las secuencias de ADN anteriores.

23. La población de secuencias de ADN variegadas según el aspecto 22, en la que las RDC1 de HC (1), (2) y (3) están presentes en la población en la relación 0,80:0,17:0,02.

24. Una población de secuencias de ADN variegadas que codifica una RDC2 de cadena pesada seleccionada del grupo que consiste en:

(1) <2>I<2><3>SGG<1>T<1>YADSVKG, 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 cada uno de los restos de aminoácidos Y, R, W, V, G y S; y <3> es una mezcla equimolar de cada uno de los restos de aminoácidos P, S y G o una mezcla equimolar de P y S;

(2) <1>I<4><1><1><G><5><1><1><1>YADSVKG, 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; <4> es una mezcla equimolar de los restos de D, I, N, S, W, Y; y <5> es una mezcla equimolar de los restos de S, G, D y N;

(3) <1>I<4><1><1>G<5><1><1>YNPSLKG, 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 <4> y <5> son como se definieron anteriormente;

(4) <1>I<8>S<1><1><1>GGYY<1>YAASVKG, 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 <8> es 0,27 R y 0,027 de cada uno de ADEFGHIKLMNPQSTVWY; y (5) mezclas de secuencias de ADN variegadas caracterizadas por alguna de las secuencias de ADN anteriores.

25. La población de secuencias de ADN variegadas según el aspecto 24, en la que está presente una mezcla de las RDC2 de HC (1)/(2) (equimolar), (3) y (4) en la población en una relación de 0,54:0,43:0,03.

26. Una población de secuencias de ADN variegadas que codifica una RDC3 de cadena pesada seleccionada del grupo que consiste en

(1) 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;

(2) 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;

(3) YYCA211111111YFDYWG, 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;

(4) 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;

(5) 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;

(6) 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;

(7) 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;

(8) 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 (9) mezclas de secuencias de ADN variegadas caracterizadas por alguna de las secuencias de ADN anteriores.

27. La población de secuencias de ADN variegadas según el aspecto 26, en la que 1 en una o todas las RDC3s de HC (1) a (8) 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.

28. La población de secuencias de ADN variegadas según los aspectos 26 ó 27, en la que las RDC3 de HC (1) a (8) están presentes en la población en las proporciones siguientes:

(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.

29. La población de secuencias de ADN variegadas según los aspectos 26 ó 27, en la que las RDC3 de HC (1) a (8) están presentes en la población en las proporciones siguientes:

(1) 0,02

(2) 0,14

(3) 0,25

(4) 0,14

(5) 0,14

(6) 0,12

(7) 0,08 y

(8) 0,11.

30. Una población de secuencias de ADN variegadas que codifica una RDC1 de cadena ligera kappa seleccionada del grupo que consiste en:

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

(2) 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 uno de ADEFGHIKLMNPQRTVWY; y <3> es 0,2 Y y 0,044 cada uno de ADEFGHIKLMNPQRSTVW; y

(3) mezclas de secuencias de ADN variegadas caracterizadas por alguna de las secuencias de ADN anteriores.

31. La población de secuencias de ADN variegadas del aspecto 30, en las que las RDC1 (1) y (2) están presentes en la población en una relación de 0,68:0,32.

32. Una población de secuencias de ADN variegadas que codifica una RDC2 de cadena ligera kappa que tiene 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 uno de ADEFGHIKLMNPQRTVWY; y <4> es 0,2 A y 0,044 de cada uno de DEFGHIKLMNPQRSTUWY.

33. Una población de secuencias de ADN variegadas que codifica una RDC3 de cadena ligera kappa seleccionadas de los grupos que consisten en:

(1) 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 uno de ADEFGHIKLMNPQRTVW;

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

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

(4) mezclas de secuencias de ADN variegadas caracterizadas por alguna de las secuencias de ADN anteriores.

34. La población de secuencias de ADN variegadas según el aspecto 33, en la que las RDC3 (1), (2) y (3) están presentes en la población en una relación de 0,65:0,1:0,25.

35. Una población de secuencias de ADN variegadas que codifica una RDC1 de cadena ligera lambda seleccionada del grupo que consiste en:

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

(2) G<2><4>L<4><4><4><3><4><4>, en la que <2> es como se definió en (1) anteriormente, y <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVWY; y (3) mezclas de secuencias de ADN variegadas caracterizadas por alguna de las secuencias de ADN anteriores.

36. La población de secuencias de ADN variegadas según el aspecto 35, en la que las RDC1 (1) y (2) están presentes en la población en una relación de 0,67:0,33.

37. Una población de secuencias de ADN variegadas que codifica una RDC2 de cadena ligera lambda tiene la secuencia: <4><4><4><2>RPS, en la que <2> es 0,27 D, 0,27 N, y 0,027 cada uno de AEFGHIKLMPQRSTVWY, y <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVW.

38. Una población de secuencias de ADN variegadas que codifica una RDC3 de cadena ligera lambda seleccionada del grupo que consiste en:

(1) <4><5><4><2><4>S<4><4><4><4>V, en la que <2> es 0,27 D, 0,27 N y 0,027 cada uno de AEFGHIKLMPQRSTVWY; <4> es una mezcla equimolar de restos de aminoácidos ADEFGHIKLMNPQRSTVW; y <5> es 0,36 S y 0,0355 cada uno de ADEFGHIKLMNPQRTVWY;

(2) <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

(3) mezclas de secuencias de ADN variegadas caracterizadas por algunas de las secuencias de ADN anteriores.

39. La población de secuencias de ADN variegadas según el aspecto 38, en la que las RDC3 (1) y (2) están presentes en la población en una mezcla equimolar.

40. La población de secuencias de ADN variegadas según el aspecto 22 ó 23, que comprende además secuencias de ADN variegadas que codifican una RDC de cadena pesada seleccionada del grupo que consiste en:

(1) una o más de las RDC2 de cadena pesada definidas en el aspecto 24ó 25;

(2) una o más de las RDC3 de cadena pesada definidas en el aspecto 26, 27, 28 ó 29; y

(3) mezclas de secuencias de ADN variegadas caracterizadas por (1) y (2).

41. La población de secuencias de ADN variegadas según el aspecto 24, que comprende además secuencias de ADN variegadas que codifican una o más RDC3 de cadena pesada seleccionadas del grupo definido en los aspectos 26, 27, 28 ó 29.

42. La población de secuencias de ADN variegadas según el aspecto 40 ó 41, que comprende además secuencias de ADN variegadas que codifican una RDC de cadena ligera seleccionada del grupo que consiste en

(1) una o más de las RDC1 de cadena ligera kappa definidas en el aspecto 30 ó 31;

(2) la RDC2 de cadena ligera kappa definida en el aspecto 32;

(3) una o más de las RDC3 de cadena ligera kappa definidas en el aspecto 33 ó 34;

(4) una o más de las RDC1 de cadena ligera kappa definidas en el aspecto 35 ó 36;

(5) la RDC2 de cadena ligera lambda definida en el aspecto 37;

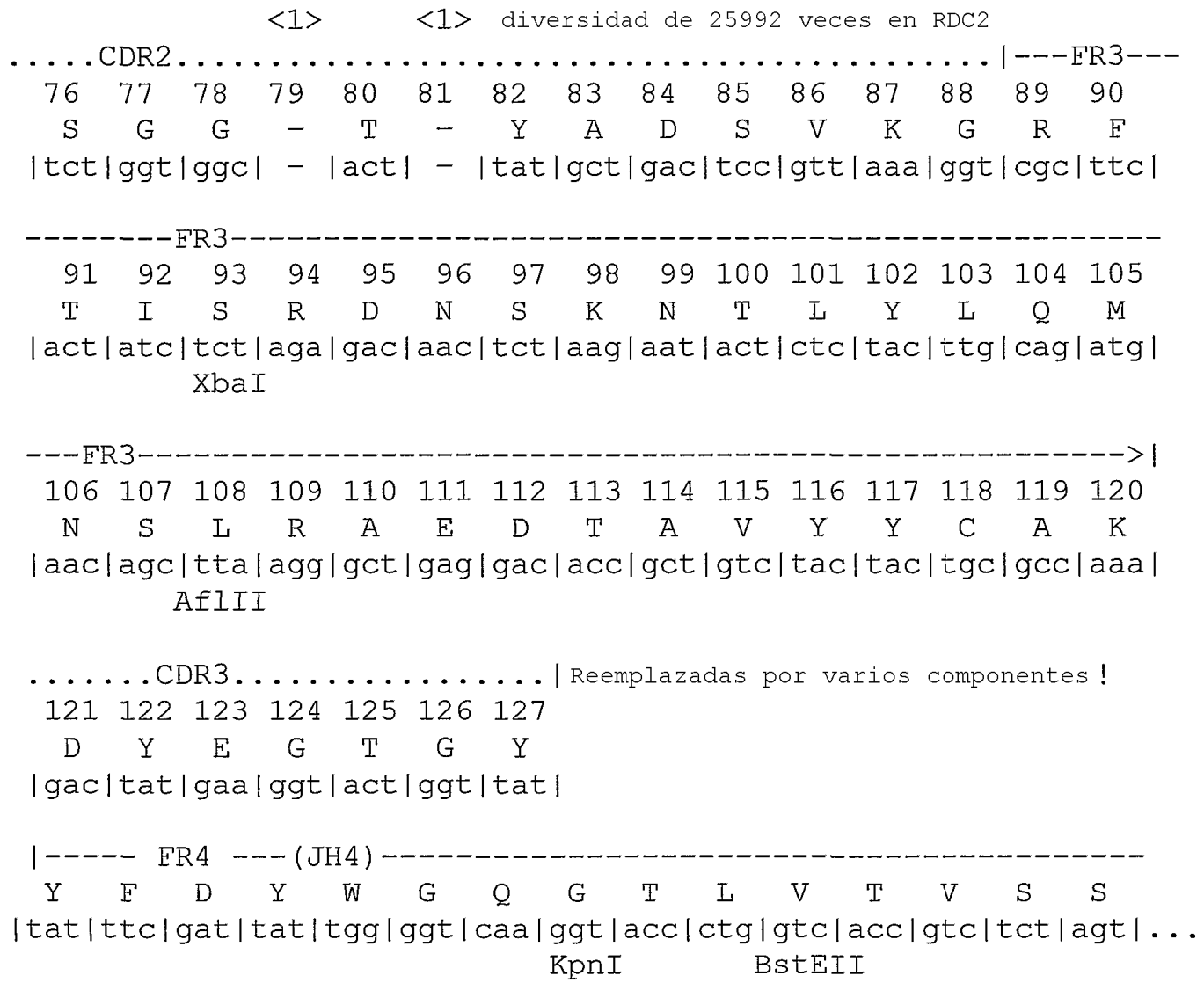
(6) una o más de las RDC3 de cadena ligera lambda definidas en el aspecto 38 ó 39; y

(7) mezclas de secuencias de ADN variegadas caracterizadas por uno o más de (1) a (6).

43. Una población de vectores que comprende las secuencias de ADN variegadas de uno cualquiera de los aspectos 22-42.

Tabla 1: 3-23:JH4 CDR1/2 diversidad = 1,78 x 10^8

```
                                    FR1(VP47/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|gtt|cag|cct|ggt|ggt|tct|tta|cgt|ctt|tct|tgc|gct|

     Puntos de variegación          <1>      <1> <1> <1>                diversidad de 6859 veces
      ----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|gtt|cgc|
             BspEI                   BsiWI                      BstXI.

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

<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>|<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 ADEFHIKLMNPQRTVWY)

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'-ggt|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 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'-ggt|ttg|gag|tgg|gtt|tct|<1>|atc|<4>|<1>|<1>|ggt|<5>|<1>|<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'-ggt|ttg|gag|tgg|gtt|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'-ggt|ttg|gag|tgg|gtt|tct|<1>|atc|<8>|agt|<1>|<1>|
<1>|ggt|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'

<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

Componente		Longitud	Complejidad	Fracción de genoteca	Fracción ajustada preferida
1	YYCA21111YFDYWG. (1=cualquier resto de aminoácidos, excepto C; 2 = K y R)	8	$2,6 \times 10^5$	,10	,02
2	YYCA2111111YFDYWG. (1=cualquier resto de aminoácidos, excepto C; 2 = K y R)	10	$9,4 \times 10^7$	,14	,14
3	YYCA211111111YFDYWG. (1=cualquier resto de aminoácidos, excepto C; 2 = K y R)	12	$3,4 \times 10^{10}$	,25	,25
4	YYCAR111S2S3111YFDYWG. (1=cualquier resto de aminoácidos, excepto C; 2 = S y G 3 = Y y W)	14	$1,9 \times 10^8$	,13	,14
5	YYCA2111CSG11CY1YFDYWG. (1=cualquier resto de aminoácidos, excepto C; 2 = K y R)	15	$9,4 \times 10^7$	,13	,14
6	YYCA211S1TIFG11111YFDYWG. (1=cualquier resto de aminoácidos, excepto C; 2 = K y R)	17	$1,7 \times 10^{10}$	,11	,12
7	YYCAR111YY2S33YY111YFDYWG. (1=cualquier resto de aminoácidos, excepto C; 2 = D o G; 3 = S y G)	18	$3,8 \times 10^8$	,04	,08
8	YYCAR1111YC2231CY111YFDYWG. (1=cualquier resto de aminoácidos, excepto C; 2 = S y G; 3 = T, D y G)	19	$2,0 \times 10^{11}$	,10	,11

Tabla 5: Oligonucleótidos utilizados para abigarrar los ocho componentes de RDC3 con HC

(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|tgg|ggc|caa|ggt|acc|ctg|gtc|acc|tcgctccacc-3'
BstEII...

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

Las 20 bases en el extremo 3' de CtprmA son idénticas a la mayoría de las 20 bases en 5' de cada una de las moléculas de vgADN

Ctop25 es idéntica a la mayoría de las 25 bases en 5' de CtprmA

La mayoría de las 23 bases en 3' de CBprmB son el complemento inverso de la mayoría de las 23 bases en 3' de cada una de las moléculas vgADN

CBot25 es idéntico a las 25 bases en el extremo 5' de CBprmB

Componente 1

(C1t08):

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

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

Componente 2

(C2t10):

5'-cc|gct|gtc|tac|tac|tgc|gcc|<2>|<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 3

(C3t12):

5'-cc|gct|gtc|tac|tac|tgc|gcc|<2>|<1>|<1>|<1>|<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 4

(C4t140):

5'-cc|gct|gtc|tac|tac|tgc|gcc|cgt|<1>|<1>|<1>|tct|<2>|tct|<3>|<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> = Y y W (mezcla equimolar)

Componente 5

(C5t15):

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

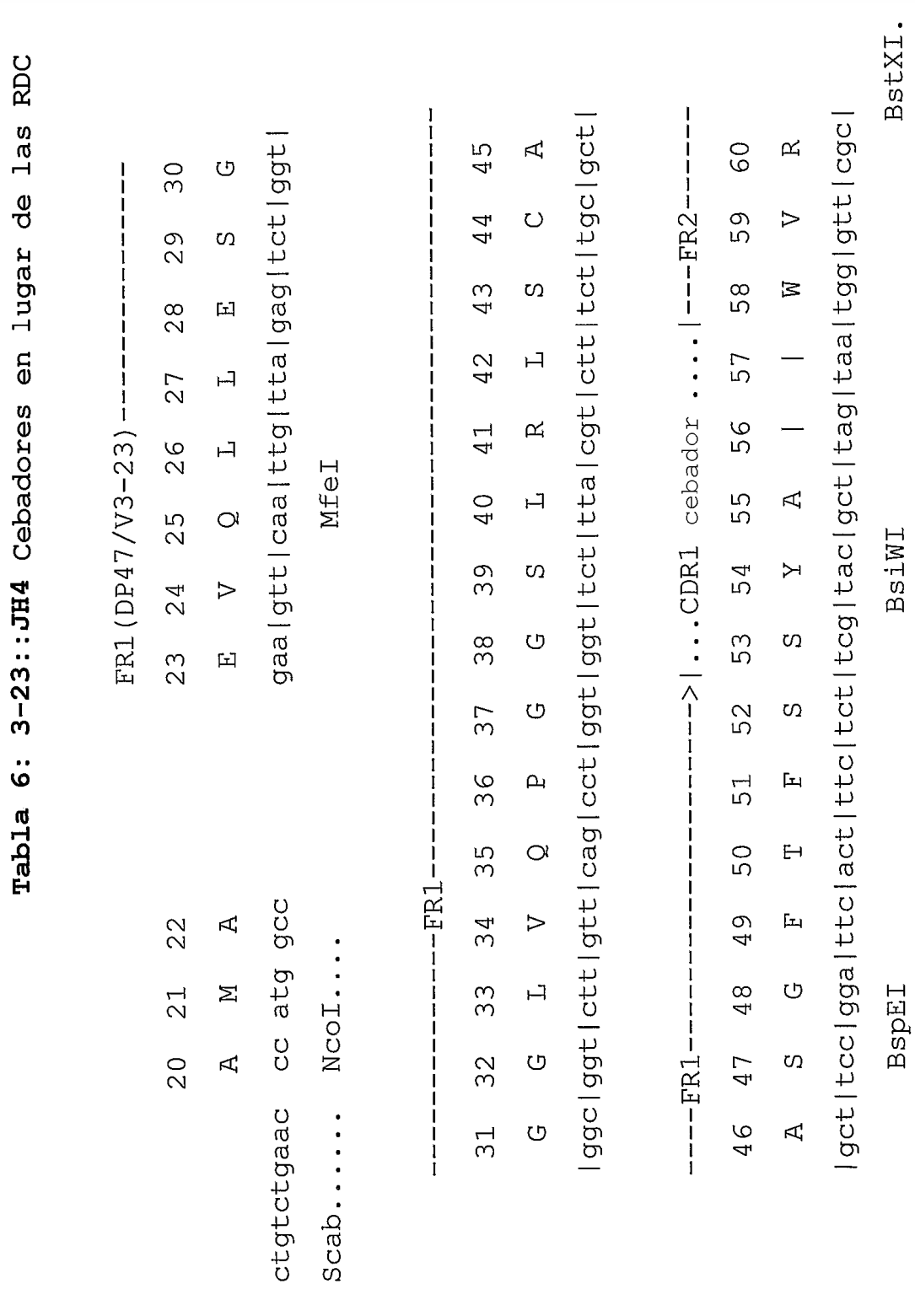
<1> = 0,095 Y + 0,095 G + 0,048 cada ADEFHIKLMNPQRSTVW, sin C; <2> = D y G (mezcla equimolar); <3> = S and G (mezcla equimolar)

Componente 8
(c8t19):

5'-cc|gct|gtc|tac|tac|tgc|gcc|cgt|<1>|<1>|<1>|<1>|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

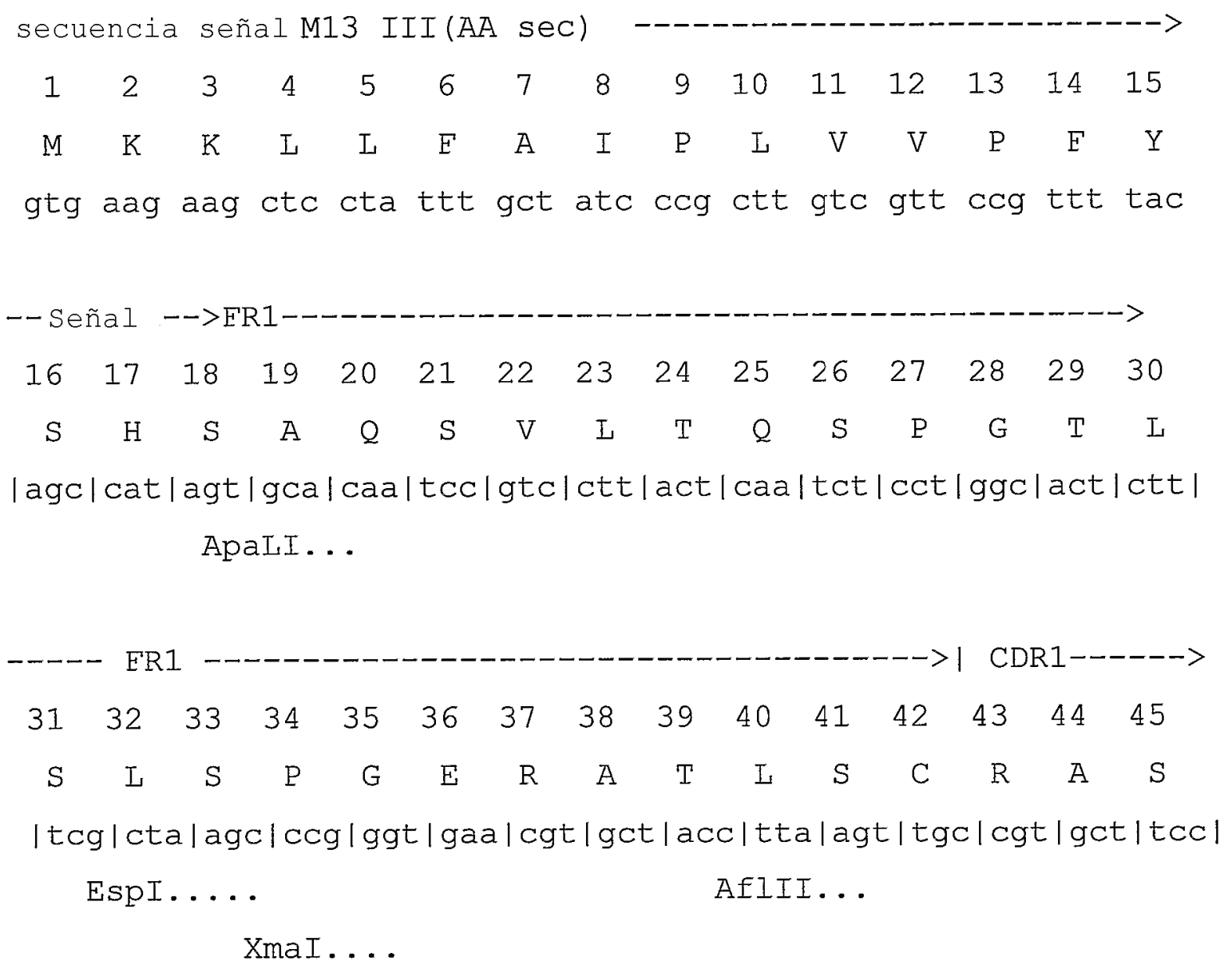


```
-------FR2--------------------------------->|...CDR2 cebador.
61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
 Q   A   P   G   K   G   L   E   W   V   S   |   p   r   |
|caa|gct|cct|ggt|aaa|ggt|ttg|gag|tgg|gtt|tct|taa|cct|agg|tag|
...BstXI                                          AvrII..
.....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 ctcactcatt aggcacccca 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 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|
ApaLI...
----- FR1 ->| CDR1------>
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
S L S P G E R A T L S C R A S
|tcg|cta|agc|ccg|ggt|gaa|cgt|gct|acc|tta|agt|tgc|cgt|gct|tcc|
EspI.....
AflII...
XmaI....

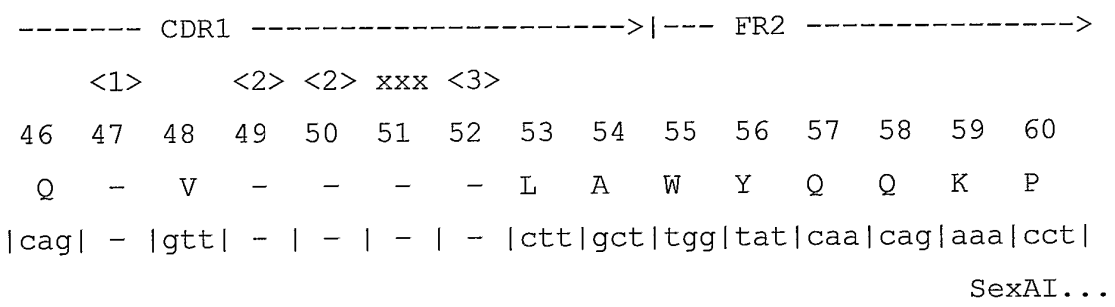
Para CDR1:

<1> ADEFGHIKLMNPQRSTVWY 1:1

<2> S(0,2) ADEFGHIKLMNPQRTVWY (0,044 cada uno)

<3> Y(0,2) ADEFGHIKLMNPQRSTVW (0,044 cada uno)

(RDC1 instalada como casete Af1II-(SexAI o KasI). Para la longitud 11 más preferida se omite el codón 51 (XXX); para la longitud 12 preferida este codón es <2>



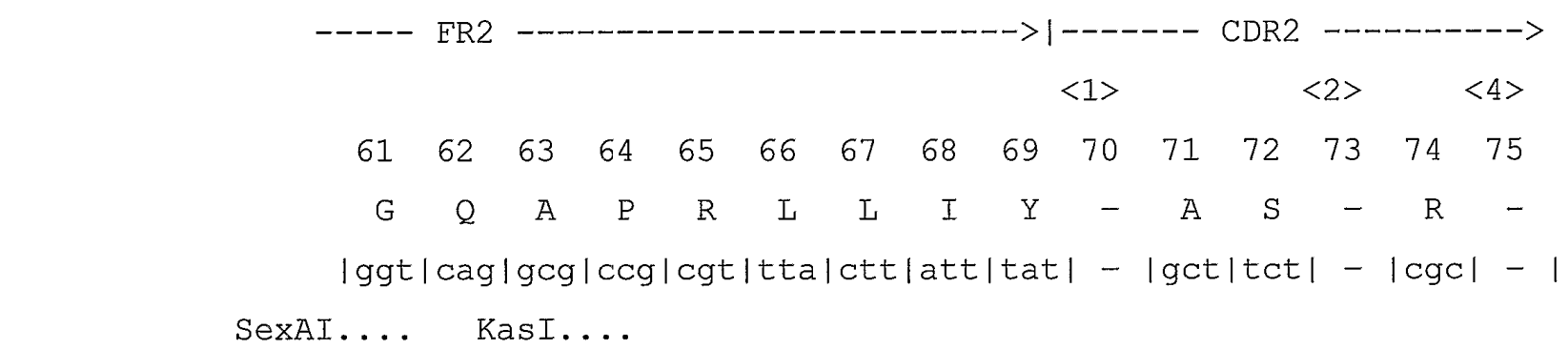
Para CDR2:

<1> ADEFGHIKLMNPQRSTVWY 1:1

<2> S(0,2) ADEFGHIKLMNPQRTVWY (0,044 cada uno)

<4> A(0,2) DEFGHIKLMNPQRSTVWY (0,044 cada uno)

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



```
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|ccg|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|gtt|
                            XbaI...
```

Para CDR3 (longitud 9):

<1> ADEFGHIKLMNPQRSTVWY 1:1

<3> Y(0,2) ADEFGHIKLMNPQRTVW (0,044 cada uno)

Para RDC3 (longitud 8): QQ33111P

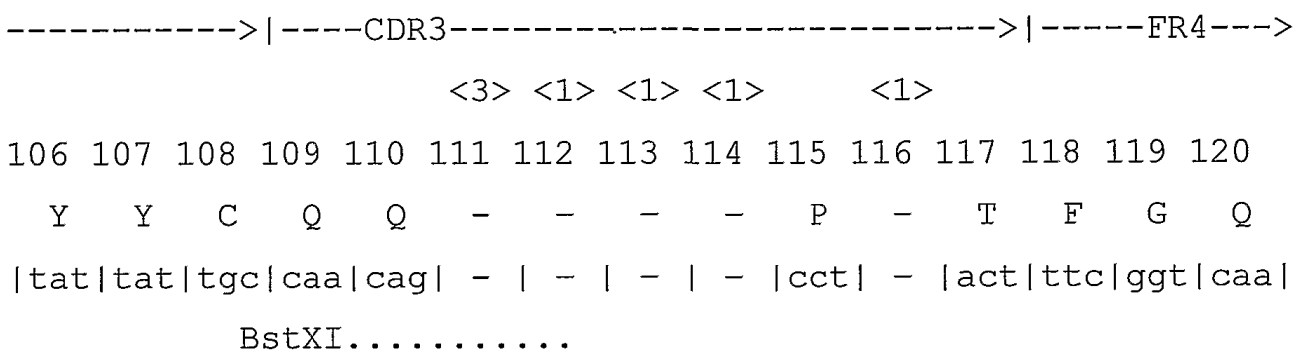
1 y 3 como se definió para la longitud 9

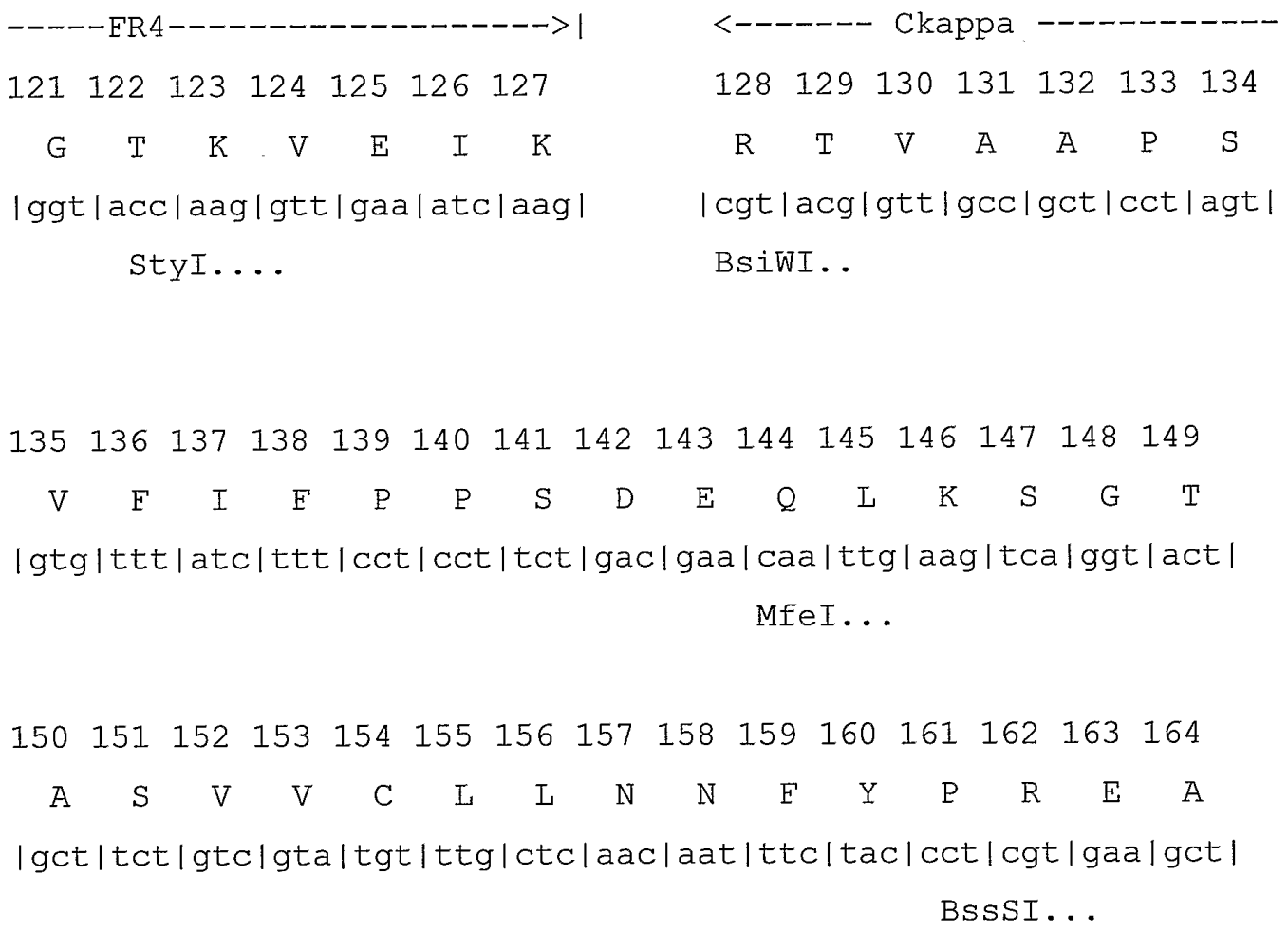
Para RDC3 (longitud 10): QQ3211PP1T

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)




-----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|gtt|gaa|atc|aag|
|cgt|acg|gtt|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   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

acgcatctctaa gcggccgc aacaggaggag
            NotI....
```

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

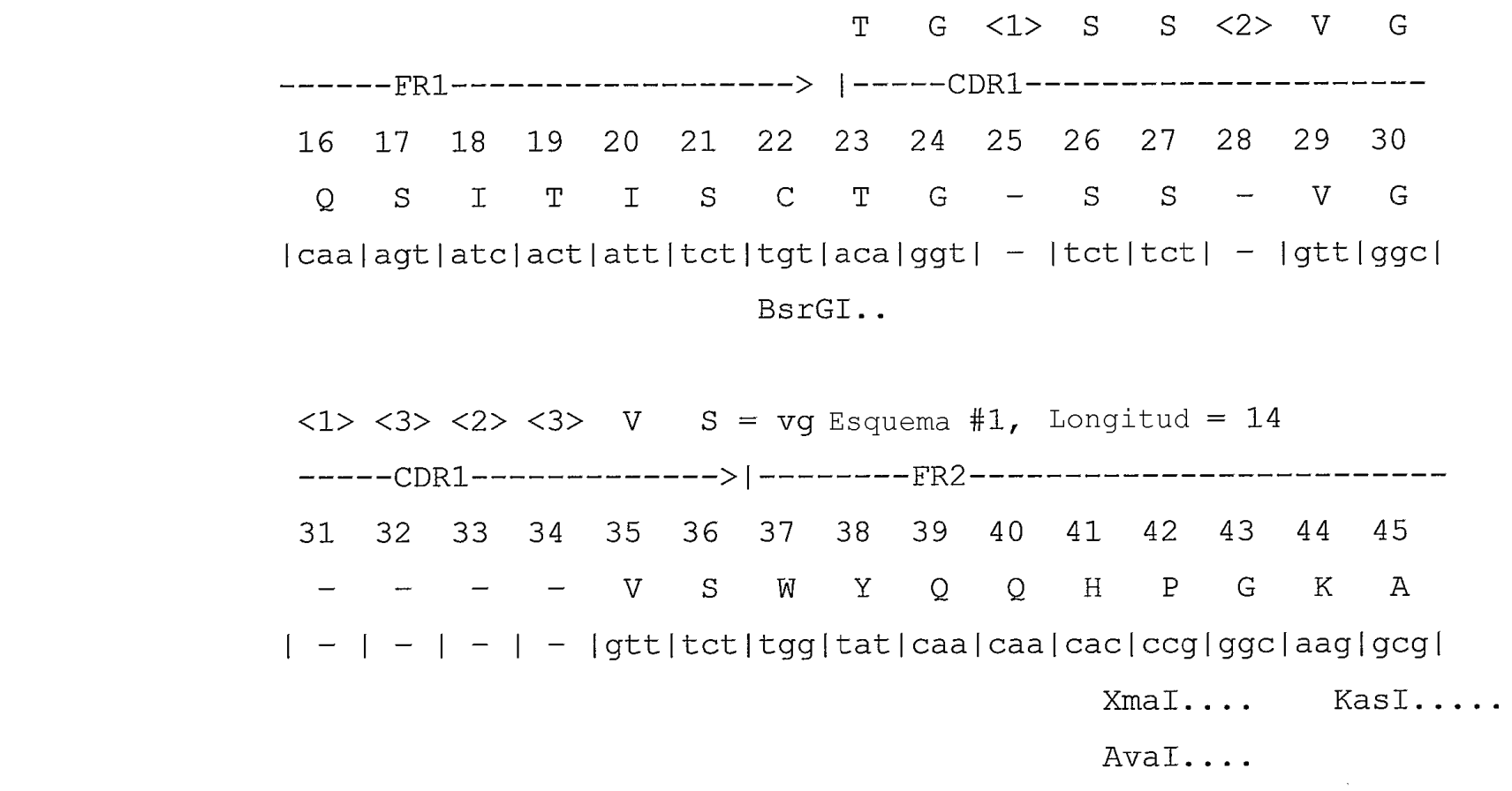
```
gaggaccatt gggcccc   ttactccgtgac
Scab...... EcoO109I
           ApaI..

         -----------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   S   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 AEFGHIKLMPQRSTVWY, sin C

<3> = 0,36 Y, 0,0355 cada uno de ADEFGHIKLMNPQRSTVW, sin C



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

T_{22}G<2><4>L<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|gtt|
KasI....

     -------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|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|ggt|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

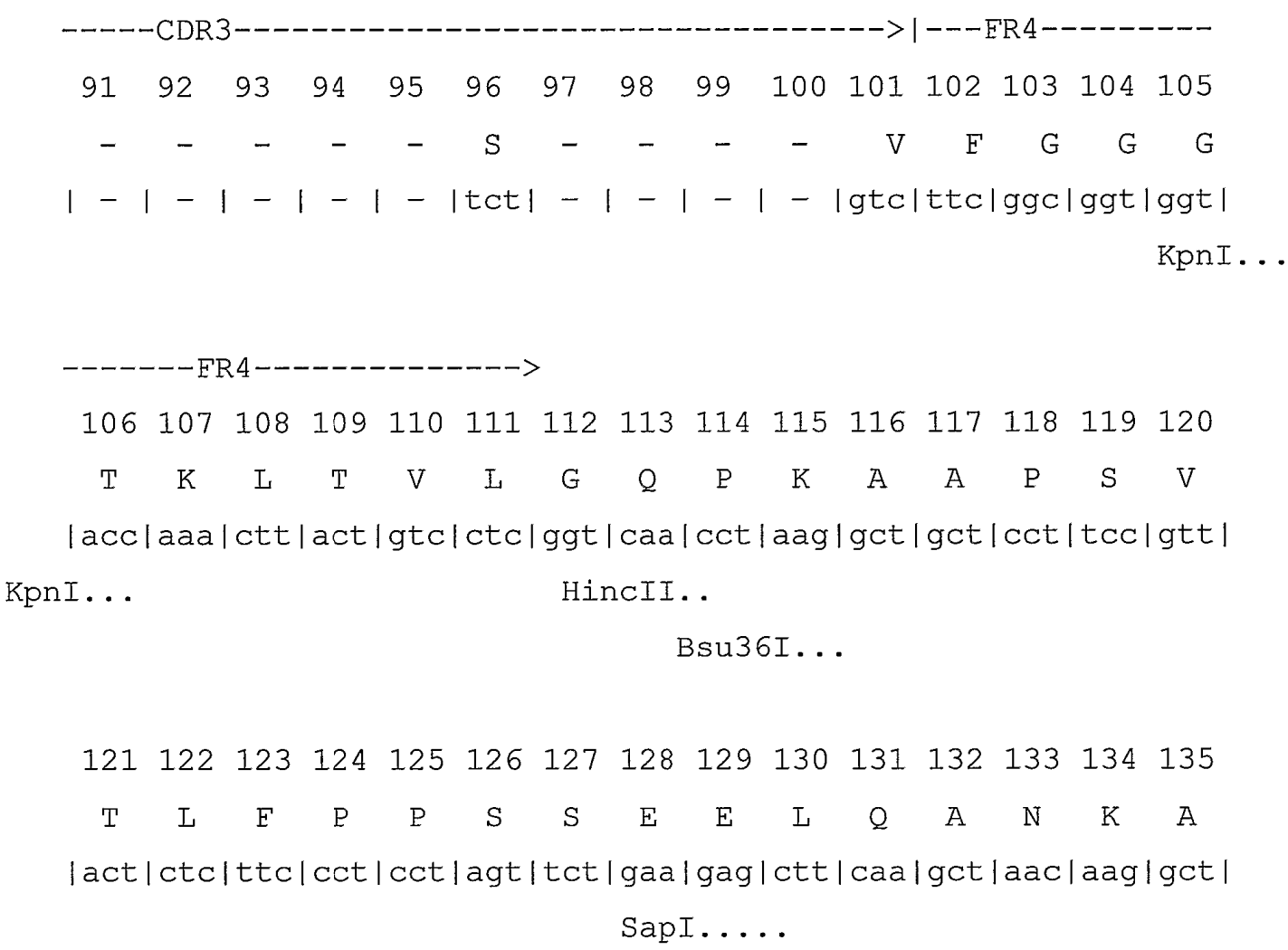
<5> = 0,36 S, 0,0355 cada uno de ADEFGHIKLMNPQRTVWY sin C

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

<1> mezcla equimolar de cada uno de ADEFGHIKLMNPQRSTVWY, 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


-----CDR3----------->|---FR4---------
91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
- - - - - S - - - - V F G G G
| - | - | - | - | - |tct| - | - | - | - |gtc|ttc|ggc|ggt|ggt|
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|ggt|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|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|gtt|act|cat|gaa|ggt|tct|acc|gtt|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|ggt|acc|ctg|gtc|acc|tcgctccacc-3'
BstEII...
(CBot25)[RC]: 5'-|ggt|acc|ctg|gtc|acc|tcgctccacc-3'

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

CDR1
(Ka1Top610): 5'-ggtctcagttg|cta|agc|ccg|ggt|gaa|cgt|gct|acc|tta|agt|tgc|cgt|gct|tcc|cag-3'
(Ka1STp615): 5'-ggtctcagttg|cta|agc|ccg|ggt|g-3'
(Ka1Bot620) [RC]: 5'-ctt|gct|tgg|tat|caa|cag|aaa|cct|ggt|cag|gcg|ccaagtcgtgtc-3'
(Ka1SB625) [RC]: 5'-cct|ggt|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|ggt|cag|gc-3'
(Ka2Tlong655): 5'-cacgagtccta|cct|ggt|cag|gcg|ccg|cgt|tta|ctt|att|tat-3'
(Ka2Bshort660):[RC]: 5'-|gac|cgt|ttc|tct|ggt|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|ggt|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|ggt|caa|ggt|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>|<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>|<1>|cct|cct|<1>|act|ttc|ggt|caa|-
|ggt|acc|aag|gtt|g-3'

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

CDR1

(Lm1TPri75): 5'-gacgagtcctgg|tca|cct|ggt|-3'
(Lm1tlo715): 5'-gacgagtcctgg|tca|cct|ggt|caa|agt|atc|act|att|tct|tgt|aca|ggt-3'
(Lm1blo724) [rc]: 5'-gtt|tct|tgg|tat|caa|caa|cac|ccg|ggc|aag|gcg|agatcttcacaggtgag-3'
(Lm1bsh737) [rc]: 5'-gc|aag|gcg|agatcttcacaggtgag-3'
(Lm1vg710b):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'

(Lm1vg710): 5'-gt|atc|act|att|tct|tgt|aca|ggt|<1>|tct|tct|<2>|gtt|ggc|-
|<1>|<3>|<2>|<3>|gtt|tct|tgg|tat|caa|caa|cac|cc-3'

CDR2

(Lm2TSh757): 5'-gagcagaggac|ccg|ggc|aag|gc-3'
(Lm2TLo753): 5'-gagcagaggac|ccg|ggc|aag|gcg|ccg|aag|ttg|atg|atc|tac|-3'
(Lm2BLo762) [RC]: 5'-cgt|cct|tct|ggt|gtc|agc|aat|cgt|ttc|tcc|gga|tcacaggtgag-3'
(Lm2BSh765) [RC]: 5'-cgt|ttc|tcc|gga|tcacaggtgag-3'
(Lm2vg750): 5'-g|ccg|aag|ttg|atg|atc|tac|-
<4>|<4>|<4>|<2>|cgt|cct|tct|ggt|gtc|agc|aat|c-3'

CDR3

(Lm3TSh822): 5'-ctg|cag|gct|gaa|gac|gag|gct|gac-3'
(Lm3TLo819): 5'-ctg|cag|gct|gaa|gac|gag|gct|gac|tac|tat|tgt|-3'
(Lm3BLo825) [RC]: 5'-gtc|ttc|ggc|ggt|ggt|acc|aaa|ctt|act|gtc|ctc|ggt|caa|cct|aag|g-
acacaggtgag-3'
(Lm3BSh832) [RC]: 5'-c|ggt|caa|cct|aag|gacacaggtgag-3'

(Lm3vg817): 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'

(Lm3vg817-10): 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.....Eco0109I

ApaI.

XhoI..

acgcaattaa tgtgagttag ctcactcatt aggcacccca ggctttacac tttatgcttc

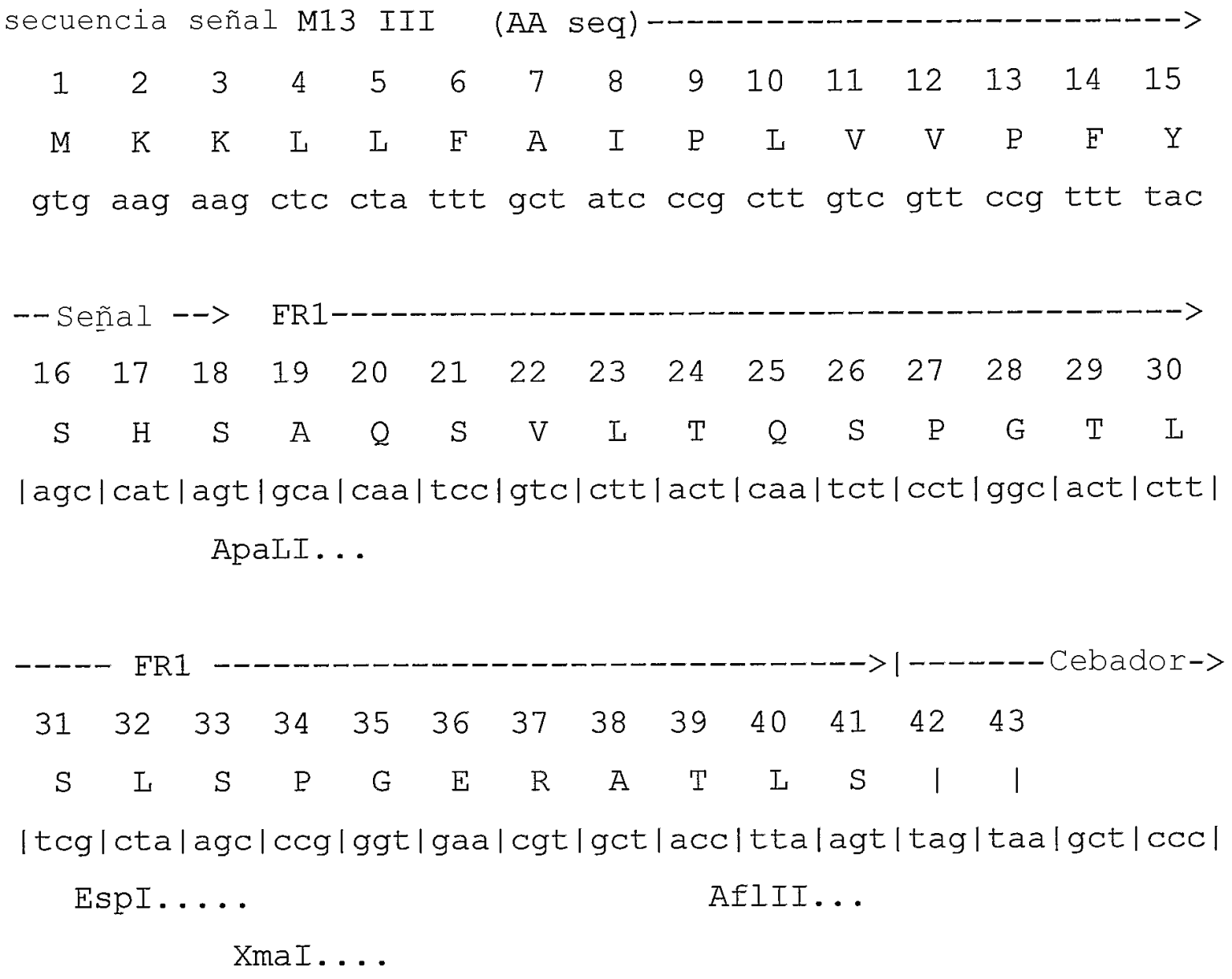
..-35.. Plac ..-10.

cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatgac

catgatta cgccaagctt tggagccttt tttttggaga ttttcaac

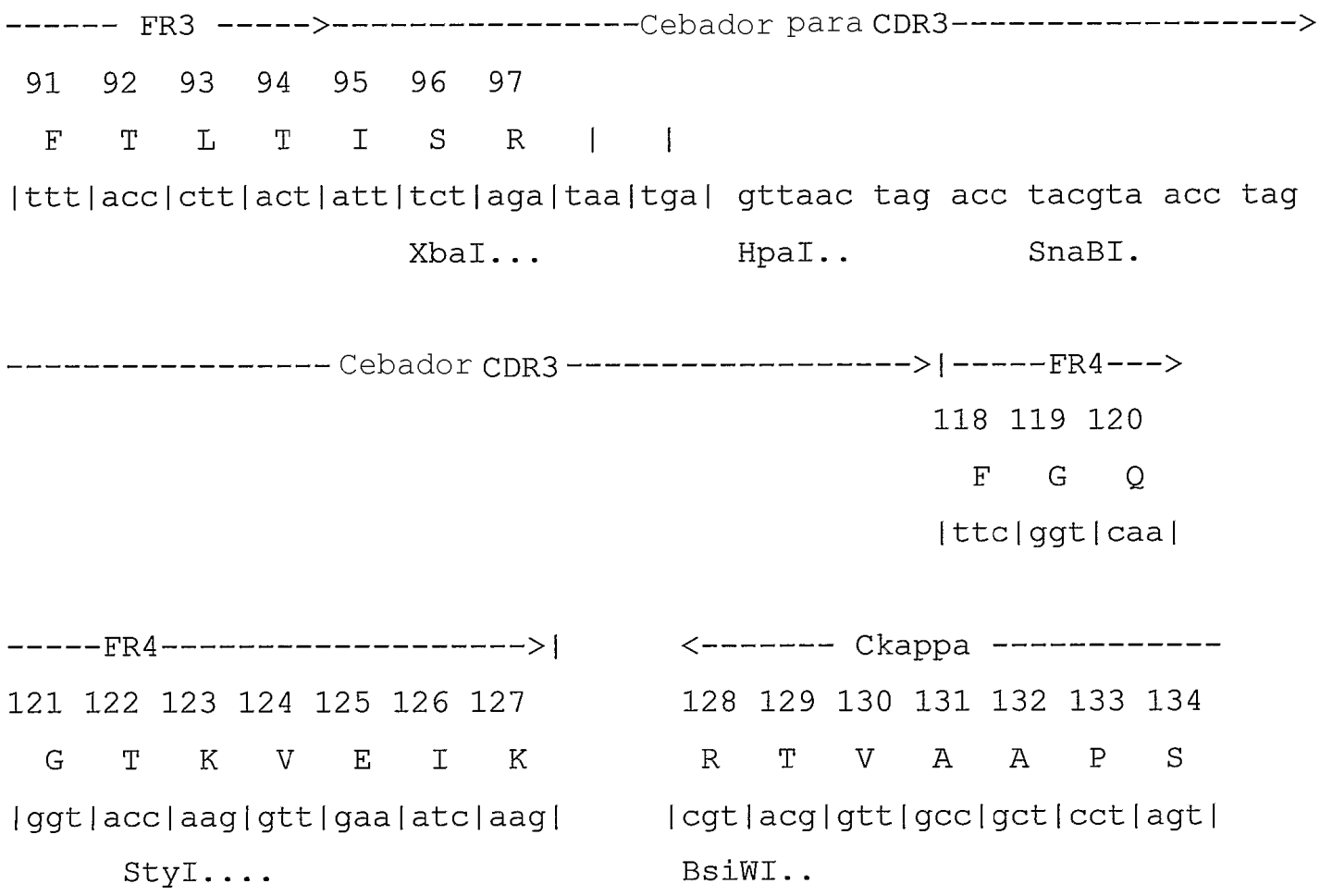
PflMI.....

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
|agc|cat|agt|gca|caa|tcc|gtc|ctt|act|caa|tct|cct|ggc|act|ctt|
ApaLI...
----- 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|ggt|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|ggt|cag|gcg|ccg|cgt|taa|tga|aagcgctaatggccaacagtg
 StuI...                 SexAI...    KasI....              AfeI..   MscI..
```

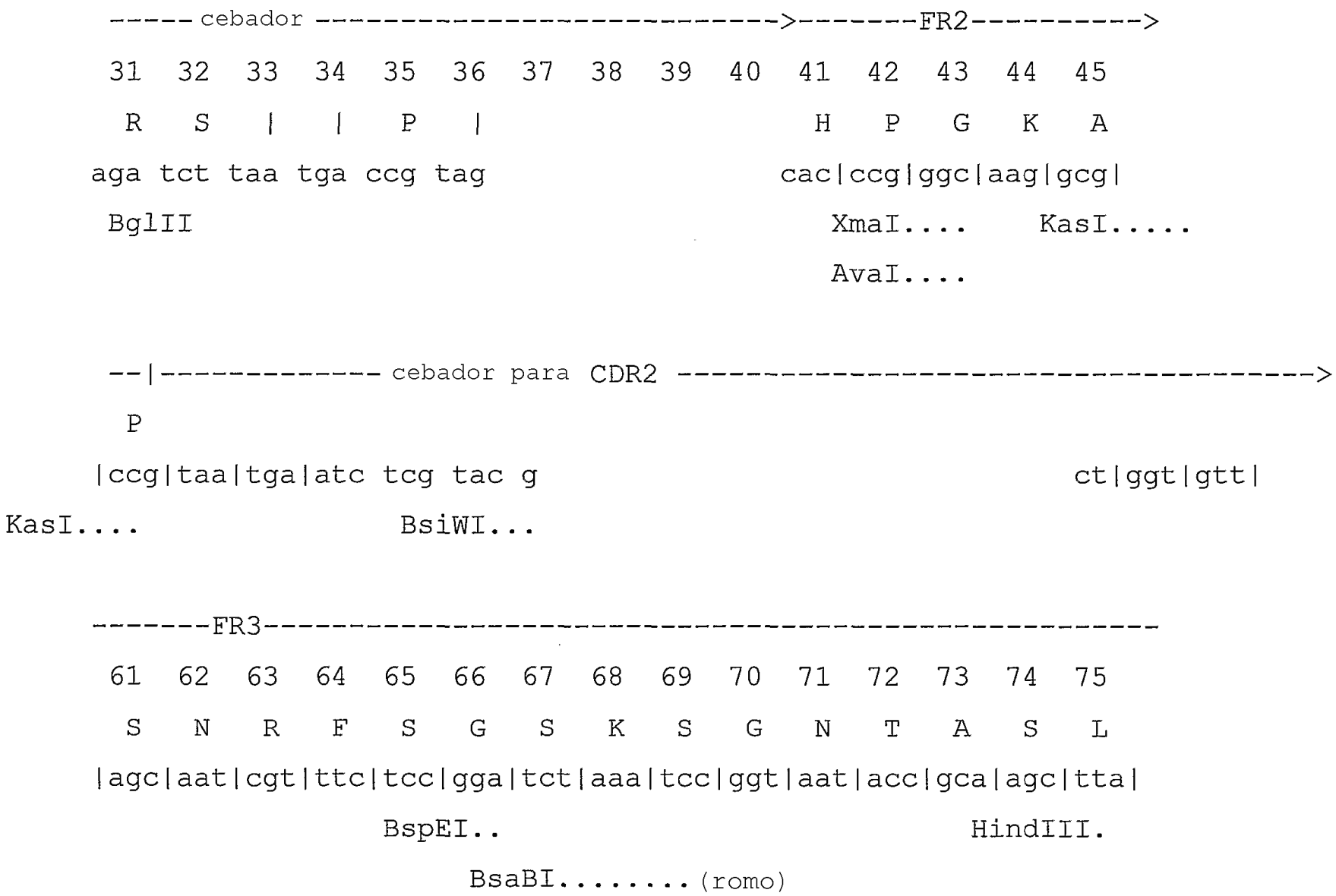
```
Cebador-->|--- FR3 ------------------------------------------------>

 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|ggt|tca|ggt|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|ggt|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
|ggt|acc|aag|gtt|gaa|atc|aag| |cgt|acg|gtt|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...

acgcatctctaa 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   S   P   G
agt|gca|caa|tcc|gct|ctc|act|cag|cct|gct|agc|gtt|tcc|ggg|tca|cct|ggt|
 ApaLI...                        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.... KasI.....
AvaI....
cebador para CDR2
P
|ccg|taa|tga|atc tcg tac g ct|ggt|gtt|
KasI.... 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|ggt|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|ggt|ctg|cag|gtt ctg tag ttc caattg ctt tag tga ccc
                 PstI...                MfeI..

----- cebador ------------------------------->|---FR4---------
                                                 103 104 105
                                                  G   G   G
                                                |ggc|ggt|ggt|
                                                          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|ggt|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|gtt|tgc|ttg|atc|agt|gac|ttt|tat|cct|ggt|gct|gtt|act|
                    BclI....
```

LISTADO DE SECUENCIAS

<110> DYAX CORPORATION

<120> BIBLIOTECAS FOCALIZADAS DE PAQUETES GENÉTICOS

<130> DYAX/004PCT

<140> PCT/US01/50297

<141> 18/12/2001

<160> 99

<170> PatentIn Ver. 2.1

<210> 1

<211> 14

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC1 de cadena pesada

<220>

<221 > MOD_RES

<222> (8)..(10)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (14)

<223> Cualquier aminoácido salvo Cys

<400> 1

```
Val Ser Gly Gly Ser Ile Ser Xaa Xaa Xaa Tyr Tyr Trp Xaa
  1               5                  10
```

<210> 2

<211> 17

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC2 de cadena pesada

<220>

<221 > MOD_RES

<222> (1)

<223> Tyr, Arg, Trp, val Gly o Ser

<220>

<221 > MOD_RES

<222> (3)

<223> Tyr, Arg, Trp, Val, Gly o Ser

<220>

<221 > MOD_RES

<222> (4)

<223> Pro, Ser o Gly

<220>

<221 > MOD_RES

<222> (8)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (10)

<223> Cualquier aminoácido salvo Cys

<400> 2

Xaa Ile Xaa Xaa Ser Gly Gly Xaa Thr Xaa Tyr Ala Asp Ser Val Lys
1 5 10 15

Gly

<210> 3

<211> 17

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC2 de cadena pesada

<220>

<221 > MOD_RES

<222> (1)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (3)

<223> Asp, Ile, Asn, Ser, Trp o Tyr

<220>

<221 > MOD_RES

<222> (4)..(5)

<223> Cualquier aminoácido salvo cys

<220>

<221 > MOD_RES

<222> (7)

<223> Ser, Gly, Asp o Asn

<220>

<221 > MOD_RES

<222> (8)..(10)

<223> Cualquier aminoácido salvo Cys

<400> 3

Xaa Ile Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Tyr Ala Asp Ser Val Lys
1 5 10 15

Gly

<210> 4

<211> 16

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC2 de cadena pesada

<220>

<221 > MOD_RES

<222> (1)

<223> Cualquier aminoácido salvo cys

<220>

<221 > MOD_RES

<222> (3)

<223> Asp, Ile, Asn, Ser, Trp o Tyr

<220>

<221 > MOD_RES

<222> (4)..(5)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (7)

<223> Ser, Gly, Asp o Asn

<220>

<221 > MOD_RES

<222> (8)..(9)

<223> Cualquier aminoácido salvo Cys

<400> 4

Xaa Ile Xaa Xaa Xaa Gly Xaa Xaa Xaa Tyr Asn Pro Ser Leu Lys Gly
1 5 10 15

<210> 5

<211> 19

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC2 de cadena pesada

<220>

<221 > MOD_RES

<222> (1)

<223> Cualquier aminoácido salvo cys

<220>

<221 > MOD_RES

<222> (3)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (5)..(7)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (12)

<223> Cualquier aminoácido salvo Cys

<400> 5

Xaa Ile Xaa Ser Xaa Xaa Xaa Gly Gly Tyr Tyr Xaa Tyr Ala Ala Ser
1 5 10 15

Val Lys Gly

<210> 6

<211> 15

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC3 de cadena pesada

<220>

<221 > MOD_RES

<222> (5)

<223> Lys o Arg

<220>

<221 > MOD_RES

<222> (6)..(9)

<223> Cualquier aminoácido salvo Cys

<400> 6

```
Tyr Tyr Cys Ala Xaa Xaa Xaa Xaa Xaa Tyr Phe Asp Tyr Trp Gly
  1               5                  10                  15
```

<210> 7

<211> 17

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC3 de cadena pesada

<220>

<221 > MOD_RES

<222> (5)

<223> Lys o Arg

<220>

<221 > MOD_RES

<222> (6)..(11)

<223> Cualquier aminoácido salvo Cys

<400> 7

```
Tyr Tyr Cys Ala Xaa Xaa Xaa Xaa Xaa Xaa Xaa Tyr Phe Asp Tyr Trp
  1               5                  10                  15

Gly
```

<210> 8

<211> 19

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC3 de cadena pesada

<220>

<221 > MOD_RES

<222> (5)

<223> Lys o Arg

<220>

<221 > MOD_RES

<222> (6)..(13)

<223> Cualquier aminoácido salvo Cys

<400> 8

Tyr Tyr Cys Ala Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Tyr Phe Asp
1 5 10 15

Tyr Trp Gly

<210> 9

<211> 21

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC3 de cadena pesada

<220>

<221 > MOD_RES

<222> (6)..(8)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (10)

<223> Ser o Gly

<220>

<221 > MOD_RES

<222> (12)

<223> Tyr o Trp

<220>

<221 > MOD_RES

<222> (13)..(15)

<223> Cualquier aminoácido salvo Cys

<400> 9

Tyr Tyr Cys Ala Arg Xaa Xaa Xaa Ser Xaa Ser Xaa Xaa Xaa Xaa Tyr
1 5 10 15

Phe Asp Tyr Trp Gly
20

<210> 10

<211> 22

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC3 de cadena pesada

<220>

<221 > MOD_RES

<222> (5)

<223> Lys o Arg

<220>

<221 > MOD_RES

<222> (6)..(8)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (12)..(13)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (16)

<223> Cualquier aminoácido salvo Cys

<400> 10

Tyr Tyr Cys Ala Xaa Xaa Xaa Xaa Cys Ser Gly Xaa Xaa Cys Tyr Xaa
1 5 10 15

Tyr Phe Asp Tyr Trp Gly
20

<210> 11

<211> 24

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC3 de cadena pesada

<220>

<221 > MOD_RES

<222> (5)

<223> Lys o Arg

<220>

<221 > MOD_RES

<222> (6)..(7)

<223> Cualquier aminoácido salvo cys

<220>

<221 > MOD_RES

<222> (9)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (14)..(18)

<223> Cualquier aminoácido salvo Cys

<400> 11

```
Tyr Tyr Cys Ala Xaa Xaa Xaa Ser Xaa Thr Ile Phe Gly Xaa Xaa Xaa
  1               5                  10                  15

Xaa Xaa Tyr Phe Asp Tyr Trp Gly
            20
```

<210> 12

<211> 25

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC3 de cadena pesada

<220>

<221 > MOD_RES

<222> (6)..(8)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (11)

<223> Asp o Gly

<220>

<221 > MOD_RES

<222> (13)..(14)

<223> Ser o Gly

<220>

<221 > MOD_RES

<222> (15)..(16)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (17)..(19)

<223> Cualquier aminoácido salvo cys

<400> 12

```
Tyr Tyr Cys Ala Arg Xaa Xaa Xaa Tyr Tyr Xaa Ser Xaa Xaa Xaa Xaa
  1               5                  10                  15

Xaa Xaa Xaa Tyr Phe Asp Tyr Trp Gly
            20                  25
```

<210> 13

<211> 26

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC3 de cadena pesada

<220>

<221 > MOD_RES

<222> (6)..(9)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (12)..(13)

<223> Ser o Gly

<220>

<221 > MOD_RES

<222> (14)

<223> Thr, Asp o Gly

<220>

<221 > MOD_RES

<222> (15)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (18)..(20)

<223> Cualquier aminoácido salvo Cys

<400> 13

```
Tyr Tyr Cys Ala Arg Xaa Xaa Xaa Xaa Tyr Cys Xaa Xaa Xaa Xaa Cys
  1               5                  10                  15

Tyr Xaa Xaa Xaa Tyr Phe Asp Tyr Trp Gly
            20                  25
```

<210> 14

<211> 11

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC1 de cadena ligera kappa

<220>

<221 > MOD_RES

<222> (5)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (7)..(8)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (9)

<223> Cualquier aminoácido salvo Cys y Ser

<400> 14

```
Arg Ala Ser Gln Xaa Val Xaa Xaa Xaa Leu Ala
  1               5                  10
```

<210> 15

<211> 12

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC1 de cadena ligera kappa

<220>

<221 > MOD_RES

<222> (5)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (7)..(9)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (10)

<223> Cualquier aminoácido salvo Cys y Ser

<400> 15

Arg Ala Ser Gln Xaa Val Xaa Xaa Xaa Xaa Leu Ala
1 5 10

<210> 16

<211> 9

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC1 de cadena ligear kappa

<220>

<221 > MOD_RES

<222> (3)

<223> Cualquier aminoácido salvo Cys o Ser

<220>

<221 > MOD_RES

<222> (4)..(6)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (8)

<223> Cualquier aminoácido salvo Cys

<400> 16

Gln Gln Xaa Xaa Xaa Xaa Pro Xaa Thr
1 5

<210> 17

<211> 10

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC3 de cadena ligera kappa

<220>

<221 > MOD_RES

<222> (3)

<223> Cualquier aminoácido salvo cys o Ser

<220>

<221 > MOD_RES

<222> (4)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (5)..(6)

<223> Cualquier aminoácido salvo cys

<220>

<221 > MOD_RES

<222> (9)

<223> Cualquier aminoácido salvo Cys

<400> 17

```
Gln Gln Xaa Xaa Xaa Xaa Pro Pro Xaa Thr
  1               5                  10
```

<210> 18

<211> 14

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC1 de cadena ligera lambda

<220>

<221 > MOD_RES

<222> (3)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (6)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (9)..(12)

<223> Cualquier aminoácido salvo Cys

<400> 18

Thr Gly Xaa Ser Ser Xaa Val Gly Xaa Xaa Xaa Xaa Val Ser
1 5 10

<210> 19

<211> 10

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Vector de RDC3 de cadena ligera lambda

<220>

<221 > MOD_RES

<222> (1)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (4)..(5)

<223> Cualquier aminoácido salvo cys

<220>

<221 > MOD_RES

<222> (7)..(8)

<223> Cualquier aminoácido salvo Cys

<220>

<221 > MOD_RES

<222> (9)

<223> Cualquier aminoácido salvo cys o Tyr

<400> 19

Xaa Ser Tyr Xaa Xaa Ser Xaa Xaa Xaa Val
1 5 10

<210> 20

<211> 14

<212> PRT

<213> Homo sapiens

<400> 20

Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Ser Ser
1 5 10

<210> 21

<211> 15

<212> PRT

<213> Homo sapiens

<400> 21

Ala Phe Asp Ile Trp Gly Gln Gly Thr Met Val Thr Val Ser Ser
1 5 10 15

<210> 22

<211> 5

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

<400> 22

Tyr Tyr Cys Ala Arg
1 5

<210> 23

<211> 323

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: 3-23: Vector de JH4 con diversidad de RDC1/2

<220>

<221 > CDS

<222> (3)..(323)

<220>

<221 > base modificada

<222> (99)..(101)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (105)..(107)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (111)..(113)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (156)..(158)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (162)..(167)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (177)..(179)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (183)..(185)

<223> a, c, t, g, otro o desconocido

<400> 23

```
 cc atg gcc gaa gtt caa ttg tta gag tct ggt ggc ggt ctt gtt cag    47
Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln
 1               5                  10                  15

cct ggt ggt tct tta cgt ctt tct tgc gct gct tcc gga ttc act ttc   95
Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe
            20                  25                  30

tct nnn tac nnn atg nnn tgg gtt cgc caa gct cct ggt aaa ggt ttg   143
Ser Xaa Tyr Xaa Met Xaa Trp Val Arg Gln Ala Pro Gly Lys Gly Leu
        35                  40                  45

gag tgg gtt tct nnn atc nnn nnn tct ggt ggc nnn act nnn tat gct   191
Glu Trp Val Ser Xaa Ile Xaa Xaa Ser Gly Gly Xaa Thr Xaa Tyr Ala


    50                  55                  60

gac tcc gtt aaa ggt cgc ttc act atc tct aga gac aac tct aag aat   239
Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn
65                  70                  75                  80

act ctc tac ttg cag atg aac agc tta agg gct gag gac acc gct gtc   287
Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val
                85                  90                  95

tac tac tgc gcc aaa gac tat gaa ggt act ggt tat                  323
Tyr Tyr Cys Ala Lys Asp Tyr Glu Gly Thr Gly Tyr
            100                 105
```

<210> 24

<211> 108

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: 3-23: Vector de JH4 con diversidad RDC1/2

<220>

<221 > MOD_RES

<222> (34)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (36)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (38)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (53)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (55)..(56)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (60)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (62)

<223> Cualquier aminoácido

<400> 24

```
Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln
  1               5                  10                  15

Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe
            20                  25                  30

Ser Xaa Tyr Xaa Met Xaa Trp Val Arg Gln Ala Pro Gly Lys Gly Leu
        35                  40                  45


Glu Trp Val Ser Xaa Ile Xaa Xaa Ser Gly Gly Xaa Thr Xaa Tyr Ala
    50                  55                  60

Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn
65                  70                  75                  80

Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val
                85                  90                  95

Tyr Tyr Cys Ala Lys Asp Tyr Glu Gly Thr Gly Tyr
            100                 105
```

<210> 25

<211> 45

<212> ADN

<213> Homo sapiens

<220>

<221 > CDS

<222> (1)..(45)

<400> 25

```
tat ttc gat tat tgg ggt caa ggt acc ctg gtc acc gtc tct agt       45
Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
  1               5                  10                  15
```

<210> 26

<211> 15

<212> PRT

<213> Homo sapiens

<400> 26

Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
1 5 10 15

<210> 27

<211> 55

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (21)..(23)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (27)..(29)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (33)..(35)

<223> a, c, t ó g

<400> 27

cttccggatt cactttctct nnntacnnna tgnnntgggt tcgccaagct cctgg 55

<210> 28

<211> 28

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 28

cctactgtct tccggattca ctttctct 28

<210> 29

<211> 30

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 29

tgggttcgcc aagctcctgg ttgctcactc 30

<210> 30

<211> 61

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (23)..(29)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (39)..(41)

<223> a, c, t ó g

<400> 30

cttccggatt cactttctct wsnnnnnnnt actactggnn ntgggttcgc caagctcctg 60
g 61

<210> 31

<211> 82

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (42)..(50)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (60)..(62)

<223> a, c, t ó g

<400> 31

cttccggatt cactttctct atcagcggtg gttctatctc cnnnnnnnnn tactactggn 60
nntgggttcg ccaagctcct gg 82

<210> 32

<211> 68

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (19)..(21)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (25)..(28)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (30)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (40)..(42)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (46)..(48)

<223> a, c, t ó g

<400> 32

ggtttggagt gggtttctnn natcnnnnsn tctggtggcn nnactnnnta tgctgactcc 60
gttaaagg 68

<210> 33

<211> 44

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 33

cttgggttcg ccaagctcct ggtaaaggtt tggagtgggt ttct 44

<210> 34

<211> 49

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 34

tatgctgact ccgttaaagg tcgcttcact atctctagat tcctgtcac 49

<210> 35

<211> 68

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (19)..(21)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (26)..(33)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (39)..(48)

<223> a, c, t ó g

<400> 35

ggtttggagt gggtttctnn natcdnnnnn nnnggtdvnn nnnnnnnnta tgctgactcc 60
gttaaagg 68

<210> 36

<211> 65

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (19)..(21)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (26)..(33)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (39)..(45)

<223> a, c, t ó g

<400> 36

ggtttggagt gggtttctnn natcdnnnnn nnnggtdvnn nnnnntataa cccttcccttaaggg 65

<210> 37

<211> 49

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 37

tataaccctt cccttaaggg tcgcttcact atctctagat tcctgtcac 49

<210> 38

<211> 74

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (19)..(21)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (25)..(27)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (31)..(39)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (52)..(54)

<223> a, c, t ó g

<400> 38

ggtttggagt gggtttctnn natcnnnagt nnnnnnnnng gtggtactac tnnntatgcc 60
gcttccgtta aggg 74

<210> 39

<211> 49

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 39

tatgccgctt ccgttaaggg tcgcttcact atctctagat tcctgtcac 49

<210> 40

<211> 25

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 40

gctctggtca acttaagggc tgagg 25

<210> 41

<211> 48

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 41

gctctggtca acttaagggc tgaggacacc gctgtctact actgcgcc 48

<210> 42

<211> 46

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 42

tacttcgatt actggggcca aggtaccctg gtcacctcgc tccacc 46

<210> 43

<211> 25

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 43

ggtaccctgg tcacctcgct ccacc 25

<210> 44

<211> 58

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (23)..(35)

<223> a, c, t ó g

<400> 44

ccgctgtcta ctactgcgcc mrnnnnnnnn nnnnntactt cgattactgg ggccaagg 58

<210> 45

<211> 64

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (23)..(41)

<223> a, c, t ó g

<400> 45

ccgctgtcta ctactgcgcc mrnnnnnnnn nnnnnnnnnn ntacttcgat tactggggcc 60
aagg 64

<210> 46

<211> 70

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (23)..(47)

<223> a, c, t ó g

<400> 46

ccgctgtcta ctactgcgcc mrnnnnnnnn nnnnnnnnnn nnnnnnntac ttcgattact 60
ggggccaagg 70

<210> 47

<211> 76

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (24)..(32)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (38)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (45)..(53)

<223> a, c, t ó g

<400> 47

ccgctgtcta ctactgcgcc cgtnnnnnnn nntctdsntc ttrbnnnnnn nnntacttcg 60
attactgggg ccaagg 76

<210> 48

<211> 79

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (23)..(32)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (42)..(47)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (54)..(56)

<223> a, c, t ó g

<400> 48

ccgctgtcta ctactgcgcc mrnnnnnnnn nntgctctgg tnnnnnntgc tatnnntact 60
tcgattactg gggccaagg 79

<210> 49

<211> 85

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (23)..(29)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (33)..(35)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (48)..(62)

<223> a, c, t ó g

<400> 49

```
ccgctgtcta ctactgcgcc mrnnnnnnnt ctnnnactat cttcggtnnn nnnnnnnnnn 60
nntacttcga ttactggggc caagg                                       85
```

<210> 50

<211> 88

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (24)..(32)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (41)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (47)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (50)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (57)..(65)

<223> a, c, t ó g

<400> 50

ccgctgtcta ctactgcgcc cgtnnnnnnn nntattacgr ntctdsndsn tactatnnnn 60
nnnnntactt cgattactgg ggccaagg 88

<210> 51

<211> 91

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (24)..(35)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (44)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (47)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (50)..(53)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (60)..(68)

<223> a, c, t ó g

<400> 51

ccgctgtcta ctactgcgcc cgtnnnnnnn nnnnntattg cdsndsnrvn nnntgctatn 60
nnnnnnnnta cttcgattac tggggccaag g 91

<210> 52

<211> 242

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: 3-23: Vector de JH4 con cebadores

<220>

<221 > CD5

<222> (3)..(107)

<220>

<221 > CDS

<222> (114)..(155)

<220>

<221 > CDS

<222> (159)..(164)

<220>

<221 > CDS

<222> (168)..(227)

<400> 52

```
 cc atg gcc gaa gtt caa ttg tta gag tct ggt ggc ggt ctt gtt cag    47
Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln
  1               5                  10                  15

cct ggt ggt tct tta cgt ctt tct tgc gct gct tcc gga ttc act ttc   95
Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe
            20                  25                  30

tct tcg tac gct tagtaa tgg gtt cgc caa gct cct ggt aaa ggt ttg   143
Ser Ser Tyr Ala        Trp Val Arg Gln Ala Pro Gly Lys Gly Leu
         35                  40                  45

gag tgg gtt tct taa cct agg tag act atc tct aga gac aac tct aag   191
Glu Trp Val Ser     Pro Arg     Thr Ile Ser Arg Asp Asn Ser Lys
            50                      55                  60

aat act ctc tac ttg cag atg aac agc tta agg gct tagtaaaggc cttaa  242
Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala
                65                  70
```

<210> 53

<211> 72

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: 3-23: Vector de JH4 con cebadores

<400> 53

```
Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln
 1               5                  10                  15

Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe
            20                  25                  30

Ser Ser Tyr Ala Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp
        35                  40                  45

Val Ser Pro Arg Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
    50                  55                  60

Leu Gln Met Asn Ser Leu Arg Ala
65                  70
```

<210> 54

<211> 952

<212> ADN

<213> Homo sapiens

<220>

<221 > CDS

<222> (206)..(907)

<220>

<221 > base modificada

<222> (344)..(346)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (350)..(361)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (413)..(415)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (422)..(424)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (428)..(433)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (536)..(547)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (551)..(553)

<223> a, c, t, g, otro o desconocido

<400> 54

```
gaggaccatt gggccccctc cgagactctc gagcgcaacg caattaatgt gagttagctc 60

actcattagg caccccaggc tttacacttt atgcttccgg ctcgtatgtt gtgtggaatt 120

gtgagcggat aacaatttca cacaggaaac agctatgacc atgattacgc caagctttgg 180

agccttttt ttggagattt tcaac gtg aag aag ctc cta ttt gct atc ccg   232
                          Met Lys Lys Leu Leu Phe Ala Ile Pro
                            1               5

ctt gtc gtt ccg ttt tac agc cat agt gca caa tcc gtc ctt act caa   280
Leu Val Val Pro Phe Tyr Ser His Ser Ala Gln Ser Val Leu Thr Gln
 10                  15                  20                  25

tct cct ggc act ctt tcg cta agc ccg ggt gaa cgt gct acc tta agt   328
Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser
                 30                  35                  40

tgc cgt gct tcc cag nnn gtt nnn nnn nnn nnn ctt gct tgg tat caa   376
Cys Arg Ala Ser Gln Xaa Val Xaa Xaa Xaa Xaa Leu Ala Trp Tyr Gln
             45                  50                  55

cag aaa cct ggt cag gcg ccg cgt tta ctt att tat nnn gct tct nnn   424
Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile Tyr Xaa Ala Ser Xaa
         60                  65                  70

cgc nnn nnn ggg atc ccg gac cgt ttc tct ggc tct ggt tca ggt act   472
Arg Xaa Xaa Gly Ile Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr
     75                  80                  85

gac ttt acc ctt act att tct aga ttg gaa cct gaa gac ttc gct gtt   520
Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu Pro Glu Asp Phe Ala Val
 90                  95                  100                 105


tat tat tgc caa cag nnn nnn nnn nnn cct nnn act ttc ggt caa ggt   568
Tyr Tyr Cys Gln Gln Xaa Xaa Xaa Xaa Pro Xaa Thr Phe Gly Gln Gly
                110                 115                 120

acc aag gtt gaa atc aag cgt acg gtt gcc gct cct agt gtg ttt atc   616
Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile
            125                 130                 135

ttt cct cct tct gac gaa caa ttg aag tca ggt act gct tct gtc gta   664
Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val
        140                 145                 150

tgt ttg ctc aac aat ttc tac cct cgt gaa gct aaa gtt cag tgg aaa   712
Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys
    155                 160                 165

gtc gat aac gcg ttg cag tcg ggt aac agt caa gaa tcc gtc act gaa   760
Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu
170                 175                 180                 185

cag gat agt aag gac tct acc tac tct ttg tcc tct act ctt act tta   808
Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu
                190                 195                 200

tca aag gct gat tat gag aag cat aag gtc tat gct tgc gaa gtt acc   856
Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr
            205                 210                 215

cac cag ggt ctg agc tcc cct gtt acc aaa agt ttc aac cgt ggt gaa   904
His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu
        220                 225                 230

tgc taatagggcg cgccacgcat ctctaagcgg ccgcaacagg aggag            952
Cys
```

<210> 55

<211> 234

<212> PRT

<213> Homo sapiens

<220>

<221 > MOD_RES

<222> (47)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (49)..(52)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (70)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (73)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (75)..(76)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (111)..(114)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (116)

<223> Cualquier aminoácido

<400> 55

```
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
 1               5                  10                  15

His Ser Ala Gln Ser Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu
            20                  25                  30

Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Xaa Val
        35                  40                  45

Xaa Xaa Xaa Xaa Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro
    50                  55                  60

Arg Leu Leu Ile Tyr Xaa Ala Ser Xaa Arg Xaa Xaa Gly Ile Pro Asp
65                  70                  75                  80

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser
                85                  90                  95

Arg Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Xaa Xaa
            100                 105                 110

Xaa Xaa Pro Xaa Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
        115                 120                 125

Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
    130                 135                 140

Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
145                 150                 155                 160

Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
                165                 170                 175

Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
            180                 185                 190

Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
        195                 200                 205

His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro
    210                 215                 220

Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
225                 230
```

<210> 56

<211> 732

<212> ADN

<213> Homo sapiens

<220>

<221 > CDS

<222> (30)..(686)

<220>

<221 > base modificada

<222> (108)..(110)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (117)..(119)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (126)..(137)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (189)..(200)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (306)..(320)

<223> a, c, t, g, otro o desconocido

<220>

<221 > base modificada

<222> (324)..(335)

<223> a, c, t, g, otro o desconocido

<400> 56

```
gaggaccatt gggcccctta ctccgtgac agt gca caa tcc gct ctc act cag   53
                                  Ser Ala Gln Ser Ala Leu Thr Gln
                                    1               5

cct gct agc gtt tcc ggg tca cct ggt caa agt atc act att tct tgt   101
Pro Ala Ser Val Ser Gly Ser Pro Gly Gln Ser Ile Thr Ile Ser Cys
     10                  15                  20

aca ggt nnn tct tct nnn gtt ggc nnn nnn nnn nnn gtt tct tgg tat   149
Thr Gly Xaa Ser Ser Xaa Val Gly Xaa Xaa Xaa Xaa Val Ser Trp Tyr
 25                  30                  35                  40

caa caa cac ccg ggc aag gcg ccg aag ttg atg atc tac nnn nnn nnn   197
Gln Gln His Pro Gly Lys Ala Pro Lys Leu Met Ile Tyr Xaa Xaa Xaa
                 45                  50                  55

nnn cgt cct tct ggt gtt agc aat cgt ttc tcc gga tct aaa tcc ggt   245
Xaa Arg Pro Ser Gly Val Ser Asn Arg Phe Ser Gly Ser Lys Ser Gly
             60                  65                  70

aat acc gca agc tta act atc tct ggt ctg cag gct gaa gac gag gct   293
Asn Thr Ala Ser Leu Thr Ile Ser Gly Leu Gln Ala Glu Asp Glu Ala
         75                  80                  85

gac tac tat tgt nnn nnn nnn nnn nnn tct nnn nnn nnn nnn gtc ttc   341
Asp Tyr Tyr Cys Xaa Xaa Xaa Xaa Xaa Ser Xaa Xaa Xaa Xaa Val Phe
     90                  95                  100

ggc ggt ggt acc aaa ctt act gtc ctc ggt caa cct aag gct gct cct   389
Gly Gly Gly Thr Lys Leu Thr Val Leu Gly Gln Pro Lys Ala Ala Pro
105                 110                 115                 120

tcc gtt act ctc ttc cct cct agt tct gaa gag ctt caa gct aac aag   437
Ser Val Thr Leu Phe Pro Pro Ser Ser Glu Glu Leu Gln Ala Asn Lys
                125                 130                 135

gct act ctt gtt tgc ttg atc agt gac ttt tat cct ggt gct gtt act   485
Ala Thr Leu Val Cys Leu Ile Ser Asp Phe Tyr Pro Gly Ala Val Thr
            140                 145                 150


gtc gct tgg aaa gcc gat tct tct cct gtt aaa gct ggt gtt gag acg   533
Val Ala Trp Lys Ala Asp Ser Ser Pro Val Lys Ala Gly Val Glu Thr
        155                 160                 165

acc act cct tct aaa caa tct aac aat aag tac gct gcg agc tct tat   581
Thr Thr Pro Ser Lys Gln Ser Asn Asn Lys Tyr Ala Ala Ser Ser Tyr
    170                 175                 180

ctt tct ctc acc cct gaa caa tgg aag tct cat aaa tcc tat tcc tgt   629
Leu Ser Leu Thr Pro Glu Gln Trp Lys Ser His Lys Ser Tyr Ser Cys
185                 190                 195                 200

caa gtt act cat gaa ggt tct acc gtt gaa aag act gtt gcc cct act   677
Gln Val Thr His Glu Gly Ser Thr Val Glu Lys Thr Val Ala Pro Thr
                205                 210                 215

gag tgt tct tagtgaggcg cgccaacgat gttcaaggcg gccgcaacag gaggag    732
Glu Cys Ser
```

<210> 57

<211> 219

<212> PRT

<213> Homo sapiens

<220>

<221 > MOD_RES

<222> (27)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (30)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (33)..(36)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (54)..(57)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (93)..(97)

<223> Cualquier aminoácido

<220>

<221 > MOD_RES

<222> (99)..(102)

<223> Cualquier aminoácido

<400> 57

```
Ser Ala Gln Ser Ala Leu Thr Gln Pro Ala Ser Val Ser Gly Ser Pro
 1               5                  10                  15

Gly Gln Ser Ile Thr Ile Ser Cys Thr Gly Xaa Ser Ser Xaa Val Gly
            20                  25                  30

Xaa Xaa Xaa Xaa Val Ser Trp Tyr Gln Gln His Pro Gly Lys Ala Pro
        35                  40                  45

Lys Leu Met Ile Tyr Xaa Xaa Xaa Xaa Arg Pro Ser Gly Val Ser Asn
    50                  55                  60


Arg Phe Ser Gly Ser Lys Ser Gly Asn Thr Ala Ser Leu Thr Ile Ser
65                  70                  75                  80

Gly Leu Gln Ala Glu Asp Glu Ala Asp Tyr Tyr Cys Xaa Xaa Xaa Xaa
                85                  90                  95

Xaa Ser Xaa Xaa Xaa Xaa Val Phe Gly Gly Gly Thr Lys Leu Thr Val
            100                 105                 110

Leu Gly Gln Pro Lys Ala Ala Pro Ser Val Thr Leu Phe Pro Pro Ser
        115                 120                 125

Ser Glu Glu Leu Gln Ala Asn Lys Ala Thr Leu Val Cys Leu Ile Ser
    130                 135                 140

Asp Phe Tyr Pro Gly Ala Val Thr Val Ala Trp Lys Ala Asp Ser Ser
145                 150                 155                 160

Pro Val Lys Ala Gly Val Glu Thr Thr Thr Pro Ser Lys Gln Ser Asn
                165                 170                 175

Asn Lys Tyr Ala Ala Ser Ser Tyr Leu Ser Leu Thr Pro Glu Gln Trp
            180                 185                 190

Lys Ser His Lys Ser Tyr Ser Cys Gln Val Thr His Glu Gly Ser Thr
        195                 200                 205

Val Glu Lys Thr Val Ala Pro Thr Glu Cys Ser
    210                 215
```

<210> 58

<211> 25

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 58

gctctggtca acttaagggc tgagg 25

<210> 59

<211> 48

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 59

gctctggtca acttaagggc tgaggacacc gctgtctact actgcgcc 48

<210> 60

<211> 46

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 60

tacttcgatt acttgggcca aggtaccctg gtcacctcgc tccacc 46

<210> 61

<211> 25

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 61

ggtaccctgg tcacctcgct ccacc 25

<210> 62

<211> 56

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 62

ggtctcagtt gctaagcccg ggtgaacgtg ctaccttaag ttgccgtgct tcccag 56

<210> 63

<211> 24

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 63

ggtctcagtt gctaagcccg ggtg 24

<210> 64

<211> 45

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 64

cttgcttggt atcaacagaa acctggtcag gcgccaagtc gtgtc 45

<210> 65

<211> 24

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 65

cctggtcagg cgccaagtcg tgtc 24

<210> 66

<211> 65

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (28)..(30)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (34)..(42)

<223> a, c, t ó g

<400> 66

gctaccttaa gttgccgtgc ttcccagnnn gttnnnnnnn nncttgcttg gtatcaacag 60
aaacc 65

<210> 67

<211> 68

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (28)..(30)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (34)..(45)

<223> a, c, t ó g

<400> 67

gctaccttaa gttgccgtgc ttcccagnnn gttnnnnnnn nnnnncttgc ttggtatcaa 60
cagaaacc 68

<210> 68

<211> 22

<212> ADN

<213> Secuencia artificial

<220>

<223> Description of Artificiai Sequence: oligonucleótido sintético

<400> 68

cacgagtcct acctggtcag gc 22

<210> 69

<211> 41

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 69

cacgagtcct acctggtcag gcgccgcgtt tacttattta t 41

<210> 70

<211> 22

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 70

gaccgtttct ctggttctca cc 22

<210> 71

<211> 76

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (25)..(27)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (34)..(36)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (40)..(45)

<223> a, c, t ó g

<400> 71

caggcgccgc gtttacttat ttatnnngct tctnnncgcn nnnnngggat cccggaccgt 60
ttctctggtt ctcacc 76

<210> 72

<211> 53

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 72

gacgagtcct tctagattgg aacctgaaga cttcgctgtt tattattgcc aac 53

<210> 73

<211> 50

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 73

actttcggtc aaggtaccaa ggttgaaatc aagcgtacgt cacaggtgag 50

<210> 74

<211> 26

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 74

gaaatcaagc gtacgtcaca ggtgag 26

<210> 75

<211> 70

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (28)..(39)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (43)..(45)

<223> a, c, t ó g

<400> 75

gacttcgctg tttattattg ccaacagnnn nnnnnnnnnc ctnnnacttt cggtcaaggt 60
accaaggttg 70

<210> 76

<211> 67

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (28)..(42)

<223> a, c, t ó g

<400> 76

gacttcgctg tttattattg ccaacagnnn nnnnnnnnnn nnccttтcgg tcaaggtacc 60
aaggttg 67

<210> 77

<211> 73

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (28)..(39)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (46)..(48)

<223> a, c, t ó g

<400> 77

gacttcgctg tttattattg ccaacagnnn nnnnnnnnnc ctcctnnnac tttcggtcaa 60
ggtaccaagg ttg 73

<210> 78

<211> 21

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 78

gacgagtcct ggtcacctgg t 21

<210> 79

<211> 48

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 79

gacgagtcct ggtcacctgg tcaaagtatc actatttctt gtacaggt 48

<210> 80

<211> 50

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 80

gtttcttggt atcaacaaca cccgggcaag gcgagatctt cacaggtgag 50

<210> 81

<211> 25

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 81

gcaaggcgag atcttcacag gtgag 25

<210> 82

<211> 67

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (24)..(29)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (33)..(50)

<223> a, c, t ó g

<400> 82

gtatcactat ttcttgtaca ggtnnnnnnc tcnnnnnnnn nnnnnnnnnn tggtatcaac 60
aacaccc 67

<210> 83

<211> 76

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (24)..(26)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (33)..(35)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (42)..(53)

<223> a, c, t ó g

<400> 83

gtatcactat ttcttgtaca ggtnnntctt ctnnngttgg cnnnnnnnnn nnngtttctt 60
ggtatcaaca acaccc 76

<210> 84

<211> 22

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 84

gagcagagga cccgggcaag gc 22

<210> 85

<211> 41

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 85

gagcagagga cccgggcaag gcgccgaagt tgatgatcta c 41

<210> 86

<211> 44

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 86

cgtccttctg gtgtcagcaa tcgtttctcc ggatcacagg tgag 44

<210> 87

<211> 23

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 87

cgtttctccg gatcacaggt gag 23

<210> 88

<211> 53

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (20)..(31)

<223> a, c, t ó g

<400> 88

gccgaagttg atgatctacn nnnnnnnnnn ncgtccttct ggtgtcagca atc 53

<210> 89

<211> 24

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 89

ctgcaggctg aagacgaggc tgac 24

<210> 90

<211> 33

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 90

ctgcaggctg aagacgaggc tgactactat tgt 33

<210> 91

<211> 57

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 91

gtcttcggcg gtggtaccaa acttactgtc ctcggtcaac ctaaggacac aggtgag 57

<210> 92

<211> 25

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 92

cggtcaacct aaggacacag gtgag 25

<210> 93

<211> 77

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (22)..(36)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (40)..(51)

<223> a, c, t ó g

<400> 93

gacgaggctg actactattg tnnnnnnnnn nnnnnnntctn nnnnnnnnnn ngtcttcggc 60
ggtggtacca aacttac 77

<210> 94

<211> 74

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<220>

<221 > base modificada

<222> (22)..(24)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (31)..(36)

<223> a, c, t ó g

<220>

<221 > base modificada

<222> (40)..(48)

<223> a, c, t ó g

<400> 94

gacgaggctg actactattg tnnnagctat nnnnnntctn nnnnnnnngt cttcggcggt 60
ggtaccaaac ttac 74

<210> 95

<211> 627

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: A27: gen de cadena ligera kappa JH1 con cebadores

<220>

<221 > CD5

<222> (206)..(328)

<220>

<221 > CD5

<222> (357)..(377)

<220>

<221 > CD5

<222> (405)..(470)

<220>

<221 > CD5

<222> (501)..(596)

<400> 95

```
gaggaccatt gggccccctc cgagactctc gagcgcaacg caattaatgt gagttagctc 60

actcattagg caccccaggc tttacacttt atgcttccgg ctcgtatgtt gtgtggaatt 120

gtgagcggat aacaatttca cacaggaaac agctatgacc atgattacgc caagctttgg 180

agccttttt ttggagattt tcaac gtg aag aag ctc cta ttt gct atc ccg   232
                           Met Lys Lys Leu Leu Phe Ala Ile Pro
                             1               5

ctt gtc gtt ccg ttt tac agc cat agt gca caa tcc gtc ctt act caa   280
Leu Val Val Pro Phe Tyr Ser His Ser Ala Gln Ser Val Leu Thr Gln
 10                  15                  20                  25

tct cct ggc act ctt tcg cta agc ccg ggt gaa cgt gct acc tta agt   328
Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser
                 30                  35                  40



 tagtaagctc ccaggcctct ttgatctg aaa cct ggt cag gcg ccg cgt        377
                               Lys Pro Gly Gln Ala Pro Arg
                                 45

 taatgaaagc gctaatggcc aacagtg act ggg atc ccg gac cgt ttc tct ggc 431
                             Thr Gly Ile Pro Asp Arg Phe Ser Gly
                               50                  55

 tct ggt tca ggt act gac ttt acc ctt act att tct aga taatgagtta    480
 Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg
          60                  65                  70

 actagaccta cgtaacctag ttc ggt caa ggt acc aag gtt gaa atc aag cgt 533
                       Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
                                 75                  80

 acg gtt gcc gct cct agt gtg ttt atc ttt cct cct tct gac gaa caa   581
 Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
              85                  90                  95

 ttg aag tca ggt act acgcatctct aagcggccgc aacaggagga g            627
 Leu Lys Ser Gly Thr
         100
```

<210> 96

<211> 102

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: A27: gen de cadena ligera kappa JH1 con cebadores

<400> 96

```
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
 1               5                  10                  15

His Ser Ala Gln Ser Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu
            20                  25                  30

Ser Pro Gly Glu Arg Ala Thr Leu Ser Lys Pro Gly Gln Ala Pro Arg
        35                  40                  45

Thr Gly Ile Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe
    50                  55                  60

Thr Leu Thr Ile Ser Arg Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
65                  70                  75                  80

Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
                85                  90                  95

Gln Leu Lys Ser Gly Thr
            100
```

<210> 97

<211> 413

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: 2a2: gen de cadena lambda humana JH2 con cebadores en el lugar de RDCS

<220>

<221 > CD5

<222> (30)..(104)

<220>

<221 > CD5

<222> (117)..(122)

<220>

<221 > CD5

<222> (177)..(239)

<220>

<221 > CD5

<222> (270)..(413)

<220>

<221 > CD5

<222> (129)..(131)

<220>

<221 > CD5

<222> (135)..(152)

<400> 97

```
gaggaccatt gggcccctta ctccgtgac agt gca caa tcc gct ctc act cag   53
                                  Ser Ala Gln Ser Ala Leu Thr Gln
                                    1               5

cct gct agc gtt tcc ggg tca cct ggt caa agt atc act att tct tgt   101
Pro Ala Ser Val Ser Gly Ser Pro Gly Gln Ser Ile Thr Ile Ser Cys
     10                  15                  20

aca tcttagtgac tc aga tct taatga ccg   tag cac ccg ggc aag gcg    149
Thr              Arg Ser        Pro       His Pro Gly Lys Ala
 25                                      30

ccg taatgaatct cgtacgctgg tgtt agc aat cgt ttc tcc gga tct aaa    200
Pro                             Ser Asn Arg Phe Ser Gly Ser Lys
                                  35                  40

tcc ggt aat acc gca agc tta act atc tct ggt ctg cag gttctgtagt    249
Ser Gly Asn Thr Ala Ser Leu Thr Ile Ser Gly Leu Gln
         45                  50                  55

tccaattgct ttagtgaccc ggc ggt ggt acc aaa ctt act gtc ctc ggt caa 302
                       Gly Gly Gly Thr Lys Leu Thr Val Leu Gly Gln
                                         60                  65

cct aag gct gct cct tcc gtt act ctc ttc cct cct agt tct gaa gag   350
Pro Lys Ala Ala Pro Ser Val Thr Leu Phe Pro Pro Ser Ser Glu Glu
             70                  75                  80

ctt caa gct aac aag gct act ctt gtt tgc ttg atc agt gac ttt tat   398
Leu Gln Ala Asn Lys Ala Thr Leu Val Cys Leu Ile Ser Asp Phe Tyr
         85                  90                  95

cct ggt gct gtt act                                               413
Pro Gly Ala Val Thr
    100
```

<210> 98

<211> 103

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: 2a2: gen de cadena lambda humana JH2 con cebadores en el lugar de RDCS

<400> 98

```
Ser Ala Gln Ser Ala Leu Thr Gln Pro Ala Ser Val Ser Gly Ser Pro
  1               5                  10                  15

Gly Gln Ser Ile Thr Ile Ser Cys Thr Arg Ser Pro His Pro Gly Lys
             20                  25                  30

Ala Pro Ser Asn Arg Phe Ser Gly Ser Lys Ser Gly Asn Thr Ala Ser
         35                  40                  45

Leu Thr Ile Ser Gly Leu Gln Gly Gly Gly Thr Lys Leu Thr Val Leu
     50                  55                  60

Gly Gln Pro Lys Ala Ala Pro Ser Val Thr Leu Phe Pro Pro Ser Ser
 65                  70                  75                  80

Glu Glu Leu Gln Ala Asn Lys Ala Thr Leu Val Cys Leu Ile Ser Asp
                 85                  90                  95

Phe Tyr Pro Gly Ala Val Thr
            100
```

<210> 99

<211> 10

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido sintético

<400> 99

ctgtctgaac 10

REIVINDICACIONES

1. Una biblioteca de ADN focalizada que comprende secuencias de ADN variegadas de anticuerpos que codifican colectivamente:

(1) regiones de RDC1 de cadena ligera kappa seleccionadas del grupo que consiste en:

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

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

en las 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 ADEFGHIKLMNPQRSTVW; y

(c) una mezcla de (a) y (b) como se exponen anteirormente;

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

(2) regiones de RDC2 de cadena ligera kappa que presentan 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 RDC2 de cadena ligera kappa presentes en la biblioteca; y

(3) regiones de RDC3 de cadena ligera kappa seleccionadas del grupo que consiste en:

(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 cada uno de 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 de cada uno de ADEFGHIKLMNPQRTVWY; y

(d) una mezcla de cualquiera de (a) a (c) expuestos anteriormente;

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

2. La biblioteca de ADN focalizada según la reivindicación 1, en la que las RDC1 (a) y (b) están presentes en la biblioteca en una relación 0,68:0,32.

3. La biblioteca de ADN focalizada según la reivindicación 1, en la que las RDC3 (a), (b) y (c) están presentes en la biblioteca en una proporción de 0,65:0,1:0,25.

4. La biblioteca de ADN focalizada de cualquiera de las reivindicaciones 1-3, que comprende además secuencias de ADN variegadas que codifican colectivamente:

(1) regiones de RDC1 de cadena ligera lambda seleccionadas del grupo que consiste en:

(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) una mezcla de (a) y (b) como se exponen anteriormente;

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

(2) regiones de RDC2 de cadena ligera lambda que presentan 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; y

(3) regiones de RDC3 de cadena ligera lambda seleccionadas del grupo que consiste en:

(a) <4><5><4><2><4>S<4><4><4><4>V,

en la que <2> es 0,27 D, 0,27 N y 0,027 cada AEFGHIKLMPQRSTVWY; <4> es una mezcla equimolar de los restos de aminoácidos ADEFGHIKLMNPQRSTVW; y <5> es 0,36 S y 0,0355 cada ADEFGHIKLMNPQRTVWY;

(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) una mezcla de (a) y (b);

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

5. La biblioteca de ADN focalizada según la reivindicación 4, en la que las RDC1 (a) y (b) están presentes en la biblioteca en una relación 0,67:0,33.

6. La biblioteca de ADN focalizada según la reivindicación 4, en la que las RDC3 (a) y (b) están presentes en una mezcla equimolar.

7. La biblioteca de ADN focalizada de cualquiera de las reivindicaciones 1 a 6, que comprende además secuencias de ADN variegadas que codifican colectivamente al menos una de:

(1) regiones de RDC1 de cadena pesada seleccionadas del grupo que consiste en:

(a) $<1>_1Y_2<1>_3M_4<1>_5$,

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;

(b) $(S/T)_1(S/G/X)_2(S/G/X)_3Y_4Y_5W_6(S/G/X)_7$,

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;

(c) $V_1S_2G_3G_4S_5I_6S_7<I>_8<I>_9<1>_{10}Y_{11}Y_{12}W_{13}<I>_{14}$,

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

(d) una mezcla de (a)-(c) como se exponen anteriormente;

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

(2) regiones de RDC2 de cadena pesada seleccionadas del grupo que consiste en:

(a) <2>I<2><3>SGG<1>T<1>YADSVKG,

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 cada uno de los restos de aminoácidos Y, R, W, V, G y S; y <3> 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) <1>I<4><1><1><G><5><1><1><1>YADSVKG,

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; <4> es una mezcla equimolar de los restos D, I, N, S, W, Y; y <5> es una mezcla equimolar de los restos S, G, D y N;

(c) <1>I<4><1><1>G<5><I><1>YNPSLKG,

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 <4> y <5> son como se definieron anteriormente;

(d) <1>I<8>S><1><1><1>GGYY<1>YAASVKG,

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

(e) una mezcla de cualquiera de (a) a (d) expuestos anteriormente;

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

(3) regiones de RDC3 de cadena pesada seleccionadas del grupo que consiste en:

(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) 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;

(c) YYCA211111111YFDYWG,

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) 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) una mezcla de cualquiera de los (a) a (h) expuesto anteriormente;

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

8. La biblioteca de ADN focalizada según la reivindicación 7, 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.

9. La biblioteca de ADN focalizada según la reivindicación 7, 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.

10. La biblioteca de ADN focalizada de cualquiera de las reivindicaciones 1-9, que es una biblioteca de vectores.

11. La biblioteca de ADN focalizada de la reivindicación 10, en la que los vectores son vectores de levaduras.

12. La biblioteca de ADN focalizada de una cualquiera de las reivindicaciones 1-9, que es una biblioteca de paquetes genéticos.

13. La biblioteca de ADN focalizada de la reivindicación 12, en la que los paquetes genéticos son fagos filamentosos o células de levadura.

14. La biblioteca de ADN focalizada de la reivindicación 13, en la que los paquetes genéticos son células de levadura que presentan los polipéptidos codificos por las secuencias de ADN variegadas en la biblioteca.

15. La biblioteca de ADN focalizada de cualquiera de las reivindicaciones 1-14, en la que la biblioteca es para uso en un cribado en busca de anticuerpos humanos, que opcionalmente son anticuerpos Fv monocatenarios (scFv), anticuerpos Fv, anticuerpos Fab, anticuerpos completos, o minicuerpos.