

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 545 792**

51 Int. Cl.:

C12N 15/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **20.05.2010 E 10731846 (1)**

97 Fecha y número de publicación de la concesión europea: **20.05.2015 EP 2432878**

54 Título: **Bancos de polipéptidos sintéticos y métodos para generar variantes de polipéptidos diversificados de forma natural**

30 Prioridad:

17.12.2009 US 287336 P
20.05.2009 US 179850 P
17.03.2010 US 314794 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
15.09.2015

73 Titular/es:

NOVIMMUNE SA (100.0%)
14 ch. Des Aulx, Plan-Les-Ouates
1228 Geneva, CH

72 Inventor/es:

FISCHER, NICOLAS;
KOSCO-VILBOIS, MARIE;
RAVN, ULLA;
GUENEAU, FRANCK y
VENET-BONNOT, SOPHIE

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 545 792 T3

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

DESCRIPCIÓN

Bancos de polipéptidos sintéticos y métodos para generar variantes de polipéptidos diversificados de forma natural

Campo de la invención

5 La invención se refiere a la generación de bancos de secuencias de ADN que codifican polipéptidos homólogos y al uso de tales bancos. Esta invención en particular se refiere a la generación de colecciones de fragmentos del anticuerpo sintético en se reemplazan una o varias regiones determinantes de complementariedad (CDR) por una colección de la CDR correspondiente capturada de una fuente natural. La invención además se refiere a la generación de colecciones de fragmentos del anticuerpo que contienen CDR derivada de un animal inmunizado y su uso como una fuente mejor para derivar fragmentos del anticuerpo de alta afinidad. La invención además se refiere a la diversificación de una porción de un polipéptido mediante la inserción de una secuencia diversificada de origen sintético o natural sin la necesidad de modificación de la original secuencia codificadora del polipéptido.

Antecedentes de la invención

15 Un anticuerpo se compone de cuatro polipéptidos: dos cadenas pesadas y dos cadenas livianas. La porción de unión al antígeno de un anticuerpo está formada por el dominio variable de la cadena liviana (VL) y el dominio variable de cadena la pesada (VH). En una extremidad de estos dominios seis bucles forman el sitio de unión al antígeno y también se denominan como regiones determinantes de complementariedad (CDR). Tres CDRs se encuentran en el dominio VH (H1, H2 y H3) y las otras tres están en el dominio VL (L1, L2 y L3). Durante el desarrollo de células B una región única de la inmunoglobulina se forma por recombinación somática conocida como recombinación V (D). 1. La región variable de la cadena pesada o liviana de la inmunoglobulina está codificada por diferentes segmentos del gen. La cadena pesada está codificada por tres segmentos llamados segmentos variable (V), de diversidad (D) y unión (J), mientras que la variable de cadena liviana está formada por la recombinación de solo dos segmentos V y J. Un gran número de paratopes de anticuerpo se puede generar por la recombinación entre una de las múltiples copias de los segmentos V, D y J que están presentes en el genoma. El segmento V codifica la CDR1 y CDR2 mientras que la CDR3 es generada por los eventos de recombinación. Durante el curso de la respuesta inmune se introduce diversidad adicional en el sitio de unión del antígeno por un proceso llamado hipermutación somática (SHM). Durante este proceso se introducen mutaciones puntuales en los genes variables de las cadenas pesada y liviana y en particular en las regiones que codifican las CDRs. Esta variabilidad adicional permite la selección y expansión de las células B que expresan variantes de anticuerpos con una mayor afinidad por su antígeno cognado.

20 En los últimos años han surgido varias tecnologías de visualización que permiten analizar grandes colecciones de proteínas o péptidos. Estas incluyen despliegue en fagos, despliegue en bacterias, despliegue en levadura y despliegue en ribosomas (Smith GP. Science. 1985 Jun 14;228(4705):1315-7; Hanes J y Pliickthun A. Proc Natl Acad Sci U S A. 1997 May 13;94(10):4937-42; Daugherty PS et al., Protein Eng. 1998 Sep; 11(9):825-32.; Boder ET y Wittrup 1(13). Nat Biotechnol. 1997 Jun; 15(6):553-7). En particular, estos métodos se han aplicado ampliamente para anticuerpos y fragmentos de los mismos. Se han descrito numerosos métodos para generar bancos de polipéptidos y seleccionar miembros con propiedades de unión deseadas.

25 Un primer enfoque es capturar por amplificación génica los genes de inmunoglobulina reordenados a partir de repertorios naturales, usando tejidos o células de seres humanos u otros mamíferos como fuente de diversidad genética. Estas colecciones de cadenas liviana y pesadas (VH y VL) reordenadas se combinan para generar bancos de pares de unión que se pueden desplegar en el bacteriófago o en otros paquetes de despliegue, tales como células de bacterias, levaduras o mamífero. En este caso se captura una gran fracción del repertorio de inmunoglobulina que se encuentra en el donante. Por lo tanto todas las estructuras codificadas por los genes de la línea germinal donante se pueden encontrar en tales repertorios, así como la diversidad generada por la recombinación de V(D)J y por hipermutación somática (Marks JD et al, J Mol Biol 05 de diciembre 1991;. 222 (3): 581-97; McCaffety patente de US N° 5.969.108).

30 Una limitación de los repertorios naturales es que los anticuerpos naturales se pueden basar en estructuras con baja estabilidad intrínseca que limitan sus niveles de expresión, vida media y su utilidad como reactivos o moléculas terapéuticas. Con el fin de superar estas limitaciones han sido desarrollados numerosos métodos para generar bancos de anticuerpos sintéticos. En estos enfoques, se seleccionan un número único o limitado de estructuras de anticuerpo seleccionadas codificadas por sus correspondientes genes de la línea germinal. La selección de estas estructuras se basa comúnmente en su estabilidad bioquímica y/o su frecuencia de expresión en repertorios de anticuerpos naturales. Con el fin de generar una colección de proteínas de unión, luego se introduce diversidad sintética en todos o un subconjunto de CDRs. Típicamente, la totalidad o parte de la CDR se diversifica usando diferentes estrategias. En algunos casos la diversidad se introdujo en posiciones seleccionadas dentro de las CDR (Knappik A et al, J Mol Biol 11 de febrero 2000; 296(1):57-86). Los residuos selectivos pueden ser los que están frecuentemente involucrados en el contacto con el antígeno, los que presentan máxima diversidad en los repertorios de anticuerpos naturales o incluso los residuos que dirigidos preferentemente por la maquinaria celular implicada en

la generación de hipermutaciones somáticas durante el proceso de maduración de afinidad natural (Balint RF, Larrick JW. *Gene*. 1993 Dec 27;137(1):109-18.).

Varios métodos se han usado para diversificar las CDR de los anticuerpos. La PCR por superposición que usa oligonucleótidos degenerados se ha utilizado ampliamente para ensamblar los elementos de la estructura y las CDR para reconstituir los genes de anticuerpos. En otro enfoque, se han manipulado genéticamente sitios de enzimas de restricción únicos en las regiones de marco en el límite de cada CDR para permitir la introducción de CDRs diversificadas por clonación mediada por la enzima de restricción. En cualquier caso, como todos los miembros de la biblioteca se basan en estructuras con características seleccionadas y preferidas, se prevé que los anticuerpos derivados de estos repertorios sean más estables y proporcionen una mejor fuente de reactivos útiles. (Knappik, US 6.696.248; Sidhu SS, et al, *Methods Enzymol* 2000; 328:333-63; Lee CV et al, *J Mol Biol* 2004 23 de julio; 340 (5): 1073-1093).

Sin embargo, una limitación importante de estos bancos sintéticos es que una proporción significativa de los miembros de la biblioteca no se expresa porque las secuencias diversificadas en forma aleatoria no permiten la expresión y/o plegamiento de adecuados de la proteína. Este problema es particularmente significativo para la CDR3 de la cadena pesada. En efecto, esta CDR a menudo contribuye la mayor parte de la energía de unión al antígeno y es altamente diversa en longitud y secuencia. Si bien las otras CDR (H1, H2, L1, L2 y L3) solo pueden adoptar un número limitado de conformaciones tridimensionales, conocidas como pliegues canónicos, el número de conformaciones que pueden ser adoptadas por la cadena pesada CDR3 aún es demasiado amplio para ser predicho (Al-Lazikani B et al., *J Mol Biol*. 1997 Nov 7; 273(4):927-48). Además, el uso de oligonucleótidos degenerados largos utilizado para cubrir CDR H3 largos a menudo introduce supresiones de pares de bases individuales. Estos factores reducen significativamente el tamaño funcional de los repertorios sintéticos.

Los repertorios naturales y sintéticos tienen ventajas y limitaciones. Por un lado, las estrategias que dependen de la captura de genes variables del anticuerpo reordenado naturalmente, no son óptimas, ya que incluyen estructuras potencialmente menos favorables dentro de la biblioteca. Un aspecto positivo es que estos genes variables reordenados incluyen CDRs que son compatibles con el plegado de dominio apropiado, ya que se han expresado en el contexto de un anticuerpo natural. Por otro lado, las estrategias basadas en la selección de estructuras y la inserción de diversidad sintética se benefician de la mejor estabilidad de las estructuras, pero están limitadas por el gran número de secuencias de CDR que no son compatibles con el plegado y/o expresión y pueden desestabilizar el dominio global (Figura 1A). Por tanto, existe una necesidad de nuevos enfoques que puedan combinar los beneficios de la utilización de estructuras seleccionadas con características deseables y combinarlos con CDR correctamente plegadas por ejemplo derivados de repertorios naturales.

Todos los enfoques descritos para generar bancos de anticuerpos, ya sea mediante la captura de secuencias de anticuerpos reordenadas de forma natural o mediante la generación de la diversidad por medios sintéticos están limitados por la aparición de mutaciones de desplazamiento de estructura que llevan a secuencias de anticuerpos no funcionales. Estas mutaciones pueden aparecer en múltiples etapas de la manipulación molecular del ADN que codifica los anticuerpos, tales como amplificación por PCR y ensamblaje del fragmento de ADN, así como la clonación molecular. La frecuencia de los miembros no funcionales en los bancos de anticuerpos típicamente varía de 15% a 45% de acuerdo con las estrategias empleadas para capturar o generar la diversidad de anticuerpos (Persson MA et al., *Proc Natl Acad Sci U S A*. 1991 Mar 15;88(6):2432-6; Schoonbroodt S, et al., *Nucleic Acids Res*. 2005 May 19;33(9):e81; Soderling E et al., *Nat Biotechnol*. 2000 Aug; 18(8):852-6 ; Rothe et al., *J Mol Biol*. 2008 Feb 29;376(4):1182-200). La frecuencia de secuencias que codifican anticuerpos no funcionales tiene un impacto importante en el proceso de identificación de anticuerpos. En primer lugar, se reduce el tamaño funcional de la biblioteca y, debido a que los clones no funcionales a menudo tienen una ventaja de crecimiento durante la propagación de los bancos, se expanden más rápido y puede comprometer el proceso de identificación de candidatos de anticuerpos (De Bruin R et al., *Nat Biotechnol* 17 de abril 1999: 397-399). Estas cuestiones se reconocen como serias limitaciones para aprovechar plenamente el potencial de las bibliotecas de anticuerpos. La generación de bancos altamente funcionales sigue siendo un desafío en el campo y ha llevado a muchos esfuerzos mejorar el proceso. Por ejemplo, múltiples estrategias de diversificación destinadas a imitar el uso de los aminoácidos encontrados en las secuencias de CDR naturales se han utilizado con el fin de probar más eficazmente la enorme diversidad de combinaciones de secuencia posibles codificadas por las CDRs sintéticas (de Kruijff J et al., *J Mol Biol*. 1995 Apr 21; 248(1):97-105; Sidhu SS et al., *J Mol Biol*. 2004 Apr 23; 338(2):299-310). Otro enfoque es para limpiar la biblioteca inicial con el fin de eliminar clones no funcionales a expensas de la pérdida de potencial de diversidad. Esto se ha aplicado a la preselección de repertorios sintéticos mediante la unión de la biblioteca de anticuerpos a un ligando genérico. Esta etapa permitió el enriquecimiento de miembros de la biblioteca que son capaces de expresar y plegar correctamente y se puede usar para recrear una biblioteca más funcional (Winter y Tomlinson, US 6.696.245 B2). Independientemente de que el enfoque de la calidad de cualquier biblioteca es dependiente de la eficiencia de los métodos de biología molecular aplicados para generar la biblioteca y generalmente producir 15% a 45% miembros no funcionales de la biblioteca. Existe por tanto una necesidad de enfoques novedosos y altamente eficientes que minimicen la frecuencia en genes no funcionales debido a los desplazamientos de la estructura introducidos durante las etapas de clonación molecular y que maximicen la funcionalidad de las bibliotecas mediante la captura de las regiones CDR que tienen una alta propensión de plegarse correctamente en estructuras de anticuerpos con propiedades deseables. Además, existe una necesidad de enfoques que permitan la captura de las secuencias de CDR de un repertorio inmune de los animales en un contexto

terapéuticamente útil tales como las estructura de anticuerpos humanos con el fin de mejorar el proceso de generación de anticuerpos de alta afinidad.

Síntesis de la invención

5 La presente invención se refiere a métodos de generación de bancos de secuencias de ácido nucleico que combinan los beneficios de la selección de estructuras estables y la inserción de las regiones determinantes de la complementariedad codificadas de forma natural (CDRs) o secuencias de aminoácidos que pueden cumplir el papel de una CDR que se han seleccionado en un contexto natural de un polipéptido funcional tal como un anticuerpo. El método permite la recuperación de CDRs largas o secuencias de aminoácidos que pueden cumplir el papel de una CDR que son muy difíciles de codificar utilizando enfoques sintéticos. Esta invención, mediante la combinación de 10 estructuras estables y CDRs correctamente plegadas o secuencias de aminoácidos que pueden cumplir el papel de un CDR, maximiza la proporción de anticuerpos funcionales en la biblioteca y por lo tanto el rendimiento del proceso de selección y la calidad de los clones seleccionados. La invención se refiere a un método para capturar CDRs expresadas naturalmente de diferentes especies ea insertarlas en una estructura de anticuerpo humano. Esto permite el uso de repertorios CDR H3 que difieren significativamente en longitud y composición cuando se compara con el repertorio humano. Por lo tanto, la invención proporciona un método de acuerdo con la reivindicación 1. La invención permite la generación de fragmentos de anticuerpos humanos que ofrecen repertorios estructurales derivados de otras especies y por lo tanto la capacidad para muestrear diferentes espacios estructurales. Los presentes métodos también se utilizan para introducir CDRs o secuencias de aminoácidos sintéticas que pueden cumplir el papel de una CDR con una mayor frecuencia de éxito que los métodos alternativos al introducir menos 15 errores que causan cambios de estructura en la secuencia codificadora. Las bancos generadas utilizando el método de la invención contienen una alta frecuencia de variantes funcionales. Las bancos de variantes se generan de acuerdo con este método utilizado para la selección y análisis con cualquier tecnología de despliegue, selección y análisis descripta.

25 El análisis de los repertorios inmunes de diferentes especies o, dentro de una especie, en diferentes etapas de desarrollo ha revelado algunas diferencias notables en las características de composición y longitud de CDR H3. Por ejemplo, la longitud media de CDRH3 en los seres humanos es más largo en adultos cuando se compara con la vida fetal o a los recién nacidos (Schroeder Jr, HW et al, 2001 Blood 98;2745-2751). Curiosamente a pesar de las grandes similitudes entre los genes de la línea germinal de anticuerpos humanos y de primates, la evolución de la CDRH-3 de longitud durante el desarrollo es diferente (Link JM et al., Molecular Immunol. 2005 42; 943-955). La comparación de las secuencias de CDR H3 se encuentran en ratones y seres humanos muestra claramente que la longitud media es significativamente más corta en ratones (Rock EP et al, J Exp Med 1994 179; 323-328). Durante el desarrollo temprano de células B en la médula ósea, la longitud media de la CDR H3 aumenta en los ratones mientras que tiende a disminuir en los seres humanos y, además, la composición de aminoácidos de los repertorios de CDRH3 murino y humano es diferente (Zemlin M et al., 2003 J Mol Biol 334; 733-749;. Ivanov I et al, 2005 J Immunol 174; 7.773-7.780). Estos ejemplos indican que las diferentes especies expresan diferentes rangos de repertorios CDR H3 a pesar del hecho de estar expuestos globalmente a clases similares de antígenos y la importancia biológica de estas observaciones aún debe ser estudiada. Se ha demostrado que la forma del sitio de combinación de los anticuerpos dirigidos contra antígenos pequeños tales como haptenos o péptidos diferentes de los dirigidos contra proteínas grandes y la formad del sitio de combinación está dictada por la longitud y composición de las CDRs (Collis A et al., J Mol Biol 2003 325; 337-354). De estos hallazgos se puede prever que el repertorio CDR H3 expresado por diferentes especies tiene diferentes propensiones a reaccionar de manera eficiente contra diferentes clases de blanco.

45 El método proporcionado en la presente está diseñado para aprovechar los diversos repertorios expresados por diferentes especies para la generación de anticuerpos terapéuticos. Estos repertorios que exploran diferentes espacios tridimensionales pueden permitir la generación de anticuerpos contra una variedad más amplia de clases de y epítomos del blanco. Los métodos para generar bancos de animales sin exposición previa o inmunizados están bien descritos y estos métodos permiten la captura de los repertorios correspondientes y la generación de anticuerpos. Sin embargo, los anticuerpos derivados de estos bancos no son de origen humano y por tanto no son muy adecuados para la terapia humana sin realizar trabajo de manipulación genética adicional tal como la humanización. En consecuencia, existe una necesidad de nuevos métodos para aprovechar la diversidad expresada en el repertorio de diferentes especies y aprovechar esta diversidad en el contexto terapéuticamente útil de un anticuerpo humano.

55 El método proporcionado en la presente aborda varias de las limitaciones descritas anteriormente y son una mejora respecto de la técnica actual. En primer lugar, los métodos proporcionados en la presente combinan los beneficios de la selección de una estructura estable y la inserción de las CDRs codificadas naturalmente que han sido seleccionadas en un contexto natural de un anticuerpo funcional. En segundo lugar, los métodos permiten una inserción altamente eficiente de secuencias CDR sintéticas o naturales en una estructura de anticuerpo que se minimiza significativamente el número de cambios de estructura en la biblioteca y por lo tanto mejora su calidad. Finalmente, la invención permite una nueva forma de utilizar la diversidad estructural del anticuerpo natural mediante 60 la captura de repertorios de CDR H3 expresados de forma natural de diferentes especies e inserción de estos en los estructuras de anticuerpos humanos. Por tanto, es posible aprovechar estos repertorios estructuralmente diversos de una manera productiva para la generación de anticuerpos para la terapia humana.

Los métodos proporcionados en la presente generan los anticuerpos que contienen una estructura estable y CDR plegadas correctamente o secuencias de aminoácidos que pueden cumplir el papel de una CDR. Los métodos capturan la diversidad natural de las secuencias en las estructuras estables.

En los métodos proporcionados en la presente, las secuencias de línea germinal para las regiones estructurales 1, 2 y 3 (FR1, FR2 y FR3) se seleccionan del organismo deseado, por ejemplo, del genoma humano (ver por ejemplo, Figuras 2 y 6). Los dominios variables del anticuerpo seleccionado se modifican mediante la introducción de una secuencia de relleno que servirán como un sitio de integración para las secuencias diversificadas. Se introduce diversidad en la secuencia fuera de la región codificadora de la inmunoglobulina mediante la introducción de los sitios de reconocimiento de la enzima de restricción, a saber sitios de restricción Tipo IIs, en la región determinante de complementariedad 3 de la cadena pesada variable (CDR H3), o la región determinante de complementariedad 3 de la cadena liviana variable (CDR L3). Si bien el método de la invención genera diversidad en la región CDHR3 (en la región de la cadena pesada variable y/o región de la cadena liviana variable), también se describe que se puede obtener diversidad en cualquier ubicación deseada, tal como, pero sin limitación, la región CDR1 (en la región de la cadena pesada variable y/o región de la cadena liviana variable) o la región CDR2 (en la región de la cadena pesada variable y/o región de la cadena liviana variable). Las secuencias de ADN diversificadas se generan con secuencias flanqueantes que incluyen los sitios de restricción Tipo IIs. En los métodos proporcionados en la presente, los extremos cohesivos generados por las enzimas de restricción son compatibles y se mantiene la estructura de lectura, de este modo permite ligar los fragmentos de ADN diversificados en una estructura aceptora.

Los métodos proporcionados en la presente también son útiles para generar secuencias de aminoácidos que tienen regiones diversificadas codificadas en las mismas. Por ejemplo, en los métodos proporcionados en la presente, las secuencias para las porciones no diversificadas del aminoácido codificado se seleccionan del organismo deseado, por ejemplo, de la secuencia humana. Una porción de la secuencia de aminoácidos modificada por la introducción de una secuencia de relleno que servirá como un sitio de integración para las secuencias diversificadas. Se introduce la diversidad en la secuencia en las ubicaciones deseadas mediante la introducción de los sitios de restricción Tipo IIs en una ubicación deseada dentro de la secuencia de aminoácidos codificada. Las secuencias de ADN diversificadas se generan con las secuencias flanqueantes que incluyen los sitios de reconocimiento Tipo IIs. En los métodos proporcionados en la presente, los extremos cohesivos generados por las enzimas de restricción son compatibles y se mantiene el marco de lectura, de este modo permite ligar los fragmentos de ADN diversificados en una estructura aceptora.

En los métodos proporcionados en la presente, una "estructura aceptora" se genera usando un "fragmento de relleno" de ADN que contiene y con preferencia limitada por dos sitios de enzimas de restricción Tipo IIs. (Ver por ejemplo, Figura 6). Con preferencia, estos dos sitios de enzimas de restricción Tipo IIs digieren las secuencias en el límite del sitio en que se desea diversidad, tal como la región CDR H3 o la región CDRL3. Como se usa en la presente, el término "estructura aceptora" se refiere a una secuencia de ácidos nucleicos que incluye las secuencias de ácidos nucleicos que codifican las regiones FR1, FR2, FR3 y FR4, las secuencias de ácidos nucleicos que codifican dos CDRs o secuencias de aminoácidos que pueden cumplir el papel de estas CDRs, y un "fragmento de relleno" que sirve como el sitio de integración para la secuencia de ácidos nucleicos diversificada. De acuerdo con la presente invención, donde se desea la diversidad en la región CDR3 (en la región de la cadena pesada variable y/o la región de la cadena liviana variable), la estructura aceptora incluye las secuencias de ácidos nucleicos que codifican las regiones FR1, FR2, FR3 y FR4, las secuencias de ácidos nucleicos que codifican las regiones de CDR1 y CDR2, y un "fragmento de relleno" que sirve como el sitio de integración para la secuencia de ácidos nucleicos diversificada.

Las expresiones "fragmento de relleno", "fragmento de ADN de relleno" y "secuencia de relleno" o cualquiera de sus variaciones gramaticales se usan de modo indistinto en la presente para referirse a una secuencia de ácidos nucleicos que incluye al menos dos sitios de reconocimiento Tipo IIs y una secuencia diversificada. La estructura aceptora puede ser una estructura aceptora de la cadena pesada variable (VH) o una estructura aceptora de la cadena liviana variable (VL). El uso de las estructurasceptoras y los fragmentos de relleno contenidas en las mismas permite la integración de una secuencia de CDR (natural o sintética) o una secuencia de aminoácidos que puede cumplir el papel de la CDR en la estructura aceptora sin nucleótidos estructurales donantes o residuos contenidos en los mismos o necesarios para la integración. Por ejemplo, el uso de las estructurasceptoras y los fragmentos de relleno contenido allí contenidos permiten la integración de una secuencia de CDR (natural o sintética) seleccionada de la CDR H3 y CDR L3 o una secuencia de aminoácidos que puede cumplir el papel de una CDR seleccionada de CDR H3 y CDR L3 en la estructura aceptora sin nucleótidos estructurales donantes o residuos contenidos en los mismos o integración. En consecuencia, después de la integración, el fragmento de relleno se elimina por completo y la región codificadora de la proteína aceptora y los fragmentos de las proteínas insertadas (es decir, las CDRs) están intactos.

Los métodos proporcionados en la presente usan cebadores que se diseñan para contener sitios de escisión para las enzimas de restricción Tipo IIs en el límite del sitio en el cual se desea diversidad, por ejemplo, la región CDR H3 o la región CDR L3. Los clones de CDR naturales aleatorias (ver por ejemplo, Figura 10) o secuencias de CDR sintéticas (ver por ejemplo, Ejemplo 6) o secuencias de aminoácidos que pueden cumplir el papel de la CDR se capturan en las estructurasceptoras, como se usa en la presente. De acuerdo con la invención, cuando se desea diversidad en la región CDR3 (en la región de la cadena pesada variable y/o la región de la cadena liviana variable),

clones de CDR3 naturales aleatorias (ver por ejemplo, Figura 10) o secuencias de CDR3 sintéticas (ver por ejemplo, Ejemplo 6) o secuencias de aminoácidos que pueden cumplir el papel de a CDR3 se capturan en la estructura aceptora usada en la presente. A modo de ejemplo, se diseñaron cebadores de oligonucleótidos específicos para las regiones flanqueantes de las secuencias de ADN que codifican la CDR H3 de las inmunoglobulinas específicas para la FR3 y FR4 de la región variable. También se pueden diseñar cebadores de oligonucleótidos específicos para las regiones flanqueantes de las secuencias de ADN que codifican otras regiones, tales como, por ejemplo, la CDR L3. Estos oligonucleótidos contienen en su extremo 5' y un sitio para una enzima de restricción Tipo IIs mientras que su porción 3' coincide con la secuencia de ADN específica.

En algunas formas de realización, el cebador es un ácido nucleico seleccionado del grupo que consiste en las SEQ ID NOs: 120-254.

Los métodos proporcionados en la presente usan enzimas de restricción Tipo IIs, tal como, por ejemplo, FokI, para insertar las secuencias de CDR sintéticas, tales como, por ejemplo, secuencias de CDR H3 o CDR L3 naturales en las estructurasceptoras descritas en la presente. Los métodos proporcionados en la presente usan enzimas de restricción Tipo IIs, tal como, por ejemplo, FokI, para insertar secuencias de CDR sintéticas, tal como, por ejemplo, secuencias de CDR H3 o CDR L3 sintéticas en las estructurasceptoras descritas en la presente. Los métodos proporcionados en la presente usan enzimas de restricción Tipo IIs, tales como, por ejemplo, FokI, para insertar secuencias de aminoácidos que pueden cumplir el papel de una región de CDR deseada, tal como, por ejemplo, una secuencia de aminoácidos que puede cumplir el papel de región de CDR H3 o CDR L3 natural o sintética en las estructurasceptoras descritas en la presente. Las enzimas de restricción Tipo IIs son enzimas que escinden fuera de su secuencia de reconocimiento en una parte. Estas enzimas son de tamaño intermedio, típicamente 400-650 aminoácidos de longitud, y reconocen secuencias que son continuas y asimétricas. Las enzimas de restricción de Tipo II adecuadas, también conoce endonucleasas de restricción y las secuencias que identifican se describen, por ejemplo, en Szybalski et al., "Class-ITS Enzimas de restricción - a Review." Gene, vol. 100: 13-26 (1991).

Los bancos primarios incluyen una estructura aceptora VH y una secuencia fija VL (también conocida como una secuencia "VL simulada") o un estructura aceptora VL y una secuencia fija VH (también conocida como una secuencia "VH simulada"). Por lo tanto, los bancos primarios exhiben diversidad solo en una de las cadenas pesadas o livianas. Los bancos secundarios se generan mediante el ligamiento de una estructura aceptora VH y una estructura aceptora VL entre sí (ver, por ejemplo, el Ejemplo 7). Los bancos secundarios tienen diversidad tanto en las cadenas pesadas como livianas.

La invención proporciona métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la cadena pesada de la inmunoglobulina humana que contiene una pluralidad de región determinante de complementariedad 3 de cadena pesada (CDR H3) aislada del repertorio del dominio variable de la inmunoglobulina de una especie no humana.

El método incluye las etapas de: (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora que codifican distintos dominios variables de la cadena pesada de la inmunoglobulina humanas, cada secuencia de ácidos nucleicos de la estructura aceptora contiene una primera región estructural (FR1), una segunda región estructural (FR2), una tercera región estructural (FR3), y una cuarta región estructural (FR4), donde las regiones FR1 y FR2 están intercaladas en una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas en una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas en una secuencia de ácidos nucleicos de relleno que incluye al menos dos sitios de reconocimiento de la enzimas de restricción Tipo IIs intercalados en una secuencia de ácidos nucleicos aleatoria; (b) proporcionar una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican las secuencias de la región determinante de complementariedad de la cadena pesada 3 (CDR H3) aislada de del repertorio de inmunoglobulina de una especie no humana donde cada una de la pluralidad de secuencias de ácidos nucleicos diversificadas incluye un sitio de reconocimiento de la enzimas de restricción Tipo IIs en cada extremidad; (c) digerir cada una de la pluralidad de secuencias de ácidos nucleicos que codifican las regiones CDR H3 usando una enzima de restricción Tipo IIs que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (b) y digerir la secuencia de ácidos nucleicos de relleno de la etapa (a) a partir de la estructura aceptora usando una enzima de restricción Tipo II que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a); y (d) ligar las secuencias de ácidos nucleicos digeridas que codifica las regiones CDR H3 o las secuencias de aminoácidos de la etapa (c) en la estructura aceptora digerida de la etapa (c) de modo que las regiones FR3 y FR4 se intercalan en las secuencias de ácidos nucleicos que codifican la región CDR H3 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR3 y se restauran secuencias que codifican el dominio variable de la inmunoglobulina completa que no contienen los sitios de reconocimiento de la enzimas de restricción Tipo IIs de las etapas (a) y (b).

En algunas formas de realización, la etapa (b) expuesta anteriormente se realiza mediante la amplificación de la secuencia de CDR H3 de una especie no humana usando cebadores de oligonucleótidos que contienen un sitio de restricción Tipo IIs. En algunas formas de realización, la etapa (b) expuesta anteriormente se realiza mediante la amplificación la secuencia de CDR H3 una especie no humana usando cebadores de oligonucleótidos que contienen un sitio de restricción FokI II. En algunas formas de realización, la especie no humana es primate no humano, roedor, canino, felino, oveja, cabra, vaca, caballo o cerdo.

- 5 En algunas formas de realización, los sitios de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a) y la etapa (b) son reconocidos por las mismas enzimas de restricción Tipo IIs. En algunas formas de realización, los sitios de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a) y la etapa (b) son reconocidos por las diferentes enzimas de restricción Tipo IIs. Por ejemplo, los sitios de reconocimiento Tipo IIs de la enzima de restricción son sitios de reconocimiento FokI, sitios de reconocimiento BsaI, y/o sitios de reconocimiento BsmBI.
- 10 En algunas formas de realización, la secuencia de ácidos nucleicos de la estructura aceptora se deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia del gen variable de la cadena pesada humana o una secuencia derivada de una secuencia del gen variable de la cadena pesada humana. En algunas formas de realización, la secuencia del gen variable de la cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5-51. En algunas formas de realización, la secuencia humana es una secuencia del gen variable de la cadena liviana kappa humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana kappa humana. Por ejemplo, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VKI-33, VK1-39, VK3-11, VK3-15, y VK3-20. En algunas formas de realización, la secuencia humana es una secuencia del gen variable de la cadena liviana lambda humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL1-44 y VL1-51.
- 15 En una forma de realización, la pluralidad de ácidos nucleicos diversificados codifica las regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye secuencias naturales o secuencias derivadas de animales inmunizados.
- 20 En una forma de realización, la pluralidad de ácidos nucleicos diversificados incluye o se deriva de secuencias seleccionadas de secuencias de CDR3 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras colecciones de polipéptidos diversificados naturales.
- 25 En una forma de realización, la pluralidad de ácidos nucleicos diversificados codifica las regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye o se deriva de secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular.
- 30 En una forma de realización, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.
- 35 En otra forma de realización, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.
- 40 En algunas formas de realización, la pluralidad de secuencias de ácidos nucleicos de la estructura aceptora incluye una mezcla de al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena pesada (VH) variable y al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena liviana variable.
- 45 En algunas formas de realización, los métodos proporcionados incluyen la etapa adicional de (e) transformar el vector de expresión de la etapa (d) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar la pluralidad de secuencias de la estructura aceptora. Por ejemplo, la célula huésped es E. coli. En algunas formas de realización, el vector de expresión es un vector fagémido. Por ejemplo, el vector fagémido es pNDS 1.
- 50 También se describen métodos para obtener un anticuerpo específico del blanco, región variable del anticuerpo o una porción de los mismos, mediante (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora que codifican dominios variables de la inmunoglobulina distintos, cada secuencia de ácidos nucleicos de la estructura aceptora que incluye una primera región estructural (FR1), una segunda región estructural (FR2), una tercera región estructural (FR3), y una cuarta región estructural (FR4), donde las regiones FR1 y FR2 están intercaladas en una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas en una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas en una secuencia de ácidos nucleicos de relleno que incluye al menos dos sitios de reconocimiento de la enzimas de restricción Tipo IIs intercalados por una secuencia de ácidos nucleicos aleatoria; (b) proporcionar una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican regiones de la región determinante de complementariedad 3 (CDR3) o que codifican secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, donde cada una de la pluralidad de secuencias de ácidos nucleicos diversificadas incluye u sitio de reconocimiento de la enzimas de restricción Tipo IIs en cada extremidad; (c) digerir cada una de la pluralidad de secuencias de ácidos nucleicos que codifican las regiones de CDR3 o secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 usando una enzima de restricción Tipo IIs que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (b) y digerir la secuencia de ácidos nucleicos de relleno de la etapa (a) usando un enzimas de restricción Tipo IIs que se une al sitio de reconocimiento de la enzima
- 55

de restricción Tipo IIs de la etapa (a); (d) clonar las secuencias de ácidos nucleicos digeridas que codifican las regiones de CDR3 o las secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 en un vector de expresión y ligar las secuencias de ácidos nucleicos digeridas que codifican las regiones de CDR3 o las secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 de la etapa (c) en la estructura aceptora de modo que las regiones FR3 y FR4 se intercalan en las secuencias de ácidos nucleicos que codifican la región CDR3 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR3 y se restaura una secuencia que codifica un gen variable de la inmunoglobulina completa; (e) transformar el vector de expresión de la etapa (e) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar la pluralidad de secuencias de la estructura aceptora; (f) poner en contacto la célula huésped con un antígeno blanco; y (g) determinar cuáles de las secuencias de la estructura aceptora expresadas se unen al antígeno blanco.

Opcionalmente, los sitios de reconocimiento Tipo IIs de la enzima de restricción de la etapa (a) y etapa (b) son reconocidos por la misma enzimas de restricción Tipo IIs. Opcionalmente, los sitios de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a) y etapa (b) son reconocidos por diferentes enzimas de restricción Tipo IIs. Por ejemplo, los sitios de reconocimiento de las enzimas de restricción Tipo IIs son sitios de reconocimiento FokI, sitios de reconocimiento BsaI, y/o sitios de reconocimiento BsmBI.

Opcionalmente, la secuencia de ácidos nucleicos de la estructura aceptora deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia del gen variable de la cadena pesada humana o una secuencia derivada de una secuencia del gen variable de la cadena pesada humana. Opcionalmente, la secuencia del gen variable de la cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5-51. Alternativamente, la secuencia humana es una secuencia del gen variable de la cadena liviana kappa humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana kappa humana. Por ejemplo, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15, y VK3-20. Alternativamente, la secuencia humana es una secuencia del gen variable de la cadena liviana lambda humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VLI-44 y VL1-51.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica las regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye secuencias naturales o secuencias derivadas de animales inmunizados.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados incluye o se deriva de secuencias seleccionadas de secuencias de CDR3 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras colecciones de polipéptidos diversificados naturales.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica las regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye o se deriva de secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular.

Alternativamente, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, y la pluralidad de ácidos nucleicos diversificados incluye las secuencias sintéticas.

En algunos casos, la pluralidad de las secuencias de ácidos nucleicos de la estructura aceptora incluyen una mezcla de al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena pesada (VH) variable y al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena liviana variable.

Opcionalmente, el vector de expresión es un vector fagémido. Por ejemplo, el vector fagémido es pNDS1. En algunas formas de realización, la célula huésped es E. coli.

Opcionalmente, el método incluye la etapa adicional de (i) secuenciar el dominio variable de la inmunoglobulina que codifica las secuencias que se unen al antígeno blanco.

También se describen métodos para obtener un anticuerpo específico del blanco, región variable del anticuerpo o una porción de los mismos, al (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora que codifican dominios variables de la inmunoglobulina distintos, cada secuencia de ácidos nucleicos de la estructura aceptora que incluye una primera región estructural (FR1), una segunda región estructural (FR2), una tercera región estructural (FR3), y una cuarta región estructural (FR4), donde las regiones FR1 y FR2 están intercaladas en una secuencia de ácidos nucleicos de relleno que incluye al menos dos sitios de reconocimiento de la enzimas de restricción Tipo IIs intercalados por una secuencia de ácidos nucleicos aleatoria, las regiones FR2 y FR3 están intercaladas en una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas en una región determinante de complementariedad 3 (CDR3); (b) proporcionar una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican las regiones de la región determinante de complementariedad (CDR1) o que codifican secuencias de aminoácidos que pueden cumplir el papel de una región CDR1, donde cada una de la pluralidad de secuencias de ácidos nucleicos diversificadas incluye un sitio de reconocimiento de la enzima de restricción Tipo IIs en cada extremidad; (c) digerir cada una de la pluralidad de

- secuencias de ácidos nucleicos que codifican las regiones CDR1 o secuencias de aminoácidos que pueden cumplir el papel de una región CDR1 usando una enzima de restricción Tipo IIs que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (b) y digerir la secuencia de ácidos nucleicos de relleno de la etapa (a) usando una enzima de restricción Tipo IIs que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a); (d) clonar las secuencias de ácidos nucleicos digeridas que codifican las regiones CDR1 o las secuencias de aminoácidos que pueden cumplir el papel de una región CDR1 en un vector de expresión y ligar las secuencias de ácidos nucleicos digeridas que codifican las regiones CDR1 o las secuencias de aminoácidos que pueden cumplir el papel de una región CDR1 de la etapa (c) en la estructura aceptora de modo que las regiones FR1 y FR2 se intercalan en las secuencias de ácidos nucleicos que codifican la región CDR1 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR1 y se restaura un gen variable de la inmunoglobulina completa que codifica la secuencia; (e) transformar el vector de expresión de la etapa (e) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar la pluralidad de secuencias de la estructura aceptora; (f) poner en contacto la célula huésped con un antígeno blanco; y (g) determinar cuáles de las secuencias de la estructura aceptora expresadas se unen al antígeno blanco.
- 15 Opcionalmente, los sitios de reconocimiento Tipo IIs de la enzima de restricción de la etapa (a) y etapa (b) son reconocidos por las mismas enzimas de restricción Tipo IIs. Opcionalmente, los sitios de reconocimiento de las enzimas de restricción Tipo IIs de la etapa (a) y etapa (b) son reconocidas por diferentes enzimas de restricción de Tipo II. Por ejemplo, los sitios de reconocimiento Tipo IIs de la enzima de restricción son sitios de reconocimiento FokI sitios de reconocimiento Bsal, y/o sitios de reconocimiento BsmBI.
- 20 En algunos casos, la secuencia de ácidos nucleicos de la estructura aceptora deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia del gen variable de la cadena pesada humana o una secuencia derivada de una secuencia del gen variable de la cadena pesada humana. En algunos casos, la secuencia del gen variable de la cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5-51. En algunos casos, la secuencia humana es una secuencia del gen variable de la cadena liviana Kappa humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana kappa humana. Por ejemplo, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15, y VK3-20. En algunos casos, la secuencia humana es una secuencia del gen variable de la cadena liviana lambda humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL1-44 y VL1-51.
- 25 Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica las regiones CDR1, y la pluralidad de ácidos nucleicos diversificados incluye secuencias naturales o secuencias derivadas de animales inmunizados.
- Opcionalmente, la pluralidad de ácidos nucleicos diversificados incluye o se deriva de las secuencias seleccionadas de las secuencias CDR1 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras colecciones de polipéptidos diversificados naturales.
- 35 Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica las regiones CDR1, y la pluralidad de ácidos nucleicos diversificados incluye o se deriva de secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular.
- 40 Alternativamente, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR1, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.
- En algunos casos, la pluralidad de secuencias de ácidos nucleicos de la estructura aceptora incluye una mezcla de al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena pesada (VH) variable y al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena liviana variable.
- 45 En algunos casos, el vector de expresión es un vector fagémido. Por ejemplo, el vector fagémido es pNDS 1. En algunos casos, la célula huésped es E. coli.
- En algunos casos, el método incluye la etapa adicional de (i) secuenciar el dominio variable de la inmunoglobulina que codifica las secuencias que se unen al antígeno blanco.
- 50 También se describen métodos para obtener un anticuerpo específico del blanco, una región variable del anticuerpo o una porción de los mismos, al (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora que codifican dominios variables de la inmunoglobulina distintos, cada secuencia de ácidos nucleicos de la estructura aceptora que incluye una primera región estructural (FR1), una segunda región estructural (FR2), una tercera región estructural (FR3), y una cuarta región estructural (FR4), donde las regiones FR1 y FR2 están intercaladas en una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas en una secuencia de ácidos nucleicos de relleno que incluye al menos dos sitios de reconocimiento Tipo IIs de la enzima de restricción intercaladas por una secuencia de ácidos nucleicos aleatoria, y las regiones FR3 y FR4 están intercaladas en una región determinante de complementariedad 3 (CDR3); (b) proporcionar una
- 55

pluralidad de secuencias de ácidos nucleicos diversificadas que codifican regiones de la región determinante de complementariedad 2 (CDR2) o que codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR2, donde cada una de la pluralidad de secuencias de ácidos nucleicos diversificadas incluye un sitio de reconocimiento de la enzima de restricción Tipo IIs en cada extremidad; (c) digerir cada una de la pluralidad de secuencias de ácidos nucleicos que codifican las regiones CDR2 o secuencias de aminoácidos que pueden cumplir el papel de una región CDR2 usando una enzima de restricción Tipo IIs que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (b) y digerir la secuencia de ácidos nucleicos de relleno de la etapa (a) a partir de la estructura aceptora usando una enzima de restricción Tipo IIs que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a); (d) ligar las secuencias de ácidos nucleicos digeridas que codifican las regiones CDR2 o las secuencias de aminoácidos que pueden cumplir el papel de una región CDR2 de la etapa (c) en la estructura aceptora digerida de la etapa (c) de modo que las regiones FR2 y FR3 se intercalan en las secuencias de ácidos nucleicos que codifican la región CDR2 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR2 y se restauran las secuencias que codifican el dominio variable de la inmunoglobulina completa que no contienen los sitios de reconocimiento Tipo IIs de la enzima de restricción de las etapas (a) y (b); (e) clonar la biblioteca de ácidos nucleicos que codifica los dominios variables de la inmunoglobulina de la etapa (d) en un vector de expresión; (f) transformar el vector de expresión de la etapa (e) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar una pluralidad de dominios variables de la inmunoglobulina codificada por la biblioteca; (g) poner en contacto la pluralidad de dominios variables de la inmunoglobulina de la etapa (f) con un antígeno blanco; y (h) determinar cuáles secuencias que codifican el dominio variable de la inmunoglobulina expresado se unen al antígeno blanco.

En algunos casos, los sitios de reconocimiento Tipo IIs de la enzima de restricción de la etapa (a) y etapa (b) son reconocidas por la misma enzima de restricción Tipo II. En algunos casos, los sitios de reconocimiento Tipo IIs de la enzima de restricción de la etapa (a) y etapa (b) son reconocidas por diferentes enzimas de restricción Tipo IIs. Por ejemplo, los sitios de reconocimiento de la enzima de restricción Tipo IIs son los sitios de reconocimiento FokI, sitios de reconocimiento BsaI, y/o sitios de reconocimiento BsmBI.

En algunos casos, la secuencia de ácidos nucleicos de la estructura aceptora deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia del gen variable de la cadena pesada humana o una secuencia derivada de una secuencia del gen variable de la cadena pesada humana. En algunos casos, la secuencia del gen variable de la cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5-51. En algunos casos, la secuencia humana es una secuencia del gen variable de la cadena liviana kappa humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana kappa humana. Por ejemplo, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15, y VK3-20. Alternativamente, la secuencia humana es una secuencia del gen variable de la cadena liviana lambda humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL1-44 y VL1-51.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica las regiones CDR2, y la pluralidad de ácidos nucleicos diversificados incluye secuencias naturales o secuencias derivadas de animales inmunizados.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados incluye o se deriva de secuencias seleccionadas de secuencias de CDR2 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras colecciones de polipéptidos diversificados naturales.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica las regiones CDR2, y la pluralidad de ácidos nucleicos diversificados incluye o se deriva de secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR2, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.

Alternativamente, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR2, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.

Opcionalmente, la pluralidad de secuencias de ácidos nucleicos de la estructura aceptora incluye una mezcla de al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena pesada variable (VH) y al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena liviana variable.

Opcionalmente, la célula huésped es E. coli. En algunas formas de realización, el vector de expresión es un vector fagémido. Por ejemplo, el vector fagémido es pNDS1.

Opcionalmente, el método incluye la etapa adicional de (i) secuenciar el dominio variable de la inmunoglobulina que codifica las secuencias que se unen antígeno blanco.

- También se describen métodos para producir una biblioteca de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina. Estos métodos incluyen las etapas de (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora de Ig en donde se introduce una fuente de diversidad en una región determinante de complementariedad única (CDR) seleccionada del grupo que consiste en una región determinante de complementariedad 1 (CDR1), región determinante de complementariedad 2 (CDR2), y región determinante de complementariedad 3 (CDR3), donde secuencia de la estructura aceptora de Ig incluye una secuencia de ácidos nucleicos de relleno que incluye al menos dos sitios de reconocimiento de la enzimas de restricción Tipo IIs, y donde la fuente de diversidad es una CDR seleccionada de secuencias de CDR naturales que contienen sitios de reconocimiento de la enzimas de restricción Tipo IIs fuera de la región de CDR, (b) introducir la fuente de diversidad dentro de cada estructura aceptora de Ig por digestión de la fuente de diversidad y las estructurasceptoras de Ig con una enzima de restricción Tipo II; y (c) ligar la fuente de diversidad digerida en la estructura aceptora de Ig de modo que se restauran las secuencias que codifican el dominio variable de la inmunoglobulina completa que no contiene los sitios de reconocimiento de las enzimas de restricción Tipo IIs de las etapas (a) y (b).
- Las secuencias de la región de CDR naturales están sustancialmente sin alterar de su tipo salvaje, es decir, estado natural. Estas secuencias de la región de CDR naturales están flanqueadas por las secuencias de aminoácidos que se han manipulado genéticamente (o manipulado artificialmente de otro modo) para contener dos sitios de reconocimiento de la enzima de restricción Tipo IIs, con un sitio de reconocimiento de la enzima de restricción Tipo IIs en cada lado de la secuencia de la región de CDR natural. Los sitios de reconocimiento de la enzima de restricción Tipo IIs están fuera de la región que codifica la CDR. La secuencia de las regiones CDR no están alteradas en los límites de la región que codifica la CDR - las enzimas de restricción reconocen y empalman en una región que está arriba del límite de la región que codifica la CDR, pero no empalma dentro de la región que codifica la CDR.
- Opcionalmente, los sitios de reconocimiento de las enzimas de restricción Tipo IIs dentro de las secuencias de ácidos nucleicos de relleno y que flanqueas las secuencias de CDR naturales son reconocidas por las mismas enzimas de restricción Tipo IIs. Opcionalmente, los sitios de reconocimiento Tipo IIs de la enzima de restricción dentro de las secuencias de ácidos nucleicos de relleno y que flanquean las secuencias de CDR naturales son reconocidas por diferentes enzimas de restricción Tipo IIs. Por ejemplo, los sitios de reconocimiento Tipo IIs de la enzima de restricción son sitios de reconocimiento FokI, sitios de reconocimiento BsaI, y/o sitios de reconocimiento BsmBI.
- Opcionalmente, la secuencia de ácidos nucleicos de la estructura aceptora de Ig deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia del gen variable de la cadena pesada humana o una secuencia derivada de una secuencia del gen variable de la cadena pesada humana. Opcionalmente, la secuencia del gen variable de la cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5-51. Opcionalmente, la secuencia humana es una secuencia del gen variable de la cadena liviana kappa humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana kappa humana. Por ejemplo, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15, y VK3-20. Opcionalmente, la secuencia humana es una secuencia del gen variable de la cadena liviana lambda humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL 1-44 y VL I -51.
- Opcionalmente, el conjunto de ácidos nucleicos naturales incluye o se deriva de secuencias seleccionadas de secuencias de CDR3 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras colecciones de polipéptidos diversificados naturales.
- Opcionalmente, el conjunto de ácidos nucleicos naturales codifica las regiones CDR3, y el conjunto de ácidos nucleicos naturales incluya las secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular.
- Opcionalmente, el conjunto de ácidos nucleicos naturales incluye o se deriva de secuencias seleccionadas de secuencias de CDR1 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras colecciones de polipéptidos diversificados naturales.
- Opcionalmente, el conjunto de ácidos nucleicos naturales codifica regiones de CDR1, y el conjunto de ácidos nucleicos naturales incluye o se deriva de secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular.
- Opcionalmente, el conjunto de ácidos nucleicos naturales incluye o se deriva de secuencias seleccionadas de secuencias de CDR2 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras

colecciones de polipéptidos diversificados naturales.

5 Opcionalmente, el conjunto de ácidos nucleicos naturales codifica las regiones CDR2, y el conjunto de ácidos nucleicos naturales incluye secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular.

Opcionalmente, la pluralidad de secuencias de ácidos nucleicos de la estructura aceptora de Ig incluye una mezcla de al menos una secuencia de ácidos nucleicos de la estructura aceptora d cadena pesada variable (VH) y al menos una secuencia de ácidos nucleicos de la estructura aceptora cadena liviana variable (VL).

10 Opcionalmente, los métodos descritos incluyen la etapas adicionales de (e) clonar la biblioteca de ácidos nucleicos que codifica dominios variables de la inmunoglobulina de la etapa (d) en un vector de expresión y (f) transformar el vector de expresión de la etapa (e) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar una pluralidad de dominio variable de la inmunoglobulina codificada por la biblioteca. Por ejemplo, la célula huésped es E. coli. En algunas formas de realización, el vector de expresión es un vector fagémido. Por ejemplo, el vector fagémido es pNDS1.

15 También se describen métodos para producir una biblioteca de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina. Estos métodos incluyen las etapas de (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora de Ig en donde se introduce una fuente de diversidad en una región determinante de complementariedad única (CDR) seleccionada del grupo que consiste en región determinante de complementariedad 1 (CDR1), región determinante de complementariedad 2 (CDR2), y región determinante de complementariedad 3 (CDR3), donde la secuencia de la estructura aceptora de Ig incluye una secuencia de ácidos nucleicos de relleno que incluye al menos dos sitios de reconocimiento de la enzima de restricción Tipo IIs, y donde la fuente de diversidad es una CDR seleccionada de secuencias de CDR producidas sintéticamente que contienen sitios de reconocimiento de la enzimas de restricción Tipo IIs fuera de la región de CDR, (b) introducir la fuente de diversidad dentro de cada estructura aceptora de Ig por digestión de la fuente de diversidad y la estructura aceptora de Ig con una enzima de restricción Tipo IIs; y (c) ligar la fuente de diversidad digerida en la estructura aceptora de Ig de modo que se restauran las secuencias que codifican el dominio variable de la inmunoglobulina completa que no contienen los sitios de reconocimiento de las enzimas de restricción Tipo IIs de las etapas (a) y (b).

30 En algunos casos, los sitios de reconocimiento de las enzimas de restricción Tipo IIs dentro de las secuencias de ácidos nucleicos de relleno y las secuencias de CDR producidas sintéticamente son reconocidas por las mismas enzimas de restricción Tipo IIs. En algunos casos, los sitios de reconocimiento de las enzimas de restricción Tipo IIs dentro de las secuencias de ácidos nucleicos de relleno y las secuencias de CDR producidas sintéticamente son reconocidas por diferentes enzimas de restricción Tipo IIs. Por ejemplo, los sitios de reconocimiento de las enzimas de restricción Tipo IIs son sitios de reconocimiento FokI, sitios de reconocimiento BsaI, y/o sitios de reconocimiento BsmBI.

40 Opcionalmente, la secuencia de ácidos nucleicos de la estructura aceptora de Ig deriva de una secuencia humana. Por ejemplo, la secuencia humana es una secuencia del gen variable de la cadena pesada humana o una secuencia derivada de una secuencia del gen variable de la cadena pesada humana. Opcionalmente, la secuencia del gen variable de la cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5-51. Opcionalmente, la secuencia humana es una secuencia del gen variable de la cadena liviana kappa humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana kappa humana. Por ejemplo, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15, y VK3-20. Alternativamente, la secuencia humana es una secuencia del gen variable de la cadena liviana lambda humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL1-44 y VL1-51.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR1, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.

50 Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica las secuencias de aminoácidos que pueden cumplir el papel de una región CDR2, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.

Opcionalmente, la pluralidad de secuencias de ácidos nucleicos de la estructura aceptora de Ig incluye una mezcla de al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena pesada (VH) variable y al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena liviana variable.

55 Opcionalmente, los métodos proporcionados incluyen las etapas adicionales de (e) clonar la biblioteca de ácidos nucleicos que codifica dominios variables de la inmunoglobulina de la etapa (d) en un vector de expresión y (f) transformar el vector de expresión de la etapa (e) en una célula huésped y cultivar la célula huésped en condiciones

suficientes para expresar una pluralidad de dominio variable de la inmunoglobulina codificado por la biblioteca. Por ejemplo, la célula huésped es E. coli. Opcionalmente, el vector de expresión es un vector fagémido. Por ejemplo, el vector fagémido es pNDS1.

5 También se describen métodos para obtener un polipéptido de la inmunoglobulina. Estos métodos incluyen las etapas de (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora de Ig en donde se introduce una fuente de diversidad en una región determinante de complementariedad única (CDR) seleccionada del grupo que consiste en región determinante de complementariedad 1 (CDR1), región determinante de complementariedad 2 (CDR2), y región determinante de complementariedad 3 (CDR3), donde secuencia de la estructura aceptora de Ig incluye una secuencia de ácidos nucleicos de relleno que incluye al menos dos sitios de reconocimiento Tipo IIs de la enzima de restricción, y donde la fuente de diversidad es una CDR seleccionada de secuencias de CDR naturales que contienen sitios de reconocimiento de la enzimas de restricción Tipo IIs fuera de la región de CDR, (b) introducir la fuente de diversidad dentro de cada estructura aceptora de Ig mediante la digestión de la fuente de diversidad y las estructurasceptoras de Ig con una enzima de restricción Tipo II; (c) ligar la fuente de diversidad digerida en la estructura aceptora de Ig de modo que se restaura una secuencia que codifica el gen variable de la inmunoglobulina completa; y (d) clonar el gen que variable de la inmunoglobulina completa que la secuencia de la etapa (c) en un vector de expresión; y (e) transformar el vector de expresión de la etapa (d) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar el gen de inmunoglobulina completo que codifica las secuencias que no contienen los sitios de reconocimiento de las enzimas de restricción Tipo IIs que se restauran.

20 Las secuencias de la región de CDR naturales están sustancialmente inalteradas de su tipo salvaje, es decir, estado natural. Estas secuencias de la región de CDR naturales están flanqueadas por secuencias de aminoácidos que se han manipulado genéticamente (o manipulado artificialmente de otro modo) para contener dos sitios de reconocimiento de la enzima de restricción Tipo IIs, con un sitio de reconocimiento de la enzima de restricción Tipo IIs en cada lado de la secuencia de la región de CDR natural.

25 Opcionalmente, los sitios de reconocimiento de las enzimas de restricción Tipo IIs dentro de las secuencias de ácidos nucleicos de relleno y que flanquean las secuencias de CDR naturales son reconocidas por la misma enzima de restricción Tipo II. Opcionalmente, los sitios de reconocimiento de la enzimas de restricción Tipo IIs dentro de las secuencias de ácidos nucleicos de relleno y que flanquean las secuencias de CDR naturales son reconocidas por diferentes enzimas de restricción Tipo IIs. Por ejemplo, los sitios de reconocimiento de las enzimas de restricción Tipo IIs son sitios de reconocimiento FokI, sitios de reconocimiento Bsal, y/o sitios de reconocimiento BsmBI.

30 En algunos casos, la secuencia de ácidos nucleicos de la estructura aceptora deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia del gen variable de la cadena pesada humana o una secuencia derivada de una la secuencia del gen variable de la cadena pesada humana. En algunos casos, la secuencia del gen variable de la cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5-51. En algunos casos, la secuencia humana es una secuencia del gen variable de la cadena liviana kappa humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana kappa humana, Por ejemplo, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3- 15, y VK3-20. En algunos casos, la secuencia humana es una secuencia del gen variable de la cadena liviana lambda humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL1-44 y VL1-51.

40 Opcionalmente, el conjunto de ácidos nucleicos naturales incluye o se deriva de secuencias seleccionadas de secuencias de CDR3 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras colecciones de polipéptidos diversificados naturales.

45 En algunos casos, el conjunto de ácidos nucleicos naturales codifica las regiones CDR3, y el conjunto de ácidos nucleicos naturales incluye secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular.

50 En algunos casos, el conjunto de ácidos nucleicos naturales incluye o se deriva de secuencias seleccionadas de secuencias de CDR1 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras colecciones de polipéptidos diversificados naturales.

55 En algunos casos, el conjunto de ácidos nucleicos naturales codifica regiones CDR1, y el conjunto de ácidos nucleicos naturales incluye o se deriva de secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular.

En algunos casos, el conjunto de ácidos nucleicos naturales incluye o se deriva de secuencias

seleccionadas de secuencias de CDR2 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras colecciones de polipéptidos diversificados naturales.

5 En algunos casos, el conjunto de ácidos nucleicos naturales codifica las regiones CDR2, y el conjunto de ácidos nucleicos naturales incluye secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular,

10 En algunos casos, la pluralidad de secuencias de ácidos nucleicos de la estructura aceptora incluye una mezcla de al menos una secuencia de ácidos nucleicos de cadena pesada variable (VH) de la estructura aceptora y al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena liviana variable.

En algunos casos, el vector de expresión es un vector fagémido. En algunos casos, la célula huésped es *E. coli*.

15 Opcionalmente, el método también incluye las etapas de poner en contacto la célula huésped con un antígeno blanco, y determinar cuáles secuencias que codifican el gen variable de Ig completa expresada se unen al antígeno blanco, de este modo se identifican los anticuerpos específicos blanco, regiones variable del anticuerpo o porciones de los mismos. Opcionalmente, el método incluye la etapa adicional de (i) secuenciar las secuencias que codifican el dominio variable de inmunoglobulina que se unen antígeno blanco.

20 También se describen métodos para obtener un polipéptido de la inmunoglobulina. Estos métodos incluyen las etapas de (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora de Ig en donde se introduce una fuente de diversidad en una región determinante de complementariedad única (CDR) seleccionada del grupo que consiste en región determinante de complementariedad 1 (CDR1), región determinante de complementariedad 2 (CDR2), y región determinante de complementariedad 3 (CDR3), donde secuencia de la estructura aceptora de Ig incluye una secuencia de ácidos nucleicos de relleno que incluye al menos dos sitios de reconocimiento de las enzimas de restricción Tipo IIs, y donde la fuente de diversidad es una CDR seleccionada de secuencias de CDR producidas sintéticamente que contienen sitios de reconocimiento de las enzimas de restricción Tipo IIs fuera de la región de CDR, (b) introducir la fuente de diversidad dentro de cada estructura aceptora de Ig mediante la digestión de la fuente de diversidad y la estructura aceptora de Ig con una enzima de restricción Tipo IIs; (c) ligar la fuente de diversidad digerida en la estructura aceptora de Ig de modo que se restaura la secuencia que codifica un gen variable de la inmunoglobulina completa; (d) clonar la estructura aceptora de Ig ligada de la etapa (c) en un vector de expresión; y (e) transformar el vector de expresión de la etapa (d) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar las secuencias que codifican el gen de inmunoglobulina completa que contienen los sitios de reconocimiento de las enzimas de restricción Tipo IIs que se restauran.

35 Opcionalmente, los sitios de reconocimiento de las enzimas de restricción Tipo IIs dentro de las secuencias de ácidos nucleicos de relleno y las secuencias de CDR producidas sintéticamente son reconocidas por la misma enzima de restricción Tipo II. Opcionalmente, los sitios de reconocimiento de las enzimas de restricción Tipo IIs dentro de las secuencias de ácidos nucleicos de relleno y las secuencias de CDR producidas sintéticamente son reconocidas por diferentes enzimas de restricción Tipo IIs. Por ejemplo, los sitios de reconocimiento de las enzimas de restricción Tipo IIs son sitios de reconocimiento FokI, sitios de reconocimiento BsaI, y/o sitios de reconocimiento BsmBI.

40 Opcionalmente, la secuencia de ácidos nucleicos de la estructura aceptora de Ig deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia del gen variable de la cadena pesada humana o una secuencia derivada de una secuencia del gen variable de la cadena pesada humana. Opcionalmente, la secuencia del gen variable de la cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5-51. Opcionalmente, la secuencia humana es una secuencia del gen variable de la cadena liviana kappa humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana kappa humana. Por ejemplo, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15, y VK3-20. Opcionalmente, la secuencia humana es una secuencia del gen variable de la cadena liviana lambda humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL1-44 y VL1-51.

50 Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR1, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.

55 Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR2, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.

Opcionalmente, la pluralidad de secuencias de ácidos nucleicos de la estructura aceptora de Ig incluye una mezcla de al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena pesada (VH) variable y al

menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena liviana variable.

Opcionalmente, el vector de expresión es un vector fagémido. En algunas formas de realización, la célula huésped es *E. coli*.

5 Opcionalmente, el método también incluye las etapas de poner en contacto la célula huésped con un antígeno blanco, y determinar cuáles secuencias que codifican el gel variables de Ig completa se unen al antígeno blanco, de este modo se identifican anticuerpos específicos del blanco, región variable del anticuerpos o porciones de los mismos. Opcionalmente, el método incluye la etapa adicional de (i) secuenciar las secuencias que codifican el dominio variable de la inmunoglobulina que se unen al antígeno blanco.

10 También se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 3 (CDR3) aisladas separadamente del repertorio del dominio variable de la inmunoglobulina de una especie de mamífero. También se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 2 (CDR2) aisladas separadamente del repertorio del dominio variable de la inmunoglobulina de una especie de mamífero. Además, se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 1 (CDR1) aisladas separadamente del repertorio del dominio variable de la inmunoglobulina de una especie de mamífero.

20 Se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 3 (CDR3) aisladas separadamente del repertorio del dominio variable de la inmunoglobulina de una especie de mamífero no humano. También se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 2 (CDR2) aisladas separadamente del repertorio del dominio variable de la inmunoglobulina de una especie de mamífero no humano.

25 También se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 1 (CDR1) aisladas separadamente del repertorio del dominio variable de la inmunoglobulina de una especie de mamífero no humano.

Opcionalmente, la especie no humana es primates no humanos, roedores, caninos, felinos, ovejas, cabras, vacas, caballos, un miembro de la familia de los camélidos, llama, camello, dromedario, o cerdo.

35 Se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 3 (CDR3) aisladas separadamente del repertorio del dominio variable de la inmunoglobulina de un ser humano. También se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 2 (CDR2) aisladas separadamente del repertorio del dominio variable de la inmunoglobulina de un ser humano. También se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 1 (CDR1) aisladas separadamente del repertorio del dominio variable de la inmunoglobulina de un ser humano.

40 También se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 3 (CDR3) aisladas separadamente del repertorio del dominio variable de la inmunoglobulina de una especie no humana.

45 Opcionalmente, estos métodos incluye las etapas de (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora que codifican distintos dominios variables de inmunoglobulina humana, cada secuencia de ácidos nucleicos de la estructura aceptora que comprende una primera región estructural (FR1), una segunda región estructural (FR2), una tercera región estructural (FR3), y a cuarta región estructural (FR4), donde las regiones FR1 y FR2 están intercaladas en una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas en una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas en una secuencia de ácidos nucleicos de relleno que comprende al menos dos sitios de reconocimiento de la enzima de restricción Tipo IIs intercaladas por una secuencia de ácidos nucleicos aleatoria; (b) proporcionar una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican secuencias de la región determinante de complementariedad 3 (CDR3) aisladas del repertorio de inmunoglobulina de la especie de mamífero donde cada una de la pluralidad de secuencias de ácidos nucleicos diversificadas comprende un sitio de reconocimiento de la enzimas de restricción Tipo IIs en cada extremidad; (c) digerir cada una de la pluralidad de

- 5 secuencias de ácidos nucleicos que codifican las regiones de CDR3 usando una enzima de restricción Tipo IIs que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (b) y digerir la secuencia de ácidos nucleicos de relleno de la etapa (a) de la estructura aceptora usando una enzima de restricción Tipo IIs que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a); y (d) ligar las secuencias de ácidos nucleicos digeridas que codifican las regiones de CDR3 o las secuencias de aminoácidos de la etapa (c) en la estructura aceptora digerida de la etapa (c) de modo que las regiones FR3 y FR4 se intercalan en las secuencias de ácidos nucleicos que codifican la región CDR3 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR3 y se restauran las secuencias que codifican un dominio variable de la inmunoglobulina completa que no contienen los sitios de reconocimiento de las enzimas de restricción Tipo IIs de las etapas (a) y (b). Estas etapas también se pueden realizar usando una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican secuencias de la región determinante de complementariedad 2 (CDR2) aisladas del repertorio de inmunoglobulina de la especie de mamífero. Estas etapas también se pueden realizar usando una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican secuencias de la región determinante de complementariedad 1 (CDR1) aisladas del repertorio de inmunoglobulina de la especie de mamífero.
- 10
- 15 La etapa (b) se puede realizar mediante la amplificación de la secuencia de CDR3 de una especie de mamífero usando cebadores de oligonucleótidos que contienen un sitio de restricción Tipo IIs. El cebador de oligonucleótido se puede diseñar para aumentar la compatibilidad entre la secuencia de CDR3 de mamífero y la estructura aceptora que codifica un dominio variable de la inmunoglobulina humana. Opcionalmente, el cebador de oligonucleótido se diseña para modificar los límites de la secuencia de CDR3 de mamífero para permitir el ligamiento eficiente por medio de extremos cohesivos compatibles en la estructura aceptora que codifica un dominio variable de la inmunoglobulina humana. Opcionalmente, las secuencias de ADN de mamífero que flanquean las regiones de CDR3 después de la escisión con las enzimas de restricción Tipo IIs podrían no generar extremos cohesivos compatibles con los extremos cohesivos de las estructurasceptoras digeridas. En tales casos los oligonucleótidos usados para la amplificación se diseñan para modificar la secuencia de mamífero blanco de modo que después de la escisión con una enzima de restricción tipo IIs, los extremos cohesivos son compatibles y se puede producir un ligamiento eficiente. Estas etapas también se pueden realizar mediante la amplificación de la secuencia de CDR2 de una especie de mamífero usando cebadores de oligonucleótidos que contienen un sitio de restricción Tipo IIs. Estas etapas también se pueden realizar mediante la amplificación de la secuencia de CDR1 de una especie de mamífero usando cebadores de oligonucleótidos que contienen un sitio de restricción Tipo IIs.
- 20
- 25
- 30 La etapa (b) se puede realizar mediante la amplificación de la secuencia de CDR3 de una especie no humana usando cebadores de oligonucleótidos que contienen un sitio de restricción FOPI IIs. Estas etapas también se pueden realizar mediante la amplificación de la secuencia de CDR2 de una especie de mamífero usando los cebadores de oligonucleótidos que contienen un sitio de restricción FokI IIs. Estas etapas también se pueden realizar mediante la amplificación de la secuencia de CDR1 de una especie de mamífero usando cebadores de oligonucleótidos que contienen un sitio de restricción FokI IIs.
- 35
- Opcionalmente, los sitios de reconocimiento de las enzimas de restricción Tipo IIs de la etapa (a) y etapa (b) son reconocidas por una enzima de restricción Tipo II diferente. Opcionalmente, los sitios de reconocimiento de las enzimas de restricción Tipo IIs son sitios de reconocimiento BsmBI, sitios de reconocimiento BsaI, sitios de reconocimiento FokI o una combinación de los mismos.
- 40
- Opcionalmente, las secuencias de ácidos nucleicos diversificadas que codifican secuencias CDR3 codifican las secuencias CDR3 de la cadena pesada (CDR H3). Opcionalmente, las secuencias de ácidos nucleicos diversificadas que codifican secuencias CDR3 codifican secuencias CDR3 (CDR L3) de cadena liviana.
- Opcionalmente, las secuencias de ácidos nucleicos diversificadas que codifican secuencias de CDR2 codifican secuencias de CDR2 de cadena pesada (CDR H2). Opcionalmente, las secuencias de ácidos nucleicos diversificadas que codifican secuencias de CDR2 codifican secuencias de CDR2 de cadena liviana (CDR L2). Opcionalmente, las secuencias de ácidos nucleicos diversificadas que codifican secuencias de CDR1 codifican secuencias de CDR1 de cadena pesada (CDR H1). Opcionalmente, las secuencias de ácidos nucleicos diversificadas que codifican secuencias de CDR1 codifican secuencias de CDR1 de cadena liviana (CDR L1).
- 45
- Opcionalmente, la secuencia de ácidos nucleicos de la estructura aceptora incluye o se deriva de al menos una porción de una secuencia del gen variable de la cadena pesada humana seleccionada de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5- 51. Opcionalmente, la secuencia de ácidos nucleicos de la estructura aceptora incluye deriva de al menos una porción de una secuencia del gen variable de la cadena liviana kappa humana. Por ejemplo, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15, y VK3-20. Opcionalmente, la secuencia de ácidos nucleicos de la estructura aceptora incluye o se deriva de al menos una porción de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL1-44 y VL1-51.
- 50
- 55
- Opcionalmente, la pluralidad de secuencias de ácidos nucleicos de la estructura aceptora comprende una mezcla de al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena pesada (VH) variable y al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena liviana variable.

Opcionalmente, los métodos descritos en la presente también incluyen las etapas de (e) clonar la biblioteca de ácidos nucleicos que codifica los dominios variables de la inmunoglobulina de la etapa (d) en un vector de expresión y (f) transformar el vector de expresión de la etapa (e) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar una pluralidad del dominio variable de la inmunoglobulina codificada por la biblioteca. Opcionalmente, el vector de expresión es un vector de fagémido o fago. Opcionalmente, la célula huésped es *E. coli*.

Se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 3 (CDR3) aisladas separadamente de dominios variables de la inmunoglobulina de un mamífero no humano inmunizado o especie no humana. También se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 2 (CDR2) aisladas separadamente de dominios variables de la inmunoglobulina de un mamífero no humano inmunizado. También se describen métodos para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que incluye una pluralidad de secuencias de la región determinante de complementariedad 1 (CDR 1) aisladas separadamente de dominios variables de la inmunoglobulina de un mamífero no humano inmunizado.

Opcionalmente, especie no humana es primates no humanos, roedores, caninos, felinos, ovejas, cabras, vacas, caballos, un miembro de la familia de los camélidos, llama, camello, dromedario, o cerdo.

Opcionalmente, los métodos incluyen las etapas de (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora que codifican distintos dominios variables de inmunoglobulina humana, cada secuencia de ácidos nucleicos de la estructura aceptora que comprenden una primera región estructural (FR1), una segunda región estructural (FR2), una tercera región estructural (FR3), y una cuarta región estructural (FR4), donde las regiones FR1 y FR2 están intercaladas en una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas en una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas en una secuencia de ácidos nucleicos de relleno que comprende al menos dos sitios de reconocimiento de la enzima de restricción Tipo IIs intercalados por una secuencia de ácidos nucleicos aleatoria; (b) proporcionar una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican secuencias de la región determinante de complementariedad 3 (CDR3) aisladas del mamífero no humano inmunizado donde cada una de la pluralidad de secuencias de ácidos nucleicos diversificadas comprende un sitio de reconocimiento de la enzima de restricción Tipo IIs en cada extremidad; (c) digerir cada una de la pluralidad de secuencias de ácidos nucleicos que codifican las regiones de CDR3 usando una enzima de restricción Tipo IIs que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (b) y digerir la secuencia de ácidos nucleicos de relleno de la etapa (a) de la estructura aceptora usando una enzima de restricción Tipo II que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a); y (d) ligar las secuencias de ácidos nucleicos digeridas que codifican las regiones de CDR3 o las secuencias de aminoácidos de la etapa (c) en la estructura aceptora digerida de la etapa (c) de modo que las regiones FR3 y FR4 se intercalan en las secuencias de ácidos nucleicos que codifican la región CDR3 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR3 y se restauran las secuencias que codifican el dominio variable de la inmunoglobulina completa que no contienen los sitios de reconocimiento de la enzima de restricción Tipo IIs de las etapas (a) y (b). Estas etapas también se pueden realizar usando una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican secuencias de la región determinante de complementariedad 2 (CDR2) aisladas del mamífero no humano inmunizado. Estas etapas también se pueden realizar usando una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican secuencias de la región determinante de complementariedad 1 (CDR1) aisladas de del mamífero no humano inmunizado.

La etapa (b) se puede realizar mediante la amplificación de la secuencia de CDR3 del mamífero no humano inmunizado usando cebadores de oligonucleótidos que contienen un sitio de restricción Tipo IIs. Opcionalmente, el cebador de oligonucleótido se puede diseñar para aumentar la compatibilidad entre la secuencia de CDR3 de mamífero y la estructura aceptora que codifica un dominio variable de la inmunoglobulina humana. Opcionalmente, el cebador de oligonucleótido se diseña para modificar los límites de la secuencia de CDR3 de mamífero para permitir el ligamiento eficiente por medio de extremos cohesivos compatibles en la estructura aceptora que codifica un dominio variable de la inmunoglobulina humana. Opcionalmente, las secuencias de ADN de mamífero que flanquean las regiones de CDR3 después de la escisión con las enzimas de restricción Tipo IIs podrían no generar extremos cohesivos compatibles con los extremos cohesivos de las estructurasceptoras digeridas. En tales casos los oligonucleótidos usados para la amplificación se diseñan para modificar la secuencia de mamífero blanco de modo que después de la escisión con una enzima de restricción tipo IIs, los extremos cohesivos son compatibles y se puede producir un ligamiento eficiente. Estas etapas también se pueden realizar mediante la amplificación de la secuencia de CDR2 de una especie de mamífero no humano inmunizado usando cebadores de oligonucleótidos que contienen un sitio de restricción Tipo IIs. Estas etapas también se pueden realizar mediante la amplificación de la secuencia de CDR1 de una especie de mamífero no humano inmunizado usando cebadores de oligonucleótidos que contienen un sitio de restricción Tipo IIs.

Opcionalmente, la etapa (b) se realiza mediante la amplificación la secuencia de CDRH3 del mamífero no humano

usando cebadores de oligonucleótidos que contienen un sitio de restricción FokI IIs. Estas etapas también se pueden realizar mediante la amplificación de la secuencia de CDR2 del mamífero no humano usando cebadores de oligonucleótidos que contienen a sitio de restricción FokI IIs. Estas etapas también se pueden realizar mediante la amplificación de la secuencia de CDR1 del mamífero no humano usando cebadores de oligonucleótidos que contienen a sitio de restricción FokI IIs.

Opcionalmente, los sitios de reconocimiento de las enzimas de restricción Tipo IIs de la etapa (a) y etapa (b) son reconocidas por diferentes enzimas de restricción Tipo IIs. Opcionalmente, los sitios de reconocimiento de las enzimas de restricción Tipo IIs son sitios de reconocimiento BsmBI, sitios de reconocimiento BsaI, sitios de reconocimiento FokI o una combinación de los mismos.

Opcionalmente, las secuencias de ácidos nucleicos diversificadas que codifican secuencias CDR3 codifican secuencias CDR3 de la cadena pesada (CDR H3). Opcionalmente, las secuencias de ácidos nucleicos diversificadas que codifican secuencias CDR3 codifican secuencias CDR3 de cadena liviana (CDR L3). Opcionalmente, las secuencias de ácidos nucleicos diversificadas que codifican secuencias de CDR2 codifican secuencias CDR2 de cadena pesada (CDR H2). Opcionalmente, las secuencias de ácidos nucleicos diversificadas que codifican secuencias de CDR2 codifican secuencias de CDR2 de cadena liviana (CDR L2). Opcionalmente, las secuencias de ácidos nucleicos diversificadas que codifican secuencias de CDR1 codifican secuencias de CDR1 de cadena pesada (CDR H1). Las secuencias de ácidos nucleicos diversificadas que codifican secuencias de CDR1 codifican secuencias de CDR1 de cadena liviana (CDR L1).

Opcionalmente, la secuencia de ácidos nucleicos de la estructura aceptora incluye o se deriva de al menos una porción de una secuencia del gen variable de la cadena pesada humana seleccionada de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5-51.

Opcionalmente, la secuencia de ácidos nucleicos de la estructura aceptora incluye o se deriva de al menos una porción de una secuencia del gen variable de la cadena liviana kappa humana. Por ejemplo, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15, y VK3-20. En algunas formas de realización, la secuencia de ácidos nucleicos de la estructura aceptora incluye o se deriva de al menos una porción de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL1-44 y VL1-51.

Opcionalmente, la pluralidad de secuencias de ácidos nucleicos de la estructura aceptora comprende una mezcla de al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena pesada (VH) variable y al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena liviana variable.

Opcionalmente, los métodos también incluyen las etapas de (e) clonar la biblioteca de ácidos nucleicos que codifica dominios variables de la inmunoglobulina de la etapa (d) en un vector de expresión y (1) transformar el vector de expresión de la etapa (e) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar una pluralidad del dominio variable de la inmunoglobulina codificado por la biblioteca. Opcionalmente, la célula huésped es E. coli. Opcionalmente, el vector de expresión es un vector de fagémido o fago.

Breve descripción de los dibujos

La Figura 1A es una representación esquemática de un dominio de la proteína con una estructura y bucles que proporcionan residuos de contacto con otra proteína o molécula. Varias situaciones se ilustran: Un dominio de proteína estable con regiones de bucles plegados apropiadamente; bucles plegados apropiadamente insertados en un dominio de estabilidad intrínseca limitada; un dominio de proteína intrínsecamente estable cuya estabilidad es afectada por las regiones del bucle.

La Figura 1B es una representación esquemática de diferentes tipos de bancos de repertorios de proteína generados usando diferentes estrategias de diversificación.

La Figura 2 es una representación esquemática de una estructura aceptora de variable del anticuerpo. Se indican las regiones estructurales, CDRs y sitio de restricción tipo IIs-RM.

La Figura 3 es una representación esquemática de una estrategia usada para capturar secuencias de CDRH3 de repertorios naturales.

La Figura 4 es una representación esquemática de el beneficio de usar cebadores que contienen enzimas de restricción Tipo IIs-RM para la amplificación e inserción de las regiones de CDR naturales en las estructurasceptoras.

La Figura 5 es una ilustración que representa las secuencias del gen de línea germinal del dominio de cadena pesada y liviana seleccionado para la generación de estructurasceptoras.

La Figura 6 es una representación esquemática de una estrategia de amplificación usada para la generación de estructurasceptoras mediante la adición a las secuencias de línea germinal de un fragmento de relleno y una

región FR4.

5 La Figura 7, panel superior, es una ilustración que representa el detalle de la secuencia de los fragmentos de relleno la estructura aceptora de VH. Las secuencias de ADN reconocidas y escindidas por la enzima de restricción BsmBI están recuadradas en rojo y negro respectivamente y se indican en el panel inferior de la figura. El marco de lectura correspondiente a la secuencia variable del anticuerpo está subrayado.

La Figura 8 es una ilustración que representa las secuencias de las estructuras aceptoras.

La Figura 9 es una representación esquemática del vector de pNDS1 solo o combinado con una región variable de la cadena pesada simulada o región variable de la cadena liviana simulada.

10 La Figura 10 es una tabla que representa las secuencias de las secuencias de CDRH3 que se recuperaron de una fuente de Adnc humana y se insertaron en las estructuras aceptoras humanas.

La Figura 11 es una tabla que representa el diseño de las secuencias de CDR sintéticas para VH, VK y VL. Las posiciones están numeradas de acuerdo con el esquema de numeración de Kabat. Se indica la diversidad teórica del CDR usando una estrategia de diversificación de codón definida (NNS, DVK, NVT, TVP). Las estrategias adoptadas para la síntesis de VH CDR están recuadradas.

15 La Figura 12 es una representación esquemática de y detalle de secuencia de la inserción de CDR sintética en una estructura aceptora.

La Figura 13 es una representación esquemática de las bancos primarias y la recombinación de cadena realizada para generar las bancos secundarias.

20 La Figura 14 es una representación esquemática de la generación de bancos de VH aceptoras combinada con las bancos de VL sintéticas y la captura de repertorios de CDRH3 de origen humano o no humano.

La Figura 15 es una representación esquemática de la generación de la biblioteca MnA, MiB y MiC usando el repertorio de CDRH3 de ratones sin exposición previa o ratones inmunizados con hIFN γ o hCCL5/RANTES como una fuente de diversidad. El tamaño de los bancos se indica en los paneles superiores. Los paneles inferiores muestran la distribución de las longitudes de CDRH3 halladas en estas bancos.

25 La Figura 16 es una serie de gráficos que ilustran la titulación de salida del fago durante la selección contra hIFN γ con las bancos secundarias AD1 y AE1.

La Figura 17 es una serie de gráficos que ilustran la titulación producida del fago durante la selección contra el anticuerpo monoclonal 5E3 con las bancos secundarias AD1 y AE1.

30 La Figura 18 es una serie de gráficos que ilustran la frecuencia de longitudes de CDR H3 halladas en los bancos de AE1 y AD1 y después de tres rondas de selección contra el anticuerpo monoclonal 5E3. Se indica la distribución de cada longitud de CDR H3 dentro de las diferentes familias de VH. Sin embargo, cuando CDR H3 tiene más de 16 aminoácidos, las secuencias de 70 pb entregadas por la plataforma de secuenciación de Illumina no cubren suficientemente la secuencia estructural para identificar de forma inequívoca la familia VH1 y por lo tanto la familia VH se indica como indeterminada.

35 La Figura 19 es una serie de gráficos que ilustra la dosis respuesta de ELISA utilizando 6 preparaciones de scFv purificadas contra anticuerpo 5E3 de ratón o un 1A6 irrelevante de ratón. Los siete clones codifican diferentes scFv. El clon A6 es un scFv específico para hIFN γ y fue utilizado como un control negativo.

La Figura 20 es un gráfico que ilustra dosis respuesta de ELISA de utilizando preparaciones de scFv purificadas contra hIFN γ y en comparación con un scFv positivo específico para hIFN γ (A6).

40 La Figura 21 es un gráfico que representa el efecto inhibitor de preparaciones de scFv purificadas en un ensayo del gen indicador de luciferasa dirigido por hIFN γ . La actividad neutralizante de los dos candidatos scFv (AD1R4P1A9 y AE14R3P2E4) se comparó con la actividad de un scFv de control positivo (G9) y un scFv control negativo (D11).

45 La Figura 22 es un gráfico que representa el efecto inhibitor de preparaciones de scFv purificadas en un ensayo de inducción de MHCII en respuesta a hIFN γ . La actividad neutralizante de los dos candidatos scFv (AD1R4P1A9 y AE14R3P2E4) se comparó con la actividad de un scFv control negativo (D11).

La Figura 23 es una serie de gráficos que ilustran el efecto inhibitor de los dos candidatos AD1R4P1A9 y AE14R3P2E4 con cambio de formato en IgG en un ensayo del gen indicador de luciferasa dirigido por hIFN γ . La actividad neutralizante de las dos Ig se comparó con la actividad de una IgG irrelevante dirigida contra RANTES humano (NI-0701).

50 La Figura 24 es una serie de gráficos que ilustran un ELISA de dosis respuesta usando IgG G11 y DA4 contra los anticuerpos 5E3 de ratón, 5E3 de rata quimérico y los correspondientes isotipos de ratón y rata.

La Figura 25 es una serie de gráficos que ilustran un ELISA para la detección de 5E3 de ratón en diferentes diluciones de suero de ratón usando IgG IgGs G11 y DA4 anti-idiotípicos como anticuerpos de captura

La Figura 26 es un gráfico que representa las relaciones de salida/entrada de fagos durante la selección contra h1FNy con los bancos MnA y MiB.

- 5 La Figura 27 es un gráfico que representa las tasas de acierto obtenidas en una prueba de ELISA de scFv con clones derivados de las bancos MNA, MiB y MiC después de cada ronda de selección en contra h1FNy. El umbral se fijó a la mitad de la señal obtenida con el scFv de control A6.

La Figura 28 es un gráfico que representa la frecuencia de distribución de scFv que da diferentes niveles de señal en experimentos de unión contra h1FNy obtenidos con clones derivados de los bancos MnA y MiB.

- 10 La Figura 29 es un gráfico que representa un ELISA de dosis respuesta usando preparaciones de scFv purificadas a partir de los clones derivados de los bancos MnA y MiB contra h1FNy y en comparación con un scFv específico positivo para h1FNy (A6).

Descripción detallada de la invención

- 15 Los bancos de proteínas sintéticas y, en particular bancos de anticuerpos sintéticos son atractivos, ya que es posible durante el proceso de generación de la biblioteca seleccionar los bloques de construcción que componen estas proteínas sintéticas e incluir las características deseadas. Una limitación importante, sin embargo, es que la aleatorización de porciones de estas proteínas sintéticas para generar una colección de variantes a menudo conduce a proteínas no funcionales y por lo tanto puede disminuir drásticamente el tamaño de la biblioteca funcional y su rendimiento. Otra limitación de la diversidad sintética es que el tamaño de la biblioteca necesario para cubrir la diversidad teórica de las extensiones de aminoácidos aleatorizados no puede ser cubierto por las limitaciones prácticas. Incluso con sistemas de despliegue tales como despliegue en ribosomas se puede generar y muestrear una diversidad de 10^{13} a 10^{14} que puede cubrir máximamente la aleatorización completa de extensiones de 9 aminoácidos. Debido a que el tamaño medio de CDR H3 natural (también denominado en la presente como CDR3 de cadena pesada o VH CDR3) está por encima de 9 y puede ser más de 20 aminoácidos de longitud, la diversidad sintética no es un enfoque viable para generar tales CDRs.

- La combinación de los métodos utilizados generalmente para la manipulación de ADN y que se usan en el curso de la generación de una biblioteca de variantes de proteínas introduce errores en las secuencias de ADN. Estos errores pueden llevar a alteraciones en el marco de lectura del ADN que ya no codificará un polipéptido funcional. Típicamente, los bancos de anticuerpos generados utilizando el ensamblaje de fragmentos de ADN por PCR y/o clonación de restricción contienen entre 15% y 45% de secuencias que no están en el marco de lectura correcto para la traducción de proteínas. Estos miembros de la biblioteca no funcionales pueden comprometer la eficiencia del proceso de selección e identificación de anticuerpos y por lo tanto se reconoce como una limitación en el campo. Los métodos descritos permiten una introducción más robusta de la diversidad en una biblioteca de anticuerpos mediante el uso de una estrategia de clonación alternativa. Típicamente, la frecuencia de secuencias en marco es de aproximadamente 90%. Otra ventaja de la invención es que combina estructura variables de anticuerpo aceptoras seleccionadas con bucles de CDR que tienen una alta probabilidad de plegamiento correcto. Se permite la captura de CDRs largas que son difíciles de cubrir con los enfoques de aleatorización sintéticos. Además, los métodos descritos no emplean ninguna modificación dentro de la región codificadora de la variable de anticuerpo aceptor para la clonación de las secuencias diversificadas. Otra ventaja de este método es que varias fuentes de diversidad se pueden capturar en el mismo conjunto de estructuras del anticuerpo aceptor. Estas fuentes incluyen, pero sin limitación: CDRs de anticuerpos naturales de origen humano u otro mamífero, CDR de anticuerpos de pollo, CDRs de moléculas tipo anticuerpos tales como VHH de camélidos, IgNARs de tiburones, bucles variables de receptores de células T. Además, las CDRs naturales se pueden derivar de animales no expuestos previamente o inmunizados. En este último caso, los CDR recuperadas se enriquecen en secuencias que estuvieron involucradas en el reconocimiento del antígeno utilizado para la inmunización.

- Una característica única de los métodos descritos en la presente es la captura eficiente de las secuencias codificadoras de CDR3 de cadena pesada de las especies no humanas y su inserción en las estructura de la inmunoglobulina humana. Mediante el uso de estos métodos, en consecuencia es posible generar diferentes sitios de combinación del anticuerpo configurados por el repertorio CDRH3 capturado de otra especie y permitir el muestreo de un espacio tridimensional diferente. Estos métodos permiten la generación de anticuerpos humanos con nuevas especificidades dirigidas a una variedad diferente de clases y epítopos blanco que los accesibles para un repertorio CDRH3 humano. Además, estos nuevos anticuerpos codifican estructuras humanas, así como las regiones CDR1 y CDR2 y por lo tanto son adecuados para la terapia humana.

- 55 En este método los dominios de la proteína seleccionados, ejemplificados en los dominios variables de anticuerpos, son modificados por la introducción de una secuencia de relleno que servirá como un sitio de integración para las secuencias diversificadas. Después de la integración, el fragmento de relleno se elimina en su totalidad, de este modo deja intacta la región codificadora de la proteína aceptor y las proteínas fragmentos insertadas (es decir, las

CDRs). Este evento de integración está mediado por el uso de una enzima de restricción Tipo IIs que reconoce un sitio definido en la secuencia de ADN pero escinde el ADN a una distancia definida desde este sitio. Este enfoque tiene dos ventajas principales: (1) permite la digestión de las estructurasceptoras sin afectar sus secuencias codificadoras (no hay necesidad de diseñar sitios de restricción silenciosos); y (2) permite la digestión y la clonación de secuencias diversificadas naturales que por definición no poseen sitios de restricción compatibles.

Como se describió anteriormente, los intentos anteriores para generar bancos y/o despliegues de secuencias de anticuerpos difieren de los métodos proporcionados en la presente. Por ejemplo, algunos métodos requieren el injerto de cada CDR, tal como se describe por ejemplo por la Patente de Estados Unidos N° 6.300.064, en la que los sitios de enzimas de restricción están manipulados genéticamente en el límite de cada CDR, no solo de la región CDR H3. En otros métodos, las secuencias de CDR procedentes de fuentes naturales se amplifican y se reordenan, como se describe en, por ejemplo, la Patente de Estados Unidos N° 6.989.250. En algunos métodos, tales como los descritos en la Publicación de Solicitud de Patente de Estados Unidos N° 20060134098, las secuencias de un ratón (u otro mamífero) se añaden a una estructura humana, de modo tal que el anticuerpo resultante tiene las regiones CDR1 y CDR2 de origen murino y una región CDR3 de origen humano. Otros métodos, tales como los descritos en la Publicación de Solicitud de Patente de Estados Unidos N° 20030232333, generan anticuerpos que tienen regiones CDR1 y/o CDR1/CDR2 sintéticas junto con una región CDR3 natural. Sin embargo, estos métodos no proporcionan bancos que contienen regiones estructurales estables y CDRs plegadas correctamente.

Los métodos proporcionados en la presente diseñan estructurasceptoras del anticuerpo por clonación de la diversidad. Se diseñó una estrategia para introducir diversidad en la CDR3 de los dominios del anticuerpo humano seleccionados que evita la modificación de la secuencia de la estructura original. La estrategia se basa en la introducción fuera de la región codificadora de la inmunoglobulina de sitios de restricción de tipo IIs. Esta clase de enzimas de restricción reconoce una secuencia asimétrica e ininterrumpida de 4-7 pares de bases de pero escinde ADN a una distancia definida de hasta 20 bases independientemente de la secuencia de ADN encontrada en el sitio de escisión. Con el fin de aprovechar este sistema para la clonación de secuencias diversificadas en los estructuras seleccionados, se diseñaron estructurasceptoras que contienen un fragmento de ADN de relleno, en lugar de la CDR3, que incluye dos sitios de restricción de tipo IIs. Del mismo modo, las secuencias de ADN diversificadas se generan con secuencias flanqueantes que incluyen Tipo IIs. Con la condición de que los extremos cohesivos generados por las enzimas de restricción sean compatibles y que se mantenga el marco de lectura, los fragmentos de ADN se pueden ligar en la estructura aceptora y restaurar la CDR3 codificada en el nuevo contexto del estructura de anticuerpo aceptora (Figura 2).

Los métodos proporcionados en la presente capturan la diversidad de CDR natural. La estrategia que se desarrolló para capturar fragmentos de proteínas de diversificados de forma natural as como fuente de diversidad también aprovecha las enzimas de restricción de tipo IIs. Como ejemplo, se diseñaron cebadores de oligonucleótidos específicos para las regiones flanqueantes de la secuencia de ADN que codifica la CDR H3 de las inmunoglobulinas, es decir, específicas para la FR3 y FR4 de la región variable. Estos oligonucleótidos contienen en su extremo 5 'un sitio para una enzima de restricción Tipo IIS mientras que su porción 3' coincide con la secuencia de ADN específica. El sitio de la enzima de restricción utilizada es preferiblemente una enzima que escinde ADN lejos del sitio de reconocimiento de ADN, tal como FokI. Este es un elemento clave del método, ya que permite la amplificación eficiente de secuencias de ADN naturales, ya que mantiene una buena correspondencia entre el extremo 3 'del cebador y el ADN que flanquea la CDR H3, mientras que permite la escisión de la secuencia codificadora de CDRH3 por la escisión de ADN en el límite entre el CDR y las regiones estructurales (Figura 3). Esta escisión precisa de la secuencia codificadora de CDR es muy difícil usando enzimas Tipo II que escinden el ADN en su sitio de reconocimiento ya que el sitio de restricción correspondiente no está presente en las secuencias de ADN naturales y que la introducción de tales sitios durante la amplificación puede ser difícil debido a la escasa hibridación del cebador. En consecuencia, este método permite la amplificación de secuencias de proteínas diversificadas y su inserción en cualquier estructura aceptora del anticuerpo, independientemente del origen de la diversidad amplificada (Figura 4).

El método de la invención produce una biblioteca de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina al: (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora que codifican dominios variables de la inmunoglobulina distintos, cada secuencia de ácidos nucleicos de la estructura aceptora que incluye una primera región estructural (FR1), una segunda región estructural (FR2), una tercera región estructural (FR3), y a cuarta región estructural (FR4), donde las regiones FR1 y FR2 están intercaladas en una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas en una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas por una secuencia de ácidos nucleicos de relleno que contiene al menos dos sitios de reconocimiento de las enzimas de restricción Tipo IIs intercalados por una secuencia de ácidos nucleicos aleatoria; (b) proporcionar una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican las regiones de la región determinante de complementariedad 3 (CDR3) o que codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, donde cada una de la pluralidad de secuencias de ácidos nucleicos diversificadas incluye un sitio de reconocimiento de la enzima de restricción Tipo IIs en cada extremidad; (c) digerir cada una de la pluralidad de secuencias de ácidos nucleicos que codifican las regiones de CDR3 o secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 usando una enzima de restricción tipo IIs que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (b) y digerir la secuencia de ácidos nucleicos de relleno de la etapa

- 5 (a) de la estructura aceptora usando una enzima de restricción Tipo IIs que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a); y (d) ligar las secuencias de ácidos nucleicos digeridas que codifican las regiones de CDR3 o las secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 de la etapa (c) en la estructura aceptora digerida de la etapa (c) de modo que las regiones FR3 y FR4 están intercaladas en las secuencias de ácidos nucleicos que codifican la región CDR3 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR3 y se restauran las secuencias que codifican el dominio variable de la inmunoglobulina completa que no contienen los sitios de reconocimiento de las enzimas de restricción Tipo IIs de las etapas (a) y (b).
- 10 En algunas formas de realización, los sitios de reconocimiento de las enzimas de restricción Tipo IIs de la etapa (a) y etapa (b) en el método expuesto anteriormente son reconocidas por una enzima de restricción Tipo II diferente. Por ejemplo, en algunas formas de realización, los sitios de reconocimiento de las enzimas de restricción Tipo IIs son sitios de reconocimiento BsmBI, sitios de reconocimiento Bsal, FokI sitios de reconocimiento o una combinación de los mismos.
- 15 En algunas formas de realización, la secuencia de ácidos nucleicos de la estructura aceptora deriva de una secuencia génica humana. Por ejemplo, en algunas formas de realización, la secuencia humana es una secuencia del gen variable de la cadena pesada humana o una secuencia derivada de una secuencia del gen variable de la cadena pesada humana. En algunas formas de realización, la secuencia del gen variable de la cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3- 48, VH3-23, y VH5-51.
- 20 En algunas formas de realización, la secuencia humana es una secuencia del gen variable de la cadena liviana kappa humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana kappa humana. Por ejemplo, en algunas formas de realización, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VH-33, VK1-39, VH-11, VK3-15, y VK3-20.
- 25 En algunas formas de realización, la secuencia humana es una secuencia del gen variable de la cadena liviana lambda humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, en algunas formas de realización, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL1-44 y VL1-51.
- 30 En algunas formas de realización, la pluralidad de ácidos nucleicos diversificados incluye o se deriva de secuencias seleccionadas de secuencias de CDR3 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras colecciones de polipéptidos diversificados naturales.
- 35 En algunas formas de realización, la pluralidad de ácidos nucleicos diversificados codifica las regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye o se deriva de secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular.
- 40 En algunas formas de realización, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.
- 45 En algunas formas de realización, la pluralidad de secuencias de ácidos nucleicos de la estructura aceptora incluye una mezcla de al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena pesada (VH) variable y al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena liviana variable.
- En algunas formas de realización, los métodos proporcionados en la presente también incluyen las etapas de (e) clonar la biblioteca de ácidos nucleicos que codifica dominios variables de la inmunoglobulina de la etapa (d) en un vector de expresión y (f) transformar el vector de expresión de la etapa (e) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar una pluralidad del dominio variable de la inmunoglobulina codificado por la biblioteca.
- En algunas formas de realización, la célula huésped es E. coli. En algunas formas de realización, el vector de expresión es un vector fagémido.
- 50 Los métodos proporcionados en la presente generan o producen de otro modo un anticuerpo específico del blanco, región variable del anticuerpo o una porción de los mismos, al: (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora que codifican distintos dominios variables de la inmunoglobulina, cada secuencia de ácidos nucleicos de la estructura aceptora que incluye una primera región estructural (FR1), una segunda región estructural (FR2), una tercera región estructural (FR3), y a cuarta región estructural (FR4), donde las regiones FR1 y FR2 están intercaladas en una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas en una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas en una secuencia de ácidos nucleicos de relleno que tienen al menos dos sitios de reconocimiento de la enzima de restricción Tipo IIs intercalados por una secuencia de ácidos nucleicos aleatoria; (b) proporcionar una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican las regiones de la región

- determinante de complementariedad 3 (CDR3) o que codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, donde cada una de la pluralidad de secuencias de ácidos nucleicos diversificadas incluye un sitio de reconocimiento de la enzima de restricción Tipo IIs en cada extremidad; (c) digerir cada una de la pluralidad de secuencias de ácidos nucleicos que codifican las regiones de CDR3 o secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 usando una enzima de restricción Tipo II que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (b) y digerir la secuencia de ácidos nucleicos de relleno de la etapa (a) de la estructura aceptora usando una enzima de restricción Tipo II que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a); (d) ligar las secuencias de ácidos nucleicos digeridas que codifican las regiones de CDR3 o las secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 de la etapa (c) en la estructura aceptora digerida de la etapa (c) de modo que las regiones FR3 y FR4 se intercalan en las secuencias de ácidos nucleicos que codifican la región CDR3 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR3 y se restauran las secuencias que codifican el dominio variable de la inmunoglobulina completa que no contienen los sitios de reconocimiento de las enzimas de restricción Tipo IIs de las etapas (a) y (b); (e) clonar la biblioteca de ácidos nucleicos que codifica dominios variables de la inmunoglobulina de la etapa (d) en un vector de expresión; (f) transformar el vector de expresión de la etapa (e) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar una pluralidad de dominios variables de la inmunoglobulina codificados por la biblioteca; (g) poner en contacto la pluralidad de los dominios de inmunoglobulina de la etapa (f) con un antígeno blanco; y (h) determinar cuáles secuencias que codifican el dominio variable de la inmunoglobulina se unen al antígeno blanco.
- En algunas formas de realización, la secuencia de ácidos nucleicos de la estructura aceptora deriva de una secuencia génica humana. Por ejemplo, en algunas formas de realización, la secuencia humana es una secuencia del gen variable de la cadena pesada humana o una secuencia derivada de una secuencia del gen variable de la cadena pesada humana. Por ejemplo, en algunas formas de realización, la secuencia del gen variable de la cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5-51.
- En algunas formas de realización, la secuencia humana es una secuencia del gen variable de la cadena liviana kappa humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana kappa humana. Por ejemplo, en algunas formas de realización, la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15, y VK3-20.
- En algunas formas de realización, la secuencia humana es una secuencia del gen variable de la cadena liviana lambda humana o una secuencia derivada de una secuencia del gen variable de la cadena liviana lambda humana. Por ejemplo, en algunas formas de realización, la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL1-44 y VL1-51.
- En algunas formas de realización, la pluralidad de ácidos nucleicos diversificados incluye o se deriva de secuencias seleccionadas de secuencias de CDR3 naturales, secuencias de Ig naturales de seres humanos, secuencias de Ig naturales de un mamífero, secuencias naturales de una región del bucle de un receptor de células T en un mamífero, y otras colecciones de polipéptidos diversificados naturales.
- En algunas formas de realización, la pluralidad de ácidos nucleicos diversificados codifica las regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye o se deriva de secuencias de inmunoglobulina que se producen naturalmente en seres humanos que han sido expuestos a un inmunógeno particular o secuencias derivadas de animales que se han identificado como que han sido expuestos a un antígeno particular.
- En algunas formas de realización, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.
- A menos que se defina lo contrario, los términos científicos y técnicos utilizados en relación con la presente invención tendrán los significados que son comúnmente entendidos por los expertos en la técnica. Además, a menos que sea requerido de otro modo por el contexto, los términos singulares incluirán pluralidades y los términos plurales incluirán el singular. Generalmente, las nomenclaturas utilizadas en relación con, y técnicas de, cultivo celular y tisular, biología molecular, y química de proteínas y oligo o polinucleótidos e hibridación descriptas en la presente son bien conocidos y comúnmente usados en la técnica. Se utilizan técnicas estándar para ADN recombinante, síntesis de oligonucleótidos, y el cultivo de tejidos y transformación (por ejemplo, electroporación, lipofección). Las reacciones enzimáticas y técnicas de purificación se realizan de acuerdo con las especificaciones del fabricante o como comúnmente se logran en la técnica o como se describe en la presente. Las técnicas y procedimientos anteriores se realizan generalmente de acuerdo con métodos convencionales bien conocidos en la técnica y como se describen en diversas referencias generales y más específicas que se citan y discuten a lo largo de la presente memoria descriptiva. Ver, por ejemplo, Sambrook et al. *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)). Las nomenclaturas utilizadas en relación con, y los procedimientos y técnicas de laboratorio, química analítica, química orgánica de síntesis y química medicinal y farmacéutica que se describen en la presente son bien conocidos y comúnmente utilizados en la materia. Se usan técnicas convencionales para síntesis químicas, análisis químicos, preparación farmacéutica, formulación, y administración y tratamiento de los pacientes.

Tal como se utiliza de acuerdo con la presente descripción, se entenderá que los siguientes términos, a menos que se indique otra cosa, tienen los siguientes significados:

- Como se usa en la presente, el término "anticuerpo" se refiere a moléculas de inmunoglobulina y porciones inmunológicamente activas de moléculas de inmunoglobulina (Ig), es decir, moléculas que contienen un sitio de unión al antígeno que se une específicamente (inmunorreacciona con) un antígeno. Por "unir específicamente" o "inmunorreacciona con" o "unir inmunoespecíficamente" se entiende que el anticuerpo reacciona con uno o más determinantes antigénicos del antígeno deseado y no reacciona con otros polipéptidos o se une con afinidad mucho menor ($K_a > 10^{-6}$). Los anticuerpos incluyen, pero sin limitación, policlonales, monoclonales, quiméricos, dAb (anticuerpo de dominio), de cadena simple, fragmentos Fab, Fab' y F(ab')₂, scFv, y una biblioteca de expresión Fab.
- La unidad estructural básica del anticuerpo es conocida por comprender un tetrámero. Cada tetrámero se compone de dos pares idénticos de cadenas de polipéptidos, cada par tiene una cadena "liviana" (aproximadamente 25 kDa) y una "pesada" (aproximadamente 50 a 70 kDa). La porción amino-terminal de cada cadena incluye una región variable de aproximadamente 100 a 110 o más aminoácidos principalmente responsables del reconocimiento de antígeno. La porción carboxi-terminal de cada cadena define una región constante responsable principalmente de la función efectora. En general, las moléculas de anticuerpos obtenidos a partir de seres humanos se refieren a cualquiera de las clases IgG, IgM, IgA, IgE e IgD, que difieren entre sí por la naturaleza de la cadena pesada presente en la molécula. Ciertas clases tienen subclases, así, tal como IgG1, IgG2, y otros. Además, en los seres humanos, la cadena liviana puede ser una cadena kappa o una cadena lambda.
- La expresión "anticuerpo monoclonal" (mAb) o "composición de anticuerpo monoclonal", como se usa en la presente, se refiere a una población de moléculas de anticuerpo que contienen solo una especie molecular de molécula de anticuerpo que consiste en un producto de gen de cadena liviana única y un producto de gen de cadena pesada única. En particular, las regiones determinantes de complementariedad (CDRs) del anticuerpo monoclonal son idénticas en todas las moléculas de la población. Los MAbs contienen un sitio de unión al antígeno capaz de inmunorreaccionar con un epítipo particular del antígeno caracterizado por una afinidad de unión única por este. El término "sitio de unión al antígeno," o "porción de unión" se refiere a la parte de la molécula de inmunoglobulina que participa en la unión al antígeno. El sitio de unión de antígeno está formado por residuos de aminoácidos de las regiones variables N-terminal ("V") de la cadenas pesada ("H") y liviana ("L"). Tres extensiones altamente divergentes dentro de las regiones V de las cadenas pesadas y livianas, denominadas como "regiones hipervariables", se "regiones estructurales" o "FR". Por lo tanto, el término "FR" se refiere a secuencias de aminoácidos que se encuentran naturalmente entre, y adyacentes a, las regiones hipervariables en inmunoglobulinas. En una molécula de anticuerpo, las tres regiones hipervariables de una cadena liviana y las tres regiones hipervariables de una cadena pesada se disponen una respecto al otro en un espacio tridimensional para formar una superficie de unión al antígeno. La superficie de unión a antígeno es complementaria a la superficie tridimensional de un antígeno unido, y las tres regiones hipervariables de cada una de las cadenas pesadas y livianas se denominan como "regiones determinantes de complementariedad" o "CDR". La asignación de aminoácidos a cada dominio está de acuerdo con las definiciones de Kabat Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md. (1987 y 1991)), o Chothia & Lesk J. Mol. Biol. 196:901-917 (1987), Chothia et al. Nature 342:878-883 (1989).
- Como se usa en la presente, el término "epítipo" incluye cualquier determinante proteico capaz de unirse a una inmunoglobulina, un scFv o un receptor de células T específico. El término "epítipo" incluye cualquier determinante proteico capaz de unirse específicamente a una inmunoglobulina o receptor de células T. Los determinantes epitípicos consisten habitualmente en agrupaciones superficiales químicamente activas de moléculas tales como cadenas laterales de aminoácidos o azúcar y normalmente tienen características estructurales tridimensionales específicas, así como características de carga específicas. Por ejemplo, se pueden generar anticuerpos contra los péptidos N-terminales o C-terminales de un polipéptido. Se dice que un anticuerpo se une específicamente a un antígeno cuando la constante de disociación es $< 1 \mu\text{M}$; por ejemplo, $\leq 100 \text{ nM}$, con preferencia $\leq 10 \text{ nM}$ y con máxima preferencia $\leq 1 \text{ nM}$.
- Como se usa en la presente, los términos "unión inmunológica" y "propiedades de unión inmunológica" se refieren a las interacciones no covalentes del tipo que se producen entre una molécula de inmunoglobulina y un antígeno para el cual la inmunoglobulina es específica. La fuerza, o afinidad de las interacciones de unión inmunológicas se pueden expresar en términos de la constante de disociación (K_d) de la interacción, en la que una K_d menor representa una afinidad mayor. Las propiedades de unión inmunológica de los polipéptidos seleccionados se pueden cuantificar usando métodos bien conocidos en la técnica. Uno de tales métodos implica la medición de las tasas de formación y disociación del sitio de unión al antígeno/complejo de antígeno, donde estas tasas dependen de las concentraciones de los compañeros del complejo, la afinidad de la interacción y los parámetros geométricos que influyen igualmente en la tasa en ambas direcciones. En consecuencia la "constante de asociación" (K_{on}) y la "constante de disociación" (K_{off}) se pueden determinar mediante el cálculo de las concentraciones y las tasas reales de asociación y disociación. (Ver Nature 361: 186-87 (1993)). La relación de K_{off}/K_{on} permite la cancelación de todos los parámetros no relacionados con afinidad, y es igual a la constante de disociación K_d . (Ver, en general, Davies et al (1990) Annual Rev Biochem. 59: 439-473). Se dice que un anticuerpo de la presente invención se une específicamente a su blanco, cuando la constante de unión (K_d) de equilibrio es $\leq 1 \mu\text{M}$, por ejemplo, $\leq 100 \text{ nM}$, con preferencia $\leq 10 \text{ nM}$, y con más preferencia $\leq 1 \text{ nM}$, como se mide por ensayos tales como ensayos de unión de

radioligando o ensayos similares conocidos por los expertos en la técnica.

El término "polinucleótido aislado" como se usa en la presente, significará un polinucleótido de origen genómico, ADNc, o sintético o alguna combinación de los mismos, que en virtud de su origen el "polinucleótido aislado" (1) no está asociado con la totalidad o una porción de un polinucleótido en el que el "polinucleótido aislado" se encuentra en la naturaleza, (2) está unido operativamente a un polinucleótido al que no está unido en la naturaleza, o (3) no se produce en la naturaleza como parte de una secuencia más grande. Los polinucleótidos de acuerdo con la invención incluyen las moléculas de ácido nucleico que codifican moléculas de inmunoglobulina de la cadena pesada, y moléculas de ácido nucleico que codifican las moléculas de inmunoglobulina de cadena liviana descritas en la presente.

5 El término "proteína aislada" mencionado en la presente significa una proteína de ADNc, ARN recombinante u origen sintético o alguna combinación de los mismos, en virtud de su origen, o fuente de derivación, la "proteína aislada" (1) no está asociada con proteínas que se encuentran en la naturaleza, (2) está libre de otras proteínas de la misma fuente, por ejemplo, libre de proteínas marinas, (3) se expresa en una célula de una especie diferente, o (4) no se produce en la naturaleza.

15 El término "polipéptido" se usa en la presente como un término genérico para referirse a proteína nativa, fragmentos, o análogos de una secuencia polipeptídica. Por lo tanto, los fragmentos de proteína nativa, y análogos son especies del género polipéptido. Los polipéptidos de acuerdo con la invención comprenden las moléculas de inmunoglobulina de cadena pesada, y las moléculas de inmunoglobulina de cadena liviana descritas en la presente, así como moléculas de anticuerpo formadas por combinaciones que comprenden las moléculas de inmunoglobulina de cadena pesada con moléculas de inmunoglobulina de cadena liviana, tales como moléculas de inmunoglobulina de cadena liviana kappa y viceversa, así como fragmentos y análogos de los mismos.

20 El término "natural" como se usa en la presente tal aplicado a un objeto se refiere al hecho de que un objeto se puede hallar en la naturaleza. Por ejemplo, una secuencia de polipéptidos o polinucleótidos que está presente en un organismo (incluyendo virus) que puede aislarse de una fuente en la naturaleza y que no ha sido modificado intencionadamente por el hombre en el laboratorio o de otro modo es natural.

25 El término "unido operativamente" como se usa en la presente se refiere a que las posiciones de los componentes así descritos están en una relación que les permite funcionar de la manera pretendida. Una secuencia de control "unida operativamente" a una secuencia codificadora está ligada de tal manera que se obtiene la expresión de la secuencia codificadora en condiciones compatibles con las secuencias de control.

30 El término "secuencia de control", como se usa en la presente, se refiere a secuencias de polinucleótidos que son necesarias para efectuar la expresión y el procesamiento de secuencias codificadoras a las que están ligadas. La naturaleza de tales secuencias de control difiere de acuerdo con el organismo huésped en procariontes, tales secuencias de control generalmente incluyen promotor, sitio de unión ribosomal, y secuencia de terminación de la transcripción en eucariotas, generalmente, tales secuencias de control incluyen promotores y secuencia de terminación de la transcripción. El término "secuencias de control" pretende incluir, como mínimo, todos los componentes cuya presencia es esencial para la expresión y procesamiento, y también puede incluir componentes adicionales cuya presencia es ventajosa, por ejemplo, secuencias líder y secuencias de compañeros de fusión. El término "polinucleótido" como se refiere en la presente significa una boro polimérico de nucleótidos de al menos 10 bases de longitud, sean ribonucleótidos o desoxinucleótidos o una forma modificada de cualquier tipo de nucleótido. El término incluye formas de ADN de cadena simple y doble.

40 Como se usa en la presente, los veinte aminoácidos convencionales y sus abreviaturas siguen el uso convencional. Ver Immunology - A Synthesis (2nd Edition, E.S. Golub y D.R. Gren, Eds., Sinauer Associates, Sunderland Mass. (1991)). Los estereoisómeros (por ejemplo, D- aminoácidos) de los veinte aminoácidos convencionales, aminoácidos no naturales tales , γ -carboxyglutamate, ϵ -N,N,N-trimetilisina, ϵ -N-acetilisina, O-fosfoserinae, N-acetilseina, N-formilmetionina, 3-metilhistidina, 5-hidroxilisina, σ -N-metilarginina, y otros aminoácidos similares e imino ácidos (por ejemplo, 4- hidroxiprolina). En la notación de polipéptidos usada en la presente, la dirección izquierda es la dirección amino terminal y la dirección derecha es la dirección carboxi-terminal, de acuerdo con la convención y el uso estándar.

45 Como se aplica a polipéptidos, el término "identidad sustancial" significa que dos secuencias peptídicas, cuando se alinean óptimamente, tales como mediante los programas GAP o BESTFIT usando ponderaciones de brecha predeterminadas, comparten al menos 80 por ciento de identidad de secuencia, con preferencia al menos 90 por ciento de identidad de secuencia, con más preferencia al menos 95 por ciento de identidad de secuencia, y con máxima preferencia al menos 99 por ciento de identidad de secuencia.

50 Con preferencia, las posiciones de residuos que no son idénticas difieren en las sustituciones conservadoras de aminoácidos.

55 Las sustituciones conservadoras de aminoácidos se refieren a la capacidad de intercambio de los residuos que tienen cadenas laterales similares. Por ejemplo, un grupo de aminoácidos que tienen cadenas laterales alifáticas es glicina, alanina, valina, leucina, e isoleucina; un grupo de aminoácidos que tienen cadenas laterales alifáticas-

hidroxilo es serina y treonina; un grupo de aminoácidos que tienen cadenas laterales que contienen amida es asparagina y glutamina; un grupo de aminoácidos que tienen cadenas laterales aromáticas es fenilalanina, tirosina, y triptofano; un grupo de aminoácidos que tienen cadenas laterales básicas es lisina, arginina, e histidina; y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre es cisteína y metionina. Los grupos de sustitución de aminoácidos conservadores preferidos son: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, valina alanina, glutámico-aspártico, y asparagina-glutamina.

Como se discute en la presente, se contemplan variaciones menores en las secuencias de aminoácidos de anticuerpos o moléculas de inmunoglobulina como comprendidas por la presente invención, siempre que las variaciones en la secuencia de aminoácidos mantengan al menos el 75%, con más preferencia al menos 80%, 90%, 95%, y con máxima preferencia 99%. En particular, se contemplan las sustituciones conservadoras de aminoácidos. Las sustituciones conservativas son aquellas que tienen lugar dentro de una familia de aminoácidos que están relacionados en sus cadenas laterales. Los aminoácidos codificados genéticamente se dividen generalmente en familias: (1) aminoácidos ácidos son aspartato, glutamato; (2) aminoácidos básicos son la lisina, arginina, histidina; (3) aminoácidos no polares son alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptofano, y (4) aminoácidos polares no cargados son glicina, asparagina, glutamina, cisteína, serina, treonina, tirosina. Los aminoácidos hidrófilos incluyen arginina, asparagina, aspartato, glutamina, glutamato, histidina, lisina, serina y treonina. Los aminoácidos hidrófobos incluyen alanina, cisteína, isoleucina, leucina, metionina, fenilalanina, prolina, triptofano, tirosina y valina. Otras familias de aminoácidos incluyen (i) serina y treonina, que son la familia hidroxialifática; (ii) asparagina y glutamina, que son la familia que contiene amida; (iii) alanina, valina, leucina e isoleucina, que son la familia alifática; y (iv) fenilalanina, triptofano, y tirosina, que son la familia aromática. Por ejemplo, es razonable esperar que un reemplazo aislado de una leucina con una isoleucina o valina, un aspartato con un glutamato, una treonina con una serina, o un reemplazo similar de un aminoácido con un aminoácido estructuralmente relacionado no tendrá un efecto importante sobre la unión o las propiedades de la molécula resultante, especialmente si el reemplazo no implica un aminoácido dentro del sitio estructural. Si un cambio de aminoácido produce un péptido funcional se puede determinar fácilmente mediante el ensayo de la actividad específica del derivado de polipéptido. Los ensayos se describen en detalle en la presente. Los fragmentos o análogos de anticuerpos o normales en la técnica. Los extremos terminales carboxi y amino preferidos de los fragmentos o análogos se producen cerca de los límites de los dominios funcionales. Los dominios estructurales y funcionales se pueden identificar por comparación de los datos de secuencia de nucleótidos y/o aminoácidos con las bases de datos de secuencias públicas o privadas. Con preferencia, se usan métodos de comparación computarizados para identificar motivos de secuencia o dominios de conformación de proteína predichos que se producen en otras proteínas de estructura y/o función conocida. Los métodos para identificar secuencias de proteínas que se pliegan en una estructura tridimensional conocida son conocidos. Bowie et al. *Science* 253: 164 (1991). Por lo tanto, los ejemplos anteriores demuestran que los expertos en la técnica pueden reconocer motivos de secuencia y conformaciones estructurales que se pueden usar para definir dominios estructurales y funcionales de acuerdo con la invención.

Las sustituciones de aminoácidos preferidas son aquellas que: (1) reducen la susceptibilidad a la proteólisis, (2) reducen la susceptibilidad a la oxidación, (3) alteran la afinidad de unión para formar complejos de proteínas, (4) alteran las afinidades de unión, y (4) confieren o modifican otras propiedades fisicoquímicas o funcionales de dichos análogos. Los análogos pueden incluir diversas mutaciones de una secuencia distinta de la secuencia peptídica natural. Por ejemplo, las sustituciones de aminoácidos únicas o múltiples (con preferencia sustituciones conservadoras de aminoácidos) se pueden realizar en la secuencia natural (con preferencia en la porción del polipéptido fuera del dominio (s) que forman los contactos intermoleculares. Una sustitución conservadora de aminoácidos no debería cambiar sustancialmente las características estructurales de la secuencia original (por ejemplo, un reemplazo de aminoácido no debería tender a romper una hélice que aparece en la secuencia original, o alterar otros tipos de estructura secundaria que caracteriza a la secuencia original). Los ejemplos de estructuras secundarias y terciarias de polipéptidos reconocidas en la técnica se *Proteins, Structures and Molecular Principles* (Creighton, Ed., W. H. Freeman y Company, New York (1984)); *Introduction to Protein Structure* (C. Branden y J. Tooze, eds., Garland Publishing, New York, N.Y. (1991)); y Thornton et al. *Nature* 354:105 (1991).

Como se usa en la presente, los términos "marca" o "marcado" se refiere a la incorporación de un marcador detectable, por ejemplo, mediante la incorporación de aminoácido radiomarcado o unión a un polipéptido de restos biotinilo que se puede detectar mediante avidina marcada (por ejemplo, un marcador fluorescente que contiene estreptavidina o actividad enzimática que puede detectarse por métodos ópticos o colorimétricos). En ciertas situaciones, la marca o marcador también pueden ser terapéuticos. Varios métodos de marcación de polipéptidos y glicoproteínas son conocidos en la técnica y se pueden usar. Los ejemplos de marcadores para polipéptidos incluyen, pero sin limitación siguientes:

radioisótopos o radionúclidos (por ejemplo, ^3H , ^{14}C , ^{15}N , ^{35}S , ^{90}Y , ^{99}Tc , ^{125}I , ^{131}I), marcas fluorescentes (por ejemplo, FITC, rodamina, fósforos de lantánidos), marcadores enzimáticos (por ejemplo, peroxidasa de rábano p-galactosidasa, luciferasa, fosfatasa alcalina), quimioluminiscente, grupos biotinilo, epitopos polipeptídicos predeterminados reconocidos por un indicador secundario (por ejemplo, secuencias de pares de cierre de leucina, sitios de unión para anticuerpos secundarios, dominios de unión a metal, marcas de epítipo). En algunas formas de realización, los marcadores se unen mediante ramas espaciadoras de diversas longitudes para reducir el potencial impedimento estérico. El término "agente farmacéutico o fármaco" como se usa en la presente, se refiere a un

compuesto químico o composición capaz de inducir un efecto terapéutico deseado cuando se administra adecuadamente a un paciente.

Otros términos químicos en la presente memoria se utilizan según el uso convencional en la técnica, como se ejemplifica por McGraw-Hill Dictionary of Chemical Terms (Parker, S., Ed., McGraw-Hill, San. Francisco (1985)).

- 5 Como se usa en la presente, "sustancialmente puro" significa que una especie objeto es la especie predominante presente (es decir, sobre una base molar es más abundante que cualquier otra especie individual en la composición), y con preferencia una fracción sustancialmente purificada es una composición en la que la especie objeto comprende al menos aproximadamente 50 por ciento (sobre una base molar) de todas las especies macromoleculares presentes.
- 10 Generalmente, una composición sustancialmente pura comprenderá más de aproximadamente el 80 por ciento de todas las especies macromoleculares presentes en la composición, con preferencia de aproximadamente 85%, 90%, 95%, y 99%. Con máxima preferencia, la especie objeto se purifica hasta homogeneidad esencial (especies contaminantes no pueden detectarse en la composición mediante métodos de detección convencionales) donde la composición consiste esencialmente en una especie macromolecular única.
- 15 El término paciente incluye sujetos humanos y veterinarios.

Los anticuerpos se purifican mediante técnicas bien conocidas, tales como cromatografía de afinidad usando proteína A o proteína G, que proporcionan principalmente la fracción IgG del suero inmune. Posteriormente, o de forma alternativa, el antígeno específico que es el blanco de la inmunoglobulina buscada, o un epítipo de este, se puede inmovilizar en una columna para purificar el anticuerpo específico inmune mediante cromatografía de inmunoafinidad. Se discute purificación de inmunoglobulinas, por ejemplo, por D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

20

La invención se describirá adicionalmente en los siguientes ejemplos, que no limitan el alcance de la invención descrita en las reivindicaciones.

25 EJEMPLOS

EJEMPLO 1: Clonación de genes de línea germinal variable de la inmunoglobulina.

Siete genes de la línea germinal variable de cadena pesada humana (VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23, VH5-51), cinco genes de la línea germinal variable de cadena liviana kappa humana (VK1-33, VK1-39, VK3-11, VK3-15, VK3-20) y dos genes de la línea germinal variable de cadena liviana lambda humana (VL1-44, VL1-51) se seleccionaron para la construcción de las bibliotecas (Lefranc, M.-P. et al, 1999 Nucleic Acids Research, 27, 209-212). Se eligieron estos genes porque se utilizan a menudo en repertorios de anticuerpos expresados humanos y las estructuras que codifican muestran perfiles de estabilidad y expresión favorables como dominios individuales o en el contexto de un par VH/VL (Ewert S et al., J Mol Biol. 2003 Jan 17;325(3):531-53). Se utilizaron dos conjuntos de cebadores específicos para amplificar estos genes a partir de ADN genómico humano por PCR anidada. Este enfoque fue necesario ya que las secuencias 5' de los genes de la línea germinal de la misma familia son idénticos o muy similares. Para cada gen, se diseñó un primer par de cebadores, llamados localizadores genómicos para ser específico de las regiones no traducidas 5' y 3' que flanquean el gen de la línea germinal. El segundo par se diseñó para ser específico para el inicio de la estructura 1 (FR1) y el extremo de la FR2. Los 14 productos de PCR independientes se clonaron en pGEMT-easy (Promega, Madison WI) y su identidad e integridad se verificaron por secuenciación. La secuencia de aminoácidos de los genes de la línea germinal seleccionados se muestra en la Figura 5.

30

35

40

Los cebadores y combinación del cebador usados se indican a continuación.

Localizadores de genómica

5 K1-33 TGTTTCTAATCGCAGGTGCCAGATG (SEQ ID NO: 120)
 3 K1-33 ATTTATGTTATGACTTGTTACACTG (SEQ ID NO: 121)
 5 K1-39 TATTTGTTTTTATGTTTTCCAATCTC (SEQ ID NO: 122)
 3 K1-39 CCTTGGAGGTTTATGTTATGACTTG (SEQ ID NO: 123)
 5 K3-11 TTATTTCCAATTTTCAGATAACCACCG (SEQ ID NO: 124)
 3 K3-11 TTGTTGGGGTTTTTGTTCATGTGG (SEQ ID NO: 125)
 5 K3-15 TATTTCCAATTTTCAGATACCACTGG (SEQ ID NO: 126)
 3 K3-15 ATGTTGAATCACTGTGGGAGGCCAG (SEQ ID NO: 127)
 5 K3-20 TTATTTCCAATCTCAGATAACCACCG (SEQ ID NO: 128)
 3 K3-20 TTTTGTTCAGCTGAATCACTGTG (SEQ ID NO: 129)
 5 L1-44 ATGTCTGTGTCTCTCTCACTTCCAG (SEQ ID NO: 130)
 3 L1-44 TTCCCATTTGGCCTGGAGCACTGTG (SEQ ID NO: 131)
 5 L1-51 GTGTCTGTGTCTCTCCTGCTTCCAG (SEQ ID NO: 132)
 3 L1-51 CTTGTCTCAGTTCCCATTTGGGCTG (SEQ ID NO: 133)
 5 H1-2 ATCTCATCCACTTCTGTGTTCTCTC (SEQ ID NO: 134)
 3 H1-2 TTGGGTTTCTGACACCCTCAGGATG (SEQ ID NO: 135)
 5 H1-18 CAGGCCAGTCATGTGAGACTTCACC (SEQ ID NO: 136)
 3 H1-18 CTGCCTCCTCCCTGGGGTTTCTGAA (SEQ ID NO: 137)
 5 H1-69 CCCCTGTGTCTCTCCACAGGTGTC (SEQ ID NO: 138)
 3 H1-69 CCGGCACAGCTGCCTTCTCCCTCAG (SEQ ID NO: 139)
 5 DP-47 GAGGTGCAGCTGTTGGAG (SEQ ID NO: 140)
 5 H3-23 TCTGACCAGGGTTTCTTTTTGTTTGC (SEQ ID NO: 141)
 3 H3-23 TTGTGTCTGGGCTCACAATGACTTC (SEQ ID NO: 142)
 5 H3-30 TGGCATTCTTGATAACGGTGTCC (SEQ ID NO: 143)
 3 H3-30 CTGCAGGGAGGTTTGTGTCTGGGCG (SEQ ID NO: 144)
 5 H3-48 ATATGTGTGGCAGTTTCTGACCTTG (SEQ ID NO: 145)
 3 H3-48 GGTTTGTGTCTGGTGTCCACTGAC (SEQ ID NO: 146)
 5 H5-a GAGTCTGTGCCGGAAGTGCAGCTGG (SEQ ID NO: 147)

Específico para secuencia codificadora

5 VH1 TATCAGGTGCAGCTGGTGCAG (SEQ ID NO: 148)
 5 VH3 TATCAGGTGCAGCTGGTGGAG (SEQ ID NO: 149)
 5 VH5 TATGAGGTGCAGCTGGTGCAG (SEQ ID NO: 150)
 3 VH1/3 ATATCTCTCGCACAGTAATACAC (SEQ ID NO: 151)

 3 VH3 ATATCTCTCGCACAGTAATATAC (SEQ ID NO: 152)
 3 VH5 ATATGTCTCGCACAGTAATACAT (SEQ ID NO: 153)
 5 VK1 TATGACATCCAGATGACCCAGTCTCCATCCTC (SEQ ID NO: 154)
 3 DPK9 ATAGGAGGGGTAAGTGTAACT (SEQ ID NO: 155)
 3 DPK1 ATAGGAGGGAGATTATCATA (SEQ ID NO: 156)
 5 DPK22_L6 TATGAAATTGTGTTGACGCAGTCT (SEQ ID NO: 157)
 3 DPK22 ATAGGAGGTGAGCTACCATACTG (SEQ ID NO: 158)
 5 DPK21 TATGAAATAGTGATGACGCAGTCT (SEQ ID NO: 159)
 3 DPK21 ATAGGAGGCCAGTTATTATACTG (SEQ ID NO: 160)
 3 L6 CAGCGTAGCAACTGGCCTCCTAT (SEQ ID NO: 161)
 5 DPL2 TACAGTCTGTGCTGACTCAG (SEQ ID NO: 162)
 3 DPL2 ATAGGACCATTCAGGCTGTCATC (SEQ ID NO: 163)
 5 DPL5 TATCAGTCTGTGTTGACGCAG (SEQ ID NO: 164)
 3 DPL5 ATAGGAGCACTCAGGCTGCTAT (SEQ ID NO: 165)

Combinaciones de cebadores usados para amplificar los genes de línea germinal seleccionados

		1ra PCR	3'	2nd PCR	3'
Familia	Línea germinal	5'	3'	5	3'
VH1	DP-&75 HV 1-2	5 H1-2	3 H1-2	5 VH1	3 VH1/3
	DP-10 HV 1-69	5 H1-69	3 H1-69	5 VH1	3 VH1/3
	DP-14 HV 1-18	5 H1-18	3 H1-18	6 VH1	2 VH1/3
VH3	DP-49 HV 3-30	5 H3-30	3 H3-30	6 VH3	3 VH1/3
	DP-51 HV 3-48	5 H3-48	3 H3-48	5 VH3	3 VH1/3
	DP-47 HV 3-23	5 H3-23	3 H3-23	5 VH3	3 VH3
VH5	HV 5a	5 H5a	3 VH5	5 VHS	3 VH5
VK1	DPK-1 KV 1.33	5 K 1-33	3 K1-33	6 VK1	3 DPK-1
	DPK-9 KV1-39	5 K 1-39	3 K1- 9	6 VKI	3 DPK-9
VKIII	L6 KV 3-11	5 K3-11	3 K3-11	5 DPK22_L6	3 L6
	DPK-21 KV 3-15	6 K3-16	3 K3-15	5 DPK21	3 DPK21
	DPK-22 KV 3.20	5 K3-20	3 K3-20	5 DPK22_L6	3 DPK22
VL1	DPL-2 LV 1-44	5 L1-44	3 L1-44	5 DPL2	3 DPL2
	DPL-5 LV 1-51	6 L1-61	3 L1-51	5 DPL5	3 DPL5

EJEMPLO 2: Generación de estructuras aceptoras

- 5 Se analizaron las secuencias de los genes de la línea germinal seleccionados para la presencia de los sitios de restricción Tipo IIs. Ningún sitio BsmBI estaba presente en los genes de la línea germinal de los anticuerpos variables seleccionados. Se hallaron dos sitios BsmBI en el esqueleto de pNDS1, el vector fagémico en el que se puede clonar la estructura aceptora. Estos dos sitios se eliminaron por mutagénesis dirigida al sitio de modo que los sitios BsmBI únicos se pueden introducir en las secuencias de ADN de relleno de las estructuras aceptoras. Cada gen de la línea germinal se amplificó por PCR anidada múltiple con el fin de añadir una secuencia de ADN de relleno en el extremo 3 'de la secuencia FR3 seguido por una secuencia que codifica FR4 que es específica para cada segmento variable correspondiente (VH, Vk, Vλ). La secuencia de aminoácidos de VH FR4 corresponde a la región FR4 codificada por los genes de la línea germinal J JH1, JH2, JH4 y JH5. La secuencia de aminoácidos de VK FR4 corresponde a la región FR4 codificada por los genes de la línea germinal J JK1. La secuencia de aminoácidos de Vλ FR4 corresponde a la región FR4 codificada por la línea germinal J genes JL2 y JL3. Dos variantes de la secuencia de Vk FR4 se generaron con una sustitución de aminoácido única en la posición 106 (arginina o glicina). Para la estructura aceptora basada en el gen de la línea germinal VH3-23, también se construyeron dos variantes que difieren en un solo aminoácido (lisina a arginina) en la posición 94, el último residuo de FR3. Durante la etapa de amplificación final se introdujeron los sitios SfiI/NcoI y XhoI en el extremo 5 'y 3' final de la VH, respectivamente.
- 10
- 15
- 20 Del mismo modo, se introdujeron los sitios Sall y NotI en el extremo 5 'y 3' final de la VL, respectivamente (Figura 6). El fragmento de relleno se diseñó de modo que se cambió el marco de lectura de traducción, de este modo se evita la expresión de cualquier proteína funcional a partir de las estructuras aceptoras (Figura 7). Los cebadores usando en este proceso se listan a continuación.

VH

5 VH1 CAGCCGGCCATGGCCAGGTGCAGCTGGTGAC (SEQ ID NO: 166)
 5 VH3-30 CAGCCGGCCATGGCCAGGTGCAGCTGGTGGAG (SEQ ID NO: 167)
 5 VH3-23 CAGCCGGCCATGGCCAGGTGCAGCTGTTGGAG (SEQ ID NO: 168)
 5 VH3-48 CAGCCGGCCATGGCCAGGTGCAGCTGGTGGAGTCTGGGGGAG (SEQ ID NO: 169)
 5 VH5-51 CAGCCGGCCATGGCCAGGTGCAGCTGGTGAC (SEQ ID NO: 170)
 3 VH1/3 CTTACCGTTATTCGTCTCATCTCGCACAGTAATACAC (SEQ ID NO: 171)
 3 VH3-23 CTTACCGTTATTCGTCTCATTTTCGCACAGTAATATAC (SEQ ID NO: 172)
 3 VH3-48 CTGCGACAGTAATACACAGCCGTGCTCGGCTCTCAGGCTG (SEQ ID NO: 173)
 3 VH5-51 CTTACCGTTATTCGTCTCATCTCGCACAGTAATACAT (SEQ ID NO: 174)
 3 VHex1 CAATACGCGTTTAAACCTGGTAAACCGCCTTACCGTTATTCGTCTCA (SEQ ID
 NO: 175)
 3 VHex2 GTTCCCTGGCCCAAGAGACGCGCCTTCCCAATACGCGTTTAAACCTG (SEQ ID
 NO: 176)
 3 VHex3 CCTCCACCGCTCGAGACTGTGACCAGGGTTCCTGGCCCAAGAG (SEQ ID
 NO: 177)

VK

5 VK1 CGGGTCGACGGACATCCAGATGACCCAGTC (SEQ ID NO: 178)
 5 VK3-11 CGGGTCGACGGAAATTGTGTTGACACAGTCTCCAGC (SEQ ID NO: 179)
 5 VK3-15 CGGGTCGACGGAAATAGTGATGACGACAGTCTCCAGC (SEQ ID NO: 180)
 5 VK3-20 CGGGTCGACGGAAATTGTGTTGACGACAGTCTCCAGG (SEQ ID NO: 181)
 3 VK1-33 CCTTACCGTTATTCGTCTCGCTGCTGACAGTAATATGTTGCAATA (SEQ ID NO: 182)
 3 VK1-39 CCTTACCGTTATTCGTCTCGCTGCTGACAGTAGTAAGTTGCAAAA (SEQ ID NO: 183)
 3 VK3 CCTTACCGTTATTCGTCTCGCTGCTGACAGTAATAAACTGCAAAATC (SEQ ID NO: 184)
 3 VKext1 CCAATACGCGTTTAAACCTGGTAAACCGCCTTACCGTTATTCGTCTC (SEQ ID
 NO: 185)

3 VKext2 GGTCCTTGGCCGAATGAGACGCGCCTTCCCAATACGCGTTTAAAC (SEQ ID
 NO: 186)
 3 Vkext3R GTGCGCCGCCCGTTTGATTTCCACCTTGGTCCCTTGGCCGAATG (SEQ ID
 NO: 187)
 3 VKext3G GTGCGCCGCCCGTTTGATTTCCACCTTGGTCCCTTGGCCGAATG (SEQ ID
 NO: 188)

VL

5 VL1-44 CGGGTCGACGACAGTCTGTGCTGACTCAGCCAC (SEQ ID NO: 189)
 5 VL1-51 CGGGTCGACGACAGTCTGTGTTGACGACGCCGC (SEQ ID NO: 190)
 3 VL1-44 CCTTACCGTTATTCGTCTCCTGCTGCACAGTAATAATC (SEQ ID NO: 191)
 3 VL1-51 CCTTACCGTTATTCGTCTCCTGTTCCGACAGTAATAATC (SEQ ID NO: 192)
 3 Vlex2 CCCTCCGCCGAACACAGAGACGCGCCTTCCCAATACGCGTTTAAAC (SEQ ID
 NO: 193)
 3 Vlex3 GTGCGCCGCCCGCTTAGGACGGTACAGTCTGGTCCCTCCGCCGAACACAGA (SEQ ID
 NO: 194)

Las secuencias de las 20 estructuras ensambladas finales se muestran en la Figura 8.

5 EJEMPLO 3: Generación de vectores aceptores fagémidos que contienen un dominio variable invariante

El vector fagémido pNDS1 utilizados para la expresión de scFv se modificó primero en retirar dos sitios BsmBI. Un dominio VH3-23 que contiene una secuencia de CDR3 definida se clonó en el pNDS1 modificado usando los sitios de restricción SfiI y XhoI para obtener el vector fagémido pNDS_VHdummy. Este dominio contenía un sitio BsmBI en la región FR4, que se corrigió por mutagénesis dirigida al sitio silente. En paralelo, luego se clonó un dominio VK1-39 contiene una secuencia de CDR3 definida en el pNDS 1 modificados usando los sitios de restricción Sall y NotI para obtener el vector fagémido pNDS_VKdummy (Figura 9). Las 8 estructurasceptoras VH se clonaron en PNDs_VHdummy utilizando los sitios de restricción Sall y NotI. Las 12 estructurasceptoras VL se clonaron en pNDS_VHdummy utilizando los sitios de restricción SfiI y XhoI. Los vectores fagémidos 20 PNDs resultantes que se enumeran a continuación en esta etapa se pueden utilizar para la clonación de CDR3 diversificada usando los sitios de BsmBI presentes en los fragmentos de ADN de relleno.

Aceptores VH: pNDS_VH1-2_VKd; pNDS_VH1-18_VKd; pNDS_VH1-69_VKd; pNDS_VH3-23R_VKd; pNDS_VH3-23KVKd; pNDS_VH3-30_VKd; pNDS_VH5-51_VKd; pNDS_VH3-48_VKd.

Aceptores VL: pNDS_VHd_VK1-33G; pNDS_VHd_VK1-33R;

pNDS_VHd_VK1-39G; pNDS_VHd_VK1-39R; pNDS_VHd_VK3-11G;

pNDS_VHd_VK3-11R; pNDS_VHd_VK3 -15 G; pNDS_VHd_VK3-15R; pNDS_VHd_VK3-20G; pNDS_VHd_VK3-20R; pNDS_VHd_VL1-44; pNDS_VHd_VKI -51.

5 EJEMPLO 4: Captura de diversidad de CDR H3 natural de los repertorios humanos

Se usaron múltiples fuentes de ADNc humano se utilizaron como molde para la amplificación de las secuencias de CDR H3. Estas fuentes incluyeron bazo fetal humano, así como mezclas de células purificadas de sangre periférica de varones y mujeres adultos normales. Se han usado varias estrategias para la amplificación con el fin de recuperar las secuencias de CDR H3 procedentes de ADNc de VH reordenado codificado por un gen de la línea germinal específico o secuencias de CDR H3 procedentes de cualquier ADNc de VH.

En primer lugar, se usaron mezclas de cebadores que se aparean con las regiones codificadoras 5' de la mayoría de las familias VH humanas en combinación con mezclas de cebadores que se aparean con todas las regiones JH humanas. Esto permitió la amplificación por PCR de la mayoría de los genes variables de inmunoglobulina de cadena pesada. Los productos de amplificación esperados de aproximadamente 400 pares de bases (pb) se aislaron por electroforesis en gel de agarosa y se purificaron. Este ADN sirvió como molde en una segunda etapa de PCR usando cebadores con una coincidencia de 13 pb y 14 pb para la región FR3 final y el comienzo de FR4, respectivamente. En la mayoría de los casos, el último residuo de la FR3 es una arginina o una lisina. Debido a que las últimas coincidencias de pb son críticas para la extensión del cebador por la polimerasa, se utilizaron dos cebadores 5' diferentes: 5 VHR_FOK (SEQ ID NO: 205 se muestra a continuación) y 5 VHK FOK (SEQ ID NO: 206 se muestra a continuación). Es importante destacar que estos cebadores también contienen un sitio de restricción FokI para la escisión de la secuencia de CDR H3 (Figura 4). Los cebadores usados en la segunda etapa de PCR se biotinilaron en su extremo 5' para facilitar etapas de purificación corriente abajo (ver el Ejemplo 5). Este enfoque de dos etapas permite una amplificación eficiente de las secuencias de CDR H3 a pesar del número limitado de coincidencias de pares de bases. Las amplificaciones se realizaron a diferentes temperaturas de apareamiento (entre 30 °C y 70 °C) y con varias ADN polimerasas termoestables para establecer condiciones óptimas. Se halló que una temperatura de apareamiento de 55-58 °C en combinación con GoTaq polimerasa (Promega) era óptimo para este conjunto de cebadores. El segundo producto de amplificación se separó en un gel de agarosa al 2% y produjo una mancha en la parte inferior del gel correspondiente a CDR H3 de diferente longitud. Se extrajo la mancha de ADN completa del gel o una región correspondiente a los fragmentos de ADN más grandes con el fin de enriquecer la CDR H3 larga.

Alternativamente, se llevó a cabo la primera etapa de amplificación usando el cebador 5' 5 VH3-23H2 (SEQ ID NO: 201 mostrado a continuación), que es específico para la secuencia que codifica la CDR H2 de la VH3-23 de la línea germinal. Como los diferentes genes de la línea germinal son diversos en esta CDR, con preferencia se pueden amplificar los ADNc de VH codificados por el gen de la línea germinal seleccionado. Las etapas de purificación y amplificación posteriores fueron idénticas. De esta manera, es posible recuperar CDRs procedentes de un ambiente estructural específico y volver a introducirlos en el mismo, una estructura similar o diferente.

A continuación se muestra una lista de los cebadores utilizados para la amplificación de los repertorios de CDR humana natural

1era etapa de PCR

5 VH1/5 CCGCACAGCCGGCCATGGCCCAGGTGCAGCTGGTGCAGTCTGG (SEQ ID NO: 195)
 5 VH3 CCGCACAGCCGGCCATGGCCGAGGTGCAGCTGGTGGAGTCTGG (SEQ ID NO: 196)
 5 VH2 CCGCACAGCCGGCCATGGCCCAGRTCACCTTGCTCGAGTCTGG (SEQ ID NO: 197)
 5 VH4 CCGCACAGCCGGCCATGGCCCAGGTGCAGCTGCAGGAGTCGGG (SEQ ID NO: 198)
 5 VH4DP64 CCGCACAGCCGGCCATGGCCCAGCTGCAGCTGCAGGAGTCCGG (SEQ ID NO: 199)
 5 VH4DP63 CCGCACAGCCGGCCATGGCCCAGGTGCAGCTACAGCAGTGGGG (SEQ ID NO: 200)
 5 VH3-23H2 TGGAGTGGGTCTCAGCTATTAGTGGTAGTGGT (SEQ ID NO: 201)
 5 HJ1/2 CGATGGGCCCTTGTTGGAGGCTGAGGAGACRGTGACCAGGGTGCC (SEQ ID NO: 202)
 3 HJ3/6 CGATGGGCCCTTGTTGGAGGCTGAAGAGACGGTGACCRKGTCCC (SEQ ID NO: 203)
 3 HJ4/5 CGATGGGCCCTTGTTGGAGGCTGAGGAGACGGTGACCAGGGTTCC (SEQ ID NO: 204)

2da etapa de PCR

5 VHR_FOK GAGCCGAGGACACGGCCGGATGTTACTGTGCGAGA (SEQ ID NO: 205)
 5 VHK_FOK GAGCCGAGGACACGGCCGGATGTTACTGTGCGAAA (SEQ ID NO: 206)
 3 JH1_FOK GAGGAGACGGTGACGGATGTGCCCTGGCCCCA (SEQ ID NO: 207)
 3 JH2_FOK GAGGAGACGGTGACGGATGTGCCACGGCCCCA (SEQ ID NO: 208)
 3 JH3456_FOK GAGGAGACGGTGACGGATGTYCCTTGGCCCCA (SEQ ID NO: 209)

EJEMPLO 5: Generación de bancos primarias por clonación de la CDR H3 humana natural en estructuras aceptoras

La CDR H3 amplificada se digirió con FokI y las extremidades escindidas, así como ADN sin digerir se eliminaron usando perlas magnéticas recubiertas con estreptavidina. En paralelo, los vectores aceptores pNDs VH se digirieron usando BsmBI. Debido a que las proyecciones generadas por estas digestiones son compatibles, la colección de CDR H3 naturales fue capaz de ligarse en la estructura aceptora VH, lo que restaura el marco de lectura apropiado. El ADN ligado se purificó y se concentró para la transformación en células competentes E. coli XL1 Blue, y los clones aleatorios se analizaron por secuenciación con el fin de comprobar que la secuencia CDR H3 había sido reconstituida y que las uniones entre la CDR y la región estructural son correctas (Figura 10). Los resultados indicaron que todos los clones contenían secuencias de CDR H3 y que se restauró el marco de lectura, de este modo se codifica una cadena pesada variable de inmunoglobulina. Además, todas las CDRs eran diferentes, lo que indica que una gran diversidad de secuencias de naturales había sido capturada por este enfoque. La longitud de la CDR H3 también fue variable y se hallaron CDR relativamente largas de 10 a 15 residuos, lo que destaca la ventaja de este enfoque para el muestreo de secuencias de CDR largas que son difíciles de cubrir usando diversidad sintética.

Usando este método, las secuencias de CDR H3 naturales, derivadas de células purificadas de de sangre periférica humana mezclada o de bazo fetal humano, se clonaron en cada una de las estructuras aceptoras VH pNDS y se transformaron en células TG1 E. coli electrocompetentes y se sembraron en placas de bioensayo 2xTYAG (medio 2x TY que contiene ampicilina 100 µg/ml y glucosa 2%). Después de la incubación durante la noche a 30 °C, se añadieron 10 ml de medio líquido 2xTYAG a las placas y las células se rasparon de la superficie y se transfirieron a un tubo de polipropileno de 50 ml. Se añadió 2xTYAG que contiene glicerol 50% a la suspensión celular para obtener una concentración final de glicerol 17%. Las alícuotas de los bancos se almacenaron a -80 °C. En este proceso, se generaron 14 bancos primarios que representa un total de 8.1×10^9 transformantes. 180 clones elegidos aleatoriamente se secuenciaron para determinar la calidad y la diversidad de los bancos. Todos los clones codificaron diferentes secuencias de VH y > 89% estaban en marco. Estos bancos primarios contienen la diversidad en el CDR H3 sólo cuando se combinan con un dominio VL simulado.

EJEMPLO 6: Generación de bancos primarias por clonación de CDR3 sintética en las estructuras aceptoras

Aunque el método es de particular interés para la recuperación de la diversidad natural, también se puede aplicar para la integración de la diversidad sintética en estructuras aceptoras. Se diseñaron secuencias CDR3 sintéticas para VH y VL. El diseño tuvo en cuenta la frecuencia de los CDR con una determinada longitud y la estrategia de diversificación (codones NNS, DVK, NVT o TVP) que permitiría una cobertura completa de la diversidad teórica dentro de un número razonable de transformantes en una biblioteca ($\sim 5 \times 10^9$ transformantes) (Figura 11). Los residuos clave para mantener la estructura canónica de la CDR se mantuvieron constantes en el diseño de CDR3 para las cadenas VK y VLs. Para la cadena pesada, solo se generaron CDR3 con hasta 10 posiciones diversificadas como el número de clones necesarios para cubrir la diversidad codificada por CDRs más largas más allá de límites prácticos de eficiencia de transformación.

Se sintetizaron oligonucleótidos degenerados de diferente longitud usando los codones NNS, NVT, DVK o TVP aleatorizados. Para cada CDR H3, se sintetizaron dos oligonucleótidos que codifican una metionina o una fenilalanina en la posición 100Z (Figura 11). Cada oligonucleótido se extendió y se amplificó con dos cebadores biotinilados externos para generar fragmentos de ADN de doble cadena que codifican las CDRs diseñadas. Estos cebadores externos contienen sitios de restricción BsmBI para su posterior escisión de la secuencia de CDR y la inserción en las estructuras aceptoras (Figura 12). Los fragmentos de ADN ensamblados se procesaron sin purificación en gel y se digirieron con BsmBI. Las extremidades escindidas, así como ADN sin digerir se eliminaron usando perlas magnéticas recubiertas con estreptavidina. Los fragmentos de ADN digeridos se concentraron por precipitación con etanol y se ligaron en los correspondientes vectores aceptores pNDS VH, VK o VL. Los productos de ligamiento se purificaron y se concentraron para su transformación en células E. coli TG 1 electrocompetentes y se sembraron en placas de bioensayo 2xTYAG (medios 2xTY contienen 100 µg/ml ampicilina y glucosa 2%). Después de la incubación durante la noche a 30 °C, se añadieron 10 ml de medio líquido 2xTYAG a las placas y las células se rasparon de la superficie y se transfirieron a un tubo de polipropileno de 50 ml. Se añadió 2xTYAG que contiene glicerol 50% a la suspensión celular para obtener una concentración final de glicerol 17%. Las alícuotas de los bancos se almacenaron a -80 °C. Se generó un total de 24 bancos de cadena pesada primarias que representa un total de 1.6×10^{10} transformantes. Del mismo modo, se generaron 13 bancos de cadena liviana primaria que representa un total de 6.9×10^9 transformantes. Estos bancos primarios contienen diversidad solo en CDR H3 cuando se combinan con un dominio VL simulado. Un total de 330 clones seleccionados aleatoriamente se secuenciaron para determinar la calidad y la diversidad de los bancos. Todos los clones codifican diferentes

5 secuencias de dominio variable y > 90% fueron en el marco. Esta baja frecuencia de secuencias que contienen cambios en el marco de lectura está en marcado contraste con los resultados obtenidos tradicionalmente durante la construcción de bancos de fragmento de anticuerpo sintético utilizando los métodos de PCR por superposición que son más propensos a la introducción de la inserción, y se ha informado con frecuencia pérdida significativa de clones funcionales (15- 45%).

L44W_4nns CGTCTCCAGCATGGGATNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 241)
 L44Y_4nns CGTCTCCAGCATATGATNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 242)
 L44W_5nns CGTCTCCAGCATGGGATNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 243)
 L44Y_5nns CGTCTCCAGCATATGATNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 244)
 L44W_6nns CGTCTCCAGCATGGGATNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 245)
 L44Y_6nns CGTCTCCAGCATATGATNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 246)

5 L51W_biot CGGTCAGTCGCAATACGTCTCGAACATGGGAT (SEQ ID NO: 247)
 5 L51Y_biot CGGTCAGTCGCAATACGTCTCGAACATATGAT (SEQ ID NO: 248)
 L51W_4nns CGTCTCGAACATGGGATNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 249)
 L51Y_4nns CGTCTCGAACATATGATNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 250)
 L51W_5nns CGTCTCGAACATGGGATNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 251)
 L51Y_5nns CGTCTCGAACATATGATNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 252)
 L51W_6nns CGTCTCGAACATGGGATNNSNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 253)
 L51Y_6nns CGTCTCGAACATATGATNNSNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEQ ID NO: 254)

EJEMPLO 7: Generación de bancos secundarias

Con el fin de generar bancos de scFv que llevan la diversidad tanto en las cadenas pesadas como livianas, las bancos de cadena liviana sintética primarias se combinaron con las bancos de cadena pesada sintéticas primarias o las bancos de cadena pesada primarias naturales (Figura 13). Se preparó el ADN del fagémido a partir de cada biblioteca primaria y se digirió con las enzimas de restricción XhoI/NotI. Los fragmentos de ADN correspondientes al ligador y las cadenas livianas de los bancos sintéticos primarios se insertaron por ligamiento en los vectores de cadena pesada naturales o sintéticos primarios digeridos. Alternativamente, la secuencia de ligador-VL también se amplificó con cebadores específicos antes de la digestión con XhoI/NotI y ligamiento. Los productos de ligamiento se purificaron por extracción con fenol/cloroformo y precipitación antes de la transformación en células E. coli TG 1 electrocompetentes y se sembraron en placas de bioensayo 2xTYAG (medios 2xTY contienen 100 µg/ml ampicilina y glucosa 2%). Después de la incubación durante la noche a 30 °C, se añadieron 10 ml de medio líquido 2xTYAG a las placas y las células se rasparon de la superficie y se transfirieron a un tubo de polipropileno de 50 ml. Se añadió 2xTYAG que contiene glicerol 50% a la suspensión celular para obtener una concentración final de glicerol 17%. Las alícuotas de los bancos se almacenaron a -80 °C. Para limitar el número de bancos para recombinar, que se agruparon por las subclases de cadena (es decir, VH1, VH3, VH5, VK1, VK3, Vλ1) y por lo tanto se realizaron 9 combinaciones de biblioteca para (es decir, VH1xVK1, VH1xVK3, VH1Vλ1, VH3xVK1, VH3xVK3, VH3xVλ1, VH5xVK1, VH5xVK3, VH5xVλ1). El tamaño total de los bancos sintéticos secundarios (que llevan la diversidad sintética en VH y VL) fue $7,3 \times 10^9$ transformantes. El tamaño total de las bancos naturales secundarias (que llevan la diversidad natural en la VH y la diversidad sintética en el VL) fueron de $1,5 \times 10^{10}$ transformantes.

EJEMPLO 8: Generación de bancos de anticuerpo humanas que presentan un repertorio de CDRH3 derivado de una especie no humana.

A fin de utilizar fuentes alternativas de diversidad que permitan explorar un espacio tridimensional diferente dentro del sitio de combinación del anticuerpo, se creó una biblioteca para la captura de la CDRH3 de los ratones y se la introdujo en una colección de estructura de anticuerpos humanos. Para este enfoque se construyó una biblioteca de aceptor que contiene una colección de genes VL sintético con diversidad CDR L3 y se combinó con una colección de secuencias aceptoras que contienen una secuencia de ADN de relleno lista adecuado para la clonación de restricción de Tipo IIs como se describe en el Ejemplo 2. Esta biblioteca representa el punto de partida punto para la rápida generación de bancos secundarias con múltiples fuentes de CDR H3 natural (humano, así como no humano) o sintéticas. En este ejemplo, la diversidad natural de CDR H3 fue capturada de ratones Balb/c sin exposición previa y ratones que habían sido inmunizados con hIFNγ o hCCL5 (hRANTES).

La primera etapa fue la generación de bancos de aceptor mediante la clonación de una colección de VL que contiene diversidad de CDR L3 sintética en vectores de estructura de VH aceptoras (Figura 14). Las secuencias de VL se derivaron de los siete bancos sintéticos primarios descriptas en el Ejemplo 6 mediante amplificación por PCR

usando cebadores 5'biot-VHdummy y 3'biot-fdtseq. Se digirieron las VL resultante que contienen fragmentos de aproximadamente 400 pb usando XhoI/NotI y se purificó en columnas de centrifugación para eliminar los cebadores y enzimas. Del mismo modo se digirieron los vectores aceptores pNDs VH que contienen una CDRH3 de relleno y una cadena liviana simulada con XhoI/NotI y Swal (Swal corta dentro de VL simulada) y se purificó en columnas Chroma Spin TE con un corte de 1000 pb para liberarse del fragmento simulado de VL. Los fragmentos VL digeridos luego se ligaron en los vectores aceptores de VH (Figura 14). Para, limitar el número de bancos para recombinar, los vectores aceptores de VH y fragmentos de VL se agruparon por las subclases de cadena (es decir, VH1, VH3, VH5, VK1, VK3, Vλ1) y por lo tanto se realizaron 9 combinaciones de biblioteca para (es decir, VH1xVK1, VH1xVK3, VH1xVλ1, VH3xVK1, VH3xVK3, VH3xVλ1, VH5xVK1, VH5xVK3, VH5xVλ1). Los productos de ligamiento se transformaron en células E. coli TG 1 electrocompetentes y se sembraron en placas de bioensayo 2xTYAG (medios 2xTY contienen 100 µg/ml ampicilina y glucosa 2%). Después de la incubación durante la noche a 30 °C, se añadieron 6 ml de medio líquido 2xTYAG a las placas y las células se rasparon de la superficie y se transfirieron a un tubo de polipropileno de 50 ml. Se añadió glicerol a la suspensión celular para obtener una concentración final de glicerol 17%. Las alícuotas de las bancos se almacenaron a -80 °C. El tamaño total de esta biblioteca aceptora que llevan la diversidad sintética en la CDR L3, fue $1,9 \times 10^9$ transformantes.

La próxima etapa fue aislar las secuencias de CDRH3 de una fuente no humana. Las células se aislaron del bazo de cinco ratones Balb/c sin exposición previa o inmunizados y se purificó el ARN total. Se obtuvo ADNc a partir del ARN extraído por RT-PCR. Este ADNc se usó como molde para aislar y amplificar VH de ratón por PCR. Se realizaron una serie de PCR usando 15 cebadores 5' diferentes (uno para cada subgrupo de VH de ratón) específicos para el comienzo de la región FR1 y una mezcla de cebadores 3' (cuatro cebadores que cubren la región JH). Estas primeras PCR se agruparon y se purificaron en un gel de agarosa 2%. El ADN purificado sirvió como molde para llevar a cabo una segunda PCR para aislar la región CDR H3 de ratón.

Los cebadores 5' y 3' para esta segunda PCR se dirigen a las regiones FR3 y FR4 de VH de ratón, respectivamente. Estos cebadores añadieron un sitio de restricción FokI con el fin de permitir la escisión precisa de la CDR H3 y la clonación en los vectores aceptores humanos.

Sin embargo, los alineamientos de secuencias de VH murinas revelaron que la secuencia en el límite 5' de CDR-H3 murina y que se encuentran en el sitio de escisión de FokI casi siempre difieren de la secuencia humana por una base, mientras que el extremo 3' coincidió entre estas dos especies. Las secuencias escindidas por FokI están recuadradas en la siguiente tabla:

	secuencias 5'		secuencias 3'		
(SEQ ID NO: 281)	Humano: T T A C T G T G C	<u>G A G A</u>	Humano: T G G G	G C C A G G G A A	(SEQ ID NO: 286)
(SEQ ID NO: 282)	ratón: V H 1	<u>A A G A</u>	ratón: J H 1	T G G G G C C G C A G G G A	(SEQ ID NO: 287)
(SEQ ID NO: 283)	V H 2	<u>A A G A</u>	J H 2	T G G G G C C A A G G C A	(SEQ ID NO: 288)
(SEQ ID NO: 284)	V H 3	<u>C A G A</u>	J H 3	T G G G G C C A G G G C A	(SEQ ID NO: 289)
(SEQ ID NO: 285)	V H 3-16	<u>A A G A</u>	J H 4	T G G G G T C A G G G C A	(SEQ ID NO: 290)

En consecuencia, la base tenía que ser corregida durante la segunda etapa de amplificación con el fin de generar extremos cohesivos que son compatibles con los extremos cohesivos generados después de la digestión de las estructuras aceptoras. Se observó la amplificación eficiente lo que sugiere que esta conversión se produjo fácilmente. En el extremo 3', las secuencias de ratón y humanas que se cortaron con las enzimas de restricción Tipo IIs son idénticas, de este modo se evita cualquier problema de corrección.

[00278] Los cebadores para la segunda amplificación se biotinizaron en sus 5' extremos para facilitar las etapas de purificación de corriente abajo. Los vectores aceptores se digirieron con BsmBI y se purificaron en columnas Chroma Spin TE que tienen un punto de corte de 1.000 pb. Después de la digestión y la purificación, las nueve combinaciones diferentes de la biblioteca se agruparon en relación equimolar para el ligamiento de la CDRH3 de ratón capturada.

El DNA ligado se purificó por extracciones de fenol /cloroformo y se concentró por precipitación antes de la transformación en células E. coli TGI competentes y se sembró en placas de bioensayo 2xTYAG (medios 2xTY contienen 100 µg/ml ampicilina y glucosa 2%). Después de la incubación durante la noche a 30 °C, se añadieron 6 ml de medio líquido 2xTYAG a las placas y las células se rasparon de la superficie y se transfirieron a un tubo de polipropileno de 50 ml. Se añadió glicerol a la suspensión celular para obtener una concentración final de glicerol 17%. Las alícuotas de las bancos se almacenaron a -80 °C. Se obtuvieron tres bancos de tamaño similar: Mna, $2,5 \times 10^8$ transformantes (que portan una diversidad estructural humana natural restringida, diversidad de ratón sin exposición previa en la CDR H3 y diversidad sintética en la CDR L3); MiB, $7,3 \times 10^7$ transformantes (que portan una diversidad estructural humana natural restringida, diversidad de ratón inmune contra h1FNy en la CDR H3 y diversidad sintética en la CDR L3) y MiC, $1,8 \times 10^8$ transformantes (que portan una diversidad estructural humana natural restringida, diversidad de ratón inmune contra hCCL5 en la CDR H3 y diversidad sintética en la CDR L3). Los clones aleatorios se analizaron por secuenciación con el fin de comprobar que la secuencia de CDR H3 había sido reconstituido y que las uniones entre la CDR y las regiones estructurales eran correctas. Los resultados indicaron

5 que todos los clones contenían secuencias de CDR H3 y que el marco de lectura fue restaurado, en consecuencia codifican una cadena pesada variable de inmunoglobulina. Todas las CDR eran diferentes y se parecían secuencias de CDR H3 de ratón típicas, lo que indica que una gran diversidad de secuencias de CDRH3 de ratón naturales fue capturada por este enfoque. Además, el análisis de los perfiles de longitud CDRH3 indicó que se obtuvo una distribución gaussiana en la biblioteca naive que se corresponde con la distribución esperada de longitudes en el repertorio normal de ratón. En contraste, en las dos bancos inmunes, los perfiles eran diferentes lo que sugiere que se ha obtenido un repertorio de CDRH3 diferente (Figura 15).

Cebadores usado para la amplificación de CDRH3 de ratones 1era PCR

10 • Cebadores 5':

m5	VH1	ATGCGGCCCAGCCGGCCATGGCCSAGGTYCAGCTBCAGCAGTC	(SEQ ID NO: 256)
m5	VH2	ATGCGGCCCAGCCGGCCATGGCCCAGGTTACCTGCAGCARTC	(SEQ ID NO: 257)
m5	VH3	ATGCGGCCCAGCCGGCCATGGCCCAGGTRCAGCTGAAGGAGTC	(SEQ ID NO: 258)
m5	VH4	ATGCGGCCCAGCCGGCCATGGCCCAGGTTCAACTVCAGCARCC	(SEQ ID NO: 259)
m5	VH5	ATGCGGCCCAGCCGGCCATGGCCCAGATCCAGTTGGTVCAGTC	(SEQ ID NO: 260)
m5	VH6	ATGCGGCCCAGCCGGCCATGGCCCAGGTGCAGCTGAAGSASTC	(SEQ ID NO: 261)
m5	VH7	ATGCGGCCCAGCCGGCCATGGCCGAGGTGCAGSKGGTGGAGTC	(SEQ ID NO: 262)
m5	VH8	ATGCGGCCCAGCCGGCCATGGCCGAAGTGAARSTTGAGGAGTC	(SEQ ID NO: 263)
m5	VH9	ATGCGGCCCAGCCGGCCATGGCCGAKGTSVAGCTTCAGGAGTC	(SEQ ID NO: 264)
m5	VH10	ATGCGGCCCAGCCGGCCATGGCCGAGGTGAASSTGGTGGAAATC	(SEQ ID NO: 265)
m5	VH11	ATGCGGCCCAGCCGGCCATGGCCGAGGTGAAGCTGRTGGARTC	(SEQ ID NO: 266)
m5	VH12	ATGCGGCCCAGCCGGCCATGGCCGARGTGAAGCTGRTGGAGTC	(SEQ ID NO: 267)
m5	VH13	ATGCGGCCCAGCCGGCCATGGCCGAAGTGCAGCTGTTGGAGAC	(SEQ ID NO: 268)
m5	VH14	ATGCGGCCCAGCCGGCCATGGCCGARGTGAAGCTTCTCSAGTC	(SEQ ID NO: 269)
m5	VH15	ATGCGGCCCAGCCGGCCATGGCCCARGTTACTCTGAAAGAGT	(SEQ ID NO: 270)

• Cebadores 3':

m3	HJ1	CCTGAACCGCCGCTCCGCTCGAGACGGTGACCGTGGTCCC	(SEQ ID NO: 271)
m3	HJ2	CCTGAACCGCCGCTCCGCTCGAGACTGTGAGAGTGGTGCC	(SEQ ID NO: 272)
m3	HJ3	CCTGAACCGCCGCTCCGCTCGAGACAGTGACCAGAGTCCC	(SEQ ID NO: 273)
m3	HJ4	CCTGAACCGCCGCTCCGCTCGAGACGGTGACTGAGGTTCC	(SEQ ID NO: 274)

1era PCR

15 • Cebadores 5':

5	VHR_FOK_biot	GAGCCGAGGACACGGCCGATGTTACTGTGCGAGA	(SEQ ID NO: 275)
---	--------------	------------------------------------	------------------

• Cebadores 3':

3'	mJH1_Fok_biot	GGGGCGCAGGGACATCCGTCACCGTCTCCTC	(SEQ ID NO: 276)
3'	mJH2_Fok_biot	GAGGAGACTGTGAGGGATGTGCCCTTGGCCCCA	(SEQ ID NO: 277)
3'	JH1_Fok	GAGGAGACGGTGACGGATGTGCCCTTGGCCCCA	(SEQ ID NO: 278)
3'	mJH4_Fok_biot	GAGGAGACGGTGACGGATGTTCTTGGACCCCA	(SEQ ID NO: 279)

EJEMPLO 9: Rescate de fagos de las bancos

20 Cada banco Primaria y Secundaria se rescató de forma independiente de acuerdo a los procedimientos de despliegue en fagos estándares que se resumen brevemente a continuación. Se añadió un volumen de célula a partir de las alícuotas de la biblioteca congeladas suficientes para cubrir por lo menos 10 veces la diversidad teórica de la biblioteca a 500 ml de 2xTYAG y se cultivó a 37 °C con agitación (240 rpm) hasta alcanzar una DO600 de 0,3 a 0,5. El cultivo se super-infectó con el fago auxiliar MK13K07 y se incubó durante una hora a 37 °C (150 rpm).
 25 continuación, el medio se cambió por centrifugación de las células a 2000 rpm durante 10 minutos, se eliminó el medio y resuspendió el sedimento en 500 ml de 2x TY-AK (100 µg/ ml ampicilina; 50 µg/ml kanamicina). A continuación, el cultivo se dejó crecer durante la noche a 30 °C (240 rpm). El cultivo se centrifugó a 4000 rpm durante 20 minutos para sedimentar las células. Se recolectó el sobrenadante y se añadió 30% (vol/vol) de PEG 8000 (NaCl (20%)/2,5 M de NaCl para precipitar las partículas de fago mediante la incubación de la mezcla 1 hora
 30 en hielo. Las partículas de fago se recolectaron por centrifugación a 10.000 rpm durante 30 minutos y se resuspendieron en 10 ml de buffer TE (10 mM de Tris-HCl pH 8,0; 1 mM de EDTA). La solución resuspendida se centrifugó a 10.000 rpm para eliminar los restos de bacterias y el procedimiento de precipitación se repitió. Después

de la resuspensión final, fago se tituló mediante infección de *E. coli* y la absorción a 280 nm. El nivel de despliegue de scFv en la superficie del fago también se evaluó mediante análisis de transferencia Western usando un anticuerpo monoclonal anti-c-myc. El fago purificado a partir de diferentes bancos se almacenó congelado a -80 °C después de la adición de glicerol a una concentración final de 15% (p/v).

- 5 Para utilizar un número manejable de las bancos durante los procedimientos de selección, el fago purificado se mezcló con 4 bancos de trabajo: AA1 - fagos de todas las bancos VH sintéticas primarias; AB1 - fago de todas las bancos VL sintéticas primarias; AC1 - fago de todas las bancos VH naturales primarias; AD1 - Fago de todas las bancos naturales secundaria; AE1 - Fago de todas las bancos sintéticas secundaria; MNA - Bancos con diversidad capturada de ratones sin exposición previa; MiB - Bancos con diversidad capturada de ratones inmunizados con hIFN γ ; MIC - Bancos con diversidad capturados de ratones inmunizados con hCCL5/RANTES.

EJEMPLO 10. Secuenciación de alto rendimiento de las bancos de anticuerpo.

15 La calidad y la diversidad de una biblioteca se pueden evaluar mediante secuenciación de ADN de los miembros de la biblioteca aleatorios. En la mayoría de los casos se secuencian unos cientos de clones que representan solo una muy pequeña fracción de la biblioteca (menos de 1 en 10 millones miembros de la biblioteca). Con el fin de analizar el rendimiento de los métodos proporcionados en la presente, se utilizó la tecnología de secuenciación de próxima generación para analizar un número más representativo de miembros de la biblioteca. El ADN aislado de la biblioteca AE1 se utilizó como molde para la secuenciación de alto rendimiento utilizando un instrumento Analizador de genoma Illumina. Este sistema de secuenciación de ADN de próxima generación permite leer miles de millones de bases en un par de días. Las lecturas de secuenciación son relativamente cortas (aproximadamente 70 bases) pero perfectamente compatible con nuestro diseño de la biblioteca. Como la diversidad se limita a las regiones CDR3 una lectura de 70 bases es suficiente para cubrir la CDRH3 y parte de la región estructural 3 para la identificación de la familia VH. Esta tecnología se ha aplicado para secuenciar varios millones de regiones CDRH3 de la biblioteca de AEI. Se obtuvieron 5.078.705 secuencias para un total de 365 666 760 bases. El análisis de los datos indicó que 5.007.022 secuencias (98,6% del total) eran únicas. Un total de 4.680.882 secuencias podría atribuirse inequívocamente a una familia VH (VH1, VH3 y VH5) y la representación de las familias VH en la biblioteca AE1 determinada (41% VH1; 30% VH3; 29% VH5). Un hallazgo importante fue que la proporción de los insertos en marco varió entre 88 y 91%. Estos datos confirmaron de una manera mucho más estadística los resultados de la secuenciación de los 24 bancos sintéticos VH primarios descritas en el Ejemplo 6. Este conjunto combinado de datos de secuenciación demuestra que el proceso de clonación de restricción tipo IIs utilizado en este método es muy robusto, lo que lleva a una la inserción eficiente y productiva en las 24 construcciones de bancos independientes realizadas para generar la diversidad VH del banco AEI.

35 La secuenciación de millones de miembros de la biblioteca representa una etapa de control de calidad sin precedentes para una biblioteca de anticuerpos. Los resultados demuestran que el método permite la generación de bancos de alta calidad y alta diversidad de una manera reproducible y robusta.

EJEMPLO 11: Selecciones de despliegue en fago usado bancos secundarias

Selecciones de fase líquida contra interferón gamma humano (hIFN γ):

40 Las alícuotas de las bancos de fagos AD1 y AEI (10^{11} - 10^{12} Pfu) se bloquearon con PBS que contiene 3% (p/v) de leche descremada durante una hora a temperatura ambiente en un mezclador rotatorio. Luego el fago bloqueado se deseleccionó sobre perlas magnéticas de estreptavidina (Dynal M-280) durante una hora a temperatura ambiente en un mezclador rotatorio. Luego el fago deseleccionado se incubó con hIFN γ biotinilado in vivo (100 nM) durante dos horas a temperatura ambiente en un mezclador rotatorio. Las perlas se capturaron utilizando un soporte magnético, seguido de cuatro lavados con PBS/Tween 20 0,1% y 3 lavados con PBS. Las perlas luego se añadieron directamente a 10 ml de células TG1 en crecimiento exponencial y se incubaron durante una hora a 37 °C con agitación lenta (100 rpm). Una alícuota de TG1 infectado 1 se diluyó en forma seriada para titular el producto de la selección. Las TG1 infectadas restantes se centrifugaron a 3000 rpm durante 15 minutos y se resuspendieron en 0,5 ml 2xTY (medio 2xTY que contienen ampicilina 100 μ g/ml y glucosa 2%) y se extendió sobre placas de bioensayo de agar 2xTYAG. Después de la incubación durante la noche a 30 °C, se añadieron 10 ml de 2xTYAG a las placas y las células se rasparon de la superficie y se transfirieron a un tubo de polipropileno de 50 ml. Se añadió 2xTYAG que contiene glicerol 50% a la suspensión celular para obtener una concentración final de glicerol 17%. Las alícuotas de la ronda de selección se mantuvieron a -80 °C. Los productos de fago se titularon después de cada ronda y el aumento progresivo de los productos indica que se está produciendo el enriquecimiento de los clones específicos para el blanco (Figura 16).

Selecciones por inmunopurificaciones contra el anticuerpo monoclonal 5E3 de rata:

55 Los inmunotubos se recubrieron con 5E3 a razón de 10 μ g/ml en PBS durante la noche a 4 °C y los inmunotubos para la desección fago se recubrieron con un anticuerpo de rata irrelevante en las mismas condiciones. Después del lavado, los inmunotubos se bloquearon con PBS que contiene 3% (p/v) de leche descremada durante una hora a temperatura ambiente. Las alícuotas de bancos de fagos de de AD1 y AE1 (10^{11} - 10^{12} Pfu) se bloquearon con PBS

que contiene 3% (p/v) de leche descremada durante una hora a temperatura ambiente en un mezclador rotatorio. Luego el fago bloqueado se deseleccionó en los inmunotubos recubiertos con un anticuerpo de rata irrelevante durante una hora a temperatura ambiente en un mezclador rotatorio. Luego el fago deseleccionado se transfirió a los inmunotubos recubiertos con 5E3 y se incubó durante dos horas a temperatura ambiente en un mezclador rotatorio.

5 Los tubos se lavaron cinco veces con PBS/0,1% de Tween 20 y 3 veces con PBS. El fago se eluyó con TEA 100 mM durante 10 minutos y se neutralizó con Tris 1M HC1 pH 7.5. El fago se añadió a 10 ml de células TG1 en crecimiento exponencial y se incubó durante una hora a 37 °C con agitación lenta (100 rpm). Una alícuota de las TG1 infectadas se diluyó en forma seriada para titular el producto de la selección. Las TG1 infectadas restantes se centrifugaron a 10 3000 rpm durante 15 minutos y se resuspendieron en 0,5 ml 2xTY (medio 2xTY que contienen ampicilina 100 µg/ml y glucosa 2%) y se extendió sobre placas de bioensayo de agar 2xTYAG. Después de la incubación durante la noche a 30 °C añadieron 10 ml de 2xTYAG a las placas y las células se rasparon de la superficie y se transfirieron a un tubo de polipropileno de 50 ml. Se añadió 2xTYAG que contiene glicerol 50% a la suspensión celular para obtener una concentración final de glicerol 17%. Las alícuotas de la ronda de selección se mantuvieron a -80 °C. Se realizaron rondas de selección alternando entre 5E3 de rata y una versión quimérica de 5E3 en los que la región variable estaba fusionada a los dominios constantes de ratón. Estas rondas alternadas se realizaron con el fin de enriquecer los clones específicos para la región variable de 5E3 y generar anticuerpos anti-idiotípicos. Los productos de fago se titularon después de cada ronda y el aumento progresivo de los productos indica que se está produciendo el enriquecimiento de los clones específicos para el blanco (Figura 17).

20 *Rescate de fago:* Se añadieron 100 µl de la suspensión celular obtenida de las rondas de selección anteriores a 20 ml de 2xTYAG y se cultivaron a 37 °C con agitación (240 rpm) hasta alcanzar una DO600 de 0,3 a 0,5. El cultivo se super-infectó con el fago auxiliar 3,3 x 10 MK13K07 y se incubó durante una hora a 37 °C (150 rpm). A continuación, el medio se cambió por centrifugación de las células a 2000 rpm durante 10 minutos, se eliminó el medio y se resuspendió el pellet en 20 ml de 2x TY-AK 100 µg/ ml ampicilina; 50 µg/ ml kanamicina). A continuación, el cultivo se dejó crecer durante la noche a 30 °C (240 rpm).

25 *Rescate del fago monoclonal por ELISA:* Se recolectaron clones individuales en una placa de microtitulación que contiene 150 µl del medio 2xTYAG (glucosa 2%) por pocillo y se cultivaron a 37 °C (100-120 rpm) durante 5-6h. El fago auxiliar M13K07 se añadió a cada pocillo para obtener una multiplicidad de infección (MOI) de 10 (es decir, 10 fagos para cada célula en el cultivo) y se incubó a 37 °C (100 rpm) durante 1 h. Después del crecimiento, las placas se centrifugaron a 3200 rpm durante 10 min. El sobrenadante se retiró cuidadosamente, las células se resuspendieron en medio 150 µl de medio 2xTYAG y se cultivaron durante la noche a 30 °C (120 rpm). Para el ELISA, el fago se bloqueó mediante la adición de 150 µl de PBS concentración 2x que contiene 5% de leche en polvo descremada seguido por una hora de incubación a temperatura ambiente. Las placas luego se centrifugaron 10 minutos a 3000 rpm y el sobrenadante que contiene el fago se usó para el ELISA.

35 *Fago ELISA:* Las placas ELISA (Maxisorb, NUNC) se recubrieron durante la noche con 2 µg/ ml h1FNy en PBS o 2 µg/ml de 5E3 de rata en PBS. Las placas de control se recubrieron con 2 µg/ml de BSA o un anticuerpo monoclonal de rata irrelevante. Después las placas se bloquearon con 3% leche descremada/PBS a temperatura ambiente durante 1 hora. Las placas se lavaron 3 veces con PBS 0,05% de Tween 20 antes de transferir los sobrenadantes de fagos pre-bloqueados y se incubaron durante una hora a temperatura ambiente. Las placas luego se lavaron 3 veces con PBS 0,05% Tween 20. Se añadieron 50 µl de leche descremada/ PBS que contiene anticuerpo (HRP) anti-M13 conjugado (HRP) (Amersham, diluido 1: 10.000) a cada pocillo. Después de la incubación a temperatura ambiente durante 1 h, las placas se lavaron 5 veces con PBS 0,05% Tween 20. El ELISA se reveló a continuación mediante la adición de 50 µl de TMB (Sigma) y 50 µl de H₂SO₄ 2N para detener la reacción. La intensidad de absorción se leyó a 450 nm. Los clones específicos para h1FNy se pueden identificar y las tasas de éxito variaron entre 10% y 30% después de la tercera ronda de selección. Los clones específicos para la región variable de 5E3 también se puede identificar y las tasas de éxito variaron entre 7 y 48% después de la tercera ronda de selección.

45 *Secuenciación del clon del fago:* Los clones individuales se cultivaron en 5 ml de medio de 2xTYAG (glucosa 2%) por pocillo y se cultivaron a 37 °C (120 rpm) durante la noche. El siguiente día el ADN fagémido se purificó y se usó para la secuenciación del ADN utilizando un cebador específico para pNDS1: mycseq, 5'-CTCTTCTGAGATGAGTTTTTG. (SEQ ID NO: 255).

50 *Purificación de scFv a gran escala:* Un cultivo iniciador de 1 ml de 2xTYAG se inoculó con una sola colonia de una placa de agar 2xTYAG recién sembrada en estrías y se incubó con agitación (240 rpm) a 37 °C durante 5 horas. Se usaron 0,9 ml de este cultivo para inocular un cultivo de 400 ml del mismo medio y se cultivó durante la noche a 30 °C con agitación vigorosa (300 rpm).

55 El día siguiente, el cultivo se indujo mediante la adición de 400 µl de IPTG 1 M y la incubación continuó durante 3 horas adicionales. Las células se recogieron por centrifugación a 5000 rpm durante 10 minutos a 4 °C. Las células sedimentadas se resuspendieron en 10 ml de buffer TES enfriado con hielo complementado con inhibidores de proteasa como se describió anteriormente. El shock osmótico se logró mediante la adición de 15 ml de buffer TES diluido 1: 5 y la incubación durante 1 hora en hielo. Las células se centrifugaron a 10.000 rpm durante 20 minutos a 4 °C para sedimentar los restos celulares. El sobrenadante se transfirió cuidadosamente a un tubo nuevo. Se añadió imidazol al sobrenadante a una concentración final de 10 mM. Se añadió 1 ml de resina Ni-NTA (Qiagen), equilibrada en PBS a cada tubo y se incubó en un mezclador rotatorio a 4 °C (20 rpm) durante 1 hora. Los tubos se

centrifugaron a 2000 rpm durante 5 minutos y se eliminó el sobrenadante cuidadosamente. La resina sedimentada se resuspendió en 10 ml de buffer de lavado 1 frío (4 °C) (50 mM de NaH₂PO₄, 300 mM de NaCl, 10 mM de imidazol, pH a 8,0). La suspensión se añadió a una columna de polyprep (Biorad). Se usaron 8 ml de buffer de lavado 2 frío (50 mM de NaH₂PO₄, 300 mM de NaCl, 10 mM de imidazol, pH a 8,0) para lavar la columna por flujo por gravedad. Los scFv se eluyeron de la columna con 2 ml de buffer de elución (50 mM de NaH₂PO₄, 300 mM de NaCl, 250 mM de imidazol, pH a 8,0). Las fracciones se analizaron por absorción a 280 nm y las fracciones que contienen proteínas se mezclaron antes de intercambio de buffer en una columna de desalinización PD10 (Amersham) equilibrada con PBS. Los scFv en PBS se analizaron por SDS-PAGE y se cuantificaron por absorción a 280 nm. Los scFv purificados se dividieron en alícuotas y se almacenaron a -20 °C y a 4 °C.

10

EJEMPLO 12. Análisis de los perfiles de CDR3 obtenidos después de la secuenciación de alto rendimiento.

Utilizando la tecnología de secuenciación de próxima generación como se describe en el Ejemplo 10, se analizó la distribución de las longitudes de CDR H3 dentro de cada familia VH en los bancos AE1 y AD1, así como en el producto obtenido después de la tercera ronda de selección. Los perfiles de los bancos AE1 y AD1 son claramente diferentes (Figura 18). La distribución de longitudes de CDR H3 en la biblioteca AE1 se corresponde con el diseño de la biblioteca prevista, con longitudes que varían entre 9-15 aminoácidos. En contraste, las CDR H3 mucho más largas de hasta 22 aminoácidos se encuentran en la biblioteca de AD1, y el perfil se corresponde con la distribución de la longitud observada en repertorios naturales humanos. Estos resultados confirman que se ha capturado un repertorio de CDR H3 natural humano en la construcción de la biblioteca de AD1. Un análisis similar realizado después de tres rondas de selección en contra de 5E3 reveló que se seleccionaron perfiles de longitud CDR H3 completamente diferentes. En particular, se puede observar un enriquecimiento muy importante de CDR H3 de 8 y 21 aminoácidos de longitud en la selección realizada con la biblioteca AD1. Este conjunto de datos demostró que diferentes perfiles de CDR H3 se enriquecieron a partir de los dos bancos después de la selección contra el mismo blanco. Además, este análisis demuestra que, usando la presente invención, las CDR H3 largas que son muy difíciles de cubrir usando diversidad sintética se pueden capturar en las estructuras seleccionadas humanas y seleccionar.

EJEMPLO 13. Evaluación de los scFvs identificados en los ensayos de unión.

Las preparaciones de scFv purificadas de los clones que tienen diferentes secuencias y que se identificaron como positivos contra la región variable de 5E3 se analizaron por la unión contra 5E3 quimérico en una ELISA de dosis-respuesta. Estas preparaciones también se analizaron contra un anticuerpo de ratón irrelevante (1A6). Las placas ELISA (Maxisorb, NUNC) se recubrieron durante la noche con 2 µg/ml de 5E3 de ratón en PBS. Las placas de control se recubrieron con 2 µg/ml de anticuerpo monoclonal 1A6. Después las placas se bloquearon con 3% de leche descremada/PBS a temperatura ambiente durante 1 hora. Las placas se lavaron 3 veces con PBS 0,05% Tween 20 antes de la adición de diferentes concentraciones de scFv purificado e incubación durante una hora a temperatura ambiente. Las placas se lavaron a continuación 3 veces con PBS 0,05% Tween 20. Se añadieron 50 µl de 3% de leche descremada/PBS que contiene anticuerpo anti-myc conjugado (HRP) a cada pocillo. Después de la incubación a temperatura ambiente durante 1 h, las placas se lavaron 5 veces con PBS 0,05% Tween 20. Luego el ELISA se reveló mediante la adición de 50 µl de sustrato fluorescente Amplex Red y la señal se leyó en un espectrofotómetro de fluorescencia. Los datos muestran que la mayoría de los clones son altamente específicos para 5E3, ya que no reconocen 1A6 y se dirigen contra las regiones variables de 5E3 (Figura 19).

Del mismo modo, las preparaciones de scFv purificadas de clones que tienen diferentes secuencias y que fueron identificados en ELISA de fagos como aglutinantes contra h1FNy se analizaron por la unión contra h1FNy en un experimento de dosis respuesta: Las placas ELISA (Maxisorb, NUNC) se recubrieron durante la noche con 2 µg/ml h1FNy en PBS o las placas de control se recubrieron con 2 µg/ml de BSA. Después las placas se bloquearon con 3% leche descremada/PBS a temperatura ambiente durante 1 hora. Las placas se lavaron 3 veces con PBS 0,05% de Tween 20 antes de añadir diferentes concentraciones de scFv purificado y se incubaron durante una hora a temperatura ambiente. Las placas luego se lavaron 3 veces con PBS 0,05% Tween 20. Se añadieron 50 µl de leche descremada/PBS que contiene anticuerpo (HRP) anti-M13 conjugado (HRP) (Amersham, diluido 1:10.000) a cada pocillo. Después de la incubación a temperatura ambiente durante 1 h, las placas se lavaron 5 veces con PBS 0,05% Tween 20. El ELISA se reveló a continuación mediante la adición de 50 µl de TMB (Sigma) y 50 µl de H₂SO₄ 2N para detener la reacción. La señal se leyó en un espectrofotómetro de absorbancia a 450 nm. Los datos muestran que los clones seleccionados se unen a h1FNy de manera dependiente de la dosis y proporcionaron una muy buena señal en comparación con un control positivo scFv A6 que tiene una alta afinidad por h1FNy (Figura 20).

EJEMPLO 14: Inhibición de ScFv de la expresión del gen indicador inducido por interferón gamma

Un panel de scFv específico seleccionado para h1FNy se produjo y purificó como se describió anteriormente y se analizó por la capacidad de bloquear la actividad biológica de h1FNy. Un gen indicador (luciferasa de luciérnaga), dirigido por el promotor GBP1 inducible por IFN-γ, se transfectó en la línea celular de melanoma humano, Me67.8.

Diversas concentraciones de scFv se incubaron con 2 ng/ml de hIFN γ y luego se añadieron al cultivo celular. Después de un tiempo de incubación de 6 horas, se realizó el ensayo de indicador de luciferasa y se midió la intensidad de la luminiscencia. La actividad se comparó con un scFv aislado a partir de otra biblioteca de anticuerpo scFv humano construida por captura tradicional de los repertorios VH/VL de donantes humanos (clon G9). Los datos muestran que scFv aislado ya sea a partir de bancos de diversidad humana sintéticas como natural (AE1 y AD1) eran capaces de neutralizar la actividad biológica de hIFN γ de una manera dependiente de la dosis (Figura 21). El potencial de neutralización de estos scFv fue superior al del clon G9 de scFv de referencia.

EJEMPLO 15: Inhibición de ScFv de la expresión de MHC Clase II inducida por interferón gamma

Un ensayo de citometría de flujo se llevó a cabo para identificar anticuerpos IgG completamente humanos, o fragmentos de los mismos, capaces de bloquear la expresión de de las moléculas de MHC de clase II inducida por IFN γ . Después del recubrimiento de las células Me67.8, se añadieron 5 ng/ml de IFN γ recombinante humano a cultivos en presencia de diversas concentraciones de anticuerpos monoclonales anti- anti-IFN γ completamente humanos candidatos. Después de 48 h en cultivo, las células se tiñeron con anticuerpo anti-MHC de clase II humano (HLA-DR) marcado con fluorescencia y se analizaron usando un FACSCalibur®. Por lo tanto, la IC₅₀ (donde se inhibe 50% de la expresión de MHC de clase II inducida por IFN γ , es decir, la concentración inhibitoria 50%), se mide para cada anticuerpo candidato.

Los scFv completamente humanos purificados se produjeron como se describe anteriormente. El efecto de scFv seleccionado sobre la expresión de MHC de clase II inducida por IFN γ en células de melanoma se evaluó usando el ensayo basado en células de citometría de flujo descrito anteriormente. Estos scFv inhibieron la expresión de MHC de clase II inducida por IFN γ en las células de melanoma (Figura 22).

EJEMPLO 16: Cambio de formato de scFv en el formato IgG

Las secuencias de VH y VL de scFv seleccionado se amplificaron con los oligonucleótidos específicos que introducen una secuencia líder y un sitio de restricción Hindi en el extremo 5'. Se introdujo un sitio Apal en el extremo 3' de la cadena pesada mientras que se introdujeron un AvrII y un sitio BsiWI en el extremo 3' de las secuencias de la cadena liviana lambda o kappa, respectivamente. Las secuencias VH amplificadas se digirieron con HindIII/Apal y se clonaron en el vector de expresión pCon_gamma (Lonza, Basilea, Suiza). Las secuencias lambda de de VL amplificadas se digirieron con HindIII/AvrII y se clonaron en el vector de expresión pCon_lambda2 y las secuencias de VL kappa amplificadas se digirieron con HindIII/BsiWI y se clonaron en el vector de expresión pCon_kappa (Lonza, Basilea, Suiza). Las construcciones se verificaron mediante secuenciación antes de transfección en células de mamífero.

Las secuencias de VH y VL de ADNc en sus vectores de expresión apropiados se transfectaron en células de mamífero utilizando el reactivo de transfección Fugene 6 (Roche, Basilea, Suiza). Brevemente, las células de los picos se cultivaron en placas de 6 pocillos a una concentración de 6×10^5 células por pocillo en 2 ml de medios de cultivo que contienen suero bovino fetal. Los vectores de expresión, que codifican las secuencias candidata de VH y VL, se co-transfectaron en las células utilizando el reactivo de transfección Fugene 6 de acuerdo con las instrucciones del fabricante. Un día después de la transfección, el medio de cultivo se aspiró y se añadieron 3 ml de medio libre de suero fresco a las células y se cultivó durante tres días a 37 °C. Después de tres días del período del cultivo, el sobrenadante se recolectó para IgG purificada en columnas de flujo rápido Proteína G-Sepharose 4B (Sigma, St. Louis, MO) de acuerdo con las instrucciones del fabricante. Brevemente, los sobrenadantes de las células transfectadas se incubaron durante la noche a 4 °C con buffer de unión de IgG ImmunoPure (G) (Pierce, Rockford IL). Las muestras luego se hicieron pasar sobre columnas de flujo rápido proteína G-Sepharose 4B y por consiguiente la IgG se purificó usando buffer de elución. La fracción IgG eluida luego se dializó frente a PBS y el contenido de IgG se cuantificó por absorción a 280 nm. La pureza y la integridad IgG se verificaron por SDS-PAGE.

EJEMPLO 17: Inhibición de IgG de la actividad biológica del interferón gamma

Dos scFv, AE1-4-R3-P2E4 (2E4) y A2-AD1-R4P1A9 (1A9), cuya actividad inhibitoria se había confirmado en contra hIFN γ en ensayos funcionales cambiaron de formato en IgG como se describe en el Ejemplo 16 y analiza en el ensayo del gen indicador inducido por interferón gamma descrito en el Ejemplo 14. Los resultados mostrados en la Figura 23 indican que en un formato de IgG tanto 1A9 y 2E4 pueden neutralizar la actividad de hIFN γ con IC₅₀ de 42 nM y 10 nM, respectivamente, mientras que un control negativo IgG (NI-0701) no tuvo ningún efecto en este ensayo. Así, estos dos candidatos aislados de las bancos de diversidad sintéticos y naturales pueden cambiar de formato IgG completa y presentar actividad neutralizante contra el blanco seleccionado.

EJEMPLO 18: Desarrollo del ensayo farmacocinético para la detección de 5E3 en suero de rata.

Dos scFv AD15E3R3P1 A4 y AD25E3R3P1 GLL candidatos que se unen específicamente a un anticuerpo monoclonal de ratón 5E3 (Figura 19) cambiaron de formato en la IgG humana completa como se describe en el Ejemplo 16. La especificidad de la correspondiente IgG DA4 y G11 se confirmó en ELISA contra 5E3 de ratón y una versión quimérica de este anticuerpo monoclonal en el que las regiones variables de ratón se han fusionado a regiones constantes de IgG de rata. Los resultados mostrados en la Figura 24 demuestran que las IgG de DA4 y G11 son específicas para la región variable de 5E3 ya que se unen 5E3 de ratón y rata quimérico y no a los controles de

isotipo de ratón y de rata. Estos dos anticuerpos monoclonales se utilizaron para desarrollar un ensayo para la cuantificación de 5E3 en suero de ratón para los estudios farmacocinéticos. Varias diluciones de suero de ratón se enriquecieron con 5 µg /ml de anticuerpo 5E3 de ratón y se diluyeron de forma seriada de tal manera que la concentración de suero se mantuvo constante durante toda la serie de dilución. Las placas Maxisorb (Nunc, Dinamarca) se recubrieron durante la noche con 1 µg/ml de IgG o IgG DA4 G11. Después de bloquear con PBS; las diluciones seriadas de BSA 1% de las preparaciones de suero enriquecidas se añadieron a los pocillos. Después de la incubación y el lavado, la señal se reveló utilizando un anticuerpo monoclonal de cadena liviana kappa anti-ratón acoplado a peroxidasa de rábano picante (HRP) y un sustrato fluorescente (rojo Amplex; Invitrogen). Los resultados muestran que ambos anticuerpos se pueden utilizar para detectar específicamente el anticuerpo monoclonal de ratón 5E3 en suero de ratón (Figura 25). El límite de detección del 5E3 de ratón en el suero fue de aproximadamente 200 ng/ml y el ensayo no fue afectado significativamente por la concentración sérica, lo que indica que las IgG de DA4 e IgG G11 son altamente específicos para el 5E3 de ratón y no se unen a otra inmunoglobulina de ratón. Estos experimentos demuestran que los anticuerpos anti-idiotípicos altamente específicos se pueden aislar de las bancos AEI y AD1 naturales o sintéticas.

15 EJEMPLO 19: Selección de fago usando bancos que contienen diversidad de CDRH3 capturada de ratones sin exposición previa e inmunizado.

Los bancos MnA, MiB y MIC descritas en los Ejemplos 8 y 9 se usaron en paralelo para las selecciones de fagos contra h1FNy siguiendo el procedimiento descrito en el Ejemplo 11. Durante el proceso de selección se observó un enriquecimiento similar de fago (Figura 26).

20 *Expresión de scFv en formato de placa de microtitulación:* Los clones individuales se recogieron en una placa de microtitulación que contienen 150 µl del medio 2xTYAG (glucosa 2) por pocillo y se cultivaron a 37 °C (100-120 rpm) durante 5-6 h. Las placas se centrifugaron a 280 rpm, el medio se descartó y los sedimentos celulares se resuspendieron en 100 µl de medio 2xTYA que contiene 1 mM de IPTG. Las placas se incubaron durante la noche a 30 °C con agitación (100 rpm). Después del crecimiento, las placas se centrifugaron a 3200 rpm durante 10 min y el sobrenadante se transfirió cuidadosamente a una placa que contiene PBS concentrado 2x que contiene 5% de leche en polvo descremada para el.

scFv ELISA: Las placas ELISA (Maxisorb, NUNC) se recubrieron durante la noche con 2 µg/ ml h1FNy en PBS. Las placas de control se recubrieron con 2 µg/ml de BSA (Sigma). Después las placas se bloquearon con 3% leche descremada/PBS a temperatura ambiente durante 1 hora. Las placas se lavaron 3 veces con PBS 0,05% de Tween 20 antes de transferir los sobrenadantes de scFv pre-bloqueados y se incubaron durante una hora a temperatura ambiente. Las placas luego se lavaron 3 veces con PBS 0,05% Tween 20. Se añadieron 50 µl de leche descremada/PBS que contiene anticuerpo (HRP) anti-cMyc conjugado (HRP) (Amersham, diluido 1: 5.000) a cada pocillo. Después de la incubación a temperatura ambiente durante 1 h, las placas se lavaron 5 veces con PBS 0,05% Tween 20. El ELISA se reveló a continuación mediante la adición de 50 µl de Amplex Red (Invitrogen). La intensidad de absorción se leyó a 590 nm después de la excitación a 530 nm.

La frecuencia de aciertos que dan una señal de la mitad de la intensidad del clon A6 de control se evaluó después de cada ronda de selección para las tres bancos (Figura 27). La tasa de aciertos obtenida con la biblioteca MiB fue marcadamente mayor en comparación con los otros dos bancos y el nivel medio de la señal fue superior para los clones derivados de la biblioteca MiB, lo que indica que los scFv de mayor afinidad estaban enriquecidos (Figura 28). Con el fin de confirmar esta observación, se secuenciaron los clones positivos, se expresaron en mayor escala y se purificaron para analizar en experimentos de unión de dosis respuesta de acuerdo con el Ejemplo 13. Los scFv derivados de la biblioteca MiB presentaron una mayor afinidad aparente por h1FNy que los aislados de la biblioteca MnA naïve (Figura 29). Los resultados indican que el repertorio CDRH3 de ratones inmunizados con una proteína puede ser capturado en contexto de estructural del anticuerpo humano de una manera productiva para generar con mayor frecuencia fragmentos de anticuerpo humano de alta afinidad. Los bancos generados usando la presente invención, por consiguiente representan un poderoso medio de generación de anticuerpos con potencial terapéutico.

Otras formas de realización

Aunque la invención se ha descrito junto con la descripción detallada, la descripción anterior pretende ilustrar y no limitar el alcance de la invención, que se define por el alcance de las reivindicaciones adjuntas. Otros aspectos, ventajas y modificaciones están dentro del alcance de las siguientes reivindicaciones.

ES 2 545 792 T3

LISTADO DE SECUENCIAS

<110> NovImmune SA, et al.
 <120> BIBLIOTECAS DE POLIPÉPTIDOS SINTÉTICOS Y MÉTODOS PARA GENERAR
 VARIANTES DE POLIPÉPTIDOS DIVERSIFICADAS NATURALMENTE
 <130> 23135-418001WO
 <150> US 61/179,850
 <151> 2009-05-20
 <150> US 61/287,336
 <151> 2009-12-17
 <150> US 61/314,794
 <151> 2010-03-17
 <160> 289
 <170> PatentIn version 3.5
 <210> 1
 <211> 98
 <212> PRT
 <213> Homo sapiens
 <400> 1

Gln	Val	Gln	Leu	Val	Gln	Ser	Gly	Ala	Glu	Val	Lys	Lys	Pro	Gly	Ala
1			5						10					15	
Ser	Val	Lys	Val	Ser	Cys	Lys	Ala	Ser	Gly	Tyr	Thr	Phe	Thr	Gly	Tyr
			20					25						30	
Tyr	Met	His	Trp	Val	Arg	Gln	Ala	Pro	Gly	Gln	Gly	Leu	Glu	Trp	Met
		35					40					45			
Gly	Trp	Ile	Asn	Pro	Asn	Ser	Gly	Gly	Thr	Asn	Tyr	Ala	Gln	Lys	Phe
		50				55						60			
Gln	Gly	Arg	Val	Thr	Met	Thr	Arg	Asp	Thr	Ser	Ile	Ser	Thr	Ala	Tyr
					70					75					80
Met	Glu	Leu	Ser	Arg	Leu	Arg	Ser	Asp	Asp	Thr	Ala	Val	Tyr	Tyr	Cys
					85				90					95	
Ala	Arg														

<210> 2
 <211> 98

ES 2 545 792 T3

<212> PRT
 <213> Homo sapiens

<400> 2

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
 20 25 30

Gly Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
 35 40 45

Gly Trp Ile Ser Ala Tyr Asn Gly Asn Thr Asn Tyr Ala Gln Lys Leu
 50 55 60

Gln Gly Arg Val Thr Met Thr Thr Asp Thr Ser Thr Ser Thr Ala Tyr
 65 70 75 80

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

Ala Arg

<210> 3
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 3

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser
 1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Gly Thr Phe Ser Ser Tyr
 20 25 30

Ala Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
 35 40 45

Gly Gly Ile Ile Pro Ile Phe Gly Thr Ala Asn Tyr Ala Gln Lys Phe
 50 55 60

Gln Gly Arg Val Thr Ile Thr Ala Asp Glu Ser Thr Ser Thr Ala Tyr
 65 70 75 80

ES 2 545 792 T3

Ala Arg

<210> 4
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 4

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
 1 5 10 15
 Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
 20 25 30
 Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45
 Ala Val Ile Ser Tyr Asp Gly Ser Asn Lys Tyr Tyr Ala Asp Ser Val
 50 55 60
 Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
 65 70 75 80
 Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95

Ala Arg

<210> 5
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 5

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
 1 5 10 15
 Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
 20 25 30

ES 2 545 792 T3

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

Ser Tyr Ile Ser Ser Ser Ser Ser Thr Ile Tyr Tyr Ala Asp Ser Val
 50 55 60

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

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

Ala Arg

<210> 6
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 6

Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
 1 5 10 15

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

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

Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
 50 55 60

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

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

Ala Lys

<210> 7
 <211> 98
 <212> PRT
 <213> Homo sapiens

ES 2 545 792 T3

<400> 7

Glu Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Glu
1 5 10 15

Ser Leu Lys Ile Ser Cys Lys Gly Ser Gly Tyr Ser Phe Thr Ser Tyr
20 25 30

Trp Ile Gly Trp Val Arg Gln Met Pro Gly Lys Gly Leu Glu Trp Met
35 40 45

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

Gln Gly Gln Val Thr Ile Ser Ala Asp Lys Ser Ile Ser Thr Ala Tyr
65 70 75 80

Leu Gln Trp Ser Ser Leu Lys Ala Ser Asp Thr Ala Met Tyr Tyr Cys
85 90 95

Ala Arg

<210> 8

<211> 98

<212> PRT

<213> Homo sapiens

<400> 8

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Ser Asn Tyr
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45

Tyr Asp Ala Ser Asn Leu Glu Thr Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Tyr Asp Asn Leu Pro Pro

ES 2 545 792 T3

Thr Val

<210> 9
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 9

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr
 20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
 50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Ser Tyr Ser Thr Pro Pro
 85 90 95

Thr Val

<210> 10
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 10

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

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile
 35 40 45

ES 2 545 792 T3

Tyr Asp Ala Ser Asn Arg Ala Thr Gly Ile Pro Ala Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Glu Pro
65 70 75 80

Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Arg Ser Asn Trp Pro Pro
85 90 95

Thr Val

<210> 11
<211> 98
<212> PRT
<213> Homo sapiens

<400> 11

Glu Ile Val Met Thr Gln Ser Pro Ala Thr Leu Ser Val Ser Pro Gly
1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Asn
20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile
35 40 45

Tyr Gly Ala Ser Thr Arg Ala Thr Gly Ile Pro Ala Arg Phe Ser Gly
50 55 60

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

Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Asn Asn Trp Pro Pro
85 90 95

Thr Val

<210> 12
<211> 99
<212> PRT
<213> Homo sapiens

<400> 12

ES 2 545 792 T3

Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly
 1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Ser
 20 25 30

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

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

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu
 65 70 75 80

Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Ser Pro
 85 90 95

Pro Thr Val

<210> 13
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 13

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

Arg Val Thr Ile Ser Cys Ser Gly Ser Ser Ser Asn Ile Gly Ser Asn
 20 25 30

Thr Val Asn Trp Tyr Gln Gln Leu Pro Gly Thr Ala Pro Lys Leu Leu
 35 40 45

Ile Tyr Ser Asn Asn Gln Arg Pro Ser Gly Val Pro Asp Arg Phe Ser
 50 55 60

Gly Ser Lys Ser Gly Thr Ser Ala Ser Leu Ala Ile Ser Gly Leu Gln
 65 70 75 80

Ser Glu Asp Glu Ala Asp Tyr Tyr Cys Ala Ala Trp Asp Asp Ser Leu
 85 90 95

ES 2 545 792 T3

Thr Gly Asp Glu Ala Asp Tyr Tyr Cys Gly Thr Trp Asp Ser Ser Leu
85 90 95

Ser Ala

<210> 15
<211> 107
<212> ADN
<213> Homo sapiens

<400> 15
tattactgtg cgagatgaga cgaataacgg taaggcgggt taccaggttt aaacgcgtat 60
tggaaggcg cgtctcttgg ggccaggaa cctgggtcac agtctcg 107

<210> 16
<211> 36
<212> PRT
<213> Homo sapiens

<400> 16

Val Leu Leu Cys Glu Met Arg Arg Ile Thr Val Arg Arg Phe Thr Arg

5

10

ES 2 545 792 T3

```

<221> misc_feature
<222> (7)..(14)
<223> n é s a, t, c o g

<400> 19
cgtctcnnnn nnnn 14

<210> 20
<211> 14
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (7)..(14)
<223> n é s a, t, c o g

<400> 20
gcagagnnnn nnnn 14

<210> 21
<211> 386
<212> ADN
<213> Homo sapiens

<400> 21
cagggtgcagc tgggtgcagtc tggggctgag gtgaagaagc ctggggcctc agtgaaggtc 60
tcctgcaagg ctctcgata caccttcacc ggctactata tgcactgggt gcgacaggcc 120
cctggacaag ggcttgatg gatgggatgg atcaacccta acagtgggtg cacaaactat 180
gcacagaagt ttcagggcag ggtcaccatg accagggaca cgtccatcag cacagcctac 240
atggagctga gcaggctgag atctgacgac acggccgtgt attactgtgc gagatgagac 300
gaataacggt aaggcggttt accaggttta aacgcgtatt gggaaggcgc gtctcttggg 360
gccagggaac cctggtcaca gtctcg 386

<210> 22
<211> 125
<212> PRT
<213> Homo sapiens

<400> 22
Pro Gly Ala Ala Gly Ala Val Trp Gly Gly Glu Glu Ala Trp Gly Leu
1 5 10 15

```

ES 2 545 792 T3

20 25 30
 Tyr Ala Leu Gly Ala Thr Gly Pro Trp Thr Arg Ala Val Asp Gly Met
 35 40 45
 Asp Gln Pro Gln Trp Trp His Lys Leu Cys Thr Glu Val Ser Gly Gln
 50 55 60
 Gly His His Asp Gln Gly His Val His Gln His Ser Leu His Gly Ala
 65 70 75 80
 Glu Gln Ala Glu Ile Arg His Gly Arg Val Leu Leu Cys Glu Met Arg
 85 90 95
 Arg Ile Thr Val Arg Arg Phe Thr Arg Phe Lys Arg Val Leu Gly Arg
 100 105 110
 Arg Val Ser Trp Gly Gln Gly Thr Leu Val Thr Val Ser
 115 120 125
 <210> 23
 <211> 125
 <212> PRT
 <213> Homo sapiens
 <400> 23
 Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15
 Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Gly Tyr
 20 25 30
 Tyr Met His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
 35 40 45
 Gly Trp Ile Asn Pro Asn Ser Gly Gly Thr Asn Tyr Ala Gln Lys Phe
 50 55 60
 Gln Gly Arg Val Thr Met Thr Arg Asp Thr Ser Ile Ser Thr Ala Tyr
 65 70 75 80
 Met Glu Leu Ser Arg Leu Arg Ser Asp Asp Thr Ala Val Tyr Tyr Cys
 85 90 95
 Ala Arg Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu

ES 2 545 792 T3

```

                100                105                110

Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser
      115                120                125

<210> 24
<211> 123
<212> PRT
<213> Homo sapiens

<400> 24

Pro Arg Cys Ser Trp Cys Ser Leu Gly Leu Arg Arg Ser Leu Gly Pro
1      5      10      15

Gln Arg Ser Pro Ala Arg Leu Leu Asp Thr Pro Ser Pro Ala Thr Ile
      20      25      30

Cys Thr Gly Cys Asp Arg Pro Leu Asp Lys Gly Leu Ser Gly Trp Asp
      35      40      45

Gly Ser Thr Leu Thr Val Val Ala Gln Thr Met His Arg Ser Phe Arg
      50      55      60

Ala Gly Ser Pro Pro Gly Thr Arg Pro Ser Ala Gln Pro Thr Trp Ser
65      70      75      80

Ala Gly Asp Leu Thr Thr Arg Pro Cys Ile Thr Val Arg Asp Glu Thr
      85      90      95

Asn Asn Gly Lys Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg
      100     105     110

Leu Leu Gly Pro Gly Asn Pro Gly His Ser Leu
      115                120

<210> 25
<211> 386
<212> ADN
<213> Homo sapiens

<400> 25
caggtgcagc tgggtcagtc tggagctgag gtgaagaagc ctggggcctc agtgaaggtc      60
tcctgcaagg cttctggtta cacctttacc agctatggtt tcagctgggt gcgacaggcc      120
cctggacaag ggcttgagtg gatgggatgg atcagcgctt acaatggtaa cacaaactat      180
gcacagaagc tccagggcag agtcaccatg accacagaca catccacgag cacagcctac      240

```

ES 2 545 792 T3

atggagctga ggagcctgag atctgacgac acggccgtgt attactgtgc gagatgagac 300
 gaataacggt aaggcgggtt accaggttta aacgcgtatt gggaaggcgc gtctcttggg 360
 gccagggaac cctggtcaca gtctcg 386

<210> 26
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 26

Pro Gly Ala Ala Gly Ala Val Trp Ser Gly Glu Glu Ala Trp Gly Leu
 1 5 10 15

Ser Glu Gly Leu Leu Gln Gly Phe Trp Leu His Leu Tyr Gln Leu Trp
 20 25 30

Tyr Gln Leu Gly Ala Thr Gly Pro Trp Thr Arg Ala Val Asp Gly Met
 35 40 45

Asp Gln Arg Leu Gln Trp His Lys Leu Cys Thr Glu Ala Pro Gly Gln
 50 55 60

Ser His His Asp His Arg His Ile His Glu His Ser Leu His Gly Ala
 65 70 75 80

Glu Glu Pro Glu Ile Arg His Gly Arg Val Leu Leu Cys Glu Met Arg
 85 90 95

Arg Ile Thr Val Arg Arg Phe Thr Arg Phe Lys Val Arg Leu Gly Arg
 100 105 110

Arg Val Ser Trp Gly Gln Gly Thr Leu Val Thr Val Ser
 115 120 125

<210> 27
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 27

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr

ES 2 545 792 T3

```

                20             25             30
Gly Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
   35                40                45
Gly Trp Ile Ser Ala Tyr Asn Gly Asn Thr Asn Tyr Ala Gln Lys Leu
   50                55                60
Gln Gly Arg Val Thr Met Thr Thr Asp Thr Ser Thr Ser Thr Ala Tyr
   65                70                75                80
Met Glu Leu Arg Ser Leu Arg Ser Asp Asp Thr Ala Val Tyr Tyr Cys
   85                90                95
Ala Arg Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu
  100                105                110
Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser
  115                120                125
<210> 28
<211> 122
<212> PRT
<213> Homo sapiens
<400> 28
Pro Arg Cys Ser Trp Cys Ser Leu Glu Leu Arg Arg Ser Leu Gly Pro
  1          5          10          15
Gln Arg Ser Pro Ala Arg Leu Leu Val Thr Pro Leu Pro Ala Met Val
  20          25          30
Ser Ala Gly Cys Asp Arg Pro Leu Asp Lys Gly Leu Ser Gly Trp Asp
  35          40          45
Gly Ser Ala Leu Thr Met Val Thr Gln Thr Met His Arg Ser Ser Arg
  50          55          60
Ala Glu Ser Pro Pro Gln Thr His Pro Arg Ala Gln Pro Thr Trp Ser
  65          70          75          80
Gly Ala Asp Leu Thr Thr Arg Pro Cys Ile Thr Val Arg Asp Glu Thr
  85          90          95
Asn Asn Gly Lys Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg

```

ES 2 545 792 T3

100 105 110

Leu Leu Gly Pro Gly Asn Pro His Ser Leu
 115 120

<210> 29
 <211> 386
 <212> ADN
 <213> Homo sapiens

<400> 29
 caggtgcagc tgggtgcagc tggggctgag gtgaagaagc ctgggtcctc ggtgaaggtc 60
 tcttgcaagg cttctggagg caccttcagc agctatgcta tcagctgggt gcgacaggcc 120
 cctggacaag ggcttgagtg gatgggagg atcatccta tctttgttac agcaaactac 180
 gcacagaagt tccagggcag agtcacgatt accgcggacg aatccacgag cacagcctac 240
 atggagctga gcagcctgag atctgaggac acggccgtgt attactgtgc gagatgagac 300
 gaataacggt aaggcggttt accaggttta aacgcgtatt gggaaggcgc gtctcttggg 360
 gccaggaac cctggtcaca gtctcg 386

<210> 30
 <211> 126
 <212> PRT
 <213> Homo sapiens

<400> 30

Pro Gly Ala Ala Gly Ala Val Trp Gly Gly Glu Glu Ala Trp Val Leu
 1 5 10 15

Gly Glu Gly Leu Leu Gln Gly Phe Trp Arg His Leu Gln Gln Leu Cys
 20 25 30

Tyr Gln Leu Gly Ala Thr Gly Pro Trp Thr Arg Ala Val Asp Gly Arg
 35 40 45

Asp His Pro Tyr Leu Trp Tyr Ser Lys Leu Arg Thr Glu Val Pro Gly
 50 55 60

Gln Ser His Asp Tyr Arg Gly Arg Ile His Glu His Ser Leu His Gly
 65 70 75 80

Ala Glu Gln Pro Glu Ile Gly His Gly Arg Val Leu Leu Cys Glu Met
 85 90 95

ES 2 545 792 T3

Arg Arg Ile Thr Val Arg Arg Phe Thr Arg Phe Lys Arg Val Leu Gly
 100 105 110

Arg Arg Val Ser Trp Gly Gln Gly Thr Leu Val Thr Val Ser
 115 120 125

<210> 31
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 31

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser
 1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Gly Thr Phe Ser Ser Tyr
 20 25 30

Ala Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
 35 40 45

Gly Gly Ile Ile Pro Ile Phe Gly Thr Ala Asn Tyr Ala Gln Lys Phe
 50 55 60

Gln Gly Arg Val Thr Ile Thr Ala Asp Glu Ser Thr Ser Thr Ala Tyr
 65 70 75 80

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

Ala Arg Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu
 100 105 110

Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser
 115 120 125

<210> 32
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 32

Pro Arg Cys Ser Trp Cys Ser Leu Gly Leu Arg Arg Ser Leu Gly Pro
 1 5 10 15

Arg Arg Ser Pro Ala Arg Leu Leu Glu Ala Pro Ser Ala Ala Met Leu

ES 2 545 792 T3

20 25 30

Ser Ala Gly Cys Asp Arg Pro Leu Asp Lys Gly Leu Ser Gly Trp Glu
 35 40 45

Gly Ser Ser Leu Ser Val Gln Gln Thr Thr His Arg Ser Ser Arg Ala
 50 55 60

Glu Ser Arg Leu Pro Arg Thr Asn Pro Arg Ala Gln Pro Thr Trp Ser
 65 70 75 80

Ala Ala Asp Leu Arg Thr Arg Pro Cys Ile Thr Val Arg Asp Glu Thr
 85 90 95

Asn Asn Gly Lys Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg
 100 105 110

Leu Leu Gly Pro Gly Asn Pro Gly His Ser Leu
 115 120

<210> 33
 <211> 386
 <212> ADN
 <213> Homo sapiens

<400> 33
 gaggtgcagc tgttgagtc tgggggagc ttggtacagc ctggggggtc cctgagactc 60
 tcctgtgcag cctctggatt cacctttagc agctatgcca tgagctgggt ccgccaggct 120
 ccaggaagg ggctggagt ggtctcagct attagtggta gtgggtgtag cacatactac 180
 gcagactcgc tgaagggccg gttcaccatc tccagagaca attccaagaa cagctgtat 240
 ctgcaaatga acagcctgag agccgaggac acggccgtat attactgtgc gagatgagac 300
 gaataacggg aaggcggttt accaggttta aacgcgtatt ggaaggcgc gtctcttggg 360
 gccaggaac cctggtcaca gtctcg 386

<210> 34
 <211> 126
 <212> PRT
 <213> Homo sapiens

<400> 34
 Arg Gly Ala Ala Val Gly Val Trp Gly Arg Leu Gly Thr Ala Trp Gly
 1 5 10 15

ES 2 545 792 T3

Val Pro Glu Thr Leu Leu Cys Ser Leu Trp Ile His Leu Gln Leu Cys
20 25 30

His Glu Leu Gly Pro Pro Gly Ser Arg Glu Gly Ala Gly Val Gly Leu
35 40 45

Ser Tyr Trp Trp Trp His Ile Leu Arg Arg Leu Arg Glu Gly Pro Val
50 55 60

His His Leu Gln Arg Gln Phe Gln Glu His Ala Val Ser Ser Ala Asn
65 70 75 80

Glu Gln Pro Glu Ser Arg Gly His Gly Arg Ile Leu Leu Cys Glu Met
85 90 95

Arg Arg Ile Thr Val Arg Arg Phe Thr Arg Phe Lys Arg Val Leu Gly
100 105 110

Arg Arg Val Ser Trp Gly Gln Gly Thr Leu Val Thr Val Ser
115 120 125

<210> 35
<211> 125
<212> PRT
<213> Homo sapiens

<400> 35

Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

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

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

Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
50 55 60

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

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

ES 2 545 792 T3

Ala Arg Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu
 100 105 110

Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser
 115 120 125

<210> 36
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 36

Pro Arg Cys Ser Cys Trp Ser Leu Gly Glu Ala Trp Tyr Ser Leu Gly
 1 5 10 15

Gly Pro Asp Ser Pro Val Gln Pro Leu Asp Ser Pro Leu Ala Ala Met
 20 25 30

Pro Ala Gly Ser Ala Arg Leu Gln Gly Arg Gly Trp Ser Gly Ser Gln
 35 40 45

Leu Leu Val Val Val Val Val Ala His Thr Thr Gln Thr Pro Arg Ala
 50 55 60

Gly Ser Pro Ser Pro Glu Thr Ile Pro Arg Thr Arg Cys Ile Cys Lys
 65 70 75 80

Thr Ala Glu Pro Arg Thr Arg Pro Tyr Ile Thr Val Arg Asp Glu Thr
 85 90 95

Asn Asn Gly Lys Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg
 100 105 110

Leu Leu Gly Pro Gly Asn Pro Gly His Ser Leu
 115 120

<210> 37
 <211> 385
 <212> ADN
 <213> Homo sapiens

<400> 37
 gaggtgcagc tgttgagtc tgggggagc ttggtacagc ctggggggtc cctgagactc 60
 tcctgtgcag cctctggatt caccttagc agctatgcca tgagctgggt ccgcccagct 120
 ccaggaagg ggctggagtg ggtctcagct attagtggta gtggtgtag cacatactac 180

ES 2 545 792 T3

gcagactccg tgaagggccg gttcacatct ccagagacaa ttccaagaac acgctgtatc 240
 tgcaaatgaa cagcctgaga gcgaggaca cggccgtata ttactgtgcg aatgagacg 300
 aataacggta aggcgggtta ccaggtttaa acgcgtattg ggaaggcgcg tctcttgggg 360
 ccaggaacc ctggtcacag tctcg 385

<210> 38
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 38

Arg Gly Ala Ala Val Gly Val Trp Gly Arg Leu Gly Thr Ala Trp Gly
 1 5 10 15
 Val Pro Glu Thr Leu Leu Cys Ser Leu Trp Ile His Leu Gln Leu Cys
 20 25 30
 His Glu Leu Gly Pro Pro Gly Ser Arg Glu Gly Ala Gly Val Gly Leu
 35 40 45
 Ser Tyr Trp Trp Trp His Ile Leu Arg Arg Leu Arg Glu Gly Pro Val
 50 55 60
 His His Leu Gln Arg Gln Phe Gln Glu His Ala Val Ser Ala Asn Glu
 65 70 75 80
 Gln Pro Glu Ser Arg Gly His Gly Arg Ile Leu Leu Cys Glu Met Arg
 85 90 95
 Arg Ile Thr Val Arg Arg Phe Thr Arg Phe Lys Arg Val Leu Gly Arg
 100 105 110
 Arg Val Ser Trp Gly Gln Gly Thr Leu Val Thr Val Ser
 115 120 125

<210> 39
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 39

Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
 1 5 10 15

ES 2 545 792 T3

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

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

Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
 50 55 60

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

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

Ala Lys Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu
 100 105 110

Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser
 115 120 125

<210> 40
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 40

Pro Arg Cys Ser Cys Trp Ser Leu Gly Glu Ala Trp Tyr Ser Leu Gly
 1 5 10 15

Gly Pro Asp Ser Pro Val Gln Pro Leu Asp Ser Pro Leu Ala Ala Met
 20 25 30

Pro Ala Gly Ser Ala Arg Leu Gln Gly Arg Gly Trp Ser Gly Ser Gln
 35 40 45

Leu Leu Val Val Val Val Val Ala His Thr Thr Gln Thr Pro Arg Ala
 50 55 60

Gly Ser Pro Ser Pro Glu Thr Ile Pro Arg Thr Arg Cys Ile Cys Lys
 65 70 75 80

Thr Ala Glu Pro Arg Thr Arg Pro Tyr Ile Thr Val Arg Asn Glu Thr
 85 90 95

ES 2 545 792 T3

Asn Asn Gly Lys Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg
 100 105 110

Leu Leu Gly Pro Gly Asn Pro Gly His Ser Leu
 115 120

<210> 41
 <211> 386
 <212> ADN
 <213> Homo sapiens

<400> 41
 caggtgcagc tgggtggagtc tgggggaggc gtggtccagc ctgggaggtc cctgagactc 60
 tcctgtgcag cctctggatt caccttcagt agctatgcta tgcactgggt ccgccaggct 120
 ccaggcaagg ggtctggagt ggtggcagtt atatcatatg atggaagtaa taaatactac 180
 gcagactccg tgaagggccg attcaccatc tccagagaca attccaagaa cacgctgtat 240
 ctgcaaatga acagcctgag agctgaggac acggctgtgt attactgtgc gagatgagac 300
 gaataacggt aaggcggttt accaggttta aacgcgtatt gggaaggcgc gtctcttggg 360
 gccagggaac cctggtcaca gtctcg 386

<210> 42
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 42

Pro Gly Ala Ala Gly Gly Val Trp Gly Arg Arg Gly Pro Ala Trp Glu
 1 5 10 15

Val Pro Glu Thr Leu Leu Cys Ser Leu Trp Ile His Leu Gln Leu Cys
 20 25 30

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

Ser Tyr Ile Ile Trp Lys Ile Leu Arg Arg Leu Arg Glu Gly Pro Ile
 50 55 60

His His Leu Gln Arg Gln Phe Gln Glu His Ala Val Ser Ala Asn Glu
 65 70 75 80

Gln Pro Glu Ser Gly His Gly Cys Val Leu Leu Cys Glu Met Arg Arg
 85 90 95

ES 2 545 792 T3

Ile Thr Val Arg Arg Phe Thr Arg Phe Lys Arg Val Leu Gly Arg Arg
 100 105 110

Val Ser Trp Gly Gln Gly Thr Leu Val Thr Val Ser
 115 120

<210> 43
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 43

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
 1 5 10 15

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

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ala Val Ile Ser Tyr Asp Gly Ser Asn Lys Tyr Tyr Ala Asp Ser Val
 50 55 60

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

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

Ala Arg Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu
 100 105 110

Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser
 115 120 125

<210> 44
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 44

Pro Arg Cys Ser Trp Trp Ser Leu Gly Glu Ala Trp Ser Ser Leu Gly
 1 5 10 15

ES 2 545 792 T3

Gly Pro Asp Ser Pro Val Gln Pro Leu Asp Ser Pro Ser Val Ala Met
 20 25 30
 Leu Cys Thr Gly Ser Ala Arg Leu Gln Ala Arg Gly Trp Ser Gly Trp
 35 40 45
 Gln Leu Tyr His Met Met Glu Val Ile Asn Thr Thr Gln Thr Pro Arg
 50 55 60
 Ala Asp Ser Pro Ser Pro Glu Thr Ile Pro Arg Thr Arg Cys Ile Cys
 65 70 75 80
 Lys Thr Ala Glu Leu Arg Thr Arg Leu Cys Ile Thr Val Arg Asp Glu
 85 90 95
 Thr Asn Asn Gly Lys Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala
 100 105 110
 Arg Leu Leu Gly Pro Gly Asn Pro Gly His Ser Leu
 115 120

<210> 45
 <211> 386
 <212> ADN
 <213> Homo sapiens

<400> 45
 gaggtgcagc tgggtggagtc tgggggagc ttggtacagc ctggggggc cctgagactc 60
 tcctgtgcag cctctggatt caccttcagt agctatagca tgaactgggt ccgccaggct 120
 ccaggggaagg ggctggagtg ggtttcatac attagtagta gtagtagtac catatactac 180
 gcagactctg tgaagggccg attcaccatc tccagagaca atgccaagaa ctcactgtat 240
 ctgcaaatga acagcctgag agccgaggac acggctgtgt attactgtgc gagatgagac 300
 gaataacggt aaggcggttt accaggttta aacgcgtatt ggggaaggcgc gtctcttggg 360
 gccaggaac cctggtcaca gtctcg 386

<210> 46
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 46
 Arg Gly Ala Ala Gly Gly Val Trp Gly Arg Leu Gly Thr Ala Trp Gly
 1 5 10 15

ES 2 545 792 T3

Val Pro Glu Thr Leu Leu Cys Ser Leu Trp Ile His Leu Gln Leu His
20 25 30

Glu Leu Gly Pro Pro Gly Ser Arg Glu Gly Ala Gly Val Gly Phe Ile
35 40 45

His Tyr His Ile Leu Arg Arg Leu Cys Glu Gly Pro Ile His His Leu
50 55 60

Gln Arg Gln Cys Gln Glu Leu Thr Val Ser Ala Asn Glu Gln Pro Glu
65 70 75 80

Ser Arg Gly His Gly Cys Val Leu Leu Cys Glu Met Arg Arg Ile Thr
85 90 95

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

Trp Gly Gln Gly Thr Leu Val Thr Val Ser
115 120

<210> 47
<211> 125
<212> PRT
<213> Homo sapiens

<400> 47

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

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

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

Ser Tyr Ile Ser Ser Ser Ser Ser Thr Ile Tyr Tyr Ala Asp Ser Val
50 55 60

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

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

ES 2 545 792 T3

Ala Arg Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu
 100 105 110

Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser
 115 120 125

<210> 48
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 48

Pro Arg Cys Ser Trp Trp Ser Leu Gly Glu Ala Trp Tyr Ser Leu Gly
 1 5 10 15

Gly Pro Asp Ser Pro Val Gln Pro Leu Asp Ser Pro Ser Val Ala Ile
 20 25 30

Ala Thr Gly Ser Ala Arg Leu Gln Gly Arg Gly Trp Ser Gly Phe His
 35 40 45

Thr Leu Val Val Val Val Val Pro Tyr Thr Thr Gln Thr Leu Arg Ala
 50 55 60

Asp Ser Pro Ser Pro Glu Thr Met Pro Arg Thr His Cys Ile Cys Lys
 65 70 75 80

Thr Ala Glu Pro Arg Thr Arg Leu Cys Ile Thr Val Arg Asp Glu Thr
 85 90 95

Asn Asn Gly Lys Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg
 100 105 110

Leu Leu Gly Pro Gly Asn Pro Gly His Ser Leu
 115 120

<210> 49
 <211> 290
 <212> ADN
 <213> Homo sapiens

<400> 49

gagggtgcagc tgggtgcagtc tggagcagag gtgaaaaaac ccgggggagtc tctgaagatc 60
 tctgttaagg gttctggata cagctttacc agctactgga tcggctgggt gcgccagatg 120
 cccgggaaag gcctggagtg gatggggatc atctatcctg gtgactctga taccagatac 180

ES 2 545 792 T3

agcccgctct tccaaggcca ggtcaccatc tcagccgaca agtccatcag caccgcctac 240
ctgcagtgga gcagcctgaa ggccctcggac accgccatgt attactgtgc 290

<210> 50
<211> 126
<212> PRT
<213> Homo sapiens

<400> 50

Arg Gly Ala Ala Gly Ala Val Trp Ser Arg Gly Glu Lys Ala Arg Gly
1 5 10 15
Val Ser Glu Asp Leu Leu Gly Phe Trp Ile Gln Leu Tyr Gln Leu Leu
20 25 30
Asp Arg Leu Gly Ala Pro Asp Ala Arg Glu Arg Pro Gly Val Asp Gly
35 40 45
Asp His Leu Ser Trp Leu Tyr Gln Ile Gln Pro Val Leu Pro Arg Pro
50 55 60
Gly His His Leu Ser Arg Gln Val His Gln His Arg Leu Pro Ala Val
65 70 75 80
Glu Gln Pro Glu Gly Leu Gly His Arg His Val Leu Leu Cys Glu Met
85 90 95
Arg Arg Ile Thr Val Arg Arg Phe Thr Arg Phe Lys Arg Val Leu Gly
100 105 110
Arg Arg Val Ser Trp Gly Gln Gly Thr Leu Val Thr Val Ser
115 120 125

<210> 51
<211> 125
<212> PRT
<213> Homo sapiens

<400> 51

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

ES 2 545 792 T3

Trp Ile Gly Trp Val Arg Gln Met Pro Gly Lys Gly Leu Glu Trp Met
 35 40 45

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

Gln Gly Gln Val Thr Ile Ser Ala Asp Lys Ser Ile Ser Thr Ala Tyr
 65 70 75 80

Leu Gln Trp Ser Ser Leu Lys Ala Ser Asp Thr Ala Met Tyr Tyr Cys
 85 90 95

Ala Arg Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu
 100 105 110

Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser
 115 120 125

<210> 52
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 52

Pro Arg Cys Ser Trp Cys Ser Leu Glu Gln Arg Lys Ser Pro Gly Ser
 1 5 10 15

Leu Arg Ser Pro Val Arg Val Leu Asp Thr Ala Leu Pro Ala Thr Gly
 20 25 30

Ser Ala Gly Cys Ala Arg Cys Pro Gly Lys Ala Trp Ser Gly Trp Gly
 35 40 45

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

Ala Arg Ser Pro Ser Gln Pro Thr Ser Pro Ser Ala Pro Pro Thr Cys
 65 70 75 80

Ser Gly Ala Ala Arg Pro Arg Thr Pro Pro Cys Ile Thr Val Arg Asp
 85 90 95

Glu Thr Asn Asn Gly Lys Ala Val Tyr Gln Val Thr Arg Ile Gly Lys
 100 105 110

ES 2 545 792 T3

Ala Arg Leu Leu Gly Pro Gly Asn Pro Gly His Ser Leu
 115 120 125

<210> 53
 <211> 374
 <212> ADN
 <213> Homo sapiens

<400> 53
 gacatccaga tgaccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
 atcacttgcc aggcgagtc ggacattagc aactatntaa attggtatca gcagaaacca 120
 gggaaagccc ctaagctcct gatctacgat gcatccaatt tggaacacagg ggtcccatca 180
 aggttcagtg gaagtggatc tgggacagat tttactttca ccatcagcag cctgcagcct 240
 gaagatattg caacatatta ctgtcagcag cgagacgaat aacggtaagg cggtttacca 300
 ggtttaaacg cgtattggga aggcgcgtct caticggcca agggaccaag gtggaaatca 360
 aaggggaggg cgca 374

<210> 54
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 54

Gly His Pro Asp Asp Pro Val Ser Ile Leu Pro Val Cys Ile Cys Arg
 1 5 10 15
 Arg Gln Ser His His His Leu Pro Gly Glu Ser Gly His Gln Leu Phe
 20 25 30
 Lys Leu Val Ser Ala Glu Thr Arg Glu Ser Pro Ala Pro Asp Leu Arg
 35 40 45
 Cys Ile Gln Phe Gly Asn Arg Gly Pro Ile Lys Val Gln Trp Lys Trp
 50 55 60
 Ile Trp Asp Arg Phe Tyr Phe His His Gln Gln Pro Ala Ala Arg Tyr
 65 70 75 80
 Cys Asn Ile Leu Leu Ser Ala Ala Arg Arg Ile Thr Val Arg Arg Phe
 85 90 95
 Thr Arg Phe Lys Arg Val Leu Gly Arg Arg Val Ser Phe Gly Gln Gly
 100 105 110

ES 2 545 792 T3

Thr Lys Val Glu Ile Lys Gly Ala Ala Ala
 115 120

<210> 55
 <211> 122
 <212> PRT
 <213> Homo sapiens
 <400> 55

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Ser Asn Tyr
 20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

Tyr Asp Ala Ser Asn Leu Glu Thr Gly Val Pro Ser Arg Phe Ser Gly
 50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Arg Asp Glu Arg Gly Gly
 85 90 95

Leu Pro Gly Leu Asn Ala Tyr Trp Glu Gly Ala Ser His Ser Ala Lys
 100 105 110

Gly Pro Arg Trp Lys Ser Lys Gly Arg Pro
 115 120

<210> 56
 <211> 120
 <212> PRT
 <213> Homo sapiens
 <400> 56

Arg Thr Ser Arg Pro Ser Leu His Pro Pro Cys Leu His Leu Glu Thr
 1 5 10 15

Glu Ser Pro Ser Leu Ala Arg Arg Val Arg Thr Leu Ala Thr Ile Ile
 20 25 30

ES 2 545 792 T3

Gly Ile Ser Arg Asn Gln Gly Lys Pro Leu Ser Ser Ser Thr Met His
 35 40 45

Pro Ile Trp Lys Gln Gly Ser His Gln Gly Ser Val Glu Val Asp Leu
 50 55 60

Gly Gln Ile Leu Leu Ser Pro Ser Ala Ala Cys Ser Leu Lys Ile Leu
 65 70 75 80

Gln His Ile Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val Tyr
 85 90 95

Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gln
 100 105 110

Gly Gly Asn Gln Arg Gly Gly Arg
 115 120

<210> 57
 <211> 374
 <212> ADN
 <213> Homo sapiens

<400> 57
 gacatccaga tgaccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
 atcacttgcc aggcgagtc ggacattagc aactatttaa attggtatca gcagaaacca 120
 gggaaagccc ctaagctcct gatctacgat gcatccaatt tggaaacagg ggtcccatca 180
 aggttcagtg gaagtggatc tgggacagat tttactttca ccatcagcag cctgcagcct 240
 gaagatattg caacatatta ctgtcagcag cgagacgaat aacggtaagg cggtttacca 300
 ggtttaaagc cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca 360
 aacgggcggc cgca 374

<210> 58
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 58
 Gly His Pro Asp Asp Pro Val Ser Ile Leu Pro Val Cys Ile Cys Arg
 1 5 10 15

Arg Gln Ser His His His Leu Pro Gly Glu Ser Gly His Gln Leu Phe
 20 25 30

ES 2 545 792 T3

Lys Leu Val Ser Ala Glu Thr Arg Glu Ser Pro Ala Pro Asp Leu Arg
 35 40 45

Cys Ile Gln Phe Gly Asn Arg Gly Pro Ile Lys Val Gln Trp Lys Trp
 50 55 60

Ile Trp Asp Arg Phe Tyr Phe His His Gln Gln Pro Ala Ala Arg Tyr
 65 70 75 80

Cys Asn Ile Leu Leu Ser Ala Ala Arg Arg Ile Thr Val Arg Arg Phe
 85 90 95

Thr Arg Phe Lys Arg Val Leu Gly Arg Arg Val Ser Phe Gly Gln Gly
 100 105 110

Thr Lys Val Glu Ile Lys Arg Ala Ala Ala
 115 120

<210> 59
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 59

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Ser Asn Tyr
 20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

Tyr Asp Ala Ser Asn Leu Glu Thr Gly Val Pro Ser Arg Phe Ser Gly
 50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Arg Asp Glu Arg Gly Gly
 85 90 95

Leu Pro Gly Leu Asn Ala Tyr Trp Glu Gly Ala Ser His Ser Ala Lys
 100 105 110

ES 2 545 792 T3

Gly Pro Arg Trp Lys Ser Asn Gly Arg Pro
 115 120

<210> 60
 <211> 120
 <212> PRT
 <213> Homo sapiens

<400> 60

Arg Thr Ser Arg Pro Ser Leu His Pro Pro Cys Leu His Leu Glu Thr
 1 5 10 15

Glu Ser Pro Ser Leu Ala Arg Arg Val Arg Thr Leu Ala Thr Ile Ile
 20 25 30

Gly Ile Ser Arg Asn Gln Gly Lys Pro Leu Ser Ser Ser Thr Met His
 35 40 45

Pro Ile Trp Lys Gln Gly Ser His Gln Gly Ser Val Glu Val Asp Leu
 50 55 60

Gly Gln Ile Leu Leu Ser Pro Ser Ala Ala Cys Ser Leu Lys Ile Leu
 65 70 75 80

Gln His Ile Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val Tyr
 85 90 95

Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gln
 100 105 110

Gly Gly Asn Gln Thr Gly Gly Arg
 115 120

<210> 61
 <211> 374
 <212> ADN
 <213> Homo sapiens

<400> 61
 gacatccaga tgaccagtc tccatctcc ctgtctgcat ctgtaggaga cacagtcacc 60
 atcacttgcc gggcaagtca gagcattagc agctatttaa attggtatca gcagaaacca 120
 gggaaagccc ctaagctcct gatctatgct gcacccagtt tgcaaagtgg ggtcccatca 180
 aggttcagtg gcagtggatc tgggacagat ttcactctca ccatcagcag tctgcaacct 240

5

10

ES 2 545 792 T3

gaagattttg caacttacta ctgtcagcag cgagacgaat aacggtaagg cggtttacca 300
 ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca 360
 aaggggcggc cgca 374

<210> 62
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 62

Gly His Pro Asp Asp Pro Val Ser Ile Leu Pro Val Cys Ile Cys Arg
 1 5 10 15
 Arg Gln Ser His His His Leu Pro Gly Lys Ser Glu His Gln Leu Phe
 20 25 30
 Lys Leu Val Ser Ala Glu Thr Arg Glu Ser Pro Ala Pro Asp Leu Cys
 35 40 45
 Cys Ile Gln Phe Ala Lys Trp Gly Pro Ile Lys Val Gln Trp Gln Trp
 50 55 60
 Ile Trp Asp Arg Phe His Ser His His Gln Gln Ser Ala Thr Arg Phe
 65 70 75 80
 Cys Asn Leu Leu Leu Ser Ala Ala Arg Arg Ile Thr Val Arg Arg Phe
 85 90 95
 Thr Arg Phe Lys Arg Val Leu Gly Arg Arg Val Ser Phe Gly Gln Gly
 100 105 110
 Thr Lys Val Glu Ile Lys Gly Ala Ala Ala
 115 120

<210> 63
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 63

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15
 Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Ser
 20 25 30

ES 2 545 792 T3

Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu
 35 40 45

Ile Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser
 50 55 60

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln
 65 70 75 80

Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Arg Asp Glu Arg Gly
 85 90 95

Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu Gly Ala Ser His Ser Ala
 100 105 110

Lys Gly Pro Arg Trp Lys Ser Lys Gly Arg Pro
 115 120

<210> 64
 <211> 120
 <212> PRT
 <213> Homo sapiens

<400> 64

Arg Thr Ser Arg Pro Ser Leu His Pro Pro Cys Leu His Leu Glu Thr
 1 5 10 15

Glu Ser Pro Ser Leu Ala Gly Gln Val Arg Ala Leu Ala Ala Ile Ile
 20 25 30

Gly Ile Ser Arg Asn Gln Gly Lys Pro Leu Ser Ser Ser Met Leu His
 35 40 45

Pro Val Cys Lys Val Gly Ser His Gln Gly Ser Val Ala Val Asp Leu
 50 55 60

Gly Gln Ile Ser Leu Ser Pro Ser Ala Val Cys Asn Leu Lys Ile Leu
 65 70 75 80

Gln Leu Thr Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val Tyr
 85 90 95

Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gln
 100 105 110

ES 2 545 792 T3

Gly Gly Asn Gln Arg Gly Gly Arg
 115 120

<210> 65
 <211> 374
 <212> ADN
 <213> Homo sapiens

<400> 65
 gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
 atcacttgcc gggcaagtca gagcattagc agctatttaa attggtatca gcagaaacca 120
 gggaaagccc ctaagctcct gatctatgct gcatccagtt tgcaaagtgg ggtcccatca 180
 aggttcagtg gcagtggatc tgggacagat ttcactetca ccatcagcag tctgcaacct 240
 gaagattttg caacttacta ctgtcagcag cgagacgaat aacggtaagg cggtttacca 300
 ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca 360
 aacgggcggc cgca 374

<210> 66
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 66

Gly His Pro Asp Asp Pro Val Ser Ile Leu Pro Val Cys Ile Cys Arg
 1 5 10 15

Arg Gln Ser His His His Leu Pro Gly Lys Ser Glu His Gln Leu Phe
 20 25 30

Lys Leu Val Ser Ala Glu Thr Arg Glu Ser Pro Ala Pro Asp Leu Cys
 35 40 45

Cys Ile Gln Phe Ala Lys Trp Gly Pro Ile Lys Val Gln Trp Gln Trp
 50 55 60

Ile Trp Asp Arg Phe His Ser His His Gln Gln Ser Ala Thr Arg Phe
 65 70 75 80

Cys Asn Leu Leu Leu Ser Ala Ala Arg Arg Ile Thr Val Arg Arg Phe
 85 90 95

Thr Arg Phe Lys Arg Val Leu Gly Arg Arg Val Ser Phe Gly Gln Gly

ES 2 545 792 T3

100 105 110

Thr Lys Val Glu Ile Lys Arg Ala Ala Ala
115 120

<210> 67
<211> 122
<212> PRT
<213> Homo sapiens

<400> 67

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Arg Asp Glu Arg Gly Gly
85 90 95

Leu Pro Gly Leu Asn Ala Tyr Trp Glu Gly Ala Ser His Ser Ala Lys
100 105 110

Gly Pro Arg Trp Lys Ser Asn Gly Arg Pro
115 120

<210> 68
<211> 120
<212> PRT
<213> Homo sapiens

<400> 68

Arg Thr Ser Arg Pro Ser Leu His Pro Pro Cys Leu His Leu Glu Thr
1 5 10 15

Glu Ser Pro Ser Leu Ala Gly Gln Val Arg Ala Leu Ala Ala Ile Ile
20 25 30

ES 2 545 792 T3

Gly Ile Ser Arg Asn Gln Gly Lys Pro Leu Ser Ser Ser Met Leu His
 35 40 45

Pro Val Cys Lys Val Gly Ser His Gln Gly Ser Val Ala Val Asp Leu
 50 55 60

Gly Gln Ile Ser Leu Ser Pro Ser Ala Val Cys Asn Leu Lys Ile Leu
 65 70 75 80

Gln Leu Thr Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val Tyr
 85 90 95

Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gln
 100 105 110

Gly Gly Asn Gln Thr Gly Gly Arg
 115 120

<210> 69
 <211> 374
 <212> ADN
 <213> Homo sapiens

<400> 69
 gaaattgtgt tgacacagtc tccagccacc ctgtctttgt ctccagggga aagagccacc 60
 ctctcctgca gggccagtc gagtggttagc agctacttag cctggtacca acagaaacct 120
 ggccaggctc ccaggctcct catctatgat gcatccaaca gggccactgg catcccagcc 180
 aggttcagtg gcagtggttc tgggacagac ttcactctca ccatcagcag cctagagcct 240
 gaagattttg cagtttatta ctgtcagcag cgagacgaat aacggtaagg cggtttacca 300
 ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca 360
 aagggcgggc cgca 374

<210> 70
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 70
 Gly Asn Cys Val Asp Thr Val Ser Ser His Pro Val Phe Val Ser Arg
 1 5 10 15

Gly Lys Ser His Pro Leu Leu Gln Gly Gln Ser Glu Cys Gln Leu Leu

ES 2 545 792 T3

20 25 30
 Ser Leu Val Pro Thr Glu Thr Trp Pro Gly Ser Gln Ala Pro His Leu
 35 40 45
 Cys Ile Gln Gln Gly His Trp His Pro Ser Gln Val Gln Trp Gln Trp
 50 55 60
 Val Trp Asp Arg Leu His Ser His His Gln Gln Pro Arg Ala Arg Phe
 65 70 75 80
 Cys Ser Leu Leu Leu Ser Ala Ala Arg Arg Ile Thr Val Arg Arg Phe
 85 90 95
 Thr Arg Phe Lys Arg Val Leu Gly Arg Arg Val Ser Phe Gly Gln Gly
 100 105 110
 Thr Lys Val Glu Ile Lys Gly Ala Ala Ala
 115 120
 <210> 71
 <211> 122
 <212> PRT
 <213> Homo sapiens
 <400> 71
 Glu Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu Ser Pro Gly
 1 5 10 15
 Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Tyr
 20 25 30
 Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile
 35 40 45
 Tyr Asp Ala Ser Asn Arg Ala Thr Gly Ile Pro Ala Arg Phe Ser Gly
 50 55 60
 Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Glu Pro
 65 70 75 80
 Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Arg Asp Glu Arg Gly Gly
 85 90 95
 Leu Pro Gly Leu Asn Ala Tyr Trp Glu Gly Ala Ser His Ser Ala Lys

ES 2 545 792 T3

```

                100                105                110

Gly Pro Arg Trp Lys Ser Lys Gly Arg Pro
      115                120

<210> 72
<211> 121
<212> PRT
<213> Homo sapiens

<400> 72

Arg Lys Leu Cys His Ser Leu Gln Pro Pro Cys Leu Cys Leu Gln Gly
 1                5                10                15

Lys Glu Pro Pro Ser Pro Ala Gly Pro Val Arg Val Leu Ala Ala Thr
 20                25                30

Pro Gly Thr Asn Arg Asn Leu Ala Arg Leu Pro Gly Ser Ser Ser Met
 35                40                45

Met His Pro Thr Gly Pro Leu Ala Ser Gln Pro Gly Ser Val Ala Val
 50                55                60

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

Leu Gln Phe Ile Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val
 85                90                95

Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp
 100                105                110

Gln Gly Gly Asn Gln Arg Gly Gly Arg
 115                120

<210> 73
<211> 374
<212> ADN
<213> Homo sapiens

<400> 73
gaaattgtgt tgacacagtc tccagccacc ctgtctttgt ctccagggga aagagccacc      60
ctctcctgca gggccagtc gagtgtagc agctacttag cctggtacca acagaaacct      120
ggccaggctc ccaggctcct catctatgat gcatccaaca gggccactgg catcccagcc      180
aggttcagtg gcagtgggtc tgggacagac ttcactctca ccatcagcag cctagagcct      240

```

ES 2 545 792 T3

gaagattttg cagtttatta ctgtcagcag cgagacgaat aacggtaagg cggtttacca 300
 ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaatca 360
 aacgggcggc cgca 374

<210> 74
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 74

Gly Asn Cys Val Asp Thr Val Ser Ser His Pro Val Phe Val Ser Arg
 1 5 10 15

Gly Lys Ser His Pro Leu Leu Gln Gly Gln Ser Glu Cys Gln Leu Leu
 20 25 30

Ser Leu Val Pro Thr Glu Thr Trp Pro Gly Ser Gln Ala Pro His Leu
 35 40 45

Cys Ile Gln Gln Gly His Trp His Pro Ser Gln Val Gln Trp Gln Trp
 50 55 60

Val Trp Asp Arg Leu His Ser His His Gln Gln Pro Arg Ala Arg Phe
 65 70 75 80

Cys Ser Leu Leu Leu Ser Ala Ala Arg Arg Ile Thr Val Arg Arg Phe
 85 90 95

Thr Arg Phe Lys Arg Val Leu Gly Arg Arg Val Ser Phe Gly Gln Gly
 100 105 110

Thr Lys Val Glu Ile Lys Arg Ala Ala Ala
 115 120

<210> 75
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 75

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

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Tyr

ES 2 545 792 T3

	100		105		110	
Gln Gly Gly Asn Gln Thr Gly Gly Arg						
	115		120			
<210> 77						
<211> 374						
<212> ADN						
<213> Homo sapiens						
<400> 77						
gaaatagtga tgacgcagtc tccagccacc ctgtctgtgt ctccagggga aagagccacc						60
ctctcctgca gggccagtca gagggttagc agcaacttag cctggtacca gcagaaacct						120
ggccaggctc ccaggctcct catctatggt gcatccacca gggccactgg tatcccagcc						180
aggttcagtg gcagtggttc tgggacagag ttcactctca ccatcagcag cctgcagtct						240
gaagattttg cagtttatta ctgtcagcag cgagacgaat aacggtaagg cggtttacca						300
ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca						360
aagggcgggc cgca						374
<210> 78						
<211> 123						
<212> PRT						
<213> Homo sapiens						
<400> 78						
Gly Asn Ser Asp Asp Ala Val Ser Ser His Pro Val Cys Val Ser Arg						
1 5 10 15						
Gly Lys Ser His Pro Leu Leu Gln Gly Gln Ser Glu Cys Gln Gln Leu						
20 25 30						
Ser Leu Val Pro Ala Glu Thr Trp Pro Gly Ser Gln Ala Pro His Leu						
35 40 45						
Trp Cys Ile His Gln Gly His Trp Tyr Pro Ser Gln Val Gln Trp Gln						
50 55 60						
Trp Val Trp Asp Arg Val His Ser His His Gln Gln Pro Ala Val Arg						
65 70 75 80						
Phe Cys Ser Leu Leu Leu Ser Ala Ala Arg Arg Ile Thr Val Arg Arg						
85 90 95						

ES 2 545 792 T3

Phe Thr Arg Phe Lys Arg Val Leu Gly Arg Arg Val Ser Phe Gly Gln
 100 105 110

Gly Thr Lys Val Glu Ile Lys Gly Ala Ala Ala
 115 120

<210> 79
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 79

Glu Ile Val Met Thr Gln Ser Pro Ala Thr Leu Ser Val Ser Pro Gly
 1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Asn
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile
 35 40 45

Tyr Gly Ala Ser Thr Arg Ala Thr Gly Ile Pro Ala Arg Phe Ser Gly
 50 55 60

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

Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Arg Asp Glu Arg Gly Gly
 85 90 95

Leu Pro Gly Leu Asn Ala Tyr Trp Glu Gly Ala Ser His Ser Ala Lys
 100 105 110

Gly Pro Arg Trp Lys Ser Lys Gly Arg Pro
 115 120

<210> 80
 <211> 120
 <212> PRT
 <213> Homo sapiens

<400> 80

Arg Lys Arg Ser Leu Gln Pro Pro Cys Leu Cys Leu Gln Gly Lys Glu
 1 5 10 15

Pro Pro Ser Pro Ala Gly Pro Val Arg Val Leu Ala Ala Thr Pro Gly

ES 2 545 792 T3

20 25 30

Thr Ser Arg Asn Leu Ala Arg Leu Pro Gly Ser Ser Ser Met Val His
35 40 45

Pro Pro Gly Pro Leu Val Ser Gln Pro Gly Ser Val Ala Val Gly Leu
50 55 60

Gly Gln Ser Ser Leu Ser Pro Ser Ala Ala Cys Ser Leu Lys Ile Leu
65 70 75 80

Gln Phe Ile Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val Tyr
85 90 95

Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gln
100 105 110

Gly Gly Asn Gln Arg Gly Gly Arg
115 120

<210> 81
<211> 374
<212> ADN
<213> Homo sapiens

<400> 81
gaaatagtga tgacgcagtc tccagccacc ctgtctgtgt ctccagggga aagagccacc 60
ctctcctgca gggccagtca gagggttagc agcaacttag cctggtacca gcagaaacct 120
ggccaggctc ccaggctcct catctatggt gcatccacca gggccactgg tatcccagcc 180
aggttcagtg gcagtgggtc tgggacagag ttcactctca ccatcagcag cctgcagtct 240
gaagattttg cagtttatta ctgtcagcag cgagacgaat aacggtaagg cggtttacca 300
ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca 360
aacgggcggc cgca 374

<210> 82
<211> 123
<212> PRT
<213> Homo sapiens

<400> 82
Gly Asn Ser Asp Asp Ala Val Ser Ser His Pro Val Cys Val Ser Arg
1 5 10 15

ES 2 545 792 T3

Gly Lys Ser His Pro Leu Leu Gln Gly Gln Ser Glu Cys Gln Gln Leu
 20 25 30

Ser Leu Val Pro Ala Glu Thr Trp Pro Gly Ser Gln Ala Pro His Leu
 35 40 45

Trp Cys Ile His Gln Gly His Trp Tyr Pro Ser Gln Val Gln Trp Gln
 50 55 60

Trp Val Trp Asp Arg Val His Ser His His Gln Gln Pro Ala Val Arg
 65 70 75 80

Phe Cys Ser Leu Leu Leu Ser Ala Ala Arg Arg Ile Thr Val Arg Arg
 85 90 95

Phe Thr Arg Phe Lys Arg Val Leu Gly Arg Arg Val Ser Phe Gly Gln
 100 105 110

Gly Thr Lys Val Glu Ile Lys Arg Ala Ala Ala
 115 120

<210> 83
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 83

Glu Ile Val Met Thr Gln Ser Pro Ala Thr Leu Ser Val Ser Pro Gly
 1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Asn
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile
 35 40 45

Tyr Gly Ala Ser Thr Arg Ala Thr Gly Ile Pro Ala Arg Phe Ser Gly
 50 55 60

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

Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Arg Asp Glu Arg Gly Gly
 85 90 95

ES 2 545 792 T3

Leu Pro Gly Leu Asn Ala Tyr Trp Glu Gly Ala Ser His Ser Ala Lys
 100 105 110

Gly Pro Arg Trp Lys Ser Asn Gly Arg Pro
 115 120

<210> 84
 <211> 119
 <212> PRT
 <213> Homo sapiens

<400> 84

Arg Lys Arg Ser Leu Gln Pro Pro Cys Leu Cys Leu Gln Gly Lys Glu
 1 5 10 15

Pro Pro Ser Pro Ala Gly Pro Val Arg Val Leu Ala Ala Thr Pro Gly
 20 25 30

Thr Ser Arg Asn Leu Ala Arg Leu Pro Gly Ser Ser Ser Met Val His
 35 40 45

Pro Pro Gly Pro Leu Val Ser Gln Pro Gly Ser Val Ala Val Gly Leu
 50 55 60

Gly Gln Ser Ser Leu Ser Pro Ser Ala Ala Cys Ser Leu Lys Ile Leu
 65 70 75 80

Gln Phe Ile Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val Tyr
 85 90 95

Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Asp Gln Gly
 100 105 110

Gly Asn Gln Thr Gly Gly Arg
 115

<210> 85
 <211> 377
 <212> ADN
 <213> Homo sapiens

<400> 85
 gaaattgtgt tgacgcagtc tccaggcacc ctgtctttgt ctccagggga aagagccacc 60
 ctctcctgca gggccagtc gagtgtagc agcagctact tagcctggta ccagcagaaa 120
 cctggccagg ctcccaggct cctcatctat ggtgcatcca gcagggccac tggcatccca 180

ES 2 545 792 T3

gacaggttca gtggcagtggt gctctgggaca gacttcactc tcaccatcag cagactggag 240
 cctgaagatt ttgcagttta ttactgtcag cagcgagacg aataacggta aggcgggtta 300
 ccaggtttaa acgcgtattg ggaaggcgcg tctcattcgg ccaagggacc aaggtggaaa 360
 tcaaaggggc ggccgca 377

<210> 86
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 86

Gly Asn Cys Val Asp Ala Val Ser Arg His Pro Val Phe Val Ser Arg
 1 5 10 15

Gly Lys Ser His Pro Leu Ile Gln Gly Gln Ser Glu Cys Gln Gln Leu
 20 25 30

Leu Ser Leu Val Pro Ala Glu Thr Trp Pro Gly Ser Gln Ala Pro His
 35 40 45

Leu Trp Cys Ile Gln Gln Gly His Trp His Pro Arg Gln Val Gln Trp
 50 55 60

Gln Trp Val Trp Asp Arg Leu His Ser His His Gln Gln Thr Gly Ala
 65 70 75 80

Arg Phe Cys Ser Leu Leu Leu Ser Ala Ala Arg Arg Ile Thr Val Arg
 85 90 95

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

Gln Gly Thr Lys Val Glu Ile Lys Gly Ala Ala Ala
 115 120

<210> 87
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 87

Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly
 1 5 10 15

ES 2 545 792 T3

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Ser
 20 25 30

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

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

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu
 65 70 75 80

Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Arg Asp Glu Arg Gly
 85 90 95

Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu Gly Ala Ser His Ser Ala
 100 105 110

Lys Gly Pro Arg Trp Lys Ser Lys Gly Arg Pro
 115 120

<210> 88
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 88

Arg Lys Leu Cys Arg Ser Leu Gln Ala Pro Cys Leu Cys Leu Gln Gly
 1 5 10 15

Lys Glu Pro Pro Ser Pro Ala Gly Pro Val Arg Val Leu Ala Ala Ala
 20 25 30

Thr Pro Gly Thr Ser Arg Asn Leu Ala Arg Leu Pro Gly Ser Ser Ser
 35 40 45

Met Val His Pro Ala Gly Pro Leu Ala Ser Gln Thr Gly Ser Val Ala
 50 55 60

Val Gly Leu Gly Gln Thr Ser Leu Ser Pro Ser Ala Asp Trp Ser Leu
 65 70 75 80

Lys Ile Leu Gln Phe Ile Thr Val Ser Ser Glu Thr Asn Asn Gly Lys
 85 90 95

ES 2 545 792 T3

Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro
 100 105 110

Arg Asp Gln Gly Gly Asn Gln Arg Gly Gly Arg
 115 120

<210> 89
 <211> 377
 <212> ADN
 <213> Homo sapiens

<400> 89
 gaaattgtgt tgacgcagtc tccaggcacc ctgtctttgt ctccagggga aagagccacc 60
 ctctcctgca gggccagtca gagtgttagc agcagctact tagcctggta ccagcagaaa 120
 cctggccagg ctcccaggct cctcatctat ggtgcatcca gcagggccac tggcatccca 180
 gacaggttca gtggcagtgg gtctgggaca gacttcactc tcaccatcag cagactggag 240
 cctgaagatt ttgcagttta ttactgtcag cagcgagacg aataacggta aggcggttta 300
 ccaggtttaa acgcgtattg ggaaggcgcg tctcattcgg ccaagggacc aaggtggaaa 360
 tcaaacgggc ggccgca 377

<210> 90
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 90

Gly Asn Cys Val Asp Ala Val Ser Arg His Pro Val Phe Val Ser Arg
 1 5 10 15

Gly Lys Ser His Pro Leu Leu Gln Gly Gln Ser Glu Cys Gln Gln Leu
 20 25 30

Leu Ser Leu Val Pro Ala Glu Thr Trp Pro Gly Ser Gln Ala Pro His
 35 40 45

Leu Trp Cys Ile Gln Gln Gly His Trp His Pro Arg Gln Val Gln Trp
 50 55 60

Gln Trp Val Trp Asp Arg Leu His Ser His His Gln Gln Thr Gly Ala
 65 70 75 80

Arg Phe Cys Ser Leu Leu Leu Ser Ala Ala Arg Arg Ile Thr Val Arg
 85 90 95

ES 2 545 792 T3

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

Gln Gly Thr Lys Val Glu Ile Lys Arg Ala Ala Ala
 115 120

<210> 91
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 91

Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly
 1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Ser
 20 25 30

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

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

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu
 65 70 75 80

Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Arg Asp Glu Arg Gly
 85 90 95

Gly Leu Pro Asp Leu Asn Ala Tyr Trp Glu Gly Ala Ser His Ser Ala
 100 105 110

Lys Gly Pro Arg Trp Lys Ser Asn Gly Arg Pro
 115 120

<210> 92
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 92

Arg Lys Leu Cys Arg Ser Leu Gln Ala Pro Cys Leu Cys Leu Gln Gly
 1 5 10 15

ES 2 545 792 T3

Lys Glu Pro Pro Ser Pro Ala Gly Pro Val Arg Val Leu Ala Ala Ala
 20 25 30

Thr Pro Gly Thr Ser Arg Asn Leu Ala Arg Leu Pro Gly Ser Ser Ser
 35 40 45

Met Val His Pro Ala Gly Pro Leu Ala Ser Gln Thr Gly Ser Val Ala
 50 55 60

Val Gly Leu Gly Gln Thr Ser Leu Ser Pro Ser Ala Asp Trp Ser Leu
 65 70 75 80

Lys Ile Leu Gln Phe Ile Thr Val Ser Ser Glu Thr Asn Asn Gly Lys
 85 90 95

Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro
 100 105 110

Arg Asp Gln Gly Gly Asn Gln Thr Gly Gly Arg
 115 120

<210> 93
 <211> 380
 <212> ADN
 <213> homo sapiens

<400> 93
 cagtctgtgt tgacgcagcc gccctcagtg tctgcggccc caggacagaa ggtcaccatc 60
 tctgtctctg gaagcagctc caacattggg aataattatg tatcctggta ccagcagctc 120
 ccaggaacag cccccaaact cctcatttat gacaataata agcgaccctc agggattcct 180
 gaccgattct ctggctccaa gtctggcaag tcagccacc tgggcatcac cggactccag 240
 actggggagc aggccgatta ttactgogga acaggagacg aataacggta aggcgggtta 300
 ccaggtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga 360
 ccgtcctagg ggcggccgca 380

<210> 94
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 94
 Ala Val Cys Val Asp Ala Ala Ala Leu Ser Val Cys Gly Pro Arg Thr
 1 5 10 15

ES 2 545 792 T3

Glu Gly His His Leu Leu Leu Trp Lys Gln Leu Gln His Trp Glu Leu
 20 25 30

Cys Ile Leu Val Pro Ala Ala Pro Arg Asn Ser Pro Gln Thr Pro His
 35 40 45

Leu Gln Ala Thr Leu Arg Asp Ser Pro Ile Leu Trp Leu Gln Val Trp
 50 55 60

His Val Ser His Pro Gly His His Arg Thr Pro Asp Trp Gly Arg Gly
 65 70 75 80

Arg Leu Leu Leu Arg Asn Arg Arg Arg Ile Thr Val Arg Arg Phe Thr
 85 90 95

Arg Phe Lys Arg Val Leu Gly Arg Arg Val Ser Val Phe Gly Gly Gly
 100 105 110

Thr Lys Leu Thr Val Leu Gly Ala Ala Ala
 115 120

<210> 95
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 95

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

Lys Val Thr Ile Ser Cys Ser Gly Ser Ser Ser Asn Ile Gly Asn Asn
 20 25 30

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

Ile Tyr Asp Asn Asn Lys Arg Pro Ser Gly Ile Pro Asp Arg Phe Ser
 50 55 60

Gly Ser Lys Ser Gly Thr Ser Ala Thr Leu Gly Ile Thr Gly Leu Gln
 65 70 75 80

Thr Gly Asp Glu Ala Asp Tyr Tyr Cys Gly Thr Gly Asp Glu Arg Gly
 85 90 95

ES 2 545 792 T3

Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu Gly Ala Ser Leu Cys Ser
 100 105 110

Ala Glu Gly Pro Ser Pro Ser Gly Arg Pro
 115 120

<210> 96
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 96

Arg Ser Leu Cys Arg Ser Arg Pro Gln Cys Leu Arg Pro Gln Asp Arg
 1 5 10 15

Arg Ser Pro Ser Pro Ala Leu Glu Ala Ala Pro Thr Leu Gly Ile Ile
 20 25 30

Met Tyr Pro Gly Thr Ser Ser Ser Gln Glu Gln Pro Pro Asn Ser Ser
 35 40 45

Phe Met Thr Ile Ile Ser Asp Pro Gln Gly Phe Leu Thr Asp Ser Leu
 50 55 60

Ala Pro Ser Leu Ala Arg Gln Pro Pro Trp Ala Ser Pro Asp Ser Arg
 65 70 75 80

Leu Gly Thr Arg Pro Ile Ile Thr Ala Glu Gln Glu Thr Asn Asn Gly
 85 90 95

Lys Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Cys Val
 100 105 110

Arg Arg Arg Asp Gln Ala Asp Arg Pro Arg Gly Gly Arg
 115 120 125

<210> 97
 <211> 380
 <212> ADN
 <213> Homo sapiens

<400> 97

cagtctgtgc tgactcagcc accctcagcg tctgggaccc cggggcagag ggtcaccatc 60
 tcttgttctg gaagcagctc caacatcgga agtaatactg taaactggta ccagcagctc 120
 ccaggaacgg cccccaaact cctcatctat agtaataatc agcggccctc aggggtccct 180

ES 2 545 792 T3

gaccgattct ctggtccaa gtctggcacc tcagcctccc tggccatcag tgggctccag 240
 tctgaggatg aggctgatta ttactgtgca gcaggagacg aataacggta aggcggttta 300
 ccaggtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga 360
 ccgtcctagg ggcggccgca 380

<210> 98
 <211> 119
 <212> PRT
 <213> Homo sapiens

<400> 98

Ala Val Cys Ala Asp Ser Ala Thr Leu Ser Val Trp Asp Pro Arg Ala
 1 5 10 15

Glu Gly His His Leu Leu Phe Trp Lys Gln Leu Gln His Arg Lys Tyr
 20 25 30

Cys Lys Leu Val Pro Ala Ala Pro Arg Asn Gly Pro Gln Thr Pro His
 35 40 45

Leu Ser Ala Ala Leu Arg Gly Pro Pro Ile Leu Trp Leu Gln Val Trp
 50 55 60

His Leu Ser Leu Pro Gly His Gln Trp Ala Pro Val Gly Gly Leu Leu
 65 70 75 80

Leu Cys Ser Arg Arg Arg Ile Thr Val Arg Arg Phe Thr Arg Phe Lys
 85 90 95

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

Thr Val Leu Gly Ala Ala Ala
 115

<210> 99
 <211> 122
 <212> PRT
 <213> Homo sapiens

<400> 99

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

ES 2 545 792 T3

Arg Val Thr Ile Ser Cys Ser Gly Ser Ser Ser Asn Ile Gly Ser Asn
20 25 30

Thr Val Asn Trp Tyr Gln Gln Leu Pro Gly Thr Ala Pro Lys Leu Leu
35 40 45

Ile Tyr Ser Asn Asn Gln Arg Pro Ser Gly Val Pro Asp Arg Phe Ser
50 55 60

Gly Ser Lys Ser Gly Thr Ser Ala Ser Leu Ala Ile Ser Gly Leu Gln
65 70 75 80

Ser Glu Asp Glu Ala Asp Tyr Tyr Cys Ala Ala Gly Asp Glu Arg Gly
85 90 95

Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu Gly Ala Ser Leu Cys Ser
100 105 110

Ala Glu Gly Pro Ser Pro Ser Gly Arg Pro
115 120

<210> 100
<211> 124
<212> PRT
<213> Homo sapiens

<400> 100

Arg Ser Leu Cys Leu Ser His Pro Gln Arg Leu Gly Pro Pro Gly Arg
1 5 10 15

Gly Ser Pro Ser Leu Val Leu Glu Ala Ala Pro Thr Ser Glu Val Ile
20 25 30

Leu Thr Gly Thr Ser Ser Ser Gln Glu Arg Pro Pro Asn Ser Ser Ser
35 40 45

Ile Val Ile Ile Ser Gly Pro Gln Gly Ser Leu Thr Asp Ser Leu Ala
50 55 60

Pro Ser Leu Ala Pro Gln Pro Pro Trp Pro Ser Val Gly Ser Ser Leu
65 70 75 80

Arg Met Arg Leu Ile Ile Thr Val Gln Gln Glu Thr Asn Asn Gly Lys
85 90 95

ES 2 545 792 T3

Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Cys Val Arg
100 105 110

Arg Arg Asp Gln Ala Asp Arg Pro Arg Gly Gly Arg
115 120

<210> 101
<211> 19
<212> PRT
<213> Homo sapiens

<400> 101

Cys Ala Lys Ala Phe Arg Pro Asn Trp Gly Ser Arg Val Leu Tyr Phe
1 5 10 15

Asp Tyr Trp

<210> 102
<211> 12
<212> PRT
<213> Homo sapiens

<400> 102

Cys Ala Lys Leu Asn Trp Gly Trp Phe Asp Pro Trp
1 5 10

<210> 103
<211> 12
<212> PRT
<213> Homo sapiens

<400> 103

Cys Ala Lys Gly Trp Glu Gly Glu Tyr Asp Tyr Trp
1 5 10

<210> 104
<211> 17
<212> PRT
<213> Homo sapiens

<400> 104

Cys Ala Lys Asp Leu Trp Leu Asp Ser Ser Ser Asn Trp Phe Asp Pro
1 5 10 15

Trp

ES 2 545 792 T3

<210> 105
<211> 11
<212> PRT
<213> Homo sapiens

<400> 105

Cys Ala Lys Asp Leu Pro Gly Asp Pro His Trp
1 5 10

<210> 106
<211> 13
<212> PRT
<213> Homo sapiens

<400> 106

Cys Ala Lys Ala Pro Pro Thr Gly Tyr Phe Asp Tyr Trp
1 5 10

<210> 107
<211> 14
<212> PRT
<213> Homo sapiens

<400> 107

Cys Ala Lys Val Gly Leu Thr Gly Val His Phe Asp Tyr Trp
1 5 10

<210> 108
<211> 16
<212> PRT
<213> Homo sapiens

<400> 108

Cys Ala Lys Asp Ser Tyr Gly Ser Gly Ser Tyr Tyr Asn Asp Tyr Trp
1 5 10 15

<210> 109
<211> 18
<212> PRT
<213> Homo sapiens

<400> 109

Cys Ala Lys Val Ala Gly Thr Trp Gly Arg Val Ala Tyr Tyr Phe Asp
1 5 10 15

Tyr Trp

ES 2 545 792 T3

<210> 110
<211> 15
<212> PRT
<213> Homo sapiens

<400> 110

Cys Ala Lys Val Thr Gly Val Phe Val Gly Asn Phe Asp Tyr Trp
1 5 10 15

<210> 111
<211> 15
<212> PRT
<213> Homo sapiens

<400> 111

Cys Ala Lys Asp Ala Arg Asn Ser Gly Ser Tyr Phe Asp Tyr Trp
1 5 10 15

<210> 112
<211> 14
<212> PRT
<213> Homo sapiens

<400> 112

Cys Ala Lys Gly Val Arg Thr Gly Val Val Phe Asp Tyr Trp
1 5 10

<210> 113
<211> 19
<212> PRT
<213> Homo sapiens

<400> 113

Cys Ala Lys Gly Val Val Asn Trp Gly Thr Arg Arg Lys Gly Trp Phe
1 5 10 15

Asp Pro Trp

<210> 114
<211> 11
<212> PRT
<213> Homo sapiens

<400> 114

Cys Ala Lys Trp Gly Gly Trp Phe Asp Pro Trp
1 5 10

ES 2 545 792 T3

<210> 115
<211> 13
<212> PRT
<213> Homo sapiens

<400> 115

Cys Ala Lys Arg Arg Asp Asn Trp Gly Ser Val Asp Trp
1 5 10

<210> 116
<211> 15
<212> PRT
<213> Homo sapiens

<400> 116

Cys Ala Lys Gly Ser Gly Phe Ser Ser Gly Trp Phe Asp Ser Trp
1 5 10 15

<210> 117
<211> 44
<212> ADN
<213> Homo sapiens

<220>
<221> misc_feature
<222> (13)..(14)
<223> n i s a , t , c o r g

<220>
<221> misc_feature
<222> (16)..(17)
<223> n i s a , t , c o r g

<220>
<221> misc_feature
<222> (19)..(20)
<223> n i s a , t , c o r g

<220>
<221> misc_feature
<222> (22)..(23)
<223> n i s a , t , c o r g

<400> 117
acgtctccga gannsnnsnn snnsatggat tattggggga gacg

44

<210> 118
<211> 14
<212> PRT
<213> Homo sapiens

ES 2 545 792 T3

<220>
 <221> MISC_FEATURE
 <222> (5)..(8)
 <223> Xaa es Cys, Trp, Arg, Ser o Gly
 <400> 118
 Thr Ser Pro Arg Xaa Xaa Xaa Xaa Met Asp Tyr Trp Gly Arg
 1 5 10

<210> 119
 <211> 14
 <212> PRT
 <213> Homo sapiens

<220>
 <221> MISC_FEATURE
 <222> (4)..(4)
 <223> Xaa es Asp o: Glu
 <220>
 <221> MISC_FEATURE
 <222> (5)..(8)
 <223> Xaa es Ser, Pro, Thr, Ala, Cys, Trp, Arg o Gly
 <400> 119

Arg Leu Arg Xaa Xaa Xaa Xaa Xaa Trp Ile Ile Gly Gly Asp
 1 5 10

<210> 120
 <211> 25
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 120
 tgtttctaataat cgcaggtgcc agatg

25

<210> 121
 <211> 25
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 121
 atttatggtta tgacttggtta cactg

25

<210> 122

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

 <220>
 <223> sintetizado químicamente

 <400> 122
 tatttgtttt tatgtttcca atctc 25

 <210> 123
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 123
 ccttgagggt ttatgttatg acttg 25

 <210> 124
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 124
 ttatttccaa ttcagatac caccg 25

 <210> 125
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 125
 ttgttggggt ttttgtttca tgtgg 25

 <210> 126
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 126
 tatttccaat ttcagatacc actgg 25

<210> 127
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 127
 atgttgaatc actgtgggag gccag 25

<210> 128
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 128
 ttatttccaa tctcagatac caccg 25

<210> 129
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 129
 ttttgtttca agctgaatca ctgtg 25

<210> 130
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 130
 atgtctgtgt ctctctcact tccag 25

<210> 131
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 131
 ttcccattg qcctqqaqca ctgtg 25

ES 2 545 792 T3

<210> 132
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 132
 gtgtctgtgt ctctcctgct tccag 25

<210> 133
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 133
 cttgtctcag ttecccattg ggctg 25

<210> 134
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 134
 atctcatcca cttctgtggt ctctc 25

<210> 135
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 135
 ttgggtttct gacaccctca ggatg 25

<210> 136
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 136
 caggccagtc atgtgagact tcacc 25

<210> 137
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 137
 ctgcctctctc cctgggggttt ctgaa 25

<210> 138
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 138
 cccctgtgtc ctctccacag gtgtc 25

<210> 139
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 139
 ccggcacagc tgccttctcc ctgag 25

<210> 140
 <211> 18
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 140
 gaggtgcagc tgttgag 18

<210> 141
 <211> 26
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 141

tctgaccagg gtttcttttt gtttgc	26
<210> 142	
<211> 25	
<212> ADN	
<213> Artificial Sequence	
<220>	
<223> sintetizado químicamente	
<400> 142	
ttgtgtctgg gctcacaatg acttc	25
<210> 143	
<211> 24	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> sintetizado químicamente	
<400> 143	
tggcattttc tgataacggt gtcc	24
<210> 144	
<211> 25	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> sintetizado químicamente	
<400> 144	
ctgcagggag gtttgtgtct gggcg	25
<210> 145	
<211> 25	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> sintetizado químicamente	
<400> 145	
atatgtgtgg cagtttctga ccttg	25
<210> 146	
<211> 25	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> sintetizado químicamente	

ES 2 545 792 T3

<p><400> 146 ggtttgtgtc tgggtgcaca ctgac</p>	25
<p><210> 147 <211> 25 <212> ADN <213> Secuencia artificial</p> <p><220> <223> sintetizado químicamente</p>	
<p><400> 147 gagtctgtgc cggaagtgca gctgg</p>	25
<p><210> 148 <211> 21 <212> ADN <213> Secuencia artificial</p> <p><220> <223> sintetizado químicamente</p>	
<p><400> 148 tatcaggtgc agctggtgca g</p>	21
<p><210> 149 <211> 21 <212> ADN <213> Secuencia artificial</p> <p><220> <223> sintetizado químicamente</p>	
<p><400> 149 tatcaggtgc agctggtgga g</p>	21
<p><210> 150 <211> 21 <212> ADN <213> Artificial Sequence</p> <p><220> <223> sintetizado químicamente</p>	
<p><400> 150 tatgaggtgc agctggtgca g</p>	21
<p><210> 151 <211> 23 <212> ADN <213> Secuencia artificial</p> <p><220> <223> sintetizado químicamente</p>	

<400> 151
 atatctctcg cacagtaata cac 23

<210> 152
 <211> 23
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 152
 atatctctcg cacagtaata tac 23

<210> 153
 <211> 23
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 153
 atatgtctcg cacagtaata cat 23

<210> 154
 <211> 32
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 154
 tatgacatcc agatgaccca gtctccatcc tc 32

<210> 155
 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 155
 ataggagggg tactgtaact 20

<210> 156
 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>

ES 2 545 792 T3

<code><223></code>	sintetizado químicamente	
<code><400></code>	156	
	ataggaggga gattatcata	20
<code><210></code>	157	
<code><211></code>	24	
<code><212></code>	ADN	
<code><213></code>	Secuencia artificial	
<code><220></code>		
<code><223></code>	sintetizado químicamente	
<code><400></code>	157	
	tatgaaattg tgttgacgca gtct	24
<code><210></code>	158	
<code><211></code>	23	
<code><212></code>	ADN	
<code><213></code>	Secuencia artificial	
<code><220></code>		
<code><223></code>	sintetizado químicamente	
<code><400></code>	158	
	ataggaggtg agctaccata ctg	23
<code><210></code>	159	
<code><211></code>	24	
<code><212></code>	DNA	
<code><213></code>	Secuencia artificial	
<code><220></code>		
<code><223></code>	sintetizado químicamente	
<code><400></code>	159	
	tatgaaatag tgatgacgca gtct	24
<code><210></code>	160	
<code><211></code>	23	
<code><212></code>	ADN	
<code><213></code>	Secuencia artificial	
<code><220></code>		
<code><223></code>	sintetizado químicamente	
<code><400></code>	160	
	ataggaggcc agttattata ctg	23
<code><210></code>	161	
<code><211></code>	23	
<code><212></code>	ADN	
<code><213></code>	Secuencia artificial	

ES 2 545 792 T3

<220>
 <223> sintetizado químicamente

 <400> 161
 cagcgtagca actggcctcc tat 23

 <210> 162
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 162
 tacagtctgt gctgactcag 20

 <210> 163
 <211> 23
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 163
 ataggaccat tcaggctgtc atc 23

 <210> 164
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 164
 tatcagtctg tgttgacgca g 21

 <210> 165
 <211> 22
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 165
 ataggagcac tcaggctgct at 22

 <210> 166
 <211> 33
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

 <400> 166
 cagccggcca tggcccaggt gcagctggtg cag 33

 <210> 167
 <211> 33
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 167
 cagccggcca tggcccaggt gcagctggtg gag 33

 <210> 168
 <211> 33
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 168
 cagccggcca tggccgaggt gcagctggtg gag 33

 <210> 169
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 169
 cagccggcca tggccgaggt gcagctggtg gagtctgggg gag 43

 <210> 170
 <211> 33
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 170
 cagccggcca tggccgaggt gcagctggtg cag 33

 <210> 171
 <211> 37
 <212> ADN

ES 2 545 792 T3

<213> Artificial Sequence
 <220>
 <223> sintetizado químicamente
 <400> 171
 cttaccggtta ttcgtctcat ctcgcacagt aatacac 37

<210> 172
 <211> 37
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 172
 cttaccggtta ttcgtctcat ttcgcacagt aatatac 37

<210> 173
 <211> 42
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 173
 ctcgcacagt aatacacagc cgtgtcctcg gctctcagc tg 42

<210> 174
 <211> 37
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 174
 cttaccggtta ttcgtctcat ctcgcacagt aatacat 37

<210> 175
 <211> 47
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 175
 caatacgcgt ttaaacctgg taaaccgcct taccggtatt cgtctca 47

<210> 176
 <211> 48

ES 2 545 792 T3

<212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 176
 gttcctggc cccaagagac ggccttccc aatagcggtt taaacctg 48

 <210> 177
 <211> 45
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 177
 cctccaccgc tcgagactgt gaccagggtt ccctggcccc aagag 45

 <210> 178
 <211> 30
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 178
 cgggtcgacg gacatccaga tgaccagtc 30

 <210> 179
 <211> 36
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 179
 cgggtcgacg gaaattgtgt tgacacagtc tccagc 36

 <210> 180
 <211> 36
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 180
 cgggtcgacg gaaatagtga tgaccagtc tccagc 36

 <210> 181

ES 2 545 792 T3

<211> 36
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 181
 cgggtcgacg gaaattgtgt tgacgcagtc tccagg 36

<210> 182
 <211> 45
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 182
 ccttaccggtt attcgtctcg ctgctgacag taatatgttg caata 45

<210> 183
 <211> 45
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 183
 ccttaccggtt attcgtctcg ctgctgacag tagtaagttg caaaa 45

<210> 184
 <211> 47
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 184
 ccttaccggtt attcgtctcg ctgctgacag taataaactg caaaatc 47

<210> 185
 <211> 47
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 185
 ccaaatacgcg tttaaacctg gtaaaccgcc ttaccgttat tegtctc 47

ES 2 545 792 T3

```

<210> 186
<211> 46
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 186
ggtcccttg cgaatgaga cgcgccttcc caatagcgt ttaaac 46

<210> 187
<211> 45
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 187
gtgcgccgc cgtttgatt tccacctgg tcccttgcc gaatg 45

<210> 188
<211> 45
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 188
gtgcgccgc cctttgatt tccacctgg tcccttgcc gaatg 45

<210> 189
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 189
cgggtcgacg cagtctgtg tgactcagcc ac 32

<210> 190
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 190
cgggtcgacg cagtctgtg tgacgcagcc gc 32

```

ES 2 545 792 T3

<210> 191
 <211> 38
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 191
 ccctaccggtt attcgtctcc tgctgcacag taataatc 38

<210> 192
 <211> 38
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 192
 ccctaccggtt attcgtctcc tgttccgcag taataatc 38

<210> 193
 <211> 46
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 193
 ccctccgcgc aacacagaga cgcgccttcc caatacgcgt ttaaac 46

<210> 194
 <211> 49
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 194
 gtgcggccgc ccctaggacg gtcagcttgg tcctccgcc gaacacaga 49

<210> 195
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 195
 ccgcacagcc ggccatggcc caggtgcagc tggcgcagtc tgg 43

<210> 196
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 196
 ccgcacagcc ggccatggcc gaggtgcagc tggaggagtc tgg 43

<210> 197
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 197
 ccgcacagcc ggccatggcc cagrtcacct tgctcgagtc tgg 43

<210> 198
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 198
 ccgcacagcc ggccatggcc caggtgcagc tgcaggagtc ggg 43

<210> 199
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 199
 ccgcacagcc ggccatggcc cagctgcagc tgcaggagtc cgg 43

<210> 200
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 200

ES 2 545 792 T3

<p>ccgcacagcc ggccatggcc caggtgcagc tacagcagtg ggg</p> <p><210> 201 <211> 32 <212> ADN <213> Secuencia artificial</p> <p><220> <223> sintetizado químicamente</p> <p><400> 201 tggagtgggt ctcagctatt agtggtagtg gt</p> <p><210> 202 <211> 45 <212> ADN <213> Secuencia artificial</p> <p><220> <223> sintetizado químicamente</p> <p><400> 202 cgatgggccc ttggtggagg ctgaggagac rgtgaccagg gtgcc</p> <p><210> 203 <211> 45 <212> ADN <213> Secuencia artificial</p> <p><220> <223> sintetizado químicamente</p> <p><400> 203 cgatgggccc ttggtggagg ctgaagagac ggtgaccrck gtccc</p> <p><210> 204 <211> 45 <212> ADN <213> Secuencia artificial</p> <p><220> <223> sintetizado químicamente</p> <p><400> 204 cgatgggccc ttggtggagg ctgaggagac ggtgaccagg gttcc</p> <p><210> 205 <211> 35 <212> ADN <213> Secuencia artificial</p> <p><220> <223> sintetizado químicamente</p>	<p>43</p> <p>32</p> <p>45</p> <p>45</p> <p>45</p> <p>45</p>
---	---

ES 2 545 792 T3

<400> 205
 gagccgagga cacggccgga tgttactgtg cgaga 35

<210> 206
 <211> 35
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 206
 gagccgagga cacggccgga tgttactgtg cgaaa 35

<210> 207
 <211> 32
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 207
 gaggagacgg tgacggatgt gccttgccc ca 32

<210> 208
 <211> 32
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 208
 gaggagacgg tgacggatgt gccacggccc ca 32

<210> 209
 <211> 32
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 209
 gaggagacgg tgacggatgt yccttgccc ca 32

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

<220>
 <223> sintetizado químicamente

ES 2 545 792 T3

<400> 210
atgatgctgc tggcacgtct ccgaga 26

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

<220>
<223> sintetizado químicamente

<400> 211
ccacgtcatc cgatccgtct cccccaataa tccat 35

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

<220>
<223> sintetizado químicamente

<400> 212
ccacgtcatc cgatccgtct cccccaataa tcaaa 35

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

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (19)..(20)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (22)..(23)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (25)..(26)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (28)..(29)
<223> n e s a , c , g , o t

<400> 213

<210> 214
 <211> 50
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

<220>
 <221> misc_feature
 <222> (19)..(20)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (22)..(23)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (25)..(26)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (28)..(29)
 <223> n e s a , c , g , o t

<400> 214
 gctggcacgt ctccgagann snnsnnsnns atggattatt gggggagacg

50

<210> 215
 <211> 53
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<220>
 <221> misc_feature
 <222> (19)..(20)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (22)..(23)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (25)..(26)
 <223> n e s a , c , g , o t

ES 2 545 792 T3

<220>
<221> misc_feature
<222> (28)..(29)
<223> n e s a, c, g, o t

<220>
<221> misc_feature
<222> (31)..(32)
<223> n e s a, c, g, o t

<400> 215
gctggcacgt ctccgagann snnsnnsnns nstttgatt attgggggag acg

53

<210> 216
<211> 53
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (19)..(20)
<223> n e s a, c, g, o t

<220>
<221> misc_feature
<222> (22)..(23)
<223> n e s a, c, g, o t

<220>
<221> misc_feature
<222> (25)..(26)
<223> n e s a, c, g, o t

<220>
<221> misc_feature
<222> (28)..(29)
<223> n e s a, c, g, o t

<220>
<221> misc_feature
<222> (31)..(32)
<223> n e s a, c, g, o t

<400> 216
gctggcacgt ctccgagann snnsnnsnns nnsatggatt attgggggag acg

53

<210> 217
<211> 56
<212> ADN
<213> Secuencia artificial

<220>

<223> sintetizado químicamente

<220>

<221> misc_feature

<222> (19)..(20)

<223> n és a, t, c o g

<220>

<221> misc_feature

<222> (22)..(23)

<223> n és a, t, c o g

<220>

<221> misc_feature

<222> (25)..(26)

<223> n és a, t, c o g

<220>

<221> misc_feature

<222> (28)..(29)

<223> n és a, t, c o g

<220>

<221> misc_feature

<222> (31)..(32)

<223> n és a, t, c o g

<220>

<221> misc_feature

<222> (34)..(35)

<223> n és a, t, c o g

<400> 217

gctggcacgt ctccgagann snnsnnsnns nnsnntttg attattgggg gagacg

56

<210> 218

<211> 56

<212> ADN

<213> Secuencia artificial

<220>

<223> sintetizado químicamente

<220>

<221> misc_feature

<222> (19)..(20)

<223> n és a, c, g, o t

<220>

<221> misc_feature

<222> (22)..(23)

<223> n és a, c, g, o t

<220>

<221> misc_feature

<222> (25)..(26)

ES 2 545 792 T3

<223> n é s a, t, c o g
 <220>
 <221> misc_feature
 <222> (28)..(29)
 <223> n é s a, t, c o g
 <220>
 <221> misc_feature
 <222> (31)..(32)
 <223> n é s a, t, c o g
 <220>
 <221> misc_feature
 <222> (34)..(35)
 <223> n é s a, t, c o g
 <400> 218
 gctggcacgt ctccgagann snnsnnsnns nnsnnsatgg attattgggg gagacg 56
 <210> 219
 <211> 56
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 219
 gctggcacgt ctccgagadv kdvkdvkdvk dvkdvktttg attattgggg gagacg 56
 <210> 220
 <211> 56
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 220
 gctggcacgt ctccgagadv kdvkdvkdvk dvkdvkatgg attattgggg gagacg 56
 <210> 221
 <211> 59
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 221
 gctggcacgt ctccgagadv kdvkdvkdvk dvkdvkdvkt ttgattattg ggggagacg 59
 <210> 222
 <211> 59

ES 2 545 792 T3

<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 222
gctggcacgt ctccgagadv kdvkdvkdvk dvkdvkdvka tggattattg ggggagacg 59

<210> 223
<211> 59
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (19)..(19)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (22)..(22)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (25)..(25)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (28)..(28)
<223> n i s a , c , g , o t

<220>
<221> misc_feature
<222> (31)..(31)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (34)..(34)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (37)..(37)
<223> n e s a , c , g , o t

<400> 223
gctggcacgt ctccgaganv tntvntvntv nvtvntvntt ttgattattg ggggagacg 59

<210> 224

ES 2 545 792 T3

<211> 59
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (19)..(19)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (22)..(22)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (25)..(25)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (28)..(28)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (31)..(31)
<223> n i s a , c , g , o t

<220>
<221> misc_feature
<222> (34)..(34)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (37)..(37)
<223> n e s a , c , g , o t

<400> 224
gctggcacgt ctccgaganv tntnvtntvt nvtntvntvta tggattattg ggggagacg

59

<210> 225
<211> 62
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (19)..(19)

```

<223> n è s a, t, c o g:

<220>
<221> misc_feature
<222> (22)..(22)
<223> n è s a, t, c o g:

<220>
<221> misc_feature
<222> (25)..(25)
<223> n e s a, c, g, o t

<220>
<221> misc_feature
<222> (28)..(28)
<223> n è s a, t, c o g

<220>
<221> misc_feature
<222> (31)..(31)
<223> n è s a, t, c o g

<220>
<221> misc_feature
<222> (34)..(34)
<223> n è s a, t, c o g

<220>
<221> misc_feature
<222> (37)..(37)
<223> n è s a, t, c o g

<220>
<221> misc_feature
<222> (40)..(40)
<223> n è s a, t, c o g

<400> 225
gotggcactg ctccgaganv tnvtnvntvnt nvtvntvntn vttttgatta ttgggggaga 60
cg 62

<210> 226
<211> 62
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (19)..(19)
<223> n e s a, c, g, o t

<220>
<221> misc_feature

```

```

<222> (22)..(22)
<223> n e s a, t, c o g

<220>
<221> misc_feature
<222> (25)..(25)
<223> n e s a, t, c o g

<220>
<221> misc_feature
<222> (28)..(28)
<223> n e s a, t, c o g

<220>
<221> misc_feature
<222> (31)..(31)
<223> n e s a, t, c o g

<220>
<221> misc_feature
<222> (34)..(34)
<223> n e s a, t, c o g

<220>
<221> misc_feature
<222> (37)..(37)
<223> n e s a, t, c o g

<220>
<221> misc_feature
<222> (40)..(40)
<223> n e s a, t, c o g

<400> 226
gctggcacgt ctccgaganv tnvtntnvt nvtntnvt vtatggatta ttgggggaga 60
cg 62

<210> 227
<211> 65
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (19)..(19)
<223> n e s a, c, g, o t

<220>
<221> misc_feature
<222> (22)..(22)
<223> n e s a, c, g, o t

<220>

```

<221> misc_feature
 <222> (25)..(25)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (28)..(28)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (31)..(31)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (34)..(34)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (37)..(37)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (40)..(40)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (43)..(43)
 <223> n e s a , c , g , o t

<400> 227
 gctggcacgt ctccgaganv tntvntvnt nvtvntvntn vtnvttttga ttattggggg 60

agacg 65

<210> 228
 <211> 65
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<220>
 <221> misc_feature
 <222> (19)..(19)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (22)..(22)
 <223> n e s a , c , g , o t

ES 2 545 792 T3

```

<220>
<221> misc_feature
<222> (25)..(25)
<223> n e s a , c , g , o . t

<220>
<221> misc_feature
<222> (28)..(28)
<223> n e s a , c , g , o . t

<220>
<221> misc_feature
<222> (31)..(31)
<223> n e s a , c , g , o . t

<220>
<221> misc_feature
<222> (34)..(34)
<223> n e s a , c , g , o . t

<220>
<221> misc_feature
<222> (37)..(37)
<223> n e s a , c , g , o . t

<220>
<221> misc_feature
<222> (40)..(40)
<223> n e s a , c , g , o . t

<220>
<221> misc_feature
<222> (43)..(43)
<223> n e s a , c , g , o . t

<400> 228
gctggcacgt ctccgaganv tnvtnvtnvt nvtvtnvtn vtnvtatgga ttattggggg 60
agacg 65

<210> 229
<211> 65
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 229
gctggcacgt ctccgagadv tdvtdvtdvt dvtvtdvtd vtdvttttga ttattggggg 60
agacg 65

<210> 230
<211> 65
<212> ADN
<213> Secuencia artificial

```

ES 2 545 792 T3

```

<220>
<223> sintetizado químicamente

<400> 230
gctggcacgt ctccgagadv tdvtdvtdvt dvtvdvtdvtd vtdvtdatgga ttattggggg      60
agacg                                                                           65

<210> 231
<211> 68
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 231
gctggcacgt ctccgagadv tdvtdvtdvt dvtvdvtdvtd vtdvtdvttt tgattattgg      60
gggagacg                                                                           68

<210> 232
<211> 68
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 232
gctggcacgt ctccgagadv tdvtdvtdvt dvtvdvtdvtd vtdvtdvttat ggattattgg      60
gggagacg                                                                           68

<210> 233
<211> 26
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 233
ccggtgtagc gaaggcgtct cagcag                                                                           26

<210> 234
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 234

```

tagggtcgcc ttgatcgtct cccgaaggtc gg

32

<210> 235
 <211> 45
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente

<220>
 <221> misc_feature
 <222> (17)..(18)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (20)..(21)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (23)..(24)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (26)..(27)
 <223> n e s a, c, g, o t

<400> 235
 gaaggcgtct cagcagnnsn nsnnsnnscc gaccttcggg agacg

45

<210> 236
 <211> 48
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<220>
 <221> misc_feature
 <222> (17)..(18)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (20)..(21)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (23)..(24)
 <223> n e s a, c, g, o t

ES 2 545 792 T3

<220>
<221> misc_feature
<222> (26)..(27)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (32)..(33)
<223> n e s a , c , g , o t

<400> 236
gaaggcgtct cagcagnnsn nsnnnsnscg gnsaccttc gggagacg

48

<210> 237
<211> 51
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (17)..(18)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (20)..(21)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (23)..(24)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (26)..(27)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (29)..(30)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (35)..(36)
<223> n e s a , c , g , o t

<400> 237
gaaggcgtct cagcagnnsn nsnnnsnscg sccgnsacc ttcgggagac g

51

<210> 238


```

<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 238
cggtcagtcg caatacgtct ccagcatggg at 32

<210> 239
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 239
cggtcagtcg caatacgtct ccagcatatg at 32

<210> 240
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 240
caggaccagt ctcgtgagga tcgtctcaac ac 32

<210> 241
<211> 46
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (18)..(19)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (21)..(22)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n e s a , c , g , o t

```

ES 2 545 792 T3

<220>
<221> misc_feature
<222> (27)..(28)
<223> n e s a , c , g , o t

<400> 241
cgtctccagc atgggatnns nnsnnsnsg tggtagagacg atcctc 46

<210> 242
<211> 46
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (18)..(19)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (21)..(22)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (27)..(28)
<223> n e s a , c , g , o t

<400> 242
cgtctccagc atatgatnns nnsnnsnsg tggtagagacg atcctc 46

<210> 243
<211> 49
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (18)..(19)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (21)..(22)

<223> n es a, c, g, o t

<220>

<221> misc_feature

<222> (24)..(25)

<223> n es a, c, g, o t

<220>

<221> misc_feature

<222> (27)..(28)

<223> n es a, c, g, o t

<220>

<221> misc_feature

<222> (30)..(31)

<223> n es a, c, g, o t

<400> 243

cgtctccagc atgggatnns nnsnnsnnsn nsgtggtgag acgatcctc

49

<210> 244

<211> 49

<212> ADN

<213> Secuencia artificial

<220>

<223> sintetizado químicamente

<220>

<221> misc_feature

<222> (18)..(19)

<223> n es a, c, g, o t

<220>

<221> misc_feature

<222> (21)..(22)

<223> n es a, c, g, o t

<220>

<221> misc_feature

<222> (24)..(25)

<223> n es a, c, g, o t

<220>

<221> misc_feature

<222> (27)..(28)

<223> n es a, c, g, o t

<220>

<221> misc_feature

<222> (30)..(31)

<223> n es a, c, g, o t

<400> 244

cgtctccagc atatgatnns nnsnnsnnsn nsgtggtgag acgatcctc

49

ES 2 545 792 T3

<210> 245
<211> 52
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (18)..(19)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (21)..(22)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (27)..(28)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (30)..(31)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (33)..(34)
<223> n e s a , c , g , o t

<400> 245
cgtctccagc atgggatnns nnsnnsnnsn nsnsgtgtt gagacgatcc tc

52

<210> 246
<211> 52
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (18)..(19)
<223> n e s a , c , g , o t

<220>
<221> misc_feature

```

<222> (21)..(22)
<223> n e s a, c, g, o t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n e s a, c, g, o t

<220>
<221> misc_feature
<222> (27)..(28)
<223> n e s a, c, g, o t

<220>
<221> misc_feature
<222> (30)..(31)
<223> n e s a, c, g, o t

<220>
<221> misc_feature
<222> (33)..(34)
<223> n e s a, c, g, o t

<400> 246
cgctccagc atatgatnns nnsnnsnnsn nsnnsrgtt gagacgatcc tc 52

<210> 247
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 247
cggtcagtcg caatagctct cgaacatggg at 32

<210> 248
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 248
cggtcagtcg caatagctct cgaacatggg at 32

<210> 249
<211> 46
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

```

```

<220>
<221> misc_feature
<222> (18)..(19)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (21)..(22)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (27)..(28)
<223> n e s a , c , g , o t

<400> 249
cgtctcgaac atgggatnns nnsnnsnsg tgttgagacg atcctc 46

<210> 250
<211> 46
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<220>
<221> misc_feature
<222> (18)..(19)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (21)..(22)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n e s a , c , g , o t

<220>
<221> misc_feature
<222> (27)..(28)
<223> n e s a , c , g , o t

<400> 250
cgtctcgaac atatgatnns nnsnnsnsg tgttgagacg atcctc 46

<210> 251

```

<211> 49
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<220>
 <221> misc_feature
 <222> (18)..(19)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (21)..(22)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (24)..(25)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (27)..(28)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (30)..(31)
 <223> n e s a, c, g, o t

<400> 251
 cgctcgaac atgggatnns nnsnnsnnsn nsgtgttgag acgacctc

49

<210> 252
 <211> 49
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<220>
 <221> misc_feature
 <222> (18)..(19)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (21)..(22)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (24)..(25)

<223> n e s a, c, g, o t
 <220>
 <221> misc_feature
 <222> (27)..(28)
 <223> n e s a, c, g, o t
 <220>
 <221> misc_feature
 <222> (30)..(31)
 <223> n e s a, c, g, o t
 <400> 252
 cgtctcgaac atatgatnns nnsnnsnnsn nsgtgttgag acgatcctc

49

<210> 253
 <211> 52
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente

<220>
 <221> misc_feature
 <222> (18)..(19)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (21)..(22)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (24)..(25)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (27)..(28)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (30)..(31)
 <223> n e s a, c, g, o t

<220>
 <221> misc_feature
 <222> (33)..(34)
 <223> n e s a, c, g, o t

<400> 253
 cgtctcgaac atgggatnns nnsnnsnnsn nsnsgtgtt gagacgatcc tc

52

ES 2 545 792 T3

<210> 254
 <211> 52
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente

<220>
 <221> misc_feature
 <222> (18)..(19)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (21)..(22)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (24)..(25)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (27)..(28)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (30)..(31)
 <223> n e s a , c , g , o t

<220>
 <221> misc_feature
 <222> (33)..(34)
 <223> n e s a , c , g , o t

<400> 254
 cgtctcgaac atgggatnns nnsnnsnnsn nsnnsgtggt gagacgatcc tc 52

<210> 255
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente

<400> 255
 ctcttctgag atgagttttt g 21

<210> 256
 <211> 43
 <212> ADN
 <213> Secuencia artificial

ES 2 545 792 T3

<220>
 <223> sintetizado químicamente
 <400> 256
 atgCGGCCCA gCGGCCatg GCCSaggtYc agctbcagca gtc 43

<210> 257
 <211> 43
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente
 <400> 257
 atgCGGCCCA gCGGCCatg GCCCaggttc acctgcagca rtc 43

<210> 258
 <211> 43
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente
 <400> 258
 atgCGGCCCA gCGGCCatg GCCCaggttc agctgaagga gtc 43

<210> 259
 <211> 43
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente
 <400> 259
 atgCGGCCCA gCGGCCatg GCCCaggttc aactvcagca rcc 43

<210> 260
 <211> 43
 <212> ADN
 <213> Artificial Sequence

<220>
 <223> sintetizado químicamente
 <400> 260
 atgCGGCCCA gCGGCCatg GCCCagatcc agttggtvca gtc 43

<210> 261
 <211> 43
 <212> ADN

<213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 261
 atgcggccca gccggccatg gcccgaggtgc agctgaagsa stc 43

<210> 262
 <211> 43
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 262
 atgcggccca gccggccatg gcccgaggtgc agskggtgga gtc 43

<210> 263
 <211> 43
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 263
 atgcggccca gccggccatg gccgaagtga arsttgagga gtc 43

<210> 264
 <211> 43
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 264
 atgcggccca gccggccatg gccgagktsv agcttcagga gtc 43

<210> 265
 <211> 43
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizado químicamente
 <400> 265
 atgcggccca gccggccatg gcccgaggtga asstggtgga atc 43

<210> 266
 <211> 43

ES 2 545 792 T3

<212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 266
 atgcggccca gccggccatg gccgaggtga agctgrtgga rtc 43

 <210> 267
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 267
 atgcggccca gccggccatg gccgargtga agctgrtgga gtc 43

 <210> 268
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 268
 atgcggccca gccggccatg gccgaagtgc agctgttgga gac 43

 <210> 269
 <211> 43
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 269
 atgcggccca gccggccatg gccgargtga agcttctcsa gtc 43

 <210> 270
 <211> 42
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 270
 atgcggccca gccggccatg gccargtta ctctgaaga gt 42

 <210> 271

ES 2 545 792 T3

<211> 41
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 271
 cctgaaccgc cgctccgct cgagacggtg accgtggtcc c 41

 <210> 272
 <211> 41
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 272
 cctgaaccgc cgctccgct cgagactgtg agagtgggtgc c 41

 <210> 273
 <211> 41
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 273
 cctgaaccgc cgctccgct cgagacagtg accagagtcc c 41

 <210> 274
 <211> 41
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 274
 cctgaaccgc cgctccgct cgagacggtg actgagggttc c 41

 <210> 275
 <211> 35
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> sintetizado químicamente

 <400> 275
 gagccgagga cacggccgga tgttactgtg cgaga 35

ES 2 545 792 T3

<210> 276
<211> 31
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 276
ggggcgcagg gacatccgtc accgtctect c 31

<210> 277
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 277
gaggagactg tgagggatgt gccttgccc ca 32

<210> 278
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 278
gaggagacgg tgacggatgt gccttgccc ca 32

<210> 279
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 279
gaggagacgg tgacggatgt tccttgacc ca 32

<210> 280
<211> 14
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (4)..(4)
<223> Xaa is Ile or Met

ES 2 545 792 T3

<220>
 <221> MISC_FEATURE
 <222> (5)..(8)
 <223> Xaa es Leu, Pro, His, Gln, Arg, Val, Ala, Asp, Glu o Gly
 <400> 280

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

<210> 281
 <211> 13
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 281
 ttactgtgcg aga 13

<210> 282
 <211> 13
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 282
 ttactgtgca aga 13

<210> 283
 <211> 13
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 283
 tttctgtgca aga 13

<210> 284
 <211> 13
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizado químicamente

<400> 284
 ctactgtgcc aga 13

<210> 285
<211> 13
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 285
tggggccagg gaa 13

<210> 286
<211> 13
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 286
tggggcgag gga 13

<210> 287
<211> 13
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 287
tggggccaag gca 13

<210> 288
<211> 13
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 288
tggggccagg gca 13

<210> 289
<211> 13
<212> ADN
<213> Secuencia artificial

<220>
<223> sintetizado químicamente

<400> 289
tggggtcagg gca 13

Reivindicaciones:

1. Un método para producir una colección de ácidos nucleicos, donde cada ácido nucleico codifica un dominio variable de la inmunoglobulina humana que comprende una pluralidad de secuencias de la región determinante de complementariedad 3 (CDR3) aisladas separadamente del repertorio del dominio variable de la inmunoglobulina de una especie de mamífero, aislado separadamente del repertorio del dominio variable de la inmunoglobulina de un ser humano, aislado separadamente del repertorio del dominio variable de la inmunoglobulina de una especie de mamífero no humano, o aislado separadamente de los dominios variables de la inmunoglobulina de un mamífero no humano humanizado, el método que comprende:
- 5
- 10 (a) proporcionar una pluralidad de secuencias de ácidos nucleicos de la estructura aceptora que codifican dominios variables de inmunoglobulina humana distintos, cada secuencia de ácidos nucleicos de la estructura aceptora que comprende una primera región estructural (FR1), una segunda región estructural (FR2), una tercera región estructural (FR3), y una cuarta región estructural (FR4), donde las regiones FR1 y FR2 están intercaladas en una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas en una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas en una secuencia de ácidos nucleicos de relleno que comprende al menos dos sitios de reconocimiento de la enzimas de restricción Tipo IIs intercalados en una secuencia de ácidos nucleicos aleatoria;
- 15
- (b) proporcionar una pluralidad de secuencias de ácidos nucleicos diversificadas que codifican las secuencias de la región determinante de complementariedad 3 (CDR3) aisladas del repertorio de inmunoglobulina de la especie de mamífero donde cada una de la pluralidad de secuencias de ácidos nucleicos diversificadas comprende un sitio de reconocimiento de las enzimas de restricción Tipo IIs en cada extremidad;
- 20
- (c) digerir cada una de la pluralidad de las secuencias de ácidos nucleicos que codifica las regiones CDR3 usando una enzima de restricción Tipo II que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (b) y digerir la secuencia de ácidos nucleicos de relleno de la etapa (a) de la estructura aceptora usando una enzima de restricción Tipo II que se une al sitio de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a); y
- 25
- (d) ligar las secuencias de ácidos nucleicos digeridas que codifican las regiones de CDR3 o las secuencias de aminoácidos de la etapa (c) en la estructura aceptora digerida de la etapa (c) de modo que las regiones FR3 y FR4 se intercalan en las secuencias de ácidos nucleicos que codifican la región CDR3 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR3 y se restauran secuencias que codifican el dominio variable de la inmunoglobulina completa que no contienen los sitios de reconocimiento de la enzimas de restricción Tipo IIs de las etapas (a) y (b).
- 30
2. El método de la reivindicación 1, donde la etapa (b) se realiza mediante amplificación de la secuencia CDR3 de una especie de mamífero usando cebadores de oligonucleótidos que contienen un sitio de restricción Tipo IIs, donde la etapa (b) se realiza mediante la amplificación de la secuencia CDR3 de una especie de mamífero no humano usando cebadores de oligonucleótidos que contienen un sitio de restricción FokI IIs, o donde la etapa (b) se realiza mediante la amplificación de la secuencia CDR3 del mamífero no humano inmunizado usando cebadores de oligonucleótidos que contienen un sitio de restricción Tipo IIs.
- 35
3. El método de la reivindicación 2, donde el cebador de oligonucleótido se diseña para aumentar la compatibilidad entre la secuencia de CDR3 de mamífero y la estructura aceptora que codifica un dominio variable de la inmunoglobulina humana.
- 40
4. El método de la reivindicación 3, donde el cebador de oligonucleótido se diseña para modificar una secuencia de ácidos nucleicos en un límite de la secuencia de CDR3 de mamífero para producir una compatible secuencia de nucleótidos cohesiva en la estructura aceptora que codifica un dominio variable de la inmunoglobulina humana.
- 45
5. El método de la reivindicación 1, donde la especie no humana es primates no humanos, roedores, caninos, felinos, ovejas, cabras, vacas, caballos, un miembro de la familia de los camélidos, llama, camello, dromedario, o cerdo.
6. El método de la reivindicación 1, donde los sitios de reconocimiento de la enzima de restricción Tipo IIs de la etapa (a) y la etapa (b) son reconocidos por una enzima de restricción Tipo IIs diferente.
- 50
7. El método de la reivindicación 6, donde los sitios de reconocimiento de la enzimas de restricción Tipo IIs son sitios de reconocimiento BsmBI, sitios de reconocimiento BsaI, sitios de reconocimiento FokI o una combinación de estos.
8. El método de la reivindicación 1, donde las secuencias de ácidos nucleicos diversificadas que codifican las secuencias CDR3 codifican secuencias CDR3 de la cadena pesada (CDR H3), o donde las secuencias de ácidos
- 55

nucleicos diversificadas que codifican las secuencias CDR3 codifican secuencias CDR3 (CDR L3) de cadena liviana.

9. El método de la reivindicación 1, donde la secuencia de ácidos nucleicos de la estructura aceptora comprende una secuencia del gen variable de la cadena pesada humana seleccionada de VH1-2, V1-11-69, VH1-18, VH3-30, VH3-48, VH3-23, y VH5-51.
- 5 10. El método de la reivindicación 1, donde la secuencia de ácidos nucleicos de la estructura aceptora comprende una secuencia del gen variable de la cadena liviana kappa humana, opcionalmente, donde la secuencia del gen variable de la cadena liviana kappa humana se selecciona de VK1-33, VKI-39, VK3-11, VK3-15, y VK3-20, o una secuencia del gen variable de la cadena liviana lambda humana, opcionalmente, donde la secuencia del gen variable de la cadena liviana lambda humana se selecciona de VL1-44 y VL1-51.
- 10 11. El método de la reivindicación 1, donde la pluralidad de las secuencias de ácidos nucleicos de la estructura aceptora comprende una mezcla de al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena pesada (VH) variable y al menos una secuencia de ácidos nucleicos de la estructura aceptora de la cadena liviana variable.
- 15 12. El método de la reivindicación 1, que además comprende las etapas de (e) clonar la biblioteca de ácidos nucleicos que codifican los dominios variables de la inmunoglobulina de la etapa (d) en un vector de expresión y (f) transformar el vector de expresión de la etapa (e) en una célula huésped y cultivar la célula huésped en condiciones suficientes para expresar una pluralidad de dominio variable de la inmunoglobulina codificado por la biblioteca.
13. El método de la reivindicación 12, donde la célula huésped es E. coli.
- 20 14. El método de acuerdo con la reivindicación 12, donde el vector de expresión es un fagémido o un vector fago.

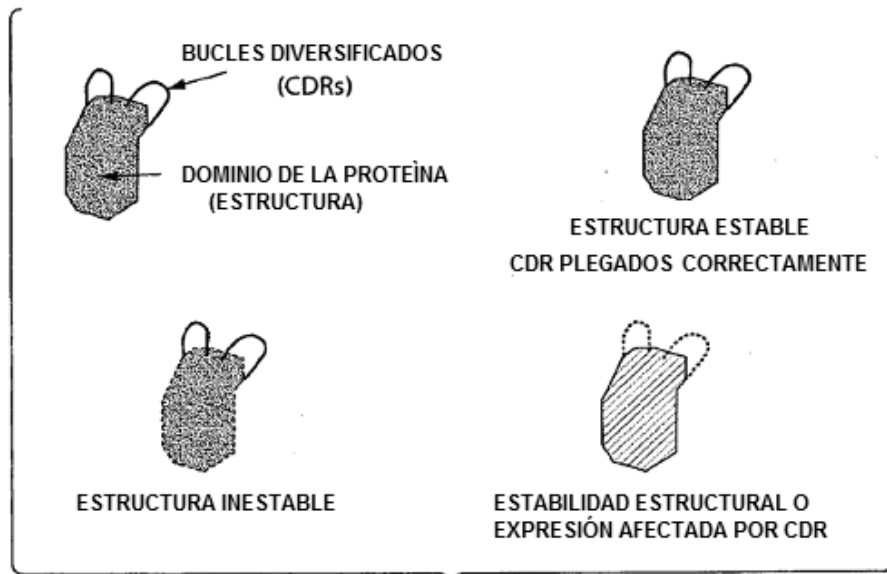


Fig. 1A

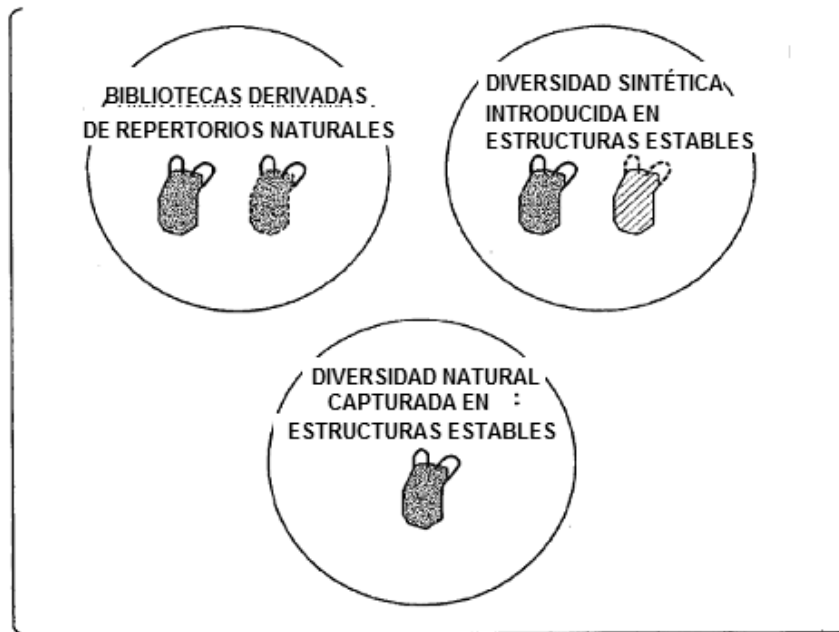


Fig. 1B

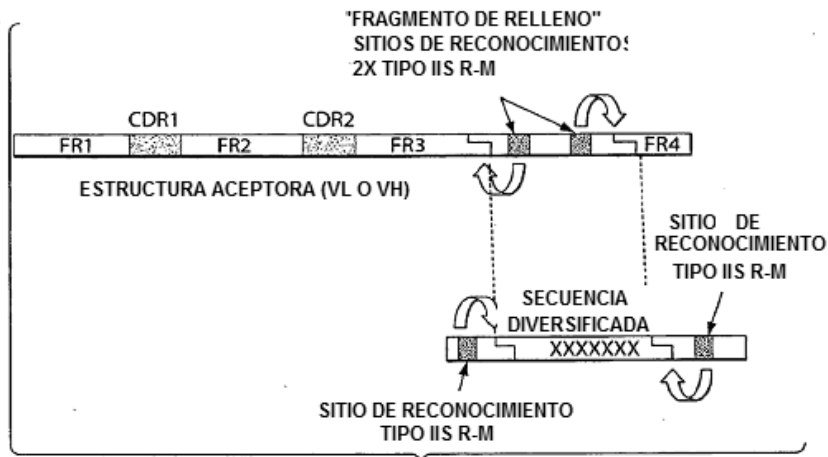


Fig. 2

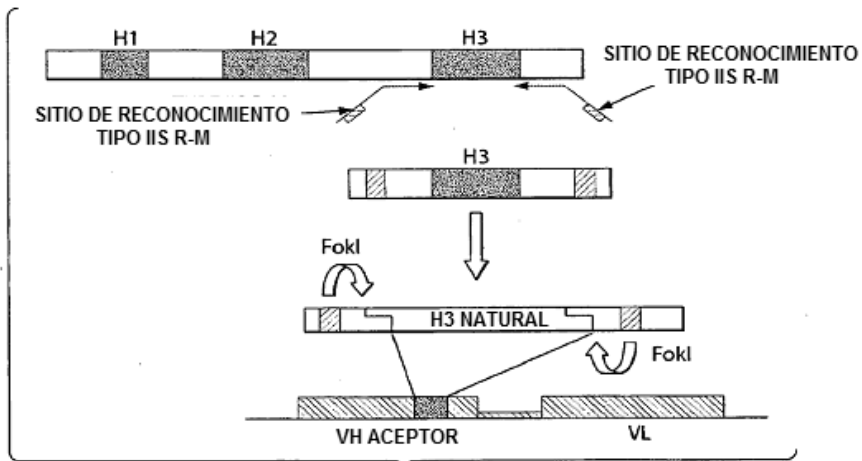


Fig. 3

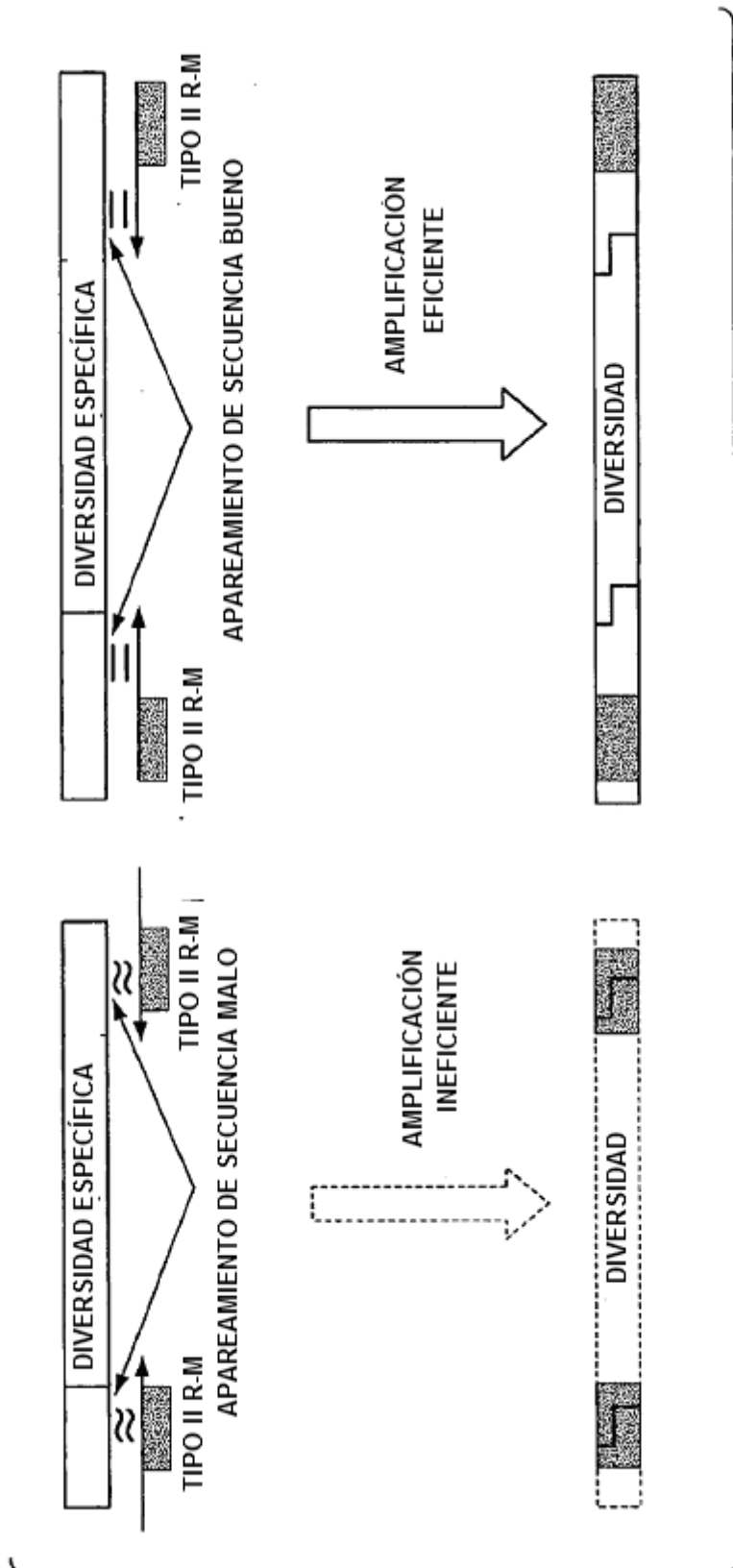


Fig. 4

VH1-2 QVQLVQSGAEVKKPGASVKVSCKASGYTFTGYMHVVRQAPCGCLEMWHINPNSGGTNYAOKFQGRVNTTRDTSISTAYMELSRRLSRDSDTAVYYCAR SEQ ID NO: 1
 VH1-18 QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGISVVRQAPCGCLEMWHISAYNGNTNYAOKLQGRVMTTDTSTAYMELSRRLSRDSDTAVYYCAR SEQ ID NO: 2
 VH1-69 QVQLVQSGAEVKKPGSSVKVSCKASGCTFSSYALISVVRQAPCGCLEMWHGIIPIFGTANYAOKFQGRVITITADESTSTAYMELSLRSEDPAVYYCAR SEQ ID NO: 3
 VH3-30 QVQLVESGGGVQPGRLSLRSLCAASGFTFSYAMHWVRQAPCKGLEWVAVTSYDGSNKYVADSVKGRFTISRDNKNTLYLOWNSLRAEDTAVYYCAR SEQ ID NO: 4
 VH3-48 EIVLVESGGGLVQPGGSLRLSCLAAAGFTFSYSMMVVRQAPCKGLEWVSYISSSSSTIYVADSVKGRFTISRDNKNSLYLOWNSLRAEDTAVYYCAR SEQ ID NO: 5
 VH3-23 EIVLVESGGGLVQPGGSLRLSCLAAAGFTFSYAMS~~VVRQAPCKGLEWVAISGGGGTYYADSVKGRFTISRDNKNTLYLOWNSLRAEDTAVYYCAK~~ SEQ ID NO: 6
 VH5-51 EIVLVQSGAEVKKPGESLKISCKGGYSFTSYWIGVVRQMPCKGLEWNGIIVPGDSDTRYSPFQGVVTSADKSI~~STAYLOWNSL~~KASDTAMYYCAR SEQ ID NO: 7
 VK1-33 DIQMTQSPSSLSASVGDRTVITCQASQDIS NYLNMVYQOKPKAPKLLIYDASNLETGVP~~SRFSGSGGTDFTFI~~SSLQPEDIAITYCQYDNLPPPTV SEQ ID NO: 8
 VK1-39 DIQMTQSPSSLSASVGDRTVITCRASQIS SYLNMVYQOKPKAPKLLIYAASSLQSGVPSR~~FSGSGGTDFTFI~~SSLQPEDFATYYCQOYSTPPTV SEQ ID NO: 9
 VK3-11 EIVLTQSPATLSLSPGERATLSCRASQSVS SYLAWYQOKPKGQAPRLLIYDASN~~RATGIPARFSGSGGTDFTFI~~SSLPEPEFAVYCCQQRSNWPPTV SEQ ID NO: 10
 VK3-15 EIVMTQSPATLSVSPGERATLSCRASQSVS SMLAWYQOKPKGQAPRLLIYGA~~STRATGIPARFSGSGGTEFTFI~~SSLQSEDFAVYCCQYNNWPPTV SEQ ID NO: 11
 VK3-20 EIVLTQSPGTLSPGERATLSCRASQSVSSYLAWYQOKPKGQAPRLLIYGA~~SSRATGIPDRFSGSGGTDFTFI~~SRLEPEFAVYCCQYGSPPPTV SEQ ID NO: 12
 VL1-44 QSVLTQPPSASGTFQQRVITISCGSSSNIGSNVWVYQQLPGTAPKLLIYNNQRP~~SGVDRFSGSKSGTSASLAISGLQSEDEADYCA~~AWDDSLANG SEQ ID NO: 13
 VL1-51 QSVLTQPPSVSAAPGQKVTISCGSSSNIGNNYVSWYQQLPGTAPKLLIYEN~~NKRP~~SGIPDRFSGSKSGTSATLGLTGLQGTGDEADYCGTWDSSLSA SEQ ID NO: 14

Fig. 5

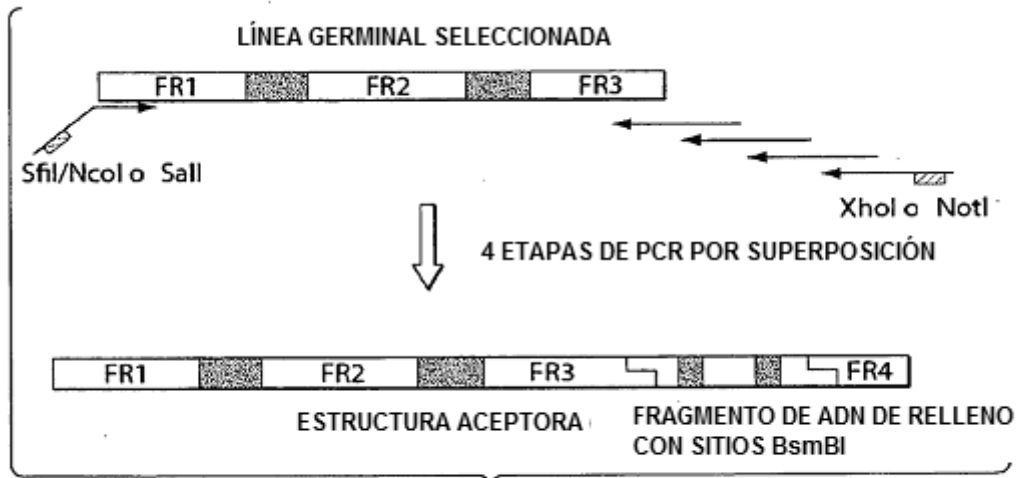


Fig. 6

550 | 2560 | 2570 | 2580 | 2590 | 2600 | 2610 | 2620 | 2630 | 2640 | 2650
 TATTACTGTCCGAGATGAGACCGAATAACCGGTTTACCAGGTTTAAACCGCTATTGGGAGGCCCGTCTCTTTGGGCGCAGGGAACCTGGTCACAGTCTCG SEQ ID NO: 15
 V L L C E N R R I T V R R F T R F K R V L G R R V S W G Q G T L V T V S SEQ ID NO: 16
 Y Y C A R . D E . R . G G L P G L N A Y W E G A S L G A R E P W S Q S R SEQ ID NO: 17
 I T V R D E T N N G K A V Y Q V . T R I G K A R L L G P G N P G H S L SEQ ID NO: 18

5' ...CGTCTC|NNNNNNNN...3' SEQ ID NO: 19
 3' ...GCAGAG|NNNNNNNN...5' SEQ ID NO: 20

Fig. 7

VH1-2 AF

caggtagc tggtagc tgggetgag gtgaagaagc ctggggcctc agtgaaggtc tcctgcaagg cttctggata caccttcacc ggctaactata tgcactgggt gcgacaggcc cctggacaag
 p g a a g a v w g - g e e a w g l s e g l l q g f w i h l h r l l y a l g a t g p w t
 q v g l v q s g a e v k k p g a s v k v s c k a s g y t f t g y y m h w v r q a p g q
 p r c s w c s l g l r - r s l g p p q - r s p a r l l d t p s p a t i c t g c d r p l d k
 >>.....VH1-2.....>>

g v q l v q s g a e v k k p g a s v k v s c k a s g y t f t g y y m h w v r q a p g q
 ggcttagtg gatggatgg atcaacccta acagtgtgg cacaaactat gcacagaagt ttcagggcag ggtcaccatg accagggaac cgtccatcag cacagcctac atggagctga gcaggctgag
 r a - v d g m d q p - q w w h k l c t e v s g q g h h d q g h v h q h s l h g a e q a e
 g l e w m g w i n p n s g g t n y a q k f q g r v t m t r d t s i s t a y m e l s r l
 g l s g w d g s t l t v v a q t m h r s f i r a g s p - p g t r p s a q p t w s - a g -
 >.....VH1-2.....>

g l e w m g w i n p n s g g t n y a q k f q g r v t m t r d t s i s t a y m e l s r l
 atctgacgac acggccgtgt attactgtgc gagatgagac gaataacgtt aaggggttt accaggttta aacgctatt gggaagggc gtctcttggg gccagggaac cctcgtcaca gtctcg SEQ ID NO: 21
 i - r h g r v l l c e m r r i t v r r f t r f k r v l g r r v s w g q g t l v t v s SEQ ID NO: 22
 r s d d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 23
 d l t t r p c i t v r d e t n n g k a v y q v - t r i g k a r l l g p g n p g h s l SEQ ID NO: 24
 >.....VH1-2.....>>

r s d d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 23

Fig. 8

VH1-18 AF

caggtcagc tggtagcagc tgaagctgag gtaagaagc ctggggccctc agtgaagtc tectgaagg ctcttggtta cacccttacc agctatggtta tcagctgggt gcgacaggcc cctggacaag
p g a a g a v w s - g e e a w g l s e g l l q g f w l h l y g l w y g l g a t g p w t
q v q l v q s g a e v k k p g a s v k v s c k a s g y t f t s y g i s w v r q a p g q
p r c s w c s l e l r - r s l g p q - r s p a r l l v t p l p a m v s a g c d r p l d k
>>.....VH1-18.....>
q v q l v q s g a e v k k p g a s v k v s c k a s g y t f t s y g i s w v r q a p g q

ggcttgatg gatggatgg atcagcgctt acaatgtaa cacaaactat gcacagaagc tccagggcag agtcaccatg accacagaca catccacgag cacagcctac atggagctga ggagcctgag
r a - v d g m d q r l q w - h k l c t e a p g g s h h d h r h i h e h s l h g a e e p e
g l e w m g w i s a l y n g n t n y a g k l q g r v t m t t d t s t s t a y m e l r s l
>.....VH1-18.....>
g l e w m g w i s a y n g n t n y a g k l q g r v t m t t d t s t s t a y m e l r s l

atctgaagc acggcctgtt attactgtc gagatgagc gaataacggt aagcggttt accaggtta aacggtatt ggaaggggc gtctctggg gccagggac cctggtcaca gtctcg SEQ ID NO: 25
i - r h g r v l l c e m r r i t v r r f t r f k r v l g r r v s w g q g t l v t v s SEQ ID NO: 26
r s d d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 27
d l t t r p c i t v r d e t n n g k a v y q v - t r i g k a r l l g p g n p g h s l SEQ ID NO: 28
>>.....VH1-18.....>
r s d d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 27

Fig. 8 Cont.

VH1-69 AF

```

caggtcagc tgggtcagtc tgggctgag gtagaagc ctgggccc ggtgaagtc tctgcaagg cttctggagg caccttcagc agctatgcta tcagctgggt gcgacagcc cctggacaag
p g a a a v w g - g e e a w v l g e g l l q g f w r h l q g l c y g l g a t g p w t
q v q l v q s g a e v k k p g s s v k v s c k a s g g t f s s y a i s w v r q a p g q
p r c s w c s l g l r - r s l g p r - r s p a r l l e a p s a a m l s a g c d r p l d k
>>.....VH1-69.....>
q v q l v q s g a e v k k p g s s v k v s c k a s g g t f s s y a i s w v r q a p g q

```

```

ggcttgatg gatggagg atcatccta tctttgtac agcaactac geacagaat tccagggcag agtcacgatt accggagc atccacag cacagcctac atggagctga gcagcctgag
r a - v d g r d h p y l w y s k l r t e v p g q s h d y r g r i h e h s l h g a e g p e
g l e w m g g i i p i f g t a n y a g k f g g r v t i t a d e s t s t a y m e l s s l
>.....VH1-69.....>
g l e w m g g i i p i f g t a n y a g k f g g r v t i t a d e s t s t a y m e l s s l

```

```

atctgaggac acggcctgt attactgtc gagatgagac gaataacggt aagcggttt accagttta aacgctatt ggggaaggc gtctcttgg gccagggac cctggtcaca gtctctg SEQ ID NO: 29
i - g h g r v l l c e m r i i t v r i f t r i f k r v l g r i v s w g q g t l v t v s SEQ ID NO: 30
r s e d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 31
d l r t r p c i t v r d e t n n g k a v y g v - t r i g k a r l l g p g n p g h s l SEQ ID NO: 32
>.....VH1-69.....>
r s e d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 31

```

Fig. 8 Cont.

VH3-23R AF

gaggtgcagc tgttgagtc tggggagc ttgtacagc ctgggggtc cctgagctc tccctgagc cctctggatt cacctttagc agctatgcca tgagctgggt ccgccagggt ccagggaagg
 rgaaavgvwgrlgtaawgvpetllcslwihl-qlichelgppgpre
 evqlllesggglvqpggslrlscaaaagftfsyaamswvrrqapgk
 prcsclsawyslggpldspvqpld splaaamp-ag sarlqgr
 >>.....VH3-23>
 evqlllesggglvqpggslrlscaaaagftfsyaamswvrrqapgk

ggttgagtg ggtctcagct attagtgtgta gtgtgtgtgtag cacatactac gcagactccg tgaaggccg gttcaaccatc tccagagaca attccaagaa cacgctgtat ctgcaaatga acagcctgag
 gaggv g l s y - w - w l l r l r e g p v h h l q r g f q e h a v s a n e q p e
 g l e w v s a i s g s g g s t y y a d s v k g r f t i s r d n s k n t l y l q m n s l
 g w s g s q l l v v v v v a h t t q t p - r a g s p s p e t i p r t r c i c k - t a -
 >.....VH3-23>
 g l e w v s a i s g s g g s t y y a d s v k g r f t i s r d n s k n t l y l q m n s l

agccgagc acggcgtat attactgtc gagatgagac gaataacggt aaggcgttt accagttta aacgctatt gggaaggccg gtctcttggg gccagggaac cctgttcaaca gtctcg SEQ ID NO: 33
 s r g h g r i l l c e m r r i t v r r f t r f k r v l g r r v s w g g g t l v t v s SEQ ID NO: 34
 r a e d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 35
 e p r t r p y i t v r d e t n n g k a v y q v - t r i g k a r l l g p g n p g h s l SEQ ID NO: 36
 >.....VH3-23>
 r a e d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 35

Fig. 8 Cont.

VH3-23K AF

```

gagtgagc tgttgagtc tggggagcc ttggtacagc ctggggggtc ctggaactc tccctgtcag cctctggatt cacctttagc agctatgcca tgagctgggt ccgccagcct ccagggaaagg
R g a a v g v w g r l g t a w g v p e t l i c s l w i h l i - q l c h e l g p p g s r e
e v q l l e s g g g l v q p g g s l r l s c a a s g f t f s s y a m s w v r g a p g k
p r c s c w s l g e a w y s l g g p - d s p v q p l d s p l a a m p - a g s a r l q g r
>>.....VH3-23.....>
e v q l l e s g g g l v q p g g s l r l s c a a s g f t f s s y a m s w v r g a p g k
ggctggagtg ggtccagct attagtggta gtgggtgtag cacatactac gcagactcog tgaagggcgg gttcaccatc tccagagaca attccaagaa caagctgtat ctgcaaatga acagctgag
g a g v g l s y - w - w w - h i l r r l r e g p v h h l q r q f q e h a v s a n e q p e
g l e w v s a i s g s g g s t y y a d s v k g r f t i s r d n s k n t l y l q m n s l
g w s g s q l l v v v v v a h t t q t p - r a g s p s p e t i p r t r c i c k - t a -
>.....VH3-23.....>
g l e w v s a i s g s g g s t y y a d s v k g r f t i s r d n s k n t l y l q m n s l
agccaggac acggccgtat attactgtc gaaatgagac gaataacggt aagggcgttt accaggttta aacgcgtatt gggaagggc gtctcttggg gccagggaac cctgggcaca gtctcg SEQ ID NO: 37
s r g h g r i l l c e m r r i t v r r f t r f k r v l g r r v s w g g g t l v t v s SEQ ID NO: 38
r a e d t a v y y c a k - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 39
e p r t r p y i t v r n e t n n g k a v y g v - t r i g k a r l l g p g n p g h s l SEQ ID NO: 40
>.....VH3-23.....>
r a e d t a v y y c a k - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 39

```

Fig. 8 Cont.

VH3-30 AF

caggtgcagc tgggggagtc tgggggagcc gtggtccacc ctgggagtc cctgagactc tccctgtcag cctctggatt caccttcagt agctatgcta tgcactgggt ccgccaggct ccaggcaagg
 p g a a g g v w g r r g p a w e v p e t l l c s l w i h l g - l c y a l g p p g s r q
 q v q l v e s g g g v v q p g r s l r l s c a a s g f t f s s y a m h w v r q a p g k
 p r c s w w s l g e a w s s l g g p - d s p v q p l d s p s v a m l c t g s a r l q a r
 >>.....VH3-30.....>
 q v q l v e s g g g v v q p g r s l r l s c a a s g f t f s s y a m h w v r q a p g k

ggctggagtg ggtggcagtt atatcatatg atggaagtaa taatactac gcagactcgg tgaaggccg attaccactc tccagagaca attccaagaa cacgctgtat ctgcaaatga acagcctgag
 g a g v g g s y i i - w k - i l r r l r e g p i h h l q r g f q e h a v s a n e q p e
 g l e w v a v i s y d g s n k y y a d s v k g r f t i s r d n s k n t l y l q m n s l
 g w s g w q l y h m m e v i n t t q t p - r a d s p s p e t i p r t r c i c k - t a -
 >.....VH3-30.....>
 g l e w v a v i s y d g s n k y y a d s v k g r f t i s r d n s k n t l y l q m n s l

agctgaggac acggctgtgt attactgtgc gagatgagac gaataacgggt aaggcggttt accaggttta aacgcgtatt gggaaggcgc gtctcttggg gccagggaac cctgggtcaca gtctcgc SEQ ID NO: 41
 s - g h g c v l l c e m r r i t v r r f t r f k r v l g r r v s w g q g t l v t v s SEQ ID NO: 42
 r a e d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 43
 e l r t r l c i t v r d e t n n g k a v y q v - t r i g k a r l l g p g n p g h s l SEQ ID NO: 44
 >.....VH3-30.....>
 r a e d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 43

Fig. 8 Cont.

VH3-48 AF

gggtgcagc tgggggagc ttggtacagc ctgggggtc cctgagctc tcctgtcag cctctgatt cacctcagt agctatagca tgaactggt ccgccaggt ccagggaggg
 r g a a g g v w g r l g t a w g v p e t l l c s l w i h l q - l - h e l g p p g s r e
 e v q l v e s g g g l v q p g g s l r l s c a a s g f t f s s y s m n w v r q a p g k
 p r c s w w s l g e a w y s l g g p - d s p v q p l d s p s v a i a - t g s a r l g g r
 >>.....VH3-48.....>
 e v q l v e s g g g l v q p g g s l r l s c a a s g f t f s s y s m n w v r q a p g k

ggctggagtg ggttccatc attagtagta gtagtagtac catatactac gcagactctg tgaaggccg attcaccatc tccagagaca atgccaaaga ctcactgtat ctgcaaatga acagcctgag
 g a g v g f i h - - - y h i l r i l c e g p i h h l g r q c g e l t v s a n e q p e
 g l e w v s y i s s s s s t i y y a d s v k g r f t i s r d n a k n s l y l q m n s l
 g w s g f h t l v v v v v p y t t q t l - r a d s p s p e t m p r t h c i c k - t a - -
 >.....VH3-48.....>
 g l e w v s y i s s s s t i y y a d s v k g r f t i s r d n a k n s l y l q m n s l

agccgagac acggtgtgt attactgtc gagatgagac gaataacgt aaggggttt accagttta aacgctatt gggaagcgc gtctcttggg gccaggaac cctggtcaca gtctcy SEQ ID NO: 45
 s r g h g c v l l c e m r i t v r r f t r f k r v l g r r v s w g q g t l v t v s SEQ ID NO: 46
 r a e d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 47
 e p r t r l c i t v r d e t n n g k a v y q v - t r i g k a r l l g p g n p g h s l SEQ ID NO: 48
 >.....VH3-48.....>>
 r a e d t a v y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 47

Fig. 8 Cont.

VH5-51 AF

```

gaggtgcagc tgggtcagtc tggagcagag gtgaaaaagc ccggggagtc tetgaagatc tctqtaagg gttctggata cagctttacc agctactgga tcggctgggt ggcaccagatg cccgggaaag
r g a a g a v w s r g e k a r g v s e d l l l - g f w i q l l y q l l d r l g a p d a r e
e v q l v q s g a e v k k p g e s l k i s c k g s g y s f t s y w i g w v r q m p g k
p r c s w c s l e q r - k s p g s l - r s p v r v l d t a l p a t g s a g c a r c p g g k
>>.....VH5-51.....>
e v q l v q s g a e v k k p g e s l k i s c k g s g y s f t s y w i g w v r q m p g k

```

```

gcctggagtg gatgggagtc atctatctctg gtgactctga taccagatc agcccgctct tccaagcca ggtcaccatc tcagccgaca agtccatcag caecgcctac ctgcagtga gcagcctgaa
r p g v d g d h l s w - l - y q i q p v l p r p q h h l s r q v h q h r l p a v e q p e
g l e w m g i i y p g d s d t r y s p s f q g q v t i s a d k s i s t a y l q w s s l
a w s g w g s s i l v t l i p d t a r p s k a r s p s g p t s p s a p p t c s g a a -
>.....VH5-51.....>
g l e w m g i i y p g d s d t r y s p s f q g q v t i s a d k s i s t a y l q w s s l

```

```

ggcctcgac accgccatgt attactgtc gagatgagac gaataacggt.aagggcgttt accagttta aacgcgtatt ggaagggcgc gtctcttggg gccagggaac cctggtcaca gtctcg SEQ ID NO: 49
g l g h r h v l l c e m r r i t v r r f t r f k r v l g r r v s w g q g t l v t v s SEQ ID NO: 50
k a s d t a m y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 51
r p r t p p c i t v r d e t n n g k a v y q v - t r i g k a r l l g p g n p g h s l SEQ ID NO: 52
>.....VH5-51.....>>
k a s d t a m y y c a r - d e - r - g g l p g l n a y w e g a s l g a r e p w s q s SEQ ID NO: 51

```

Fig. 8 Cont.

VK1-33G AF

```

gacatccaga taccaccagtc tccatccctcc ctgtctgcat ctgtaggaga cagagtcacc atcacttqcc aggcagtgca ggacattagc aactatttaa attggtatca gcagaaacca gggaaagccc
ghp d d p v s i l p v c i c r r q s h h l p g e s g h - q l f k l v s a e t r e s
d i q m t q s p s s l s a s v g d r v t i t c q a s g d i s n y l n w y q q k p g k a
r t s r - p s l h p p c l h l - e t e s p s l a r i v r t l a t i - i g i s r n q g k p
>>.....VK 1-33G.....>
d i q m t q s p s s l s a s v g d r v t i t c q a s g d i s n y l n w y q q k p g k a

ctaagctcct gatctacgat gcatccaatt tggaaacagg ggtcccatac aggttcaqgt gaagtggtatc tgggacagat ttactttca ccatcagcag cctgcagcct gaagatattg caacatatta
p - a p d l r c i q f g n r g p i k v g w k w i w d r f y f h h g g p a a - r y c n i l
p k l l i y d a s n l e t g v p s r f s g s g t d f t f t i s s l q p e d i a t y
l s s - s t m h p i w k q g s h q g s v e v d l g q i l l s p s a a c s l k i l q h i
>>.....VK 1-33G.....>
p k l l i y d a s n l e t g v p s r f s g s g t d f t f t i s s l q p e d i a t y

ctgtcagcag cyagacgaat aacggttaagg cgtttacca ggtttaaacg cgtattggga aggcggtct cattoggcca agggaccaag gtggaaatea aagggcgcc cgca SEQ ID NO: 53
l s a a r r i t v r r f t r f k r v l g r r v s f g q g t k v e i k g a a a SEQ ID NO: 54
y c g q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s k g r p SEQ ID NO: 55
t v s s e t n n g k a v y q v - t r i g k a r l i r p r d q g g n q r g g r
>>.....VK 1-33G.....>
y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s k g r p SEQ ID NO: 56

```

Fig. 8 Cont.

VH1-33R AF

gacatccaga tgaccacagtc tccatctctcc ctgtctgcac ctgttaggaga cagagtcacc atcaattgcc aggcagatca ggacattagc aactatttaa attggtatca gcagaaacca gggaaagccc
 ghpd d d p v s i l p v c i c r r q s h h h l p g e s g h - g l f k l v s a e t r e s
 d i q m t q s p s s l s a s v g d r v t i t c g a s g d i s n y l n w y q q k p g k a
 r t s r - p s l h p p c l h l - e t e s p s l a r r v r t l a t i - i g i s r n q g k p
 >>.....VK 1-33R.....>
 d i q m t q s p s s l s a s v g d r v t i t c g a s g d i s n y l n w y q q k p g k a

ctaagctct gattacgat gcaccaatt tggaaacagg ggtccatca aggttcaqtg gaagtgatc tggacagat tttactttca ccatcagcag cctgcagcct gaagatattg caacatatta
 p - a p d l r c i q f g n r g p i k v q w k w i w d r f y f h h q q p a a - r y c n i l
 p k l l i y d a s n l e t g v p s r f s g s g t d f t f t i s s l q p e d i a t y
 l s s - s t m h p i w k q g s h q g s v e v d l g q i l l s p s a a c s l k i l q h l
 >>.....VK 1-33R.....>
 p k l l i y d a s n l e t g v p s r f s g s g t d f t f t i s s l q p e d i a t y

ctgtcagcag cgagacgaat aacggttaagg cggtttacca ggtttaaagc cgtattgga aggcgctct cattoggeca agggaccaag gtggaatca aacggcgccg cgca SEQ ID NO: 57
 l s a a r r i t v r r f t r f k r v l g r r v s f g q g t k v e i k r a a a SEQ ID NO: 58
 y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s n g r p SEQ ID NO: 59
 t v s s e t n n g k a v y q v - t r i g k a r l i r p r d q g g n q t g g r
 >.....VK 1-33R.....>>>
 y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s n g r p SEQ ID NO: 59

Fig. 8 Cont.

VK1-39G AF

gacatccaga tgaccacagtc tccatccctcc ctgtctgcat ctgtaggaga cagagtcacc atcaactgcc gggeaagtca gagcattagc agctatctaa attggtatca gcagaaatca gggaaagccc
 ghpd d p v s i l p v c i c r r q s h h h l p g k s e h - q l f k l v s a e t i e s
 d i q m t q s p s s l s a s v g d r v t i t c r a s q s i s s y l n w y q g k p g k a
 r t s r - p s l h p p c l h l - e t e s p s l a g q v r a l a a i - i g i s r n q g k p
 >>.....VK1-39G.....>

d i q m t q s p s s l s a s v g d r v t i t c r a s q s i s s y l n w y q g k p g k a
 ctaagetctt gatctatgct gcatccagtt tgcaaaagtg ggtcccatca aggttcaagt ggagtgatc tgggacagat tteactctca ccatcagcag tctgcaacct gaagattttg caacttacta
 p - a p d l c c i q f a k w g p i k v q w q w i w d r f h s h h q q s a t - r f c n l l
 p k l l i y a a s s l q s g v p s r f s g s g t d f t l t i s s l q p e d f a t y
 l s s - s m l h p v c k v g s h q g s v a v d l g q i s l s p s a v c h l l k i l q l t
 >.....VK1-39G.....>

p k l l i y a a s s l q s g v p s r f s g s g t d f t l t i s s l q p e d f a t y
 ctgtcagcag cgagacgaat aacggttaagg cggttacca ggtttaaag cgtattgga aggcgctct cattcgcca agggaccaag gtggaatca aagggcggc cgca SEQ ID NO: 61
 l s a a r i i t v i r f t r f k r v l g r v s f g q g t k v e i k g a a a SEQ ID NO: 62
 y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s k g r p SEQ ID NO: 63
 t v s s e t n n g k a v y q v - t r i g k a r l i r p r d q g g n q r g g r
 >.....VK1-39G.....>
 y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s k g r p SEQ ID NO: 64

Fig. 8 Cont.

VH1-39R AF

gacatccaga tgaccocagtc tccatctctc ctgtctgcat ctgttaggaga cagagtcaccc atccattccc gggcaatca gagcattagc agctatttaa attggtatca gcagaaacca gggaaagccc
 ghpdppv silpvc icrir qshhhllpgkseh-qllfkllvsaetres
 diqmtgspsslsasvgsdirtvtitcraasqsis syl nwy qqkppgka
 >>.....VK 1-39R.....>>
 diqmtgspsslsasvgsdirtvtitcraasqsis syl nwy qqkppgka

ctaagctct gatctatgct geatccagtt tgcaagtgg ggtcccatca aggttcagtg gcagtggatc tgggacagat ttcactctca ccatcagcag tctgcaacct gaagattttg caacttacta
 p-apidlcicqfakwgpikvqwgqwiwdrfhs h h q q s a t - r f c n l l l
 plliyaasslqs gvpss rfs gsgsgtd f t l t i s s l q p e d f a t y
 l s s - s m l h p v c k v g s h q g s v a v d l g q i s l s p s a v c n l k i l q l l t
 >.....VK 1-39R.....>
 plliyaasslqs gvpss rfs gsgsgtd f t l t i s s l q p e d f a t y

ctgtcagcag cgagacgaat aacggtaagg cggtttaacca ggtttaaacg cgtattggga agggcgctct cattcgcca agggaccaag gtggaaatca aacggggggc cgca SEQ ID NO: 65
 lsarritvritftrfk rvlgrvrvsfgqgqtkveikraaa SEQ ID NO: 66
 ycqqqrde-rr-ggllpqlnawyegas h s a k g p r w k s n g r p SEQ ID NO: 67
 tvssetnngkavyqv-t r i g k a r l i r p r d q g g n g t g g r
 >.....VK 1-39R.....>>
 ycqqqrde-rr-ggllpqlnawyegas h s a k g p r w k s n g r p SEQ ID NO: 67

Fig. 8 Cont.

VK3-11G AF

```

gaaattgtgt tgacacagtc tccagccacc ctgtctttgt ctccaggga aagaccacc ctctcttqca gggccagta gagtgtagc agtacttag cctggtacca acagaaacct ggcaggtc
gncvdtv ssh p v f v s r g k s h p l l q g q s e c - g l l s l v p t e t w p g
e i v l t q s p a t l s l s p g e r a t l s c r a s q s v s s y l a w y q q k p g q a
r k l c - h s l q p p c l c l q g k e p p s p a g p v r v l a a t - p g t n r n l a r l
>>.....VK 3-11G.....>
e i v l t q s p a t l s l s p g e r a t l s c r a s q s v s s y l a w y q q k p g q a

ccaggctcct catctatgat gcatcaaca gggccactgg catcccagcc aggttcagtg gcagtggtc tggcagac ttcacttca ccatcagcag cctagagcct gaagattttg cagtttatta
s g a p h l - c i q q g g h w h p s q v q w q w v w d r l h s h h q q p r a - r f c s l l
p r l l i y d a s n r a t g i p a r f s g s g s g t d f t l t i s s l e p e d f a v y
p g s s s m m h p t g p l a s q p g s v a v g l g q t s l s p s a a - s l k l l q f i
>.....VK 3-11G.....>
p r l l i y d a s n r a t g i p a r f s g s g s g t d f t l t i s s l e p e d f a v y

ctgtcagcag cgagacgaat aacggttaagg cggtttacca ggtttaaacg cgtattgga aggcgctct cattcgcca agggaccaag gtggaataca aagggcggc cgca SEQ ID NO: 69
l s a a r r i t v r r f t r f k r v l g r r v s f g q g t k v e i k g a a a SEQ ID NO: 70
y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s k g r p SEQ ID NO: 71
t v s s e t n n g k a v y q v - t r i g k a r l i r p r d q g g n q r g g r
>.....VK 3-11G.....>
y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s k g r p SEQ ID NO: 72

```

Fig. 8 Cont.

VH3-11R AF

```

gaaattgtg tgcacagtc tccagccacc ctgtctttgt ctccaggga aagagccacc ctctctgca gggcagtc gagtgtag agctacttag cctggtacca acagaaacct ggcaggctc
gncvdvtv sshp vfvsvrg kshpll qgqs ec - qlslsvptetwppg
eivltqspat lslspg erat lsc ras qsvs syl awy qqkppgga
rklc - hslqpplc lqgk eppspagpvr vlvlaat - pgt n r n l a r l
>>.....VK 3-11R.....>
eivltqspat lslspg erat lsc ras qsvs syl awy qqkppgga
ccaggctct catctatgat gcatccaaca gggccactgg catcccagc aggttcagtg gagtggttc tggcacaagc ttaactca ccaatcagcag cctagagcct gaagattttg cagtttatta
sqaphl - ciqqqhwhpsqvqvwvwdrlhshhqqqpra - rfcslll
prlliydasn rat giparfs gsgsgtdftltis slepe dfavy
p g s s s m h p t g p l a s q p g s v a v q l g q t s l s p s a a - s l k i l q f i
>>.....VK 3-11R.....>
prlliydasn rat giparfs gsgsgtdftltis slepe dfavy
ctgtcagcag cgagacgaat aacggttaagg cggtttaacca ggitttaaacg cgtattggga aggcgctct caticggcca agggacaag gtggaaatca aacggggcgc cgca SEQ ID NO: 73
lsaarritvritrft rfk rvlgrivrsvfg qgtkveikraaa SEQ ID NO: 74
ycqqqrde - r - gglp gl n a y w e g a s h s a k g p r w k s n g r r p SEQ ID NO: 75
tvs setn ngk avy gv - t r i g k a r l i r p r d q g g n q t g g r
>>.....VK 3-11R.....>>
ycqqqrde - r - gglp gl n a y w e g a s h s a k g p r w k s n g r r p SEQ ID NO: 76

```

Fig. 8 Cont.

VK3-15G AF

```

gaaatagtga tgacgcagtc tccagccacc ctgtctgtgt ctccaggaga aagagccacc ctctcttga gggccagtca gagtgtagc agcaacttag cctggtacca gcagaaacct ggcaggtctc
gns d d a v s s h p v c v s r g k s h p l l q g q s e c - g q l s l v p a e t w p g
e i v m t q s p a t l s v s p g e r a t l s c r a s q s v s s n l a w y q q k p g q a
r k - - - r s l q p p c l c l q g k e p p s p a g p v r v l a a t - p g t s r n l a r l
>>.....VK 3-15G.....>
e i v m t q s p a t l s v s p g e r a t l s c r a s q s v s s n l a w y q q k p g q a

ccaggctcct catctatgtt gcattcacc gggccactgg tatccagcc aggttcagtg gcagtggttc tggcacagag ttcactctca ccatcagcag cctgcagtct gaagattttg cagtttatta
s g a p h l w c i h q g h w y p s q v q w q w v w d r v h s h h q q p a v - r f c s l l
p r l l i y g a s t r a t g i p a r f s g s g s g t e f t l t i s s l q s e d f a v y
p g s s s m v h p p g p l v s q p g s v a v g l g q s s l s p s a a c s l k i l q f i
>.....VK 3-15G.....>
p r l l i y g a s t r a t g i p a r f s g s g s g t e f t l t i s s l q s e d f a v y

ctgtcagcag cgagacgaat aacggtaagg cggtttacc ggtttaaacg cgtattgga aggcggtct cattcggcca agggaccaag gtggaatca aagggggcgc cgca SEQ ID NO: 77
l s a a r i t v r f t r f k r v l g r r v s f g q g t k v e i k g a a a SEQ ID NO: 78
y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s k g r p SEQ ID NO: 79
t v s s e t n n g k a v y q v - t r i g k a r l i r p r d q g g n q r g g r SEQ ID NO: 80
>.....VK 3-15G.....>
y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s k g r p SEQ ID NO: 79

```

Fig. 8 Cont.

VH3-15R AF

```

gaatatgta tgaagcagtc tcaagcacc ctgtctgtgt ctccagggga aagagcacc ctctctcca gggcagtcga gactgttagc agcaacttag cctggtacca gcagaaacct ggcacggctc
gnsd'avsshpvcvsrhkshpvlqgqsesc-qqlslvpaetwpg
eivmtgspatlsvsvpgeratlsclscrasqsvssnlawyqqkppgqa
rk--rslqppclclqlqgkepsspa gppvrvl aat-pgtsrnlarl
>>.....VK 3-15R.....>
eivmtgspatlsvsvpgeratlsclscrasqsvssnlawyqqkppgqa

```

```

ccaggctct catctatggt gcaccaaca gggccactgg tatcccagc aggttcaagt ggaagggtc tgggacagag ttcaictca ccatacagcag cctgcagctc gaagattttg cagtttatta
sqaphlwcihqghwypsqvqvwqwwdrrvhs h h q q p a v - r f c s l l
p r l l i y g a s t r a t g i p a r f s g s g t e f t l t i s s l q s e d f a v y i
p g s s m v h p p g p l v s q p g s v a v g l g q s s l s p s a a c s l k i l q f i
>.....VK 3-15R.....>
p r l l i y g a s t r a t g i p a r f s g s g t e f t l t i s s l q s e d f a v y

```

```

ctgtcagcag cgagacgaat aacggttaagg cggtttacc ggtttaaacg cgtattggga aggcgctct caticggcca agggaccaag gtggaaatca aacggggggc cgca SEQ ID NO: 81
lsaarritvrriftrfkrvlgrvsvfgqgqtkveikraa SEQ ID NO: 82
ycqqrde-r-gglp glnaywegas h s a k g p r w k s n g r p SEQ ID NO: 83
tvssetnngkavyqv-triqkarlirp r d q g g n q t g g r
>.....VK 3-15R.....>>
ycqqrde-r-gglp glnaywegas h s a k g p r w k s n g r p SEQ ID NO: 83

```

Fig. 8 Cont.

VK3-20G AF

gaaattgtg tgaacgagtc tccaggaacc ctgtctttgt ctccaggga aagagcacc ctctctca gggccagtca ggtgttagc agcagctact tagcctggtta ccagcagaaa cctggccagg
 gncvda v s r h p v f v s r g k s h p l l q g q s e c - q g l l s l v p a e t w p
 e i v l t q s p g t l s l s p g e r a t l s c r a s q s v s s s y l a w y q q k p g q
 r k l c - r s l q a p c l c l q g k e p p s p a g p v r v l a a a t - p g t s r n l a r
 >>.....VK 3-20G.....>
 e i v l t q s p g t l s l s p g e r a t l s c r a s q s v s s s y l a w y q q k p g q

ctccaggct cctcatctat ggtgatcca gcaggccac tggatcca gacagttca gtggagttgg gtctggaca gacttcactc tcaccatcag cagactggag cctgagatt ttgagttta
 g s q a p h l w c i q q g h w h p r q v q w q w v w d r l h s h h q q t g a - r f c s l
 a p r l l i y g a s s r a t g i p d r f s g s g t d f t l t i s r l e p e d f a v
 l p g s s s m v h p a g p l a s q t g s v a v g l g q t s l s p s a d w s l k i l q f
 >.....VK 3-20G.....>
 a p r l l i y g a s s r a t g i p d r f s g s g s g t d f t l t i s r l e p e d f a v

ttactgtcag cagcgagacg aataacggta agcgggttta ccagggttaa acggtattg ggaaggcgc tctcattcgg ccaaggacc aagtggaat tcaaaagggc ggccgca SEQ ID NO: 85
 l l s a a r r i t v r r f t r f k r v l g r r v s f g q g t k v e i k g a a a SEQ ID NO: 86
 y y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s k g r p SEQ ID NO: 87
 i t v s s e t n n g k a v y q v - t r i g k a r l i r p r d q g g n q r g g r SEQ ID NO: 88
 >.....VK 3-20G.....>
 y y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s k g r p SEQ ID NO: 87

Fig. 8 Cont.

VH3-20R AF

gaaatttgt tgacgcagtc tccaggcacc ctgtctttgt ctccagggga aagagccacc ctctcttqpa gggccagtca gagtgttagc agcagctact tagctctgta ccagcagaaa cctggccagg
 gncvda v s r h p v f v s r g k s h p l l q g q s e c - g q l l s l v p a e t w p
 e i v l t q s p g t l s l s p g e r a t l s c r a s q s v s s s y l a w y q q k p g q
 r k l c - r s l q a p c l c l q g g k e p p s p a g p v r v l a a a t - p g t s i n l a r
 >>.....VK 3-20R.....>
 e i v l t q s p g t l s l s p g e r a t l s c r a s q s v s s s y l a w y q q k p g q

ctccaggct cctcatctat ggtgatcca gcaggccac tggcatocca gacaggttca gtgcagtggt gtctgggaca gaattcactc tcaccatcag cagactggag cctgaagatt ttgcagttta
 g s q a p h l w c i q q g h w h p r g v q w q w v w d r l h s h h g q t g a - r f c s l
 a p r l l i y g a s s r a t g i p d r f s g s g s g t d f t l t i s r l e p e d f a v
 l p g s s s m v h p a g p l a s q t g s v a v g l g q t s l s p s a d w s l k i l q f
 >.....VK 3-20R.....>
 a p r l l i y g a s s r a t g i p d r f s g s g s g t d f t l t i s r l e p e d f a v

ttactgtcag cagcagagc aataacggta aggcggttta ccagggttaa acgcgtattg gaaaggcgg tctcattcgg ccaaggacc aaggtggaaa tcaaacggc ggcgcga	SEQ ID NO: 89
l l s a a r i r i t v r i r f t r f k r v l g r r v s f g q g t k v e i k r a a a	SEQ ID NO: 90
y y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s n g r p	SEQ ID NO: 91
i t v s s e t n n g k a v y q v - t r i g k a r l i r p r d q g g n g t g g r	SEQ ID NO: 92
>.....VK 3-20R.....>	
y y c q q r d e - r - g g l p g l n a y w e g a s h s a k g p r w k s n g r p	SEQ ID NO: 91

Fig. 8 Cont.

VL1-44 AF

```

caqctgtgt tgacgcagcc gcctcagtg tctgggccc caggacagaa ggtaccatc tctctctctg gaagcagctc caacattggg aataattatg tatctctgta ccagcagctc ccaggaacag
avc vdaa als vcg prte ghhl llwkg lqhwe - lcllv paa pirn
qsv l tq ppsv saa pgg kvtkvti scs gss snig nny vsw yqql pgt
rslc - rsrppq clirp qdr rsp spal eaa p t l g i i m y p g t s s s s q e q
>>.....VL1-51.....>
qsv l tq ppsv saa pgg kvtkvti scs gss snig nny vsw yqql pgt

```

```

ccccaaact cctcatttat gacaataata agcaccctc agggatttct gaccattct ctgctccaa gctctggcag tggcacc ccgactccag actggggagc aggcgatta
spqt phl - q - - a t l r d s - p i l w l q v w h v s h p g h h r t p d w g r g r l
a p k l l i y d n n k r p s g i p d r f s g s k s g t s a t l g i t g l q t g d e a d
p p n s s f m t i i s d p q g f l t d s l a p s l a r q p p w a s p d s r l g t r p i
>.....VL1-51.....>
a p k l l i y d n n k r p s g i p d r f s g s k s g t s a t l g i t g l q t g d e a d

```

```

ttactgcgga acaggagacg aatacggta aggcggtta ccaggttaa acgcgtattg ggaaggcgg tctctgttt cggcggagg accaagctga ccgtcttagg ggcggccgca SEQ ID NO: 93
llr nr r i r i t v i r f t r f k r v l g r r v s v f g g g t k l t v l g a a a SEQ ID NO: 94
y y c g t g d e - r - g g l p g l n a y w e g a s l c s a e g p s - p s - g r p SEQ ID NO: 95
i t a e q e t n n g k a v y q v - t r i g k a r l c v r r r d q a d r p r g g r SEQ ID NO: 96
>.....VL1-51.....>
y y c g t g d e - r - g g l p g l n a y w e g a s l c s a e g p s - p s - g r p SEQ ID NO: 95

```

Fig. 8 Cont.

VL1-51 AF

caftctgtgc tgactcagcc accctcagcg tctgggacc ccgggacag ggtcaccatc tctgtttctg gaagcagctc caacatcggg agtaactctg taaacttgta ccagcagctc ccagggaacgg
 avc adsaatlslsvwd p r a e g h h l l f w k q l g h r k - y c k l v p a a p r n
 q s v l t q p p s a s g t p g q r v t i s c s g s s s n i g s n t v n w y q q l p g t
 r s l c - l s h p q r l g p p g r g s p s l v l e a a p t s e v i l - t g t s s s q e r
 >>.....VL1-44.....>
 q s v l t q p p s a s g t p g q r v t i s c s g s s s n i g s n t v n w y q q l p g t

ccccaaact cctcatctat agtaataatc aggggacctc aggggtccct gacgattct ctggtccaa gctcggacc tcagctcc tcggcatcag tggctccag tctgagatg aggtgatta
 g p q t p h l - - - s a a l r g p - p i l w l q v w h l s l p g h q w a p v - g - g - l
 a p k l l i y s n n q r p s g v p d r f s g s k s g t s a s l a i s g l q s e d e a d
 p p n s s s i v i i s g p q g s l t d s l a p s l a p q p p w p s v g s s l r m r l i
 >.....VL1-44.....>
 a p k l l i y s n n q r p s g v p d r f s g s k s g t s a s l a i s g l q s e d e a d

ttactgtca gcaggagacg aataacgta aggcggttta ccaggttta acgctattg ggaagcgcg tctctgttt cggcgaggg accaagctga cagctcagg ggcggcgca	SEQ ID NO: 97
l l c s r r r i t v r r f t r f k r v l g r r v s v f g g g t k l t v l g a a a	SEQ ID NO: 98
Y Y c a a g d e - r - g g l p g l n a y w e g a s l c s a e g p s - p s - g r p	SEQ ID NO: 99
i t v q q e t n n g k a v y q v - t r i g k a r l c v r r r d q a d r p r g g r >.....VL1-44.....>>	SEQ ID NO: 100
Y Y c a a g d e - r - g g l p g l n a y w e g a s l c s a e g p s - p s - g r p	SEQ ID NO: 99

Fig. 8 Cont.

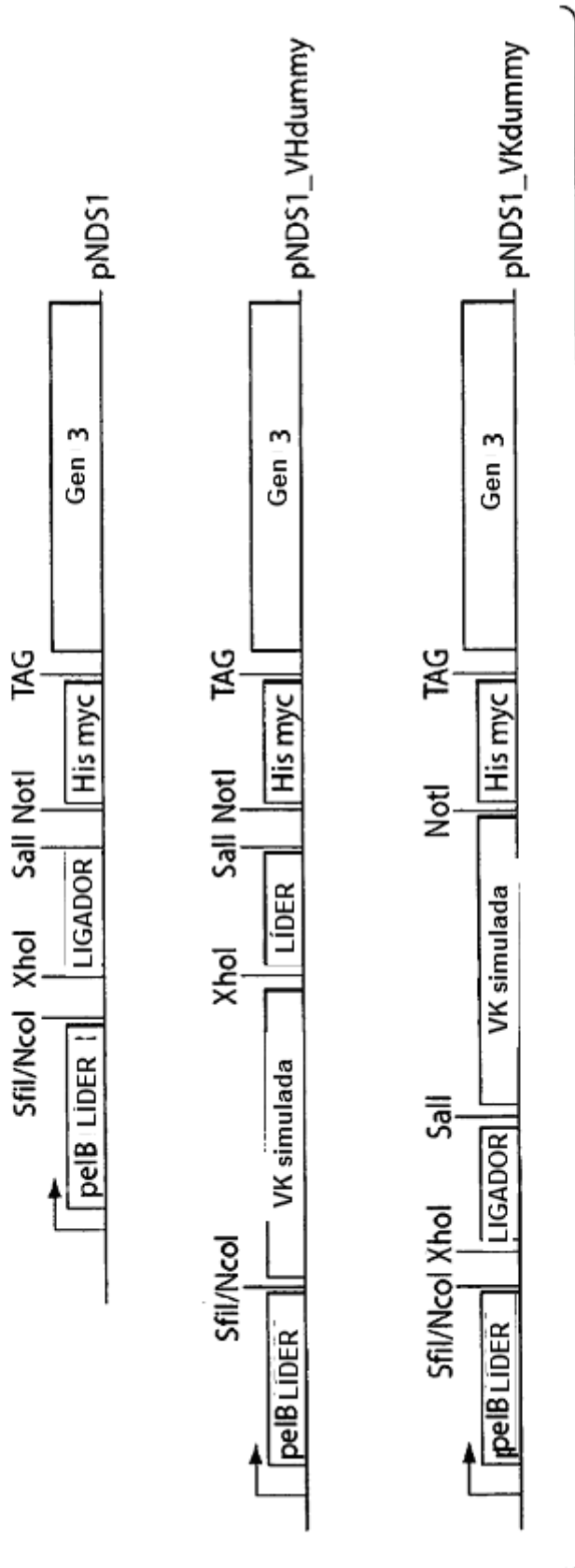


Fig. 9

MUESTRA	FW3	H3	H3/FW4	LONG CDR3	
1a	CAK	AFRPNWGSRVLYF	DYW	15	SEQ ID NO: 101
1b	CAK	LNWGWFF	DPW	8	SEQ ID NO: 102
1c	CAK	GWEGEY	DYW	8	SEQ ID NO: 103
1d	CAK	DLWLDSSSNWF	DPW	13	SEQ ID NO: 104
1e	CAK	DLPG	DPHW	7	SEQ ID NO: 105
1f	CAK	APPTGYF	DYW	9	SEQ ID NO: 106
1g	CAK	VGLTG VHF	DYW	10	SEQ ID NO: 107
1h	CAK	DSYGSGSYYN	DYW	12	SEQ ID NO: 108
2a	CAK	VAGTWGRVAYYF	DYW	14	SEQ ID NO: 109
2b	CAK	VTGVFVGNF	DYW	11	SEQ ID NO: 110
2c	CAK	DARNSGSYF	DYW	11	SEQ ID NO: 111
2d	CAK	GVRTGVVF	DYW	10	SEQ ID NO: 112
2e	CAK	GVVNWGTRRKGWF	DPW	15	SEQ ID NO: 113
2f	CAK	WGGWF	DPW	7	SEQ ID NO: 114
2g	CAK	RRDNWGSV	DW	9	SEQ ID NO: 115
2h	CAK	GSGFSSGWFF	DSW	11	SEQ ID NO: 116

Fig. 10

LONG FRECUENCIA		CDR3										DIVERSIDAD TEÓRICA																																			
WH		32	93	94	95	96	97	98	99	100	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	z	101	102	103	104	105	106	POSICIONES RND	NNS	DVK	NWT	DVT									
7	25	C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	4	1.0E+06	1.0E+05	2.1E+04	6.6E+03
8	4	C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	5	3.4E+07	1.9E+06	2.5E+05	5.9E+04
9	6	C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	6	1.1E+09	3.4E+07	3.0E+06	5.3E+05
10	8	C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	7	3.4E+10	6.1E+08	3.6E+07	4.8E+06
11	9	C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	8	1.1E+12	1.1E+10	4.3E+08	4.3E+07
12	11	C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	9	3.5E+13	2.0E+11	5.2E+09	3.9E+08
13	10	C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	10	1.1E+15	3.6E+12	6.2E+10	3.5E+09
14	9	C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	11	3.6E+16	6.4E+13	7.4E+11	3.1E+10
15	8	C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	12	1.2E+18	1.2E+15	8.9E+12	2.8E+11
16		C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	13	3.7E+19	2.1E+16	1.1E+14	2.5E+12
17		C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	14	1.2E+21	3.7E+17	1.3E+15	2.3E+13
18		C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	15	3.8E+22	6.7E+18	1.5E+16	2.1E+14
19		C	A	R	K	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	F	M	D	Y	W	G	Q	G	T	L	V	T	V	S	S	16	1.2E+24	1.2E+20	1.8E+17	1.9E+15

WL	LONG FRECUENCIA	BUCLE ESTRUCTURAL										DIVERSIDAD TEÓRICA						
		89	90	91	92	93	94	95	96	96a	97	PLEGADO CANÓNICO	POSICIONES ALEATORIAS	NNS	DVK	NWT	DVT	
4		C	Q	X	X	X	X	X	X	X	X	5 F G Q G	4	4				
5	+	C	Q	X	X	X	X	X	X	X	P	- T F G Q G	3	4	1.0E+06	1.0E+05	2.1E+04	6.6E+03
6	+++	C	Q	X	X	X	X	X	X	X	P	- T F G Q G	1	5	3.4E+07	1.9E+06	2.5E+05	5.9E+04
7	+	C	Q	X	X	X	X	X	X	X	P	X T F G Q G	5	6	1.1E+09	3.4E+07	3.0E+06	5.3E+05
6		C	Q	X	X	X	X	X	X	X	P	X - T F G Q G	2	5				

VL	LONG FRECUENCIA	BUCLE ESTRUCTURAL										DIVERSIDAD TEÓRICA						
		89	90	91	92	93	94	95	95a	95b	96	97	PLEGADO CANÓNICO	POSICIONES ALEATORIAS	NNS	DVK	NWT	DVT
VL1-44	6	C	A	W	Y	D	X	X	X	X	X	- X V F G G G	1	4	2.1E+06	2.1E+05	4.1E+04	1.3E+04
	7	C	A	W	Y	D	X	X	X	X	X	- X V F G G G		5	6.7E+07	3.8E+06	5.0E+05	1.2E+05
	8	C	A	W	Y	D	X	X	X	X	X	X X V F G G G	2	6	2.1E+09	6.8E+07	6.0E+06	1.1E+06
VL1-51	6	C	G	T	W	Y	D	X	X	X	X	- X V F G G G		4	2.1E+06	2.1E+05	4.1E+04	1.3E+04
	7	C	G	T	W	Y	D	X	X	X	X	- X V F G G G		5	6.7E+07	3.8E+06	5.0E+05	1.2E+05
	8	C	G	T	W	Y	D	X	X	X	X	X X V F G G G	2	6	2.1E+09	6.8E+07	6.0E+06	1.1E+06

Fig. 11

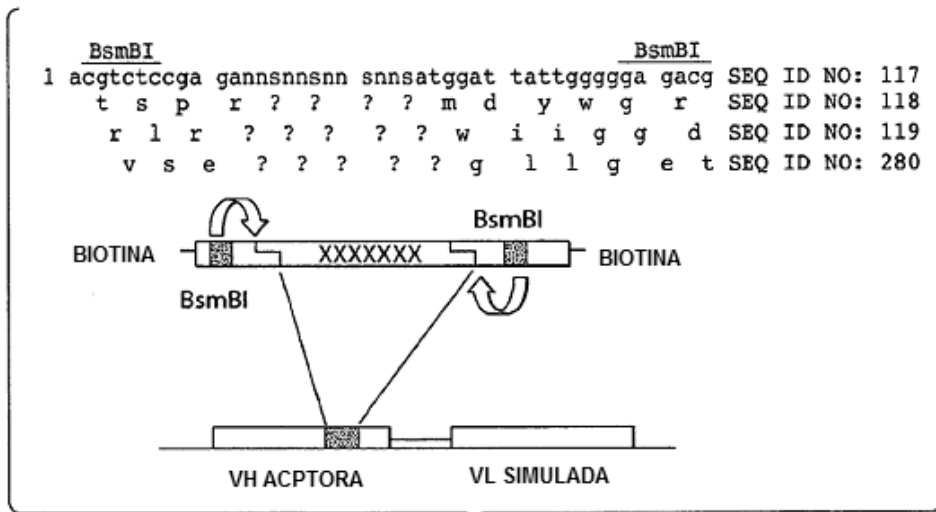


Fig. 12

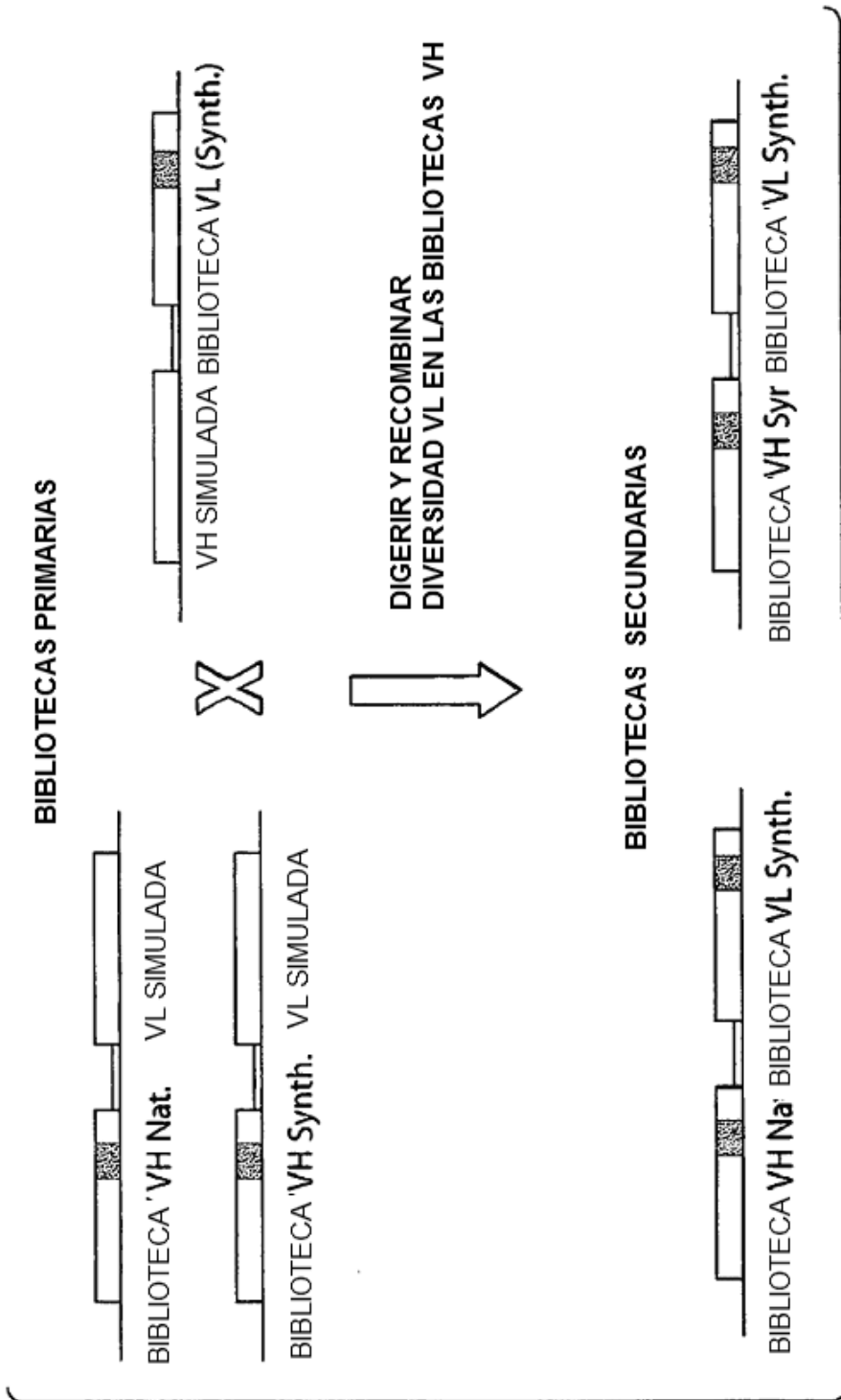


Fig. 13

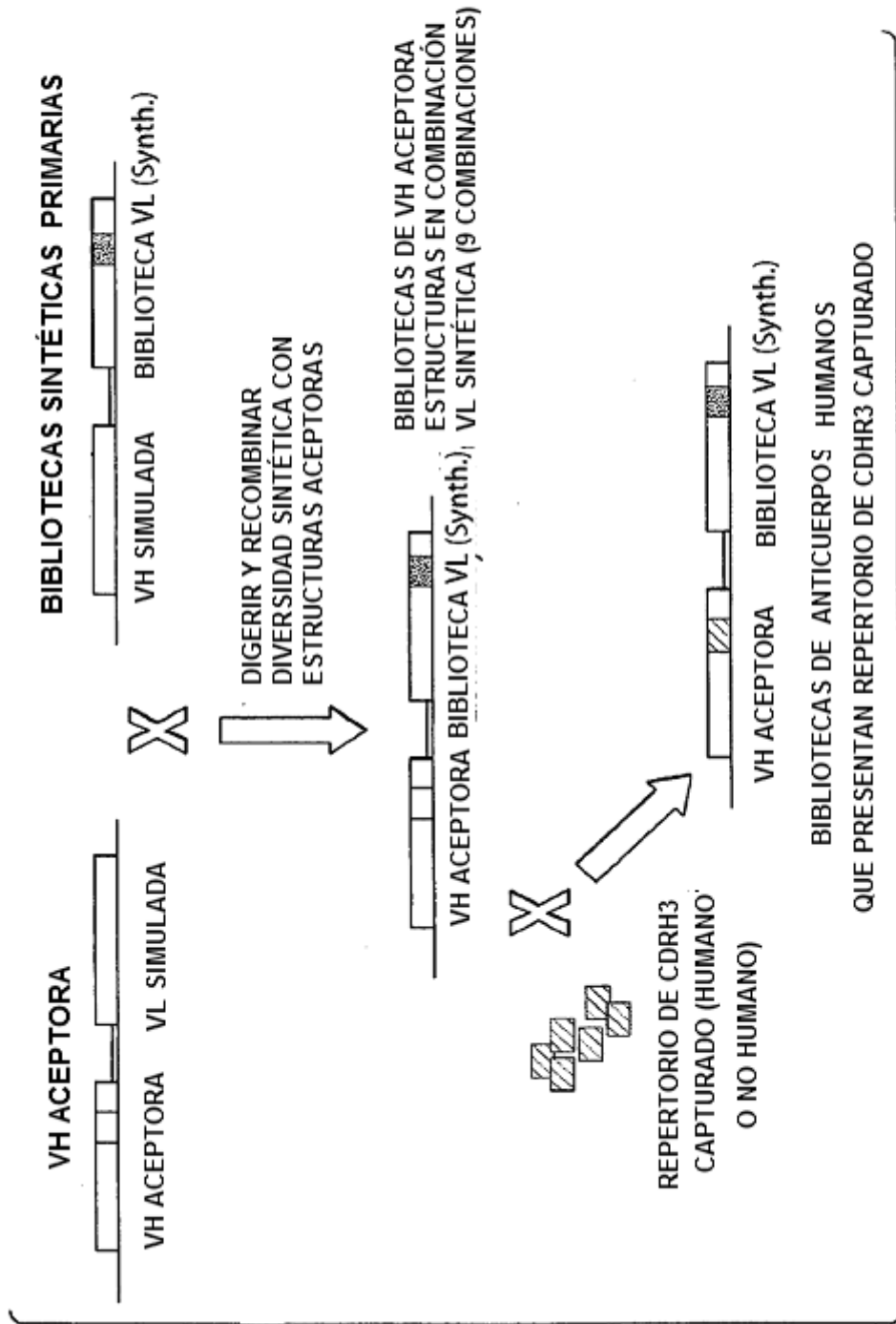


Fig. 14

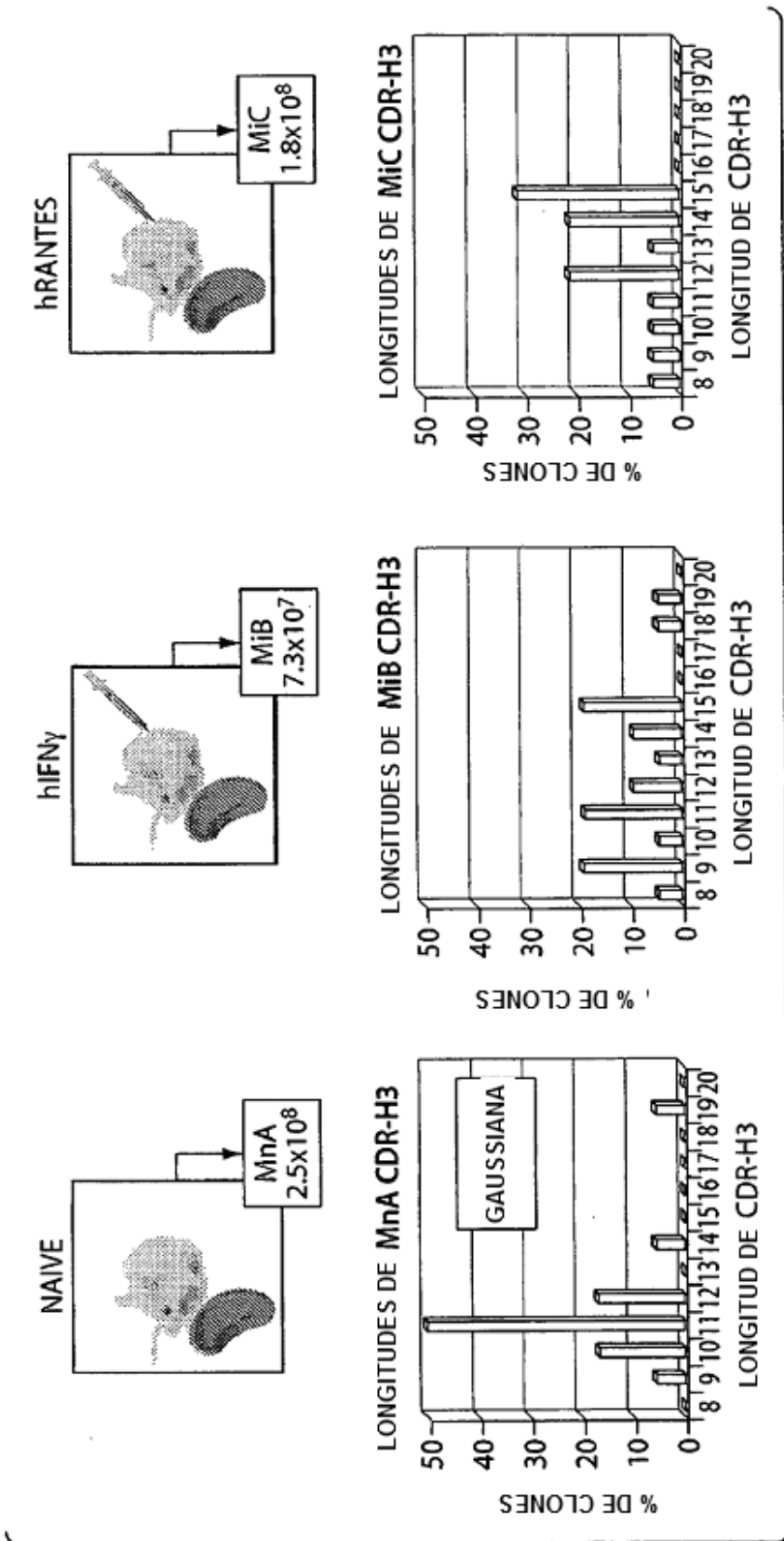


Fig. 15

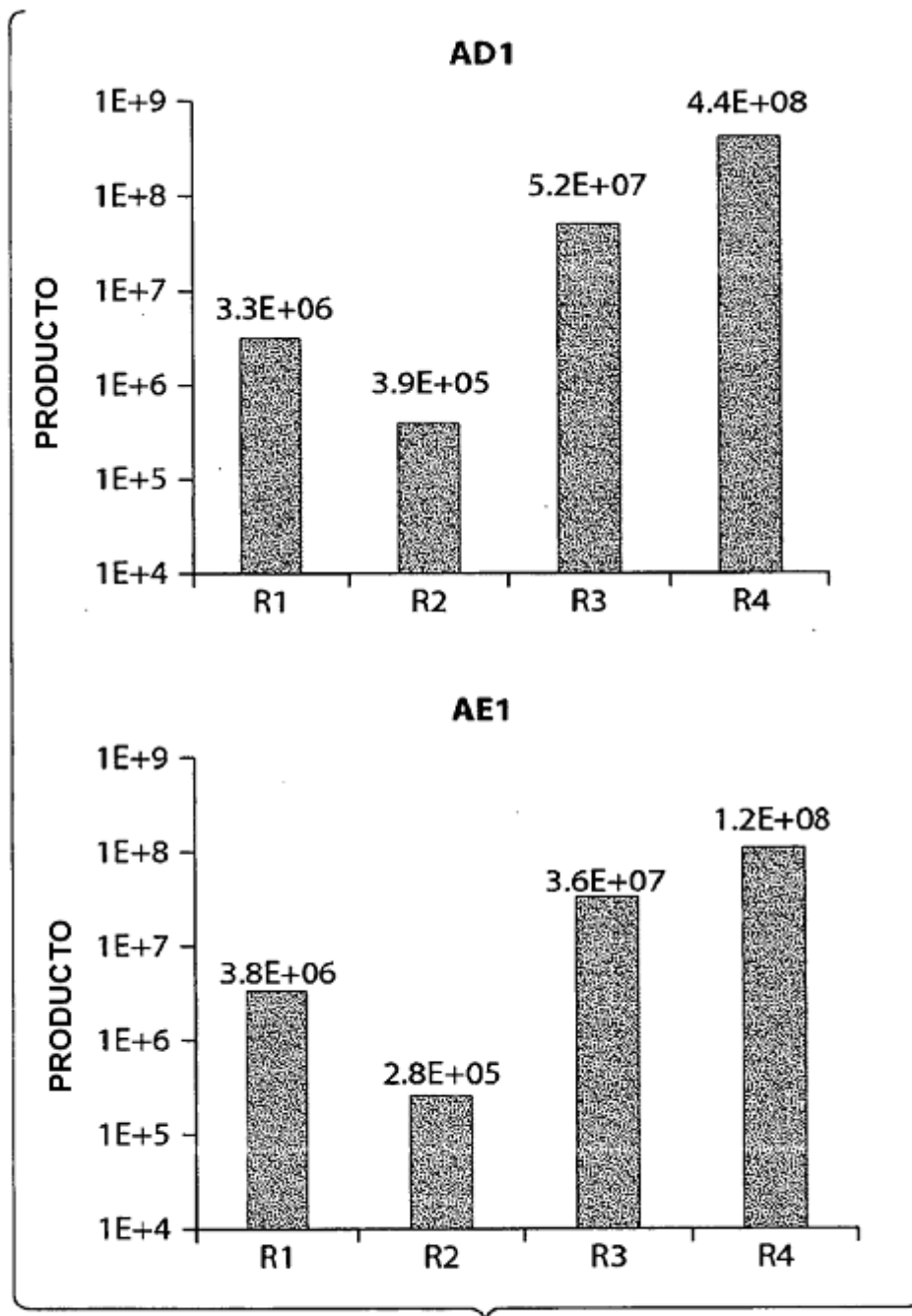


Fig. 16

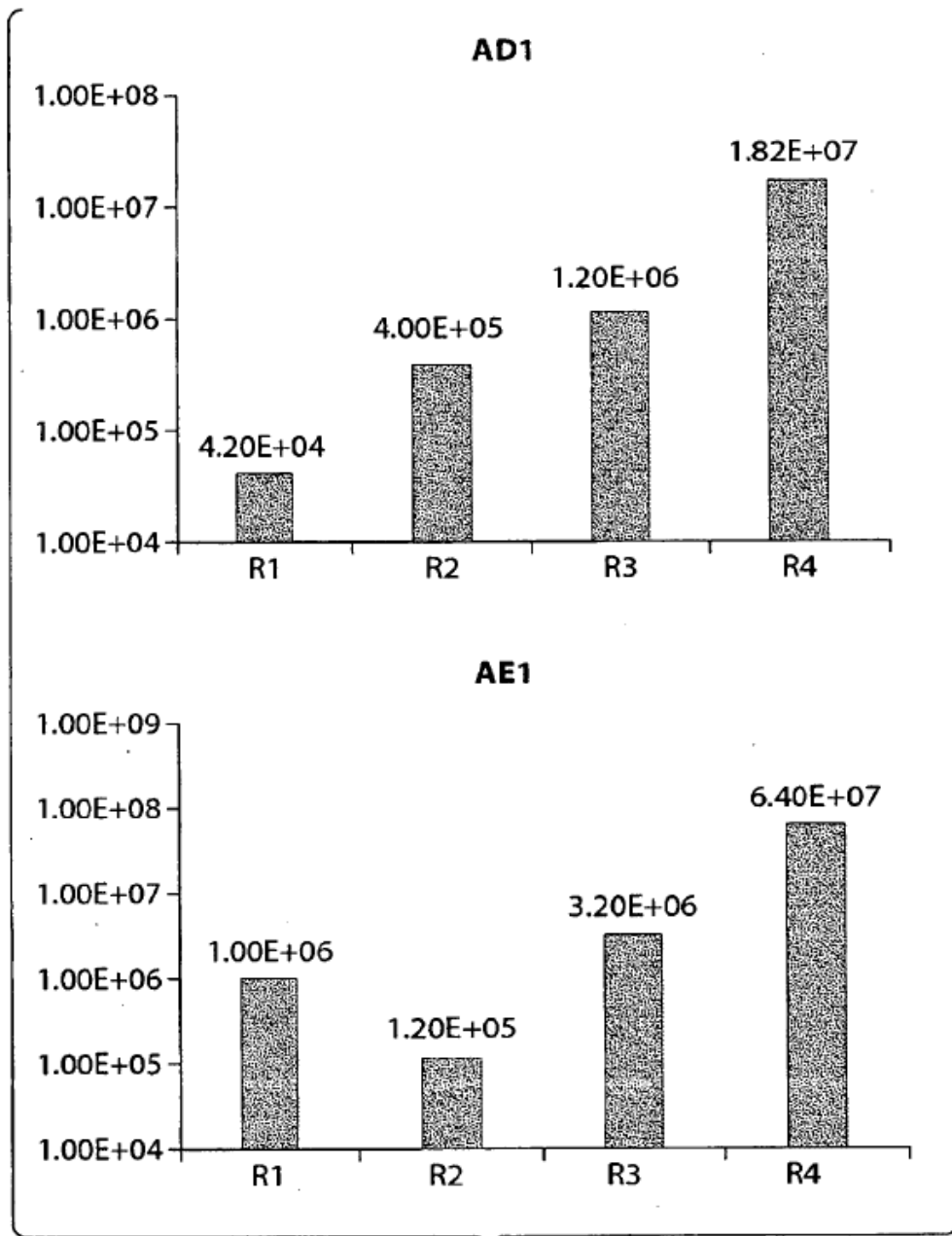


Fig. 17

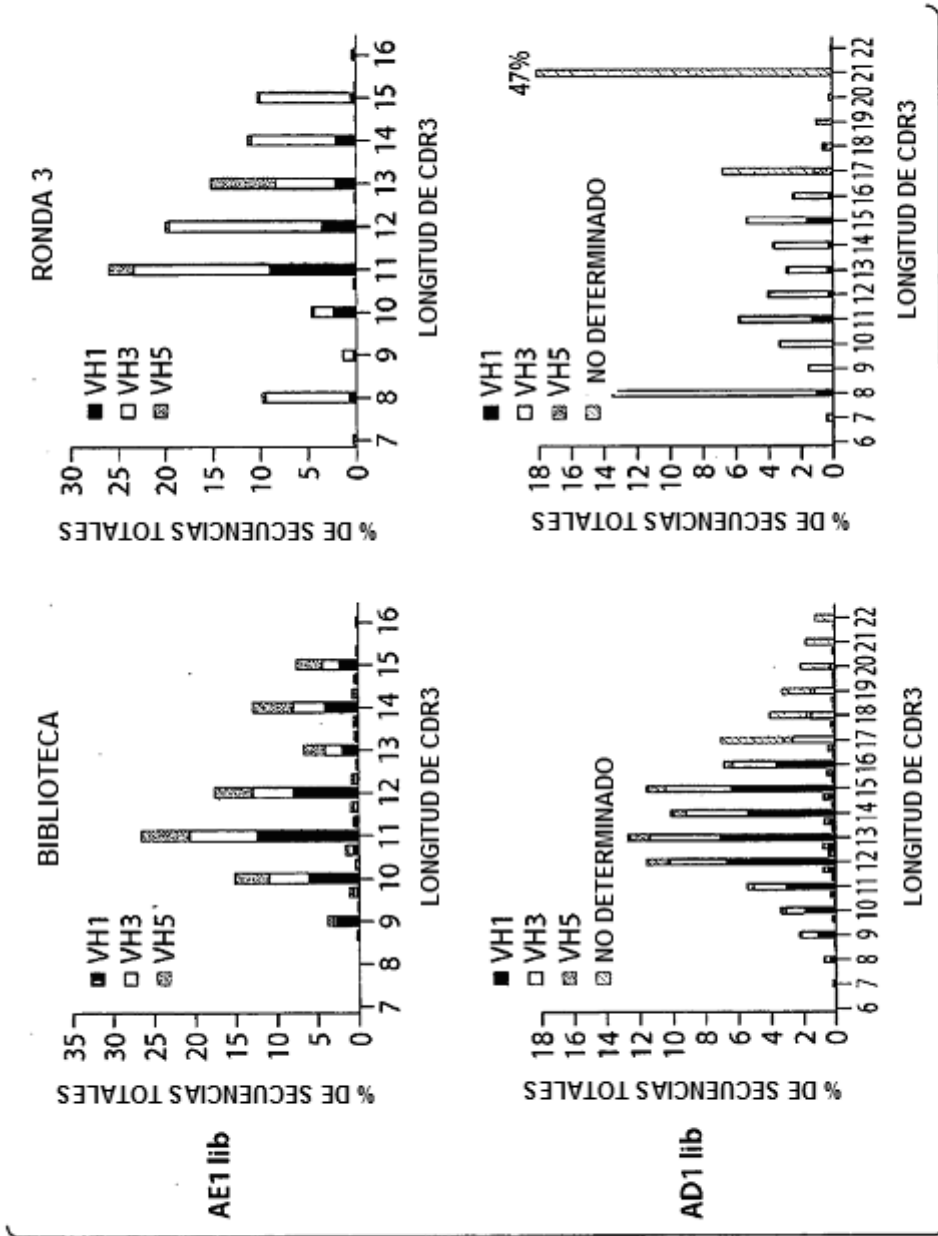


Fig. 18

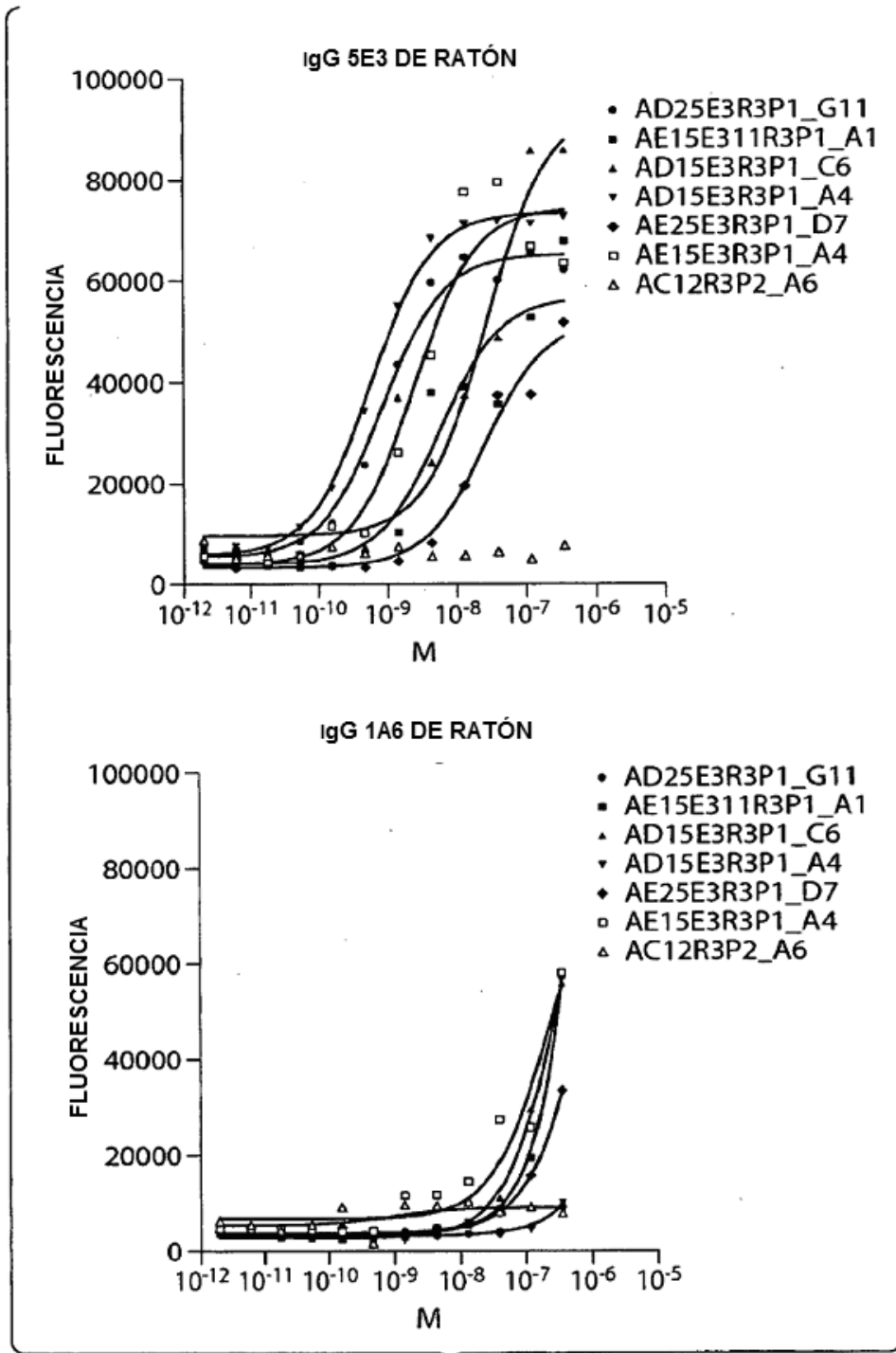


Fig. 19

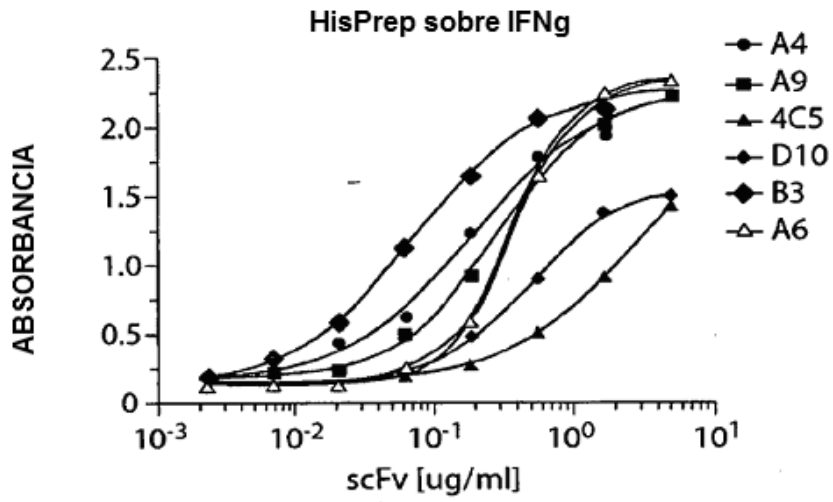


Fig. 20

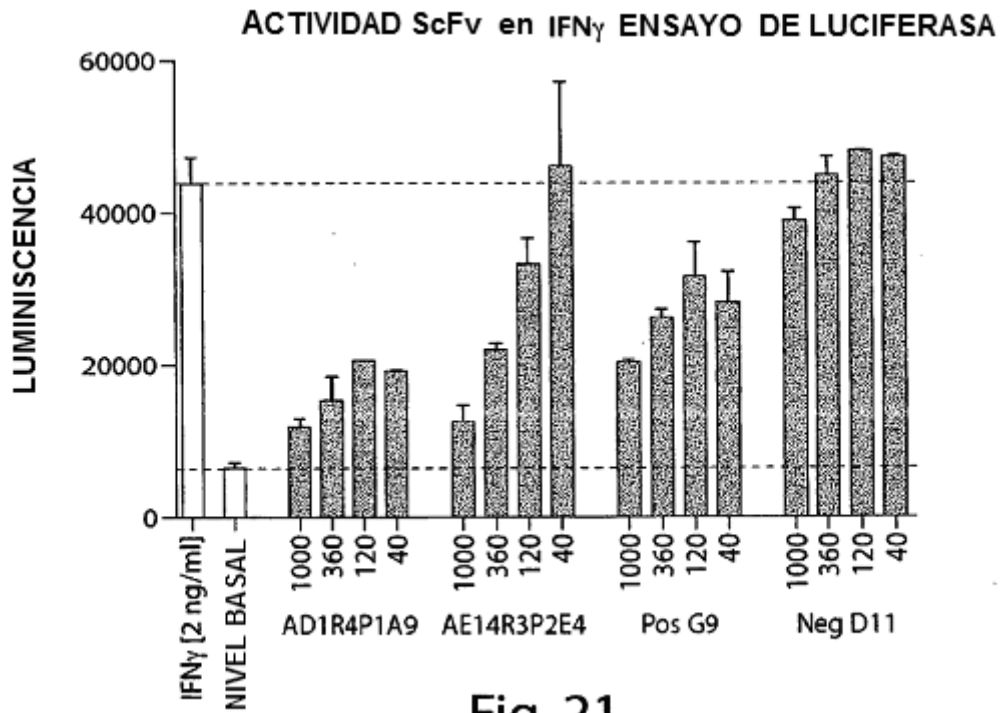


Fig. 21

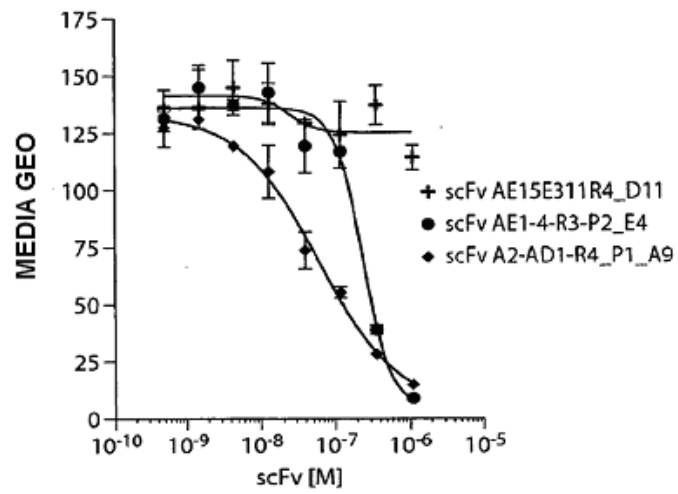


Fig. 22

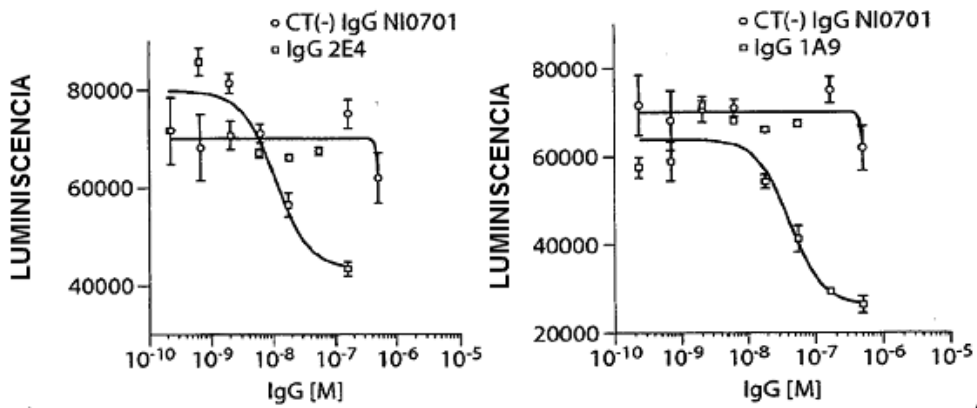


Fig. 23

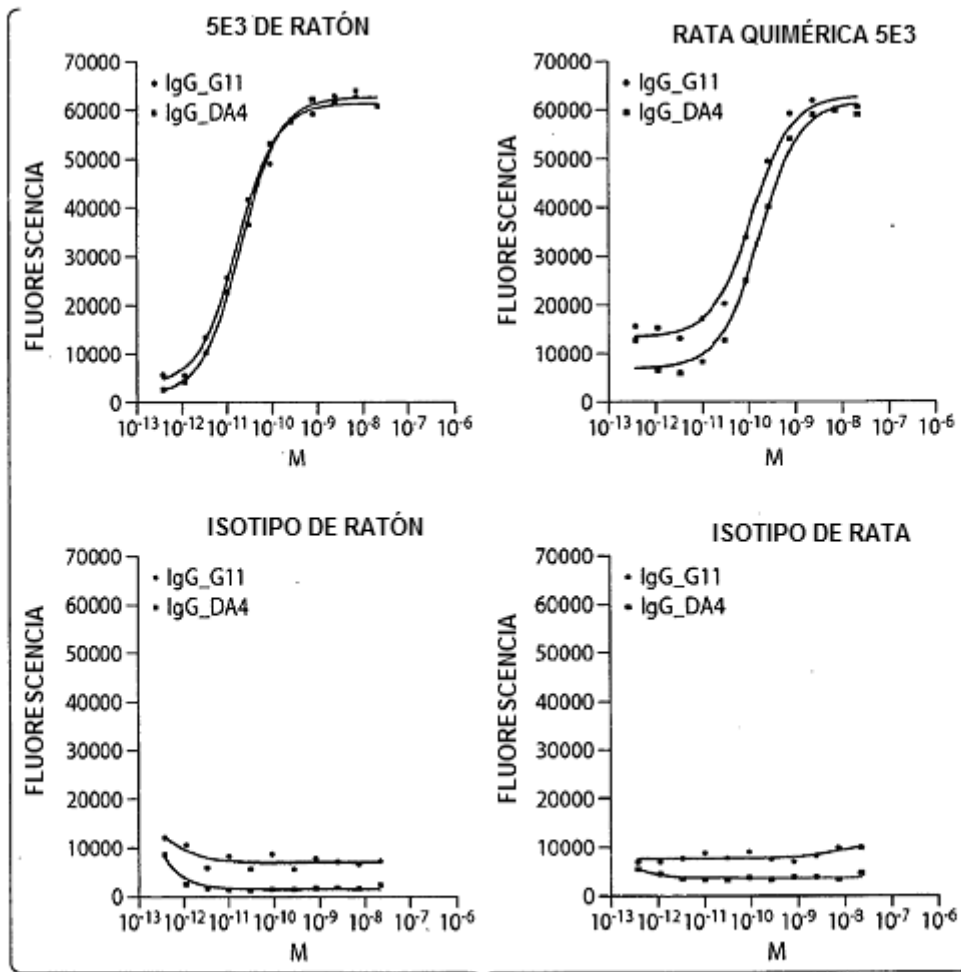


Fig. 24

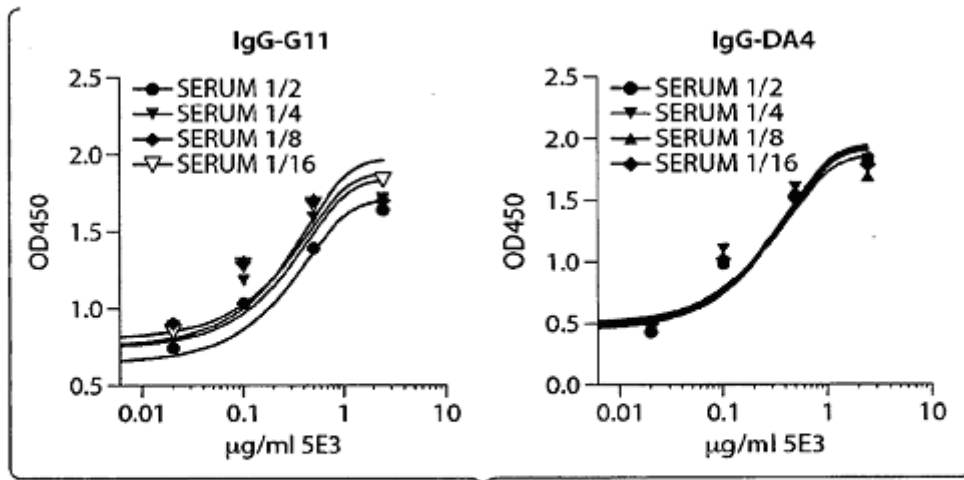


Fig. 25

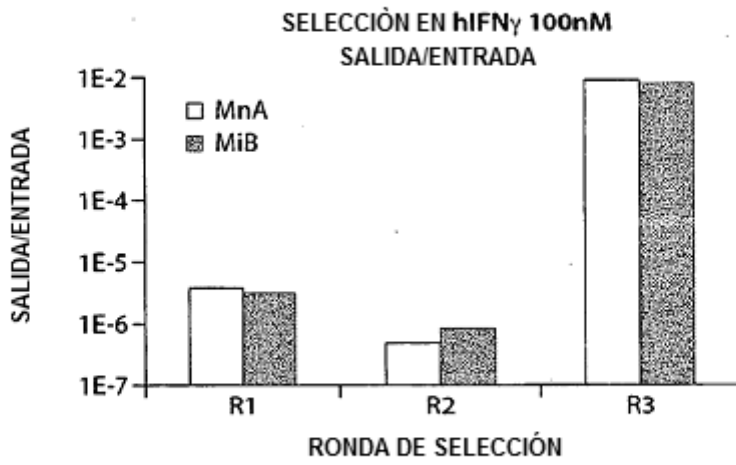


Fig. 26

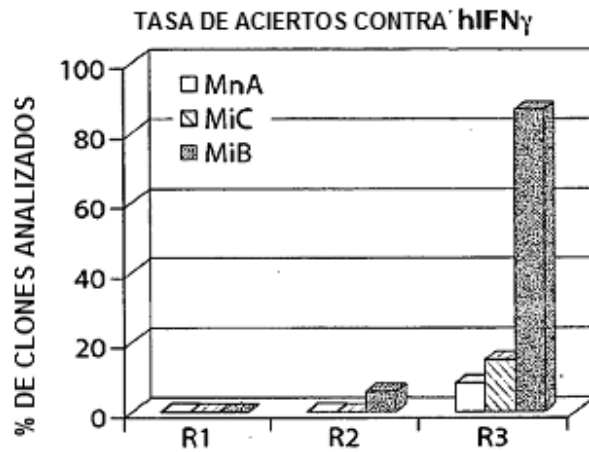


Fig. 27

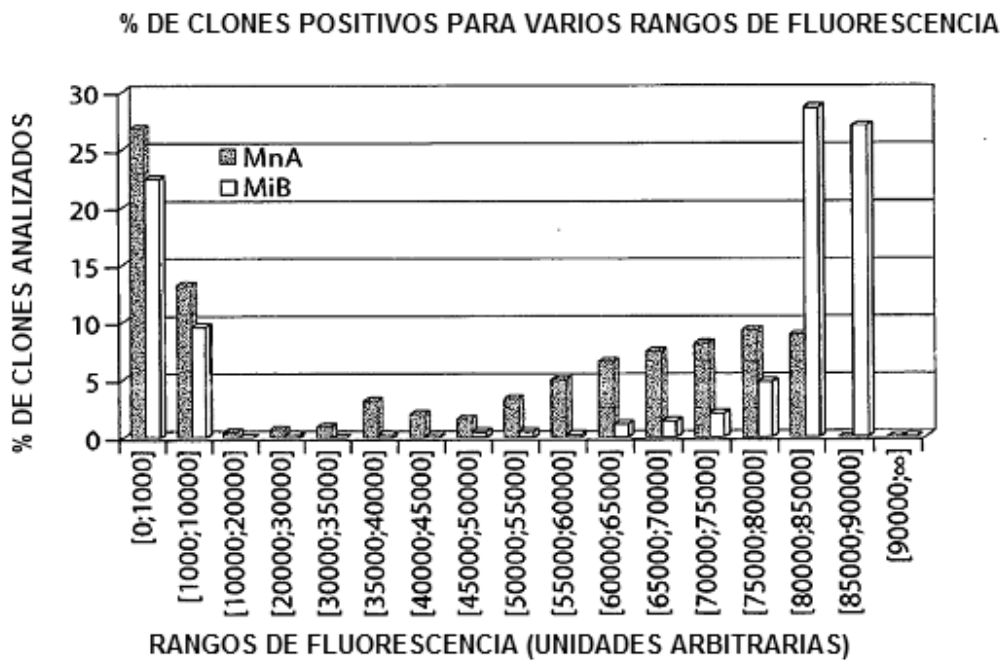


Fig. 28

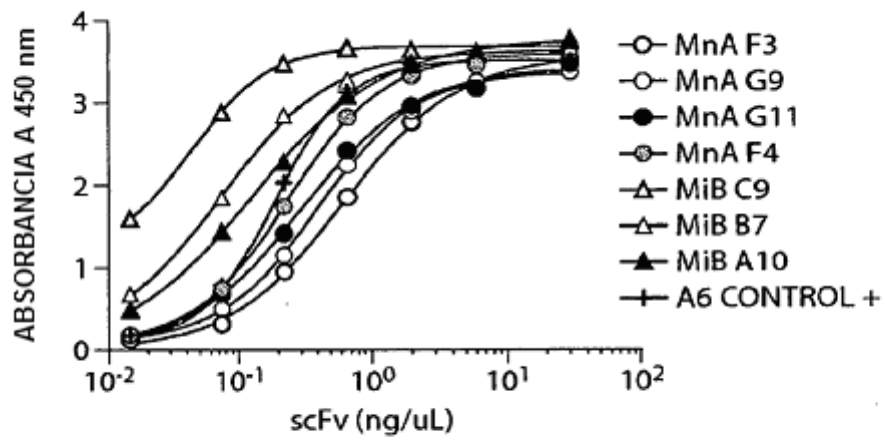


Fig. 29