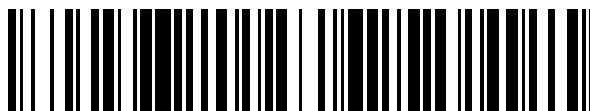


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 542 744**

51 Int. Cl.:

C12N 15/10

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.11.2010** **E 10842428 (4)**

97 Fecha y número de publicación de la concesión europea: **18.03.2015** **EP 2513312**

54 Título: **Bibliotecas de polipéptidos sintéticos y métodos para generar variantes polipeptídicas diversificadas de forma natural**

30 Prioridad:

17.03.2010 US 314794 P

20.05.2010 US 784190

17.12.2009 US 287336 P

02.09.2010 US 379571 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.08.2015

73 Titular/es:

NOVIMMUNE SA (100.0%)

14 ch. des Aulx 1228 Plan-Les-Ouates
Geneva, CH

72 Inventor/es:

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

74 Agente/Representante:

SUGRAÑES MOLINÉ, Pedro

ES 2 542 744 T3

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

DESCRIPCIÓN

Bibliotecas de polipéptidos sintéticos y métodos para generar variantes polipeptídicas diversificadas de forma natural

5 **Campo de la invención**

La invención se refiere a la generación de bibliotecas de secuencias de ADN que codifican polipéptidos homólogos y al uso de dichas bibliotecas. La presente invención se refiere en particular a la generación de colecciones de fragmentos de anticuerpos sintéticos en las que una o varias reacciones determinantes de complementariedad (CDR) se reemplazan por una colección de las CDR correspondientes capturadas a partir de una fuente natural. La invención se refiere además a la generación de colecciones de fragmentos de anticuerpo que contienen CDR derivadas de un animal inmunizado y su uso como una mejor fuente para derivar fragmentos de anticuerpo de alta afinidad. La invención se refiere además a la dosificación de una parte de un polipéptido insertando una secuencia diversificada de origen sintético o natural sin la necesidad de modificación de la secuencia codificante del polipéptido original.

Antecedentes de la invención

Un anticuerpo está compuesto de cuatro polipéptidos: dos cadenas pesadas y dos cadenas ligeras. La parte de unión a antígeno de un anticuerpo está formada por el dominio variable de cadena ligera (VL) y el dominio variable de cadena pesada (VH). En un extremo de estos dominios seis bucles forman el sitio de unión al antígeno y también se denominan las regiones determinantes de complementariedad (CDR). Tres CDR se localizan 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 linfocitos B se forma una región de inmunoglobulina única por recombinación somática conocida como recombinación V(D)J. La región variable de la cadena pesada o ligera de inmunoglobulina está codificada por diferentes segmentos génicos. La cadena pesada está codificada por tres segmentos denominados segmentos variables (V), de diversidad (D) y de unión (J) mientras que la cadena ligera variable está formada por la recombinación de solamente dos segmentos V y J. Un gran número de parátomos de anticuerpos pueden generarse por 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 se genera por los acontecimientos de recombinación. Durante el transcurso de la respuesta inmunitaria se introduce más diversidad en el sitio de unión a antígeno por un proceso denominado hipermutación somática (SHM). Durante este proceso se introducen mutaciones puntuales en los genes variables de las cadenas pesadas y ligeras y en particular en las regiones que codifican las CDR. Esta variabilidad adicional permite la selección y expansión de linfocitos B que expresan variantes de anticuerpo con afinidad mejorada por su antígeno afín.

En años recientes han surgido varias tecnologías de presentación y estas permiten la exploración de grandes colecciones de proteínas o péptidos. Estas incluyen presentación en fagos, presentación bacteriana, presentación en levadura y presentación en ribosomas (Smith GP. *Science*. 14 Jun 1985; 228(4705): 1315-7; Hanes J y Plückthun A. *Proc Natl Acad Sci USA*. 13 May 1997; 94(10): 4937-42.; Daugherty PS *et al.*, *Protein Eng*. Sep 1998; 11(9): 825-32.; Boder ET y Wittrup KD. *Nat Biotechnol*. Jun 1997; 15(6):553-7). En particular estos métodos se han aplicado extensivamente a anticuerpos y fragmentos de los mismos. Se han descrito varios métodos para generar bibliotecas de polipéptidos y para explorar con respecto a miembros con propiedades de unión deseadas.

Un primer enfoque es capturar por amplificación génica genes de inmunoglobulina reordenados de repertorios naturales usando tejidos o células de seres humanos u otros mamíferos como una fuente de diversidad genética. Estas colecciones de cadenas pesadas y ligeras reordenadas (VH y VL) se combinan después para generar bibliotecas de pares de unión que pueden presentarse en bacteriófagos o en otros paquetes de presentación tales como bacterias, células de levadura o de mamífero. En este caso se captura una fracción grande del repertorio de inmunoglobulina hallado en el donante. Por lo tanto todos los marcos conservados codificados por los genes de línea germinal del donante pueden encontrarse en dichos repertorios así como diversidad generada tanto por recombinación de V(D)J como por hipermutación somática (Marks JD *et al.*, *J Mol Biol*. 5 Dic 1991; 222(3): 581-97; Patente de Estados Unidos N° 5.969.108 de McCafferty).

Una limitación de los repertorios naturales es que los anticuerpos de origen natural pueden basarse en marcos conservados con baja estabilidad intrínseca que limitan sus niveles de expresión, periodo de validez y su utilidad como reactivos o moléculas terapéuticas. Para superar estas limitaciones se han desarrollado varios métodos para generar bibliotecas de anticuerpos sintéticos. En estos enfoques, se seleccionan un único o un número limitado de marcos conservados de anticuerpos seleccionados codificados por sus genes de línea germinal correspondientes. La selección de estos marcos conservados se basa habitualmente en su estabilidad bioquímica y/o su frecuencia de expresión en repertorios de anticuerpos naturales. Para generar una colección de proteínas de unión, se introduce después diversidad sintética en todas o subconjunto de CDR. Típicamente se diversifica bien el total de o parte de las CDR usando diferentes estrategias. En algunos casos se introdujo diversidad en posiciones seleccionadas dentro de las CDR (Knappik A *et al.*, *J Mol Biol*. 11 Feb 2000; 296(1): 57-86). Los restos diana pueden ser los implicados frecuentemente en el contacto con antígenos, los que presentan diversidad máxima en repertorios de anticuerpos naturales o incluso restos que serían preferentemente dianas de la maquinaria celular implicada en la

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

Se han usado varios métodos para diversificar los CDR de anticuerpos. Se ha usado extensivamente PCR solapante usando oligonucleótidos degradados para ensamblar elementos de marco conservado y CDR para reconstituir genes de anticuerpo. En otro enfoque, se han generado sitios de enzimas de restricción únicos en las regiones marco conservadas en el límite de cada CDR permitiendo la introducción de CDR diversificadas por clonación mediada por enzimas de restricción. En cualquier caso, como todos los miembros de la biblioteca se basan en marcos conservados con características seleccionadas y preferidas, se anticipa que los anticuerpos derivados de estos repertorios son más estables y proporcionan una mejor fuente de reactivos útiles (Knappik, documento US 6696248; Sidhu SS, *et al.*, *Methods Enzymol.* 2000; 328: 333-63; Lee CV *et al.*, *J Mol Biol.* 23 Jul 2004; 340(5): 1073-93).

Soderlind *et al* (*Nature Biotechnology* 18(8): p 852-856) desvela un método para producir una colección de ácidos nucleicos en la que dicha secuencia de ácido nucleico contiene secuencias de FR1, CDR1, FR2, CDR2, FR3, CDR3 y FR4 ensambladas aleatoriamente de regiones variables de cadena pesada o ligera.

Sidhu *et al* (*J Mol Biol* 338(2): p 299-310) desvela un método para producir una colección de ácidos nucleicos en la que cada ácido nucleico codifica un dominio variable de inmunoglobulina que comprende una pluralidad de secuencias de CDR, en el que se introduce diversidad con diversidad sintética y mutagénesis dirigida.

El documento WO 2008/112640 desvela un método para producir una colección de ácidos nucleicos en el que cada ácido nucleico codifica un dominio variable de inmunoglobulina que comprende una pluralidad de secuencias de CDR, en el que se introduce diversidad con diversidad sintética y mutagénesis dirigida.

Sin embargo, una limitación importante de estas bibliotecas sintéticas es que una proporción significativa de los miembros de bibliotecas nos se expresan porque las secuencias diversificadas aleatoriamente no permiten una expresión y/o un plegamiento apropiados de la proteína. Este problema es particularmente significativo para la CDR3 de la cadena pesada. De hecho, esta CDR contribuye con frecuencia a la mayoría de la energía de unión para el antígeno y es altamente diversa en longitud y secuencia. Aunque la otra CDR (H1, H2, L1, L2 y L3) puede adoptar solamente un número limitado de conformaciones tridimensionales, conocidas como pliegues canónicos, el número de conformaciones que pueden adoptarse por la CDR3 de cadena pesada sigue siendo demasiado diverso para predecir (Al-Lazikani B *et al.*, *J Mol Biol.* 7 Nov 1997; 273(4): 927-48). Además, el uso de oligonucleótidos degradados largos usados para abarcar CDR H3 largas introduce con frecuencia delecciones de pares de bases individuales. Estos factores reducen significativamente el tamaño funcional de los repertorios sintéticos.

Los repertorios tanto naturales como sintéticos tienen ventajas y limitaciones. Por un lado, las estrategias que se basan en la captura de genes variables de anticuerpo reordenados de forma natural no son óptimas ya que incluyen potencialmente marcos conservados menos favorables dentro de la biblioteca. Un aspecto positivo es que estos genes variables reordenados incluyen CDR que son compatibles con un plegamiento 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 marcos conservados e inserción de diversidad sintética se benefician de la estabilidad mejorada de los marcos conservados pero está limitada por el gran número de secuencias de CDR que no son compatibles con el plegamiento y/o la expresión y pueden desestabilizar el dominio general (Figura 1A). Existe por tanto la necesidad de enfoques nuevos que puedan combinar los beneficios de usar marcos conservados seleccionados con características deseables y combinarlos con CDR plegadas de forma apropiada por ejemplo derivadas de repertorios naturales.

Todos los enfoques descritos para generar bibliotecas de anticuerpos bien por captura de secuencias de anticuerpo reordenadas de forma natural o bien generando diversidad por medios sintéticos están limitados por la aparición de mutaciones de desplazamiento de fase que conducen 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 de fragmentos de ADN así como clonación molecular. La frecuencia de miembros no funcionales en bibliotecas de anticuerpo varía típicamente del 15 % al 45 % dependiendo de las estrategias empleadas para captura o generar la diversidad de anticuerpos (Persson MA *et al.*, *Proc Natl Acad Sci U S A.* 15 Mar 1991; 88(6): 2432-6; Schoonbroodt S, *et al.*, *Nucleic Acids Res.* 19 May 2005; 33(9): e81; Söderling E *et al.*, *Nat Biotechnol.* Ago 2000; 18(8): 852-6; Rothe *et al.*, *J Mol Biol.* 29 Feb 2008; 376(4): 1182-200). La frecuencia de las secuencias que codifican anticuerpos no funcionales tiene una gran influencia en el proceso de identificación de anticuerpos. En primer lugar, el tamaño funcional de la biblioteca se reduce y, debido a que los clones no funcionales tienen con frecuencia una ventaja de crecimiento durante la propagación de las bibliotecas, se expanden más rápidamente y pueden comprometer el proceso de identificación de candidatos a anticuerpos (De Bruin R *et al.*, *Nat Biotechnol* 17 de abril de 1999: 397-399). Estos problemas se reconocen como limitaciones graves para aprovechar completamente el potencial de las bibliotecas de anticuerpos. La generación de bibliotecas altamente funcionales sigue siendo un reto en el campo y ha impulsado muchos intentos de mejorar el proceso. Por ejemplo, se han usado múltiples estrategias de diversificación que se dirigen a imitar el uso de aminoácidos hallado en secuencias de CDR naturales para muestrear más eficazmente la enorme diversidad de posibles combinaciones de

secuencia codificadas por CDR sintéticas (de Kruif J *et al.*, J Mol Biol. 21 Abr 1995; 248(1):97-105; Sidhu SS *et al.*, J Mol Biol. 23 Abr 2004; 338(2): 299-310). Otro enfoque es limpiar la biblioteca inicial para retirar clones no funcionales con el coste potencial de pérdida de diversidad. Esto se ha aplicado a la preselección de repertorios sintéticos uniendo la biblioteca de anticuerpos con un ligando genérico. Esta etapa permitió el enriquecimiento de miembros de bibliotecas que son capaces de expresarse y plegarse de forma apropiada y pueden usarse para recrear una biblioteca más funcional (Winter y Tomlinson, documento US 6.696.245 B2). Independientemente del enfoque de la calidad de cualquier biblioteca depende de la eficacia de los métodos de biología molecular aplicados para generar la biblioteca y en general conducen al 15 % a 45 % de miembros no funcionales de la biblioteca. Por lo tanto existe la necesidad de nuevos enfoques altamente eficaces que minimicen la frecuencia en genes no funcionales debido a desplazamientos de fase introducidos durante las etapas de clonación molecular y que maximicen la funcionalidad de bibliotecas capturando regiones CDR que tengan una alta propensión a plegarse correctamente en marcos conservados de anticuerpos con propiedades deseables. Además, existe la necesidad de enfoques que permitan la captura de las secuencias de CDR de un repertorio inmunitario animal en un contexto terapéuticamente útil tal como marcos conservados de anticuerpos humanos para mejorar el proceso de generación de anticuerpos de alta afinidad.

Sumario de la invención

La presente invención proporciona métodos para generar bibliotecas de secuencias de ácido nucleico que combinen los beneficios de la selección de un marco conservado estable y la inserción de regiones determinantes de complementariedad (CDR) codificadas de forma natural o secuencias de aminoácidos que puedan 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 CDR largas o secuencias de aminoácidos que puedan cumplir el papel de una CDR que son muy difíciles de codificar usando enfoques sintéticos. La presente invención, combinando marcos conservados estables y CDR plegadas de forma apropiada o secuencias de aminoácidos que puedan cumplir el papel de una CDR, maximiza la proporción de anticuerpos funcionales de la biblioteca y por lo tanto el rendimiento del proceso de selección y la calidad de los clones seleccionados. La invención proporciona un método para capturar CDR expresadas de forma natural de diferentes especies y para insertarlas en un marco conservado de anticuerpo humano. Esto permite el uso de repertorios de CDR H3 que difieran significativamente en su longitud y composición en comparación con el repertorio humano. La invención permite la generación de fragmentos de anticuerpo humano que presentan repertorios estructurales derivados de otras especies y por lo tanto la capacidad para tomar muestras de espacios estructurales diferentes. Los presentes métodos también se usan para introducir CDR de origen sintético o secuencias de aminoácidos que puedan cumplir el papel de una CDR con una mayor frecuencia de éxito que métodos alternativos que introducen menos errores que provocan desplazamientos de fase en la secuencia codificante. Las bibliotecas generadas usando los presentes métodos contienen una alta frecuencia de variantes funcionales. Se usan bibliotecas de variantes generadas de acuerdo con este método para selección y exploración con cualquier tecnología de presentación, selección y exploración descrita.

El análisis de repertorios inmunitarios de diferentes especies o, dentro de una especie, en diferentes estadios del desarrollo ha revelado algunas diferencias sorprendentes en las características de la composición y longitud de CDR H3. Por ejemplo la longitud de CDR H3 promedio en seres humanos es más larga en adultos en comparación con la vida fetal o con neonatos (Schroeder Jr, HW *et al.*, 2001 Blood 98; 2745-2751). Resulta interesante que a pesar de grandes similitudes entre genes de línea germinal de anticuerpos de primates y humanos, la evolución de la longitud de CDR H3 durante el desarrollo difiere (Link JM *et al.*, Molecular Immunol. 2005 42; 943-955). La comparación de las secuencias de CDR H3 halladas en ratones y seres humanos claramente muestra que la longitud promedio es significativamente más corta en ratones (Rock EP *et al.*, J Exp Med 1994 179; 323-328). Durante el desarrollo de linfocitos B temprano en la médula ósea, la longitud de CDR H3 promedio aumenta en ratones mientras que tiende a reducirse en seres humanos y además la composición de aminoácidos de los repertorios de CDR H3 murinos y humanos difiere (Zemlin M *et al.*, 2003 J Mol Biol 334; 733-749; Ivanov I *et al.*, 2005 J Immunol 174; 7773-7780). Estos ejemplos indican que diferentes especies expresan diferentes series de repertorios de CDR H3 a pesar del hecho de que se exponen globalmente a clases similares de antígenos y la importancia biológica de estas observaciones aún debe estudiarse adicionalmente. Se ha demostrado que la forma del sitio de combinación de anticuerpos dirigidos contra antígenos pequeños tales como haptenos o péptidos difiere de los dirigidos contra proteínas grandes y la forma del sitio de combinación está dictada por la longitud y composición de las CDR (Collis A *et al.*, J Mol Biol 2003 325; 337-354). A partir de estos hallazgos se puede anticipar que el repertorio de CDR H3 expresado por diferentes especies tiene diversas propensiones a reaccionar eficazmente contra diferentes clases diana.

Los métodos proporcionados y las bibliotecas de anticuerpos descritas en el presente documento se diseñan para aprovechar los diversos repertorios expresados por diferentes especies para la generación de anticuerpos terapéuticos. Estos repertorios que exploran diferentes espacios tridimensionales podrían permitir la generación de anticuerpos contra una mayor diversidad de clases y epítopos diana. Se han descrito bien métodos para generar bibliotecas de animales sin tratamiento previo o inmunizados y estos métodos permiten la captura de los repertorios correspondientes y la generación de anticuerpos. Sin embargo, los anticuerpos derivados de estas bibliotecas no son de origen humano y por lo tanto no están bien adaptados para terapia humana sin realizar trabajo técnico adicional, tal como humanización. Existe por lo tanto una necesidad de nuevos métodos para aprovechar la

diversidad expresada en el repertorio de diferentes especies y para aprovechar esta diversidad en el contexto terapéuticamente útil de un anticuerpo humano.

Los métodos proporcionados y las bibliotecas de anticuerpo descritas en el presente documento abordan varias de las limitaciones descritas anteriormente y son una mejora frente a la técnica actual. En primer lugar, los métodos proporcionados en el presente documento combinan los beneficios de la selección de marco conservado estable y la inserción de CDR codificadas de forma natural que se ha seleccionado en un contexto natural de un anticuerpo funcional. En segundo lugar, los métodos permiten una inserción altamente eficaz de secuencias de CDR sintéticas o naturales en un marco conservado de anticuerpo que minimiza significativamente el número de desplazamientos de fase en la biblioteca y por lo tanto mejora su calidad. Finalmente, la invención permite un modo nuevo de usar la diversidad estructural de anticuerpos de origen natural capturando repertorios de CDR H3 expresados de forma natural de diferentes especies e insertarlos en marcos conservados de anticuerpos humanos. Es por lo tanto posible aprovechar estos repertorios estructuralmente diversos de una manera productiva para la generación de anticuerpos para terapia humana.

Los métodos proporcionados en el presente documento generan anticuerpos que contienen un marco conservado estable y CDR plegadas correctamente o secuencias de aminoácidos que pueden cumplir el papel de una CDR. Los métodos capturan la diversidad natural de secuencias en marcos conservados estables.

En los métodos proporcionados en el presente documento, las secuencias de línea germinal para regiones marco conservadas 1, 2 y 3 (FR1, FR2 y FR3) se seleccionan del organismo deseado, por ejemplo, del genoma humano (véase por ejemplo, Figuras 2 y 6). Las regiones FR3 y FR4 se separan por una secuencia de relleno que comprende al menos dos sitios de restricción de Tipo II intercalados con una secuencia de ácido nucleico aleatoria que realiza la función de una región CD3, que actuará como un sitio de integración para secuencias diversificadas. Se introduce diversidad de la secuencia fuera de la región codificante de inmunoglobulina introduciendo sitios de restricción de Tipo II, en la región determinante de complementariedad 3 de cadena pesada variable (CDR H3) o la región determinante de complementariedad 3 de cadena ligera variable (CDR L3). La invención proporciona por lo tanto un método de acuerdo con la reivindicación 1.

Aunque los ejemplos proporcionados en el presente documento demuestran diversidad en la región CDR3 (en la región de cadena pesada variable y/o región de cadena ligera variable), también se desvela que puede conseguirse diversidad en cualquier localización deseada, tal como, pero sin limitación, la región CDR1 (en la región de cadena pesada variable y/o región de cadena ligera variable) o la región CDR2 (en la región de cadena pesada variable y/o región de cadena ligera variable). Se generan secuencias de ADN diversificadas con secuencias flanqueantes que incluyen sitios de restricción de Tipo II. En los métodos proporcionados en el presente documento, los extremos cohesivos generados por las enzimas de restricción son compatibles y se mantiene la fase de lectura, permitiendo de este modo que se ligan fragmentos de ADN diversificados en un marco conservado aceptor.

Los métodos proporcionados en el presente documento también son útiles para generar secuencias de aminoácidos que tienen regiones diversificadas codificadas en las mismas. Por ejemplo, en los métodos proporcionados en el presente documento, las secuencias para las partes no diversificadas del aminoácido codificado se seleccionan del organismo deseado, por ejemplo, de la secuencia humana. Una parte de la secuencia de aminoácidos codificada se modifica introduciendo una secuencia de relleno que actuará como un sitio de integración para secuencias diversificadas. Se introduce diversidad en la secuencia en la localización o las localizaciones deseadas introduciendo sitios de restricción de Tipo II en la secuencia de ácido nucleico de relleno que intercala FR3 y FR4. Se generan secuencias de ADN diversificadas con secuencias flanqueantes que incluyen sitios de reconocimiento de Tipo II. En los métodos proporcionados en el presente documento, los extremos cohesivos generados por las enzimas de restricción son compatibles y la fase de lectura se mantiene, permitiendo de este modo que los fragmentos de ADN diversificados se ligan en un marco conservado aceptor.

Los métodos proporcionados en el presente documento también son útiles para generar bibliotecas de ácidos nucleicos diversos que codifican un alto porcentaje de polipéptidos que pueden plegarse de forma apropiada y expresarse como una entidad funcional tal como, por ejemplo, una inmunoglobulina.

Varios factores pueden influir significativamente en la calidad de un repertorio de polipéptidos, tal como una biblioteca de anticuerpos, y por lo tanto la probabilidad de identificar polipéptidos con propiedades deseadas. El tamaño y la diversidad del repertorio son obviamente críticos, y los estudios han demostrado la correlación entre el tamaño de un repertorio de anticuerpos y la afinidad de los anticuerpos aislados de ese repertorio (Griffiths *et al.*, EMBO J. 15 Jul 1994; 13(14): 3245-60). El tamaño de una biblioteca se determina típicamente por el número de transformantes obtenido durante la construcción y la diversidad se estima secuenciando un número limitado de miembros de la biblioteca. Este tipo de análisis proporciona solamente una evaluación superficial de la calidad de la biblioteca. En particular, la información de secuencia no puede indicar con fiabilidad si un polipéptido diversificado puede plegarse de forma apropiada y expresarse como una entidad funcional. Por lo tanto, dependiendo de la fuente de diversidad o de la estrategia que se aplica para diversificar un polipéptido, el tamaño funcional del repertorio puede diferir significativamente de su tamaño teórico (basado en la evaluación de tamaño y diversidad). Idealmente, un repertorio debería contener solamente miembros funcionales que puedan producir un polipéptido que tenga

potencialmente las características deseadas. Además, los miembros de la biblioteca que codifican polipéptidos no funcionales representan no solamente diversidad inútil sino que también tienen una influencia negativa importante durante el proceso de selección.

Como se ha descrito anteriormente, la calidad de cualquier biblioteca depende de la eficacia de los métodos de biología molecular aplicados para generar la biblioteca y muchos métodos generalmente conducen a aproximadamente 15 % a 45 % de miembros no funcionales de la biblioteca. Por lo tanto es importante durante las etapas de clonación o diversificación de la construcción de bibliotecas maximizar el número de secuencias que están en fase e idealmente codificar polipéptidos que pueden plegarse en un polipéptido funcional. Se han descrito métodos basados en la preselección de miembros de una biblioteca para plegamiento apropiado mediante unión con proteínas tales como Proteína A o Proteína L. (Véase, por ejemplo, Winter y Tomlinson, documento US 6.696.245 B2). Además, como pueden introducirse errores que conducen a desplazamientos de fase en la secuencia codificante en cada etapa de clonación, es importante minimizar el número de etapas de ensamblaje de ADN o clonación y desarrollar estrategias de clonación eficaces. El enfoque de clonación de restricción de Tipo IIS descrito en la invención conduce a un gran número de insertos en fase (>90 %) pero no asegura que las secuencias de ADN diversificadas que se clonan codifiquen un polipéptido que permita el plegamiento apropiado de un dominio variable de inmunoglobulina y puedan cumplir la función de una CDR.

Por lo tanto, la invención proporciona métodos para abordar estas limitaciones y generar bibliotecas de ácidos nucleicos diversificados que codifican un mayor porcentaje de miembros funcionales. El método de la invención permite la selección de diversidad funcional introducida en uno de los dominios variables de anticuerpo expresando los dominios variables de cadena pesada o ligera diversificados en el contexto de un dominio constante pesado o variable (cadenas de prueba) y seleccionando miembros de la biblioteca que puedan expresarse y presentarse en la superficie de un sistema de presentación tal como fagos. Esta etapa de preselección se consigue expresando el repertorio de polipéptidos diversificado usando un fago auxiliar que no codifica una proteína pIII de tipo silvestre. En este sistema, el ensamblaje de fagos se basa en la proteína de fusión polipéptido-pIII que por lo tanto tiene que tener la capacidad de expresarse y plegarse de forma suficiente para integrarse en una partícula de fago. Este fago auxiliar deficiente en pIII denominado "Hiperfago" se ha descrito como un modo de seleccionar fases de lectura abiertas (véase por ejemplo, Hust M *et al.*, *Biotechniques* Sep 2006; 41(3): 335-42). Una limitación de esta técnica, sin embargo, es que, después de la preselección, la cadena variable común que se expresó junto con el repertorio diversificado tiene que reemplazarse por otro repertorio variable para obtener una biblioteca con cadenas pesadas y ligeras diversificadas usando clonación de restricción convencional del dominio variable completo.

Para combinar el beneficio de la invención para diversificación de la región CDR3 por captura de diferentes fuentes de diversidad natural o sintética usando una enzima de restricción de Tipo IIS y el uso de una cadena común para preselección de repertorio, se describen métodos para identificar dominios variables comunes, o de prueba, que contienen un fragmento de ADN de relleno usado para clonación de diversidad que también puede cumplir la función de una CDR3 funcional. Esto permite la generación de bibliotecas de Aceptores que contienen dominios variables de cadena ligera preseleccionados y diversificados funcionales que pueden usarse directamente para la inserción de CDR H3 de captura como se muestra en la Figura 30. Los ejemplos proporcionados en el presente documento describen métodos para identificar dichas secuencias, así como varios ejemplos de dichos fragmentos de ADN de relleno que deben cumplir tres restricciones principales: 1) incluir dos sitios de restricción de Tipo IIS; 2) mantener la fase de lectura entre las regiones FR3 y FR4 y 3) codificar una CDR3 de dominio variable pesado que permita el plegamiento y la expresión de un dominio variable de anticuerpo.

Las bibliotecas generadas usando el método proporcionado en el presente documento tienen una mayor frecuencia de miembros potencialmente funcionales reduciendo o eliminando secuencias fuera de fase. Dichas bibliotecas preseleccionadas contienen al menos 90 % de secuencias que están en fase y por lo tanto tienen el potencial de codificar un polipéptido funcional.

En los métodos proporcionados en el presente documento, un "Marco conservado Aceptor" se genera usando un "fragmento de relleno" de ADN que contiene y está limitado, preferentemente, por dos sitios de enzimas de restricción de Tipo IIS (véase por ejemplo, Figura 6). Preferentemente, estos dos sitios de enzimas de restricción de Tipo IIS digieren secuencias en el límite del sitio en el que se desea diversidad, tal como, por ejemplo, la región CDR H3 o la región CDR L3.

Como se usa en el presente documento, la expresión "Marco conservado Aceptor" se refiere a una secuencia de ácido nucleico que incluye las secuencias de ácido nucleico que codifican las regiones FR1, FR2, FR3 y FR4, las secuencias de ácido nucleico que codifican dos CDR o secuencias de aminoácidos que pueden cumplir el papel de estas CDR, y un "fragmento de relleno" que actúa como el sitio de integración para secuencia de ácido nucleico diversificada. Se desea diversidad en la región CDR3 (en la región de cadena pesada variable y/o la región de cadena ligera variable), de modo que en el método de la invención el Marco conservado Aceptor incluye las secuencias de ácido nucleico que codifican las regiones FR1, FR2, FR3 y FR4, las secuencias de ácido nucleico que codifican las regiones CDR1 y CDR2, y un "fragmento de relleno" que actúa como el sitio de integración para la secuencia de ácido nucleico diversificada.

Las expresiones “fragmento de relleno”, “fragmento de ADN de relleno” y “secuencia de relleno” o cualquier variación gramatical de las mismas se usan de forma intercambiable en el presente documento para hacer referencia a una secuencia de ácido nucleico que incluye al menos dos sitios de reconocimiento de Tipo IIs y una secuencia diversificada. El Marco conservado Aceptor puede ser un Marco conservado Aceptor de cadena pesada variable (VH) o un Marco conservado Aceptor de cadena ligera variable (VL). El uso de los Marcos conservados Aceptores y los fragmentos de relleno contenidos en los mismos permiten la integración de una secuencia de CDR (natural o sintética) o una secuencia de aminoácidos que puede cumplir el papel de la CDR en el Marco conservado Aceptor sin nucleótidos o restos de marco conservado donantes contenidos en la misma o necesarios para integración. Por ejemplo, el uso de los Marcos conservados Aceptores y los fragmentos de relleno contenidos en los mismos permiten la integración de una secuencia de CDR (natural o sintética) seleccionada de CDR H3 y CDR L3, o una secuencia de aminoácidos que pueda cumplir el papel de una CDR seleccionada de CDR H3 y CDR L3 en el Marco conservado Aceptor sin nucleótidos o restos de marco conservado donante contenidos en los mismos o necesarios para integración. Por lo tanto, tras la integración, el fragmento de relleno se retira completamente, y la región codificante de la proteína aceptora y los fragmentos proteicos insertados (es decir, las CDR) están intactos.

En algunas realizaciones, el fragmento de relleno incluye dos sitios de restricción de Tipo IIs, mantiene la fase de lectura entre las regiones FR3 y FR4 y codifica una CDR3 de dominio variable pesado que permite el plegamiento y la expresión de un dominio variable de anticuerpo.

Los métodos proporcionados en el presente documento usan cebadores que están diseñados para contener sitios de escisión para contener sitios de escisión para enzimas de restricción de Tipo IIs en el límite del sitio en el que se desea diversidad, por ejemplo, la región CDR H3 o la región CDR L3. Se capturan clones de CDR de origen natural, aleatorios, (véase por ejemplo, Figura 10) o secuencias de CDR sintéticas (véase por ejemplo, Ejemplo 6) o secuencias de aminoácidos que puedan cumplir el papel de la CDR en los Marcos conservados Aceptores usados en el presente documento. Por ejemplo, se capturan clones de CDR3 de origen natural, aleatorios (véase por ejemplo, Figura 10) o secuencias de CDR3 sintéticas (véase por ejemplo, Ejemplo 6) o secuencias de aminoácidos que puedan cumplir el papel de una CDR3 en los marcos conservados aceptores usados en el presente documento. Como ejemplo, se diseñaron cebadores oligonucleotídicos específicos para regiones flanqueantes de la secuencia de ADN que codifica la CDR H3 de inmunoglobulinas, es decir, específicos para las FR3 y FR4 de la región variable. También pueden diseñarse cebadores oligonucleotídicos específicos para 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' un sitio para una enzima de restricción de Tipo IIs mientras que su parte 3' coincide con la secuencia de ADN diana.

En algunas realizaciones, el cebador es un ácido nucleico seleccionado del grupo que consiste en SEC ID N°: 120-254.

Los métodos proporcionados en el presente documento usan enzimas de restricción de Tipo IIs, tales como, por ejemplo, FokI, para insertar secuencias de CDR naturales, tales como, por ejemplo, secuencias de CDR H3 o CDR L3 naturales en los Marcos conservados Aceptores descritos en el presente documento. Los métodos proporcionados en el presente documento usan enzimas de restricción de Tipo IIs, tales como, por ejemplo, FokI, para insertar secuencias de CDR sintéticas, tales como, por ejemplo, secuencias de CDR H3 o CDR L3 sintéticas en los Marcos conservados Aceptores descritos en el presente documento. Los métodos proporcionados en el presente documento usan enzimas de restricción de Tipo IIs, tales como, por ejemplo, FokI, para insertar secuencias de aminoácidos que puedan cumplir el papel de una región CDR deseada, tal como, por ejemplo, una secuencia de aminoácidos que pueda cumplir el papel de una región CDR H3 o CDR L3 natural o sintética en los Marcos conservados Aceptores descritos en el presente documento. Las enzimas de restricción de Tipo IIs son enzimas que escinden fuera de su secuencia de reconocimiento en un lateral. Estas enzimas son de tamaño intermedio, típicamente 400-650 aminoácidos de longitud, y reconocen secuencias que son continuas y asimétricas. Se describen enzimas de restricción de Tipo IIs adecuadas, también conocidas como endonucleasas de restricción de Tipo IIs, y las secuencias que identifican, por ejemplo, en Szybalski *et al.*, “Class-IIS Restriction Enzymes - a Review.” *Gene*, vol. 100: 13-26 (1991).

Las Bibliotecas Primarias incluyen un Marco conservado Aceptor VH y una secuencia VL fija (también denominada secuencia “VL de prueba”) o un Marco conservado Aceptor VL y una secuencia VH fija (también denominada secuencia “VH de prueba”). Por lo tanto, las Bibliotecas Primarias muestran diversidad en solamente una de las cadenas pesadas o ligeras. Se generan Bibliotecas Secundarias ligando un Marco conservado Aceptor VH y un Marco conservado Aceptor VL entre sí (véase por ejemplo, Ejemplo 7). Las Bibliotecas Secundarias tienen diversidad en las cadenas tanto pesadas como ligeras.

La invención proporciona métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de cadena pesada de inmunoglobulina humana que contiene una pluralidad de región determinante de complementariedad de cadena pesada 3 (CDR H3) aislada del repertorio de dominio variable de inmunoglobulina de una especie de mamífero o una especie no humana inmunizada. El método incluye las etapas de: (a) proporcionar una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor que codifican dominios variables de cadena pesada de inmunoglobulina humana distintos, conteniendo cada secuencia de ácido

nucleico de Marco conservado Aceptor una primera región marco conservada (FR1), una segunda región marco conservada (FR2), una tercera región marco conservada (FR3) y una cuarta región marco conservada (FR4), en la que las regiones FR1 y FR2 están intercaladas con una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas con una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas con una secuencia de ácido nucleico de relleno que incluye al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs intercalados con una secuencia de ácido nucleico aleatoria; (b) proporcionar una pluralidad de secuencias de ácido nucleico diversificadas que codifican secuencias de región determinante de complementariedad de cadena pesada 3 (CDR H3) aisladas de un repertorio de inmunoglobulina de especie no humana en el que cada una de la pluralidad de secuencias de ácido nucleico diversificadas incluye un sitio de reconocimiento de enzimas de restricción de Tipo IIs en cada extremo; (c) digerir cada una de la pluralidad de secuencias de ácido nucleico que codifican las regiones CDR H3 usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzima de restricción de Tipo IIs de la etapa (b) y digerir la secuencia de ácido nucleico de relleno de la etapa (a) del Marco conservado Aceptor usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzima de restricción de Tipo IIs de la etapa (a); y (d) ligar las secuencias de ácido nucleico digeridas que codifican las regiones CDR H3 o las secuencias de aminoácidos de la etapa (c) en el Marco conservado Aceptor digerido de la etapa (c) de modo que las regiones FR3 y FR4 estén intercaladas con las secuencias de ácido nucleico que codifican la región CDR H3 o la secuencia de aminoácidos que pueda cumplir el papel de una región CDR3 y se restauran secuencias que codifican un dominio variable de inmunoglobulina completo que no contienen los sitios de reconocimiento de enzimas de restricción de Tipo IIs de las etapas (a) y (b).

En algunas realizaciones, la etapa (b) como se ha expuesto anteriormente se realiza amplificando la secuencia de CDR H3 de una especie no humana usando cebadores oligonucleotídicos que contienen un sitio de restricción de Tipo IIs. En algunas realizaciones, se realiza la etapa (b) como se ha expuesto anteriormente amplificando la secuencia de CDR H3 de una especie no humana usando cebadores oligonucleotídicos que contienen un sitio de restricción FokI IIs. En algunas realizaciones, la especie no humana es un primate no humano, roedor, canino, felino, oveja, cabra, vaca, caballo o cerdo.

En algunas realizaciones, los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y la etapa (b) se reconocen por la misma enzima de restricción de Tipo IIs. En algunas realizaciones, los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y la etapa (b) se reconocen por diferentes enzimas de restricción de Tipo IIs. Por ejemplo, los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de FokI, sitios de reconocimiento de BsaI y/o sitios de reconocimiento de Bsmbl.

En algunas realizaciones, la secuencia de ácido nucleico del Marco conservado Aceptor deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia de gen variable de cadena pesada humana o una secuencia derivada de una secuencia de gen variable de cadena pesada humana. En algunas realizaciones, la secuencia de gen variable de cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23 y VH5-51. En algunas realizaciones, la secuencia humana es una secuencia de gen variable de cadena ligera kappa humana o una secuencia derivada de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo, la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20. En algunas realizaciones, la secuencia humana es una secuencia de gen variable de cadena ligera lambda humana o una secuencia derivada de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la secuencia de gen variable de cadena ligera lambda humana se selecciona de VL1-44 y VL1-51.

En una realización, la pluralidad de ácidos nucleicos diversificados codifica regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye secuencias de origen natural o secuencias derivadas de animales inmunizados.

En una realización, la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias seleccionadas de secuencias de CDR3 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificados de forma natural.

En una realización, la pluralidad de ácidos nucleicos diversificados codifica regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se han identificado como expuestos a un antígeno particular.

En una 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.

En otra 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.

En algunas realizaciones, la pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor incluye una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable.

- 5 En algunas realizaciones, los métodos proporcionados incluyen la etapa adicional de (e) transformar el vector de expresión de la etapa (d) en una célula hospedadora y cultivar la célula hospedadora en condiciones suficientes para expresar la pluralidad de secuencias de Marco conservado Aceptor. Por ejemplo, la célula hospedadora es *E. coli*. En algunas realizaciones, el vector de expresión es un vector fagémido. Por ejemplo, el vector fagémido es pNDS1.
- 10 También se describen métodos para realizar un anticuerpo específico de diana, región variable de anticuerpo o una parte de la misma, (a) proporcionando una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor que codifican distintos dominios variables de inmunoglobulina, incluyendo cada secuencia de ácido nucleico de Marco conservado Aceptor una primera región marco conservada (FR1), una segunda región marco conservada (FR2), una tercera región marco conservada (FR3) y una cuarta región marco conservada (FR4), en la que las regiones FR1 y FR2 están intercaladas con una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas con una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas con una secuencia de ácido nucleico de relleno que incluye al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs intercalados con una secuencia de ácido nucleico aleatoria; (b) proporcionar una pluralidad de secuencias de ácido nucleico diversificadas que codifican secuencias de regiones de región determinante de complementariedad 3 (CDR3) o que codifican secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, en las que cada una de la pluralidad de secuencias de ácido nucleico diversificadas incluye un sitio de reconocimiento de enzima de restricción de Tipo IIs en cada extremo; (c) digerir cada una de la pluralidad de secuencias de ácido nucleico que codifican las regiones CDR3 o secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzima de restricción de Tipo IIs de la etapa (b) y digerir la secuencia de ácido nucleico de relleno de la etapa (a) usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de la enzima de restricción de Tipo IIs de la etapa (a); (d) clonar las secuencias de ácido nucleico digeridas que codifican las regiones 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 ácido nucleico digeridas que codifican las regiones CDR3 o las secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 de la etapa (c) en el Marco conservado Aceptor de modo que las regiones FR3 y FR4 estén intercaladas con las secuencias de ácido nucleico 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 inmunoglobulina completo; (e) transformar la expresión del vector de la etapa (e) en una célula hospedadora y cultivar la célula hospedadora en condiciones suficientes para expresar la pluralidad de secuencias Marco conservadas Aceptoras; (f) poner en contacto la célula hospedadora con un antígeno diana; y (g) determinar qué secuencias Marco conservadas Aceptoras expresadas se unen con el antígeno diana.
- 40 Opcionalmente, los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y la etapa (b) se reconocen por la misma enzima de restricción de Tipo IIs. En algunas realizaciones, los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y la etapa (b) se reconocen por diferentes enzimas de restricción de Tipo IIs. Por ejemplo, los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de FokI, sitios de reconocimiento de BsaI y/o sitios de reconocimiento de Bsmbl.
- 45 Opcionalmente, la secuencia de ácido nucleico del Marco conservado Aceptor deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia de gen variable de cadena pesada humana o una secuencia derivada de una secuencia de gen variable de cadena pesada humana. Opcionalmente, la secuencia de gen variable de cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23 y VH5-51. En algunas realizaciones, la secuencia humana es una secuencia de gen variable de cadena ligera kappa humana o una secuencia derivada de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo, la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20. Opcionalmente, la secuencia humana es una secuencia de gen variable de cadena ligera lambda humana o una secuencia derivada de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la secuencia de gen variable de cadena ligera lambda humana se selecciona de VL1-44 y VL1-51.
- 55 Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye secuencias de origen natural o secuencias derivadas de animales inmunizados.
- 60 Opcionalmente, la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias seleccionadas de secuencias de CDR3 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificadas de forma natural.
- 65 Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se ha

identificado que han estado expuestos a un antígeno particular.

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 secuencias de ácido nucleico de un Marco conservado Aceptor incluyen una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable.

Opcionalmente, el vector de expresión es un vector fagémido. Por ejemplo, el vector fagémido es pNDS1. Opcionalmente, la célula hospedadora *E. coli*.

Opcionalmente el método incluye la etapa adicional de (i) secuenciar las secuencias que codifican dominio variable de inmunoglobulina que se unen con el antígeno diana.

También se describen en el presente documento métodos para realizar un anticuerpo específico de diana, región variable de anticuerpo o una parte del mismo, (a) proporcionando una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor que codifican distintos dominios variables de inmunoglobulina, incluyendo cada secuencia de ácido nucleico de Marco conservado Aceptor una primera región marco conservada (FR1), una segunda región marco conservada (FR2), una tercera región marco conservada (FR3) y una cuarta región marco conservada (FR4), en las que las regiones FR1 y FR2 están intercaladas con una secuencia de ácido nucleico de relleno que incluye al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs intercaladas con una secuencia de ácido nucleico aleatoria, las regiones FR2 y FR3 están intercaladas con una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas con una región determinante de complementariedad 3 (CDR3); (b) proporcionando una pluralidad de secuencias de ácido nucleico diversificadas que codifican regiones de región determinante de complementariedad 1 (CDR1) o que codifican secuencias de aminoácidos que pueden cumplir el papel de una región CDR1, en las que cada una de la pluralidad de secuencias de ácido nucleico diversificadas incluye un sitio de reconocimiento de enzimas de restricción de Tipo IIs en cada extremidad; (c) digiriendo cada una de la pluralidad de secuencias de ácido nucleico que codifican las regiones CDR1 o secuencias de aminoácidos que puedan cumplir el papel de una región CDR1 usando una enzima de restricción de Tipo IIs que se una con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (b) y digerir la secuencia de ácido nucleico de relleno de la etapa (a) usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a); (d) clonando las secuencias de ácido nucleico 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 ligando las secuencias de ácido nucleico 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 el Marco conservado Aceptor de modo que las regiones FR1 y FR2 estén intercaladas con las secuencias de ácido nucleico que codifican la región CDR1 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR1 y se restaura una secuencia que codifica un gen variable de inmunoglobulina completo; (e) transformando el vector de expresión de la etapa (e) en una célula hospedadora y cultivar la célula hospedadora en condiciones suficientes para expresar la pluralidad de secuencias Marco conservadas Aceptoras; (f) poniendo en contacto la célula hospedadora con un antígeno diana; y (g) determinando qué secuencias Marco conservadas Aceptoras expresadas se unen con el antígeno diana.

Opcionalmente, los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y la etapa (b) se reconocen por la misma enzima de restricción de Tipo IIs. En algunas realizaciones, los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y la etapa (b) se reconocen por diferentes enzimas de restricción de Tipo IIs. Por ejemplo, los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de FokI, sitios de reconocimiento de BsaI y/o sitios de reconocimiento de Bsmbl.

Opcionalmente, la secuencia de ácido nucleico del Marco conservado Aceptor deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia de gen variable de cadena pesada humana o una secuencia derivada de una secuencia de gen variable de cadena pesada humana. Opcionalmente, la secuencia de gen variable de 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 de gen variable de cadena ligera kappa humana o una secuencia derivada de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo, la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20. Opcionalmente, la secuencia humana es una secuencia de gen variable de cadena ligera lambda humana o una secuencia derivada de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la secuencia de gen variable de cadena ligera lambda humana se selecciona de VL1-44 y VL1-51.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica regiones CDR1, y la pluralidad de ácidos nucleicos diversificados incluye secuencias de origen natural o secuencias derivadas de animales inmunizados.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias seleccionadas de secuencias de CDR1 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de

origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificados de forma natural.

Opcionalmente, la pluralidad de ácidos nucleicos diversificados codifica regiones CDR1, y la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se ha identificado que han estado expuestos a un antígeno particular.

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.

Opcionalmente, la pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor incluyen una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable.

Opcionalmente, el vector de expresión es un vector fagémido. Por ejemplo, el vector fagémido es pNDS1. En algunas realizaciones, la célula hospedadora es *E. coli*.

Opcionalmente el método incluye la etapa adicional de (i) secuenciar las secuencias codificantes de dominio variable de inmunoglobulina que se unen con el antígeno diana.

También se desvelan métodos para preparar un anticuerpo específico de diana, región variable de anticuerpo o una parte de la misma, (a) proporcionando una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor que codifican distintos dominios variables de inmunoglobulina, incluyendo cada secuencia de ácido nucleico de Marco conservado Aceptor una primera región marco conservada (FR1), una segunda región marco conservada (FR2), una tercera región marco conservada (FR3) y una cuarta región marco conservada (FR4), en la que las regiones FR1 y FR2 están intercaladas con una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas con una secuencia de ácido nucleico de relleno que incluye al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs intercaladas con una secuencia de ácido nucleico aleatoria, y las regiones FR3 y FR4 están intercaladas con una región determinante de complementariedad 3 (CDR3); (b) proporcionando una pluralidad de secuencias de ácido nucleico diversificadas que codifican regiones de región determinante de complementariedad 2 (CDR2) o que codifican secuencias de aminoácidos que pueden cumplir el papel de una región CDR2, en las que cada una de la pluralidad de secuencias de ácido nucleico diversificadas incluye un sitio de reconocimiento de enzimas de restricción de Tipo IIs en cada extremo; (c) digiriendo cada una de la pluralidad de secuencias de ácido nucleico que codifican las regiones CDR2 o secuencias de aminoácidos que pueden cumplir el papel de una región CDR2 usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (b) y digiriendo la secuencia de ácido nucleico de relleno de la etapa (a) del Marco conservado Aceptor usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a); (d) ligando las secuencias de ácido nucleico 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 el Marco conservado Aceptor digerido de la etapa (c) de modo que las regiones FR2 y FR3 están intercaladas con las secuencias de ácido nucleico que codifican la región CDR2 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR2 y se restauran secuencias codificantes de dominio variable de inmunoglobulina completo que no contienen los sitios de reconocimiento de enzimas de restricción de Tipo IIs de las etapas (a) y (b); (e) clonando la biblioteca de ácidos nucleicos que codifican dominios variables de inmunoglobulina de la etapa (d) en un vector de expresión; (f) transformando el vector de expresión de la etapa (e) en una célula hospedadora y cultivando la célula hospedadora en condiciones suficientes para expresar una pluralidad de dominios variables de inmunoglobulina codificados por la biblioteca; (g) poniendo en contacto la pluralidad de dominios variables de inmunoglobulina de la etapa (f) con un antígeno diana; y (h) determinando qué secuencias codificantes de dominio variable de inmunoglobulina expresadas se unen con el antígeno diana.

Opcionalmente, los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y la etapa (b) se reconocen por la misma enzima de restricción de Tipo IIs. En algunas realizaciones, los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y la etapa (b) se reconocen por diferentes enzimas de restricción de Tipo IIs. Por ejemplo, los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de FokI, sitios de reconocimiento de BsaI y/o sitios de reconocimiento de Bsmbl.

Opcionalmente, la secuencia de ácido nucleico del Marco conservado Aceptor deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia de gen variable de cadena pesada humana o una secuencia derivada de una secuencia de gen variable de cadena pesada humana. Opcionalmente, la secuencia de gen variable de 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 de gen variable de cadena ligera kappa humana o una secuencia derivada de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo, la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20. Opcionalmente, la secuencia humana es una secuencia de gen variable de cadena ligera lambda humana o una secuencia derivada de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la secuencia

de gen variable de cadena ligera lambda humana se selecciona de VL1-44 y VL1-51.

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

En un caso, la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias seleccionadas de secuencias de CDR2 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificados de forma natural.

En un caso, la pluralidad de ácidos nucleicos diversificados codifica regiones CDR2, y la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se ha identificado que han estado expuestos a un antígeno particular.

En un caso, 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.

En otro caso, 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.

En algunos ejemplos, la pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor incluye una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable.

Opcionalmente, la célula hospedadora es *E. coli*. En algunas realizaciones, 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 las secuencias que codifican dominio variable de inmunoglobulina que se unen con el antígeno diana.

También se desvelan métodos para producir una biblioteca de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina. Estos métodos incluyen las etapas de (a) proporcionar una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor de Ig en las que ha introducido una fuente de diversidad en una única región determinante de complementariedad (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), en las que la secuencia de Marco conservado Aceptor de Ig incluye una secuencia de ácido nucleico de relleno que incluye al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs, y en las que la fuente de diversidad es una CDR seleccionada de secuencias de CDR de origen natural que contienen sitios de reconocimiento de enzimas de restricción de Tipo IIs fuera de la región CDR, (b) introducir la fuente de diversidad dentro de cada Marco conservado Aceptor de Ig dirigiendo tanto la fuente de diversidad como los Marcos conservados Aceptores de Ig con una enzima de restricción de Tipo IIs, y (c) ligar la fuente de diversidad digerida en el Marco conservado Aceptor de Ig de modo que se restauren secuencias codificantes de un dominio variable de inmunoglobulina completo que no contengan los sitios de reconocimiento de enzimas de restricción de Tipo IIs de las etapas (a) y (b).

Las secuencias de región CDR de origen natural están sustancialmente inalteradas a partir de su tipo silvestre, es decir, estado natural. Estas secuencias de región CDR de origen natural están flanqueadas por secuencias de aminoácidos que se han modificado técnicamente (o se han manipulado de otro modo artificialmente) para contener dos sitios de reconocimiento de enzimas de restricción de Tipo IIs, con un sitio de reconocimiento de enzimas de restricción de Tipo IIs en cada uno de los laterales de la secuencia de región CDR de origen natural. Los sitios de reconocimiento de enzimas de restricción de Tipo IIs están fuera de la región codificante de CDR. La secuencia de regiones CDR está inalterada en los extremos de la región codificante de CDR, las enzimas de restricción reconocen y cortan en una región que está hasta el límite de la región codificante de CDR, pero no cortan dentro de la región codificante de CDR.

Opcionalmente, los sitios de reconocimiento de enzimas de restricción de Tipo IIs dentro de las secuencias de ácido nucleico de relleno y que flanquean las secuencias de CDR de origen natural se reconocen por la misma enzima de restricción de Tipo IIs. Opcionalmente, los sitios de reconocimiento de enzimas de restricción de Tipo IIs dentro de las secuencias de ácido nucleico de relleno y que flanquean las secuencias de CDR de origen natural se reconocen por enzimas de restricción de Tipo IIs diferentes. Por ejemplo, los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de FokI, sitios de reconocimiento de BsaI y/o sitios de reconocimiento de Bsmbl.

En algunos casos, la secuencia de ácido nucleico de Marco conservado Aceptor de Ig deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia de gen variable de cadena pesada humana o

- una secuencia derivada de una secuencia de gen variable de cadena pesada humana. En algunos casos, la secuencia de gen variable de 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 de gen variable de cadena ligera kappa humana o una secuencia derivada de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo,
- 5 la secuencia de gen variable de cadena ligera 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 de gen variable de cadena ligera lambda humana o una secuencia derivada de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la secuencia de gen variable de cadena ligera lambda humana se selecciona de VL1-44 y VL1-51.
- 10 Opcionalmente, el conjunto de ácidos nucleicos de origen natural incluye o deriva de secuencias seleccionadas de secuencias de CDR3 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificadas de forma natural.
- 15 Opcionalmente, el conjunto de ácidos nucleicos de origen natural codifica regiones CDR3, y el conjunto de ácidos nucleicos de origen natural incluye secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se ha identificado que han estado expuestos a un antígeno particular.
- 20 Opcionalmente, el conjunto de ácidos nucleicos de origen natural incluye o deriva de secuencias seleccionadas de secuencias de CDR1 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificadas de forma natural.
- 25 Opcionalmente, el conjunto de ácidos nucleicos de origen natural codifica regiones CDR1, y el conjunto de ácidos nucleicos de origen natural incluye o deriva de secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se ha identificado que han estado expuestos a un antígeno particular.
- 30 En algunos casos, el conjunto de ácidos nucleicos de origen natural incluye o deriva de secuencias seleccionadas de secuencias de CDR2 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificados de forma natural.
- 35 En algunos casos, el conjunto de ácidos nucleicos de origen natural codifica regiones CDR2, y el conjunto de ácidos nucleicos de origen natural incluye secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se ha identificado que han estado expuestos a un antígeno particular.
- 40 En algunos casos, la pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor de Ig incluye una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable (VL).
- 45 En algunos casos, los métodos descritos incluyen las etapas adicionales de (e) clonar la biblioteca de ácidos nucleicos que codifican dominios variables de inmunoglobulina de la etapa (d) en un vector de expresión y (f) transformar el vector de expresión de la etapa (e) en una célula hospedadora y cultivar la célula hospedadora en condiciones suficientes para expresar una pluralidad de dominios variables de inmunoglobulina codificados por la biblioteca. Por ejemplo, la célula hospedadora es *E. coli*. Opcionalmente, el vector de expresión es un vector fagémido. Por ejemplo, el vector fagémido es pNDS1.
- 50 También se describen métodos para producir una biblioteca de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina. Estos métodos incluyen las etapas de (a) proporcionar una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor de Ig en las que se introduce una fuente de diversidad en una única región determinante de complementariedad (CDR) seleccionada del grupo que consiste
- 55 en región determinante de complementariedad 1 (CDR1), región determinante de complementariedad 2 (CDR2) y región determinante de complementariedad 3 (CDR3), en las que la secuencia de Marco conservado Aceptor de Ig incluye una secuencia de ácido nucleico de relleno que incluye al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs, y en las que la fuente de diversidad es una CDR seleccionada de secuencias de CDR producidas de forma sintética que contienen sitios de reconocimiento de enzimas de restricción de Tipo IIs fuera de
- 60 la región CDR, (b) introducir la fuente de diversidad dentro de cada Marco conservado Aceptor de Ig digiriendo tanto la fuente de diversidad como el Marco conservado Aceptor de Ig con una enzima de restricción de Tipo IIs; y (c) ligar la fuente digerida de diversidad en el Marco conservado Aceptor de Ig de modo que se restauren secuencias codificantes de dominio variable de inmunoglobulina completo que no contienen los sitios de reconocimiento de enzimas de restricción de Tipo IIs de las etapas (a) y (b).
- 65

Opcionalmente, los sitios de reconocimiento de enzimas de restricción de Tipo IIs dentro de las secuencias de ácido nucleico de relleno y las secuencias de CDR producidas de forma sintética se reconocen por la misma enzima de restricción de Tipo IIs. Opcionalmente, los sitios de reconocimiento de enzimas de restricción de Tipo IIs dentro de las secuencias de ácido nucleico de relleno y las secuencias de CDR producidas de forma sintética se reconocen por diferentes enzimas de restricción de Tipo IIs. Por ejemplo, los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de FokI, sitios de reconocimiento de BsaI y/o sitios de reconocimiento de Bsmbl.

Opcionalmente, la secuencia de ácido nucleico de Marco conservado Aceptor de Ig deriva de una secuencia humana. Por ejemplo, la secuencia humana es una secuencia de gen variable de cadena pesada humana o una secuencia derivada de una secuencia de gen variable de cadena pesada humana. Opcionalmente, la secuencia de gen variable de 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 de gen variable de cadena ligera kappa humana o una secuencia derivada de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo, la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20. Opcionalmente, la secuencia humana es una secuencia de gen variable de cadena ligera lambda humana o una secuencia derivada de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la secuencia de gen variable de cadena ligera 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.

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 ácido nucleico de Marco conservado Aceptor de Ig incluye una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable (VL).

Opcionalmente, los métodos proporcionados incluyen las etapas adicionales de (c) clonar la biblioteca de ácidos nucleicos que codifican dominios variables de inmunoglobulina de la etapa (d) en un vector de expresión y (f) transformar el vector de expresión de la etapa (e) en una célula hospedadora y cultivar la célula hospedadora en condiciones suficientes para expresar una pluralidad de dominio variable de inmunoglobulina codificado por la biblioteca. Por ejemplo, la célula hospedadora es *E. coli*. Opcionalmente, el vector de expresión es un vector fagémido. Por ejemplo, el vector fagémido es pNDS1.

También se describen en el presente documento métodos para preparar un polipéptido de inmunoglobulina. Estos métodos incluyen las etapas de (a) proporcionar una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor de Ig en las que se introduce una fuente de diversidad en una única región determinante de complementariedad (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), en las que la secuencia de Marco conservado Aceptor de Ig incluye una secuencia de ácido nucleico de relleno que incluye al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs, y en las que la fuente de diversidad es una CDR seleccionada de secuencias de CDR de origen natural que contienen sitios de reconocimiento de enzimas de restricción de Tipo IIs fuera de la región CDR, (b) introducir la fuente de diversidad dentro de cada Marco conservado Aceptor de Ig digiriendo tanto la fuente de diversidad como los Marcos conservados Aceptores de Ig con una enzima de restricción de Tipo IIs; y (c) ligar la fuente de diversidad digerida en el Marco conservado Aceptor de Ig de modo que se restaure una secuencia codificante de gen variable de inmunoglobulina completa; y (d) clonar la secuencia codificante de gen variable de inmunoglobulina completa de la etapa (c) en un vector de expresión; y (e) transformar el vector de expresión de la etapa (d) en una célula hospedadora y cultivar la célula hospedadora en condiciones suficientes para expresar las secuencias codificantes de genes de inmunoglobulina completa que no tienen los sitios de reconocimiento de enzimas de restricción de Tipo IIs que se restauran.

Las secuencias de región CDR de origen natural están sustancialmente inalteradas con respecto a su tipo silvestre, es decir, estado natural. Estas secuencias de región CDR de origen natural están flanqueadas por secuencias de aminoácidos que se han modificado técnicamente (o se han manipulado artificialmente de otro modo) para contener dos sitios de reconocimiento de enzimas de restricción de Tipo IIs, con un sitio de reconocimiento de enzimas de restricción de Tipo IIs en cada lado de la secuencia de región CDR de origen natural.

Opcionalmente, los sitios de reconocimiento de enzimas de restricción de Tipo IIs dentro de las secuencias de ácido nucleico de relleno y que flanquean las secuencias de CDR de origen natural se reconocen por la misma enzima de restricción de Tipo IIs. Como alternativa, los sitios de reconocimiento de enzimas de restricción de Tipo IIs dentro de

las secuencias de ácido nucleico de relleno y que flanquean las secuencias de CDR de origen natural se reconocen por diferentes enzimas de restricción de Tipo IIs. Por ejemplo, los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de FokI, sitios de reconocimiento de BsaI y/o sitios de reconocimiento de Bsmbl.

Opcionalmente, la secuencia de ácido nucleico de Marco conservado Aceptor deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia de gen variable de cadena pesada humana o una secuencia derivada de una secuencia de gen variable de cadena pesada humana. Opcionalmente, la secuencia de gen variable de 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 de gen variable de cadena ligera kappa humana o una secuencia derivada de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo, la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20. Como alternativa, la secuencia humana es una secuencia de gen variable de cadena ligera lambda humana o una secuencia derivada de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la secuencia de gen variable de cadena ligera lambda humana se selecciona de VL1-44 y VL1-51.

Opcionalmente, el conjunto de ácidos nucleicos de origen natural incluye o deriva de secuencias seleccionadas de secuencias de CDR3 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificadas de forma natural.

Opcionalmente, el conjunto de ácidos nucleicos de origen natural codifica regiones CDR3, y el conjunto de ácidos nucleicos de origen natural incluye secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se ha identificado que han estado expuestos a un antígeno particular.

Opcionalmente, el conjunto de ácidos nucleicos de origen natural incluye o deriva de secuencias seleccionadas de secuencias de CDR1 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificadas de forma natural.

Opcionalmente, el conjunto de ácidos nucleicos de origen natural codifica regiones CDR1, y el conjunto de ácidos nucleicos de origen natural incluye o deriva de secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se ha identificado que han estado expuestos a un antígeno particular.

Opcionalmente, el conjunto de ácidos nucleicos de origen natural incluye o deriva de secuencias seleccionadas de secuencias de CDR2 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificados de forma natural.

Opcionalmente, el conjunto de ácidos nucleicos de origen natural codifica regiones CDR2, y el conjunto de ácidos nucleicos de origen natural incluye secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se ha identificado que han estado expuestos a un antígeno particular.

Opcionalmente, la pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor incluye una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable.

Opcionalmente, el vector de expresión es un vector fagémido. En algunas realizaciones, la célula hospedadora es *E. coli*.

Opcionalmente, el método también incluye las etapas de poner en contacto la célula hospedadora con un antígeno diana, y determinar qué secuencias codificantes de genes variables de Ig completa expresadas se unen con el antígeno diana, identificando de este modo anticuerpos específicos de diana, regiones variables de anticuerpo o partes de los mismos. Opcionalmente, el método incluye la etapa adicional de (i) secuenciar el dominio variable de inmunoglobulina que codifica secuencias que se unen con el antígeno diana.

También se desvelan en el presente documento métodos para preparar un polipéptido de inmunoglobulina. Estos métodos incluyen las etapas de (a) proporcionar una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor de Ig en las que se introduce una fuente de diversidad en una única región determinante de complementariedad (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), en las que la secuencia de Marco conservado Aceptor de Ig incluye una secuencia de ácido nucleico de relleno que incluye al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs, y en las que la

fente de diversidad es una CDR seleccionada de secuencias de CDR producidas de forma sintética que contienen sitios de reconocimiento de enzimas de restricción de Tipo IIs fuera de la región CDR, (b) introducir la fuente de diversidad dentro de cada Marco conservado Aceptor de Ig digiriendo tanto la fuente de diversidad como el Marco conservado Aceptor de Ig con una enzima de restricción de Tipo IIs; (c) ligar la fuente de diversidad digerida en el Marco conservado Aceptor de Ig de modo que se restaure una secuencia codificante de gen variable de inmunoglobulina completa; (d) clonar el Marco conservado Aceptor de Ig ligado de la etapa (c) en un vector de expresión; y (e) transformar el vector de expresión de la etapa (d) en una célula hospedadora y cultivar la célula hospedadora en condiciones suficientes para expresar las secuencias codificantes de gen de inmunoglobulina completa que no contienen los sitios de reconocimiento de enzimas de restricción de Tipo IIs que se restauran.

Opcionalmente, los sitios de reconocimiento de enzimas de restricción de Tipo IIs dentro de las secuencias de ácido nucleico de relleno y las secuencias de CDR producidas de forma sintética se reconocen por la misma enzima de restricción de Tipo IIs. Opcionalmente, los sitios de reconocimiento de enzimas de restricción de Tipo IIs dentro de las secuencias de ácido nucleico de relleno y las secuencias de CDR producidas de forma sintética se reconocen por diferentes enzimas de restricción de Tipo IIs. Por ejemplo, los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de FokI, sitios de reconocimiento de BsaI y/o sitios de reconocimiento de Bsmbl.

Opcionalmente, la secuencia de ácido nucleico de Marco conservado Aceptor de Ig deriva de una secuencia génica humana. Por ejemplo, la secuencia humana es una secuencia de gen variable de cadena pesada humana o una secuencia derivada de una secuencia de gen variable de cadena pesada humana. Opcionalmente, la secuencia de gen variable de 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 de gen variable de cadena ligera kappa humana o una secuencia derivada de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo, la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20. Opcionalmente, la secuencia humana es una secuencia de gen variable de cadena ligera lambda humana o una secuencia derivada de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la secuencia de gen variable de cadena ligera 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.

En algunos casos, la pluralidad de ácidos nucleicos diversificados codifica secuencias de aminoácidos que pueden cumplir el papel de una región CDR, y la pluralidad de ácidos nucleicos diversificados incluye secuencias sintéticas.

En algunos casos, 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.

En algunos casos, la pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor de Ig incluye una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable.

En algunos ejemplos, el vector de expresión es un vector fagémido. En algunas realizaciones, la célula hospedadora es *E. coli*.

Opcionalmente, el método también incluye las etapas de poner en contacto la célula hospedadora con un antígeno diana, y determinar qué secuencias codificantes de gen variable de Ig completa expresadas se unen con el antígeno diana, identificando de este modo anticuerpos específicos de diana, regiones variables de anticuerpo o partes de los mismos. Opcionalmente, el método incluye la etapa adicional de (i) secuenciar las secuencias codificantes del dominio variable de inmunoglobulina que se unen con el antígeno diana.

También se desvelan en el presente documento métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que incluye una pluralidad de secuencias de región determinante de complementariedad 3 (CDR3) aisladas por separado del repertorio de dominios variables de inmunoglobulina de una especie de mamífero. Además, se desvelan métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que incluye una pluralidad de secuencias de región determinante de complementariedad 2 (CDR2) aisladas por separado del repertorio de dominios variables de inmunoglobulina de una especie de mamífero.

También se desvelan métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que incluye una pluralidad de secuencias de región determinante de complementariedad 1 (CDR1) aisladas por separado del repertorio de dominio variable de inmunoglobulina de una especie de mamífero.

Se desvelan métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que incluye una pluralidad de secuencias de región determinante de

complementariedad 3 (CDR3) aisladas por separado del repertorio de dominios variables de inmunoglobulina de una especie de mamífero no humana. También se desvelan métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que incluye una pluralidad de secuencias de región determinante de complementariedad 2 (CDR2) aisladas por separado del repertorio de dominios variables de inmunoglobulina de una especie de mamífero. Se desvelan también métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que incluye una pluralidad de secuencias de región determinante de complementariedad 1 (CDR1) aisladas por separado del repertorio de dominios variables de inmunoglobulina de una especie de mamífero.

- 10 Opcionalmente, la especie no humana es primate no humano, roedor, canino, felino, oveja, cabra, vaca, caballo, un miembro de la familia Camelidae, llama, camello, dromedario o cerdo.

La invención proporciona métodos de acuerdo con la reivindicación 1 para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina que incluye una pluralidad de secuencias de región determinante de complementariedad 3 (CDR3) aisladas por separado del repertorio de dominios variables de inmunoglobulina humana de un ser humano. Además, se desvelan métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina que incluye una pluralidad de secuencias de región determinante de complementariedad 2 (CDR2) aisladas por separado del repertorio de dominios variables de inmunoglobulina de un ser humano. También se desvelan métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que incluye una pluralidad de secuencias de región determinante de complementariedad 1 (CDR1) aisladas por separado del repertorio de dominios variables de inmunoglobulina de un ser humano.

La invención proporciona métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que incluye una pluralidad de secuencias de región determinante de complementariedad 3 (CDR3) aisladas por separado del repertorio de dominios variables de inmunoglobulina de una especie no humana.

Estos métodos incluyen las etapas de (a) proporcionar una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor que codifican distintos dominios variables de inmunoglobulina humana, conteniendo cada secuencia de ácido nucleico de Marco conservado Aceptor una primera región marco conservada (FR1), una segunda región marco conservada (FR2), una tercera región marco conservada (FR3) y una cuarta región marco conservada (FR4), en las que las regiones FR1 y FR2 están intercaladas con una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas con una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas con una secuencia de ácido nucleico de relleno que comprende al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs intercalados con una secuencia de ácido nucleico aleatoria; (b) proporcionar una pluralidad de secuencias de ácido nucleico diversificadas que codifican secuencias de región determinante de complementariedad 3 (CDR3) del repertorio de inmunoglobulinas de especies de mamíferos en el que cada una de la pluralidad de secuencias de ácido nucleico diversificadas comprende un sitio de reconocimiento de enzimas de restricción de Tipo IIs en cada extremo; (c) digerir cada una de la pluralidad de secuencias de ácido nucleico que codifican las regiones CDR3 usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (b) y que digiere la secuencia de ácido nucleico de relleno de la etapa (a) del Marco conservado Aceptor usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzima de restricción de Tipo IIs de la etapa (a); y (d) ligar las secuencias de ácido nucleico digeridas que codifican las regiones CDR3 de las secuencias de aminoácidos de la etapa (c) en el Marco conservado Aceptor digerido de la etapa (c) de modo que las regiones FR3 y FR4 estén intercaladas con las secuencias de ácido nucleico que codifican la región CDR3 de la secuencia de aminoácidos que puede cumplir el papel de una región CDR3 y se restauran secuencias codificantes de dominio variable de inmunoglobulina completa que no contienen los sitios de reconocimiento de enzimas de restricción de Tipo IIs de las etapas (a) y (b).

En algunas realizaciones la etapa (b) se realiza amplificando la secuencia de CDR3 de una especie de mamífero usando cebadores oligonucleotídicos que contienen un sitio de restricción de Tipo IIs. En algunas realizaciones, el cebador oligonucleotídico se diseña para potenciar la compatibilidad entre la secuencia de CDR3 de mamífero y el Marco conservado Aceptor que codifica un dominio variable de inmunoglobulina humana. En algunas realizaciones, el cebador oligonucleotídico se diseña para modificar la secuencia en los límites de las secuencias de CDR3 de mamífero para permitir un ligamiento eficaz mediante extremos cohesivos compatibles en el Marco conservado Aceptor que codifica un dominio variable de inmunoglobulina humana. En algunas realizaciones las secuencias de ADN de mamífero que flanquean las regiones CDR3 podrían, tras la escisión por enzimas de restricción de Tipo IIs, no generar extremos cohesivos compatibles con los extremos cohesivos de los Marcos conservados Aceptores digeridos. En dichos casos los oligonucleótidos usados para amplificación se diseñan para modificar la secuencia del mamífero diana de modo que tras la escisión con una enzima de restricción de Tipo IIs, los extremos cohesivos sean compatibles y pueda producirse ligamiento eficaz.

En algunas realizaciones, la etapa (b) se realiza amplificando la secuencia de CDR3 de una especie no humana usando cebadores oligonucleotídicos que contienen un sitio de restricción IIs de FokI. Estas etapas también pueden

realizarse amplificando la secuencia de CDR2 de una especie de mamífero usando cebadores oligonucleotídicos que contienen un sitio de restricción IIs de FokI. Estas etapas también pueden realizarse amplificando la secuencia de CDR1 de una especie de mamífero usando cebadores oligonucleotídicos que contienen un sitio de restricción IIs de FokI.

En algunas realizaciones, los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y la etapa (b) se reconocen por una enzima de restricción de Tipo IIs diferente. En algunas realizaciones, los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de BsmBI, sitios de reconocimiento de BsaI, sitios de reconocimiento de FokI o una combinación de los mismos.

En algunas realizaciones, las secuencias de ácido nucleico diversificadas que codifican secuencias de CDR3 codifican secuencias de CDR3 de cadena pesada (CDR H3). En algunas realizaciones, las secuencias de ácido nucleico diversificadas que codifican secuencias de CDR3 codifican secuencias de CDR3 de cadena ligera (CDR L3).

En algunas realizaciones, la secuencia de ácido nucleico de Marco conservado Aceptor incluye o deriva de al menos una parte de una secuencia de gen variable de cadena pesada humana seleccionada de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23 y VH5-51. En algunas realizaciones, la secuencia de ácido nucleico de Marco conservado Aceptor incluye o deriva de al menos una parte de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo, la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20. En algunas realizaciones, la secuencia de ácido nucleico de Marco conservado Aceptor incluye o deriva de al menos una parte de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la secuencia de gen variable de cadena ligera lambda humana se selecciona de VL1-44 y VL1-51.

En algunas realizaciones, la pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor comprende una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable.

En algunas realizaciones, los métodos descritos en el presente documento también incluyen las etapas de (e) clonar la biblioteca de ácidos nucleicos que codifican dominios variables de inmunoglobulina de la etapa (d) en un vector de expresión y (f) transformar el vector de expresión de la etapa (e) en una célula hospedadora y cultivar la célula hospedadora en condiciones suficientes para expresar una pluralidad de dominios variables de inmunoglobulina codificados por la biblioteca. En algunas realizaciones, el vector de expresión es un fagémido o vector de fago. En algunas realizaciones, la célula hospedadora es *E. coli*.

La invención proporciona métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que incluye una pluralidad de secuencias de región determinante de complementariedad 3 (CDR3) aisladas por separado de dominios variables de inmunoglobulina de un mamífero no humano inmunizado o especie no humana. También se desvela la producción de una colección de ácidos nucleicos, en la que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que incluye una pluralidad de secuencias de región determinante de complementariedad 2 (CDR2) aisladas por separado de dominios variables de inmunoglobulina de un mamífero no humano inmunizado. La invención también proporciona métodos para producir una colección de ácidos nucleicos, en los que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que incluye una pluralidad de secuencias de región determinante de complementariedad 1 (CDR1) aisladas por separado de dominios variables de inmunoglobulina de un mamífero no humano inmunizado.

Opcionalmente, la especie no humana es primate no humano, roedor, canino, felino, oveja, cabra, vaca, caballo, un miembro de la familia Camelidae, llama, camello, dromedario o cerdo.

Opcionalmente, los métodos incluyen las etapas de (a) proporcionar una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor que codifican dominios variables de inmunoglobulina humana distintos, comprendiendo cada secuencia de ácido nucleico de Marco conservado Aceptor una primera región marco conservada (FR1), una segunda región marco conservada (FR2), una tercera región marco conservada (FR3) y una cuarta región marco conservada (FR4), en las que las regiones FR1 y FR2 están intercaladas con una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas con una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas con una secuencia de ácido nucleico de relleno que comprende al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs intercalados con una secuencia de ácido nucleico aleatoria; (b) proporcionar una pluralidad de secuencias de ácido nucleico diversificadas que codifican secuencias de región determinante de complementariedad 3 (CDR3) aisladas del mamífero no humano inmunizado en las que cada una de la pluralidad de secuencias de ácido nucleico diversificadas comprende un sitio de reconocimiento de enzimas de restricción de Tipo IIs en cada extremo; (c) digerir cada una de la pluralidad de secuencias de ácido nucleico que codifican las regiones CDR3 usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (b) y digerir la secuencia de ácido nucleico de relleno de la etapa (a) del Marco conservado Aceptor usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs

de la etapa (a); y (d) ligar las secuencias de ácido nucleico digeridas que codifican las regiones CDR3 o las secuencias de aminoácidos de la etapa (c) en el Marco conservado Aceptor digerido de la etapa (c) de modo que las regiones FR3 y FR4 estén intercaladas con las secuencias de ácido nucleico que codifican la región CDR3 o la secuencia de aminoácidos que puede cumplir el papel de una región CDR3 y se restauran secuencias codificantes de dominio variable de inmunoglobulina completa que no contienen los sitios de reconocimiento de enzimas de restricción de Tipo IIs de las etapas (a) y (b). Estas etapas también pueden realizarse usando una pluralidad de secuencias de ácido nucleico diversificadas que codifican secuencias de región determinante de complementariedad 2 (CDR2) aisladas del mamífero no humano inmunizado. Estas etapas también pueden realizarse usando una pluralidad de secuencias de ácido nucleico diversificadas que codifican secuencias de región determinante de complementariedad 1 (CDR1) aisladas del mamífero no humano inmunizado.

La etapa (b) puede realizarse amplificando la secuencia de CDR3 del mamífero no humano no humano inmunizado usando cebadores oligonucleotídicos que contienen un sitio de restricción de Tipo IIs. El cebador oligonucleotídico puede diseñarse para potenciar la compatibilidad entre la secuencia de CDR3 de mamífero y el Marco conservado Aceptor que codifica un dominio variable de inmunoglobulina humana. El cebador oligonucleotídico puede diseñarse para modificar la secuencia en los límites de las secuencias de CDR3 de mamífero para permitir el ligamiento eficaz mediante extremos cohesivos compatibles en el Marco conservado Aceptor que codifica un dominio variable de inmunoglobulina humana. Las secuencias de ADN de mamífero que flanquean las regiones CDR3 podrían, tras la escisión por enzimas de restricción de Tipo IIs, no generar extremos cohesivos compatibles con los extremos cohesivos de los Marcos conservados Aceptores digeridos. En dichos casos los oligonucleótidos usados para amplificación se diseñan para modificar la secuencia del mamífero diana de modo que después de escisión con una enzima de restricción de Tipo IIs, los extremos cohesivos sean compatibles y pueda producirse un ligamiento eficaz. Estas etapas también pueden realizarse amplificando la secuencia CDR2 del mamífero no humano inmunizado usando cebadores oligonucleotídicos que contienen un sitio de restricción de Tipo IIs. Estas etapas también pueden realizarse amplificando la secuencia de CDR1 del mamífero no humano inmunizado usando cebadores oligonucleotídicos que contienen un sitio de restricción de Tipo IIs.

La etapa (b) puede realizarse amplificando la secuencia de CDR H3 del mamífero no humano usando cebadores oligonucleotídicos que contienen un sitio de restricción IIs de FokI. Estas etapas también pueden realizarse amplificando la secuencia de CDR2 del mamífero no humano usando cebadores oligonucleotídicos que contienen un sitio de restricción IIs de FokI. Estas etapas también pueden realizarse amplificando la secuencia de CDR1 del mamífero no humano usando cebadores oligonucleotídicos que contienen un sitio de restricción IIs de FokI.

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

Opcionalmente, las secuencias de ácido nucleico diversificadas que codifican secuencias de CDR3 codifican secuencias de CDR3 de cadena pesada (CDR H3). Opcionalmente, las secuencias de ácido nucleico diversificadas que codifican secuencias de CDR3 codifican secuencias de CDR3 de cadena ligera (CDR L3). Opcionalmente, las secuencias de ácido nucleico diversificadas que codifican secuencias de CDR2 codifican secuencias de CDR2 de cadena pesada (CDR H2). Opcionalmente, las secuencias de ácido nucleico diversificadas que codifican secuencias de CDR2 codifican secuencias de CDR2 de cadena ligera (CDR L2). Opcionalmente, las secuencias de ácido nucleico diversificadas que codifican secuencias de CDR1 codifican secuencias de CDR1 de cadena pesada (CDR H1). Opcionalmente, las secuencias de ácido nucleico diversificadas que codifican secuencias de CDR1 codifican secuencias de CDR1 de cadena ligera (CDR L1).

Opcionalmente, la secuencia de ácido nucleico de Marco conservado Aceptor incluye o deriva de al menos una parte de una secuencia de gen variable de cadena pesada humana seleccionada de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23 y VH5-51.

Opcionalmente, la secuencia de ácido nucleico de Marco conservado Aceptor incluye o deriva de al menos una parte de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo, la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20. Como alternativa, la secuencia de ácido nucleico de Marco conservado Aceptor incluye o deriva de al menos una parte de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la secuencia de gen variable de cadena ligera lambda humana se selecciona de VL1-44 y VL1-51.

Opcionalmente, la pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor comprende una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable.

Opcionalmente, los métodos también incluyen las etapas de (e) clonar la biblioteca de ácidos nucleicos que codifican dominios variables de inmunoglobulina de la etapa (d) en un vector de expresión y (f) transformar el vector de expresión de la etapa (e) en una célula hospedadora y cultivar la célula hospedadora en condiciones suficientes para

expresar una pluralidad de dominios variables de inmunoglobulina codificados por la biblioteca. Opcionalmente, la célula hospedadora es *E. coli*. Opcionalmente, el vector de expresión es un vector fagémido o de fago.

Breve descripción de los dibujos

La Figura 1A es una representación esquemática de un dominio proteico con un Marco conservado y bucles que proporcionan restos de contacto con otra proteína o molécula. Se representan varias situaciones: un dominio de proteína estable con regiones de bucle plegadas de forma apropiada; bucles plegados de forma apropiada insertados en un dominio de estabilidad intrínseca limitada; un dominio proteico intrínsecamente estable cuya estabilidad se ve afectada por las regiones de bucle.

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

La Figura 2 es una representación esquemática de un Marco conservado Aceptor variable de anticuerpos. Se indican regiones Marco conservadas, CDR y sitios de restricción de 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 del beneficio de usar cebadores que contienen enzimas de restricción de Tipo IIS-RM para la amplificación e inserción de regiones CDR naturales en Marcos conservados Aceptores.

La Figura 5 es una ilustración que representa las secuencias de genes de línea germinal del dominio variable de cadena pesada y ligera seleccionados para la generación de Marcos conservados Aceptores.

La Figura 6 es una representación esquemática de una estrategia de amplificación usada para la generación de Marcos conservados Aceptores mediante la adición a las secuencias de línea germinal de un fragmento de relleno y una región FR4.

La Figura 7, panel superior, es una ilustración que representa el detalle de secuencia de fragmentos de Relleno de Marco conservado aceptor VH. Las secuencias de ADN reconocidas y escindidas por la enzima de restricción BsmBI se encuadran en rojo y negro respectivamente y se indican en el panel inferior de la figura. La fase de lectura correspondiente a la secuencia variable de anticuerpo está subrayada.

La Figura 8 es una ilustración que representa las secuencias de los 20 Marcos conservados Aceptores.

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

La Figura 10 es una tabla que representa las secuencias de secuencias de CDRH3 que se recuperaron de una fuente de ADNc humana y se insertaron en Marcos conservados Aceptores humanos.

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

La Figura 12 es una representación esquemática y detalle de secuencia de inserción de CDR sintética en un Marco conservado Aceptor.

La Figura 13 es una representación esquemática de bibliotecas Primarias y la recombinación de cadenas realizada para generar bibliotecas Secundarias.

La Figura 14 es una representación esquemática de la generación de bibliotecas de VH Aceptoras combinadas con bibliotecas sintéticas de VL y la captura de repertorios de CDRH3 de origen humano o no humano.

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

La Figura 16 es una serie de gráficas que representan la titulación de producción de fagos durante selección contra hIFN γ con las bibliotecas secundarias AD1 y AE1.

La Figura 17 es una serie de gráficas que representan la titulación de producción de fagos durante la selección contra el anticuerpo monoclonal 5E3 con las bibliotecas secundarias AD1 y AE1.

La Figura 18 es una serie de gráficas que representan la frecuencia de longitudes de CDR H3 halladas en las bibliotecas de AE1 y AD1 y después de tres ciclos 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 las CDR H3 son más largas de 16 aminoácidos, las secuencias de 70 pb suministradas por la plataforma de Secuenciación Illumina no abarcan suficiente secuencia de Marco conservado para identificar inequívocamente la familia de VH1 y por lo tanto la familia de VH se indica como indeterminada.

La Figura 19 es una serie de gráficas que representan ELISA de respuesta a dosis usando 6 preparaciones de scFv purificadas contra 5E3 de ratón o un anticuerpo de ratón irrelevante 1A6. Los siete clones codifican diferentes scFv. El clon A6 es un scFv específico para hIFN γ y se usó como un control negativo.

La Figura 20 es una gráfica que representa ELISA de respuesta a dosis usando preparaciones de scFv purificadas contra hIFN γ y en comparación con un scFv positivo específico para hIFN γ (A6).

La Figura 21 es una gráfica que representa el efecto inhibitorio de preparaciones de scFv purificadas en un ensayo de gen indicador de luciferasa conducido por hIFN γ . La actividad neutralizante de dos candidatos a scFv (AD1R4P1A9 y AE14R3P2E4) se comparó con la actividad de un scFv de control positivo (G9) y un scFv de

control negativo (D11).

La Figura 22 es una gráfica 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 dos candidatos a scFv (AD1R4P1A9 y AE14R3P2E4) se comparó con la actividad de un scFv de control negativo (D11).

La Figura 23 es una serie de gráficas que representan el efecto inhibitor de los dos candidatos AD1R4P1A9 y AE14R3P2E4 reformateados en IgG en un ensayo de gen indicador de luciferasa conducido por hIFN γ . La actividad neutralizante de dos IgG se comparó con la actividad de un IgG irrelevante dirigido contra RANTES humano (NI-0701).

La Figura 24 es una serie de gráficas que representan un ELISA de respuesta a dosis usando el IgG G11 y DA4 contra 5E3 de ratón, 5E3 de rata quimérico y los anticuerpos de isotipo de ratón y rata correspondientes.

La Figura 25 es una serie de gráficas que representan un ELISA para la detección de 5E3 de ratón en diferentes diluciones de suero de ratón usando los IgG antiidiotípicos G11 y DA4 como anticuerpos de captura.

La Figura 26 es una gráfica que representa las relaciones de producción/aportación de fagos durante la selección contra hIFN γ con las bibliotecas MnA y MiB.

La Figura 27 es una gráfica que representa las tasas de aciertos obtenidas en una exploración de ELISA de scFv con clones derivados de las bibliotecas de MnA, MiB y MiC después de cada ciclo de selección contra hIFN γ . El umbral se estableció a la mitad de la señal obtenida con el scFv de control A6.

La Figura 28 es una gráfica que representa la frecuencia de distribución de scFv que proporciona diferentes niveles de señal en experimentos de unión frente a hIFN γ obtenido con clones derivados de las bibliotecas MnA y MiB.

La Figura 29 es una gráfica que representa ELISA de respuesta a dosis usando preparaciones de scFv purificadas de clones derivados de las bibliotecas de MnA y MiB frente a hIFN γ y comparados con un scFv positivo específico para hIFN γ (A6).

La Figura 30 es una representación esquemática de métodos para generar bibliotecas de Aceptores que contienen dominios variables de cadena ligera diversificados funcionales y preseleccionados que pueden usarse directamente para la inserción de regiones CDRH3 capturadas.

La Figura 31 es una ilustración que representa oligonucleótidos que se diseñaron para sintetizar una colección de fragmentos de relleno que contenían dos sitios de restricción de BsmBI e introducir diversidad en uno o dos codones.

La Figuras 32 y 33 son ilustraciones que representan las secuencias oligonucleotídicas identificadas en los clones seleccionados.

Descripción detallada de la invención

Son atractivas bibliotecas de proteínas sintéticas y en particular bibliotecas de anticuerpos sintéticos ya que es posible durante el proceso de generación de bibliotecas seleccionar los componentes básicos que componen estas proteínas sintéticas e incluyen características deseadas. Una limitación importante, sin embargo, es que la selección aleatoria de partes de estas proteínas sintéticas para generar una colección de variantes con frecuencia conduce a proteínas no funcionales y por lo tanto puede reducir 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 abarcar la diversidad teórica de tramos de aminoácidos aleatorios no puede abarcarse debido a limitaciones prácticas. Incluso con sistemas de presentación tales como presentación en ribosomas puede generarse una diversidad de 10^{13} a 10^{14} y tomarse muestras que pueden abarcar como máximo la selección aleatoria completa de tramos de 9 aminoácidos. Como el tamaño medio de CDR H3 natural (también denominada en el presente documento CDR3 de cadena pesada o VH CDR3) es de más de 9 y puede ser de más de 20 aminoácidos de longitud, la diversidad sintética no es un enfoque practicable para generar dichas CDR.

La combinación de métodos usados en general para manipulación de ADN y que se usan en el transcurso de la generación de una biblioteca de variantes proteicas introduce errores en las secuencias de ADN. Estos errores pueden conducir a alteraciones en la fase de lectura del ADN que ya no codifica un polipéptido funcional. Típicamente, las bibliotecas de anticuerpos generadas usando ensamblaje de fragmentos de ADN por PCR y/o clonación de restricción contienen entre el 15 % y 45 % de secuencias que no están en la fase de lectura correcta para la traducción de proteínas. Estos miembros de bibliotecas no funcionales pueden comprometer la eficacia del proceso de selección e identificación de anticuerpos y por lo tanto se reconocen como una limitación del campo. Los métodos descritos permiten una introducción más robusta de diversidad en una biblioteca de anticuerpos usando una estrategia de clonación alternativa. Típicamente, la frecuencia de secuencias en fase es de aproximadamente el 90 %. Otra ventaja de la invención es que combina marcos conservados variables de anticuerpos aceptores seleccionados con bucles de CDR que tienen una alta probabilidad de plegamiento incorrecto. Permite la captura de CDR largas que son difíciles de abarcar con enfoques de selección aleatoria sintéticos. Además los métodos descritos no emplean ninguna modificación dentro de la región codificante de anticuerpo aceptor variable para clonación de las secuencias diversificadas. Otra ventaja de este método es que varias fuentes de diversidad puede capturarse en el mismo conjunto de marcos conservados de anticuerpos aceptores. Estas fuentes incluyen pero sin limitación: CDR de anticuerpos naturales de origen humano o de otro mamífero, CDR de anticuerpos de pollo, CDR de moléculas de tipo anticuerpo tales como VHH de camélidos, IgNAR de tiburones, bucles variables de receptores de linfocitos T. Además, pueden derivarse CDR naturales de animales sin tratamiento previo o inmunizados. En este

último caso, las CDR recuperadas se enriquecen en secuencias que estaban implicadas en el reconocimiento del antígeno usado para inmunización.

Una característica única de los métodos descritos en el presente documento es la captura eficaz de secuencias codificantes de CDR3 de cadena pesada de especies no humanas y su inserción en marcos conservados de inmunoglobulina humana. Usando estos métodos, es por lo tanto posible generar diferentes sitios de combinación de anticuerpos que se moldean por el repertorio de CDRH3 capturado de la especie y permitir la toma de muestras de un espacio tridimensional diferente. Estos métodos permiten la generación de anticuerpos humanos con nuevas especificidades que se dirigen a una serie diferente de clases y epítomos diana que los accesibles para un repertorio de CDRH3 humano. Además, estos nuevos anticuerpos codifican regiones marco conservadas humanas así como CDR1 y CDR2 y son por lo tanto adecuados para terapia humana.

En este método se modifican dominios de proteínas seleccionadas, como se ejemplifica por dominios variables de anticuerpo, introduciendo una secuencia de relleno que actuará como un sitio de integración para secuencias diversificadas. Tras la integración, el fragmento de relleno se retira completamente, dejando de este modo intacta la región codificante de la proteína aceptora y los fragmentos de proteínas insertadas (es decir, las CDR). Este acontecimiento de integración está mediado por el uso de una enzima de restricción de Tipo IIs que reconoce un sitio definido en la secuencia de ADN pero escinde el ADN a una distancia definida de este sitio. Este enfoque tiene dos ventajas principales: (1) permite la digestión de marcos conservados aceptores sin afectar a sus secuencias codificantes (sin necesidad de modificar técnicamente sitios de restricciones silenciosos); y (2) permite la digestión y clonación de secuencias diversificadas de forma natural que por definición no poseen sitios de restricción compatibles.

Como se ha descrito anteriormente, intentos anteriores para generar bibliotecas y/o presentaciones de secuencias de anticuerpo difieren de los métodos proporcionados en el presente documento. Por ejemplo, algunos métodos requieren el injerto de cada CDR, como se describe por ejemplo en la Patente de Estados Unidos Nº 6.300.064, en la que se introducen técnicamente sitios de enzimas de restricción en el límite de cada CDR, no solamente la región CDR H3. En otros métodos, se amplifican y reordenan secuencias de CDR de fuentes naturales, como se describe, por ejemplo, en la Patente de Estados Unidos Nº 6.989.250. En algunos métodos, tales como los descritos en la Publicación de Solicitud de Patente de Estados Unidos Nº 20060134098, se añaden secuencias de un ratón (u otro mamífero) a un marco conservado humano, de modo que el anticuerpo resultante tenga regiones CDR1 y CDR2 de origen humano y una región CDR3 de origen humano. Otros métodos, tales como los descritos en la Publicación de Solicitud de Patente de Estados Unidos Nº 20030232333, generan anticuerpos que tienen regiones CDR1 y/o CDR1/CDR2 sintéticas junto con una región CDR3 natural. Sin embargo, estos métodos no consiguen proporcionar bibliotecas que contengan regiones marco conservadas estables y CDR plegadas correctamente.

Los métodos proporcionados en el presente documento diseñan los marcos conservados efectores de anticuerpos para clonación de diversidad. Se diseñó una estrategia para introducir diversidad en la CDR3 de dominios de anticuerpo humano seleccionados que evita la modificación de la secuencia del marco conservado original. La estrategia se basa en la introducción fuera de la región codificante de 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 pero escinden ADN a una distancia definida de hasta 20 bases independientemente de la secuencia de ADN hallada en el sitio de escisión. Para aprovechar este sistema para clonación de secuencias diversificadas en marcos conservados seleccionados, se diseñaron marcos conservados aceptores que contenían un fragmento de ADN de relleno, en lugar de la CDR3, que incluye los sitios de restricción de Tipo IIs. De forma similar, se generan secuencias de ADN diversificadas con secuencias flanqueantes que incluyen Tipo IIs. Siempre que los extremos cohesivos generados por las enzimas de restricción sean compatibles y que la fase de lectura se mantenga, los fragmentos de ADN pueden ligarse en el marco conservado aceptor y restaurar la CDR3 codificada en el nuevo contexto del marco conservado de anticuerpo aceptor (Figura 2).

Los métodos proporcionados en el presente documento capturan diversidad de CDR natural. La estrategia que se desarrolló para capturar fragmentos proteicos diversificados de forma natural como una fuente de diversidad también aprovecha las enzimas de restricción de Tipo IIs. Como ejemplo, se diseñaron cebadores oligonucleotídicos específicos para regiones flanqueantes de la secuencia de ADN que codifica la CDR3 H3 de inmunoglobulinas, es decir, específica para la FR3 y FR4 de la región variable. Estos oligonucleótidos contienen en su extremo 5' un sitio para una enzima de restricción de Tipo IIs mientras que su parte 3' coincide con la secuencia de ADN diana. El sitio de enzima de restricción usado es preferentemente 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 eficaz de secuencias de ADN naturales ya que mantiene una buena coincidencia entre el extremo 3' del cebador y el ADN que flanquea la CDR H3 permitiendo al mismo tiempo la escisión de la secuencia codificante de CDR H3 por escisión de ADN en el límite entre las regiones CDR y marco conservadas (Figura 3). Esta escisión precisa de la secuencia codificante de CDR es muy difícil usando enzimas de Tipo II que escinden 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 dichos sitios durante la amplificación sería difícil debido a una escasa hibridación de cebadores. Por lo tanto este método permite la amplificación de secuencias proteicas diversificadas y su inserción en cualquiera de los marcos conservados de anticuerpos aceptores independientemente del origen de la diversidad amplificada (Figura 4).

Los métodos descritos en el presente documento producen una biblioteca de ácidos nucleicos, en la que cada ácido nucleico codifica un dominio variable de inmunoglobulina: (a) proporcionando una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor que codifican dominios variables de inmunoglobulina distintos, incluyendo cada secuencia de ácido nucleico de Marco conservado Aceptor una primera región marco conservada (FR1), una segunda región marco conservada (FR2), una tercera región marco conservada (FR3) y una cuarta región marco conservada (FR4), en la que las regiones FR1 y FR2 se intercalan con una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 se intercalan con una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 se intercalan con una secuencia de ácido nucleico de relleno que contiene al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs intercaladas con una secuencia de ácido nucleico aleatoria; (b) proporcionando una pluralidad de secuencias de ácido nucleico diversificadas que codifican regiones de la región determinante de complementariedad 3 (CDR H3) o que codifican secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, en la que cada una de la pluralidad de secuencias de ácido nucleico diversificadas incluye un sitio de reconocimiento de enzimas de restricción de Tipo IIs en cada extremidad; (c) digiriendo cada una de la pluralidad de secuencias de ácido nucleico que codifican las regiones CDR3 o secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (b) y digiriendo la secuencia de ácido nucleico de relleno de la etapa (a) del Marco conservado Aceptor usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a); y (d) ligando las secuencias de ácido nucleico digeridas que codifican las regiones CDR3 o las secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 de la etapa (c) en el Marco conservado Aceptor digerido de la etapa (c) de modo que las regiones FR3 y FR4 se intercalan con las secuencias de ácido nucleico 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 un dominio variable de inmunoglobulina completo que no contienen los sitios de reconocimiento de enzimas de restricción de Tipo IIs de las etapas (a) y (b).

En algunas realizaciones, los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y etapa (b) en los métodos expuestos anteriormente se reconocen por una enzima de restricción de Tipo IIs diferente. Por ejemplo, en algunas realizaciones, los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de BsmBI, sitios de reconocimiento de BsaI, sitios de reconocimiento de FokI o una combinación de los mismos.

En algunas realizaciones, la secuencia de ácido nucleico de Marco conservado Aceptor deriva de una secuencia génica humana. Por ejemplo, en algunas realizaciones, la secuencia humana es una secuencia de gen variable de cadena pesada humana o una secuencia derivada de una secuencia de gen variable de cadena pesada humana. En algunas realizaciones, la secuencia de gen variable de cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23 y VH5-51.

En algunas realizaciones, la secuencia humana es una secuencia de gen variable de cadena ligera kappa humana o una secuencia derivada de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo, en algunas realizaciones, la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20.

En algunas realizaciones, la secuencia humana es una secuencia de gen variable de cadena ligera lambda humana o una secuencia derivada de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, en algunas realizaciones, la secuencia de gen variable de cadena ligera lambda humana se selecciona de VL1-44 y VL1-51.

En algunas realizaciones, la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias seleccionadas de secuencias de CDR3 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificados de forma natural.

En algunas realizaciones, la pluralidad de ácidos nucleicos diversificados codifica regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se ha identificado que se han expuesto a un antígeno particular.

En algunas realizaciones, 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.

En algunas realizaciones, la pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor incluye una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable.

En algunas realizaciones, los métodos proporcionados en el presente documento incluyen además las etapas de (e) clonar la biblioteca de ácidos nucleicos que codifican dominios variables de inmunoglobulina de la etapa (d) en un vector de expresión y (f) transformar el vector de expresión de la etapa (e) en una célula hospedadora y cultivar la célula hospedadora en condiciones suficientes para expresar una pluralidad de dominio variable de inmunoglobulina codificado por la biblioteca.

En algunas realizaciones, la célula hospedadora es *E. coli*. En algunas realizaciones, el vector de expresión es un vector fagémido.

Los métodos generan o producen de otro modo un anticuerpo específico de diana, región variable de anticuerpo o una parte del mismo,; (a) proporcionando una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor que codifican dominios variables de inmunoglobulina distintos, incluyendo cada secuencia de ácido nucleico de Marco conservado Aceptor una primera región marco conservada (FR1), una segunda región marco conservada (FR2), una tercera región marco conservada (FR3) y una cuarta región marco conservada (FR4), en las que las regiones FR1 y FR2 se intercalan con una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 se intercalan con una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 se intercalan con una secuencia de ácido nucleico de relleno que tiene al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs intercalados con una secuencia de ácido nucleico aleatoria; (b) proporcionando una pluralidad de secuencias de ácido nucleico diversificadas que codifican regiones de región determinante de complementariedad 3 (CDR3) o que codifican secuencias de aminoácidos que pueden cumplir el papel de una región CDR3, en las que cada una de la pluralidad de secuencias de ácido nucleico diversificadas incluye un sitio de reconocimiento de enzimas de restricción de Tipo IIs en cada extremo; (c) digiriendo cada una de la pluralidad de secuencias de ácido nucleico que codifican las regiones CDR3 o secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (b) y digiriendo la secuencia de ácido nucleico de relleno de la etapa (a) del Marco conservado Aceptor usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a); (d) ligando las secuencias de ácido nucleico digeridas que codifican las regiones CDR3 o las secuencias de aminoácidos que pueden cumplir el papel de una región CDR3 de la etapa (c) en el Marco conservado Aceptor digerido de la etapa (c) de modo que las regiones FR3 y FR4 se intercalan con las secuencias de ácido nucleico 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 dominios variables de inmunoglobulina completos que no contienen los sitios de reconocimiento de enzimas de restricción de Tipo IIs de las etapas (a) y (b); (e) clonando la biblioteca de ácidos nucleicos que codifican dominios variables de inmunoglobulina de la etapa (d) en un vector de expresión; (f) transformando el vector de expresión de la etapa (e) en una célula hospedadora y cultivando la célula hospedadora en condiciones suficientes para expresar una pluralidad de dominios variables de inmunoglobulina codificados por la biblioteca; (g) poniendo en contacto la pluralidad de dominios de inmunoglobulina de la etapa (f) con un antígeno diana; y (h) determinando qué secuencias codificantes de dominios variables de inmunoglobulina expresadas se unen con el antígeno diana.

En algunas realizaciones, los métodos proporcionados en el presente documento incluyen además la etapa de (i) secuenciar las secuencias que codifican dominios variables de inmunoglobulina que se unen con el antígeno diana.

En algunas realizaciones, los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y la etapa (b) se reconocen por una enzima de restricción de Tipo IIs diferente.

En algunas realizaciones, los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de Bsmbl, sitios de reconocimiento de Bsal, sitios de reconocimiento de FokI o una combinación de los mismos.

En algunas realizaciones, la secuencia de ácido nucleico de Marco conservado Aceptor deriva de una secuencia génica humana. Por ejemplo, en algunas realizaciones, la secuencia humana es una secuencia de gen variable de cadena pesada humana o una secuencia derivada de una secuencia de gen variable de cadena pesada humana. Por ejemplo, en algunas realizaciones, la secuencia de gen variable de cadena pesada humana se selecciona de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23 y VH5-51.

En algunas realizaciones, la secuencia humana es una secuencia de gen variable de cadena ligera kappa humana o una secuencia derivada de una secuencia de gen variable de cadena ligera kappa humana. Por ejemplo, en algunas realizaciones, la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20.

En algunas realizaciones, la secuencia humana es una secuencia de gen variable de cadena ligera lambda humana o una secuencia derivada de una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, en algunas realizaciones, la secuencia de gen variable de cadena ligera lambda humana se selecciona de VL1-44 y VL1-51.

En algunas realizaciones, la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias seleccionadas de secuencias de CDR3 de origen natural, secuencias de Ig de origen natural de seres humanos, secuencias de Ig de origen natural de un mamífero, secuencias de origen natural de una región de bucle de un receptor de linfocitos T en un mamífero, y otras colecciones de polipéptidos diversificados de forma natural.

En algunas realizaciones, la pluralidad de ácidos nucleicos diversificados codifica regiones CDR3, y la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias de inmunoglobulina que aparecen de forma natural en seres humanos que se han expuesto a un inmunógeno particular o secuencias derivadas de animales que se ha identificado que se han expuesto a un antígeno particular.

En algunas realizaciones, 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.

En algunas realizaciones, la pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor incluye una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable.

En algunas realizaciones, el vector de expresión es un vector fagémido. En algunas realizaciones, la célula hospedadora es *E. coli*.

A no ser que se defina de otro modo, los términos científicos y técnicos usados en relación con la presente invención tendrán los significados que se entienden habitualmente por los expertos habituales en la materia. Además, a no ser que se requiera de otro modo por el contexto, los términos singulares incluirán plurales y los términos plurales incluirán el singular. En general, 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 descritas en el presente documento son las que se conocen bien y se usan habitualmente en la técnica. Se usan técnicas convencionales para ADN recombinante, síntesis de oligonucleótidos y cultivo y transformación tisular (por ejemplo, electroporación, lipofección). Se realizan reacciones enzimáticas y técnicas de purificación de acuerdo con especificaciones del fabricante o como se realiza habitualmente en este campo o como se describe en el presente documento. Las técnicas y procedimientos anteriores se realizan en general de acuerdo con métodos convencionales bien conocidos en este campo y como se describe en diversas referencias generales y más específicas que se citan y analizan en toda la presente memoria descriptiva. Véase, por ejemplo, Sambrook *et al.* *Molecular Cloning: A Laboratory Manual* (2ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (1989)). Las nomenclaturas utilizadas en relación con, y los procedimientos de laboratorio y técnicas de, química analítica, química orgánica sintética, y química médica y farmacéutica descritas en el presente documento son las que se conocen bien y se usan habitualmente en la técnica. Se usan técnicas convencionales para síntesis química, análisis químicos, preparación farmacéutica, formulación y suministro, y tratamiento de pacientes.

Como se utiliza de acuerdo con la presente divulgación, se entenderá que los siguientes términos, a no ser que se indique de otro modo, tienen los siguientes significados:

Como se usa en el presente documento, el término “anticuerpo” se refiere a moléculas de inmunoglobulina y partes inmunológicamente activas de moléculas de inmunoglobulina (Ig), es decir, moléculas que contienen un sitio de unión a antígeno que se une específicamente (inmunorreacciona con) un antígeno. Por “unirse específicamente” o “inmunorreacciona con” o “unirse 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 mucha menor afinidad ($K_d > 10^{-6}$). Los anticuerpos incluyen, pero sin limitación, fragmentos policlonales, monoclonales, quiméricos dAb (anticuerpo de dominio), monocatenarios, F_{ab} , $F_{ab'}$ y $F_{(ab')_2}$, scFv y una biblioteca de expresión de F_{ab} .

Se sabe que la unidad estructural de anticuerpo básica comprende un tetrámero. Cada tetrámero está compuesto de dos pares idénticos de cadenas polipeptídicas, teniendo cada par una cadena “ligera” (aproximadamente 25 kDa) y una “pesada” (aproximadamente 50-70 kDa). La parte amino terminal de cada cadena incluye una región variable de aproximadamente 100 a 110 o más aminoácidos principalmente responsables del reconocimiento de antígenos. La parte carboxilo terminal de cada cadena define una región constante responsable principalmente de la función efectora. En general, las moléculas de anticuerpo obtenidas de seres humanos están relacionadas con 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 también subclases, tales como IgG₁, IgG₂ y otras. Además, en seres humanos, la cadena ligera puede ser una cadena kappa o una cadena lambda.

La expresión “anticuerpo monoclonal” (MAb) o “composición de anticuerpo monoclonal”, como se usa en el presente documento, se refiere a una población de moléculas de anticuerpo que contiene solamente una especie molecular de molécula de anticuerpo que consiste en un único producto de gen de cadena ligera y un único producto de gen de cadena pesada. En particular, las regiones determinantes de complementariedad (CDR) del anticuerpo monoclonal son idénticas en todas las moléculas de la población. Los MAb contienen un sitio de unión a antígeno capaz de inmunorreaccionar con un epítipo particular del antígeno caracterizado por una única afinidad de unión por

el.

La expresión "sitio de unión a antígeno" o "parte de unión" se refiere a la parte de la molécula de inmunoglobulina que participa en la unión a antígeno. El sitio de unión a antígeno está formado por restos de aminoácidos de las regiones variables N terminales ("V") de las cadenas pesada ("H") y ligera ("L"). Tres tramos altamente divergentes dentro de las regiones V de las cadenas pesada y ligera, denominados "regiones hipervariables", se intercalan entre tramos flanqueantes más conservados conocidos como "regiones marco conservadas" o "FR". Por lo tanto, el término "FR" se refiere a secuencias de aminoácidos que se encuentran de forma natural entre, y adyacentes a, regiones hipervariables en inmunoglobulinas. En una molécula de anticuerpo, las tres regiones hipervariables de una cadena ligera y las tres regiones hipervariables de una cadena pesada se disponen entre sí en espacio tridimensional para formar una superficie de unión a antígeno. La superficie de unión a antígeno es complementaria de la superficie tridimensional de un antígeno unido, y las tres regiones hipervariables de cada una de las cadenas pesadas y ligeras se denominan "regiones determinantes de complementariedad" o "CDR". La asignación de aminoácidos a cada dominio es de acuerdo con las definiciones de Kabat Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md. (1987 y 1991)), o Chothia y Lesk J. Mol. Biol. 196: 901-917 (1987), Chothia *et al.* Nature 342: 878-883 (1989).

Como se usa en el presente documento, el término "epítipo" incluye cualquier determinante proteico capaz de unirse específicamente con una inmunoglobulina, un scFv o un receptor de linfocitos T. El término "epítipo" incluye cualquier determinante proteico capaz de unirse específicamente con una inmunoglobulina o receptor de linfocitos T. Los determinantes epitópicos consisten habitualmente en agrupamientos de superficie químicamente activos de moléculas tales como aminoácidos o cadenas laterales de azúcares que habitualmente tienen tres características estructurales tridimensionales específicas, así como características de carga específicas. Por ejemplo, pueden inducirse anticuerpos contra péptidos N terminales o C terminales de un polipéptido. Se dice que un anticuerpo se une específicamente con un antígeno cuando la constante de disociación es $\leq 1 \mu\text{M}$; por ejemplo, $\leq 100 \text{ nM}$, preferentemente $\leq 10 \text{ nM}$ y más preferentemente $\leq 1 \text{ nM}$.

Como se usa en el presente documento, las expresiones "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 que es específica la inmunoglobulina. La fuerza, o afinidad de las interacciones de unión inmunológicas pueden expresarse con respecto a la constante de disociación (K_d) de la interacción, en la que una K_d más pequeña representa una mayor afinidad. Las propiedades de unión inmunológica de polipéptidos seleccionados pueden cuantificarse usando métodos bien conocidos en la técnica. Uno de dichos métodos implica medir las velocidades de formación y disociación de complejo de sitio y unión a antígeno/antígeno, en el que dichas velocidades dependen de las concentraciones de los compañeros complejos, la afinidad de la interacción y parámetros geométricos que influyen igualmente en la velocidad de ambas direcciones. Por lo tanto, tanto la "constante de velocidad de asociación" (K_{on}) como la "constante de velocidad de disociación" (K_{off}) pueden determinarse calculando las concentraciones y las velocidades reales de asociación y disociación (Véase Nature 361: 186-87 (1993)). La relación de K_{off}/K_{on} permite la cancelación de todos los parámetros no relacionados con la afinidad, y es igual a la constante de disociación K_d . (Véase, en general, Davies *et al.* (1990) Annual Rev Biochem 59: 439-473). Se dice que un anticuerpo de la presente invención se une específicamente con su diana, cuando la constante de unión en equilibrio (K_d) es $\leq 1 \mu\text{M}$, por ejemplo, $\leq 100 \text{ nM}$, preferentemente $\leq 10 \text{ nM}$, y más preferentemente $\leq 1 \text{ nM}$, como se mide por ensayos tales como ensayos de unión a radioligando o ensayos similares conocidos por los expertos en la materia.

La expresión "polinucleótido aislado" como se usa en el presente documento significará un polinucleótido de origen genómico, de ADNc o sintético o alguna combinación de los mismos, que según su origen el "polinucleótido aislado" (1) no está asociado con todo o una parte de un polinucleótido en el que el "polinucleótido aislado" se encuentra en la naturaleza, (2) está unido operativamente con un polinucleótido con el que no está unido en la naturaleza, o (3) no aparece en la naturaleza como parte de una secuencia mayor. Los polinucleótidos de acuerdo con la invención incluyen las moléculas de ácido nucleico que codifican las moléculas de inmunoglobulina de cadena pesada, y moléculas de ácido nucleico que codifican las moléculas de inmunoglobulina de cadena ligera descritas en el presente documento.

La expresión "proteína aislada" a la que se hace referencia en el presente documento significa una proteína de ADNc, ARN recombinante, o de origen sintético o alguna combinación de los mismos, que según su origen, o fuente de derivación, la "proteína aislada" (1) no está asociada con proteínas halladas en la naturaleza, (2) está libre de otras proteínas de la misma fuente, por ejemplo, libre de proteínas marinas, (3) se expresa por una célula de una especie diferente, o (4) no aparece en la naturaleza.

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

moléculas de inmunoglobulina de cadena ligera kappa, y viceversa, así como fragmentos y análogos de las mismas.

La expresión “de origen natural” como se usa en el presente documento como se aplica a un objeto se refiere al hecho de que un objeto puede hallarse en la naturaleza. Por ejemplo, una secuencia polipeptídica o polinucleotídica que está presente en un organismo (incluyendo virus) que puede aislarse de una fuente en la naturaleza y que no se ha modificado intencionadamente por el hombre en el laboratorio o de otro modo es de origen natural.

La expresión “unido operativamente” como se usa en el presente documento se refiere a que posiciones de componentes así descritos están en una relación que les permite actuar de su manera pretendida. Una secuencia de control “unida operativamente” con una secuencia codificante está ligada de tal manera que la expresión de la secuencia codificante se consigue en condiciones compatibles con las secuencias de control.

La expresión “secuencia de control” como se usa en el presente documento se refiere a secuencias polinucleotídicas que son necesarias para efectuar la expresión y procesamiento de secuencias codificantes con las que se ligan. La naturaleza de dichas secuencias de control difiere dependiendo del organismo hospedador en procariotas, dichas secuencias de control incluyen en general promotor, sitio de unión a ribosomas y secuencias de terminación de la transcripción en eucariotas, en general, dichas secuencias de control incluye promotores y secuencia de terminación de la transcripción. Se entiende que la expresión “secuencias de control” incluye, como mínimo, todos los componentes cuya presencia es esencial para expresión y procesamiento, y también pueden incluir componentes adicionales cuya presencia es ventajosa, por ejemplo, secuencias líderes y secuencias de compañeros de fusión. El término “polinucleótido” como se denomina en el presente documento significa un boro polimérico de nucleótidos de al menos 10 bases de longitud, bien ribonucleótidos o bien desoxinucleótidos o una forma modificada de uno de los tipos de nucleótido. El término incluye formas mono y bicatenarias de ADN.

Como se usa en el presente documento, los veinte aminoácidos convencionales y sus abreviaturas siguen el uso convencional. Véase *Immunology - A Synthesis* (2ª Edición, 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 como aminoácidos α , α disustituidos, N-alquil aminoácidos, ácido láctico y otros aminoácidos no convencionales también pueden ser componentes adecuados para polipéptidos de la presente invención. Los ejemplos de aminoácidos no convencionales incluyen: 4 hidroxiprolina, γ -carboxiglutamato, ϵ -N,N,N-trimetilisina, ϵ -N-acetilisina, O-fosfoserina, N-acetilserina, N-formilmietionina, 3-metilhistidina, 5-hidroxilisina, σ -N-metilarginina, y otros aminoácidos similares e iminoácidos (por ejemplo, 4- hidroxiprolina). En la notación de polipéptidos usada en el presente documento, la dirección izquierda es la dirección amino terminal y la dirección derecha es la dirección carboxilo terminal, de acuerdo con el uso habitual y la convención.

Como se aplica a los polipéptidos, la expresión “identidad sustancial” significa que dos secuencias peptídicas, cuando se alinean de forma óptima, tal como por los programas GAP o BESTFIT usando pesos de huecos por defecto, comparten al menos 80 por ciento de la identidad de secuencia, preferentemente al menos 90 por ciento de la identidad de secuencia, más preferentemente al menos 95 por ciento de la identidad de secuencia, y más preferentemente al menos 99 por ciento de la identidad de secuencia.

Preferentemente, las posiciones de restos que no son idénticas difieren en sustituciones de aminoácidos conservativas.

Las sustituciones de aminoácidos conservativas se refieren a la intercambiabilidad de restos 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 de hidroxilo-alifáticas 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 triptófano; 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. Son grupos de sustitución de aminoácidos conservativa preferidos: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina valina, ácido glutámico-ácido aspártico y asparagina-glutamina.

Como se analiza en el presente documento, se contemplan variaciones menores en las secuencias de aminoácidos de anticuerpos o moléculas de inmunoglobulina que están abarcadas por la presente invención, siempre que las variaciones en la secuencia de aminoácidos mantengan al menos 75 %, más preferentemente al menos 80 %, 90 %, 95 % y más preferentemente 99 %. En particular, se contemplan reemplazos de aminoácidos conservativos. Son reemplazos conservativos los que tienen lugar dentro de una familia de aminoácidos que están relacionados en sus cadenas laterales. Los aminoácidos codificados genéticamente se dividen en general en familias: (1) son aminoácidos ácidos aspartato, glutamato; (2) son aminoácidos básicos lisina, arginina, histidina; (3) son aminoácidos no polares alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano y (4) son aminoácidos polares sin carga 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, triptófano, tirosina y valina. Otras familias de aminoácidos incluyen (i) serina y treonina, que son la familia de hidroxilo-alifática; (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, triptófano 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 relacionado estructuralmente no tendrá un efecto importante en la unión o propiedades de la molécula resultante, especialmente si el reemplazo no implica un aminoácido dentro de un sitio marco conservado. Si un cambio de aminoácido da como resultado un péptido funcional se puede determinar fácilmente ensayando la actividad específica del derivado polipeptídico. Se describen ensayos en detalle en el presente documento. Los expertos habituales en la técnica pueden preparar fácilmente fragmentos o análogos de anticuerpos o moléculas de inmunoglobulina. Los extremos amino y carboxilo terminal de fragmentos o análogos aparecen cerca de los límites de los dominios funcionales. Pueden identificarse dominios estructurales y funcionales por comparación de los datos de secuencia de nucleótidos y/o aminoácidos con bases de datos de secuencias públicas o patentadas. Preferentemente, se usan métodos de comparación computarizados para identificar motivos de secuencia o dominios de conformación de proteína predichos que aparecen en otras proteínas de estructura y/o función conocidas. Se conocen métodos para identificar secuencias proteicas que se pliegan en una estructura tridimensional conocida. Bowie *et al.* Science 253: 164 (1991). Por lo tanto, los ejemplos anteriores demuestran que los expertos en la materia pueden reconocer motivos de secuencia y conformaciones estructurales que pueden usarse para definir dominios estructurales y funcionales de acuerdo con la invención.

Son sustituciones de aminoácidos preferidas las que: (1) reducen la susceptibilidad a proteólisis, (2) reducen la susceptibilidad a oxidación, (3) alteran la afinidad de unión para formar complejos proteicos, (4) alteran las afinidades de unión y (4) confieren o modifican otras propiedades físico químicas o funcionales de dichos análogos. Los análogos pueden incluir diversas muteínas de una secuencia distinta de la secuencia peptídica de origen natural. Por ejemplo, pueden prepararse sustituciones de aminoácidos individuales o múltiples (preferentemente sustituciones de aminoácidos conservativas) en la secuencia de origen natural (preferentemente en la parte del polipéptido fuera del dominio o los dominios que forman contactos intermoleculares. Una sustitución de aminoácidos conservativa no debería cambiar sustancialmente las características estructurales de la secuencia parental (por ejemplo, un aminoácido de reemplazo no debería tender a romper una hélice que aparezca en la secuencia parental, o alterar otros tipos de estructuras secundarias que caracterizan la secuencia parental). Se describen ejemplos de estructuras secundarias y terciarias de polipéptidos reconocidas en la técnica en Proteins, Structures and Molecular Principles (Creighton, Ed., W. H. Freeman and Company, Nueva York (1984)); Introduction to Protein Structure (C. Branden y J. Tooze, eds., Garland Publishing, Nueva York, N. Y. (1991)); y Thornton *et al.* Nature 354: 105 (1991).

Como se usa en el presente documento, los términos “marcador” o “marcado” se refieren a la incorporación de un marcador detectable, por ejemplo, mediante la incorporación de un aminoácido radiomarcado o unión con un polipéptido de restos de biotínulo que pueden detectarse por avidina marcada (por ejemplo, estreptavidina que contiene un marcador fluorescente o actividad enzimática que puede detectarse por métodos ópticos o calorimétricos). En ciertas situaciones, el marcador o la etiqueta también pueden ser terapéuticos. Se conocen en la técnica y pueden usarse diversos métodos para marcar polipéptidos y glucoproteínas. Los ejemplos de marcadores para polipéptidos incluyen, pero sin limitación, los siguientes: radioisótopos o radionúclidos (por ejemplo, ^3H , ^{14}C , ^{15}N , ^{35}S , ^{90}Y , ^{99}Tc , ^{111}In , ^{125}I , ^{131}I), marcadores fluorescentes (por ejemplo, FITC, rodamina, fósforos de lantánido), marcadores enzimáticos (por ejemplo, peroxidasa de rábano rusticano, p-galactosidasa, luciferasa, fosfatasa alcalina), grupos quimioluminiscentes, de biotínulo, epítomos polipeptídicos predeterminados reconocidos por un indicador secundario (por ejemplo, secuencias de pares de cremallera de leucina, sitios de unión para anticuerpos secundarios, dominios de unión a metales, marcadores epítópicos). En algunas realizaciones, los marcadores se unen con ramas espaciadoras de diversas longitudes para reducir el impedimento estérico potencial. La expresión “agente farmacéutico o fármaco” como se usa en el presente documento se refiere a un compuesto químico o composición capaz de inducir un efecto terapéutico deseado cuando se administra de forma apropiada a un paciente.

Otros términos de química del presente documento se usan de acuerdo con el uso convencional en la técnica, como se ejemplifica en The McGraw-Hill Dictionary of Chemical Terms (Parker, S., Ed., McGraw-Hill, San Francisco (1985)).

Como se usa en el presente documento, “sustancialmente puro” significa que una especie objeto es la especie predominante presente (es decir, en medida molar es más abundante que cualquier otra especie individual en la composición), y preferentemente una fracción sustancialmente purificada es una composición en la que la especie objeto comprende al menos aproximadamente 50 por ciento (en una medida molar) de todas las especies macromoleculares presentes.

En general, una composición sustancialmente pura comprenderá más de aproximadamente el 80 por ciento de todas las especies macromoleculares presentes en la composición, más preferentemente más de aproximadamente 85 %, 90 %, 95 % y 99 %. Más preferentemente, la especie objeto se purifica hasta homogeneidad esencial (la especie contaminante no puede detectarse en la composición por métodos de detección convencionales) en la que la composición consiste esencialmente en una única especie macromolecular.

El término paciente incluye sujetos humanos y veterinarios.

Se purifican anticuerpos por técnicas bien conocidas, tales como cromatografía de afinidad usando proteína A o proteína G, que proporcionan principalmente la fracción IgG de suero inmunitario. Posteriormente, o como alternativa, el antígeno específico que es la diana de la inmunoglobulina que se busca, o un epítipo del mismo, puede inmovilizarse en una columna para purificar el anticuerpo específico inmunitario por cromatografía de inmovilización. Se analiza la purificación de inmunoglobulinas, por ejemplo, en D. Wilkinson (The Scientist, publicado en The Scientist, Inc., Filadelfia PA, Vol. 14, Nº 8 (17 de abril de 2000), pp. 25-28).

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

Ejemplos

Ejemplo 1: clonación de genes de línea germinal variables de inmunoglobulina

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

Los cebadores y combinación de cebadores usados se indican a continuación.

Localizadores genómicos

5	K1-33 TGTTTCTAATCGCAGGTGCCAGATG	(SEC ID Nº: 120)
3	K1-33 ATTTATGTTATGACTTGTTACACTG	(SEC ID Nº: 121)
5	K1-39 TATTTGTTTTATGTTTCCAATCTC	(SEC ID Nº: 122)
3	K1-39 CCTTGGAGGTTTATGTTATGACTTG	(SEC ID Nº: 123)
5	K3-11 TTATTTCCAATTTCCAGATACCACCG	(SEC ID Nº: 124)
3	K3-11 TTGTTGGGGTTTTTGTTCATGTGG	(SEC ID Nº: 125)
5	K3-15 TATTTCCAATTTCCAGATACCACTGG	(SEC ID Nº: 126)
3	K3-15 ATGTTGAATCACTGTGGGAGGCCAG	(SEC ID Nº: 127)
5	K3-20 TTATTTCCAATCTCAGATACCACCG	(SEC ID Nº: 128)
3	K3-20 TTTTGTTCAGCTGAATCACTGTG	(SEC ID Nº: 129)
5	L1-44 ATGTCTGTGTCTCTCTCACTTCCAG	(SEC ID Nº: 130)
3	L1-44 TTCCCCATTGGCCTGGAGCACTGTG	(SEC ID Nº: 131)
5	L1-51 GTGTCTGTGTCTCTCTCTGTTCCAG	(SEC ID Nº: 132)
3	L1-51 CTTGTCTCAGTTCCCCATTGGGCTG	(SEC ID Nº: 133)
5	H1-2 ATCTCATCCACTTCTGTGTTCTCTC	(SEC ID Nº: 134)
3	H1-2 TTGGGTTTCTGACACCCTCAGGATG	(SEC ID Nº: 135)
5	H1-18 CAGGCCAGTCATGTGAGACTTCACC	(SEC ID Nº: 136)
3	H1-18 CTGCCTCCTCCCTGGGGTTTCTGAA	(SEC ID Nº: 137)
5	H1-69 CCCCTGTGTCTCTCCACAGGTGTC	(SEC ID Nº: 138)
3	H1-69 CCGGCACAGCTGCCTTCTCCCTCAG	(SEC ID Nº: 139)
5	DP-47 GAGGTGCAGCTGTTGGAG	(SEC ID Nº: 140)
5	H3-23 TCTGACCAGGGTTTCTTTTGTGTTGC	(SEC ID Nº: 141)
3	H3-23 TTGTGTCTGGGCTCACAATGACTTC	(SEC ID Nº: 142)
5	H3-30 TGGCATTTTCTGATAACGGTGTC	(SEC ID Nº: 143)
3	H3-30 CTGCAGGGAGGTTTGTGTCTGGGCG	(SEC ID Nº: 144)

Localizadores genómicos

5	H3-48 ATATGTGTGGCAGTTTCTGACCTTG	(SEC ID N°: 145)
3	H3-48 GGTTTGTGTCTGGTGTACACTGAC	(SEC ID N°: 146)
5	H5-a GAGTCTGTGCCGGAAGTGCAGCTGG	(SEC ID N°: 147)

Específico para secuencia codificante

5	VH1 TATCAGGTGCAGCTGGTGCAG	(SEC ID N°: 148)
5	VH3 TATCAGGTGCAGCTGGTGGAG	(SEC ID N°: 149)
5	VH5 TATGAGGTGCAGCTGGTGCAG	(SEC ID N°: 150)
3	VH1/3 ATATCTCTCGCACAGTAATACAC	(SEC ID N°: 151)
3	VH3 ATATCTCTCGCACAGTAATATAC	(SEC ID N°: 152)
3	VH5 ATATGTCTCGCACAGTAATACAT	(SEC ID N°: 153)
5	VK1 TATGACATCCAGATGACCCAGTCTCCATCCTC	(SEC ID N°: 154)
3	DPK9 ATAGGAGGGGTACTGTAACT	(SEC ID N°: 155)
3	DPK1 ATAGGAGGGAGATTATCATA	(SEC ID N°: 156)
5	DPK22_L6 TATGAAATTGTGTTGACGCAGTCT	(SEC ID N°: 157)
3	DPK22 ATAGGAGGTGAGCTACCATACTG	(SEC ID N°: 158)
5	DPK21 TATGAAATAGTGATGACGCAGTCT	(SEC ID N°: 159)
3	DPK21 ATAGGAGGCCAGTTATTATACTG	(SEC ID N°: 160)
3	L6 CAGCGTAGCAACTGGCCTCCTAT	(SEC ID N°: 161)
5	DPL2 TACAGTCTGTGCTGACTCAG	(SEC ID N°: 162)
3	DPL2 ATAGGACCATTGAGGCTGTCATC	(SEC ID N°: 163)
5	DPL5 TATCAGTCTGTGTTGACGCAG	(SEC ID N°: 164)
3	DPL5 ATAGGAGCACTCAGGCTGCTAT	(SEC ID N°: 165)

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

		1ª PCR		2ª PCR	
Familia	línea germinal	5'	3'	5'	3'
VH1	DP-8/75 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	5 VH1	3 VH1/3
VH3	DP-49 HV 3-30	5 H3-30	3 H3-30	5 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 H223	3 H3-23	5 VH3	3 VH3
VH5	HV 5a	5 H5a	3 VH5	5 VH5	3 VH5
VKI	DPK-1 KV 1-33	5 K 1-33	3 K 1-33	5 VK1	3 DPK-1
	DPK-9 KV 1-39	5 K 1-39	3 K 1-39	5 VK	3 DPK-9
VKIII	L6 KV 3-11	5 K3-11	3 K3-11	5 DPK22_L6	3 L6
	DPK-21 KV 3-15	5 K3-15	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	5 L1-51	3 L1-51	5 DPL5	3 DPL5

5 Ejemplo 2: generación de marcos conservados aceptores

Las secuencias de los genes de línea germinal seleccionados se analizaron con respecto a la presencia de sitios de restricción de Tipo IIs. No estaba presente ningún sitio BsmBI en los genes de línea germinal variables de

anticuerpos seleccionados. Se encontraron dos sitios BsmBI en la cadena principal de pNDS1, el vector fagémido en el que se clonaría el Marco conservado Aceptor. Estos dos sitios se retiraron por mutagénesis dirigida de modo que pudieran introducirse sitios de BsmBI únicos en las secuencias de ADN de relleno de los Marcos conservados Aceptores. Cada gen de línea germinal se amplificó por PCR anidada múltiple para añadir una secuencia de ADN de relleno en el extremo 3' de la secuencia de FR3 seguida de 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 J de línea germinal JH1, JH3, JH4 y JH5. La secuencia de aminoácidos de VK FR4 corresponde a la región FR4 codificada por los genes J de línea germinal JK1. La secuencia de aminoácidos de Vλ FR4 corresponde a la región FR4 codificada por los genes J de línea germinal JL2 y JL3. Se generaron dos variantes de la secuencia de Vk FR4 con una sustitución de un único aminoácido en la posición 106 (Arginina o Glicina). Para el Marco conservado Aceptor basado en el gen de línea germinal VH3-23, también se construyeron dos variantes que diferían en un único aminoácido (Lisina a Arginina) en la posición 94, el último resto de FR3. Durante la etapa de amplificación final se introdujeron sitios SfiI/NcoI y XhoI en los extremos 5' y 3' del VH, respectivamente.

De forma similar, se introdujeron sitios Sall y NotI en el extremo 5' y 3' de VL, respectivamente (Figura 6). El fragmento de relleno se diseñó de modo que la fase de lectura de traducción se desplazara evitando de este modo la expresión de cualquier proteína funcional de los Marcos conservados Aceptores (Figura 7). Los cebadores usados en este proceso se enumeran a continuación.

VH

- 5 VH1 CAGCCGGCCATGGCCCAGGTGCAGCTGGTGCAG (SEC ID N°: 166)
- 5 VH3-30 CAGCCGGCCATGGCCCAGGTGCAGCTGGTGGAG (SEC ID N°: 167)
- 5 VH3-23 CAGCCGGCCATGGCCGAGGTGCAGCTGTTGGAG (SEC ID N°: 168)
- 5 VH3-48 CAGCCGGCCATGGCCGAGGTGCAGCTGGTGGAGTCTGGGGGAG (SEC ID N°: 169)
- 5 VH5-51 CAGCCGGCCATGGCCGAGGTGCAGCTGGTGCAG (SEC ID N°: 170)
- 3 VH1/3 CTTACCGTTATTCGTCTCATCTCGCACAGTAATACAC (SEC ID N°: 171)
- 3 VH3-23 CTTACCGTTATTCGTCTCATTTTCGCACAGTAATATAC (SEC ID N°: 172)
- 3 VH3-48 CTCGCACAGTAATACACAGCCGTGTCCTCGGCTCTCAGGCTG (SEC ID N°: 173)
- 3 VH5-51 CTTACCGTTATTCGTCTCATCTCGCACAGTAATACAT (SEC ID N°: 174)
- 3 VHex1 CAATACGCGTTTAAACCTGGTAAACCGCCTTACCGTTATTCGTCTCA (SEC ID N°: 175)
- 3 VHex2 GTTCCCTGGCCCCAAGAGACGCGCCTTCCCAATACGCGTTTAAACCTG (SEC ID N°: 176)
- 3 VHex3 CCTCCACCGCTCGAGACTGTGACCAGGGTCCCTGGCCCCAAGAG (SEC ID N°: 177)

VK

- 5 VK1 CGGGTCGACGGACATCCAGATGACCCAGTC (SEC ID N°: 178)
- 5 VK3-11 CGGGTCGACGGAAATTGTGTTGACACAGTCTCCAGC (SEC ID N°: 179)
- 5 VK3-15 CGGGTCGACGGAAATAGTGATGACGCAGTCTCCAGC (SEC ID N°: 180)
- 5 VK3-20 CGGGTCGACGGAAATTGTGTTGACGCAGTCTCCAGG (SEC ID N°: 181)

VK

- 3 VK1-33 CCTTACCGTTATTCGTCTCGCTGCTGACAGTAATATGTTGCAATA (SEC ID N°: 182)
- 3 VK1-39 CCTTACCGTTATTCGTCTCGCTGCTGACAGTAGTAAGTTGCAAAA (SEC ID N°: 183)
- 3 VK3 CCTTACCGTTATTCGTCTCGCTGCTGACAGTAATAAACTGCAAAATC (SEC ID N°: 184)
- 3 VKext1 CCAATACGCGTTTAAACCTGGTAAACCGCCTTACCGTTATTCGTCTC (SEC ID N°: 185)
- 3 VKext2 GGTCCCTTGGCCGAATGAGACGCGCCTTCCCAATACGCGTTTAAAC (SEC ID N°: 186)
- 3 VKext3R GTGCGGCCGCCCCGTTTGATTTCACCTTGGTCCCTTGGCCGAATG (SEC ID N°: 187)
- 3 VKext3G GTGCGGCCGCCCCGTTTGATTTCACCTTGGTCCCTTGGCCGAATG (SEC ID N°: 188)

Vλ

- 5 VL1-44 CGGGTCGACGCAGTCTGTGCTGACTCAGCCAC (SEC ID N°: 189)
- 5 VL1-51 CGGGTCGACGCAGTCTGTGTTGACGCAGCCGC (SEC ID N°: 190)
- 3 VL1-44 CCTTACCGTTATTCGTCTCCTGCTGCACAGTAATAATC (SEC ID N°: 191)

- 3 VL1-51 CCTTACCGTTATTCTGTCCTGTTCCGCGAGTAATAATC (SEC ID N°: 192)
 3 Vlex2 CCCTCCGCCGAACACAGAGACGCGCCTTCCCAATACGCGTTTAAAC (SEC ID N°: 193)
 3 Vlex3 GTGCGGCCGCCCTAGGACGGTCAGCTTGGTCCCTCCGCCGAACACAGA (SEC ID N°: 194)

Las secuencias de los 20 Marcos conservados Aceptores ensamblados finales se muestran en la Figura 8.

Ejemplo 3: generación de vectores aceptores de fagémidos que contienen un dominio variable invariante

El vector fagémido pNDS1 usado para la expresión de scFv se modificó en primer lugar para retirar dos sitios de BsmBI. Se clonó un dominio VH3-23 que contenía una secuencia de CDR3 definida en el pNDS1 modificado usando los sitios de restricción SfiI y XhoI para obtener el vector fagémido pNDS_VHdummy. Este dominio contenía un sitio de BsmBI en la región FR4, que se corrigió por mutagénesis dirigida silenciosa. En paralelo, se clonó después un dominio VK1-39 que contenía una secuencia de CDR3 definida en el pNDS1 modificado usando los sitios de restricción de Sall y NotI para obtener el vector fagémido pNDS_VKdummy (Figura 9). Los 8 Marcos conservados Aceptores de VH se clonaron en pNDS_VKdummy usando los sitios de restricción de Sall y NotI. Los 12 Marcos conservados Aceptores VL se clonaron en pNDS_VHdummy usando los sitios de restricción de SfiI y XhoI. Los 20 vectores fagémidos pNDS resultantes que se enumeran a continuación podrían usarse en este estadio para clonar 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_VH3-23K_VKd; 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_VK3-15G; pNDS_VHd_VK3-15R; pNDS_VHd_VK3-20G; pNDS_VHd_VK3-20R; pNDS_VHd_VL1-44; pNDS_VHd_VK1-51.

Ejemplo 4: captura de diversidad de CDR H3 natural de repertorios humanos

Se usaron múltiples fuentes de ADNc humano como un molde para amplificación de secuencias de CDR H3. Estas fuentes incluían bazo fetal humano así como grupos de células purificadas de sangre periférica de adulto normal hombre y mujer. Se han usado varias estrategias para amplificación para recuperar secuencias de CDR H3 que se originan de ADNc VH reordenado codificado por un gen de línea germinal específico o secuencias de CDR H3 que se originan de cualquier ADNc VH.

En primer lugar, se usaron mezclas de cebadores que coincidían con las regiones codificantes 5' de la mayoría de las familias de VH humanas en combinación con mezclas de cebadores que coincidían con todas las regiones JH humanas. Esto permitió la amplificación por PCR de una 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. Ese ADN actuó como un molde en una segunda etapa de PCR usando cebadores con una coincidencia de 13 pb y 14 pb para el final de la región FR3 y el comienzo de FR4, respectivamente. En la mayoría de los casos, el último resto de la FR3 es una arginina o una lisina. Como las últimas coincidencias de pb son críticas para la extensión de cebadores por la polimerasa, se usaron dos cebadores 5' diferentes: 5 VHR_FOK (SEC ID N°: 205 mostrada posteriormente) y 5 VHK_FOK (SEC ID N°: 206 mostrada posteriormente). Resulta importante que estos cebadores también contienen un sitio de restricción de FokI para escisión de la secuencia de CDR H3 (Figura 4). Los cebadores usados en la segunda etapa de PCR se biotinizaron en el extremo 5' para facilitar etapas de purificación corriente abajo (véase ejemplo 5). Este enfoque de dos etapas permite una amplificación eficaz de las secuencias de CDR H3 a pesar del número limitado de coincidencias de pares de bases. Se realizaron amplificaciones a diversas temperaturas de hibridación (entre 30 °C y 70 °C) y con varias ADN polimerasas termoestables para establecer condiciones óptimas. Se descubrió que una temperatura de hibridación de 55-58 °C en combinación con GoTaq polimerasa (Promega) era óptima para este conjunto de cebadores. El segundo producto de amplificación se separó en un gel de agarosa al 2 % y dio como resultado una mancha en la parte inferior del gel correspondiente a CDR H3 de diferente longitud. La mancha de ADN completa se extrajo del gel o una región correspondiente a fragmentos de ADN mayores para enriquecer con respecto a CDR H3 larga.

Como alternativa, la primera etapa de amplificación se realizó usando el cebador 5' 5 VH3-23H2 (SEC ID N°: 201 mostrado posteriormente), que es específico para la secuencia que codifica la CDR H2 de la línea germinal VH3-23. Como los genes de línea germinal diferentes son diversos en esta CDR, pueden amplificarse preferentemente ADNc de VH codificados por el gen de línea germinal seleccionado. Las etapas de purificación y amplificación posteriores fueron idénticas. De esta manera, es posible recuperar CDR que se originan de un ambiente de marco conservado específico y reintroducirlas en el mismo marco conservado, uno similar o uno diferente.

A continuación hay una lista de cebadores usados para la amplificación de repertorios de CDR H3 humanos naturales.

1ª etapa de PCR

5 VH1/5 CCGCACAGCCGGCCATGGCCCAGGTGCAGCTGGTGCAGTCTGG (SEC ID Nº: 195)
 5 VH3 CCGCACAGCCGGCCATGGCCCAGGTGCAGCTGGTGGAGTCTGG (SEC ID Nº: 196)
 5 VH2 CCGCACAGCCGGCCATGGCCCAGRTACCTTGCTCGAGTCTGG (SEC ID Nº: 197)
 5 VH4 CCGCACAGCCGGCCATGGCCCAGGTGCAGCTGCAGGAGTCGGG (SEC ID Nº: 198)
 5 VH4DP64 CCGCACAGCCGGCCATGGCCCAGGTGCAGCTGCAGGAGTCCGG (SEC ID Nº: 199)
 5 VH4DP63 CCGCACAGCCGGCCATGGCCCAGGTGCAGCTACAGCAGTGGGG (SEC ID Nº: 200)
 5 VH3-23H2 TGGAGTGGGTCTCAGCTATTAGTGGTAGTGGT (SEC ID Nº: 201)
 3 HJ1/2 CGATGGGGCCCTTGGTGGAGGCTGAGGAGACRGTGACCAGGGTGCC (SEC ID Nº: 202)
 3 HJ3/6 CGATGGGGCCCTTGGTGGAGGCTGAAGAGACGGTGACCRKTGTCCTC (SEC ID Nº: 203)
 3 HJ4/5 CGATGGGGCCCTTGGTGGAGGCTGAGGAGACGGTGACCAGGGTTCC (SEC ID Nº: 204)

2ª etapa de PCR

5 VHR_FOK GAGCCGAGGACACGGCCGGATGTTACTGTGCGAGA (SEC ID Nº: 205)
 5 VHK_FOK GAGCCGAGGACACGGCCGGATGTTACTGTGCGAAA (SEC ID Nº: 206)
 3 JH1_FOK GAGGAGACGGTGACGGATGTGCCCTGGCCCCA (SEC ID Nº: 207)
 3 JH2_FOK GAGGAGACGGTGACGGATGTGCCACGGCCCCA (SEC ID Nº: 208)
 3 JH3456_FOK GAGGAGACGGTGACGGATGTGYCCTTGGCCCCA (SEC ID Nº: 209)

Ejemplo 5: generación de bibliotecas primarias por clonación de CDR H3 humanas naturales en marcos conservados aceptores

5 La CDR H3 amplificada se digirió con FokI, y los extremos escindidos así como el ADN no digerido se retiraron usando perlas magnéticas recubiertas con estreptavidina. En paralelo, se digirieron vectores Aceptores de VH pNDS usando BsmBI. Como los salientes generados por estas digestiones son compatibles, la colección de CDR H3 naturales fue capaz de ligarse en el Marco conservado Aceptor de VH restaurando la fase de lectura apropiada. El ADN ligado se purificó y se concentró para transformación en células de *E. coli* XL1 Blue competentes, y se
 10 analizaron clones aleatorios por secuenciación para comprobar que la secuencia de CDR H3 se había reconstituido y que los puntos de unión entre la región CDR y la marco conservada eran correctos (Figura 10). Los resultados indicaron que todos los clones contenían secuencias de CDR H3 y que la fase de lectura se había restaurado, codificando de este modo una cadena pesada variable de inmunoglobulina. Además, todas las CDR fueron diferentes, lo que indicó que se habían capturado por este enfoque una gran diversidad de secuencias de origen
 15 natural. La longitud de la CDR H3 también fue variable y se encontraron CDR relativamente largos de 10 a 15 restos, subrayando de este modo la ventaja de este enfoque para tomar muestras de secuencias de CDR largas que son difíciles de abarcar usando diversidad sintética.

Usando este método, se clonaron secuencias de CDR H3 naturales, derivadas de células purificadas de sangre
 20 periférica humana agrupadas o bazo fetal humano, en cada uno de los Marcos conservados Aceptores de VH de pNDS y se transformaron en células *E. coli* TG1 electrocompetentes y se sembraron en placas de Bioensayo 2xTYAG (medio 2xTY que contiene ampicilina 100 µg/ml y glucosa 2 %). Después de incubación durante una 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 contenía glicerol al 50 % a la suspensión
 25 celular para obtener una concentración final de glicerol al 17 %. Las alícuotas de las bibliotecas se almacenaron a -80 °C. En este proceso, se generaron 14 bibliotecas primarias que representaban un total de $8,1 \times 10^9$ transformantes. Se secuenciaron 180 clones seleccionados aleatoriamente para determinar la calidad y diversidad de las bibliotecas. Todos los clones codificaban secuencias de VH diferentes y más del 89 % estaban en fase. Estas bibliotecas primarias contienen diversidad en la CDR H3 solamente ya que se combinan con un dominio VL de
 30 prueba.

Ejemplo 6: generación de bibliotecas primarias clonando CDR3 sintéticas en marcos conservados aceptores

Aunque el método es particularmente interesante para recuperar diversidad natural, también puede aplicarse para la
 35 integración de diversidad sintética en Marcos conservados Aceptores. Se diseñaron secuencias de CDR3 sintéticas tanto para VH como para VL. Se tuvo en cuenta en el diseño la frecuencia de CDR con una longitud dada y la estrategia de diversificación (codones NNS, DVK, NVT o DVT) que permitirían 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 restos clave para mantener la estructura canónica de las CDR se mantuvieron constantes en el
 40 diseño de CDR3 para cadenas VK y VL. Para la cadena pesada, solamente se generaron CDR3 con hasta 10 posiciones diversificadas ya que el número de clones requerido para abarcar la diversidad codificada por CDR más largas está más allá de los límites prácticos de la eficacia de transformación.

Se sintetizaron oligonucleótidos degradados de diferente longitud usando codones aleatorios NNS, NVT, DVK o
 45 DVT. Para cada CDR H3, se sintetizaron dos oligonucleótidos que codificaban 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 bicatenarios que codifican las CDR diseñadas. Estos cebadores externos

contienen sitios de restricción de BsmBI para escisión posterior de la secuencia de CDR e inserción en los Marcos conservados Aceptores (Figura 12). Los fragmentos de ADN ensamblados se procesaron sin purificación en gel y se digirieron con BsmBI. Las extremidades escindidas así como ADN no digerido se retiraron usando perlas magnéticas recubiertas con estreptavidina. Los fragmentos de ADN digeridos se concentraron por precipitación con etanol y se ligaron en los vectores Aceptores de VH, VL o V λ pNDS correspondientes. Se purificaron productos de ligamiento y se concentraron para transformación en células *E. coli* TG1 electrocompetentes y se sembraron en placas de Bioensayo 2xTYAG (medio 2xTY que contiene ampicilina 100 μ g/ml y glucosa al 2 %). Después de incubación durante una 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 contenía glicerol al 50 % a la suspensión celular para obtener una concentración final de glicerol al 17 %. Se almacenaron alícuotas de la biblioteca a -80 °C. Se generó un total de 24 bibliotecas de cadena pesada primarias que representaban un total de $1,6 \times 10^{10}$ transformantes. De forma similar, se generaron 13 bibliotecas de cadena ligera primarias que representaban un total de $6,9 \times 10^9$ transformantes. Estas bibliotecas primarias contienen diversidad en la CDR H3 solamente ya que se combinan con un dominio VL de prueba. Se secuenció un total de 330 clones seleccionados aleatoriamente para determinar la calidad y diversidad de las bibliotecas. Todos los clones codificaban secuencias de dominio variable diferentes y >90 % estaban en fase. Esta baja frecuencia de secuencias que contenían desplazamientos en la fase de lectura está en claro contraste con resultados obtenidos tradicionalmente durante la construcción de bibliotecas de fragmentos de anticuerpos sintéticos usando enfoques de PCR solapante que son más propensos a la introducción de inserción, y se ha presentado frecuentemente pérdida significativa de clones funcionales (15-45 %).

La diversidad en estas bibliotecas primarias se restringió a la CDR H3 o CDR L3 ya que se combinan con una cadena VL o VH de prueba, respectivamente.

Se enumeran a continuación los cebadores usados para ensamblaje de CDR sintética.

5 H3_R_biot ATGATGTCTGGTCCAGCTCTCCGAGA (SEC ID N°: 210)
3 H3_M_biot CCACGTCATCCGATCCGTCTCCCCAATAATCCAT (SEC ID N°: 211)
3 H3_F_biot CCACGTCATCCGATCCGTCTCCCCAATAATCAAA (SEC ID N°: 212)
H3_4nnsF GCTGGCACGTCTCCGAGANNSNNSNNSNNSNTTTGATTATTGGGGGAGACG (SEC ID N°: 213)
H3_4nnsM GCTGGCACGTCTCCGAGANNSNNSNNSNNSNNSATGGATTATTGGGGGAGACG (SEC ID N°: 214)
H3_5nnsF GCTGGCACGTCTCCGAGANNSNNSNNSNNSNNSNTTTGATTATTGGGGGAGACG (SEC ID N°: 215)
H3_5nnsM GCTGGCACGTCTCCGAGANNSNNSNNSNNSNNSNNSATGGATTATTGGGGGAGACG (SEC ID N°: 216)
H3_6nnsF GCTGGCACGTCTCCGAGANNSNNSNNSNNSNNSNNSNTTTGATTATTGGGGGAGACG (SEC ID N°: 217)
H3_6nnsM GCTGGCACGTCTCCGAGANNSNNSNNSNNSNNSNNSNNSATGGATTATTGGGGGAGACG (SEC ID N°: 218)
H3_6dvkF GCTGGCACGTCTCCGAGADV KDV KDV KDV KDV KDV KDV KDTT TTTGATTATTGGGGGAGACG (SEC ID N°: 219)
H3_6dvkM GCTGGCACGTCTCCGAGADV KDV KDV KDV KDV KDV KDV KDKATGGATTATTGGGGGAGACG (SEC ID N°: 220)
H3_7dvkF GCTGGCACGTCTCCGAGADV KDV KDV KDV KDV KDV KDV KDV KDTT TTTGATTATTGGGGGAGACG (SEC ID N°: 221)
H3_7dvkM GCTGGCACGTCTCCGAGADV KDV KDV KDV KDV KDV KDV KDV KDKATGGATTATTGGGGGAGACG (SEC ID N°: 222)
H3_7nvtF GCTGGCACGTCTCCGAGANVTNV TNV TNV TNV TNV TNV TNV TTTTGATTATTGGGGGAGACG (SEC ID N°: 223)
H3_7nvtM GCTGGCACGTCTCCGAGANVTNV TNV TNV TNV TNV TNV TNV TNV TATGGATTATTGGGGGAGACG (SEC ID N°: 224)
H3_8nvtF GCTGGCACGTCTCCGAGANVTNV TNV TNV TNV TNV TNV TNV TNV TTTTGATTATTGGGGGAGACG (SEC ID N°: 225)
H3_8nvtM GCTGGCACGTCTCCGAGANVTNV TNV TNV TNV TNV TNV TNV TNV TATGGATTATTGGGGGAGACG (SEC ID N°: 226)
H3_9nvtF GCTGGCACGTCTCCGAGANVTNV TNV TNV TNV TNV TNV TNV TNV TTTTGATTATTGGGGGAGACG (SEC ID N°: 227)
H3_9nvtM GCTGGCACGTCTCCGAGANVTNV TNV TNV TNV TNV TNV TNV TNV TATGGATTATTGGGGGAGACG (SEC ID N°: 228)
H3_9dvtF GCTGGCACGTCTCCGAGADV TDV TDV TDV TDV TDV TDV TDV TTTTGATTATTGGGGGAGACG (SEC ID N°: 229)
H3_9dvtM GCTGGCACGTCTCCGAGADV TDV TDV TDV TDV TDV TDV TDV TATGGATTATTGGGGGAGACG (SEC ID N°: 230)
H3_10dvtF GCTGGCACGTCTCCGAGADV TDV TDV TDV TDV TDV TDV TDV TDV TTTTGATTATTGGGGGAGACG (SEC ID N°: 231)
H3_10dvtM CTGGCACGTCTCCGAGADV TDV TDV TDV TDV TDV TDV TDV TDV TATGGATTATTGGGGGAGACG (SEC ID N°: 232)
5 KL3_biot CCGGTGTAGCGAAGGCGTCTCAGCAG (SEC ID N°: 233)
3 KL3_biot TAGGGTGCCTTGATCGTCTCCCGAAGGTCGG (SEC ID N°: 234)

K_4nns GAAGGCGTCTCAGCAGNNSNNSNNSNNSCCGACCTTCGGGAGACG (SEC ID Nº: 235)
 K_5nns GAAGGCGTCTCAGCAGNNSNNSNNSNNSCCGNSACCTTCGGGAGACG (SEC ID Nº: 236)
 K_6nns GAAGGCGTCTCAGCAGNNSNNSNNSNNSNNSCCGNSACCTTCGGGAGACG (SEC ID Nº: 237)
 5 L44W_biot CGGTCAAGTCGCAATACGTCTCCAGCATGGGAT (SEC ID Nº: 238)
 5 L44Y_biot CGGTCAAGTCGCAATACGTCTCCAGCATATGAT (SEC ID Nº: 239)
 3 L_biot CAGGACCAGTCTCGTGAGGATCGTCTCAACAC (SEC ID Nº: 240)
 L44W_4nns CGTCTCCAGCATGGGATNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 241)
 L44Y_4nns CGTCTCCAGCATATGATNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 242)
 L44W_5nns CGTCTCCAGCATGGGATNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 243)
 10 L44Y_5nns CGTCTCCAGCATATGATNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 244)
 L44W_6nns CGTCTCCAGCATGGGATNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 245)
 L44Y_6nns CGTCTCCAGCATATGATNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 246)
 5 L51W_biot CGGTCAAGTCGCAATACGTCTCGAACATGGGAT (SEC ID Nº: 247)
 5 L51Y_biot CGGTCAAGTCGCAATACGTCTCGAACATATGAT (SEC ID Nº: 248)
 15 L51W_4nns CGTCTCGAACATGGGATNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 249)
 L51Y_4nns CGTCTCGAACATATGATNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 250)
 L51W_5nns CGTCTCGAACATGGGATNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 251)
 L51Y_5nns CGTCTCGAACATATGATNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 252)
 L51W_6nns CGTCTCGAACATGGGATNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 253)
 20 L51Y_6nns CGTCTCGAACATATGATNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 254)

Ejemplo 7: generación de bibliotecas secundarias

Para generar bibliotecas de scFv que portan diversidad en las cadenas tanto pesadas como ligeras, las bibliotecas de cadena ligera sintéticas Primarias se combinaron con las bibliotecas de cadenas pesadas sintéticas Primarias o las bibliotecas de cadenas pesadas naturales Primarias (Figura 13). Se preparó ADN de fagémido de cada biblioteca primaria y se dirigió con enzimas de restricción XhoI/NotI. Los fragmentos de ADN correspondientes a las cadenas ligeras y enlazadoras de las bibliotecas sintéticas Primarias se insertaron por ligamiento en los vectores de cadena pesada sintética o natural primaria digeridos. Como alternativa la secuencia de enlazador-VL también se amplificó con cebadores específicos antes de digestión con XhoI/NotI y ligamiento. Los productos de ligamiento se purificaron por extracción con fenol/cloroformo y precipitación antes de transformación en células *E. coli* TG1 electrocompetentes y siembra en placas de Bioensayo 2xTYAG (medio 2xTY que contiene ampicilina 100 µg/ml y glucosa al 2 %). Después de incubación durante una 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 contenía glicerol al 50 % a la suspensión celular para obtener una concentración final de glicerol del 17 %. Las alícuotas de las bibliotecas se almacenaron -80 °C. Para limitar el número de bibliotecas para recombinar, se agruparon por subclases de cadena (es decir, VH1, VH3, VH5, VK1, VK3, VL) y de este modo se evaluaron 9 combinaciones de biblioteca (es decir, VH1xVK1, VH1xVK3, VH1xVL1, VH3xVK1, VH3xVK3, VH3xVL1, VH5xVK1, VH5xVK3, VH5xVL1). El tamaño total de las bibliotecas sintéticas Secundarias (que portaban diversidad sintética tanto en el VH como en el VL) fueron $7,3 \times 10^9$ transformantes. El tamaño total de las bibliotecas naturales secundarias (que portan diversidad natural en el VH y diversidad sintética en el VL) fue de $1,5 \times 10^{10}$ transformantes.

Ejemplo 8: generación de bibliotecas de anticuerpos humanos que presentan un repertorio de CDRH3 derivado de una especie no humana.

Para utilizar fuentes alternativas de diversidad que permitirían explorar un espacio tridimensional diferente dentro del sitio de combinación de anticuerpos, se creó una biblioteca capturando las CDRH3 de ratones e introduciéndolas en una colección de marcos conservados de anticuerpos humanos. Para este enfoque se construyó una biblioteca de aceptores que contenía una colección de genes de VL con diversidad de CDR L3 sintética y se combinó con una colección de secuencias aceptoras que contenían una secuencia de ADN de relleno lista adecuada para clonación de restricción de Tipo IIS como se ha descrito en el Ejemplo 2. Esta biblioteca representa el punto de partida para la generación rápida de bibliotecas secundarias con múltiples fuentes de CDR H3 naturales (humanas así como no humanas) o sintéticas. En este ejemplo, se capturó la diversidad de CDR H3 natural a partir de ratones Balb/c sin tratamiento previo y ratones que se habían inmunizado con hIFN γ o hCCL5 (hRANTES).

La primera etapa fue la generación de bibliotecas aceptoras clonando una colección de VL que contenían diversidad de CDR L3 sintética en vectores de marco conservado de VH aceptores (Figura 14). Las secuencias de VL derivaron de las siete bibliotecas sintéticas Primarias descritas en el Ejemplo 6 por amplificación por PCR usando los cebadores 5'biot-VHdummy y 3'biot-fdtseq. Los fragmentos que contienen VL resultantes de aproximadamente 400 pb se digirieron usando XhoI/NotI y se purificaron en columnas de centrifugación para retirar cebadores y enzimas. De forma similar los vectores aceptores de VH pNDS que contenían un relleno de CDRH3 y una cadena ligera de prueba se digirieron con XhoI/NotI y Swal (cortando Swal dentro del VL de prueba) y se purificaron en columnas Chroma Spin TE con un punto de corte de 1000 pb para deshacerse del fragmento de VL de prueba. Los fragmentos de VL digeridos se ligaron después en los vectores aceptores de VH (Figura 14). Para limitar el número de bibliotecas para recombinar, se agruparon los vectores aceptores de VH y fragmentos de VL por subclases de

cadena (es decir, VH1, VH3, VH5, V κ 1, V κ 3, V λ 1) y por lo tanto se realizaron nueve combinaciones de bibliotecas (es decir, VH1xV κ 1, VH1xV κ 3, VH1xV λ 1, VH3xV κ 1, VH3xV κ 3, VH3xV λ 1, VH5xV κ 1, VH5xV κ 3, VH5xV λ 1). Los productos de ligamiento se transformaron en células *E. coli* TG1 electrocompetentes y se sembraron en placas de Bioensayo 2xTYAG (medio 2xTY que contenía ampicilina 100 µg/ml y glucosa al 2 %). Después de incubación durante una 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 al 50 % a la suspensión celular para obtener una concentración final de glicerol del 17 %. Se almacenaron alícuotas de las bibliotecas a -80 °C. El tamaño total de esta biblioteca de aceptores, que porta diversidad sintética en las CDR L3, fue de 1,9x10⁹ transformantes.

La siguiente etapa fue aislar secuencias de CDRH3 de una fuente no humana. Las células se aislaron del bazo de cinco ratones Balb/c sin tratamiento previo o inmunizados y se purificó el ARN total. Se obtuvo ADNc del ARN extraído por RT-PCR. Este ADNc se usó como molde para aislar y amplificar VH de ratón por PCR. Se realizó una serie de PCR usando 15 cebadores 5' diferentes (uno para cada subgrupo de VH de ratón) específicos para el inicio de la región FR1 y un grupo de cebadores 3' (cuatro cebadores que abarcan la región JH). Estas primeras PCR se agruparon y se purificaron en un gel de agarosa al 2 %. El ADN purificado actuó como un molde para realizar 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 de FokI para permitir la escisión precisa de la CDR H3 y clonarla 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 localiza en el sitio de escisión de FokI casi siempre difieren de la secuencia humana en una base, mientras que el extremo 3' coincidió entre estas dos especies. Las secuencias escindidas por FokI se encuadran en la Tabla 1 a continuación:

	Secuencias 5'			Secuencias 3'		
(SEC ID Nº: 281)	Humana:	GAGA		Humana:	IGGG	(SEC ID Nº: 285)
(SEC ID Nº: 282)	TTACTGTGC				GCCAGGGAA	
(SEC ID Nº: 282)	Ratón:	AAGA		Ratón:	TGGG	(SEC ID Nº: 286)
(SEC ID Nº: 283)	VH1 TTACTGTGC			JH1	GCGCAGGGA	
(SEC ID Nº: 283)	TTCTGTGC	AAGA		JH2	TGGG	(SEC ID Nº: 287)
(SEC ID Nº: 284)	VH2 CTACTGTGC	CAGA		JH3	TGGG	(SEC ID Nº: 288)
(SEC ID Nº: 282)	VH3- 16	AAGA		JH4	TGGG	(SEC ID Nº: 289)
	TTACTGTGC				GTCAGGGCA	

En consecuencia hubo que corregir la base durante la segunda de amplificación para generar extremos cohesivos que son compatibles con los extremos cohesivos generados tras la digestión de los Marcos conservados Aceptores. Se observó amplificación eficaz lo que sugiere que esta conversión se produjo fácilmente. En el extremo 3', las secuencias de ratón y humanas que se cortarán por las enzimas de restricción de Tipo IIS son idénticas, evitando de este modo cualquier problema de corrección.

Los cebadores para la segunda amplificación se biotinilaron en sus extremos 5' para facilitar las etapas de purificación corriente abajo. Los vectores aceptores se digirieron con BsmBI y se purificaron en columnas de Chroma Spin TE que tienen un punto de corte de 1000 pb. Después de digestión y purificación, las nueve combinaciones de biblioteca diferentes se agruparon en una relación equimolar para ligamiento de la CDR H3 de ratón capturada.

El ADN ligado se purificó por extracciones con fenol/cloroformo y se concentró por precipitación antes de transformación en células *E. coli* TG1 competentes y se sembraron en placas de Bioensayo 2xTYAG (medio 2xTY que contiene ampicilina 100 µg/ml y glucosa al 2 %). Después de incubación durante una 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 al 50 % a la suspensión celular para obtener una concentración final de glicerol del 17 %. Se almacenaron alícuotas de las bibliotecas a -80 °C. Se obtuvieron tres bibliotecas de tamaño similar: MnA, 2,5x10⁹ transformantes (que portaban una diversidad de marco conservado humano natural restringida, diversidad de ratones sin tratamiento previo en la CDR H3 y diversidad sintética en el CDR L3); MiB, 7,3x10⁷ transformantes (que portaban una diversidad de marco conservado humano natural restringida, diversidad inmunitaria de ratón contra hIFN γ en la CDR H3 y diversidad sintética en la CDR L3) y MiC, 1,8x10⁸ transformantes (que portan una diversidad de marco conservado humano natural restringida, diversidad inmunitaria de ratón contra hCCL5 en la CDR H3 y diversidad sintética en la CDR L3). Se analizaron clones aleatorios secuenciando para comprobar que la secuencia de CDR H3 se había reconstituido y que los puntos de unión entre las regiones CDR y las Marco conservadas eran correctos. Los resultados indicaron que todos los clones contenían secuencias de CDR H3 y que la fase de lectura estaba restaurada, codificando de este modo una cadena pesada variable de inmunoglobulina. Todas las CDR fueron diferentes y se asemejaron a las secuencias de CDR H3 de ratón típicas lo que indica que una gran diversidad de secuencias de CDR H3 de ratón de origen natural se había capturado por este enfoque. Además, el análisis de los perfiles de longitudes CDR H3 indicó que se capturó una

distribución Gaussiana en la biblioteca sin tratamiento previo que corresponde a la distribución esperada de longitudes en el repertorio de ratón normal. Por lo contrario, en las dos bibliotecas inmunitarias los perfiles fueron diferentes lo que sugiere que se había capturado un repertorio de CDR H3 diferente (Figura 15).

5 Cebadores usados para la amplificación de CDR H3 de ratones

1ª PCR

• cebadores 5':

m5 VH1	ATGCGGCCCAGCCGGCCATGGCCSAGGTTCAGCTBCAGCAGTC (SEC ID N°: 256)
m5 VH2	ATGCGGCCCAGCCGGCCATGGCCCAGGTTACCTGCAGCARTC (SEC ID N°: 257)
m5 VH3	ATGCGGCCCAGCCGGCCATGGCCCAGGTRCAGCTGAAGGAGTC (SEC ID N°: 258)
m5 VH4	ATGCGGCCCAGCCGGCCATGGCCCAGGTCCAACCTVCAGCARCC (SEC ID N°: 259)
m5 VH5	ATGCGGCCCAGCCGGCCATGGCCCAGATCCAGTTGGTVAGTC (SEC ID N°: 260)
m5 VH6	ATGCGGCCCAGCCGGCCATGGCCCAGGTGCAGCTGAAGSASTC (SEC ID N°: 261)
m5 VH7	ATGCGGCCCAGCCGGCCATGGCCGAGGTGCAGSKGGTGGAGTC (SEC ID N°: 262)
m5 VH8	ATGCGGCCCAGCCGGCCATGGCCGAAGTGAARSTTGAGGAGTC (SEC ID N°: 263)
m5 VH9	ATGCGGCCCAGCCGGCCATGGCCGAKGTSVAGCTTCAGGAGTC (SEC ID N°: 264)
m5 VH10	ATGCGGCCCAGCCGGCCATGGCCGAGGTGAASSTGGTGAATC (SEC ID N°: 265)
m5 VH11	ATGCGGCCCAGCCGGCCATGGCCGAGGTGAAGCTGRTGGARTC (SEC ID N°: 266)
m5 VH12	ATGCGGCCCAGCCGGCCATGGCCGARGTGAAGCTGRTGGAGTC (SEC ID N°: 267)
m5 VH13	ATGCGGCCCAGCCGGCCATGGCCGAAGTGCAGCTGTTGGAGAC (SEC ID N°: 268)
m5 VH14	ATGCGGCCCAGCCGGCCATGGCCGARGTGAAGCTTCTCSAGTC (SEC ID N°: 269)
m5 VH15	ATGCGGCCCAGCCGGCCATGGCCCARGTTACTCTGAAAGAGT (SEC ID N°: 270)

• cebadores 3':

m3 HJ1	CCTGAACCGCCGCCTCCGCTCGAGACGGTGACCGTGGTCCC (SEC ID N°: 271)
m3 HJ2	CCTGAACCGCCGCCTCCGCTCGAGACTGTGAGAGTGGTGCC (SEC ID N°: 272)
m3 HJ3	CCTGAACCGCCGCCTCCGCTCGAGACAGTGACCAGAGTCCC (SEC ID N°: 273)
m3 HJ4	CCTGAACCGCCGCCTCCGCTCGAGACGGTGACTGAGGTTCC (SEC ID N°: 274)

2ª PCR

• cebadores 5':

5 VHR_FOK_biot	GAGCCGAGGACACGGCCGGATGTTACTGTGCGAGA (SEC ID N°: 275)
----------------	--

• cebadores 3':

3'mJH1_Fok_biot	GGGGCGCAGGGACATCCGTCACCGTCTCCTC (SEC ID N°: 276)
3'mJH2_Fok_biot	GAGGAGACTGTGAGGGATGTGCCTTGCCCCA (SEC ID N°: 277)
3'JH1_Fok	GAGGAGACGGTGACGGATGTGCCCTGGCCCCA (SEC ID N°: 278)
3'mJH4_Fok_biot	GAGGAGACGGTGACGGATGTTCCTTGACCCCA (SEC ID N°: 279)

Ejemplo 9: rescate de fagos de las bibliotecas

- 10 Cada biblioteca Primaria y Secundaria se rescató independientemente de acuerdo con procedimientos de presentación en fagos convencionales brevemente resumidos a continuación. Se añadió un volumen de células de las alícuotas de biblioteca congeladas suficientes para abarcar al menos 10 veces la diversidad teórica de la biblioteca a 500 ml de 2xTYAG y se cultivaron a 37 °C con agitación (240 rpm) hasta que se alcanzó una DO600 de 0,3 a 0,5. El cultivo se superinfectó después con fago auxiliar MK13K07 y se incubó durante una hora a 37 °C (150 rpm). El medio se cambió después centrifugando las células a 2000 rpm durante 10 minutos, retirando el medio y resuspendiendo el sedimento en 500 ml de 2xTY-AK (ampicilina 100 µg/ml; kanamicina 50 µg/ml). El cultivo se dejó crecer después durante una noche a 30 °C (240 rpm). El cultivo se centrifugó a 4000 rpm durante 20 minutos para sedimentar las células. El sobrenadante se recogió y se añadió 30 % (v/v) de PEG 8000 (20 %)/NaCl 2,5 M para precipitar las partículas de fago incubando la mezcla 1 hora en hielo. Las partículas de fago se recogieron por centrifugación a 10.000 rpm durante 30 minutos y se resuspendieron en 10 ml de tampón TE (tris-HCl 10 mM pH 8,0; EDTA 1 mM). La solución resuspendida se centrifugó a 10.000 rpm para eliminar los residuos bacterianos y se repitió el procedimiento de precipitación. Después de la resuspensión final, el fago se valoró por infección de *E. coli* y absorción a 280 nm. El nivel de presentación de scFv en la superficie del fago también se evaluó por análisis de transferencia de Western usando un anticuerpo monoclonal anti c-myc. Se almacenó fago purificado de diferentes bibliotecas a -80 °C después de adición de glicerol a una concentración final de 15 % (p/v).

- Para usar un número manejable de bibliotecas durante los procedimientos de selección, el fago purificado se agrupó en 4 bibliotecas de trabajo: AA1 - Fagos de todas las bibliotecas de VH sintéticas Primarias; AB1 - Fagos de todas las bibliotecas de VL sintéticas Primarias; AC1 - Fagos de todas las bibliotecas de VH naturales primarias; AD1 - Fagos de todas las bibliotecas naturales Secundarias; AE1 - Fagos de todas las bibliotecas sintéticas Secundarias; MnA

- Bibliotecas con diversidad capturada de ratones sin tratamiento previo; MiB - Bibliotecas con diversidad capturada de ratones inmunizados con hIFN γ ; MiC - Bibliotecas con diversidad capturada de ratones inmunizados con hCCL5/RANTES.

5 **Ejemplo 10. Secuenciación de alto rendimiento de bibliotecas de anticuerpos.**

La calidad y diversidad de una biblioteca puede evaluarse por secuenciación de ADN de miembros de la biblioteca aleatorios. En la mayoría de los casos se secuencian varios cientos de clones que representan solamente una fracción muy pequeña de la biblioteca (menos de 1 de cada 10.000.000 de miembros de la biblioteca). Para analizar el rendimiento de los métodos proporcionados en el presente documento, se usó tecnología de secuenciación de siguiente generación para analizar un número más representativo de miembros de la biblioteca. Se usó ADN aislado de la biblioteca AE1 como un molde para secuenciación de alto rendimiento usando un instrumento Analizador de Genoma de Illumina. Este sistema de secuenciación de ADN de siguiente generación permite que se lean miles de millones de bases en varios días. Las lecturas de secuenciación son relativamente cortas (aproximadamente 70 bases) pero perfectamente compatibles con el diseño de la biblioteca de los inventores. Como la diversidad está restringida a las regiones CDR3 una lectura de 70 bases es suficiente para abarcar la CDRH3 y parte de la región marco conservada 3 para identificación de la familia de VH. Esta tecnología se ha aplicado para secuenciar varios millones de regiones CDRH3 de la biblioteca AE1. 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 podrían asignarse inequívocamente a una familia de VH (VH1, VH3 y VH5) y determinarse la representación de las familias de VH en la biblioteca AE1 (41 % VH1; 30 % VH3; 29 % VH5). Un hallazgo importante fue que la proporción de insertos en fase varió entre el 88 y el 91 %. Estos datos confirmaron de una manera mucho más estadística los resultados de secuenciación de las 24 bibliotecas sintéticas de VH primarias descritas en el Ejemplo 6. Este conjunto combinado de datos de secuenciación demuestra que el proceso de clonación de restricción de Tipo IIs usado en este método es muy robusto, lo que conduce a una inserción eficaz y productiva en las 24 construcciones de biblioteca independientes realizadas para generar la diversidad de VH de la biblioteca AE1.

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 bibliotecas de alta calidad y alta diversidad de una manera reproducible y robusta.

Ejemplo 11: selecciones de presentación en fagos usando bibliotecas secundarias

Selecciones de fase líquida frente a interferón gamma humano (hIFN): se bloquearon alícuotas de bibliotecas de fagos AD1 y AE1 (10^{11} - 10^{12} Ufp) con PBS que contenía leche desnatada al 3 % (p/v) durante una hora a temperatura ambiente en un mezclador rotatorio. Después se deseleccionaron fagos bloqueados en perlas magnéticas de estreptavidina (Dynal M-280) durante una hora a temperatura ambiente en un mezclador rotatorio. Los fagos deseleccionados se incubaron después con hIFN γ biotinilado *in vivo* (100 nM) durante dos horas a temperatura ambiente en un mezclador rotatorio. Las perlas se capturaron usando un soporte magnético seguido de cuatro lavados con PBS/Tween 20 0,1 % y 3 lavados con PBS. Las perlas se añadieron después directamente a 10 ml de células TG1 con crecimiento exponencial y se incubaron durante una hora a 37 °C con agitación lenta (100 rpm). Una alícuota de las TG1 infectadas se diluyó en serie para valorar la producción de selección. Las TG1 infectadas restantes se centrifugaron a 3000 rpm durante 15 minutos y se resuspendieron en 0,5 ml de 2xTYAG (medio 2xTY que contenía ampicilina 100 μ g/ml y glucosa al 2 %) y se extendieron en placas de Bioensayo de agar 2xTYAG. Después de incubación durante una noche a 30 °C, se añadieron 10 ml de 2xTYAG a las placas y las células se rasparon de las superficies y se transfirieron a un tubo de polipropileno de 50 ml. Se añadió 2xTYAG que contenía glicerol al 50 % a la suspensión celular para obtener una concentración final de glicerol al 17 %. Se mantuvieron alícuotas del ciclo de selección a -80 °C. Las producciones de fagos se titularon después de cada ciclo y el aumento progresivo de las producciones indicó que se estaba produciendo el enriquecimiento de clones específicos para la diana (Figura 16).

Selecciones por panning frente al anticuerpo monoclonal de rata 5E3:

Se recubrieron inmunotubos con 5E3 a 10 μ g/ml en PBS durante una noche a 4 °C y se recubrieron inmunotubos para selección de fagos con un anticuerpo de rata irrelevante en las mismas condiciones. Después de lavar los inmunotubos se bloquearon con PBS que contenía leche desnatada al 3 % (p/v) durante una hora a temperatura ambiente. Se bloquearon alícuotas de bibliotecas de fagos AD1 y AE1 (10^{11} - 10^{12} Ufp) con PBS que contenían leche desnatada al 3 % (p/v) durante una hora a temperatura ambiente en un mezclador rotatorio. Después se deseleccionaron fagos bloqueados en los inmunotubos recubiertos con un anticuerpo de rata irrelevante durante una hora a temperatura ambiente en un mezclador rotatorio. Después se transfirieron fagos deseleccionados a los inmunotubos recubiertos con 5E3 y se incubaron durante dos horas a temperatura ambiente en un mezclador rotatorio. Los tubos se lavaron cinco veces con PBS/Tween 20 0,1 % y 3 veces con PBS. Los fagos se eluyeron con TEA 100 mM durante 10 minutos y se neutralizaron con Tris HCl 1 M pH 7,5. Se añadieron fagos a 10 ml de células TG1 con crecimiento exponencial y se incubó durante una hora a 37 °C con agitación lenta (100 rpm). Se diluyó en serie una alícuota de las TG1 infectadas para valorar la producción de selección. Las TG1 infectadas restantes se centrifugaron a 3000 rpm durante 15 minutos y se resuspendieron en 0,5 ml de 2xTYAG (medio

2xTY que contenía ampicilina 100 µg/ml y glucosa al 2 %) y se extendió en placas de Bioensayo de agar 2xTYAG. Después de incubación durante una 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 contenía glicerol al 50 % a la suspensión celular para obtener una concentración final de glicerol del 17 %. Se mantuvieron alícuotas del ciclo de selección a -80 °C. Se realizaron ciclos de selección alternando entre 5E3 de rata y una versión quimérica de 5E3 en los que la región variable se fusionó con dominios constantes de ratón. Estos ciclos alternantes se realizaron para enriquecer con respecto a clones específicos para la región variable de 5E3 y generar anticuerpos antiidiotípicos. Se trituraron producciones de fagos después de cada ciclo y el aumento progresivo de las producciones indicó que se estaba produciendo el enriquecimiento de clones específicos para la diana (Figura 17).

Fagos rescatados: se añadieron 100 µl de suspensión celular obtenida de ciclos de selección previos a 20 ml de 2xTYAG y se cultivaron a 37 °C con agitación (240 rpm) hasta que se alcanzó una DO600 de 0,3 a 0,5. El cultivo se superinfectó después con $3,3 \times 10^{10}$ fagos auxiliares MK13K07 y se incubó durante una hora a 37 °C (150 rpm). El medio se cambió después centrifugando las células a 2000 rpm durante 10 minutos, retirando el medio y resuspendiendo el sedimento en 20 ml de 2xTY-AK (ampicilina 100 µg/ml; kanamicina 50 µg/ml). El cultivo se dejó crecer después durante una noche a 30 °C (240 rpm).

Rescate de fagos monoclonales para ELISA: se seleccionaron clones individuales en una placa de microtitulación que contenía 150 µl de medio 2xTYAG (glucosa al 2 %) por pocillo y se cultivaron a 37 °C (100-120 rpm) durante 5-6 h. Se añadió fago auxiliar M13K07 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 de cultivo, las placas se centrifugaron a 3.200 rpm durante 10 minutos. El sobrenadante se retiró cuidadosamente, las células se resuspendieron en 150 µl de medio 2xTYAK y se cultivaron durante una noche a 30 °C (120 rpm). Para el ELISA, los fagos se bloquean añadiendo 150 µl de PBS a concentración 2x que contiene leche en polvo desnatada al 5 % seguido de una hora de incubación a temperatura ambiente. Las placas se centrifugaron después 10 minutos a 3000 rpm y el fago que contenía sobrenadante se usó para el ELISA.

ELISA de fagos: se recubrieron placas de ELISA (Maxisorb, NUNC) durante una noche con hIFN γ 2 µg/ml en PBS o 5E3 de rata 2 µg/ml en PBS. Se recubrieron placas de control con BSA 2 µg/ml o un anticuerpo monoclonal de rata irrelevante. Las placas se bloquearon después con leche desnatada al 3 % / PBS a temperatura ambiente durante 1 h. Las placas se lavaron 3 veces con PBS Tween 20 0,05 % antes de transferir los sobrenadantes de fago prebloqueado e incubación durante una hora a temperatura ambiente. Las placas se lavaron después 3 veces con PBS Tween 20 0,05 %, 50 µl de leche desnatada al 3 % / PBS que contenía anticuerpo anti M13 conjugado con (HRP) (Amersham, diluido 1:10.000) a cada pocillo. Después de incubación a temperatura ambiente durante 1 h, las placas se lavaron 5 veces con PBS Tween 20 0,05 %. El ELISA se reveló después añadiendo 50 µl de TMB (Sigma) y 50 µl de H₂SO₄ 2 N para detener la reacción. Se leyó la intensidad de la absorción a 450 nm. Pudieron identificarse clones específicos para hIFN γ y las tasas de acierto variaron entre el 10 % y el 30 % después del tercer ciclo de selección. También pudieron identificarse clones específicos para la región variable de 5E3 y las tasas de aciertos variaron entre el 7 y el 48 % después del tercer ciclo de selección.

Secuenciación de clones de fagos: se cultivaron clones individuales en 5 ml de medio 2xTYAG (glucosa al 2 %) por pocillo y se cultivaron a 37 °C (120 rpm) durante una noche. Al día siguiente se purificó ADN de fagémido y se usó para secuenciación de ADN usando un cebador específico para pNDS1: mycseq, 5'-CTCTTCTGAGATGAGTTTGTG. (SEC ID N°: 255).

Purificación de scFv a gran escala: se inoculó un cultivo de partida de 1 ml de 2xTYAG con una única 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 dejó crecer durante una noche a 30 °C con agitación vigorosa (300 rpm).

Al día siguiente el cultivo se indujo añadiendo 400 µl de IPTG 1 M y se continuó la incubación durante 3 horas adicionales. Las células se recogieron por centrifugación a 5.000 rpm durante 10 minutos a 4 °C. Las células sedimentadas se resuspendieron en 10 ml de tampón TES helado complementado con inhibidores de proteasa como se ha descrito anteriormente. Se consiguió choque osmótico añadiendo 15 ml de tampón TES diluido 1:5 e incubación durante 1 hora en hielo. Las células se centrifugaron a 10.000 rpm durante 20 minutos a 4 °C para sedimentar residuos celulares. El sobrenadante se transfirió cuidadosamente a un tubo nuevo. Se añadió imidazol al sobrenadante hasta 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 2.000 rpm durante 5 minutos y el sobrenadante se retiró cuidadosamente. La resina sedimentada se resuspendió en 10 ml de tampón de lavado 1 frío (4 °C) (NaH₂PO₄ 50 mM, NaCl 300 mM, imidazol 10 mM, pH hasta 8,0). La suspensión se añadió a una columna polyprep (Biorad). Se usaron 8 ml de Tampón de Lavado 2 frío (NaH₂PO₄ 50 mM, NaCl 300 mM, imidazol 20 mM, pH hasta 8,0) para lavar la columna por flujo gravitatorio. Los scFv se eluyeron de la columna con 2 ml de tampón de Elución (NaH₂PO₄ 50 mM, NaCl 300 mM, imidazol 250 mM, pH hasta 8,0). Las fracciones se analizaron por absorción a 280 nm y se agruparon fracciones que contenían proteínas antes del

intercambio de tampón en una columna de desalación 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 separaron en alícuotas y se almacenaron a -20 °C y a 4 °C.

5 Ejemplo 12. Análisis de perfiles de CDR3 obtenidos después de selección usando secuenciación de alto rendimiento.

Usando tecnología de secuenciación de siguiente generación como se describe en el Ejemplo 10, se analizó la distribución de las longitudes de CDR H3 dentro de cada familia de VH en las bibliotecas AE1 y AD1 así como en la producción obtenida después del tercer ciclo de selección. Los perfiles de las bibliotecas AE1 y AD1 son claramente diferentes (Figura 18). La distribución de longitudes de CDR H3 en la biblioteca AE1 corresponde al diseño de biblioteca pretendido, variando las longitudes entre 9 y 15 aminoácidos. Por el contrario, se han descubierto CDR H3 mucho más largas de hasta 22 aminoácidos en la biblioteca AD1, y el perfil corresponde a la distribución de longitudes observada en repertorios naturales humanos. Estos resultados confirman que se ha capturado un repertorio de CDR H3 natural humano durante la construcción de la biblioteca AD1. Un análisis similar realizado después de tres ciclos de selección frente a 5E3 reveló que se seleccionaron perfiles de longitudes de CDR H3 completamente diferentes. En particular, se pudo observar un enriquecimiento drástico 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 las dos bibliotecas después de selección contra la misma diana. Además, este análisis demuestra que, usando la presente invención, podrían capturarse CDR H3 largas que son muy difíciles de abarcar usando diversidad sintética en marcos conservados humanos seleccionados y seleccionarse.

25 Ejemplo 13. Evaluación de scFv identificados en ensayos de unión.

Se ensayaron preparaciones de scFv purificadas de clones que tenían diferentes secuencias y que se identificaron como positivos contra la región variable de 5E3 con respecto a unión frente a 5E3 quimérico en un ELISA de respuesta a dosis. Estas preparaciones también se ensayaron frente a un anticuerpo de ratón irrelevante (1A6). Se recubrieron placas de ELISA (Maxisorb, NUNC) durante una noche con 5E3 de ratón 2 µg/ml en PBS. Se recubrieron placas de control con anticuerpo monoclonal 1A6 2 µg/ml. Las placas se bloquearon después con leche desnatada al 3 % / PBS a temperatura ambiente durante 1 h. Las placas se lavaron tres veces con PBS Tween 20 0,05 % antes de añadir diferentes concentraciones de scFv purificado e incubación durante una hora a temperatura ambiente. Las placas se lavaron después 3 veces con PBS Tween 20 0,05 %, 50 µl de leche desnatada al 3 % / PBS que contenía anticuerpo anti myc conjugado con (HRP) a cada pocillo. Después de incubación a temperatura ambiente durante 1 h, las placas se lavaron 5 veces con PBS Tween 20 al 0,05 %. El ELISA se reveló después añadiendo 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 que se dirigen contra las regiones variables de 5E3 (Figura 19).

De forma similar, se ensayaron preparaciones de scFv purificadas de clones que tienen secuencias diferentes y que se identificaron en ELISA de fago como agentes de unión contra hIFNγ con respecto a unión frente a hIFNγ en un experimento de respuesta a dosis. Se recubrieron placas de ELISA (Maxisorb, NUNC) durante una noche con hIFNγ 2 µg/ml en PBS y se recubrieron placas de control con BSA 2 µg/ml en PBS. Después se bloquearon placas con leche desnatada al 3 % / PBS a temperatura ambiente durante 1 h. Las placas se lavaron tres veces con PBS Tween 20 0,05 % antes de añadir una concentración diferente de scFv purificado e incubación durante una hora a temperatura ambiente. Las placas se lavaron después 3 veces con PBS Tween 20 0,05 %, 50 µl de leche desnatada al 3 % / PBS que contenía anticuerpo anti myc conjugado con (HRP) a cada pocillo. Después de incubación a temperatura ambiente durante 1 h, las placas se lavaron 5 veces con PBS Tween 20 0,05 %. El ELISA se reveló después añadiendo 50 µl de sustrato de TMB y 50 µl de H₂SO₄ 2 N 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 con hIFNγ de una manera dependiente de dosis y proporcionaron una señal muy buena en comparación con un scFv A6 de control positivo que tiene una alta afinidad por hIFNγ (Figura 20).

55 Ejemplo 14: inhibición por scFv de la expresión del gen indicador inducida por interferón gamma

Se produjo un panel de scFv seleccionados específicos para hIFNγ y se purificó como se ha descrito anteriormente y se ensayó con respecto a la capacidad para bloquear la actividad biológica de hIFNγ. Se transfectó un gen indicador (luciferasa de luciérnaga), conducido por el promotor GBP1 inducible por IFNγ, en la línea celular de melanoma humano, Me67.8. Se incubaron diversas concentraciones de scFv con hIFNγ 2 ng/ml y después se añadieron al cultivo celular. Después de un tiempo de incubación de 6 horas, se realizó el ensayo indicador de luciferasa y se midió la intensidad de la luminiscencia. La actividad se comparó con un scFv aislado de otra biblioteca de anticuerpos scFv humana construida por captura tradicional de los repertorios de VH/VL de donantes humanos (clon G9). Los datos muestran que scFv aislado de bibliotecas de diversidad humana sintéticas o naturales (AE1 y AD1) fueron capaces de neutralizar la actividad biológica de hIFNγ de una manera dependiente de dosis (Figura 21). El potencial de neutralización de estos scFv fue superior al clon G9 de scFv de referencia.

Ejemplo 15: inhibición por scFv de la expresión de MHC de clase II inducida por interferón gamma

Se implantó un ensayo citométrico de flujo para identificar anticuerpos IgG completamente humanos, o fragmentos de los mismos, capaces de bloquear la expresión de moléculas del MHC de clase II inducidas por IFN γ . Después de la siembra de células Me67.8 en placas, se añadió IFN γ humano recombinante 5 ng/ml a cultivos en presencia de diversas concentraciones de anticuerpos monoclonales anti IFN γ completamente humanos candidatos. Después de 48 horas en cultivo, las células se tiñeron con anticuerpo anti MHC de clase II humano marcado con fluorescencia (HLA-DR) y se analizaron usando un FACSCalibur®. Por lo tanto, se midió la CI₅₀ (en la que el 50 % de la expresión del MHC de clase II inducida por IFN γ está inhibida, es decir, concentración inhibidora al 50 %) para cada anticuerpo candidato.

Se produjeron scFv purificados completamente humanos como se ha descrito anteriormente. El efecto de scFv seleccionados en la expresión de MHC de clase II inducida por IFN γ en células de melanoma se evaluó usando el ensayo basado en células citométrico de flujo descrito anteriormente. Estos scFv inhibieron la expresión del MHC II inducido por IFN γ en células de melanoma (Figura 22).

Ejemplo 16: cambio de formato de scFv a formato de IgG

La secuencia de V_H y V_L de scFv seleccionados se amplificaron con oligonucleótidos específicos introduciendo una secuencia líder y un sitio de restricción HindIII en el extremo 5'. Se introdujo un sitio Apal en el extremo 3' de la cadena pesada mientras que se introdujeron un sitio AvrII y uno BsiWI en el extremo 3' de la secuencia de cadena ligera lambda o kappa, respectivamente. Las secuencias de V_H amplificadas se digirieron con HindIII/Apal y se clonaron en el vector de expresión de pCon_gammal (LONZA, Basilea, Suiza). Las secuencias de V_L lambda amplificadas se digirieron con HindIII/AvrII y se clonaron en el vector de expresión pCon_lambda2 y las secuencias de V_L kappa amplificadas se digirieron con HindIII/BsiWI y se clonaron en el vector de expresión pCon_kappa (LONZA, Basilea, Suiza). Las construcciones se verificaron secuenciando antes de la transfección en células de mamífero.

Las secuencias de ADNc de V_H y V_L en sus vectores de expresión apropiados se transfectaron a células de mamífero usando el Reactivo de Transfección Eugene 6 (Roche, Basilea, Suiza). Brevemente, las células Pico se cultivaron en placas de 6 pocillos a una concentración de 6×10^5 células por pocillo en 2 ml de medio de cultivo que contenía suero bovino fetal. Los vectores de expresión, que codificaban las secuencias de V_H y V_L candidatas, se cotransfectaron en las células usando el Reactivo de Transfección Eugene 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 sin suero nuevo a las células y se cultivaron durante tres días a 37 °C. Después de un periodo de cultivo de tres días, el sobrenadante se recogió para IgG purificada en columnas de flujo rápido de proteína G-Sepharose 4B (Sigma, St. Louis, MO) de acuerdo con las instrucciones del fabricante. Brevemente, se incubaron sobrenadantes de células transfectadas durante una noche a 4 °C con tampón de unión de IgG ImmunoPure (G) (Pierce, Rockford IL). Después se pasaron muestras sobre las columnas de flujo rápido de Proteína G-Sepharose 4B y el IgG se purificó posteriormente usando tampón de elución. La fracción de IgG eluida se dializó después contra PBS y el contenido de IgG se cuantificó por absorción a 280 nm. Se verificaron la pureza e integridad de IgG por SDS-PAGE.

Ejemplo 17: inhibición de IgG de la actividad biológica de interferón gamma

Se cambió el formato de dos scFv, AE1-4-R3-P2E4 (2E4) y A2-AD1-R4P1A9 (1A9), que tenían actividad inhibidora confirmada contra hIFN γ en ensayos funcionales en IgG como se ha descrito en el Ejemplo 16 y se ensayaron en el ensayo de 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 como 2E4 podrían neutralizar la actividad de hIFN γ con CI₅₀ de 42 nM y 10 nM, respectivamente mientras que un IgG de control negativo (NI-0701) no tuvo ningún efecto en este ensayo. Por lo tanto estos dos candidatos aislados de bibliotecas de diversidad tanto sintética como natural pudieron cambiarse de formato a IgG completo y presentan actividad neutralizante contra la diana seleccionada.

Ejemplo 18: desarrollo de un ensayo farmacocinético para la detección de 5E3 en suero de ratón.

Se cambiaron dos formatos candidatos a scFv AD15E3R3P1_A4 y AD25E3R3P1_G11 que se unen específicamente con anticuerpo monoclonal de ratón 5E3 (Figura 19) a IgG humano completo como se ha descrito en el Ejemplo 16. La especificidad de los IgG correspondientes DA4 y G11 se confirmó en ELISA frente a 5E3 de ratón y una versión química de este anticuerpo monoclonal en la que las regiones variables de ratón se han fusionado con regiones IgG constantes de rata. Los resultados mostrados en la Figura 24 demuestran que los IgG DA4 y G11 son específicos para la región variable de 5E3 ya que se unen tanto con ratón como con 5E3 de rata química y no con controles de isotipo de rata y de ratón. Estos dos anticuerpos monoclonales se usaron para desarrollar un ensayo para la cuantificación de 5E3 en suero de ratón para estudios farmacocinéticos. Se añadió a varias diluciones de suero de ratón anticuerpo 5E3 de ratón 5 μ g/ml y se diluyó en serie de tal manera que la concentración en suero se mantuvo constante en toda la serie de diluciones. Se recubrieron placas Maxisorb (Nunc, Dinamarca) durante una noche con 1 μ g/ml de IgG DA4 o IgG G11. Después de bloquear con PBS; se añadió una serie de diluciones de BSA

1 % de las preparaciones de suero con adiciones a los pocillos. Después de incubar y lavar, la señal se reveló usando un anticuerpo monoclonal de cadena ligera Kappa anti ratón acoplado con peroxidasa de rábano rusticano (HRP) y un sustrato fluorescente (Amplex red; Invitrogen). Los resultados muestran que ambos anticuerpos pueden usarse para detectar específicamente el anticuerpo 5E3 monoclonal de ratón en suero de ratón (Figura 25). El límite de detección de 5E3 de ratón en suero fue de aproximadamente 200 ng/ml y el ensayo no se vio afectado significativamente por la concentración en suero lo que indica que IgG DA4 e IgG G11 son altamente específicos para 5E3 de ratón y no se unen con otra inmunoglobulina de ratón. Estos experimentos demuestran que podrían aislarse anticuerpos antiidiotípicos altamente específicos de las bibliotecas naturales o sintéticas AE1 y AD1.

Ejemplo 19: selección de fagos usando bibliotecas que contienen diversidad de CDRH3 capturada de ratones sin tratamiento previo e inmunizados.

Las bibliotecas MnA, MiB y MiC descritas en los Ejemplos 8 y 9 se usaron en paralelo para selecciones de fagos frente a hIFN γ siguiendo el procedimiento descrito en el Ejemplo 11. Durante el proceso de selección se observó un enriquecimiento de fagos similar (Figura 26).

Expresión de scFv en formato de placa de microtitulación: se seleccionaron clones individuales en una placa de microtitulación que contenía 150 μ l de 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 μ l de medio 2xTYA que contenía IPTG 1 mM. Las placas se incubaron durante una noche a 30 °C con agitación (100 rpm). Después del cultivo, las placas se centrifugaron a 3.200 rpm durante 10 minutos y el sobrenadante se transfirió cuidadosamente a una placa que contenía PBS concentrado 2x que contenía leche en polvo desnatada al 5 % para bloqueo.

ELISA de scFv: se recubrieron placas de ELISA (Maxisorb, NUNC) durante una noche con hIFN γ 2 μ g/ml en PBS. Se recubrieron placas de control con BSA recombinante 2 μ g/ml (Sigma). Las placas se bloquearon después con leche desnatada al 3 % / PBS a temperatura ambiente durante 1 h. Las placas se lavaron 3 veces con PBS Tween 20 0,05 % antes de transferir los sobrenadantes de scFv prebloqueados e incubación durante una hora a temperatura ambiente. Las placas se lavaron después 3 veces con PBS Tween 20 0,05 %, 50 μ l de leche desnatada al 3 % / PBS que contenía anticuerpo anti cMyc conjugado con (HRP) (diluido 1:5.000) a cada pocillo. Después de incubación a temperatura ambiente durante 1 h, las placas se lavaron 5 veces con PBS Tween 20 0,05 %. El ELISA se reveló después añadiendo 50 μ l de Amplex Red (Invitrogen). Se midió la intensidad de fluorescencia a 590 nm tras excitación a 530 nm. La frecuencia de aciertos que proporcionan una señal de la mitad de la intensidad del clon A6 de control se evaluó después de cada ciclo de selección para las tres bibliotecas (Figura 27). La tasa de aciertos obtenida con la biblioteca de MiB fue drásticamente superior en comparación con las otras dos bibliotecas y el nivel promedio de la señal fue superior para los clones derivados de la biblioteca de MiB, lo que indica que se enriquecieron scFv de mayor afinidad (Figura 28). Para confirmar esta observación, se secuenciaron clones positivos, se expresaron a mayor escala y se purificaron para ensayarse en experimentos de unión de respuesta a dosis de acuerdo con el Ejemplo 13. Todos los scFv derivados de la biblioteca de MiB tuvieron una mayor afinidad aparente por hIFN γ que los aislados de la biblioteca de MnA sin tratamiento previo (Figura 29). Los resultados indican que el repertorio de CDRH3 de ratones inmunizados con una proteína podría capturarse en un contexto de marco conservado de anticuerpo humano de una manera productiva para generar a mayor frecuencia fragmentos de anticuerpo humanos de alta afinidad. Las bibliotecas generadas usando la presente invención representan por lo tanto un medio potente para generar anticuerpos con potencial terapéutico.

Ejemplo 20: identificación de fragmentos de ADN de relleno que pueden codificar CDRH3 funcional

Se usó un enfoque combinatorio para identificar fragmentos de ADN de relleno que cumplen los siguientes criterios: 1) incluyen dos sitios de restricción de Tipo IIS; 2) mantienen la fase de lectura entre las regiones FR3 y FR4 y 3) codifican una CDR3 de dominio variable pesado que permite el plegamiento y la expresión de un dominio variable de anticuerpo. La presencia de los dos sitios de enzimas de restricción define parcialmente la secuencia de la CDRH3 al nivel de proteínas. Para maximizar las probabilidades de encontrar secuencias que pudieran ajustarse a esta restricción, se diseñaron oligonucleótidos para sintetizar una colección de fragmentos de relleno que contenían dos sitios de restricción de BsmBI e introducir diversidad en uno o dos codones para explorar múltiples soluciones para dos longitudes de CDRH3 definidas (Figura 31). Estas dos colecciones de fragmentos de relleno en fase (SIF) se generaron por PCR de ensamblaje usando los siguientes cebadores:

5 VHstufIF1 ATTACTGTGCGAGAGGAGACGNSNCGTCTCTTGGGGCCAGGGAAC (SEC ID N°: 290)
 5 VHstufIF2 ATTACTGTGCGAGAGGAGACGNCGTCTCTTGGGGCCAGGGAACCCT (SEC ID N°: 291)
 3 VHIF ttatgtgtataggGTTCCCTGGCCCCAAGAGACG (SEC ID N°: 292)
 5 VHIF gtgatctgtacctATTACTGTGCGAGAGGAGACG (SEC ID N°: 293)

Los SIF1 y SIF2 amplificados se digirieron con BsmBI y se clonaron en el marco conservado aceptor de vector fagémido pNDS_VH3-23-VK dummy previamente digerido con BsmBI. Los productos de ligamiento se transformaron en células *E. coli* TG1 electrocompetentes y se sembraron en placas de Bioensayo 2xTYAG (medio 2xTY que

contenía ampicilina 100 µg/ml y glucosa al 2 %). Después de incubación durante una 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. Estas pequeñas bibliotecas de diversidad llamadas IF1 e IF2 se rescataron usando Hiperfago (Hust M *et al.*, Biotechniques Sep 2006; 41(3): 335-42) de modo que solamente los miembros de la biblioteca que codificaban scFv compatibles con la expresión como una proteína de fusión pIII y ensamblaje en una partícula de fago pueden conducir a la producción de fagos. El fago rescatado se usó directamente para infectar células TG1 que después se sembraron en placas de Bioensayo 2xTYAG. Después de raspar las células, se realizó un segundo ciclo de rescate e infección y se secuenciaron colonias individuales para identificar secuencias que se enriquecieron en este proceso de selección. Se identificó un total de 8 secuencias independientes de SIF2 y 15 de SIF1 en los clones seleccionados (Figuras 32 y 33). Cada clon se expresó y se purificó independientemente como un scFv usando purificación de scFv a gran escala como se ha descrito en el Ejemplo 11 para confirmar que la secuencia de SIF era compatible con la producción de un scFv. El rendimiento de la producción de scFv se determinó y la integridad de la proteína se evaluó por SDS-PAGE. Usando estos parámetros se seleccionaron 6 clones que contenían secuencias de SIF diferentes y el ADN de vector correspondiente se preparó para ensayar la eficacia de clonación de Tipo IIS y la capacidad de BsmBI para digerir ambos sitios en el contexto de estas secuencias de SIF. La secuencia del clon SIF_2b8 se seleccionó y se integró en todas las secuencias Aceptoras de marcos conservados de VH.

Ejemplo 21: generación y limpieza de bibliotecas aceptoras que contienen un SIF

Los aceptores de VH SIF se combinaron después con bibliotecas primarias sintéticas de VL como se ha descrito en el Ejemplo 8 para generar bibliotecas aceptoras en las que la diversidad de CDRH3 puede introducirse por digestión del SIF. Las secuencias de VL derivaron de las siete Bibliotecas Sintéticas Primarias descritas en el Ejemplo 6 por amplificación por PCR usando los cebadores 5'biot-VHdummy y 3'biot-fdtseq. Los fragmentos que contienen VL resultantes de aproximadamente 400 pb se digirieron usando XhoI/NotI y se purificaron en columnas de centrifugación para retirar cebadores y enzimas. De forma similar, los vectores aceptores de VH pNDS que contenían un relleno de SIF y una cadena ligera de prueba se digirieron con XhoI/NotI y Swal (cortando Swal dentro del VL de prueba) y se purificaron en columnas Chroma Spin TE con un punto de corte de 1000 pb para deshacerse del fragmento de prueba de VL. Los fragmentos de VL digeridos se ligaron después en los vectores aceptores de VH SIF. Los productos de ligamiento se transformaron en células *E. coli* TG1 electrocompetentes y se sembraron en placas de Bioensayo 2xTYAG (medio 2xTY que contenía ampicilina 100 µg/ml y glucosa al 2 %). Después de incubación durante una 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 al 50 % a la suspensión celular para obtener una concentración final de glicerol del 17 %. Las alícuotas de las bibliotecas aceptoras se almacenaron a -80 °C. El tamaño total de esta biblioteca aceptora, que portaba diversidad sintética en la CDR L3, fue de $4,3 \times 10^9$.

Las bibliotecas se rescataron usando Hiperfago y se usaron para infección de TG1 como se ha descrito anteriormente para retirar secuencias fuera de fase y por lo tanto enriquecer las bibliotecas aceptoras para insertos en fase. Para evaluar la eficacia del proceso, se seleccionaron 30 clones individuales de tres bibliotecas y se secuenciaron antes y después del procedimiento de limpieza y se determinó la frecuencia de las secuencias en fase. Los resultados mostrados posteriormente en la Tabla 2 indican que la frecuencia de las secuencias en fase aumentó significativamente por este proceso y en dos bibliotecas las 30 secuencias estuvieron en fase. Este proceso y el uso de SIF en las bibliotecas Aceptoras aumentaron la funcionalidad de la biblioteca Aceptora haciéndola un mejor receptáculo para diversidad de CDRH3.

Tabla 2. Frecuencia de secuencias en fase de las bibliotecas aceptoras de SIF antes y después del proceso de limpieza

	Antes de la limpieza	Después de la limpieza
Bibliotecas	secuencia en fase (%)	
VH1-2-VK1	76	100
VH1-2-VK3	77	100
VH1-2-Vλ	87	94

Otras realizaciones

Aunque la invención se ha descrito junto con la descripción detallada de la misma, se pretende que la descripción anterior ilustre y no limite 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 reivindicaciones adjuntas.

LISTADO DE SECUENCIAS

5 <110> NovImmune SA, *et al.* Fischer, Nicolas Kosco-Vilbois, Marie Ravn, Ulla Gueneau, Franck Venet-Bonnot, Sophie

<120> BIBLIOTECAS DE POLIPÉPTIDOS SINTÉTICOS Y MÉTODOS PARA GENERAR VARIANTES POLIPEPTÍDICAS DIVERSIFICADAS DE FORMA NATURAL

10 <130> 23135-418002WO

<150> US 61/379.571
<151> 02-09-2010

15 <150> US 12/784.190
<151> 20-05-2010

<150> US 61/179.850
<151> 20-05-2009

20 <150> US 61/287.336
<151> 17-12-2009

<150> US 61/314.794
<151> 17-03-2010

25 <160> 366

<170> PatentIn versión 3.5

30 <210> 1
<211> 98
<212> PRT
<213> *Homo sapiens*

35 <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
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

<210> 2
 <211> 98
 <212> PRT
 <213> *Homo sapiens*

5

<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

10

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

15

<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

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

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

ES 2 542 744 T3

<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

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

```

5

<210> 6

<211> 98

<212> PRT

10 <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*

<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

ES 2 542 744 T3

<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
          85           90           95

```

5 Thr Val

<210> 9

<211> 98

<212> PRT

10 <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

```

ES 2 542 744 T3

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

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

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

ES 2 542 744 T3

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

5

<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

Asn Gly

10

<210> 14
 <211> 98
 <212> PRT
 <213> *Homo sapiens*

15

<400> 14

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 Glu 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 Trp Asp Ser Ser Leu
85 90 95

Ser Ala

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

<400> 15

tattactgtg cgagatgaga cgaataacgg taaggcgggtt taccaggttt aaacgcgtat 60

tgggaaggcg cgtctcttgg ggccagggaa ccctggtcac 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
1 5 10 15

Phe Lys Arg Val Leu Gly Arg Arg Val Ser Trp Gly Gln Gly Thr Leu
20 25 30

Val Thr Val Ser
35

<210> 17
<211> 33
<212> PRT
<213> *Homo sapiens*

<400> 17

Tyr Tyr Cys Ala Arg Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala
1 5 10 15

Tyr Trp Glu Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser
20 25 30

Arg

5 <210> 18
<211> 35
<212> PRT
<213> *Homo sapiens*

10 <400> 18

Ile Thr Val Arg Asp Glu Thr Asn Asn Gly Lys Ala Val Tyr Gln Val
1 5 10 15

Thr Arg Ile Gly Thr Lys Ala Arg Leu Leu Gly Pro Gly Asn Pro Gly
20 25 30

His Ser Leu
35

15 <210> 19
<211> 14
<212> ADN
<213> Secuencia artificial

20 <220>
<223> sintetizada químicamente

25 <220>
<221> misc_feature
<222> (7)..(14)
<223> n es a, t, c o g

<400> 19
cgtctcnnnn nnnn 14

30 <210> 20
<211> 14
<212> ADN
<213> Secuencia artificial

35 <220>
<223> sintetizada químicamente

40 <220>
<221> misc_feature
<222> (7)..(14)
<223> n es a, t, c o g

<400> 20
gcagagnnnn nnnn 14

45 <210> 21
<211> 386
<212> ADN

<213> *Homo sapiens*

<400> 21

```

cagggtgcagc tgggtgcagtc tggggctgag gtgaagaagc ctggggcctc agtgaaggtc      60
tcctgcaagg cttctggata caccttcacc ggctactata tgcactgggt gcgacaggcc      120
cctggacaag ggcttgagtg gatgggatgg atcaacccta acagtgggtg cacaaactat      180
gcacagaagt ttcagggcag ggtcaccatg accagggaaca cgtccatcag cacagcctac      240
atggagctga gcaggctgag atctgacgac acggccgtgt attactgtgc gagatgagac      300
gaataacggt aaggcggttt accagggttta 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

Ser Glu Gly Leu Leu Gln Gly Phe Trp Ile His Leu His Arg Leu Leu
      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

ES 2 542 744 T3

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
 100 105 110
 Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser
 115 120 125

5 <210> 24
 <211> 123
 <212> PRT
 <213> *Homo sapiens*

10 <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

cagggtgcagc tgggtgcagtc tggagctgag gtgaagaagc ctggggcctc agtgaaggtc 60
tcttgcaagg cttctgggta cacctttacc agctatggta tcagctgggt gcgacaggcc 120
cctggacaag ggcttgagtg gatgggatgg atcagcgctt acaatggtaa cacaaactat 180
gcacagaagc tccagggcag agtcaccatg accacagaca catccacgag cacagcctac 240
atggagctga ggagcctgag atctgacgac acggccgtgt attactgtgc gagatgagac 300
gaataacggt aaggcgggttt accaggttta aacgcgtatt gggaaggcgc gtctcttggg 360
gccaggggaac cctggtcaca gtctcg 386

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

<400> 26

ES 2 542 744 T3

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
		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

ES 2 542 744 T3

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
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 tgggtgcagtc tggggctgag gtgaagaagc ctgggtcctc ggtgaaggtc 60
tcctgcaagg cttctggagg caccttcagc agctatgcta tcagctgggt gcgacaggcc 120
cctggacaag ggcttgagtg gatgggaggg atcatcccta tctttggtac agcaaactac 180
gcacagaagt tccagggcag agtcacgatt accgcggacg aatccacgag cacagcctac 240
atggagctga gcagcctgag atctgaggac acggcctgtg attactgtgc gagatgagac 300
gaataacggg aaggcgggtt accagggtta aacgcgtatt gggaaggcgc gtctcttggg 360
gccagggaac cctggtcaca gtctcg 386

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

<400> 30

ES 2 542 744 T3

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

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

ES 2 542 744 T3

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

5

ES 2 542 744 T3

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
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 tgggggaggc ttggtacagc ctggggggtc cctgagactc 60
tcctgtgcag cctctggatt caccttttagc agctatgcc a tgagctgggt ccgccaggct 120
ccaggggaagg ggctggagtg ggtctcagct attagtggta gtggtggtag cacatactac 180
gcagactccg tgaagggccg gttcaccatc tccagagaca attccaagaa cacgctgtat 240
ctgcaaataga acagcctgag agccgaggac acggccgtat attactgtgc gagatgagac 300
gaataacggt aaggcggttt accaggttta aacgcgtatt gggaaggcgc gtctcttggg 360
gccaggggaac cctggtcaca gtctcg 386

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

<400> 34

ES 2 542 744 T3

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	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
 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

ES 2 542 744 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> 37
<211> 385
<212> ADN
<213> *Homo sapiens*

<400> 37

gaggtgcagc	tggtggagtc	tgggggaggc	ttggtacagc	ctgggggggc	cctgagactc	60
tcctgtgcag	cctctggatt	cacctttagc	agctatgcc	tgagctgggt	ccgccaggct	120
ccaggggaagg	ggctggagtg	ggtctcagct	attagtggta	gtggtggtag	cacatactac	180
gcagactccg	tgaagggccg	gttcacatct	ccagagacaa	ttccaagaac	acgctgtatc	240
tgcaaatgaa	cagcctgaga	gccgaggaca	cggccgtata	ttactgtgcg	aaatgagacg	300
aataacggta	aggcggttta	ccagggttaa	acgcgtattg	ggaaggcgcg	tctcttgggg	360
ccaggggaacc	ctggtcacag	tctcg				385

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

<400> 38

ES 2 542 744 T3

```

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

ES 2 542 744 T3

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 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

ES 2 542 744 T3

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

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 ggctggagtg ggtggcagtt atatcatatg atggaagtaa taaatactac 180
gcagactccg tgaagggccg attcaccatc tccagagaca attccaagaa cacgctgtat 240
ctgcaaatac acagcctgag agctgaggac acggctgtgt attactgtgc gagatgagac 300
gaataacggt aaggcggttt accaggttta aacgcgtatt gggaaggcgc gtctcttggg 360
gccaggggaa cctggtcaca gtctcg 386

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

<400> 42

ES 2 542 744 T3

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

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

ES 2 542 744 T3

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

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

ES 2 542 744 T3

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 tggggggaggc ttggtacagc ctgggggggtc cctgagactc	60
tctctgtgcag cctctggatt caccttcagt agctatagca tgaactgggt ccgccaggct	120
ccaggggaagg ggctggagtg ggtttcatac attagtagta gtagtagtac catatactac	180
gcagactctg tgaagggccg attcaccatc tccagagaca atgccaagaa ctactgtat	240
ctgcaaatga acagcctgag agccgaggac acggctgtgt attactgtgc gagatgagac	300
gaataacggt aaggcggttt accaggttta aacgcgtatt gggaaggcgc gtctcttggg	360
gccagggaac cctggtcaca gtctcg	386

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

<400> 46

ES 2 542 744 T3

```

Arg Gly Ala Ala Gly 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 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

ES 2 542 744 T3

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	
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

ES 2 542 744 T3

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

gaggtgcagc	tggtgcagtc	tggagcagag	gtgaaaaagc	coggggagtc	tctgaagatc	60
tctgtgaagg	gttctggata	cagctttacc	agctactgga	toggctgggt	gcgccagatg	120
cccgggaaag	gcctggagtg	gatggggatc	atctatcctg	gtgactctga	taccagatac	180
agcccgtcct	tccaaggcca	ggtcaccatc	tcagccgaca	agtccatcag	cacogcctac	240
ctgcagtgga	gcagcctgaa	ggcctcggac	accgccatgt	attactgtgc		290

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

<400> 50

ES 2 542 744 T3

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

ES 2 542 744 T3

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	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

ES 2 542 744 T3

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

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 aactatttaa attggtatca gcagaaacca 120
gggaaagccc ctaagctcct gatctacgat gcatccaatt tggaacagg ggtcccatca 180
aggttcagtg gaagtggatc tgggacagat tttactttca ccatcagcag cctgcagcct 240
gaagatattg caacatatta ctgtcagcag cgagacgaat aacggttaagg cggtttacca 300
ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca 360
aaggggcggc cgca 374

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

<400> 54

ES 2 542 744 T3

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

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

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

<400> 55

ES 2 542 744 T3

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		
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	aggcgagtca	ggacattagc	aactatttaa	attggtatca	gcagaaacca	120
gggaaagccc	ctaagctoct	gatctaogat	gcatccaatt	tggaacacag	ggtcccatca	180
aggttcagt	gaagtggatc	tgggacagat	tttactttca	ccatcagcag	cctgcagcct	240
gaagatattg	caacatatta	ctgtcagcag	cgagacgaat	aacggtaag	cggtttacca	300
ggtttaaacg	cgtattggga	aggcgogtct	cattcggcc	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

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

ES 2 542 744 T3

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 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 tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
 atcacttgcc gggcaagtca gagcattagc agctatttaa attggtatca gcagaaacca 120
 gggaaagccc ctaagctcct gatctatgct gcatccagtt tgcaaagtgg ggtcccatca 180
 aggttcagtg gcagtggatc tgggacagat ttcaactetca ccatcagcag tctgcaacct 240
 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

ES 2 542 744 T3

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	
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

ES 2 542 744 T3

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
 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 ttcactctca ccatcagcag tctgcaacct 240
 gaagattttg caacttacta ctgtcagcag cgagacgaat aacggtaagg cggtttacca 300
 gggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca 360
 aacgggcggc cgca 374

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

<400> 66

ES 2 542 744 T3

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 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

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 2 542 744 T3

Gly Gly Asn Gln Thr Gly Gly Arg
115 120

5
<210> 69
<211> 374
<212> ADN
<213> *Homo sapiens*

<400> 69

gaaattgtgt tgacacagtc tccagccacc ctgtctttgt ctccagggga aagagccacc 60
ctctcctgca gggccagtca gagtgttagc agctacttag cctggtacca acagaaacct 120
ggccaggctc ccaggctcct catctatgat gcatccaaca gggccactgg catcccagcc 180
aggttcagtg gcagtgggtc tgggacagac ttcactctca ccatcagcag cctagagcct 240
gaagattttg cagtttatta ctgtcagcag cgagacgaat aacggtaagg cggtttacca 300
ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca 360
aaggggcggc cgca 374

10

15
<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
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

ES 2 542 744 T3

<210> 71
 <211> 122
 <212> PRT
 <213> *Homo sapiens*

5

<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
 100 105 110

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

10 <210> 72
 <211> 121
 <212> PRT
 <213> *Homo sapiens*

15 <400> 72

ES 2 542 744 T3

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	gggccagtca	gagtgttagc	agctacttag	cctggtacca	acagaaacct	120
ggccaggctc	ccaggctcct	catctatgat	gcacccaaca	gggcactgg	catccagcc	180
aggttcagt	gcagtgggtc	tgggacagac	ttcactctca	ccatcagcag	cctagagcct	240
gaagattttg	cagtttatta	ctgtcagcag	cgagacgaat	aacggtaagg	cggtttacca	300
ggtttaaacg	cgtattggga	aggcgcgtct	cattcgcca	agggaccaag	gtggaaatca	360
aacgggcggc	cgca					374

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

<400> 74

ES 2 542 744 T3

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

5

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
 100 105 110

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

<210> 76
 <211> 121
 <212> PRT
 <213> *Homo sapiens*

<400> 76

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

ES 2 542 744 T3

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	Thr	Gly	Gly	Arg							
		115					120								

<210> 77
 <211> 374
 <212> ADN
 <213> *Homo sapiens*
 <400> 77

gaaatagtga	tgaacgcagtc	tccagccacc	ctgtctgtgt	ctccagggga	aagagccacc	60
ctctcctgca	gggccagtc	gagtgtagc	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
aaggggcggc	cgca					374

<210> 78
 <211> 123
 <212> PRT
 <213> *Homo sapiens*
 <400> 78

ES 2 542 744 T3

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	
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

ES 2 542 744 T3

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

ES 2 542 744 T3

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 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 gggccagtc gagtgtagc agcaacttag cctggtacca gcagaaacct 120
ggccaggctc ccaggctcct catctatggt gcatccacca gggccactgg tatccagcc 180
aggttcagtg gcagtgggtc tgggacagag ttcactctca ccatcagcag cctgcagtct 240
gaagattttg cagtttatta ctgtcagcag cgagacgaat aacggtaagg cgttttacca 300
ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca 360
aacgggcggc cgca 374

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

<400> 82

ES 2 542 744 T3

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	
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

ES 2 542 744 T3

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	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	gagtgttagc	agcagctact	tagcctggta	ccagcagaaa	120
cctggccagg	ctcccaggct	cctcatctat	ggtgcatcca	gcagggccac	tggcatocca	180
gacaggttca	gtggcagtg	gtctgggaca	gacttcactc	tcaccatcag	cagactggag	240
cctgaagatt	ttgcagttta	ttactgtcag	cagcgagacg	aataacggta	aggcggttta	300
ccagggttaa	acgcgtattg	ggaaggcgcg	tctcattcgg	ccaagggacc	aaggtggaaa	360
tcaaaggggc	ggccgca					377

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

<400> 86

ES 2 542 744 T3

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

5

ES 2 542 744 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 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

ES 2 542 744 T3

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
 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 gagggttagc 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
 ccaggtttta 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

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

ES 2 542 744 T3

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

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
tcttgctctg gaagcagctc caacattggg aataattatg tatcctggta ccagcagctc      120
ccaggaacag cccccaaact cctcatttat gacaataata agcgaccctc agggattcct      180
gaccgattct ctggctccaa gtctggcagc tcagccaccc tgggcatcac cggactccag      240
actgggggacg aggccgatta ttactgcgga acaggagacg aataacggta aggcgggttta      300
ccagggtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga      360
ccgtcctagg ggcggcgca                                     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

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

ES 2 542 744 T3

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	
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

ES 2 542 744 T3

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 ccgggcagag ggtcaccatc 60
 tcttgttctg gaagcagctc caacatcgga agtaatactg taaactggta ccagcagctc 120
 ccaggaacgg cccccaaact cctcatctat agtaataatc agcggccctc aggggtccct 180
 gaccgattct ctggctccaa gtctggcacc tcagcctccc tggccatcag tgggctccag 240
 tctgaggatg aggctgatta ttactgtgca gcaggagacg aataacggta aggcgggtta 300
 ccagggttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcgaggag accaagctga 360
 cgtcctagg ggcggcgcga 380

<210> 98
 <211> 119
 <212> PRT
 <213> *Homo sapiens*

<400> 98

ES 2 542 744 T3

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 2 542 744 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

ES 2 542 744 T3

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

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

5 <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

15 <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

25 <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

35 <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

40 <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

50 <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

5 Tyr Trp

<210> 110

<211> 15

<212> PRT

10 <213> *Homo sapiens*

<400> 110

Cys Ala Lys Val Thr Gly Val Phe Val Gly Asn Phe Asp Tyr Trp
1 5 10 15

15 <210> 111

<211> 15

<212> PRT

20 <213> *Homo sapiens*

<400> 111

Cys Ala Lys Asp Ala Arg Asn Ser Gly Ser Tyr Phe Asp Tyr Trp
1 5 10 15

25 <210> 112

<211> 14

<212> PRT

<213> *Homo sapiens*

30 <400> 112

Cys Ala Lys Gly Val Arg Thr Gly Val Val Phe Asp Tyr Trp
1 5 10

35 <210> 113

<211> 19

<212> PRT

<213> *Homo sapiens*

<400> 113

40

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

45 <212> PRT

<213> *Homo sapiens*

<400> 114

		Cys	Ala	Lys	Trp	Gly	Gly	Trp	Phe	Asp	Pro	Trp	
		1				5					10		

5

<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			

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	

20

<210> 117
 <211> 44
 <212> ADN
 <213> *Homo sapiens*

25

<220>
 <221> misc_feature
 <222> (13)..(14)
 <223> n es a, t, c o g

30

<220>
 <221> misc_feature
 <222> (16)..(17)
 <223> n es a, t, c o g

35

<220>
 <221> misc_feature
 <222> (19)..(20)
 <223> n es a, t, c o g

40

<220>
 <221> misc_feature
 <222> (22)..(23)
 <223> n es a, t, c o g

45

<400> 117
 acgtctccga gannsnnsnn snnsatggat tattggggga gacg 44

50

<210> 118
 <211> 14
 <212> PRT
 <213> *Homo sapiens*

55

<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				

5 <210> 119
 <211> 14
 <212> PRT
 <213> *Homo sapiens*

10 <220>
 <221> MISC_FEATURE
 <222> (4)..(4)
 <223> Xaa es Asp o Glu

15 <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				

20 <210> 120
 <211> 25
 <212> ADN
 25 <213> Secuencia artificial

<220>
 <223> sintetizada químicamente

30 <400> 120
 tgttctaataat cgcaggtgcc agatg 25

35 <210> 121
 <211> 25
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> sintetizada químicamente

<400> 121
 atttatgtta tgactgtta cactg 25

45 <210> 122
 <211> 25
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> sintetizada químicamente

<400> 122
 tatttgtttt tatgtttcca atctc 25

55 <210> 123
 <211> 25
 <212> ADN
 <213> Secuencia artificial

60 <220>
 <223> sintetizada químicamente

<400> 123

	ccttgagggt ttatgttatg acttg	25
5	<210> 124 <211> 25 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 124 ttatttccaa ttcagatac caccg	25
15	<210> 125 <211> 25 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 125 ttgttgggggt tttgtttca tgtgg	25
30	<210> 126 <211> 25 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
40	<400> 126 tatttccaat ttcagatacc actgg	25
45	<210> 127 <211> 25 <212> ADN <213> Secuencia artificial	
50	<220> <223> sintetizada químicamente	
55	<400> 127 atgttgaatc actgtgggag gccag	25
60	<210> 128 <211> 25 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<400> 128 ttatttccaa ttcagatac caccg	25
	<210> 129 <211> 25 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
	<400> 129	

	ttttgtttca agctgaatca ctgtg	25
5	<210> 130 <211> 25 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 130 atgtctgtgt ctctctcact tccag	25
15	<210> 131 <211> 25 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
	<400> 131 ttccccattg gcctggagca ctgtg	25
25	<210> 132 <211> 25 <212> ADN <213> Secuencia artificial	
30	<220> <223> sintetizada químicamente	
	<400> 132 gtgtctgtgt ctctctgtgt tccag	25
35	<210> 133 <211> 25 <212> ADN <213> Secuencia artificial	
40	<220> <223> sintetizada químicamente	
45	<400> 133 ctgtctcag ttccccattg ggctg	25
50	<210> 134 <211> 25 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
55	<400> 134 atctcatcca ctctgtgtt ctctc	25
60	<210> 135 <211> 25 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<400> 135	

	ttgggtttct gacaccctca ggatg	25
5	<210> 136 <211> 25 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 136 caggccagtc atgtgagact tcacc	25
15	<210> 137 <211> 25 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 137 ctgcctctct cctgggggtt ctgaa	25
30	<210> 138 <211> 25 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
35	<400> 138 cccctgtgtc ctctccacag gtgtc	25
40	<210> 139 <211> 25 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
45	<400> 139 ccggcacagc tgccttctcc ctgag	25
50	<210> 140 <211> 18 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
55	<400> 140 gaggtgcagc tgttgag	18
60	<210> 141 <211> 26 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
65	<400> 141	

	tctgaccagg gtttctttt gtttgc	26
5	<210> 142 <211> 25 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 142 ttgtgtctgg gctcacaatg acttc	25
15	<210> 143 <211> 24 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 143 tggcattttc tgataacggt gtcc	24
30	<210> 144 <211> 25 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
40	<400> 144 ctgcagggag gtttgtgtct gggcg	25
45	<210> 145 <211> 25 <212> ADN <213> Secuencia artificial	
50	<220> <223> sintetizada químicamente	
55	<400> 145 atatgtgtgg cagtttctga ccttg	25
60	<210> 146 <211> 25 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<400> 146 ggtttgtgtc tgggtgcaca ctgac	25
	<210> 147 <211> 25 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
	<400> 147	

	gagtctgtgc cggaagtgca gctgg	25
5	<210> 148 <211> 21 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 148 tattcaggtgc agctgggtgca g	21
15	<210> 149 <211> 21 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 149 tattcaggtgc agctgggtgga g	21
30	<210> 150 <211> 21 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
40	<400> 150 tattgaggtgc agctgggtgca g	21
45	<210> 151 <211> 23 <212> ADN <213> Secuencia artificial	
50	<220> <223> sintetizada químicamente	
55	<400> 151 atatctctcg cacagtaata cac	23
60	<210> 152 <211> 23 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<400> 153	

	atatgtctcg cacagtaata cat	23
5	<210> 154 <211> 32 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 154 tatgacatcc agatgaccca gtctccatcc tc	32
15	<210> 155 <211> 20 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 155 ataggagggg tactgtaact	20
30	<210> 156 <211> 20 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
40	<400> 156 ataggaggga gattatcata	20
45	<210> 157 <211> 24 <212> ADN <213> Secuencia artificial	
50	<220> <223> sintetizada químicamente	
55	<400> 157 tatgaaattg tgttgacgca gtct	24
60	<210> 158 <211> 23 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<400> 158 ataggagggtg agctaccata ctg	23
	<210> 159 <211> 24 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
	<400> 159	

	tatgaaatag tgatgacgca gtct	24
5	<210> 160 <211> 23 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 160 ataggaggcc agttattata ctg	23
15	<210> 161 <211> 23 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
	<400> 161 cagcgtagca actggcctcc tat	23
25	<210> 162 <211> 20 <212> ADN <213> Secuencia artificial	
30	<220> <223> sintetizada químicamente	
35	<400> 162 tacagtctgt gctgactcag	20
40	<210> 163 <211> 23 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
45	<400> 163 ataggaccat tcaggctgtc atc	23
50	<210> 164 <211> 21 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
55	<400> 164 tatcagtctg tggtagcga g	21
60	<210> 165 <211> 22 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<400> 165	

	ataggagcac tcaggctgct at	22	
5	<210> 166 <211> 33 <212> ADN <213> Secuencia artificial		
10	<220> <223> sintetizada químicamente		
	<400> 166 cagccggcca tggccaggt gcagctggtg cag	33	
15	<210> 167 <211> 33 <212> ADN <213> Secuencia artificial		
20	<220> <223> sintetizada químicamente		
	<400> 167 cagccggcca tggccaggt gcagctggtg gag	33	
25	<210> 168 <211> 33 <212> ADN <213> Secuencia artificial		
30	<220> <223> sintetizada químicamente		
	<400> 168 cagccggcca tggccaggt gcagctgtg gag	33	
35	<210> 169 <211> 43 <212> ADN <213> Secuencia artificial		
40	<220> <223> sintetizada químicamente		
45	<400> 169 cagccggcca tggccaggt gcagctggtg gagtctgggg gag		43
50	<210> 170 <211> 33 <212> ADN <213> Secuencia artificial		
	<220> <223> sintetizada químicamente		
55	<400> 170 cagccggcca tggccaggt gcagctggtg cag	33	
60	<210> 171 <211> 37 <212> ADN <213> Secuencia artificial		
65	<220> <223> sintetizada químicamente		
	<400> 171		

	cttaccggtta ttcgttcat ctgcacagt aatacac	37
5	<210> 172 <211> 37 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 172 cttaccggtta ttcgttcat ctgcacagt aatacac	37
15	<210> 173 <211> 42 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 173 ctgcacagt aatacacagc cgtgtctcg gcttcaggc tg	42
	<210> 174 <211> 37 <212> ADN <213> Secuencia artificial	
30	<220> <223> sintetizada químicamente	
35	<400> 174 cttaccggtta ttcgttcat ctgcacagt aatacat	37
40	<210> 175 <211> 47 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
45	<400> 175 caatacgcgt ttaaacctgg taaaccgcct taccgttatt cgtctca	47
50	<210> 176 <211> 48 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
55	<400> 176 gttccctggc cccaagagac ggccttccc aatacgcgtt taaacctg	48
60	<210> 177 <211> 45 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<400> 177	

	cctccaccgc tcgagactgt gaccagggtt ccctggcccc aagag	45
5	<210> 178 <211> 30 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 178 cgggtcgacg gacatccaga tgaccagtc	30
15	<210> 179 <211> 36 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 179 cgggtcgacg gaaattgtgt tgacacagtc tccagc	36
30	<210> 180 <211> 36 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
40	<400> 180 cgggtcgacg gaaatagtga tgacgcagtc tccagc	36
45	<210> 181 <211> 36 <212> ADN <213> Secuencia artificial	
50	<220> <223> sintetizada químicamente	
55	<400> 181 cgggtcgacg gaaattgtgt tgacgcagtc tccagg	36
60	<210> 182 <211> 45 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<400> 182 ccttaccgtt attcgtctcg ctgctgacag taatatgttg caata	45
	<210> 183 <211> 45 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
	<400> 183	

	ccttaccgtt attcgtctcg ctgctgacag tagtaagttg caaaa	45
5	<210> 184 <211> 47 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 184 ccttaccgtt attcgtctcg ctgctgacag taataaactg caaaatc	47
15	<210> 185 <211> 47 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 185 ccaatacgcg ttaaacctg gtaaccgcc ttaccgttat tcgtctc	47
30	<210> 186 <211> 46 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
35	<400> 186 ggtcccttgg ccgaatgaga cgcgccttcc caatacgctt taaac	46
40	<210> 187 <211> 45 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
45	<400> 187 gtgcggccgc ccgttgatt tccacctgg tccctggcc gaatg	45
50	<210> 188 <211> 45 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
55	<400> 188 gtgcggccgc cccttgatt tccacctgg tccctggcc gaatg	45
60	<210> 189 <211> 32 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
65	<400> 189	

	cgggtcgacg cagtctgtgc tgactcagcc ac	32
5	<210> 190 <211> 32 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 190 cgggtcgacg cagtctgtgc tgacgcagcc gc	32
15	<210> 191 <211> 38 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 191 ccttaccgtt attcgtctcc tgctgcacag taataatc	38
30	<210> 192 <211> 38 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
35	<400> 192 ccttaccgtt attcgtctcc tgtccgcag taataatc	38
40	<210> 193 <211> 46 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
45	<400> 193 ccctccgccc aacacagaga cgcgcctcc caatacgcgt taaac	46
50	<210> 194 <211> 49 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
55	<400> 194 gtgcggccgc ccctaggacg gtcagcttgg tcctccgcc gaacacaga	49
60	<210> 195 <211> 43 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<400> 195	

	ccgcacagcc ggccatggcc caggtgcagc tggcgcagtc tgg	43
5	<210> 196 <211> 43 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 196 ccgcacagcc ggccatggcc gaggtgcagc tggcgcagtc tgg	43
15	<210> 197 <211> 43 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 197 ccgcacagcc ggccatggcc cagrtcacct tgctcgagtc tgg	43
30	<210> 198 <211> 43 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
40	<400> 198 ccgcacagcc ggccatggcc caggtgcagc tgcaggagtc ggg	43
45	<210> 199 <211> 43 <212> ADN <213> Secuencia artificial	
50	<220> <223> sintetizada químicamente	
55	<400> 199 ccgcacagcc ggccatggcc cagctgcagc tgcaggagtc cgg	43
60	<210> 200 <211> 43 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<400> 201	

	tggagtgggt ctcagctatt agtggtagtg gt	32
5	<210> 202 <211> 45 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 202 cgatgggccc ttggtggagg ctgaggagac rgtgaccagg gtgcc	45
15	<210> 203 <211> 45 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 203 cgatgggccc ttggtggagg ctgaagagac ggtgaccrkt gtccc	45
30	<210> 204 <211> 45 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
	<400> 204 cgatgggccc ttggtggagg ctgaggagac ggtgaccagg gtccc	45
40	<210> 205 <211> 35 <212> ADN <213> Secuencia artificial	
45	<220> <223> sintetizada químicamente	
	<400> 205 gagccgagga cacggccgga tgttactgtg cgaga	35
50	<210> 206 <211> 35 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente	
	<400> 206 gagccgagga cacggccgga tgttactgtg cgaaa	35
60	<210> 207 <211> 32 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<400> 207	

	gaggagacgg tgacggatgt gccctggccc ca	32
5	<210> 208 <211> 32 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 208 gaggagacgg tgacggatgt gccacggccc ca	32
15	<210> 209 <211> 32 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 209 gaggagacgg tgacggatgt yccttggccc ca	32
30	<210> 210 <211> 26 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
	<400> 210 atgatgctgc tggcacgtct ccgaga	26
40	<210> 211 <211> 35 <212> ADN <213> Secuencia artificial	
45	<220> <223> sintetizada químicamente	
	<400> 211 ccacgtcatc cgatccgtct cccccaataa tccat	35
50	<210> 212 <211> 35 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente	
	<400> 212 ccacgtcatc cgatccgtct cccccaataa tcaaa	35
60	<210> 213 <211> 50 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<220>	

	<221> misc_feature <222> (19)..(20) <223> n es a, c, g, o t	
5	<220> <221> misc_feature <222> (22)..(23) <223> n es a, c, g, o t	
10	<220> <221> misc_feature <222> (25)..(26) <223> n es a, c, g, o t	
15	<220> <221> misc_feature <222> (28)..(29) <223> n es a, c, g, o t	
20	<400> 213 gctggcacgt ctccgagann snnsnnsnns ttgattatt gggggagacg	50
25	<210> 214 <211> 50 <212> ADN <213> Secuencia artificial	
30	<220> <223> sintetizada químicamente	
35	<220> <221> misc_feature <222> (19)..(20) <223> n es a, c, g, o t	
40	<220> <221> misc_feature <222> (22)..(23) <223> n es a, c, g, o t	
45	<220> <221> misc_feature <222> (25)..(26) <223> n es a, c, g, o t	
50	<220> <221> misc_feature <222> (28)..(29) <223> n es a, c, g, o t	
55	<400> 214 gctggcacgt ctccgagann snnsnnsnns atggattatt gggggagacg	50
60	<210> 215 <211> 53 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
	<220> <221> misc_feature <222> (19)..(20) <223> n es a, c, g, o t	

5	<220> <221> misc_feature <222> (22)..(23) <223> n es a, c, g, o t	
10	<220> <221> misc_feature <222> (25)..(26) <223> n es a, c, g, o t	
15	<220> <221> misc_feature <222> (28)..(29) <223> n es a, c, g, o t	
20	<220> <221> misc_feature <222> (31)..(32) <223> n es a, c, g, o t	
25	<400> 215 gctggcacgt ctccgagann snnsnnsnns nnssttgatt attgggggag acg	53
30	<210> 216 <211> 53 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
40	<220> <221> misc_feature <222> (19)..(20) <223> n es a, c, g, o t	
45	<220> <221> misc_feature <222> (22)..(23) <223> n es a, c, g, o t	
50	<220> <221> misc_feature <222> (25)..(26) <223> n es a, c, g, o t	
55	<220> <221> misc_feature <222> (28)..(29) <223> n es a, c, g, o t	
60	<220> <221> misc_feature <222> (31)..(32) <223> n es a, c, g, o t	
65	<400> 216 gctggcacgt ctccgagann snnsnnsnns nnsatggatt attgggggag acg	53
	<210> 217 <211> 56 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	

5	<220> <221> misc_feature <222> (19)..(20) <223> n es a, c, g, o t	
10	<220> <221> misc_feature <222> (22)..(23) <223> n es a, c, g, o t	
15	<220> <221> misc_feature <222> (25)..(26) <223> n es a, c, g, o t	
20	<220> <221> misc_feature <222> (28)..(29) <223> n es a, c, g, o t	
25	<220> <221> misc_feature <222> (31)..(32) <223> n es a, c, g, o t	
30	<220> <221> misc_feature <222> (34)..(35) <223> n es a, c, g, o t	
35	<400> 217 gctggcacgt ctccgagann snnsnnsnns nnsnnsttg attattgggg gagacg	56
40	<210> 218 <211> 56 <212> ADN <213> Secuencia artificial	
45	<220> <221> misc_feature <222> (19)..(20) <223> n es a, c, g, o t	
50	<220> <221> misc_feature <222> (22)..(23) <223> n es a, c, g, o t	
55	<220> <221> misc_feature <222> (25)..(26) <223> n es a, c, g, o t	
60	<220> <221> misc_feature <222> (28)..(29) <223> n es a, c, g, o t	
65	<220> <221> misc_feature <222> (31)..(32) <223> n es a, c, g, o t	

5	<220> <221> misc_feature <222> (34)..(35) <223> n es a, c, g, o t	
10	<400> 218 gctggcacgt ctccgagann snnsnnsnns nnsnnsatgg attattgggg gagacg	56
15	<210> 219 <211> 56 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 219 gctggcacgt ctccgagadv kdvkdvkdvk dvkdvkttg attattgggg gagacg	56
30	<210> 220 <211> 56 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
40	<400> 220 gctggcacgt ctccgagadv kdvkdvkdvk dvkdvkatgg attattgggg gagacg	56
45	<210> 221 <211> 59 <212> ADN <213> Secuencia artificial	
50	<220> <223> sintetizada químicamente	
55	<400> 221 gctggcacgt ctccgagadv kdvkdvkdvk dvkdvkdvkt ttgattattg ggggagacg	59
60	<210> 222 <211> 59 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
70	<220> <221> misc_feature <222> (19)..(19) <223> n es a, c, g, o t	

5	<220> <221> misc_feature <222> (22)..(22) <223> n es a, c, g, o t	
10	<220> <221> misc_feature <222> (25)..(25) <223> n es a, c, g, o t	
15	<220> <221> misc_feature <222> (28)..(28) <223> n es a, c, g, o t	
20	<220> <221> misc_feature <222> (31)..(31) <223> n es a, c, g, o t	
25	<220> <221> misc_feature <222> (34)..(34) <223> n es a, c, g, o t	
30	<220> <221> misc_feature <222> (37)..(37) <223> n es a, c, g, o t	
	<400> 223 gctggcacgt ctccgaganv tntvntvnt nvtvntvnt ttgattattg ggggagacg	59
35	<210> 224 <211> 59 <212> ADN <213> Secuencia artificial	
40	<220> <223> sintetizada químicamente	
45	<220> <221> misc_feature <222> (19)..(19) <223> n es a, c, g, o t	
50	<220> <221> misc_feature <222> (22)..(22) <223> n es a, c, g, o t	
55	<220> <221> misc_feature <222> (25)..(25) <223> n es a, c, g, o t	
60	<220> <221> misc_feature <222> (28)..(28) <223> n es a, c, g, o t	
65	<220> <221> misc_feature <222> (31)..(31) <223> n es a, c, g, o t	

	<220>		
	<221> misc_feature		
	<222> (34)..(34)		
5	<223> n es a, c, g, o t		
	<220>		
	<221> misc_feature		
	<222> (37)..(37)		
10	<223> n es a, c, g, o t		
	<400> 224		
	gctggcacgt ctccgaganv tntvntvntv ntntvntvta tggattattg ggggagacg	59	
15	<210> 225		
	<211> 62		
	<212> ADN		
	<213> Secuencia artificial		
20	<220>		
	<223> sintetizada químicamente		
	<220>		
	<221> misc_feature		
25	<222> (19)..(19)		
	<223> n es a, c, g, o t		
	<220>		
	<221> misc_feature		
30	<222> (22)..(22)		
	<223> n es a, c, g, o t		
	<220>		
	<221> misc_feature		
35	<222> (25)..(25)		
	<223> n es a, c, g, o t		
	<220>		
	<221> misc_feature		
40	<222> (28)..(28)		
	<223> n es a, c, g, o t		
	<220>		
	<221> misc_feature		
45	<222> (31)..(31)		
	<223> n es a, c, g, o t		
	<220>		
	<221> misc_feature		
50	<222> (34)..(34)		
	<223> n es a, c, g, o t		
	<220>		
	<221> misc_feature		
55	<222> (37)..(37)		
	<223> n es a, c, g, o t		
	<220>		
	<221> misc_feature		
60	<222> (40)..(40)		
	<223> n es a, c, g, o t		
	<400> 225		
	gctggcacgt ctccgaganv tntvntvntv ntntvntvtn vttttgatta ttggggggaga	60	
	cg	62	

5 <210> 226
 <211> 62
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizada químicamente

10 <220>
 <221> misc_feature
 <222> (19)..(19)
 <223> n es a, c, g, o t

15 <220>
 <221> misc_feature
 <222> (22)..(22)
 <223> n es a, c, g, o t

20 <220>
 <221> misc_feature
 <222> (25)..(25)
 <223> n es a, c, g, o t

25 <220>
 <221> misc_feature
 <222> (28)..(28)
 <223> n es a, c, g, o t

30 <220>
 <221> misc_feature
 <222> (31)..(31)
 <223> n es a, c, g, o t

35 <220>
 <221> misc_feature
 <222> (34)..(34)
 <223> n es a, c, g, o t

40 <220>
 <221> misc_feature
 <222> (37)..(37)
 <223> n es a, c, g, o t

45 <220>
 <221> misc_feature
 <222> (40)..(40)
 <223> n es a, c, g, o t

50 <400> 226

gctggcacgt ctccgaganv tnvtnvtnvt nvtvtnvtn vtatggatta ttgggggaga 60

cg 62

55 <210> 227
 <211> 65
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> sintetizada químicamente

60 <220>
 <221> misc_feature

	<222> (19)..(19)	
	<223> n es a, c, g, o t	
5	<220> <221> misc_feature <222> (22)..(22) <223> n es a, c, g, o t	
10	<220> <221> misc_feature <222> (25)..(25) <223> n es a, c, g, o t	
15	<220> <221> misc_feature <222> (28)..(28) <223> n es a, c, g, o t	
20	<220> <221> misc_feature <222> (31)..(31) <223> n es a, c, g, o t	
25	<220> <221> misc_feature <222> (34)..(34) <223> n es a, c, g, o t	
30	<220> <221> misc_feature <222> (37)..(37) <223> n es a, c, g, o t	
35	<220> <221> misc_feature <222> (40)..(40) <223> n es a, c, g, o t	
40	<220> <221> misc_feature <222> (43)..(43) <223> n es a, c, g, o t	
45	<400> 227	
	gctggcacgt ctccgaganv tnvtntnvt nvtntnvt nvtnttttga ttattggggg	60
	agacg	65
50	<210> 228 <211> 65 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente	
60	<220> <221> misc_feature <222> (19)..(19) <223> n es a, c, g, o t	
	<220> <221> misc_feature <222> (22)..(22)	

	<223> n es a, c, g, o t	
	<220>	
5	<221> misc_feature	
	<222> (25)..(25)	
	<223> n es a, c, g, o t	
	<220>	
10	<221> misc_feature	
	<222> (28)..(28)	
	<223> n es a, c, g, o t	
	<220>	
15	<221> misc_feature	
	<222> (31)..(31)	
	<223> n es a, c, g, o t	
	<220>	
20	<221> misc_feature	
	<222> (34)..(34)	
	<223> n es a, c, g, o t	
	<220>	
25	<221> misc_feature	
	<222> (37)..(37)	
	<223> n es a, c, g, o t	
	<220>	
30	<221> misc_feature	
	<222> (40)..(40)	
	<223> n es a, c, g, o t	
	<220>	
35	<221> misc_feature	
	<222> (43)..(43)	
	<223> n es a, c, g, o t	
	<400> 228	
	gctggcacgt ctccgaganv tnvtnvtnvt nvtvtnvtn vtnvtatgga ttattggggg	60
40	agacg	65
	<210> 229	
	<211> 65	
	<212> ADN	
45	<213> Secuencia artificial	
	<220>	
	<223> sintetizada químicamente	
50	<400> 229	
	gctggcacgt ctccgagadv tdvtdvtdvt dvtvtdvtd vtdvttttga ttattggggg	60
	agacg	65
	<210> 230	
55	<211> 65	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
60	<223> sintetizada químicamente	

	<400> 230		
	gctggcacgt ctccgagadv tdvtdvtdvt dvtvdvtdvtd vtdvtatgga ttattggggg	60	
	agacg	65	
5	<210> 231 <211> 68 <212> ADN <213> Secuencia artificial		
10	<220> <223> sintetizada químicamente		
15	<400> 231		
	gctggcacgt ctccgagadv tdvtdvtdvt dvtvdvtdvtd vtdvtdvttt tgattattgg	60	
	gggagacg	68	
20	<210> 232 <211> 68 <212> ADN <213> Secuencia artificial		
25	<220> <223> sintetizada químicamente		
	<400> 232		
	gctggcacgt ctccgagadv tdvtdvtdvt dvtvdvtdvtd vtdvtdvtat ggattattgg	60	
	gggagacg	68	
30	<210> 233 <211> 26 <212> ADN <213> Secuencia artificial		
35	<220> <223> sintetizada químicamente		
40	<400> 233 ccggtgtagc gaaggcgtct cagcag 26		
45	<210> 234 <211> 32 <212> ADN <213> Secuencia artificial		
	<220> <223> sintetizada químicamente		
50	<400> 234 tagggtgcc ttgatcgtct cccgaaggtc gg	32	
55	<210> 235 <211> 45 <212> ADN <213> Secuencia artificial		
	<220> <223> sintetizada químicamente		

5	<220> <221> misc_feature <222> (17)..(18) <223> n es a, c, g, o t	
10	<220> <221> misc_feature <222> (20)..(21) <223> n es a, c, g, o t	
15	<220> <221> misc_feature <222> (23)..(24) <223> n es a, c, g, o t	
20	<220> <221> misc_feature <222> (26)..(27) <223> n es a, c, g, o t	
25	<400> 235 gaaggcgtct cagcagnnsn nsnsnnscc gacctcggg agacg <210> 236 <211> 48 <212> ADN <213> Secuencia artificial	45
30	<220> <223> sintetizada químicamente	
35	<220> <221> misc_feature <222> (17)..(18) <223> n es a, c, g, o t	
40	<220> <221> misc_feature <222> (20)..(21) <223> n es a, c, g, o t	
45	<220> <221> misc_feature <222> (23)..(24) <223> n es a, c, g, o t	
50	<220> <221> misc_feature <222> (26)..(27) <223> n es a, c, g, o t	
55	<220> <221> misc_feature <222> (32)..(33) <223> n es a, c, g, o t	
60	<400> 236 gaaggcgtct cagcagnnsn nsnsnnscc gnnsacctc gggagacg <210> 237 <211> 51 <212> ADN <213> Secuencia artificial	48
65	<220>	

	<223> sintetizada químicamente	
	<220>	
5	<221> misc_feature	
	<222> (17)..(18)	
	<223> n es a, c, g, o t	
	<220>	
10	<221> misc_feature	
	<222> (20)..(21)	
	<223> n es a, c, g, o t	
	<220>	
15	<221> misc_feature	
	<222> (23)..(24)	
	<223> n es a, c, g, o t	
	<220>	
20	<221> misc_feature	
	<222> (26)..(27)	
	<223> n es a, c, g, o t	
	<220>	
25	<221> misc_feature	
	<222> (29)..(30)	
	<223> n es a, c, g, o t	
	<220>	
30	<221> misc_feature	
	<222> (35)..(36)	
	<223> n es a, c, g, o t	
35	<400> 237 gaaggcgtct cagcagnnsn nsnnsnnsn sccgnnsacc ttcgggagac g	51
	<210> 238	
	<211> 32	
	<212> ADN	
40	<213> Secuencia artificial	
	<220>	
	<223> sintetizada químicamente	
45	<400> 238 cggtcagtcg caatacgtct ccagcatggg at	32
	<210> 239	
	<211> 32	
50	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> sintetizada químicamente	
55	<400> 239 cggtcagtcg caatacgtct ccagcatatg at	32
	<210> 240	
	<211> 32	
60	<212> ADN	
	<213> Secuencia artificial	
	<220>	
65	<223> sintetizada químicamente	
	<400> 240	

	caggaccagt ctcgtgagga tcgtctcaac ac	32
5	<210> 241 <211> 46 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
15	<220> <221> misc_feature <222> (18)..(19) <223> n es a, c, g, o t	
20	<220> <221> misc_feature <222> (21)..(22) <223> n es a, c, g, o t	
25	<220> <221> misc_feature <222> (24)..(25) <223> n es a, c, g, o t	
30	<220> <221> misc_feature <222> (27)..(28) <223> n es a, c, g, o t	
	<400> 241 cgtctccagc atgggatnns nnsnnsnsg tgttgagacg atcctc	46
35	<210> 242 <211> 46 <212> ADN <213> Secuencia artificial	
40	<220> <223> sintetizada químicamente	
45	<220> <221> misc_feature <222> (18)..(19) <223> n es a, c, g, o t	
50	<220> <221> misc_feature <222> (21)..(22) <223> n es a, c, g, o t	
55	<220> <221> misc_feature <222> (24)..(25) <223> n es a, c, g, o t	
60	<220> <221> misc_feature <222> (27)..(28) <223> n es a, c, g, o t	
65	<400> 242 cgtctccagc atatgatnns nnsnnsnsg tgttgagacg atcctc	46
	<210> 243 <211> 49	

<212> ADN
 <213> Secuencia artificial

 5 <220>
 <223> sintetizada químicamente

 10 <220>
 <221> misc_feature
 <222> (18)..(19)
 <223> n es a, c, g, o t

 15 <220>
 <221> misc_feature
 <222> (21)..(22)
 <223> n es a, c, g, o t

 20 <220>
 <221> misc_feature
 <222> (24)..(25)
 <223> n es a, c, g, o t

 25 <220>
 <221> misc_feature
 <222> (27)..(28)
 <223> n es a, c, g, o t

 30 <220>
 <221> misc_feature
 <222> (30)..(31)
 <223> n es a, c, g, o t

 <400> 243
 cgctccagc atgggatnns nnsnnsnnsn nsgtgttgag acgatactc 49

 35 <210> 244
 <211> 49
 <212> ADN
 <213> Secuencia artificial

 40 <220>
 <223> sintetizada químicamente

 45 <220>
 <221> misc_feature
 <222> (18)..(19)
 <223> n es a, c, g, o t

 50 <220>
 <221> misc_feature
 <222> (21)..(22)
 <223> n es a, c, g, o t

 55 <220>
 <221> misc_feature
 <222> (24)..(25)
 <223> n es a, c, g, o t

 60 <220>
 <221> misc_feature
 <222> (27)..(28)
 <223> n es a, c, g, o t

 65 <220>
 <221> misc_feature
 <222> (30)..(31)
 <223> n es a, c, g, o t

	<400> 244 cgctccagc atatgatnns nnsnnsnnsn nsgtgtgag acgatcctc	49
5	<210> 245 <211> 52 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
15	<220> <221> misc_feature <222> (18)..(19) <223> n es a, c, g, o t	
20	<220> <221> misc_feature <222> (21)..(22) <223> n es a, c, g, o t	
25	<220> <221> misc_feature <222> (24)..(25) <223> n es a, c, g, o t	
30	<220> <221> misc_feature <222> (27)..(28) <223> n es a, c, g, o t	
35	<220> <221> misc_feature <222> (30)..(31) <223> n es a, c, g, o t	
40	<220> <221> misc_feature <222> (33)..(34) <223> n es a, c, g, o t	
45	<400> 245 cgctccagc atgggatnns nnsnnsnnsn nsnnsgtgtt gagacgatcc tc	52
50	<210> 246 <211> 52 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente	
60	<220> <221> misc_feature <222> (18)..(19) <223> n es a, c, g, o t	
65	<220> <221> misc_feature <222> (21)..(22) <223> n es a, c, g, o t	
70	<220> <221> misc_feature <222> (24)..(25)	

	<223> n es a, c, g, o t		
	<220>		
5	<221> misc_feature		
	<222> (27)..(28)		
	<223> n es a, c, g, o t		
	<220>		
10	<221> misc_feature		
	<222> (30)..(31)		
	<223> n es a, c, g, o t		
	<220>		
15	<221> misc_feature		
	<222> (33)..(34)		
	<223> n es a, c, g, o t		
	<400> 246		
20	cgtctccagc atatgatnns nnsnnsnnsn nsnnsgtgtt gagacgatcc tc	52	
	<210> 247		
	<211> 32		
	<212> ADN		
25	<213> Secuencia artificial		
	<220>		
	<223> sintetizada químicamente		
	<400> 247		
30	cggtcagtcg caatacgtct cgaacatggg at	32	
	<210> 248		
	<211> 32		
	<212> ADN		
35	<213> Secuencia artificial		
	<220>		
	<223> sintetizada químicamente		
	<400> 248		
40	cggtcagtcg caatacgtct cgaacatggg at	32	
	<210> 249		
	<211> 46		
45	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> sintetizada químicamente		
50			
	<220>		
	<221> misc_feature		
	<222> (18)..(19)		
55	<223> n es a, c, g, o t		
	<220>		
	<221> misc_feature		
	<222> (21)..(22)		
60	<223> n es a, c, g, o t		
	<220>		
	<221> misc_feature		
	<222> (24)..(25)		
65	<223> n es a, c, g, o t		
	<220>		

	<221> misc_feature <222> (27)..(28) <223> n es a, c, g, o t	
5	<400> 249 cgctcgaac atgggatnns nnsnnsnsg tgtgagacg atcctc	46
	<210> 250 <211> 46 <212> ADN <213> Secuencia artificial	
10		
	<220> <223> sintetizada químicamente	
15		
	<220> <221> misc_feature <222> (18)..(19) <223> n es a, c, g, o t	
20		
	<220> <221> misc_feature <222> (21)..(22) <223> n es a, c, g, o t	
25		
	<220> <221> misc_feature <222> (24)..(25) <223> n es a, c, g, o t	
30		
	<220> <221> misc_feature <222> (27)..(28) <223> n es a, c, g, o t	
35		
	<400> 250 cgctcgaac atatgatnns nnsnnsnsg tgtgagacg atcctc	46
40		
	<210> 251 <211> 49 <212> ADN <213> Secuencia artificial	
45		
	<220> <223> sintetizada químicamente	
	<220> <221> misc_feature <222> (18)..(19) <223> n es a, c, g, o t	
50		
	<220> <221> misc_feature <222> (21)..(22) <223> n es a, c, g, o t	
55		
	<220> <221> misc_feature <222> (24)..(25) <223> n es a, c, g, o t	
60		
	<220> <221> misc_feature <222> (27)..(28) <223> n es a, c, g, o t	
65		

	<220> <221> misc_feature <222> (30)..(31) <223> n es a, c, g, o t	
5	<400> 251 cgctcgaac atgggatnns nnsnnsnnsn nsgtgttgag acgatcctc	49
10	<210> 252 <211> 49 <212> ADN <213> Secuencia artificial	
15	<220> <223> sintetizada químicamente	
20	<220> <221> misc_feature <222> (18)..(19) <223> n es a, c, g, o t	
25	<220> <221> misc_feature <222> (21)..(22) <223> n es a, c, g, o t	
30	<220> <221> misc_feature <222> (24)..(25) <223> n es a, c, g, o t	
35	<220> <221> misc_feature <222> (27)..(28) <223> n es a, c, g, o t	
40	<220> <221> misc_feature <222> (30)..(31) <223> n es a, c, g, o t	
45	<400> 252 cgctcgaac atatgatnns nnsnnsnnsn nsgtgttgag acgatcctc	49
50	<210> 253 <211> 52 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente	
60	<220> <221> misc_feature <222> (18)..(19) <223> n es a, c, g, o t	
65	<220> <221> misc_feature <222> (21)..(22) <223> n es a, c, g, o t	
70	<220> <221> misc_feature <222> (24)..(25) <223> n es a, c, g, o t	

5	<220> <221> misc_feature <222> (27)..(28) <223> n es a, c, g, o t	
10	<220> <221> misc_feature <222> (30)..(31) <223> n es a, c, g, o t	
15	<220> <221> misc_feature <222> (33)..(34) <223> n es a, c, g, o t	
20	<400> 253 cgtctcgaac atgggatnns nnsnnsnnsn nsnnsgtggt gagacgatcc tc	52
25	<210> 254 <211> 52 <212> ADN <213> Secuencia artificial	
30	<220> <221> misc_feature <222> (18)..(19) <223> n es a, c, g, o t	
35	<220> <221> misc_feature <222> (21)..(22) <223> n es a, c, g, o t	
40	<220> <221> misc_feature <222> (24)..(25) <223> n es a, c, g, o t	
45	<220> <221> misc_feature <222> (27)..(28) <223> n es a, c, g, o t	
50	<220> <221> misc_feature <222> (30)..(31) <223> n es a, c, g, o t	
55	<220> <221> misc_feature <222> (33)..(34) <223> n es a, c, g, o t	
60	<400> 254 cgtctcgaac atgggatnns nnsnnsnnsn nsnnsgtggt gagacgatcc tc	52
65	<210> 255 <211> 21 <212> ADN <213> Secuencia artificial <220>	

	<223> sintetizada químicamente	
5	<400> 255 ctctctgag atgagtttt g	21
	<210> 256 <211> 43 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
15	<400> 256 atgcgggcca gccggccatg gccsaggttc agctbcagca gtc	43
20	<210> 257 <211> 43 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
25	<400> 257 atgcgggcca gccggccatg gccsaggttc acctgcagca rtc	43
30	<210> 258 <211> 43 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
35	<400> 258 atgcgggcca gccggccatg gccsaggttc agctgaagga gtc	43
40	<210> 259 <211> 43 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
45	<400> 259 atgcgggcca gccggccatg gccsaggttc aactvcagca rcc	43
50	<210> 260 <211> 43 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente	
60	<400> 260 atgcgggcca gccggccatg gccsaggttc agttggtvca gtc	43
65	<210> 261 <211> 43 <212> ADN <213> Secuencia artificial	
	<220>	

	<223> sintetizada químicamente	
5	<400> 261 atgcgggcca gccggccatg gccaggtgc agctgaagsa stc	43
	<210> 262 <211> 43 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
15	<400> 262 atgcgggcca gccggccatg gccaggtgc agskggtgga gtc	43
	<210> 263 <211> 43 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 263 atgcgggcca gccggccatg gccgaagtga arsttgagga gtc	43
	<210> 264 <211> 43 <212> ADN <213> Secuencia artificial	
30	<220> <223> sintetizada químicamente	
35	<400> 264 atgcgggcca gccggccatg gccgakgtsv agcttcagga gtc	43
	<210> 265 <211> 43 <212> ADN <213> Secuencia artificial	
40	<220> <223> sintetizada químicamente	
45	<400> 265 atgcgggcca gccggccatg gccaggtga asstggtgga atc	43
	<210> 266 <211> 43 <212> ADN <213> Secuencia artificial	
50	<220> <223> sintetizada químicamente	
55	<400> 266 atgcgggcca gccggccatg gccaggtga agctgrtgga rtc	43
60	<210> 267 <211> 43 <212> ADN <213> Secuencia artificial	
65	<220>	

	<223> sintetizada químicamente	
5	<400> 267 atgcggccca gccggccatg gccgargtga agctgrtgga gtc	43
	<210> 268 <211> 43 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
15	<400> 268 atgcggccca gccggccatg gccgaagtgc agctgttgga gac	43
20	<210> 269 <211> 43 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
25	<400> 269 atgcggccca gccggccatg gccgargtga agcttctsa gtc	43
30	<210> 270 <211> 42 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
35	<400> 270 atgcggccca gccggccatg gcccgargtga ctctgaaaga gt	42
40	<210> 271 <211> 41 <212> ADN <213> Secuencia artificial	
45	<220> <223> sintetizada químicamente	
50	<400> 271 cctgaaccgc cgcctccgct cgagacggtg accgtggtcc c	41
55	<210> 272 <211> 41 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
60	<400> 272 cctgaaccgc cgcctccgct cgagactgtg agagtgggtgc c	41
65	<210> 273 <211> 41 <212> ADN <213> Secuencia artificial	
	<220>	

	<223> sintetizada químicamente	
5	<400> 273 cctgaaccgc cgctccgct cgagacagtg accagagtc c	41
	<210> 274 <211> 41 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
15	<400> 274 cctgaaccgc cgctccgct cgagacggtg actgaggtc c	41
	<210> 275 <211> 35 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 275 gagccgagga cacggccgga tgttactgtg cgaga	35
30	<210> 276 <211> 31 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
	<400> 276 ggggcgagg gacatccgtc accgtctct c	31
40	<210> 277 <211> 32 <212> ADN <213> Secuencia artificial	
45	<220> <223> sintetizada químicamente	
	<400> 277 gaggagactg tgagggatgt gcctggccc ca	32
50	<210> 278 <211> 32 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente	
60	<400> 278 gaggagacgg tgacggatgt gccctggccc ca	32
	<210> 279 <211> 32 <212> ADN <213> Secuencia artificial	
65	<220>	

<223> sintetizada químicamente
 <400> 279
 5 gaggagacgg tgacgatgt tcctgaccc ca 32
 <210> 280
 <211> 14
 <212> PRT
 <213> *Homo sapiens*
 10 <220>
 <221> MISC_FEATURE
 <222> (4)..(4)
 <223> Xaa es Ile o Met
 15 <220>
 <221> MISC_FEATURE
 <222> (5)..(8)
 <223> Xaa es Leu, Pro, His, Gln, Arg, Val, Ala, Asp, Glu o Gly
 20 <400> 280
 Val Ser Glu Xaa Xaa Xaa Xaa Xaa Gly Leu Leu Gly Glu Thr
 1 5 10
 25 <210> 281
 <211> 13
 <212> ADN
 <213> Secuencia artificial
 30 <220>
 <223> sintetizada químicamente
 <400> 281
 35 ttactgtgca aga 13
 <210> 282
 <211> 13
 <212> ADN
 <213> Secuencia artificial
 40 <220>
 <223> sintetizada químicamente
 <400> 282
 45 ttactgtgca aga 13
 <210> 283
 <211> 13
 <212> ADN
 50 <213> Secuencia artificial
 <220>
 <223> sintetizada químicamente
 55 <400> 283
 ttctgtgca aga 13
 <210> 284
 <211> 13
 60 <212> ADN
 <213> Secuencia artificial
 <220>
 65 <223> sintetizada químicamente

	<400> 284 ctactgtgcc aga	13
5	<210> 285 <211> 13 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 285 tggggccagg gaa	13
15	<210> 286 <211> 13 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 286 tggggcgag gga	13
30	<210> 287 <211> 13 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
35	<400> 287 tggggccaag gca	13
40	<210> 288 <211> 13 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
45	<400> 288 tggggccagg gca	13
50	<210> 289 <211> 13 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente	
	<400> 289 tggggtcagg gca	13
60	<210> 290 <211> 46 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	

5	<220>		
	<221> misc_feature		
	<222> (22)..(22)		
	<223> n es a, t, c o g		
10	<220>		
	<221> misc_feature		
	<222> (24)..(25)		
	<223> n es a, t, c o g		
15	<400> 290		
	attactgtgc gagaggagac gnsnncgtct cttggggcca gggaac	46	
	<210> 291		
	<211> 46		
20	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> sintetizada químicamente		
25	<220>		
	<221> misc_feature		
	<222> (22)..(22)		
	<223> n es a, t, c o g		
30	<400> 291		
	attactgtgc gagaggagac gncgtctctt ggggccaggg aaccct	46	
	<210> 292		
	<211> 34		
35	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> sintetizada químicamente		
40	<400> 292		
	ttatgtgtat agggttccct ggccccaaga gacg	34	
	<210> 293		
	<211> 34		
45	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> sintetizada químicamente		
50	<400> 293		
	gtgatctgta cctattactg tgcgagagga gacg	34	
	<210> 294		
	<211> 46		
55	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> sintetizada químicamente		
60	<220>		
	<221> misc_feature		
	<222> (22)..(22)		
	<223> n es a, t, c o g		
65	<220>		
	<221> misc_feature		

ES 2 542 744 T3

<222> (24)..(25)
 <223> n es a, t, c o g

5 <400> 294
 taatgacacg ctctcctctg cnsnngcaga gaaccccggt cccttg 46

10 <210> 295
 <211> 14
 <212> PRT
 <213> Secuencia artificial

15 <220>
 <223> sintetizada químicamente

20 <220>
 <221> MISC_FEATURE
 <222> (7)..(7)
 <223> X se selecciona entre Ala, Asp, Glu, Gly y Val

25 <220>
 <221> MISC_FEATURE
 <222> (8)..(8)
 <223> X se selecciona entre Ala, Arg, Asn, Asp, Cys, Glu, Gly, His, Ile, Leu, Phe, Pro, Ser, Tyr y Val

30 <400> 295
 Tyr Cys Ala Arg Gly Asp Xaa Xaa Val Ser Trp Gly Gln Gly
 1 5 10

35 <210> 296
 <211> 46
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> sintetizada químicamente

45 <220>
 <221> misc_feature
 <222> (22)..(22)
 <223> n es a, t, c o g

50 <400> 296
 taatgacacg ctctcctctg cngcagagaa ccccggtccc ttggga 46

55 <210> 297
 <211> 14
 <212> PRT
 <213> Secuencia artificial

60 <220>
 <223> sintetizada químicamente

65 <220>
 <221> MISC_FEATURE
 <222> (7)..(7)
 <223> X se selecciona entre Ala, Asp, Gly y Val

70 <400> 297
 Tyr Cys Ala Arg Gly Asp Xaa Val Ser Trp Gly Gln Gly Thr
 1 5 10

75 <210> 298
 <211> 24

	<212> ADN		
	<213> Secuencia artificial		
5	<220>		
	<223> sintetizada químicamente		
	<400> 298		
	gcgaaaggag acgccccgt ctct	24	
10	<210> 299		
	<211> 24		
	<212> ADN		
	<213> Secuencia artificial		
15	<220>		
	<223> sintetizada químicamente		
	<400> 299		
20	cgcttcctc tgcgggggca gaga	24	
	<210> 300		
	<211> 8		
	<212> PRT		
	<213> Secuencia artificial		
25	<220>		
	<223> sintetizada químicamente		
30	<400> 300		
	Ala Lys Gly Asp Ala Pro Val Ser		
	1 5		
35	<210> 301		
	<211> 24		
	<212> ADN		
	<213> Secuencia artificial		
40	<220>		
	<223> sintetizada químicamente		
	<400> 301		
	gcgagaggag acgccttcgt ctct	24	
45	<210> 302		
	<211> 24		
	<212> ADN		
	<213> Secuencia artificial		
50	<220>		
	<223> sintetizada químicamente		
	<400> 302		
	cgcttcctc tgcggaagca gaga	24	
55	<210> 303		
	<211> 8		
	<212> PRT		
	<213> Secuencia artificial		
60	<220>		
	<223> sintetizada químicamente		
	<400> 303		

	Ala	Arg	Gly	Asp	Ala	Phe	Val	Ser
	1				5			
5	<210> 304 <211> 24 <212> ADN <213> Secuencia artificial							
10	<220> <223> sintetizada químicamente							
	<400> 304 gcgagaggag acgagtacgt ctct 24							
15	<210> 305 <211> 24 <212> ADN <213> Secuencia artificial							
20	<220> <223> sintetizada químicamente							
	<400> 305 cgctctctc tgctatgca gaga 24							
25	<210> 306 <211> 8 <212> PRT <213> Secuencia artificial							
30	<220> <223> sintetizada químicamente							
	<400> 306							
	Ala	Arg	Gly	Asp	Glu	Tyr	Val	Ser
	1				5			
35	<210> 307 <211> 24 <212> ADN <213> Secuencia artificial							
40	<220> <223> sintetizada químicamente							
45	<400> 307 gcgagaggag acgagctcgt ctct 24							
50	<210> 308 <211> 24 <212> ADN <213> Secuencia artificial							
	<220> <223> sintetizada químicamente							
55	<400> 308 cgctctctc tgctcgagca gaga 24							
60	<210> 309 <211> 8 <212> PRT <213> Secuencia artificial							

<220>
 <223> sintetizada químicamente
 <400> 309
 5
 Ala Arg Gly Asp Glu Leu Val Ser
 1 5
 <210> 310
 <211> 24
 10 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizada químicamente
 15
 <400> 310
 gcgagaggag acggctgcgt ctct 24
 <210> 311
 20 <211> 24
 <212> ADN
 <213> Secuencia artificial
 <220>
 25 <223> sintetizada químicamente
 <400> 311
 cgctctctc tgccgacga gaga 24
 30 <210> 312
 <211> 8
 <212> PRT
 <213> Secuencia artificial
 35 <220>
 <223> sintetizada químicamente
 <400> 312
 Ala Arg Gly Asp Gly Cys Val Ser
 40 1 5
 <210> 313
 <211> 24
 45 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizada químicamente
 50
 <400> 313
 gcgagaggag acgagcccgt ctct 24
 <210> 314
 <211> 24
 55 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> sintetizada químicamente
 60
 <400> 314
 cgctctctc tgctcgggca gaga 24

5	<210> 315 <211> 8 <212> PRT <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente <400> 315	Ala Arg Gly Asp Glu Pro Val Ser 1 5
15	<210> 316 <211> 24 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente <400> 316 gcgagaggag acgggatcgt ctct	24
25	<210> 317 <211> 24 <212> ADN <213> Secuencia artificial	
30	<220> <223> sintetizada químicamente <400> 317 cgctctctc tgcctagca gaga	24
35	<210> 318 <211> 8 <212> PRT <213> Secuencia artificial	
40	<220> <223> sintetizada químicamente <400> 318	
45		Ala Arg Gly Asp Gly Ile Val Ser 1 5
50	<210> 319 <211> 24 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente <400> 319 gcgaaaggag acgggcgcgt ctct	24
60	<210> 320 <211> 24 <212> ADN <213> Secuencia artificial	

	<220>		
	<223> sintetizada químicamente		
5	<400> 320 cgctttcctc tgcccgcgca gaga	24	
	<210> 321 <211> 8 <212> PRT <213> Secuencia artificial		
10	<220> <223> sintetizada químicamente		
15	<400> 321		
	Ala Lys Gly Asp Gly Arg Val Ser		
		1	5
20	<210> 322 <211> 24 <212> ADN <213> Secuencia artificial		
25	<220> <223> sintetizada químicamente		
	<400> 322 gcgagaggag acgacgccgt ctct	24	
30	<210> 323 <211> 24 <212> ADN <213> Secuencia artificial		
35	<220> <223> sintetizada químicamente		
40	<400> 323 cgctctcctc tgctgcggca gaga	24	
	<210> 324 <211> 8 <212> PRT <213> Secuencia artificial		
45	<220> <223> sintetizada químicamente		
50	<400> 324		
	Ala Arg Gly Asp Asp Ala Val Ser	1	5
55	<210> 325 <211> 24 <212> ADN <213> Secuencia artificial		
60	<220> <223> sintetizada químicamente		

	<400> 325 gcgagaggag acgggtccgt ctct	24
5	<210> 326 <211> 24 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 326 cgctctctc tgcccaggca gaga	24
15	<210> 327 <211> 8 <212> PRT <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente <400> 327	
25	Ala Arg Gly Asp Gly Ser Val Ser 1 5	
30	<210> 328 <211> 24 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
35	<400> 328 gcgagaggag acgccgtcgt ctct	24
40	<210> 329 <211> 24 <212> ADN <213> Secuencia artificial	
45	<220> <223> sintetizada químicamente	
	<400> 329 cgctctctc tgcggcagca gaga	24
50	<210> 330 <211> 8 <212> PRT <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente <400> 330	
60	Ala Arg Gly Asp Ala Val Val Ser 1 5	
	<210> 331 <211> 24	

	<212> ADN		
	<213> Secuencia artificial		
5	<220>		
	<223> sintetizada químicamente		
	<400> 331		
	gcgaaaggag acgtgtccgt ctct	24	
10	<210> 332		
	<211> 24		
	<212> ADN		
	<213> Secuencia artificial		
15	<220>		
	<223> sintetizada químicamente		
	<400> 332		
20	cgctttcttc tgcacaggca gaga	24	
	<210> 333		
	<211> 8		
	<212> PRT		
	<213> Secuencia artificial		
25	<220>		
	<223> sintetizada químicamente		
	<400> 333		
30			
	Ala Lys Gly Asp Val Ser Val Ser		
	1 5		
	<210> 334		
	<211> 24		
35	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> sintetizada químicamente		
40	<400> 334		
	gcgagaggag acgggcacgt ctct	24	
	<210> 335		
45	<211> 24		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
50	<223> sintetizada químicamente		
	<400> 335		
	cgctctcttc tgcccgtgca gaga	24	
55	<210> 336		
	<211> 8		
	<212> PRT		
	<213> Secuencia artificial		
60	<220>		
	<223> sintetizada químicamente		
	<400> 336		

		Ala Arg Gly Asp Gly His Val Ser	
		1 5	
5	<210> 337		
	<211> 24		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
10	<223> sintetizada químicamente		
	<400> 337		
	gcgagaggag acgagaccgt ctct	24	
15	<210> 338		
	<211> 24		
	<212> PRT		
	<213> Secuencia artificial		
	<220>		
20	<223> sintetizada químicamente		
	<400> 338		
	Cys Gly Cys Thr Cys Thr Cys Cys Thr Cys Thr Gly Cys Thr Cys Thr		
	1 5 10 15		
	Gly Gly Cys Ala Gly Ala Gly Ala		
	20		
25	<210> 339		
	<211> 8		
	<212> PRT		
	<213> Secuencia artificial		
30	<220>		
	<223> sintetizada químicamente		
	<400> 339		
35			
	Ala Arg Gly Asp Glu Thr Val Ser		
	1 5		
40	<210> 340		
	<211> 24		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
45	<223> sintetizada químicamente		
	<400> 340		
	gcgagaggag acgccatcgt ctct	24	
50	<210> 341		
	<211> 24		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
55	<223> sintetizada químicamente		

	<400> 341 cgctctctc tgcggtagca gaga	24
5	<210> 342 <211> 8 <212> PRT <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
	<400> 342	
15	Ala Arg Gly Asp Ala Ile Val Ser 1 5	
	<210> 343 <211> 21 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
25	<400> 343 gcgagaggag acggcgtctc t	21
30	<210> 344 <211> 21 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
35	<400> 344 cgctctctc tgcgcagag a	21
40	<210> 345 <211> 7 <212> PRT <213> Secuencia artificial	
45	<220> <223> sintetizada químicamente	
	<400> 345	
	Ala Arg Gly Asp Gly Val Ser 1 5	
50	<210> 346 <211> 21 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente	
60	<400> 346 gcgagaggag acgacgtctc t	21
	<210> 347 <211> 21 <212> ADN	

	<213> Secuencia artificial		
	<220>		
5	<223> sintetizada químicamente		
	<400> 347		
	cgctctctc tgctgcagag a	21	
10	<210> 348		
	<211> 7		
	<212> PRT		
	<213> Secuencia artificial		
15	<220>		
	<223> sintetizada químicamente		
	<400> 348		
		Ala Arg Gly Asp Asp Val Ser	
20		1 5	
	<210> 349		
	<211> 21		
	<212> ADN		
25	<213> Secuencia artificial		
	<220>		
	<223> sintetizada químicamente		
30	<400> 349		
	gcgagaggag acgccgtctc t	21	
	<210> 350		
	<211> 21		
35	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> sintetizada químicamente		
40	<400> 350		
	cgctctctc tgcggcagag a	21	
	<210> 351		
45	<211> 7		
	<212> PRT		
	<213> Secuencia artificial		
	<220>		
50	<223> sintetizada químicamente		
	<400> 351		
		Ala Arg Gly Asp Ala Val Ser	
		1 5	
55	<210> 352		
	<211> 21		
	<212> ADN		
	<213> Secuencia artificial		
60	<220>		
	<223> sintetizada químicamente		

ES 2 542 744 T3

	<400> 352 gcgaaaggag acgacgtctc t	21	
5	<210> 353 <211> 21 <212> ADN <213> Secuencia artificial		
10	<220> <223> sintetizada químicamente		
	<400> 353 cgcttcctc tgctgcagag a	21	
15	<210> 354 <211> 7 <212> PRT <213> Secuencia artificial		
20	<220> <223> sintetizada químicamente		
	<400> 354		
25			Ala Lys Gly Asp Asp Val Ser 1 5
30	<210> 355 <211> 21 <212> ADN <213> Secuencia artificial		
	<220> <223> sintetizada químicamente		
35	<400> 355 gcgagaggag acgtcgtctc t	21	
40	<210> 356 <211> 21 <212> ADN <213> Secuencia artificial		
45	<220> <223> sintetizada químicamente		
	<400> 356 cgctctctc tgcagcagag a	21	
50	<210> 357 <211> 7 <212> PRT <213> Secuencia artificial		
55	<220> <223> sintetizada químicamente		
	<400> 357		
60			Ala Arg Gly Asp Val Val Ser 1 5
	<210> 358 <211> 21		

	<212> ADN		
	<213> Secuencia artificial		
5	<220>		
	<223> sintetizada químicamente		
	<400> 358		
	gcgaaaggag acggcgtctc t	21	
10	<210> 359		
	<211> 21		
	<212> ADN		
	<213> Secuencia artificial		
15	<220>		
	<223> sintetizada químicamente		
	<400> 359		
20	cgctttctc tgccgcagag a	21	
	<210> 360		
	<211> 7		
	<212> PRT		
25	<213> Secuencia artificial		
	<220>		
	<223> sintetizada químicamente		
30	<400> 360		
	Ala Lys Gly Asp Gly Val Ser		
	1 5		
35	<210> 361		
	<211> 21		
	<212> ADN		
	<213> Secuencia artificial		
40	<220>		
	<223> sintetizada químicamente		
	<400> 361		
	gcgaaaggag acgccgtctc t	21	
45	<210> 362		
	<211> 21		
	<212> ADN		
	<213> Secuencia artificial		
50	<220>		
	<223> sintetizada químicamente		
	<400> 362		
	cgctttctc tgccgcagag a	21	
55	<210> 363		
	<211> 7		
	<212> PRT		
	<213> Secuencia artificial		
60	<220>		
	<223> sintetizada químicamente		
	<400> 363		

	Ala	Lys	Gly	Asp	Ala	Val	Ser
	1				5		
5	<210> 364 <211> 21 <212> ADN <213> Secuencia artificial						
10	<220> <223> sintetizada químicamente						
	<400> 364 gcgaaaggag acgtcgtctc t						
				21			
15	<210> 365 <211> 21 <212> ADN <213> Secuencia artificial						
20	<220> <223> sintetizada químicamente						
	<400> 365 cgcttcctc tgcagcagag a						
				21			
25	<210> 366 <211> 7 <212> PRT <213> Secuencia artificial						
30	<220> <223> sintetizada químicamente						
	<400> 366						
35	Ala	Lys	Gly	Asp	Val	Val	Ser
	1				5		

REIVINDICACIONES

1. Un método para producir una colección de ácidos nucleicos, en el que cada ácido nucleico codifica un dominio variable de inmunoglobulina humana que comprende una pluralidad de secuencias de región determinante de complementariedad 3 (CDR3) aisladas por separado del repertorio de dominio variable de inmunoglobulina de una especie de mamífero o un mamífero no humano inmunizado, comprendiendo el método:

(a) proporcionar una pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor que codifican dominios variables de inmunoglobulina humana distintos, comprendiendo cada secuencia de ácido nucleico de Marco conservado Aceptor una primera región marco conservada (FR1), una segunda región marco conservada (FR2), una tercera región marco conservada (FR3) y una cuarta región marco conservada (FR4), en las que las regiones FR1 y FR2 están intercaladas con una región determinante de complementariedad 1 (CDR1), las regiones FR2 y FR3 están intercaladas con una región determinante de complementariedad 2 (CDR2), y las regiones FR3 y FR4 están intercaladas con una secuencia de ácido nucleico de relleno que comprende al menos dos sitios de reconocimiento de enzimas de restricción de Tipo IIs intercalados con una secuencia de ácido nucleico aleatoria que codifica un polipéptido que realiza la función de una región CDR3 de inmunoglobulina variable;

(b) proporcionar una pluralidad de secuencias de ácido nucleico diversificadas que codifican secuencias de región determinante de complementariedad 3 (CDR3) del repertorio de inmunoglobulina de especie de mamífero o el mamífero no humano inmunizado en el que cada una de la pluralidad de secuencias de ácido nucleico diversificadas comprende un sitio de reconocimiento de enzimas de restricción de Tipo IIs en cada extremo;

(c) digerir cada una de la pluralidad de secuencias de ácido nucleico que codifican las regiones CDR3 usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (b) y digerir la secuencia de ácido nucleico de relleno de la etapa (a) del Marco conservado Aceptor usando una enzima de restricción de Tipo IIs que se une con el sitio de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a); y

(d) ligar las secuencias de ácido nucleico digeridas que codifican las regiones CDR3 o las secuencias de aminoácidos de la etapa (c) en el Marco conservado Aceptor digerido de la etapa (c) de modo que las regiones FR3 y FR4 estén intercaladas con las secuencias de ácido nucleico 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 un dominio variable de inmunoglobulina completo que no contienen los sitios de reconocimiento de enzimas de restricción de Tipo IIs de las etapas (a) y (b).

2. El método de la reivindicación 1, en el que la etapa (b) se realiza amplificando la secuencia de CDR3 de una especie de mamífero o el mamífero no humano inmunizado usando cebadores oligonucleotídicos que contienen un sitio de restricción de Tipo IIs.

3. El método de la reivindicación 2, en el que el cebador oligonucleotídico se diseña para potenciar la compatibilidad entre la secuencia de CDR3 de mamífero y el Marco conservado Aceptor que codifica un dominio variable de inmunoglobulina humana.

4. El método de la reivindicación 3, en el que el cebador oligonucleotídico se diseña para modificar una secuencia de ácido nucleico en un extremo de la secuencia de CDR3 de mamífero para producir una secuencia de nucleótidos cohesiva compatible en el Marco conservado Aceptor que codifica un dominio variable de inmunoglobulina humana.

5. El método de la reivindicación 1, en el que la especie de mamífero es ser humano, primate no humano, roedor, canino, felino, oveja, cabra, vaca, caballo, un miembro de la familia Camelidae, llama, camello, dromedario o cerdo; o en el que el mamífero no humano es primate no humano, roedor, canino, felino, oveja, cabra, vaca, caballo, llama, camello, dromedario o cerdo.

6. El método de la reivindicación 1, en el que los sitios de reconocimiento de enzimas de restricción de Tipo IIs de la etapa (a) y la etapa (b) se reconocen por una enzima de restricción de Tipo IIs diferente, preferentemente en el que los sitios de reconocimiento de enzimas de restricción de Tipo IIs son sitios de reconocimiento de BsmBI, sitios de reconocimiento de BsaI, sitios de reconocimiento de FokI o una combinación de los mismos.

7. El método de la reivindicación 1, en el que las secuencias de ácido nucleico diversificadas que codifican secuencias de CDR3 codifican secuencias de CDR3 de cadena pesada (CDR H3), secuencias de CDR3 de cadena ligera (CDR L3) o una combinación de las mismas.

8. El método de la reivindicación 1, en el que la secuencia de ácido nucleico de Marco conservado Aceptor comprende una secuencia de gen variable de cadena pesada humana seleccionada de VH1-2, VH1-69, VH1-18, VH3-30, VH3-48, VH3-23 y VH5-51.

9. El método de la reivindicación 1, en el que la secuencia de ácido nucleico de Marco conservado Aceptor comprende una secuencia de gen variable de cadena ligera kappa humana, preferentemente en el que la secuencia de gen variable de cadena ligera kappa humana se selecciona de VK1-33, VK1-39, VK3-11, VK3-15 y VK3-20.

10. El método de la reivindicación 1, en el que la secuencia de ácido nucleico de Marco conservado Aceptor comprende una secuencia de gen variable de cadena ligera lambda humana, preferentemente en el que la secuencia de gen variable de cadena ligera lambda humana se selecciona de VL1-44 y VL1-51.
- 5 11. El método de la reivindicación 1, en el que la pluralidad de secuencias de ácido nucleico de Marco conservado Aceptor comprende una mezcla de al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena pesada variable (VH) y al menos una secuencia de ácido nucleico de Marco conservado Aceptor de cadena ligera variable.
- 10 12. El método de la reivindicación 1, que comprende además las etapas de (e) clonar la biblioteca de ácidos nucleicos que codifican dominios variables de inmunoglobulina de la etapa (d) en un vector de expresión y (f) transformar el vector de expresión de la etapa (e) en una célula hospedadora y cultivar la célula hospedadora en condiciones suficientes para expresar una pluralidad de dominios variables de inmunoglobulina codificados por la biblioteca.
- 15 13. El método de la reivindicación 12, en el que la célula hospedadora es *E. coli*.
14. El método de la reivindicación 13, en el que el vector de expresión es un fagémido o un vector de fago.
- 20 15. El método de la reivindicación 1, en el que la etapa (b) se realiza amplificando la secuencia de CDR H3 del mamífero no humano usando cebadores oligonucleotídicos que contienen un sitio de restricción IIs de FokI.

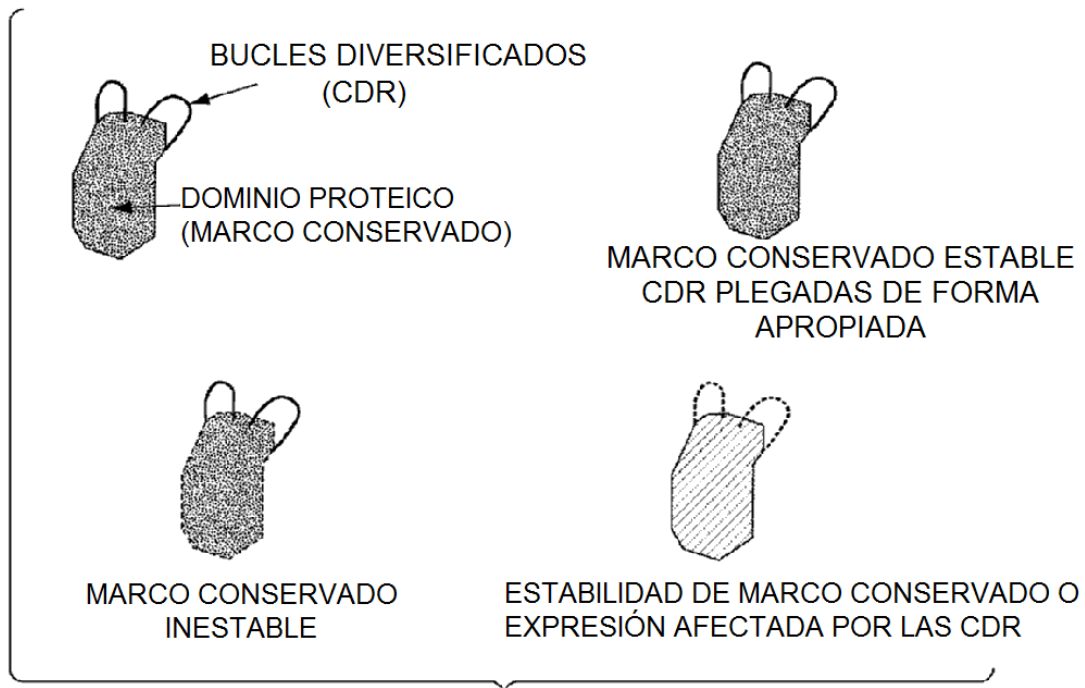


Fig. 1A

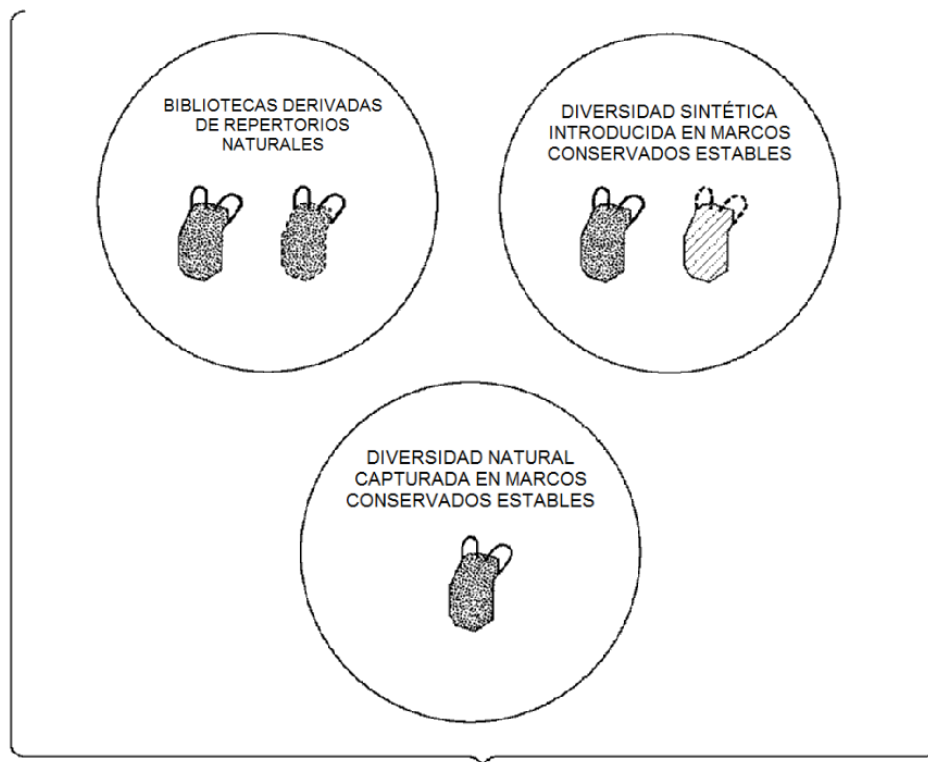


Fig. 1B

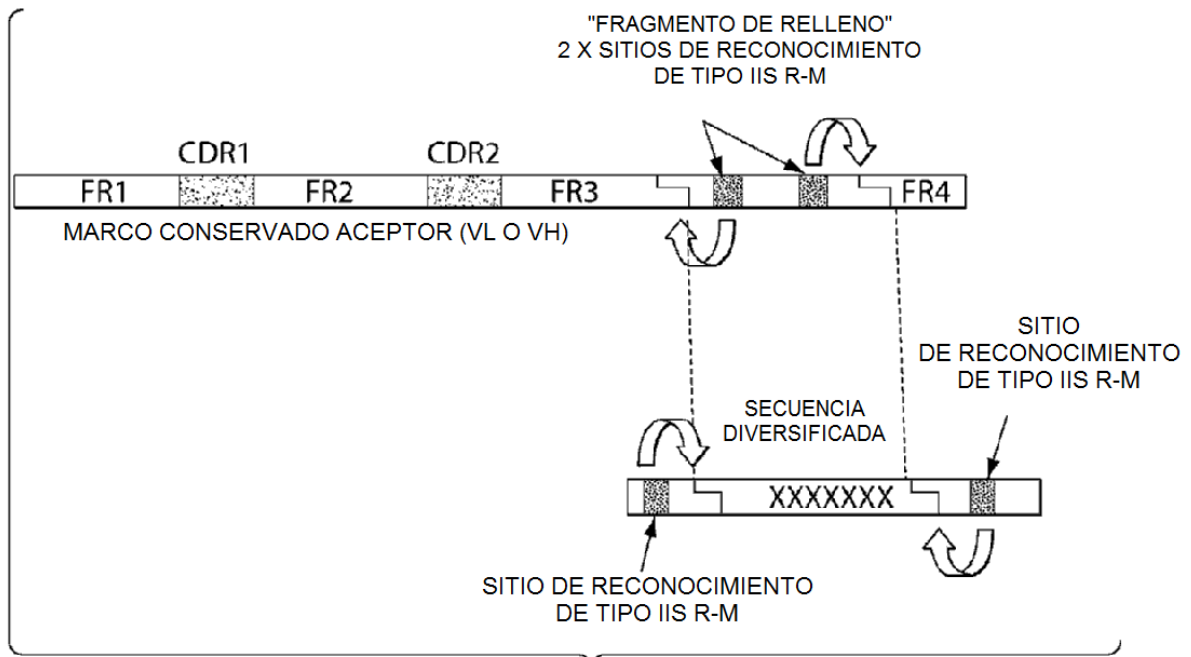


Fig. 2

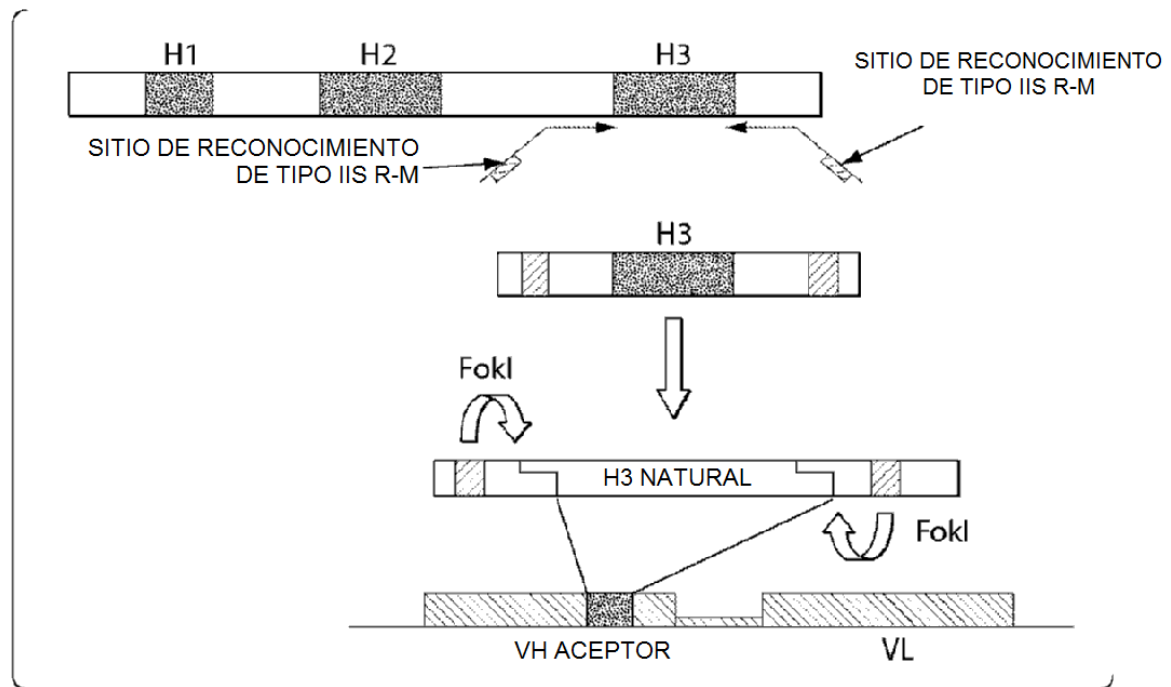
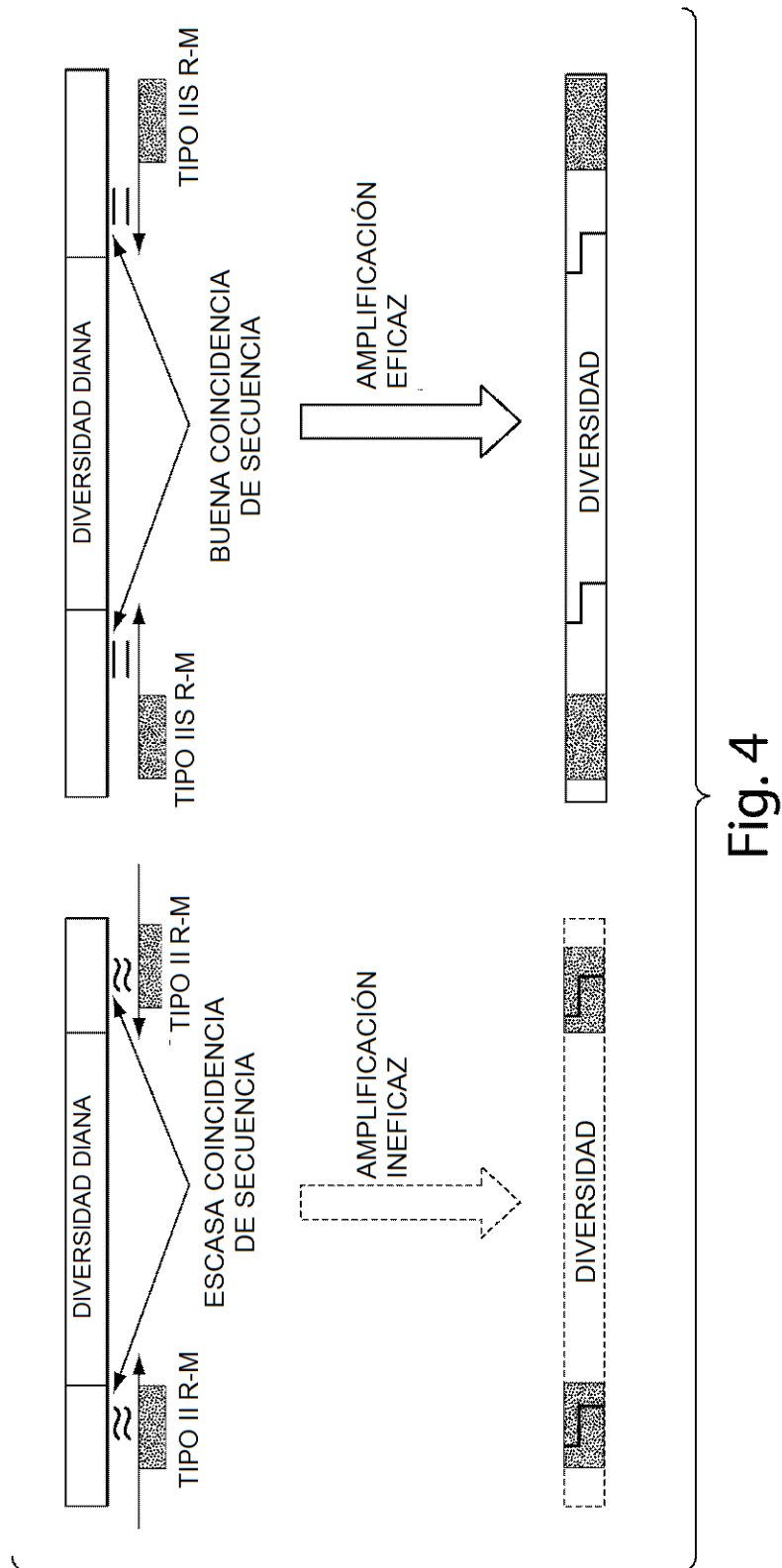


Fig. 3



VH1-2 QVQLVQSGAEVKKPGASVKVSCKASGYTFTGYYMHVRQAPGQGLEWMGWNPNSGGTNYAQKFOGRVTMTTRDTSISTAYMELSLRLSDDTAVYYCAR SEC ID N° : 1
 VH1-18 QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGISWVRQAPGQGLEWMGWISAYNGNTNYAQKLGQGRVTMTTDTSTSTAYMELSLRLSDDTAVYYCAR SEC ID N° : 2
 VH1-69 QVQLVQSGAEVKKPGSSVKVSCKASGGTFSYAISSWRQAPGQGLEWMGGIIPILFGTANYAQKFOGRVTITADESTSTAYMELSSLRSEDTAVYYCAR SEC ID N° : 3
 VH3-30 QVQLVESGGGVQPGRSLRLSCAASGFTFSYAMHWVRQAPGKGLEWVAVISYDGSNKYYADSVKGRFTISRDNKNTLYLOMNSLRAEDTAVYYCAR SEC ID N° : 4
 VH3-48 EVQLVESGGGLVQPGGSLRLSCAASGFTFSYSMMWVRQAPGKGLEWVSYSSTIYYADSVKGRFTISRDNKNTLYLOMNSLRAEDTAVYYCAR SEC ID N° : 5
 VH3-23 EVQLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVRQAPGKGLEWVAISGSGSTIYYADSVKGRFTISRDNKNTLYLOMNSLRAEDTAVYYCAR SEC ID N° : 6
 VH5-51 EVQLVQSGAEVKKPGESLKISCKGSGYSFTSYWIGWVRQMPGKGLEWMGLIYPGSDSDTRYSPSFGQGVTTISADKSIISTAYLQWSSSLKASDTAMYYCAR SEC ID N° : 7
 VK1-33 DIQMTQSPSSLSASVGRVTITCQASQDIS NYLNWYQQKPKGAPKLLIYDASNLETGVPSPRFSGSGSGTDFTFTISSLOPEDIAITYYCOQYDNLPTTV SEC ID N° : 8
 VK1-39 DIQMTQSPSSLSASVGRVTITCRASQIS SYLNWYQQKPKGAPKLLIYAASLQSGVPSRFSGSGSGTDFTLTITSSLOPEDFATYYCOQSYSTPTTV SEC ID N° : 9
 VK3-11 EIVLTQSPATLSLSPGERATLSCRASQSVS SYLAWYQQKPGQAPRLIYDASNRATGIPARFSGSGSGTDFTLTITSSLEPEDFVAVYCOQQRNWPPTV SEC ID N° : 10
 VK3-15 EIVMTQSPATLSVSPGERATLSCRASQSVS SNLAWYQQKPGQAPRLIYTGASTRATGIPARFSGSGSGTEFTLTITSSLOQSEDFVAVYCOQYNNWPPTV SEC ID N° : 11
 VK3-20 EIVLTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQKPGQAPRLIYGCASSRATGIPDRFSGSGSGTDFTLTISRLEPEDFVAVYCOQYGSSTPTV SEC ID N° : 12
 VL1-44 QSVLTQPPSASGTPGQRVTITSCGSSSNIGSNVTNWTNQQLPGTAPKLLIYNNORPSGVDPDRFSGSKGTSASLAISGLQSEDEADVYCAAWDDSLNG SEC ID N° : 13
 VL1-51 QSVLTQPPSVSAAPGQKVTITSCGSSSNIGNNVSWYQQLPGTAPKLLIYENMKRPSSGIPDRFSGSKGTSATLGTGLTGTGDEADYVCGTWDSSLSA SEC ID N° : 14

Fig. 5

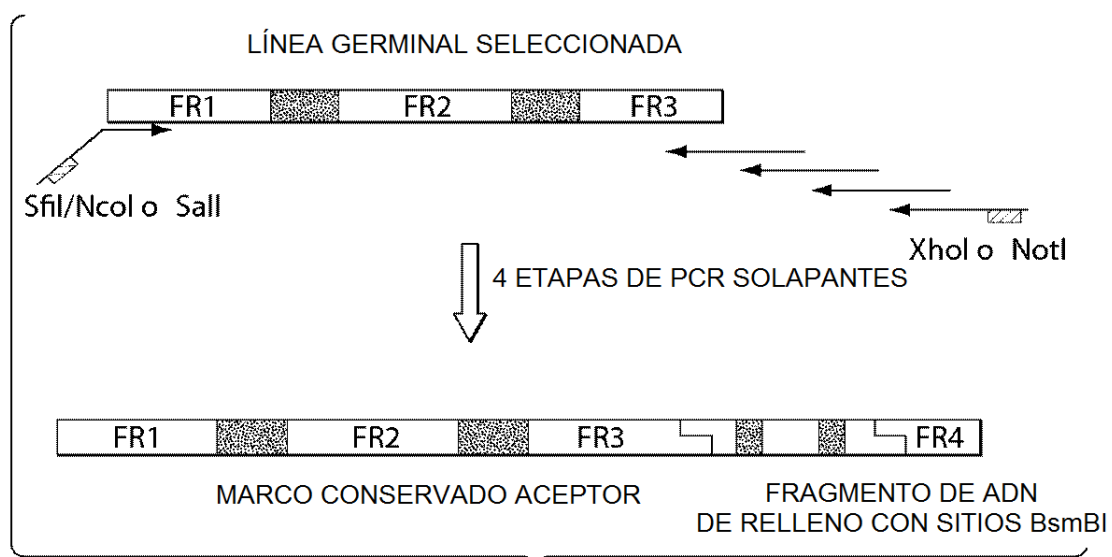


Fig. 6

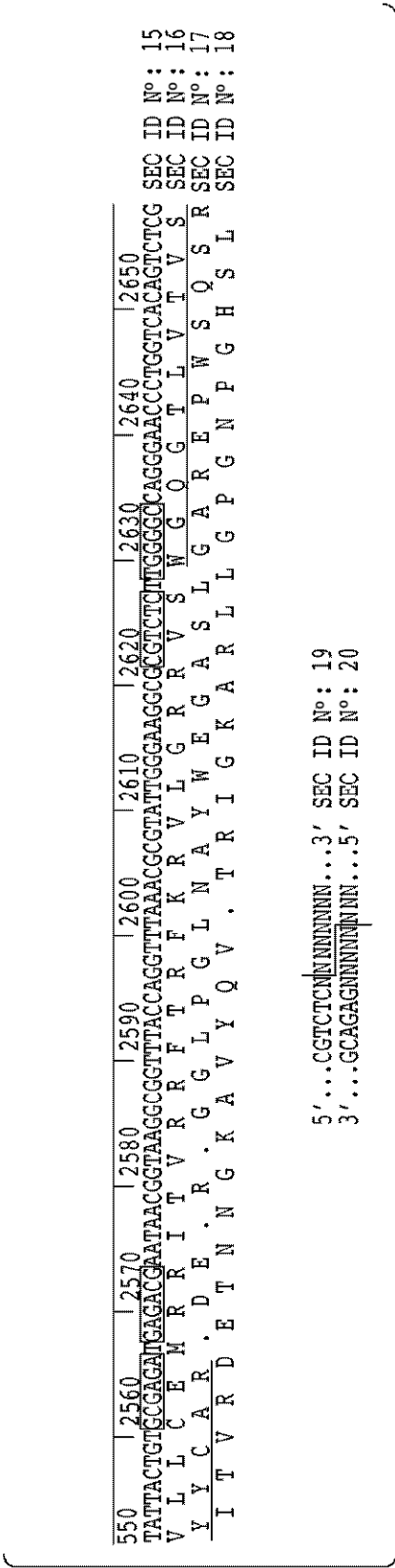


Fig. 7

VH1-2 AF

```

caggtgcagc tgggtcagtc ggaagaagc ctggggcctc agtgaaggtc tcttgcagg cttcgggata caccttcacc ggctactata tgcactgggt ggaagggcc 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 l h l h r l l y a 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 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 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.....>
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 g y y m h w v r q a p g q

ggcttgagt gattggatgg atcaacctc acagtgggtg cacaaactat gcacagaadt ttccaggcag ggtcacccatg accagggaac cgtccatcag cacaggctac atggagctga gcaggctgag
r a ~ v d g m d q p - q w w h k l c t e v s g g q g h h d g 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 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

atctgaagc agggcggtgt attactgtc gagatgagac gaataacggt aaggcggttt accaggttt aacgggtatt gggaaggcc gtccttggg gccagggaac cctgggtcaca gtctcg SEC ID N°: 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 SEC ID N°: 22
r s d d t a v y c a r - d e - r - g g l p g l n a y w e g a a s l g a r e p w s q s SEC ID N°: 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 SEC ID N°: 24
>.....VH1-2.....>
r s d d t a v 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 SEC ID N°: 23

```

Fig. 8

VH1-18 AF

caggtgcagc tggtagctgc gtgaagaagc ctgggggctc agtgaagtc tccgaagtc cttctggtta cactttacc agctatgta 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 q l w y q 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

gccttgatg gatggatgg atcagcgtt acaatgtaa cacaactat gcacagaagc tccaggcag agtcaacctg accacagaca catcaagag cacagctac atggagctga ggagcctgag
r a - v d g m d g r l q w - h k l c t e a p g q 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 q k l q g r v t m t t d t s t s t a y m e l r s l
g l s g w d g s a l t m v t q t m h r s s r a e s p - p q t h p r a q p t w s - g a -
>.....VH1-18.....>
g l e w m g w i s a y n g n t n y a q k l q g r v t m t t d t s t s t a y m e l r s l

atctgacgac acggcogtgt attactgtc gagatgagac gaataacggt aaggcggtt accaggttta aacgggtatt gggaaggcgc gtctcttggg gccagggaac cctgttcaca gtctcg SEC ID N°: 25
i - r h g r 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 SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 27

Fig. 8 Cont.

VH1-69 AF

caggtgcagc tgggtgaagtc tggggctgag gtaagaagc ctggggcctc ggtgaagtc tctgaagc ctttgaagc agctatgcta tcagctgggt ggcacaggcc cctggacaag
 p g a a g a v w g - g e a w v l g e g l l q g f w r h l q q l c y q 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
 ggcctgagtg gatggaggg atcatcccta tccttggtac agcaaaactac gcacagaagt tcacaggcag agtcaagatt accggcgagc aatccacagc 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 q p e
 g l e w m g g i i p i f g t a n y a q k f q g r v t i t a d e s t s t a y m e l s s l
 g l s g w e g s s l s l v q q t t h r s s r a e s r l p r t n p r a q p t w s - a a a -
 >.....VH1-69.....>
 g l e w m g g i l p i f g t a n y a q k f q g r v t i t a d e s t s t a y m e l s s l
 atctgaggac accggcgtgt attactgtc gagatgagac gaataacggt aaggcggttt accaggttta aacgcgtatt gggaaggcgc gtctcttggg gccagggaac cctgtgtaca gtctcg SEC ID N°: 29
 i - g h g r 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 SEC ID N°: 30
 r s e d t a v 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 SEC ID N°: 31
 d l r 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 SEC ID N°: 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 SEC ID N°: 31

Fig. 8 Cont.

VH3-23R AF

gagtgacgc tgttgaagtc tgggggagc ttgtacagc ctggggggctc cctgagactc tctgtgcag cctctgatt cactttagc agctatgca tgagctgggt ccgccaggct ccagggaaagg
 rgaa avgv wgr lgt a wgv pet l l c s l w i h l l - q l c h e l g p p g s r e
 evq l l e s g g g l v q p g g s l r l s c a a a g f t f s s y a m s w v r q 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.....>
 evq l l e s g g g l v q p g g s l r l s c a a a g f t f s s y a m s w v r q a p g k
 ggcggagtg ggtctagct attagtgtta gtgtgttag cacatactac gcagactccg tgaaggccg gttcacatc tccagagaca attccaagaa caagctgtat ctgcaaatga acagcctgag
 g a g v g l s y - w - w w - h i l 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
 agccgagac acgcccgat attactgtgc gagatgagac gaataacggt aaggcggttt accaggttta aacgggtatt gggaaggcgc gtccttggg gccagggaac cctgtgtcaca gtctcg SEC ID N°: 33
 s r g h g r i 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 SEC ID N°: 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 SEC ID N°: 35
 e p r t r p y i t v i 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 SEC ID N°: 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 SEC ID N°: 35

Fig. 8 Cont.

VH3-23K AF

gaggtgcagc tgttgagtc tggggaggc ttgttacgc ctgggggtc octgagactc tctgtgcag octctggatt caoctttagc agttatgca tgagctgggt ccgccaggct ccagggaagg
r g a a v g v w g r l g t a w g v p e t l l c s l w i h l l - g 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 q 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 q a p g k

ggctggagtg gttctcagct attagtgtg gtggtgtag cacatactac gcagactccg tgaaggccg gttcacatc tccagagaca attccaaga caagctgtat ctgcaaatga acagcctgag
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

agccgagac acggccgtat attactgtc gaatgagac gaataacggt aaggcggtt accaggttta aacgcgtatt gggaaggcgc gtctctggg gccagggaac cctggteaca gtctcg SEC ID N°: 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 q g t l v t v s SEC ID N°: 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 SEC ID N°: 39
e p r t r p y i t v r n 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 SEC ID N°: 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 SEC ID N°: 39

Fig. 8 Cont.

VH3-30 AF

```

cagtgacgc tgggggagc gtgtccagc ctggagctc cctgaactc tctgtcag cctctggatt caccttcagt agctatgcta tgcactgggt cggccaggct 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 q - l c y a l g p p g s r q
q v q l v e s g g g v 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 gcagactcog tgaaggcgc attcaccatc tccagagaca attccaagaa cagcctgtat ctgcaaatga acagccctgag
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 q 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

agctgagac acgctgtgtt attactgtc gagatgagac gaataacggt aaggcgttt accaggttta aacgggtatt ggaaggcgc gtctcttggg gccagggaac cctggtcaca gtctog SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 43

```

Fig. 8 Cont.

VH3-48 AF

```

gaggtgcagc tgggggagc ttggtacagc ctgggggggc cotgagactc tctgttcag cotctggatt cacottcagt agctatagca tgaactgggt ccgccagggt ccagggaagg
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 g - 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 q g r
>>.....VH3-48.....>
e v q l v e s 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
>>.....VH3-48.....>

ggctggagtg ggttccatcac attagtagta gtagtagtac catatactac gcagactctg tgaaggcgcg attcaccatc tccagagaca atgccaagaa ctcactgtat ctgcaaatga acagcctgag
g a g v g f i h - - - y h i l r r l c e g p i h h l q r q c q 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 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
>>.....VH3-48.....>

agccgaggac acgctctgtt attactgtgc gagatgagac gaataacggtt aagcgggtt accaggttta aacgggtatt ggaaggcgc gtctcttggg gccagggaac cctggtcaca gtctcg SEC ID N°: 45
s r 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 SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 47
>>.....VH3-48.....>

```

Fig. 8 Cont.

VH5-51 AF

```

gaggtgcagc tggtagcagc gtgaaagc ccggggagtc ttgtagatc ttctgtaagg gttctggata cagcttiacc agctactgga tccggtgggt gcgccagatg 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 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 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 gatgggagc atctatctg gtgactctga taccagatc agccggtctt tccaaggcca ggtcacatc tcaagcaga agtccatcag caccgctac ctgactgga gcagctgaa
r p g v d g d h l s w - l - y q i q p v l p r p g 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 q p t s p s a 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

ggctcggagc accgccatgt attactgtc gagatgagac gaataacggt aaggcggttt accaggttta aacgggtatt gggaaggccg gtctcttggg gccagggaac cctggtcaca gtctcg SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 51

```

Fig. 8 Cont.

VK1-33G AF

```

gacatccaga tgaccacgac tcactctcc ctgtctgcac cagactcacc atcacttccc aggcagatca ggcattagc aactatttaa attgtatca ggaagaacca gggaaagccc
g h p 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 - 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 q 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
>>.....VK1-33G.....>
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 q d i s n y l n w y q q k p g k a
>>.....VK1-33G.....>

ctaagctcct gatataagat gcattccaatt tggaaacagg ggtccaatca agtttcagtg gaagtgcac tgggacagat ttactttca ccatcagcag octgcagcct gaagatattg caaatatta
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 g 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
>>.....VK1-33G.....>
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 cgagacgaat aacggtaagg cggtttacc ggtttaaag cgtattgga aggcggtct catcggcca aggaaccaag gtggaatca aaggggggc cgca SEC ID N°: 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 SEC ID N°: 54
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 a k g p r w k s k g r p SEC ID N°: 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 SEC ID N°: 56
>.....VK1-33G.....>
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 a k g p r w k s k g r p SEC ID N°: 55

```

Fig. 8 Cont.

VH1-33R AF

```

gacatccaga tgaccagatc tcatctgcat ctgttagaga cagagtcacc atacttgc aggcagtgca aggcattagc aactatttaa attggtatca gcagaacca gggaagccc
ghp 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 - 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 q 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 q a s q d i s n y l n w y q q k p g k a
>>.....VK 1-33R.....>

ctaagctcct gatatagat geatccaatt tggaacacagg ggtcccatca aggttcagtg gaagtggatc tgggacagat ttatcttca ccateagtag cctgcagcct gaagatatg 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 h 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 l 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
>>.....VK 1-33R.....>

ctgtcagcag cgagacgaat aacggttaagg cggtttacca ggtttaaaag cgtattgga agcgcggtct catcggcca agggaccaag gtggaaatca aacggggggc cgca SEC ID N°: 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 SEC ID N°: 58
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 n g r p SEC ID N°: 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 SEC ID N°: 60
>.....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 SEC ID N°: 59

```

Fig. 8 Cont.

VK1-39G AF

```

gacatccaga tgaccacagtc tccatccctcc ctgtctgcat ctgtaggaga cagagtcacc atcaacttccc gggcaagtca gagcattagc agctatttaa attggtatca gcagaaacca gggaaagccc
ghp d d p v s i l l p v c i c r r q s h h h l p g k s e 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 r a s q s i s s 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 g q v r a l a a i - i g i s r n q g k p
>>.....VK 1-39G.....>
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 q k p g k a
>>.....VK 1-39G.....>

ctaagctcct gatctatgct gcacacagtt tgcgaagtgg ggtcccatca aggttcagtg gcagtgatc tggcacagat ticaacttca ccatcagcag totgcaacct 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 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 n l k i l q l t
>.....VK 1-39G.....>
p k l l i y a a s s l q s g v p s r f s g s g s g t d f t l t i s s l q p e d f a t y
>>.....VK 1-39G.....>

ctgtcagcag cggagcgaat aacggttaagg cggtttaccg gttttaaacg cgtattggga aggcgctct cattcgcca agggaccaag gtgaaatca aagggcgccg cga SEC ID N°: 61
l s a a 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 SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 64
>>.....VK 1-39G.....>
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 SEC ID N°: 63

```

Fig. 8 Cont.

VH1-39R AF

gacatccaga tgaccacgac tcactctctc ctgtctgpat ctgttaggaga cagagtcacc atcaattgcc gggcaagtcga gagcattagc agctatttaa attgttatca gcagaacca gggaaagccc
g h p d 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 r e s
d i q m t g 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 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 g q v r a l a a i - i g i s r n q g k p
>>.....VK 1-39R.....>
d i q m t g 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 q k p g k a
ctaagctcct gatctatgct gcattccagtt tgcaaatgg ggttcacgat gcaatgcacat ttactctca ccatacagag 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 k i l g l t
>.....VK 1-39R.....>
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 cggtttaoca ggtttaaag cgtattggga aggcggtct cattoggcca agggaccaag gtggaaatca aacggggggc cgca SEC ID N°: 65
l s a a r r i t v r i 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 SEC ID N°: 66
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 SEC ID N°: 67
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 SEC ID N°: 68
>.....VK 1-39R.....>
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 SEC ID N°: 67

Fig. 8 Cont.

VK3-11G AF

```

gaaatttgt tgacacagtc tccagccacc ctgtctttgt ctccaggga aagagccacc ctctcttga gggccagtca gagtgttagc agctacttag ccttggtacca acagaaacct ggcaggttc
gncvdtvssh pvhfvsrg ksh pll qgqs ec - ql lslvp 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
>>.....VK 3-11G.....>

ccaggctct catctatgat gcatacaaa gggccactgg catcccagcc aggttcagtg gcagtggttc tgggacagac ttactctca ccatcagcag cctagaccct gaagattttg cagtttatta
sqaphl - ci q q 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 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 i 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 t d f t l t i s s l e p e d f a v y
>>.....VK 3-11G.....>

ctgtcagcag cgagacgaat aacggtaagg cggttacca ggtttaaag cgtattgga aggcgggtct cattcgcca agggaccaag gtggaatca aagggcggc cgca SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 72
>>.....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 SEC ID N°: 71

```

Fig. 8 Cont.

VH3-11R AF

```

gaaattgtg tgacacagtc tccagccacc ctgtctttgt ciccagggga aagagccacc ctctctctga gggccagtca gactgttagc agctacttag cctgttacc acagaaact ggcaggctc
gnc'vdtvssh p v f v s r g k s h p l l g g q s e c - q 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 g k p g q a
r k l c - h s l q p p c l c l q g 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 3-11R.....>
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 g k p g q a
>>.....VK 3-11R.....>

ccaggctct catctatgat gcatacaaca gggccactgg catccagcc aggttcactg gcaatgggtc tggccagac ttaactctca ccatcagcag cctagagcct gaagattttg cagtttatta
s q a p h l - c i q g g h w h p s q v q w g w v w d r l h s h h g 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 q l g q t s l s p s a a - s l k l l q f i
>>.....VK 3-11R.....>
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
>>.....VK 3-11R.....>

ctgtcagcag cgagacgaat aacggttaag cggtttacca ggtttaaacg cgtattgga aggcgcgtct cattcgcca agggaccacag gtgaaatca aacgggcggc cgca SEC ID N°: 73
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 SEC ID N°: 74
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 SEC ID N°: 75
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 SEC ID N°: 76
>>.....VK 3-11R.....>
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 SEC ID N°: 75

```

Fig. 8 Cont.

VK3-15G AF

```

gaaatagtga tgaagcagtc tccagccacc ctgtctgtgt ctccaggga aagagccacc ctctcttcca gggccagtica gagtgttagc agcaacttag cctggtacca gcagaacct gggcaggtc
g n s 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 - q 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
>>.....VK 3-15G.....>

ccaggctcct catctatgtt gcattccacca gggccactgg tatccagcc aggttcagtg gcagtggttc tgggacagag ttaacttca ccaatcagag cctgcagct gaagatttg cagttatta
s q a p h l w c i h q g h w y p s g 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
>>.....VK 3-15G.....>

cgtgcagag cgagacgaat aacggttaag cgttttacca ggtttaaacg cgtattgga aggcggtct cattcgcca agggaccaag gtggaatca aaggggggc cgca
l s a a 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
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
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 3-15G.....>
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
>>.....VK 3-15G.....>

```

SEC ID N°: 77
SEC ID N°: 78
SEC ID N°: 79
SEC ID N°: 80
SEC ID N°: 79

Fig. 8 Cont.

VH3-15R AF

gaaatagtga tgaagcagtc tccagccacc ctgtctgtgt ctccagggga aagagccacc ctctcttca gggccagtca gagtgttagc agcaacttag cctgttacc gcagaaacct ggcagggtc
g n s 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 - q 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 3-15R.....>
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 gcatccaaca gggccactgg tatccacacc aggttcagtg gcagtggttc tgggacagag ttcaacttca ccatcacag cctgcagttc gaagattttg cagtttatta
s q 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 q 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 s g t e f t l t i s s l q s e d f a v y

ctgtcacag cgagacgaat aacggttaagg cagtttacc ggtttaaacg cgtattggga aggcgtctct cattoggcca agggaccaag gtgaaatca aacgggcggc cgca SEC ID N°: 81
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 SEC ID N°: 82
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 SEC ID N°: 83
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 SEC ID N°: 84
>.....VK 3-15R.....>
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 SEC ID N°: 83

Fig. 8 Cont.

VK3-20G AF

```

gaatgtgt tgagcagtc tcacggcacc ctgttttgt ctocaggga aagagccacc ctctctgca gggcagtica gagtgtagc agcagctact tagcctgta ccagcagaaa cctggccagg
gnc v d a v s r h p v f v s r g k s h p l l q g q s e c - q 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 g 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 g k p g q
>>.....VK 3-20G.....>

ctocagggt octcatctat ggtgcaccca gcaggccac tggcatocca gacaggtica gtggcagtg gtctgggaca gacttactc tcaccatcag cagactggag cctgaagatt ttgcagttta
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 g 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 t d f t l t i s r l e p e d f a v
>>.....VK 3-20G.....>

ttactgtcag cagcagacg aataacgta aggggutta ccaggittaa accgtattg ggaaggcgg totcattogg ccaaggacc aaggtggaaa tcaaggggc ggccgca SEC ID N°: 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 SEC ID N°: 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 SEC ID N°: 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 g r g r SEC ID N°: 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 SEC ID N°: 87

```

Fig. 8 Cont.

```
ttactgtcag cagcgagacg aataacggtta aggcggttta ccaggtttaa acgcgtattg ggaaggcgcg tctcatcgg ccaaggagacc aagttggaaa tcaaacggcg ggccgca
lllsaarritvrrftrfkrrvlgrrvsfgqgtkv ekraaa
yyycqqrde-r-ggglpglnaywegasshsakgprwk sngrp
itvsssetnnngkav yqv-t r i g k a r l i r p r d q g g n q t g g r
>.....VK 3-20R.....>>>
yyccqqrde-r-ggglpglnaywegasshsakgprwk sngrp
```

Fig. 8 Cont.

VL1-44 AF

cagtctgtgt tgacgagac gccctcaatgt tctgcggccc caggacagaa ggtcaccatc tctgtgtgtg gaacagctc cacattgg aataattatg tatcttggtta ccagcagctc ccaggaacag
 avc vdaa alsv c g p r t e g h h l l w k q l q h w e - l c i l v p a a p r n
 q s v l t q p p s v s a a p g q k v t i s c s g s s s n i g n n y v s w y q q l p g t
 r s l c - r s r p q c l r p q d r r s p s p a l e a a p t l g i i m y p g t s s s q e q
 >>.....VL1-51.....>
 q s v l t q p p s v s a a p g q k v t i s c s g s s s n i g n n y v s w y q q l p g t

ccccaaact cctcatttat gacaataata agcacctc agggattct gacgattct ctggtccaa gtctggacg tccgaccc tggcaccac cgaactccag actggggacg agccgatta
 s p q t p h l - 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

ttactggga acaggagacg aataacggtta aggggttta ccaggtttaa acggtattg ggaaggcgg tctgtgtt cggcgagg accaagctga ccgtctagg ggcgcgca SEC ID N°: 93
 l l r n r r i t v r r f t r f k r v l g r r v s v f g g t k l t v l g a a a SEC ID N°: 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 SEC ID N°: 95
 i t a e g 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 SEC ID N°: 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 SEC ID N°: 95

Fig. 8 Cont.

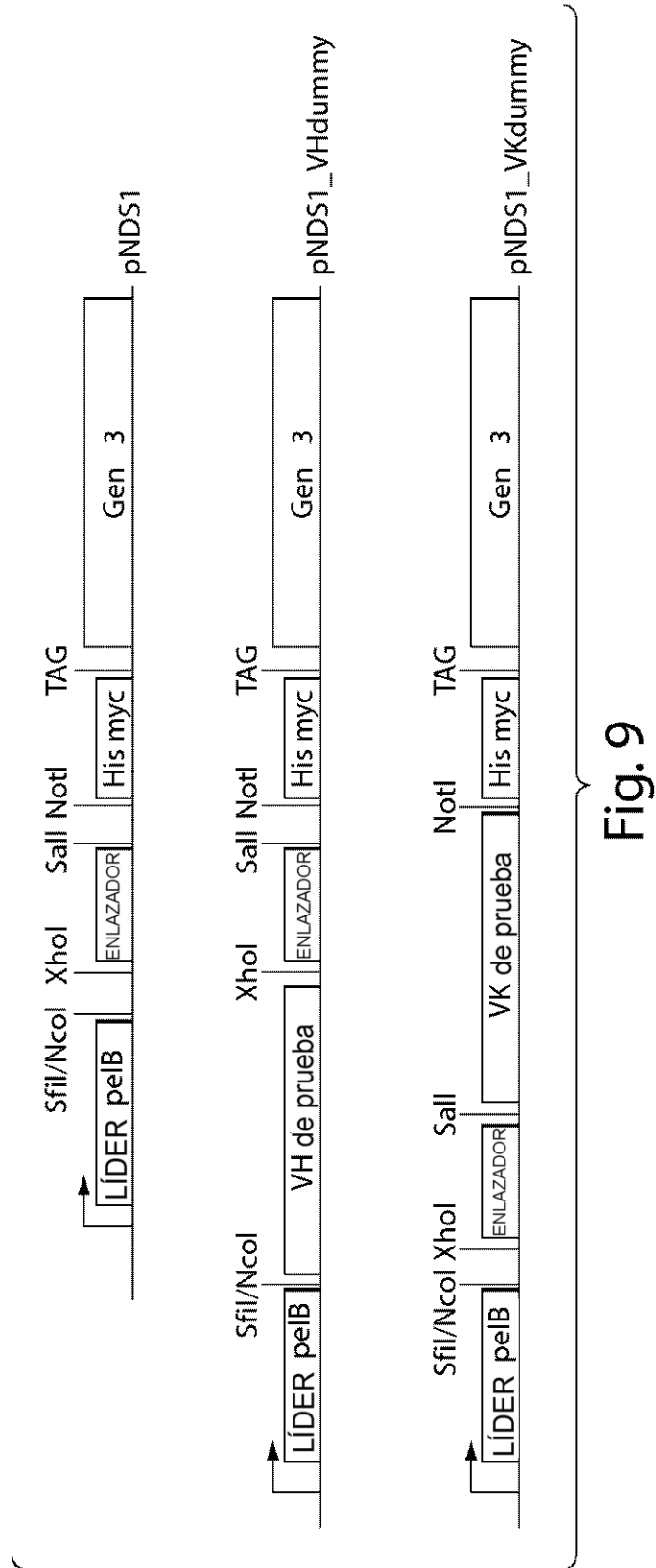
VL1-51 AF

cagttgtgc tgactcagcc acctcagcg tctgggaccc ccgggagag ggtcacctc tctgttctg gaagcagtc caacatcgga agtaatactg taaactgta ccagcagtc ccaggaacgg
a v c a d s a t l s v w d p r a e g h l l f w k q l q h i 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

ccccaaaot octatctat agtaataatc aggggcctc aggggtccot gaccgattct ctggctccaa gtctggcacc tcagctccc tgggcatcag tggtctcag tctgaggtg 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 - 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

ttactgtga gcaggagacg aataacgta aggggttta ccagggtta accggtattg ggaaggcgg tctctgtgtt cggcgaggg accaagctga ccgtctctagg ggcggccgca	SEC 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	SEC 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	SEC 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	SEC ID No: 100
>.....VL1-44.....>>	
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	SEC ID No: 99

Fig. 8 Cont.



MUESTRA	FW3	H3	H3/FW4	LONGITUD DE CDR3	
1a	CAK	AFRPNWGSRVLYF	DYW	15	SEC ID N° : 101
1b	CAK	LNWGWFF	DPW	8	SEC ID N° : 102
1c	CAK	GWEGEY	DYW	8	SEC ID N° : 103
1d	CAK	DLWLDSSSNWF	DPW	13	SEC ID N° : 104
1e	CAK	DLPG	DPHW	7	SEC ID N° : 105
1f	CAK	APPTGYF	DYW	9	SEC ID N° : 106
1g	CAK	VGLTGCVHF	DYW	10	SEC ID N° : 107
1h	CAK	DSYGSGSYYN	DYW	12	SEC ID N° : 108
2a	CAK	VAGTWGRVAYYF	DYW	14	SEC ID N° : 109
2b	CAK	VTGVFVGNF	DYW	11	SEC ID N° : 110
2c	CAK	DARNSGSYF	DYW	11	SEC ID N° : 111
2d	CAK	GVRTGVVF	DYW	10	SEC ID N° : 112
2e	CAK	GVVNWGTRRKGVF	DPW	15	SEC ID N° : 113
2f	CAK	WGGWF	DPW	7	SEC ID N° : 114
2g	CAK	RRDNWGSV	DW	9	SEC ID N° : 115
2h	CAK	GSGFSSGVF	DSW	11	SEC ID N° : 116

Fig. 10

Fig. 11

BsmBI BsmBI
 1 acgtctccga gannsnnsnn snnsatggat tattggggga gacg SEC ID N°: 117
 t s p r ? ? ? ? m d y w g r SEC ID N°: 118
 r l r ? ? ? ? ? w i i g g d SEC ID N°: 119
 v s e ? ? ? ? ? g l l g e t SEC ID N°: 280

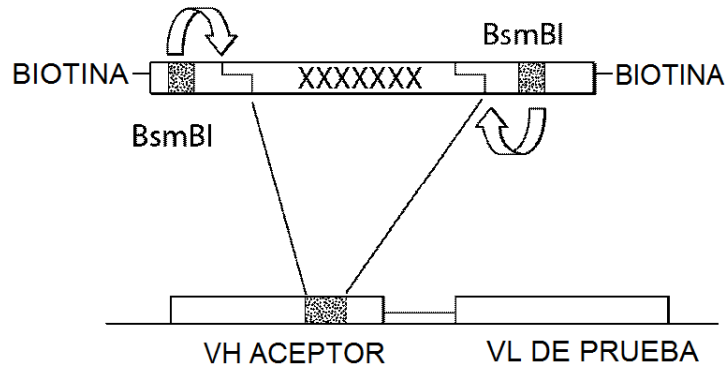
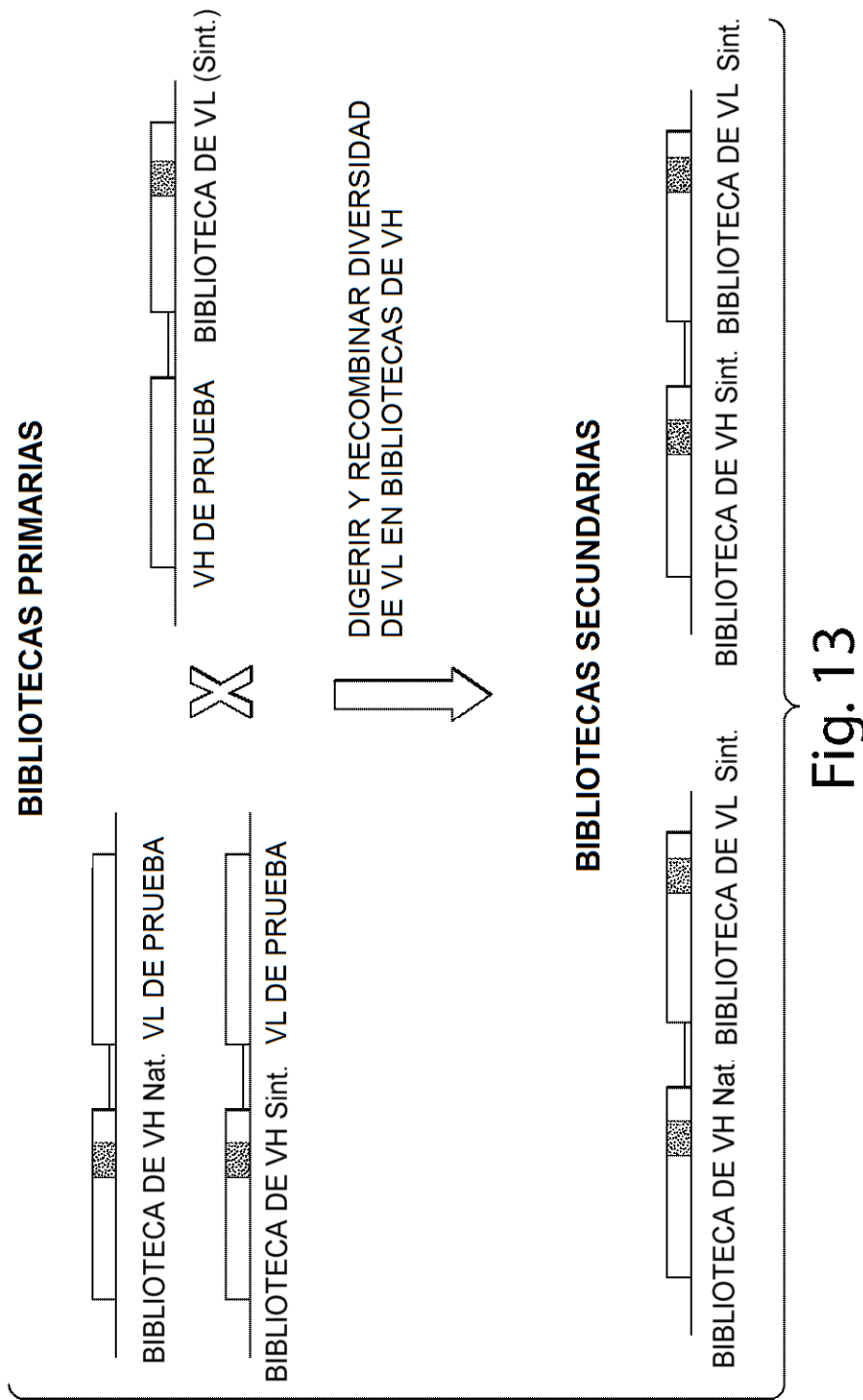


Fig. 12



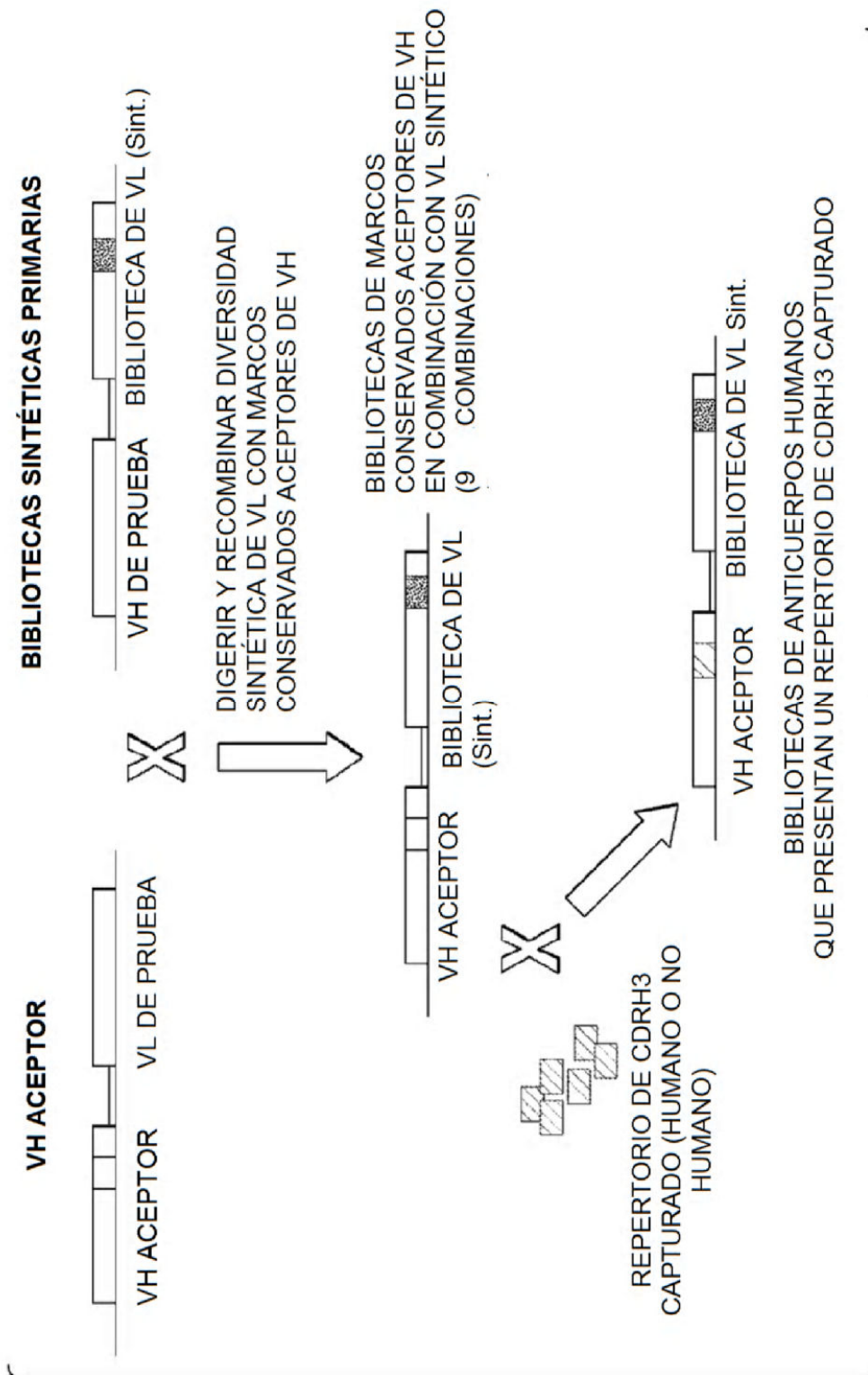


Fig. 14

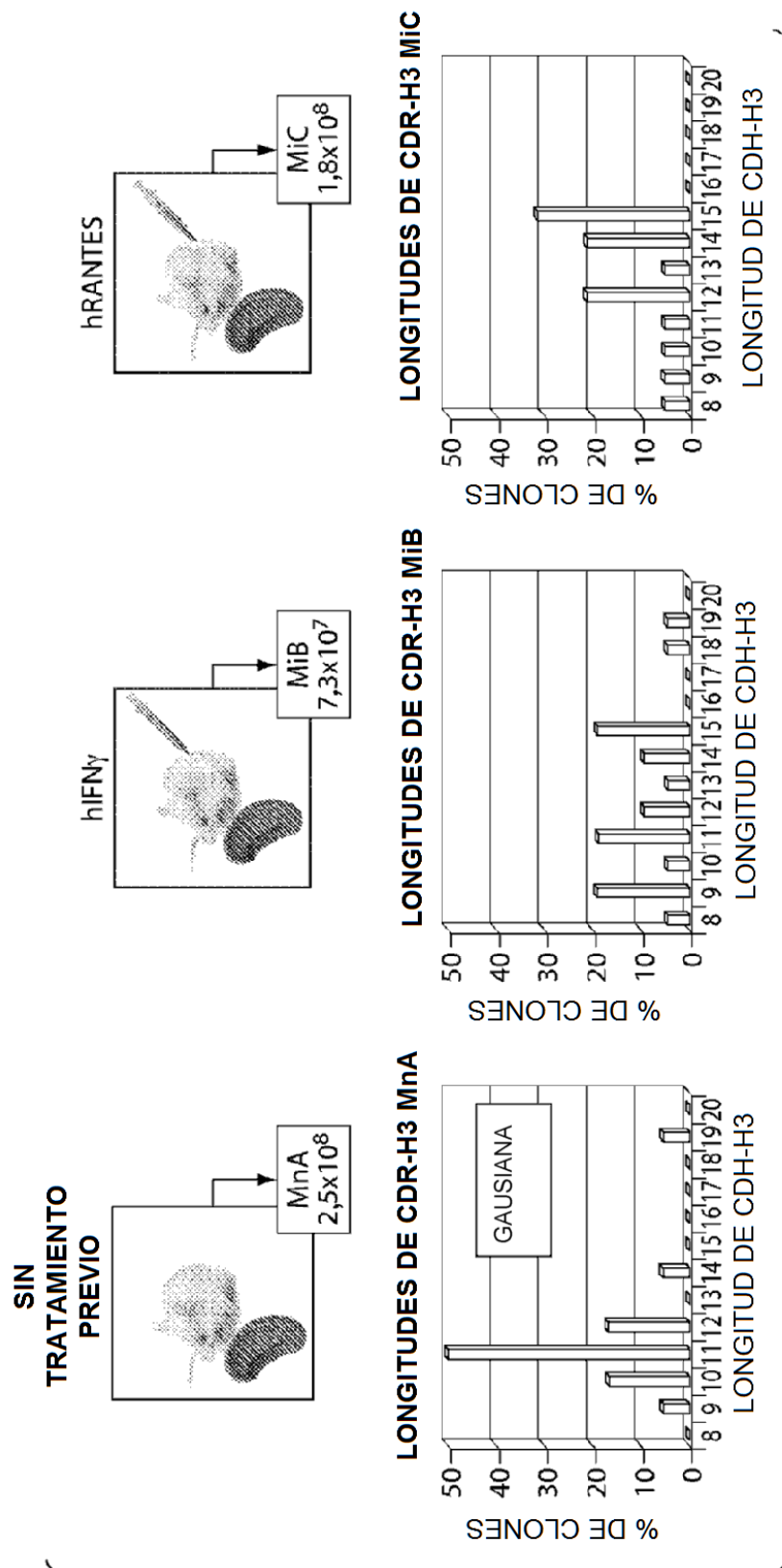


Fig. 15

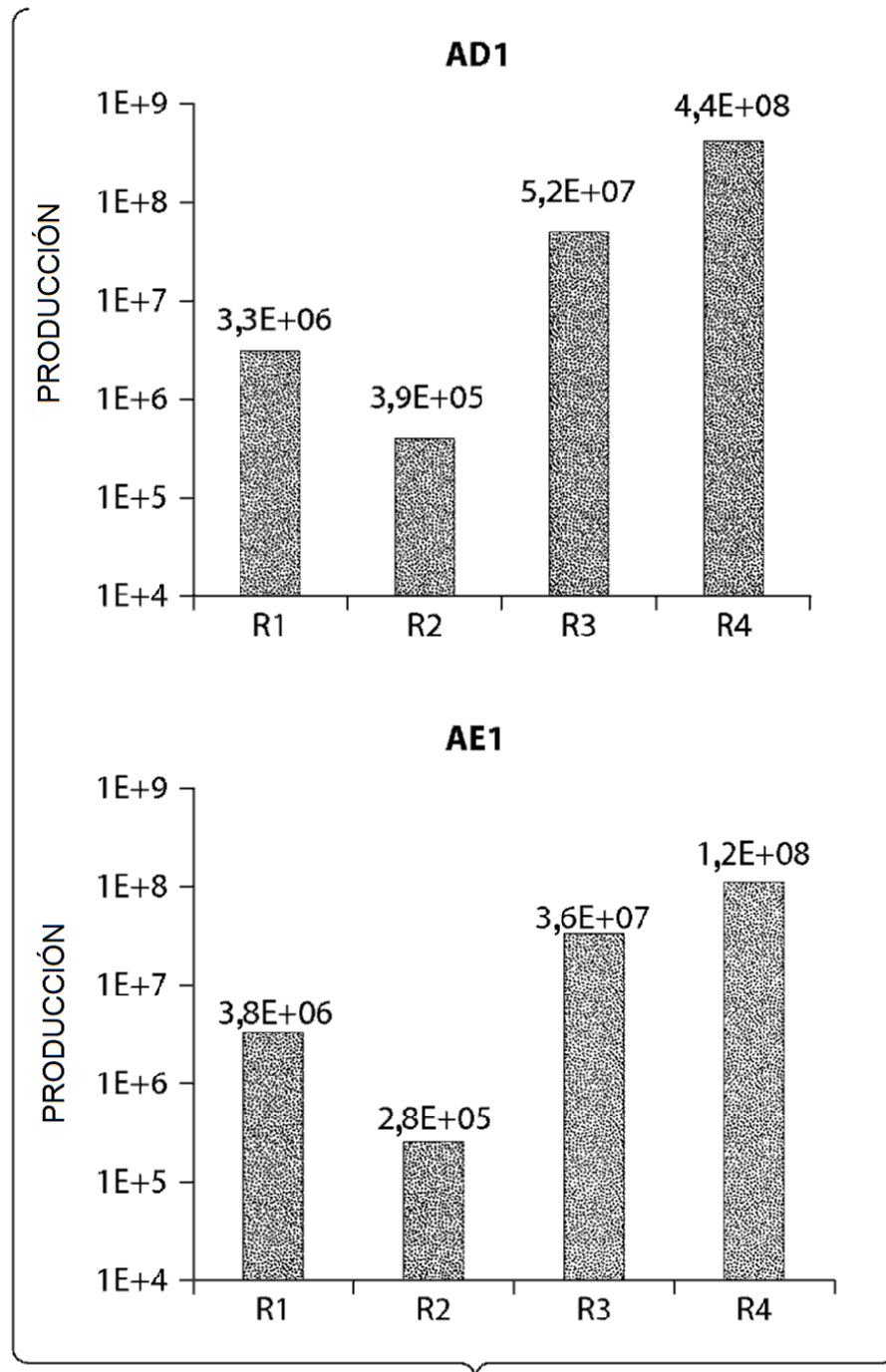


Fig. 16

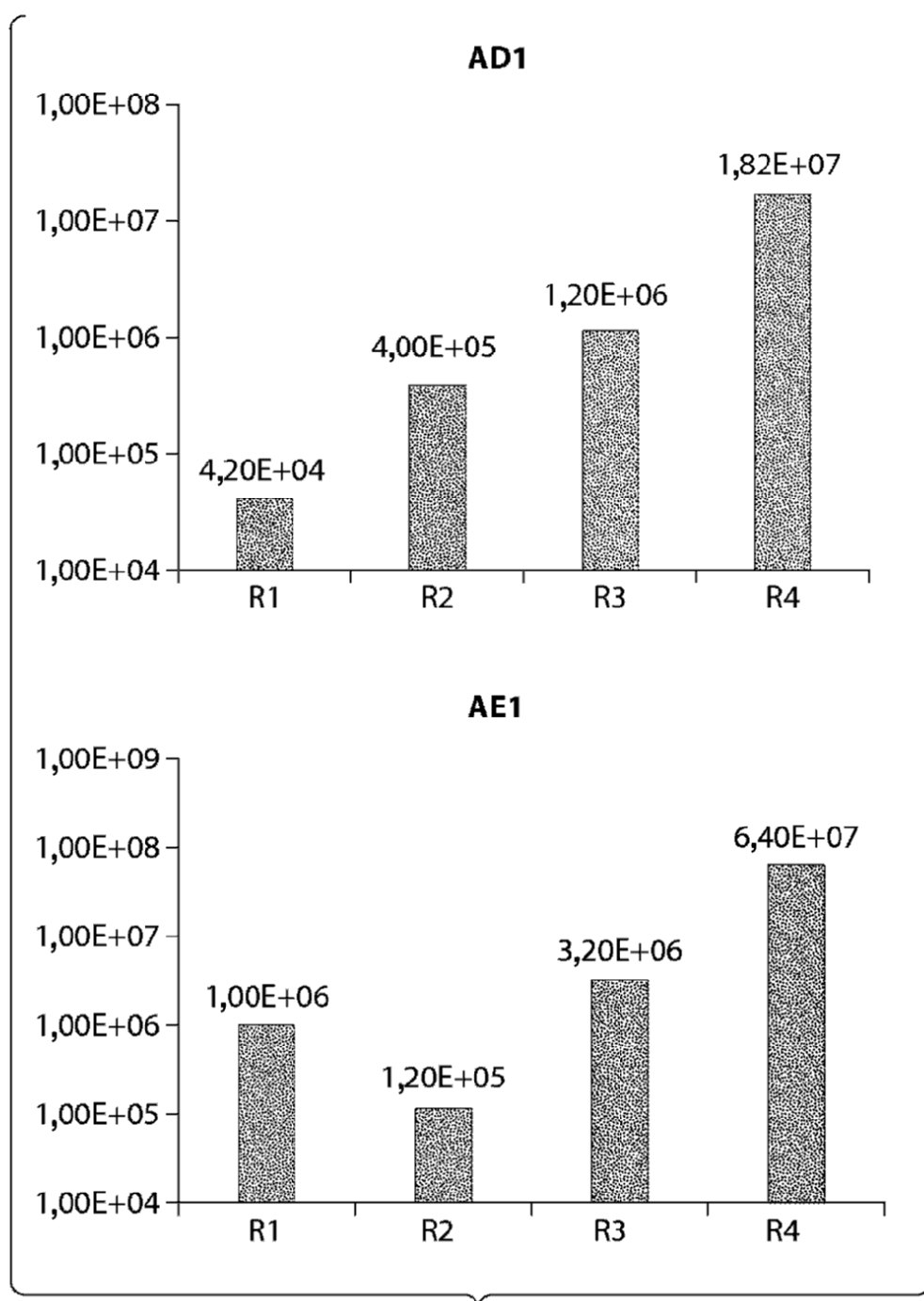


Fig. 17

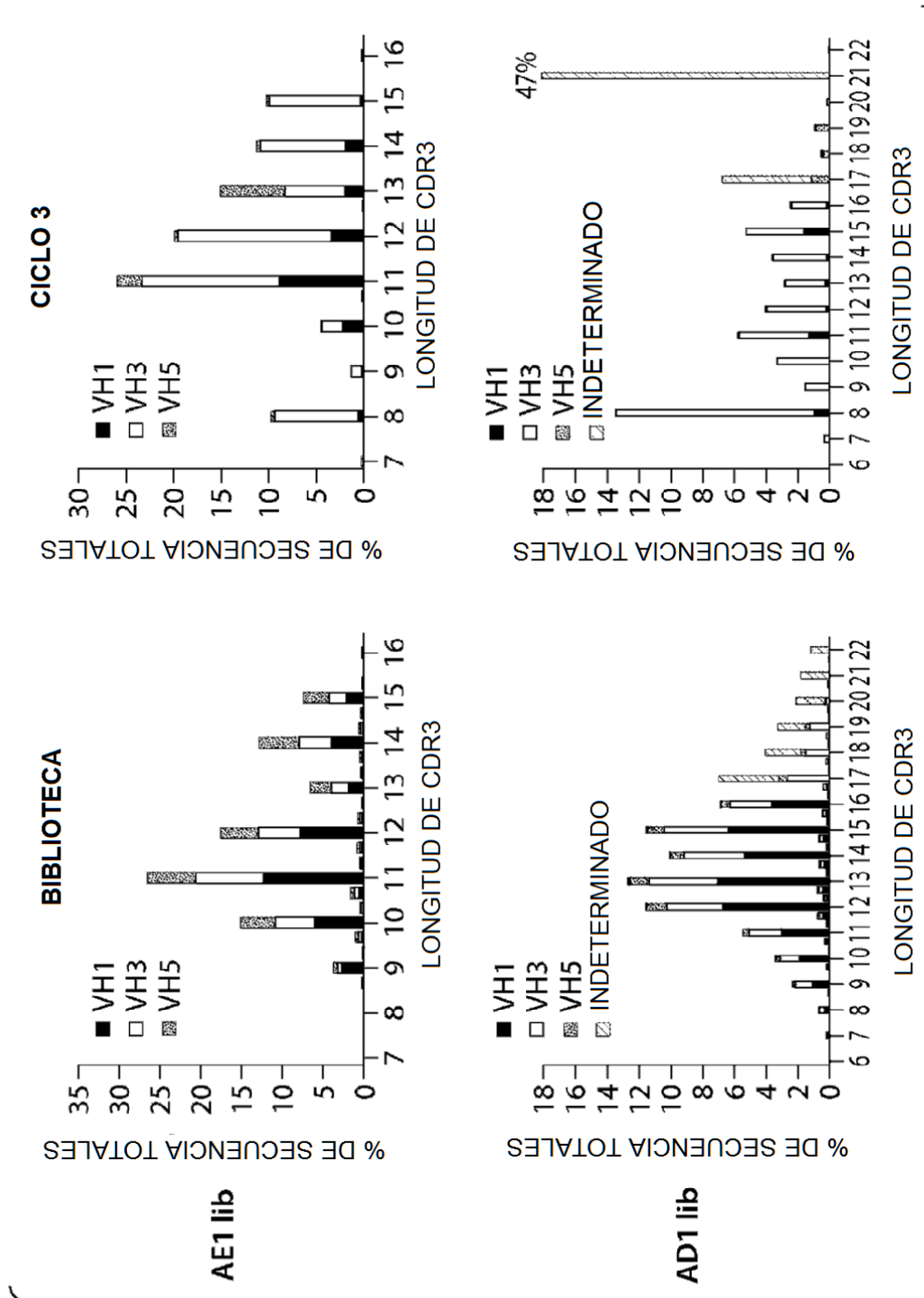


Fig. 18

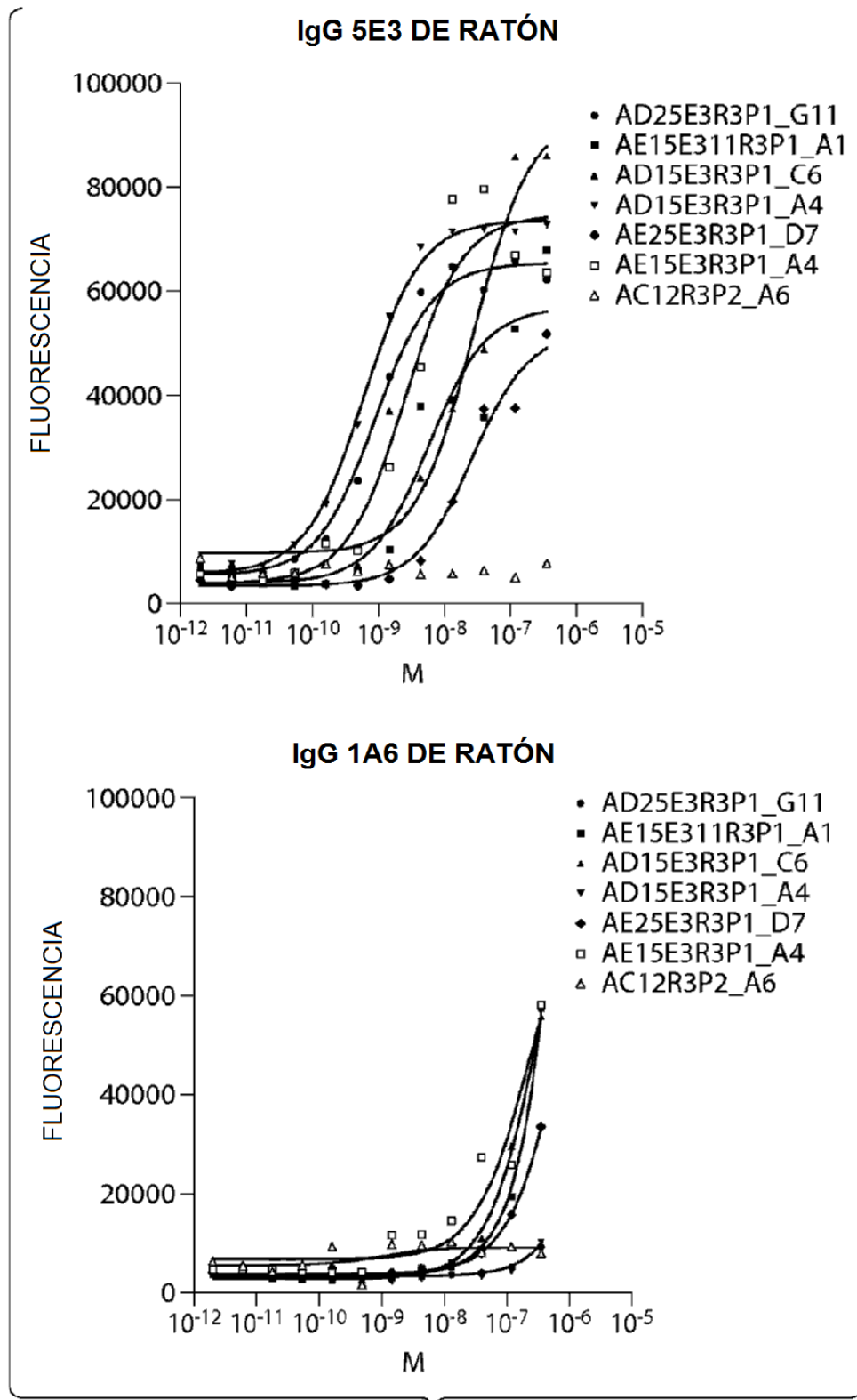
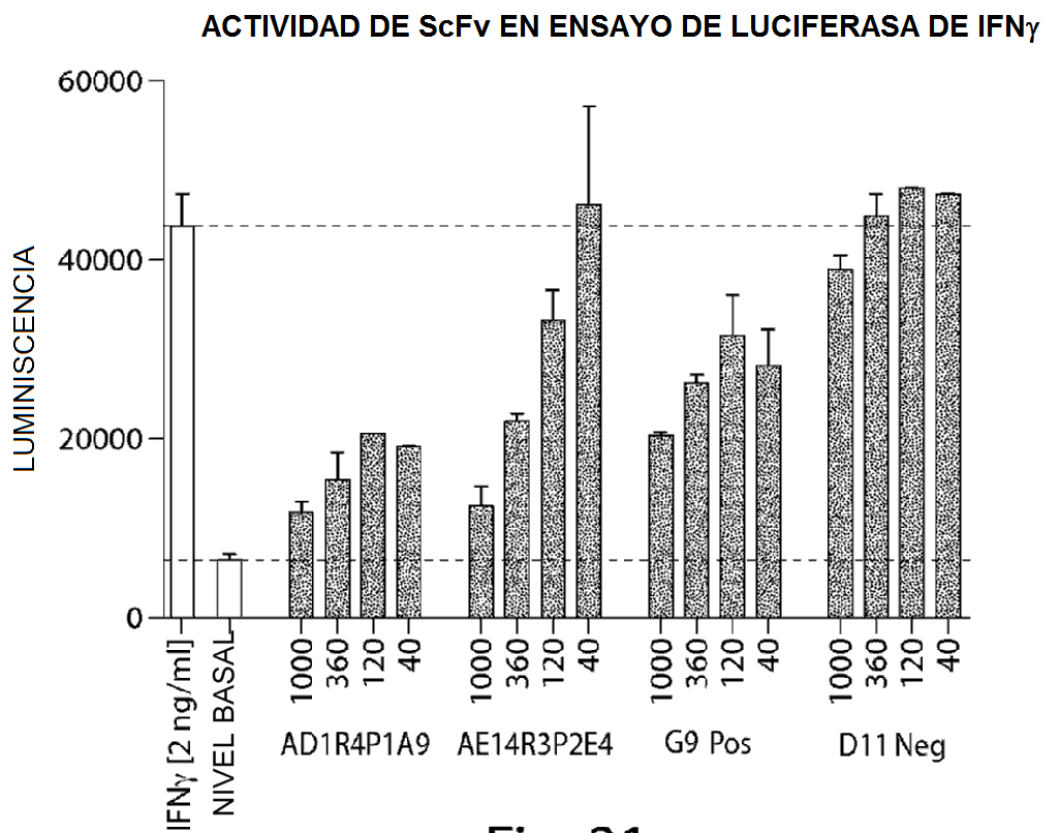
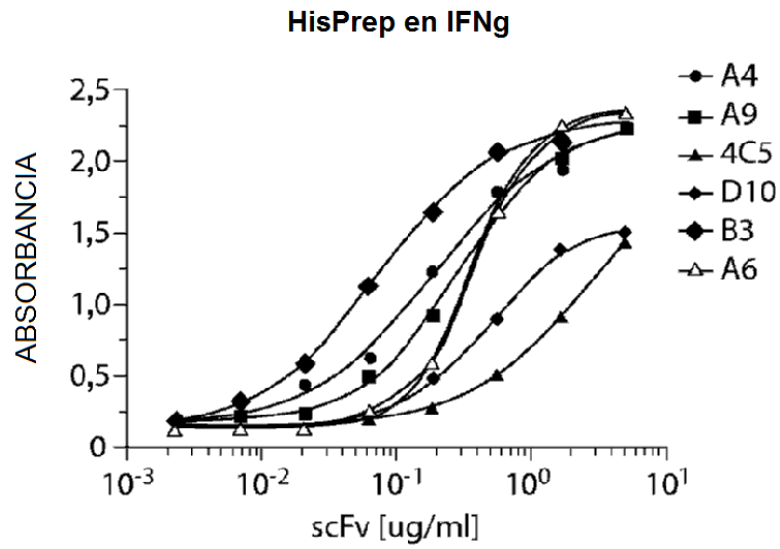


Fig. 19



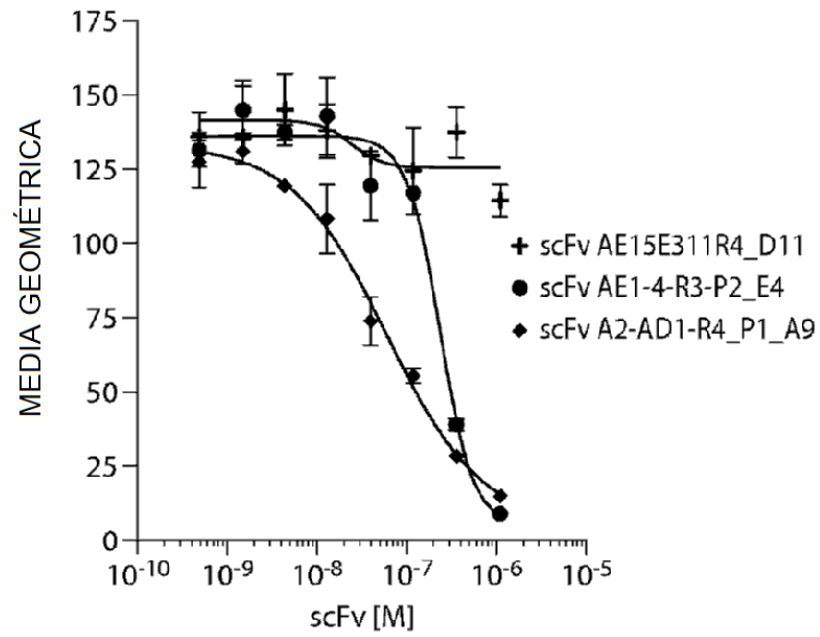


Fig. 22

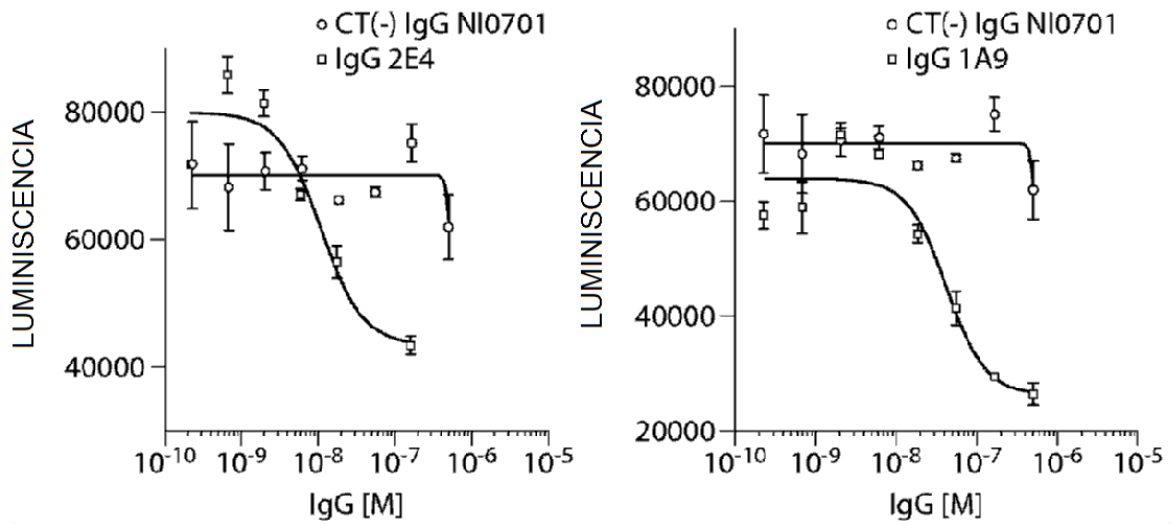


Fig. 23

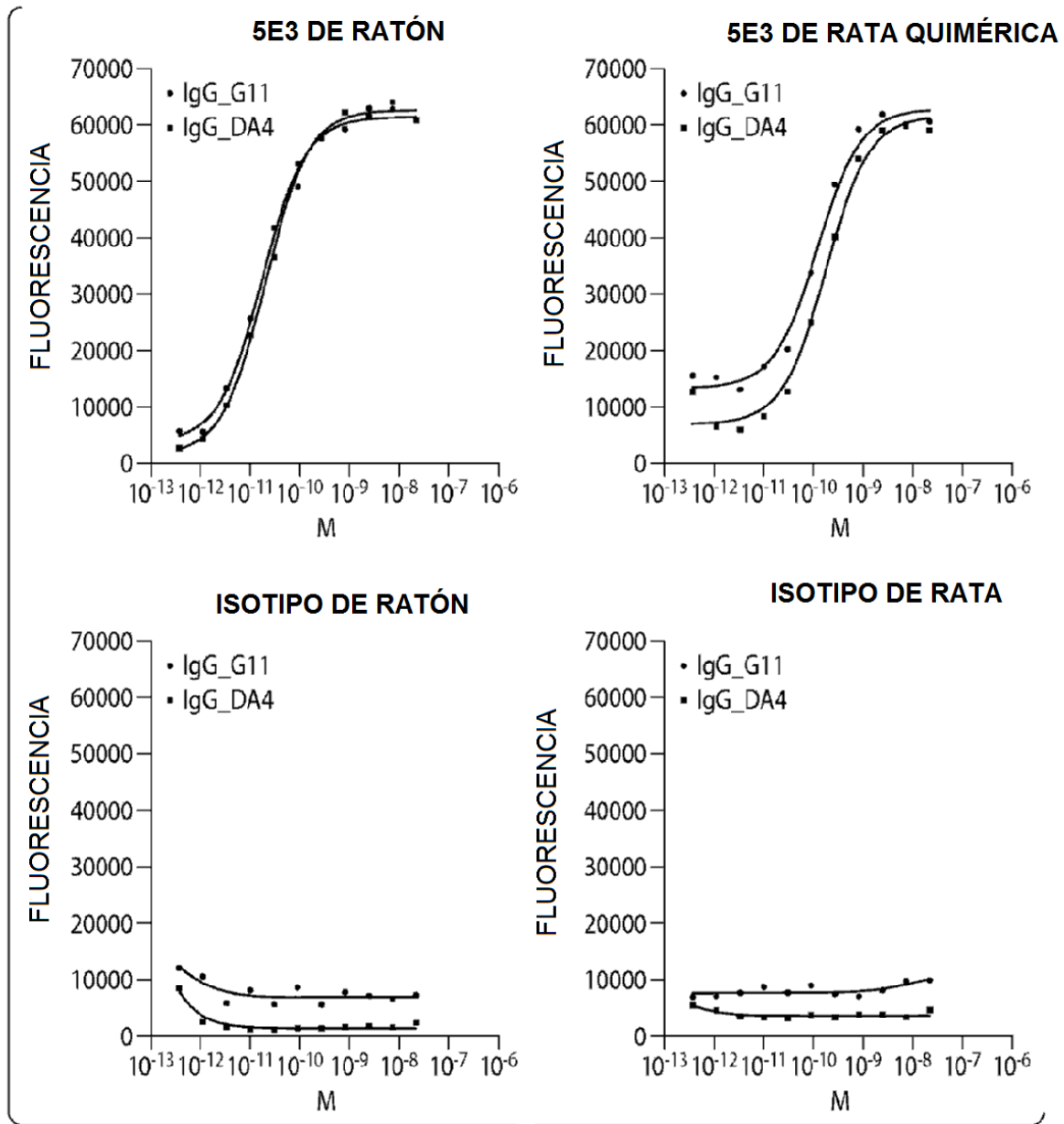


Fig. 24

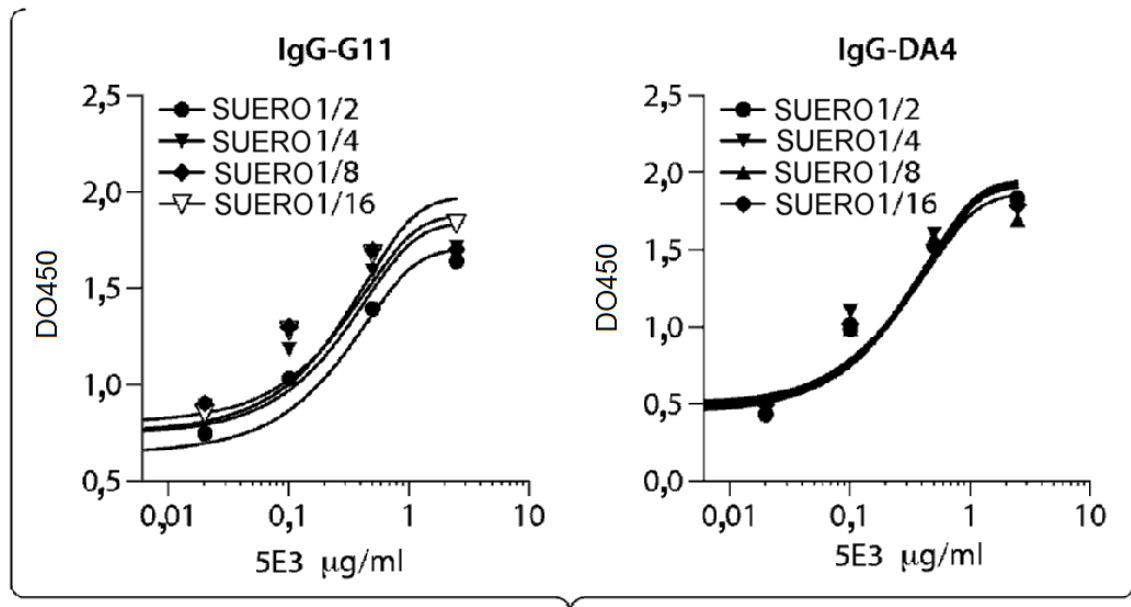


Fig. 25

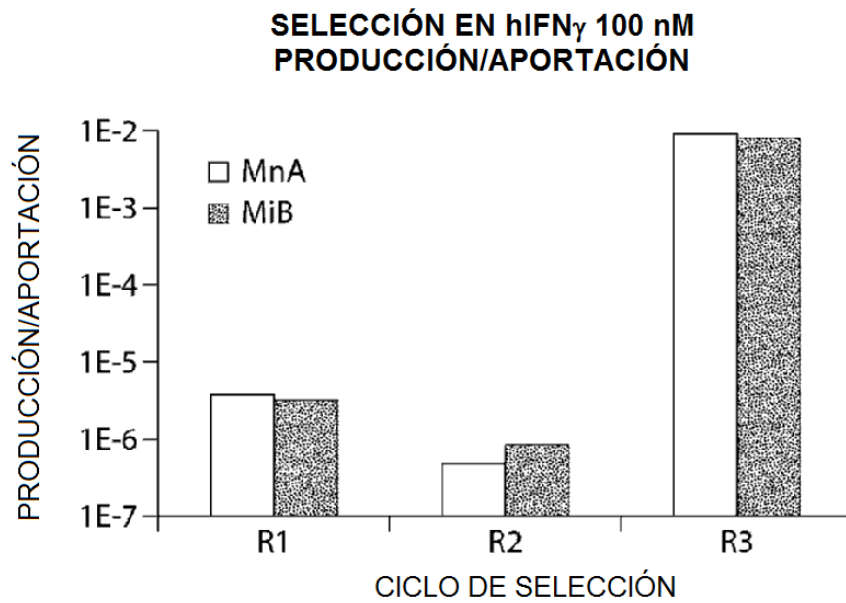


Fig. 26

TASA DE ACIERTOS FRENTE A hIFN γ

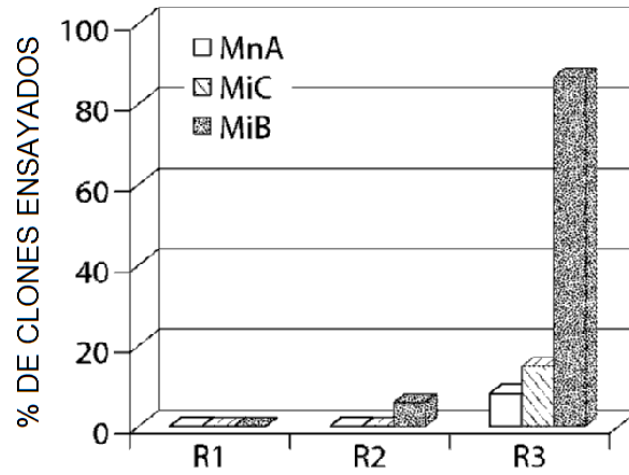


Fig. 27

% DE CLONES POSITIVOS PARA DIVERSOS INTERVALOS DE FLUORESCENCIA

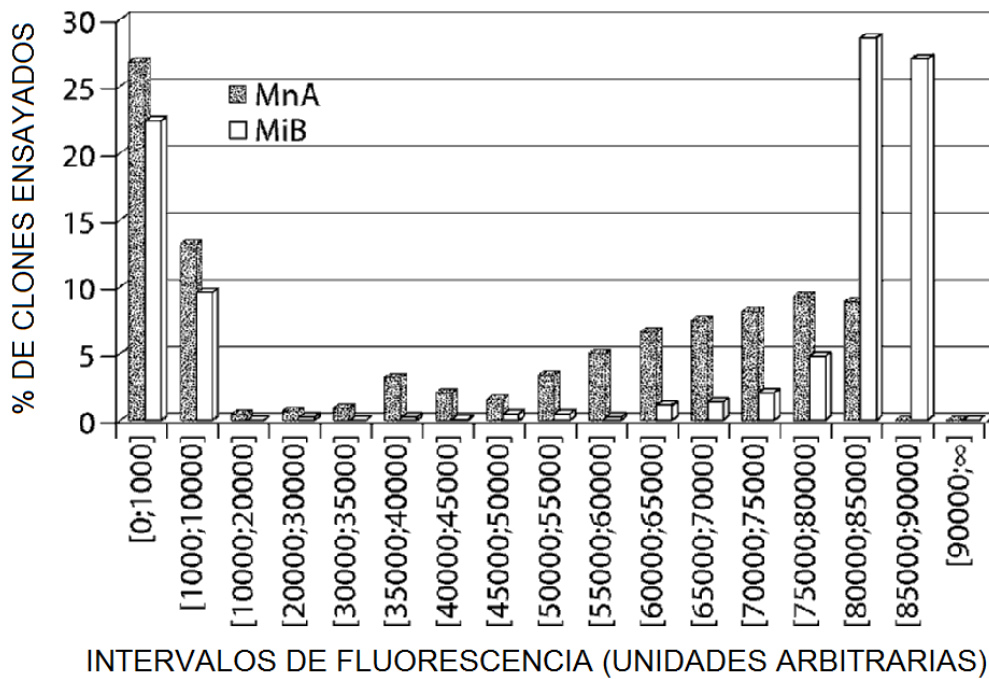


Fig. 28

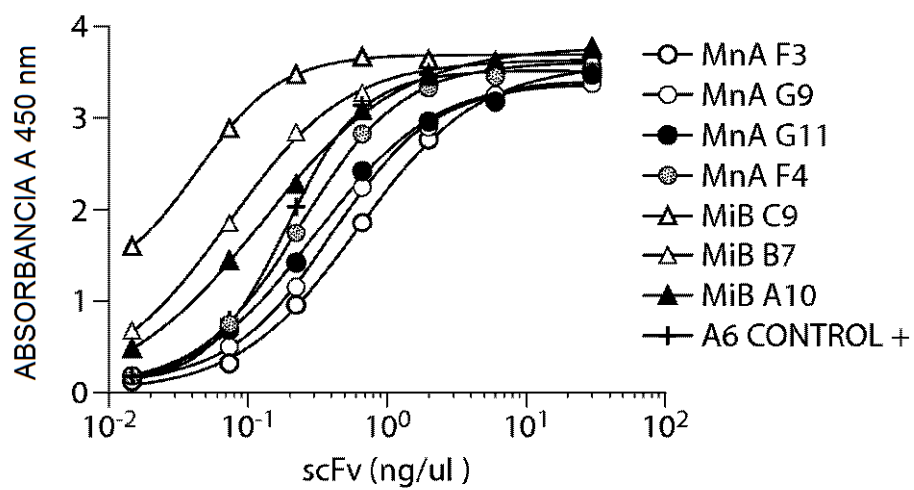


Fig. 29

Figura 30

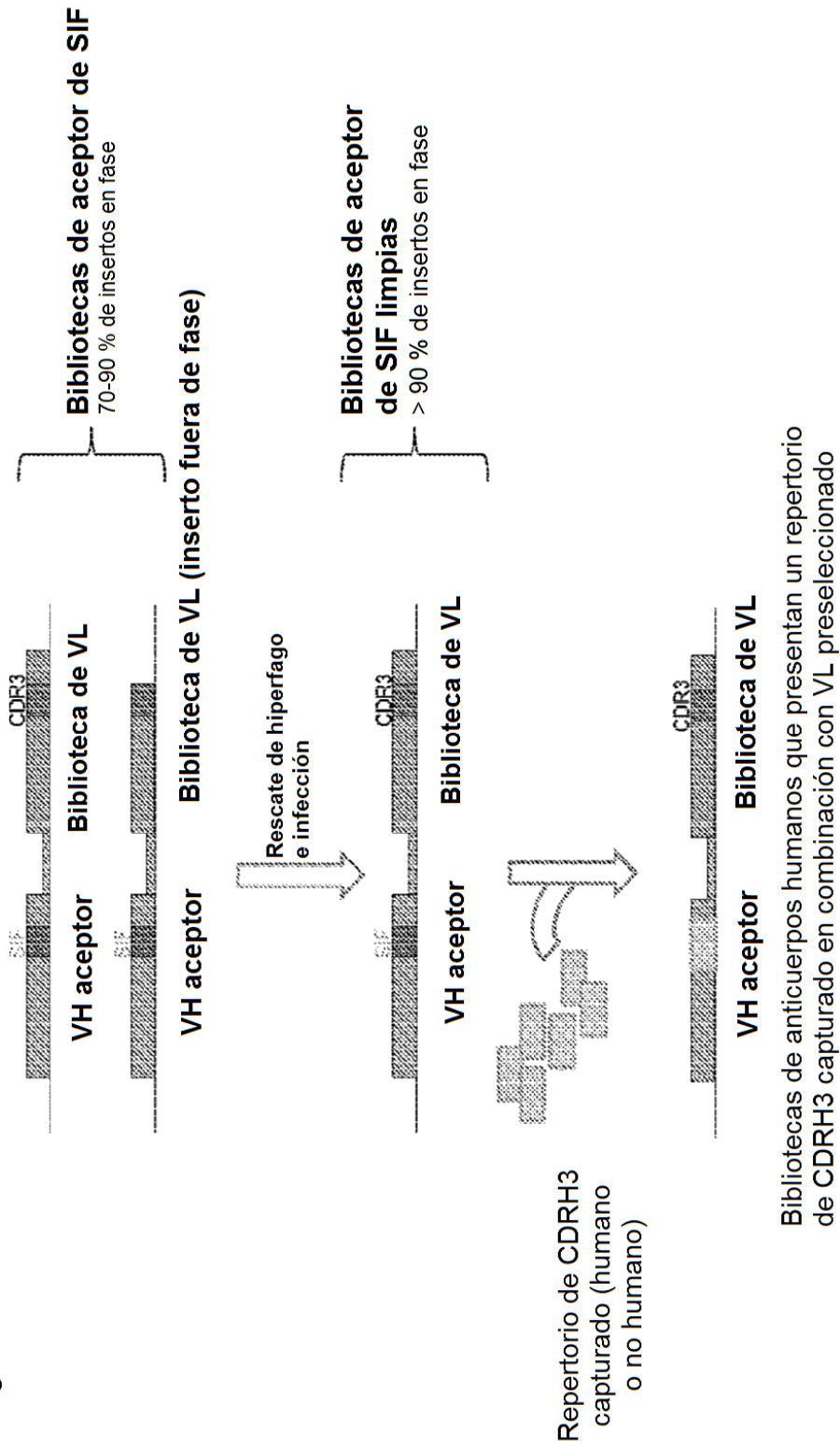


Figura 31

SIF1 sec ID: 290	ATTACTGTGTC	GAGAGGAGAC	GNHNHCSTCT	CTTGGGCCCCA	GCGAAC	SEC ID Nº: 290
	TAATGACACG	CTCTCCCTCTG	GNHNHCACGA	GAGCCCCGCT	CCCTTG	SEC ID Nº: 294
	y c a z q d ? ? y ? ? q q q					SEC ID Nº: 295
-----ErmB1-----						
SIF2 sec ID: 291	ATTACTGTGTC	GAGAGGAGAC	GNCGTCTCTT	GCGGCCAGGG	AACCCCT	SEC ID Nº: 291
	TAATGACACG	CTCTCCCTCTG	GNCGACAGAA	CCCCGATCCC	TTGGGA	SEC ID Nº: 296
	y c a z q d ? ? y ? ? q q q					SEC ID Nº: 297
-----ErmB1-----						

Figura 32

SEC.IDNº:	IF1b 2
298	
299	
300	
	
SEC.IDNº:	IF1b 6
301	
302	
303	
	
SEC.IDNº:	IF1b 10
304	
305	
306	
	
SEC.IDNº:	IF1b 12
307	
308	
309	
	
SEC.IDNº:	IF1b 18
310	
311	
312	
	
SEC.IDNº:	IF1b 20
313	
314	
315	
	
SEC.IDNº:	IF1b 26
316	
317	
318	
	
SEC.IDNº:	IF1b 29
319	
320	
321	

SEC.IDNº:	IF1b 31
322	
323	
324	
	
SEC.IDNº:	IF1b 36
325	
326	
327	
	
SEC.IDNº:	IF1b 38
328	
329	
330	
	
SEC.IDNº:	IF1b 41
331	
332	
333	
	
SEC.IDNº:	IF1b 43
334	
335	
336	
	
SEC.IDNº:	IF1b 47
337	
338	
339	
	
SEC.IDNº:	IF1b 48
340	
341	
342	

