



OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(97) Fecha y número de publicación de la concesión europea:



11 Número de publicación: 2 542 744

51 Int. Cl.:

C12N 15/10 (2006.01)

18.03.2015

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

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

(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

EP 2513312

(72) Inventor/es:

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

(74) Agente/Representante:

SUGRAÑES MOLINÉ, Pedro

DESCRIPCIÓN

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

5 Campo de la invención

10

15

40

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 20 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 25 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 parátopos 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 30 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 35

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 Nº 5.969.108 de McCaffety).

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.

10

20

25

30

35

40

45

50

55

60

65

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 deleciones 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 limitadas 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

10

15

20

25

30

35

40

45

50

55

60

65

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

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.

5

10

15

40

45

50

55

60

65

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 liquen 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. Un 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 IIs. 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 liquen 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 secuenciado 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.

5 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 10 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 15 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.

20

25

30

35

40

45

50

55

60

65

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.

15

5

10

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.

20 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 25 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. 30 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.

35 En algunas realizaciones, el cebador es un ácido nucleico seleccionado del grupo que consiste en SEC ID №: 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.

60

65

55

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 Fokl IIs. En algunas realizaciones, la especie no humana es un primate no humano, roedor, canino, felino, oveia, 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 Pokl, sitios de reconocimiento de Bsal 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. En algunas realizaciones, la 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 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. 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.
- 60 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.

65

10

15

20

25

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 incluyen la etapa adicional de (e) transformar el vector de expresión de la etapa (d) en una célula hospedadora y cultivar la célula hospedador 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 15 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 20 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 25 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 30 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 35 hospedadora con un antígeno diana; y (g) determinar 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 Fokl, sitios de reconocimiento de Bsal 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. 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 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. Por ejemplo, la secuencia de gen variable de cadena ligera lambda humana es una secuencia de gen variable de cadena ligera lambda humana.

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.
- Opcionalmente, la pluralidad de ácidos nucleicos diversificados incluye o deriva de secuencias seleccionadas de secuencias de CDR3 de origen natural, secuencias de lg de origen natural de seres humanos, secuencias de lg 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, 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.

5

15

20

25

30

35

40

55

60

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 Fokl, sitios de reconocimiento de Bsal 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. Opcionalmente, la 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 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. 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.

5

15

20

25

30

35

40

45

50

55

60

65

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 Bsal 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 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. 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.

5

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.

10

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.

15

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.

20

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.

25

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.

30

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.

35

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 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 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).

45

50

40

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.

55

60

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 Fokl, sitios de reconocimiento de Bsal y/o sitios de reconocimiento de Bsmbl.

65

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, 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.

- 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.
- 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.
- 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.
- 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).
- 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.

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 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 Ils, 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 Ils 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 Ils; 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 Ils de las etapas (a) y (b).

65

50

55

60

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 Fokl, sitios de reconocimiento de Bsal 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 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. Por ejemplo, la secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la secuencia de gen variable de cadena ligera lambda humana.
- 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.

25

35

60

- 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.
- 40 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 45 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 50 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 55 de genes de inmunoglobulina completa que no tienen los sitios de reconocimiento de enzimas de restricción de Tipo Ils 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 Fokl, sitios de reconocimiento de Bsal y/o sitios de reconocimiento de Bsmbl.

5

10

15

20

25

30

35

40

45

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. Opcionalmente, la 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 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. 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.

50

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 lg 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

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 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.

10

15

5

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 Fokl, sitios de reconocimiento de Bsal y/o sitios de reconocimiento de Bsmbl.

20

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 lambda humana o una secuencia de gen variable de cadena ligera lambda humana o una secuencia de gen variable de cadena ligera lambda humana. Por ejemplo, la 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.

30

25

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.

35

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.

00

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.

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.

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

45

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 lg 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.

50

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.

60

55

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.

65

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.

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.

15

20

25

30

35

40

45

50

55

60

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.

65 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 Fokl. 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 Fokl. 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 Fokl.

5

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.

10

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).

15

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.

20

25

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.

30

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 hospedador 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*.

35

40

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.

45

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.

50

55

60

65

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 secuencia de ácido nucleico de relleno que comprende al menos dos sitios de reconocimiento de enzimas de restricción de Tipo Ils 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 Ils 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 Ils que se une con el sitio de reconocimiento de enzimas de restricción de Tipo Ils 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 Ils que se une con el sitio de reconocimiento de enzimas de restricción de Tipo Ils que se une con el sitio de reconocimiento de enzimas de restricción de Tipo Ils que se une con el sitio de reconocimiento de enzimas de restricción de Tipo Ils que se une con el sitio de reconocimiento de enzimas de restricción de Tipo Ils

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.

10

15

20

25

30

55

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 Fokl. 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 Fokl. 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 Fokl.

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 ácido nucleico diversificadas que codifican 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, VK3-39, 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

5

10

55

60

65

- 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 seleccionado 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 BsmBl 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 Vλ. 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.
- 40 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 hIFNγ o hCCL5/RANTES como una fuente de diversidad. El tamaño de las bibliotecas se indica en los paneles superiores. Los panales 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 hIFNγ 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 hIFNγ 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 hIFNγ y en comparación con un scFv positivo específico para hIFNγ (A6).
 - La Figura 21 es una gráfica que representa el efecto inhibidor de preparaciones de scFv purificadas en un ensayo de gen indicador de luciferasa conducido por hIFNy. La actividad neutralizante de dos candidatos a scFv (AD1R4P1A9 y AE14R3P2E4) se comparó con la activad de un scFv de control positivo (G9) y un scFv de

control negativo (D11).

10

20

50

55

60

65

La Figura 22 es una gráfica que representa el efecto inhibidor de preparaciones de scFv purificadas en un ensayo de inducción de MHCII en respuesta a hIFNγ. La actividad neutralizante de dos candidatos a scFv (AD1R4P1A9 y AE14R3P2E4) se comparó con la activad de un scFv de control negativo (D11).

- 5 La Figura 23 es una serie de gráficas que representan el efecto inhibidor de los dos candidatos AD1R4P1A9 y AE14R3P2E4 reformateados en IgG en un ensayo de gen indicador de luciferasa conducido por hIFNγ. 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 hIFNy 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 hIFNγ. 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 hIFNγ 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.
- 30 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 10¹³ a 10¹⁴ 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 epítopos 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.

5

10

15

20

25

30

35

40

45

50

55

60

65

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 200302323333, 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 Fokl. 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 Ils 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 BsmBI, 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 lg de origen natural de seres humanos, secuencias de lg 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.

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.

65

5

10

15

20

25

30

35

40

45

50

55

60

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 10 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 15 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 20 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 25 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 30 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 sitos de reconocimiento de enzimas de restricción de Tipo Ils 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 35 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.
- 40 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 Fokl 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.

65

60

45

5

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 lg 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.

5

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.

10

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.

15

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.

20

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 material. Además, a no ser que se requiera de otro modo por el contexto, los términos singulares incluirán plurales y los términos plurales 25 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 30 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

35

40

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:

farmacéutica, formulación y suministro, y tratamiento de pacientes.

45

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 policionales, monocionales, quiméricos dAb (anticuerpo de dominio), monocatenarios, Fab, Fab y F(ab)2, scFv y una biblioteca de expresión de Fab.

50

55

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.

60

65

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 inmunorreacionar con un epítopo particular del antígeno caracterizado por una única afinidad de unión por

el.

10

15

20

25

30

35

40

45

50

65

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 "epítopo" incluye cualquier determinante proteico capaz de unirse específicamente con una inmunoglobulina, un scFv o un receptor de linfocitos T. El término "epítopo" 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 μ M; por ejemplo, \leq 100 nM, preferentemente \leq 1 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 (Kd) de la interacción, en la que una Kd 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" (Kon) como la "constante de velocidad de disociación" (Koff) 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 Kd. (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 $\leq 1~\mu M$, por ejemplo, $\leq 100~n M$, preferentemente $\leq 10~n M$, y más preferentemente ≤ 1 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.

5

10

15

20

35

40

55

60

65

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-trimetillisina, ε-N-acetillisina, O-fosfoserina, N-acetilserina, N-formilmetionina, 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.

10

15

20

25

30

35

40

45

55

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 biotinilo 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, 14C, ¹⁵N, ³⁵S, ⁹⁰Y, ⁹⁹Tc, ¹¹¹In, ¹²⁵I, ¹³¹I), 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 biotinilo, epítopos 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 epitópicos). 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 epítopo del mismo, puede inmovilizarse en una columna para purificar el anticuerpo específico inmunitario por cromatografía de inmunoafinidad. Se analiza la purificación de inmunoglobulinas, por ejemplo, en D. Wilkinson (The Scientist, publicado en The Scientist, Inc., Filadelfía 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.

Eiemplos

5

10

30

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

15 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 20 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 25 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

Localiza	dores genomicos	
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 TTATTTCCAATTTCAGATACCACCG	(SEC ID Nº: 124)
3	K3-11 TTGTTGGGGTTTTTGTTTCATGTGG	(SEC ID Nº: 125)
5	K3-15 TATTTCCAATTTCAGATACCACTGG	(SEC ID Nº: 126)
3	K3-15 ATGTTGAATCACTGTGGGAGGCCAG	(SEC ID Nº: 127)
5	K3-20 TTATTTCCAATCTCAGATACCACCG	(SEC ID Nº: 128)
3	K3-20 TTTTGTTTCAAGCTGAATCACTGTG	(SEC ID Nº: 129)
5	L1-44 ATGTCTGTGTCTCTCACTTCCAG	(SEC ID Nº: 130)
3	L1-44 TTCCCCATTGGCCTGGAGCACTGTG	(SEC ID Nº: 131)
5	L1-51 GTGTCTGTGTCTCCTGCTTCCAG	(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 CTGCCTCCCTGGGGTTTCTGAA	(SEC ID Nº: 137)
5	H1-69 CCCCTGTGTCCTCTCCACAGGTGTC	(SEC ID Nº: 138)
3	H1-69 CCGGCACAGCTGCCTTCTCCCTCAG	(SEC ID Nº: 139)
5	DP-47 GAGGTGCAGCTGTTGGAG	(SEC ID Nº: 140)
5	H3-23 TCTGACCAGGGTTTCTTTTTGTTTGC	(SEC ID Nº: 141)
3	H3-23 TTGTGTCTGGGCTCACAATGACTTC	(SEC ID Nº: 142)
5	H3-30 TGGCATTTTCTGATAACGGTGTCC	(SEC ID Nº: 143)
3	H3-30 CTGCAGGGAGGTTTGTGTCTGGGCG	(SEC ID Nº: 144)

Localizadores genómicos

5	H3-48 ATATGTGTGGCAGTTTCTGACCTTG	(SEC ID Nº: 145)
3	H3-48 GGTTTGTGTCTGGTGTCACACTGAC	(SEC ID Nº: 146)
5	H5-a GAGTCTGTGCCGGAAGTGCAGCTGG	(SEC ID Nº: 147)

Específico para secuencia codificante

5	VH1 TATCAGGTGCAGCTGGTGCAG	(SEC ID Nº: 148)
5	VH3 TATCAGGTGCAGCTGGTGGAG	(SEC ID Nº: 149)
5	VH5 TATGAGGTGCAGCTGGTGCAG	(SEC ID Nº: 150)
3	VH1/3 ATATCTCTCGCACAGTAATACAC	(SEC ID Nº: 151)
3	VH3 ATATCTCTCGCACAGTAATATAC	(SEC ID Nº: 152)
3	VH5 ATATGTCTCGCACAGTAATACAT	(SEC ID Nº: 153)
5	VK1 TATGACATCCAGATGACCCAGTCTCCATCCTC	(SEC ID Nº: 154)
3	DPK9 ATAGGAGGGTACTGTAACT	(SEC ID Nº: 155)
3	DPK1 ATAGGAGGGAGATTATCATA	(SEC ID Nº: 156)
5	DPK22_L6 TATGAAATTGTGTTGACGCAGTCT	(SEC ID Nº: 157)
3	DPK22 ATAGGAGGTGAGCTACCATACTG	(SEC ID Nº: 158)
5	DPK21 TATGAAATAGTGATGACGCAGTCT	(SEC ID Nº: 159)
3	DPK21 ATAGGAGGCCAGTTATTATACTG	(SEC ID Nº: 160)
3	L6 CAGCGTAGCAACTGGCCTCCTAT	(SEC ID Nº: 161)
5	DPL2 TACAGTCTGTGCTGACTCAG	(SEC ID Nº: 162)
3	DPL2 ATAGGACCATTCAGGCTGTCATC	(SEC ID Nº: 163)
5	DPL5 TATCAGTCTGTGTTGACGCAG	(SEC ID Nº: 164)
3	DPL5 ATAGGAGCACTCAGGCTGCTAT	(SEC ID Nº: 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 BsmBl 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 BsmBl 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 Sfil/Ncol y Xhol en los extremos 5' y 3' del VH, respectivamente.

15

10

De forma similar, se introdujeron sitios Sall y Notl 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.

20

VΗ

- 5 VH1 CAGCCGGCCATGGCCCAGGTGCAGCTGGTGCAG (SEC ID №: 166)
- 5 VH3-30 CAGCCGGCCATGGCCCAGGTGCAGCTGGTGGAG (SEC ID №: 167)
- 5 VH3-23 CAGCCGGCCATGGCCGAGGTGCAGCTGTTGGAG (SEC ID №: 168)
- 5 VH3-48 CAGCCGGCCATGGCCGAGGTGCAGCTGGTGGAGTCTGGGGGAG (SEC ID №: 169)
- 5 VH5-51 CAGCCGGCCATGGCCGAGGTGCAGCTGGTGCAG (SEC ID №: 170)
- 3 VH1/3 CTTACCGTTATTCGTCTCATCTCGCACAGTAATACAC (SEC ID №: 171)
- 3 VH3-23 CTTACCGTTATTCGTCTCATTTCGCACAGTAATATAC (SEC ID Nº: 172)
- 3 VH3-48 CTCGCACAGTAATACACAGCCGTGTCCTCGGCTCTCAGGCTG (SEC ID №: 173)
- 3 VH5-51 CTTACCGTTATTCGTCTCATCTCGCACAGTAATACAT (SEC ID №: 174)
- 3 VHext1 CAATACGCGTTTAAACCTGGTAAACCGCCTTACCGTTATTCGTCTCA (SEC ID №: 175)
- 3 VHext2 GTTCCCTGGCCCCAAGAGACGCGCCTTCCCAATACGCGTTTAAACCTG (SEC ID №: 176)
- 3 VHext3 CCTCCACCGCTCGAGACTGTGACCAGGGTTCCCTGGCCCCAAGAG (SEC ID Nº: 177)

٧K

- 5 VK1 CGGGTCGACG*GACATCCAGATGACCCAGTC* (SEC ID №: 178)
- 5 VK3-11 CGGGTCGACGGAAATTGTGTTGACACAGTCTCCAGC (SEC ID №: 179)
- 5 VK3-15 CGGGTCGACG*GAAATAGTGATGACGCAGTCTCCAGC* (SEC ID №: 180)
- 5 VK3-20 CGGGTCGACGGAAATTGTGTTGACGCAGTCTCCAGG (SEC ID №: 181)

٧K

- 3 VK1-33 CCTTACCGTTATTCGTCTCGCTGCTGACAGTAATATGTTGCAATA (SEC ID №: 182)
- 3 VK1-39 CCTTACCGTTATTCGTCTCGCTGCTGACAGTAGTAAGTTGCAAAA (SEC ID №: 183)
- 3 VK3 CCTTACCGTTATTCGTCTCGCTGCTGACAGTAATAAACTGCAAAATC (SEC ID №: 184)
- 3 VKext1 CCAATACGCGTTTAAACCTGGTAAACCGCCTTACCGTTATTCGTCTC (SEC ID №: 185)
- 3 VKext2 GGTCCCTTGGCCGAATGAGACGCGCCTTCCCAATACGCGTTTAAAC (SEC ID №: 186)
- 3 Vkext3R GTGCGGCCGCTTTGATTTCCACCTTGGTCCCTTGGCCGAATG (SEC ID №: 187)
- 3 VKext3G GTGCGGCCGCCCTTTGATTTCCACCTTGGTCCCTTGGCCGAATG (SEC ID №: 188)

Vλ.

- 5 VL1-44 CGGGTCGACGCAGTCTGTGCTGACTCAGCCAC (SEC ID №: 189)
- 5 VL1-51 CGGGTCGACGCAGTCTGTGTTGACGCAGCCGC (SEC ID №: 190)
- 3 VL1-44 CCTTACCGTTATTCGTCTCCTGCTGCACAGTAATAATC (SEC ID №: 191)

- 3 VL1-51 CCTTACCGTTATTCGTCTCCTGTTCCGCAGTAATAATC (SEC ID №: 192)
- 3 VIext2 CCCTCCGCCGAACACAGAGACGCGCCTTCCCAATACGCGTTTAAAC (SEC ID Nº: 193)
- 3 VIext3 GTGCGGCCGCCCTAGGACGGTCAGCTTGGTCCCTCCGCCGAACACAGA (SEC ID №: 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 BsmBl. Se clonó un dominio VH3-23 que contenía una secuencia de CDR3 definida en el pNDS1 modificado usando los sitios de restricción Sfil y Xhol para obtener el vector fagémido pNDS_VHdummy. Este dominio contenía un sitio de BsmBl 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 Notl 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 Notl. Los 12 Marcos conservados Aceptores VL se clonaron en pNDS_VHdummy usando los sitios de restricción de Sfil y Xhol. Los 20 vectores fagémidos pNDS resultantes que se enumeran a continuación podrían usarse en este estadio para clonar CDR3 diversificada usando los sitios de BsmBl 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_VH3-23R_VKd; pNDS_VH3-30_VKd; pNDS_VH3-30_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_VK3-20R

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

10

15

25

30

35

40

45

50

55

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 Nº: 205 mostrada posteriormente) y 5 VHK_FOK (SEC ID Nº: 206 mostrada posteriormente). Resulta importante que estos cebadores también contienen un sitio de restricción de Fokl para escisión de la secuencia de CDR H3 (Figura 4). Los cebadores usados en la segunda etapa de PCR se biotinilaron 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 ºC y 70 ºC) y con varias ADN polimerasas termoestables para establecer condiciones óptimas. Se descubrió que una temperatura de hibridación de 55-58 ºC en combinación con GoTag 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

Como alternativa, la primera etapa de amplificación se realizó usando el cebador 5' 5 VH3-23H2 (SEC ID Nº: 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 №: 195)
- 5 VH3 CCGCACAGCCGGCCATGGCCGAGGTGCAGCTGGTGGAGTCTGG (SEC ID №: 196)
- 5 VH2 CCGCACAGCCGGCCATGGCCCAGRTCACCTTGCTCGAGTCTGG (SEC ID №: 197)
- 5 VH4 CCGCACAGCCGGCCATGGCCCAGGTGCAGCTGCAGGAGTCGGG (SEC ID Nº: 198)
- 5 VH4DP64 CCGCACAGCCGGCCATGGCCCAGCTGCAGCAGCAGCTCCGG (SEC ID №: 199)
- 5 VH4DP63 CCGCACAGCCGGCCATGGCCCAGGTGCAGCTACAGCAGTGGGG (SEC ID №: 200)
- 5 VH3-23H2 TGGAGTGGGTCTCAGCTATTAGTGGTAGTGGT (SEC ID №: 201)
- 3 HJ1/2 CGATGGCCCTTGGTGGAGGCTGAGGAGACRGTGACCAGGGTGCC (SEC ID №: 202)
- 3 HJ3/6 CGATGGGCCCTTGGTGGAGGGCTGAAGAGACGTGACCRTKGTCCC (SEC ID Nº: 203)
- 3 HJ4/5 CGATGGCCCTTGGTGGAGGCTGAGGAGGCGGTGACCAGGGTTCC (SEC ID №: 204)

2ª etapa de PCR

20

25

30

- 5 VHR FOK GAGCCGAGGACACGGCCGGATGTTACTGTGCGAGA (SEC ID Nº: 205)
- 5 VHK FOK GAGCCGAGGACACGCCGGATGTTACTGTGCGAAA (SEC ID Nº: 206)
- 3 JH1 FOK GAGGAGACGGTGACGGATGTGCCCTGGCCCCA (SEC ID №: 207)
- 3 JH2 FOK GAGGAGACGGTGACGGATGTGCCACGGCCCCA (SEC ID Nº: 208)
- 3 JH3456 FOK GAGGAGACGGTGACGGATGTYCCTTGGCCCCA (SEC ID Nº: 209)

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

La CDR H3 amplificada se digirió con Fokl, y los extremos escindidos así como el ADN no digerido se retiraron 5 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 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 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 8,1x109 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 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 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 (~5x109 transformantes) (Figura 11). Los restos clave para mantener la estructura canónica de las CDR se mantuvieron constantes en el diseño de CDR3 para cadenas VK y Vλ. 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 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 BsmBl 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 BsmBl. 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 Vλ 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 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 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 1,6x10¹⁰ transformantes. De forma similar, se generaron 13 bibliotecas de cadena ligera primarias que representaban un total de 6,9x109 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.

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

5

10

15

20

40

50

60

Nº: 226)

5 H3_R_biot ATGATGCTGCACGTCTCCGAGA (SEC ID Nº: 210)
3 H3_M_biot CCACGTCATCCGATCCGTCTCCCCCAATAATCCAT (SEC ID Nº: 211)
3 H3 F biot CCACGTCATCCGATCCGTCTCCCCCAATAATCAAA (SEC ID Nº: 212)

- H3_4nnsF GCTGGCACGTCTCCGAGANNSNNSNNSNNSTTTGATTATTGGGGGAGACG (SEC ID №: 213)
 H3_4nnsM GCTGGCACGTCTCCGAGANNSNNSNNSNNSNTGGATTATTGGGGGAGACG (SEC ID №: 214)
 H3_5nnsF GCTGGCACGTCTCCGAGANNSNNSNNSNNSNNSTTTGATTATTGGGGGAGACG (SEC ID №: 215)
 H3_5nnsM GCTGGCACGTCTCCGAGANNSNNSNNSNNSNNSNNSNTGGATTATTGGGGGAGACG (SEC ID №: 216)
 H3_6nnsF GCTGGCACGTCTCCGAGANNSNNSNNSNNSNNSNNSTTTGATTATTGGGGGAGACG (SEC ID №: 35
 - H3_6nnsM GCTGGCACGTCTCCGAGANNSNNSNNSNNSNNSNNSNNSATGGATTATTGGGGGAGACG (SEC ID N° : 218)
 - H3_6dvkF GCTGGCACGTCTCCGAGADVKDVKDVKDVKDVKDVKTTTGATTATTGGGGGAGACG (SEC ID Nº: 219) H3_6dvkM GCTGGCACGTCTCCGAGADVKDVKDVKDVKDVKDVKATGGATTATTGGGGGAGACG (SEC ID Nº: 220)
 - H3_7dvkF GCTGGCACGTCTCCGAGADVKDVKDVKDVKDVKDVKDVKTTTGATTATTGGGGGAGACG (SEC ID №: 221)
 - $H3_7dvkM$ GCTGGCACGTCTCCGAGADVKDVKDVKDVKDVKDVKDVKATGGATTATTGGGGGAGACG (SEC ID N^{0} : 222)
- 45 H3_7nvtF GCTGGCACGTCTCCGAGANVTNVTNVTNVTNVTNVTNVTTTTGATTATTGGGGGAGACG (SEC ID №: 223)
 H3 7nvtM GCTGGCACGTCTCCGAGANVTNVTNVTNVTNVTNVTNVTATGGATTATTGGGGGAGACG (SEC ID №:
 - 224)
 H3 8nvtF GCTGGCACGTCTCCGAGANVTNVTNVTNVTNVTNVTNVTNVTTTTGATTATTGGGGGAGACG (SEC ID
 - №: 225) H3_8nvtM GCTGGCACGTCTCCGAGANVTNVTNVTNVTNVTNVTNVTNVTATGGATTATTGGGGGAGACG (SEC ID
 - H3_9nvtF GCTGGCACGTCTCCGAGANVTNVTNVTNVTNVTNVTNVTNVTNVTTTTGATTATTGGGGGAGACG (SEC ID №: 227)
- 55 H3_9nvtM GCTGGCACGTCTCCGAGANVTNVTNVTNVTNVTNVTNVTNVTNVTNVTATGGATTATTGGGGGAGACG (SEC ID №: 228)
 - H3_9dvtF GCTGGCACGTCTCCGAGADVTDVTDVTDVTDVTDVTDVTDVTDVTTTTGATTATTGGGGGAGACG (SEC ID №: 229)
 - H3_9dvtM GCTGGCACGTCTCCGAGADVTDVTDVTDVTDVTDVTDVTDVTDVTATGGATTATTGGGGGAGACG (SEC ID №: 230)
 - H3_10dvtF GCTGGCACGTCTCCGAGADVTDVTDVTDVTDVTDVTDVTDVTDVTDVTDTTTGATTATTGGGGGAGACG (SEC ID Nº: 231)
 - H3_10dvtM CTGGCACGTCTCCGAGADVTDVTDVTDVTDVTDVTDVTDVTDVTDVTDVTATGGATTATTGGGGGAGACG (SEC ID \mathbb{N}° : 232)
- 65 5 KL3_biot CCGGTGTAGCGAAGGCGTCTCAGCAG (SEC ID Nº: 233) 3 KL3 biot TAGGGTCGCCTTGATCGTCTCCCGAAGGTCGG (SEC ID Nº: 234)

K 4nns GAAGGCGTCTCAGCAGNNSNNSNNSNNSCCGACCTTCGGGAGACG (SEC ID Nº: 235) K 5nns GAAGGCGTCTCAGCAGNNSNNSNNSNNSCCGNNSACCTTCGGGAGACG (SEC ID Nº: 236) K_6nns GAAGGCGTCTCAGCAGNNSNNSNNSNNSNNSNNSCCGNNSACCTTCGGGAGACG (SEC ID №: 237) 5 L44W biot CGGTCAGTCGCAATACGTCTCCAGCATGGGAT (SEC ID №: 238) 5 L44Y biot CGGTCAGTCGCAATACGTCTCCAGCATATGAT (SEC ID №: 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) L44Y 5nns CGTCTCCAGCATATGATNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 244) 10 L44W 6nns CGTCTCCAGCATGGGATNNSNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 245) L44Y 6nns CGTCTCCAGCATATGATNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 246) 5 L51W_biot CGGTCAGTCGCAATACGTCTCGAACATGGGAT (SEC ID Nº: 247) 5 L51Y biot CGGTCAGTCGCAATACGTCTCGAACATATGAT (SEC ID Nº: 248) 15 L51W 4nns CGTCTCGAACATGGGATNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID №: 249) L51Y 4nns CGTCTCGAACATATGATNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 250) L51W 5nns CGTCTCGAACATGGGATNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID №: 251) L51Y 5nns CGTCTCGAACATATGATNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 252) L51W 6nns CGTCTCGAACATGGGATNNSNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 253) L51Y 6nns CGTCTCGAACATATGATNNSNNSNNSNNSNNSNNSGTGTTGAGACGATCCTC (SEC ID Nº: 254) 20

Ejemplo 7: generación de bibliotecas secundarias

25

30

35

40

45

50

55

60

65

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 digirió con enzimas de restricción Xhol/Notl. 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 Xhol/Notl 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, Vλ) y de este modo se evaluaron 9 combinaciones de biblioteca (es decir, VH1xVK1, VH1xVK3, VH1xVX1, VH3xVK1, VH3xVK3, VH3xVλ1, VH5xVK1, VH5xVK3. VH5xVλ1). El tamaño total de las bibliotecas sintéticas Secundarias (que portaban diversidad sintética tanto en el VH como en el VL) fueron 7,3x10⁹ transformantes. El tamaño total de las bibliotecas naturales secundarias (que portan diversidad natural en el VH y diversidad sintética en el VL) fue de 1,5x10¹⁰ 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 hIFNγ 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 Xhol/Notl 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 Xhol/Notl 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 κ 1, VH3xV κ 1, VH3xV κ 3, VH3xVX 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 μ g/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 1,9x10° transformantes.

10

15

5

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.

20

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 Fokl 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 Fokl casi siempre difieren de la secuencia humana en una base, mientras que el extremo 3' coincidió entre estas dos especies. Las secuencias escindidas por Fokl se encuadran en la Tabla 1 a continuación:

25

	Secuencias 5'		Secuencias 3'			
(SEC ID №: 281)	Humana: TTACTGTGC	<u>G</u> AGA	Humana:	IGGG	GCCAGGGAA	(SEC ID Nº: 285)
(SEC ID Nº: 282)	Ratón: VH1 TTACTGTGC	<u>A</u> AGA	Ratón: JH1	TGGG	GCGCAGGGA	(SEC ID Nº: 286)
(SEC ID Nº: 283)	TTCTGTGC	<u>A</u> AGA	JH2	TGGG	GCCAAGGCA	(SEC ID Nº: 287)
(SEC ID Nº: 284)	VH2 CTACTGTGC	<u>C</u> AGA	JH3	TGGG	GCCAGGGCA	(SEC ID Nº: 288)
(SEC ID Nº: 282)	VH3- 16 TTACTGTGC	<u>A</u> AGA	JH4	TGGG	GTCAGGGCA	(SEC ID Nº: 289)

30

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.

35

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 BsmBl 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.

45

50

55

40

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 ºC, 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 hIFNy 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

30

• cebadores 5': m5 VH1 ATGCGGCCCAGCCGGCCATGGCCSAGGTYCAGCTBCAGCAGTC (SEC ID №: 256) ATGCGGCCCAGCCGGCCATGGCCCAGGTTCACCTGCAGCARTC (SEC ID №: 257) m5 VH2 m5 VH3 ATGCGGCCCAGCCGGCCATGGCCCAGGTRCAGCTGAAGGAGTC (SEC ID Nº: 258) m5 VH4 ATGCGGCCCAGCCGGCCATGGCCCAGGTCCAACTVCAGCARCC (SEC ID Nº: 259) ATGCGGCCCAGCCGGCCATGGCCCAGATCCAGTTGGTVCAGTC (SEC ID №: 260) m5 VH5 m5 VH6 ATGCGGCCCAGCCGGCCATGGCCCAGGTGCAGCTGAAGSASTC (SEC ID Nº: 261) m5 VH7 ATGCGGCCCAGCCGGCCATGGCCGAGGTGCAGSKGGTGGAGTC (SEC ID Nº: 262) ATGCGGCCCAGCCGGCCATGGCCGAAGTGAARSTTGAGGAGTC (SEC ID №: 263) m5 VH8 ATGCGGCCCAGCCGGCCATGGCCGAKGTSVAGCTTCAGGAGTC (SEC ID №: 264) m5 VH9 ATGCGGCCCAGCCGGCCATGGCCGAGGTGAASSTGGTGGAATC (SEC ID №: 265) m5 VH10 m5 VH11 ATGCGGCCAGCCGGCCATGGCCGAGGTGAAGCTGRTGGARTC (SEC ID Nº: 266) ATGCGGCCCAGCCGGCCATGGCCGARGTGAAGCTGRTGGAGTC (SEC ID №: 267) m5 VH12 m5 VH13 ATGCGGCCAGCCGCCATGGCCGAAGTGCAGCTGTTGGAGAC (SEC ID Nº: 268) m5 VH14 ATGCGGCCCAGCCGCCATGCCGARGTGAAGCTTCTCSAGTC (SEC ID Nº: 269) m5 VH15 ATGCGGCCAGCCGGCCATGGCCCARGTTACTCTGAAAGAGT (SEC ID №: 270) · cebadores 3': m3 HJ1 CCTGAACCGCCGCCTCGAGACGGTGACCGTGGTCCC (SEC ID №: 271) m3 HJ2 CCTGAACCGCCGCCTCGAGACTGTGAGAGTGGTGCC (SEC ID №: 272) m3 HJ3 CCTGAACCGCCGCCTCGGCTCGAGACAGTGACCAGAGTCCC (SEC ID Nº: 273) CCTGAACCGCCGCTCCGCTCGAGACGGTGACTGAGGTTCC (SEC ID №: 274) m3 HJ4 2ª PCR • cebadores 5': 5 VHR FOK biot GAGCCGAGGACACGGCCGGATGTTACTGTGCGAGA (SEC ID Nº: 275) · cebadores 3': GGGGCGCAGGGACATCCGTCACCGTCTCCTC (SEC ID №: 276) 3'mJH1 Fok biot 3'mJH2 Fok biot GAGGAGACTGTGAGGGATGTGCCTTGGCCCCA (SEC ID №: 277) 3'JH1 Fok GAGGAGACGGTGACGGATGTGCCCTGGCCCCA (SEC ID №: 278) 3'mJH4 Fok biot GAGGAGACGGTGACGGATGTTCCTTGACCCCA (SEC ID №: 279)

Ejemplo 9: rescate de fagos de las bibliotecas

Cada biblioteca Primaria y Secundaria se rescató independientemente de acuerdo con procedimientos de 10 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 ºC (150 15 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 20 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). 25

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 bibliotecas sintéticas Secundarias; MnA

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

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

10

15

20

30

55

60

65

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 35 fagos AD1 y AE1 (1011-1012 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 hIFNy 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 40 ml de células TG1 con crecimiento exponencial y se incubaron durante una hora a 37 ºC 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 μg/ml y glucosa al 2 %) y se extendieron en placas de Bioensayo de agar 45 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 ºC. 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 50 específicos para la diana (Figura 16).

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

Se recubrieron inmunotubos con 5E3 a 10 μg/ml en PBS durante una noche a 4 °C y se recubrieron inmunotubos para deselección 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 (10¹¹-10¹² 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).

5

10

15

30

35

40

45

50

55

60

65

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 $^{\circ}$ 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 x 10 10 fagos auxiliares MK13K07 y se incubó durante una hora a 37 $^{\circ}$ 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 $^{\circ}$ 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 M13KO7 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 hIFNγ 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 hIFNγ 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'-CTCTTCTGAGATGAGTTTTTG. (SEC ID Nº: 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 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.

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

10

15

20

25

30

35

55

60

65

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 μ g/ml en PBS. Se recubrieron placas de control con anticuerpo monoclonal 1A6 2 μ g/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 μ l 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 μ l 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).

40 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 hIFNy con respecto a unión frente a hIFNy en un experimento de respuesta a dosis. Se recubrieron placas de ELISA (Maxisorb, NUNC) durante una noche con hIFNy 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 45 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 ul 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 50 espectrofotómetro de absorbancia a 450 nm. Los datos muestran que los clones seleccionados se unen con hIFNy 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 hIFNy (Figura 20).

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 IFNγ. Después de la siembra de células Me67.8 en placas, se añadió IFNγ humano recombinante 5 ng/ml a cultivos en presencia de diversas concentraciones de anticuerpos monoclonales anti IFNγ 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 Cl₅₀ (en la que el 50 % de la expresión del MHC de clase II inducida por IFNγ 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 IFNy 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 IFNy en células de melanoma (Figura 22).

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

5

10

15

30

35

40

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 Fugene 6 (Roche, Basilea, Suiza). Brevemente, las células Pico se cultivaron en placas de 6 pocillos a una concentración de 6 x 10⁵ 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 Fugene 6 de acuerdo con las instrucciones del fabricante. Un día después de la transfección, el medio de cultivo se aspiró, y se añadieron 3 ml de medio 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 hIFNγ 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 hIFNγ 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 quimérica 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 quimérica 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 μg/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 μg/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.

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

15

20

25

30

35

40

45

50

55

60

65

Las bibliotecas MnA, MiB y MiC descritas en los Ejemplos 8 y 9 se usaron en paralelo para selecciones de fagos frente a hIFNγ 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 μl 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 μl 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 hIFNγ 2 μg/ml en PBS. Se recubrieron placas de control con BSA recombinante 2 µg/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 µl 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 µl 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 hIFNy 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 sitos 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 VHstuflF1 ATTACTGTGCGAGAGGAGACGNSNNCGTCTCTTGGGGCCAGGGAAC (SEC ID №: 290)
- 5 VHstufIF2 ATTACTGTGCGAGAGGAGACGNCGTCTCTTGGGGCCAGGGAACCCT (SEC ID №: 291)
- 3 VHIF ttatgtgtataggGTTCCCTGGCCCCAAGAGACG (SEC ID №: 292)
- 5 VHIF gtgatctgtacctATTACTGTGCGAGAGGAGACG (SEC ID №: 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 BsmBl 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 20 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 Xhol/Notl y se purificaron en columnas de 25 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 Xhol/Notl 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 30 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 35 la CDR L3, fue de 4,3x109.

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.

55

50

40

45

5

10

LISTADO DE SECUENCIAS

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

5

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

<130> 23135-418002WO

10

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

<1 15 <1

<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> 9 <212> F <213> H	8 PRT	sapier	ns													
5	<400> 2																
		Gln 1	Val	Gln	Leu	Val 5	Gl n	Ser	Gly	Ala	Glu 10	Val	Lys	Lys	Pro	Gly 15	Ala
		Ser	Val	Lys	Val 20	Ser	Cys	Lys	Ala	Ser 25	Gly	Tyr	Thr	Phe	Thr 30	Ser	Tyr
		Gly	Ile	Ser 35	Trp	Val	Arg	Gln	Ala 40	Pro	Gly	Gln	Gly	Leu 45	Glu	Trp	Met
		Gly	Trp 50	Ile	Ser	Ala	Tyr	Asn 55	Gly	Asn	Thr	Asn	Tyr 60	Ala	Gln	Lys	Leu
		Gl n 65	Gly	Arg	Val	Thr	Met 70	Thr	Thr	Asp	Thr	Ser 75	Thr	Ser	Thr	Ala	Tyr 80
		Met	Glu	Leu	Arg	Ser 85	Leu	Arg	Ser	Asp	Asp 90	Thr	Ala	Val	Tyr	Tyr 95	Сув
		Ala	Arg														
10	<210> 3 <211> 9 <212> F <213> H	8 PRT	sapier	າຣ													

15

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser

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

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

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

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

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys

Ala Arg

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

<400> 4

5

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

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr

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 70 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

10

<212> PRT

<213> Homo sapiens

<400>5

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

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

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 $50 \\ 60$

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

		Glu 1	Val	Gln	Leu	Leu 5	Glu	Ser	Gly	Gly	Gly 10	Leu	Val	Gln	Pro	Gly 15	Gly
		Ser	Leu	Arg	Leu 20	Ser	Cys	Ala	Ala	Ser 25	Gly	Phe	Thr	Phe	Ser 30	Ser	Tyr
		Ala	Met	Ser 35	Trp	Val	Arg	Gln	Ala 40	Pro	Gly	Lys	Gly	Leu 45	Glu	Trp	Val
		Ser	Ala 50	Ile	Ser	Gly	Ser	Gly 55	Gly	Ser	Thr	Tyr	Tyr 60	Ala	Asp	Ser	Val
		Lys 65	Gly	Arg	Phe	Thr	Ile 70	Ser	Arg	Asp	Asn	Ser 75	Lys	Asn	Thr	Leu	Tyr 80
		Leu	Gln	Met	Asn	Ser 85	Leu	Arg	Ala	Glu	Asp 90	Thr	Ala	Val	Туг	Туг 95	Cys
		Ala	Lys														
5	<210> 7 <211> 98 <212> PF <213> Ho	RT	apiens	;													
	<400> 7																
		Glu 1	Val	Gln	Leu	Val 5	Gln	Ser	Gly	Ala	Glu 10	Val	Lys	Lys	Pro	Gly 15	Glu
		Ser	Leu	Lys	Ile 20	Ser	Cys	Lys	Gly	Ser 25	Gly	Tyr	Ser	Phe	Thr 30	Ser	Tyr
10		Trp	Ile	Gly 35	Trp	Val	Arg	Gln	Met 40	Pro	Gly	Lys	Gly	Leu 45	Glu	Trp	Met
		Gly	Ile 50	Ile	Tyr	Pro	Gly	Asp 55	Ser	Asp	Thr	Arg	Туг 60	Ser	Pro	Ser	Phe
		G1 n 65	Gly	Gl n	Val	Thr	I l e 70	Ser	Ala	Asp	Lys	Ser 75	Ile	Ser	Thr	Ala	Tyr 80
		Leu	Gln	Trp	Ser	Ser 85	Leu	Lys	Ala	Ser	Asp 90	Thr	Ala	Met	Tyr	Tyr 95	Cys
		Ala	Arg														
	<210>8																

<211> 98 <212> PRT

<213> Homo sapiens

<400>8

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

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 Cys Gln Gln Tyr Asp Asn Leu Pro Pro 85 90 95

Thr Val

<210>9

5

<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

					20					25					30		
		Leu	Asn	Trp 35	Tyr	Gln	Gln	Lys	Pro 40	Gly	Lys	Ala	Pro	Lys 45	Leu	Leu	Ile
		Tyr	Ala 50	Ala	Ser	Ser	Leu	Gln 55	Ser	Gly	Val	Pro	Ser 60	Arg	Phe	Ser	Gly
		Ser 65	Gly	Ser	Gly	Thr	Asp 70	Phe	Thr	Leu	Thr	Ile 75	Ser	Ser	Leu	Gln	Pro 80
		Glu	Asp	Phe	Ala	Thr 85	Tyr	Tyr	Cys	Gln	Gln 90	Ser	Tyr	Ser	Thr	Pro 95	Pro
		Thr	Val														
5	<210> 10 <211> 90 <212> P <213> H	8 RT	apiens	5													
	<400> 10	0															
		Glu 1	Ile	Val	Leu	Thr 5	Gln	Ser	Pro	Ala	Thr 10	Leu	Ser	Leu	Ser	Pro 15	Gly
		Glu	Arg	Ala	Thr 20	Leu	Ser	Cys	Arg	Ala 25	Ser	Gl n	Ser	Val	Ser 30	Ser	Tyr
		Leu	Ala	Trp 35	Tyr	Gln	Gl n	Lys	Pro 40	Gly	Gln	Ala	Pro	Arg 45	Leu	Leu	Ile
		Tyr	Asp 50	Ala	Ser	Asn	Arg	Ala 55	Thr	Gly	Ile	Pro	Ala 60	Arg	Ph€	Ser	Gly
		Ser 65	Gly	Ser	Gly	Thr	Asp 70	Phe	Thr	Leu	Thr	Ile 75	Ser	Ser	Leu	Glu	Pro 80
		Glu	Asp	Phe	Ala	Val 85	Tyr	Tyr	Cys	Gl n	Gl n 90	Arg	Ser	Asn	Trp	Pro 95	Pro
10		Thr	Val														
	<210> 1 <211> 98	8															
15	<212> P <213> H		apiens	5													

15

	Glu 1	Ile	Val	Met	Thr 5	Gln	Ser	Pro	Ala	Thr 10	Leu	Ser	Val	Ser	Pro 15	Gly
	Glu	Arg	Ala	Thr 20	Leu	Ser	Сув	Arg	Ala 25	Ser	Gln	Ser	Val	Ser 30	Ser	Asn
	Leu	Ala	Trp 35	Tyr	Gln	Gln	Lys	Pro 40	Gly	Gln	Ala	Pro	Arg 45	Leu	Leu	Ile
	Tyr	Gly 50	Ala	Ser	Thr	Arg	Ala 55	Thr	Gly	Ile	Pro	Ala 60	Arg	Phe	Ser	Gly
	Ser 65	Gly	Ser	Gly	Thr	Glu 70	Phe	Thr	Leu	Thr	Ile 75	Ser	Ser	Leu	Gln	Ser 80
	Glu	Asp	Phe	Ala	Val 85	Tyr	Tyr	Cys	Gln	G1n 90	Tyr	Asn	Asn	Trp	Pro 95	Pro
	Thr	Val														
<210> 1 <211> 9																
<212> P <213> <i>H</i>		apien	s													
	lomo s	apien	s													
<213> F	lomo s 2			Leu	Thr 5	Gln	Ser	Pro	Gly	Thr 10	Leu	Ser	Leu	Ser	Pro 15	Gly
<213> F	omo s 2 Glu 1	Ile	Val							10					15	
<213> F	Glu 1	Ile Arg	Val Ala	Thr 20	5	Ser	Cys	Arg	Ala 25	10 Ser	Gln	Ser	Val	Ser 30	15 Ser	Ser
<213> F	Glu Glu Tyr	Ile Arg	Val Ala Ala 35	Thr 20 Trp	5 Leu	Ser Gln	Cys	Arg Lys 40	Ala 25	10 Ser Gly	Gln Gln	Ser Ala	Val Pro 45	Ser 30 Arg	15 Ser Leu	Ser
<213> F	Glu Glu Tyr	Ile Arg Leu Tyr 50	Val Ala Ala 35 Gly	Thr 20 Trp	5 Leu Tyr	Ser Gln Ser	Cys Gln Arg 55	Arg Lys 40	Ala 25 Pro	Ser Gly	Gln Gln Ile	Ser Ala Pro	Val Pro 45	Ser 30 Arg	Ser Leu Phe	Ser Leu Ser

5

10

Pro Thr Val

<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 10 Arg Val Thr Ile Ser Cys Ser Gly Ser Ser Ser Asn Ile Gly Ser Asn 25 Thr Val Asn Trp Tyr Gln Gln Leu Pro Gly Thr Ala Pro Lys Leu Leu Ile Tyr Ser Asn Asn Gln Arg Pro Ser Gly Val Pro Asp Arg Phe Ser 50 55 Gly Ser Lys Ser Gly Thr Ser Ala Ser Leu Ala Ile Ser Gly Leu Gln 70 Ser Glu Asp Glu Ala Asp Tyr Tyr Cys Ala Ala Trp Asp Asp Ser Leu 85 90 Asn Gly 10 <210> 14

<211>98

<212> PRT

<213> Homo sapiens

Gln Ser Val Leu Thr Gln Pro Pro Ser Val Ser Ala Ala Pro Gly Gln

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

		Tyr	Val	Ser 35	Trp	Tyr	Gln	Gln	Leu 40	Pro	Gly	Thr	Ala	Pro 45	Lys	Leu	Leu	
		Ile	Tyr 50	Glu	Asn	Asn	Lys	Arg 55	Pro	Ser	Gly	Ile	Pro 60	Asp	Arg	Phe	Ser	
		Gly 65	Ser	Lys	Ser	Gly	Thr 70	Ser	Ala	Thr	Leu	Gl y 75	Ile	Thr	Gly	Leu	Gl n 80	
		Thr	Gly	Asp	Glu	Ala 85	Asp	Tyr	Tyr	Сув	Gly 90	Thr	Trp	Asp	Ser	Ser 95	Leu	
		Ser	Ala															
5	<210> 15 <211> 10 <212> A <213> H	07 DN	apiens	5														
	<400> 1	5																
	tat	tacto	rtg c	gaga	tgaga	a cga	ataa	cgg	taag	gcggt	tt ta	.ccag	gttt	aaa	gcgt	at		60
10	tgg	gaago	ica c	gtct	cttg	g ggd	cago	gaa	ccct	ggtca	ac ag	tctc	g				:	107
15	<210> 10 <211> 30 <212> P <213> H	3 RT	apiens	6														
	<400> 10	6																
		Val 1	Leu	Leu	Cys	Glu 5	Met	Arg	Arg	Ile	Thr 10	Val	Arg	Arg	Phe	Thr 15	Arg	
		Phe	Lys	Arg	Val 20	Leu	Gly	Arg	Arg	Val 25	Ser	Trp	Gly	Gln	Gly 30	Thr	Leu	
20		Val	Thr	Val 35	Ser													
	<210> 1 <211> 3 <212> P <213> H	3 RT	apiens	6														
25	<400> 1	7																

Tyr Tyr Cys Ala Arg Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala

```
Tyr Trp Glu Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser
                                                        25
                Arg
        <210> 18
 5
        <211>35
        <212> PRT
        <213> Homo sapiens
        <400> 18
10
                Ile Thr Val Arg Asp Glu Thr Asn Asn Gly Lys Ala Val Tyr Gln Val
                                   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
        <210>19
        <211>14
15
        <212> ADN
        <213> Secuencia artificial
        <223> sintetizada químicamente
20
        <220>
        <221> misc_feature
        <222> (7)..(14)
        <223> n es a, t, c o g
25
        <400>19
        cgtctcnnnn nnnn
                                  14
        <210> 20
30
        <211> 14
        <212> ADN
        <213> Secuencia artificial
        <220>
35
        <223> sintetizada químicamente
        <221> misc_feature
        <222> (7)..(14)
40
        <223> n es a, t, c o g
        <400> 20
        gcagagnnnn nnnn
                                  14
        <210> 21
45
        <211>386
        <212> ADN
```

<213> Homo sapiens

	<400> 21															
	caggt	gcagc	tggt	gcagt	ta t g	l d ååc	tgag	gtga	agaa	gc ct	gggg	cctc	agtq	aagg	tc	60
	tcctg	caagg	cttc	tggat	ta ca	cctt	cacc	ggct	acta	ta tç	gcact	gggt	gcga	cagg	cc	120
	cctgg	acaag	ggct	tgagt	tg ga	tggg	atgg	atca	accci	ta ad	cagtg	gtgg	caca	aact	at	180
	gcaca	gaagt	ttca	gggca	ag gç	rtcac	catg	acca	.ggga	ca co	tcca	tcag	caca	gcct	ac	240
	atgga	gctga	gcag	gctga	ag at	ctga	cgac	acgg	ccgt	gt at	tact	gtgc	gaga	tgag	ac	300
	gaata	acggt	aagg	cggti	tt ac	cagg	ttta	aacg	rcgta	tt gç	ggaag	gcgc	gtct	cttg	gg	360
5	gccag	ggaac	cctg	gtca	ca gt	ctcg										386
10	<210> 22 <211> 125 <212> PRT <213> Homo	sapier	าร													
	<400> 22															
	Pro 1	Gly	Ala	Ala	Gly 5	Ala	Val	Trp	Gly	Gly 10	Glu	Glu	Ala	Trp	Gly 15	Leu
	Se	r Glu	Gly	Leu 20	Leu	Gln	Gly	Phe	Trp 25	Ile	His	Leu	His	Arg 30	Leu	Leu
	ТY	r Ala	Leu 35	Gly	Ala	Thr	Gly	Pro 40	Trp	Thr	Arg	Ala	Val 45	Asp	Gly	Met
	Asj	50	Pro	Gln	Trp	Trp	His 55	Lys	Leu	Cys	Thr	Glu 60	Val	Ser	Gly	Gln
	G1 ; 65	y His	His	Asp	Gln	Gly 70	His	Val	His	Gln	His 75	Ser	Leu	His	Gly	Ala 80
	G1:	ı Gln	Ala	Glu	Ile 85	Arg	His	Gly	Arg	Val 90	Leu	Leu	Cys	Glu	Met 95	Arg
	Arc	j Ile	Thr	Val 100	Arg	Arg	Phe	Thr	Arg 105	Phe	Lys	Arg	Val	Leu 110	Gly	Arg
4.5	Are	y Val	Ser 115	Trp	Gly	Gln	Gly	Thr 120	Leu	Val	Thr	Val	Ser 125			
15	<210> 23 <211> 125 <212> PRT															
20	<213> Homo <400> 23	sapier	าร													

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

5 <210> 24 <211> 123 <212> PRT

<213> Homo sapiens

	Pro 1	Arg	Суз	Ser	Trp 5	Суз	Ser	Leu	Gly	Leu 10	Arg	Arg	Ser	Leu	Gly 15	Pro	
	Gln	Arg	Ser	Pro 20	Ala	Arg	Leu	Leu	Asp 25	Thr	Pro	Ser	Pro	Ala 30	Thr	Ile	
	Cys	Thr	Gly 35	Cys	Asp	Arg	Pro	Leu 40	Asp	Lys	Gly	Leu	Ser 45	Gly	Trp	Asp	
	Gly	Ser 50	Thr	Leu	Thr	Val	Val 55	Ala	Gln	Thr	Met	His 60	Arg	Ser	Phe	Arg	
	Ala 65	Gly	Ser	Pro	Pro	Gly 70	Thr	Arg	Pro	Ser	Ala 75	Gl n	Pro	Thr	Trp	Ser 80	
	Ala	Gly	Asp	Leu	Thr 85	Thr	Arg	Pro	Cys	Ile 90	Thr	Val	Arg	Asp	Gl u 95	Thr	
	Asn	Asn	Gly	Lys 100	Ala	Val	Tyr	Gln	Val 105	Thr	Arg	Ile	Gly	Lys 110	Ala	Arg	
	Leu	Leu	Gly 115	Pro	Gly	Asn	Pro	Gly 120	His	Ser	Leu						
<212	> 25 > 386 > ADN > <i>Homo s</i>	sapien	s														
<400	> 25																
c	aggtgc	agc t	ggtg	cagt	c tgç	gaget	gag	gtga	agaaq	gc ct	: gg gg	rcctc	agt	gaagq	gt c		60
t	.cctgca	agg c	ttct	ggtt	a cad	cattt	acc	agct	atggl	ta to	caget	gggt	geg	acago	gcc		120
c	ctggac	aag ç	gctt	gagt	g gat	ggga	itgg	atca	geget	tt ac	caatg	gtaa	cac	aaact	at		180
g	rcacaga	agc t	ccag	ggca	g agt	cacc	atg	acca	caga	ca ca	atcca	cgag	cac	agect	ac		240
а	tggagc	tga ç	gago	ctga	g ato	tgac	gac	acgg	ccgt	gt at	tact	gtgc	gag	atgaç	gac .		300
g	raataac	ggt a	aggo	ggtt	t acc	caggt	tta	aacg	cgtai	tt gç	ggaag	igege	gtc	tatta	1 3 3		360
g	rccaggg	aac c	ctgg	tcac	a gto	eteg											386
<212	> 26 > 125 > PRT > <i>Homo s</i>	sapien	s														
<400	> 26																

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

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

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

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

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

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

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

<210> 27

5

<211> 125

<212> PRT

<213> Homo sapiens

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

5

<213> Homo sapiens

	Pro 1	Arg	Cys	Ser	Trp 5	Cys	Ser	Leu	Glu	L eu 10	Arg	Arg	Ser	Leu	Gl y 15	Pro
	Gln	Arg	Ser	Pro 20	Ala	Arg	Leu	Leu	Val 25	Thr	Pro	Leu	Pro	Ala 30	Met	Val
	Ser	Ala	Gly 35	Cys	Asp	Arg	Pro	Leu 40	Asp	Lys	Gly	Leu	Ser 45	Gly	Trp	Asp
	Gly	Ser 50	Ala	Leu	Thr	Met	Val 55	Thr	Gln	Thr	Met	His 60	Arg	Ser	Ser	Arg
	Ala 65	Glu	Ser	Pro	Pro	Gln 70	Thr	His	Pro	Arg	Ala 75	Gln	Pro	Thr	Trp	Ser 80
	Gly	Ala	Asp	Leu	Thr 85	Thr	Arg	Pro	Cys	Ile 90	Thr	Val	Arg	Asp	Gl u 95	Thr
	Asn	Asn	Gly	Lys 100	Ala	Val	Tyr	Gl n	Val 105	Thr	Arg	Ile	Gly	Lys 110	Ala	Arg
	Leu	Leu	Gly 115	Pro	Gly	Asn	Pro	His 120	Ser	Leu						
<210: <211: <212: <213:	386	sapien	s													
<400>	- 29															
	caggtg	cagc	tggt	gcag	tc tç	jgggc	tgag	gtga	agaaq	gc ct	gggt	cctc	ggtg	aaggl	te	60
	teetge	aagg	ctto	tgga	gg ca	cctt	cage	agct	atge	ta to	eaget	gggt	gcga	caggo	cc	120
	cctgga	caag	ggct	tgag	tg ga	tggg	aggg	atca	tece	ta to	tttg	gtac	agca	aacta	ac	180
	gcacag	aagt	tcca	igggc	ag ag	gtcac	gatt	accg	cgga	cg aa	tcca	cgag	caca	gccta	ac	240
	atggag	ctga	gcag	rcctg	agrat	ctga	ggac	acgg	ccgt	gt at	tact	gtgc	gaga	tgaga	ac	300
	gaataa	cggt	aagç	cggt	tt ac	cagg	ttta	aacg	cgta	tt go	gaag	gege	gtct	cttg	39	360
	gecagg	gaac	cctç	gtca	ca gt	ctcg										386
<210><211><211><212><213>	126	sapien	s													
<400>	- 30															

	Pro 1	Gly	Ala	Ala	Gly 5	Ala	Val	Trp	Gly	Gly 10	Glu	Glu	Ala	Trp	Val 15	Leu
	Gly	Glu	Gly	Leu 20	Leu	Gln	Gly	Phe	Trp 25	Arg	His	Leu	Gln	Gln 30	Leu	Cys
	Tyr	Gl n	Leu 35	Gly	Ala	Thr	Gly	Pro 40	Trp	Thr	Arg	Ala	Val 45	Asp	Gly	Arg
	Asp	His 50	Pro	туг	Leu	Trp	Tyr 55	Ser	Lys	Leu	Arg	Thr 60	Glu	Val	Pro	Gly
	Gln 65	Ser	His	Asp	Tyr	Arg 70	Gly	Arg	Ile	His	Glu 75	His	Ser	Leu	His	Gly 80
	Ala	Glu	Gln	Pro	Gl u 85	Ile	Gly	His	Gly	Arg 90	Val	Leu	Leu	Cys	Gl u 95	Met
	Arg	Arg	Ile	Thr 100	Val	Arg	Arg	Phe	Thr 105	Arg	Phe	Lys	Arg	Val 110	Leu	Gly
	Arg	Arg	Val 115	Ser	Trp	Gly	Gln	Gly 120	Thr	Leu	Val	Thr	Val 125	Ser		
<210> 3 <211> 1 <212> P <213> <i>H</i>	25 RT	sapien	s													
<400> 3	1															

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser

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

10

10

5

	20	25	30
Ala Ile Ser	Trp Val Arg Gln	Ala Pro Gly Gln G	ly Leu Glu Trp Met
35		40	45
Gly Gly Ile	Ile Pro Ile Phe	Gly Thr Ala Asn T	yr Ala Gln Lys Phe
50	55		O
Gln Gly Arg	Val Thr Ile Thr	Ala Asp Glu Ser T	hr Ser Thr Ala Tyr
65	70	75	80
Met Glu Leu	Ser Ser Leu Arg 85	Ser Glu Asp Thr A	la Val Tyr Tyr Cys 95
Ala Arg Asp	Glu Arg Gly Gly 100	Leu Pro Gly Leu A	sn Ala Tyr Trp Glu 110

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

<210> 32

5

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

	Pro 1	Arg	Cys	Ser	Trp 5	Cys	Ser	Leu	Gly	Leu 10	Arg	Arg	Ser	Leu	Gly 15	Pro
	Arg	Arg	Ser	Pro 20	Ala	Arg	Leu	Leu	G1u 25	Ala	Pro	Ser	Ala	Ala 30	Met	Leu
	Ser	Ala	Gly 35	Cys	Asp	Arg	Pro	Leu 40	Asp	Lys	Gly	Leu	Ser 45	Gly	Trp	Glu
	Gly	Ser 50	Ser	Leu	Ser	Val	Gln 55	Gln	Thr	Thr	His	Arg 60	Ser	Ser	Arg	Ala
	Glu 65	Ser	Arg	Leu	Pro	Arg 70	Thr	Asn	Pro	Arg	Ala 75	Gln	Pro	Thr	Trp	Ser 80
	Ala	Ala	Asp	Leu	Arg 85	Thr	Arg	Pro	Cys	Ile 90	Thr	Val	Arg	Asp	Glu 95	Thr
	Asn	Asn	Gly	Lys 100	Ala	Val	Tyr	Gln	Val 105	Thr	Arg	Ile	Gly	Lys 110	Ala	Arg
	Leu	Leu	Gly 115	Pro	Gly	Asn	Pro	Gly 120	His	Ser	Leu					
<211>	115 120 <210> 33 <211> 386 <212> ADN <213> Homo sapiens															
<400>	33															
	gaggtg	cagc	tgtt	ggag	to to	3333 3	aggc	ttg	rtaca	ga at	tgggg	ggtc	cctç	gagac	tc	60
	tectgt	gcag	ccto	etgga	tt c	acctt	tage	agct	atgc	ca t	gagct	.gggt	ccg	ccagg	ct	120
	ccaggg	aagg	ggct	ggag	tg g	gtete	agct	atta	agtgg	ta g	t ggt g	gtag	caca	atact	ac	180
	gcagac	tecg	tgaa	agggc	cg gt	ttcac	catc	tcca	agaga	ca a	ttcca	agaa	cacç	getgt	at	240
	ctgcaa	atga	acaç	geetg	ag ag	geega	ggac	acgo	geegt	at a	ttact	gtgc	gaga	atgag	ac	300
	gaataa	cggt	aagg	geggt	tt a	ccago	ttta	aacç	gcgta	tt g	ggaag	làcàc	gtct	cttg	āā	360
	gccagg	gaac	cctç	gtca	ca gr	tatag	Ì									386
<210><211><211><212><213>	126	sapien	s													
<400>	34															

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 $35 \hspace{1.5cm} 40 \hspace{1.5cm} 45$

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

	Glu 1	Val	Gln	Leu	Leu 5	Glu	Ser	Gly	Gly	Gly 10	Leu	Val	Gln	Pro	Gly 15	Gly
	Ser	Leu	Arg	Leu 20	Ser	Cys	Ala	Ala	Ser 25	Gly	Phe	Thr	Phe	Ser 30	Ser	Tyr
	Ala	Met	Ser 35	Trp	Val	Arg	Gln	Ala 40	Pro	Gly	Lys	Gly	Leu 45	Glu	Trp	Val
	Ser	Ala 50	Ile	Ser	Gly	Ser	Gly 55	Gly	Ser	Thr	Tyr	Tyr 60	Ala	Asp	Ser	Val
	Lys 65	Gly	Arg	Phe	Thr	Ile 70	Ser	Arg	Asp	Asn	Ser 75	Lys	Asn	Thr	Leu	Tyr 80
	Leu	Gln	Met	Asn	Ser 85	Leu	Arg	Ala	Glu	Asp 90	Thr	Ala	Val	Tyr	Tyr 95	Cys
	Ala	Arg	Asp	Gl u 100	Arg	Gly	Gly	Leu	Pro 105	Gly	Leu	Asn	Ala	Tyr 110	Trp	Gl u
	_	Ala	Ser 115	Leu	Gly	Ala	Arg	Gl u 120	Pro	Trp	Ser	Gln	Ser 125			
<210> 3 <211> 1 <212> F <213> F	23 PRT	sapien	ıs													
<400> 3	36															
	Pro 1	Arg	Cys	Ser	Cys 5	Trp	Ser	Leu	Gly	Glu 10	Ala	Trp	Tyr	Ser	Leu 15	Gly
	Gly	Pro	Asp	Ser 20	Pro	Val	Gln	Pro	Leu 25	Asp	Ser	Pro	Leu	Ala 30	Ala	Met
	Pro	Ala	Gly 35	Ser	Ala	Arg	Leu	Gln 40	Gly	Arg	Gly	Trp	Ser 45	Gly	Ser	Glr

Leu Leu Val Val Val Val Ala His Thr Thr Gln Thr Pro Arg Ala

Gly Ser Pro Ser Pro Glu Thr Ile Pro Arg Thr Arg Cys Ile Cys Lys

Thr Ala Glu Pro Arg Thr Arg Pro Tyr Ile Thr Val Arg Asp Glu Thr

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
5	<212> ADN
	<213> Homo sapiens

<400> 37

gaggtgcagc	tgttggagtc	tgggggaggc	ttggtacagc	ctggggggtc	cctgagactc	60
tcctgtgcag	cctctggatt	cacctttagc	agctatgcca	tgagctgggt	cegecagget	120
ccagggaagg	ggctggagtg	ggtctcagct	attagtggta	gtggtggtag	cacatactac	180
gcagactccg	tgaagggccg	gttcacatct	ccagagacaa	ttccaagaac	acgctgtatc	240
tgcaaatgaa	cagcctgaga	gccgaggaca	cggccgtata	ttactgtgcg	aaatgagacg	300
aataacggta	aggcggttta	ccaggtttaa	acgcgtattg	ggaaggcgcg	tctcttgggg	360
ccagggaacc	ctggtcacag	tctcg				385

10

15

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

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 His Ile Leu Arg Arg Leu Arg Glu Gly Pro Val 50 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

5

<212> PRT

<213> Homo sapiens

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr 30		Glu 1	Val	Gln	Leu	Leu 5	Glu	Ser	Gly	Gly	Gly 10	Leu	Val	Gln	Pro	Gly 15	Gly
Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val 50 Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr 65 Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys 95 Ala Lys Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu 100 Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser 115 <210> 40 <211> 123 <213> Homo sapiens <400> 40 Fro Arg Cys Ser Cys Trp Ser Leu Gly Glu Asp Ser Pro Leu Ala Ala Met Gly Pro Asp Ser Pro Val Gln Pro Leu Asp Ser Pro Leu Ala Ala Met		Ser	Leu	Arg		Ser	Cys	Ala	Ala		Gly	Phe	Thr	Phe		Ser	Tyr
Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr 70		Ala	Met		Trp	Val	Arg	Gln		Pro	Gly	Lys	Gly		Glu	Trp	Val
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys 95 Ala Lys Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu 100 Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser 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 10 Gly Pro Asp Ser Pro Val Gln Pro Leu Asp Ser Pro Leu Ala Ala Met		Ser		Ile	Ser	Gly	Ser		Gly	Ser	Thr	Tyr		Ala	Asp	Ser	Val
### Ala Lys Asp Glu Arg Gly Gly Leu Pro Gly Leu Asn Ala Tyr Trp Glu 100 ### Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser 125 ### 115 ### Cys Ser Cys Trp Ser Leu Gly Glu Ala Trp Tyr Ser Leu Gly 10 ### Pro Arg Cys Ser Cys Trp Ser Leu Gly Glu Ala Trp Tyr Ser Leu Gly 10 ### Gly Pro Asp Ser Pro Val Gln Pro Leu Asp Ser Pro Leu Ala Ala Meters #### Pro Arg Cys Ser Pro Val Gln Pro Leu Asp Ser Pro Leu Ala Ala Meters ###################################			Gly	Arg	Phe	Thr		Ser	Arg	Asp	Asn		Lys	Asn	Thr	Leu	
Gly Ala Ser Leu Gly Ala Arg Glu Pro Trp Ser Gln Ser 115		Leu	Gln	Met	Asn		Leu	Arg	Ala	Glu		Thr	Ala	Val	Tyr		Cys
<pre>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 10 Gly Pro Asp Ser Pro Val Gln Pro Leu Asp Ser Pro Leu Ala Ala Met</pre>		Ala	Lys	Asp		Arg	Gly	Gly	Leu		Gly	Leu	Asn	Ala		Trp	Glu
<pre><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 Gly Pro Asp Ser Pro Val Gln Pro Leu Asp Ser Pro Leu Ala Ala Met</pre>		Gly	Ala		Leu	Gly	Ala	Arg		Pro	Trp	Ser	Gln				
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	<211> 12 <212> P	23 RT	apiens	6													
1 5 10 15 Gly Pro Asp Ser Pro Val Gln Pro Leu Asp Ser Pro Leu Ala Ala Met	<400> 40)															
			Arg	Cys	Ser	_	Trp	Ser	Leu	Gly		Ala	Trp	Tyr	Ser		Gly
		Gly	Pro	Asp		Pro	Val	Gln	Pro		Asp	Ser	Pro	Leu		Ala	Met

Pro Ala Gly Ser Ala Arg Leu Gln Gly Arg Gly Trp Ser Gly Ser Gln

Leu Leu Val Val Val Val Ala His Thr Thr Gln Thr Pro Arg Ala

	Gly 65	Ser	Pro	Ser	Pro	Glu 70	Thr	Ile	Pro	Arg	Thr 75	Arg	Суз	Ile	Cys	Lys 80	
	Thr	Ala	Glu	Pro	Arg 85	Thr	Arg	Pro	Tyr	Ile 90	Thr	Val	Arg	Asn	Gl u 95	Thr	
	Asn	Asn	Gly	Lys 100	Ala	Val	Tyr	Gln	Val 105	Thr	Arg	Ile	Gly	Lys 110	Ala	Arg	
	Leu	Leu	Gly 115	Pro	Gly	Asn	Pro	Gly 120	His	Ser	Leu						
5	<210> 41 <211> 386 <212> ADN <213> <i>Homo</i>	sapier	າຣ														
	<400> 41																
	caggtgo	cage	tggt	ggagt	ic to	igggg	aggc	gtgg	jtcca	gc c	tggga	aggto	cct	gaga	ctc		60
	tectgto	gcag	cctc	tggat	t ca	cctt	cagt	aget	atgo	ta t	gcact	tgggt	ccg	ccag	gct		120
	ccaggca	agg	ggct	ggagt	a go	rtggc	agtt	atat	cata	tg a	tggaa	agtaa	taa	atac	tac		180
	gcagact	ccg	tgaa	gggco	egrat	tcac	catc	tcca	ıgaga	ca a	ttcc	aagaa	cac	gctg	tat		240
	ctgcaaa	atga	acag	cctga	ag ag	ıctga	ggac	acgo	gctgt	gt a	ttaci	tgtgd	gag	atga	gac		300
	gaataad	eggt	aagg	cggtt	t ac	cagg	ttta	aacç	jcgta	tt g	ggaaq	ggege	gto	tctt	ggg		360
10	gccaggg	gaac	cctg	gtcac	ca gt	ctcg											386
15	<210> 42 <211> 124 <212> PRT <213> Homo	sapier	าร														

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

5

<213> Homo sapiens

Gln 1	Val	Gln	Leu	Val 5	Glu	Ser	Gly	Gly	Gly 10	Val	Val	Gln	Pro	Gly 15	Arg
Ser	Leu	Arg	Leu 20	Ser	Cys	Ala	Ala	Ser 25	Gly	Phe	Thr	Phe	Ser 30	Ser	Tyr
Ala	Met	His 35	Trp	Val	Arg	Gln	Ala 40	Pro	Gly	Lys	Gly	Leu 45	Glu	Trp	Val
Ala	Val 50	Ile	Ser	Туг	Asp	Gly 55	Ser	Asn	Lys	Tyr	Tyr 60	Ala	Asp	Ser	Val
Lys 65	Gly	Arg	Phe	Thr	Ile 70	Ser	Arg	Asp	Asn	Ser 75	Lys	Asn	Thr	Leu	Tyr 80
Leu	Gln	Met	Asn	Ser 85	Leu	Arg	Ala	Glu	Asp 90	Thr	Ala	Val	Tyr	Tyr 95	Cys
Ala	Arg	Asp	Glu 100	Arg	Gly	Gly	Leu	Pro 105	Gly	Leu	Asn	Ala	Tyr 110	Trp	Glu
Gly	Ala	Ser 115	Leu	Gly	Ala	Arg	Glu 120	Pro	Trp	Ser	Gln	Ser 125			
44 124 PRT Homo s	sapiens	s													
44															
Pro 1	Arg	Cys	Ser	Trp 5	Trp	Ser	Leu	Gly	Gl u 10	Ala	Trp	Ser	Ser	Leu 15	Gly
Gly	Pro	Asp	Ser 20	Pro	Val	Gln	Pro	L eu 25	Asp	Ser	Pro	Ser	Val 30	Ala	Met

<210> <211> <212>

<213>

<400>

5

	Leu	Cys	Thr 35	Gly	Ser	Ala	Arg	Leu 40	Gln	Ala	Arg	Gly	Trp 45	Ser	Gly	Trp
	Gln	Leu 50	Tyr	His	Met	Met	Gl u 55	Val	Ile	Asn	Thr	Thr 60	Gl n	Thr	Pro	Arg
	Ala 65	Asp	Ser	Pro	Ser	Pro 70	Glu	Thr	Ilę	Pro	Arg 75	Thr	Arg	Cys	Il∉	Cys 80
	Lys	Thr	Ala	Glu	Leu 85	Arg	Thr	Arg	Leu	Cys 90	Ile	Thr	Val	Arg	Asp 95	Glu
	Thr	Asn	Asn	Gly 100	Lys	Ala	Val	Tyr	Gln 105	Val	Thr	Arg	Ile	Gly 110	Lys	Ala
	Arg	Leu	Leu 115	Gly	Pro	Gly	Asn	Pro 120	Gly	His	Ser	Leu				
<211> <212>	x210> 45 x211> 386 x212> ADN x213> Homo sapiens															
<400>	45															
	gaggto	gcagc	tggt	ggag	tc to	3 3 3 39	aggc	ttgg	tacaç	gc ct	gggg	ggtc	cctg	agact	c	60
	tcctgt	gcag	cct	etgga	tt ca	acctt	cagt	agct	atago	ca to	raact	gggt	ccgc	caggo	et	120
	ccaggo	gaagg	ggct	ggag	tg g	gttto	atac	atta	gtagi	ta gt	agta	gtac	cata	tacta	c	180
	gcagad	ctctg	tgaa	aggge	cg at	ttcac	catc	tcca	gagad	ca at	gcca	agaa	ctca	ctgta	ıt	240
	ctgcaa	aatga	acaç	gcctg	ag ag	geega	ggac	acgg	ctgt	gt at	tact	gtgc	gaga	tgaga	ac	300
	gaataa							aacg	cgtai	tt gg	gaag	gcgc	gtct	cttgg	ją	360
	gccago	ggaac	cctq	ggtca	ca gt	teteg	Ī									386
<210><211><212><213>	122	apiens	5													
<400>	46															

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

5

<213> Homo sapiens

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

5

<213> Homo sapiens

Pro Arg Cys Ser Trp Trp Ser Leu Gly Glu Ala Trp Tyr Ser Leu Gly

	1				5					10					15		
	Gly	Pro	Asp	Ser 20	Pro	Val	Gln	Pro	Leu 25	Asp	Ser	Pro	Ser	Val 30	Ala	Ile	
	Ala	Thr	Gly 35	Ser	Ala	Arg	Leu	Gln 40	Gly	Arg	Gly	Trp	Ser 45	Gly	Phe	His	
	Thr	Leu 50	Val	Val	Val	Val	Val 55	Pro	Tyr	Thr	Thr	Gln 60	Thr	Leu	Arg	Ala	
	Asp 65	Ser	Pro	Ser	Pro	Glu 70	Thr	Met	Pro	Arg	Thr 75	His	Cys	Ile	Cys	Lys 80	
	Thr	Ala	Glu	Pro	Arg 85	Thr	Arg	Leu	Cys	Ile 90	Thr	Val	Arg	Asp	Glu 95	Thr	
	Asn	Asn	Gly	Lys 100	Ala	Val	Tyr	Gln	Val 105	Thr	Arg	Ile	Gly	Lys 110	Ala	Arg	
	Leu	Leu	Gly 115	Pro	Gly	Asn	Pro	Gly 120	His	Ser	Leu						
<210> 4 <211> 2 <212> A <213> F	90 DN	sapien	ıs														
<400> 4	9																
ga	ggtgd	cago	tggt	gcagt	c tg	gagca	agag	gtga	aaaaq	je ee	aaaa	agtc	tctg	aagat	c	60)
to	ctgta	agg	gttc	tggat	a ca	gatti	tacc	aget	actgo	ga to	ggct	gggt	gege	cagat	g	120)
CC	cggga	aaag	gcct	ggagt	g ga	tggg	gatc	atct	atcct	g gt	gact	ctga	tacc	agata	ac.	180)
ag	cccgt	cct	teca	aggco	a gg	tcac	catc	tcag	ccgac	ca ag	tcca	tcag	cacc	gccta	ac	240)
ct	gcagt	gga	gcag	cctga	a gg	cctc	ggac	accg	ccato	gt at	tact	gtgc				290)
<210> 5 <211> 1 <212> F <213> <i>F</i>	26 PRT	sapieri	ıs														
<400> 5	0																

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

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

5

<213> Homo sapiens

	Pro 1	Arg	Сув	Ser	Trp 5	Cys	Ser	Leu	Glu	Gln 10	Arg	Lys	Ser	Pro	Gly 15	Ser
	Leu	Arg	Ser	Pro 20	Val	Arg	Val	Leu	Asp 25	Thr	Ala	Leu	Pro	Ala 30	Thr	Gly
	Ser	Ala	Gly 35	Cys	Ala	Arg	Cys	Pro 40	Gly	Lys	Ala	Trp	Ser 45	Gly	Trp	Gly
	Ser	Ser 50	Ile	Leu	Val	Thr	Leu 55	Ile	Pro	Asp	Thr	Ala 60	Arg	Pro	Ser	Lys
	Ala 65	Arg	Ser	Pro	Ser	Gl n 70	Pro	Thr	Ser	Pro	Ser 75	Ala	Pro	Pro	Thr	Cys 80
	Ser	Gly	Ala	Ala	Arg 85	Pro	Arg	Thr	Pro	Pro 90	Cys	Ile	Thr	Val	Arg 95	Asp
	Glu	Thr	Asn	Asn 100	Gly	Lys	Ala	Val	Tyr 105	Gln	Val	Thr	Arg	Ile 110	Gly	Lys
	Ala	Arg	Leu 115	Leu	Gly	Pro	Gly	Asn 120	Pro	Gly	His	Ser	Leu 125			
<210> <211> <212> <213>	374 ADN	apien:	s													
<400>	53															
gac	atcca	ga t	gacco	agto	tcc	atcci	toc (etgte	tgca	t ct	gtago	gaga	caga	gtca	cc	60
ato	acttg	cc a	ggcga	igtca	gga	catta	agc a	acta	ttta	a ati	tggta	atca	gcag	aaac	ca	120
ggg	aaagc	cc c	taago	etect	gat	ctac	gat q	gcato	caat	t tg	gaaa	cagg	ggtc	ccat	ca	180
agg	ttcag	tg g	aagto	gato	tgg	gaca	gat t	ttac	tttc	a cc	atcaç	gcag	cctg	cagc	ct	240
gaa	gatat	tg c	aacat	atta	ctg	tcag	cag (gaga	ıcgaa	t aa	cggta	aagg	cggt	ttac	ca	300
ggt	ttaaa	cg c	gtatt	ggga	agg	cgcgt	tat d	catto	ggcc	a ag	ggaco	caag	gtgg	aaat	ca	360
aag	gggcg	gc c	gca													374
<210> <211> <212> <213>	122 PRT	apien:	S													
<400>	54															

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

Arg Gln Ser 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 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

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

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

	Arg 1	Thr	Ser	Arg	Pro 5	Ser	Leu	His	Pro	Pro 10	Cys	Leu	His	Leu	Glu 15	Thr	
	Glu	Ser	Pro	Ser 20	Leu	Ala	Arg	Arg	Val 25	Arg	Thr	Leu	Ala	Thr 30	Ile	Ile	
	Gly	Ile	Ser 35	Arg	Asn	Gln	Gly	Lys 40	Pro	Leu	Ser	Ser	Ser 45	Thr	Met	His	
	Pro	Ile 50	Trp	Lys	Gln	Gly	Ser 55	His	Gln	Gly	Ser	Val 60	Glu	Val	Asp	Leu	
	Gly 65	Gln	Ile	Leu	Leu	Ser 70	Pro	Ser	Ala	Ala	Cys 75	Ser	Leu	Lys	Ile	Leu 80	
	Gln	His	Ile	Thr	Val 85	Ser	Ser	Glu	Thr	Asn 90	Asn	Gly	Lys	Ala	Val 95	Tyr	
	Gln	Val	Thr	Arg 100	Ile	Gly	Lys	Ala	Arg 105	Leu	Ile	Arg	Pro	A rg 110	Asp	Gln	
	Gly	Gly	Asn 115	Gln	Arg	Gly	Gly	Arg 120									
<212)> 57 > 374 ?> ADN 8> <i>Homo s</i>	apiens	S														
<400)> 57																
	gacatco	aga	tgac	ccagt	c tc	catco	ctcc	ctgt	ctgca	at ct	gtag	gaga	caga	igtca	cc		60
	atcactt	gcc	aggc	gagto	a gg	acati	tagc	aact	attta	aa at	tggt	atca	gcag	raaac	ca	1	120
	gggaaag	lccc	ctaaq	gatac	t ga	tctac	cgat	gcat	ccaat	tt to	gaaa	cagg	ggto	ccat	ca	1	180
	aggttca	ıgtg	gaagt	tggat	c tg	ggaca	agat	ttta	cttt	ca co	atca	gcag	cctg	cago	ct	2	240
	gaagata	ttg	caaca	atatt	a ct	gtcaq	gcag	cgag	acgaa	at aa	ıcggt	aagg	cggt	ttac	ca	3	300
	ggtttaa	_	_	tggg	ra ag	gege	gtct	catt	cggc	ca aç	ggac	caag	gtgg	aaat	ca		360
	aacgggc	ggc	cgca													:	374
<212)> 58 > 122 !> PRT !> <i>Homo s</i>	apiens	S														
<400	> 58																

	Gly 1	His	Pro	Asp	Asp 5	Pro	Val	Ser	Ile	Leu 10	Pro	Val	Cys	Ile	Cys 15	Arg
	Arg	Gln	Ser	His 20	His	His	Leu	Pro	Gly 25	Glu	Ser	Gly	His	Gln 30	Leu	Phe
	Lys	Leu	Val 35	Ser	Ala	Glu	Thr	Arg 40	Glu	Ser	Pro	Ala	Pro 45	Asp	Leu	Arg
	Cys	Ile 50	Gln	Phe	Gly	Asn	Arg 55	Gly	Pro	Ile	Lys	Val 60	Gln	Trp	Lys	Trp
	Ile 65	Trp	Asp	Arg	Phe	Tyr 70	Phe	His	His	Gln	Gln 75	Pro	Ala	Ala	Arg	Tyr 80
	Cys	Asn	Ile	Leu	Leu 85	Ser	Ala	Ala	Arg	Arg 90	Ile	Thr	Val	Arg	Arg 95	Phe
	Thr	Arg	Phe	Lys 100	Arg	Val	Leu	Gly	Arg 105	Arg	Val	Ser	Phe	Gly 110	Gln	Gly
	Thr	Lys	Val 115	Glu	Ile	Lys	Arg	Ala 120	Ala	Ala						
<210> 50 <211> 10 <212> P <213> F	22 RT	sapien	s													
<400> 5	9															
	Asp 1	Ile	Gln	Met	Thr 5	Gln	Ser	Pro	Ser	Ser 10	Leu	Ser	Ala	Ser	Val 15	Gly
	Asp	Arg	Val	Thr 20	Ile	Thr	Cys	Gln	Al a 25	Ser	Gl n	Asp	Ile	Ser 30	Asn	Tyr

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

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile

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

Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Arg Asp Glu Arg Gly Gly

Leu Pro Gly Leu Asn Ala Tyr Trp Glu Gly Ala Ser His Ser Ala Lys 105 Gly Pro Arg Trp Lys Ser Asn Gly Arg Pro 115 120 <210> 60 <211> 120 <212> PRT 5 <213> Homo sapiens <400>60 Arg Thr Ser Arg Pro Ser Leu His Pro Pro Cys Leu His Leu Glu Thr Glu Ser Pro Ser Leu Ala Arg Arg Val Arg Thr Leu Ala Thr Ile Ile Gly Ile Ser Arg Asn Gln Gly Lys Pro Leu Ser Ser Ser Thr Met His Pro Ile Trp Lys Gln Gly Ser His Gln Gly Ser Val Glu Val Asp Leu Gly Gln Ile Leu Leu Ser Pro Ser Ala Ala Cys Ser Leu Lys Ile Leu Gln His Ile Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val Tyr Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gln Gly Gly Asn Gln Thr Gly Gly Arg 115 10 <210> 61 <211>374 <212> ADN 15 <213> Homo sapiens <400>61

gacatccaga	tgacccagtc	tccatcctcc	ctgtctgcat	ctgtaggaga	cagagtcacc	60
atcacttgcc	gggcaagtca	gagcattagc	agctatttaa	attggtatca	gcagaaacca	120
gggaaagece	ctaagctcct	gatetatget	gcatccagtt	tgcaaagtgg	ggtcccatca	180
aggttcagtg	gcagtggatc	tgggacagat	ttcactctca	ccatcagcag	totgcaacct	240
gaagattttg	caacttacta	ctgtcagcag	cgagacgaat	aacggtaagg	cggtttacca	300
ggtttaaacg	cgtattggga	aggcgcgtct	cattcggcca	agggaccaag	gtggaaatca	360
aaggggcggc	cgca					374

<210>62

5

<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 5

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

Lys Leu Val Ser Ala Glu Thr Arg Glu Ser Pro Ala Pro Asp Leu Cys

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

Ile Trp Asp Arg Phe His Ser His His Gln Gln Ser Ala Thr Arg Phe

Cys Asn Leu Leu Ser Ala Ala Arg Arg Ile Thr Val Arg Arg Phe

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

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

<210> 63

10

<211> 123

<212> PRT

15 <213> Homo sapiens

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

5

<212> PRT

<213> Homo sapiens

Glu Ser Pro Ser Leu Ala Gly Gln Val Arg Ala Leu Ala Ala Ile Il 25 Gly Ile Ser Arg Asn Gln Gly Lys Pro Leu Ser Ser Ser Met Leu Hi 35 Pro Val Cys Lys Val Gly Ser His Gln Gly Ser Val Ala Val Asp Le 50 Gly Gln Ile Ser Leu Ser Pro Ser Ala Val Cys Asn Leu Lys Ile Le 65 Gln Leu Thr Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val Ty 90 Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gl 100 Gly Gly Asn Gln Arg Gly Gly Arg 120 C210>65 C211>374 C212>ADN C13-Homo sapiens C400>65 gacatecaga tgacecagte tecatectee etgetegeat etgtaggaga cagagteace atcaettgee gggaaagee ctaageteet gacetatatge gcatecagte ttgeaacet gaggatttg caacttacta etgteageag etggacaga agggaceaa gggaceaag gtggaaatea gggtttaaaag eggtttaaaag eggtttacaa etgteagaga aggegeteet catteggeca agggaceaag gtggaaatea aacggtaagg eggtttaacaa aacggaceag etggaaatea aacggacaag etggaaatea etggaacaagacaagacaacaacaacaacaacaacaacaaca		1	Inr	Ser	Arg	Pro 5	Ser	Lęu	His	Pro	Pro 10	Cys	Leu	His	Leu	Glu 15	Thr
Pro Val Cys Lys Val Gly Ser His Gln Gly Ser Val Ala Val Asp Le 50 Gly Gln Ile Ser Leu Ser Pro Ser Ala Val Cys Asn Leu Lys Ile Le 65 Gln Leu Thr Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val Ty 90 Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gl 105 Gly Gly Asn Gln Arg Gly Gly Arg 115 C210> 65 C211> 374 C212> ADN C213> Homo sapiens C400> 65 gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc atcacttgcc gggcaagtca gagcattagc agctattaa attggtatca gcagaaacca gggaaagccc ctaagctcct gatctatgct gcatccagtt tgcaaagtg ggtccatca aggttcagtg gcagtggatc tgggacagat ttcactctca ccatcagcag tctgcaacct gaagattttg caacttacta ctgtcagcag cgagacgaat aacggtaagg cggtttacca ggtttaaacc cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca		Glu	Ser	Pro		Leu	Ala	Gly	Gln		Arg	Ala	Leu	Ala		Ile	Ile
Gly Gln Ile Ser Leu Ser Pro Ser Ala Val Cys Asn Leu Lys Ile Le 65 Gln Leu Thr Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val Ty 90 Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gl 100 Gly Gly Asn Gln Arg Gly Gly Arg 115 Gla Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gl 110 Gly Gly Asn Gln Arg Gly Gly Arg 120 <210>65 <211>374 <212>ADN <213>Homo sapiens <400>65 gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc atcacttgcc gggcaagtca gagcattagc agctatttaa attggtatca gcagaaacca gggaaagccc ctaagctcct gatctatgct gcatccagtt tgcaaagtgg ggtccatca aggttcagtg gcagtggatc tgggacagat ttcactctca ccatcagcag tctgcaacct gaagattttg caacttacta ctgtcagcag cgagacgaat aacggtaagg cggtttacca ggtttaaacg cgtattggga aggcgctct cattcggcca agggaccaag gtggaaatca		Gly	Ile		Arg	Asn	Gln	Gly		Pro	Leu	Ser	Ser		Met	Leu	His
Gln Leu Thr Thr Val Ser Ser Glu Thr Asn Asn Gly Lys Ala Val Ty 85 Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gl 100 Gly Gly Asn Gln Arg Gly Gly Arg 120 <210>65 <211>374 <212>ADN <213>Homo sapiens <400>65 gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc atcacttgcc gggcaagtca gagcattagc agctatttaa attggtatca gcagaaacca gggaaagccc ctaagctcct gatctatgct gcatccagtt tgcaaagtg ggtccatca aggttcagtg gcagtggatc tgggacagat ttcactctca ccatcagcag tctgcaacct gaagattttg caacttacta ctgtcagcag cgagacgaat aacggtaagg cggtttaacca ggtttaaacg cgtattgga aggcgcgtct cattcggcca agggaccaag gtggaaatca		Pro		Cys	Lys	Val	Gly		His	Gln	Gly	Ser		Ala	Val	Asp	Leu
Gln Val Thr Arg Ile Gly Lys Ala Arg Leu Ile Arg Pro Arg Asp Gl Gly Gly Asn Gln Arg Gly Gly Arg 115 120 210>65 2211>374 2212> ADN 2213> Homo sapiens <400>65 gacatccaga tgacccagtc tecatectec etgtetgeat etgtaggaga cagagteace ateaettgee gggeaagtea gageattage agetattaa attggtatea geagaaacea gggaaagee etaageteet gatetatget geatecagtt tgeaaagtgg ggteecatea aggtteagtg geagtggate tgggacagat tteaetetea ecateageag tetgeaacet gaagattttg caacttacta etgteageag egagacgaat aaeggtaagg eggtttacea ggtttaaaeg eggtattgga aggeegetet eatteggea agggaceaag gtggaaatea			Gln	Ile	Ser	Leu		Pro	Ser	Ala	Val	_==	Asn	Leu	Lys	Ile	Leu 80
Gly Gly Asn Gln Arg Gly Gly Arg 110 **C210> 65 **C211> 374 **C212> ADN **C213> Homo sapiens** **C400> 65 gacatecaga tgacecagte tecatectee etgtetgeat etgtaggaga cagagteace atcaettgee gggcaagtea gageattage agetatttaa attggtatea geagaaacea gggaaageee etaageteet gatetatget geatecagtt tgeaaagtgg ggteecatea aggtteagtg geagtggate tgggacagat tteactetea ecateageag tetgeaacet gaagattttg caacttacta etgteageag eggacagaat aaeggtaagg eggtttaeea ggtttaaaeg eggtttaaea gggttaaeg eggtttaeea ggtttaaea eggttaaeg eggtttaeea ggtttaaea eggttaaeg eggtttaeea ggtttaaea eggttaaeg eggtttaeea ggtttaaeae eggttaaeg eggtttaeea gggttaaeae eggttaaeg eggtttaeea ggtttaaeae eggttaaeg eggtttaeea gggttaaeeg eggtttaeea ggttaaeeg eggtttaeea gggttaaeeg eggtttaeea gggttaaeeg eggtttaeea gggttaaeeg eggttaaeea ggggaceaeg gtggaaatea		Gln	Leu	Thr	Thr		Ser	Ser	Glu	Thr		Asn	Gly	Lys	Ala		Tyr
<pre>210> 65 <211> 374 <212> ADN <213> Homo sapiens <400> 65 gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc atcacttgcc gggcaagtca gagcattagc agctatttaa attggtatca gcagaaacca gggaaagccc ctaagctcct gatctatgct gcatccagtt tgcaaagtgg ggtcccatca aggttcagtg gcagtggatc tgggacagat ttcactctca ccatcagcag tctgcaacct gaagattttg caacttacta ctgtcagcag cgagacgaat aacggtaagg cggtttacca ggtttaaacg cgtattggga aggcggtct cattcggcca agggaccaag gtggaaatca</pre>		Gln	Val	Thr	_	Ile	Gly	Lys	Ala	_	Leu	Ile	Arg	Pro	_	Asp	Glr
<pre><211> 374 <212> ADN <213> Homo sapiens <400> 65</pre>		G1	G1) Non	G1n	Ara	G1	C1	Ara								
atcacttgcc gggcaagtca gagcattagc agctatttaa attggtatca gcagaaacca gggaaagccc ctaagctcct gatctatgct gcatccagtt tgcaaagtgg ggtcccatca aggttcagtg gcagtggatc tgggacagat ttcactctca ccatcagcag tctgcaacct gaagattttg caacttacta ctgtcagcag cgagacgaat aacggtaagg cggtttacca ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca		GIŢ	GIĀ		GIII	my	GIY	GIĀ									
gggaaagece etaageteet gatetatget geatecagtt tgeaaagtgg ggteecatea aggtteagtg geagtggate tgggacagat tteaetetea ceateageag tetgeaacet gaagattttg caacttacta etgteageag egagacgaat aacggtaagg eggtttacea ggtttaaacg egtattggga aggegegtet eatteggeea agggaceaag gtggaaatea	<211> 3 <212> 7 <213> 7	65 374 ADN <i>Homo s</i>		115	GIII	mry.	GIY	GIŸ									
aggiteagig geagiggate igggaeagat iteaetetea ceateageag ietigeaacet gaagatiitig caacitacia eigieageag egagaegaat aaeggiaagg eggitiaeea ggittaaaeg egiatiggga aggegegiet eaiteggeea agggaeeaag giggaaatea	<211> 3 <212> 7 <213> 7 <400> 6	65 374 ADN <i>Homo s</i> 65	apiens	115					120	ctgc	at c	cgtaç	ıgaga	caga	agtca	.cc	
gaagattttg caacttacta ctgtcagcag cgagacgaat aacggtaagg cggtttacca ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca	<211> 3 <212> 7 <213> 7 <400> 6	65 374 ADN <i>Homo s</i> 65	apien: caga	115	ccag	tc to	ccato	ctcc	120	_				_	_		
ggtttaaacg cgtattggga aggcgcgtct cattcggcca agggaccaag gtggaaatca	<211> (<212> / <212> / <213> / <400> (65 374 ADN <i>Homo s</i> 65 gacatc	apien: caga tgcc	tgac	:ccag	to to	ccato	ctcc tage	ctgt	attt	aa at	tggt	atca	gcaq	gaaac	ca	
	<211> 3 <212> 7 <213> 7 <400> 6 g	65 374 ADN Homo s 65 gacate	apien: caga tgcc gccc	tgac gggc ctaa	ccag aagt	to to ca ga ct ga	ccato agcat atcta	etec tage	ctgt aget geat	attt	aa at	tggt gcaaa	atca	gcaq	gaaac	ca	
aacgggcggc cgca	<211> 3 <212> 4 <213> 4 <400> 6 g	65 374 ADN Homos 65 gacate	apiens caga tgcc gccc agtg	tgac gggc ctaa gcag	ccag aagt gctc tgga	to to ca ga ct ga to to	ccato agcat atcta gggac	etec tage tget	ctgt aget geat	attt ccag	aa at tt to	tggt gcaaa	atca igtgg igcag	gcaq ggtq tctq	gaaac ccat gcaac	ca ca	
	<211> 3	65 374 ADN Homos 65 gacate atcact gggaaa aggtte	apiens caga tgcc gccc agtg	tgac gggc ctaa gcag	ccag aagt gctc tgga ettac	to to ca ga ct ga to to	ccato agcat atcta gggac cgtca	ctcc tagc tgct agat	ctgt aget geat ttea	attt ccag ctct	aa at tt to ca co at aa	tggt gcaaa catca acggt	atca igtgg igcag aagg	gcaq ggtq tctq cggt	gaaac ccat gcaac ttac	ca ca ct	

	Asp 1	Ile	Gln	Met	Thr 5	Gln	Ser	Pro	Ser	Ser 10	Leu	Ser	Ala	Ser	Val 15	Gly
67	7															
ΡI	22 RT	sapien	s													
	Thr	Lys	Val 115	Glu	Ile	Lys	Arg	Ala 120	Ala	Ala						
	Thr	Arg	Phe	Lys 100	Arg	Val	Leu	Gly	Ar g 105	Arg	Val	Ser	Phe	Gly 110	Gln	Gly
	Cys	Asn	Leu	Leu	Leu 85	Ser	Ala	Ala	Arg	Arg 90	Ile	Thr	Val	Arg	Arg 95	Phe
	Ile 65	Trp	Asp	Arg	Phe	His 70	Ser	His	His	Gln	Gln 75	Ser	Ala	Thr	Arg	Phe 80
	Cys	Ile 50	Gln	Phe	Ala	Lys	Trp 55	Gly	Pro	Ile	Lys	Val 60	Gln	Trp	Gln	Trp
	Lys	Leu	Val 35	Ser	Ala	Glu	Thr	Arg 40	Glu	Ser	Pro	Ala	Pro 45	Asp	Leu	Cys
	Arg	Gln	Ser	His 20	His	His	Leu	Pro	Gly 25	Lys	Ser	Glu	His	Gln 30	Leu	Phe
	Gly 1	His	Pro	Asp	Asp 5	Pro	Val	Ser	Ile	Leu 10	Pro	Val	Cys	Ile	Cys 15	Arg

<210> <211> <212> <213>

<400>

5

10

Asp	Arg	Val	Thr	Ile	Thr	Сув	Arg	Ala	Şer	Gln	Şer	Ile	Ser	Şer	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 $35 \hspace{1.5cm} 40 \hspace{1.5cm} 45$

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

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

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

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

Gly Gly Asn Gln Thr Gly Gly Arg 115 120

5	<210> 6 <211> 3 <212> A <213> B	74 ADN	apien	s														
	<400> 6	9																
	gaaa	ttgt	gt to	gacac	cagto	tcc	agcc	acc ·	ctgt	ettte	gt ct	ccag	ggga	aag	agcca	acc		60
	ctct	cctgo	ca go	ggcca	igtca	gag	tgtt	age :	agcta	actta	ng co	tggt	acca	aca	gaaa	cct		120
	ggcc	aggct	te e	caggo	ctcct	cat	ctat	gat (gcat	ccaac	a gg	gcca	ctgg	cat	ccca	gcc		180
		tcagt					-	-										240
		rattt	_					_										300
		taaad			ggga	agg	cgcg	tct ·	catto	egged	a ag	rggac	caag	gtg	gaaai	tca		360
10	aagg	lààcad	ac c	gca														374
ıe	<210> 7 <211> 1 <212> F	22 PRT																
15	<213> F		apien	S														
	<400> /																	
		Gly 1	Asn	Сув	Val	Asp 5	Thr	Val	Ser	Ser	His 10	Pro	Val	Phe	Val	Ser 15	Arg	
		Gly	Lys	Ser	His 20	Pro	Leu	Leu	Gln	Gly 25	Gln	Ser	Glu	Cys	Gl n 30	Leu	Leu	
		Ser	Leu	Val 35	Pro	Thr	Glu	Thr	Trp 40	Pro	Gly	Ser	Gln	Ala 45	Pro	His	Leu	
		Cys	Ile 50	Gln	Gln	Gly	His	Trp 55	His	Pro	Ser	Gln	Val 60	Gln	Trp	Gln	Trp	
		Val 65	Trp	Asp	Arg	Leu	His 70	Ser	His	His	Gln	Gl n 75	Pro	Arg	Ala	Arg	Phe 80	
		Cys	Ser	Leu	Leu	Leu 85	Ser	Ala	Ala	Arg	Arg 90	Ile	Thr	Val	Arg	Arg 95	Ph€	
		Thr	Arg	Phe	Lys 100	Arg	Val	Leu	Gly	Arg 105	Arg	Val	Ser	Phe	Gly 110	Gln	Gly	

20

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

	<210><211><211><212><213>	122	sapier	ıs													
5	<400>	71															
		Glu 1	Ile	Val	Leu	Thr 5	Gln	Ser	Pro	Ala	Thr 10	Leu	Ser	Leu	Ser	Pro 15	Gly
		Glu	Arg	Ala	Thr 20	Leu	Ser	Cys	Arg	Ala 25	Ser	Gln	Ser	Val	Ser 30	Ser	Tyr
		Leu	Ala	Trp 35	Tyr	Gln	Gln	Lys	Pro 40	Gly	Gln	Ala	Pro	Arg 45	Leu	Leu	Ile
		Tyr	Asp 50	Ala	Ser	Asn	Arg	Ala 55	Thr	Gly	Ile	Pro	Ala 60	Arg	Phe	Ser	Gly
		Ser 65	Gly	Ser	Gly	Thr	Asp 70	Phe	Thr	Leu	Thr	Ile 75	Ser	Ser	Leu	Glu	Pro 80
		Glu	Asp	Phe	Ala	Val 85	Tyr	Tyr	Cys	Gln	Gln 90	Arg	Asp	Glu	Arg	Gly 95	Gly
		Leu	Pro	Gly	Leu 100	Aşn	Ala	Tyr	Trp	Glu 105	Gly	Ala	Ser	His	Ser 110	Ala	Lys
		Gly	Pro	Arg 115	Trp	Lys	Ser	Lys	Gly 120	Arg	Pro						
10	<210><211><211><212><213>	121	sapier	ıs													
15	<400>	72															

	Arg 1	Lys	Leu	Cys	His 5	Ser	Leu	Gln	Pro	Pro 10	Cys	Leu	Cys	Leu	Gln 15	Gly	
	Lys	Glu	Pro	Pro 20	Ser	Pro	Ala	Gly	Pro 25	Val	Arg	Val	Leu	Ala 30	Ala	Thr	
	Pro	Gly	Thr 35	Asn	Arg	Asn	Leu	Ala 40	Arg	Leu	Pro	Gly	Ser 45	Ser	Ser	Met	
	Met	His 50	Pro	Thr	Gly	Pro	Leu 55	Ala	Ser	Gln	Pro	Gly 60	Ser	Val	Ala	Val	
	Gly 65	Leu	Gly	Gln	Thr	Ser 70	Leu	Ser	Pro	Ser	Ala 75	Ala	Ser	Leu	Lys	Ile 80	
	Leu	Gln	Phe	Ile	Thr 85	Val	Ser	Ser	Glu	Thr 90	Asn	Asn	Gly	Lys	Ala 95	Val	
	Tyr	Gln	Val	Thr 100	Arg	Ile	Gly	Lys	Al a 105	Arg	Leu	Ile	Arg	Pro 110	Arg	Asp	
	Gln	Gly	Gly 115	Asn	Gln	Arg	Gly	Gly 120	Arg								
<212	> 73 > 374 > ADN > <i>Homo</i>	sapier	ns														
<400	> 73																
	gaaatt	gtgt	tgaca	acagt	c tc	cagc	cacc	ctgt	ctttç	gt ct	ccag	ggga	aaga	igeca	cc	6	0
	ctctcct	tgca	gggc	cagto	a ga	gtgti	tagc	agct	actta	ag co	tggt	acca	acaç	raaac	ct	12	0
	ggccag	gata	ccag	gataa	t ca	tctai	tgat	gcat	ccaa	ca gg	gcca	ctgg	cato	ccag	cc	18	0
;	aggttca	agtg	gcagt	tgggt	c tg	ggac	agac	ttca	ctct	ca co	atca	gcag	ccta	gage	ct	24	0
	gaagatt	tttg	cagt	ttatt	a ct	gtca	gcag	cgag	acgaa	at aa	cggt	aagg	cggt	ttac	ca	30	0
•	ggtttaa	aacg	cgtai	ttggg	ja ag	gcgc	gtct	catt	cggc	ca ag	ggac	caag	gtgg	raaat	ca	36	0
,	aacggg	egge	cgca													37	4
<212	> 74 > 122 > PRT > <i>Homo</i>	sapier	าร														
<400°	~ 74																

G]	Ly	Asn	Сув	Val	Asp 5	Thr	Val	Ser	Ser	His 10	Pro	Val	Phe	Val	Ser 15	Arg
G1	Ly	Lys	Ser	His 20	Pro	Leu	Leu	Gln	Gly 25	Gln	Ser	Glu	Cys	Gln 30	Leu	Leu
Se	èr	Leu	Val 35	Pro	Thr	Glu	Thr	Trp 40	Pro	Gly	Ser	Gln	Ala 45	Pro	His	Leu
C?	/S	Ile 50	Gln	Gln	Gly	His	Trp 55	His	Pro	Ser	Gln	Val 60	Gln	Trp	Gln	Trp
Va 65		Trp	Asp	Arg	Leu	His 70	Ser	His	His	Gln	Gln 75	Pro	Arg	Ala	Arg	Phe 80
C?	/S	Ser	Leu	Leu	Leu 85	Ser	Ala	Ala	Arg	Arg 90	Ile	Thr	Val	Arg	Arg 95	Phe
Tì	ır	Arg	Phe	Lys	Arg	Val	Leu	Gly	Arg	Arg	Val	Ser	Phe	Gly	Gln	Gly
					100					105					110	

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

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

<400> 75

	Glu 1	Ile	Val	Leu	Thr 5	Gln	Ser	Pro	Ala	Thr 10	Leu	Ser	Leu	Ser	Pro 15	Gly
	Glu	Arg	Ala	Thr 20	Leu	Ser	Cys	Arg	Ala 25	Ser	Gln	Ser	Val	Ser 30	Ser	Tyr
	Leu	Ala	Trp 35	Tyr	Gln	Gln	Lys	Pro 40	Gly	Gln	Ala	Pro	Arg 45	Leu	Leu	Ile
	Tyr	Asp 50	Ala	Ser	Asn	Arg	Ala 55	Thr	Gly	Ile	Pro	Ala 60	Arg	Phe	Ser	Gly
	Ser 65	Gly	Ser	Gly	Thr	Asp 70	Phe	Thr	Leu	Thr	Ile 75	Ser	Ser	Leu	Glu	Pro 80
	Glu	Asp	Phe	Ala	Val 85	Tyr	Tyr	Cys	Gln	Gln 90	Arg	Asp	Glu	Arg	Gly 95	Gly
	Leu	Pro	Gly	Leu 100	Asn	Ala	Tyr	Trp	Glu 105	Gly	Ala	Ser	His	Ser 110	Ala	Lys
	Gly	Pro	Arg 115	Trp	Lys	Ser	Asn	Gly 120	Arg	Pro						
<210> 76 <211> 12 <212> PI <213> H	21 RT	apiens	5													
<400 > 76	3															

<213> H <400> 76

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

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

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

10

	Met	His 50	Pro	Thr	Gly	Pro	Leu 55	Ala	Ser	Gln	Pro	Gly 60	Ser	Val	Ala	Val	
	Gly 65	Leu	Gly	Gln	Thr	Ser 70	Leu	Ser	Pro	Ser	Ala 75	Ala	Ser	Leu	Lys	Ile 80	
	Leu	Gln	Phe	Ile	Thr 85	Val	Ser	Ser	Glu	Thr 90	Asn	Asn	Gly	Lys	Ala 95	Val	
	Tyr	Gln	Val	Thr 100	Arg	Ile	Gly	Lys	Ala 105	Arg	Leu	Ile	Arg	Pro 110	Arg	Asp	
	Gln	Gly	Gly 115	Aşn	Gln	Thr	Gly	Gly 120	Arg								
<210> 7 <211> 3 <212> A <213> F	74 DN	apien	s														
<400> 7	7																
ga	aaata	gtga	tgac	gcagi	to to	cagc	cacc	ctgt	ctgt	gt ct	ccag	ggga	aaga	igcca	.cc	(60
ct	ctcc	tgca	gggc	cagt	ca ga	ıgtgt	tagc	agca	actt	ag co	ctggt	acca	gcag	gaaac	ct	12	20
gç	gccag	gete	ccag	gctc	ct ca	itcta	tggt	gcat	ccac	ca go	ggeca	ctgg	tato	ccag	cc	18	80
aç	gtte	agtg	gcag	tgggt	tc to	ggac	agag	ttca	ctct	ca co	catca	gcag	ccto	gcagt	ct	24	40
ga	agati	tttg	cagt	ttati	ta ct	gtca	gcag	cgag	racga	at aa	acggt	aagg	cggt	ttac	ca	30	00
gg	gttta	aacg	cgta	ttgg	ga aç	làcàc	gtct	catt	cggc	ca aç	ggac	caag	gtg	gaaat	ca	36	60
aa	agggg (egge	cgca													31	74
<210> 7 <211> 1 <212> P <213> F	23 PRT	apiens	s														
<400> 7	8																

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

<210> 79

<211> 122

5

<212> PRT <213> Homo sapiens

	Arg 1	Lys	Arg	Ser	Leu 5	Gln	Pro	Pro	Cys	Leu 10	Сув	Leu	Gln	Gly	Lys 15	Glu
<400> 8	0															
<210> 8 <211> 1 <212> P <213> F	20 PRT	sapien	s													
	Gly	Pro	Arg 115	Trp	Lys	Ser	Lys	Gly 120	Arg	Pro						
	Leu	Pro	Gly	Leu 100	Asn	Ala	Tyr	Trp	Glu 105	Gly	Ala	Ser	His	Ser 110	Ala	Lys
	Glu	Asp	Phe	Ala	Val 85	Tyr	Tyr	Cys	Gln	Gl n 90	Arg	Asp	Glu	Arg	Gly 95	Gly
	Ser 65	Gly	Ser	Gly	Thr	Glu 70	Phe	Thr	Leu	Thr	Ile 75	Ser	Ser	Leu	Gln	Ser 80
	Tyr	Gly 50	Ala	Ser	Thr	Arg	Ala 55	Thr	Gly	Ile	Pro	Ala 60	Arg	Phe	Ser	Gly
	Leu	Ala	Trp 35	Tyr	Gl n	Gln	Lys	Pro 40	Gly	Gln	Ala	Pro	Arg 45	Leu	Leu	Ile
	Glu	Arg	Ala	Thr 20	Leu	Ser	Cys	Arg	Ala 25	Ser	Gln	Ser	Val	Ser 30	Ser	Asn
	Glu 1	Ile	Val	Met	Thr 5	Gln	Ser	Pro	Ala	Thr 10	Leu	Ser	Val	Ser	Pro 15	Gly

5

10

	Pro	Pro	Ser	Pro 20	Ala	Gly	Pro	Val	Arg 25	Val	Leu	Ala	Ala	Thr 30	Pro	Gly
	Thr	Ser	Arg 35	Asn	Leu	Ala	Arg	Leu 40	Pro	Gly	Ser	Ser	Ser 45	Met	Val	His
	Pro	Pro 50	Gly	Pro	Leu	Val	Ser 55	Gln	Pro	Gly	Ser	Val 60	Ala	Val	Gly	Leu
	Gly 65	Gln	Ser	Ser	Leu	Ser 70	Pro	Ser	Ala	Ala	Cys 75	Ser	Leu	Lys	Ile	Le u 80
	Gln	Phe	Ile	Thr	Val 85	Ser	Ser	Glu	Thr	Asn 90	Asn	Gly	Lys	Ala	Val 95	Tyr
	Gln	Val	Thr	Arg 100	Ile	Gly	Lys	Ala	Arg 105	Leu	Ile	Arg	Pro	Arg 110	Asp	Gln
	Gly	Gly	Asn 115	Gl n	Arg	Gly	Gly	Arg 120								
<210> 8 <211> 3 <212> A <213> H	374 ADN Homo s	apien:	5													
<211> 3 <212> A <213> <i>F</i> <400> 8	374 ADN Homo s 31							at a			+ a a a .					
<211> 3 <212> A <213> F <400> 8	874 ADN Homo s 81	ngtga	tgad													
<211> 3 <212> A <213> F <400> 8	374 ADN Homo s 31 gaaata	igtga :tgca	tga:	ccagt	ca g	agtgi	tage	agc	aactt	ag c	ctggi	tacca	gca	gaaad	cct	
<211> 3 <212> A <213> F <400> 8 g	874 ADN Homo s 81	igtga etgca ggete	tgad gggd ccad	ccagt ggctc	ca g	agtgl	tage atggt	agca	aactt tccac	ag c	ctgg:	tacca actgg	gca tat	gaaad cccaq	ect gec	
<211> 3 <212> A <213> F <400> 8 g g a	374 ADN Homo s 31 gaaata gtctcc	igtga etgca ggctc eagtg	tgad gggd ccaq gcaq	ccagt ggctc gtggg	ca g ect c jtc t	agtgi atcta gggad	tago atggt cagag	age gear	aactt tccad actct	ag c ca g	ctgg ggcca	tacca actgo agcag	gca tat	gaaad cccaq gcagt	eet gee tet	
<211> 3 <212> A <213> F <400> 8 g a g	374 ADN Homo s 31 gaaata gtctcc ggccag	ngtga etgca ggctc eagtg	tgad gggd ccad gcad cagt	ccagt ggctc gtggg tttat	ca g ct c ftc t	agtgi atcta gggad tgtca	ctago atggt cagag agcag	gca ttc	aactt tccad actct gacga	ag c ca g ca c	ctggf ggcca catca acggf	tacca actgo agcao taago	gca tat cet	gaaad cccaq gcagt tttad	gee tet	
<211> 3 <212> A <213> F <400> 8 g g g g g	a74 ADN Homo s a1 gaaata gccag ggccag	ngtga etgca ggctc eagtg etttg	tgad gggd ccaq gcaq cagt	ccagt ggctc gtggg tttat attgg	ca g ct c ftc t	agtgi atcta gggad tgtca	ctago atggt cagag agcag	gca ttc	aactt tccad actct gacga	ag c ca g ca c	ctggf ggcca catca acggf	tacca actgo agcao taago	gca tat cet	gaaad cccaq gcagt tttad	gee tet	
<211> 3 <212> A <213> F <400> 8 g g g g g	araata gaaata gacag ggccag ggttc gaagat ggttta gacggg 22 23 PRT	agtga etgca ggctc eagtg etttg aacg	tgad gggd ccad gcad cagl egta	ccagt ggctc gtggg tttat attgg	ca g ct c ftc t	agtgi atcta gggad tgtca	ctago atggt cagag agcag	gca ttc	aactt tccad actct gacga	ag c ca g ca c	ctggf ggcca catca acggf	tacca actgo agcao taago	gca tat cet	gaaad cccaq gcagt tttad	gee tet	

Gly Asn Ser Asp Asp Ala Val Ser Ser His Pro Val Cys Val Ser Arg $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

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 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 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

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

5

<213> Homo sapiens

	Arg 1	Lys	Arg	Ser	Leu 5	Gln	Pro	Pro	Cys	Leu 10	Суз	Leu	Gln	Gly	Lys 15	Glu	
	Pro	Pro	Ser	Pro 20	Ala	Gly	Pro	Val	Arg 25	Val	Leu	Ala	Ala	Thr 30	Pro	Gly	
	Thr	Ser	Arg 35	Asn	Leu	Ala	Arg	L eu 4 0	Pro	Gly	Ser	Ser	Ser 45	Met	Val	His	
	Pro	Pro 50	Gly	Pro	Leu	Val	Ser 55	Gln	Pro	Gly	Ser	Val 60	Ala	Val	Gly	Leu	
	Gly 65	Gln	Ser	Ser	Leu	Ser 70	Pro	Ser	Ala	Ala	Cys 75	Ser	Leu	Lys	Ile	Leu 80	
	Gln	Phe	Ile	Thr	Val 85	Ser	Ser	Glu	Thr	Asn 90	Asn	Gly	Lys	Ala	Val 95	Tyr	
	Gln	Val	Thr	Arg 100	Ile	Gly	Lys	Ala	Arg 105	Leu	Ile	Arg	Pro	Asp 110	Gln	Gly	
	Gly	Asn	Gln 115	Thr	Gly	Gly	Arg										
		sapien	ıs														
<400>	> 85																
g	aaattg	tgt 1	tgacç	gcagt	c tc	caggo	cacc	ctgt	cttt	gt c	tecaç	jggga	aag	agcca	acc		60
c	tetect	gca (ggged	cagto	a ga	gtgti	agc	agca	gcta	ct ta	agcct	ggta	cca	gcaga	aaa		120
c	ctggcc	agg (ctccc	aggo	t cc	tcato	ctat	ggtg	catc	ca g	caggo	rccac	tgg	catc	cca		180
g	acaggt	tca (gtggd	agtg	g gt	ctgg	gaca	gact	tcac	tc to	cacca	itcag	cag	actg	gag		240
c	ctgaag	att 1	ttgca	igttt	a tt	actgt	cag	cage	gaga	cg a	ataac	ggta	agg	cggti	ta		300
c	caggtt	taa a	acgco	gtatt	g gg	aaggo	egeg	tctc	atto	gg c	caago	gacc	aag	gtgga	aaa		360
t	caaagg	ggc (ggceç	jca													377
<210: <211: <212: <213:	124	sapien	ıs														
<400>	× 86																

Gly Asn Cys Val Asp Ala Val Ser Arg His Pro Val Phe Val Ser Arg 1 5 10 15 Gly Lys Ser His Pro Leu Ile Gln Gly Gln Ser Glu Cys Gln Gln Leu 20 25 Leu Ser Leu Val Pro Ala Glu Thr Trp Pro Gly Ser Gln Ala Pro His 35 40 Leu Trp Cys Ile Gln Gln Gly His Trp His Pro Arg Gln Val Gln Trp 50 55 Gln Trp Val Trp Asp Arg Leu His Ser His His Gln Gln Thr Gly Ala 70 75 Arg Phe Cys Ser 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 Gln Gly Thr Lys Val Glu Ile Lys Gly Ala Ala Ala 115 120

<210>87

<211> 123

<212> PRT

5

<213> Homo sapiens

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 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

5

<212> PRT

<213> Homo sapiens

	Arq 1	, Lys	Leu	Суз	Arg 5	Ser	Leu	Gln	Ala	Pro 10	Cys	Leu	Суз	Leu	Gln 15	Gly
	Lys	Glu	Pro	Pro 20	Ser	Pro	Ala	Gly	Pro 25	Val	Arg	Val	Leu	Ala 30	Ala	Ala
	Thi	Pro	Gly 35	Thr	Ser	Arg	Asn	Leu 40	Ala	Arg	Leu	Pro	Gly 45	Ser	Ser	Ser
	Met	. Val 50	His	Pro	Ala	Gly	Pro 55	Leu	Ala	Ser	Gln	Thr 60	Gly	Ser	Val	Ala
	Va.1 65	. Gly	Leu	Gly	Gln	Thr 70	Ser	Leu	Ser	Pro	Ser 75	Ala	Asp	Trp	Ser	L eu 80
	Lys	: Ile	Leu	Gln	Phe 85	Ile	Thr	Val	Ser	Ser 90	Glu	Thr	Asn	Asn	Gly 95	Lys
	Ala	. Val	Tyr	Gln 100	Val	Thr	Arg	Ile	Gly 105	Lys	Ala	Arg	Leu	Ile 110	Arg	Pro
	Arç	, Asp	Gln 115	Gly	Gly	Asn	Gln	Arg 120	Gly	Gly	Arg					
<212)> 89 > 377 ?> ADN 3> <i>Homo</i>	sapien	s													
<400	> 89															
	gaaatt	gtgt	tgacç	rcagt	c tco	aggo	acc	ctgto	stttg	rt ct	ccago	gga	aaga	gccac	ec .	60
	ctctcc	tgca	gggcc	agtc	a gaç	gtgtt	agc	agcaç	gctac	t ta	geetç	gta	ccag	cagaa	aa	120
	cetgge	cagg	ctccc	aggc	t cct	cato	tat	ggtgd	catco	a gc	agggo	cac	tggc	ateco	a	180
	gacagg	ttca	gtggc	agtg	g gto	ctggg	aca	gactt	cact	c tc	accat	cag	caga	ctgga	ng	240
	cctgaa	gatt	ttgca	igttt	a tta	actgt	cag	cagco	gagac	g aa	taacq	gta	aggo	ggttt	:a	300
	ccaggt	ttaa	acgcg	tatt	g gga	aggo	gcg .	totoa	attcg	ig cc	aaggg	gacc	aagg	tggaa	a	360
	tcaaac	gggc	ggccg	rca.												3 77
<212)> 90 > 124 ?> PRT 3> <i>Homo</i> :	sapien	s													
<400	> 90															

		Gly 1	Aşn	Суз	Val	Asp 5	Ala	Val	Ser	Arg	His 10	Pro	Val	Phę	Val	Ser 15	Arg
		Gly	Lys	Ser	His 20	Pro	Leu	Leu	Gln	Gly 25	Gl n	Ser	Gl u	Cys	Gl n 30	Gl n	Leu
		Leu	Ser	Leu 35	Val	Pro	Ala	Glu	Thr 40	Trp	Pro	Gly	Ser	Gln 45	Ala	Pro	His
		Leu	Trp 50	Cys	Ile	Gln	Gln	Gly 55	His	Trp	His	Pro	Arg 60	Gln	Val	Gln	Trp
		Gl n 65	Trp	Val	Trp	Asp	Arg 70	Leu	His	Ser	His	His 75	Gl n	Gln	Thr	Gly	Ala 80
		Arg	Phe	Суз	Ser	Leu 85	Leu	Leu	Ser	Ala	Ala 90	Arg	Arg	Ile	Thr	Va l 95	Arg
		Arg	Phe	Thr	Arg 100	Phe	Lys	Arg	Val	Leu 105	Gly	Arg	Arg	Val	Ser 110	Phe	Gly
		Gln	Gly	Thr 115	Lys	Val	Glu	Ile	Lys 120	Arg	Ala	Ala	Ala				
5	<210> 9 <211> 12 <212> P <213> H	23 RT	apien:	s													
	<400> 9	1															
		Glu 1	Ile	Val	Leu	Thr 5	Gln	Ser	Pro	Gly	Thr 10	Leu	Ser	Leu	Ser	Pro 15	Gly
		Glu	Arg	Ala	Thr 20	Leu	Ser	Cys	Arg	Ala 25	Ser	Gln	Ser	Val	Ser 30	Ser	Ser
		Tyr	Leu	Ala 35	Trp	Tyr	Gln	Gln	Lys 40	Pro	Gly	Gln	Ala	Pro 45	Arg	Leu	Leu
		Ile	Tyr 50	Gly	Ala	Ser	Ser	Arg 55	Ala	Thr	Gly	Ile	Pro 60	Asp	Arg	Phę	Ser
10		Gly 65	Ser	Gly	Ser	Gly	Thr 70	Asp	Phe	Thr	Leu	Thr 75	Ile	Ser	Arg	Leu	Glu 80

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

		Gly	Leu	Pro	Asp 100	Leu	Asn	Ala	Tyr	Trp 105	Glu	Gly	Ala	Ser	His 110	Ser	Ala
		Lys	Gly	Pro 115	Arg	Trp	Lys	Ser	Asn 120	Gly	Arg	Pro					
5	<210> 9 <211> 1 <212> F <213> F	23 PRT	apien:	s													
	<400> 9	2															
		Arg 1	Lys	Leu	Cys	Arg 5	Ser	Leu	Gln	Ala	Pro 10	Cys	Leu	Cys	Leu	Gln 15	Gly
		Lys	Glu	Pro	Pro 20	Ser	Pro	Ala	Gly	Pro 25	Val	Arg	Val	Leu	Ala 30	Ala	Ala
		Thr	Pro	Gly 35	Thr	Ser	Arg	Asn	Leu 40	Ala	Arg	Leu	Pro	Gly 45	Ser	Ser	Ser
		Met	Val 50	His	Pro	Ala	Gly	Pro 55	Leu	Ala	Ser	Gln	Thr 60	Gly	Ser	Val	Ala
		Val 65	Gly	Leu	Gly	Gln	Thr 70	Ser	Leu	Ser	Pro	Ser 75	Ala	Asp	Trp	Ser	Leu 80
		Lys	Ile	Leu	Gln	Phe 85	Ile	Thr	Val	Ser	Ser 90	Glu	Thr	Asn	Asn	Gly 95	Lys
		Ala	Val		Gln 100	Val	Thr	Arg	Ile	Gly 105	Lys	Ala	Arg	Leu	Ile 110	Arg	Pro
10		Arg	Asp	Gln 115	Gly	Gly	Asn	Gln	Thr 120	Gly	Gly	Arg					
15	<210> 9 <211> 3 <212> A <213> B	80 NDN	apien:	s													
	<400> 9	3															

	cagtete	gtgt	tgac	gcago	ec go	cctc	agtg	tctç	gegge	ec e	agga	cagaa	ggt	cacc	atc		60
	tectget	ctg	gaag	cagct	c ca	acat	tggg	aata	atta	tg t	atec	tggta	сса	gcag	ctc		120
	ccaggaa	acag	cccc	caaac	et co	tcat	ttat	gaca	ataa	ta a	gcga	ccctc	agg	gatt	cct		180
	gaccgat	tct	ctgg	ctcca	aa gt	ctgg	cacg	tcaç	gccac	cc t	gggc	atcac	. cgg	acto	cag		240
	actgggg	jacg	aggc	cgatt	a tt	actg	cgga	acaç	gaga	cg a	ataa	cggta	agg	cggt	tta		300
	ccaggtt	taa	acgc	gtatt	g gg	raagg	cgcg	tctc	tgtg	tt c	ggcg	gaggg	acc	aagc	tga		360
	ccgtcct	agg	ggcg	geege	a												380
<21 <21 <21	0> 94 1> 122 2> PRT 3> <i>Homo</i> s	sapien	ıs														
	Ala 1	Val	Cys	Val	Asp 5	Ala	Ala	Ala	Leu	Ser 10	Val	Cys	Gly	Pro	Arg 15	Thr	
	Glu	Gly	His	His 20	Leu	Leu	Leu	Trp	Lys 25	Gln	Leu	Gln	His	Trp 30	Glu	Leu	
	Cys	Ile	Leu 35	Val	Pro	Ala	Ala	Pro 40	Arg	Asn	Ser	Pro	Gln 45	Thr	Pro	His	
	Leu	Gln 50	Ala	Thr	Leu	Arg	Asp 55	Ser	Pro	Ile	Leu	Trp 60	Leu	Gln	Val	Trp	
	His 65	Val	Ser	His	Pro	Gly 70	His	His	Arg	Thr	Pro 75	Asp	Trp	Gly	Arg	Gly 80	
	Arg	Leu	Leu	Leu	Arg 85	Asn	Arg	Arg	Arg	Ile 90	Thr	Val	Arg	Arg	Phe 95	Thr	
	Arg	Phe	Lys	Arg 100	Val	Leu	Gly	Arg	Arg 105	Val	Ser	Val	Phe	Gly 110	Gly	Gly	

10

15

5

<210>95 <211> 122 <212> PRT <213> Homo sapiens

115

<400>95

120

Thr Lys Leu Thr Val Leu Gly Ala Ala Ala

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 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

5

<212> PRT

<213> Homo sapiens

	Arg 1	Ser	Leu	Cys	Arg 5	Ser	Arg	Pro	Gl n	Cys 10	Leu	Arg	Pro	Gln	Asp 15	Arg
	Arg	Ser	Pro	Ser 20	Pro	Ala	Leu	Gl u	Al a 25	Ala	Pro	Thr	Leu	Gly 30	Ile	Ile
	Met	Tyr	Pro 35	Gly	Thr	Ser	Ser	Ser 40	Gln	Glu	Gln	Pro	Pro 45	Asn	Ser	Ser
	Phe	Met 50	Thr	Ile	Ile	Ser	Asp 55	Pro	Gl n	Gly	Phe	Leu 60	Thr	Asp	Ser	Leu
	Al a 65	Pro	Ser	Leu	Ala	Arg 70	Gl n	Pro	Pro	Trp	Ala 75	Ser	Pro	Asp	Ser	Arg 80
	Leu	Gly	Thr	Arg	Pro 85	Ile	Ile	Thr	Ala	Glu 90	Gln	Glu	Thr	Asn	Asn 95	Gly
	Lys	Ala	Val	Tyr 100	Gln	Val	Thr	Arg	Ile 105	Gly	Lys	Ala	Arg	Leu 110	Cys	Val
	Arg	Arg	Arg 115	Asp	Gln	Ala	Asp	Arg 120	Pro	Arg	Gly	Gly	Arg 125			
5	<210> 97 <211> 380 <212> ADN <213> <i>Homo</i>	sapien	s													
	<400> 97															
	cagtctg	tgc t	gacto	cagoo	acc	ctca	gog t	.ctgg	gacc	e eeg	ıggca	gag	ggtca	ccat	c	60
	tettgtt	ctg g	aagca	agete	caa	cate	gga a	gtaa	tact	g taa	actg	gta	ccago	agct	c	120
	ccaggaa	cgg c	cccc	aaact	cct	catc	tat a	igtaa	taat	e ago	ggce	ctc .	agggç	jtece	t	180
	gaccgati															240
	tetgagg															300 360
	ccaggtt: ccgtcct					aggc	ycy c	.ccc	gugu	c cyç	legga	ggg	accas	igety	a	380
10		-99 9	9099	J0900												555
	<210> 98 <211> 119															
15	<212> PRT <213> <i>Homo</i>	sapien	s													
	<400> 98															

	Ala 1	Val	Cys	Ala	Asp 5	Ser	Ala	Thr	Leu	Ser 10	Val	Trp	Asp	Pro	Arg 15	Ala
	Glu	Gly	His	His 20	Leu	Leu	Phe	Trp	Lys 25	Gln	Leu	Gln	His	Arg 30	Lys	Tyr
	Cys	Lys	Leu 35	Val	Pro	Ala	Ala	Pro 40	Arg	Asn	Gly	Pro	Gln 45	Thr	Pro	His
	Leu	Ser 50	Ala	Ala	Leu	Arg	Gly 55	Pro	Pro	Ile	Leu	Trp 60	Leu	Gln	Val	Trp
	His 65	Leu	Ser	Leu	Pro	Gl y 70	His	Gln	Trp	Ala	Pro 75	Val	Gly	Gly	Leu	Leu 80
	Leu	Cys	Ser	Arg	Arg 85	Arg	Ile	Thr	Val	Arg 90	Arg	Phe	Thr	Arg	Phe 95	Lys
	Arg	Val	Leu	Gly 100	Arg	Arg	Val	Ser	Val 105	Phe	Gly	Gly	Gly	Thr 110	Lys	Leu
	Thr	Val	Leu 115	Gly	Ala	Ala	Ala									
<210> 99 <211> 13 <212> P <213> H	22 RT	sapien	s													
<400> 9	9															
	Gl n 1	Ser	Val	Leu	Thr 5	Gln	Pro	Pro	Ser	Ala 10	Ser	Gly	Thr	Pro	Gly 15	Gln

10

5

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 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

5

<213> Homo sapiens

<400> 100

		Arg 1	Ser	Leu	Cys	Leu 5	Ser	His	Pro	Gln	Arg 10	Leu	Gly	Pro	Pro	Gly 15	Arg
		Gly	Ser	Pro	Ser 20	Leu	Val	Leu	Glu	Ala 25	Ala	Pro	Thr	Ser	Gl u 30	Val	Ile
		Leu	Thr	Gly 35	Thr	Ser	Ser	Ser	Gln 40	Glu	Arg	Pro	Pro	Asn 45	Ser	Ser	Ser
		Ile	Val 50	Ile	Ile	Ser	Gly	Pro 55	Gln	Gly	Ser	Leu	Thr 60	Asp	Ser	Leu	Ala
		Pro 65	Ser	Leu	Ala	Pro	Gl n 70	Pro	Pro	Trp	Pro	Ser 75	Val	Gly	Ser	Ser	Leu 80
		Arg	Met	Arg	Leu	Ile 85	Ile	Thr	Val	Gln	Gl n 90	Glu	Thr	Asn	Asn	Gly 95	Lys
		Ala	Val	Tyr	Gln 100	Val	Thr	Arg	Ile	Gly 105	Lys	Ala	Arg	Leu	Cys 110	Val	Arg
		Arg	Arg	Asp	Gln	Ala	Asp	Arg	Pro	Arg	Gly	Gly	Arg				
							115					120					
5	<210> <211> <212> <213>	19 PRT	sapien	ıs													
	<400>	101															
		Cys 1	Ala	Lys	Ala	Phe 5	Arg	Pro	Asn	Trp	Gly 10	Ser	Arg	Val	Leu	Tyr 15	Phe
10		Asp	Tyr	Trp													
15	<210> <211> <212> <213>	12 PRT	sapien	ıs													
	<400>	102															
20				Cys 1	: Ala	Lys	Leu	Asn 5	Trp	Gly	Trp	Phe	Asp 10	Pro	Trp	ı	
-	<210> <211> <212> <213>	12 PRT	sapien	ıs													

<400> 103 Cys Ala Lys Gly Trp Glu Gly Glu Tyr Asp Tyr Trp 5 <210> 104 <211> 17 <212> PRT <213> Homo sapiens 10 <400> 104 Cys Ala Lys Asp Leu Trp Leu Asp Ser Ser Ser Asn Trp Phe Asp Pro 10 Trp <210> 105 15 <211>11 <212> PRT <213> Homo sapiens <400> 105 20 Cys Ala Lys Asp Leu Pro Gly Asp Pro His Trp 5 10 <210> 106 <211> 13 25 <212> PRT <213> Homo sapiens <400> 106 Cys Ala Lys Ala Pro Pro Thr Gly Tyr Phe Asp Tyr Trp 30 <210> 107 <211> 14 <212> PRT 35 <213> Homo sapiens <400> 107 Cys Ala Lys Val Gly Leu Thr Gly Val His Phe Asp Tyr Trp 40 <210> 108 <211> 16 <212> PRT <213> Homo sapiens 45 <400> 108 Cys Ala Lys Asp Ser Tyr Gly Ser Gly Ser Tyr Tyr Asn Asp Tyr Trp 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
               Tyr Trp
5
        <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
15
        <210> 111
        <211> 15
        <212> PRT
        <213> Homo sapiens
20
        <400> 111
                 Cys Ala Lys Asp Ala Arg Asn Ser Gly Ser Tyr Phe Asp Tyr Trp
25
        <210> 112
        <211> 14
        <212> PRT
        <213> Homo sapiens
        <400> 112
30
                   Cys Ala Lys Gly Val Arg Thr Gly Val Val Phe Asp Tyr Trp
        <210> 113
        <211> 19
35
        <212> PRT
        <213> Homo sapiens
        <400> 113
40
               Cys Ala Lys Gly Val Val Asn Trp Gly Thr Arg Arg Lys Gly Trp Phe
                                                        10
                                                                              15
               Asp Pro Trp
        <210> 114
        <211> 11
        <212> PRT
45
        <213> Homo sapiens
        <400> 114
```

Cys Ala Lys Trp Gly Gly Trp Phe Asp Pro Trp

<210> 115 <211> 13 <212> PRT 5 <213> Homo sapiens <400> 115 Cys Ala Lys Arg Arg Asp Asn Trp Gly Ser Val Asp Trp 10 10 <210> 116 <211> 15 <212> PRT 15 <213> Homo sapiens <400> 116 Cys Ala Lys Gly Ser Gly Phe Ser Ser Gly Trp Phe Asp Ser Trp 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 acgtctccga gannsnnsnn snnsatggat tattggggga gacg 44 <210> 118 50 <211> 14 <212> PRT <213> Homo sapiens <220> <221> MISC_FEATURE 55 <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 10

```
<210>119
         <211> 14
 5
         <212> PRT
         <213> Homo sapiens
         <220>
         <221> MISC_FEATURE
10
         <222> (4)..(4)
         <223> Xaa es Asp o Glu
         <220>
         <221> MISC FEATURE
15
         <222> (5)..(8)
         <223> Xaa es Ser, Pro, Thr, Ala, Cys, Trp, Arg o Gly
         <400> 119
                      Arg Leu Arg Xaa Xaa Xaa Xaa Trp Ile Ile Gly Gly Asp
                                            5
                                                                       10
20
         <210> 120
         <211>25
         <212> ADN
25
         <213> Secuencia artificial
         <220>
         <223> sintetizada químicamente
30
         <400> 120
                                              25
         tgtttctaat cgcaggtgcc agatg
         <210> 121
         <211> 25
         <212> ADN
35
         <213> Secuencia artificial
         <220>
         <223> sintetizada químicamente
40
         <400> 121
         atttatgtta tgacttgtta cactg
                                              25
         <210> 122
45
         <211> 25
         <212> ADN
         <213> Secuencia artificial
         <220>
50
         <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
```

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

	ttttgtttca agctgaatca ctgtg	25
5	<210> 130 <211> 25 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
10	<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	
0.5	<400> 132 gtgtctgtgt ctctcctgct tccag	25
35	<210> 133 <211> 25 <212> ADN <213> Secuencia artificial	
40	<220> <223> sintetizada químicamente	
45	<400> 133 cttgtctcag ttccccattg ggctg	25
50	<210> 134 <211> 25 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
55	<400> 134 atotcatcca cttctgtgtt ctctc	25
60	<210> 135 <211> 25 <212> ADN <213> Secuencia artificial	
C.E.	<220> <223> sintetizada químicamente	
65	<400 > 135	

	ttgggtttct gacaccctca ggatg	25
5	<210> 136 <211> 25 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
10	<400> 136 caggccagtc atgtgagact tcacc	25
15	<210> 137 <211> 25 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
	<400> 137 ctgcctcctc cctggggttt ctgaa	25
25	<210> 138 <211> 25 <212> ADN <213> Secuencia artificial	
30	<220> <223> sintetizada químicamente	
0.5	<400> 138 cccctgtgtc ctctccacag gtgtc	25
35 40	<210> 139 <211> 25 <212> ADN <213> Secuencia artificial	
40	<220> <223> sintetizada químicamente	
45	<400> 139 ccggcacage tgccttctcc ctcag	25
50	<210> 140 <211> 18 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
55	<400> 140 gaggtgcagc tgttggag 18	
60	<210> 141 <211> 26 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
55	<400> 141	

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

	gagtctgtgc cggaagtgca gctgg	25
5	<210> 148 <211> 21 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
10	<400> 148 tatcaggtgc agctggtgca g	21
15	<210> 149 <211> 21 <212> ADN <213> Secuencia artificial	
20	<220> <223> sintetizada químicamente	
	<400> 149 tatcaggtgc agctggtgga g	21
25	<210> 150 <211> 21 <212> ADN <213> Secuencia artificial	
30	<220> <223> sintetizada químicamente	
35	<400> 150 tatgaggtgc agctggtgca g	21
40	<210> 151 <211> 23 <212> ADN <213> Secuencia artificial	
40	<220> <223> sintetizada químicamente	
45	<400> 151 atateteteg cacagtaata cac	23
50	<210> 152 <211> 23 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
55	<400> 152 atateteteg cacagtaata tac	23
60	<210> 153 <211> 23 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
30	<400> 153	

	atatgtctcg cacagtaata cat		23				
5	<210> 154 <211> 32 <212> ADN <213> Secuencia artificial						
10	<220> <223> sintetizada químicamente						
10	<400> 154 tatgacatcc agatgaccca gtctccate	cc tc		32			
15	<210> 155 <211> 20 <212> ADN <213> Secuencia artificial						
20	<220> <223> sintetizada químicament	e					
	<400> 155 ataggagggg tactgtaact	20					
25	<210> 156 <211> 20 <212> ADN <213> Secuencia artificial						
30	<220> <223> sintetizada químicament	:e					
O.E.	<400> 156 ataggaggga gattatcata	20					
35	<210> 157 <211> 24 <212> ADN <213> Secuencia artificial						
40	<220> <223> sintetizada químicament	te					
45	<400> 157 tatgaaattg tgttgacgca gtct		24				
50	<210> 158 <211> 23 <212> ADN <213> Secuencia artificial						
	<220> <223> sintetizada químicament	:e					
55	<400> 158 ataggaggtg agctaccata ctg		23				
60	<210> 159 <211> 24 <212> ADN <213> Secuencia artificial						
65	<220> <223> sintetizada químicament	te					
00	<400> 159						

	tatgaaatag tgatgacgca gtct	24
5	<210> 160 <211> 23 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
10	<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	
0.E	<400> 162 tacagtctgt gctgactcag 20	
35	<210> 163 <211> 23 <212> ADN <213> Secuencia artificial	
40	<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 tgttgacgca g 21	
60	<210> 165 <211> 22 <212> ADN <213> Secuencia artificial	
65	<220> <223> sintetizada químicamente	
55	<400> 165	

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

	cttaccgtta ttcgtctcat ctcgcacagt aatacac	37		
5	<210> 172 <211> 37 <212> ADN <213> Secuencia artificial			
10	<220> <223> sintetizada químicamente			
	<400> 172 cttaccgtta ttcgtctcat ttcgcacagt aatatac	37		
15	<210> 173 <211> 42 <212> ADN <213> Secuencia artificial			
20	<220> <223> sintetizada químicamente			
	<400> 173 ctcgcacagt aatacacagc cgtgtcctcg gctctcaggc tg		42	
25	<210> 174 <211> 37 <212> ADN <213> Secuencia artificial			
30	<220> <223> sintetizada químicamente			
25	<400> 174 cttaccgtta ttcgtctcat ctcgcacagt aatacat	37		
35	<210> 175 <211> 47 <212> ADN <213> Secuencia artificial			
40	<220> <223> sintetizada químicamente			
45	<400> 175 caatacgcgt ttaaacctgg taaaccgcct taccgttatt cgtctc	a		47
50	<210> 176 <211> 48 <212> ADN <213> Secuencia artificial			
	<220> <223> sintetizada químicamente			
55	<400> 176 gttccctggc cccaagagac gcgccttccc aatacgcgtt taaa	ıcctg		48
60	<210> 177 <211> 45 <212> ADN <213> Secuencia artificial			
65	<220> <223> sintetizada químicamente			
55	<400> 177			

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

	ccttaccgtt attcgtctcg ctgctgacag tagtaagttg caaaa	45	
5	<210> 184 <211> 47 <212> ADN <213> Secuencia artificial		
10	<220> <223> sintetizada químicamente		
10	<400> 184 ccttaccgtt attcgtctcg ctgctgacag taataaactg caaaatc		47
15	<210> 185 <211> 47 <212> ADN <213> Secuencia artificial		
20	<220> <223> sintetizada químicamente		
	<400> 185 ccaatacgcg tttaaacctg gtaaaccgcc ttaccgttat tcgtctc		47
25	<210> 186 <211> 46 <212> ADN <213> Secuencia artificial		
30	<220> <223> sintetizada químicamente		
0.5	<400> 186 ggtcccttgg ccgaatgaga cgcgccttcc caatacgcgt ttaaac		46
35	<210> 187 <211> 45 <212> ADN <213> Secuencia artificial		
40	<220> <223> sintetizada químicamente		
45	<400> 187 gtgcggccgc ccgtttgatt tccaccttgg tcccttggcc gaatg	45	
50	<210> 188 <211> 45 <212> ADN <213> Secuencia artificial		
	<220> <223> sintetizada químicamente		
55	<400> 188 gtgcggccgc ccctttgatt tccaccttgg tcccttggcc gaatg	45	
60	<210> 189 <211> 32 <212> ADN <213> Secuencia artificial		
65	<220> <223> sintetizada químicamente		
00	<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 cagtctgtgt tgacgcagcc gc	32	
15	<210> 191 <211> 38 <212> ADN <213> Secuencia artificial		
20	<220> <223> sintetizada químicamente		
	<400> 191 cettacegtt attegtetee tgetgeacag taataate	38	
25	<210> 192 <211> 38 <212> ADN <213> Secuencia artificial		
30	<220> <223> sintetizada químicamente		
35	<400> 192 cettacegtt attegtetee tgttccgcag taataate	38	
	<210> 193 <211> 46 <212> ADN <213> Secuencia artificial		
40	<220> <223> sintetizada químicamente		
45	<400> 193 ccctccgccg aacacagaga cgcgccttcc caatacgcgt tta	aac	46
50	<210> 194 <211> 49 <212> ADN <213> Secuencia artificial		
	<220> <223> sintetizada químicamente		
55	<400> 194 gtgcggccgc ccctaggacg gtcagcttgg tccctccgcc gaa	ıcacaga	49
60	<210> 195 <211> 43 <212> ADN <213> Secuencia artificial		
65	<220> <223> sintetizada químicamente		
00	<400> 195		

	ccgcacagcc ggccatggcc caggtgcagc tggtgcagtc tgg		43
5	<210> 196 <211> 43 <212> ADN <213> Secuencia artificial		
10	<220> <223> sintetizada químicamente		
	<400> 196 ccgcacagcc ggccatggcc gaggtgcagc tggtggagtc tgg		43
15	<210> 197 <211> 43 <212> ADN <213> Secuencia artificial		
20	<220> <223> sintetizada químicamente		
	<400> 197 ccgcacagcc ggccatggcc cagrtcacct tgctcgagtc tgg	43	
25	<210> 198 <211> 43 <212> ADN <213> Secuencia artificial		
30	<220> <223> sintetizada químicamente		
35	<400> 198 ccgcacagcc ggccatggcc caggtgcagc tgcaggagtc ggg		43
33	<210> 199 <211> 43 <212> ADN <213> Secuencia artificial		
40	<220> <223> sintetizada químicamente		
45	<400> 199 ccgcacagcc ggccatggcc cagctgcagc tgcaggagtc cgg		43
50	<210> 200 <211> 43 <212> ADN <213> Secuencia artificial		
	<220> <223> sintetizada químicamente		
55	<400> 200 ccgcacagcc ggccatggcc caggtgcagc tacagcagtg ggg		43
60	<210> 201 <211> 32 <212> ADN <213> Secuencia artificial		
65	<220> <223> sintetizada químicamente		
00	<400> 201		

	tggagtgggt ctcagctatt agtggtagtg gt 3	2	
5	<210> 202 <211> 45 <212> ADN <213> Secuencia artificial		
10	<220> <223> sintetizada químicamente		
. •	<400> 202 cgatgggccc ttggtggagg ctgaggagac rgtgaccagg gtg	cc	45
15	<210> 203 <211> 45 <212> ADN <213> Secuencia artificial		
20	<220> <223> sintetizada químicamente		
	<400> 203 cgatgggccc ttggtggagg ctgaagagac ggtgaccrtk gtcc	С	45
25	<210> 204 <211> 45 <212> ADN <213> Secuencia artificial		
30	<220> <223> sintetizada químicamente		
35	<400> 204 cgatgggccc ttggtggagg ctgaggagac ggtgaccagg gttd	cc	45
	<210> 205 <211> 35 <212> ADN <213> Secuencia artificial		
40	<220> <223> sintetizada químicamente		
45	<400> 205 gagccgagga cacggccgga tgttactgtg cgaga	35	
50	<210> 206 <211> 35 <212> ADN <213> Secuencia artificial		
	<220> <223> sintetizada químicamente		
55	<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		
10	<400> 208 gaggagacgg tgacggatgt gccacggccc ca		32
15	<210> 209 <211> 32 <212> ADN <213> Secuencia artificial		
20	<220> <223> sintetizada químicamente		
	<400> 209 gaggagacgg tgacggatgt yccttggccc ca		32
25	<210> 210 <211> 26 <212> ADN <213> Secuencia artificial		
30	<220> <223> sintetizada químicamente		
0.E	<400> 210 atgatgctgc tggcacgtct ccgaga	26	
35	<210> 211 <211> 35 <212> ADN <213> Secuencia artificial		
40	<220> <223> sintetizada químicamente		
45	<400> 211 ccacgtcate cgatccgtct cccccaataa tccat		35
50	<210> 212 <211> 35 <212> ADN <213> Secuencia artificial		
	<220> <223> sintetizada químicamente		
55	<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 tttgattatt gggggagacg
                                                                          50
          <210> 214
          <211>50
25
          <212> ADN
          <213> Secuencia artificial
          <223> sintetizada químicamente
30
          <220>
          <221> misc_feature
          <222> (19)..(20)
          <223> n es a, c, g, o t
35
          <220>
          <221> misc feature
          <222> (22)..(23)
          <223> n es a, c, g, o t
40
          <220>
          <221> misc_feature
          <222> (25)..(26)
          <223> n es a, c, g, o t
45
          <220>
          <221> misc feature
          <222> (28)..(29)
          <223> n es a, c, g, o t
50
          gctggcacgt ctccgagann snnsnnsnns atggattatt gggggagacg
                                                                                   50
          <210> 215
55
          <211>53
          <212> ADN
          <213> Secuencia artificial
          <220>
60
          <223> sintetizada químicamente
          <220>
          <221> misc feature
          <222> (19)..(20)
65
          <223> n es a, c, g, o t
```

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

```
<220>
          <221> misc_feature
          <222> (19)..(20)
 5
          <223> n es a, c, g, o t
          <220>
          <221> misc feature
          <222> (22)..(23)
10
          <223> n es a, c, g, o t
          <221> misc_feature
          <222> (25)..(26)
15
          <223> n es a, c, g, o t
          <220>
          <221> misc feature
          <222> (28)..(29)
20
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
          <222> (31)..(32)
25
          <223> n es a, c, g, o t
          <220>
          <221> misc feature
          <222> (34)..(35)
30
          <223> n es a, c, g, o t
          gctggcacgt ctccgagann snnsnnsnns nnsnnstttg attattgggg gagacg
                                                                                    56
35
          <210> 218
          <211>56
          <212> ADN
          <213> Secuencia artificial
40
          <223> sintetizada químicamente
          <220>
          <221> misc_feature
45
          <222> (19)..(20)
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
50
          <222> (22)..(23)
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
55
          <222> (25)..(26)
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
60
          <222> (28)..(29)
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
65
          <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		
	<400> 218 gctggcacgt ctccgagann snnsnnsnns nnsnnsatgg attattgggg gagacg	56	
10	<210> 219 <211> 56 <212> ADN <213> Secuencia artificial		
15	<220> <223> sintetizada químicamente		
20	<400> 219 gctggcacgt ctccgagadv kdvkdvkdvk dvkdvktttg attattgggg gagacg	56	
	<210> 220 <211> 56 <212> ADN <213> Secuencia artificial		
25	<220> <223> sintetizada químicamente		
30	<400> 220 gctggcacgt ctccgagadv kdvkdvkdvk dvkdvkatgg attattgggg gagacg	56	
35	<210> 221 <211> 59 <212> ADN <213> Secuencia artificial		
	<220> <223> sintetizada químicamente		
40	<400> 221 gctggcacgt ctccgagadv kdvkdvkdvk dvkdvkdvkt ttgattattg ggggagacg		59
45	<210> 222 <211> 59 <212> ADN <213> Secuencia artificial		
	<220> <223> sintetizada químicamente		
50	<400> 222 gctggcacgt ctccgagadv kdvkdvkdvk dvkdvkdvka tggattattg ggggagacg		59
55	<210> 223 <211> 59 <212> ADN <213> Secuencia artificial		
60	<220> <223> sintetizada químicamente		
65	<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
 5
          <220>
          <221> misc_feature
          <222> (25)..(25)
          <223> n es a, c, g, o t
10
          <220>
          <221> misc feature
          <222> (28)..(28)
          <223> n es a, c, g, o t
15
          <220>
          <221> misc_feature
          <222> (31)..(31)
          <223> n es a, c, g, o t
20
          <220>
          <221> misc_feature
          <222> (34)..(34)
          <223> n es a, c, g, o t
25
          <220>
          <221> misc feature
          <222> (37)..(37)
          <223> n es a, c, g, o t
30
          <400> 223
          gctggcacgt ctccgaganv tnvtnvtnvt nvtnvtnvtt ttgattattg ggggagacg
                                                                                    59
          <210> 224
35
          <211>59
          <212> ADN
          <213> Secuencia artificial
          <220>
40
          <223> sintetizada químicamente
          <220>
          <221> misc_feature
          <222> (19)..(19)
45
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
          <222> (22)..(22)
50
          <223> n es a, c, g, o t
          <220>
          <221> misc feature
          <222> (25)..(25)
55
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
          <222> (28)..(28)
60
          <223> n es a, c, g, o t
          <220>
          <221> misc feature
          <222> (31)..(31)
65
          <223> n es a, c, g, o t
```

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

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 nvtnvtnvtn vtatggatta ttgggggaga	60
	cg	62
55	<210> 227 <211> 65 <212> ADN <213> Secuencia artificial	
60	<220> <223> sintetizada químicamente	
	<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 tnvtnvtnvt nvtnvtnvtn vtnvttttga ttattggggg	60
	agacg	65
50	<210> 228 <211> 65 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente	
	<220> <221> misc_feature <222> (19)(19)	
60	<223> n es a, c, g, o t	
	<220> <221> misc_feature <222> (22)(22)	

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

	<400> 230	
	gctggcacgt ctccgagadv tdvtdvtdvt dvtdvtdvtd vtdvtatgga ttattggggg	60
	agacg	65
5		
	<210>231	
	<211> 68 <212> ADN	
	<213> Secuencia artificial	
10	12 102 Coccondia artificial	
	<220> <223> sintetizada químicamente	
15	<400> 231	
.0	gctggcacgt ctccgagadv tdvtdvtdvt dvtdvtdvtd vtdvtdvttt tgattattgg	60
	gggagacg	68
	<210> 232	
	<211>68	
20	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> sintetizada químicamente	
25	<400> 232	
	gctggcacgt ctccgagadv tdvtdvtdvt dvtdvtdvtd vtdvtdvtat ggattattgg	60
	gggagacg	68
00	040, 000	
30	<210> 233 <211> 26	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> sintetizada químicamente	
	<400> 233	
	ccggtgtagc gaaggcgtct cagcag 26	
40		
	<210>234	
	<211> 32 <212> ADN	
	<213> Secuencia artificial	
45		
	<220>	
	<223> sintetizada químicamente	
	<400> 234	
50	tagggtcgcc ttgatcgtct cccgaaggtc gg 32	
	<210> 235	
	<211> 45	
55	<212> ADN <213> Secuencia artificial	
	CE 107 Goodonida artinolar	
	<220>	
	<223> sintetizada químicamente	

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

```
<223> sintetizada químicamente
          <220>
          <221> misc_feature
 5
          <222> (17)..(18)
          <223> n es a, c, g, o t
          <221> misc_feature
10
          <222> (20)..(21)
          <223> n es a, c, g, o t
          <220>
          <221> misc feature
15
          <222> (23)..(24)
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
20
          <222> (26).. (27)
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
25
          <222> (29)..(30)
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
30
          <222> (35)..(36)
          <223> n es a, c, g, o t
          <400> 237
          gaaggegtet cageagnnsn nsnnsnnsnn seegnnsace ttegggagae g
                                                                                  51
35
          <210> 238
          <211>32
          <212> ADN
          <213> Secuencia artificial
40
          <220>
          <223> sintetizada químicamente
          <400> 238
45
                                                         32
          cggtcagtcg caatacgtct ccagcatggg at
          <210> 239
          <211>32
          <212> ADN
50
          <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
          <223> sintetizada químicamente
65
          <400> 240
```

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

```
<212> ADN
          <213> Secuencia artificial
          <220>
 5
          <223> sintetizada químicamente
          <220>
          <221> misc feature
          <222> (18)..(19)
10
          <223> n es a, c, g, o t
          <221> misc_feature
          <222> (21)..(22)
15
          <223> n es a, c, g, o t
          <220>
          <221> misc feature
          <222> (24)..(25)
20
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
          <222> (27)..(28)
25
          <223> n es a, c, g, o t
          <220>
          <221> misc feature
          <222> (30)..(31)
30
          <223> n es a, c, g, o t
          <400> 243
          cgtctccagc atgggatnns nnsnnsnnsn nsgtgttgag acgatcctc
                                                                           49
35
          <210> 244
          <211>49
          <212> ADN
          <213> Secuencia artificial
40
          <223> sintetizada químicamente
          <220>
          <221> misc_feature
45
          <222> (18)..(19)
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
50
          <222> (21)..(22)
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
55
          <222> (24)..(25)
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
60
          <222> (27)..(28)
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
65
          <222> (30)..(31)
          <223> n es a, c, g, o t
```

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

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

```
<221> misc feature
          <222> (27)..(28)
          <223> n es a, c, g, o t
 5
          <400> 249
          cgtctcgaac atgggatnns nnsnnsnnsg tgttgagacg atcctc
                                                                           46
          <211>46
          <212> ADN
10
          <213> Secuencia artificial
          <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
          cgtctcgaac atatgatnns nnsnnsnnsg tgttgagacg atcctc
                                                                           46
          <210> 251
40
          <211>49
          <212> ADN
          <213> Secuencia artificial
          <220>
45
          <223> sintetizada químicamente
          <220>
          <221> misc_feature
          <222> (18)..(19)
50
          <223> n es a, c, g, o t
          <220>
          <221> misc feature
          <222> (21)..(22)
55
          <223> n es a, c, g, o t
          <220>
          <221> misc_feature
          <222> (24)..(25)
60
          <223> n es a, c, g, o t
          <220>
          <221> misc feature
          <222> (27)..(28)
65
          <223> n es a, c, g, o t
```

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

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

	<223> sintetizada químicamente
5	<400> 255 ctcttctgag atgagttttt g 21
	<210> 256 <211> 43 <212> ADN <213> Secuencia artificial
10	<220> <223> sintetizada químicamente
15	<400> 256 atgcggccca gccggccatg gccsaggtyc agctbcagca gtc 43
20	<210> 257 <211> 43 <212> ADN <213> Secuencia artificial
	<220> <223> sintetizada químicamente
25	<400> 257 atgcggccca gccggccatg gcccaggttc acctgcagca rtc 43
30	<210> 258 <211> 43 <212> ADN <213> Secuencia artificial
35	<220> <223> sintetizada químicamente
00	<400> 258 atgcggccca gccggccatg gcccaggtrc agctgaagga gtc 43
40	<210> 259 <211> 43 <212> ADN <213> Secuencia artificial
45	<220> <223> sintetizada químicamente
	<400> 259 atgcggccca gccggccatg gcccaggtcc aactvcagca rcc 43
50	<210> 260 <211> 43 <212> ADN <213> Secuencia artificial
55	<220> <223> sintetizada químicamente
60	<400> 260 atgcggccca gccggccatg gcccagatcc agttggtvca gtc 43
	<210> 261 <211> 43 <212> ADN <213> Secuencia artificial
65	<220>

	<223> sintetizada químicamente	
5	<400> 261 atgcggccca gccggccatg gcccaggtgc agctgaagsa stc	43
	<210> 262 <211> 43 <212> ADN <213> Secuencia artificial	
10	<220> <223> sintetizada químicamente	
15	<400> 262 atgcggccca gccggccatg gccgaggtgc agskggtgga gtc	43
20	<210> 263 <211> 43 <212> ADN <213> Secuencia artificial	
	<220> <223> sintetizada químicamente	
25	<400> 263 atgcggccca gccggccatg gccgaagtga arsttgagga gtc	43
30	<210> 264 <211> 43 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintetizada químicamente	
00	<400> 264 atgcggccca gccggccatg gccgakgtsv agcttcagga gtc	43
40	<210> 265 <211> 43 <212> ADN <213> Secuencia artificial	
45	<220> <223> sintetizada químicamente	
	<400> 265 atgcggccca gccggccatg gccgaggtga asstggtgga atc	43
50	<210> 266 <211> 43 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintetizada químicamente	
60	<400> 266 atgcggccca gccggccatg gccgaggtga agctgrtgga rtc	43
	<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 agcttctcsa gtc	43	
30	<210> 270 <211> 42 <212> ADN <213> Secuencia artificial		
35	<220> <223> sintetizada químicamente		
00	<400> 270 atgcggccca gccggccatg gcccargtta ctctgaaaga gt	42	
40	<210> 271 <211> 41 <212> ADN <213> Secuencia artificial		
45	<220> <223> sintetizada químicamente		
	<400> 271 cctgaaccgc cgcctccgct cgagacggtg accgtggtcc c	41	
50	<210> 272 <211> 41 <212> ADN <213> Secuencia artificial		
55	<220> <223> sintetizada químicamente		
60	<400> 272 cctgaaccgc cgcctccgct cgagactgtg agagtggtgc c	41	
	<210> 273 <211> 41 <212> ADN <213> Secuencia artificial		
65	<220>		

	<223> sintetizada químicamente		
5	<400> 273 cctgaaccgc cgcctccgct cgagacagtg accagagtcc c		41
3	<210> 274 <211> 41 <212> ADN <213> Secuencia artificial		
10	<220> <223> sintetizada químicamente		
15	<400> 274 cctgaaccgc cgcctccgct cgagacggtg actgaggttc c		41
20	<210> 275 <211> 35 <212> ADN <213> Secuencia artificial		
	<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		
00	<400> 276 ggggcgcagg gacatccgtc accgtctcct c	31	
40	<210> 277 <211> 32 <212> ADN <213> Secuencia artificial		
45	<220> <223> sintetizada químicamente		
	<400> 277 gaggagactg tgagggatgt gccttggccc 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
         gaggagacgg tgacggatgt tccttgaccc ca
                                                     32
 5
         <210> 280
         <211> 14
         <212> PRT
         <213> Homo sapiens
10
         <220>
         <221> MISC FEATURE
         <222> (4)..(4)
         <223> Xaa es lle 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 Gly Leu Leu Gly Glu Thr
                                             5
25
         <210> 281
         <211>13
         <212> ADN
         <213> Secuencia artificial
30
         <220>
         <223> sintetizada químicamente
         <400> 281
         ttactgtgcg aga
                              13
35
         <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
         <223> sintetizada químicamente
55
         <400> 283
         tttctgtgca aga
                              13
         <210> 284
         <211> 13
         <212> ADN
60
         <213> Secuencia artificial
         <223> sintetizada químicamente
65
```

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

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

```
<222> (24)..(25)
         <223> n es a, t, c o g
         <400> 294
 5
         taatgacacg ctctcctctg cnsnngcaga gaaccccggt cccttg
                                                                      46
         <210> 295
         <211>14
         <212> PRT
         <213> Secuencia artificial
10
         <220>
         <223> sintetizada químicamente
15
         <220>
         <221> MISC FEATURE
         <222> (7)..(7)
         <223> X se selecciona entre Ala, Asp, Glu, Gly y Val
20
         <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
25
         <400> 295
                       Tyr Cys Ala Arg Gly Asp Xaa Xaa Val Ser Trp Gly Gln Gly
                                            5
         <210> 296
30
         <211>46
         <212> ADN
         <213> Secuencia artificial
         <220>
35
         <223> sintetizada químicamente
         <220>
         <221> misc_feature
         <222> (22)..(22)
40
         <223> n es a, t, c o g
         <400> 296
         taatgacacg ctctcctctg cngcagagaa ccccggtccc ttggga
                                                                      46
45
         <210> 297
         <211>14
         <212> PRT
         <213> Secuencia artificial
50
         <220>
         <223> sintetizada químicamente
         <220>
         <221> MISC_FEATURE
55
         <222> (7)..(7)
         <223> X se selecciona entre Ala, Asp, Gly y Val
         <400> 297
                       Tyr Cys Ala Arg Gly Asp Xaa Val Ser Trp Gly Gln Gly Thr
60
         <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					
20	<400> 299 cgctttcctc tgcgggggca gaga	24				
20	<210> 300 <211> 8 <212> PRT <213> Secuencia artificial					
25	<220> <223> sintetizada químicamente					
30	<400> 300					
		Ala Lys 1	Gly Asp	Ala Pro 5	Val	Ser
35	<210> 301 <211> 24 <212> ADN <213> Secuencia artificial					
	(210) Goodonola artinolar					
40	<220> <223> sintetizada químicamente					
40	<220>	24				
40 45	<220> <223> sintetizada químicamente <400> 301	24				
	<220> <223> sintetizada químicamente <400> 301 gcgagaggag acgccttcgt ctct <210> 302 <211> 24 <212> ADN	24				
45	<220> <223> sintetizada químicamente <400> 301 gcgagaggag acgccttcgt ctct <210> 302 <211> 24 <212> ADN <213> Secuencia artificial <220>	24				
45	<220> <223> sintetizada químicamente <400> 301 gcgagaggag acgccttcgt ctct <210> 302 <211> 24 <212> ADN <213> Secuencia artificial <220> <223> sintetizada químicamente <400> 302					
45 50	<220> <223> sintetizada químicamente <400> 301 gcgagaggag acgccttcgt ctct <210> 302 <211> 24 <212> ADN <213> Secuencia artificial <220> <223> sintetizada químicamente <400> 302 cgctctcctc tgcggaagca gaga <210> 303 <211> 8 <212> PRT					

Ala Arg Gly Asp Ala Phe Val Ser 1 5

5	<210> 304 <211> 24 <212> ADN <213> Secuencia artificial					
4.0	<220> <223> sintetizada químicamente					
10	<400> 304 gcgagaggag acgagtacgt ctct	24				
15	<210> 305 <211> 24 <212> ADN <213> Secuencia artificial					
20	<220> <223> sintetizada químicamente					
	<400> 305 cgctctcctc tgcctatgca gaga	24				
25	<210> 306 <211> 8 <212> PRT <213> Secuencia artificial					
30	<220> <223> sintetizada químicamente					
	<400> 306					
		Ala Arg Gl	Ly Asp	Glu T	yr Val	Ser
35	<210> 307 <211> 24 <212> ADN	_				
40	<213> Secuencia artificial					
10	<213> Secuencia artificial <220> <223> sintetizada químicamente					
45	<220>	24				
	<220> <223> sintetizada químicamente <400> 307	24				
45	<220> <223> sintetizada químicamente <400> 307 gcgagaggag acgagctcgt ctct <210> 308 <211> 24 <212> ADN	24				
45	<220> <223> sintetizada químicamente <400> 307 gcgagaggag acgagctcgt ctct <210> 308 <211> 24 <212> ADN <213> Secuencia artificial <220>	24				

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

F	<210> 315 <211> 8 <212> PRT <213> Secuencia artificial								
5	<220> <223> sintetizada químicamente								
10	<400> 315								
		Ala 1	Arg	Gly	Asp	Glu 5	Pro	Val	Ser
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 cgctctcctc tgccctagca gaga		24						
35	<210> 318 <211> 8 <212> PRT <213> Secuencia artificial								
40	<220> <223> sintetizada químicamente								
	<400> 318								
45		Ala 1	Arg	Gly	Asp	Gly 5	Ile	Val	Ser
50	<210> 319 <211> 24 <212> ADN <213> Secuencia artificial								
	<220> <223> sintetizada químicamente								
55	<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	2	24						
10	<210> 321 <211> 8 <212> PRT <213> Secuencia artificial								
	<220> <223> sintetizada químicamente								
15	<400> 321								
		Ala I	.ys	Gly	Asp	Gly	Arg	Val	Ser
			1	L				5	
20	<210> 322 <211> 24 <212> ADN <213> Secuencia artificial								
25	<220> <223> sintetizada químicamente								
	<400> 322 gcgagaggag acgacgccgt ctct	2	24						
30	<210> 323 <211> 24 <212> ADN <213> Secuencia artificial								
35	<220> <223> sintetizada químicamente								
40	<400> 323 cgctctcctc tgctgcggca gaga	2	24						
40	<210> 324 <211> 8 <212> PRT								
45	<213> Secuencia artificial <220> <223> sintetizada químicamente								
50	<400> 324								
30		Ala Ai 1	rg (Gly	Asp	Asp 5	Ala	Val	Ser
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 cgctctcctc tgcccaggca gaga		24						
15	<210> 327 <211> 8 <212> PRT <213> Secuencia artificial								
20	<220> <223> sintetizada químicamente								
	<400> 327								
25		Ala 1	Arg	Gly	Asp	Gly 5	Ser	Val	Ser
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								
TO	<400> 329 cgctctcctc tgcggcagca gaga		24						
50	<210> 330 <211> 8 <212> PRT <213> Secuencia artificial								
55	<220> <223> sintetizada químicamente								
	<400> 330								
60		Ala 1	Arg	Gly	Asp	Ala 5	Val	Val	Ser
60	<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						
20	<400> 332 cgctttcctc tgcacaggca gaga	24					
	<210> 333 <211> 8 <212> PRT <213> Secuencia artificial						
25	<220> <223> sintetizada químicamente						
30	<400> 333						
00		Ala Lys 1	Gly A	sp Val	Ser	Val	Ser
35	<210> 334 <211> 24 <212> ADN <213> Secuencia artificial						
35	<211> 24 <212> ADN						
	<211> 24 <212> ADN <213> Secuencia artificial <220> <223> sintetizada químicamente	24					
40	<211> 24 <212> ADN <213> Secuencia artificial <220> <223> sintetizada químicamente <400> 334 gcgagaggag acgggcacgt ctct <210> 335 <211> 24 <212> ADN	24					
40 45	<211> 24 <212> ADN <213> Secuencia artificial <220> <223> sintetizada químicamente <400> 334 gcgagaggag acgggcacgt ctct <210> 335 <211> 24 <212> ADN <213> Secuencia artificial <220>	24					
40 45	<211> 24 <212> ADN <213> Secuencia artificial <220> <223> sintetizada químicamente <400> 334 gcgagaggag acgggcacgt ctct <210> 335 <211> 24 <212> ADN <213> Secuencia artificial <220> <223> sintetizada químicamente <400> 335						
40 45 50	<211> 24 <212> ADN <213> Secuencia artificial <220> <223> sintetizada químicamente <400> 334 gcgagaggag acgggcacgt ctct <210> 335 <211> 24 <212> ADN <213> Secuencia artificial <220> <223> sintetizada químicamente <400> 335 cgetetecte tgecegtgea gaga <210> 336 <211> 8 <212> PRT						

Ala Arg Gly Asp Gly His Val Ser 1 5

```
<210> 337
         <211>24
         <212> ADN
 5
         <213> Secuencia artificial
         <220>
         <223> sintetizada químicamente
10
         <400> 337
         gcgagaggag acgagaccgt ctct
                                           24
         <210> 338
15
         <211>24
         <212> PRT
         <213> Secuencia artificial
20
         <223> sintetizada químicamente
         <400> 338
                 Cys Gly Cys Thr Cys Thr Cys Cys Thr Cys Thr Gly Cys Thr Cys Thr
                                                               10
                 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
                                                          5
         <210> 340
         <211>24
         <212> ADN
40
         <213> Secuencia artificial
         <220>
         <223> sintetizada químicamente
45
         <400> 340
                                           24
         gcgagaggag acgccatcgt ctct
         <210> 341
50
         <211>24
         <212> ADN
         <213> Secuencia artificial
55
         <223> sintetizada químicamente
```

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

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

	<400> 352 gcgaaaggag acgacgtctc t	21				
5	<210> 353 <211> 21 <212> ADN <213> Secuencia artificial					
10	<220> <223> sintetizada químicamente					
	<400> 353 cgctttcctc tgctgcagag a 21					
15	<210> 354 <211> 7 <212> PRT <213> Secuencia artificial					
20	<220> <223> sintetizada químicamente					
	<400> 354					
25		Ala Lys Gly	y Asp	Asp 5	Val	Ser
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 cgctctcctc tgcagcagag a	21				
50	<210> 357 <211> 7 <212> PRT <213> Secuencia artificial					
55	<220> <223> sintetizada químicamente					
	<400> 357					
		Ala Arg Gly 1	Asp	Val 5	Val	Ser
60	<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 cgctttcctc tgccgcagag a	21	
20	<210> 360 <211> 7 <212> PRT <213> Secuencia artificial		
25	<220> <223> sintetizada químicamente		
30	<400> 360	Ala Lys Gly Asp Gl	. Val Sor
		1 5	y vai sei
35	<210> 361 <211> 21 <212> ADN <213> Secuencia artificial		
40	<220> <223> sintetizada químicamente		
	<223> Sintetizada quimicamente		
	<400> 361 gcgaaaggag acgccgtctc t	21	
45	<400> 361	21	
45 50	<400> 361 gcgaaaggag acgccgtctc t <210> 362 <211> 21 <212> ADN	21	
	<400> 361 gcgaaaggag acgccgtctc t <210> 362 <211> 21 <212> ADN <213> Secuencia artificial <220>	21	
	<400> 361 gcgaaaggag acgccgtctc t <210> 362 <211> 21 <212> ADN <213> Secuencia artificial <220> <223> sintetizada químicamente <400> 362		
50	<400> 361 gcgaaaggag acgccgtctc t <210> 362 <211> 21 <212> ADN <213> Secuencia artificial <220> <223> sintetizada químicamente <400> 362 cgctttcctc tgcggcagag a <210> 363 <211> 7 <212> PRT		

Ala Lys Gly Asp Ala Val Ser

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

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:

5

10

15

20

25

30

40

45

50

55

- (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 FR3 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, 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.
- 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.
 - 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 Fokl.

172

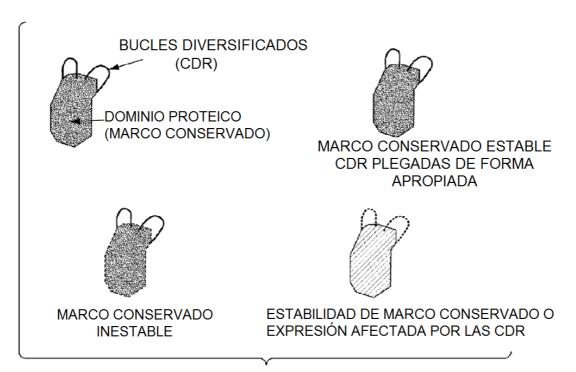


Fig. 1A

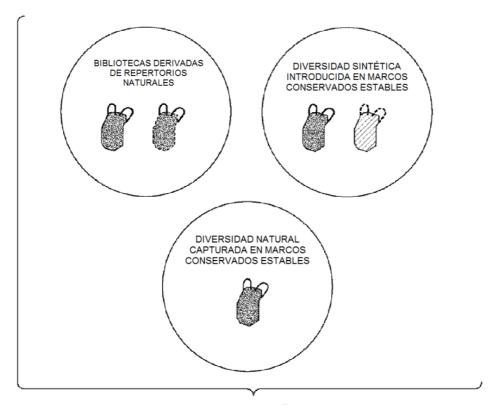
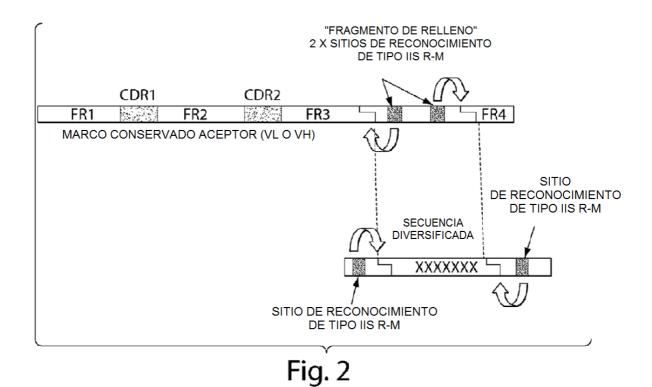
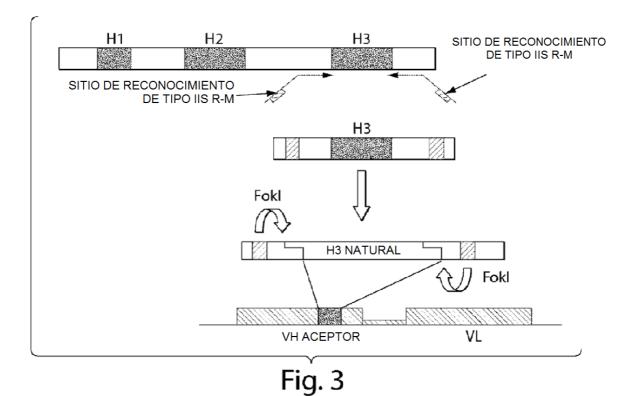
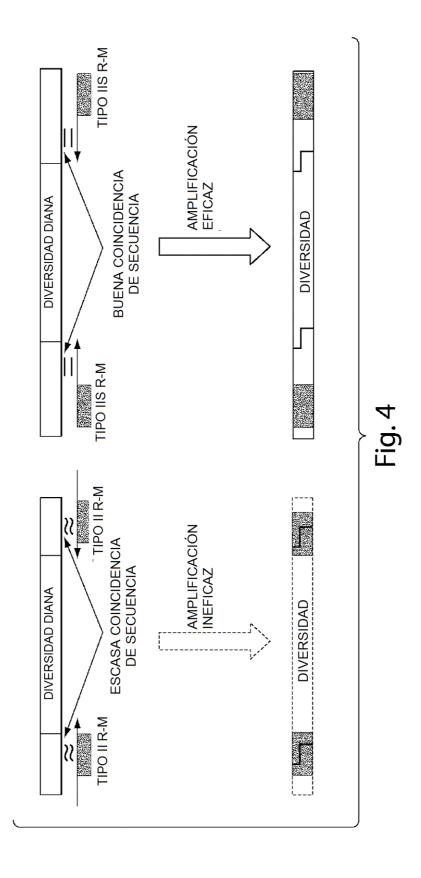


Fig. 1B







```
No: 10
No: 11
                                                                                                                                                                                                                                                                                                                                            N^{\circ}: 12
 [--
                                                                                                                                                                                                                                                                                                                                                                                    N %
....
2 2 2
                                                                                                                                                                                                                   2 2
                                                                               2 % &
                                                                                                                                                                         2
                                                                                                                                                                                                                  88
                                                                                                                                                                                                                                                                                                                                             SEC ID
 888
                                                                                                                                                                                                                  SEC
                                                                                                                                                                                                                                                                                SEC
                                                                                    SEC
SEC
SEC
                                                                                                                                                                                                                                                                                                                                                                                  QSVITQPPSASGTPGQRVTISCSGSSSNIGSNTVNWYQQLPGTAPKLLIYSNNQRPSGVPDRFSGSKSGTSASLAISGLQSEDEADYYCAAWDDSLNG SEC
QSVITQPPSVSAAPGQKVTISCSGSSSNIGNNYVSWYQQLPGTAPKLLIYENNKRPSGIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDSSLSA SEC
QVQLVQSGAEVKKPGASVKVSCKASGYTFT<u>GYYMH</u>WVRQAPGQGLEWMGW<u>INPNSGGTNYAQKFQG</u>RVTMTRDTSISTAYMELSRLRSDDTAVYYCAR SEC
QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGISWVRQAPGQGLEWMGWISAYNGNTNYAQKLQGRVTMTTDTSTSTAYMELRSLRSDDTAVYYCAR SEC
QVQLVQSGAEVKKPGSSVKVSCKASGGTFS<u>SYAIS</u>WVRQAPGQGLEWMG<u>GIIPIFGTANYAQKFQG</u>RVTITADESTSTAYMELSSLRSEDTAVYYCAR SEC
                                                                                                                                                                         EVQLVQSGAEVKRPCESLKISCKGSGYSFTSYWIGNVRQMPGKGLEWMGIIYPGDSDTRYSPSFQGQVTISADKSISTAYLQWSSLKASDTAMYYCAR SEC
                                                                                                                                                                                                                  DIQMTQSPSSLSASVGDRVTITCQASQDIS NYLNWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQYDNLPPTV
DIQMTQSPSSLSASVGDRVTITC<u>RASQSIS SYLN</u>WYQQKPGKAPKLLIY<u>AASSLQS</u>GVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTPPTV
                                                                                                                                                                                                                                                                              SYLAWYQQKPGQAPRILIYDASNRATGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQQKSNWPPTV
SNLAWYQQKPGQAPRILTYGASTRATGIPARFSGSGSGTEFTLTISSLQSEDFAVYYCQQYNNWPPTV
                                                                                                                                                                                                                                                                                                                                             EIVILTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCQQYGSSPPTV
                                                                                    QVQLVESGGGVVQPGRSLRLSCAASGFTFSSYAMHWVRQAPGKGLEWVAVISYDGSNKYYADSVKGRFTISRDNSKNTLYLOMNSLRAEDTAVYYCAR
EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYSMNWVRQAPGKGLEWVSYISSSSSTIYYADSVKGRFTISRDNAKNSLYLOMNSLRAEDTAVYYCAR
EVQLLESGGGLVQPGGSLRLSCAASGFTFS<u>SYAMS</u>WVRQAPGKGLEWVS<u>AISGSGGSTYYADSVKG</u>RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK
                                                                                                                                                                                                                                                                                EIVITOSPATLSLSPGERATLSCRASOSVS
EIVMTOSPATLSVSPGERATLSCRASOSVS
VH1-2 (
VH1-18 (
                                               VH1-69 (
                                                                                 VH3-30 (VH3-48 IVH3-23 I
                                                                                                                                                                                                                VK1-33 I
                                                                                                                                                                                                                                                                                VK3-11 1
VK3-15 1
                                                                                                                                                                       VH5-51
                                                                                                                                                                                                                                                                                                                                                                                                             VL1-51
```

Fig. 5

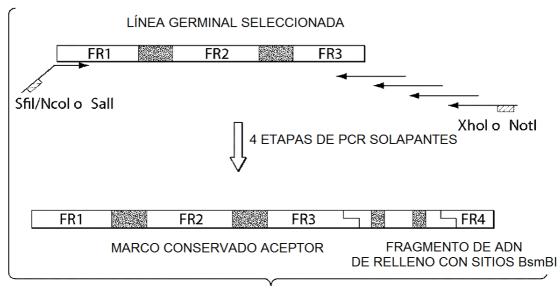
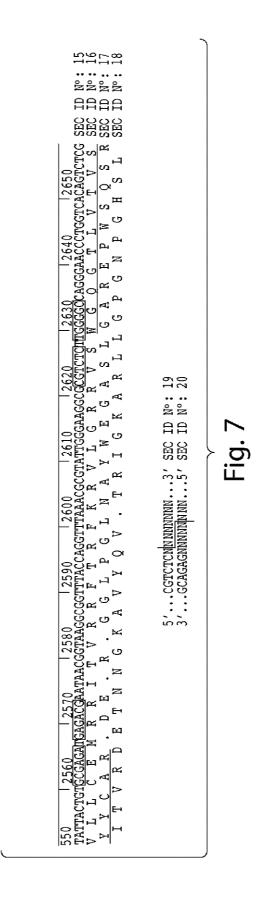


Fig. 6



			m)
			2 2222
			8222 2
	ag 7 × 4	gottgag a e r l g - r l	SEC SEC SEC
	g - 5		. w.a
	£433 1 3 3 1 3 3 1 3 1 3 1 3 1 3 1 3 1 3 1	3ggg	o in the sector of the sector
	octi	ල ය ස ය	ge og g
	0 d d d d d d d d d d d d d d d d d d d	a a a	हु र द
	8° 5" 5"		cctggtcaca gtctcg t 1 v t v s p w s q s p g h s 1
	r: dr t	о · 🦡	±gg_1 ≥ ° : ≥
	gegaci a t c d v r	atg	5 + c = -
	tgggt g	Se 7	agc e e e
	t did	gcctg	19999
	Sact h t h t	4. A	ر م م م م م
	Sy E S	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	ဦး ဗီး ဗေး
	tata y	2 cag	ttggg gccagggaac S W g q g l g a r e l g p g n l g a r e
	acta Y Y	# 주 ·	s L
	ggcta T 1 G N	rcc.	2 s 3 s 5
	p p p	t t t	ਲੂ ਸ਼ਿਲ੍ਹੇ
	ည္ _သ ု ္	accagggaca d q g t r d - p g t t r d	gggaaggcgc gt. V e g a g k a r W e g a
	transfer of the second		5 7 5 E
	ᇦᇃᆠᅆᇉᇕ	gg d	og o
	a cacct	tg accaggga n t r - p g n t r	200 200 300 300 300 300 300 300 300 300
	t	S. = . =	ogtatt or i i i a y i a
	ggat g g	ggtcaccal g h h g s p g s p	gtatt a v i
	244	tca V V	8 de 1
	TT Badie	B 1 5 1	aacgcgtatt k r v n a Y - t r i n a y
		ttcagggcag s g g g f r g f r a .VH1-2	. 영
	c k	5 - 7 - 3 dd	gttt g
	cctgca 1 1 (s c p a VH1-2 s c	SS P III	ccagg t 1 V 4 W1-2.
	ചാ ∾ <u>.</u>	# *	
	agtgaaggtc s v k v q - r s v k v	A B N	# 5 5
	Se o × H	cacagaaa a q b h r	T T T G G G G G G G G G G G G G G G G G
	s e s v v v v v v v v v v v v v v v v v	aca h h	as agge
		5	80 > 1
	a re	actat n y t t	r. gragt
	BO' E.	k act	ataacggt aagg i t v r e - r - n n g k e - r -
	B B B B B B B B B B B B B B B B B B B	8 7 . t	o = to
	5 ๙ 유 : 드	g g	g +
	86 W \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	d w	o car
	e gage	B >	g = ' _ : '
	7 - 1 dag	acagtggi n s g t v t	rie gat
	₽p. 7 :	og e e	g
	9 dag	orth and	\$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$
	عَلِيْ مِ وَأَلِيْ الْمُ		\$7 ± 1
	rggggctg w g a l g a s g a	85 -4 -4	T. L. L.
	დ ფ	# b :	to > ~ : ~
	gtc g s	tgg g	tigt so so
	ggtgcagt g a a w c s	ggat a a a a	ggccgtc g r t a r r p
fr.	£ 5 × 1 × 1	gatgg	4: 4 4 7
AF	\$2° d : d	g × ; ×	8 + :
N	ည်း _{စွေ} တွင်း	gtg v e s	d t d d
[tgcagc g a v q c s		ctgacgac i - r s d d l l t s d d
VH1-2	caggtgcagc tggtgcagtc tggggctgag gtgaagaagc p g a a g a v w g - g e e e q v q l v q s g a e v k k p r c s w c s l g l r - r s >>	ggcttgagtg gatgggatgg atcaaccta acagtggtgg ra - v d g m d q p - q w w g l e w m g w i n p n s g g l e w m g w i n p n s g g l e w m g w i n p n s g	atctgacgac acggccgtgt attactgtgc gagatgagac gas 1
>	8 4 4 €	P. P. V.	면 거, \

Fig. 8

			25 27 28 27 27
			No N
			SEC 1
		gag L Legg	
	5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	agcct e p s s	D 20 - C - C
	ctg P	ട്ടെട്ട	gtctc q s s 1 s 1 >
	2 e e		g + _ t
	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	igotga f a l s - l	S S S S S S S S S S S S S S S S S S S
			cctggtcaca t l v t p w s p g h
		ac atggac 1 h g y m e t w	
	gggt gcgac W V r g c d	a y ctac	g g e e
	igottg	55° a 2: a	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
	1	s t	gcca g g g g
	ta 7	Cgag cache to the transfer of	11
	atggta L w Y g m v	8 4 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ttg = T
	gcta s s	troca L s	ottct S s s
	ge of :	8-4 : 1	gege gtetette rrvs gasl garll
	8 7 : 1	d t t	3000 E
	ccttt	T P BB	agge b. k
	ta caccttt. I h l I v t f v t p v t f	2 + - 2 + - 2 g	gggaaggcgc gtctcttggg gccagggaac cctggtcaca gtctcg l g r r v s w g q g t l v t v s w e g a s l g a r e p w s q s g k a r l l g p g n p g h s l w e g a s l g a r e p w s q s
	g > :_	catg accac th d t m t P - P	aacgcgtatt gggaagg k r v l g r n a y w e - t r i g k n a y w e
	ctggt s g	22 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2	acgedte r r v t r t r n a
	# " " " " " " " " " " " " " " " " " " "	gtc.	acg a se
	о в т	о и .н Бът ::	
		ည္ထိမည့္သိုင္တ	accaggttta as t r f k l p g l y q v - VHI-18
	S Ctg	Cag P P E	Cag t t p TI-II
	2	2	•
	aggte te	Sport :	ggt aaggeggtt t v r f r g g g n g k a v
	2 da a	gcacagaa c t e a q m h r	ය ද ය
	agg s d		aag ' ' ' ' ' ' '
	a a a a	ctat n y t y	
	8 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	a ract	, , , , , , , , , , , , , , , , , , ,
	4 4 4 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	ت ب ب عق	
	o s	8 L	ت ب ق
	gaagaagg	W W g	jàg∂ - e - e
	og a l		the state of the s
	e g	r 1 g n a y n a 1 t r r r r r r r r r r r r r r r r r r	9 d d
	gag • • • •	a: lart	ittactgtgc 1 c 2 y y c c i t v v v c v v v c v v v v v v v v v v v
	GCtg	ggg d g	v. Y.
	56 8 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5		atta Y
	gtgcagtc tg g a v v q s w c s]	56 ± ≥	rtgt a
	5	g a date	a r a r a r a r a r a r a r a r a r a r
AFI	1946 1 9 1 4 1		acggcc h g t r
VH1-18 AF	caggigcagc tggigcagtc tggagctgag gigaagaagc ct p g a a g a v w s - g e e a a q v q l v q s g a e v k k p p r c s w c s l e l r - r s >>	gottgagtg gatgggatgg atcagcgctt acaatggtaa ca g l e w m g w i s a y n g n g l s g w d g s a l t m v t	atctgacgac acggccgtgt attactgtgc gagatgagac gail - r h g r v l l c e m r r r s d d t a v y y c a r - d d l t t r p c i t v r d e t con r s d d t a v y y c a r - d s con r s d d t a v y y c a r - d con r s d d t a v y y c a r - d
<u>-</u>	gtgcagc g a l v g r c s	Jagte	cga d t
11-	ggtg g v g v	acttg	atctgacgac a i - r r s d d d l t t ' r s d d
Z	88 a	8 5 × 5	# 12 ' 1

Fig. 8 Cont.

						31	31
					No.	o o	
						自自	
	B	` ; o	ag e		SE SE	SEC	SEC
	g † g † d d		gcagcctgag e q p e s s l a a -	. m	လ ကို	_ ^	
	P Stage	ь а	တ္တီ တို့ တို့ ထို	62	tct.	s d ⊗	. p.
	8 m	• rcs	සි <u>1</u> 1	:	გ 1	. <u> </u>	-
	20 20 T	т Б	igetga a e 1 s -	o)	cetggtcaca gteteg	ر م	က
	cgac a r	: 🛏	tgga n g m	E S	Ctgc T	თ გ გ	
	cagctgggt gcgacaggcc q 1 g a t g i s w v r q a s a g c d r	-	gcctac atggagctga gcagcctgag s l h g a e q p e a y m e l s s l g p t w s - a a -		ပ္သ	a =	ө Си
	tggg L ™ a g	3	S S	t s t a	gtotottggg gocagggaac v s w g g g	์ มู่ ธา	H
	S sage	· · ·	acag h t	<u>.</u>	cag J q	а в д в	സ
	gcta tc c Y a 1	- ro	ല വ	വ	р. Б. О.	p	b
	tgct " a	*103	aatccacgag cacag riheh estst nprag	. 	ttge s		ა _ დ
	ctat L	<i>S</i>	T T T	യ	ctc.	8 H	က
	tcagc agctatgc	· · · · · · · · · · · · · · · · · · ·	e ag	-	r g	ત્વ	<i>ਵ</i> ਹ
	ည်း ကို	: 4	igac g	ರ	ggogc r r	м 9 д 8 д	p,
	t light	D D	00 to	. " : .	वृष्ठे व	ക്	e B
	gg caccttca r h l g g t f e a p s	57	Ď		1 33	≫	:
	B D	: 57	gatt r	rvtita	ltati v	>- _ ⊱ı	کر خ
	fetgg f	မ	tcac s t v	. >	0 1 1		. ຜ
	ઇ등~	ris.	ਜੂ ਕਰ੍ਹੀ ਜਾ		aga 국		Inay wegasigare p
	c tectgeaagg	; <u>~</u>	naactac gcacagaagt tccagggcag agtcacggacg aatccacgag cacagcctac atggagc k l r t e v p g q s h d y r g r i h e h s l h g l n y a q k f q g r v t i t a d e s t s t a y m e g t t h r s s r a e s r l p r t n p r a q p t w s $\frac{1}{2}$	f g g	ggttta aacgogtatt gggaaggcgc gtctcttggg gccagggaac cctggtcaca gtctc rfkrvlgtvet vswggg gtlvtv	_ 6 _ 8	- 9 9 1 p 9 1
		0 0 0 0	P G G	1 b	cagg I I	다 요성	ြ
	S S	i s	t t	-	t accag f t	T D V	_
	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		aagt e k	n y a g k	gtt	ь ^{>}	5 7
	7 daa	: >	acag t p d	. O	ggcg r	ه م	p,
	gg s	်	96 + 3		da.	ı —	
	Sctc s p	ဟ	otac L	: ~	aacggt aaggog i t v r	ម ⁶¹	ы
	. g . g . g	. b	saaar k k a	. a		- I	
	الم مرث	വ	ag t	<u>ب</u>	gag r	L.	a
	agg e k	<u>:~</u>	ytac % y g v	ာ	agac r	- a	7
	gg a _ i	-24	ttgc 1	44	latg:	ָׁ ס	5 -4
	7 9 gtg	:	tct.	 -	ලිසු ම	정보	• rcs
	gg	о . ко	octa P P	Ω. -⊣	stgc . c	yycar-d itvrdet	уусаг
	जु जु जु जु जु जु जु जु जु जु ज	•	atcc h l] <u>1</u>	acte 1	> P	>
	tgg 1		a de de		att v	>-T	>
	igtc g «	5 '	1999 r 1 9	6 b m a	r r	A 0	
ш	tgca	:>	ggga m g		9000	ซา	ಸ
A	tgg 1	:	gat. « d	: >	acg h	ىد	<u>.</u> +>
69	် ရ	: 57	gtg v e s	െ	gac	יס א	o o
VH1-69 AF	1 4 4 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	ם ס	l l	: —	atctgaggac acggccgtgt attactgtgc gagatgagac gaat $i-g$ h g r v l l c e m r r	⊢ ⊕	rsedtavyycar-de
ΛH	caggtgcagc tggtgcagtc tgggggctgag gtgaagaagc ctgg p g a a g a v w g - g e e a w q v g 1 v q s g a e v k p p r c s w c s 1 g 1 r - r s 1	· 5'	<pre>ggcttgagtg gatgggaggg atcatccta tctttggtac agca r a - v d g r d h p y l w y s g l e w m g g i i p i f g t a g l s g w e g s s l s l v g </pre>	5	atc: i	rsedtavy dlrtrpc >	. _H

Fig. 8 Cont.

			35 35
			8 2 2 8
	A A A A	cctgag	SEC SEC SEC SEC
	10	trgag	മ് യ ്
	999 4 9 1	acagcc e q n s t a	gtctcg v s q s q s z 1 s 1 c 2
	25 d d	acago	gto g
	ه: ها		g + _ :
	B	aatga n q m q m	g > v _ ; v
	gg g a	gg a D	gg * s
	20 20 20 20 20 20 20 20 20 20 20 20 20 2	ctgcaaatga s a n i c k - 1 g m	cctggtcaca t l v t p w s p g h
	igt o	tat v v v v	
	55 ≥ 6 . ≥	gta 1 . c	gccagggaac g q g g p g n a r e
	s s	F 의 및	2 d d i
	gg = ' =	Gagge	ر م م م م
	4 C.		tggg g
	ه <u>ا</u> ه کړي	agaa A r r	£ 1
	*	ള് ഫ് മ	
	ಬ್ಲಿ ಜ	# _ # · · · · ·	gtctcttggg v s w a s l g r l l
	യ് ര	.α.O., •	
	<u> </u>	တ် ့ တြင့္မွ်ပွဲ	ggcgc r r g d
	ttt	ੂ ਦੇ ਦੇ ਕੋੜ੍ਹੇ	rd (1) "(1)
	+	ccagagaca 1 q r s r d s p e t	9999a
	Ö ₩ ° ₩	ъ .	•
	d: 6 7 3 4 5 5 6 7 3 5 5 6 7 5	# # # # # # # # # # # # # # # # # # #	aacgcgtatt k r v n a y - t r i n a y
	5° ₹	28 4 7 6 : 1)	a: La Light
	م الم	tca 	ار الم
	ه الله الله	유 비 : 1	aac K
	a	g	aggttta a:
	24 c 1.23	B. B. H. H. J.	P H D HO
	S VHJ	MI e	P 3-2 P
	S→ %:	ដ្ឋិក > : >	
	F 7 7 4 3	8 Cd 8	## # # : s
	E H G	ਰ ⁺ ਰ ਰ	8. T
		gg	aaggco
	S &	8 +	aag v ' ' '
	ည္ ^စ ု ့	²	r; g r
		act	8
	39999 ₩ 1 9	4; pt 13;	e p i
	E C	ენე დ დ დ დ	8 H .:
		® ' _ > :	igac d d
	tgggggaggc ttggtacagc s g g g l v g l g l v g s g g l v g s g g g g g g g g g g g g g g g g g	gg v s	rdag
	1gt 	19t × 9 × 9.	gagatç a r r d
	# e:	19 N N	р н . В
	0 E	\$ 5 b	
	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	. v . w	tg tg
	3636 5 . 5	Б — : <u> </u>	8 - 7 - 7
	1 8 4 66 m	T A St	Z Z
	s d	க்_ உ _{ல்} பூ	a
Ť.,	tggagtc v g v l e ; w s	Saget S S G	a B B B B B B B B B B B B B B B B B B B
AI	.tg.	ctc	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
γ.	٠ ٠ د	9 4 9 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	å t t t t t t t t t t t t t t t t t t t
\sim	S	igtg ggt	ਰ ਰੂ ਯੂ
7	Cago C G	agt agt	gaggac r g e C
$\ddot{\varsigma}$	16 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Jottagaag a d d g k g l	ggggggggggggggggggggggggggggggggggggg
VH3-23R AF	gaggtgcagc tg r g a a e v g l p r c s >>	ggotggagtg ggtotcagct attagtggta gtggtggtag g a g v g l s V w w - w w - g l e w v s a l s g s g g g w s g s g l l v v v v v g l e w v s a i s g s g g	agccgaggac acggccgtat attactgtgc gagatgagac s r g h g r i l l c e m r r a e d t a v y y c a r - d e p r t r p y i t v r d e >
-	<u>.</u> , ,,	J. J. J.	/ -

Fig. 8 Cont.

	<pre>ggatt cacctttagc agctatgcca tgagctggt ccgccaggct ccagggaagg w i h l - q l c h e l g p p g s r e g f t f s s y a m s w v r g a p g k l d s p l a a m p - a g s a r l q g r g f t f s s y a m s w v r g a p g k g f t f s s y a m s w v r g a p g k</pre>	gccg gttcaccatc tccagagaca attccaagaa cacgctgtat ctgcaaatga acagctgag g p v h h l g r q f q e h a v s a n e q p e g r f t i s r d n s k n t l y l q m n s l a g s p s p e t i p r t r c i c k - t a - 33	trfkrvlgrrvsgetettggggetettggggeteagggaac ettggteaca gteteg SEC ID N°: 37 trfkrvlgrrvsg getettggggeteagggaac ettggteaca gteteg SEC ID N°: 38 trfkrvlgrrvsgeteggaac ettggteaca gteteggrsgetegggaac etggrsgetegggaac etggrsgetegggaac etggggaac etgggaac etgg
	tcctgtgcag cctctgga 1	actecg tgaagggeeg gtteaceate teeagger 1 r e g p v h h 1 g d s v k g r f t i s r t p - r a g s p s p r t t i s r d s v k g r f t i s r d s v k g r f t i s r	
	sc ttggtacagc ctggggggtc cctgagactc r 1 g t a W g v p e t g l v q p g g s l r l a w y s l g g p - d g l v q p g g s l r l g g l v g p g g s l r l	tactac gcagac i 1 r r 1 Y Y a d h t t q Y Y a d	cgtat attactgtgc gaaatgagac gaataacggt aaggcggttt
VH3-23K AF	gaggtgcagc tgttggagtc tggggggaggc ttggtacagc ctggr r g a a v g v w g r l g t a v g v w g r l g t a v g p p r c s c w s l g e a w y s l s s g g g l v g p e v g l l e s g g g l v g p	ggctggagtg ggtctcagct attagtggta gtggtggtag caca g a g v g l s y - w - w w - h g l e w v s a i s g s g g s t g w s g g l l v v v v a s g g s t s g w s g g s t s g s z g g s t s g s g s s t s g s g s s t s g s s s s	agccgaggac acggccgtat attactgtgc gaaatgagac gaars r g h g r i l l c e m r r r a e d t a v y y c a k - d e e p r t r p y i t v r n e t ''' a e d t a v y y c a k - d e r r a e d t a v y y c a k - d e

Fig. 8 Cont.

			7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
		ex	SEC SEC
	8 Ph 1. 7	7. 1 6 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
	55 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	ည္ က က မ	s s
	ရွိတ္တင္းမ	acaç e q	tete
	# m	eg _ :	g 4) _ :
	င္တိုင္က ဗ ဗ	caaatga a n q m c k -	s b caca
	SS T a F1	85 a : 1	cotggtcaca gtctcg t l v t v s p w s q s p g h s l
	200 × ×	cat ctgcae v s a y l q y l c	
	8 8 8 8 G		d d t
	acto h t t	in a sector	66 J 6 J 7
	E B	g	gcca g a g e
	5	K K e K	9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
	E N	attccaagaa c q f q e n s k n i p r n s k n	gtctcttggg gccagggaac v s w g q g a s l g a r e r l l g p g n a s l g a r e
	S L	f too	s 1 s 1 s 1 s 1
	a ag	# 5 5	a r a
	က် ရောင် ရ ရ ရ ရ ရ ရ ရ ရ ရ ရ ရ ရ ရ ရ ရ ရ ရ ရ ရ	tocagagaca 1 q r s r d s p e t s p r d s r d	Jgcgc r r g a
		agaga d 1	a sign
	cac L s	tccac 1 c s p	999939
	a a tt	h h h	gtatt gggaagg v l g r a y w e r i g k a y w e
	2 S S S S S S S S S S S S S S S S S S S	٠: الله الله الله الله الله الله الله الل	ه : ۲ ه رُ رَجُّ
	a P	r f f	a aacgcgtatt k r v l n a y - t r i l n a y
	- E	p.o.	
	gtgcac	tgaagggccg r e g p v k g r 1 aVH3-30	r f t r f k r g l b g l b g l b g l b g l b g l b l b
	s cut of the second sec	aage k k	ccage t r t r y d y d H3-30
	. α ;	t d H tg	R T T THE
	gagactc e t - d I r l) 	3 tt
	gag. - 1	cagact r r a d g t	ည္ႏုတ္ႏွင္း
	s P S	gça r a	gt aaggoog r r r g r r r g r r r r r r r r r
	agtc r g	sctac L l Y Y t	ъ н ъ : н
	gaagg	ract i y y y	
	8° - :	[본 - 국] : 국	و : ۵ و تو
	~ ~	ו מ די די די	ntgagac ga m r r c - d c d e t
	စ္ပြဲသည့္က တ	s v	aga r
	ir rgp gvvggtccago gvvg gvvg	atggaagtaa - w k - d g s m e v d g s	igatga e m r d
	gtg ' ' ' '	d atg	gag a : k
	caggtgcagc tggtggagtc tgggggaggg gtggtccagc p g a a g g v w g r r g p q v q l v e s g g g v v q p r c s w w s l g e a w s s >> q v q l v e s g g g v v q	•	agotgaggac acggotgtgt attactgtgc gagatgagac s - g h g c v l c e m r r a e d t a v y y c a r - d e l r t r l c i t v r d e '
	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	catatg	tgt L
	tgggggaggc	tat ''	ttac
	2 >	g ggtggcagtt atatc	cgtgt attactgtgc c v l l c a v y y c l c i t v a v y y c
	gagte gagte	rtggcagt g g s v a v a v a v a v a	tgtg
ΞŢ	gtggagt g g g v e e v e c	d tag	traget fraget
KŲ.	ا ا ا ا	gg × g	gac acco
30	2003G	8 949	
ω 1	gtgcagc tg g a a I v g l r c s	ggga T	ctgagga s - g a e e
VH3-30 AF	86 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	<u> </u>	9
	(.) ()		

Fig. 8 Cont.

			24 48 74 74
	s r e p g k d d d d d d d d d d d d d d d d d d	acagcctgag e q p e n s l t a -	SEC SEC SEC SEC SEC SEC
	2 g 5 g	o p	ρ, ω . `
	CCagg	တ္ထိုင္း _	S S S
	გ <u> </u>	g g g	to the transfer of the transfe
	ಕ್ಲಿ ಕಿ	g E E	caca v t s
	Signary of the signar	gg Sg Sg	٠ ١ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢ ٢
	Dg La St	ctgcaaat s a n l a l i c k	cctggtcaca gtctcg t l v t v s p w s q s p g h s l
	ΩΩ Λ Λ	т г	
	1.03 to 3.03	gtat C V V	gccagggaac c g q g t a r e g p g n
	ST	±, 0 + 4 <u>4</u>	86 J. J.
	g -	ctcact L t L t t h t h n s	р в д в в в в в в в в в в в в в в в в в
		gird + id	g, b, e, .
	t agctatagca t	aagaa q e k k	gtctttggg gc v s w g a s l g r l l g a s l g
	T ag	® o . : ₩	
	yctata - 1 s y a i	<u>ນ</u> ວ ຮື່ຮ	totett v s s l c l s l
	bg s	atgo q o n n	a r a
	ting s	g g	
	Ö	r r gaga	ည္တို႕ ဗ်ာ္ဂ်ီး ဗ်ာ
	# P + E	CCagg	iggaagg
	the state of the s	22 - 22 - 22 - 22 - 22 - 22 - 22 - 22	gggaaggcgc l g r r w e g g k a
		요 그	
	ter a	t: Pot	1 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
	25 T S S S S S S S S S S S S S S S S S S	Cac Cac	9 L P 2
	cctctggatt s l w i a s g p l d a s g	## T_7 : []	aacgcgtatt k r v n a y - t r i n a y
		9000 9000 9000 9000	accaggttta aacgcgt t r f k r l p g l n a y g v - t r VH3-48
	gtgcag c a c a 3-48	ည္ကို မည <i>္က</i> ်ဆီး မ	ggtt accaggttta $\begin{bmatrix} r & t & r & f \\ g & 1 & p & g & 1 \\ x & v & y & q & v \\ & v & VH3-48 & \dots \\ g & 1 & p & g & 1 \end{bmatrix}$
	ച്ച് മഹ	A A A H	
	troch S S S	tga	acca(
		p 2 2	<u>н</u> ч ,:
	ctgagactc p e t l r l p - d p - d	_: t;	r r r
	g, _ ' :	a r c a	ους πους το
	S P S	ਕੂਜ਼ ਜ਼ਿਲ੍ਹ	аад к
	9 > B	2 × ×	gt aaggegg v r r r - g g k a r - g
	3 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	# # # E	8 +
	55 5 5	ate in	atagataga
	ctggc	tga:	gaat r e
	ည်း ဦး	g > :	တ် တြင့်မှု
	, t	agtac - Y s v	gatgagac e m r r - d c d e
	56 8	gtagta s s v v	gatg r r
	# " " " " " " " " " " " " " " " " " " "	gta s v	9 8 1 . 6 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
	ည္ ဗ ဗ	g > :,	cactgtgc gagat 1 1 c e y c a r i t v r c y c a r
	<u>6</u> 5 ° 5 ° 5 ° 5 ° 5 ° 5 ° 5 ° 5 ° 5 ° 5	agt.	gtgc c c
	85 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	agt.	act 1 X
	c tgggggaggg v w g r s g g g g l g e g	#4	# 57 : >
	8 >		acgotigt attactific gar h g c v l l c a t a v y y c a t r l c l t v r r d t a v y y c a d t a v y y c a d t a v y y c a
	agtc g v e s	8; p. 1.24	tgtg
Щ	4 d d d d d d d d d d d d d d d d d d d	ti _ b tic	dct a d d
AF	್್ ಇಡಿ	9 8 9 8 8 9 9 8 9 9 9 9 9 9 9 9 9 9 9 9	5 4 7 +
∞	ည္ ၂ က : _	5.5	ວ ີ ອຸ ູວ
4	tgcago	9 9 9	1 dd
m	9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	# ± 438	gccgag S r s r e p
VH3-48	gaggtgcagc tggtggagtc ttggtacagc r g a a g g v w g r l g t e v g l v e s g g g l v q p r c s w w s l g e a w y s s y g l v e s g g g l v g e v g l v e s g g g l v g	ggotggagtg ggtttcatac attagtagta gtagtagtac c g a g v g f i h Y g l e w v s y i s s s s s t g w s g f h t l v v v v v v v v v v v v v v v v v v	agocgaggac acggctgtgt attactgtgc gagatgagac s r g h g c v l l c e m r r a e d t a v y y c a r - d e p r t r l c i t v r d e >
	- had /		

Fig. 8 Cont.

VH5-51 AF 9aggtpoagt tyggagagte tygagagagte tetgaagate tetgaagag tetegaata cagettace agetactgga teggetaggy tygecagata energy of a refer of a g a v w s r g e k a r g v s e d l l - g f w i g l y g l l d r l g a p d a re 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 g m p g k p r c s w c s l e g r - k s p g s l - r s p v r v l d t a l p a t g s g a g c a r c p g k p r c s v g l v g l v g 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 g m p g k e v g l v g l v g l v g l v l p r p g k l s r s v l g w v r g m p g k l s v d g d h l s w - l - V g i q p v l p r p g h l s r g v h g 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 g g v t i s a d k s i s t a y l g w s s l l v v v s v g g g t l v t v s s r l l v t v s s r l l v t v r r i t v r r i t v r r f t r f k r v l g r r v s w g g g t l v t v s s s l l l v r v r d e t n n g k a v l r r r g v r r g r v r r g r r g v r r g r r g r r g r r g r r g r r g r				49 50 51 51 51
cogggaage tectgaagate tectgaaga gitetggata cagettace agetactgga teggetgggt gegecaging a region of the region of th				
coggggagt tertgaagate teetgaagg gitetggata eageittace agetaetgga teggetgggt gegecagatg eeegggaaag a r g v s e d l l - g f w i g l y q l l d r l g a r d m p g k 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 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 taccagatac ageccotect tecaaggea ggtcaccate teagecgaca agtcactag caccgetac etgcagtgga gcagcctgaa g r s p s f g g v t i s a d k s i s t a y l g w s s l d t r y s p s f g g v t i s a d k s i s t a y l g w s s l d t r y s p s f g g v t i s a d k s i s t a y l g w s s l gaataacggt aaggeggttt accaggitta aacggtatt gggaaggeg gtcttggg gccagggaac cetggtcaca gtctcg sEcrit v r f t r f k r v l g r r v s w g g g t l v t v s SEC r i t v r f t r f k r v l g r r v s w g g g t l v t v s SEC r n n n g k a v y g v t r r f x r f y w e g a s l g a r e p w s g s s c c r g g l p g l n a y w e g a s l g a r e p w s g s s c				
coggggagt tectgaagate tectgaagg gitetggata eagetitace agetaetgga teggetaggt gegecagatg eeegggaa a r g v s e d l l - g f w i q l y q l l d r l g w v r q m p g b g s g s l x i s c k g s g y s f t s y w i g w v r q m p g b g s g s l x i s c k g s g y s f t s y w i g w v r q m p g g s g s l x i s c k g s g y s f t s y w i g w v r q m p g g t x coagatac agecegtect tecaaggeca ggteaceate teagecgaea agtecateag caecegectae etgeaggag geagectgg i p d v l p r p g h h l s r g v h q h r l p g w s g l g g v t i s a d k s i s t a y l g w s s l d t r y s p s f g g v t i s a d k s i s t a y l g w s s l d t r y s p s f g g v t i s a d k s i s t a y l g w s s l d t r y s p s f g g v t i s a d k s i s t a y l g w s s l d t r y s p s f g g v t i s a d k s i s t a y l g w s s l d t r y s p s f g g v t i s a d k s i s t a y l g w s s l gaataacggt aaggeggtt accaggttta aacgegtatt gggaaggegg gtetttggg gecagggaac eetggteaca gteteg r i t v r f k r v l g r r v s w g g g t l v t v s e n n g k a v y g v - t r i g k a r l g g a r e p w s g s t n n g w v y g g v - t r i g k a r l g a r e p w s g s s l n n a y w e g a s l g a r e p w s g s s				
coggggagt totgaagate teetgaagg gitetggata cagetiace agetactgga teggetgggt gegecagatg eeeggaa a r g v s e d l l - g f w l q l y q l l d r l g w v r q m p g p g s l - r s p v r v l d t a l p a t g s g g v r r r r p g g g v r r v l d r s y w i g w v r q m p g 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 t ceagatac agecegtect tecaaggeca getcaecate teagecgaca agtecateag caecagectae ctgcagtgga geagectu y g i g p v l p r p g h h l s r q v h g h r l p a v e q p t c s g s i p d v r i s a d k s i s t a y l q w s s i v r y s p s f q g q v r i s a d k s i s t a y l q w s s s i t r y s p s f q g q v r i s a d k s i s t a y l q w s s s i t r v r r f k r v l g r r v s w g q g r l v r v s e - r - g g l p g l n a y w e g a s l g a r e p w s g s t n n a y w e g a s l g a r e p w s g s t n n a g w w e g a s l g a r e p w s g s s e - r - g g l p g l n a y w e g a s l g a r e p w s g s		B. ~ ~ ~ ~	g e ^ .	
coggggagt tetgaagate teetgaagg gitetggata eageillace ageilacigga teggetgggt gegecagatg eeegg a r g v s e d l l - g f w i q l y q l l d r l g a p d a r p p g s g y s f t s y w i g w v r q m p 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 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 t 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 t g v d i g p v l p r p g h h l s r q v h q h r l p a v e g d t r y s p s f q g v t i s a d k s i s t a y l g w s s i p d t r y s p s f q g v t i s a d k s i s t a y l g w s s d t r y s p s f q g v t i s a d k s i s t a y l g w s s d t r y s p s f q g v t i s a d k s i s t a y l g w s s d t r r y s p s f q g v t i s a d k s i s t a y l g w s s d t r r y r r f k r v l g r r v s w g q g t l v t v r r r t t r f k r v l g r r v s w g q g t l v t v r r r r t r r r r r r r r r r r r			£ 6	PTS ESS
coggggagtc tctgaagatc tcctgtaagg gttctggata cagctttacc agctactgga tcggctgggt gcgccagatg a r g v s e d l l - g f w i q l y q l l d r l g a p d p d p g s l - r s p v r v l d t a l p a t g s a g v r q m 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 taccagatac agcccgtcct tccaaggcca ggtcaccatc tcagccgaca agtccatcag caccgcctac ctgcagtgga d f r y s p s f g g v t i s a d k s i s t a y l g w v r d w r l r y s p s f g g v t i s a d k s i s t a y l g w gaataacggt aaggcgttta aacgcgtatt gggaaggcg gtctcttggg gccagggaac cctggtcaca r i t v r r f t r f k r v l g r r v s w g g g t l v t e - r - g g l p g l n a y w e g a s l g a r e p w s t n n g k a v y g v - t r i g k a r l l g p g n p g h c - r - g g l p g l n a y w e g a s l g a r e p w s		£		Ç s d ⊗ s
coggggagtc tctgaagatc tcctgtaagg gttctggata cagctttacc agctactgga tcggctgggt gcgccagatg a r g v s e d l l - g f w i q l y q l l d r l g a p d p d p g s l - r s p v r v l d t a l p a t g s a g v r q m 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 taccagatac agcccgtcct tccaaggcca ggtcaccatc tcagccgaca agtccatcag caccgcctac ctgcagtgga d f r y s p s f g g v t i s a d k s i s t a y l g w v r d w r l r y s p s f g g v t i s a d k s i s t a y l g w gaataacggt aaggcgttta aacgcgtatt gggaaggcg gtctcttggg gccagggaac cctggtcaca r i t v r r f t r f k r v l g r r v s w g g g t l v t e - r - g g l p g l n a y w e g a s l g a r e p w s t n n g k a v y g v - t r i g k a r l l g p g n p g h c - r - g g l p g l n a y w e g a s l g a r e p w s		7. C	လုံ့ မှ လုံင်း လုံ့ မှ လုံင်း	7; s.g. c.g.
coggggagtc tctgaagatc tcctgtaagg gttctggata cagctttacc agctactgga tcggctgggt gcgcc a r g v s e d l l - g f w i g r s r s y v r v l d t a l p a t g s a g c c e p g s l k i s c k g s g y s f t s y w i g w v r v r v l s c k g s g y s f t s y w i g w v r taccagatac agcocgtcct tccaaggcca ggtcaccatc tcagccgaca agtccatcag caccgcctac ctgcs y g i g p v l p r p g h h l s r g v h g h r l p e d t r y s p s f g g v t i s a d k s i s t a y l i p d t r y s p s f g g v t i s a d k s i s t a y l g t r y s p s f g g v t i s a d k s i s t a y l g t r r y s p s f g g v t i s a d k s i s t a y l g t r r v s w g g g t l e - r - g g l p g l n a y w e g a s l g a r e p w t n n g k a v y g v - t r i g k a r l l g p g n p t n n g k a v y g v - t r i g k a r l l g p g n p t w e g a s l g a r e p w v v y g v - t r i g k a r l l g g r r v v v v v v v v v v v v v v v v		Bo ■ : ■		
coggggagtc tctgaagatc tcctgtaagg gttctggata cagctttacc agctactgga tcggctgggt gcgcc a r g v s e d l l - g f w i g r s r s y v r v l d t a l p a t g s a g c c e p g s l k i s c k g s g y s f t s y w i g w v r v r v l s c k g s g y s f t s y w i g w v r taccagatac agcocgtcct tccaaggcca ggtcaccatc tcagccgaca agtccatcag caccgcctac ctgcs y g i g p v l p r p g h h l s r g v h g h r l p e d t r y s p s f g g v t i s a d k s i s t a y l i p d t r y s p s f g g v t i s a d k s i s t a y l g t r y s p s f g g v t i s a d k s i s t a y l g t r r y s p s f g g v t i s a d k s i s t a y l g t r r v s w g g g t l e - r - g g l p g l n a y w e g a s l g a r e p w t n n g k a v y g v - t r i g k a r l l g p g n p t n n g k a v y g v - t r i g k a r l l g p g n p t w e g a s l g a r e p w v v y g v - t r i g k a r l l g g r r v v v v v v v v v v v v v v v v		జ్ఞా చా : రా	3. s. a.	S S S
coggggagtc tergaagate teetqtaagg gitetggata cagetitaee agetaetgga teggetgggt go a r g v s e d 1 1 - g f w i q 1 y q 1 1 d r 1 g w v v p g s s 1 - r s p v r v l d t a 1 p a t g s a g c c w g s g y s f t s y w i g w v v v v v i s c k g s g y s f t s y w i g w v v v q i g p v l p r p g h h l s r g v h g h r l p d t r y s p s f g g v t i s a d k s i s t a y t d t r y s p s f g g v t i s a d k s i s t a y t d t r y s p s f g g v t i s a d k s i s t a y d t r r r f t r f k r v l g r r v s w g g g t t n n g k a v y g v - t r i g k a r l l g p g n l t t n n g k a v y g v - t r i g k a r l l g p g n l e r r c g g l p g l n a y w e g a s l g a r e p t n n g k a v y g v - t r i g k a r l l g p g n l e r r c g g l p g l n a y w e g a s l g a r e p t n n g k a v y g v - t r i g k a r l l g g a r e p t n n g k a v y g v - t r i g k a r l l g a r e p t n n g k a v y g v - t r i g k a r l l g a r e p t n n g k a v y g v - t r i g k a r l l g a r e p t n n g k a v y g v - t r i g k a r l l g a r e p t n n g k a v y g v - t r i g k a r l l g a r e p t c n n g k a v y g v - t r i g k a r l l g a r e p t c n n g k a v y g v - t r i g k a r l l g a r e p t c n n g k a v y g v - t r i g k a r l l g a r e p t c n n g k a v y g v - t r i g k a r l l g a r e p p t c n n g k a v y g v - t r i g k a r l l g a r e p p t n l s c r c - g g l p g l n a w w e g a a s l g a r e p p t n l c r c - g g l p g l n a w w e g a a s l g a r e p p t n l c r c - g g l p g l n a w w e g a a s l g a r e p p t n l c r c - g g l p g l n a w w e g a a s l g a r e p p c p p c p p c p c p p c p c p p c p c p p c p c p p c p		છું _ન જ : ન	g e	£
coggggagtc tectgaaggt tectgaagg gitetggata cagetiace agetactgga teggetggggaar r g v s e d 1 1 - g f w 1 q 1 y q 1 l d r 1 g w s g s g v s f t s y w i g w v vH5-51 r s p v r v l d t a l p a t g s a g g v s f t s y w i g w vaccagatac agecegicet tecaaggeca ggicaccate teagecgaca agicaccateag caccgectac accagatac agecegicet tecaaggeca ggicaccate teagecgaca ageceateag caccgectac accagatac agecegic tecaaggeca ggicaccate teagecgaca ageceateag caccgectac r v s p s f g g q v t i s a d k s i s t a g v d t r y s p s f g g q v t i s a d k s i s t a gaataaccggt aaggeggitt accaggitta aacgegatt gggaaaggege gitettiggg gecagggaac r i t v r f t r f k r v l g r r v s w g g g e - r - g g l p g l n a y w e g a s l g a r e t n n g k a v y q v - t r i g k a r l l g p g r e - r - g g l p g l n a y w e g a s l g a r e e - r - g g l p g l n a y w e g a s l g a r e			\$ aT . :	· B. L. C. C.
ccggggagtc tctgaagatc tcctgtaagg gttctggata cagcttacc agctactgga a r g v s e d l l s c k g s g v s f t s v w g s g s l r s v w r v l d t a l p a t g p v r v v l d t a l p a t g p v r v v l d t s r s v w r a d t r y s p s f g g v t i s a d k s i s d t r y s p s f g g v t i s a d k s i gaataacggt aagccgttta aacgcgtatt gggaaggcg gtcttggg r i t v r f t r f k r v l g r v s v s w t n g k a v y g v v t i g k a r l l s r t n n g k a v y g v v t r i g k a r l l s e r r g g l p g l n a y w e g a s l g t n n g k a v y g v v v r r i g r r v s w s c r r r g l b g l n a y w e g a s l g e r r r g l p g l n a y w e g a s l g e r r g l p g l n a y w e g a s l g e r r g l p g l n a y w e g a s l g		E 50.	일 : >	2 <u></u>
ccggggagtc tctgaagatc tcctgtaagg gttctggata cagcttacc agctactgga a r g v s e d l l s c k g s g v s f t s v w g s g s l r s v w r v l d t a l p a t g p v r v v l d t a l p a t g p v r v v l d t s r s v w r a d t r y s p s f g g v t i s a d k s i s d t r y s p s f g g v t i s a d k s i gaataacggt aagccgttta aacgcgtatt gggaaggcg gtcttggg r i t v r f t r f k r v l g r v s v s w t n g k a v y g v v t i g k a r l l s r t n n g k a v y g v v t r i g k a r l l s e r r g g l p g l n a y w e g a s l g t n n g k a v y g v v v r r i g r r v s w s c r r r g l b g l n a y w e g a s l g e r r r g l p g l n a y w e g a s l g e r r g l p g l n a y w e g a s l g e r r g l p g l n a y w e g a s l g		BH ≥	5 8 a	ဗ္ဗိတ္တိတ 🚼
ccggggagtc tctgaagatc tcctgtaagg gttctggata cagcttacc agctactgga a r g v s e d l l s c k g s g v s f t s v w g s g s l r s v w r v l d t a l p a t g p v r v v l d t a l p a t g p v r v v l d t s r s v w r a d t r y s p s f g g v t i s a d k s i s d t r y s p s f g g v t i s a d k s i gaataacggt aagccgttta aacgcgtatt gggaaggcg gtcttggg r i t v r f t r f k r v l g r v s v s w t n g k a v y g v v t i g k a r l l s r t n n g k a v y g v v t r i g k a r l l s e r r g g l p g l n a y w e g a s l g t n n g k a v y g v v v r r i g r r v s w s c r r r g l b g l n a y w e g a s l g e r r r g l p g l n a y w e g a s l g e r r g l p g l n a y w e g a s l g e r r g l p g l n a y w e g a s l g		, m	£ 1 2 3	agg
ccggggagtc tctgaagatc tcctgtaagg gttctggata cagcttacc agctactgga a r g v s e d l l s c k g s g v s f t s v w g s g s l r s v w r v l d t a l p a t g p v r v v l d t a l p a t g p v r v v l d t s r s v w r a d t r y s p s f g g v t i s a d k s i s d t r y s p s f g g v t i s a d k s i gaataacggt aagccgttta aacgcgtatt gggaaggcg gtcttggg r i t v r f t r f k r v l g r v s v s w t n g k a v y g v v t i g k a r l l s r t n n g k a v y g v v t r i g k a r l l s e r r g g l p g l n a y w e g a s l g t n n g k a v y g v v v r r i g r r v s w s c r r r g l b g l n a y w e g a s l g e r r r g l p g l n a y w e g a s l g e r r g l p g l n a y w e g a s l g e r r g l p g l n a y w e g a s l g		ੂੰ ਜਾਂਦਾ ਦੇ ਉੱਦਾ ਜਾਂਦਾ ਹੈ। ਜਾਂਦਾ ਜਾਂਦਾ ਜਾਂਦਾ ਜਾਂਦਾ ਹੈ।	29 g : g	25 2 a L i a
cegggagtc tetgaagate teetgtaagg gitetggata cagetitace agetaetga a r g v s e d l l - g f w i q l y q l l p g e s l k i s c k g s g y s f t s y v v vH5-51. p g e s l k i s c k g s g y s f t s y v v v v l g s g y s f t s y v v g i g p v l p r p g h h l s r g v h l p r p g h h l s r g v h l p r p g h h l s r g v h l p r p g h h l s r g v h l p r p g v t i s a d k s i l p d t r y s p s f g g v t i s a d k s i g a t r y s p s f g g v t i s a d k s i g a t r y s p s f g g v t i s a d k s i g a t r y s p s f g g v t i s a d k s i g a t r y s p s f g g v t i s a d k s i g a t r y s p s f g g v t i s a d k s i g a t r r f t r f k r v l g r r v s v v r r f t r f k r v l g r r v s v v r r f t r f t r r g v r r g t r r g v r r r f t r r f v r r g r r v s v r r r r r r r r r r r r r r r		g _ b : a	•	ы ы ; ы
coggggagt tetgaagate teetgaagg gitetggata cagetitace ageta a form of the second of the		5 · :	ੂੰ ਜ਼ਰੂ ਜ਼ਰੂ ਜ਼ਰੂ ਜ਼ਰੂ ਜ਼ਰੂ	<u>ე</u> — :
ccgggagtc tctgaagatc tcctgtaagg gttttggata cagctttacc age a r g v s e d l l s c k g s g v s f v g g p v r v l d t a l p p g s l - r s p v r v l d t a l p p v r v l d t s f t s s g s g v s f t s s g s g v s f t s s g s g v s f t s s g s g v s f t s s g s g v s f t s s g s g v s f t s s g s g v s f t s s g s g v s f t s s g s g v s f t s s g s g v s f t s s g s g v s f s s g s g v s f s s g s g v s f s s g s g v s f s s g s g v s f s s g s g v s f s s g s g v s f s s g s g v s f s s g s f g g v s f s s g s f g g v s f s s g s f g g v s f s s f s s f g g v s f s s f s f s s f s f s s f s f s s f s f s f s s f		g-1	္ န လည္သန္တ	ე "⊢: <u>"</u>
ceggggagt tetgaagate teetgaagg gttetggata cagetttace a r g v s e d l l g f w l q l y b g s l k i s c k g s g y s f t l p g s l - r s p v r v l d t a l l p g e s l k i s c k g s g y s f t t l s c k g s g y s f t t l p g e s l k i s c k g s g y s f t t accagatac agcecgtect tecaaggeca ggtcaccate teagcegaca y q i q p v l p r p g h h l s r d i p d t a r p s k a r s p s g p t l p d t r y s p s f q g q v t i s a d d t r y s p s f q g q v t i s a d d t r y s p s f q g q v t i s a d gaataacggt aaggeggtt accaggtta aacgegtatt gggaaggeg r i t v r r f t r f k r v l g r r e - r - g g l p g l n a y w e g t n n g k a v y g q v - t r i g k a e - r - g g l p g l n a y w e g e - r - g g l p g l n a y w e g		ဥက္က _တ ႏအ	prints in	r 4 tr
ccgggagtc tctgaagatc tcctgtaagg gttctggata cagctti a r g v s e d l l g c k g s g y s l p g s l - r s p v r v l d t a p g s l - r s p v r v l d t a taccagatac agcccgtcct tccaaggcca ggtcaccatc tcagccc y g i g p v l p r p g h h l s a i p d t a r p s k a r s p s g g t r y s p s f g g g v t i s a gaataacggt aaggcggttt accaggttta aacgcgtatt gggaagg r i t v r f t r f k r v l g r t n n g k a v y g v - t r i g k t n n g k a v y g v - t r i g k e - r - g g l p g l n a y w e t n n g k a v y g v - t r i g k e - r - g g l p g l n a y w e		•		=
cegggagtc tetgaagate teetgtaagg gttetggata caget a r g v s e d l l e g r w l q p g e s l k l s c k g s g y s p g e s l k l s c k g s g y s p g e s l k l s c k g s g y s taccagatac agccegtect tecaaggeca ggteaceate teage y q i q p v l p r p g h h l s d t r y s p s f q g q v t l s i p d t a r p s k a r s p s g d t r y s p s f q g q v t l s gaataacggt aaggeggttt accaggttta aacgcgtatt gggaa r i t v r f t r f k r v l g t n n g k a v y g v - t r l g t n n g k a v y g v - t r l g e - r - g g l p g l n a y w e t n n g k a v y g q v - t r l g e - r - g g l p g l n a y w e				ည္တို႕ဗုိ ဗု
cegggagtc tetgaagatc teetgtaagg gttetggata caga a r g v s e d l s c k g s g v g g s g v s s l - r s p v r v l d t v v l d t v v l d t v v l d t v s c s g s g v v d d t r v s p s f d g g d v t i s l b d t r v s p s f d g g v t i s l b d t r v s p s f d g g v t i s d t r v s p s f d g g v t i s d t r v s p s f d g g v t i s d t r v s p s f d g g v t i s d t r v s p s f d g g v t i s d t r v s p s f d g g v t i s d t r v s p s f d g g v t i s d t r v s p s f d g g v t i s d t r v s p s f d g g v t i s d t r v s p s f d g g v t i s d t r v r r f t r f k r v l c r r f t r r f r r r r r r r r r r r r r		s: a ct	D 0 0 0	ညီး မြန္ မ
ccggggagtc tctgaagatc tcctgtaagg gttctggata a r g v s e d l s c k g s g v s g s s l - r s p v r v l d d s s l - r s p v r v l d d s s l k i s c k g s g s g s g s s l - r s p v r v l d d s s g s g s g s g s g s g s g s g s		<u> </u>	ος του	
ccgggagtc tctgaagatc tcctgtaagg gttctggaar of s s s s s s s s s s s s s s s s s s		ю <u>-</u>	ਤਜ‴:ਜ	+
ccggggagtc tctgaagatc tcctgtaagg gttct a r g v s l k l s c k g s p v r v v v v v v v v v v v v v v v v v		ga aga aga aga aga aga aga aga aga aga	tint in	t d t
ccgggagtc tctgaagatc tcctgtaagg gtt a r g v s d l l - g g p g s l - r s p v r v p g s l - r s p v r v p g e s l k i s c k g taccagatac agcccgtcct tccaaggcca ggt d t r y s p s f g g g i p d t a r p s f g g g d t r y s p s f g g g t r v r f t r f k e - r - g g l p g l n t n n g k a v y g v - t n n g k a v y g v - e - r - g g l p g l n e - r - g g l p g l n		S S S	gg s	FX H :
ccgggagtc tctgaagatc tcctgtaagg a r g v s l k i s c k p g s l - r s p v r vH3-51 p g e s l k i s c k r k g g g g g g g g g g g g g g g g g		# 5 : 5 · 5	₽ ₀ 7 .	žaž r
ccggggagtc tctgaagatc tcctgt a r g v s l k l s c p g s l - r s p c p g s l - r s p c p g s l k l s c p g e s l k l s c taccagatac agcccgtcct tccaag v g l g p v l p c l p d t a r p s p s f g t r y s p s f g gaataacgt aaggcggtt accagg r l t v r r f t r t n n g k a v g l p t n n g k a v y g t n n g k a v y g e - r - g g l p e - r - g g l p		Press Sud 7	•	
cegggagte tetgaagate teetgaagate teetgaagate teetgaagate s 1 k 1 s p p g s 1 - r s p p g g s 1 k 1 s p g g g g g g g g g g g g g g g g g g		8 5	7 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	
cegggagtc tetgaagatc tc a r g v s e d l b g e s l k i p g e s l k i p g e s l k i taccagatac agcccgtcct tc v g i g p v l d t r y s p s f i p d t a r p r d t r y s p s f i r v r r f r i t v r r f t n n g k a v y t n n g k a v y t n n g k a v y t n n g k a v y t n n g k a v y		5 04	8 7 2 E	agga T T T T T T T T T T T T T T T T T T T
ccggggagtc tctgaaga a r g v s s e p g s s s s s s s s s s s s s s s s s s		()		
ccggggagtc tctgaaga a r g v s s e p g s s s s s s s s s s s s s s s s s s		ე _ი	#	## %:
cegggagtc tctga a r g v s p g e s l - p g e s l - p g e s l taccagatac agcco taccagatac agcco v g i g p d t r y s l d t r y s s d t r y s r d t r y s r e - r - g t n n g k e		g o ~ ` · ~	2° 4:	ညာ မြင်္က ့ ဗ်
ccggggagtc a r g v p g s p g s p g e taccagatac taccagatac t r y i p d i r y i p d t r y t r y i t e - r t n n g			8 H & :	ည္တြင္း ဗ
ccggggagtc a r g v p g s p g s p g e taccagatac taccagatac t r y i p d i r y i p d t r y t r y i t e - r t n n g		·	agg a t	aag ' ' ' ' '
ccggggag a r g p g e p g e p g e taccagat i p d d t r d t r d t r e - t n n		ဌာန တွင်း	Z . A	
· + ° ° ° • .		ညီ မင္းမ	High Tar	
· + ° ° ° • .		Du	ئ ئو ئ	t
		ള് പ്ല	d: Lay	755 <u>1</u> 3
WH5-51 AF gaggtgcagtc tggggcagag gtgaaaaa r g a a g a v w s r g e k e v g l v g s g a e v k b r c s w c s l e g r - k >> v g l v g s g a e v k e v g l v g s g a e v k e v g l v g s g a e v k gcctggagtg gatggggatc atctatcctg gtgactct r p g v d g d h l s w - l g l e w m g i i y p g d s a w s g w g s s i l v t g d s g l e w m g i i y p g d s g l e w m g i i y p g d s r b g v d g d s r c s w c s d t a m y y c a r - r p r t p p c i t v r d e k a s d t a m y y c a r - r p r t p p c i t v r d e r s d t a m y y c a r -				
VH5-51 AF gaggtgcagc tggtgcagtc tggagcagag gtgaa,		8 7 7 : 7 8 4 7 4 : 7	5-4 %	© = 0 : 1
Gaggtgcagtc tggagcagag gtc r g a a g a v w s r g e v g l v g s g a e v g r c s w c s l e g r >> v g l v g s g a e v e v g l v g s g a e v gcctggagtg gatgggatc atctatcctg gtc r p g v d g d h l s w g l e w m g i i y p g a w s g w g s s l l v > v g v d d d h l s w g l e w m g i i y p g a s g w g s s l l v > v g v d g d h l s w g l e w m g i i y p g g l e w m g i i i y p g g l e w m g i i i y p g g l e w m g i i i y p g g l e w m g i i i v p g g l e w m g i i i v p g g l e w m g i i i v p g g l e w m g i i i v p g g l e w			ਰੂ ਰਾਜ਼	H d H
gagggcagc tggtgcagtc tggagcagag r g a a g a v w s r e v q l v q s q a e q e q s e q e v q l v q s q a e e v q l v q s q a e e v q l v q s g a e e v q l v q s g a e e v q l v q s g a e e v q l v q s g a e e ctggagtg gatgggatc atctatcctg r p g v d g d h l s g v d g d h l s a w s g w g s s l l v p e gcctggac accgccatgt attactgtgc g l e w m g i i y p ggcctcggac accgccatgt attactgtgc g l g h r h v l l c k a s d t a m y y c r r p r t p p c i t v k a s d t a m y y c		£2, 7 : 1,	£ 2 5	8 H B BB
yH5-51 AF gaggtgcagc tggtgcagtc tggagcag; r g a a g a v w s e v q a v c s l g a e v q l v q s g a e v q l v q s g a e v q l v q s g a gcctggagtg gatggggatc atctatcc r p g v d g d h l s g l e w m g i i y l g l e w i v l g l e w i v l g l e w i v l g l e w i v l g l e w i v l g l e w i v l g l e w i v l g l e w i v l g l e w i v l g l e w i v l g l e w i v l g l e		တ္မွားစာ မေ	F at : a.	2 ° '
Gaggtgcagc tggtgcagtc tggaggc r g a a g a v w w w g c v g l v q s g l e v q l v q s g g gcctggagtg gatggggatc atctat r p g v d g d h l g l e w m g i i y g l e w m g i i l g g l e w m g i i l g g l e w m g i i l g g l e w m g i i l g g l e w m g i i l g g g gcctcggac accgccatgt attact g l e w m g i i l g l e w m g i i l g g l e w m g i i l g g l e w m g i i l g g l e w m g i i l g g l e w m g i i l g g g l e w m g i i l g g g g g g g g g g g g g g g g		6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	5 6 7 7	# o .
Gagggcagc tggtgcagtc tgg r g a a g a v w e v q l v q s p r c s w c s l >> v q l v q s e v q l v q s e v q l v q s e v q l v q s e v q l v q s gcctggagtg gatggggatc atc r p g v d g d h g l e w m g i i g l e w m g i g l e w m g i g l e w m g i g l e w m g i g l e w m g i g l e w m g i g		တို့ ဗေ	tat L	act T
yH5-51 AF gaggtgcagc tggtgcagtc r g a a g a v e v q l v q p r c s w c s >> v q l v q q r v q q v q q v q l v q q v q l v q q v q l v q q v q l v q q v q l v q q l v q v q d q l v d d d q l v m g i a w s g w g v g l e w m g i g l e w m g i g l e w m g i g l e w m g i r r r r r r r r k a s d t a m r r r r r r r k a s d t a m		۳: ۳ ۳: ۳: ۳	#	att v v v v v v v v v v v v v v v v v v
VH5-51 AF gaggtgcagc tggtgcag r g a a g a e v a l v q p r c s w c >> v q l v q gctggagtg gatggggar r p g v d g a g l e w m g a w s g w g g l e w m g g l e w m g g l e w m g g l e w m g g l e w m g g l e w m g g l e w m g g l e w m g		\$ \$ tt	~ : ~ 등	th = 0 :=
VH5-51 AF gaggtgcagc tggtgg r g a a g e v g l v e v g l v e v g l v e v g l v e v g l v e v g l v gcctggagtg gatggg r p g v d g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m g l e w m		B, a D, . D.		te de de
WH5-51 P gaggtgcagc tgg r r g a e v g l p r c s l s s s gcctggagtg gat r g l e w g l	丏	\$600	É E	ğ, ¹
ydagtgcagc r g a e v g g r c s s y s y c c g gcctggagtg r p g v g v g r g v g r g r s s d g r e s s d g r e s s d r e s d r e s s d r e s s d r e s d r e s d r e s s d r e s d r e s s d r e s d r	Æ	ੜ੍ਹੇ.°	gg ≥ p:≥	36 1 + + + + + + + + + + + + + + + + + + +
yeagatgca r g a r g a r g a r g a r c c gagag r c c g a r g g g r c g g g r c g g r c g g r c g g r c g r c g r c r c	H	ဥ္က ဗွာ ^ဏ ႏွာ	ა :	ರ್ ರ್ಡ್ನ
VHS gaggti r g r g r g gcctg g l g l g l k a r p k a r p k a r p k a r p k a r p k a	 (万	ege s	e dag	လ မ လ
V S S S S S S S S S S S S S S S S S S S	T2	Tie at	Par I	ع: ۵ م تو
······································	VI	8. ~ ~	ద్ది _అ . అ	9g × 4 × 4

Fig. 8 Cont.

	ი გ^	r0 - 4 ∧	55 55 55 55 55
	gcc k a	atta L'.	
	aga p	t: Pt Bet	
	66 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	og T e or e	
	ದ ದ	Б :	SEC SEC SEC
	e a 3 CC	tat d d	
	gcagaga a d c d	e	g
	ဗို့ ဗုိ ဗ	5° -	cgca a a a b r r
	atca v s v v	D P P	a a r
	w w	20 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	
	n i	s l	aaggggcggc k g a k g r g r g g
	<u>~</u>	20 m m	
	ortatttaa e g l f n y l t i - n y l	gcag g g g s	# * * *
	t v	55	igaaal w k w k
	a a c	ب ب ب الم	gtgg.
	gacattagc aactatttaa al g h - g l f k d i s n y l l r t l a t i -	g 41 m	დ.⊶ <u> :</u>
	ggacattagc g h - q d i s r t l		caa t b
	d d d	tac tac	agggac q g k g r C
	p. p. ; c.	₽₩ -a :™	
	aggcgagtca p g e s q a s r r v q a s	o: Toright	cattoggcca of formal can be so a lar be lar
	gagti e r s	d d d	r s
	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	6 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	s ttc
		ρ. τ . ω . τ.	•
	tgcc 1 336.	936.99	a r
	actt	8 1 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	
	at of it.	aga sa	aggegegtet rrv egas kar
	s h v t s	s s	w. ire.
	gtcaco	0 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	cgtattggga r v l g a y w r i g VX 1-33G
	Bron H	r g r	a y K
	ું તે દુ	సో రా∶	
	ρ. υ υ υ υ υ υ υ υ υ υ υ υ υ υ υ υ υ υ υ	ratea P s P s	k k l n l n
	per v		
	ctgtô	9 9 P	ggtt
	o a p	tree Sand	t r p p
	ctgcat v c s a l h	acag r t t	f t l p c l
	gtc c	gaga G G F G G	_ ± _ m = : m
	ਹੈ ਕ	\$44,54	တိုင်္က ဗ
	tcc s	# 5 = = = = = = = = = = = = = = = = = =	agg
	t t	s s	gta r g
	cca L L	gcatccaatt c i g a s n m h p i	- = - + gg
	catcocaga tgacccagtc tccatcctcc i h p d d p v s i l d i q m t q s p s s t s r - p s l h p d i q m t q s p s	p ∰ .	gaat aacggtaagg ccrife - r - g t n n g k 6
r_	og og	rcga Y Y	4 - 4 da
AE	ط ب و طور	۳.: « ۲. <u>ت</u>	lagacg r d e t
r.	tgaccca d d d l m t m t - p	gat I I da	cga r s r
33	gg a : b	o but	් ය ශ්ශ
1	S	2	ည်း တွင်္က တွင်္က လွ
VK1-33G AF	gacatccaga tgacccagtc tccatcctcc ctgt g h p d d p v s i l p d i q m t q s p s s l r t s r - p s l h p p c >>	ctaagctcct gatctacgat gcatccaatt tggaa p - a p d l r c i g f g p k l l i y d a s n l e l s s - s t m h p i w >> t l i y d a s n l e	ctgtcagcag cgagacgaat aacggtaagg cggt l s a a r i t v r r y c g g r d e - r - g g t v s s e t n n g k a >
>	80. 13	5 4 × 4	k, y

Fig. 8 Cont.

	a gggaaagccc r r e s g g k r g g k r		59 60 58
	00 g e	ratta 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
	8 2 3	+: = + = #	
	65 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	caacatatta a t N 1 q h 1 a t V 2 a t V 3	
	ნი" ნ". ო ი⊿ •Ω-	ittg caacate y c n i a t i a t i a t	SEC SEC SEC SEC
	r r e s k b d k d k d k d k d k d k d k d k d k	¥	
	T e gag	r d d	
	a gcagaaacca gg	age o	cgca a a p r v
	a attggtatca gcagaaacca gg k l v s a e t i n w y g g k p - i g i s r n g	الا الله الله الله الله الله الله الله	•
	v v Y	000	aacgggcggck ra a b g r g d d g g d g d g g d g g g g g g g
	55-1 ≥ _3 ≥	250° - 2° - 1	acgggc k r k n g t g
	# = =	S S S	aac T P
	ictatttaa attggta 1 1 f k l v n y l n w t i - i g	ccatcagcag cctgcagcct gaagatath h q q p a a - r y t is s l q p e d p s a c s l k i t i s s l q p e d t i s s l q p e d	gtggaaatca aa v e i l r w k s g g n g
		9 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	tg o × E ×
	T t	p h	gtggaaatca v e i r w k s g g n r w k s
	S 33	8 9	gtggaaatc v e i r w k : g g n r w k
	ago 1	f g	caag gtggaaatca aacc t k v e i k p r w k s n l q g g n q t p r w k s n
	### ##################################	T T	م ب هو ا
	2 dag	the state of the s	26 26 H
	s g d s d c s g d c s d d c d c d d c d d c d d c d d d c d d c d	gat tttactttc r f y f i i i l l i	ca agggacc a k g p r d a k g
	s s	1 d t 1 d d d d d	igcc 1 g
	aggogag d a r r d a	96 9 9 5 8	atteggees for a second in the
	8 7 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	တင္း မွာ	s s l
	atcacttgcc aggcgagtca ggacattagc aactatttaa h h l p g e s g h - g l f i t c g a s g d i s n y l p s l a r r v r t l a t i - VK 1-33R	grggate tgggacagat tttactttea ceateag k w i w d r f f h h g s g t d f t f t i	cgtattggga aggcgcgtct cattcggcca agggaccaag rvlgrrvsfgggg gtk a ywegashsakgp rigkarlirprdg .VKl-33Rsrkgp
	## T # T # T # T # T # T # T # T # T #	tgg;	cgcgtc r v g a r a r g a a
	8 - E	gaagtgg w k w g s e v e v VK 1-33	2000 H
		37.9 9. 4. 4. 4. 4. 4. 4. 4. 4. 4. 4. 4. 4. 4.	cgtattggga agg rvlgr aywer rigk VKl-33R
	tgtctgcat ctgtaggaga cagagtcacc p v c i c r r g s h l s a s v g d r v t c l h l - e t e s l s a s v g d r v t	ည် မိုင္က	ggga 1 g w 1-33
	R G R	ttc g	att. V V V
	رة م ب م يورو	9 2 5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	cgtatt r v a y r i r i a y a y a y
	ggaga r r g G	യ് വ	c r f k r p g l n d q v - t p g l n p
	agga r v g	cccatca y p i v p s s h	aga T I I
	s v	22± ° ° ° ≥ ° 1	### ##################################
	a s s	90 0 B	£ 4 , p
	a g	n r g p i t g v p i g g s h	ttacca ggtttaaacg f t r f k l p g l n v y g v - t
	s s	aac e	
	o : 1	. 1 1	a ggtt
	ο ν :«	원 · · · · · ·	о в в в н
	tectec i l s s h p	scaatt 1 s n i s n	a a a a a a a a a a a a a a a a a a a
		catco	# _ # _ # _ # _ # _ # _ # _ # _ # _ # _
	to tecat	ညီပြီး မြ	gaat aacggtaacri t v e e - r - e t n n g l
	of to	d rat	कः क का
بتا	tgacccagt d d p m t g - p g m t q	tacgat 1 r y d s t	tcagcag cgagacgaat aacggtaagg . s a r r i t v r c q q r d e - r - v s s e t n n g k c q q r d e - r -
VH1-33R AF	gaç 	ad de la	r r r
3,2	ا ا ا	, ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' '	egtcagcag cg
'n	සින් සින් කි. කි. කි.	taagctcct P k 1 P s s	ctgtcagcag of I s a g g g g g g g g g g g g g g g g g g
	# 1 : T	3gC × 8 × ×	c cac
ΛH	gacatccaga tgacccagtc tccatcctcc ctgtct g h p d d p v s i l p v d i g m t g s p s s l s r t s r - p s l h p p c >> d i g m t g s p s s s s d i g m t g s p s s l s	ctaagctcct gatctacgat gcatccaatt tggaag p - a p d l r c i q f g p k l l i y d a s n l e l s s - s t m h p i w k >>	ctgtcagcag cgagacgaat aacggtaagg cggtt 1
-	/\	. — /	

Fig. 8 Cont.

	ം മ^		^ Lø	61 63 64 63
	್ದಿ ಜಿ	ന	acta 1 1 1 4 1 4 1 4 1 7	
	gg o b :	~	‡ + : +	
	9939 4	b	o o o	
	8 D	Ω_{i}	fg T. I.	SEC
	aacce e t k	<u>~</u>	# 	
	E T	ರ್	gaagat r r R r	rgca P a a
	ဗို ဗုိ	ט		cgca a a r r
	atca v s v s	⊳ ₁	b t	1 d d d
	ata M	ja:	88 8 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	gggcggc g a g r r g g
	A Lu	ca ·	totgcars s 1 s 1 c c c c c c c c c c c c c c c	pp 7 x x x
	m l			യ്യ മ
	## :	y	20 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	: P C
	tatt 1 a y	ž s	trag	gaaa w k g w
	aggrange မ	ťΩ	ccatcal h h t p s t i	gtgggggr
	ည္ ဗ	on.	es la sil	2 4 H
	tra i	- - 1	ictete h s t	ccaag t k d p d q
	gcar e]	ťΩ	r Pact	<u> </u>
	g Pr	ਨਾ	¥ :	agga A A A A A A A A A A A A A A A A A A
	tca s v	va	d ad	d d a
	Egy o	rcs	d d d	965 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
	9990 4 9990	ы	996 × 6	cattoggcca s f g h s a l i r p
	ρ.α. 		ე . ა . ა . ა . ა ა ა ა ა ა ა ა	•
	. 55 396.	ပ	18 30 m	a dect
	- Farana	د	8 - 4 a	ည်း မ မ မ
	atc h i VX	1	S & P S S	aggegegtet rrv egas kar kar egas
	8,4 ⁺	ىب	gtg s <	- 1-
	tca V S	>	28 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	cgtattggga r v l g a y w r i g VK 1-39C
	gagi G	5 —4	1 9 t	gtat v r r r.vr a y
	g p	o	8° ≥ :	_
	aga r g e	57	tca i s	taaacg (
	agg v	>	ದ್ದಿ ದ	tttaaa f k g l v _ g
	tgt L	co.	agtco	# 5° : 5
	r r		O. O	t r P q
	် ပို့ ၁ ၁ ၁ ၁	Q	igtgg k w s v	ttacca f t l p v y v y
	tct.	Ω		
	Ctg D		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
	ည္း	ശ	# % > %	B H :
	ಕ್ಷಿ ಕ್ಷಿ	603	S C C C C C C C C C C C C C C C C C C C	igtaag r r - r - r -
	cat P	Ω,	atc	acgg
	್ದ ಜ್ಞ	co.	8,0 7,	aacggtaagg cgg i t v r r - r - g ı n g k a
	gtc g v	ਨਾ	a a a	gaat r r t e
Ŀ	cca t p	ப	tat 1 s y	cgagacgaat a r d e r d e r n r d e r n r d e r n r d e r n r d e r n r d e r d
Fili	ည်း ရောင် မျ	E	2	gagacc a r r d e t
96	क <u>म</u>		5 7 .7	υ υ υ υ υ υ υ υ υ υ υ υ υ υ υ υ υ υ υ
m	ည်း တို့ ၁	ਨਾ	a la	908 G
VK1-39G AF	atc h t l	dig mtg spss 1	gel z s	ctgtcagcag cgagacgaat aacggtaagg cggt 1 s a r r i t v r r y c q q r d e - r - g g t v s s e t n n g k a s >
ζ	<pre>gacatccaga tgacccagtc tccatcctcc ctgf g h p d d p v s i l p d i g m t q s p s s l r t s r - p s l h p p c >></pre>	0	ctaagctcct gatctatgct gcatccagtt tgcaar p - a p d l c c i q f a p k l l i y a a s s l q l s s - s m l h p v c k > p k l l i y a a s s l q	ctgtcagcag cge 1 s a a y c q q I t v s s >

Fig. 8 Cont.

			,
	^		65 67 67
	S & C	は一つかっ	
	ည္သိုင္သ	T Z Z	
	ള്ക്		
	gggaaag r e g k g g j	ည်း မှာ မှာ	
	8 G G	00 :	SEC SEC SEC
	8 - 1 - 1	f a t y t a t y t a t y t a t y	
	g o ×	t ~~ ~	
	g	გე ი _™ : ი	დ დ ∧
	attggtatca gcagaaacca gggaaagccck I v s a e t r e s n w y q q k p g k a i g i s r n q g k p n w y q q k p g k a	gaagattt	cgca a a T V
	atca c	acct ga	ပ္တ ၂ ဗ
	atc √ v oft	స్ట్ _{ర్} : రా	cggc a r g g
	attggta k l v n w i g	tctgcag	aacgggcg k r a n g q t g
		totgo	acgg a t
	α × .		o :
	r	3 d d	tca s :
	T T T	33 33 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	e a x
	S S L	tid . d	86 × 5 : ≥
	g g	847-:1	gtgga r w g g
	atcacttgcc gggcaagtca gagcattagc agctatttaa h h l p g k s e h - q l f i t c r a s q s i s s y l p s l a g q v r a l a a i - v x 1-39Ri t c r a s q s i s s y l i t c r a s q s i s s y l	cactotca coatcag h s h h q f t l t i s l s p s f t l t i	agggaccaag gtggaaatca aacgggcggc cgca g g t k v e i k r a a a k g p r w k s n g r p r d g g n g t g g r k g p r w k s n g r p
		₹ a	ള്ച്ച്
	s s	1 pg	3 d d d d d
	رم طور برم الم	f h s f t s l t t t t t t t t t t t t t t t t t	agggac a g r g k g
	в в в	#	c b b c a
	S V S	Tage Hard	gg a a a a a a a a a a a a a a a a a a
	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	tgggacaga w d r s g t l g q	විද්ධ ස
	T a a	8 6 8 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	#
	ည္တတ္ ု		t cattegges s h s g l r r r r r r r r r r r r r r r r r r
	o b b b c c d	g gad	0 >
	1-3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1-3 4 5 8 5 8 5 8 5 8 5 8 5 8 5 8 5 8 5 8 5	a a r
	28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28. 1 28.		r r r g
	A P	S S S S S S S S S S S S S S S S S S S	igga agg . g r . g k . g k -39x
	cagagtcacc of r v t t t t t t t t t t t t t t t t t t	s d	tggga agg 1 g 1 w e 1 i g k 1-39R
	agtcacc r r v t c t e s e s r r v t r v v t r v v v r v v v r v v v r v v v r v v v v r v	g, o	r v l g a y w r i g . r i g
	E o I	12	X X X
	ද අත්ත	k v	gtattg v 1 x i r i v.VK 1.
	0 0 0	დ თ : თ ი თ :	acca ggtttaaacg c t r f k r l p g l n y q v - t l p g l n l p g l n
	ggaga c	₽~~ <u>_</u> ~	56 7 = _ : =
	agg	ggtcccatc g v p g v p	8 ' :
	s s	Σ _δ δ	# 5 ° 5 ° 5
	ರ್ಷ : "	isagtgg ggtccci	tacca ggtttaaacg f t r f k l p g l n y q v - t l p g l n
	a p	s s	tacca (f t l p y y
	5 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	# ~ ~	
	ر ج د را	ര്രേഗ് : ഗ	4 d 1 ft
		; ° ;	cggttt g g g a v
	ပ က ့်က	+ > :	acggtaagg cggtt t v r r - r - g g n g k a v
	: 유 : **	# p	Bg
	¥, , , , , , , , , , , , , , , , , , ,	ည္တိ _{ုက္} တ ^{ညာ} ႏွတ	gte r.gr
	වෙදි යු . අ	# : ™ ; ; ; ; ;	ຼິ ₂
	ມັ ຜ່ <u>.</u> ພ	ъ ~	તે ⊣
	20 A C C C C C C C C C C C C C C C C C C	್ಟ್ ಇನ್ನ	ം പ്രൂപ്പ
بتآ	စ္က် <u>က</u> ့ က	Z Z Z	gagacgaat aacggtaag r r i t v r d e - r - e t n n g k r d e - r -
⋖		합 ⁸ :	r e r
24	caga tgacccagtc tccatcctcc p d d p v s i l c m t q s p s s s r - p s l h p s r q m t q s p s s	g : _	ည္ ဖ
9	g pr	성요 g	တ္ထုတ္ ေတ
(1)	8 c	₽; _" ¬ " ; ¬	ក្តី ភ្នំ ភ្នំ
VH1-39R AF	# - 1 = E	taagctcct gatctatgct gcatccagtt tgcaa - a p d l c c i g f a p k l l i y a a s s l g l s s - s m l h p v c k p k l l i y a a s s l g	ctgtcagcag cgagacgaat aacggtaagg cggttt 1 s a r r i t v r r i y c g g r d e - r - g g t v s s e t n n g k a v
VF	<pre>gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtagc g h p d d p v s i l p v c i c r d i q m t q s p s s l s a s v r t s r - p s l h p p c l h l - >></pre>	ctaagctcct gatctatgct gcatccagtt tgcaar p - a p d l c c i q f a p d l s s l q s s l q s s l q s s l q s s l q s s l l i y a a s s l q s s l l i y a a s s l q p k l l i y a a s s l q	ctgtcagcag cgagacgaat ad s s s s t r i s t v s s e t n s c g g r d e s s e t n s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s s c g g r d e s c g g r d e s c g g r d e s c g g r d e s c g g r d e s c g g r d e s c g g r d e s c g g r d e s c g g r d e s c g g r d e s c g g r d e s c
,	***		

Fig. 8 Cont.

			60 T Z T
		mt i A	69 17 1
	ggccaggctc W p g g q a l a r g q a g q a	# x ? x	
	ragg P d d d	tta 1 v v	8 8888
	3gCC3 ▼ 1 9 9	agti a q	SEC
	•	້ ບໍ່	8 888 8
	الم برنج	\$ 4-1:4-1	
	g 9 7 7	# To : 0	
	acagaaacct t e t g g g k p n r n e g g k p	ва е ж е да	g o ^
	ညီ _တ ြင်	हु । स	cgca a a cyca
	s I v p q b g t g w y q d a w y q d	D P Ct	
	w v v v v v v v v v v v v v v v v v v v	n add	8 8° 8
	tdgt	tag 1. 1	8 4 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
	a a	s s	g - 5
	be de la company	ය ය ය ය ය ග	latca ack i k s h s de n de
	Z: t	B 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	A Salat
	S. a.	catc h p	ygaa w g
	р О	8 4 7 4 1	agggaccaag gtggaaatca q g t k v e i k g p r w k s r d q g g n k g p r w k s
	ည္မ		t k P G
	o ' ' ' '	1 s t	accase d p d p
	gagtgi e c r v r v	Sg. H sg. H	33 ag
	ਹੈ ਜ਼ਿਲ੍ਹ	# ·	
	gccagtca g g s g p v	ਹਾਂ ਹਾਂ	a a a
	cagt	ب م ^ب و رق	2000 E 2 T 2
	20 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	الطوعة م م م م	T T T
	8557 : H	tc tgggi	ct cat
	tectgea 1 1 s c s p a x 3-116.	gtc g g g	25 × 25 : 12 × 12 + 12 × 12 + 12 × 12 × 12 × 12 ×
	S S-1	7499 8 % 3-1	r gat
	ctctc	Gaga G Kanga	ပို့မှုတ _ိ ု့တ
	t c	5 3	96 9 30 5 4 4 50
	없. ^다 요.	D C S S S S S S S S S S S S S S S S S S	9998 1 9 W 9
	င္တင္က	t t cac	gtattg
	9 9 9 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1	14 2 4 5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	cgtatt r v a y r jVK
	a a . a	ര മ.	, , , , , , , , , , , , , , , , , , ,
	් යු ය යු යු	Zagcc P a P a	aaacg ck 1 n 1 n t t t t t t t t t t t t t t t t
	. agg	55 d	\$ 4 - ; -
	S S	क व्य	ggtttaaacg cgtattggga aggcgcgtct cattcggcca r f k r v l g r r v s f g g l n a y w e g a s h s a g v - t r i g k a r l i r p
	- S	ပိ တိုင်းတ တနား	
	ttgt 1 c	4 1 2	acca g
	1123 × S	ğ	
	te de la capacita de	Aggece L a L a L a	ggttt a g v
	о	മാമ് യ ച:	ന് ത്ത
	25 d	20 D T T	gag → · ·
	တို့ _အ က မ	25.7 × 1 : ×	H. G. H.
	25.2.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1	gcatcci c i c i m a s	acg
	gtgt tgacacagtc tccagccacc c v d t v s s h v l t q s p a t l c - h s l q p v l t q s p a t	o, a ;	cgaat aacggtaag r i t v d e - r - t n n g k
	d sagt	ctatgat 1 - i y d s m	cgaar d e d e
AF	t 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2		D R
r m	T. L.	s p	R R R R R R R R R R R R R R R R R R R
16	+		0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
=	943 C	1: 3 T	ر م م به م
က်	e k	aggetect cat q a p h r l l l	tgtcagcag cgagacgaat aacggtaagg 1 s a a r i t v r c q q r d e - r - t v s s e t n n g k c q g r d e - r -
VK3-11G AF	<pre>gaattgtgt tgacacagtc tccagccacc ctgtct g n c v d t v s s h p v e i v l t q s p a t l s r k l c - h s l q p p c >> v l t q s p a t l s e i v l t q s p a t l s</pre>	ccaggctcct catctatgat gcatccaaca gggcca s q a p h l - c i q q g p r l l i y d a s n r a p g s s m m h p t g p	ctgtcagcag cgagacgaat aacggtaagg cggttt 1 s a a r i t v r r f y c q q r d e - r - g g t v s s e t n n g k a v >
			A - 1

Fig. 8 Cont.

			6440 K
	5 - 3	# 7 5 7 5	2 4 2 5 5 E
	ggccaggctc W P g g g a l a r l g g g a	ttat 1 V V	No: 73 No: 75 No: 76 No: 75
	g	11 s 1 s 1 s 1 s 1 s 1 s 1 s 1 s 1 s 1	
		tg cagttfatta c s 1 l f a v y l l g f 1 l l f 1 l l l l l l l l l l l l l l	SEC SEC
	the late	¥. 44-	01 02 02 01
	8 9 7 . 7	₹ ~ ~ . ~ ~ . ~ ~ . ~ ~ . ~ ~ . ~ ~ ~ ~	
	م ل م ل م الم الم الم الم الم الم الم ال	gg a × a	cgca p a a
	cca acagaa I p t e I n r	cctagagcct gaaga Prarr Slepe a-Slk a-Slk	cgca p a a
	Z T Z	cct S B	agg a
	ag ag	gåg	B
	a a e	To the	acg n t
	0 M	0 % · · · · · ·	gtggaaatca aacgggcggc v e i k r a r w k s n g r g g n g t g g r w k s n g r
	tt tt	gc ag s	\$ -
	s y	1	
	ည္က က မ က မ	25 T P T T P	\$ 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
	ည္မ	β _ ω : _	aag gtggaaatca aa k v e i k p r w k s q g g n g p r w k s
	gagtgttagc agctacttag cctggtacca acagaaacct e c - g l l s l v p t e t g s v s s y l a w y g g k p r v l a a t - p g t n r n g s v s s y l a w y g g k p	ttcagtg gcagtgggtc tgggacagac ttcactctca ccatcagcag cctagagcct gaagattt v q w q w v w d r l h s h h q q p r a - r f f s g s g s g t d f t l t i s s l e p e d g s v a v g l g q t s l s p s a a - s l k i vvx 3-1R f s g s g s g t d f t l t i s s l e p e d	t cattoggcca agggaccaag gtggaatca aacgggcggc s f g q g t k v e i k r a s h s a k g p r w k s n g r l i r p r d q g g n q t g g s h s a k g p r w k s n g r
	s c		
	g 5 5 5	# : "	gg. P. x , ; x
	യ വ	8 T D	gcca agggacc g q g g r p r d r p r d
	or a s s s s s s s s s s s s s s s s s s	1 6 6 6 8 8 4 8 6 8 8 4 8 8 8 8 8 8 8 8 8	
	99 P H B B H B B B B H B B B B B B B B B B	66 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	# 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
	8.5.	ე, ო <u>_</u> ო	S 2 -
	ర్జ్లు స్ట్రాల్లో	\$ <u>`</u> \$	rgtct cattcggc r v s f g a s h s r l l r l r r l s r s r s r s s s s s
	Series Series	rtgg s - x - x - x - x - x - x - x - x - x -	cgcg
	ctccagggga aagagccacc ctctcctgca gggccag v s r g k s h p l l g g g s p g e r a t l s c r a l g g k e p p s p a g p v s p g e r a t l s c r a	cccagcc aggttcagtg gcag 1 p s q v q w q 1 p a r f s g s q p g s v a i p a r f s g	aggegegtet catteggeea rrvsfg egashsaa karlirp R
	S - 1 - 1	5 , > :	acca ggtttaaacg cgtattggga agg t r f k r v l g r l p g l n a y w e r y g v - t r i g k v g v - t r i g k v g l n a y w e
	8 . a	\$ 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	tggga 1 g w 1 g 3-111
	E H		att v v
	ctttgt ctccagggga aagagccacc v f v s r g k s h s l s p g e r a t l c l q g k e p s l s p g e r a t	ccactgg catcccagcc aggttcacgg h w h p s g v c a t g i p a r f p l a s g p g s s a t g i p a r f	ttacca ggtttaaacg cgtattggga f t r f k r v l g l p g l n a y w v y g v - t r i g v y g v - t r i g l p g l n a y w
	s r g s p g g s p	ည္အ ေ _ : ဇ	tacca ggtttaaacg f t r f k l p g l n r y q v - t
	aggggga r g r g g g	, a	agg
		ਨੂੰ ਕੂ ਜ਼ੂਜ਼ ਰਾਜ਼	## # * * * * * * * * * * * * * * * * *
	5 s	ပြီး တုိင္ဘေ	. р.;
	\$	ctgg catc	tttacca ggt f t r g l p g v y q g l p c
	S T S T S S T S S S S S S S S S S S S S	00 _ # J : #	# > :
	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	256 4 5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	aggt aggt
	μ. μ. μ. υ	ого ; ч р:	ပ် ဗ ^က ဗေ
	್ಣ _ಇ ಜ	aac d d :	ge > ' : '
	gg 2 5 : 4	Satccaa a s h p	r d t
	S L S	25 0 E E	i i i i
	Hattgtgt tgacacagtc tccagccacc ctgron control control control control ctground ctgro	d: d	gcag cgagacgaat aacggtaagg cg a a r i t v r q q r d e - r - g s s e t n n g k a q r d e - r - g
ш		atg ✓: ™ ✓	d: t d ga
A	ر م <u>رو</u> ق	rtct r s r	r e r aga
ĸ	\$ 7	ra Las	ටි ශ .
11	tg > ° : >	cct s	ອ ອ ອ ອ ອ ອ ອ ອ ອ ອ ອ ອ ອ ອ ອ ອ ອ ອ ອ
3	\$ -d :-d	gct r a	tgtcagg
VH3-11R AF	<pre>gaattgtgt tgacacagtc tccagccacc ctgt, g n c v d t v s s h p v e i v l t g s p a t l r k l c - h s l g p p c >> e i v l t g s p a t l e i v l t g s p a t l</pre>	ccaggctcct catctatgat gcatccaaca gggccs s q a p h l - c i q q g p r l l i y d a s n r a p g s s s m m h p t g y r l l i y d a s n r a p r l l i y d a s n r a	ctgtcagcag cgagacgat aacggtaagg cggt 1 s a a r i t v r r y c q q r d e - r - g g t v s s e t n n g k a s >
-	ou AX	ο α	O PA A PA

Fig. 8 Cont.

			F 8 5 8 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
	a	מָּשְֹי	
	ggccaggctc w p g g q a l a r l g q a	ttg cagtttatta f c s l l f a v y i f g f i ,	
	0 P P P	S 2 . >	
	9go 1	og o	SEC
	n h	الم الم	
	r e dage	∄ ~თ~ : ი	
	gcagaaacct a e t g g k p s r n g g k p	8 9 4 : 0	o m
		g, Ц:	r r a
	tggtacca gcagaa W y q q q p g t s r w y q q q w y q q q q q q q q q q q q q		2 L 2 C
	gtac v v v g g	age a co	E
		cctgcagt p a s 1 q a c s	99999 7 9
		ე აც . ა	සි ප
	ta ta	agcag a a a a a s	ggaaatca r e i w k s g n w k s
	gg gg i e	20.00 1.00 1.00 1.00	cggaaat w k w k
	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	ccatcagg	gtggaaa v e r w k g g g
	დ . დ ი . დ	ອ ທ ໍ	
	u s v s	s	ccaag t k d q d q
		₽⊒ + ' : +	
	da d	trea f f f	aggga K d
	က် က	ജൂ ം വ	8 _ a a . a
	agto g s s	٠ . م نايق	မ္မာ မေ
	9 1 0 1 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 1 0	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	f f i s i s
	86 P. H.	t to	
	55.0 56.0	gggtc w v g -15G.	r o
	cctgca 1 1 s c p a 3-15G.	54 € × × × × × × × × × × × × × × × × × ×	a a
	L A S	a ∰a dd	& 4 ° ' ; ° '
		တ်'≋ ် က >	
	8년 a : _	agtg	cgtattggga r v l g a y w r i g VK 3-155
	မ္က်င္မ မ	, , , , , , , , , , , , , , , , , , ,	tattgc
	a adga k k r r r	Pp P C:	a r cgta
		ກ ຂ ີ່ ຂ	
	gggga r g q g	합교교: 교	ttaaacg f k l n v - t
	~ ~ "O.	Š _ ~ ° :	4 5 th
	ctcccv s s s s s s s s s s s s s s s s s s	tateccag Y P g i P v s q	Б н Б :
	ytgt c v v	ب المرقع	tacca ggt f t r l p g y q l p g
		E >CT	tage
	tgtctc 1 s c c	d d d d d d d d d d d d d d d d d d d	cggttt r r g a g
	5 ^{ch} . ch :	557 :	
	864 T a : 1	+; pr	gtaagg
	မ္ဘီး ႏွင့္ ကို ကို ႏိုင္ငံ	22 -	r g r
	cccagccac s s h s P a l q p	a late	gg 1
	2 8	g . g	+
Fr.	gg gg : b	rtggt v g m y g	gagacgaat r r r r d e t n
AI	# # # # # # # # # # # # # # # # # # #	χςτ. 1	55 T O
Ō	gtga tgacgcagtc tccagccacc s d d a v s s h v m t q s p a t r s l q p v m t q s p a t	tg T s :-	т т го п
<u>L</u> 5	g > ' :>	cct c	g g c : 5
Ĭ	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1]. 9 . a	tcagcag cgagacgaat aacggtaagg s a a r i t v r c q q r d e - r - v s s e t n n g k c q q r d e - r -
VK3-15G AF	gaaatagtga tgacgcagtc ctgtct g n s d d a v s s h p v e i v m t g s p a t l s r k r s l q p p c >>	ccaggctcct catctatggt gcatccacca gggcca s q a p h l w c i h q g l p r l l i y g a s t r a p g s s s m v h p p g p p r l l i y g a s t r a	ctytcagcag cgagacgaat aacggtaagg cggttt 1
\triangleright	en a ∨	5 2 3 3 4	ち みずべか

Fig. 8 Cont.

	,, <u> </u>	or .	22 22 22 22 E
	agger and a graph	y: v	
	2 P 20	tte	8 8888
	55 × 6 × 6	a a a	SEC SEC SEC SEC
	P L Ct	g 0	യയയയ
	aacc k		
	2	~ :	رم به به م
	B	gg dg	cgca a a p r r
	oca y t y t	tot s	a a r
	gtac w y	gg gg . 5	C) 500 +
	\$ _ a : _	cctgcag p a s 1 q a c s	§ 7 = 1 ; =
	ະ :		•
		00 00 00 00 00 00 00 00 00 00 00 00 00	atc x
	age of the section of		ය ල ම ව
	8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	t p t	gtggaaatca v e i r w k s g g n r w k s
	යි ය	g	
	世。	1, 1, 2, 2, 2, 1, 1, 2, 2, 1, 1, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,	accase d t d p
	gagtç g s g s	frage frage frage	agggaccaag q g t k k g p r d q r d q
	88 . 	ည်း ဖ တေလ •လ	
	gtca g s v v	tgggacagag w d r s g t e 1 g g	00 p
	20 20 a g	25 g g	tcg
	999 4 4 5		cattoggoca s f g h s a l i r p
	5. 2 ° 5.	gggtc w v g g -15k.	2 < Ct
	S 3-15	2 × 2 × 2 × 2 × 2 × 2 × 2 × 2 × 2 × 2 ×	D + 8 : 8
	eteti P VX	7089 9 9 9	5 7 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
	•	8 - 8 - E	
	a	ttcag v q f s	cgtattggga r v l g a y w r i g r i g vvK 3-15.
	e a a a a a a a a a a a a a a a a a a a		tatt v v v v v v v v v v v v v v v v v v v
		a agg	•
	8 p p p	ည္က် ⁸ က်	tttaaacg f k g l n y - t
	S T G	2000 1 s 1	T T
	55 ° – ° ° °	tat de la tat	4
	c c c c c c c c c c c c c c c c c c c	ნ,≽	രൂച്ച വ
	s 1	1 2 3 3 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	tag T T T
	₩	gggcca g g g p	a datt
	5 + :+ 5 4 :	B. D. I.	တို့ တက္ခ
	s b t a t	: L L Back	13 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
	ည်း က က က က က က က က က က က က က က က က က က က	toca i s s:	1
	S S S S S S S S S S S S S S S S S S S	gg o d : m	at aacggtaagg i t v r e - r - n n g k n r e - r -
	ਹੈ;	अर्थ के किया के किया किया किया किया किया किया किया किया	cgaat aacggtaac
أتتأ	де д г г г г	y y Y	d t d K
Æ	da d	i gradit	cgagac a r r c s e
VH3-15R AF	<pre>gaaatagtga tgacgagtc tccagccacc ctgtct g n s d d a v s s h p v e i v m t q s p a t l s r k r s l q p p c >> e i v m t q s p a t l e i v m t q s p a t l e i v m t q s p a t l e i v m t q s p a t l</pre>		ctgtcagcag cgagacgaat aacggtaagg cggttt 1
ij	gtg v	aggetect q a p r l p g s	tcagcag s a c q q v s
13,	a k n n n n n n n n n n n n n n n n n n	ਹੁੰਦੀ ਜ਼ੁਰੂ ਜ਼ੁਰੂ	otcago c d c q c q
VE	gga	ccaggctcct catctatggt gcatccaaca gggccs s q a p h l w c i h q g p r l l i y g a s t r a p g s s s m v h p p g p p r l l i y g a s t r a	ctg

Fig. 8 Cont.

			8 8 8 8 8 7 8 8 7 8
	to HA	gcagttta c s 1 a v 1 q f	
	တ္တြင္း ေတ	# > _ >	
	55 B 3	g	
	cctg t p. p.	ttgcagttt f c s l a v f a v	SEC SEC SEC SEC SEC
	ida cctggccagg e t w p k p g q n l a n		
	B D H D	aagatt t	
	D D D	cctgaaga a - r p e s 1 k	gccgca a a r p g r g r
	تع ب أي ي		ggcc a a r g
		D 0	சு எ
	£ * . *	ottggag 1	9 9 9 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
	5 g g	o të	ge *
	Pp. 1 1	gg or	caaaggg
	+ + :	accatcag cagactggag cctgaag h h q q t g a - t i s r l e p e g p s a d w s l l k t i s r l e p e	aaggtggaaa tcaaaggggckveikgbrwkskg prwkskg qggnngrg
	a y Lac	t g	gg , 6
	5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	ر بـ بـ روة بـ بـ بـ بـ ا	<u> </u>
	ပ္ဆိုင္မွာ ေလ	tcac 1 1	gg × g · ·
	<u>യ</u> മ	to teaceateag cagactggag cetga s h h q q t g a - t l t i s r l e p l s p s a d w s l t l t i s r l e p	υ μ . μ . μ . μ
	tag 1	다 다 다 다 다 다 다 다 다 다 다 다 다 다 다 다 다 다 다	g d
	ئے دی اور ا	tt tt	핥
	gagtgttagc agcagctact tagcctggta ccagcagaaa cctggccagge c - q q l l s l v p a e t w p q s v b s s s y l a w y q g k p g q r v l a a a t - p g t s r n l a q s v s s s y l a w y q g k p g q q s v s s s y l a w y q g k p g g	d L	g ccaaggacc a g g t g g t r d r d r d r d r d r d r d r d r d r
	agtca gagtgttagc agcagctact tagcctggta ccagcs g s e c - g g l l s l v p s s g s v s s s y l a w y g p v r v l a a a t - p g t s s g s v s s s y l a w y g	ca gacttca d r l t d f g t s t d f	tcattcgg ccaagggacc aaggtg s h s a k g p r l i r p r d q g s h s a k g p r
	हुत तुर्दि इ. इ. इ.	98 6 9 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	¥
	9 9 9 e	25 L S	rtcati
	9886 p.	9 6 G	tct
	ctccagggga aagagccacc ctctcctqca gggccag / s r g k s h p l g g g s p g e r a t l s c r a l q g k e p p s p a g p v v 3-206	ca gtggcagtgg gtctgggaca gacttcact q w q w v w d r l h f s g s g s g t d f s v a v g l g q t s VK 3-206	k r v l g r r v s f g g g t r v s f g g g t l n a y w e g a s h s a k g g r r r i g k a r l i r p r d s v s s s s s d g s s s s s s s d s s d s s d s s d s s d s s d s s d s
	rectt	3-2 d d	ρου μ ο ο ο ο ο ο ο ο ο ο ο ο ο ο ο ο ο ο ο
	ctctc p 1 1 s P s VK 3	orgage orways orways	9 d da
	aagagccacc ctctc	ca gtgg	cattg ggaag v l g y w e r i g k r 3-20G
	25 T T T	The state of the s	acgcgtattg k r v l n a y t r i t r iVK 3-2 n a y
	င္ဆိုင္က မ	B P H C H	a r a
	BA A	tocca gacaggtt h p r g v i p d r s g t g	ttaa acgcgt f k r l n a a r r r r r l n a
	99999 B B B B B B B B B B B B B B B B B	g <u>a</u> . a	© <u>.</u>
	ctccagggga v s r g s p g l q g	200	ftttaa g l g l g l
	വ് പ് പ	ع م ال	ccaggttl
	ctccac	tggcat w h t g l a	a ccaggttt t r f l p g y q v
		iccac tg	g
	+	9 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	ir f g l g l
	<u>`</u> 2, ≥ " ∪	ğ b d	1 d d d d d d d d d d
	చ్ _ద ్	d d d d d d d d d d d d d d d d d d d	ا : ۱ مرور
		တ္တည္း အ	ta d
	ဦး ေ	atc	acggta t v n g
	င်အ ရှင်းရေ	60 p	ataa L e l e l
	\$ ₂₀ ; 2	tat ggtgcatc 1 w c i y g a s m v h	е н • • • •
	9tc	tat 1 y	r g
AF	95 P P P	atc.	ਮੂੰ ਮੂ
	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2 L Ctc	ر م م
00	aattgtgt tgacgcagtc tccaggcacc ctgtctt y n c v d a v s r h p v e i v l t q s p g t l s k l c - r s l q a p c l e i v l t q s p g t l s	5	actgtcag cagcgagacg aataacggta aggcgg 1 1 s a a r r i t v r r y c q q r d e - r - g i t v s s e t n n g k a y c q q r d e - r - g
7	or control of the con	д 1 1	S S
က်	#	gg a d	類一、上、
VK3-20G	<pre>gaaattgtgt tgacgcagtc tccaggcacc ctgtct g n c v d a v s r h p v e i v l t q s p g t l s r k l c - r s l q a p c >> e i v l t q s p g t l s e i v l t q s p g t l s</pre>	ctcccaggct cctcatctat ggtgcatcca gcaggg g s q a p h l w c i q q a p r l l i y g a s s r l p g s s s m w h p a g	ttactgtcag cagcgagacg aataacggta aggcgg 1 1 s a a r r i t v r r y y c q q r d e - r - g i t v s s e t n n g k a >
-	J. 47	3 G. X	+

Fig. 8 Cont.

			92 92 91 92
			N N N N
	B	tt L L V	
	2002 P	agttt	
	\$ 2 : 4	ttgcagttta f c s l f a v l q f	SEC SEC
	agcagaaa cc q q k g q k s r n		
	T eas	te in the second	
	टिंडवु व व	otgaaga P e r P e r	cgca a a d d d
	y P	cctgaaag a - a s 1 e p e	ggccgca a a a r p g r
		9 d d	
	1.23 · · · · · · · · · · · · · · · · · · ·	86 - * · -	cgggc r r r g t g
	agcctg	r t	· · ·
	tag	cagactgg	tcaa i s n q
	ctact tagcctgg l l s l s y l a w a t - p s y l a v	c teaceateag eagact s h h q q t l t i s r l s p s a d	ida tcaa e i k s r n q
	. w - w	atc	aggtggak k v k r w q q g g
	70 70 70	caccat. h h l t s p	r g
	9.0.	5 ° _	aaggt k k p r q g
			ggacc aaggtggaaa to g t k v e i g p r w k r d q g g n g p r w k
	gtto	t t t	999 K 9
	gagtgttag e c - q s v r v l	gacttcact r l h d f q t s d f	ccaaggg
	ω α · · · · · · · · · · · · · · · · · ·	a	ρ. α μ α μ
	gtca g s s v	විසි ව විසි විසි	trage in the second sec
	aggcca r a a r a b	2 w S S W S S W S S W S S W S S S S S S S	stratto s h s h
		gtctgggaca gai v w d r g s g t g l g q g s g t	g totcattogg ccaagggacc aaggtggaaa v s f g q g t k v e a s h s a k g p r w k r l i r p r d q g g a s h s a k g p r w k
	acc ctctctgca h p l l t l s c p p s p a v y 3-20R.	.gg 	ggogog teteatter r r v s j k a s h k a r l i k a s h e g a s h e g a s h
	3-2-2 3-2-2	gcagtgg w g w g s a v (3-20R.	Igaga I I I I I I
	Ctctc P L	gtggca g w g v a v v a v v a s g	. g 28
	6 4 :n	a gtgg g v s v s	~:
	ccacc ctctcctgca s h l l a t l s c t p p s p a p avk 3-20R.	ttg: # .	acgcgtattg k r v l n a y t r i t r i n v y n a y
	B 0 .	aggtt	regta r v a r r r v r v r v
	ര് ധ്ര	ਰ ਦੂ ਦੂ ਰ ਦੂ ਦੂ ਰ ਦੂ ਦੂ	acgcgt k r n a t r r r r r r n a n a n a n a
	60 p p	മെ പ	aggtttaa a r f k g l d d l d l d l d l d l d l d l d l d
		1	# 6 : 5
	ಜ್ಜಿ ದಿ	မွ်း နွေး ရွေး နေး ရေး ရေး	28 d d
		T 1	۲ م ر د وو
	\$4-°	a p	gtta cca g l t g y y g l I
	s s	26 £ 5 ± 1	9 4 4 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 6
	than or in	် ရှင်းမှ ရေး	999 r g
	ני עני	о о.;	ल । • । ल >
	2 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	tcci	ggta t v r g
	agga L d L	დენე	dtaacg i t e - n n
	tccaggcacc c s r h s r h d s t l d a E s p g t s r s p g t s r s p g t s p g t s r s p g t s	29g ± 19g ±	e e adr
	2 4 1	tat ggtgcatc W c i Y g a S m V h y g a	d t d
压	cag r d	tct 1 s 1 s	ਦ ਦੇ ਸ਼ੁਰੂ
K	gacgcagtc tccaggcacc d a v s r h l t q s p g t - r s l q a l t q s p g t	rtca 1	gc g
ᄶ	\$ - :-	20 T	actgicag cagcgagacg aataacggta aggcggf 1 1 s a a r r i t v r r Y c q q r d e - r - g g t v s s e t n n g k a Y c q q r d e - r - g g
20	tgtgt t	igct r g	Cag
ω _	¥ = 7 : -7	teccagget s q a a p r l p g a p r	\$ - 0 - 5
VH3-20R AF	<pre>gaattgtgt tgacgcagtc tccaggcacc ctgtctt g n c v d a v s r h p v e i v l t q s p g t l s r k l c - r s l q a p c l >> e i v l t q s p g t l s t l d s p g t l s</pre>	cteccagget ceteatetat ggtgeateca geaggge g s q a p h l w c i q q g a p r l l i y g a s s r l p g s s s m v h p a g	ttactgtcag cagcgagacg aataacggta aggcggf
-	υ, Η Λ	J 55 A	
			-

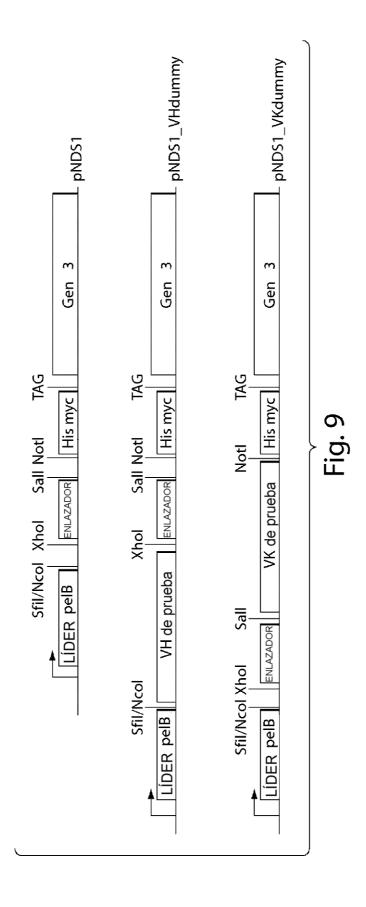
Fig. 8 Cont.

	aacaa g t t vag t	aggccgatta r g r l e a d r p i r p i e a d	SEC ID N°: 93 SEC ID N°: 94 SEC ID N°: 95 SEC ID N°: 96
	cottggta ccagcagctc ccaggaacag i I v p a a p r n s w y q q l p g t i p g t s s s q e q s w y q q l p g t	actggggacg d w g t g d r l g t	ccgtcctagg ggcggccgca t v 1 g a a a p s - g r p d r p r g g r
	acattggg aataattatg tatcct q h w e - 1 c i n i g n n y v s t 1 g i i m y p n i g n n y v s	agccaccc tgggcatcac cgg s h p g h h r s a t l g i t g p p w a s p s a t l g i t	ggg accaagctga g t k 1 e g p s - r d g a e g p s -
	c tectgetetg gaageagete caacattggg aataattatg tateetggta ecageage: 1	ctggctccaa gtctggcacg tcagccacc tgggcatcac cggactccag a la l	rtg ggaaggege tetetgtgtt eggegga 7
	caggacagaa ggtcaccatc p r t e g h h p g q r r s p q d r r s p p g q k v t i	agggatteet gacegattet r d s - p i s g i p d r f g g f l t d s s g i p d r f	ccaggtttaa acgcgta t r f k r v p g l n a y g v - t r p g l n a
4 AF	cagtctgtgt tgacgcagcc gcctcagtg tctgcggccc a a a 1 s v c g g v 1 t q p p s v s a a r s 1 c - r s r p q c 1 r p s v 1 t q p p s v s a a r s 1 c - r s a a a	ccccaaact cctcatttat gacaataata agcgacctc s p q t p h 1 - q a t 1 a p k 1 l i y d n n k r p p p n s s f m t i i s d p > y p n l i y d n n k r p a p k l l i y d n n k r p	ttactgogga acaggagacg aataacggta aggoggtta 1
VL1-44	cagtctgtc a v c q s v r s l >>	cccccaaact s p q t a p k p p n y p n a p k	ttactgcgg

Fig. 8 Cont.

captchycy tyactcagac accetcagog tetaggacac cegggacaga gatcaccac tettaggacac cacatega agraatety taaactgga accetcagog tetaggacac cegggacaga gatcaccac tettagacacac cectcagacacac at 1 s v w d p r a s v v t i s c s g s s n i g s n t v n w y q q l p q r v t i s c s g s s n i g s n t v n w y q q l p q t v v t i s c s g s s n i g s n t v n w y q q l p q t v v t i s c s g s s n i g s n t v n w y q q l p q t v v t i s c s g s s n i g s n t v n w y q q l p q t v v t i s c s g s s n i g s n t v n w y q q l p q t v v t i s 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 n i g s n t v n w y q q l p g t v v l v v v l l s l p p q r p g q r v t i s c s g s s n i g s n t v n w y q q l p q t v v v v l l s l p p q r p s s g v p d r f s g s k s g t s s n i s s n t v n w y q q l p g t v v l l s l p p q r p s g g v p d r f s g s k s g t s s a s l r m r l l l l l l l l l l l l l l l l				-
gocagae gyteaceate tettqttetg gaageagete caacategga aqtaatactg taaactgqta ecageagete ceaggaacgg r r v t i s c s g s s n i g s n t v n w y q q l p g t g r y v t i s c s g s s n i g s n t v n w y q q l p g t r y t i s c s g s s s n i g s n t v n w y q g l p g t r y t i s c s g s s s n i g s n t v n w y q g l p g t r y t i s c s g s s s n i g s n t v n w y q g l p g t r y t i s c s g s s s n i g s n t v n w y q g l p g t r y b d r f s g t s a s l a i s g l g s s l r n r l i s v b d r f s g s s s n i g s n t v n w y g g l p g t r n r l i s v b d r f s g s s s n i g s n t v n w g g s c f c a d d g s s l a p g p p w p s v g s s l r n r l i s v b d r f s g s s s n i g s n t v n g r v s v f g g g t k l t v l g a a s s c l n n r l i s g l n a y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s - g r p s c l n n y w e g a s l c s a e g p s - p s - g r p s - g r p s c l n n y w e g l n n a y w e g a s l c s a e g p s - p s - g r p s - g r p s - g r p s c l n n y s c l n n y s c l n n s c l c s a e g p s - b s - g r p s - g r p s - g r p s c l n n n y l n s c l c s a e g p s - b s - g r p s - g r p s - g r p s c l n n n y l n s c l c s a e g p s - b s - g r p s - g r p s - g r p s - g r p s - g r p s - g r p s - g r p s - g r p s - g r p s - g r p s - g r p s - g r p s - g r p s - g r p s - g r p s - g r p				F 8000 0
ggcagag ggtcaccatc tettqttetg gaagcagete caacategga agtaatactg taaactggta ceagcagete ceaggaacgg from $\frac{1}{8}$ 1				
ggcagag ggtcaccatc tettqttetg gaagcagete caacategga agtaatactg taaactggta ceagcagete ceaggaacgg from $\frac{1}{8}$ 1				No No No No
ggcagag ggtcaccatc tcitgtictg gaagcagctc caacaccgga agtaatactg taaactggta ccagcagctc ccaggaacgg r a b l l f w k q l q h r k - y c k l v p a a p r n g q r v t i s c s g s s n i g s n t v n w y q q l p g t s s 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 s 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 g q r v t i s c s g s s n i g s n t v n w y q q l p g t g q r v t i s c s g s s n i g s n t v n w y q q l p g t g g c 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 s g r v p d r f s g s k s g t s a s l a i s g l g s e d e a d g s l t d s g s k s g t s a s l a i s g l g s e d e a d g s r v p d r f s g s k s g t s a s l a i s g l g s e d e a d g s t k r v l g r r v s v f g g g t k l t v l g a a a s s c 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 s s c f f r v r r r d g a r r r r r d g a r r r r d g a r r r r r d g a r r r r r d g a r r r r r d g a r r r r r d g a r r r r r r r r r r r r r r r r r r				
ggcagag ggtcaccatc tcitgfictg gaagcagctc caacatcgga agtaatactg taaactggta ccagcagctc ccaggaaccagc tcitgfictg gaagcagctc caacatcgga agtaatactg taaactggta ccagcagctc ccaggaaccag 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 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 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 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 g p - p i l w l q v w h l s l p g h q w a p v - g - g - g - g - g 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 g s l t d s l a p g p p w p s v g s s l r m r l l 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 g t k r v l g r r v s v f g g t k l t v l g a a a g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p g l n a y w e g a s l c s a e g p s - p s - g r p r p g l n a y w e g a s l c s a e g p s - p s - g r p r p s - g r p		brs. 54 A	ert	ខ្លួខ្លួខ ខ
ggcagag ggtcaccatc tcttgttctg gaagcagctc caacatcgga agtaatactg ta a e g h h l l f w k g l g h r k - y c g g r v t i s c s g s s n i g s n t v v v l -44. g r g s p s l v l e a a p t s e v i l v v l s c s g s s s n i g s n t v g g r v t i s c s g s s s n i g s n t v g p - p i l w l g v w h l s l p g h g s g s s l a i s g s s l t d s l a g s b s l a p g p p w p s v v l d s l a g s s s l a i s s gtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga cc r f k r v l g r v s v f g g g t k l t g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l v l l a y w e g a s l c s a e g p s - l s l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a s l c s a e g p p s - l c l s a s l c s s a s l s l c s a s l c s a s l s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s		4 E.G.	g: Tant	S SS SS
ggcagag ggtcaccatc tcttgttctg gaagcagctc caacatcgga agtaatactg ta a e g h h l l f w k g l g h r k - y c g g r v t i s c s g s s n i g s n t v v v l -44. g r g s p s l v l e a a p t s e v i l v v l s c s g s s s n i g s n t v g g r v t i s c s g s s s n i g s n t v g p - p i l w l g v w h l s l p g h g s g s s l a i s g s s l t d s l a g s b s l a p g p p w p s v v l d s l a g s s s l a i s s gtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga cc r f k r v l g r v s v f g g g t k l t g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l v l l a y w e g a s l c s a e g p s - l s l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a s l c s a e g p p s - l c l s a s l c s s a s l s l c s a s l c s a s l s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s		ع . و م . و	gg	
ggcagag ggtcaccatc tcttgttctg gaagcagctc caacatcgga agtaatactg ta a e g h h l l f w k g l g h r k - y c g g r v t i s c s g s s n i g s n t v v v l -44. g r g s p s l v l e a a p t s e v i l v v l s c s g s s s n i g s n t v g g r v t i s c s g s s s n i g s n t v g p - p i l w l g v w h l s l p g h g s g s s l a i s g s s l t d s l a g s b s l a p g p p w p s v v l d s l a g s s s l a i s s gtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga cc r f k r v l g r v s v f g g g t k l t g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l v l l a y w e g a s l c s a e g p s - l s l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a s l c s a e g p p s - l c l s a s l c s s a s l s l c s a s l c s a s l s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s		96. ^a p a	ਹੁਣ ਨੂੰ ਜ਼ਰੂ	
ggcagag ggtcaccatc tcttgttctg gaagcagctc caacatcgga agtaatactg ta a e g h h l l f w k g l g h r k - y c g g r v t i s c s g s s n i g s n t v v v l -44. g r g s p s l v l e a a p t s e v i l v v l s c s g s s s n i g s n t v g g r v t i s c s g s s s n i g s n t v g p - p i l w l g v w h l s l p g h g s g s s l a i s g s s l t d s l a g s b s l a p g p p w p s v v l d s l a g s s s l a i s s gtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga cc r f k r v l g r v s v f g g g t k l t g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l v l l a y w e g a s l c s a e g p s - l s l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a s l c s a e g p p s - l c l s a s l c s s a s l s l c s a s l c s a s l s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s		გ _ლ ა :	g . a . a	
ggcagag ggtcaccatc tcttgttctg gaagcagctc caacatcgga agtaatactg ta a e g h h l l f w k g l g h r k - y c g g r v t i s c s g s s n i g s n t v v v l -44. g r g s p s l v l e a a p t s e v i l v v l s c s g s s s n i g s n t v g g r v t i s c s g s s s n i g s n t v g p - p i l w l g v w h l s l p g h g s g s s l a i s g s s l t d s l a g s b s l a p g p p w p s v v l d s l a g s s s l a i s s gtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga cc r f k r v l g r v s v f g g g t k l t g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l v l l a y w e g a s l c s a e g p s - l s l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a s l c s a e g p p s - l c l s a s l c s s a s l s l c s a s l c s a s l s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s]: Tag	d. ga tg	ਲੂ ਪੂਨ੍ਹੇ
ggcagag ggtcaccatc tcttgttctg gaagcagctc caacatcgga agtaatactg ta a e g h h l l f w k g l g h r k - y c g g r v t i s c s g s s n i g s n t v v v l -44. g r g s p s l v l e a a p t s e v i l v v l s c s g s s s n i g s n t v g g r v t i s c s g s s s n i g s n t v g p - p i l w l g v w h l s l p g h g s g s s l a i s g s s l t d s l a g s b s l a p g p p w p s v v l d s l a g s s s l a i s s gtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga cc r f k r v l g r v s v f g g g t k l t g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l v l l a y w e g a s l c s a e g p s - l s l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a s l c s a e g p p s - l c l s a s l c s s a s l s l c s a s l c s a s l s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s		ರ್ ರ [ಿ] ಇತ್ತ	8 H 6 G	E P P P P P P P P P P P P P P P P P P P
ggcagag ggtcaccatc tcttgttctg gaagcagctc caacatcgga agtaatactg ta a e g h h l l f w k g l g h r k - y c g g r v t i s c s g s s n i g s n t v v v l -44. g r g s p s l v l e a a p t s e v i l v v l s c s g s s s n i g s n t v g g r v t i s c s g s s s n i g s n t v g p - p i l w l g v w h l s l p g h g s g s s l a i s g s s l t d s l a g s b s l a p g p p w p s v v l d s l a g s s s l a i s s gtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga cc r f k r v l g r v s v f g g g t k l t g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l v l l a y w e g a s l c s a e g p s - l s l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a s l c s a e g p p s - l c l s a s l c s s a s l s l c s a s l c s a s l s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s		g	g	gg a h
ggcagag ggtcaccatc tcttgttctg gaagcagctc caacatcgga agtaatactg ta a e g h h l l f w k g l g h r k - y c g g r v t i s c s g s s n i g s n t v v v l -44. g r g s p s l v l e a a p t s e v i l v v l s c s g s s s n i g s n t v g g r v t i s c s g s s s n i g s n t v g p - p i l w l g v w h l s l p g h g s g s s l a i s g s s l t d s l a g s b s l a p g p p w p s v v l d s l a g s s s l a i s s gtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga cc r f k r v l g r v s v f g g g t k l t g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l v l l a y w e g a s l c s a e g p s - l s l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a s l c s a e g p p s - l c l s a s l c s s a s l s l c s a s l c s a s l s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s		Z T T	2 × 2 × 1	999
ggcagag ggtcaccatc tcttgttctg gaagcagctc caacatcgga agtaatactg ta a e g h h l l f w k g l g h r k - y c g g r v t i s c s g s s n i g s n t v v v l -44. g r g s p s l v l e a a p t s e v i l v v l s c s g s s s n i g s n t v g g r v t i s c s g s s s n i g s n t v g p - p i l w l g v w h l s l p g h g s g s s l a i s g s s l t d s l a g s b s l a p g p p w p s v v l d s l a g s s s l a i s s gtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga cc r f k r v l g r v s v f g g g t k l t g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l v l l a y w e g a s l c s a e g p s - l s l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a s l c s a e g p p s - l c l s a s l c s s a s l s l c s a s l c s a s l s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s		ته دي نط دي	بر م م	<u>Б</u>
ggcagag ggtcaccatc tcttgttctg gaagcagctc caacatcgga agtaatactg ta a e g h h l l f w k g l g h r k - y c g g r v t i s c s g s s n i g s n t v v v l -44. g r g s p s l v l e a a p t s e v i l v v l s c s g s s s n i g s n t v g g r v t i s c s g s s s n i g s n t v g p - p i l w l g v w h l s l p g h g s g s s l a i s g s s l t d s l a g s b s l a p g p p w p s v v l d s l a g s s s l a i s s gtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga cc r f k r v l g r v s v f g g g t k l t g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l v l l a y w e g a s l c s a e g p s - l s l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a s l c s a e g p p s - l c l s a s l c s s a s l s l c s a s l c s a s l s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s		B ≥	1 s 1	r lag
ggcagag ggtcaccatc tcttgttctg gaagcagctc caacatcgga agtaatactg ta a e g h h l l f w k g l g h r k - y c g g r v t i s c s g s s n i g s n t v v v l -44. g r g s p s l v l e a a p t s e v i l v v l s c s g s s s n i g s n t v g g r v t i s c s g s s s n i g s n t v g p - p i l w l g v w h l s l p g h g s g s s l a i s g s s l t d s l a g s b s l a p g p p w p s v v l d s l a g s s s l a i s s gtttaa acgcgtattg ggaaggcgcg tctctgtgtt cggcggaggg accaagctga cc r f k r v l g r v s v f g g g t k l t g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l g l n a y w e g a s l c s a e g p s - l v l l a y w e g a s l c s a e g p s - l s l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l n a y w e g a s l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c s a e g p p s - l c l s a e g p p s - l c l s a e g p p s - l c l s a s l c s a e g p p s - l c l s a s l c s s a s l s l c s a s l c s a s l s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s a s l c s		gg × a r	ور ع ع ع ع	tcc : s d
		taa ' '	tgg « « « «	
		g	p.p.	g
		act y	بَــٰ عَنْ الْحَادِينَ الْحَادِينَ الْحَادِينَ الْحَادِينَ الْحَادِينَ الْحَادِينَ الْحَادِينَ الْحَادِينَ الْ	ctg 1 a s
		aat _ n v v	a . p	A A
		s s	.gg 1	d
		α α α α	, u	a p _ p
		3gg 1 2 3 1	tcc s	e H e 33
			S S S S S S S S S S S S S S S S S S S	933 a r
		2 d	l s	ر الالالالالالالالالالالالالالالالالالال
		5		04
		s s	28 - L B	o oft
		S B S	386	tgt 1
		gg ga b	tct v s	otc N S S
		£ 3		H G
		15. S. 14. S.	8 5 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	96 r g
		10 110	9 1-a 9 156	999 r g k
		ctt 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Fr. 1	gaa 4. e
		5 - T - T - T	5	3 9 W
		ga	+ - + + + + + + + + + + + + + + + + + +	tt V 1
		32 T	gat r	gta r a
		940	م: _ب م ا	1 t 1 t 3
		р н :н	р П	a z
		gaç g	S. P. CC	tag 1
		عرب م م م م	gtc , ,	9 t
cagtctgtgc tgactcagcc accctcagcg tctgggaccc ccaa d s a t l s v w d p r s a s g t r s l c - l s h p g r l g p r s a s g t r s l c - l s h p g r l g p r l g p r l g p r l g p r l g p r l g p p s a s g t p r l d s r l l l i y s n n g r p s r l b b l l l i y s n n g r p s r l tactgtgca gcaggagacg aataacggta aggcggttta cc rtactgtgca gcaggagacg aataacggta aggcggttta cc l r r g g l r r r r r r r r r r r r r r				P P P
cagtctgtgc tgactcagcc accctcagcg tctgggaccc a v c a d s a t l s v w d c s v l t q p p s a s g t r s l c - l s h p q r l g p >> q s v l t q p p s a s g t q s v l t q p p s a s g t ccccaaact cctcatctat agtaataatc agcggccctc g p q t l l i y s n n q r p p p n s s s i v i i s g p a p k l l i y s n n q r p ractgtgca gcaggagacg aataacggta aggcggttta l c s r r i t v r f y y c a a g d e - r - g g l i t v g q e t n n g k a v y c a a g d e - r - g g l		5 a	α · · · · · · · · · · · · · · · · · · ·	ς.
cagtctgtgc tgactcagcc accctcagcg tctggga a v c a d s a t l s v w q q s v l t q p p s a s g r l g s v l t q p p s a s g q s v l t q p p s a s g q s v l t q p p s a s g q s v l t q p p s a s g q ccccaaact cctcatctat agtaataatc agcggcc g p q t l l i y s n n q r g p p n s s s i v i i s q r s s p k l l i y s n n q r g r l s p k l l i y s n n q r g r l l c s r r i t v r r v r r v q q e t n n g k a l l v q q e t n n g k a l l v q q e t n n g k a l l v q q e t n n g k a l v v r r v v v v q q e t n n g k a l v v v r v v v v v v v v v v v v v v v		t h t gg	ctc P P	tta f v
cagtctgtgc tgactcagcc accctcagcg tctgg a v c a d s a t l s v v q s v l t q p p s a s r s l c - l s h p q r l >> v l t q p p s a s q s v l t q p p s a s q s v l t q p p s a s q v l t q p p s a s cccccaact cctcatctat agtaataatc agcgg g p q t p l l i y s n n q r p p n s s s i v i i s a p k l l i y s n n q r ttactgtgca gcaggagacg aataacggta aggcg l c s r r r t v r y v c a a g d e - r - g i t v q q e t n n g k a y y c a a g d e - r - g i t v q q e t n n g k a		g. g. g.	ος σ	r g
cagtctgtgc tgactcagcc accetcagcg tc a v c a d s a t l s v g s v l t q p p s a r s l c - l s h p q r y v l t q p p s a cocccaact cetcatctat agtaataatc ag g p q t p h l - s a p k l l i y s n n q p p n s s s i v i i n a p k l l i y s n n q ttactgtgca gcaggagacg aataacggta ag l c s r r i t v y v c a a g d e - r - i t v q q e t n n g k y r a a g d e - r - i t v q q e t n n g k y y c a a g d e - r - i t v q q e t n n g k		tgo S S	S a s	900 9 9 9
cagtctgtgc tgactcagcc accetcagcg a v c a d s a t l s q s v l t q p p s a r s l c - l s h p q >> v l t q p p s a ccccaaact cctcatctat agtaataatc g p q t p l l i y s n n p p n s s s i v i i n p p n s s s i v i i a p k l l i y s n n r a p k l l i y s n n r a p k l l i y s n n y y c a a g d e - r i t v q q e t n n g y y c a a g d e - r		r r	g 2 5	ga ' 🛪 : ı
cagtctgtgc tgactcagcc accctcaga v c a d s a t l r q p p s r l t q p p p s l r l r q p p p s l r l r q p p p s l r l r q p p p p p p p p p p p p p p p p p p		ය සියිල්	n n	gta g
cagtctgtgc tgactcagcc accct a v c a d s a t r g v l t q p p r s l c - l s h p >>		တို့ ကို	r tag	cgg
cagtctgtgc tgactcagcc ac a v c a d s a r s l c - l s h y l t q p r s l c - l s h y q q t l l i y p p n s s s i p p l s s s i p p l s s s i p p l s s s i p p q t l l i y p p q q e t ttactgtgca gcaggagacg aar ttactgtgca gcaggagacg aar l c s r r r y y c a a g d e i t v q q e t y y c a a g d e i t v q q e t y y c a a g d e i t v q q e t		55 7 m : u	taa s	taa '''''
cagtctgtgc tgactcagcc a v c a d s a q s v l t q r s l c - l s >>		ac d d d d d d d d d d d d d d d d d d d	gg 1 :	о н в в н в
cagtctgtgc tgactcar a v c a d s a v c a d s r s l c - l s l c - l s v l t c ccccaaact cctcatct g p q l l l p p n s s s a p k l l l i p p n s s s a p k l l i r q e g s y c a a g i t v q q e y y c a a g i t v a a g		3 g	rat v	d: talligg
cagtctgtgc tgact a v c a d q s v l t r s l c - >> q s v l t q s v l t g p l t l p p n s s a p k l l p p n s s l l c s r y v c a a i t v q q q y v c a a i t v q q q y v c a a		1	i s i i	gg gg
cagtctgtgc tg a v c a g s v l r s l c y v l y v l a p k l b p n s a p k l a p k l b p n s a p k l b v q ccccaaact cc g p t l a p k l b v q ccccaaact cc g v l a p q t l c s l c s l c s l v q c a c l t v q c y v c a c l t v q c	Ä	t: t agt	tca P 1 s	agg a d
cagtctgtgc a v c q s v r s l c >> q s v ccccaaact g p q t a p k p p n > p p n a p k y v c a l c y y c a l t v	7	₽° d	$\frac{1}{s}$	ညီ အ
cagtetgt a v c g s r s 1 >> g s ccccaas g p q a p q p p > a p k > a p k > y v c i t v	51	ος ν : _ν	ا ا ا ا	ຊິລ ຊື່ ຜ
cagtc a v r g r s y v g q q g p a p y v t t t t t t t t t v y v y v y v y v	ĭ	tgt L S	88 P. J. J.	gtg c c
Seary Seary att Mark	$\Gamma 1$	gtc g v	ე	act Y Y
	\triangleright	Sa ⊢ ∖	0 5 ° ° ° ° °	Y. Y

Fig. 8 Cont.



MUESTRA	FW3	Н3	H3/FW4	LONGITUD DE CDR3				
1a	CAK	AFRPNWGSRVLYF	DYW	15	SEC	ID	N°:	101
1b	CAK	LNWGWF	DPW	8	SEC	ID	N°:	102
1c	CAK	GWEGEY	DYW	8	SEC	ID	N°:	103
1d	CAK	DLWLDSSSNWF	DPW	13	SEC	ID	N°:	104
le	CAK	DLPG	DPHW	7	SEC	ID	N°:	105
1f	CAK	APPTGYF	DYW	9	SEC	ID	N°:	106
1g	CAK	VGLTGVHF	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	GVVNWGTRRKGWF	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	GSGFSSGWF	DSW	11	SEC	ID	N°:	116

Fig. 10

					_
DVT	66F-03 53F-04 48F-05 43F-07 33F-08 33F-08 35F-08	2,84-11 2,55-12 2,35-13 2,16-14 1,95-15			
RICA	21F-04 25F-05 30F-06 30F-06 43E-08 52F-10 62F-10 74F-11	9,94-12 1,16-14 1,36-15 1,56-16 1,56-16			
DIVERSIDAD TEÓRICA NNS DVK NV	1,0E+05 1,9E+06 3,4E+07 6,1E+08 1,1E+10 2,0E+11 3,6E+12 6,4E+13	1,2E+15 2,1E+16 3,7E+17 6,7E+18 1,2E+20			
DIVERS	3,45+10 3,45+10 3,45+10 3,55+13 3,55+13 3,64-10	3,7E+19 1,2E+21 3,8E+22 1,2E+24			
POSICIONES	4 v v v v v v v v v v v v v v v v v v v	525456			
k m n ob z 101 102 103 104 105 106		<pre><pre><pre></pre></pre></pre>	ALEATORIAS NNS DVK NVT DVT 4 1,0E+06 1,0E+05 2,1E+04 6,6E+03 5 3,4E+07 1,9E+06 2,5E+05 5,9E+04 6 1,1E+09 3,4E+07 3,0E+06 5,3E+05 5	ALEATORIAS DIVERSIDAD TEÓRICA ALEATORIAS NNS DVK NVT DVT 4 2,1E+06 2,1E+05 4,1E+04 1,3E+04 5 6,7E+07 3,8E+06 5,0E+05 1,2E+05 6 2,1E+09 6,8E+07 6,0E+06 1,1E+06	4 2,1E+06 2,1E+05 4,1E+04 1,3E+04 5,7E+07 3,8E+06 5,0E+05 1,2E+05 6,1E+09 6,8E+07 6,0E+06 1,1E+06
CDR3 CDR3 S 99 99 99 99 6 d e f a h i i	C A RK X X X X X X X X X X X X X X X X X X	A BKK X X X X X X X X X X X X X X X X X X	89 90 91 92 93 94 95 96 96a 97 CQ Q X X X X C 5 F G Q G 4 CQ Q X X X X P T F G Q G 3 CQ Q X X X X P X - T F G Q G 1 CQ Q X X X X P X - T F G Q G 1 CQ Q X X X X X P X - F G Q G 5 CQ Q X X X X X P X - F G Q G 5 CQ Q X X X X X X P X T F G Q G 5	BUCLE ESTRUCTURAL PLIEGUE	C G T W/Y D X X X X Y F G G G I W/Y D X X X X - X Y F G G G C G T W/Y D X X X X X X Y F G G G 2
WH LONGITUD FRECUENCIA	7 25 9 6 6 10 8 6 11 11 9 13 10	+ 11 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	VK LONGITUD FRECUENCIA 4 + + 5 + + 7 + + 6	VL LONGITUD FRECUENCIA VL1-44 6 8	VL1-51 6 7 8

Fig. 11

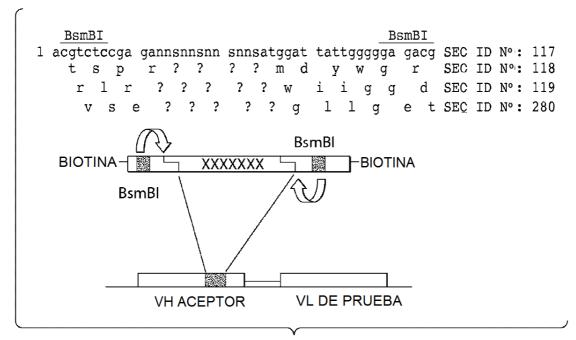
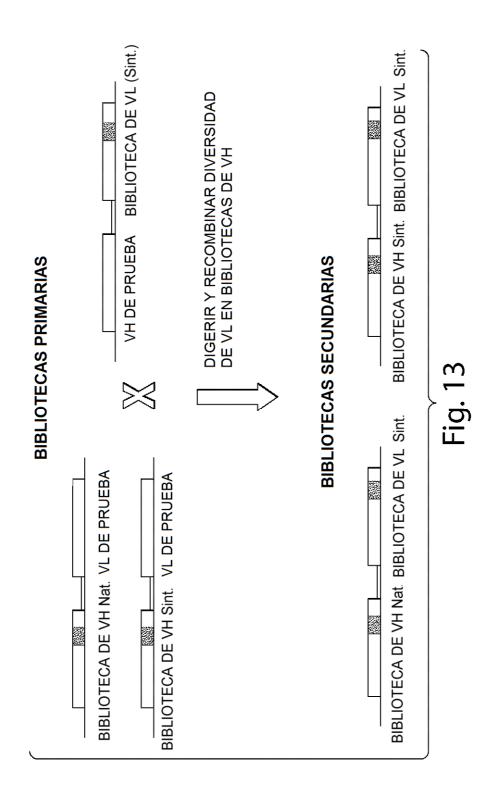
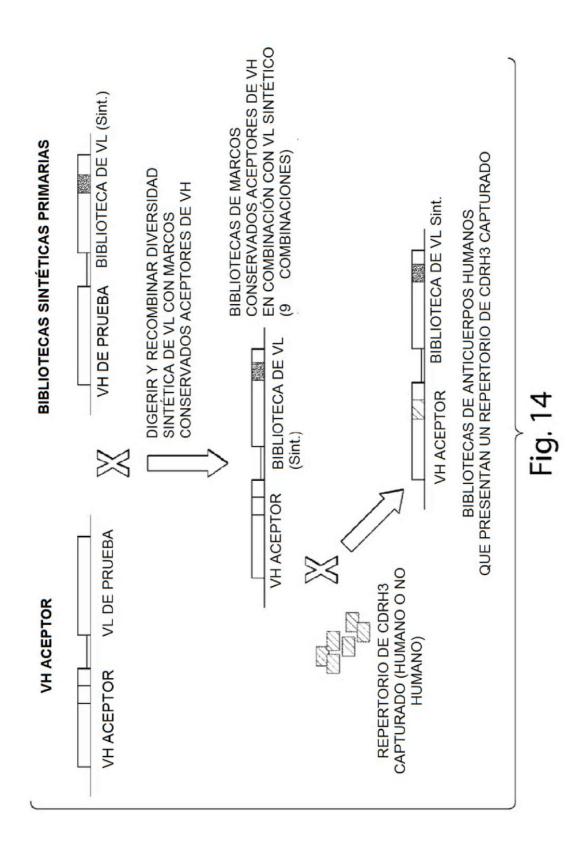
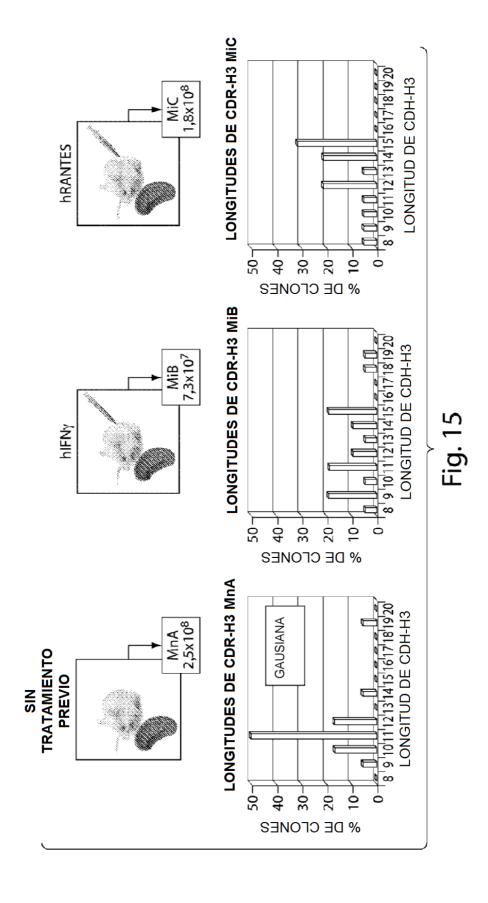


Fig. 12







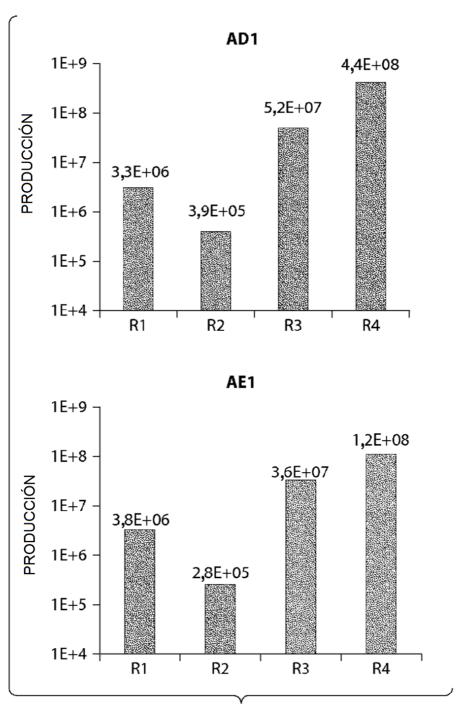


Fig. 16

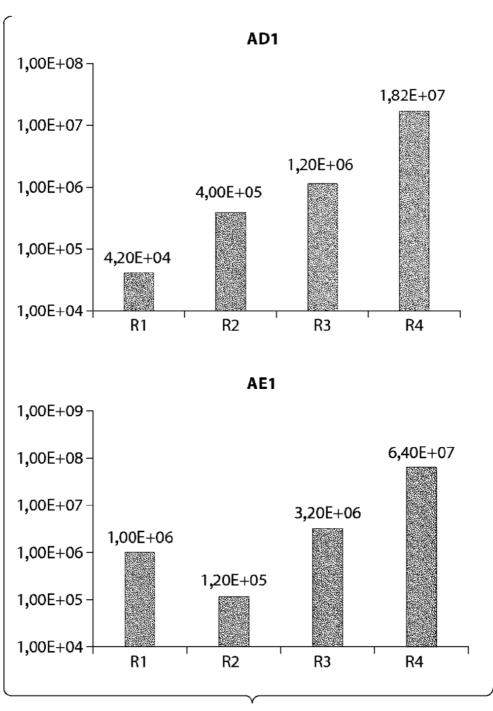
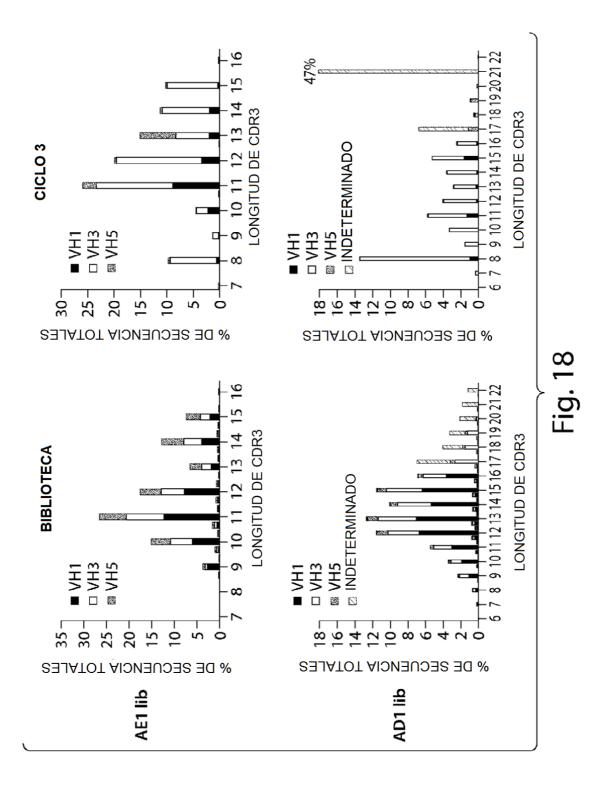


Fig. 17



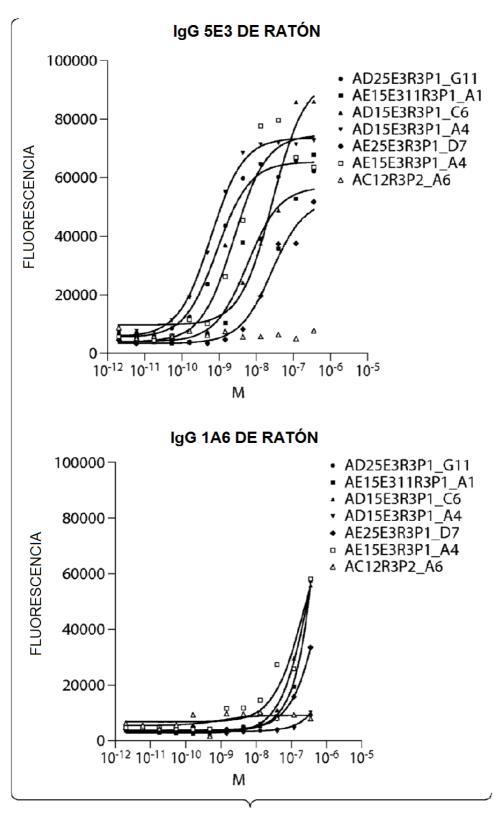
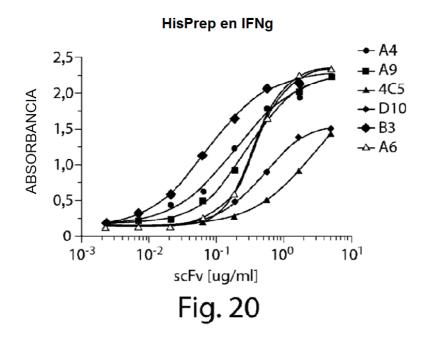
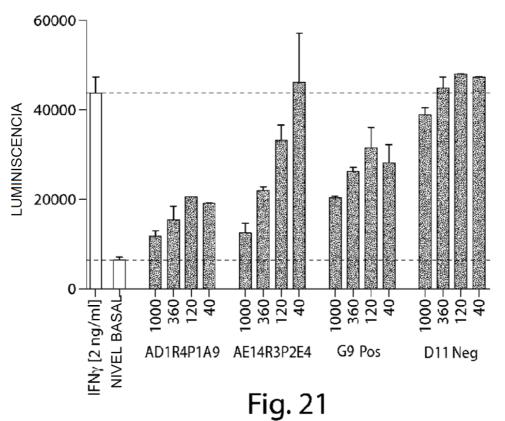
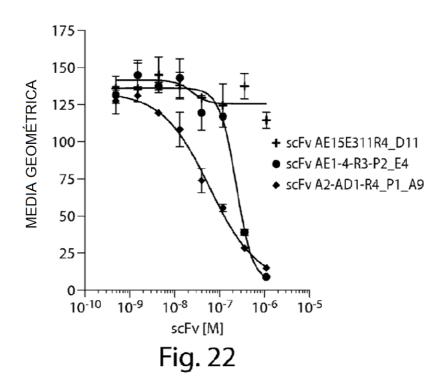


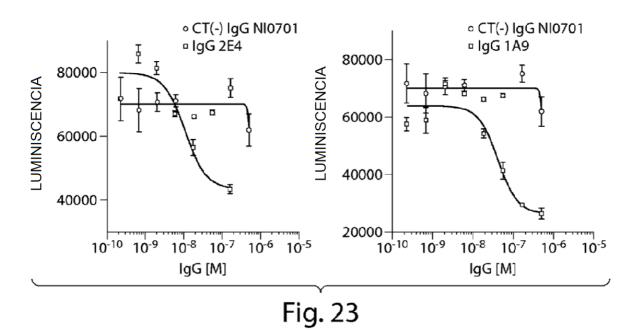
Fig. 19



ACTIVIDAD DE ScFv EN ENSAYO DE LUCIFERASA DE IFN γ







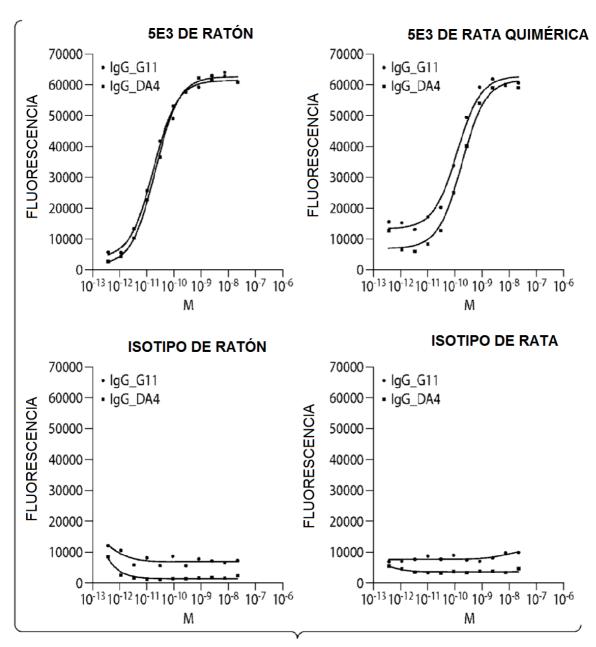


Fig. 24

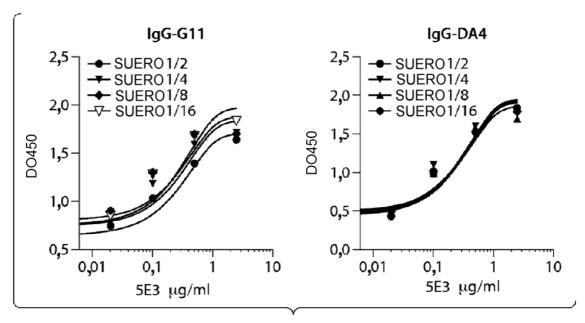


Fig. 25



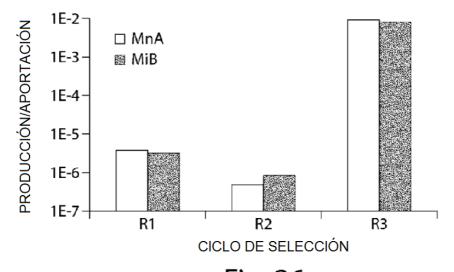
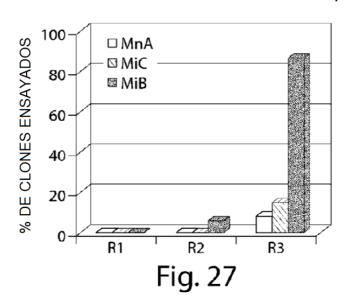
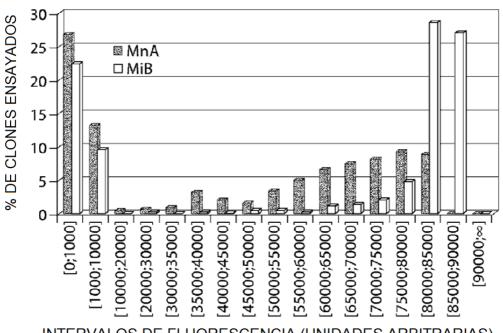


Fig. 26

TASA DE ACIERTOS FRENTE A hIFNY

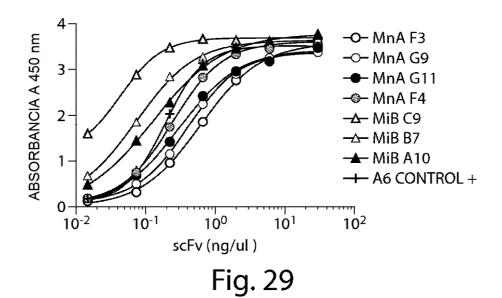


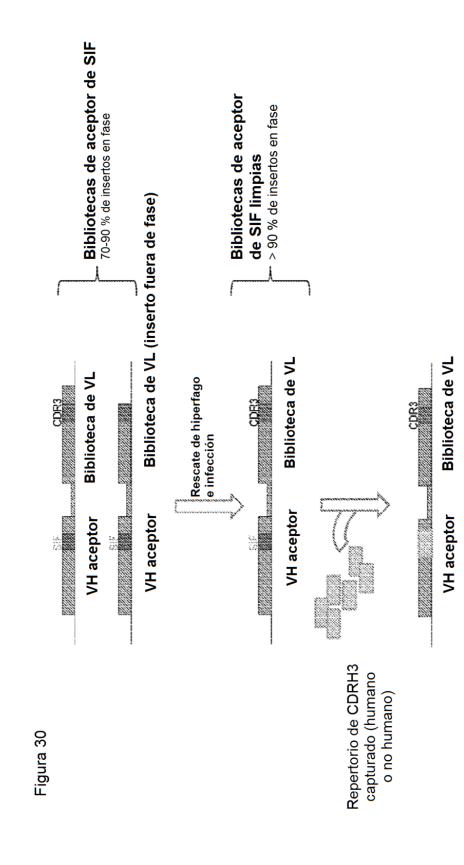
% DE CLONES POSITIVOS PARA DIVERSOS INTERVALOS DE FLUORESCENCIA



INTERVALOS DE FLUORESCENCIA (UNIDADES ARBITRARIAS)

Fig. 28





Bibliotecas de anticuerpos humanos que presentan un repertorio de CDRH3 capturado en combinación con VL preseleccionado

Figura 31

SIF1 sec ID: 290	ATTACTSTGC TAATSACACS	CAGAGORGAC G CTCTCCTCTG G	ATTACTOTOC CAGAGGAGG CANSMACGTCT CTTUGGGGCL GGGAAC TAATGACAGG CTCTTCCTCTG CANSMAGCAGA GAACGCGGGT CCCTTG y c	CTTGGGGCCA.	0000\$ 000770 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	SEC ID Nº : 290 SEC ID Nº : 294 SEC ID Nº : 295
SIF2 sec ID: 291	ATTACTGTGC TAATGACACG	ESSER	######################################	GOGGCAGGG A	**CCCT TTGGG&	SEC ID Nº: 291 SEC ID Nº: 296 SEC ID Nº: 297

F1b41

Department recently con-

\$ X & &

IF1b38

Series Series Series Google Court Co

SECIDNO:

328 329 330

F1b31

egotétocta tgolgoggea gaga o v g d d a v g

322 323 324

gagegengeg sageannagt atat

F1b36

Social Seeds Socialization to the companies and seeds a r s r s r s r s r s

F1b 43

degradagias acquicacque cocr equicacores turcoqueacque casa a c g d y b v o

SEC IDNº

334 33**5** 336

F1b47

gryadangski sogaspuregt eret egeterere tysteregora gaga a z g å e e v s

AURS ...

3/11/83

F1b48

mage subcessor corr ogatekacta tgoggtagoa gaga a x a a a x x v s

SEC IDNO 340 341 342

Figura 32

IF1b2	IF1b6	IF1b 10	IF1b12	IF1b18	IF1b 20	IF1b 26	IF1b 29
SEC DNo	SECIDNO 301 grigadagag kngrutrogt elect 302 egeteletet igegramgem gogm 303 % % % % % % % % %	SECIDNO 304 graphyggsky scypagiscyn stut 305 cynteiteur tytektyda ysga 306 k k g d k y v x	SECIDNO. SANS. SAN	SECIDNO 310 GOGBAGAGGAG BACGGGTGGGT CLGT 311 OUGDOTFORD AGGGGTGGGA GAGGA 312 R R G G G G G	SECIDNO 313 SOURSEAGERS SOURSECUL VIOL 314 SQUICCLOCK LYCKNOGYGOM GROSS 315 % % G G G G G G	SECIDNO. 316 Geographication of the state o	SECIDNO Sensi Banda 319 gognamangang mengapanget elet 320 engetetenta typeregoupen ganga 321 m x g d g v v g

F2b 20

egettteete tyrageagag a grosmacysk arktogroto t degaaaggad accondicte t agatteete tgaggaagaa gogenaagaa acygognoto to ogotitototo to ogotitototo tynogoagaa a SEC IDNO 364 365 366 SECIDNO: 358 359 360 361 362 363 GROBANGAN BOOMCONTON TO CHARLOS OF SACRETON OF SACRETO gognesages enganges cognitions touristics as a second touristic touristics as a second touristic touristic touristics as a second touristic touri gegagaggag aeggegtete t egetetete recedengag a 1 DIDIOCOCCE BEBOSCOCO egenerate tgeogeaga a Norwell . Figura 33 SEC IDNº 346 347 348 343 344 345 350 351 352 353 354

Survey.